Mathematisches Forschungsinstitut Oberwolfach

Report No. 37/2009

DOI: 10.4171/OWR/2009/37

Linear and Nonlinear Eigenproblems for PDEs

Organised by
Andrew Knyazev, Denver
Volker Mehrmann, Berlin
John Osborn, College Park
Jinchao Xu, University Park

August 9th – August 15th, 2009

Abstract. The workshop discussed the numerical solution of linear and nonlinear eigenvalue problems for partial differential equations. It included the theoretical analysis the development of new (adaptive) methods, the iterative solution of the algebraic problems as well as the application in many areas of science and engineering.

Mathematics Subject Classification (2000): 65N25, 65F15, 15A18.

Introduction by the Organisers

The workshop *Linear and Nonlinear Eigenproblems for PDEs*, organised by Andrew Knyazev (University of Colorado, Denver), Volker Mehrmann (Technical University, Berlin), John E. Osborn (University of Maryland, College Park) and Jinchao Xu (Pennsylvania State University, University Park) was held August 9th – August 15th, 2009. This meeting was well attended with over 40 participants with broad geographic representation from all continents.

Numerical solution of linear and nonlinear eigenvalue problems for partial differential equations is an important task in many application areas such as:

- dynamics of electromagnetic fields;
- electronic structure calculations;
- band structure calculations in photonic crystals;
- vibration analysis of heterogeneous material structures;
- particle accelerator simulations;

- vibrations and buckling in mechanics, structural dynamics;
- neutron flow simulations in nuclear reactors.

It involves research in several different areas of mathematics ranging from matrix computation to modern numerical treatment of partial differential equations. Major new research developments in the area of PDE eigenvalue problems that have taken place in recent years include the following:

- meshless and generalized finite element method approximation methods;
- adaptive finite element methods;
- methods for polynomial and other nonlinear eigenvalue problems;
- a priori and a posteriori eigenvalue and eigenvector error estimation;
- convergence theory for preconditioned and inexact eigensolvers;
- multigrid, domain decomposition and incomplete factorization based preconditioning for eigenproblems;
- public software implementing efficient eigensolvers for parallel computers.

Furthermore, important progress has been made for some of the challenging problems in this area, e.g., efficient solvers have been developed for some classes of non-selfadjoint and non-linear eigenvalue problems. On the other hand, many difficult questions remain open even for linear eigenvalue problems as practically interesting engineering applications demand high accuracy of the models, which often leads to ill-conditioned problems of enormous sizes.

We feel that the workshop summed up the recent developments in the different communities in this area and indicated new and fruitful directions. The purpose of this meeting was to bring together researchers with diverse background in matrix computations, numerical methods for PDEs, and relevant application areas, to present the state of the art in the area, and to exchange ideas and approaches for further development.

Senior investigators as well as a significant number of young researchers participated the workshop which format was ideal to stimulate research in this increasingly important and rapidly developing area.

Workshop: Linear and Nonlinear Eigenproblems for PDEs

Table of Contents

Abstracts

Eigenvalue Tensor Calculation

Wolfgang Hackbusch

This contribution is about the use of the tensor calculus for the computation of eigenvalues. We consider eigenvalue problems $Lu = \lambda u$ for a second order elliptic differential operator L in a product domain Ω, e.g., $\Omega = (0,1)^3$, more generally $I_1 \times I_2 \times I_3$ (I_j intervals) or $\Omega = \mathbb{R}^3$. The boundary conditions may be Dirichlet homogeneous values. After discretisation (Galerkin, difference method) the system matrix is denoted by L_h and the eigenvalue problem becomes $L_h u_h = \lambda_h u_h$ or $L_h u_h = \lambda_h M_h u_h$ (M_h: mass matrix).

The true aim of the consideration is the solution of DFT-like formulations in Quantum Chemistry, e.g., the Hartree-Fock or Kohn-Sham equations. The corresponding eigenvalue problem is of the form

$$H\varphi_\alpha = \lambda_\alpha \varphi_\alpha \qquad (\alpha = 1, \dots, \ell) \qquad \text{in } \mathbb{R}^3$$

with the operator depending on the eigenfunctions: $H = H(\varphi_1, \dots, \varphi_\ell)$.

For simplicity, we assume $\Omega = (0,1)^3$ and discretise by a regular grid with n points per direction. The arising dimension n^3 – e.g. for $n = 1000$ – becomes too large for standard eigen solvers. Instead we want to use the tensor calculus. $\mathbb{R}^{n^3}$ is the tensor product $\mathbb{R}^n \otimes \mathbb{R}^n \otimes \mathbb{R}^n$, hence grid functions from $\mathbb{R}^{n^3}$ have the form

$$u = \sum_{\rho=1}^{r} u^{\rho,1} \otimes u^{\rho,2} \otimes u^{\rho,3} \qquad \text{with } u^{\rho,j} \in \mathbb{R}^n. \tag{1}$$

The minimal r in (1) is the *tensor rank*. Fixing r, we consider the set T_r of tensors prepresentable by (1). In that case, r is called the *representation rank* of $u \in T_r$. The obvious advantage is the storage cost of $3rn$. If r is moderate (say $r \le 50$), a tensor with large n is easy to handle.

A first argument why the tensor format is fitting to eigenvalue problems, follow from a separable differential operator $L = L_1 + L_2 + L_3$ with

$$L_j = \frac{\partial}{\partial x_j} a_j(x_j) \frac{\partial}{\partial x_j} + b_j(x_j) \frac{\partial}{\partial x_j} + c_j(x_j).$$

Then all eigenfunctions $u(x_1, x_2, x_3) = u^{\nu_1,1}(x_1) u^{\nu_2,2}(x_2) u^{\nu_3,3}(x_3)$ have tensor format with tensor rank $r = 1$, where $u^{\nu,j}(x_j)$ is the νth eigenfunction of the 1D eigenvalue problem $L_j u^{\nu,j} = \lambda_{\nu,j} u^{\nu,j}$ with $u^{\nu,j}|_{\partial[0,1]} = 0$. These properties are inherited by the discrete eigenvalue problem, if e.g. the Galerkin method is used with a tensor basis functions, i.e., $b_i(x_1, x_2, x_3) = b_{i_1}(x_1) b_{i_2}(x_2) b_{i_3}(x_3)$ for $i = (i_1, i_2, i_3)$.

In the general (non-separable) case, we have two different sources of errors. The *discretisation error* measures the difference between the eigenfunction u from the continuous level and the eigenvector u_h of the discretisation. Considering $u_h \in \mathbb{R}^{n^3}$ as a tensor (with possible high rank), we need a *tensor approximation* $u_{h,r} \in T_r$ of representation rank r, which introduces a second approximation error. While

the discretisation error is well-studied, the discussion of the error for the rank r tensor approximation is new.

The tensor calculus can be used to perform matrix-vector multiplications efficiently. Let the system matrix L_h be of Kronecker tensor format, i.e., $M = \sum_{\nu=1}^{m} M_{\nu,1} \otimes M_{\nu,2} \otimes M_{\nu,3}$ with $M_{\rho,j} \in \mathbb{R}^{n\times n}$. Then a multiplication Mu with u from (1) reduces to 1D products $M_{\nu,j}u_{\rho,j}$ $(1 \le j \le 3)$. However, as other operations the resulting number of terms (representation rank) is increased. Therefore, it is essential that after performing the operations exactly, a suitable truncation T_r from rank r' to r is applied. In general, when $u \mapsto \Phi u$ describes any iteration step, is will be replaced by the truncated iteration $u \mapsto T_r(\Phi u)$. If the power iteration is to be preconditioned, one should choose a positive definite $M_0 = A$ of the form $A = A_1 \otimes I \otimes I + I \otimes A_2 \otimes I + I \otimes I \otimes A_3$. Then the approximation

$$A^{-1} \approx \sum\nolimits_{\nu=1}^{k} \omega_\nu \exp(\alpha_\nu A_1) \otimes \exp(\alpha_\nu A_2) \otimes \exp(\alpha_\nu A_3)$$

has exponential accuracy w.r.t. k (cf. [3]).

An alternative to the representation (1) of tensors from $\mathbb{V} = V^1 \otimes V^2 \otimes V^3$ is the subspace tensor representation, where for suitable r-dimensional subspaces $U^j \subset V^j$ $(\dim U^j = r)$ one approximates $v \in \mathbb{V}$ by $\tilde{v} \in \mathbb{U} := U^1 \otimes U^2 \otimes U^3$. Fixing (orthonormal) bases $(b_{\nu,j})_{\nu=1,\dots,r}$ of U^j, the representation of $\tilde{v}$ is of the form

$$u = \sum\nolimits_{\rho\in\{1,\dots,r\}^3} a_\rho b_{\rho_1,1} \otimes b_{\rho_2,2} \otimes b_{\rho_3,3} \in \mathbb{U}. \tag{2}$$

To estimate the tensor approximation error, the differential operator $Lu := -\mathrm{div}(A \operatorname{grad} u) + b \operatorname{grad} u + cu$ with analytic coefficients has been considered in the paper [5], which is based on [8]. It turns out that the analytic eigenfunction can be approximated by polynomials with exponential accuracy w.r.t. the degree. Hence $\mathbb{U}$ in (2) uses U^j being univariate polynomials of degree $r-1$.

In the case of the Hartree-Fock equation $H\psi_\nu(\mathbf{y}) = \lambda_\nu \psi_\nu(\mathbf{y})$ with $H = -\Delta + V_{\mathrm{nucl}} + V_{\mathrm{Hartree}} + K_{xc}$, the potential V_{nucl} of the nuclii is singular at the location of the atoms. Hence, the coefficients and consequently the solution is not globally analytic. Nevertheless, tensor approximations work well not only for the eigenfunctions ψ_ν, but also for the electron density $\rho = 2\sum_{\nu=1}^{N/2} \psi_\nu^2$, which appears in the Hartree potential $V_{\mathrm{Hartree}}(\mathbf{x}) = \int \frac{\rho(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}\, d\mathbf{y}$. This is demonstrated by numerical examples in [9].

As mentioned above, the truncation to a certain rank (either in the format (1) or (2)) is essential. Unfortunately, in both cases the truncation is a complicated nonlinear iteration which may converge to a local but not global optimum (see [1], [2], [7]). To overcome these difficulties, one may replace the tensor formats (1) or (2) but another one which also allows efficient operations and additionally a fast truncation. Such a new tensor format is described in [6]. Details of this format are given in the lecture of L. Grasedyck [4].

References

[1] M. Espig, *Effiziente Bestapproximation mittels Summen von Elementartensoren in hohen Dimensionen*, doctoral thesis, Universität Leipzig, 2008.

[2] H.-J. Flad, W. Hackbusch, Boris N. Khoromskij, and R. Schneider: *Concept of data-sparse tensor-product approximation in many-particle modelling.* In: Matrix Methods - Theory, Algorithms, Applications (V. Olshevsky, E. Tyrtyshnikov, eds.), World Scientific Publishing, 2008, pages 313-343.

[3] L. Grasedyck, *Existence and computation of low Kronecker-rank approximation to the solution of a tensor system with tensor right-hand side*, Computing **72** (2004), 247-265.

[4] L. Grasedyck, *Hierarchical singular value decomposition of tensors*, this Oberwolfach report.

[5] W. Hackbusch, B.N. Khoromskij, S. Sauter, E.E. Tyrtyshnikov, *Use of Tensor Formats in Elliptic Eigenvalue Problems.* Preprint 78, Max-Planck-Institut für Mathematik, Leipzig, 2008 (submitted).

[6] W. Hackbusch and S. Kühn, *A new scheme for the tensor representation.* Preprint 3, Max-Planck-Institut für Mathematik, Leipzig, 2009 (accepted for J. Fourier Analysis Appl.)

[7] B. Khoromskij, Structured rank $(r_1, \ldots, r_d)$ decomposition of function-related tensors in $\mathbb{R}^d$, Comp. Meth. Appl. Math. **6** (2006), 194-220.

[8] J.M. Melenk, *hp-finite element methods for singular perturbations.* Springer Lecture Notes in Mathematics, **1796**, (2002).

[9] S. Rao Chinnamsetty, M. Espig, B. Khoromskij, W. Hackbusch, and H-J. Flad: *Tensor product approximation with optimal rank in quantum chemistry.* J. Chem. Physics **127** (2007) 08411

Hierarchical Singular Value Decomposition of Tensors

Lars Grasedyck

We define the hierarchical singular value decomposition (SVD) for tensors of order $d \geq 2$. This hierarchical SVD has properties like the matrix SVD (and collapses to the SVD in $d = 2$), and we prove these. In particular, one can find low rank (almost) best approximations in a hierarchical format ($\mathcal{H}$-Tucker) which requires only $\mathcal{O}((d-1)k^3 + dnk)$ data, where d is the order of the tensor, n the size of the modes and k the rank. The $\mathcal{H}$-Tucker format is a specialization of the Tucker format and it contains as a special case all (canonical) rank k tensors. Based on this new concept of a hierarchical SVD we present algorithms for hierarchical tensor calculations allowing for a rigorous error analysis. The complexity of the truncation (finding lower rank approximations to hierarchical rank k tensors) is in $\mathcal{O}((d-1)k^4 + dnk^2)$ and the attainable accuracy is just 2–3 digits less than machine precision.

Let $A \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ be an order d tensor. The hierarchical Tucker format is a multilevel variant of the Tucker format — multilevel in terms of the order of the tensor. In order to define the format we have to introduce a hierarchy among the modes $\{1, \ldots, d\}$.

(Dimension tree) A dimension tree $T_{\mathcal{I}}$ for dimension $d \in \mathbb{N}$ is a tree with root $\{1, \ldots, d\}$ and depth $p = \lceil \log_2(d) \rceil := \min\{i \in \mathbb{N}_0 \mid i \geq \log_2(d)\}$ such that each node $t \in T_d$ is either

(1) a leaf and singleton $t = \{\mu\}$ on level $\ell \in \{p-1, p\}$ or
(2) the disjoint union of two successors $S(t) = \{s_1, s_2\}$: $t = s_1 \dot\cup s_2$.

The level ℓ of the tree is defined as the set of all nodes having a distance of exactly ℓ to the root, cf. Figure 1. We denote the level ℓ of the tree by

$$T_{\mathcal{I}}^{\ell} := \{t \in T_{\mathcal{I}} \mid \text{level}(t) = \ell\}.$$

The set of leaves of the tree is denoted by $\mathcal{L}(T_{\mathcal{I}})$ and the set of interior (non-leaf) nodes is denoted by $\mathcal{I}(T_{\mathcal{I}})$. A node of the tree is a so-called mode cluster.

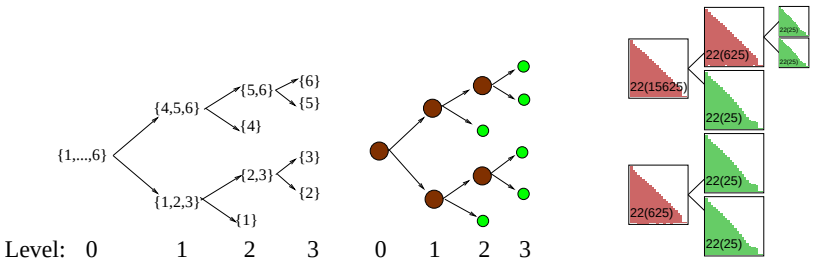

FIGURE 1. Left: A dimension tree for $d = 6$. Right: A hierarchical SVD for tensor with respect to another dimension tree.

For each mode cluster t of the dimension tree we define the **matricization** by introducing the complementary cluster $t' := \{1, \ldots, d\} \setminus t$,

$$\mathcal{I}_t := \bigotimes_{\mu \in t} \mathcal{I}_\mu, \qquad \mathcal{I}_{t'} := \bigotimes_{\mu \in t'} \mathcal{I}_\mu,$$

and the corresponding *t-matricization*

$$\mathcal{M}_t : \mathbb{R}^{\mathcal{I}} \to \mathbb{R}^{\mathcal{I}_t \times \mathcal{I}_{t'}}, \qquad (\mathcal{M}_t(A))_{(i_\mu)_{\mu \in t},(i_\mu)_{\mu \in t'}} := A_{(i_1,\ldots,i_d)},$$

where the special case is $\mathcal{M}_\emptyset(A) := \mathcal{M}_{\{1,\ldots,d\}}(A) := A$. We use the short notation $A^{(t)} := \mathcal{M}_t(A)$.

Now we can define the set of hierarchical ($\mathcal{H}$-) Tucker tensors [3, 4]

$$\mathcal{H}\text{-Tucker}((k_t)_{t \in T_{\mathcal{I}}}) := \{A \in \mathbb{R}^{\mathcal{I}} \mid \forall t \in T_{\mathcal{I}} : \ \text{rank}(A^{(t)}) \le k_t\}$$

and the corresponding hierarchical SVD by the SVDs of the nodes in the tree, cf. Figure 1. This set contains all tensors of canonical rank k for $k_t \equiv k$, it is thus a rather rich manifold, nonetheless it allows a fast arithmetic (truncation to smaller rank) and an efficient storage by exploiting a nested (hierarchical) structure among the matricizations $A^{(t)}$. It avoids the shortcomings of the canonical format [2] and has the same nice properties as the Tucker format [1].

(Theorem)[4] Let $A \in \mathcal{H}\text{-Tucker}((k_t)_{t \in T_{\mathcal{I}}}), k_t \equiv k$ and let $n_j \equiv n$. Then the storage complexity for A is in $\mathcal{O}(dk^3 + dnk)$.

Let U_t denote the matrix that contains as columns the first $\widetilde{k}_t$ left singular vectors of $A^{(t)}$ (corresponding to the largest singular values), and define the projections $\pi_t(A) := \mathcal{M}_t^{-1}(U_t U_t^T A^{(t)})$.

(Theorem)[4] Let $A \in \mathcal{H}\text{-Tucker}((k_t)_{t \in T_{\mathcal{I}}}), k_t \equiv k$ and let $n_j \equiv n$. Define $A_{\mathcal{H}} := \prod_{t \in T_{\mathcal{I}}} \pi_t A$. Then

(1) $A_{\mathcal{H}} \in \mathcal{H}\text{-Tucker}((\widetilde{k}_t)_{t\in T_{\mathcal{I}}})$
(2) $\|A - A_{\mathcal{H}}\| \leq \sqrt{2d-3}\min_{B\in\mathcal{H}\text{-Tucker}((\widetilde{k}_t)_{t\in T_{\mathcal{I}}})} \|A - B\|$
(3) $A_{\mathcal{H}}$ can be computed in $\mathcal{O}(dk^4 + dnk^2)$

References

[1] L. De Lathauwer, B. De Moor and J. Vandewalle. *A multilinear singular value decomposition*, *SIAM* Journal on Matrix Analysis and Applications **21** (2000), 1253–1278.
[2] V. de Silva and L.-H. Lim. *Tensor rank and the ill-posedness of the best low-rank approximation problem*, *Technical Report* SCCM-06-06 (2006), Stanford University.
[3] S. Kühn and W. Hackbusch. *A new scheme for the tensor representation*, *Technical Report* **2** (2009), Max Planck Institute for Mathematics in the Sciences, Leipzig. (submitted to JFAA)
[4] L. Grasedyck. *Hierarchical Singular Value Decomposition of Tensors*, *Technical Report* **27** (2009), Max Planck Institute for Mathematics in the Sciences, Leipzig. (submitted to SIMAX)

Shift-Invert Iteration for Pure Imaginary Eigenvalues for Large Sparse Matrices

Alastair Spence
(joint work with Karl Meerbergen)

In this talk we discuss a numerical procedure for the determination of the smallest λ for which the large, sparse eigenvalue problem

$$(A + \lambda B)x = \mu M x, \tag{1}$$

has a pair of pure imaginary μ's. Here λ is a physical parameter and μ is the eigenvalue of the generalised eigenvalue problem (1).

This work is motivated by the bifurcation analysis of the non-linear dynamical system

$$\frac{du}{dt} = f(u,\lambda) \quad , \quad u(0) = u_0,$$

where f is an operator in $(\mathbf{R}^n, \mathbf{R}) \mapsto \mathbf{R}^n$ with n large. Such analysis includes the computation of bifurcation diagrams, and more particularly, the stability analysis and detection of Hopf bifurcations. In many situations, for example, in nonlinear finite element discretizations, we have an dynamical system of the form

$$M\frac{du}{dt} = f(u,\lambda),$$

where M is a large sparse symmetric positive definite mass matrix. In the case of steady state solutions, i.e. $du/dt = 0$, often the values of λ are sought for which the solution u loses stability. In a linearized stability analysis, the steady state is said to be stable when the eigenvalues μ of

$$J(\lambda)x = \mu M x, \tag{2}$$

have strictly negative real parts, with $J(\lambda)$ denoting the Jacobian matrix evaluated at the steady state $u(\lambda)$, namely, $J(\lambda) = \frac{\partial f}{\partial u}(u(\lambda),\lambda)$. Values of λ where

eigenvalues of (2) cross the imaginary axis indicate a transition from a stable to unstable regime. When stability is lost due to a real eigenvalue μ passing through zero there are many techniques available to determine the critical value of λ. In contrast, at a Hopf bifurcation on a path of stable steady states, (2) has two pure imaginary eigenvalues and their detection is a particularly difficult task for large scale dynamical systems. On the other hand, if close starting values for λ and μ are known then there are good methods, usually based on Newton's Method, for their accurate determination. The contribution of this paper is the determination of good starting values with which to initiate the Hopf calculation.

Perhaps the most straightforward method to detect Hopf points is to monitor the right-most eigenvalues of (2) for a discrete set of λ's. This requires the solution of an eigenvalue problem for each selected λ, which can be very expensive, especially when the system size is large. The next most obvious method is the shift-invert Arnoldi method with zero shift. This is an attractive approach since a matrix factorization of $J(\lambda)$ may well be available from the steady state computation. All these methods are quite reliable for computing eigenvalues near a point or target, but sometimes fail to compute the right-most eigenvalue. This is likely to be the case when there are many eigenvalues lying near zero but the eigenvalues that cross the Imaginary axis have large imaginary part.

For small scale problems Guckenheimer and co-workers (SINUM, 34 (1997) pp1-21) introduced a novel technique to detect Hopf bifurcations based on the use of the bialternate product of $J(\lambda)$, defined as $(J(\lambda) \otimes I + I \otimes J(\lambda))/2$, which is an $n^2 \times n^2$ matrix that has a pair of zero eigenvalues when $J(\lambda)$ has a pair of pure imaginary eigenvalues. This construction forms the first theoretical step in our method, but we emphasise that we do not compute with the Kronecker product forms.

In this talk we build on Guckenheimer's idea, and develop a method applicable to large scale problems. We use an inverse iteration approach to solve the resulting Lyapunov equations.

We consider the situation where $J(\lambda)$ has the form

$$J(\lambda) = (A + \lambda B) \ .$$

In the bifurcation setting, A and $B \in \mathbf{R}^{n\times n}$ arise from a linearization of $f(u, \lambda)$ as follows: Assume that $u(0)$ is a known steady state at $\lambda = 0$, then $J(0)$ is known and $J(\lambda) \approx J(0) + \lambda J'(0)$, which we write as $A + \lambda B$. Here, A and B are usually nonsymmetric matrices and B can be singular. In this talk, we assume the following situation: $\lambda = 0$ corresponds to a stable steady state solution, i.e. all eigenvalues of $Ax = \mu Mx$ lie in the stable half plane. The goal is to compute the smallest λ for which the eigenvalue problem (1) has pure imaginary μ's. Generically, the μ's will be continuous functions of λ and the first λ for which there are μ's on the imaginary axis (including the case when $\mu = 0$), will approximate the value of λ that corresponds to a transition from a stable to unstable steady state, or to a Hopf point.

First, note that the bialternate product of $(A+\lambda B)$ has the form (dropping the factor of $\frac{1}{2}$)

$$(A + \lambda B) \otimes M + M \otimes (A + \lambda B). \tag{3}$$

If (1) has pure imaginary eigenvalues then has a double eigenvalue zero. This translates into the following $n^2 \times n^2$ linear eigenvalue problem

$$(A \otimes M + M \otimes A)z + \lambda(B \otimes M + M \otimes B)z = 0 \ , \tag{4}$$

whose solution gives the values of λ for which (1) has pure imaginary eigenvalues μ. Although this is a nice mathematical property, it should only be used in this form for problems of small size. To produce a method that is applicable to large scale problems we exploit the well-known equivalence between equations involving the Kronecker product and Lyapunov equations. This shows that the eigenvalue problem (4) is equivalent to the eigenvalue problem

$$MZA^T + AZM^T + \lambda(MZB^T + BZM^T) = 0 \ , \tag{5}$$

where $\mathrm{vec}(Z) = z$, with z as in (4). Here Z is an unknown $n \times n$ matrix, which we also call an 'eigenvector' of (5). A new element in this talk is our discussion of the rank of Z and its symmetry properties. A clear understanding of these two aspects turns out to be key in obtaining an efficient algorithm for large scale problems.

Numerical examples, including two physical applications, support the theory in this paper.

Adaptive FE eigenvalue approximation (with application to hydrodynamic stability)

Rolf Rannacher

(joint work with Vincent Heuveline and Antje Westenberger, Winnifried Wollner)

This talk presents an adaptive finite element method for the solution of eigenvalue problems associated with the linearized stability analysis of nonlinear operators in the context of hydrodynamic stability theory. The general framework is the Dual Weighted Residual (DWR) method for local mesh adaptation which is driven by residual-based and sensitivity-controlled a posteriori information (see Becker & Rannacher [6, 7]). The basic idea is to embed the eigenvalue approximation into the general framework of Galerkin methods for nonlinear variational equations for which the DWR method is already well developed (see Rannacher [17] and the monograph Bangerth & Rannacher [2]). The evaluation of these error representations results in a posteriori error bounds for approximate eigenvalues reflecting the errors by discretization of the eigenvalue problem as well as those by linearization about an only approximately known base solution. From these error estimates local error indicators are derived by which economical meshes can be constructed. Some parts of the underlying theory are developed within a more abstract setting suggesting the application to other kinds of nonlinear problems.

On a bounded domain $\Omega \subset \mathbb{R}^d$ with $d = 2$ or $d = 3$, we consider elliptic eigenvalue problems of increasing complexity, such as symmetric diffusion problems

$$-\Delta v = \lambda v, \quad \text{in } \Omega, \quad v_{|\partial\Omega} = 0, \tag{1}$$

nonsymmetric convection-diffusion problems

$$-\Delta v + b \cdot \nabla v = \lambda v, \quad \text{in } \Omega, \quad v_{|\partial\Omega} = 0, \tag{2}$$

and eigenvalue problems in hydrodynamic stability theory,

$$\begin{aligned} &-\nu\Delta v + \hat{v} \cdot \nabla v + v \cdot \nabla \hat{v} + \nabla q = \lambda v, \;\; \nabla \cdot v = 0, \quad \text{in } \Omega, \\ &v|_{\Gamma_{\text{rigid}}} = 0, \;\; v|_{\Gamma_{\text{in}}} = 0, \;\; \nu\partial_n v - pn|_{\Gamma_{\text{out}}} = 0. \end{aligned} \tag{3}$$

Here, $\hat{u} := \{\hat{v}, \hat{p}\}$ is a stationary "base flow", that is a stationary solution of the Navier-Stokes equations

$$\begin{aligned} &-\nu\Delta v + v \cdot \nabla v + \nabla p = f, \;\; \nabla \cdot v = 0, \quad \text{in } \Omega, \\ &v|_{\Gamma_{\text{rigid}}} = 0, \;\; v|_{\Gamma_{\text{in}}} = v^{\text{in}}, \;\; \nu\partial_n v - pn|_{\Gamma_{\text{out}}} = 0, \end{aligned} \tag{4}$$

where v is the velocity vector field of the flow, p is its hydrostatic pressure, ν is the kinematic viscosity (density $\rho \equiv 1$), and f is a prescribed volume force. The goal is to investigate the stability of the base flow $\{\hat{v}, \hat{p}\}$ under small perturbations, which leads us to consider the eigenvalue problem (3). If an eigenvalue of (3) has $\operatorname{Re}\lambda < 0$, the base flow is unstable, otherwise it is said to by "linearly stable". That means that the solution of the linearized nonstationary perturbation problem

$$\partial_t w - \nu\Delta w + \hat{v} \cdot \nabla w + w \cdot \nabla \hat{v} + \nabla q = 0, \quad \nabla \cdot w = 0, \quad \text{in } \Omega, \tag{5}$$

corresponding to an initial perturbation $w_{|t=0} = w^0$ satisfies a bound of the form

$$\sup_{t \geq 0} \|w(t)\| \leq A\|w^0\|, \tag{6}$$

for some constant $A \geq 1$, where $\|\cdot\|$ denotes the L^2-norm over Ω and $(\cdot,\cdot)$ the corresponding inner product. However, *linear* stability does not guarantee full *nonlinear* stability due to effects caused by the non-normality of problem (3) making the constant A large (see Landahl [13] and the monograph Trefethen & Embree [18]).

The finite element discretization of these eigenvalue problems is based on variational formulations. It uses finite element spaces V_h consisting of piecewise polynomial functions on certain decompositions $\mathbb{T}_h$ of the domain $\bar{\Omega}$ into cells $T \in \mathbb{T}_h$ (triangle, quadrilateral, etc.) with diameter $h_T := \operatorname{diam}(T)$. Our primal goal is to derive computable *a posteriori* estimates for the error $\lambda - \lambda_h$ in terms of the "cell residuals" of the computed approximations $\hat{v}_h$ and $\{v_h, \lambda_h\}$. For symmetric eigenvalue problems of the kind (1) *a posteriori* error estimates of this type have been derived by Nystedt [15], Larson [14], and Verfürth [19]. Their proofs largely exploit symmetry and also require the H^2-regularity of the problem which excludes domains with reentrant corners. In Heuveline & Rannacher [10, 11] nonsymmetric eigenvalue problems of the kind (2) have been treated by employing duality techniques from optimal control theory. In this approach the central idea is to

simultaneously consider the approximation of the the "primal" eigenvalue problem $Au = \lambda u$ and its associated "dual" analogue $A^*u^* = \bar{\lambda}u^*$. The combined problem is embedded into the general optimal-control framework of Galerkin approximations of nonlinear variational equations developed in Becker & Rannacher [6, 7]. This results in an *a posteriori* error estimate of the form

$$|\lambda - \lambda_h| \approx \sum_{T\in\mathbb{T}_h} h_T^2 \Big\{ \hat{\rho}_T(\hat{u}_h)\hat{\omega}_T^* + \hat{\rho}_T^*(\hat{u}_h^*)\hat{\omega}_T + \rho_T(u_h,\lambda_h)\omega_T^* + \rho_T^*(u_h^*,\bar{\lambda}_h)\omega_T \Big\}, \tag{7}$$

involving the cell residuals $\hat{\rho}_T$, $\hat{\rho}_T^*$, and ρ_T, ρ_T^* of the computed primal and dual base solutions $\hat{u}_h$, $\hat{u}_h^*$, and eigenpairs $\{u_h,\lambda_h\}$, $\{u_h^*,\bar{\lambda}_h\}$, and corresponding local weights (sensitivity factors) $\hat{\omega}_T,\hat{\omega}_T^*$ and ω_T,ω_T^*. These weights, though depending on the unknown "exact" solutions, can be cheaply approximated from the computed discrete solutions by local postprocessing. The error estimate (7) accomplishes simultaneous control of the error in the linearization, $\hat{v} - \hat{v}_h$, and the error in the resulting eigenvalues, $\lambda - \lambda_h$. The cellwise error indicators can then guide the mesh refinement process. Further, we devise a simple computational criterion based on the inner products (v_h, v_h^*) which can detect possible instability caused by non-normality effects. The practical features of this method are illustrated for several model situations including a scalar convection-diffusion problem and a problem arising in stationary optimal flow control, drag minimization by stationary boundary control (see Becker [3, 4] and Becker, Heuveline & Rannacher [5]).

We note that the approach of simultaneously considering the primal and dual eigenvalue problems is also crucial for constructing efficient multigrid solvers for the discrete problems. This has been successfully exploited for nonsymmetric elliptic eigenvalue problems in Heuveline & Bertsch [9].

The *a priori* error analysis for the nonsymmetric eigenvalue problem is well developed in the literature, Bramble & Osborn [8], Osborn [16], Babuska & Osborn [1], and the literature cited therein). Particularly to mention is Osborn [16], where the eigenvalue problem of the linearized Navier-Stokes equations is considered, though neglecting the additional error due to the approximative linearization. These studies usually employ the heavy machinery of resolvent integral calculus as described in Kato [12]. The *a posteriori* error analysis only needs arguments from elementary calculus since it is based on the assumption that the approximation is sufficiently accurate on the considered meshes. The justification of this assumption is supplied by the *a priori* error analysis.

References

[1] I. Babuska and J. E. Osborn. *Eigenvalue problems.* Handbook of Numerical Analysis, Vol. 2, Chapter Finite Element Methods, Part 1 (P.G. Ciarlet and J.L. Lions, eds), pp. 641–792, Elsevier, 1991.

[2] W. Bangerth and R. Rannacher. *Adaptive Finite Element Methods for Differential Equations.* Birkhäuser: Basel-Boston-Berlin, 2003.

[3] R. Becker. *An optimal-control approach to a posteriori error estimation for finite element discretizations of the Navier-Stokes equations*. East-West J. Numer. Math. 9, 257–274 (2000).
[4] R. Becker. *Mesh adaptation for stationary flow control*. J. Math. Fluid Mech. 3, 317–341 (2001).
[5] R. Becker, V. Heuveline and R. Rannacher. *An optimal control approach to adaptivity in computational fluid mechanics*, Int. J. Numer. Meth. Fluids. 40, 105-120 (2002).
[6] R. Becker and R. Rannacher. *A feed-back approach to error control in finite element methods: basic analysis and examples*. East-West J. Numer. Math. 4, 237–264 (1996).
[7] R. Becker and R. Rannacher. *An optimal control approach to error control and mesh adaptation in finite element methods*. Acta Numerica 2001 (A. Iserles, ed.), Cambridge University Press, 2001.
[8] J.H. Bramble and J.E. Osborn. *Rate of convergence estimates for nonselfadjoint eigenvalue approximations*. Math. Comp. 27, 525–545 (1973).
[9] V. Heuveline and C. Bertsch. *On multigrid methods for the eigenvalue computation of nonselfadjoint elliptic operators*. East-West J. Numer. Math. 8, 275–297 (2000).
[10] V. Heuveline and R. Rannacher. *A posteriori error control for finite element approximations of elliptic eigenvalue problems*. J. Comput. Math. Appl. 15, 107–138 (2001).
[11] V. Heuveline, R. Rannacher: *Adaptive FE eigenvalue approximation with application to hydrodynamic stability analysis*. Proc. Int. Conf. Advances in Numerical Mathematics, Moscow, Sept. 16-17, 2005 (W. Fitzgibbon et al., eds), pp. 109–140, Institute of Numerical Mathematics RAS, Moscow, 2006.
[12] T. Kato. *Perturbation Theory for Linear Operators*, Springer, Berlin-Heidelberg-New York, 1966.
[13] M. Landahl. *A note on an algebraic instability of viscous parallel shear flows*, J. Fluid Mech. 98, 243 (1980).
[14] M.G. Larson. *A posteriori and a priori error analysis for finite element approximations of self-adjoint elliptic eigenvalue problems*. SIAM J. Numer. Anal. 38, 608–625 (2000).
[15] C. Nystedt. *A priori and a posteriori error estimates and adaptive finite element methods for a model eigenvalue problem*. Technical Report N0 1995-05, Department of Mathematics, Chalmers University of Technology, 1995.
[16] J. E. Osborn. *Spectral approximation for compact operators*. Math. Comp. 29, 712–725 (1975).
[17] R. Rannacher. *Lectures on numerical flow simulation: Finite element discretization, numerical boundary conditions and automatic mesh adaptation*. In Hemodynamical Flows: Aspects of Modeling, Analysis and Simulation (G.P. Galdi et al., eds), Birkhäuser, Basel-Boston-Berlin, 2008.
[18] L. N. Trefethen and M. Embree. *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators*, Princeton University Press, Princeton and Oxford, 2005.
[19] R. Verfürth. *A Review of a posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*. John Wiley/Teubner, New York-Stuttgart, 1996.

Nonconforming Maxwell Eigensolvers

Susanne C. Brenner

(joint work with Jintao Cui, Fengyan Li, Jiangguo Liu and Li-yeng Sung)

Let $\Omega \subset \mathbb{R}^2$ be a bounded polygonal domain. We consider three nonconforming eigensolvers for the following Maxwell eigenproblem with perfectly conducting boundary condition: Find $\lambda \in \mathbb{R}$ such that there exists a nontrivial $\mathbf{u} \in$

$H_0(\mathrm{curl};\Omega)\cap H(\mathrm{div}^0;\Omega)$ satisfying

$$(\nabla\times\mathbf{u},\nabla\times\mathbf{v})=\lambda(\mathbf{u},\mathbf{v})\qquad\forall\,\mathbf{v}\in H_0(\mathrm{curl};\Omega)\cap H(\mathrm{div}^0;\Omega).$$

Let $c_1,\ldots,c_L$ be the corners of Ω and ω_ℓ be the interior angle at c_ℓ. We use a triangulation $\mathcal{T}_h$ of Ω with the property that $h_T\approx\Phi_\mu(T)$, where the weight $\Phi_\mu(T)$ is defined by

$$\Phi_\mu(T)=\Pi_{\ell=1}^{L}|c_\ell-c_T|^{\mu_\ell},$$

and the grading parameters $\mu_1,\ldots,\mu_L$ are chosen according to

$$\begin{aligned}\mu_\ell&=1 &&\text{if}\quad \omega_\ell\le\frac{\pi}{2}\\ \mu_\ell&\le\frac{\pi}{2\omega_\ell} &&\text{if}\quad \omega_\ell>\frac{\pi}{2}\end{aligned}$$

Let $V_{1,h}$ be the space of Crouzeix-Raviart locally divergence-free weakly continuous P_1 vector fields associated with $\mathcal{T}_h$ whose tangential components vanish at the midpoints of the boundary edges. The first eigensolver computes $\lambda_h\in\mathbb{R}$ such that there exists a nontrivial $\mathbf{u}_h\in V_{1,h}$ satisfying

$$a_h(\mathbf{u}_h,\mathbf{v})=\lambda_h(\mathbf{u}_h,\mathbf{v})\qquad\forall\,\mathbf{v}\in V_{1,h},$$

where

$$\begin{aligned}a_h(\mathbf{w},\mathbf{v})=&\sum_{T\in\mathcal{T}_h}\int_T(\nabla\times\mathbf{w})(\nabla\times\mathbf{v})dx+\sum_{e\in\mathcal{E}_h}\frac{[\Phi_\mu(e)]^2}{|e|}\int_e[\![\mathbf{n}\times\mathbf{w}]\!]\,[\![\mathbf{n}\times\mathbf{v}]\!]\,ds\\&+\sum_{e\in\mathcal{E}_h^i}\frac{[\Phi_\mu(e)]^2}{|e|}\int_e[\![\mathbf{n}\cdot\mathbf{w}]\!][\![\mathbf{n}\cdot\mathbf{v}]\!]ds,\end{aligned}$$

$\mathcal{E}_h$ (resp. $\mathcal{E}_h^i$) is the set of edges (resp. interior edges) of $\mathcal{T}_h$, $|e|$ denotes the length of the edge e, $\mathbf{n}$ is a unit normal of e, and $[\![\mathbf{n}\times\mathbf{v}]\!]$ (resp. $[\![\mathbf{n}\cdot\mathbf{v}]\!]$) denotes the jump of the tangential (resp. normal) component of $\mathbf{v}$. The weight $\Phi_\mu(e)$ is defined by

$$\Phi_\mu(e)=\Pi_{\ell=1}^{L}|c_\ell-m_e|^{1-\mu_\ell},$$

where m_e is the midpoint of e.

For domains that are not simply connected, the Crouzeix-Raviart space does not have a completely local basis. This difficulty can be avoided by replacing $V_{1,h}$ by $V_{2,h}$, the space of weakly continuous P_1 vector fields, without the local divergence-free condition. The second eigensolver computes $\lambda_h\in\mathbb{R}$ such that there exists a nontrivial $\mathbf{u}_h\in V_{2,h}$ satisfying

$$a_{2,h}(\mathbf{u}_h,\mathbf{v})=\lambda_h(\mathbf{u}_h,\mathbf{v})\qquad\forall\,\mathbf{v}\in V_{2,h},$$

where

$$a_{2,h}(\mathbf{w},\mathbf{v})=a_{1,h}(\mathbf{w},\mathbf{v})+\sum_{T\in\mathcal{T}_h}h_T^{-2}\int_T(\nabla\cdot\mathbf{w})(\nabla\cdot\mathbf{v})dx.$$

For the third eigensolver we take $V_{3,h}$ to be the space of discontinuous locally divergence-free P_1 vector fields and compute $\lambda_h\in\mathbb{R}$ such that there exists a nontrivial $\mathbf{u}_h\in V_{3,h}$ satisfying

$$a_{3,h}(\mathbf{u}_h,\mathbf{v})=\lambda_h(\mathbf{u}_h,\mathbf{v})\qquad\forall\,\mathbf{v}\in V_{3,h},$$

where

$$a_{3,h}(\mathbf{w},\mathbf{v}) = a_{1,h}(\mathbf{w},\mathbf{v}) + h^{-2}\sum_{e\in\mathcal{E}_h}\frac{1}{|e|}\int_e \big(\Pi_e^0[\![\mathbf{n}\times\mathbf{w}]\!]\big)\big(\Pi_e^0[\![\mathbf{n}\times\mathbf{v}]\!]\big)\,ds$$
$$+ h^{-2}\sum_{e\in\mathcal{E}_h^i}\frac{1}{|e|}\int_e \big(\Pi_e^0[\![\mathbf{n}\cdot\mathbf{w}]\!]\big)\big(\Pi_e^0[\![\mathbf{n}\cdot\mathbf{v}]\!]\big)\,ds,$$

and Π_e is the orthogonal projection from $L_2(e)$ onto $P_0(e)$. This eigensolver can be implemented on nonconforming meshes since the vector fields in $V_{3,h}$ are discontinuous.

The analysis of the eigensolvers are based on the L_2 error estimates for the corresponding nonconforming methods for the source problem, which imply that the solution operators of the discrete problems converge in the L_2 operator norm to the solution operator of the continuous problem. Since the space $H_0(\mathrm{curl};\Omega)\cap H(\mathrm{div}^0;\Omega)$ is compactly embedded in $[L_2(\Omega)]^2$, the convergence of the eigensolvers then follows immediately from the classical theory of spectral approximation. In particular, no spurious eigenvalues are generated by any of these eigensolvers.

We also present preliminary 3D numerical results for the unit cube showing that the discrete eigenvalues are well separated with the correct multiplicity.

Details for the eigensolvers and related results for the source problem can be found in the papers listed in the references.

References

[1] S.C. Brenner, J. Cui, F. Li and L.-Y. Sung, *A nonconforming finite element method for a two-dimensional curl-curl and grad-div problem*, Numer. Math. **109** (2008), 509–533.

[2] S.C. Brenner, F. Li and L.-Y. Sung, *A locally divergence-free nonconforming finite element method for the time-harmonic Maxwell equations*, Math. Comp. **76** (2007), 573–595.

[3] S.C. Brenner, F. Li and L.-Y. Sung, *A locally divergence-free interior penalty method for two dimensional curl-curl problems*, SIAM J. Numer. Anal. **46** (2008), 1190–1211.

[4] S.C. Brenner, F. Li and L.-Y. Sung, *A nonconforming penalty method for a two dimensional curl-curl problem*, M3AS **19** (2009), 651–668.

[5] S.C. Brenner, F. Li and L.-Y. Sung, *Nonconforming Maxwell Eigensolvers*, J. Sci. Comput. **40** (2009), 51–85.

[6] S.C. Brenner, J. Cui and L.-Y. Sung, *An interior penalty method for a two dimensional curl-curl and grad-div Problem*, ANZIAM **50** (2009), C947–C975.

An eigenvalue problem coming from the computation of resonances in open systems using a perfectly matched layer

Joseph E. Pasciak

(joint work with Seungil Kim)

In this talk, we consider the computation of resonance values in open systems using approximations coming from the PML (perfectly matched layer) technique. Problems involving resonances in open systems result from many applications, including the modeling of slat and flap noise from an airplane wing, gravitational

waves in astrophysics, and quantum mechanical systems. We consider the approximation of resonances in the frequency domain.

As a model problem, we consider a resonance problem in three dimensional space which results from a compactly supported perturbation of the Laplacian, i.e.,

$$Lu = -\Delta u + L_1 u$$

where L_1 is symmetric and lives on a bounded domain $\Omega \subset \mathbb{R}^3$. We then seek k such that there are non-trivial "eigenfunctions" ψ satisfying

$$L\psi = k^2\psi.$$

The function ψ is required to satisfy an outgoing condition corresponding to the wave number k. A resonance value k corresponds to an improper eigenvalue problem as the corresponding eigenvector grows exponentially. This exponential growth makes the problem difficult to formulate and even more difficult to approximate numerically.

The PML technique was introduced by Bérenger for scattering problems [3, 4]. The idea was to surround the area of computational interest by an absorbing media (layer) which was perfectly free of reflection. Bérenger's original formulation involved splitting the equations in the absorbing layer. Subsequent formulations avoid this splitting and often can be interpreted as a complex stretching, c.f., [8]. A properly devised complex stretching (or change of variable) preserves the solution inside the layer while introducing exponential decay at infinity. Thus, it is natural to truncate the problem to a finite domain, introduce a convenient boundary condition on the artificial boundary and apply the finite element method. This is a very successful technique which has been investigated both theoretically and computationally.

For scattering problems, one traditionally introduces a PML stretching which depends on the wave number yielding wave number independent decay [14]. In contrast, for PML resonance problems, we use a stretching independent of the wave number. This results in wave number dependent decay. This is important in that when this decay is stronger than the exponential growth of the resonance eigenfunction, this eigenfunction is transformed into a proper eigenfunction for the PML equation (still on the infinite domain). In fact, the resulting PML eigenfunction still has residue exponential decay. Because of this decay, it is natural to consider truncating to a finite domain and subsequently applying the finite element method. The goal of this talk is to study the eigenvalue behavior when the eigenvalues of the (infinite) PML problem are approximated by those corresponding to a finite element method on a truncated computational domain.

The application of PML for the computation of resonances is not new. It turns out that PML is related to what is called "spectral deformation theory" developed by Aguilar, Balslev, Combes and Simon [1, 2, 11, 17]. This theory shows that the PML eigenvalues for the infinite domain problem coincide with the resonance values of the original problem for rather general scaling schemes. The experimental behavior of truncated PML approximations to resonances in open systems was

investigated in [9, 10]. Our paper provides the first theoretical analysis of the convergence of the truncated PML eigenvalue approximation.

In many applications, the convergence of approximate eigenvalues is the result of classical perturbation theory (cf. [12]). Indeed, one is able to conclude basic eigenvalue convergence provided that one can show that approximate operators converge in norm to the appropriate continuous operator. The PML eigenvalue problem is interesting as norm convergence does not hold in this case and so eigenvalue convergence has to be proved in a more basic way.

The convergence of the eigenvalues obtained by finite element discretization has had a long history of research. Osborn [16] and Bramble-Osborn [7] consider the case of second order elliptic problems on bounded domains. In this case, the solution operator is compact and it is possible to prove convergence in norm of the discrete approximation. More recent work has been done on the Maxwell eigenvalue problem [5, 6, 13, 15]. This problem is more difficult than the uniformly elliptic problem mentioned above due to a non-compact inverse however, in this case, the only trouble is due to gradient fields and results in a non-discrete spectrum consisting of only one point, the origin. In contrast, the PML problem is posed on an infinite domain and its inverse has essential spectrum while the inverse of the truncated PML problem is compact. Accordingly, there cannot be norm convergence and the analysis of the convergence of eigenvalues/eigenvector must proceed in a non-standard fashion. Of course, the critical property here is the decay of the approximated eigenfunctions and it is this property which is central to our analysis.

The question of spurious eigenvalues is delicate. First of all, numerical methods at any level of discretization usually generate some numerical eigenvalues which have nothing to do with the actual eigenvalues. For example, high frequency eigenvector components in a Galerkin approximation to say the Dirichlet problem on a bounded domain exhibit little eigenvalue accuracy. Fortunately, the corresponding eigenvalues remain far away from the lower eigenvalues which have the desired convergence. Thus, the goal is to have a method where the eigenvalues of interest are approximated well and can be easily identified from the extraneous numerical ones.

The PML resonance case is much more complicated. Computational eigenvalues not related to the resonance values result from the essential spectrum associated with the infinite domain PML problem, the domain truncation, the mesh size and the placement of the PML layer. At least in the model problem which we are considering, the location of the essential spectrum is *apriori* known so the numerical eigenvalues associated with it are easy to identify. Other "spurious like" eigenvalues can be moved away by changing the layer placement, truncation domain size and mesh parameter. Our theorems guarantee convergence at the resonance values provided that the mesh parameter is sufficiently small and the computational domain is sufficiently large. How to choose all of these parameters in such a way to guarantee that the resonances of interest in an actual application have been located remains a difficult question.

References

[1] J. Aguilar and J. M. Combes. A class of analytic perturbations for one-body Schrödinger Hamiltonians. *Comm. Math. Phys.*, 22:269–279, 1971.

[2] E. Balslev and J. M. Combes. Spectral properties of many-body Schrödinger operators with dilatation-analytic interactions. *Comm. Math. Phys.*, 22:280–294, 1971.

[3] J.-P. Berenger. A perfectly matched layer for the absorption of electromagnetic waves. *J. Comput. Phys.*, 114(2):185–200, 1994.

[4] J.-P. Berenger. Three-dimensional perfectly matched layer for the absorption of electromagnetic waves. *J. Comput. Phys.*, 127(2):363–379, 1996.

[5] D. Boffi. Fortin operator and discrete compactness for edge elements. *Numer. Math.*, 87(2):229–246, 2000.

[6] D. Boffi, P. Fernandes, L. Gastaldi, and I. Perugia. Computational models of electromagnetic resonators: analysis of edge element approximation. *SIAM J. Numer. Anal.*, 36(4):1264–1290 (electronic), 1999.

[7] J. Bramble and J. Osborn. Rate of convergence estimates for nonselfadjoint eigenvalue approximations. *Math. Comp.*, 27:525–549, 1973.

[8] W. Chew and W. Weedon. A 3d perfectly matched medium for modified Maxwell's equations with stretched coordinates. *Microwave Opt. Techno. Lett.*, 13(7):599–604, 1994.

[9] S. Hein, T. Hohage, and W. Koch. On resonances in open systems. *J. Fluid Mech*, 506:225–284, 2004.

[10] S. Hein, T. Hohage, W. Koch, and J. Schöberl. Acoustic resonances in a high lift configuration. *J. Fluid Mech*, 582:179–202, 2007.

[11] P. D. Hislop and I. M. Sigal. *Introduction to spectral theory with applications to Schrödinger operators*, volume 113 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1996.

[12] T. Kato. *Perturbation theory for linear operators*. Springer-Verlag, Berlin, second edition, 1976.

[13] F. Kukuchi. On a discrete compactness property for the Nédélec finite elements. *J. Fac. Sci. Univ. Tokyo*, Sect. 1A, Math, 36:479–490, 1989.

[14] P. Monk. *Finite element methods for Maxwell's equations*. Numerical Mathematics and Scientific Computation. Oxford University Press, New York, 2003.

[15] P. Monk and L. Demkowicz. Discrete compactness and the approximation of Maxwell's equations in $\mathbb{R}^3$. *Math. Comp.*, 70(234):507–523, 2001.

[16] J. E. Osborn. A note on a perturbation theorem for the matrix eigenvalue problem. *Numer. Math.*, 13:152–153, 1969.

[17] B. Simon. Resonances in n-body quantum systems with dilatation analytic potentials and the foundations of time-dependent perturbation theory. *Ann. of Math. (2)*, 97:247–274, 1973.

Quasi-Optimal Convergence and Computational Complexity of the Adaptive Finite Element Method for Symmetric Elliptic Eigenvalue Problems

Joscha Gedicke

(joint work with Carsten Carstensen)

The eigenvalue problems for second order elliptic boundary value problems involve the discretisation error of some adaptive finite element method as well as the error left from some iterative solver for the algebraic eigenvalue problem. This talk presents the first adaptive finite element eigenvalue solver (AFEMES) of overall quasi-optimal complexity.

The quasi-optimal convergence and computational complexity of the adaptive finite element method (AFEM) for the source problem has recently been established [3, 11]. Convergence of the AFEM for the symmetric eigenvalue problem has been proven in [2, 6, 7, 10]. In [6] the inner node property and an extra refinement of oscillations is needed. Convergence for more general eigenvalue problems with possibly jumping coefficients and without the inner node property has been derived in [7]. In [2] a convergence proof for a reduced edge residual estimator without oscillations is given. These works on convergence as well as on quasi-optimal convergence [5] of the AFEM for the eigenvalue problem do assume unrealistically the exact knowledge of algebraic eigenpairs. While a second optimality result for linear symmetric operator eigenvalue problems [4] is based on coarsening. Assuming a saturation assumption, [8, 9] present combined adaptive finite element and linear algebra algorithms.

In this talk the quasi-optimal convergence for the simple Laplace eigenproblem is first shown for exact algebraic eigenvalue solutions without using the inner node property or any oscillations: Suppose that (λ_ℓ, u_ℓ) is a discrete eigenpair to the continuous eigenpair (λ, u). Let $u \in \mathcal{A}_s$, $\mathcal{A}_s$ denotes an approximation space, and $(\mathcal{T}_\ell)_\ell$ be a sequence of nested regular triangulations. Then (λ_ℓ, u_ℓ) converges quasi-optimal,

$$|\lambda - \lambda_\ell| + |||u - u_\ell|||^2 \lesssim (|\mathcal{T}_\ell| - |\mathcal{T}_0|)^{-2s}.$$

The notation $a \lesssim b$ abbreviates the inequality $a \leq Cb$ with a constant $C > 0$ which does not depend on the mesh-size. Finally, $|\mathcal{T}_\ell|$ denotes the cardinality of $\mathcal{T}_\ell$. In contrast to [5] the proofs are based on the eigenvalue formulation and not on a relation to its corresponding source problem. Hence, no additional oscillations arising from the corresponding source problem have to be treated. In a second step this result is extended to the case of inexact algebraic eigenvalue solutions: Suppose (λ, u) with $u \in \mathcal{A}_s$ is an exact eigenpair and (λ_ℓ, u_ℓ) and $(\lambda_{\ell+1}, u_{\ell+1})$ corresponding discrete eigenpairs on level ℓ and $\ell + 1$. Let the iterative approximations (μ_ℓ, w_ℓ) on $\mathcal{T}_\ell$ and $(\mu_{\ell+1}, w_{\ell+1})$ on $\mathcal{T}_{\ell+1}$ satisfy

$$|\lambda_{\ell+1} - \mu_{\ell+1}| + |\lambda_\ell - \mu_\ell| + |||u_{\ell+1} - w_{\ell+1}|||^2 + |||u_\ell - w_\ell|||^2 \leq \omega \eta_\ell^2(\mu_\ell, w_\ell)$$

for sufficiently small $\omega > 0$. Then the iterative solutions μ_ℓ and w_ℓ converge optimal up to positive constants,

$$|\lambda - \mu_\ell| + |||u - w_\ell|||^2 \lesssim (|\mathcal{T}_\ell| - |\mathcal{T}_0|)^{-2s}.$$

Finally this result leads to an optimal combined AFEM and numerical linear algebra algorithm AFEMES which does not need any coarsening or saturation assumption [1].

References

[1] C. Carstensen and J. Gedicke, *An adaptive finite element eigenvalue solver of quasi-optimal computational complexity*, MATHEON-Preprint **662** (2009).

[2] C. Cartensen and J. Gedicke, *An oscillation-free adaptive FEM for symmetric eigenvalue problems*, MATHEON-Preprint **489** (2008).

[3] J.M. Cascon, C. Kreuzer, R.H. Nochetto, and K.G. Siebert, *Quasi-optimal convergence rate for an adaptive finite element method*, SIAM J. Numer. Anal. **46** (2008), 2524–2550.
[4] W. Dahmen, T. Rohwedder, R. Schneider, and A. Zeiser, *Adaptive eigenvalue computation: complexity estimates*, Numer. Math. **110** (2008), 277–312.
[5] X. Dai, J. Xu, and A. Zhou, *Convergence and optimal complexity of adaptive finite element eigenvalue computations*, Numer. Math. **110** (2008), 313–355.
[6] S. Giani and I.G. Graham, *A convergent adaptive method for elliptic eigenvalue problems*, SIAM J. Numer. Anal. **47** (2009), 1067–1091.
[7] E.M. Garau, P. Morin, and C. Zuppa, *Convergence of adaptive finite element methods for eigenvalue problems*, arXiv-Preprint **0803.0365v1** (2008).
[8] V. Mehrmann and A. Międlar, *Adaptive solution of elliptic PDE- eigenvalue problems. part I: Eigenvalues*, MATHEON-Preprint **565** (2009).
[9] K. Neymeyr, *A posteriori error estimation for elliptic eigenproblems*, Numer. Linear Algebra Appl. **9** (2002), 263–279.
[10] S. Sauter, *Finite elements for elliptic eigenvalue problems in the preasymptotic regime*, Preprint University of Zurich (2007).
[11] R. Stevenson, *Optimality of a standard adaptive finite element method*, Found. Comp. Math. **7** (2007), 245–269.

Adaptive solution of elliptic PDE-eigenvalue problems

Agnieszka Międlar
(joint work with Volker Mehrmann)

Many modern technological applications, e.g. vibration of structures, computation of acoustic fields or energy levels in quantum mechanics, involve the solution of the eigenvalue problems for partial differential equations (PDEs). It has been an important research topic in the last 30 years to design the adaptive finite element methods (AFEM) for PDE-eigenvalue problems with meshes that reduce the computational complexity, while retaining the overall accuracy.

In most AFEM approaches it is assumed that the resulting finite dimensional algebraic problem (linear system or eigenvalue problem) is solved exactly and the computational costs for this part of the method as well as the fact that this problem is solved in finite precision arithmetic are typically ignored. However, in particular, in the context of eigenvalue problems, often the costs for the solution of the algebraic eigenvalue problem dominate the total costs, and (in particular for nonsymmetric problems) the desired accuracy may not be achievable due to ill-conditioning.

We introduce a new adaptive finite element algorithm called AFEMLA which incorporates the adaptation in the algebraic approach. It allows us to avoid the usual assumption that the solution of the discrete problem is exact, to assure a good approximation via AFEM. Since the accuracy of the computed eigenvalue cannot be better than the quality of the discretization, there is no need to solve the algebraic eigenvalue problem up to very high precision if the discretization scheme guarantees only small precision. The goal of the adaptive method is to achieve a desired accuracy with minimal computational effort. In order to determine the

error estimates, we only solve the algebraic eigenvalue problem on the current coarse grid.

However, the algebraic adaptation alone cannot assure that the refinement procedure will lead to convergence. If the analytic approach with the standard assumption of solving the algebraic problems exactly does not converge, then neither does our extended approach. However, the AFEMLA approach makes the adaptation process much more efficient with guaranteed computable bounds in the algebraic part.

In summary, we have obtained fully computable bounds between the exact eigenvalues and computed eigenvalue approximations. Under the saturation assumption, for problems with well-conditioned eigenvalues, a small residual vector is equivalent to a good accuracy of the computed eigenvalues. Thus, if the saturation assumption holds and the problem is well-conditioned, then these bounds and the corresponding residuals can be used to control the adaptation process via the computed coarse mesh eigenvalues and eigenvectors.

Additionally we will introduce computable error bounds for the eigenfunctions and discuss possible extensions of our approach to the non-symmetric problems.

References

[1] V. Mehrmann and A. Miedlar Adaptive solution of elliptic PDE-eigenvalue problems. Part I: Eigenvalues, DFG Research Center Matheon Preprint 565, 2009.

Convergence of adaptive finite element methods for elliptic eigenvalue problems with applications to photonic crystals

Stefano Giani
(joint work with Ivan G. Graham)

In the last decades, mesh adaptivity has been widely used to improve the accuracy of numerical solutions to many scientific problems. The basic idea is to refine the mesh only where the error is high, with the aim of achieving an accurate solution using an optimal number of degrees of freedom. There is a large amount of numerical analysis literature on adaptivity, in particular on reliable and efficient a posteriori error estimates. Recently, the question of convergence of adaptive methods has received intensive interest and a number of convergence results for the adaptive solution of boundary value problems have appeared.

We proved the convergence of an adaptive linear finite element algorithm for computing eigenvalues and eigenvectors of scalar symmetric elliptic partial differential operators in bounded polygonal or polyhedral domains, subject to Dirichlet boundary data. Such problems arise in many applications, e.g. resonance problems and nuclear reactor criticality, to name but two.

Throughout, Ω will denote a bounded domain in $\mathbb{R}^d$ ($d = 2$ or 3). In fact Ω will be assumed to be a polygon ($d = 2$) or polyhedron ($d = 3$). We will be concerned

with the problem of finding an eigenvalue $\lambda \in \mathbb{R}$ and eigenfunction $0 \neq u \in H_0^1(\Omega)$ satisfying

$$a(u,v) := \lambda\, b(u,v)\ , \quad \text{for all} \quad v \in H_0^1(\Omega)\ , \tag{1}$$

where, for real valued functions u and v,

$$a(u,v) = \int_\Omega \nabla u(x)^T \mathcal{A}(x) \nabla v(x) dx \quad \text{and} \quad b(u,v) = \int_\Omega \mathcal{B}(x) u(x) v(x) dx\ . \tag{2}$$

Here, the matrix-valued function $\mathcal{A}$ is required to be symmetric and uniformly positive definite, i.e.

$$0 \ < \ \underline{a} \ \leq \ \xi^T \mathcal{A}(x) \xi \ \leq \ \overline{a} \quad \text{for all} \quad \xi \in \mathbb{R}^d \quad \text{with} \quad |\xi| = 1 \quad \text{and all} \quad x \in \Omega. \tag{3}$$

The scalar function $\mathcal{B}$ is required to be bounded above and below by positive constants for all $x \in \Omega$, i.e.

$$0 \ < \ \underline{b} \ \leq \ \mathcal{B}(x) \ \leq \ \overline{b} \quad \text{for all} \quad x \in \Omega. \tag{4}$$

We will assume that $\mathcal{A}$ and $\mathcal{B}$ are both piecewise constant on Ω and that any jumps in $\mathcal{A}$ and $\mathcal{B}$ are aligned with the meshes $\mathcal{T}_n$ (introduced below), for all n.

Throughout the report, for any polygonal (polyhedral) subdomain of $D \subset \Omega$, and any $s \in [0,1]$, $\|\cdot\|_{s,D}$ will denote the standard norm. We also define the energy norm induced by the bilinear form a and the weighted L^2 norm:

$$|||u|||_\Omega^2 := a(u,u) \quad \text{and} \quad \|u\|_{0,\mathcal{B},\Omega}^2 \ := \ b(u,u) \quad \text{for all } u \in H_0^1(\Omega)\ .$$

Rewriting the eigenvalue problem (1) in standard normalised form, we seek $(\lambda, u) \in \mathbb{R} \times H_0^1(\Omega)$ such that

$$\left.\begin{array}{lcl} a(u,v) & = & \lambda\, b(u,v), \quad \text{for all} \quad v \in H_0^1(\Omega) \\ \|u\|_{0,\mathcal{B},\Omega} & = & 1 \end{array}\right\} \tag{5}$$

To approximate problem (5) we use the piecewise linear finite element method. Accordingly, let $\mathcal{T}_n\ , n = 1, 2, \ldots$ denote a family of conforming triangular ($d = 2$) or tetrahedral ($d = 3$) meshes on Ω. Each mesh consists of elements denoted $\tau \in \mathcal{T}_n$. We assume that for each n, $\mathcal{T}_{n+1}$ is a refinement of $\mathcal{T}_n$. For a typical element τ of any mesh, its diameter is denoted H_τ and the diameter of its largest inscribed ball is denoted ρ_τ. For each n, let H_n denote the piecewise constant mesh function on Ω, whose value on each element $\tau \in \mathcal{T}_n$ is H_τ and let $H_n^{\max} = \max_{\tau \in \mathcal{T}_n} H_\tau$. Throughout we will assume that the family of meshes $\mathcal{T}_n$ is shape regular, i.e. there exists a constant $C_{\rm reg}$ such that: $H_\tau \leq C_{\rm reg} \rho_\tau$, for all $\tau \in \mathcal{T}_n$ and all $n = 1, 2, \ldots$.
The $\mathcal{T}_n$ meshes are produced by an adaptive process which ensures shape regularity.

We let V_n denote the usual finite dimensional subspace of $H_0^1(\Omega)$, consisting of all continuous piecewise linear functions with respect to the mesh $\mathcal{T}_n$. Then the discrete formulation of problem (5) is to seek the eigenpairs $(\lambda_n, u_n) \in \mathbb{R} \times V_n$ such that

$$\left.\begin{array}{lcl} a(u_n,v_n) & = & \lambda_n\, b(u_n,v_n), \quad \text{for all} \quad v_n \in V_n \\ \|u_n\|_{0,\mathcal{B},\Omega} & = & 1\ . \end{array}\right\} \tag{6}$$

Our refinement procedure is based on two locally defined quantities, firstly a standard a posteriori error estimator and secondly a measure of the variability (or "oscillation") of the computed eigenfunction. (Measures of "data oscillation" appear in the theory of adaptivity for boundary value problems, e.g. [9]. In the eigenvalue problem the computed eigenvalue and eigenfunction on the present mesh play the role of "data" for the next iteration of the adaptive procedure.) Our algorithm performs local refinement on all elements on which the minimum of these two local quantities is sufficiently large. The refinement procedure is based on three recursive applications of the newest node algorithm [8] to each marked triangle, first creating two sons, then four grandsons and finally bisecting two of the grandsons. This well-known algorithm stated without name in [9, §5.1], is called "bisection5" in [1] and is called "full refinement" in [10]. This technique creates a new node in the middle of each marked edge and also a new node in the interior of each marked element. It follows from [8] that this algorithm yields shape regular conforming meshes in 2D. In the 3D-case we use a suitable refinement that creates a new node on each marked face and a node in the interior of each marked element. We prove that the adaptive method converges, provided that the initial mesh is sufficiently fine. The latter condition, while absent for adaptive methods for linear symmetric elliptic boundary value problems, commonly appears for nonlinear problems and can be thought of as a manifestation of the nonlinearity of the eigenvalue problem.

Our error estimator is obtained by adapting standard estimates for source problems to the eigenvalue problem. Analogous eigenvalue estimates can be found in [3] (for the Laplace problem) and [11] (for linear elasticity). Recalling the mesh sequence $\mathcal{T}_n$ defined above, we let $\mathcal{S}_n$ denote the set of all the interior edges (or the set of interior faces in 3D) of the elements of the mesh $\mathcal{T}_n$. For each $S \in \mathcal{S}_n$, we denote by $\tau_1(S)$ and $\tau_2(S)$ the elements sharing S (i.e. $\tau_1(S) \cap \tau_2(S) = S$) and we write $\Omega(S) = \tau_1(S) \cup \tau_2(S)$. We let $\vec{n}_S$ denote the unit normal vector to S, orientated from $\tau_1(S)$ to $\tau_2(S)$. Furthermore we denote the diameter of S by H_S.

For a function g, which is piecewise continuous on the mesh $\mathcal{T}_n$, we introduce its jump across an interior edge (face) $S \in \mathcal{S}_n$ by:

$$[g]_S(x) := \left(\lim_{\substack{\tilde{x} \in \tau_1(S) \\ \tilde{x} \to x}} g(\tilde{x}) - \lim_{\substack{\tilde{x} \in \tau_2(S) \\ \tilde{x} \to x}} g(\tilde{x}) \right), \quad \text{for } x \in \text{int}(S) \ .$$

The error estimator η_n on the mesh $\mathcal{T}_n$ is defined as

$$\eta_n^2 := \sum_{S \in S_n} \eta_{S,n}^2 \ , \tag{7}$$

where, for each $S \in \mathcal{S}_n$,

$$\eta_{S,n}^2 := H_n^2 \|\nabla \cdot (\mathcal{A}\nabla u) + \lambda_n \mathcal{B} u_n\|_{0,\Omega(S)}^2 + H_S \|[\vec{n}_S \cdot \mathcal{A}\nabla u]_S\|_{0,S}^2 \ . \tag{8}$$

It is possible to prove, by adapting the usual arguments for linear source problems, that the error estimator η_n is reliable for both eigenvalues and eigenvectors and that it also efficient.

For mesh refinement based on the local contributions to η_n, we use the same marking strategy as in [2] and [9]. The idea is to refine a subset of the elements of $\mathcal{T}_n$ whose side residuals sum up to a fixed proportion of the total residual η_n.

Definition 1 (Marking Strategy 1)**.** *Given a parameter $0 < \theta < 1$, the procedure is: mark the sides in a minimal subset $\hat{\mathcal{S}}_n$ of $\mathcal{S}_n$ such that*

$$\left(\sum_{S \in \hat{\mathcal{S}}_n} \eta_{S,n}^2 \right)^{1/2} \geq \theta \eta_n \ . \tag{9}$$

To compute $\hat{\mathcal{S}}_n$, we compute all the "local residuals" $\eta_{S,n}$, then insert edges (faces) into $\hat{\mathcal{S}}_n$ in order of non-increasing magnitude of $\eta_{S,n}$, until (9) is satisfied. A minimal subset $\hat{\mathcal{S}}_n$ may not be unique. After this is done, we construct another set $\hat{\mathcal{T}}_n$, containing all the elements of $\mathcal{T}_n$ which contain at least one edge (face) belonging to $\hat{\mathcal{S}}_n$.

In order to prove our convergence theory, we require an additional marking strategy based on oscillations (Definition 2 below). This also appears in some theories of adaptivity for source problems (e.g. [2], [9], [7] and [1]), but to our knowledge has not yet been used in connection with eigenvalue problems. The concept of "oscillation" is just a measure of how well a function may be approximated by piecewise constants on a particular mesh. For any function $v \in L^2(\Omega)$, and any mesh $\mathcal{T}_n$, we introduce its orthogonal projection $P_n v$ onto piecewise constants defined by:

$$(P_n v)|_\tau = \frac{1}{|\tau|} \int_\tau v_n, \quad \text{for all} \quad \tau \in \mathcal{T}_n. \tag{10}$$

Then we define on a mesh $\mathcal{T}_n$:

$$\operatorname{osc}(v, \mathcal{T}_n) := \|H_n(v - P_n v)\|_{0,\mathcal{B},\Omega} \ . \tag{11}$$

Definition 2 (Marking Strategy 2)**.** *Given a parameter $0 < \tilde{\theta} < 1$: mark the elements in a minimal subset $\tilde{\mathcal{T}}_n$ of $\mathcal{T}_n$ such that*

$$\operatorname{osc}(u_n, \tilde{\mathcal{T}}_n) \ \geq \ \tilde{\theta} \ \operatorname{osc}(u_n, \mathcal{T}_n) \ . \tag{12}$$

Analogously to (9), we compute $\tilde{\mathcal{T}}_n$ by inserting elements τ into $\tilde{\mathcal{T}}_n$ according to non-increasing order of their local contributions $H_\tau^2 \|(u_n - P_n u_n)\|_{0,\mathcal{B},\tau}^2$ until (12) is satisfied.

The main result is Theorem 3 below which proves convergence of the adaptive method and also demonstrates the decay of oscillations of the sequence of approximate eigenfunctions under the assumption that the eigenvalue λ is simple. The proof of the convergence result can be found in [5].

Theorem 3. *Provided the initial mesh $\mathcal{T}_0$ is chosen so that $H_0^{\max}$ is small enough, there exists a constant $p \in (0,1)$ such that the recursive application of our adaptive*

algorithm yields a convergent sequence of approximate eigenvalues and eigenvectors, with the property:

$$|||u - u_n|||_\Omega \ \leq \ B_0 p^n, \qquad |\lambda - \lambda_n| \ \leq \ B_0^2 p^{2n}, \tag{13}$$

and

$$\lambda_n \ \mathrm{osc}(u_n, \mathcal{T}_n) \ \leq \ B_1 p^n, \tag{14}$$

where B_0 and B_1 are positive constants.

Finally, in [4] we extended the convergence result for our adaptive method to Hermitian elliptic partial differential operators with discontinuous coefficients in bounded polygonal, subject to periodic boundary conditions. We are particularly interested in applying our converging method to photonic crystals, which are used for optical communications, filters, lasers, switchers and optical transistors. More specifically we used our convergent adaptive finite element method to localize the spectral band gaps of crystals and also to compute accurately and efficiently the modes trapped in the defects of the crystals. These modes are very important in applications because they decay exponentially away from the defects, and so they can remain well confined within the crystal. Numerous numerical experiments on photonic crystals can be found in [6, 4].

References

[1] C. Carstensen and R. H. W. Hoppe. Error reduction and convergence for an adaptive mixed finite element method. *Math. Comp.*, 75(255):1033–1042, 2006.

[2] W. Dörfler. A convergent adaptive algorithm for Poisson's equation. *SIAM J. Numer. Anal.*, 33:1106–1124, 1996.

[3] R.G. Durán, C. Padra and R. Rodríguez, A posteriori estimates for the finite element approximation of eigenvalue problems, *Math. Mod. Meth. Appl. Sci.* 13(8):1219-1229, 2003.

[4] S. Giani. Convergence of Adaptive Finite Element Methods for Elliptic Eigenvalue Problems with Application to Photonic Crystals, PhD Thesis, University of Bath, 2008.

[5] S. Giani and I.G. Graham, A convergent adaptive method for elliptic eigenvalue problems, *SIAM J. Numer. Anal.* 47(8):1067-1091, 2009.

[6] S. Giani and I.G. Graham, A convergent adaptive method for elliptic eigenvalue problems and numerical experiments, Research Report 14/08, Bath Institute for Complex Systems, 2008. `http://www.bath.ac.uk/math-sci/BICS/`

[7] K. Mekchay and R. H. Nochetto, Convergence of adaptive finite element methods for general second order linear elliptic pdes. *SIAM J. Numer. Anal.*, 43:1803–1827, 2005.

[8] W. Mitchell, Optimal multilevel iterative methods for adaptive grids. *SIAM J. Sci. Stat. Comput.* 13:146-167, 1992.

[9] P. Morin, R. H. Nochetto and K. G. Siebert. Data oscillation and convergence of adaptive fem. *SIAM J. Numer. Anal.*, 38:466–488, 2000.

[10] R. Stevenson, Optimality of a standard adaptive finite element method, Found. Comput. Math. 7(2): 245-269, 2007.

[11] T.F. Walsh, G.M. Reese and U.L. Hetmaniuk, Explicit a posteriori error estimates for eigenvalue analysis of heterogeneous elastic structures *Comput. Methods Appl. Mech. Engrg.* 196: 3614-3623, 2007.

A robust hierarchical eigenvalue/eigenvector enhancement

Luka Grubišić
(joint work with Randolph Bank and Jeffrey S. Ovall)

We prove the reliability and efficiency of a variant hierarchical basis eigenvalue/ eigenvector approximation's estimator for a general symmetric and elliptic operator. The reliability estimates hold without any further assumptions save the well-posedness of the associated variational boundary value problem(s). The new estimator extends to general symmetric divergence type operators our earlier result on hierarchical error estimators from [2]. The extension is based on the recent work [5].

Further, we present a framework for incorporating the information contained in the hierarchical error estimator to enhance the convergence properties of the eigenvalue approximation. Also, a robust method for enhancing eigenvector approximations is presented. Let us restate that our theory of the convergence enhancement for both eigenvalues and eigenvectors is presented under the sole assumption that associated boundary value problem is well-posed. This is our definition of robust.

An alternative eigenvalue enhancement technique can be realized with gradient recovery approach as is reported in [4]. In comparison with [4] we offer a theory which is better amenable to treating eigenvalue multiplicity. Furthermore, our theory uses less regularity assumptions on the operator as well as on the triangular mesh (triangulation) and it treats eigenvector enhancement within the same framework.

To simplify the discussion, let us consider a simplest symmetric and elliptic divergence type operator. We consider the formal differential expression

$$\mathbb{A}u = \nabla \cdot (A \nabla u) \tag{1}$$

as the self-adjoint operator which is defined by the quadratic form $\mathfrak{a}(u,u) = \int_{\mathcal{R}} A\nabla u \cdot \nabla u$, $u \in H_0^1(\mathcal{R})$ in the sense of Kato. Here $\mathcal{R} \subset \mathbb{R}^2$ is a bounded polygonal domain, A is an appropriate $L^\infty(\mathcal{R})$ matrix valued function and $H_0^1(\mathcal{R})$ is the Sobolev space of functions with zero trace on the boundary $\partial\mathcal{R}$.

To compute discrete approximations of the eigenvalues of $\mathbb{A}$ we proceed in the standard way. Let $\mathcal{T}_d$ be a *conforming triangulation* of $\mathcal{R}$. For a given $\mathcal{T}_d$ we define the finite dimensional function spaces:

$$\mathcal{L}_d = \{u \in H_0^1(\mathcal{R}) \mid \text{ for } T \in \mathcal{T}_d,\, u|_T \text{ is a linear function}\}, \tag{2}$$

$$\mathfrak{Q}_d = \{u \in H_0^1(\mathcal{R}) \mid \text{ for } T \in \mathcal{T}_d,\, u|_T \text{ is a quadratic function}\}\ . \tag{3}$$

We will also make use of the space $\mathfrak{B}_d$ of edge bubble functions, which are those functions from $\mathfrak{Q}_d$ which vanish at the vertices of all triangles in $\mathcal{T}_d$.

A discrete variant of (1) now reads: Find nonzero $u \in \mathcal{L}_d$ such that

$$\begin{cases} \int_{\mathcal{R}} A\nabla u \cdot \nabla \psi &= \lambda \int_{\mathcal{R}} u\psi, \qquad \text{for all } \psi \in \mathcal{L}_d \\ \int_{\mathcal{R}} |u|^2 &= 1. \end{cases} \tag{4}$$

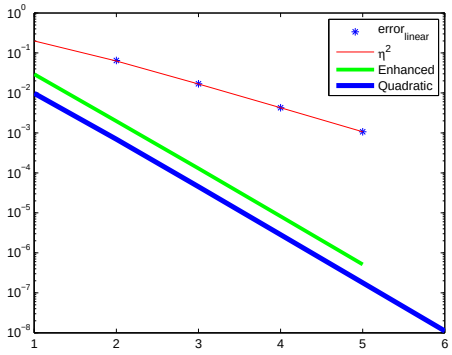

FIGURE 1. Six refinements—error for the linear approximation, error for the quadratic approximation, error for the enhanced approximation.

The problem (4) is attained by a finite number of discrete eigenvalues $\lambda_i(\mathcal{L}_d)$ and discrete eigenvectors $u_i(\mathcal{L}_d)$, $i = 1, \ldots, \dim \mathcal{L}_d$. The discrete eigenvalues $\lambda_i(\mathcal{L}_d)$ (discrete eigenvectors $u_i(\mathcal{L}_d)$) are often called the Ritz values/vectors.

Let us assume that we want to approximate the eigenvalue λ_q, $q \in \mathbb{N}$ of multiplicity $m \in \mathbb{N}$ of the self-adjoint elliptic operator $\mathbb{A}$ from (1). This is to say we assume that

$$\lambda_{q-1} < \lambda_q = \lambda_{q+1} = \cdots = \lambda_{q+m-1} < \lambda_{q+m},$$

holds for the standard numeration of lowermost eigenvalues of $\mathbb{A}$. We also assume that $q + m < \dim \mathfrak{L}_d$. By P_d we denote the orthogonal projection onto the linear span of

$$\{u_q(\mathcal{L}_d), \ldots, u_{q+m-1}(\mathcal{L}_d)\}.$$

For $\psi \in \mathsf{Ran}(P_d)$ we consider the functions $u_{\mathrm{lin}}(\psi, \mathcal{T}_d) \in \mathcal{L}_d$ and $\varepsilon(\psi, \mathcal{T}_d) \in \mathfrak{B}_d$ defined by:

$$\int_{\mathcal{R}} A\nabla u_{\mathrm{lin}}(\psi, \mathcal{T}_d) \cdot \nabla v = \int_{\mathcal{R}} \psi v \quad \text{for all } v \in \mathcal{L}_d$$
$$\int_{\mathcal{R}} A\nabla \varepsilon(\psi, \mathcal{T}_d) \cdot \nabla v = \int_{\mathcal{R}} \psi v - \nabla u_{\mathrm{lin}}(\psi, \mathcal{T}_d) \cdot \nabla v \quad \text{for all } v \in \mathfrak{B}_d\ .$$

As enhanced eigenvectors we propose

$$u_i^{\mathrm{en}}(\mathcal{T}_d) = u_i(\mathcal{L}_d) + \sqrt{\lambda_i(\mathcal{L}_d)}\, \varepsilon(u_i(\mathcal{L}_d), \mathcal{T}_d), \qquad u_i^{\mathrm{en}}(\mathcal{T}_d) \in \mathfrak{Q}_d.$$

Roughly speaking we establish that, assuming $\lambda_i^{\mathrm{en}}(\mathcal{T}_d) = \frac{\mathfrak{a}(u_i^{\mathrm{en}}(\mathcal{T}_d), u_i^{\mathrm{en}}(\mathcal{T}_d))}{\|u_i^{\mathrm{en}}(\mathcal{T}_d)\|^2}$ is the enhanced eigenvalue, we have

$$\frac{|\lambda_i - \lambda_i^{\mathrm{en}}(\mathcal{T}_d)|}{\lambda_i} \approx \frac{|\lambda_i - \lambda_i(\mathfrak{Q}_d)|}{\lambda_i}.$$

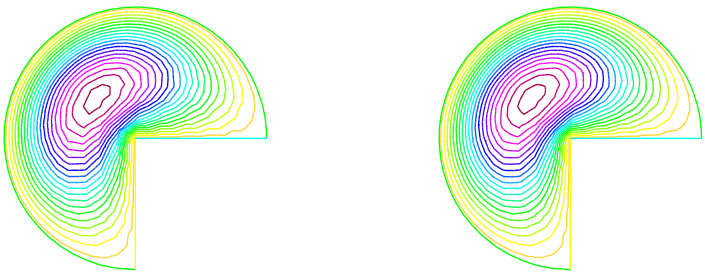

FIGURE 2. Enhanced and non-enhanced eigenvector on a rough mesh

Here $\lambda_i(\mathfrak{Q}_d)$ is the piecewise quadratic eigenvalue approximation obtained by substituting the test space $\mathfrak{Q}_d$ into the variational formulation (4) and $\|\cdot\|$ is the norm on $L^2(\mathcal{R})$.

We also present a formula which computes the enhanced eigenvalues $\lambda_i^{\text{en}}(\mathcal{T}_d)$ solely from $u_i^{\text{en}}(\mathcal{L}_d)$ and the hierarchical error estimator $\mathfrak{a}(\varepsilon(u_i(\mathcal{L}_d), \varepsilon(u_i(\mathcal{L}_d))$, see also [2, Theorem 3.1]. This formula is to be preferred for practical purposes, since a direct computation involving $\mathfrak{a}(u_i^{\text{en}}(\mathcal{T}_d), u_i^{\text{en}}(\mathcal{T}_d))$ might be to expensive to be competitive.

The results of preliminary FreeFEM tests can be seen on Figure 1. In short, we claim that for a little more than Linear–Linear eigenvalue solution $\lambda_i(\mathcal{T}_d), u_i(\mathcal{T}_d)$—by say a call to an ARPACK procedure—we obtain both an eigenvalue as well as eigenvector approximations which satisfy the same estimates as do the Quadratic–Quadratic eigenvalue/vector approximations $\lambda_i(\mathfrak{Q}_d), u_i(\mathfrak{Q}_d)$. For more on the assessment of the computational cost for computing $\varepsilon(u_i(\mathcal{L}_d))$ see [2, Section 5]. The theoretical results which we have just described have been tested in various experiments which include testing in MATLAB®, FreeFEM and PLTMG.

REFERENCES

[1] Randolph E. Bank. PLTMG 10.0. available from `http://ccom.ucsd.edu/~reb/software.html`.

[2] L. Grubišić and J. S. Ovall. On estimators for eigenvalue/eigenvector approximations. *Mathematics of Computation*, 78(266):739–770, 2009.

[3] Rich Lehoucq, Kristi Maschhoff, Danny Sorensen and Chao Yang ARPACK–Arnoldi package available from `http://www.caam.rice.edu/software/ARPACK/`.

[4] Ahmed Naga, Zhimin Zhang, and Aihui Zhou. Enhancing eigenvalue approximation by gradient recovery. *SIAM J. Sci. Comput*, 28(4):1289-1300, 2006.

[5] J. S. Ovall. An Efficient and Reliable Hierarchical Basis Error Estimate Requiring only H^1-regularity of the Exact Solution. *submitted*, preprint available from `http://www.acm.caltech.edu/~jovall/research.html`, 2009.

[6] J. S. Ovall. A Robust Error Estimator for Elliptic Eigenvalue Problems. *MAFELAP 09 talk*, available from `http://www.acm.caltech.edu/~jovall/research.html`, 2009.

[7] Olivier Pironneau, Frédéric Hecht, Jacques Morice. FreeFem++. available from `http://www.freefem.org/ff++/index.htm`, 2009.

The calculation of the distance to instability by the computation of a Jordan block

Melina Freitag
(joint work with Alastair Spence)

We introduce a new method for computing the distance of a stable matrix to the set of unstable matrices. Let A be a complex $n \times n$ matrix with all its eigenvalues in the open left half plane. In this case A is called a stable matrix. Stability is a very important property for physical and engineering applications (see, for example, [6] for a collection of examples). However, a small perturbation E to the matrix A may lead to eigenvalues of $A+E$ crossing the imaginary axis and hence the matrix $A+E$ being unstable. Papers which deal with the problem of finding the smallest perturbation E which makes $A+E$ unstable are those of van Loan [8] and Byers [1].

The smallest singular value of $A \in \mathbb{C}^{n\times n}$ satisfies

$$\sigma_{\min}(A) = \min\{\|E\|_F \,|\, \det(A+E) = 0,\, E \in \mathbb{C}^{n\times n}\}, \tag{1}$$

which leads to the minimising $E_{\min} = -\sigma v u^H$ where σ is the smallest singular value of A and v, u the corresponding right and left singular vectors. The distance of a matrix A to instability can be described as

$$\beta(A) = \min\{\|E\|_F \,|\, \eta(A+E) = 0,\, E \in \mathbb{C}^{n\times n}\},$$

where $\eta(A) = \max\{\mathrm{Re}(\lambda) \,|\, \lambda \in \Lambda(A)\}$. If $\eta(A)$ is negative, A is stable and if $A+E$ has an eigenvalue on the imaginary axis then E is a destabilising perturbation. In this case $(A+E-\omega iI)z = 0$ for some $\omega \in \mathbb{R}$ and $z \in \mathbb{C}^n$. Using (1) this leads to a measure of the distance to instability of a stable matrix A as defined in [8],

$$\beta(A) = \min_{\omega\in\mathbb{R}} \sigma_{\min}(A - \omega iI),$$

where $\sigma_{\min}(A-\omega iI)$ is the smallest singular value of $A - \omega iI$. For any $\omega \in \mathbb{R}$ upper and lower bounds on $\beta(A)$ (see [8]) are given by

$$\beta(A) \le \sigma_{min}(A - \omega iI). \quad \text{and} \quad \frac{1}{2}\mathrm{sep}(A) \le \beta(A),$$

where

$$\mathrm{sep}(A) = \min\{\|AY + YA^H\|_F,\, Y \in \mathbb{C}^{n\times n},\, \|Y\|_F = 1\}$$

is the separation of A and $-A^H$, A^H denotes the complex conjugate transpose of A and $\|\cdot\|_F$ the Frobenius norm. Clearly, $\mathrm{sep}(A) = 0$ if and only if A has an eigenvalue on the imaginary axis.

In [1] a bisection method for computing $\beta(A)$ was introduced. The method provides lower and upper bounds on $\beta(A)$ but requires the solution of a sequence of eigenvalue problems for the $2n \times 2n$ Hamiltonian matrix

$$H(\alpha) = \begin{bmatrix} A & -\alpha I \\ \alpha I & -A^H \end{bmatrix}$$

for a positive real α. In [1, Theorem 1] it has been shown that $H(\alpha)$ has a pure imaginary eigenvalue if and only if $\alpha \ge \beta(A)$. It is clear that the eigenvalues of

$H(0)$ are the union of the eigenvalues of A with the eigenvalues of $-A^H$, where the latter are the eigenvalues of A mirrored in the imaginary axis. Hence, in order to find the distance to instability one needs to find the minimum value of α such that $H(\alpha)$ has two identical imaginary eigenvalues. We show that in this case and for the critical value of α, $H(\alpha)$ has a Jordan block of dimension 2, and we shall exploit this fact in our algorithm.

He and Watson [3] built on the ideas in [1] and used a method based on inverse iteration for singular values in order to find a stationary point of $f(\omega) = \sigma_{min}(A - \omega iI)$ and then solved an eigenvalue problem for $H(\alpha)$ in order to check if this point is a global minimum.

We introduce a new algorithm to find the minimum value of α such that $H(\alpha)$ has a pure imaginary eigenvalue. Our method is based on the implicit determinant method of [7] but is extended to find the value α such that $H(\alpha) - \omega iI$ has a zero eigenvalue corresponding to a 2-dimensional Jordan block. Numerical experiments indicate that this method proves to be significantly faster than the methods discussed in [1] and [3].

In the talk we show that in the generic case $H(\alpha) - \omega iI$ has a 2-dimensional Jordan block at the required value of α. We extend the implicit determinant method of [7] to the case of a 2-dimensional Jordan block. Then we describe the Newton and Gauß-Newton methods to solve the nonlinear equations obtained from the implicit determinant method. We finally give numerical examples to illustrate the theory and which show that the method is competitive with and in almost all cases outperforms earlier algorithms.

References

[1] R. Byers, *A bisection method for measuring the distance of a stable matrix to the unstable matrices*, SIAM J. Sci. Statist. Comput., 9 (1988), pp. 875–881.

[2] A. Griewank and G. W. Reddien, *Characterization and computation of generalized turning points*, SIAM J. Numer. Anal., 21 (1984), pp. 176–185.

[3] C. He and G. A. Watson, *An algorithm for computing the distance to instability*, SIAM J. Matrix Anal. Appl., 20 (1998), pp. 101–116.

[4] H. B. Keller, *Numerical solution of bifurcation and nonlinear eigenvalue problems*, in Applications of Bifurcation Theory, P. H. Rabinowitz, ed., Academic Press, New York, 1977, pp. 359–384.

[5] D. Kressner, *Finding the distance to instability of a large sparse matrix*, in Computer Aided Control System Design, IEEE Int. Conf. on Contr. Appl., Munich, Germany, IEEE, 2006, pp. 31–36.

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

[7] A. Spence and C. Poulton, *Photonic band structure calculations using nonlinear eigenvalue techniques*, J. Comput. Phys., 204 (2005), pp. 65–81.

[8] C. F. Van Loan, *How near is a stable matrix to an unstable matrix?*, in Linear algebra and its role in systems theory (Brunswick, Maine, 1984), vol. 47 of Contemp. Math., Amer. Math. Soc., Providence, RI, 1985, pp. 465–478.

Nonlinear eigenvalue problems in Density Functional Theory calculations

Jean-Luc Fattebert

Developed in the 1960's by W. Kohn and coauthors[1, 2], Density Functional Theory (DFT) is a very popular quantum model for First-Principles simulations in chemistry and material sciences. It allows calculations of systems made of hundreds of atoms. Indeed DFT reduces the 3N-dimensional Schroedinger electronic structure problem to the search for a ground state electronic density in 3D. In practice it leads to the search for N electronic wave functions solutions of an energy minimization problem in 3D, or equivalently the solution of an eigenvalue problem with a non-linear operator.

We consider the Density Functional Theory energy functional written as a functional of N orthonormal electronic wave functions ψ_i (Kohn-Sham formulation)

$$\begin{aligned} (1)\ E_{KS}[\{\psi_i\}_{i=1}^N] &= \sum_{i=1}^N \int_\Omega \psi_i(x)\left(-\triangle\psi_i\right)(x)dx + \int_\Omega\int_\Omega \frac{\rho(x_1)\rho(x_2)}{|x_1-x_2|}dx_1dx_2 \\ &+ E_{XC}[\rho] + \sum_{i=1}^N \int_\Omega \psi_i(x)(V_{ext}\psi_i)(x)dx \end{aligned}$$

where ρ is the electronic density defined by

$$\rho(x) = \sum_{i=1}^N |\psi_i(x)|^2 \tag{2}$$

(see for example [3]). E_{KS} is made of the sum of the kinetic energy of the electrons, the Coulomb interaction between electrons, the exchange and correlation electronic energy, and the energy of interaction of the electrons with the potential generated by all the atomic cores V_{ext}. Given an external potential V_{ext} — defined by the various atomic species present in the problem, their respective positions and pseudopotentials — the ground state of the physical system is obtained by minimizing the energy functional (1) under the orthonormality constraints

$$\int_\Omega \psi_i(x)\psi_j(x) = \delta_{ij}, i, j = 1, \ldots, N. \tag{3}$$

To avoid mathematical difficulties irrelevant in the discussion of the numerical solver, let us assume that we have a problem in a finite dimensional space of dimension M resulting from the discretization of the above equations. To be concrete, suppose that we have a finite difference discretization on a uniform mesh with periodic boundary conditions, and thus the functions ψ_i are M-dimensional vectors with components corresponding to their values at the mesh points, $\psi_{i,k} = \psi_i(x_k)$. Let L_h be a finite difference approximation of the Laplacian operator. Without restriction of generality, wave functions are assumed to take real values only.

One can derive the Euler-Lagrange equations associated to the minimization problem (1) with N^2 Lagrange parameters corresponding to the orthonormality

constraints (3). One obtains the so-called Kohn-Sham (KS) equations in their usual form for the particular choice of the functions $\{\psi_i\}_{i=1}^N$ which diagonalizes the matrix of the Lagrange parameters,

$$\left\{ \begin{array}{ccc} -L_h\psi_i + V_{KS}[\rho]\psi_i & = & \lambda_i\psi_i \\ \rho(x_k) & = & \sum_{i=1}^N |\psi_i(x_k)|^2 \\ \sum_{k=1}^M \psi_i(x_k)\psi_j(x_k) & = & \delta_{ij} \end{array} \right.$$

where V_{KS} is a discretized nonlinear effective potential operator (see *e.g.* [3]). In this approach, one has to find the N lowest eigenvalues $\lambda_i, i = 1, \ldots, N$ and the corresponding eigenfunctions. We assume that $\lambda_{N+1} - \lambda_N > 0$.

We can represent the solution of the discretized problem as an M by N matrix

$$\Psi = (\psi_1, \ldots, \psi_N)\,.$$

Ψ represents the invariant subspace spanned by the eigenvectors associated with the N lowest eigenvalues. Other representations of that same subspace by N linearly independent vectors can be written as

$$\Phi = (\varphi_1, \ldots, \varphi_N)\,.$$

One can find an $N \times N$ matrix C such that $\Psi = \Phi C$. C satisfies $CC^T = S^{-1}$ where $S = \Phi^T\Phi$, the Gram matrix, is of rank N. Using the previous relations, the electronic density can be written in terms of the matrix elements of Φ and S^{-1},

$$\rho_k = \sum_{i,j=1}^{N} \left(S^{-1}\right)_{ij} \Phi_{ki}\Phi_{kj}$$

where ρ_k denotes the value of the electronic density at the mesh point x_k. Also the KS equations for Φ can be rewritten as

$$-L_h\Phi + V_{KS}[\rho]\Phi = \Phi S^{-1} H_\Phi \tag{4}$$

where $H_\Phi = \Phi^T(-L_h + V_{KS})\Phi$. In this formulation, unlike more traditional approaches which include an additional equation to enforces orthonormality, the columns of Φ constitute a general nonorthogonal basis of the trial invariant subspace.

In first-principles molecular dynamics simulations, the equations of Density Functional Theory need to be solved at every step to determine the forces acting on individual atoms and responsible for the dynamics. To obtain meaningful statistics, $O(10^5)$ steps are required. Thus an efficient solver is needed for the KS equations.

For finite differences approaches — or pseudo-spectral methods also very popular in the field —, about 1% of the eigenpairs of large sparse KS Hamiltonian operators are needed. In this context, the nonlinear problem — Eq.(4) — can be efficiently solved using a subspace preconditioned inverse iteration [4]. A good preconditioner can be obtained based on the observation that the Hamiltonian operator is a Laplacian plus a perturbation (potential) [5]. Close to convergence,

subspace preconditioned inverse iterations can be accelerated using a simple extrapolation scheme for Φ, using the approximations obtained at the previous m steps of the iterative solver,

$$\bar{\Phi}_k := \Phi_k + \sum_{j=1}^{m} \theta_j (\Phi_{k-j} - \Phi_k).$$

The columns of Φ are assumed to be normalized at each step. The coefficients $\theta_j, j = 1, ..., m$ are obtained from a residual minimization condition. Small values of m (1 or 2) work well in practice. Numerical results demonstrate excellent convergence rate for large scale applications involving over a 1000 eigenpairs. Such an algorithm is compatible with recently proposed O(N) complexity approaches where the searched subspace is represented by a set of functions strictly localized in real-space [6].

Financial support for this research was provided by the Office of Science, U.S. Department of Energy, SciDAC Grant DE-FC02-06ER46262. This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

References

[1] P. Hohenberg and W. Kohn, *Inhomogeneous electron gas*, Phys. Rev. B, **136** (1964), 864–871.

[2] W. Kohn, and L. J. Sham, *Self-consistent equations including exchange and correlation effects*, Phys. Rev. A **140** (1965), 1133–1138.

[3] E. Cances, M. Defranceschi, W. Kutzelnigg, C. LeBris, and Y. Maday, *Computational chemistry: a primer*, Handbook of Numerical Analysis (P.G. Ciarlet and C. Le Bris, eds.), vol. X, Elsevier Science, 2003, 3–270.

[4] Klaus Neymeyr, *A geometric theory for preconditioned inverse iteration applied to a subspace*, Math. Comp. **71** (2002), no. 237, 197–216.

[5] J.-L. Fattebert and M. Buongiorno Nardelli, *Finite difference methods for ab initio electronic structure and quantum transport calculations of nanostructures*, Handbook of Numerical Analysis (P.G. Ciarlet and C. Le Bris, eds.), vol. X, Elsevier Science, 2003, 571.

[6] J.-L. Fattebert and F. Gygi, *Linear scaling first-principles molecular dynamics with controlled accuracy*, Comput. Phys. Commun. **162** (2004), 24–36.

Local Finite Element Discretizations for Quantum Eigenvalue Problems

Aihui Zhou

(joint work with Xiaoying Dai, Xingao Gong, Lihua Shen, Dier Zhang)

It is known that for quantum many-particle systems, there is a general principle of locality or "nearsightedness" that the properties at one point may be considered independent of what happens at distant points [5, 7, 9]. It is also shown that the wavefunction of a quantum many-particle system is somehow smooth and oscillates in the region of the system only (c.f., e.g., [1, 6, 13]). In modern electronic structure computations, the so-called density functional theory is fundamental, with which Kohn-Sham equations need to be solved [2, 8, 11]. Thus an efficient discretization

scheme for solving the Kohn-Sham equation should be good at handling the locality of the system.

The central computation in solving Kohn-Sham equations is the repeated solution of the following model eigenvalue problem, posed on a convex polygonal domain $\Omega \subset R^3$:

$$\left\{ \begin{array}{rcl} -\Delta u + Vu & = & \lambda u \text{ in } \Omega, \\ u & = & 0 \text{ on } \partial\Omega, \end{array} \right. \tag{1}$$

where V is some effect potential and smooth in the pseudopotential setting.

As we know, a significant strength of the finite element method lies in the ability to place adaptive/local refinements in regions where the desired functions vary rapidly while treating the distant zones with a coarser description. Based on these observations and our understanding of local behaviors of finite element solutions to some elliptic partial differential problems, we have proposed some local finite element discretizations for a class of quantum eigenvalue problems. The main idea of our approach is to use a lower order finite element to approximate the low frequency of the solution and then to use some linear algebraic systems to correct the residual (which contains mostly high frequencies) in the higher order finite element space by some local and parallel procedure.

Let us use (1) to give a description of the main idea. Let $S_0^{h,1}(\Omega)$ and $S_0^{h,2}(\Omega)$, satisfying $S_0^{h,1}(\Omega) \subset S_0^{h,2}(\Omega) \subset H_0^1(\Omega)$, be the linear finite element space and the quadratic finite element space associated with a finite element grid $T^h(\Omega)$, respectively. We may employ the following algorithm to discretize (1) to obtain eigenpair approximations:

(1) Solve an eigenvalue problem in the linear finite element space: Find $\lambda_{h,1} \in R^1, u_{h,1} \in S_0^{h,1}(\Omega)$ such that $\|u_{h,1}\|_{L^2(\Omega)} = 1$ and

$$\int_\Omega (\nabla u_{h,1} \cdot \nabla v + V u_{h,1} v) = \lambda_{h,1} \int_\Omega u_{h,1} v \quad \forall v \in S_0^{h,1}(\Omega).$$

(2) Solve a linear boundary value problem in the quadratic finite element space locally: Find $e_{h,2} \in S_0^{h,2}(\Omega_0)$ such that

$$\begin{array}{rl} & \displaystyle\int_{\Omega_0} (\nabla e_{h,2} \cdot \nabla v + V e_{h,2} v) \\ = & \displaystyle \lambda_{h,1} \int_{\Omega_0} u_{h,1} v - \int_{\Omega_0} (\nabla u_{h,1} \cdot \nabla v + V u_{h,1} v) \quad \forall v \in S_0^{h,2}(\Omega_0). \end{array}$$

(3) Set:

$$u^h = \left\{ \begin{array}{l} u_{h,1} + e_{h,2} \text{ on } \bar{\Omega}_0, \\ u_{h,1} \text{ in } \Omega \setminus \bar{\Omega}_0 \end{array} \right.$$

and

$$\lambda^h = \frac{a(u^h, u^h)}{\| u^h \|_{0,\Omega}^2}.$$

If, for example, $\lambda_{h,1}$ is the first eigenvalue of the problem at the first step, then under some reasonable assumption we can establish the following result [3]

$$\|\nabla(u - u^{h,2})\|_{1,D} = O(h^2), \tag{2}$$

where u is an eigenfunction associated with the first eigenvalue λ of (1) that satisfies $\|u\|_{L^2(\Omega)} = 1$ and $D \subset\subset \Omega_0$ that means $D \subset \Omega_0$ and dist $(\partial D \setminus \partial\Omega, \partial\Omega_0 \setminus \partial\Omega) > 0$. Our numerical examples show that λ^h is much accurate than $\lambda_{h,1}$.

Consequently, we may also design a local computational scheme as follows: first, solve the eigenvalue problem in the linear finite element space over a finite element mesh $T^h(\Omega)$, then solve some linear boundary value problems on a collection of overlapped subdomains in the quadratic finite element space associated with the same mesh $T^h(\Omega)$ locally, and finally compute some Rayleigh quotient to obtain a new eigenvalue approximation $\lambda^{h,2}$ which asymptotically has the same accuracy as that of the quadratic finite element eigenvalue approximation. We have applied these new schemes to several typical electronic structure computations successfully.

References

[1] S. Agmon, *Lectures on the Exponential Decay of Solutions of Second-Order Elliptic Operators*, Princeton University Press, Princeton, NJ, 1981.

[2] T. L. Beck, *Real-space mesh techniques in density-function theory*, Rev. Modern Phys. **72** (2000), 1041-1080.

[3] X. Dai, L. Shen, and A.Zhou, *A local computational scheme for higher order finite element eigenvalue approximations*, Internat. J Numer. Anal. Model. **5** (2008), 570–589.

[4] X. Dai and A. Zhou, *Three-scale finite element discretizations for quantum eigenvalue problems*, SIAM J. Numer. Anal. **46** (2008), 295–324.

[5] S. Goedecker, *Linear scaling methods for the solution of Schrödinger's equation*, In: Handbook of Numerical Analysis, Vol. X: Computational Chemistry, edited by C. Le Bris, Elsevier, Amsterdam, 537-570, 2003.

[6] X. Gong, L. Shen, D. Zhang, and A. Zhou, *Finite element approximations for Schrödinger equations with applications to electronic structure computations*, J Comput. Math. **26** (2008), 310–323.

[7] W. Kohn, *Density functional and density matrix method scaling linearly with the number of atoms*, Phys. Rev. Lett. **76** (1996), 3168-3171.

[8] C. Le Bris (ed.), *Handbook of Numerical Analysis, Vol. X. Special issue: Computational Chemistry*, North-Holland, Amsterdam, 2003.

[9] R. M. Martin, *Electronic Structure: Basic Theory and Practical Method*, Cambridge University Press, Cambridge, 2004.

[10] J. E. Pask and P. A. Sterne, *Finite element methods in ab initio electronic structure calculations*, Model. Simul. Mater. Sci. Eng. **13** (2005), R71-R96.

[11] R. G. Parr and W. T. Yang, *Density-Functional Theory of Atoms and Molecules*, Oxford University Press, New York, Clarendon Press, Oxford, 1994.

[12] T. Torsti, T. Eirola, J. Enkovaara, T. Hakala, P. Havu, V. Havu, T. Hoynalanmaa, J. Ignatius, M. Lyly, I. Makkonen, T. T. Rantala, J. Ruokolainen, K. Ruotsalainen, E. Rasanen, H. Saarikoski, and M. J. Puska, *Three real-space discretization techniques in electronic structure calculations*, Phys. Status Solidi B **243** (2006), 1016-1053.

[13] H. Yserentant, *On the regularity of the electronic Schrödinger equation in Hilbert space of mixed derivatives,* Numer. Math. **98** (2004), 731-759.

Decay estimates for spectral projectors with applications to electronic structure

Michele Benzi
(joint work with Paola Boito and Nader Razouk)

In quantum chemistry and solid state physics, one is interested in determining the *electronic structure* of (possibly large) atomic and molecular systems. The problem amounts to computing the ground state (smallest eigenvalue and corresponding eigenfunction) of the many-body quantum-mechanical Hamiltonian (Schrödinger operator), $\mathcal{H}$. Variationally, we want to minimize the Rayleigh quotient:

$$E_0 = \min_{\Psi \neq 0} \frac{\langle \mathcal{H}\Psi, \Psi \rangle}{\langle \Psi, \Psi \rangle} \quad \text{and} \quad \Psi_0 = \text{argmin}_{\Psi \neq 0} \frac{\langle \mathcal{H}\Psi, \Psi \rangle}{\langle \Psi, \Psi \rangle}$$

where $\langle \cdot, \cdot \rangle$ denotes the L^2 inner product. In the Born–Oppenheimer approximation, the many-body Hamiltonian is given (in atomic units) by

$$\mathcal{H} = \sum_{i=1}^{N_e} \left(-\frac{1}{2}\Delta_i - \sum_{j=1}^{M} \frac{Z_j}{|\mathbf{x}_i - \mathbf{r}_j|} + \sum_{j \neq i}^{N_e} \frac{1}{|\mathbf{x}_i - \mathbf{x}_j|} \right)$$

where N_e = number of electrons and M = number of nuclei in the system. The electron positions are denoted by $\mathbf{x}_i$, those of the nuclei by $\mathbf{r}_j$; as usual, the charges are denoted by Z_j. The operator $\mathcal{H}$ acts on a suitable subspace of $H^1(\mathbb{R}^{3N_e})$ consisting of anti-symmetric functions (as a consequence of Pauli's Exclusion Principle for Fermions). Here, the spin is neglected in order to simplify the presentation.

Unless N_e is very small, the "curse of dimensionality" makes this problem almost intractable. In order to make the problem more tractable, various approximations have been devised, most notably:

- Wavefunction methods (e.g., Hartree-Fock)
- Density Functional Theory (e.g., Kohn-Sham)
- Hybrid methods (e.g., B3LYP)

In these approximations the original, linear eigenproblem $\mathcal{H}\Psi = E\Psi$ for the many-electrons Hamiltonian is replaced by a non-linear one-particle eigenproblem:

$$\mathcal{F}\psi_i = \lambda_i \psi_i, \quad \langle \psi_i, \psi_j \rangle = \delta_{ij}, \quad 1 \leq i, j \leq N_e$$

where $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_{N_e}$. Here $\mathcal{F}\psi_i = \left(-\frac{1}{2}\Delta + V(\rho)\right)\psi_i$, with $\rho = \sum_{i=1}^{N_e} |\psi_i(\mathbf{x})|^2$. The nonlinear problem can be solved by a 'self-consistent field' (SCF) iteration, leading to a sequence of *linear* eigenproblems

$$\mathcal{F}^{(k)}\psi_i^{(k)} = \lambda_i^{(k)}\psi_i^{(k)}, \quad \langle \psi_i^{(k)}, \psi_j^{(k)} \rangle = \delta_{ij}, \quad k = 1, 2, \ldots$$

$(1 \leq i, j \leq N_e)$, where each $\mathcal{F}^{(k)} = -\frac{1}{2}\Delta + V^{(k)}$ is a one-electron (Fock, or *effective*) Hamiltonian:

$$V^{(k)} = V^{(k)}(\rho^{(k-1)}), \quad \rho^{(k-1)} = \sum_{i=1}^{N_e} |\psi_i^{(k-1)}(\mathbf{x})|^2.$$

Solution of each of the (discretized) linear eigenproblems above leads to a typical $O(N_e^3)$ cost per SCF iteration. However, the actual eigenpairs $(\psi_i^{(k)}, \lambda_i^{(k)})$ are unnecessary; hence, diagonalization of the (discretized) one-particle Hamiltonians can be avoided. Indeed, all one really needs is the orthogonal projector P onto the invariant subspace

$$V_{occ} = \text{span}\{\psi_1, \ldots, \psi_{N_e}\}$$

corresponding to the N_e lowest eigenvalues $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_{N_e}$ ('occupied states'). At the kth SCF cycle, an approximation to the orthogonal projector $P^{(k)}$ onto the occupied subspace $V_{occ}^{(k)}$ needs to be computed. All quantities of interest in electronic structure theory can be computed from P. For example, the expected value for the energy of the system described by the Hamiltonian H is given by $\langle E \rangle = \text{Trace}(PH)$, and similarly for other observables. The forces can also be computed once P is known. P is called the *density operator* of the system (von Neumann, 1927). At equilibrium, $[H, P] = 0$ and P is a function of the Hamiltonian: $P = f(H)$. In practice, operators are replaced by matrices by Galerkin projection onto a finite-dimensional subspace spanned by a set of basis functions $\{\varphi_i\}_{i=1}^N$, where N is a multiple of N_e. Typically, $N = N_b \cdot N_e$ where $N_b \geq 2$ is a moderate constant when linear combinations of Gaussian-type orbitals are used; often, $N_b \approx 10$.

For simplicity of exposition, here we use orthonormal bases. We also assume that the basis functions are localized, so that the resulting discrete Hamiltonians are, up to some small truncation tolerance, sparse: their pattern/bandwidth is determined by the range of the interactions. Non-orthogonal bases are easily accommodated into the theory by a congruence transformation to an orthogonal basis, which can be accomplished via an inverse-Cholesky factorization; the transformed Hamiltonian is $\hat{H} = Z^T H Z$ where $S^{-1} = ZZ^T$ is the inverse factorization of the overlap matrix. Since S is localized and well-conditioned independent of N, its inverse (and therefore its Cholesky factor Z) decays exponentially with a rate independent of N [4, 3]. Hence, up to a small truncation tolerance Z is sparse and so is $\hat{H}$. From here on, we assume that the transformation has already been performed and we denote the representation of the discrete Hamiltonian in the orthogonal basis by H instead of $\hat{H}$.

Thus, the fundamental problem of (zero-temperature) electronic structure theory is the approximation of the spectral projector P onto the subspace spanned by the N_e lowest eigenfunctions of H (occupied states):

$$P = \psi_1 \otimes \psi_1 + \cdots + \psi_{N_e} \otimes \psi_{N_e}$$

where $H\psi_i = \lambda_i \psi_i,\ i = 1, \ldots, N_e$. Clearly we can write $P = f(H)$ where f is the step function

$$f(x) = \begin{cases} 1 & \text{if} \quad x < \mu \\ 0 & \text{if} \quad x > \mu \end{cases}$$

with $\lambda_{N_e} < \mu < \lambda_{N_e+1}$ (μ is the "Fermi level"). Alternatively, we can write $P = (I - \text{sign}(H - \mu I))/2$. If the spectral ('HOMO-LUMO') gap $\Gamma = \lambda_{N_e+1} - \lambda_{N_e}$

is not too small, f can be well approximated by the Fermi–Dirac function

$$f(x) = \frac{1}{1 + \mathrm{e}^{\beta(x-\mu)}},$$

which tends to a step function as the parameter β increases.

Physicists have observed long ago that for 'gapped systems' (like insulators and, to some extent, semiconductors) the entries of the density matrix P decay exponentially fast away from the main diagonal, reflecting the fact that interaction strengths decrease rapidly with the distance. (For metallic systems, the decay is only algebraic.) Such decay property is crucial as it provides the basis for so-called *linear scaling* (i.e., $O(N)$) methods for electronic structure calculations; indeed, rapidly decaying matrices can be approximated by banded (or sparse) matrices, uniformly in N. In this contribution we apply general results on functions of sparse matrices from [1, 2] to obtain a rigorous and rather elementary proof of this property for spectral projectors corresponding to localized Hamiltonians with nonvanishing spectral gap. For simplicity, we state our main result for the banded case only; extension to general sparsity patterns is fairly straightforward (see [5]).

Theorem. *Let N_b be a fixed positive integer and $N = N_b \cdot N_e$ where the integers N_e form a monotonically increasing sequence. Let $\{H_N\}$ be a sequence of Hermitian $N \times N$ matrices with the following properties:*

(1) *Each H_N has bandwidth m independent of N;*
(2) *There exist two (fixed) intervals $I_1 = [a, b]$, $I_2 = [c, d] \subset \mathbf{R}$ with $\gamma = c - b > 0$ such that for all $N = N_b \cdot N_e$, I_1 contains the smallest N_e eigenvalues of H_N (counted with their multiplicities) and I_2 contains the remaining $N - N_e$ eigenvalues.*

Let P_N denote the $N \times N$ spectral projector onto the subspace spanned by the eigenvectors associated with the N_e smallest eigenvalues of H_N, for each N. Then for any $\varepsilon > 0$ there is a matrix $P_N^{(\hat{m})}$ of bandwidth $\hat{m}$ independent of N such that $\|P_N - P_N^{(\hat{m})}\| < \varepsilon$, for all N.

In addition, we show how to estimate the bandwidth $\hat{m}$ in the approximant $P_N^{(\hat{m})}$ in terms of the relative spectral gap γ and bandwidth m (more generally, sparsity pattern) of the discrete Hamiltonians; as expected, $\hat{m}$ increases as $\gamma \to 0$.

References

[1] M. Benzi and G. H. Golub, *Bounds for the entries of matrix functions with applications to preconditioning*, BIT **39** (1999), 417–438.

[2] M. Benzi and N. Razouk, *Decay rates and $O(n)$ algorithms for approximating functions of sparse matrices*, Electronic Transactions on Numerical Analysis, **28** (2007), 16–39.

[3] M. Benzi and M. Tůma, *Orderings for factorized sparse approximate inverse preconditioners*, SIAM J. Scientific Computing, **21** (2000), 1851–1868.

[4] S. Demko, W. F. Moss, and P. W. Smith, *Decay rates for inverses of band matrices*, Mathematics of Computation, **43** (1984), pp. 491–499.

[5] N. Razouk, *Localization Phenomena in Matrix Functions: Theory and Algorithms*, PhD Thesis, Emory University, Atlanta, GA, 2008.

Solving a 12 dimensional Schroedinger equation to compute energy levels of CH_5^+

Tucker Carrington
(joint work with Xiaogang Wang)

It is difficult to compute vibrational spectra of molecules because of the dimensionality of the associated Schroedinger equation. For CH_5^+ the equation is 12-dimensional. Basis functions are used to discretize the Schroedinger equation to obtain a matrix eigenvalue problem. To compute numerically exact solutions for CH_5^+ it is essential to choose basis functions that have significant amplitude in all the 120 wells among which the protons move. [1] This, in conjunction with the fact that there are 6 atoms, means that the required basis is huge. It is impossible to use a basis of products of functions of a single coordinate. [2] Even if only 10 functions per coordinate were sufficient one would need 8 terabytes of memory for a single vector.

Some sort of contraction scheme is therefore imperative. We make contracted functions from eigenvectors of reduced-dimension Hamiltonians extracted from the full Hamiltonian operator by setting some coordinates to reference values. [3] In several papers we have shown that it is possible to efficiently do matrix-vector products using a basis of simply contracted bend and stretch functions. [3, 4, 5, 6] It is always possible to write the full Hamiltonian

$$H = H^{(b)} + H^{(s)} + \Delta T + \Delta V \tag{1}$$

where

$$\Delta V(r,\theta) = V(r,\theta) - V_{bref}(\theta) - V_{sref}(r) \tag{2}$$

and

$$\Delta T = \sum_i \Delta B_i(r) T_b^{(i)}(\theta) \ . \tag{3}$$

with

$$\Delta B_i(r) = B_i(r) - B_i(r_e) \ . \tag{4}$$

$H^{(b)}$ is the bend Hamiltonian obtained from the full Hamiltonian by discarding the stretch kinetic energy term, choosing an appropriate reference bend potential $V_{bref}(\theta)$, and evaluating the B in the bend kinetic energy operator (KEO), Eq. (5), at a reference value for the stretch coordinates (denoted r_e). The bend KEO is,

$$T_{ben} = \sum_i B_i(r) T_b^{(i)}(\theta) \ , \tag{5}$$

where $T_b^{(i)}(\theta)$ are differential operators. Note that θ represents all of the bend coordinates and r represents all of the stretch coordinates. $H^{(s)}$ is the stretch Hamiltonian obtained from the full Hamiltonian by discarding the bend kinetic energy term and choosing an appropriate reference stretch potential $V_{sref}(r)$. The product contracted (PC) basis functions we use are products of eigenfunctions of

$H^{(b)}$ and $H^{(s)}$. For both the stretch and the bend we retain eigenfunctions with eigenvalues below some threshold.

Matrix elements of $H^{(b)}$ and $H^{(s)}$ in the PC basis are trivial and matrix-vector products with these terms are simple. PC matrix-vector products with ΔT and ΔV can be done without storing vectors in the primitive bases used to solve the bend and stretch problems. Matrix-vector products for ΔT are done by taking advantage of the fact that it is diagonal in any stretch discrete variable representation (DVR) basis. [7] The product structure of Eq. (3) can be easily exploited. The most difficult part of the ΔT matrix-vector product is the computation of $T_b^{(i)}$ matrix elements from eigenvectors of the bend Hamiltonian.

Matrix-vector products for ΔV are more difficult. They are done using the **F** matrix idea of Ref [3] The matrix **F** is defined by,

$$F_{b'b,\alpha} = \sum_{\beta} \tilde{C}_{\beta b'} \tilde{C}_{\beta b} \Delta V_{\beta\alpha} \tag{6}$$

with

$$\tilde{C}_{\beta b} \equiv \sum_{l} T_{l\beta} C_{lb} \ . \tag{7}$$

where the eigenfunctions of $H^{(b)}$ are

$$X_b(\theta) = \sum_{l} C_{lb} f_l(\theta) \ . \tag{8}$$

Bend quadrature points are labelled by β, stretch quadrature points are labelled by α, $V_{\beta\alpha}$ is a potential value on the full grid, The $f_l(\theta)$ are primitive bend functions used to solve the bend problem. $T_{l\beta}$ is a matrix used to compute quadratures in the primitive bend basis. $\tilde{C}_{\beta b}$ is a grid representation of the bend eigenfunctions. Note that l, β, and α are composite indices, i.e., $l \equiv \{l_1, l_2, m_2, l_3, m_3, l_4, m_4\}$, $\beta \equiv \{\gamma_1, \gamma_2, \beta_2, \gamma_3, \beta_3, \gamma_4, \beta_4\}$, where γ_k and β_k label θ_k and φ_k quadrature points, and α labels a point on the 5d stretch grid. Eigenvalues and eigenfunctions of $H^{(b)}$ and $H^{(b)}$ are computed using the Lanczos algorithm with no re-orthogonalization.[8, 9]

References

[1] XIAO-GANG WANG and TUCKER CARRINGTON JR., *Vibrational energy levels of* CH_5^+ J. Chem. Phys. **129**, 234102-1 – 234102-13 (2008)

[2] M.J. BRAMLEY and T. CARRINGTON JR., *A general discrete variable method to calculate vibrational energy levels of three- and four-atom molecules*, J. Chem. Phys. **99**, 8519-8541 (1993).

[3] XIAO-GANG WANG and TUCKER CARRINGTON JR., *New ideas for using contracted basis functions with a Lanczos eigensolver for computing vibrational spectra of molecules with four or more atoms* J. Chem. Phys. **117**, 6923-6934 (2002).

[4] JEAN CHRISTOPHE TREMBLAY and TUCKER CARRINGTON JR. *Calculating vibrational energies and wavefunctions of vinylidene using a contracted basis with a locally re-orthogonalized coupled two-term Lanczos eigensolver* J. Chem. Phys. **125**, 094311-1–094311-12 (2006)

[5] TUCKER CARRINGTON JR., *Methods for calculating vibrational energy levels* Encyclopedia of Computational Chemistry, editor-in-chief, Paul von Ragu Schleyer, 2nd edition, (John Wiley and Sons) (2003).
[6] XIAO-GANG WANG and T. CARRINGTON JR., *A contracted basis-Lanczos calculation of the vibrational levels of methane: solving the Schroedinger equation in nine dimensions* J. Chem. Phys. **119**, 101-117 (2003).
[7] J. C. LIGHT and T. CARRINGTON JR., *Discrete variable representations and their utilization* Advances in Chemical Physics **114**, 263-310 (2000).
[8] C.C. Paige, *Computational variants of the Lanczos method for the eigenvalue problem* J. Inst. Math. Appl., **10**, 373–381 (1972)
[9] J.K. Cullum and R.A. Willoughby, Lanczos algorithms for large symmetric eigenvalue computations. Volume 1, theory, Birkhauser, Boston, 1985.

Discrete Compactness for Spectral Edge Elements

Ralf Hiptmair
(joint work with D. Boffi, M. Costabel, M. Dauge, L. Demkowicz)

1. Introduction.

Given a Lipschitz polyhedron $\Omega \subset \mathbb{R}^3$, identifying *spectrally correct* and *spurious free* [9, Sect. 4] conforming Galerkin approximations of the Maxwell eigenvalue problem [11]: seek $\mathbf{u} \in \boldsymbol{H}(\mathbf{curl}, \Omega)$ and $\omega > 0$, such that

$$\left(\boldsymbol{\mu}^{-1}\,\mathbf{curl}\,\mathbf{u}, \mathbf{curl}\,\mathbf{v}\right)_{L^2(\Omega)} = \omega^2\,(\boldsymbol{\varepsilon}\mathbf{u}, \mathbf{v})_{L^2(\Omega)} \quad \forall \mathbf{v} \in \boldsymbol{H}(\mathbf{curl}, \Omega)\;, \tag{1}$$

($\boldsymbol{\varepsilon}, \boldsymbol{\mu} \in (L^\infty(\Omega))^{3,3}$ uniformly positive definite material tensors), is a challenging problem, because the underlying operator possesses the infinite-dimensional null space $\boldsymbol{H}(\mathbf{curl}\,0, \Omega)$, leading to the essential spectrum $\{0\}$ beside a point spectrum $\subset \mathbb{R}^+$.As a consequence, asymptotic density of Galerkin trial/test spaces for (1) will not guarantee convergence of discrete eigenvalues and eigenvectors, *cf.* the spurious modes reported in [6].

It was soon recognized that the Galerkin discretization of (1) requires suitable finite element spaces which are generally termed *edge* finite elements (see [27, 28, 7]). Pioneering analysis for lowest order edge elements (h-version) was conducted in [23, 24], where the so-called *discrete compactness property* (see below) had been identified as a key concept. Other relevant works on the subject are [6, 1, 9, 26, 11, 2], and a review is available in [20, Sect. 4], [25, Sect. 7.3.2]. The analysis presented in the references above covers the h-version for basically all known families of edge finite elements.

It soon turned out that the analysis of the p- and hp-versions of edge finite elements needed tools different from those developed for the h-version. Partial results could only be achieved in 2D [5, 3]. Only recently, thanks to the invention of new and powerful tools, the proof of discrete compactness for the spectral families of edge elements (p-version) could be achieved [4, 21].

2. Discrete Compactness.

Let $\left(\mathcal{W}_p^1(\mathcal{M})\right)_{p\in\mathbb{N}}$ be a family of finite dimensional subspaces of $\boldsymbol{H}(\mathbf{curl},\Omega)$. As discussed in [9, 8] based on [17] they ought to enjoy the following property in order provide a convergent Galerkin discretization of (1).

Definition 1. *The* discrete compactness property *holds for an asymptotically dense family* $\left(\mathcal{W}_p^1(\mathcal{M})\right)_{p\in\mathbb{N}}$ *of finite dimensional subspaces of* $\boldsymbol{H}(\mathbf{curl},\Omega)$*, if any* bounded *sequence* $(\mathbf{u}_p)_{p\in\mathbb{N}} \subset \boldsymbol{H}(\mathbf{curl},\Omega)$ *with* $\mathbf{u}_p \in \mathcal{X}_p^1(\mathcal{M})$ *contains a subsequence that* converges in $\boldsymbol{L}^2(\Omega)$.

Here, $\mathcal{X}_p^1(\mathcal{M})$ is a subspace of $\mathcal{W}_p^1(\mathcal{M})$ to which discrete eigenfunctions associated with non-zero eigenvalues will belong:

$$\mathcal{X}_p^1(\mathcal{M}) := \{\mathbf{w}_p \in \mathcal{W}_p^1(\mathcal{M}) : (\boldsymbol{\varepsilon}\mathbf{w}_p, \mathbf{v}_p)_{L^2(\Omega)} = 0 \ \forall \mathbf{v}_p \in \mathrm{Ker}(\mathbf{curl}) \cap \mathcal{W}_p^1(\mathcal{M})\} \ .$$

To elucidate the term "discrete compactness" we recall the compact embedding [29, 22]

$$X := \{\mathbf{v} \in \boldsymbol{H}(\mathbf{curl},\Omega) : \ (\mathbf{v},\mathbf{z})_{L^2(\Omega)} = 0 \ \forall \mathbf{z} \in \mathrm{Ker}(\mathbf{curl})\} \overset{c}{\hookrightarrow} \boldsymbol{L}^2(\Omega) \ .$$

When restricting (1) to X, it becomes an eigenvalue problem for an s.p.d. operator with compact resolvent and Riesz-Schauder theory can be applied. The compact embedding means that any bounded sequence in X, which is equipped with the $\boldsymbol{H}(\mathbf{curl},\Omega)$-norm possesses a subsequence that converges in $\boldsymbol{L}^2(\Omega)$. Definition 1 states this very feature for the spaces $\mathcal{X}_p^1(\mathcal{M})$, which can be regarded as discrete counterparts of X.

3. Abstract Theory

As proved in [21, Sect. 6] the discrete compactness property follows from the existence of special projectors that preserve **curl**-free vectorfields:

Assumption 1. *There are linear projectors* $\Pi_p : \mathcal{D}(\Pi_p) \subset \boldsymbol{H}(\mathbf{curl},\Omega) \mapsto \mathcal{W}_p^1(\mathcal{M})$ *that satisfy*

- $\mathbf{v} \in \boldsymbol{H}(\mathbf{curl}\,0,\Omega) \cap \mathcal{D}(\Pi_p) \quad \Rightarrow \quad \mathbf{curl}\,\Pi_p\mathbf{v} = 0$,
- $\mathcal{U}_p := \{\mathbf{v} \in H^{1/2+s}(\Omega),\ \mathbf{curl}\,\mathbf{v} \in \mathbf{curl}\,\mathcal{W}_p^1(\mathcal{M})\} \subset \mathcal{D}(\Pi_p)$,
- $\mathbf{u} \in \mathcal{U}_p : \quad \|\mathbf{u} - \Pi_p\mathbf{u}\|_{L^2(\Omega)} \lesssim \delta(p)\left(\|\mathbf{u}\|_{H^{1/2+s}(\Omega)} + \|\mathbf{curl}\,\mathbf{u}\|_{L^2(\Omega)}\right)$, *with* $\lim\limits_{p\to\infty}\delta(p) = 0$.

Here, $\lesssim$ hints at estimates with constants independent of p and any function involved.

Theorem 1. *Under Assumption* 1 *the family* $\left(\mathcal{W}_p^1(\mathcal{M})\right)_{p\in\mathbb{N}}$ *features the discrete compactness property of Definition* 1.

The existence of such projectors can be reduced to more concrete assumptions for finite element spaces built on a *fixed* mesh $\mathcal{M} = \{T\}$ of Ω.

Assumption 2. *The projectors* $\Pi_p : \mathcal{D}(\Pi_p) \subset \boldsymbol{H}(\mathbf{curl}, \Omega) \mapsto \mathcal{W}_p^1(\mathcal{M})$ *are local in the sense that* $(\Pi_p \mathbf{v})_{|T} = \Pi_{p,T}(\mathbf{v}_{|T})$ *for all cells* $T \in \mathcal{M}$.

Assumption 3. *There are local projectors* $\mathsf{I}_{p,T} : H^{3/2+s}(T) \mapsto \mathcal{S}_p(T)$, $T \in \mathcal{M}$, *such that*

- $\mathbf{grad} \circ \mathsf{I}_{p,T} = \Pi_{p,T} \circ \mathbf{grad}$ *on* $H^{3/2+s}(T)$ *(commuting diagram property)*
- $|\varphi - \mathsf{I}_{p,T}\varphi|_{H^1(T)} \lesssim \delta(p)|\varphi|_{H^{3/2+s}(T)}$, *with* $\lim_{p\to\infty} \delta(p) = 0$.

Assumption 4. *There are* continuous *local liftings* $\mathsf{R}_T : (L^2(T))^3 \mapsto (H^1(T))^3$ *such that for any* $T \in \mathcal{M}$

- $\mathbf{curl} \circ \mathsf{R}_T \circ \mathbf{curl} = \mathbf{curl}$ *on* $\boldsymbol{H}(\mathbf{curl}, T)$,
- $\mathsf{R}_T \, \mathbf{curl}\, \mathcal{W}_p^1(T) \subset \mathcal{W}_p^1(T)$, *(compatibility with FE spaces)*

Theorem 2. *Assumptions* 2 *through* 4 *ensure that the projectors* Π_p *assembled from the local projectors* $\Pi_{p,T}$ *comply with Assumption* 1.

4. Tools

We restrict ourselves to $\mathcal{W}_p^1(\mathcal{M})$, $p \in \mathbb{N}$, generated by the first family of Nedelec's tetrahedral edge elements [27].

(I) Commuting projection based interpolation operators. This class of projectors that fit Assumptions 2 and 3 were devised by L. Demkowicz and others in [14, 15, 13, 16], see also [20, Sect. 3.5] for a general construction and [21, Sect. 4.2] for technical details. The main building blocks are projections (of traces) onto function spaces on edges and faces of tetrahedra, and polynomial preserving extension operators. The projection based interpolation operators from [16] allow the choice $\delta(p) = (\log p)^{3/2} p^{-1/2-\varepsilon}$.

(II) Smoothed Poincaré liftings. The local liftings postulated in Assumption 4 are to provide right inverses of **curl**, when restricted to its range. Such operator is known as Poincaré lifting in differential geometry [10, Sect. 2.13]. For $D \subset \mathbb{R}^3$ star-shaped with respect to $\boldsymbol{a} \in D$ it reads [18, Prop. 2.1]

$$\mathsf{R}_{\boldsymbol{a}}(\mathbf{u})(\boldsymbol{x}) := \int_0^1 t\mathbf{u}(\boldsymbol{x} + t(\boldsymbol{x} - \boldsymbol{a}))\,\mathrm{d}t \times (\boldsymbol{x} - \boldsymbol{a}) \;, \quad \boldsymbol{x} \in D \;.$$

Preservation of $\mathcal{W}_p(T)$ under the action of $\mathsf{R}_{\boldsymbol{a}}$ can be shown [19]. However, this operator fails to meet continuity requirements of Assumption 4. As discovered in the breakthrough paper [12], a smoothing trick often used in the theory of function spaces offers a remedy: For D star-shaped w.r.t. ball $B \subset D$ pick $\Phi \in C^\infty(\mathbb{R}^3)$ with $\operatorname{supp}\Phi \subset B$, $\int_B \Phi(\boldsymbol{a})\,\mathrm{d}\boldsymbol{a} = 1$ and define

$$\mathsf{R}(\mathbf{u}) := \int_B \Phi(\boldsymbol{a}) \mathsf{R}_{\boldsymbol{a}}(\mathbf{u})\,\mathrm{d}\boldsymbol{a} \;,$$

which turns out to be a pseudo-differential operator of order -1. With $D = T$ this operator qualifies as a lifting R_T in Assumption 4.

References

[1] D. Boffi, *Fortin operator and discrete compactness for edge elements*, Numer. Math., 87 (2000), pp. 229–246.

[2] ———, *Approximation of eigenvalues in mixed form, discrete compactness property, and application to hp mixed finite elements*, Computer Methods in Applied Mechanics and Engineering, 196 (2007), pp. 3672–3681.

[3] D. Boffi, M. Costabel, M. Dauge, and L. Demkowicz, *Discrete compactness for the hp version of rectangular edge finite elements*, SIAM J. Numer. Anal., 44 (2006), pp. 979–1004.

[4] D. Boffi, M. Costabel, M. Dauge, L. Demkowicz, and R. Hiptmair, *Discrete compactness for the p-version of discrete differential forms*, report, SAM, ETH Zürich, Zürich, Switzerland, August 2009.

[5] D. Boffi, L. Demkowicz, and M. Costabel, *Discrete compactness for p and hp 2D edge finite elements*, Math. Models Methods Appl. Sci., 13 (2003), pp. 1673–1687.

[6] D. Boffi, P. Fernandes, L. Gastaldi, and I. Perugia, *Computational models of electromagnetic resonators: Analysis of edge element approximation*, SIAM J. Numer. Anal., 36 (1999), pp. 1264–1290.

[7] A. Bossavit, *Whitney forms: A class of finite elements for three-dimensional computations in electromagnetism*, IEE Proc. A, 135 (1988), pp. 493–500.

[8] S. Caorsi, P. Fernandes, and M. Rafetto, *Spurious-free approximations of electromagnetic eigenproblems by means of nedelec-type elements*, Math. Model. Numer. Anal., 35 (2001), pp. 331–354.

[9] S. Caorsi, P. Fernandes, and M. Raffetto, *On the convergence of Galerkin finite element approximations of electromagnetic eigenproblems*, SIAM J. Numer. Anal., 38 (2000), pp. 580–607.

[10] H. Cartan, *Formes Différentielles*, Hermann, Paris, 1967.

[11] M. Costabel and M. Dauge, *Computation of resonance frequencies for Maxwell equations in non-smooth domains*, in Computational Methods in Wave Propagation, M. Ainsworth, ed., Springer, New York, 2003, pp. 127–164. In print.

[12] M. Costabel and A. McIntosh, *On Bogovskii and regularized Poincaré integral operators for de Rham complexes on Lipschitz domains*, Math. Z., (2009). DOI 10.1007/s00209-009-0517-8, http://arxiv.org/abs/0808.2614v1.

[13] L. Demkowicz, *Projection-based interpolation*, Report 04-03, ICES, University of Texas at Austin, Austin, TX, 2004.

[14] ———, *Polynomial exact sequences and projection-based interpolation with applications to Maxwell equations*, in Mixed Finite Elements, Compatibility Conditions, and Applications, D. Boffi, F. Brezzi, L. Demkowicz, R. Duran, R. falk, and M. Fortin, eds., vol. 1939 of Lecture Notes in Mathematics, Springer, Berlin, 2008, pp. 101–158.

[15] L. Demkowicz and I. Babuška, *p interpolation error estimates for edge finite elements of variable order in two dimensions*, SIAM J. Numer. Anal., 41 (2003), pp. 1195–1208.

[16] L. Demkowicz and A. Buffa, *H^1, $\mathbf{H}(\mathbf{curl})$, and $\mathbf{H}(\mathrm{div})$-conforming projection-based interpolation int three dimensions. Quasi-optimal p-interpolation estimates*, Comput. Meth. Appl. Mech. Engr., 194 (2005), pp. 267–296.

[17] J. Descloux, N. Nassaf, and J. Rappaz, *On spectral approximation. Part I. The problem of convergence*, R.A.I.R.O. Numerical Analysis, 12 (1978), pp. 97–112.

[18] J. Gopalakrishnan and L. Demkowicz, *Quasioptimality of some spectral mixed method*, K. Comput. Appl. Math., 167 (2004), pp. 163–182.

[19] R. Hiptmair, *Canonical construction of finite elements*, Math. Comp., 68 (1999), pp. 1325–1346.

[20] ———, *Finite elements in computational electromagnetism*, Acta Numerica, 11 (2002), pp. 237–339.

[21] ———, *Discrete compactness for p-version of tetrahedral edge elements*, Report 2008-31, SAM, ETH Zurich, Zürich, Switzerland, 2008. http://arxiv.org/abs/0901.0761.

[22] F. Jochmann, *A compactness result for vector fields with divergence and curl in $L^q(\Omega)$ involving mixed boundary conditions*, Appl. Anal., 66 (1997), pp. 189–203.
[23] F. Kikuch, *Mixed and penalty formulations for finite element analysis of an eigenvalue problem in electromagnetism*, Comput. Methods Appl. Mech. Eng., 64 (1987), pp. 509–521.
[24] F. Kikuchi, *On a discrete compactness property for the Nédélec finite elements*, J. Fac. Sci., Univ. Tokyo, Sect. I A, 36 (1989), pp. 479–490.
[25] P. Monk, *Finite Element Methods for Maxwell's Equations*, Clarendon Press, Oxford, UK, 2003.
[26] P. Monk and L. Demkowicz, *Discrete compactness and the approximation of Maxwell's equations in $\mathbb{R}^3$*, Math. Comp., 70 (2001), pp. 507–523. Published online February 23, 2000.
[27] J. Nédélec, *Mixed finite elements in $\mathbb{R}^3$*, Numer. Math., 35 (1980), pp. 315–341.
[28] J. Nédélec, *A new family of mixed finite elements in R^3*, Numer. Math., 50 (1986), pp. 57–81.
[29] R. Picard, *An elementary proof for a compact imbedding result in generalized electromagnetic theory*, Math. Z., 187 (1984), pp. 151–161.

Computation of resonances for Maxwell's equations: transparent boundary conditions and goal oriented error estimation

Lin Zschiedrich
(joint work with Frank Schmidt)

We address the numerical computation of resonances in open optical devices. The imaginary part of the resonance eigenvalue corresponds to radiation losses and is of major importance in practice. Typically this imaginary part is orders of magnitude smaller than the real part. Therefore, a precise computation of radiation losses by means of an adaptive finite element method requires a specially constructed goal oriented error estimator in order to guarantee optimum convergence in the leading digits of the eigenvalue's imaginary part irrespective of the magnitude of the eigenvalue. Following Rannacher et. al [1, 3] we develop such an estimator and demonstrate the accuracy of the method for a challenging leaky waveguide mode computation benchmark problem, see Bienstman et al. [2].

References

[1] R. Becker and R. Rannacher, *An optimal control approach to a posteriori error estimation in finite element methods.* In A. Iserles, editor, *Acta Numerica 2000*, pages 1–102. Cambridge Univerity Press.
[2] P. Bienstman and S. Selleri and L. Rosa, *Modelling lossy photonics wires: a mode solver comparison.* OWTNM, 2006.
[3] V. Heuveline and R. Rannacher, *A posteriori error control for finite element approximations of elliptic eigenvalue problems. J. Adv. Comp. Math.*, 15:107, 2001.

On gradient solvers for elliptic eigenvalue problems

Klaus Neymeyr

The topic of the talk is the numerical solution of large eigenvalue problems for self-adjoint elliptic partial differential operators by means of (preconditioned) gradient iterations. These iterative solvers are also called "matrix-free" iterations, which allows to distinguish them from transformation methods; the gradient iterations use the discretization and the mass matrix only in terms of matrix-vector multiplications.

In an introductory part (preconditioned) gradient iterations are systematically derived by a comparison of the boundary value and of the eigenvalue problem - both problems are considered for self-adjoint elliptic partial differential operators. For the boundary value problem a highly efficient/optimal solver is the (multigrid) preconditioned conjugate gradient scheme. The corresponding most efficient scheme for the eigenvalue problem (for the same partial differential operator) appears to be the Locally Optimal Preconditioned Conjugate Gradient (LOPCG) scheme which can reach an effectiveness comparable to that of the conjugate gradient scheme.

We present a new approach to the convergence analysis of the simplest preconditioned eigensolver for the generalized hermitian positive semi-definite eigenproblem [1]. This solver is the preconditioned gradient scheme with fixed step length; the proof reduces the derivation of the convergence rate bound to a two-dimensional subspace by analysing the extremal behavior of the gradient flow for the Rayleigh quotient.

Next a general hierarchy of gradient-type eigensolvers is suggested. These eigensolvers are based on an application of the Rayleigh-Ritz procedure (as an effective minimizer) to subspaces which are spanned by the current preconditioned residual and a fixed number of preceding iterates. The convergence behavior of these solvers can be understood very well by considering the limit of a preconditioning with the exact inverse of the system matrix. In this limit these solvers are Invert-Lanczos processes. This link between these (ideally) preconditioned eigensolvers and the Lanczos process allows to apply the Kaniel-Paige convergence theory of the Lanczos process in order to derive convergence estimates for a general class of iterative schemes [2].

Furthermore, a necessary condition is derived for the poorest convergence of these Invert-Lanczos processes with respect to level set of initial iterates having a fixed Rayleigh quotient. It is shown that the poorest convergence is attained in low dimensional eigenspaces. This analysis provides the justification for "mini-dimensional" approaches to the convergence analysis of these eigensolvers; sharp convergence estimates are derived for steepest ascent/steepest descent both in $\mathrm{span}\{x, Ax\}$ and $\mathrm{span}\{x, A^{-1}x\}$ [3].

References

[1] A.V. Knyazev, K. Neymeyr, *Gradient flow approach to geometric convergence analysis of preconditioned eigensolvers*, SIAM J. Matrix Analysis **31** (2009), 621–628.

[2] K. Neymeyr, *On preconditioned eigensolvers and Invert-Lanczos processes*, Linear Algebra Appl. **430** (2009), 1039–1056.
[3] K. Neymeyr, M. Zhou, *On gradient type solvers for the symmetric eigenvalue problem*, under preparation (2009).

An adaptive algorithm for symmetric eigenvalue problems - Convergence and complexity

Thorsten Rohwedder
(joint work with Wolfgang Dahmen, Reinhold Schneider, Andreas Zeiser)

In this talk, we present our results from [DRSZ] and [RSZ] concerned with the analysis of convergence and complexity of an adaptive iteration scheme for the calculation of the smallest eigenvalue of an elliptic operator eigenvalue problem. Let $V \hookrightarrow H \cong H^* \hookrightarrow V^*$ a Gelfand triple with a dual pairing $\langle \cdot, \cdot \rangle$ given on $V^* \times V$. For given bounded, symmetric bilinear form $a : V \times V \to \mathbb{R}$ which is strongly positive (i.e. $a(u,u) \geq \kappa \|u\|_V^2$), we are concerned with finding an eigenpair (λ^*, u^*) of the weak eigenvalue problem

$$a(u,v) = \lambda \langle u, v \rangle \quad \text{for all } v \in V, \tag{1}$$

where λ^* is the lowest eigenvalue of a. For simplicity, we assume that λ^* is a simple eigenvalue, and that the rest of the spectrum is bounded from below by $\Lambda > \lambda^*$. Our problem can alternatively be recast in the (weak) operator EVP

$$Au = \lambda^* Eu \quad \text{in } V^*.$$

with $A : V \to V^*, Au = a(u,.), E : H \to H^*, Eu = \langle u, . \rangle$. As an example, we mention the Poisson eigenvalue, where we have $V = H_0^1(\Omega)$, $H = L_2(\Omega)$, $a(u,v) = \langle \nabla u, \nabla v \rangle$, $A = \Delta : V \to V^*$.

The standard procedure for the treatment of (1) is to discretize the equation, e.g. by finite elements or finite differences, which gives rise to a finite dimensional discrete problem. In contrast to this, our main attention is to avoid the separation of the discretization and solution process. Rather, we tackle the problem from a different perspective inspired by the work of [CDD]: Instead of solving a sequence of finite dimensional eigenvalue problems, we stick to the *infinite dimensional* formulation. On this level we can construct an ideal iterative algorithm (formulated in the continuous space H) basing on a gradient-type solver,

PPINVIT (Perturbed Preconditioned Inverse Iteration)

Starting vector $u^0 \in V$, $||u^0||_H = 1$.

$$\begin{aligned} \tilde{u}_\xi^{n+1} &= u^n - P^{-1}(Au^n - R(u^n)Eu^n) + \xi^{n+1}, \\ u^{n+1} &= \tilde{u}_\xi^{n+1} / ||\tilde{u}_\xi^{n+1}||_H, \end{aligned}$$

In *PPINVIT*, the n-th step allows for a perturbation ξ^n which will later in the numerical realization of the algorithm be associated with the error invoked by

finite-dimensional approximation of the in principle infinite dimensional residual $r^n := P^{-1}(Au^n - R(u^n)Eu^n)$.

Sharp results for the convergence of the eigenvalues for the discretized preconditioned inverse iteration (*PINVIT*) have recently been reproven by a elegant, short geometric proof [KN]. In [RSZ], we generalize this proof to the operator case (including above perturbations): If one assumes that the preconditioner is scaled such that $||I - P^{-1}A||_A = \gamma < 1$, that the Rayleigh quotient $R(v)$ of the current iterate v fulfills $\lambda_k \leq R(v) < \lambda_{k+1}$, and that the perturbation ξ is bounded by

$$\|\xi\|_A \;\leq\; \gamma_\xi \cdot \|Av - R(v)Ev\|_{A^{-1}}, \tag{2}$$

where $\tilde{\gamma} := \gamma + \gamma_\xi < 1$, the next *PPINVIT* step gives $R(v') < \lambda_k$ or $\lambda_k \leq R(v') < \lambda_{k+1}$. In the latter case,

$$\frac{R(v') - \lambda_k}{\lambda_{k+1} - R(v')} \;\leq\; q^2(\tilde{\gamma}, \lambda_k, \lambda_{k+1})\; \frac{R(v) - \lambda_k}{\lambda_{k+1} - R(v)},$$

where $q(\tilde{\gamma}, \lambda_k, \lambda_{k+1}) = 1 - (1 - \tilde{\gamma})(1 - \lambda_k/\lambda_{k+1})$. A corresponding result from [DRSZ] for the subspace error measured by $\sin_P \angle(v, u^*) = ||(I - P_{P,u^*})v||_P$ states that asymptotically, we can maintain linear convergence of the error if the perturbation is bounded by a multiple of the current residual $\|A - \lambda(v)Ev\|_{A^{-1}}$.

We now have to realize the in principle infinite dimensional updates r^n (up to the perturbation ξ) in a finite dimensional setting, particularly directing our attention to carrying out this task at possibly low computational cost. In [RSZ], we devise an evaluation scheme *PPINVIT_STEP* based on the residual estimator $\rho_\varepsilon =: \|A_\varepsilon(v) - \mu_\varepsilon(v)E_\varepsilon(v)\|_{P^{-1}}/\|v\|_P$, where the ε indicates that the respective quantities are not evaluated exactly. We show that by *PPINVIT_STEP*, the condition (2) can be fulfilled in each step while the accuracy ε needed for applications of A and E stays bounded by $\varepsilon \geq C \cdot \max(\tau, \rho(v))$.

We turn to the matter of discretization now and show that in the context of stable wavelet discretization, our adaptive scheme exhibits in some sense asymptotically quasi-optimal complexity: For the coefficient vector $\mathbf{u} \in \ell_2(\mathcal{I})$ belonging to the discretization of $u \in V$ with respect to a chosen basis, the best accuracy obtainable with N degrees of freedom in $\ell_2(\mathcal{I})$ is

$$\sigma_N(\mathbf{u}) \;=\; \inf\{\|\mathbf{u} - \mathbf{w}\|_{\ell_2},\; \mathrm{supp}(\mathbf{w}) \leq N\}. \tag{3}$$

We define related subspaces $\mathcal{A}^s$ of $\ell_2(\mathcal{I})$ by saying that $\mathbf{u} \in \mathcal{A}^s$ if it can be approximated to any accuracy $\varepsilon > 0$ by a vector $\mathbf{u}_\varepsilon \in \ell_2(\mathcal{I})$ with

$$\mathrm{supp}(\mathbf{u}_\varepsilon) \;\leq\; \varepsilon^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s}$$

(see [CDD] for details). A *quasi-optimal approximation* to u is a sequence of vectors for which $|\mathbf{u}^n - \mathbf{u}||_{\ell_2} < \varepsilon(n) \to 0$, while $\mathrm{supp}(\mathbf{u}^n) \leq C \cdot \varepsilon(n)^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}$, with constants independent of n. By using a nonlinear application scheme *APPLY* for

computation of $\mathbf{Ax}$ we may compute an output $\mathbf{a}$ that, for given δ, satisfies

$$||\mathbf{Ax}-\mathbf{a}|| < \delta, \quad ||\mathbf{b}||_{\mathcal{A}^s} \leq C \cdot ||\mathbf{x}||_{\mathcal{A}^s}, \quad \mathrm{supp}(\mathbf{b}) \leq C \cdot \mathrm{supp}(\mathbf{x}),$$

while the computation of $\mathbf{a}$ may be performed in linear complexity in $\mathrm{supp}(\mathbf{u})$, $\varepsilon^{-1/s}$ for wide class of discretized operators (quasi-sparsity), see [CDD] for details. Analogous results are shown for the remaining involved quantities in [RSZ]. We now construct an algorithm *MINIEIG* from the above *PPINVIT_STEP* and an additional coarsening step *COARSE* taken from [CDD], see [DRSZ] for an algorithmic formulation. *MINIEIG* has following properties:

Theorem. *For compressible operators* $\mathbf{A}$ *and* $\mathbf{E}$*, and a prescribed target accuracy* δ*,* $MINIEIG$ *computes a vector* $\mathbf{u}_\delta$ *for which*

$$\angle(\mathbf{u}_\delta, \mathbf{u}^*) \ \leq C \cdot \ \delta, \quad |\mu_\delta(\mathbf{u}_\delta) - \lambda^*| \ \leq C \cdot \ \delta.$$

If the eigenvector $\mathbf{u}^* \in \mathcal{A}^s$*, then*

$$\mathrm{supp}(\mathbf{u}_\delta) \leq C \cdot \delta^{-1/s}$$

given sufficiently compressible operators. The number of floating point operations remains bounded by $C \cdot \delta^{-1/s}$*.*

To end with, we turn to numerical examples: For the application of stable wavelet bases, the compressibility of functions is related with their Besov regularity. Comparison of an analysis of the eigenfunctions of the Poisson eigenvalue problem on a polygonal domain with a numerical example shows that the predicted behaviour is reproduced for the model problem of the Laplacian on the L-shaped domain, see [RSZ].
A second application is given by a Galerkin ansatz for the Schrödinger equation $H\Psi = E\Psi$, (called CISD by quantum chemists), where $V \subseteq H^1(\mathbb{R}^3 \times \{\pm\frac{1}{2}\})^N$ and $H : V \to V^*$ is the electronic Hamiltonian. In this case, V is discretized in advance, but the resulting tensor product basis of functions $\varphi_i \in H^1(\mathbb{R}^3)$ of size $\binom{I}{N}$ is extremely large. We tested the use of *COARSE*-steps to keep the iteration data-sparse, using a threshold adjusted by the current residual measured in the dual norm. The complexity thus reduces from N^6 for usual methods to $N^2 \cdot \#\mathrm{supp}(\mathbf{u})$. Although the convergence behaviour seems to be quite robust, the compressed supports are unfortunately still extremely large, mainly because the standard bases of quantum chemistry are non-local. A utilization of localized orbitals to be realized in the future looks more promising.

References

[CDD] A. Cohen, W. Dahmen, R. DeVore, *Adaptive wavelet methods for elliptic operator equations: Convergence rates*, Math. Comp. **70** (2001), 27-75,

[KN] A. Knyazev, K. Neymeyr, *Gradient flow approach to geometric convergence analysis of precond. eigensolvers*, Accepted to SIMAX. Published as a tech. report http://arxiv.org/abs/0801.3099 (2009).

[DRSZ] W. Dahmen, T. Rohwedder, R. Schneider, A. Zeiser,*Adaptive Eigenvalue Computation – Complexity Estimates* Num. Math. 110, **3**, (2008), 277-312.

[RSZ] T. Rohwedder, R. Schneider, A. Zeiser, *Perturbed preconditioned inverse iteration for operator eigenvalue problems with applications to adaptive wavelet discretizations*, to appear in AiCM (2009).

Automated multilevel substructuring for elastodynamic eigenvalue problems

Jeffrey K. Bennighof

Automated multilevel substructuring (AMLS)[1] is a method for approximating many (e.g., 10^4) eigenpairs of high-dimension (e.g., 10^7 d.o.f.) finite element (FE) discretizations of structures, for use in frequency response calculations. In the continuous setting, the problem domain is recursively partitioned into many subdomains by interfaces. The Sobolev space over which the vibration differential eigenvalue problem is defined is decomposed into a set of subspaces associated with subdomains, and subspaces of minimum-energy extensions of trace functions on interfaces. All of these subspaces are orthogonal in the bilinear form associated with strain energy. Differential eigenvalue problems are defined over the subspaces, and a basis for the global Sobolev space is constructed in terms of eigenfunctions from all of the subspaces. All of the subspace eigenproblems are of similar form, so their eigenspaces can be truncated consistently.

This approach can be applied to a large FE discretization by first defining a multilevel tree of substructures, using nested dissection on the union of the graphs of the stiffness and mass matrices. Subdomain eigensolutions are obtained using submatrices associated with leaf nodes in the substructure tree. Interface eigenvalue problems first require construction of interface submatrices through definition of minimum-energy extensions in the discrete FE subspace. Partial eigensolutions of both subdomain and interface eigenvalue problems establish a subspace of lower dimension from which the global eigensolution is approximated.

This strategy has been found to be much more economical for producing many approximate eigenpairs than solving the global eigenvalue problem using a Krylov subspace approach. The frequency response error associated with the dimensional reduction performed in AMLS is ordinarily much smaller than the FE discretization error. AMLS enables localized dimensional reduction of the FE discretization, on inexpensive computer hardware, in a manner that respects matrix sparsity and handles ill-conditioned system matrices effectively.

References

[1] J. K. Bennighof and R. B. Lehoucq, *An automated multilevel substructuring method for eigenspace computation in linear elastodynamics*, SIAM J. Sci. Comput. **25** (2004), 2084–2106.

Spectral information and Krylov subspace methods

Zdeněk Strakoš

Spectral information naturally plays an important role in analysis and understanding of many phenomena in computational mathematics. It often reflects some properties of the underlying real world problems that are approximately solved via mathematical modeling, discretization and subsequent computation. These individual stages should be considered parts of a single solution process which couples together knowledge from, e.g., theory of partial differential equations, numerical analysis and numerical methods, including numerical linear algebra. In particular, the computational stage should not be separated from modeling, numerical analysis and discretization. This is useful for achieving computational efficiency as well as for getting an insight into various subproblems in the individual stages.

Behaviour of Krylov subspace methods for solving linear algebraic problems may serve as an example. Unlike in the classical iterative methods, in the Lanczos method and the conjugate gradient method (CG), or in the Arnoldi method and the generalized minimal residual method (GMRES), the optimality at each step over Krylov subspaces of increasing dimensionality makes a linearized description inadequate. By their nature they represent *model reductions based on matching moments*. Such view complements, to our opinion, the standard description using the projection processes framework, and it reveals in an instructive way the *nonlinear character* of Krylov subspace methods.

Consider first the Hermitian case. The Lanczos and CG behaviour (including the effects of rounding errors) is fully determined by the spectral decomposition of the problem, but the relationship between convergence and the spectral information is nothing but simple. This can be demonstrated by using a modification of the classical Stieltjes moment problem, and by showing that its solution is given:

- in the language of orthogonal polynomials by the Gauss-Christoffel quadrature;
- in the algebraic matrix form by the Lanczos and CG method.

As a consequence, the Lanczos and CG method can be considered as a *matrix formulation of the Gauss-Christoffel quadrature.* In order to allow straightforward generalizations, we use the Vorobyev method of moments, and present some basic Krylov subspace iterations as moment matching model reduction.

In the non-Hermitian case the spectral information is not generally sufficient for description of behaviour of Krylov subspace methods. In particular, given an arbitrary prescribed convergence history of GMRES and an arbitrary prescribed spectrum of the system matrix, there is always a system $Ax = b$ such that GMRES follows the prescribed convergence while A has the prescribed spectrum. From the practical point of view, discretization of convection-diffusion problems indeed leads to algebraic systems where the link between the spectral information and convergence of GMRES is not apparent. Moreover, the subproblem of stabilization of the discrete solution in the convection-dominated case can not be separated from the question on efficiency of the algebraic solver.

Given a (non-increasing) convergence curve and an arbitrary spectrum, we know the complete parametrization of the set of *all* matrices and right hand sides such that GMRES follows the prescribed convergence while A has the prescribed spectrum. We believe that this offers an interesting tool which can be used in further theoretical and experimental investigations.

This talk recalls results published in the papers [1], [2], [3], [4], [5], [6], [7], [8], and [9]. Most of them contain extensive lists of further references. The Vorobyev method of moments is presented in the monograph [10] (with the Russian original published in 1958) that remained almost unknown and it is very rarely quoted.

References

[1] A. Greenbaum and Z. Strakoš, *Matrices that generate the same Krylov residual spaces*, in Recent Advances in Iterative Methods, G. H. Golub, A. Greenbaum, and M. Luskin, eds., vol. 60 of The IMA volumes in mathematics and its applications, Springer-Verlag, New York, 1994, 95–118.

[2] A. Greenbaum, V. Pták, and Z. Strakoš, *Any nonincreasing convergence curve is possible for GMRES*, SIAM J. Matrix Anal. Appl. **17** (1996), 465–469.

[3] M. Arioli, V. Pták, and Z. Strakoš, Krylov sequences of maximal length and convergence of GMRES, BIT **38** (1998), 636–643.

[4] Z. Strakoš and P. Tichý, On error estimation in the conjugate gradient method and why it works in finite precision computations, Electron. Trans. Numer. Anal. **13** (2002), 56–80.

[5] J. Liesen and Z. Strakoš, *GMRES Convergence analysis for a convection-diffusion model problem*, SIAM J. Sci. Comput. **26** (2005), 1989–2009.

[6] G. Meurant and Z. Strakoš, *The Lanczos and conjugate gradient algorithms in finite precision arithmetic*, Acta Numerica **15** (2006), 471–542.

[7] D. P. O'Leary, Z. Strakoš, and P. Tichý, *On sensitivity of Gauss-Christoffel quadrature*, Numerische Mathematik **107** (2007), 147–174.

[8] Z. Strakoš, Model reduction using the Vorobyev momentproblem, Numerical Algorithms **51** (2009), 363–379.

[9] Z. Strakoš and P. Tichý, On efficient numerical approximation of the scattering amplitude $c^*A^{-1}b$ via matching moments, submitted to SIAM J. Sci. Comput. (2009).

[10] Y. V. Vorobyev, *Methods of Moments in Applied Mathematics*, Translated from the Russian by Bernard Seckler, Gordon and Breach Science Publishers, New York, 1965.

Comparative analysis of inner-outer eigenvalue solvers

Yvan Notay

(joint work with Michiel E. Hochstenbach)

We consider inverse iteration-based eigensolvers, which require at each step solving an "inner" linear system. We assume that this linear system is solved by some (preconditioned) Krylov subspace method. In this framework, several approaches are possible, which differ by the linear system to be solved and/or the way the preconditioner is used. This includes methods such as inexact shift-and-invert, inexact Rayleigh quotient iteration, Jacobi-Davidson and generalized preconditioned inverse iteration. In this talk, we discuss and compare them, focusing on the evolution of the "outer" convergence (towards the desired eigenpair) according

to the numerical effort spent in inner iterations. The discussion also includes the preconditioned Lanczos method although it is of slightly different type.

The analysis is based on the observation that all these methods extract the approximate eigenvector from a given subspace. Then the discussion is conducted in two phases. Firstly, we investigate the quality of this extraction; that is, considering separately each method, we examine how far is the extracted eigenvector from the best approximation in the subspace at hand. It turns out that all method perform remarkably well from that viewpoint.

Now, this good behavior can be, to a large extent, explained by looking more deeply into the involved processes.

Consider first IRQI. Let V be a $n \times m$ matrix whose columns form a basis of the search space explored by the RQI method. The approximate eigenvector $\mathbf{z}$ is then computed according to

$$\begin{array}{rcl} \mathbf{z} & = & V\,y \\ \big(V^*(A-\eta I)V\big)\,y & = & V^*\mathbf{u}\;. \end{array} \tag{1}$$

On the other hand, when performing Rayleigh-Ritz extraction from the same subspace, one solves

$$V^*A\,V\,y \;=\; \mu\,V^*V\,y\;. \tag{2}$$

Letting δ and $\mathbf{s}$ be such that $\mu = \eta - \delta$ and $Vy = \mathbf{u} + \mathbf{s}$, one has then

$$V^*(A-\eta I)V\,y \;=\; \delta\,V^*\mathbf{u} + \delta\,V^*\mathbf{s}\;. \tag{3}$$

Now, δ and $\mathbf{s}$ are both expected to be small in a relative sense when $\mathbf{u}$ is already somehow close to an eigenvector. Hence, the last term in the right hand side is a second order term, and the comparison of (3) with (1) indicates that IRQI will indeed perform similarly to full Rayleigh-Ritz extraction as long as this term can be neglected.

Further, this reasoning carries over Jacobi-Davidson. Let V be a $n \times m$ matrix whose columns form a basis of the search space in which the correction $\mathbf{t}$ is searched; $\mathbf{t}$ is computed according to

$$\begin{array}{rcl} \mathbf{t} & = & V\,y \\ (V^*(I-\mathbf{u}\mathbf{u}^*)(A-\eta I)V)\;y & = & -V^*\mathbf{r}\;. \end{array} \tag{4}$$

Now, let

$$\widetilde{V} \;=\; \begin{pmatrix} \mathbf{u} & V \end{pmatrix}\;,$$

so that the columns of $\widetilde{V}$ span the extended subspace in which the approximate eigenvector is searched. Letting

$$\widetilde{y} \;=\; \begin{pmatrix} 1 \\ y \end{pmatrix}\;,$$

$\mathbf{z} = \mathbf{u} + \mathbf{t}$ satisfies

$$\mathbf{z} \;=\; \widetilde{V}\,\widetilde{y}\;,$$

whereas the second equation (4) amounts to

$$\Big(V^*(I-\mathbf{u}\mathbf{u}^*)(A-\eta I)\widetilde{V}\Big)\;\widetilde{y} \;=\; 0\;;$$

that is,

$$\left(V^*(A-\eta I)\widetilde{V}\right)\,\widetilde{y} \;=\; \left(V^*\mathbf{u}\right)\left(\mathbf{u}^*(A-\eta I)\widetilde{V}\,\widetilde{y}\right) \;=\; \alpha\, V^*\mathbf{u}\;, \tag{5}$$

where we have let $\alpha = \mathbf{u}^*(A-\eta I)\widetilde{V}\widetilde{y}$. Since $\mathbf{u}^*\mathbf{u} = 1$, one has also $\mathbf{u}^*(A-\eta I)\widetilde{V}\widetilde{y} = \alpha\,\mathbf{u}^*\mathbf{u}$, and, grouping the latter relation with (5), one obtains

$$\begin{aligned} \mathbf{z} &= \widetilde{V}\,\widetilde{y} \\ \left(\widetilde{V}^*(A-\eta I)\widetilde{V}\right)\widetilde{y} &= \alpha\,\widetilde{V}^*\mathbf{u} \\ \widetilde{y}_1 &= 1\;. \end{aligned} \tag{6}$$

This is the same as (1), except that the right-hand-side in the middle equation is scaled so as to satisfy the last additional equation, which is unimportant since the scaling of the approximate eigenvector does not matter. Hence, Jacobi-Davidson acts as an IRQI method on the extended search space. Accordingly, the similarity between Rayleigh-Ritz and IRQI when second order effects are negligible applies as well to Jacobi-Davidson.

This allows to compare the methods on the basis of the related search space, leaving asides their other peculiarities. As will be seen during the talk, this viewpoint allows to explain why some approaches are more efficient.

Computing eigenvalues of diagonally dominant matrices accurately with applications to differential operators

Qiang Ye

Consider a general symmetric positive definite matrix A and let $0 < \lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$ be its eigenvalues. Conventional dense matrix eigenvalue algorithms (such as the QR algorithm) are normwise backward stable, i.e., the computed eigenvalues $\widehat{\lambda}_i$ are the exact eigenvalues of $A+E$ with $\|E\|_2 = \mathbf{O}(\mathbf{u})\|A\|_2$, where $\mathbf{u}$ is the machine roundoff unit. Eigenvalues of large (sparse) matrices are typically computed by an iterative method (such as the Lanczos algorithm), which produces an approximate eigenvalue $\widehat{\lambda}_i$ and an approximate eigenvector $\widehat{x}_i$ whose residual $\|A\widehat{x}_i - \widehat{\lambda}_i\widehat{x}_i\|_2$ is at best (at convergence) $\mathbf{O}(\mathbf{u})\|A\|_2\|\widehat{x}_i\|_2$ In both cases, we have $|\widehat{\lambda}_i - \lambda_i| \le \mathbf{O}(\mathbf{u})\|A\|_2$. This is sufficient to guarantee good relative accuracy for larger eigenvalues (i.e. for $\lambda_i \approx \lambda_n$), but for smaller eigenvalue (i.e. for $\lambda_i \approx \lambda_1$), we have

$$\frac{|\widehat{\lambda}_i - \lambda_i|}{\lambda_i} \le \mathbf{O}(\mathbf{u})\frac{\lambda_n}{\lambda_i} \approx \mathbf{O}(\mathbf{u})\kappa_2(A)$$

where $\kappa_2(A) = \|A\|_2\|A^{-1}\|_2$. Therefore, the best relative accuracy of the smaller eigenvalues that one can compute depends on the condition number of A.

For the matrix eigenvalue problem arising in discretizations of differential operators, it is usually smaller eigenvalues that well approximate the eigenvalues of the differential operators and are of interest. The finite difference discretization leads to a standard eigenvalue problem $Ax = \lambda x$ and the finite element method results in a generalized eigenvalue problem $Ax = \lambda Bx$, where A (and B) are often

diagonally dominant. With the condition number for the discretized problem A (or $B^{-1}A$) typically large, smaller eigenvalues computed are expected to have a relative accuracy of order $\kappa_2(A)\mathbf{u}$ (or $\kappa_2(B^{-1}A)\mathbf{u}$).

For a second order elliptic differential operator, the condition number of a discretization matrix is typically of order $\mathbf{O}(h^{-2})$, where h is the mesh size. Then the relative errors for smaller eigenvalues computed are expected to be $\mathbf{O}(h^{-2})\mathbf{u}$. On the other hand, for a fourth order elliptic differential operator such as the bi-harmonic operator, the condition number of a discretization matrix is typically of order $\mathbf{O}(h^{-4})$. Their discretization matrices are extremely ill-conditioned even for a modestly small h (i.e. $h = 10^{-4}$) and the smaller eigenvalues computed may have little or no relative accuracy at all.

In this talk, we present our recent works on high relative accuracy algorithms for computing eigenvalues of diagonally dominant matrices. We present a relative perturbation bound for the eigenvalues of symmetric positive definite diagonally dominant matrices demonstrating that that all eigenvalues are determined to the same relative accuracy as in the off-diagonal entries and the diagonally dominant parts [2]. We then describe an algorithm that computes all singular values of a diagonally dominant matrix (and hence all eigenvalues if A is also symmetric positive definite) to high relative accuracy [1]. We further consider using the algorithm in iterative methods for large scale eigenvalue problem and we show smaller eigenvalues of finite difference discretizations of differential operators can be computed accurately. Numerical examples are presented to demonstrate both the accuracy issue of traditional algorithm as well as the high accuracy achieved by the new algorithm.

References

[1] Q. Ye, *Computing SVD of Diagonally Dominant Matrices to high relative accuracy*, Mathematics of Computations, 77 (2008): 2195-2230.

[2] Q. Ye, *Relative Perturbation Bounds for Eigenvalues of Diagonally Dominant Matrices*, SIAM Journal on Matrix Analysis and Applications, 2009

Scaling, Sensitivity and Stability in the Numerical Solution of Quadratic Eigenvalue Problems

Françoise Tisseur

(joint work with N. J. Higham, D. S. Mackey and S. D. Garvey)

The purpose of this work is to emphasize the importance of *scaling* the coefficient matrices of second order systems before numerically computing the eigenvalues via linearization. Our discussion is illustrated with a simple but nontrivial example consisting of a slender beam simply supported at both ends and damped at the midpoint. The equation of motion governing the transverse displacement $u(\mathsf{x}, t)$ of the beam has the form

$$\rho A \frac{\partial^2 u}{\partial t^2} + c(\mathsf{x})\frac{\partial u}{\partial t} + EI\frac{\partial^4 u}{\partial \mathsf{x}^4} = 0, \tag{1}$$

where ρA is the mass per unit length, $c(\mathsf{x}) \geq 0$ represents the external damping and EI is the bending stiffness. The boundary conditions are $u(0,t) = u''(0,t) = 0$ and $u(\mathsf{L},t) = u''(\mathsf{L},t) = 0$, where L is the length of the beam. Making the separation hypothesis $u(\mathsf{x},t) = e^{\lambda t}v(\mathsf{x})$ yields the boundary-value problem for the free vibrations

$$\lambda^2 \rho A v(\mathsf{x}) + \lambda c(\mathsf{x}) v(\mathsf{x}) + EI \frac{d^4}{d\mathsf{x}^4} v(\mathsf{x}) = 0, \tag{2}$$
$$v(0) = v''(0) = v(\mathsf{L}) = v''(\mathsf{L}) = 0.$$

We discretize the boundary-value problem (2) by finite elements using cubic Hermite polynomials as interpolation shape functions. This gives the finite-dimensional quadratic eigenvalue problem (QEP)

$$Q(\lambda)x = (\lambda^2 M + \lambda D + K)x = 0. \tag{3}$$

The resulting mass matrix M and stiffness matrix K are symmetric positive definite by construction and D is symmetric positive semidefinite. As a consequence, the roots of the quadratic equation $x^* Q(\lambda)x = 0$ have nonpositive real parts for all vectors x. This implies that all the eigenvalues of (3) lie in the closed left half plane and the beam problem is (weakly) stable.

The standard approach to the numerical solution of the QEP is to convert the quadratic $Q(\lambda) = \lambda^2 M + \lambda D + K$ into a linear polynomial $L(\lambda) = \lambda X + Y$ of twice the dimension of Q but with the same spectrum. The resulting generalized eigenproblem $L(\lambda)z = 0$ is usually solved by the QZ algorithm for small to medium size problems or by a Krylov method for large sparse problems [1], [9]. A common choice of L in practice is the first companion form, given by

$$C_1(\lambda) = \lambda \begin{bmatrix} M & 0 \\ 0 & I \end{bmatrix} + \begin{bmatrix} D & K \\ -I & 0 \end{bmatrix}. \tag{4}$$

It can be shown that $C_1(\lambda)$ is always a linearization in the sense that it satisfies $E(\lambda)C_1(\lambda)F(\lambda) = \begin{bmatrix} Q(\lambda & 0 \\ 0 & I \end{bmatrix}$ for some $E(\lambda)$ and $F(\lambda)$ with constant, nonzero determinants. This implies that C_1 and Q have the same spectrum. When K and M, respectively, are nonsingular the two pencils

$$L_1(\lambda) = \lambda \begin{bmatrix} M & 0 \\ 0 & -K \end{bmatrix} + \begin{bmatrix} D & K \\ K & 0 \end{bmatrix}, \; L_2(\lambda) = \lambda \begin{bmatrix} 0 & M \\ M & D \end{bmatrix} + \begin{bmatrix} -M & 0 \\ 0 & K \end{bmatrix} \tag{5}$$

are other possible linearizations [7], [8].

We generated the matrices in (3) using the NLEVP MATLAB toolbox [2] via `nlevp('damped_beam',nele)`. We used `nele` $= 100$ beam elements, which result in matrices M, D and K of dimension $n = 200$. The eigenvalues were computed by calling MATLAB's function `eig`, which implements the QZ algorithm, on each of the three linearizations (4)–(5). The first three plots in Figure 1 display those computed eigenvalues having real parts in the interval $[-16, 4]$. Since (4)–(5) are all linearizations of Q, the first three plots should be identical, but in fact they are very different, and none of them correctly displays the real part of the spectrum of Q to visual accuracy. These three plots have one feature in common:

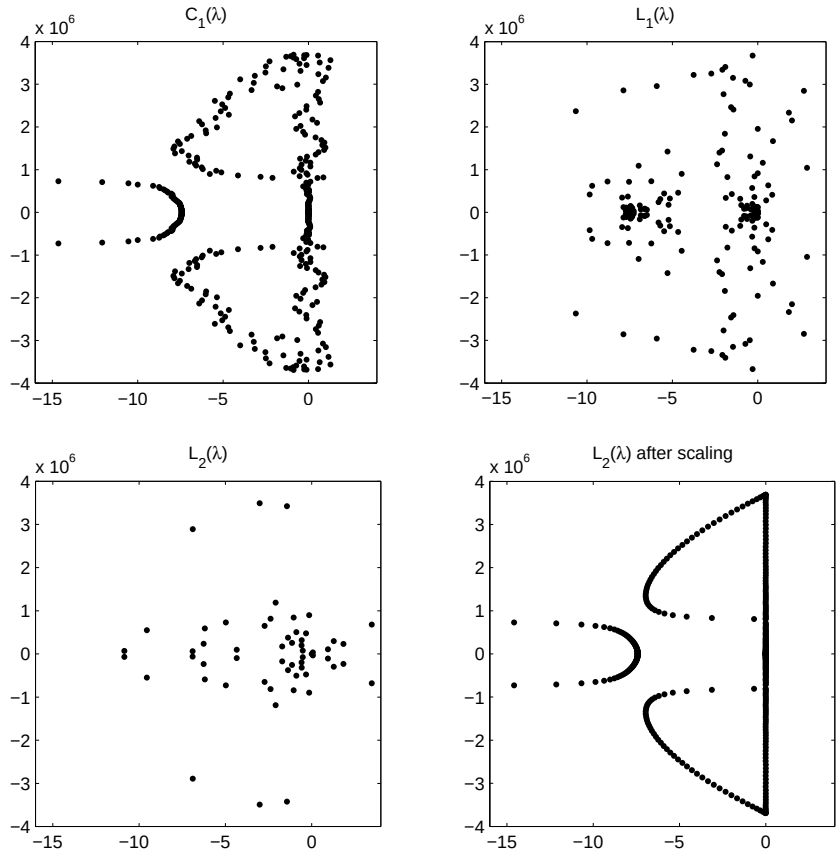

FIGURE 1. Beam problem discretized with 100 finite elements. Computed eigenvalues λ with $\mathrm{Re}(\lambda) \in [-16, 4]$ of the linearizations $C_1(\lambda), L_1(\lambda)$ and $L_2(\lambda)$ defined in (4)–(5).

some eigenvalues lie in the right half plane, therefore implying that the discretized beam problem is unstable and seeming to contradict the theory.

Now let us convert $Q(\lambda) = \lambda^2 M + \lambda D + K$ to $\widetilde{Q}(\mu) = \delta Q(\gamma\mu) = \mu^2 \widetilde{M} + \mu \widetilde{D} + \widetilde{K}$, where $\lambda = \gamma\mu$, $\widetilde{M} = \gamma^2 \delta M$, $\widetilde{D} = \gamma\delta D$, $\widetilde{K} = \delta K$ with $\gamma = \sqrt{\|K\|_2 / \|M\|_2}$, $\delta = 2/(\|K\|_2 + \|D\|_2 \gamma)$. This scaling was proposed by Fan, Lin, and Van Dooren [3]. The eigenvalues μ of $\widetilde{Q}(\mu)$ are then computed by calling `eig` on each of the three linearizations (4)–(5) with the scaled matrices $\widetilde{M}$, $\widetilde{D}$ and $\widetilde{K}$ in place of M, D and K. The eigenvalues of $Q(\lambda)$ are recovered from those of $\widetilde{Q}(\mu)$ via $\lambda = \gamma\mu$. The last plot in Figure 1 shows the spectrum of $Q(\lambda)$ computed using the linearization L_2 after scaling; all three linearizations (4)–(5) yield similar plots after scaling. Note that all the eigenvalues are now in the left half-plane and that many of the eigenvalues appear to be pure imaginary. Indeed we prove in [6] that for the

discretized beam problem half of the eigenvalues of Q are pure imaginary and that they coincide with half of the eigenvalues of the undamped quadratic $\lambda^2 M + K$. It is therefore reasonable to believe — and we are able to conclude from our analysis in [4], [5], [6]— that the fourth plot in Figure 1 is a good approximation of the spectrum of Q, unlike the first three plots.

Our aim is to give a theoretical explanation of these somewhat surprising numerical results and to convince the scientific community of the importance of scaling quadratic eigenvalue problems before computing their eigenvalues via linearizations. Our results are not confined to the beam problem but apply to any quadratic eigenvalue problem.

References

[1] Z. Bai, J. W. Demmel, J. J. Dongarra, A. Ruhe, and H. A. Van der Vorst, eds., *Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide*, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2000.

[2] T. Betcke, N. J. Higham, V. Mehrmann, C. Schröder, and F. Tisseur, *NLEVP: A collection of nonlinear eigenvalue problems.* `http://www.mims.manchester.ac.uk/research/numerical-analysis/nlevp.htm% l`.

[3] H.-Y. Fan, W.-W. Lin, and P. Van Dooren, *Normwise scaling of second order polynomial matrices*, SIAM J. Matrix Anal. Appl., 26 (2004), pp. 252–256.

[4] N. J. Higham, R.-C. Li, and F. Tisseur, *Backward error of polynomial eigenproblems solved by linearization*, SIAM J. Matrix Anal. Appl., 29 (2007), pp. 1218–1241.

[5] N. J. Higham, D. S. Mackey, and F. Tisseur, *The conditioning of linearizations of matrix polynomials*, SIAM J. Matrix Anal. Appl., 28 (2006), pp. 1005–1028.

[6] N. J. Higham, D. S. Mackey, F. Tisseur, and S. D. Garvey, *Scaling, sensitivity and stability in the numerical solution of quadratic eigenvalue problems*, Internat. J. Numer. Methods Eng., 73 (2008), pp. 344–360.

[7] P. Lancaster, *Lambda-Matrices and Vibrating Systems*, Pergamon Press, Oxford, 1966. Reprinted by Dover, New York, 2002.

[8] D. S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann, *Vector spaces of linearizations for matrix polynomials*, SIAM J. Matrix Anal. Appl., 28 (2006), pp. 971–1004.

[9] F. Tisseur and K. Meerbergen, *The quadratic eigenvalue problem*, SIAM Rev., 43 (2001), pp. 235–286.

Some simple error estimates for the Rayleigh-Ritz method

Harry Yserentant

The Rayleigh-Ritz method is a variational method to compute the eigenvalues below the essential spectrum and the corresponding eigenvectors. It has the advantage of being based on minimal, very general assumptions and produces optimal solutions in terms of the approximation properties of the underlying trial spaces. The theory of the Rayleigh-Ritz method has to a large extent been developed in the context of finite element methods, see [1], [2], or [6]. In a recent paper [5], Knyazev and Osborn derived error estimates into which, in contrast to the older theory, only the best approximation error of the just considered eigenfunction enters. Here we show that such estimates can be derived in a very simple way utilizing the stability of certain projection operators, in the case of second order

elliptic problems the H^1-stability of the L_2-projection onto the given ansatz space. We start from the usual abstract framework with two real Hilbert spaces $\mathcal{H}_0$ and $\mathcal{H} \subseteq \mathcal{H}_0$ and a symmetric, coercive, and bounded bilinear form $a : \mathcal{H} \times \mathcal{H} \to \mathbb{R}$. The inner product on $\mathcal{H}_0$ is denoted by (u, v) and the induced norm by $\|u\|_0$. We equip $\mathcal{H}$ with the energy norm $\|u\|$ induced by the bilinear form and assume for simplicity that $\|u\|_0 \leq \|u\|$ for all $u \in \mathcal{H}$. The problem is then to approximate the solutions $(u, \lambda) \in \mathcal{H} \times \mathbb{R}$, $u \neq 0$, of the eigenvalue problem

$$a(u, v) \;=\; \lambda\,(u, v), \quad v \in \mathcal{H}. \tag{1}$$

For this, we select an n-dimensional subspace $\mathcal{S}$ of $\mathcal{H}$. Then there exist pairwise orthogonal normed vectors $u'_1, \ldots, u'_n \in \mathcal{S}$ and real numbers $\lambda'_1, \ldots, \lambda'_n$ with

$$a(u'_k, v) \;=\; \lambda'_k\,(u'_k, v), \quad v \in \mathcal{S}. \tag{2}$$

Without restriction, let $\lambda'_1 \leq \ldots \leq \lambda'_n$. As will be shown, the discrete eigenvalues λ'_k approximate then the original eigenvalues λ and the discrete eigenvectors u'_k the corresponding eigenvectors u in a sense explained later. This already fixes the method, which replicates the weak form of the eigenvalue problem and is determined by the choice of the subspace $\mathcal{S}$ replacing the original solution space. In the case of a second order elliptic eigenvalue problem over the domain Ω, $\mathcal{H}_0 = L_2(\Omega)$, and $\mathcal{H}$ is a subspace of the Sobolev space $H^1(\Omega)$, depending on the boundary conditions. Typical approximation spaces $\mathcal{S}$ are in this case finite element spaces. The approximation properties of the space $\mathcal{S}$ are measured in terms of the a-orthogonal projection operator $P : \mathcal{H} \to \mathcal{S}$ defined by

$$a(Pu, v) = a(u, v), \quad v \in \mathcal{S}. \tag{3}$$

With respect to the energy norm the projection Pu is the best approximation of $u \in \mathcal{H}$ by an element of $\mathcal{S}$, which means that for all $v \in \mathcal{S}$

$$\|u - Pu\| \;\leq\; \|u - v\|. \tag{4}$$

For boundary value problems, Pu is the Ritz-approximation of the solution u.

The classical error estimates are largely based on the min-max principle. Assume for simplicity that the spectrum is purely discrete and that the original eigenvalues are $0 < \lambda_1 \leq \lambda_2 \leq \ldots$. Let $u_1, u_2, u_3, \ldots$ be the corresponding eigenvectors. A typical old-style error estimate reads then as follows:

$$0 \;\leq\; \frac{\lambda'_k - \lambda_k}{\lambda_k} \;\leq\; \frac{4}{\lambda_1}\,\sup\Big\{\,\|u - Pu\| \;\Big|\; u \,=\, \sum_{i=1}^{k} \alpha_i u_i,\ \|u\|_0 = 1\,\Big\}^2. \tag{5}$$

Such estimates depend not only on the quality of the approximation of the eigenvectors assigned to the considered eigenvalue λ_k but also on the size of the approximation error of all eigenvectors for eigenvalues below it. This is not the case for the following error estimate for the eigenvectors that essentially originates from [6].

Theorem 1. *Let $u \in \mathcal{H}$ be an eigenvector for the eigenvalue λ. Then*

$$\Big\|\, u \;-\; \sum_{|\mu'_k - \mu| < r} (u, u'_k)\, u'_k \,\Big\|_0 \;\leq\; \frac{1}{r\lambda}\,\|u - Pu\|_0, \tag{6}$$

where $\mu = 1/\lambda$ and $\mu'_k = 1/\lambda'_k$ has been set and $0 < r \le 1/\lambda$ is arbitrary.

Proof. We first represent the difference to be estimated in the form

$$u \;-\; \sum_{|\mu'_k-\mu|<r} (u,u'_k)\,u'_k \;=\; \sum_{|\mu'_k-\mu|\ge r} (u,u'_k)\,u'_k \;+\; u \;-\; \sum_{k=1}^{n} (u,u'_k)\,u'_k$$

and replace the inner products in the first sum on the right hand side by

$$(u,u'_k) \;=\; \frac{\mu}{\mu-\mu'_k}\,(u-Pu,u'_k).$$

This is possible as u is an eigenvector and the u'_k are discrete eigenvectors. Thus

$$(u,u'_k) \;=\; \lambda^{-1}a(u,u'_k) \;=\; \lambda^{-1}a(u'_k,Pu) \;=\; \lambda^{-1}\lambda'_k\,(Pu,u'_k).$$

The resulting error representation reads in abbreviated form

$$u \;-\; \sum_{|\mu'_k-\mu|<r} (u,u'_k)\,u'_k \;=\; \frac{1}{\lambda}\,R\,(u-Pu) \;+\; (I-P_0)(u-Pu),$$

where the operator R and the $\mathcal{H}_0$-orthogonal projection P_0 onto $\mathcal{S}$ are given by

$$Rf \;=\; \sum_{|\mu'_k-\mu|\ge r} \frac{1}{\mu-\mu'_k}\,(f,u'_k)\,u'_k, \quad P_0 f \;=\; \sum_{k=1}^{n} (f,u'_k)\,u'_k.$$

Expressing the norms in terms of the expansion coefficients in the orthonormal basis of $\mathcal{S}$ consisting of the discrete eigenvectors $u'_1,\dots,u'_n$, one finds

$$\|Rf\|_0^2 \;=\; \sum_{|\mu'_k-\mu|\ge r} \Big|\,\frac{1}{\mu-\mu'_k}\,(f,u'_k)\,\Big|^2 \;\le\; \frac{1}{r^2}\,\|P_0 f\|_0^2.$$

This estimate is used to estimate the first term in the error representation. The proposition follows from the orthogonality properties of the different terms. □

The larger r is chosen, the more discrete eigenvectors u'_k are used to approximate the given eigenvector u and the smaller the error is, but the less specific the relation between the original and the discrete eigenvectors becomes. If the considered eigenvalue λ is sufficiently well separated from its neighbors λ', one can set

$$r \;=\; \frac{1}{2}\,\min_{\lambda'\ne\lambda}\Big|\,\frac{1}{\lambda}-\frac{1}{\lambda'}\,\Big|, \quad \text{i.e.,} \quad \frac{1}{r\lambda} \;=\; 2\,\max_{\lambda'\ne\lambda}\Big|\,\frac{\lambda'}{\lambda'-\lambda}\,\Big|. \tag{7}$$

The smaller r then is, the better the given eigenvalue λ is separated from its neighbors. If the eigenvalue λ belongs to a cluster of closely neighbored eigenvalues, the parameter r should be chosen accordingly and (6) be interpreted as a result on the approximation by an element in the corresponding discrete invariant subspace.

To prove a corresponding estimate for the $\mathcal{H}$-norm of the error, we assume that the $\mathcal{H}_0$-orthogonal projection P_0 onto the ansatz space $\mathcal{S}$ is stable in the energy norm, that is, that there is a constant κ with

$$\|P_0 v\| \;\le\; \kappa\,\|v\|, \quad v\in\mathcal{H}. \tag{8}$$

This constant should be independent of hidden discretization parameters which holds, for example, for certain spectral methods, for wavelets, and in the finite element case, there at least under some restrictions on the underlying grids [3], [4].

Theorem 2. *Let $u \in \mathcal{H}$ be an eigenvector for the eigenvalue λ. Then*

$$\Big\| u \; - \sum_{|\mu'_k - \mu| < r} (u, u'_k)\, u'_k \Big\| \;\le\; \frac{2\kappa + 1}{r\lambda}\, \|u - Pu\|, \tag{9}$$

where $\mu = 1/\lambda$ and $\mu'_k = 1/\lambda'_k$ has been set and $0 < r \le 1/\lambda$ is arbitrary.

Proof. The proof of (9) is based on the same error representation as that of Theorem 1 and transfers almost literally. Particularly it uses the norm estimate

$$\|Rf\|^2 \;=\; \sum_{|\mu'_k - \mu| \ge r} \lambda'_k \Big| \frac{1}{\mu - \mu'_k}\, (f, u'_k) \Big|^2 \;\le\; \frac{1}{r^2}\, \|P_0 v\|^2.$$

The only exception is that in the final step one can no longer argue using the orthogonality properties of the different terms. At this point the bound for the norm of the operator P_0 enters in form of the estimate

$$\|P_0(u - Pu)\| \;\le\; \kappa\, \|u - Pu\|$$

for the projection of the approximation error. □

A similar error estimate holds for the higher eigenvalues, at least for those that are sufficiently well separated from the eigenvalues below them:

Theorem 3. *Let $u \in \mathcal{H}$ be a normed eigenvector for the eigenvalue λ. Assume that $\lambda'_k \ge \lambda$ for all discrete eigenvalues λ'_k in the neighborhood of λ fixed by the condition $|\mu'_k - \mu| < r$, where again $\mu = 1/\lambda$, $\mu'_k = 1/\lambda'_k$, and $0 < r \le 1/\lambda$. Then*

$$\min_{\lambda'_k \ge \lambda} (\lambda'_k - \lambda) \;\le\; \Big(\frac{2\kappa + 1}{r\lambda}\Big)^2\, \|u - Pu\|^2, \tag{10}$$

provided that there is already a discrete eigenvalue $\lambda'_k \ge \lambda$ for which $\lambda'_k - \lambda \le \lambda$.

Proof. Denoting by u' the projection of u from Theorem 1 or Theorem 2 onto the chosen span of discrete eigenvectors,

$$\|u - u'\|^2 \;=\; \lambda\, \|u - u'\|_0^2 \;+ \sum_{|\mu'_k - \mu| < r} (\lambda'_k - \lambda)(u, u'_k)^2.$$

Since the given differences $\lambda'_k - \lambda$ are by assumption nonnegative, this implies

$$\|u - u'\|^2 \;\ge\; \lambda\, \|u - u'\|_0^2 \;+ \min_{\lambda'_k \ge \lambda} (\lambda'_k - \lambda)\, \|u'\|_0^2.$$

Since u' and $u - u'$ are by definition $\mathcal{H}_0$-orthogonal and $\|u\|_0 = 1$, this means

$$\|u - u'\|^2 \;\ge\; \min_{\lambda'_k \ge \lambda} (\lambda'_k - \lambda) \;+\; \lambda \left(1 - \min_{\lambda'_k \ge \lambda} \frac{\lambda'_k - \lambda}{\lambda}\right) \|u - u'\|_0^2.$$

As the second term on the right-hand side of this inequality is by assumption nonnegative, the proposition follows from Theorem 2. □

One can even get rid of the assumption that there is already a discrete eigenvalue $\lambda'_k \geq \lambda$ for which $\lambda'_k - \lambda \leq \lambda$ at the price of a slightly more complicated expression on the right hand side of the error estimate. If there is a discrete eigenvalue $\lambda'_k < \lambda$, the best possible choice for the parameter r is given by

$$\frac{1}{r\lambda} = \max\left\{1, \max_{\lambda'_k < \lambda} \frac{\lambda'_k}{\lambda - \lambda'_k}\right\}. \tag{11}$$

Assuming the energy norm stability (8) of the $\mathcal{H}_0$-orthogonal projection onto the ansatz space, the Rayleigh-Ritz method can thus take full advantage of a higher regularity of the considered eigenvector or eigenfunction compared to the other ones, particularly compared to those for lower eigenvalues. It should further be noted that in the finite-element context one gains, depending on the regularity of the problem, up to one order of approximation in the $\mathcal{H}_0$-norm compared to the $\mathcal{H}$-norm. By Theorem 1 this property transfers to the approximate eigenfunctions.

References

[1] I. Babuška and J.E. Osborn, *Finite element-Galerkin approximation of the eigenvalues and eigenvectors of selfadjoint problems*, Math. Comp. **52** (1989), 275–297.

[2] I. Babuška and J.E. Osborn, *Eigenvalue problems, In: Handbook of Numerical Analysis, Vol. II, Finite Element Methods (Part 1), P.G. Ciarlet and J.L. Lions eds.*, 641–792, Elsevier, Amsterdam 1991.

[3] J.H. Bramble, J.E. Pasciak, and O. Steinbach, *On the stability of the L_2-projection in $H^1(\Omega)$*, Math. Comp. **71** (2001), 147–156.

[4] C. Carstensen, *Merging the Bramble-Pasciak-Steinbach and the Crouzeix-Thomée criterion for H^1-stability of the L_2-projection onto finite element spaces*, Math. Comp. **71** (2001), 157–163.

[5] A.V. Knyazev and J.E. Osborn, *New a priori FEM error estimates for eigenvalues*, SIAM J. Numer. Anal. **43** (2006), 2647–2667.

[6] P.A. Raviart and J.M. Thomas, *Introduction à L'Analyse Numérique des Équations aux Dérivées Partielles*, Masson, Paris, 1983.

Inverse Spectral Computations for Damped Wave Operators

Mark Embree

(joint work with Steven J. Cox)

Can one determine the damping properties of a vibrating string from its spectrum? We investigate this question in the context of the viscously damped wave equation

$$u_{tt}(x,t) = u_{xx}(x,t) - 2a(x)u_t(x,t),$$

posed on $x \in (0,1)$ with $t > 0$, with homogeneous Dirichlet boundary conditions $u(0,t) = u(1,t) = 0$ for all t, and initial data for $u(x,0)$ and $u_t(x,0)$. This equation is conventionally expressed in the form of the evolution problem

$$\begin{bmatrix} u \\ u_t \end{bmatrix}_t = \begin{bmatrix} 0 & I \\ \mathrm{d}^2/\mathrm{d}x^2 & -2a \end{bmatrix} \begin{bmatrix} u \\ u_t \end{bmatrix},$$

which we write as $U_t = AU$. Here $\mathrm{Dom}(A) = H^2(0,1) \cup H_0^1(0,1) \times H_0^1(0,1) \subset H_0^1(0,1) \times L^2(0,1)$, and we endow this space with the energy inner product

$$\left\langle \begin{bmatrix} u \\ v \end{bmatrix}, \begin{bmatrix} f \\ g \end{bmatrix} \right\rangle = \int_0^1 u'(x)\overline{f'(x)} + v(x)\overline{g(x)}\,\mathrm{d}x.$$

Considerable analysis of this operator, its spectrum and evolution behavior, has been reported by Cox and Zuazua [1]. In the absence of damping, A is skew-adjoint with eigenvalues and orthonormal eigenfunctions

$$\lambda_{\pm n} = \pm n\pi i, \qquad V_{\pm n} = g_n(x) \begin{bmatrix} 1 \\ \lambda_{\pm n} \end{bmatrix}, \quad g_n(x) = \sin(n\pi x)$$

for $n = 1, 2, \ldots$. Damping makes the operator nonnormal, and moves eigenvalues into the left half plane. Constant damping, $a(x) \equiv a_0$, changes the eigenvalues to $\lambda_{\pm n} = -a_0 \pm \sqrt{a_0^2 - n^2\pi^2}$; the eigenfunctions maintain the same form.

We consider the variable damping $a(x)$, a modest perturbation of the mean

$$a_0 \equiv \int_0^1 a(x)\,\mathrm{d}x.$$

When a_0 is not an integer multiple of π, all damping functions in a neighborhood of a_0 will give operators A with simple eigenvalues. Extending formulas of Cox and Zuazua [1], we arrive at asymptotic expansions for the eigenvalues and eigenfunctions:

$$\begin{aligned} \lambda_{\pm n} &= \pm in\pi - a_0 \mp \frac{ia_2}{2n\pi} + O(1/n^2) \\ g_n(x) &= \frac{\sinh(\alpha(x) + in\pi x)}{in\pi} + \frac{(a(x) + a(0) - 2a_0)\sinh(\alpha(x) + in\pi x)}{2\pi^2 n^2} \\ &\quad + \frac{(\alpha_2(x) - a_2 x)\cosh(\alpha(x) + in\pi x)}{2\pi^2 n^2} + O(1/n^3), \end{aligned}$$

where $a_2 = \int_0^1 a(t)^2\,\mathrm{d}t$, $\alpha(x) = \int_0^x (a(t) - a_0)\,\mathrm{d}t$, and $\alpha_2(x) = \int_0^x \alpha(t)^2\,\mathrm{d}t$.

These formulas provide the primary vehicle for approximate spectral inversion. We closely follow the strategy described for (self-adjoint) Sturm–Liouville problems by Pöschel and Trubowitz [2]. Consider a parameterized family of damping functions, $a_t(x) = a_0 + t\widetilde{a}(x)$, where $\widetilde{a}(x) \equiv a(x) - a_0$. Formally we have

$$\lambda_n(a_0 + \widetilde{a}) - \lambda_n(a_0) = \int_0^1 \dot{\lambda}_n(a_0 + t\widetilde{a})\,\mathrm{d}t.$$

Implicitly differentiating $AV_n = \lambda_n V_n$ with respect to t and taking the inner product with the left eigenfunction

$$W_n = \overline{g_n(x)} \begin{bmatrix} 1 \\ -\overline{\lambda_n} \end{bmatrix}$$

yields the formula

$$\lambda_n(a_0 + \widetilde{a}) - \lambda_n(a_0) = \int_0^1 \frac{\langle \dot{A}(a + t\widetilde{a})V_n(a_0 + t\widetilde{a}), W_n(a_0 + t\widetilde{a})\rangle}{\langle V_n(a_0 + t\widetilde{a}), W_n(a_0 + t\widetilde{a})\rangle}\,\mathrm{d}t, \qquad (*)$$

where

$$\dot{A}(a + t\widetilde{a}) = \begin{bmatrix} 0 & 0 \\ 0 & \widetilde{a} \end{bmatrix}.$$

Given the spectrum for $a(x) = a_0 + \widetilde{a}(x)$, one can recover the constant a_0 as the dominant real part in the asymptotic formula for λ_n, and hence one knows the left-hand side of $(*)$. We substitute asymptotic formulas for V_n and W_n into the right-hand side of $(*)$ to obtain the rough approximation

$$\mathrm{Re}(\lambda_n(a) - \lambda_n(a_0)) \approx \int_0^1 \widetilde{a}(x) \frac{\sinh(2\alpha(x))}{2\alpha(x)} \cos(2n\pi x)\, \mathrm{d}x.$$

If $a(x)$ only deviates modestly from its mean, a_0, then $\alpha(x)$ will be small, and we can further approximate

$$\mathrm{Re}(\lambda_n(a) - \lambda_n(a_0)) \approx \int_0^1 \widetilde{a}(x) \cos(2n\pi x)\, \mathrm{d}x.$$

If additionally $a(x)$ is even, then so too is $\widetilde{a}$: in this case, this last integral is simply a scaled Fourier coefficient for $\widetilde{a}$, and we can approximate

$$a(x) \approx a_0 + 2 \sum_{n=1}^{\infty} \mathrm{Re}(\lambda_n(a) - \lambda_n(a_0)) \cos(2n\pi x).$$

Computational examples illustrate the ability of this procedure to recover even damping functions from spectral data to decent accuracy, even in the presence of a fair degree of noise.

References

[1] Steven Cox and Ernique Zuazua, *The rate at which energy decays in a damped string*, Comm. Partial Diff. Eqns. **19** (1994), 213–243.

[2] Jürgen Pöschel and Eugene Trubowitz, *Inverse Spectral Theory*, Academic Press, Boston, 1987.

Computation of Interior Eigenvalues: Restarted Nonlinear Arnoldi

Marta M. Betcke

(joint work with Vera Lochmann, Heinrich Voss)

According to [2] the main source of noise generated by vehicles moving at speeds above 50 km/h is the sound radiated by their rolling tires. Therefore, recently a lot of effort has been directed into the development of methods allowing for the simulation of the effect of tire-road surface interaction, with the aim to improve the design of the tires by for instance optimization of the tread pattern.

Simulations of the sound radiation require the solution of the Helmholtz equation

$$\Delta u + k^2 u = 0 \tag{1}$$

on the exterior of the domain occupied by the tire for a particular set of wave numbers $k = \omega_k/c$, where c denotes the speed of sound. The frequencies ω_k are

obtained by Fourier analysis of the transient excitation forces $f(t) \approx \sum_k f_k e^{i\omega_k t}$ such as an impact of the roughness of the road surface or the tread pattern. As the domain is infinite, the Sommerfeld radiation condition is assumed for large r. On the outer boundary of the tire the normal velocities of the tire $\partial u_k/\partial n$ for the particular frequency ω_k and the excitition force f_k

$$-\omega_k^2 M u_k + i\omega_k G u_k + K u_k = f_k \tag{2}$$

are prescribed as a boundary condition. To reduce the cost associated with the solution of (2) for many ω_k, under the assumption of modal superposition, $u_k \approx \tilde{u}_k = X z_k$, an approximation to the normal velocities for the particular frequency ω_k is extracted from the projected equation. In practice only the eigenmodes X with frequencies close to the sought frequency ω_k are used which corresponds to the range up to 2000 Hz.

The extraction of the corresponding eigenspace, X, requires solution of the gyroscopic eigenvalue problem of the form

$$T(\lambda)x := -\lambda^2 M x + i\lambda G x + K x = 0, \tag{3}$$

where the matrices M and K are symmetric positive definite and the skew-symmetric matrix $G = -G^T$ accounts for the presence of gyroscopic forces. Since the system is conservative all its eigenvalues are purely imaginary and of the form $i\lambda$. Therefore all the eigenvalues of $T(\cdot)$ are real and moreover $T(\cdot)$ admits a variational characterisation [5]. Hence, iterative projection methods like Nonlinear Arnoldi or Jacobi Davidson [3, 4] can be used to compute the eigenvalues safely one after another. Such methods thought hit their limitations if a large number of eigenvalues or a set of some subsequent eigenvalues in the interior of the spectrum is required. In order to preserve the numbering the dimension of the search space has to be at least as large as the number of eigenvalues preceding the sought one. Therefore the size of the projected problem is growing with the number of the wanted eigenvalue, which results in increasing computation time consumed by the nonlinear solver and increasing storage requirements.

To overcome these difficulties we propose a restarted variants of these methods using a local numbering, which does not require to include the entire set of preceding eigenvectors or the corresponding part of the invariant subspace of $T(\lambda)$ into the search subspace after a restart [1].

Assume that we are given an eigenvalue $\hat{\lambda}$ of the nonlinear eigenproblem (3), which we call an anchor, and a corresponding eigenvector $\hat{x}$. Let $\mathcal{V}$ be a subspace of $\mathbb{C}^n$ that contains $\hat{x}$, and let the columns of V form a basis of $\mathcal{V}$. Then $\hat{\lambda}$ is also an eigenvalue of the projected problem

$$T_V(\hat{\lambda})y := V^H T(\hat{\lambda}) V y = 0, \tag{4}$$

and since $T_V(\cdot)$ has the same structure that $T(\cdot)$ it also admits variational characterisation i.e. we can assign to $\hat{\lambda}$ a local number $\ell = \ell(\mathcal{V})$ in the following way: $\hat{\lambda}$ is an ℓth eigenvalue of problem (4) if $\mu(\hat{\lambda}) = 0$ is the ℓ largest eigenvalue of the

linear problem

$$V^H T(\hat{\lambda}) V y = \mu(\hat{\lambda}) y. \tag{5}$$

Starting with $\mathcal{V} =: \mathcal{V}_0$ we determine approximations to the eigenvalue subsequent to the anchor $\hat{\lambda}$ projecting problem (3) to a sequence of subspaces $\mathcal{V}_0 \subset \mathcal{V}_1 \subset \mathcal{V}_2 \subset \ldots$ which are generated in course of an iterative projection method aiming at the $(\ell(\mathcal{V}_k)+1)$th eigenvalue in the kth iteration step. Explicitly stating the dependence of ℓ on $\mathcal{V}_k$ we emphasize that the number $\ell(\mathcal{V}_k)$ of the anchor may change in the course of the iteration.

After convergence we may continue the iterative projection method aiming at the $(\ell(\mathcal{V}_k)+2)$th eigenvalue or we may replace the anchor by the newly converged eigenpair. Since the current search space contains useful information about further eigenvalues it is advisable to continue expanding the search spaces until the convergence has become too slow or the dimension exceeds a given bound.

Once we have the local numbering there is no necessity any more to include all the eigenvectors corresponding to the preceding eigenvalues or the invariant subspace of $T(\hat{\lambda})$ corresponding to its nonnegative eigenvalues into the search space after a restart. All that we need to set up the new search subspace is an eigenvector $\hat{x}$ corresponding to an anchor $\hat{\lambda}$ and an approximation v_1 to the next eigenvector (or a random vector if such an approximation is not at hand). This leads to the restart framework in Algorithm 1.

Algorithm 1 Restart Framework

Require: Preconditioner $M \approx T(\sigma)^{-1}$ for a suitable shift σ,
Require: (λ_i, x_i) an (approximate) eigenpair of $T(\cdot)$
Require: v_1 an approximation to x_{i+1}
$V = [x_i, v_1]$;
j = 1;
while Restart condition not satisfied **do**
  repeat
    Determine largest eigenvalues $\mu_1(\lambda_i) \geq \cdots \geq \mu_k(\lambda_i) > 0 \geq \mu_{k+1}(\lambda_i)$ of (5)
    Set $\ell := k$ if $\mu_k \leq -\mu_{k+1}$, and else $\ell := k+1$
    Compute $(\ell+j)$th eigenpair $(\tilde{\lambda}_{\ell+j}, y_{\ell+j})$ of $T_V(\cdot)$
    Expand V aiming at $(\lambda_{\ell+j}, x_{\ell+j})$
  until Eigenpair $(\tilde{\lambda}_{\ell+j}, V y_{l+j}) =: (\lambda_{i+j}, x_{i+j})$ converged
  j = j+1;
end while

The described restart technique suffers from spurious eigenvalues temporarily arising in the search subspace and being linear combinations of the eigenvalues outside of the part of the spectrum covered by the actual search subspace. This phenomenon is intrinsic to higher frequency computations, when not the entire eigensubspace corresponding to all eigenvalues from the lower end of the spectrum can be maintained in the search subspace. The occurrence of the spurious eigenvalues surfaces through a repeated convergence of the algorithm to the same

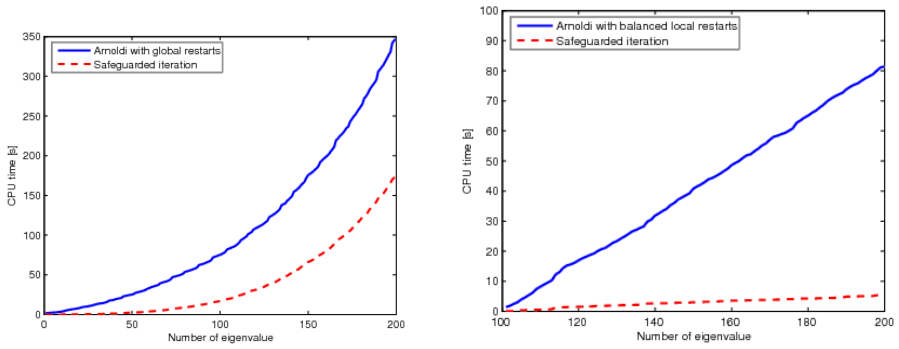

FIGURE 1. Nonlinear Arnoldi (left), Restarted Nonlinear Arnoldi (right) for a model rubber wheel problem 1728 DOF.

eigenvalue. Though this can happen for other reasons as well, with simple tests it is possible to isolate the case of arising of a spurious eigenvalue and drive it out from the search subspace by a dedicated iteration.

REFERENCES

[1] M.M. Betcke, *Iterative Projection Methods for Symmetric Nonlinear Eigenvalue Problems with Applications*, PhD Thesis Hamburg University of Technology **2007**, ISBN-10: 3866242832

[2] M. Brinkmeier, U. Nackenhorst , O. von Estorff, S. Petersen, *Physically Based Modeling of Tire-Rolling-Noise by a Finite Element Approach*, Proceedings of Inter-noise 2004

[3] T. Betcke, H. Voss, *A Jacobi-Davidson-type projection method for nonlinear eigenvalue problems*, Future Generation Computer Systems, **20(3)** (2004), 363 – 372

[4] H. Voss, *An Arnoldi method for nonlinear symmetric eigenvalue problems*, Online Proceedings of the SIAM Conference on Applied Linear Algebra, Williamsburg, `http://www.siam.org/meetings/laa03/`

[5] H. Voss, B. Werner, *A Minimax Principle for Nonlinear Eigenvalue Problems with Applications to Nonoverdamped Systems*, Math. Meth. Appl. Sci., **4** (1982), 415–424

Eigenvalue Computations with Nonpolynomial Basis Functions

TIMO BETCKE
(joint work with A. H. Barnett, L. N. Trefethen)

Let $\Omega \subset \mathbb{R}^2$ be a bounded, simply connected domain. The accurate computation of Laplacian eigenvalues λ_j and corresponding eigenfunctions u_j satisfying

$$-\Delta u_j = \lambda_j u_j \quad \text{in } \Omega \tag{1a}$$

$$u_j = 0 \quad \text{on } \partial\Omega \tag{1b}$$

is an important and fascinating problem in a wide range of areas, such as energy states in quantum systems, acoustic computations or recently in the fingerprinting of images. In pure mathematics much attention has been devoted to this problem

in the last century influenced by Kac's famous 1966 article "Can one hear the shape of a drum?" [12], in which it was asked whether there exist two nontrivially distinct domains that have the same spectrum. Indeed, this question was first answered by Gordon, Webb and Wolpert in 1992 [9] by constructing such isospectral drums. Beautiful computations of the first few eigenvalues and eigenfunctions of isospectral drums are given by Driscoll in [7]. A survey about Laplacian eigenvalue problems is given by Kuttler and Sigillito in [14].

But to date the accurate computation of eigenvalues and eigenfunctions remains a challenge. This is due to the fact that the eigenfunctions u_j become highly oscillatory compared with the global length scale of Ω as j becomes large. Standard linear finite element methods need to resolve this oscillatory behavior leading to discrete systems that grow at least with $O(\lambda)$. Furthermore, the computation of large eigenvalues for the discretized problem is still a major challenge for numerical solvers.

In recent years nonpolynomial finite element methods have become more and more attractive for the solution of Helmholtz problems with large wavenumbers. The idea is to use locally on each element $\Omega_i \subset \Omega$ a basis of functions that already satisfy the Helmholtz equation, such as Fourier-Bessel functions or plane waves. This is closely related to the method of particular solutions, which for eigenvalue problems has been popularized in the Numerical Analysis Community in the 1966 paper by Fox, Henrici and Moler [8]. A brief overview of the history of this method and a stable numerical implementation can be found in [3].

The principle idea is the following. Denote by $\mathcal{A}(\lambda)$ a space of basis functions $g_\ell \in \mathcal{C}^2(\Omega) \cap \mathcal{C}(\overline{\Omega})$, $\ell = 1, \ldots, N$ satisfying $-\Delta g_\ell = \lambda g_\ell$ in Ω. Now define the function

$$t(\lambda) = \min_{u \in \mathcal{A}(\lambda)} \frac{\|u\|_{L^2(\partial\Omega)}}{\|u\|_{L^2(\Omega)}}.$$

This function measures the smallest achievable boundary error from a space of particular solutions for a given value λ. In [13] it was shown that there exists $C > 0$ and an eigenvalue λ_j of (1) such that

$$\frac{|\lambda - \lambda_j|}{\lambda_j} \leq C t(\lambda).$$

Hence, by minimizing $t(\lambda)$ with respect to λ we can obtain accurate eigenvalue estimates. In the physics literature this approach is also known in the context of numerical methods for quantum billiards [10, 11, 1, 2].

The evaluation of $t(\lambda)$ can be efficiently and numerically stably done using the generalized singular value decomposition [3, 5]. The great advantage compared to finite elements is that the computational effort of this method only grows with $O(\sqrt{\lambda})$ since it is a boundary based approach. Furthermore, it is simple to approximate all eigenvalues in a given interval $[\lambda_{min}, \lambda_{max}]$ just by evaluating $t(\lambda)$ for sufficiently many points in that interval. The disadvantage compared to finite elements is that each function evaluation makes one GSVD computation necessary, while finite elements can find several eigenvalues with one matrix decomposition.

For more complicated geometries it is often of advantage not to use global basis functions but to do a domain decomposition and to work with a set of basis functions on each subdomain. Let $\overline{\Omega} = \bigcup_i \overline{\Omega_i}$ be a decomposition of Ω into subdomains Ω_i. Define local approximation spaces

$$\mathcal{A}_i(\lambda) := \left\{ \sum_{\ell=1}^{N_i} c_\ell g_\ell^{(i)} |\ c_\ell \in \mathbb{R}, g_\ell \in \mathcal{C}^2(\Omega_i) \cap \mathcal{C}(\overline{\Omega_i}) \right\}.$$

We define the global space $\mathcal{A}(\lambda)$ such that $u \in \mathcal{A}(\lambda)$ if and only if $u|_{\Omega_i} \in \mathcal{A}_i(\lambda)$. Define $\mathcal{T}(\lambda, u)$ by

$$\mathcal{T}(\lambda, u) = \sum_{i<j} \int_{\Gamma_{ij}} \lambda |[u]|^2 ds + |[\partial_\nu u]|^2 ds + \sum_i \int_{\partial\Omega_i \cap \partial\Omega} |u|^2 ds,$$

where Γ_{ij} is the interface between two elements Ω_i and Ω_j and $[u]$, respectively $[\partial_\nu u]$ are the jumps of u and its normal derivative across Γ_{ij}. Furthermore, define

$$\mathcal{M}(\lambda) := \sum_i \|u\|^2_{L^2(\Omega_j)}.$$

Then a generalization of the case of global basis functions is given by the error functional

$$t(\lambda) := \min_{u \in \mathcal{A}(\lambda)} \left(\frac{\mathcal{T}(\lambda, u)}{\mathcal{M}(\lambda, u)} \right)^{1/2}.$$

One again has an error bound of the form $\frac{|\lambda - \lambda_j|}{\lambda_j} \leq C t(\lambda)$ for some $C > 0$.

This idea goes back to Descloux and Tolley in [6]. In [4] a stable implementation based on the generalized singular value decomposition and a convergence analysis based on Vekua theory are given. If basis functions are used that are adapted to potential corner singularities of the eigenfunctions then these nonpolynomial methods can be designed to have excellent exponential convergence properties leading to accurate eigenvalue computations on a wide range of planar domains.

References

[1] A. H. Barnett, *Dissipation in deforming chaotic billiards*, Ph.D thesis, Department of Physics, Harvard University, 2000.

[2] A. H. Barnett and T. Betcke, *Quantum mushroom billiards*, Chaos **17** (2008), 043125.

[3] T. Betcke and L. N. Trefethen, *Reviving the method of particular solutions*, SIAM Rev. **47** (2005), 469–491.

[4] T. Betcke, *A GSVD formulation of a domain decomposition method for planar eigenvalue problems*, IMA J. Numer. Anal. **27** (2007), 451–478.

[5] T. Betcke, *The generalized singular value decomposition and the method of particular solutions*, SIAM J. Sci. Comp. **30** (2008), 1278–1295.

[6] J. Descloux and M. Tolley, *An accurate algorithm for computing the eigenvalues of a polygonal membrane*, Comput. Methods Appl. Mech. Engrg. **39** (1983), 37–53.

[7] T. A. Driscoll, *Eigenmodes of isospectral drums*, SIAM Rev. **39** (1997), 1–17.

[8] L. Fox, P. Henrici, and C. B. Moler, *Approximations and bounds for eigenvalues of elliptic operators*, SIAM J. Numer. Anal. **4** (1967), 89–102.

[9] C. Gordon, G. Webb, and S. Wolpert, *Isospectral plane domains and surfaces via Riemannian orbifolds*, Invent. Math. **110** (1992), 1–22.

[10] E. J. Heller, *Bound-state eigenfunctions of classically chaotic Hamiltonian systems: Scars of periodic orbits*, Phys. Rev. Lett. **53** (1984), 1515–1518.
[11] E. J. Heller, *Wavepacket dynamics and quantum chaology*, In M. J. Giannoni, A. Voros, and J. Zinn-Justin, editors, *Proceedings of the 1989 Les Houches summer school on "Chaos and Quantum Physics"*, pp. 547–663, North Holland, 1991. Elsevier Science Publishers B. V.
[12] M. Kac, *Can one hear the shape of a drum?*, Amer. Math. Monthly **73** (1966), 1–23.
[13] J. R. Kuttler and V. G. Sigillito, *Bounding eigenvalues of elliptic operators*, SIAM J. Math. Anal. **9** (1978), 768–773.
[14] J. R. Kuttler and V. G. Sigillito, *Eigenvalues of the Laplacian in two dimensions*, SIAM Rev. **26** (1984), 163–193.

Compatible Discretization of the Helmholtz Equation

Joachim Schöberl
(joint work with Maria Rechberger, Thorsten Hohage, Mark Ainswoth, Werner Koch, Stefan Hain)

We consider the Helmholtz equation on unbounded domains. A radiation condition is formulated by means of the PML method. We are looking for resonances, i.e., discrete points of the spectrum. It is well known that the PML method causes artificial eigenvalues already on the continuous level. We show that additional spurious eigenvalues occur from the finite element discretization. We propose a finite element method involving projection operators to avoid this spurious eigenvalues on structured mesh.

A computer-assisted existence proof for photonic band gaps

Michael Plum
(joint work with Vu Hoang and Christian Wieners)

Photonic crystals are optical materials in which light with frequency in certain ranges - the so-called photonic band gaps - cannot propagate, while other frequencies can. Mathematically they are usually modelled by Maxwell's equations (in the frequency domain)

$$\begin{aligned} &\operatorname{curl} E = -i\omega H,\ \operatorname{curl} H = i\omega\varepsilon E, \\ &\operatorname{div}(\varepsilon E) = 0,\ \operatorname{div} H = 0 \ \ (\text{on } \mathbb{R}^3) \end{aligned} \tag{1}$$

for the electric field E and the magnetic field H, with $\varepsilon : \mathbb{R}^3 \to \mathbb{R}$ denoting the electric permittivity, a bounded function bounded positively away from zero, which for a photonic crystal is moreover periodic on $\mathbb{R}^3$. ω is the frequency of a monochromatic light wave sent into the crystal, and mathematically it takes the role of a spectral parameter in (1). If ω^2 is in the *spectrum* of (a suitable self-adjoint operator theoretical realization of) problem (1), the light wave can propagate in the crystal; if ω^2 is in the resolvent set, the light wave is absorbed.

Applying curl $(\cdot)$ to the first and curl $\left(\frac{1}{\varepsilon}\cdot\right)$ to the second equation in (1), we obtain the second-order problems

$$\operatorname{curl}\operatorname{curl} E = \lambda\varepsilon E,\ \operatorname{div}(\varepsilon E) = 0 \text{ on } \mathbb{R}^3 \tag{2}$$

and

$$\operatorname{curl}\left(\frac{1}{\varepsilon}\operatorname{curl} H\right) = \lambda H,\ \operatorname{div} H = 0 \text{ on } \mathbb{R}^3, \tag{3}$$

where $\lambda = \omega^2$. Specializing to the 2D-situation $\varepsilon = \varepsilon(x_1, x_2)$ and to TM-polarized waves, where $E = (0, 0, u)$, the second equation in (2) tells that $u = u(x_1, x_2)$, and the first reads

$$-\Delta u = \lambda\varepsilon u \text{ on } \mathbb{R}^2. \tag{4}$$

Assume now for simplicity that the periodicity cell of ε is $\Omega = (0, 1)^2$. By Floquet-Bloch theory, the spectrum σ of (4) is given as a union

$$\sigma = \bigcup_{n=1}^{\infty} I_n \tag{5}$$

of compact real intervals I_n, and in turn each I_n is characterized by

$$I_n = \left\{\lambda_{k,n} \ :\ k \in [-\pi, \pi]^2\right\}, \tag{6}$$

where $\lambda_{k,n}$ is the n-th eigenvalue of the semiperiodic eigenvalue problem

$$\begin{aligned} -\Delta u &= \lambda\varepsilon u \text{ in } \Omega, \\ u(x + z) &= e^{ik\cdot z} u(x)\ (x \in \Omega,\ z \in \mathbb{Z}^2) \end{aligned} \tag{7}$$

depending on $k \in [-\pi, \pi]^2$.

Usually, the *spectral bands* I_n in (5) overlap, but it may happen that gaps open between them. These are the band gaps of prohibited frequencies mentioned earlier, and it is of great practical interest to know if they are present or not. Analytically, this question is very difficult to decide.

We propose a computer-assisted approach for proving the existence of photonic band gaps for specific materials, i.e. for specific permittivity functions ε. We use the characterization (5) - (7) for this purpose, but we have to overcome the problem that *infinitely many* eigenvalue problems (7) have to be treated due to their dependence on $k \in [-\pi, \pi]^2$.

For this purpose, we first choose a finite *grid* G in $[-\pi, \pi]^2$, and compute enclosures for $\lambda_{k,1}, \cdots \lambda_{k,N}$ (with $N \in \mathbb{N}$ chosen fixed) for all $k \in G$: First we use a Finite Element discretization and Knyazev's LOBPCG method, with multigrid preconditioning, to compute approximations $\tilde{u}_{k,1}, \ldots, \tilde{u}_{k,N}$ to the first N eigenfunctions. Next we apply variational eigenvalue enclosure techniques, more particularly the Rayleigh-Ritz method (for upper bounds) and the Lehmann-Goerisch method (for lower bounds), supplemented by a homotopy algorithm connecting problem (7) to a simple constant coefficients problem which can be solved in closed form.

Note that only the *small*-scale $(N \times N)$ matrix eigenvalue problems arising in the Rayleigh-Ritz and in the Lehmann-Goerisch method need to be solved in verified form (using e.g. interval arithmetic).

Next, we use an explicit *perturbation* argument for self-adjoint operators to capture *all* $k \in [-\pi, \pi]^2$. Here we briefly describe the idea. Let $[a, b] \subset \mathbb{R}$ be an interval such that, for some $n \in \mathbb{N}$, some $k \in G$, and some $\delta_k > 0$,

$$\lambda_{k,n} + \delta_k \leq a < b \leq \lambda_{k,n+1} - \delta_k, \tag{8}$$

whence $[a, b]$ is contained in the resolvent set of (7). Moreover let $k + h$ be a perturbation of k, with $|h| < r_k$, where

$$r_k = \sqrt{\varepsilon_{\min}} \left[\sqrt{\lambda_{k,n+1} + \delta_k} - \sqrt{\lambda_{k,n+1}} \right].$$

Then, it can be shown that $[a, b]$ is contained in the resolvent set of the perturbed problem (7), with k replaced by $k + h$.

Let (8) hold for *all* gridpoints $k \in G$, and suppose that

$$\bigcup_{\text{gridpoints } k \in G} \text{Ball}(k, r_k) \supset [-\pi, \pi]^2.$$

Then, it follows directly that $[a, b]$ is contained in a spectral gap.

We applied the method to the case where $\varepsilon(x) = 1$ for $x \in \left[\frac{1}{16}, \frac{15}{16}\right]^2$ (corresponding to vacuum in this part of Ω) and $\varepsilon(x) = 5$ in the rest of Ω. We could prove that the interval

$$[18.2\ ,\ 18.25]$$

is contained in a gap between the third and the fourth spectral band; see Figure 1 for illustration. The proof required to choose about 100 grid points k in the Brillouin zone $[-\pi, \pi]^2$, and altogether about 5000 eigenvalue enclosures (many of which are needed within the homotopy algorithm mentioned above), which needed about 90 hours computing time.

References

[1] M. Plum, V. Hoang, Ch. Wieners: A computer-assisted proof for photonic band gaps. Z. angew. Math. Phys., DOI 10.1007/s00033-008-8021-2 (2009); www.birkhauser.ch/ZAMP.

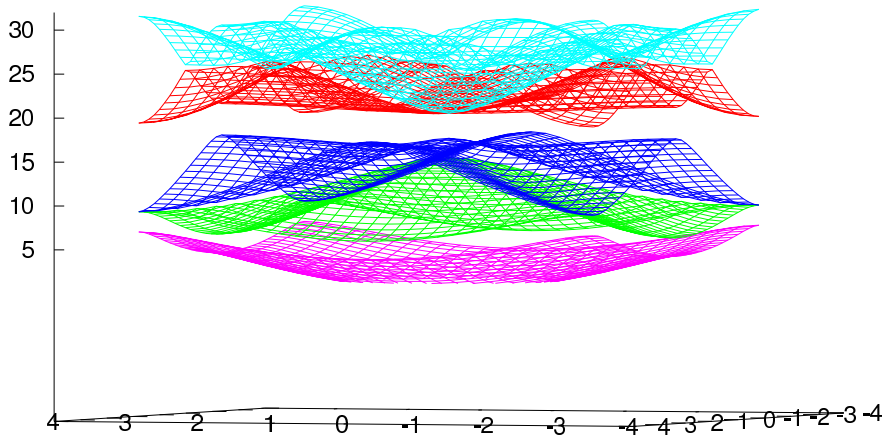

FIGURE 1. Eigenvalues $\lambda_{k,1}, ..., \lambda_{k,5}$ as functions of $k \in [-\pi, \pi]^2$.

Dealing with several eigenvalues simultaneously in nonlinear eigenvalue problems

DANIEL KRESSNER

We consider eigenvalue problems that are nonlinear in the eigenvalue parameter:

$$T(\lambda)x = 0, \tag{1}$$

where the entries of the $n \times n$ matrix T are analytic functions in λ. In most applications, it is of interest to compute a few eigenvalues in a specified region of the complex plane along with the corresponding eigenvectors. For example, when computing band gap diagrams for photonic crystals a modest number of real or nearly real eigenvalues need to be determined for several, smoothly varying nonlinear eigenvalue problems.

When dealing with several eigenvalues of (1) simultaneously it is crucial to find a good representation for these eigenvalues and the corresponding eigenvectors. This task turns out to be surprisingly intricate. In contrast to the linear case, there may be eigenvector/eigenvalue pairs (λ_1, x_1), ..., (λ_k, x_k) of (1), for which the eigenvalues $\lambda_1, \ldots, \lambda_k$ are pairwise distinct but $\{x_1, \ldots, x_k\}$ is linearly dependent. This possibility is already evident from the fact that k can be larger than n. Another well-known example is given by

$$T(\lambda) = \begin{bmatrix} 0 & 12 \\ -2 & 14 \end{bmatrix} + \lambda \begin{bmatrix} -1 & -6 \\ 2 & -9 \end{bmatrix} + \lambda^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \tag{2}$$

for which the eigenvalues 3 and 4 share the same eigenvector $\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]$. The occurrence of such linear dependencies is an annoyance when attempting to develop numerical methods for computing more than one eigenvalue of (1). For example, standard Newton methods for the simultaneous computation of several eigenvalues crucially

depend on the existence of a basis for the invariant subspace belonging to the eigenvalues of interest. In methods that determine several eigenvalues subsequently, such as Krylov subspace or Jacobi-Davidson methods, repeated convergence towards an eigenvalue is usually avoided by reorthogonalization against converged eigenvectors. If such an idea was directly applied to nonlinear eigenvalue problems, eigenvalues could be missed due to linear dependencies among eigenvectors.

Extending joint work with Timo Betcke (U Reading) on polynomial eigenvalue problems, we will show how the concept of invariant pairs can be used to address the difficulties mentioned above. For this purpose, it will be more convenient to reformulate the nonlinear eigenvalue problem (1) in the form

$$\big(f_1(\lambda)A_1 + f_2(\lambda)A_2 + \cdots + f_m(\lambda)A_m\big)x = 0. \tag{3}$$

for analytic functions $f_1, \ldots, f_m : \Omega \to \mathbb{C}$ and constant matrices $A_1, \ldots, A_m \in \mathbb{C}^{n\times n}$. A pair $(X, S) \in \mathbb{C}^{n\times k} \times \mathbb{C}^{k\times k}$ is called *invariant* w.r.t. the nonlinear eigenvalue problem (3) if

$$A_1 X f_1(S) + A_2 X f_2(S) + \cdots + A_m X f_m(S) = 0. \tag{4}$$

As a normalization condition we impose

$$W^H V_l(X, S) - I = 0 \tag{5}$$

for some matrix W and

$$V_l(X, S) = \begin{bmatrix} X \\ XS \\ \vdots \\ XS^{l-1} \end{bmatrix} \tag{6}$$

for some sufficiently large integer l. The crucial point is to note that $V_l(X, S)$ has full rank under rather mild assumptions and hence imposing a condition of the form (5) is reasonable (in contrast to imposing a condition of the form $W^H X = I$). The main result of our work shows that (X, S) is a regular point of the set of nonlinear matrix equations (4)–(6). This implies that (X, S) is a well-posed object and can be computed numerically. In fact Newton's method can be applied in a rather straightforward manner to (4)–(6). The use of the resulting numerical method is illustrated for the computation of band gap diagrams mentioned above.

References

[1] T. Betcke and D. Kressner, *Perturbation, Computation and Refinement of Invariant Pairs for Matrix Polynomials*, Technical report 2009-21, Seminar for applied mathematics, ETH Zurich, July 2009.

[2] D. Kressner. *A block Newton method for nonlinear eigenvalue problems.* Technical report 2009-05, Seminar for applied mathematics, ETH Zurich, February 2009. Revised July 2009. To appear in Numerische Mathematik.

Eigenvalues of fuzzy symmetric matrices

Karl Meerbergen
(joint work with Jeroen De Vlieger)

This talk reports work in progress on the eigenvalues of parameterized matrices.

Complex mathematical PDE models are used for simulating physical models. Most simulations are currently deterministic, i.e. all physical parameters (thickness, temperature, ...) are fixed. In practice, however, uncertainties about the precise values of parameters make deterministic solutions less attractive. In particular, one may want to validate a solution chosen by an automatic method, that minimizes a cost function. Often, the engineers still want to have an idea on the behaviour of the system near the optimal parameter values.

Uncertainties can be modeled in various ways. Stochastic methods are used for dealing with irreducible uncertainties due to the production process e.g. [5]. For an earlier stage in the design process, stochastic data about the parameters are not known. The parameter ranges can be large and a statistical analysis is not appropriate. The goal is to reduce the level of uncertainty. In order to study the impact of the parameters, recently, fuzzy numbers are used [4]. A sensitivity analysis of the parameter values (using derivatives) is only an option for small changes of the parameters. In this talk, we study the situation where parameters may vary significantly w.r.t. the chosen parameter values. The goal is to evaluate the variation of the solution among the given parameter values.

In the study of vibrations, sources of uncertainties can be geometric or material parameters, such as thickness, Young modulus, etc. An eigenvalue analysis is often performed in order to understand vibrations. The first eigenfrequency (smallest eigenvalue) should lie outside the frequency range for which the system is used, in order to avoid undesired vibrations. We therefore focus on the computation of the smallest eigenvalue. After finite element discretization, the algebraic eigenvalue problem takes the form $Ku = \lambda Mu$, where K and M are symmetric, and M is positive definite. In this talk, we only consider standard eigenvalue problems. As we now from the literature, $Ku = \lambda Mu$ can be written as a symmetric standard eigenvalue problem.

Pseudospectra can be used for a sensitivity analysis. The bounds provided by the pseudospectrum might be too large, since the specific parameter structure is not taken into account. When only one parameter changes, the structured pseudospectrum is useful. In this talk, we use fuzzy numbers as a generalization of a single parameter perturbation.

The notion fuzzy number is an extension of fuzzy set theory [6]. We use a tilde to denote fuzzy variables. We represent a fuzzy number $\widetilde{x}$ by its membership function $\mu_{\widetilde{x}}(x) : \mathbb{R} \to [0, 1]$. It defines the level of membership of x to the fuzzy set: $\mu_{\widetilde{x}}(x) = 1$ implies that x belongs to the set and $\mu_{\widetilde{x}}(x) = 0$ means that x does not belong to the set. For example, the concept "x is near 1" could be represented by the fuzzy number with the membership function that is one for $x = 1$, zero for $x < 0.5$ and $x > 1.5$ and values between 0 and 1 for $0.5 \leq x \leq 1.5$. The α cut is

defined as the set of x's for which $\mu_{\widetilde{x}}(x) \geq \alpha$. When all α-cuts are intervals, $\widetilde{x}$ is called a convex fuzzy number.

The extension principle is used to extend a (deterministic) function $f(x)$ with $x \in \mathbb{R}^N$ to fuzzy numbers. Let $\widetilde{x}$ be an N-vector of convex fuzzy numbers and f be continuous, then the extremes of the α-cuts of $\widetilde{y} = f(\widetilde{x})$ can be computed as:

$$[\widetilde{y}]_\alpha^- = \min_{x_i \in [\widetilde{x}_i]_\alpha} f(x_1, \ldots, x_n) \tag{1a}$$

$$[\widetilde{y}]_\alpha^+ = \max_{x_i \in [\widetilde{x}_i]_\alpha} f(x_1, \ldots, x_n)\ . \tag{1b}$$

There are three classes of methods for computing these extremes. The most popular one is the montecarlo method, which evaluates random samples. This method is not only expensive, it may also be inaccurate. The most accurate class solve the extremes by numerical optimization.

Fuzzy eigenvalues have been mentioned first for non-negative matrices [1][2], where all matrix entries can be independent fuzzy numbers. In this talk, we consider real symmetric matrices of the following form. Define

$$A(x) = B + \sum_{i=1}^{N} x_i E_i$$

where B, and E_i, $i = 1, \ldots, N$ are real symmetric. If $\widetilde{x}$ is a fuzzy number, we call $A(\widetilde{x})$ a fuzzy matrix. Following the extension principle, the eigenvalues of $A(\widetilde{x})$ are also fuzzy numbers.

We talk about the computation of the smallest or the largest eigenvalues (which is basically the same operation). Several approaches can be used for validating a selection of parameter values: the pseudo-spectrum computes the variation for all parameters together regardless the form of the perturbation. With the structure preserving pseudo-spectrum, we can handle one parameter (or a linear combination of parameters) at a time; this corresponds to the vertex method for fuzzy numbers, which can only be used in specific cases. In this talk, we analyse the cost needed for solving structured perturbations in the context of symmetric eigenvalue problems.

In order to solve the optimization problems (1) efficiently, we derive properties of the eigenvalues of A as a function of x where x lies in an α-cut of $\widetilde{x}$. Indeed, since function evaluations are eigenvalue computations, it may be important to reduce the number of function evaluations as much as possible. The key observation is that the largest eigenvalue of $A(x)$ is a convex function in x. This also holds when eigenvalues are multiple. If the elements of $\widetilde{x}$ are independent from each other, the α-cuts are hypercubes. The consequence is that the maximum is attained in one of the cornerpoints of the hypercube. This implies the solution of 2^N eigenvalue problems for each α-cut. The minimum is the local minimum of a (non-smooth) convex optimization problem. A similar conclusion holds for the smallest eigenvalue.

Details on the theory and the optimization algorithm can be found in [3].

References

[1] J. J. Buckley. Fuzzy eigenvalues and input-output analysis. *Fuzzy Sets Syst.*, 34(2):187–195, 1990.

[2] K.-P. Chiao. Generalized fuzzy eigenvalue problems. *Tamsui Oxford Journal of Mathematical Sciences*, 14:31–37, 1998.

[3] J. De Vlieger. *De berekening van de grootste vage eigenwaarde van onzekere symmetrische matrices.* Master thesis, K.U.Leuven, Department of Computer Science, Leuven, Belgium, 2009, 2009. (In English: The computation of the largest fuzzy eigenvalue of uncertain symmetric matrices).

[4] W. Oberkampf, S. DeLand, B. Rutherford, K. Diegert, and K. Alvin. A new methodology for the estimation of the total uncertainty in computational simulation. In *proceedings of 40th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*, pages 3061–3083, St. Louise, MO, USA, 1999.

[5] G.I. Schuëller. A state-of-the-art report on computational stochastic mechanics. *Probabilistic Engineering Mechanics*, 12(4):197–321, 1997.

[6] L.A. Zadeh. Fuzzy sets. *Information and Control*, 8(3):338–353, 1965.

Dynamical Systems and Non-Hermitian Iterative Eigensolvers

Richard B. Lehoucq

(joint work with Mark Embree)

Suppose we seek a small number of eigenvalues (and the associated eigenspace) of the non-Hermitian matrix A, having at our disposal a nonsingular matrix B that approximates A. Given a starting vector p_0, compute

$$p_{j+1} = p_j + B^{-1}(\theta_j - A)p_j, \tag{1}$$

where $\theta_j - A$ is shorthand for $I\theta_j - A$, and

$$\theta_j = \frac{(Ap_j, p_j)}{(p_j, p_j)}$$

for some inner product $(\cdot, \cdot)$. Knyazev, Neymeyr, and others have studied this iteration for Hermitian positive definite A; see [13, 14] and references therein for convergence analysis and numerical experiments.

Clearly the choice of N will influence the behavior of this iteration. With $N = A$, the method (1) reduces to (scaled) inverse iteration:

$$p_{j+1} = As^{-1}p_j\theta_j.$$

We are interested in the case where N approximates A, yet one can apply N^{-1} to a vector much more efficiently than A^{-1} itself. Such a N acts as a preconditioner for A, and, hence, (1) represents a preconditioned iteration.

This method contrasts with a different class of algorithms, based on inverse iteration (or the shift-invert Arnoldi algorithm), that apply a preconditioner to accelerate an "inner iteration" that approximates the solution to a linear system at each step; see, e.g., [16, 8, 10]. For numerous practical large-scale non-Hermitian eigenvalue problems, such as those described in [17, 24], these inner iterations can be extremely expensive and highly dependent on the quality of the preconditioner.

In contrast, as we shall see, the iteration (1) can converge to a leftmost eigenpair even when N is a suitable multiple of the identity.

This paper provides a rigorous convergence theory that establishes sufficient conditions for (1) to converge to the leftmost eigenpair for non-Hermitian A. We obtain these results by viewing this iteration as the forward Euler discretization of the autonomous nonlinear differential equation

$$\dot{p} = N^{-1}\left(p\frac{(Ap,p)}{(p,p)} - Ap\right) \tag{2}$$

with a unit step size. Here A and N are fixed but p depends on a parameter, t; $\dot{p}$ denotes differentiation with respect to t. In the absence of preconditioning, the differential equation (2) has been studied in connection with power iteration [5, 20], as described in more detail below. The nonzero steady-states of this system correspond to (right) eigenvectors of A, and, hence, one might attempt to compute eigenvalues by driving this differential equation to steady-state as swiftly as possible. Properties of the preconditioner determine which of the eigenvectors is an attracting steady-state.

The differential equation (2) enjoys a distinguished property, observed, for example, in [5, 20] with $N = I$. Suppose that p solves (2), $\theta = (p,p)^{-1}(Ap,p)$, and N is self-adjoint and invertible (A may be non-self-adjoint). Then for all t,

$$\begin{aligned}\frac{d}{dt}(p,Np) &= \big(N^{-1}(p\theta - Ap), Np\big) + \big(p, NN^{-1}(p\theta - Ap)\big)\\ &= (p\theta,p) - (Ap,p) + (p,p\theta) - (p,Ap)\\ &= 0. \end{aligned} \tag{3}$$

Thus, (p,Np) is an *invariant* (or *first integral*), as its value is independent of time; see [12, section 1.3] for a discussion of the unpreconditioned case ($N = I$), and, e.g., [11] for a general introduction to invariant theory and geometric integration.

The invariant describes a manifold in n-dimensional space, $(p,Np) = (p_0,Np_0)$, on which the solution to the differential equation with $p(0) = p_0$ must fall. Simple discretizations, such as Euler's method (1), do not typically respect such invariants, giving approximate solutions that drift from the manifold. Invariant-preserving alternatives (see, e.g., [11]) generally require significantly more computation per step (though a tractable method for the unpreconditioned, Hermitian case has been proposed by Nakamura, Kajiwara, and Shiotani [19]). Our goal is to explain the relationship between convergence and stability of the continuous and discrete dynamical systems. In particular, the quadratic invariant is a crucial property of the continuous system, and plays an important role in the convergence theory of the corresponding discretization, even when that iteration does not preserve the invariant.

The present study draws upon this body of work, but takes a different perspective: we seek a better understanding of iterations such as (1) that provide only *approximate* solutions (with a truncation error due to discretization) to continuous time systems such as (2). The distinction is significant: for example, a continuous-time generalization of the power method will converge, with mild caveats, to the

largest magnitude eigenvalue, whereas the related systems we study can potentially converge to the leftmost eigenvalue at a shift-independent rate with little more work per iteration than the power method.

References

[1] P.-A. Absil, *Continuous-time systems that solve computational problems*, Int. J. Uncov. Comput., 2 (2006), pp. 291–304.

[2] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*, Princeton University Press, Princeton, NJ, 2008.

[3] P.-A. Absil, R. Sepulchre, and R. Mahony, *Continuous-time subspace flows related to the symmetric eigenproblem*, Pacific J. Optim., 4 (2008), pp. 179–194.

[4] Z. Bai, D. Day, and Q. Ye, *ABLE: An adaptive block Lanczos method for non-Hermitian eigenvalue problems*, SIAM J. Matrix Anal. Appl., 20 (1999), pp. 1060–1082.

[5] M. T. Chu, *Curves on s^{n-1} that lead to eigenvalues or their means of a matrix*, SIAM J. Alg. Disc. Math., 7 (1986), pp. 425–432.

[6] M. T. Chu, *On the continuous realization of iterative processes*, SIAM Rev., 30 (1988), pp. 375–387.

[7] M. Embree, *The Arnoldi eigenvalue iteration with exact shifts can fail*, SIAM J. Matrix Anal. Appl., to appear.

[8] M. A. Freitag and A. Spence, *Convergence theory for inexact inverse iteration applied to the generalised nonsymmetric eigenvalue problem*, Electron. Trans. Numer. Anal., 28 (2007), pp. 40–64.

[9] G. H. Golub and L.-Z. Liao, *Continuous methods for extreme and interior eigenvalue problems*, Linear Algebra Appl., 415 (2006), pp. 31–51.

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

[11] E. Hairer, C. Lubich, and G. Wanner, *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*, 2nd ed., Springer-Verlag, Berlin, 2006.

[12] U. Helmke and J. B. Moore, *Optimization and Dynamical Systems*, Springer, London, 1994.

[13] A. V. Knyazev, *Preconditioned eigensolvers—an oxymoron?*, Elec. Trans. Numer. Anal., 7 (1998), pp. 104–123.

[14] A. V. Knyazev and K. Neymeyr, *Efficient solution of symmetric eigenvalue problems using multigrid preconditioners in the locally optimal block conjugate gradient method*, Elec. Trans. Numer. Anal., 7 (2003), pp. 38–55.

[15] A. V. Knyazev and K. Neymeyr, *A geometric theory for preconditioned inverse iteration. III: A short and sharp convergence estimate for generalized eigenvalue problems*, Linear Algebra Appl., 358 (2003), pp. 95–114.

[16] Y.-L. Lai, K.-Y. Lin, and W.-W. Lin, *An inexact inverse iteration for large sparse eigenvalue problems*, Numer. Linear Algebra Appl., 1 (1997), pp. 1–13.

[17] R. B. Lehoucq and A. J. Salinger, *Large-scale eigenvalue calculations for stability analysis of steady flows on massively parallel computers*, Internat. J. Numer. Methods Fluids, 36 (2001), pp. 309–327.

[18] R. Mahony and P.-A. Absil, *The continuous time Rayleigh quotient flow on the sphere*, Linear Algebra Appl., 368 (2003), pp. 343–357.

[19] Y. Nakamura, K. Kajiwara, and H. Shiotani, *On an integrable discretization of the Rayleigh quotient gradient system and the power method with a shift*, J. Comput. Appl. Math., 96 (1998), pp. 77–90.

[20] T. Nanda, *Differential equations and the QR algorithm*, SIAM J. Numer. Anal., 22 (1985), pp. 310–321.

[21] B. N. Parlett, *The Symmetric Eigenvalue Problem*, no. 20 in Classics in Applied Mathematics, SIAM, Philadelphia, 1998. Amended reprint of 1980 Prentice-Hall edition.

[22] B. T. Polyak, *Introduction to Optimization*, Translation Series in Mathematics and Engineering, Optimization Software, Inc., New York, 1987.
[23] W. W. Symes, *The QR algorithm and scattering for the finite nonperiodic Toda lattice*, Physica D, 4 (1982), pp. 275–280.
[24] J. S. Warsa, T. A. Wareing, J. E. Morel, J. M. McGhee, and R. B. Lehoucq, *Krylov subspace iterations for deterministic k-eigenvalue calculations*, Nuc. Sci. Engrg., 147 (2004), pp. 26–42.
[25] D. S. Watkins, *Isospectral flows*, SIAM Rev., 26 (1984), pp. 379–391.

Computing the Eigenvalues of Hierarchical Matrices by LR-Cholesky Transformations

Peter Benner
(joint work with Thomas Mach)

We investigate the use of Rutishauser's LR-Cholesky transformation [8] to compute all eigenvalues of symmetric hierarchical ($\mathcal{H}$-) matrices. Historically, the LR transformation can be considered to be the first algorithm of the class of GR algorithms [11] and is based on the iteration

$$\begin{aligned} L_{m+1}L_{m+1}^T &= M_m - \mu_m I, \\ M_{m+1} &= L_{m+1}^T L_{m+1} = L_{m+1}^{-1} M_m L_{m+1}. \end{aligned} \tag{1}$$

The hierarchical matrix format, see, e.g., [6, 7, 4], allows storing a variety of dense matrices from certain applications in a special data-sparse way with linear-polylogarithmic complexity. Most $\mathcal{H}$-arithmetic operations, including the $\mathcal{H}$-Cholesky decomposition and the $\mathcal{H}$-QR decomposition [1], have linear-polylogarithmic complexity, too. As for simpler structured matrices, like semiseparable matrices, GR algorithms exist [10], it is plausible that an eigenvalue algorithm of bulge-chasing type might also exist for $\mathcal{H}$-matrices. For GR algorithms, typically $\mathcal{O}(n)$ iterations are sufficient, so we expect to find an algorithm for $\mathcal{H}$-matrices with a complexity of $\mathcal{O}(n^2 \log^\alpha n)$. The quest for such an algorithm poses merely an academic challenge, but it might become useful in a variety of applications requiring many (small, inner) eigenvalues of differential operators discretized by boundary- or finite-element methods.

The first straightforward approach is to substitute the Cholesky decomposition in the LR-Cholesky transformation (1) by the $\mathcal{H}$-Cholesky decomposition. This yields a meaningful LR-Cholesky algorithm in $\mathcal{H}$-matrix arithmetic, since the limit of the iterative process is a diagonal matrix and hence an $\mathcal{H}$-matrix, too. But some of the intermediate iterates may not have good $\mathcal{H}$-matrix approximations. If we multiply two $\mathcal{H}$-matrices, the rank of the product is bounded by (see [5])

$$k_{\mathcal{H}\cdot\mathcal{H}} \leq C_{id} C_{sp} (p+1) k_{\mathcal{H}},$$

with C_{id} the idempotency constant, C_{sp} the sparsity constant and p the depth of the $\mathcal{H}$-tree of the involved $\mathcal{H}$-matrices. This means, that the maximal rank of the matrix M_m can grow really fast with each iteration step and they actually do in general, see Figure 1. This can be regarded as the $\mathcal{H}$-arithmetic analog to fill-in.

Large blocks of full rank destroy the data sparsity of the hierarchical matrix and increase the complexity to $\mathcal{O}(n^3)$.

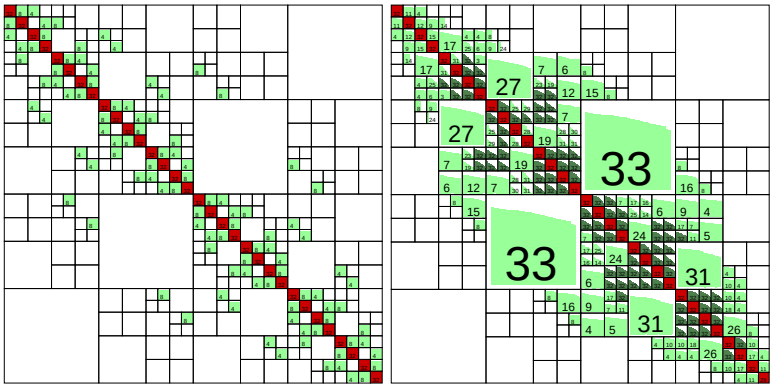

FIGURE 1. Matrix FEM32 (left) and the same matrix after 10 steps (with shift) of the LR Cholesky transformation (right), dark green and red (dark grey) blocks have full rank.

Matrix	Dimension	Time	Accuracy
FEM8	64	0.04	9.4e-6
FEM16	256	3.20	1.1e-2
FEM32	1024	243.05	4.2e-2
FEM64	4096	11,116.96	2.6e-1

TABLE 1. Computation times for LR Cholesky transformation.

In GR algorithms it is necessary to shift the matrix M_m to accelerate the convergence. In the case of LR-Cholesky transformations, the shift must preserve the positive definiteness of M_m. There are shift strategies ensuring this by Rutishauser [9] and Wilkinson [12]. We use some steps of inverse power iteration to find an approximation of the smallest eigenvalue and subtract a safety margin to be sure having a shift smaller than the smallest eigenvalue.

If the smallest eigenvalue λ_n is found, we have to deflate this eigenvalue. After the deflation we can increase the shift, since the smallest eigenvalue of the deflated matrix is $\lambda_{n-1} \geq \lambda_n$. There is a second type of deflation. In the QR algorithm for Hessenberg matrices, this occurs if a subdiagonal element becomes small enough. In our case there is no Hessenberg or band structure, so this deflation occurs if the submatrix $M(j+1:n, 1:j)$ is (close to) zero. Then we can divide the matrix into two smaller matrices. The spectrum of M_m is the union of the spectrum of the two smaller matrices $M(1:j, 1:j)$ and $M(j+1:n, j+1:n)$.

Table 1 shows the needed CPU-time if we use the LR-Cholesky transformation together with the described shift and deflation strategies for four $\mathcal{H}$-matrices of

different size. These matrices present the finite-element discretisations of the 2D-Laplacian with 8 to 64 discretisation points in each direction. We see that the CPU-time grows slightly slower than $\mathcal{O}(n^3)$, but faster than $\mathcal{O}(n^2 \log^2 n)$. Also, the accuracy degrades already for small n. Hence, current research focuses on improving both the shift and the deflation strategies in order to improve both aspects of the LR-Cholesky transformation for $\mathcal{H}$-matrices in order to find a sequence of iterates that have better $\mathcal{H}$-matrix approximations than those obtained from the current implementation. The existence of such a "path of iterates" within the class of $\mathcal{H}$-matrices is an open problem at this point.

The hierarchical semi-separable matrices (HSS matrices, [2]) form a subset of the $\mathcal{H}$-matrices. The $\mathcal{H}$-tree of an HSS matrices have only admissible knots or knots of the form $s \times s$. Further the admissible blocks satisfy the condition, that $A_{t\times s}$ is in the linear span of $[0; A_{t_2\times t_1}]$, with $s < t$ and $t = t_1 \cup t_2$. The LR-Cholesky transformation for HSS matrices produces no $\mathcal{H}$-fill-in which suggests the existence of a GR-type algorithm for this subset of the $\mathcal{H}$-matrices.

For practical purposes eigenvalue algorithms for computing a small subset of the spectrum are interesting, too. There is a well known $\mathcal{H}$-matrix-vector product of linear-polylogarithmic complexity. This product is sufficient to implement the power method:

$$y_k = Ax_k, \quad x_k = \frac{y_k}{\|y_k\|_2},$$

or the Jacobi-Davidson algorithm [3]. As in standard arithmetic, the Jacobi-Davidson algorithm converges faster than the power method and permits the computation of inner eigenvalues. Current and future work focus on a detailed analysis of these methods and will also include the use of $\mathcal{H}$-matrices in preconditioned eigensolvers.

Acknowledgments. The work of Thomas Mach is supported by a grant of the Free State of Saxony (501-G-209).

References

[1] P. Benner and T. Mach. On the QR decomposition of $\mathcal{H}$-matrices. Chemnitz Scientific Computing Preprints CSC/09-04, Technische Universität Chemnitz, 2009.

[2] S. Chandrasekaran, M. Gu, and W. Lyons. A fast adaptive solver for hierarchically semiseparable representations. *Calcolo*, 42:171–185, 2005.

[3] D. Fokkema, G. Sleijpen, and H. V. der Vorst. Jacobi-Davidson style QR and QZ algorithms for the reduction of matrix pencils. *SIAM J. Sci. Comput.*, 20(1):94–125, 1998.

[4] S. Goreinov, E. Tyrtyshnikov, and A. Yeremin. Matrix-free iterative solution strategies for large dense linear systems. *Numer. Lin. Alg. Appl.*, 4(4):273–294, 1997.

[5] L. Grasedyck. *Theorie und Anwendungen Hierarchischer Matrizen*. PhD thesis, Mathematisch-Naturwissenschaftlichen Fakultät, Christian-Albrechts-Universität Kiel, July 2001.

[6] W. Hackbusch. A sparse matrix arithmetic based on $\mathcal{H}$-matrices. Part I: Introduction to $\mathcal{H}$-matrices. *Computing*, 62(2):89–108, 1999.

[7] W. Hackbusch and B. Khoromskij. A sparse $\mathcal{H}$-matrix arithmetic Part II: Application to multi-dimensional problems. *Computing*, 64:21–47, 2000.

[8] H. Rutishauser. Solution of eigenvalue problems with the LR-transformation. *Nat. Bur. Standards Appl. Math. Ser.*, 49:47–81, 1958.

[9] H. Rutishauser. Über eine kubisch konvergente Variante der LR-Transformation. *ZAMM Z. Ange. Math. Mech.*, 40(1-3):49–54, 1960.
[10] R. Vandebril, M. Van Barel, and N. Mastronardi. *Matrix Computations and Semiseparable Matrices. Vol. II. Eigenvalue and Singular Value Methods.* Johns Hopkins University Press, Baltimore, MD, 2008.
[11] D. Watkins. QR-like algorithms for eigenvalue problems. *J. Comput. Appl. Math.*, 123:67–83, 2000.
[12] J. Wilkinson. *The Algebraic Eigenvalue Problem.* Clarendon Press, Oxford, first edition, 1965.

Reporter: Agnieszka Międlar

Participants

Prof. Dr. Peter Arbenz
Chair of Computational Science
ETH Zürich
CAB G 69.3
Universitätsstr. 6
CH-8092 Zürich

Prof. Dr. Uday Banerjee
Dept. of Mathematics
Syracuse University
Syracuse , NY 13244-1150
USA

Prof. Dr. Peter Benner
Fakultät für Mathematik
Technische Universität Chemnitz
09107 Chemnitz

Prof. Dr. Jeffrey K. Bennighof
University of Texas at Austin
Department of Aerospace Engineering &
Engineering Mechanics
1 University Station C0600
Austin , TX 78712
USA

Prof. Dr. Michele Benzi
Dept. of Mathematics and
Computer Science
Emory University
400, Dowman Dr.
Atlanta , GA 30322
USA

Dr. Marta M. Betcke
School of Mathematics
The University of Manchester
Oxford Road
GB-Manchester M13 9PL

Dr. Timo Betcke
Department of Mathematics
University of Reading
Whiteknights
P.O.Box 220
GB-Reading RG6 6AX

Prof. Dr. Susanne C. Brenner
Dept. of Mathematics
Louisiana State University
Baton Rouge LA 70803-4918
USA

Prof. Dr. Tucker Carrington
Chemistry Department
Queen's University
Kingston, Ontario K7L 3N6
CANADA

Prof. Dr. Carsten Carstensen
Institut für Mathematik
Humboldt-Universität zu Berlin
Unter den Linden 6
10099 Berlin

Prof. Dr. Mark Embree
Computational and Applied Math.
Rice University
6100 Main Street
Houston , TX 77005-1892
USA

Dr. Jean-Luc Fattebert
Center for Applied Scientific Computing
Lawrence Livermore National Laboratory
P.O.Box 808, L-561
Livermore CA 94550
USA

Dr. Melina Freitag
Department of Mathematical Sciences
University of Bath
Claverton Down
GB-Bath BA2 7AY

Dipl.-Math. Joscha Gedicke
Institut für Mathematik
Humboldt-Universität Berlin
Rudower Chaussee 25
10099 Berlin

Dr. Stefano Giani
Dept. of Mathematics
The University of Nottingham
University Park
GB-Nottingham NG7 2RD

Dr. Lars Grasedyck
Max-Planck-Institut für Mathematik
in den Naturwissenschaften
Inselstr. 22 - 26
04103 Leipzig

Dr. Luka Grubisic
Department of Mathematics
University of Zagreb
Bijenicka 30
10000 Zagreb
CROATIA

Prof. Dr. Wolfgang Hackbusch
Max-Planck-Institut für Mathematik
in den Naturwissenschaften
Inselstr. 22 - 26
04103 Leipzig

Prof. Dr. Ulrich Hetmaniuk
Department of Applied Mathematics
Box 352420
University of Washington
Seattle , WA 98195-2420
USA

Prof. Dr. Ralf Hiptmair
Seminar für Angewandte Mathematik
ETH-Zentrum
Rämistr. 101
CH-8092 Zürich

Prof. Dr. Yunqing Huang
Department of Mathematics
Xiangtan University
411105 Xiangtan
P.R.China

Prof. Dr. Andrew Knyazev
Department of Mathematics
University of Colorado at Denver
Campus Box 170
Denver , CO 80217-3364
USA

Prof. Dr. Daniel Kressner
Departement Mathematik
ETH-Zentrum
Rämistr. 101
CH-8092 Zürich

Prof. Dr. Mats G. Larson
Dept. of Mathematics
University of Umea
S-901 87 Umea

Prof. Dr. Richard B. Lehoucq
Sandia National Laboratories
Computational Mathematics and
Algorithms Department
MS 1320, P.O.Box 5800
Albuquerque , NM 87185-1320
USA

Dr. Karl Meerbergen
Departement Computerwetenschappen
Katholieke Universiteit Leuven
Celestijnenlaan 200A
B-3001 Heverlee

Prof. Dr. Volker Mehrmann
Institut für Mathematik
Technische Universität Berlin
Sekr. MA 4-5
Strasse des 17. Juni 136
10623 Berlin

Agnieszka Miedlar
Institut für Mathematik
Sekr. MA 4-5
Technische Universität Berlin
Straße des 17. Juni 136
10623 Berlin

Prof. Dr. Klaus Neymeyr
Fachbereich Mathematik
Universität Rostock
Universitätsplatz 1
18055 Rostock

Prof. Dr. Yvan Notay
Universite Libre de Bruxelles
Service de Metrologie Nucleaire
(CP) 165-84)
50 av. F.D.Roosevelt
B-1050 Bruxelles

Prof. Dr. John Osborn
Department of Mathematics
University of Maryland
College Park , MD 20742-4015
USA

Prof. Dr. Beresford Parlett
Department of Mathematics
University of California
Berkeley , CA 94720-3840
USA

Prof. Dr. Joseph E. Pasciak
Department of Mathematics
Texas A & M University
College Station , TX 77843-3368
USA

Prof. Dr. Michael Plum
Universität Karlsruhe
Institut für Analysis
76128 Karlsruhe

Prof. Dr. Rolf Rannacher
Institut für Angewandte Mathematik
Universität Heidelberg
Im Neuenheimer Feld 294
69120 Heidelberg

Thorsten Rohwedder
Fakultät II -Institut f. Mathematik
Technische Universität Berlin
Sekr. MA 5-3
Straße des 17. Juni 136
10623 Berlin

Prof. Dr. Joachim Schöberl
Center for Computational
Engineering Science
RWTH Aachen
Pauwelstraße 19
52065 Aachen

Dr. Imbo Sim
Ecole Polytechnique Federale de Lausanne
Institut d'analyse et de calcul scientif
MAC 2614
CH-1015 Lausanne

Prof. Dr. Danny C. Sorensen
Department of Computational &
Applied Mathematics
Rice University
6100 Main St., MS 134
Houston TX 77005-1892
USA

Prof. Dr. Alastair Spence
Department of Mathematical Sciences
University of Bath
Claverton Down
GB-Bath BA2 7AY

Prof. Dr. Zdenek Strakos
Institute of Computer Science
Czech Academy of Sciences
Pod vodarenskou vezi 2
18207 Prague 8
CZECH REPUBLIC

Prof. Dr. Francoise Tisseur
Department of Mathematics
The University of Manchester
Oxford Road
GB-Manchester M13 9PL

Prof. Dr. David S. Watkins
Department of Mathematics
Washington State University
Pullman , WA 99164-3113
USA

Prof. Dr. Jinchao Xu
Department of Mathematics
Pennsylvania State University
University Park , PA 16802
USA

Prof. Dr. Qiang Ye
Department of Mathematics
University of Kentucky
715 Patterson Office Tower
Lexington , KY 40506-0027
USA

Prof. Dr. Harry Yserentant
Institut für Mathematik
Technische Universität Berlin
Straße des 17. Juni 136
10623 Berlin

Prof. Dr. Aihui Zhou
Institute of Computational Mathematics
and Scientific/Engineering Computing
Academy of Mathematics and Systems
Sc.
Chinese Academy of Sciences
Beijing 100190
P.R.CHINA

Dr. Lin Zschiedrich
Konrad-Zuse-Zentrum für
Informationstechnik Berlin (ZIB)
Takustr. 7
14195 Berlin