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

Tomáš Gergelits* Bjørn Fredrik Nielsen† Zdeněk Strakoš‡

Abstract

We consider operator preconditioning $\mathcal{B}^{-1}\mathcal{A}$, which is employed in the numerical solution of boundary value problems. Here, the self-adjoint operators $\mathcal{A}, \mathcal{B} : H_0^1(\Omega) \to H^{-1}(\Omega)$ are the standard integral/functional representations of the partial differential operators $-\nabla \cdot (k(x)\nabla u)$ and $-\nabla \cdot (g(x)\nabla u)$, respectively, and the scalar coefficient functions $k(x)$ and $g(x)$ are assumed to be continuous throughout the closure of the solution domain. The function $g(x)$ is also assumed to be uniformly positive. When the discretized problem, with the preconditioned operator $\mathcal{B}_n^{-1}\mathcal{A}_n$, is solved with Krylov subspace methods, the convergence behavior depends on the distribution of the eigenvalues. Therefore it is crucial to understand how the eigenvalues of $\mathcal{B}_n^{-1}\mathcal{A}_n$ are related to the spectrum of $\mathcal{B}^{-1}\mathcal{A}$. Following the path started in the two recent papers published in SIAM J. Numer. Anal. [57 (2019), pp. 1369-1394 and 58 (2020), pp. 2193-2211], the first part of this paper addresses the open question concerning the distribution of the eigenvalues of $\mathcal{B}_n^{-1}\mathcal{A}_n$ formulated at the end of the second paper. The approximation of the spectrum studied in the present paper differs from the eigenvalue problem studied in the classical PDE literature which addresses individual eigenvalues of compact (solution) operators.

In the second part of this paper we generalize some of our results to general bounded and self-adjoint operators $\mathcal{A}, B : V \to V^{\#}$, where $V^{\#}$ denotes the dual of V. More specifically, provided that $\mathcal{B}$ is coercive and that the standard Galerkin discretization approximation properties hold, we prove that the whole spectrum of $\mathcal{B}^{-1}\mathcal{A} : V \to V$ is approximated to an arbitrary accuracy by the eigenvalues of its finite dimensional discretization $\mathcal{B}_n^{-1}\mathcal{A}_n$.

*Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic. Email: gergelits@karlin.mff.cuni.cz. This author was supported by Lawrence Livermore National Security, LLC Subcontract Award B639388 under Prime Contract No. DE-AC52-07NA27344.

†Faculty of Science and Technology, Norwegian University of Life Sciences, P.O. Box 5003, NO-1432 Ås, Norway. Email: bjorn.f.nielsen@nmbu.no. Nielsen's work was supported by The Research Council of Norway, project number 239070.

‡Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic. Email: strakos@karlin.mff.cuni.cz. Supported by the Grant Agency of the Czech Republic under the contract No. 17-04150J.

Keywords: Second order PDEs, bounded non-compact operators, generalized spectrum, numerical approximation, preconditioning.

1 Introduction.

Any linear self-adjoint operator $\mathcal{G} : V \to V$ defined on an infinite dimensional real Hilbert space V can be expressed in terms of the Riemann-Stieltjes integral as

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

where the spectral function $E(\lambda)$ of $\mathcal{G}$ represents a family of orthogonal projections (projection-valued measure), analogous to the family of projectors onto individual invariant subspaces for symmetric matrices; see [29, Chapter II]. This analogy is straightforward for compact self-adjoint operators with infinite range: They can be regarded as limits in norm of finite Hermitian matrices and have a spectrum composed of an infinite number of isolated eigenvalues with zero being the only accumulation point. Furthermore, in such cases the Riemann-Stieltjes integral, presented above, can be written as an infinite sum. In this paper we will, however, consider the case when the operator spectrum contains a continuous part.

Consider a Krylov subspace method, such as the conjugate gradient method (CG) or the minimal residual method (MINRES), applied for solving a linear system associated with the operator $\mathcal{G}$. Then it is beneficial to consider the infinite-dimensional distribution function[1] $\omega(\lambda)$ which is determined by the integral representation

$$(\mathcal{G}r, r) = \int \lambda \, d(E(\lambda)r, r) = \int \lambda \, d\omega(\lambda),$$

where r denotes the initial residual normalized to one (cf. [24, section 5.2]). This is analogous to using an orthonormal eigenvector basis to obtain the diagonal representation of symmetric matrices. Being able to describe the convergence of the Gauss quadrature approximations $\omega^{(n)}(\lambda)$ to $\omega(\lambda)$ (cf. [30, Theorem IX, p. 61]), we would have complete information about the computational behavior of CG or MINRES (cf. [22, Chapter 5] and [24, section 5.2]) and the motivating example in [14, sections 2 and 4.4]. Such information is in practice not available[2], but it would certainly be highly beneficial to have information about the complete spectrum of the discretized operators and of its convergence to the infinite-dimensional spectrum. This will be investigated below.

[1] Equivalently, we can also consider the closely related spectral measure; see, e.g., [9].

[2] Recent beautiful results published in [9, 8] show how to compute smoothed approximations of spectral measures for infinite dimensional self-adjoint operators.

Extending the path of research started in [26, 14, 15], this paper will first consider the differential operators $-\nabla\cdot(k(x)\nabla u)$ and $-\nabla\cdot(g(x)\nabla u)$ on the open and bounded Lipschitz domain $\Omega\subset\mathbb{R}^2$, where the scalar functions $g(x)$ and $k(x)$ are continuous throughout the closure $\overline{\Omega}$, and $g(x)$ is, in addition, uniformly positive. The associated operator representations $\mathcal{A}$, $\mathcal{B}: H_0^1(\Omega)\to H^{-1}(\Omega)$ are given by[3]

$$\langle \mathcal{A}u, v\rangle = \int_\Omega k(x)\nabla u\cdot\nabla v, \quad u,v\in H_0^1(\Omega), \tag{1}$$

$$\langle \mathcal{B}u, v\rangle = \int_\Omega g(x)\nabla u\cdot\nabla v, \quad u,v\in H_0^1(\Omega). \tag{2}$$

In the first part of this paper, i.e., in sections 2 and 3, we will start with characterizing the spectrum of the preconditioned operator

$$\mathcal{B}^{-1}\mathcal{A}: H_0^1(\Omega)\to H_0^1(\Omega), \tag{3}$$

defined as the complement of the resolvent set, i.e.,

$$\mathrm{sp}(\mathcal{B}^{-1}\mathcal{A}) := \left\{\lambda\in\mathbb{C};\ \lambda\mathcal{I}-\mathcal{B}^{-1}\mathcal{A} \text{ does not have a bounded inverse}\right\}. \tag{4}$$

More specifically, we prove that

$$\mathrm{sp}(\mathcal{B}^{-1}\mathcal{A}) = \left[\inf_{x\in\overline{\Omega}}\ \frac{k(x)}{g(x)}, \sup_{x\in\overline{\Omega}}\ \frac{k(x)}{g(x)}\right].$$

Consider a sequence of finite dimensional subspaces $\{V_n\}$ of $H_0^1(\Omega)$ defined via the nodal polynomial basis functions[4] $\phi_1,\ldots,\phi_n$ with the local supports

$$\mathcal{T}_j = \mathrm{supp}(\phi_j), \quad j=1,\ldots,n. \tag{5}$$

Throughout this paper we assume that

$$\overline{\Omega} = \cup_{j=1}^n \mathcal{T}_j,$$
$$V_n = \mathrm{span}\,\{\phi_1,\ldots,\phi_n\}.$$

Typically, $\mathcal{T}_1,\ldots,\mathcal{T}_n$ are composed in $\mathbb{R}^2$ of triangles, which only share edges and vertices.

The standard Galerkin finite element discretization of the operators $\mathcal{A}$ and $\mathcal{B}$ gives the matrix representations of the discretised operators $\mathcal{A}_n$ and

[3] Here we consider homogeneous Dirichlet boundary conditions. The results will remain valid (with an appropriate choice of the associated spaces) also for homogeneous Neumann boundary conditions; see [15, section 5] and the numerical experiments section below.

[4] As in [14], we consider conforming Finite Element (FE) methods using Lagrange elements.

$\mathcal{B}_n$ in terms of the basis $\phi_1, \ldots, \phi_n$,

$$[\mathbf{A}_n]_{ij} = \int_\Omega k(x) \nabla \phi_j \cdot \nabla \phi_i, \quad i, j = 1, \ldots, n. \tag{6}$$

$$[\mathbf{B}_n]_{ij} = \int_\Omega g(x) \nabla \phi_j \cdot \nabla \phi_i, \quad i, j = 1, \ldots, n. \tag{7}$$

Part one of this paper then also contains an investigation of the approximation of the whole spectrum of $\mathcal{B}^{-1}\mathcal{A}$ by the eigenvalues of the preconditioned matrices $\mathbf{B}_n^{-1}\mathbf{A}_n$ as $n \to \infty$.

The second part, i.e., section 4, generalizes the results obtained for (1) and (2). More precisely, the spectral approximation is explored in terms of an abstract setting where $\mathcal{A}, \mathcal{B} : V \to V^\#$ are assumed to be bounded and self-adjoint[5] linear operators, with $\mathcal{B}$ being also coercive. Here, $V^\#$ denotes the dual of V consisting of all linear bounded functionals from the infinite dimensional Hilbert space V to $\mathbb{R}$. Considering a Galerkin discretization, using a sequence $\{V_n\}$ of subspaces $V_n \subset V$ satisfying the standard approximation property

$$\lim_{n\to\infty} \inf_{v \in V_n} \|w - v\| = 0 \quad \text{for all } w \in V, \tag{8}$$

we prove that the entire spectrum of $\mathcal{B}^{-1}\mathcal{A} : V \to V$ is approximated, as $n \to \infty$, to an arbitrary accuracy by the eigenvalues of the finite dimensional discretizations $\mathcal{B}_n^{-1}\mathcal{A}_n$. Since this also includes the case of $\mathcal{B}^{-1}\mathcal{A}$ being continuously invertible and V is of infinite dimension, $\mathcal{B}^{-1}\mathcal{A}$ is not compact. Such an investigation can then not be based on the uniform (normwise) convergence, and it relies instead upon the pointwise (strong) convergence of $\mathcal{B}_n^{-1}\mathcal{A}_n$ to $\mathcal{B}^{-1}\mathcal{A}$.

The kind of problems we consider can be analyzed using an abstract operator theoretical framework or techniques closely related to the numerical PDE literature. We chose to do the latter due to our competence and because this is more appropriate from a numerical point of view. We also believe that our approach can make the related spectral theory results more accessible to a wider audience. For this sake we provide detailed references to the related operator theory results, and present a concise reproduction of the proof of Theorem 4.1 on the approximation of the spectrum of self-adjoint operators based on pointwise (strong) convergence.

The results in this paper deal with operator preconditioning, and most of the analyses presented in sections 2 and 3 can be applied to the case when the piecewise continuous function $k(x)$ is locally approximated by some easier-to-handle function $g(x)$ (with the goal to solve linear equations

[5]Self-adjoint in the sense that $\langle \mathcal{A}u, v\rangle = \langle \mathcal{A}v, u\rangle, \langle \mathcal{B}u, v\rangle = \langle \mathcal{B}v, u\rangle$ for all $u, v \in V$, where $\langle \cdot, \cdot \rangle : V^\# \times V \to \mathbb{R}$ is the duality pairing. Equivalently, $\tau\mathcal{A}$, $\tau\mathcal{B} : V \to V$ are self-adjoint, where τ is the Riesz map.

efficiently with the preconditioner $\mathcal{B}$). This can happen, e.g., for a piecewise-constant preconditioning; see [2, section 5.1.2], [18, 3]. The aim of this paper is, however, neither to study computationally a particular preconditioning technique nor to compare different approaches. It is also worth to point out that most of the literature on operator preconditioning is based on the spectral (or norm) equivalence of operators. This can lead to independence of the bound for the condition number of the preconditioned matrices wrt. the discretization parameter (h-refinement of the mesh) and eventually wrt. other problem-specific parameters; see, e.g., [12, 2, 25, 17] and the recent paper [21] that also indicates possible benefits of approximating the whole spectrum of the preconditioned operators. Unless the resulting bound is sufficiently small, this approach can not, however, guarantee computational efficiency for a particular mesh (and for the particular values of the problem-specific parameters); see [12, 16, 3].

Following [26, 14, 15], and keeping in mind that the efficiency of preconditioning depends on the whole spectrum of the associated infinite dimensional preconditioned operator, our goal is to point out that complete information about this spectrum, for an important class of problems, is available almost for free. In addition, we will describe its link to the eigenvalues of the preconditioned matrices that determine the behavior of methods such as CG or MINRES.

2 Preconditioning by Laplacian ($g(x) = 1$).

Considering the case $g(x) = 1$, i.e., the preconditioner equals the operator representation of the Laplacian $\mathcal{B} = \mathcal{L}$, the paper [15] determines the spectrum of the preconditioned operator $\mathcal{L}^{-1}\mathcal{A}$ in the following way (for brevity we use a bit stronger assumptions than in [15]):

Theorem 2.1 (cf. [15], Theorem 1.1)**.** *Consider an open and bounded Lipschitz domain $\Omega \subset \mathbb{R}^2$. Assume that the scalar function $k(x)$ is continuous throughout the closure $\overline{\Omega}$. Then the spectrum of the operator $\mathcal{L}^{-1}\mathcal{A}$ equals the interval*

$$\mathrm{sp}(\mathcal{L}^{-1}\mathcal{A}) = \left[\inf_{x\in\overline{\Omega}} k(x), \sup_{x\in\overline{\Omega}} k(x) \right]. \tag{9}$$

In other words, for a continuous function $k(x)$, the spectrum of $\mathcal{L}^{-1}\mathcal{A}$ equals the range $k(\overline{\Omega})$.

Eigenvalues of the discretized operator $\mathcal{L}_n^{-1}\mathcal{A}_n$, that is represented by the matrix $\mathbf{L}_n^{-1}\mathbf{A}_n$, can be approximated using the following theorem from [14]. (Here we again use assumptions conforming to the setting in the current paper.)

Theorem 2.2 (cf. [14], Theorem 3.1)**.** *Let* $0 < \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$ *be the eigenvalues of* $\mathbf{L}_n^{-1}\mathbf{A}_n$*, where* $\mathbf{A}_n$ *and* $\mathbf{L}_n$ *are defined in* (6) *and* (7)[6]*, respectively. Let* $k(x)$ *be continuous throughout the closure* $\overline{\Omega}$*. Then there exists a (possibly non-unique) permutation* π *such that the eigenvalues* $\lambda_{\pi(j)}$ *of the matrix* $\mathbf{L}_n^{-1}\mathbf{A}_n$ *satisfy*

$$\lambda_{\pi(j)} \in \left[\inf_{x\in\mathcal{T}_j} k(x), \sup_{x\in\mathcal{T}_j} k(x)\right], \quad j = 1, \ldots, n, \tag{10}$$

where $\mathcal{T}_j$ *is defined in (5).*

Consequently, there is a one-to-one mapping (possibly non-unique) between the eigenvalues of $\mathbf{L}_n^{-1}\mathbf{A}_n$ and the ranges of $k(x)$ over the supports of the individual basis functions. With an appropriate grid refinement of the discretization, the size of the intervals containing the individual eigenvalues of $\mathbf{L}_n^{-1}\mathbf{A}_n$ converge linearly to zero.[7]

The paper [15] formulates a dual open question about the *distribution of the eigenvalues of the discretized operators within the interval (9).* Assuming in addition that $k(x) \in \mathcal{C}^2(\overline{\Omega})$, we will now show that theorems 2.1 and 2.2 yield that *any point in the spectrum* of the infinite dimensional operator $\mathcal{L}^{-1}\mathcal{A}$ is approximated, as the maximal diameter h of the associated finite elements vanish with $n \to \infty$, to an arbitrary accuracy by the eigenvalues (10) of the matrices $\mathbf{L}_n^{-1}\mathbf{A}_n$.

Consider an arbitrary point μ in the spectrum of the operator $\mathcal{L}^{-1}\mathcal{A}$. It should be noted that μ may not be an eigenvalue, and that our investigation differs from the eigenvalue problem studied in the numerical PDE literature which is based on approximations of the eigenvalues within the framework of infinite dimensional compact solution operators.

Using theorem 2.1, μ is the image under $k(x)$ of some point $y \in \overline{\Omega}$, i.e., $\mu = k(y)$. We first consider the case $y \in \Omega$. The case $y \in \partial\Omega$ will be resolved later by a simple limiting argument. Let $\varepsilon > 0$ be an arbitrarily small positive constant, and let

$$\delta = \frac{\varepsilon}{2\sup_{x\in\Omega} \|\nabla k(x)\|},$$

provided that $\sup_{x\in\Omega} \|\nabla k(x)\| > 0$. (The case $\sup_{x\in\Omega} \|\nabla k(x)\| = 0$ is uninteresting because then $\mathcal{A} = c\mathcal{L}$ for some constant c.) Consider further a Galerkin discretization such that the support of at least one of the nodal basis functions[8] that contains the point y is itself contained in the disc with center y and radius δ. Denote this support $\mathcal{T}_j$ and the associated eigenvalue

[6] With $g(x) = 1$.

[7] An interesting application inspired by this result that uses a different approach is presented in [20]

[8] Supports of all discretization functions are contained in $\overline{\Omega}$.

of the discretized operator given by Theorem 2.2 as $\lambda_{\pi(j)}$. Using Corollary 3.2 in [14] (with $y = \hat{x}_j$)

$$\begin{aligned} |\lambda_{\pi(j)} - k(y)| &\leq \sup_{x \in \mathcal{T}_j} \|k(x) - k(y)\| \\ &\leq \delta \|\nabla k(y)\| + \frac{1}{2}\delta^2 \sup_{x \in \mathcal{T}_j} \|D^2 k(x)\|, \end{aligned} \tag{11}$$

where $D^2k(x)$ denotes the second order derivative of the function $k(x)$[9]. For ε sufficiently small we thus get, after a simple manipulation,

$$|\lambda_{\pi(j)} - \mu| \leq \varepsilon.$$

If $\mu = k(y)$ and $y \in \partial\Omega$, the same conclusion can be obtained using the previous derivation and the continuity of $k(x)$ throughout $\overline{\Omega}$.

We also note that, using the notation $h_j = \operatorname{diam}(\mathcal{T}_j)$ and $h = \max_j h_j$,

$$\begin{aligned} |\lambda_{\pi(j)} - \mu| &= |\lambda_{\pi(j)} - k(y)| \\ &\leq \sup_{x \in \mathcal{T}_j} \|k(x) - k(y)\| \\ &\leq h_j \sup_{x \in \mathcal{T}_j} \|\nabla k(x)\| \qquad (12) \\ &\leq h \sup_{x \in \Omega} \|\nabla k(x)\|. \qquad (13) \end{aligned}$$

The second last inequality shows that the individual points in the infinite dimensional spectrum can be approximated with different "speed" as $h_j \to 0$. Furthermore, the last inequality expresses that this "speed" is at least linear, with respect to h, and uniformly bounded from zero because $\sup_{x \in \Omega} \|\nabla k(x)\| < \infty$.

Summing up, this proves the following theorem:

Theorem 2.3 (Approximation of the spectrum by matrix eigenvalues)**.** *Let $k(x)$ be twice continuously differentiable throughout the closure $\overline{\Omega}$. Let the maximal diameter h of the local supports of the basis functions used in the Galerkin discretization (5)-(7) vanish as $n \to \infty$. Then any point in the spectrum of the operator $\mathcal{L}^{-1}\mathcal{A}$ is for $n \to \infty$ approximated to an arbitrary accuracy by the eigenvalues of the matrices $\mathbf{L}_n^{-1}\mathbf{A}_n$ representing the discretized preconditioned operators. Moreover, since the size of $\|\nabla k(y)\|$ is uniformly bounded from above throughout $\overline{\Omega}$, the speed of convergence towards the individual spectral points, as $n \to \infty$ and $h \to 0$, is uniformly bounded from zero by* (13) *throughout the whole spectrum of $\mathcal{B}^{-1}\mathcal{A}$. It can, however, differ for different spectral points; see* (12)*.*

[9] See [6, Section 1.2] for the definition of the second order derivative.

If one assumes that $k(x)$ is Hölder continuous, then one would get bounds in the form $C_1 h_j^\alpha$ and $C_2 h^\alpha$ for $|\lambda_{\pi(j)} - \mu|$ instead of (12) and (13), respectively, where $0 < \alpha < 1$ denotes the exponent of the Hölder condition. This only leads to improved bounds when $h_j > 1$ or $h > 1$. In fact, the numerical experiments presented in [26, tables 1-3] suggest that the approximation is of order $O(h)$ (also when $k(x)$ is twice continuously differentiable throughout $\overline{\Omega}$).

3 Generalization to (piecewise) continuous and uniformly positive $g(x)$.

The purpose of this section is to generalize the results presented above to preconditioners in the form (2). We first present the theorems and a corollary, and thereafter their proofs are discussed.

The content of the present section is motivated by the desire to increase our knowledge about second order differential operators and preconditioning issues. In particular, it can contribute towards a better understanding of piecewise constant preconditioners.

Theorem 3.1 (Spectrum of the infinite dimensional preconditioned operator)**.**
Consider an open and bounded Lipschitz domain $\Omega \subset \mathbb{R}^2$. 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 $\mathcal{B}^{-1}\mathcal{A}$, defined in (4)*, equals*

$$\mathrm{sp}(\mathcal{B}^{-1}\mathcal{A}) = \left[\inf_{x\in\overline{\Omega}} \frac{k(x)}{g(x)}, \sup_{x\in\overline{\Omega}} \frac{k(x)}{g(x)} \right]. \tag{14}$$

Note that (14) shows that the spectrum of $\mathcal{B}^{-1}\mathcal{A}$ is identical to the spectrum of the simple multiplication operator defined point wise as

$$(\mathcal{M}u)(x) = \frac{k(x)}{g(x)} u(x).$$

The next theorem deals with the localization of the eigenvalues of the preconditioned matrix arising from the discretization. It does not consider the approximation of the spectrum of the infinite dimensional operator $\mathcal{B}^{-1}\mathcal{A}$. Analogously to [14], we can therefore relax the assumptions about the continuity of the coefficient functions $k(x)$ and $g(x)$.

Theorem 3.2 (Eigenvalues of the preconditioned matrix)**.** *Let $0 < \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$ be the eigenvalues of $\mathbf{B}_n^{-1}\mathbf{A}_n$, where $\mathbf{A}_n$ and $\mathbf{B}_n$ are defined in* (6) *and* (7)*, respectively. Let $g(x)$ and $k(x)$ be bounded and* piecewise continuous *functions, and $g(x)$ is, in addition, uniformly positive. Then*

there exists a (possibly non-unique) permutation π such that the eigenvalues of the matrix $\mathbf{B}_n^{-1}\mathbf{A}_n$ satisfy

$$\lambda_{\pi(j)} \in \left[\inf_{x\in\mathcal{T}_j} \frac{k(x)}{g(x)}, \sup_{x\in\mathcal{T}_j} \frac{k(x)}{g(x)} \right], \quad j = 1, \ldots, n, \tag{15}$$

where $\mathcal{T}_j$ is defined in (5).

Consequently, if $k(x)/g(x)$ is continuous on $\mathcal{T}_j$, then there exists at least one point $y_j \in \mathcal{T}_j$ such that $\lambda_{\pi(j)} = k(y_j)/g(y_j)$. Otherwise, we only know that $\lambda_{\pi(j)}$ belongs to the interval (15), which may be large. In order to prove Theorem 3.2, see below, we approximate $k(x)$ locally with a constant and thereafter employ a standard result concerning eigenvalue perturbations of Hermitian matrices. We do not know how to modify this argument to the infinite dimensional case. In fact, the proof of Theorem 3.1, which we present below, is indirect and relies on the continuity of $k(x)$ and $g(x)$.

Corollary 3.3 (Pairing the eigenvalues and the nodal values)**.** *Using the notation and the assumptions of Theorem 3.2, consider any point $\hat{x}_j \in \mathcal{T}_j$. Then the associated eigenvalue $\lambda_{\pi(j)}$ of the matrix $\mathbf{B}_n^{-1}\mathbf{A}_n$ satisfies*

$$|\lambda_{\pi(j)} - r(\hat{x}_j)| \le \sup_{x\in\mathcal{T}_j} |r(x) - r(\hat{x}_j)|, \quad j = 1, \ldots, n, \tag{16}$$

where

$$r(x) \equiv \frac{k(x)}{g(x)}.$$

If, in addition, $k(x)$ and $g(x) \in \mathcal{C}^2(\mathcal{T}_j)$, then

$$\begin{aligned} |\lambda_{\pi(j)} - r(\hat{x}_j)| &\le \sup_{x\in\mathcal{T}_j} |r(x) - r(\hat{x}_j)| \\ &\le h_j \|\nabla r(\hat{x}_j)\| + \tfrac{1}{2} {h_j}^2 \sup_{x\in\mathcal{T}_j} \|D^2 r(x)\|, \quad j = 1, \ldots, n, \end{aligned} \tag{17}$$

where $h_j = \operatorname{diam}(\mathcal{T}_j)$ and $D^2 r(x)$ denotes the second order derivative of $r(x)$. In particular, (16) *and* (17) *hold for any discretization mesh node $\hat{x}_j \in \mathcal{T}_j$.*

Invoking theorems 3.1 and 3.2 we find, analogously to (12) and (13), that

$$|\lambda_{\pi(j)} - \mu| \le h_j \sup_{x\in\mathcal{T}_j} \|\nabla r(x)\| \tag{18}$$

$$\le h \sup_{x\in\Omega} \|\nabla r(x)\|, \tag{19}$$

where $h = \max_j h_j$, μ is an arbitrary point in the spectrum of $\mathcal{B}^{-1}\mathcal{A}$ and $\lambda_{\pi(j)}$ is its approximation which is an eigenvalue of $\mathbf{B}_n^{-1}\mathbf{A}_n$.

Summing up, this proves the following theorem analogous to Theorem 2.3. This result is also of interest in relation to the fact that in other problems it was proved that the uniformly converging numerical approximation to the whole spectrum is not possible; see [10].

Theorem 3.4 (Approximation of the spectrum by matrix eigenvalues)**.** *Consider the notation and assumptions of Theorem 3.2 and Corollary 3.3. In addition, let $r(x)$ be twice continuously differentiable throughout the closure $\overline{\Omega}$. Let the maximal diameter h of the local supports of the basis functions used in the Galerkin discretization (5)-(7) vanish as $n \to \infty$. Then any point in the spectrum of the operator $\mathcal{B}^{-1}\mathcal{A}$ is for $n \to \infty$ approximated to an arbitrary accuracy by the eigenvalues of the matrices $\mathbf{B}_n^{-1}\mathbf{A}_n$ representing the discretized preconditioned operators. Moreover, since the size of $\|\nabla r(x)\|$ is uniformly bounded from above throughout $\overline{\Omega}$, the speed of convergence towards the individual spectral points, as $n \to \infty$ and $h \to 0$, is uniformly bounded from zero by* (19) *throughout the whole spectrum of $\mathcal{B}^{-1}\mathcal{A}$. It can, however, differ for different spectral points; see* (18)*.*

While the proofs of Theorem 3.1 and Theorem 3.2 are presented below, Corollary 3.3 follows immediately by applying [14, Corollary 3.2] to the ratio function $r(x)$.

Proof of Theorem 3.1.

Recall the definition (2) of the operator $\mathcal{B}$, and let us introduce the following notation for the inner product and norm induced by $\mathcal{B}$,

$$\begin{aligned}(u,v)_{\mathcal{B}} &= \langle \mathcal{B}u, v\rangle, \quad u, v \in H_0^1(\Omega),\\ \|u\|_{\mathcal{B}}^2 &= (u,u)_{\mathcal{B}}, \quad u \in H_0^1(\Omega).\end{aligned}$$

1. The proof of the fact that

$$\mathrm{sp}(\mathcal{B}^{-1}\mathcal{A}) \subset \left[\inf_{x\in\overline{\Omega}} \frac{k(x)}{g(x)}, \sup_{x\in\overline{\Omega}} \frac{k(x)}{g(x)}\right]$$

is analogous to the proof in [15, Section 3], employing the inner product induced by $\mathcal{B}$ instead of that induced by the Laplacian $\mathcal{L}$. More precisely, using the self-adjointness of the operator $\mathcal{B}^{-1}\mathcal{A}$ with respect to the inner product $(\cdot,\cdot)_{\mathcal{B}}$, the spectrum of $\mathcal{B}^{-1}\mathcal{A}$ is real and it is contained in the interval

$$\begin{aligned}\mathrm{sp}(\mathcal{B}^{-1}\mathcal{A}) &\subset \left[\inf_{u\in H_0^1(\Omega)} \frac{(\mathcal{B}^{-1}\mathcal{A}u,u)_{\mathcal{B}}}{(u,u)_{\mathcal{B}}}, \sup_{u\in H_0^1(\Omega)} \frac{(\mathcal{B}^{-1}\mathcal{A}u,u)_{\mathcal{B}}}{(u,u)_{\mathcal{B}}}\right]\\ &= \left[\inf_{u\in H_0^1(\Omega)} \frac{\langle \mathcal{A}u,u\rangle}{\langle \mathcal{B}u,u\rangle}, \sup_{u\in H_0^1(\Omega)} \frac{\langle \mathcal{A}u,u\rangle}{\langle \mathcal{B}u,u\rangle}\right]. \qquad (20)\end{aligned}$$

Moreover, the endpoints of this interval are contained in the spectrum. It remains to bound

$$\frac{\langle \mathcal{A}u,u\rangle}{\langle \mathcal{B}u,u\rangle} \qquad (21)$$

in terms of the values of the scalar functions $g(x)$ and $k(x)$. Let $\|\cdot\|$ denote the standard Euclidean norm. Then,

$$\sup_{u\in H_0^1(\Omega)} \frac{\langle \mathcal{A}u, u\rangle}{\langle \mathcal{B}u, u\rangle} = \sup_{u\in H_0^1(\Omega)} \frac{\int_\Omega k(x)\|\nabla u\|^2}{\int_\Omega g(x)\|\nabla u\|^2} = \sup_{u\in H_0^1(\Omega)} \frac{\int_\Omega \frac{k(x)}{g(x)} g(x)\|\nabla u\|^2}{\int_\Omega g(x)\|\nabla u\|^2}$$
$$\leq \sup_{x\in\Omega} \frac{k(x)}{g(x)}, \tag{22}$$

where we have used the assumption that $k(x)$ and $g(x)$ are continuous on $\overline{\Omega}$ and that $g(x)$ is uniformly positive. Similarly,

$$\inf_{u\in H_0^1(\Omega)} \frac{\langle \mathcal{A}u, u\rangle}{\langle \mathcal{L}u, u\rangle} \geq \inf_{x\in\Omega} \frac{k(x)}{g(x)}. \tag{23}$$

2. The proof of the converse inclusion

$$\left[\inf_{x\in\overline{\Omega}} \frac{k(x)}{g(x)}, \sup_{x\in\overline{\Omega}} \frac{k(x)}{g(x)}\right] \subset \mathrm{sp}(\mathcal{B}^{-1}\mathcal{A})$$

is analogous to the proof of [26, Theorem 3.1].

- For an arbitrary $x_0 \in \Omega$, consider $\lambda_0 = k(x_0)/g(x_0)$.
- Let $\{v_r\}_{r\in\mathbb{R}_+}$ be a set of functions satisfying[10]

$$\mathrm{supp}(v_r) \subset x_0 + U_r, \tag{24}$$
$$\|v_r\|_{\mathcal{B}} = 1, \tag{25}$$

  where $U_r = \{z \in \mathbb{R}^2 |\ \|z\| \leq r\}$.
- Next, let

$$u_r = (\lambda_0\mathcal{I} - \mathcal{B}^{-1}\mathcal{A})v_r, \tag{26}$$

  and observe that

$$\mathcal{B}u_r = (\lambda_0\mathcal{B} - \mathcal{A})v_r.$$

  Consequently (see (2)),

$$\begin{aligned}
\|u_r\|_{\mathcal{B}}^2 &= \langle(\lambda_0\mathcal{B} - \mathcal{A})v_r, u_r\rangle \\
&= \int_{x_0+U_r} (g(x)\lambda_0 - k(x))\,\nabla v_r \cdot \nabla u_r \\
&\leq \left(\int_{x_0+U_r} (g(x)\lambda_0 - k(x))^2\,|\nabla v_r|^2\right)^{1/2} \left(\int_{x_0+U_r} |\nabla u_r|^2\right)^{1/2} \\
&\leq \sup_{x\in x_0+U_r} \left|g(x)\frac{k(x_0)}{g(x_0)} - k(x)\right|\ g_{\min}^{-1}\,\|v_r\|_{\mathcal{B}}\|u_r\|_{\mathcal{B}},
\end{aligned}$$

[10] Note that no limit of v_r, as $r \to 0$, is needed in this proof. Only the existence of a set of functions satisfying (24) and (25) is required.

where

$$g_{\min} = \min_{x \in x_0 + U_r} g(x).$$

Employing (25),

$$\|u_r\|_{\mathcal{B}} \leq \sup_{x \in x_0 + U_r} \left| g(x) \frac{k(x_0)}{g(x_0)} - k(x) \right| g_{\min}^{-1},$$

and from the continuity of $g(x)$ and $k(x)$ we conclude that

$$\lim_{r \to 0} \|u_r\|_{\mathcal{B}} = 0. \tag{27}$$

- Assume that $\lambda_0 \notin \mathrm{sp}(\mathcal{B}^{-1}\mathcal{A})$. Then $\lambda_0 \mathcal{I} - \mathcal{B}^{-1}\mathcal{A}$ has a bounded inverse, and (26) and (27) imply that

  $$\|v_r\|_{\mathcal{B}} \leq \|(\lambda_0 \mathcal{I} - \mathcal{B}^{-1}\mathcal{A})^{-1}\|_{\mathcal{B}} \, \|u_r\|_{\mathcal{B}} \longrightarrow 0$$

  as $r \to 0$, which contradicts (25). We conclude that

  $$\lambda_0 = \frac{k(x_0)}{g(x_0)} \in \mathrm{sp}(\mathcal{B}^{-1}\mathcal{A}).$$

- Since $x_0 \in \Omega$ was arbitrary in this argument, g and k are continuous and g is uniformly positive on $\overline{\Omega}$, it follows that

  $$\left(\inf_{x \in \overline{\Omega}} \frac{k(x)}{g(x)}, \sup_{x \in \overline{\Omega}} \frac{k(x)}{g(x)} \right) \subset \mathrm{sp}(\mathcal{B}^{-1}\mathcal{A}). \tag{28}$$

  The endpoints of the interval (28) belong to $\mathrm{sp}(\mathcal{B}^{-1}\mathcal{A})$ because the spectrum is a closed set.

Proof Theorem 3.2.

The proof of Theorem 3.2 is analogous to the proof of Theorem 3.1 in [14]. As was explained in detail in [13, Section 3.2], due to the use of the Hall's theorem for bipartite graphs (see, e.g., [4, Theorem 5.2]), it is sufficient to prove the statement formulated in the following lemma (cf. [14, Lemma 3.3]).

Lemma 3.5. *Assume that $k(x)$ and $g(x)$ are bounded and piecewise continuous functions, and that $g(x)$ is uniformly positive. Let the matrix $\mathbf{B}_n^{-1}\mathbf{A}_n$ be given by* (6) *and* (7)*. Let, moreover, $\mathcal{J} \subset \{1, \ldots, n\}$ and let $\mathcal{T}_{\mathcal{J}} := \bigcup_{j \in \mathcal{J}} \mathcal{T}_j$ be the union of the supports of the basis functions ϕ_j, $j \in \mathcal{J}$. Then there are at least $|\mathcal{J}|$ eigenvalues of the matrix $\mathbf{B}_n^{-1}\mathbf{A}_n$ located in the interval*

$$\left[\inf_{x \in \mathcal{T}_{\mathcal{J}}} \frac{k(x)}{g(x)}, \sup_{x \in \mathcal{T}_{\mathcal{J}}} \frac{k(x)}{g(x)} \right]. \tag{29}$$

Proof. Following the proof of Lemma 3.3 in [14], consider, for any set of indices $\mathcal{J} \subset \{1, \ldots, n\}$, the (local) perturbation $\tilde{k}_{\mathcal{J}}(x)$ of the coefficient function $k(x)$,

$$\tilde{k}_{\mathcal{J}}(x) = \begin{cases} K \cdot g(x) & \text{for} \quad x \in \mathcal{T}_{\mathcal{J}}, \\ k(x) & \text{elsewhere}, \end{cases} \tag{30}$$

where K is a an appropriate scalar to be determined below. Analogously to (6), the matrix $\tilde{\mathbf{A}}_{\mathcal{J},n}$ obtained by the discretization of the associated perturbed operator $\tilde{\mathcal{A}}_{\mathcal{J}}$ is given by

$$[\tilde{\mathbf{A}}_{\mathcal{J},n}]_{lj} = \left\langle \tilde{\mathcal{A}}_{\mathcal{J},n}\phi_j, \phi_l \right\rangle = \int_\Omega \tilde{k}_{\mathcal{J}}(x) \nabla\phi_j \cdot \nabla\phi_l \, .$$

The simple observation (here e_j is the j-th vector of the standard Euclidean basis)

$$\tilde{\mathbf{A}}_{\mathcal{J},n}\mathbf{e}_j = K\,\mathbf{B}_n\mathbf{e}_j, \quad j \in \mathcal{J},$$

shows that K is an eigenvalue of the matrix $\mathbf{B}_n^{-1}\tilde{\mathbf{A}}_{\mathcal{J},n}$ with multiplicity of at least $|\mathcal{J}|$.

By similarity transformations, the spectrum of $\mathbf{B}_n^{-1}\mathbf{A}_n$ equals the spectrum of $\mathbf{B}_n^{-1/2}\mathbf{A}_n\mathbf{B}_n^{-1/2}$, and the spectrum of $\mathbf{B}_n^{-1}\tilde{\mathbf{A}}_{\mathcal{J},n}$ is equal to the spectrum of $\mathbf{B}_n^{-1/2}\tilde{\mathbf{A}}_{\mathcal{J},n}\mathbf{B}_n^{-1/2}$. Using the standard perturbation result for symmetric matrices (see, e.g., [28, Corollary 4.9, p. 203]), there are at least $|\mathcal{J}|$ eigenvalues of $\mathbf{B}_n^{-1}\mathbf{A}_n$ in the interval

$$[K + \theta_{min}, K + \theta_{max}] \subset [K - \Theta, K + \Theta], \tag{31}$$

where $\Theta = \max\{|\theta_{min}|, |\theta_{max}|\}$ and θ_{min} and θ_{max} denote the smallest and largest eigenvalues, respectively, of the perturbation matrix $\mathbf{B}_n^{-1}(\mathbf{A}_n - \tilde{\mathbf{A}}_{\mathcal{J},n})$.

The Rayleigh quotient for an eigenvalue-eigenvector pair $(\theta, \mathbf{q})$ of the perturbation matrix, with the associated eigenfunction $q = \sum_{j=1}^N \nu_j\phi_j$, $\mathbf{q}^T = [\nu_1, \ldots, \nu_N]$, satisfies

$$\begin{aligned} \theta &= \frac{\mathbf{q}^T(\mathbf{A}_n - \tilde{\mathbf{A}}_{\mathcal{J},n})\mathbf{q}}{\mathbf{q}^T\mathbf{B}_n\mathbf{q}} = \frac{\langle (\mathcal{A} - \tilde{\mathcal{A}}_{\mathcal{J}})q, q\rangle}{\langle \mathcal{B}q, q\rangle} \\ &= \frac{\int_\Omega (k(x) - \tilde{k}_{\mathcal{J}}(x))\nabla q \cdot \nabla q \, dx}{\int_\Omega g(x)\|\nabla q\|^2 \, dx} = \frac{\int_{\mathcal{T}_{\mathcal{J}}} (k(x) - Kg(x))\|\nabla q\|^2 \, dx}{\int_\Omega g(x)\|\nabla q\|^2 \, dx} \\ &= \frac{\int_{\mathcal{T}_{\mathcal{J}}} \left(\frac{k(x)}{g(x)} - K\right) g(x)\|\nabla q\|^2 \, dx}{\int_\Omega g(x)\|\nabla q\|^2 \, dx}, \end{aligned}$$

where we used the fact that $\tilde{k}_{\mathcal{J}}(x) = k(x)$ for $x \in \Omega \setminus \mathcal{T}_{\mathcal{J}}$; see (30). Using the uniform positivity of $g(x)$,

$$|\theta| \leq \sup_{x \in \mathcal{T}_{\mathcal{J}}} \left| \frac{k(x)}{g(x)} - K \right| . \tag{32}$$

Substituting (32) into (31) yields the existence of at least $|\mathcal{J}|$ eigenvalues of $\mathbf{B}_n^{-1}\mathbf{A}_n$ in the interval

$$\left[K - \sup_{x\in\mathcal{T}_\mathcal{J}} \left| \frac{k(x)}{g(x)} - K \right|, \; K + \sup_{x\in\mathcal{T}_\mathcal{J}} \left| \frac{k(x)}{g(x)} - K \right| \right]. \tag{33}$$

Setting $K = \frac{1}{2}\left(\inf_{x\in\mathcal{T}_\mathcal{J}} \frac{k(x)}{g(x)} + \sup_{x\in\mathcal{T}_\mathcal{J}} \frac{k(x)}{g(x)} \right)$ finishes the proof. □

4 Abstract setting.

This section investigates numerical approximations of the spectrum of preconditioned linear operators within an abstract Hilbert space setting; see, e.g., [24, 17]. Let V be an infinite dimensional real Hilbert space with the inner product

$$(\cdot,\cdot) : V \times V \to \mathbb{R}. \tag{34}$$

Throughout this text, $V^\#$ denotes the dual of V consisting of all linear bounded functionals from V to $\mathbb{R}$, with the associated duality pairing

$$\langle\cdot,\cdot\rangle : V^\# \times V \to \mathbb{R},$$

and the Riesz map

$$\langle\cdot,\cdot\rangle =: (\tau\cdot,\cdot), \quad \tau : V^\# \to V.$$

Consider two bounded linear operators $\mathcal{A}, \mathcal{B} : V \to V^\#$ that are self-adjoint with respect to the duality pairing, and let the operator $\mathcal{B}$ be also coercive. We will investigate whether all the points in the spectrum of the preconditioned operator $\mathcal{B}^{-1}\mathcal{A} : V \to V$ are approximated to an arbitrary accuracy by the eigenvalues of the finite dimensional operators in a sequence $\{\mathcal{B}_n^{-1}\mathcal{A}_n\}$ determined via the Galerkin discretization.

For the problems considered in sections 2 and 3, we obtained concrete expressions for the approximations of the spectrum of $\mathcal{B}^{-1}\mathcal{A}$ in terms of the coefficient functions $g(x)$ and $k(x)$. Such detailed information is, of course, not obtainable for the present abstract setting. We must be content with analyzing whether, or in what sense, the set of eigenvalues of the discretized mapping converges toward the spectrum of the corresponding infinite dimensional operator. Such knowledge is crucial for understanding the convergence behavior of Krylov subspace methods applied to discretized self-adjoint problems. It has not been, in our opinion, addressed in the literature. From the point of view of the spectral theory of self-adjoint operators, the questions can be formulated within the framework of the lower and upper spectral semicontinuity; see, e.g., the detailed references to [19] and [5] given below.

Since $\mathcal{B}^{-1}\mathcal{A}$ is self-adjoint with respect to the inner product

$$(\cdot,\cdot)_{\mathcal{B}} := \langle \mathcal{B}\,\cdot,\cdot\rangle : V \times V \to \mathbb{R}, \tag{35}$$

it is convenient to use this inner product instead of (34) whenever appropriate[11]. The associated norm $\|\cdot\|_{\mathcal{B}}$ is (topologically) equivalent to the norm $\|\cdot\|$ defined by the inner product (34), and the Riesz map $\tau_{\mathcal{B}}$ representing the operator preconditioning is determined by

$$\langle\cdot,\cdot\rangle =: (\tau_{\mathcal{B}}\,\cdot,\cdot)_{\mathcal{B}} = \langle \mathcal{B}\,\tau_{\mathcal{B}}\,\cdot,\cdot\rangle, \quad \text{i.e., } \tau_{\mathcal{B}} = \mathcal{B}^{-1}.$$

This setting covers also the case of the investigated preconditioned operator $\mathcal{B}^{-1}\mathcal{A}$ being continuously invertible on the Hilbert space V of infinite dimension. Therefore its finite dimensional approximations can not converge to it in norm (uniformly). We will instead use Theorem 4.1 (below) that assumes the pointwise (strong) convergence. Its statement reformulates a theorem presented in [19, chapter VIII, § 1.2, Theorem 1.14, p. 431], which is reproduced also in [5, section 5.4, Theorem 5.12, pp. 239-240]. The second monograph also provides several references to related results of J. Descloux and collaborators published earlier; see, in particular, the seminal paper [11]. In terms of the spectral representation of self-adjoint operators, a bit stronger statements were proved in the context of the problem of moments in [30, section III.3, Theorem IX, p. 61] and, more generally, in [19, chapter VIII, § 1.2, Theorem 1.15, p. 432]. Concerning the spectral pollution when computing Galerkin approximations; see, e.g., [27]. The formulations in [19] and [5] require a careful study of parts of the books. We will therefore, for the sake of convenience, include a proof of the following theorem in appendix A together with a brief review of the related concepts of convergence of operators.

Theorem 4.1 (Approximation of the spectrum of self-adjoint operators)**.** *Let $\mathcal{Z}$ be a linear self-adjoint operator on a Hilbert space V and let $\{\mathcal{Z}_n\}$ be a sequence of linear self-adjoint operators on V converging to $\mathcal{Z}$ pointwise (strongly). Then, for any point $\lambda \in \mathrm{sp}(\mathcal{Z})$ in the spectrum of $\mathcal{Z}$, and for any of its neighbourhoods, the intersection of the spectrum $\mathrm{sp}(\mathcal{Z}_j)$ with this neighbourhood is, for j sufficiently large, nonempty.*

Using this theorem and the Hilbert space V equipped with the inner product (35), it remains, within our setting, to show that the self-adjoint operators $\mathcal{Z}_n$, which arise from $\mathcal{B}_n^{-1}\mathcal{A}_n$ by extending it to the whole space V, converge pointwise to the original self-adjoint operator[12]

$$\mathcal{Z} := \mathcal{B}^{-1}\mathcal{A}.$$

[11] Since $\mathcal{B}^{-1}\mathcal{A} = (\tau\mathcal{B})^{-1}(\tau\mathcal{A})$, one can as an alternative investigate approximations of the spectrum of the symmetrized operator $(\tau\mathcal{B})^{-1/2}\tau\mathcal{A}(\tau\mathcal{B})^{-1/2}$ that is self-adjoint with respect to the inner product (34).

[12] Recall the assumptions on $\mathcal{A}$ and $\mathcal{B}$, which guarantee that $\mathcal{B}$ is continuously invertible and that $\mathcal{B}^{-1}\mathcal{A}$ is self-adjoint with respect to the inner product (35).

This can be done using standard tools. The discretization will be based on a sequence of subspaces $\{V_n\}$, $V_n \subset V$, satisfying the approximation property (8)[13]

$$\lim_{n\to\infty} \inf_{v\in V_n} \|w - v\| = 0 \quad \text{for all } w \in V, \tag{36}$$

see, e.g., [1, relation (8)]. Note that (36) typically yields that Galerkin discretizations of boundary value problems are convergent; see also [24, chapter 9, relation (9.8)]. For a concrete problem one could, for example, use Lagrange elements of appropriate order to construct the subspaces $\{V_n\}$ such that (36) holds.

Consider a basis $\Phi_n = (\phi_1^{(n)}, \ldots, \phi_n^{(n)})$ of the n-dimensional subspace $V_n \subset V$. Then the Galerkin discretizations $\mathcal{A}_n$ and $\mathcal{B}_n$ of the operators $\mathcal{A}$ and $\mathcal{B}$ are determined by (see [17, section 4.1] and [24, chapter 6]),

$$\langle \mathcal{A}_n w, v\rangle := \langle \mathcal{A} w, v\rangle \quad \text{and} \quad \langle \mathcal{B}_n w, v\rangle := \langle \mathcal{B} w, v\rangle, \quad \text{for all } w, v \in V_n. \tag{37}$$

Their matrix representations are given by

$$\mathbf{A}_n = \left(\langle \mathcal{A}\phi_j^{(n)}, \phi_i^{(n)}\rangle\right)_{i,j=1,\ldots,n}, \tag{38}$$

and

$$\mathbf{B}_n = \left(\langle \mathcal{B}\phi_j^{(n)}, \phi_i^{(n)}\rangle\right)_{i,j=1,\ldots,n}. \tag{39}$$

The spectrum of the discretized operator $\mathcal{B}_n^{-1}\mathcal{A}_n : V_n \to V_n$ is given by the eigenvalues of its matrix representation $\mathbf{B}_n^{-1}\mathbf{A}_n$. The operator $\mathcal{B}_n^{-1}\mathcal{A}_n$ is self-adjoint with respect to the inner product (35), and the matrix $\mathbf{B}_n^{-1}\mathbf{A}_n$ is self-adjoint with respect to the algebraic inner product $(\mathbf{x}, \mathbf{y})_{\mathbf{B}_n} := \mathbf{y}^* \mathbf{B}_n \mathbf{x}$.

Using the orthogonal projection

$$\Pi_{\mathcal{B}}^n : V \to V_n,$$

where the orthogonality is determined by the inner product (35), $\mathcal{B}_n^{-1}\mathcal{A}_n$ is extended to the whole space V,

$$\mathcal{Z}_n := \mathcal{B}_n^{-1}\mathcal{A}_n \,\Pi_{\mathcal{B}}^n : V \to V_n \subset V. \tag{40}$$

For all $w, v \in V_n$ we obtain

$$\begin{aligned}
(\mathcal{Z}_n w, v)_{\mathcal{B}} &= (\mathcal{B}_n^{-1}\mathcal{A}_n \Pi_{\mathcal{B}}^n w, v)_{\mathcal{B}} = (\mathcal{B}_n^{-1}\mathcal{A}_n w, v)_{\mathcal{B}} \\
&= \langle \mathcal{B}\mathcal{B}_n^{-1}\mathcal{A}_n w, v\rangle \\
&= \langle \mathcal{B}_n\mathcal{B}_n^{-1}\mathcal{A}_n w, v\rangle \\
&= \langle \mathcal{A}_n w, v\rangle \\
&= \langle \mathcal{A} w, v\rangle \\
&= (\mathcal{B}^{-1}\mathcal{A} w, v)_{\mathcal{B}} \\
&= (\mathcal{Z} w, v)_{\mathcal{B}},
\end{aligned}$$

[13] Since the norms $\|\cdot\|_{\mathcal{B}}$ and $\|\cdot\|$ are equivalent, it does not matter which of these we use in (36).

which implies that

$$(\mathcal{Z}w - \mathcal{Z}_n w, v)_{\mathcal{B}} = 0.$$

Therefore

$$\mathcal{Z}_n = \Pi_{\mathcal{B}}^n \mathcal{Z} \Pi_{\mathcal{B}}^n$$

and it represents the Galerkin discretization of the operator $\mathcal{Z}$ trivially extended to the whole space V, where the Galerkin orthogonality is defined by the inner product (35).

The spectrum of $\mathcal{Z}_n$ consists of the spectrum of $\mathcal{B}_n^{-1}\mathcal{A}_n$ and the point $\{0\}$ that plays no role in the further considerations.

We need to show that $\mathcal{Z}_n$ is self-adjoint with respect to the inner product (35). Using, for any $u, w \in V$, the associated orthogonal decompositions $u = \Pi_{\mathcal{B}}^n u + u^\perp$ and $w = \Pi_{\mathcal{B}}^n w + w^\perp$, we can write

$$\begin{aligned}
\langle \mathcal{B}\mathcal{Z}_n u, w\rangle &= \langle \mathcal{B}\mathcal{B}_n^{-1}\mathcal{A}_n \Pi_{\mathcal{B}}^n u, w\rangle = \langle \mathcal{B}\mathcal{B}_n^{-1}\mathcal{A}_n \Pi_{\mathcal{B}}^n u, \Pi_{\mathcal{B}}^n w\rangle + \langle \mathcal{B}\mathcal{B}_n^{-1}\mathcal{A}_n \Pi_{\mathcal{B}}^n u, w^\perp\rangle \\
&= \langle \mathcal{B}_n\mathcal{B}_n^{-1}\mathcal{A}_n \Pi_{\mathcal{B}}^n u, \Pi_{\mathcal{B}}^n w\rangle = \langle \mathcal{A}_n \Pi_{\mathcal{B}}^n u, \Pi_{\mathcal{B}}^n w\rangle = \langle \mathcal{A}_n \Pi_{\mathcal{B}}^n w, \Pi_{\mathcal{B}}^n u\rangle \\
&= \langle \mathcal{B}_n\mathcal{B}_n^{-1}\mathcal{A}_n \Pi_{\mathcal{B}}^n w, \Pi_{\mathcal{B}}^n u\rangle = \langle \mathcal{B}\mathcal{B}_n^{-1}\mathcal{A}_n \Pi_{\mathcal{B}}^n w, \Pi_{\mathcal{B}}^n u\rangle \\
&= \langle \mathcal{B}\mathcal{B}_n^{-1}\mathcal{A}_n \Pi_{\mathcal{B}}^n w, \Pi_{\mathcal{B}}^n u\rangle + \langle \mathcal{B}\mathcal{B}_n^{-1}\mathcal{A}_n \Pi_{\mathcal{B}}^n w, u^\perp\rangle \\
&= \langle \mathcal{B}\mathcal{B}_n^{-1}\mathcal{A}_n \Pi_{\mathcal{B}}^n w, u\rangle = \langle \mathcal{B}\mathcal{Z}_n w, u\rangle,
\end{aligned}$$

which gives the self-adjointness.

Summarizing, the sequence of subspaces $\{V_n\}$ determines a sequence of self-adjoint operators $\{\mathcal{Z}_n\}$ defined on the whole space V. The dimension of the ranges of these operators is finite, but increases as n increases. It remains to prove that $\{\mathcal{Z}_n\}$ converges pointwise to $\mathcal{Z}$.

As pointed out by Miroslav Bačák, the following theorem is valid for an arbitrary linear bounded operator Z and for its (extended) Galerkin discretization Z_n. Therefore, here $\mathcal{A}$ is not assumed to be self-adjoint. This assumption is, however, needed in the subsequent corollary.

Theorem 4.2 (Pointwise convergence)**.** *Let $\mathcal{Z} = \mathcal{B}^{-1}\mathcal{A}$ be a bounded linear operator on a Hilbert space V, where $\mathcal{A}, \mathcal{B} : V \to V^\#$ are bounded linear operators and $\mathcal{B}$ is, in addition, self-adjoint with respect to the duality pairing and coercive. Let $\{\mathcal{Z}_n\}$ be the sequence of linear operators defined in (40). Assume that the sequence of subspaces $\{V_n\}$ satisfies the approximation property* (36)*. Then the sequence $\{\mathcal{Z}_n\}$ converges pointwise (strongly) to $\mathcal{Z}$, i.e., for all $w \in V$*

$$\lim_{n\to\infty} \|\mathcal{Z}w - \mathcal{Z}_n w\| = 0.$$

Proof. For an arbitrary $w \in V$,

$$\begin{aligned}
\|\mathcal{Z}w - \mathcal{Z}_n w\| &= \|\mathcal{Z}w - \Pi_{\mathcal{B}}^n \mathcal{Z} \Pi_{\mathcal{B}}^n w\| \\
&\le \|\mathcal{Z}w - \Pi_{\mathcal{B}}^n \mathcal{Z} w\| + \|\Pi_{\mathcal{B}}^n \mathcal{Z}w - \Pi_{\mathcal{B}}^n \mathcal{Z} \Pi_{\mathcal{B}}^n w\| \\
&\le \|\mathcal{Z}w - \Pi_{\mathcal{B}}^n \mathcal{Z} w\| + \|\mathcal{Z}\|\|w - \Pi_{\mathcal{B}}^n w\|.
\end{aligned}$$

Using the equivalent form of the approximation property (36),

$$\lim_{n\to\infty} \|v - \Pi_{\mathcal{B}}^n v\| = 0 \quad \text{for all } v \in V,$$

the proof is finished. □

Theorems 4.1 and 4.2 immediately give the final corollary.

Corollary 4.3 (Spectral approximation)**.** *Consider an infinite dimensional Hilbert space V, its dual $V^\#$, and bounded linear operators $\mathcal{A}, \mathcal{B} : V \to V^\#$ that are self-adjoint with respect to the duality pairing, with $\mathcal{B}$ being also coercive. Consider further a sequence of subspaces $\{V_n\}$ of V satisfying the approximation property* (36)*.*

Let the sequences of matrices $\{\mathbf{A}_n\}$ and $\{\mathbf{B}_n\}$ be defined by (37) - (39). Then all points in the spectrum of the preconditioned operator

$$\mathcal{B}^{-1}\mathcal{A} : V \to V$$

are approximated to an arbitrary accuracy by the eigenvalues of the preconditioned matrices in the sequence $\{\mathbf{B}_n^{-1}\mathbf{A}_n\}$. That is, for any point $\lambda \in \mathrm{sp}(\mathcal{B}^{-1}\mathcal{A})$ and any $\epsilon > 0$, there exists n^* such that for all $n \geq n^*$ the preconditioned matrix $\mathbf{B}_n^{-1}\mathbf{A}_n$ has an eigenvalue $\lambda_{j(n)}$ satisfying $|\lambda - \lambda_{j(n)}| < \epsilon$.

5 Numerical experiments

This paper primarily deals with the numerical approximation of the spectrum of infinite-dimensional self-adjoint operators by the eigenvalues of the preconditioned matrices arising from discretizations. The numerical experiments will aim to illustrate our main theoretical result stating that, within the given class of PDE problems, the whole continuous spectrum is approximated with arbitrary accuracy, although the individual spectral points are approached with different speed. We are not going to focus on the approximation of the matrix eigenvalues by easy-to-compute intervals or nodal values associated with the involved coefficient functions; see Theorem 3.2 and Corollary 3.3. These issues were numerically illustrated in detail in [14]. Nevertheless, the nodal values will appear in the easy-to-compute approximations of the *cumulative spectral density* (see [23, Appendix C]) that we will discuss next.

As mentioned in the introduction with references to the literature, the behavior of Krylov subspace methods like CG or MINRES is determined by the distribution function $\omega(\lambda)$, and it can be described via the convergence of the Gauss-quadrature approximations $\omega^{(n)}(\lambda)$ to $\omega(\lambda)$. For a given preconditioned matrix, $\omega(\lambda)$ is a staircase function with the points of increase at the matrix eigenvalues and the height of the stairs equal to the squared size of the components of the normalized residual in the associated

(orthogonal) invariant subspaces; see, e.g., [14, relation (2.8)]. Given only a matrix without a residual vector, we can consider an auxiliary normalized residual with all its components in the invariant subspaces of equal size. Following [23, Appendix C], we denote the resulting staircase function $\psi(\lambda)$. This function, called *the cumulative spectral density*, is of great importance in physics dealing with the so called *density of states*; see [23] for a survey of numerical approximations, as well as for references to literature addressing applications. With this connection in mind, we use the cumulative spectral density $\psi(\lambda)$ for the illustration of our results.

We use the differential operators $-\nabla\cdot(k(x)\nabla u)$ and $-\nabla\cdot(g(x)\nabla u)$ on the square domain $\Omega = (-1,1)\times(-1,1)$, and we exploit Matlab's PDE-Toolbox to compute scalars λ satisfying

$$\mathbf{A}_n\mathbf{v} = \lambda\mathbf{B}_n\mathbf{v}, \tag{41}$$

where $\mathbf{A}_n$ and $\mathbf{B}_n$ denote the stiffness matrices defined in (6) and (7), respectively. The discretization uses grids based on standard uniform triangulations with three levels of refinements. This will result in linear algebraic systems with 54, 163, and 724 degrees of freedom. For the sake of convenient language, we will use the words "fine" and "very fine" for the refinements even though one could argue that the refined meshes are still rather coarse. Using further refinements makes it difficult to present informative plots.

Whereas our theoretical study concerns problems with homogeneous Dirichlet boundary conditions, we employed homogeneous Neumann boundary conditions in the numerical experiments below. This was done for the sake of completeness. One can show that the appropriately reformulated results presented in sections 2 and 3 also hold in the case of the homogeneous Neumann boundary conditions. While the operators $\mathcal{A}$ and $\mathcal{B}$ are, with the right choice of the associated space, continuously invertible (see [15, section 5]), the matrices representing their discretizations are singular, with

$$\mathbf{0} = \mathbf{A}_n\mathbf{c} = \lambda\mathbf{B}_n\mathbf{c} = \mathbf{0}$$

for any constant vector $\mathbf{c}$ and any scalar λ. Therefore we can not use the inverses and must reformulate the statements concerning the eigenvalues of the preconditioned matrices in terms of the generalized eigenvalue problem. Matlab handles this matter, i.e., (41) is solved subject to the constraint that $\mathbf{v}$ must not belong to the intersection of the nullspaces of $\mathbf{A}_n$ and $\mathbf{B}_n$.

We will consider the following two examples:

Example 1

$$\begin{aligned} k(x,y) &= (1+50\exp(-5(x^2+y^2)))(2+\sin(x+y)),\\ g(x,y) &= 1+50\exp(-5(x^2+y^2)),\\ r(x,y) &= k(x,y)/g(x,y) = 2+\sin(x+y). \end{aligned}$$

Example 2

Compared with Example 1, we interchange the formulas/expressions for $g(x, y)$ and $r(x, y)$, i.e.,

$$\begin{aligned} k(x, y) &= (1 + 50\exp(-5(x^2 + y^2)))(2 + \sin(x + y)), \\ g(x, y) &= 2 + \sin(x + y), \\ r(x, y) &= k(x, y)/g(x, y) = 1 + 50\exp(-5(x^2 + y^2)). \end{aligned}$$

The spectrum of the infinite dimensional preconditioned operator $\mathrm{sp}(\mathcal{B}^{-1}\mathcal{A})$ is given for Example 1 by the interval $[1, 3]$ and for Example 2 by the interval $[1 + 50\exp(-10), 51] \approx [1.0023, 51]$; see Theorem 3.1.

The results for Example 1 are presented in Figure 1 and Table 1, and for Example 2 in Figure 2 and Table 2. The left parts of the figures illustrate how the the cumulative spectral density $\psi(\lambda)$ refines with grid refinement. It clearly covers the whole spectrum. Example 2 shows that different parts of the spectrum indeed can be approximated with different speed, as suggested by our theoretical results. The tables illustrate how the intervals in the spectrum that contain no matrix eigenvalue shrink with grid refinement. We can observe that the rate of this shrinking is roughly the same for both the maximal and the average interval, which is again to be expected.

The right parts of these figures illustrate approximations of the preconditioned matrix eigenvalues by the nodal values of the function $r(x) = k(x)/g(x)$ (both are sorted in the increasing order to respect the link to the cumulative spectral density). These results are as one could have anticipated from Theorem 3.1, Theorem 3.2 and Corollary 3.3; for more detailed illustrations of the convergence of the nodal values to the eigenvalues we refer to [14].

Here we emphasize that, for the given class of PDE problems, the cumulative spectral density can be approximated at a negligible cost (in comparison to the approaches reviewed in [23]) using the nodal values of the function $r(x)$. Whether or not this approximation can be competitive in terms of accuracy and whether it can be useful in physics, studying density of states, goes behind the scope of this paper, but it seems worth further

Mesh	coarse	fine	very fine
Length of the longest interval	0.1593	0.0524	0.0134
Average length of the intervals	0.0357	0.0121	0.0028

Table 1: Example 1. The length of the longest, as well as the average length, of the intervals not containing an eigenvalue.

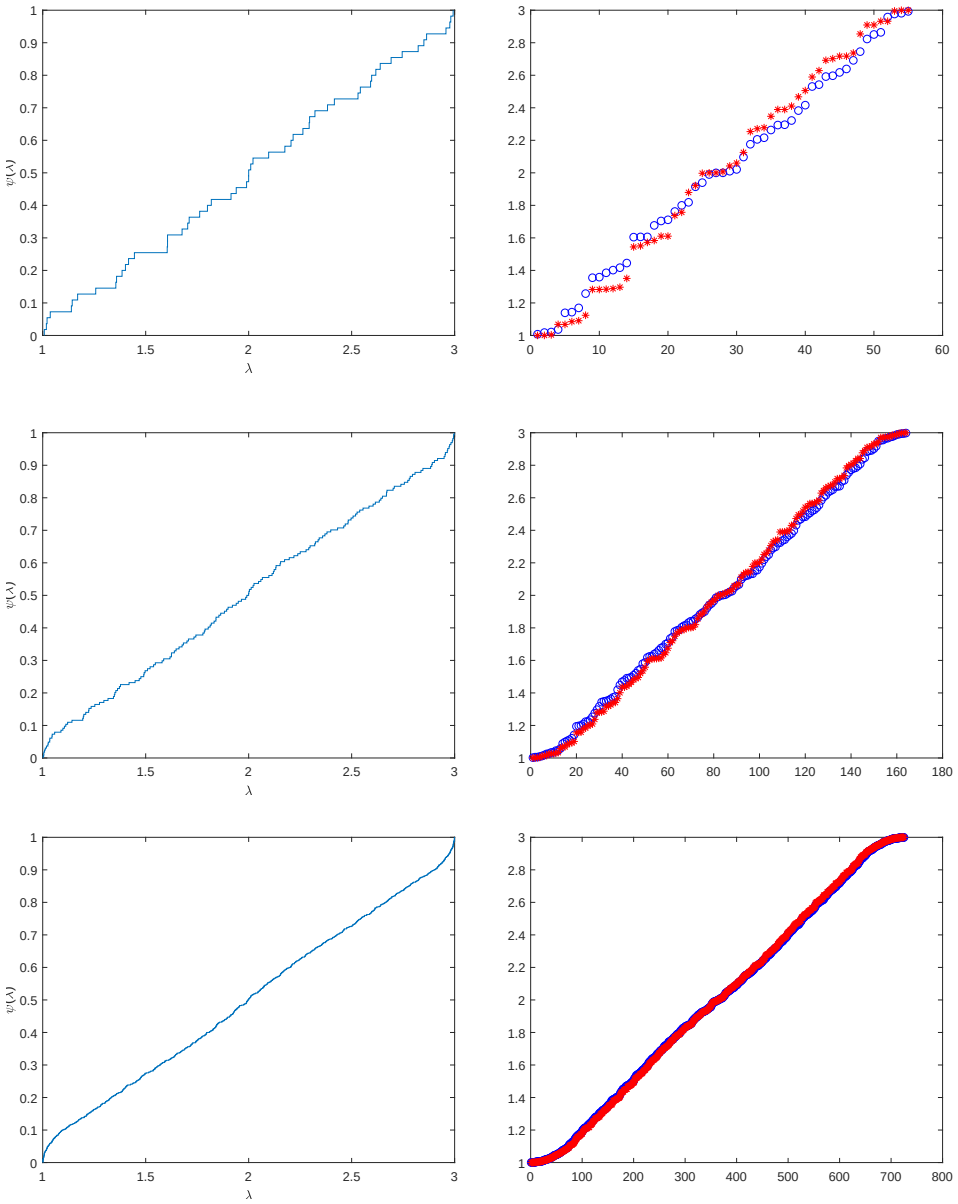

Figure 1: Example 1. Left: cumulative spectral densities. Right: generalized eigenvalues (blue circles) and nodal values of $r(x, y)$ (red asterisks). The top, middle and lower rows show results obtained with the coarse, fine and very fine meshes, respectively.

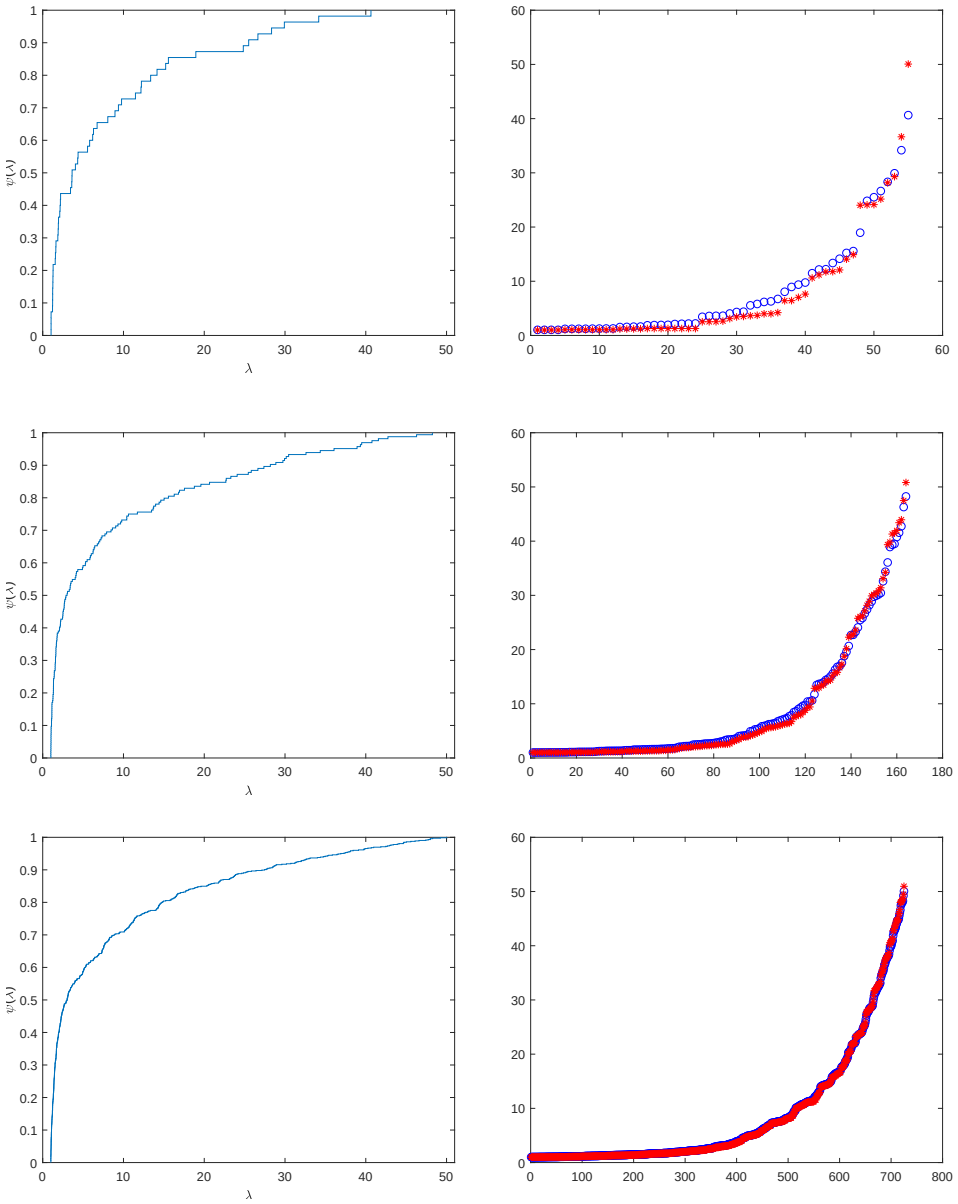

Figure 2: Example 2. Left: cumulative spectral densities. Right: generalized eigenvalues (blue circles) and nodal values of $r(x, y)$ (red asterisks). The top, middle and lower rows show results obtained with the coarse, fine and very fine meshes, respectively.

Mesh	coarse	fine	very fine
Length of the longest interval	10.3584	3.5443	0.9580
Average length of the intervals	0.8928	0.3030	0.0689

Table 2: Example 2. The length of the longest, as well as the average length, of the intervals not containing an eigenvalue.

investigations.

6 Conclusions and further work.

We have extended our earlier results [26, 14, 15], addressing Laplacian preconditioning, to preconditioners defined in terms of more general elliptic differential operators, and proved that the *entire* spectrum of any operator in the form $\mathcal{B}^{-1}\mathcal{A}$ can be approximated with arbitrary accuracy by the eigenvalues of the associated discretized mappings. In the language of the spectral approximation theory, Corollary 4.3 provides the lower semicontinuity of the spectrum of the preconditioned operator $\mathcal{B}^{-1}\mathcal{A}$. Here $\mathcal{A}, B : V \to V^{\#}$ are linear, bounded and self-adjoint operators defined on an infinite dimensional Hilbert space V, and $V^{\#}$ denotes the dual space. Moreover, the operator $\mathcal{B}$ is also coercive. The result summarized in Theorem 3.4 shows both the lower and upper semicontinuity within the given particular finite element setting. Our analysis differs significantly from the classical investigations of the point spectrum of second order differential operators, which is typically done within the framework of compact (solution) operators.

In our opinion, these results yield a new perspective on the importance of the continuous spectrum of preconditioned second order differential operators and the computational issues concerning preconditioning and the use of Krylov subspace methods for self-adjoint problems. There are several unanswered questions: For example, are the results presented in section 3 also valid if the coefficient functions $k(x)$ and $g(x)$ are replaced by symmetric conductivity tensors $K(x)$ and $G(x)$ (where $G(x)$ is, in addition, uniformly positive definite), respectively? The way of how to use the presented ideas to speed-up practical calculations is also a subject of further work.

Acknowledgments.

The authors are indebted to Miroslav Bačák for strengthening Theorem 4.2. We also wish to thank Roland Herzog, David Krejčiřík, Václav Kučera, Josef Málek and Ivana Pultarová for stimulating discussions during the work on this paper. Finally, we would like to thank the anonymous referee for many suggestions that has led to improvements of the text.

A Approximations of the spectrum of self-adjoint operators.

Using [5, Chapter 3] and [19, Chapter 8], we first recall several results concerning the convergence of linear self-adjoint operators defined on infinite dimensional Hilbert spaces. By the Hellinger-Toeplitz Theorem (see [7, Theorem 5.7.2, p. 260]), any linear self-adjoint operator $\mathcal{Z} : V \to V$ on a Hilbert space V is closed and, according to the Banach closed-graph theorem, bounded (continuous).

Consider a bounded linear operator $\mathcal{G} : V \to V$ (not necessarily self-adjoint, therefore we for the moment change the notation) and a sequence of its bounded linear approximations $\{\mathcal{G}_n\}, \mathcal{G}_n : V \to V$, that can converge to $\mathcal{G}$ in different ways:

- pointwise (strongly), i.e., $\mathcal{G}_n \xrightarrow{p} \mathcal{G}$

$$\text{iff for all } x \in V, \quad \lim_{n\to\infty} \|\mathcal{G}x - \mathcal{G}_n x\| = 0\,;$$

- uniformly (in norm), iff $\lim_{n\to\infty} \|\mathcal{G} - \mathcal{G}_n\| = 0\,;$
- stably, i.e. $\mathcal{G}_n \xrightarrow{s} \mathcal{G}$ iff
  - $\mathcal{G}_n \xrightarrow{p} \mathcal{G}$, and
  - the inverse operators $\{\mathcal{G}_n^{-1}\}$ are uniformly bounded, i.e., for some $C > 0$, $\|\mathcal{G}_n^{-1}\| \le C$ for all n.

Clearly, uniform convergence implies pointwise convergence, but the converse implication does not hold. Since the class of compact operators is closed with respect to uniform convergence, the uniform convergence concept can not be used to investigate the convergence of compact to non-compact operators, such as to bounded continuously invertible operators defined on infinite dimensional Hilbert spaces.

The spectral theory for bounded linear operators is based on the concept of the operator resolvent

$$\mathcal{R}(\mu) := (\mu\mathcal{I} - \mathcal{G})^{-1}$$

and on the resolvent set

$$\rho(\mathcal{G}) := \{\mu \in \mathbb{C};\, \mu\mathcal{I} - \mathcal{G} \text{ has a bounded inverse}\}\,. \tag{42}$$

It is interesting to notice that, for any $\mu \in \rho(\mathcal{G})$, a sequence of shifted operators $\{\mu\mathcal{I} - \mathcal{G}_n\}$ converge to $\mu\mathcal{I} - \mathcal{G}$ stably if, and only if, $\{\mathcal{G}_n\}$ and

the resolvents $\{\mathcal{R}_n(\mu)\}, \mathcal{R}_n(\mu) := (\mu\mathcal{I} - \mathcal{G}_n)^{-1}$, converge to $\mathcal{G}$ and $\mathcal{R}(\mu)$ pointwise, respectively, i.e.,[14]

$$\mu\mathcal{I} - \mathcal{G}_n \xrightarrow{s} \mu\mathcal{I} - \mathcal{G} \quad \text{iff} \quad \mathcal{G}_n \xrightarrow{p} \mathcal{G} \text{ and } \mathcal{R}_n(\mu) \xrightarrow{p} \mathcal{R}(\mu). \tag{43}$$

Indeed, using the resolvent identity

$$\mathcal{R}_n(\mu) - \mathcal{R}(\mu) = \mathcal{R}_n(\mu)\,(\mathcal{G}_n - \mathcal{G})\,\mathcal{R}(\mu),$$

the right implication follows immediately from the definition of stable convergence. Conversely, from the pointwise convergence of $\mathcal{R}_n(\mu)$ and the uniform boundedness principle (Banach–Steinhaus theorem) we conclude that $\{\mathcal{R}_n(\mu)\} = \{(\mu\mathcal{I} - \mathcal{G}_n)^{-1}\}$ is uniformly bounded and the result follows.

We will now present the proof of Theorem 4.1; cf. also [19, chapter VIII, § 1.2, Theorem 1.14, p. 431] and [5, chapter 5, section 4, Theorem 5.12, p. 239-240].

Theorem A.1 (Approximation of the spectrum of self-adjoint operators)**.** *Let $\mathcal{Z}$ be a linear self-adjoint operator on a Hilbert space V and let $\{\mathcal{Z}_n\}$ be a sequence of linear self-adjoint operators on V converging to $\mathcal{Z}$ pointwise (strongly). Then, for any point $\lambda \in \mathrm{sp}(\mathcal{Z})$ in the spectrum of $\mathcal{Z}$, and for any of its neighbourhoods, the intersection of the spectrum $\mathrm{sp}(\mathcal{Z}_j)$ with this neighbourhood is, for j sufficiently large, nonempty.*

Proof. Consider any point $\lambda \in \mathrm{sp}(\mathcal{Z}) \subset \mathbb{R}$ in the spectrum of $\mathcal{Z}$. Then, for any $\varepsilon > 0$, the point $\mu := \lambda + \iota\varepsilon$, where ι is the complex unit, belongs to the resolvent set $\rho(\mathcal{Z})$ (because $\mathcal{Z}$ is self-adjoint). For self-adjoint operators, the norm of the resolvent at any point in the resolvent set is equal to the inverse of the distance of the given point to the spectrum (see, e.g., [5, Proposition 2.32]). Therefore, for all n,

$$\|\mathcal{R}(\mu)\| := \|(\mu\mathcal{I} - \mathcal{Z})^{-1}\| = \frac{1}{\mathrm{dist}(\mu, \mathrm{sp}(\mathcal{Z}))} = \frac{1}{\varepsilon}\,, \tag{44}$$

$$\|\mathcal{R}_n(\mu)\| := \|(\mu\mathcal{I} - \mathcal{Z}_n)^{-1}\| = \frac{1}{\mathrm{dist}(\mu, \mathrm{sp}(\mathcal{Z}_n))} \leq \frac{1}{\varepsilon}\,. \tag{45}$$

The inequality in (45) follows from the assumption that $\{\mathcal{Z}_n\}$ is a sequence of self-adjoint operators, i.e., $\mathrm{sp}(\mathcal{Z}_n) \subset \mathbb{R}$. Note that inequality (45) provides the uniform boundedness of $\{\mathcal{R}_n(\mu)\}$, which, together with the pointwise convergence of $\{\mathcal{Z}_n\}$, yields that

$$\mu\mathcal{I} - \mathcal{Z}_n \xrightarrow{s} \mu\mathcal{I} - \mathcal{Z}.$$

Using (43), we thus have the pointwise convergence of $\{\mathcal{R}_n(\mu)\}$, i.e., for any $x \in V, ||x|| = 1$,

$$\|\mathcal{R}(\mu)x\| = \lim_{n\to\infty} \|\mathcal{R}_n(\mu)x\|.$$

[14] See [5, Lemma 3.16].

Consider a fixed $x \in V, ||x\| = 1$, such that

$$\frac{1}{2\varepsilon} \leq \|\mathcal{R}(\mu)x\| \leq \frac{1}{\varepsilon}.$$

(The existence of such a $x \in V$ follows from (44).) Then, from the pointwise convergence, there must exist n^* such that for all $n \geq n^*$

$$\|\mathcal{R}_n(\mu)\| \geq \|\mathcal{R}_n(\mu)x\| \geq \frac{1}{3\varepsilon}.$$

Recall that $\|\mathcal{R}_n(\mu)\| = [\mathrm{dist}(\mu, \mathrm{sp}(\mathcal{Z}_n))]^{-1}$. Therefore there exists a point $\lambda_n \in \mathrm{sp}(\mathcal{Z}_n)$ such that

$$|\lambda_n - \mu| \leq 3\varepsilon \quad \text{and, consequently,} \quad |\lambda_n - \lambda| \leq 4\varepsilon,$$

which finishes the proof. □

References

[1] D. N. Arnold, R. S. Falk, and R. Winther. Finite element exterior calculus: from Hodge theory to numerical stability. *Bull. Amer. Math. Soc. (N.S.)*, 47:281–354, 2010.

[2] O. Axelsson and J. Karátson. Equivalent operator preconditioning for elliptic problems. *Numer. Algorithms*, 50(3):297–380, 2009.

[3] R. Blaheta, S. Margenov, and M. Neytcheva. Uniform estimate of the constant in the strengthened CBS inequality for anisotropic non-conforming FEM systems. *Numer. Lin. Alg. with Appl.*, 11:309–326, 2004.

[4] J. A. Bondy and U. S. R. Murty. *Graph theory with applications*. American Elsevier Publishing Co., Inc., New York, 1976.

[5] F. Chatelin. *Spectral approximation of linear operators*. Academic Press, New York, 1983.

[6] P. G. Ciarlet. *The finite element method for elliptic problems*, volume 40 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002. Reprint of the 1978 original [North-Holland, Amsterdam; MR0520174 (58 #25001)].

[7] P. G. Ciarlet. *Linear and nonlinear functional analysis with applications*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 2013.

[8] M. Colbrook and A. Horning. Specsolve: Spectral methods for spectral measures. *arXiv preprint 2201.01314*, 2022.

[9] M. Colbrook, A. Horning, and A. Towsend. Computing spectral measures of self-adjoint operators. *SIAM Review*, 63:489–524, 2021.

[10] M. Colbrook, B. Roman, and A. Hansen. How to compute spectra with error control. *Phys. Rev. Lett.*, 122:250201, Jun 2019.

[11] J. Descloux, N. Nassif, and J. Rappaz. On spectral approximation. part 1. the problem of convergence. *RAIRO - Analyse Numérique*, 12:97–112, 1978.

[12] V. Faber, Thomas A. Manteuffel, and Seymour V. Parter. On the theory of equivalent operators and application to the numerical solution of uniformly elliptic partial differential equations. *Adv. in Appl. Math.*, 11(2):109–163, 1990.

[13] T. Gergelits. *Krylov Subspace Methods: Analysis and Applications.* PhD thesis, Charles University, 2020.

[14] T. Gergelits, K. A. Mardal, B. F. Nielsen, and Z. Strakoš. Laplacian preconditioning of elliptic PDEs: Localization of the eigenvalues of the discretized operator. *SIAM Journal on Numerical Analysis*, 57(3):1369–1394, 2019.

[15] T. Gergelits, B. F. Nielsen, and Z. Strakoš. Generalized spectrum of second order differential operators. *SIAM Journal on Numerical Analysis*, 58(4):2193–2211, 2020.

[16] R. Hiptmair. Operator preconditioning. *Computers & Mathematics with Applications. An International Journal*, 52(5):699–706, 2006.

[17] J. Hrnčíř, I. Pultarová, and Z. Strakoš. Decomposition of subspaces preconditioning: abstract framework. *Numerical Algorithms*, 83:57–98, 2020.

[18] J. Karátson. Operator preconditioning with efficient applications for elliptic problems. *Cent. Eur. J. Math.*, 10(3):231–249, 2012.

[19] T. Kato. *Perturbation Theory for Linear Operators.* Springer-Verlag, Berlin Heidelberg, 1980.

[20] M. Ladecký, I. Pultarová, and J. Zeman. Guaranteed two-sided bounds on all eigenvalues of preconditioned diffusion and elasticity problems solved by the finite element method. *Applications of Mathematics*, 2020.

[21] R. J. Leute, M. Ladecký, A. Falsafi, I. Jödicke, I. Pultarová, J. Zeman, T. Junge, and L. Pastewka. Elimination of ringing artifacts by finite-element projection in fft-based homogenization. *Journal of Computational Physics*, 453:110931, 2022.

[22] J. Liesen and Z. Strakoš. *Krylov subspace methods: principles and analysis*. Numerical Mathematics and Scientific Computation. Oxford University Press, Oxford, 2012.

[23] L. Lin, Y. Saad, and C. Yang. Approximating spectral densities of large matrices. *SIAM Review*, 58:34–65, 2016.

[24] J. Málek and Z. Strakoš. *Preconditioning and the conjugate gradient method in the context of solving PDEs*. SIAM Spotlight Series, December 2014.

[25] K. A. Mardal and R. Winther. Preconditioning discretizations of systems of partial differential equations. *Numerical Linear Algebra with Applications*, 18(1):1–40, 2011.

[26] B. F. Nielsen, A. Tveito, and W. Hackbusch. Preconditioning by inverting the Laplacian; an analysis of the eigenvalues. *IMA Journal of Numerical Analysis*, 29(1):24–42, 2009.

[27] J. Rappaz, J. Sanchez Hubert, E. Sanchez Palencia, and D. Vassiliev. On spectral pollution in the finite element approximation of thin elastic "membrane" shells. *Numer. Math.*, 75:473–500, 1997.

[28] G. W. Stewart and J. G. Sun. *Matrix perturbation theory*. Computer Science and Scientific Computing. Academic Press, Inc., Boston, MA, 1990.

[29] J. von Neumann. *Mathematical Foundations of Quantum Mechanics*. Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, 1996. Translated from the 1932 German original and with a preface by R. T. Beyer.

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