Numerical approximation
of the spectrum of self-adjoint operators
and operator preconditioning

Zdeněk Strakoš
Charles University, Prague
Jindřich Nečas Center for Mathematical Modelling
Based on a joint work with
Tomáš Gergelits, Kent-André Mardal, and Bjørn Fredrik Nielsen

GATIPOR, Paris, June 2022
A problem with bounded invertible operator $\mathcal{G}$ on an infinite dimensional Hilbert space $V$

$$\mathcal{G} u = f$$

is approximated on a finite dimensional subspace $V_n \subset V$ by a problem with the finite dimensional operator

$$\mathcal{G}_n u_n = f_n,$$

represented, using an appropriate basis of $V_n$, by the matrix problem

$$A x = b.$$
A problem with bounded invertible operator $G$ on an infinite dimensional Hilbert space $V$

$$G u = f$$

is approximated on a finite dimensional subspace $V_n \subset V$ by a problem with the finite dimensional operator

$$G_n u_n = f_n,$$

represented, using an appropriate basis of $V_n$, by the matrix problem

$$A \mathbf{x} = \mathbf{b}.$$

“There is a continuous operator equation posed in infinite-dimensional spaces that underlines the linear system of equations [...] awareness of this connection is key to devising efficient solution strategies for the linear systems.” Hiptmair (2006)
(Infinite dimensional) Krylov subspace methods implicitly construct at the step \( j \) the finite dimensional approximation \( G_j \) of \( G \) which determines the desired approximate solution \( u_j \in u_0 + \mathcal{K}_j(G, r), \quad r = f - Gu_0 \)

\[
u_j := u_0 + p_{j-1}(\lambda) r \approx u = G^{-1}f.\]

Here \( p_{j-1}(\lambda) \) is the associated polynomial of degree at most \( j - 1 \) and \( G_j \) is obtained by restricting and projecting \( G \) onto the \( j \)th Krylov subspace

\[
\mathcal{K}_j(G, r) := \text{span}\{r, Gr, \ldots, G^{j-1}r\}.
\]

A.N. Krylov (1931), Gantmakher (1934), Hestenes and Stiefel (1952), Lanczos (1952-53); Karush (1952), Hayes (1954), Stesin (1954), Vorobyev (1958)
From
\[ r_j = f - \mathcal{G} u_j = r - \mathcal{G} p_{j-1}(\mathcal{G}) r =: \varphi_j(\mathcal{G}) r \]

we get the approximation polynomial

\[ \varphi_j(\lambda) = 1 - \lambda p_{j-1}(\lambda), \]

which is nonlinear both in \( \mathcal{G} \) (obvious) and \( f \) (through the orthogonality/optimality property defining the particular method). Clearly

\[ \varphi_n^M(0) = 1. \]

Preconditioning goes much beyond conditioning

Operator preconditioning

\[ G = B^{-1} A \]

where \( A, B : V \to V^\# \) are bounded linear operators on an infinite dimensional Hilbert space \( V \), with its dual \( V^\# \), and \( B \) is, in addition, also self-adjoint with respect to the duality pairing and coercive.
Preconditioning goes much beyond conditioning

Operator preconditioning

\[ \mathcal{G} = \mathcal{B}^{-1} \mathcal{A} \]

where \( \mathcal{A}, \mathcal{B} : V \rightarrow V^\# \) are bounded linear operators on an infinite dimensional Hilbert space \( V \), with its dual \( V^\# \), and \( \mathcal{B} \) is, in addition, also self-adjoint with respect to the duality pairing and coercive.

Spectral and norm equivalence may guarantee mesh (parameter) independence, but they do not necessarily provide computational efficiency.

Faber, Manteuffel and Parter (1990), ... , Hiptmair (2006), Málek and S (2015)
Consider an infinite dimensional Hilbert space $V$, its dual $V^\#$, and bounded linear operators $A, B : V \to V^\#$ that are self-adjoint with respect to the duality pairing, and $B$ is, in addition, also coercive. Consider further a sequence of subspaces $\{V_n\}$ of $V$ satisfying the standard approximation property

$$\lim_{n \to \infty} \inf_{v \in V_n} \|w - v\| = 0 \quad \text{for all } w \in V.$$ 

Note that this typically yields that Galerkin discretizations of boundary value problems are convergent.
Theorem.

Let the sequences of matrices \( \{A_n\} \) and \( \{B_n\} \) be defined via the standard Galerkin discretization. Then all points in the spectrum of the preconditioned operator

\[ B^{-1}A : V \rightarrow V \]

are approximated to an arbitrary accuracy by the eigenvalues of the preconditioned matrices in the sequence \( \{B_n^{-1}A_n\} \).

That is, for any point \( \lambda \in \text{sp}(B^{-1}A) \) and any \( \epsilon > 0 \), there exists \( n^* \) such that for all \( n \geq n^* \) the preconditioned matrix \( B_n^{-1}A_n \) has an eigenvalue \( \lambda_{j(n)} \) satisfying

\[ |\lambda - \lambda_{j(n)}| < \epsilon. \]

1. Spectral information and convergence of the conjugate gradient method.


6. Spectral approximation of operators and/or PDE eigenvalue problem.
Any self-adjoint operator $\mathcal{G}$ defined on $V$ can be expressed in terms of the Riemann-Stieltjes integral as

$$\mathcal{G} = \int \lambda \, dE(\lambda), \quad \text{i.e.} \quad (\mathcal{G}u, v) = \int \lambda \, d(E(\lambda)u, v) \text{ for all } u, v \in V,$$

The spectrum of $\mathcal{G}$ is defined as the complement of the resolvent set, i.e.,

$$\text{sp}(\mathcal{G}) = \{ \lambda \in \mathbb{R}; \lambda I - \mathcal{G} \text{ does not have a bounded inverse} \}.$$

The distribution function $\omega(\lambda)$ is defined by $\mathcal{G}$ and the normalized initial residual $r, \|r\| = 1$ as

$$(\mathcal{G}r, r) = \int \lambda \, d(E(\lambda)r, r) = \int \lambda \, d\omega(\lambda).$$
\{\lambda_i, y_i\} \text{ are the eigenpairs of } \mathbf{A}, \quad \omega_i = |(y_i, w_1)|^2, \quad (w_1 = r_0/\|r_0\|)
The key concept: CG as the Gauss-Christoffel quadrature.

At any iteration step \( j \), CG represents the matrix formulation of the \( j \)-point Gauss quadrature of the Riemann-Stieljes integral determined by \( A \) and \( r_0 \),

\[
\int_0^\infty \phi(\lambda) \, d\omega(\lambda) = \sum_{i=1}^{j} \omega_i^{(j)} \phi(\theta_i^{(j)}) + R_j(\phi).
\]

For the function \( \phi(\lambda) \equiv \lambda^{-1} \),

\[
\frac{\|x - x_0\|_A^2}{\|r_0\|_A^2} = \text{\( j \)-point Gauss quadrature} + \frac{\|x - x_j\|_A^2}{\|r_0\|_A^2}.
\]

Consequence: For the discretized problem, CG convergence behavior is determined by the approximations of the distribution function given by \( A, r_0 \) via the sequence of the Gauss-Christoffel step-wise distribution functions \( \{\omega^{(j)}(\lambda)\} \).

But this should be linked to the infinite dimensional distribution function \( \omega(\lambda) \) determined by \( G, r \).
Rounding errors seemingly irreparably destroy the underlying mathematical structure that is based on orthogonality, and therefore the link with Gauss-Christoffel quadrature seems to be irreparably lost as well. However,

Lanczos (with small inaccuracy also CG) in finite precision arithmetic can be seen as the exact arithmetic Lanczos (CG) for the problem with the slightly modified distribution function with single eigenvalues replaced by tight clusters.

Paige (1971-80), Greenbaum (1989), Parlett (1990), S (1991), Greenbaum and S (1992), Notay (1993), ... , Druskin, Kniznermann, Zemke, Wülling, Meurant, ...

What is the relationship between the distribution function of the problem defined on the infinite dimensional Hilbert space and the stepwise distribution functions defined by the associated discretized problems?

Or, what is, at least, the relationship between the spectrum of the infinite dimensional (non-compact) operator $\mathcal{G}$ and the spectra of the associated (sequence of) matrices arising from (adaptively refined) discretizations?

Can we \textit{a priori} say anything about the spectra of these matrices arising from discretizations?
2 Stimulating work that formulated an open problem (2009)

1. Spectral information and convergence of the conjugate gradient method.


6. Spectral approximation of operators and/or PDE eigenvalue problem.
Theorem.

Consider open and bounded Lipschitz domain $\Omega \in \mathbb{R}^2$ and the operator $\nabla \cdot (k(x)\nabla u)$, where $k(x) : \Omega \to \mathbb{R}$ is a scalar real valued bounded and uniformly positive function. Then for all $x \in \Omega$ at which $k(x)$ is continuous,

$$k(x) \in \text{sp}(L^{-1}A),$$

i.e., the image of the domain under a continuous coefficient function $k(x)$ is a subset of the spectrum of the preconditioned operator $L^{-1}A$, where

$$\mathcal{A} : H^1_0(\Omega) \mapsto H^{-1}(\Omega), \quad \langle \mathcal{A}u, v \rangle = \int_{\Omega} k(x)\nabla u \cdot \nabla v, \quad u, v \in H^1_0(\Omega),$$

$$\mathcal{L} : H^1_0(\Omega) \mapsto H^{-1}(\Omega), \quad \langle \mathcal{L}u, v \rangle = \int_{\Omega} \nabla u \cdot \nabla v, \quad u, v \in H^1_0(\Omega).$$
Numerical experiments suggest that:

1. $k(\Omega)$ yields a good approximation of the whole spectrum of $\mathcal{L}^{-1}\mathcal{A}$;

2. this infinite dimensional spectrum (possibly including its continuous part) is well approximated by the eigenvalues of matrices arising from discretization.
Spectral information and convergence of the conjugate gradient method.


Spectral approximation of operators and/or PDE eigenvalue problem.
Theorem.

Consider discretization via conforming FEM with the basis functions \( \phi_j, j = 1, \ldots, N \). Let \( A, L \) be the matrix representations of the discrete operators. Let \( 0 < \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_N \) be the eigenvalues of \( L^{-1}A \). Let \( k(x) \) be uniformly positive, bounded and piecewise continuous.

Then there exists a (possibly non-unique) permutation \( \pi \) such that the eigenvalues of the matrix \( L^{-1}A \) satisfy

\[
\lambda_{\pi(j)} \in k(T_j), \quad j = 1, \ldots, N,
\]

where

\[
k(T_j) \equiv [\inf_{x \in T_j} k(x), \sup_{x \in T_j} k(x)], \quad T_j = \text{supp}(\phi_j), \quad j = 1, \ldots, N.
\]
Theorem.

Consider discretization via conforming FEM with the basis functions \( \phi_j, \ j = 1, \cdots, N \). Let \( A, L \) be the matrix representations of the discrete operators. Let \( 0 < \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_N \) be the eigenvalues of \( L^{-1}A \). Let \( k(x) \) be uniformly positive, bounded and piecewise continuous.

Then there exists a (possibly non-unique) permutation \( \pi \) such that the eigenvalues of the matrix \( L^{-1}A \) satisfy

\[
\lambda_{\pi(j)} \in k(T_j), \quad j = 1, \ldots, N,
\]

where

\[
k(T_j) \equiv [ \inf_{x \in T_j} k(x), \sup_{x \in T_j} k(x) ], \quad T_j = \text{supp}(\phi_j), \quad j = 1, \ldots, N.
\]

Proof:
Constructive perturbation argument and the Hall’s theorem on bipartite graphs.
3 Numerical illustration, 4 problems, nodal values, N = 81

[Graphs showing eigenvalues and nodal values for different scales.]
Let $k(\mathcal{T}_j)$ be constant over a patch of the discretization supports. Then we know the associated eigenvalue exactly including the multiplicity.

Other approach by Ladecký, Pultarová and Zeman (Appl. of Math., 2020).
Let $k(T_j)$ be constant over a patch of the discretization supports. Then we know the associated eigenvalue exactly including the multiplicity.

Other approach by Ladecký, Pultarová and Zeman (Appl. of Math., 2020).

Open questions:

- Can the whole spectrum of the infinite dimensional preconditioned operator $L^{-1}A$ be determined as $k(\Omega)$?
- Generalizations to tensors, more general preconditioning, indefinite problems?
4 Determining a priori the operator spectrum

1. Spectral information and convergence of the conjugate gradient method.


6. Spectral approximation of operators and/or PDE eigenvalue problem.
Consider the operator \( \nabla \cdot (K(x) \nabla u) \) with the real valued tensor function \( K(x) : \Omega \to \mathbb{R}^{2 \times 2} \) being symmetric with its entries being bounded Lebesgue integrable functions, and with the spectral decomposition

\[
K(x) = Q(x) \Lambda(x) Q^T(x), \quad x \in \Omega,
\]

where

\[
\Lambda(x) = \begin{bmatrix}
\kappa_1(x) & 0 \\
0 & \kappa_2(x)
\end{bmatrix}, \quad QQ^T = Q^T Q = I.
\]
Theorem.

Let the symmetric tensor $K(x)$ be continuous throughout the closure $\overline{\Omega}$. Then the spectrum of the operator $\mathcal{L}^{-1}A$ is given by the interval

$$\text{sp}(\mathcal{L}^{-1}A) = \text{Conv}(\kappa_1(\overline{\Omega}) \cup \kappa_2(\overline{\Omega})),$$

where

$$\text{Conv}(\kappa_1(\overline{\Omega}) \cup \kappa_2(\overline{\Omega})) = \left[ \inf_{x \in \overline{\Omega}} \min_{i=1,2} \kappa_i(x), \sup_{x \in \overline{\Omega}} \max_{i=1,2} \kappa_i(x) \right].$$
Theorem.

Let the symmetric tensor $K(x)$ be continuous throughout the closure $\bar{\Omega}$. Then the spectrum of the operator $\mathcal{L}^{-1}A$ is given by the interval

$$\text{sp}(\mathcal{L}^{-1}A) = \text{Conv}(\kappa_1(\bar{\Omega}) \cup \kappa_2(\bar{\Omega})),$$

where

$$\text{Conv}(\kappa_1(\bar{\Omega}) \cup \kappa_2(\bar{\Omega})) = \left[ \inf_{x \in \bar{\Omega}} \min_{i=1,2} \kappa_i(x) \right, \sup_{x \in \bar{\Omega}} \max_{i=1,2} \kappa_i(x) \right].$$

Assuming only that the symmetric tensor $K(x)$ is continuous at least at a single point in $\Omega$ and (WLOG) $\sup_{x \in \Omega} \kappa_1(x) < \inf_{x \in \Omega} \kappa_2(x)$, then the following closed interval belongs to the spectrum of $\mathcal{L}^{-1}A$,

$$\left[ \sup_{x \in \Omega} \kappa_1(x), \inf_{x \in \Omega} \kappa_2(x) \right] \subset \text{sp}(\mathcal{L}^{-1}A).$$
4 Eigenvalues of the discretized problems P1 – P3 in the paper

P1: constant $\kappa_1 \neq \kappa_2$

P2: non overlapping $\kappa_1(\Omega)$ and $\kappa_2(\Omega)$

P3: overlapping $\kappa_1(\Omega)$ and $\kappa_2(\Omega)$
4 Remaining questions

- Spectrum of the infinite dimensional preconditioned operator is the complement of the resolvent set. How do the spectra of matrices that represent discretized preconditioned operators approximate the spectral interval of the infinite dimensional preconditioned operator?

- More general preconditioning? (Instead of approximating the distribution function, here we deal only with approximating the spectrum).

Here we do not ask about numerical approximation of the eigenvalues of the infinite dimensional (PDE) operator, which represents a fundamentally different problem.
Spectral information and convergence of the conjugate gradient method.


Spectral approximation of operators and/or PDE eigenvalue problem.
Infinite-dimensional spectrum:

**Theorem.**

Consider an open and bounded Lipschitz domain $\Omega \subset \mathbb{R}^2$, and the operators $\nabla \cdot (k(x)\nabla u)$, and $\nabla \cdot (g(x)\nabla u)$. Assume that the scalar functions $g(x)$ and $k(x)$ are continuous throughout the closure $\overline{\Omega}$ and that $g(x)$ is, in addition, uniformly positive. Then the spectrum of the operator $B^{-1}A$ equals

$$
\text{sp}(B^{-1}A) = \left[ \inf_{x \in \overline{\Omega}} \frac{k(x)}{g(x)}, \sup_{x \in \overline{\Omega}} \frac{k(x)}{g(x)} \right].
$$
Eigenvalues of the discretized matrices:

**Theorem.**

Let $0 < \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$ be the eigenvalues of $B_n^{-1}A_n$. Let $g(x)$ and $k(x)$ be bounded and *piecewise continuous* functions, and $g(x)$ be, in addition, uniformly positive. Then there exists a (possibly non-unique) permutation $\pi$ such that the eigenvalues of the matrix $B_n^{-1}A_n$ satisfy

$$
\lambda_{\pi(j)} \in \left[ \inf_{x \in T_j} \frac{k(x)}{g(x)}, \sup_{x \in T_j} \frac{k(x)}{g(x)} \right], \quad j = 1, \ldots, n,
$$

where $T_j$ represents the support of the $j$th FEM basis function.
For the specific cases we get both lower and upper semicontinuity of the spectrum. For the abstract setting we get only lower semicontinuity (see Theorem in part 1).

Here we approximate the spectrum of the bounded and continuously invertible operator $B^{-1}A : V \rightarrow V$ on the infinite dimensional Hilbert space.
For the specific cases we get both lower and upper semicontinuity of the spectrum. For the abstract setting we get only lower semicontinuity (see Theorem in part 1).

Here we approximate the spectrum of the bounded and continuously invertible operator $B^{-1}A : V \rightarrow V$ on the infinite dimensional Hilbert space.

Puzzling question:

When the whole spectrum of the infinite dimensional operator is in the limit approximated by the eigenvalues of the associated matrices, and the whole spectrum is a large interval, does it mean that for refined discretizations the performance of CG applied to the discretized problems significantly deteriorates with the mesh refinement? Not necessarily! Motivating example in Gergelits, Mardal, Nilsen and S (2019) offers an explanation.
1 Spectral information and convergence of the conjugate gradient method.


6 Spectral approximation of operators and/or PDE eigenvalue problem.
PDE eigenvalue problem is based on construction of *compact solution operators*. Babuška - Osborn theory.

The set of compact operators is closed wrt the norm-wise (uniform) convergence.

Spectrum of an infinite dimensional compact operator is composed of isolated eigenvalues with a single accumulation point.
PDE eigenvalue problem is based on construction of compact solution operators. Babuška - Osborn theory.

The set of compact operators is closed wrt the norm-wise (uniform) convergence.

Spectrum of an infinite dimensional compact operator is composed of isolated eigenvalues with a single accumulation point.

Bounded continuously invertible operator on an infinite dimensional Hilbert space is not compact.

Convergence of matrix eigenvalues to eigenvalues of a compact operator is a different problem than approximation of the whole spectrum of invertible operators. The later, not the former, is relevant to the operator preconditioning.
The presented line of development does not allow to approximate the distribution function \( \omega(\lambda) \). Assuming that all eigenspaces contribute to the finite dimensional distribution functions equally, we get the so-called *cumulative spectral density*, which is important in physics dealing with the so-called *density of states*; see, e.g., Lin, Saad and Yang, (SIREV, 2016). For the given class of problems we can cheaply approximate this, but the infinite dimensional case is approached only as a limit of the refinements of the discrete cases.

An amazingly beautiful results that do allow to compute (not only) the cumulative spectral density of wide class of infinite dimensional operators are presented in the PhD Thesis by Colbrook (Cambridge U, 2020) and in the several recent related papers; see, in particular, the paper by Colbrook, Horning and Townsend (SIREV, 2021).
“We will go on pondering and meditating, the great mysteries still ahead of us, we will err and stumble on the way, and if we win a little victory, we will be jubilant and thankful, without claiming, however, that we have done something that can eliminate the contribution of all the millenia before us.”
“There remains this: we beech the skilled in these things, that we thought worth showing, they will think openly receiving, an whatever it hides, worth imparting more properly by themselves to the wider mathematical community.”
“There remains this: we beech the skilled in these things, that we thought worth showing, they will think openly receiving, an whatever it hides, worth imparting more properly by themselves to the wider mathematical community.”

Thank you for your kind attention!