On the convergence of second-order spectra and multiplicity

By Lyonell Boulton1,* and Michael Strauss2

1 Department of Mathematics and Maxwell Institute for Mathematical Sciences, Heriot-Watt University, Edinburgh EH14 4AS, UK
2 Department of Mathematics and Statistics, University of Strathclyde, Glasgow G1 1XH, UK

The notion of second-order relative spectrum of a self-adjoint operator acting on a Hilbert space has been studied recently in connection with the phenomenon of spectral pollution in the Galerkin method. In this paper we examine how the second-order spectrum encodes precise information about the multiplicity of the isolated eigenvalues of the underlying operator. Our theoretical findings are supported by various numerical experiments on the computation of guaranteed eigenvalue inclusions via finite element bases.

Keywords: second-order spectrum; spectral pollution; spectral exactness

1. Introduction

Let \( A \) be a self-adjoint operator acting on an infinite-dimensional Hilbert space \( \mathcal{H} \) and let \( \lambda \) be an isolated eigenvalue of \( A \). For \( \mathcal{I} \subset \mathbb{R} \) let \( \mathbb{1}_\mathcal{I}(A) = \int_\mathcal{I} \, dE_\mu \), where \( E_\mu \) is the spectral measure associated to \( A \). The numerical estimation of \( \lambda \), whenever \( \text{Tr}\, \mathbb{1}_{(-\infty,\lambda]}(A) = \text{Tr}\, \mathbb{1}_{(\lambda,\infty)}(A) = \infty \), constitutes a serious challenge in computational spectral theory. Indeed, it is well established that classical approaches, such as the Galerkin method, suffer from variational collapse under no further restrictions on the approximating space and therefore might lead to spectral pollution (see Rappaz et al. 1997; Dauge & Suri 2002; Lewin & Séré 2010 and references therein).

The notion of second-order relative spectrum, originated from Davies (1998), has recently allowed the formulation of a general pollution-free strategy for eigenvalue computation. This was proposed by Shargorodsky (2000) and subsequently examined by Levitin & Shargorodsky (2004), Boulton (2007), Boulton & Strauss (2007) and Strauss (in press). Numerical implementations of the general principle very much preserve the spirit of the Galerkin method and have presently been tested on applications from Stokes systems (Levitin & Shargorodsky 2004), solid-state physics (Boulton & Levitin 2007), magnetohydrodynamics (Strauss in press) and relativistic quantum mechanics (Boulton & Boussaid 2010).

In this paper we further examine the potential role of this pollution-free technique for robust computation of spectral inclusions. Our goal is twofold. On the one hand, we establish various abstract properties of limit sets of second-order...
relative spectra. On the other hand, we report on the outcomes of various numerical experiments. Both our theoretical and practical findings indicate that second-order spectra provide reliable information about the multiplicity of any isolated eigenvalue of $A$.

Section 2 is devoted to reformulating some of the concepts around the notion of second-order spectra in order to allow a more general setting. In this framework, we consider the natural notion of algebraic and geometric multiplicity of second-order spectral points (definition 2.2) and establish a ‘second-order’ spectral mapping theorem (lemma 2.6).

In §3 we pursue a detailed analysis of accumulation points of the second-order relative spectra on the real line. Our main contribution (theorem 3.4) is a significant improvement upon similar results previously found. It allows calculation of rigorous convergence rates when the test subspaces are generated by a non-orthogonal basis. Concrete applications include the important case of a finite-element basis, which was not covered by theorem 2.1 of Boulton (2007). Our present approach relies upon a homotopy argument, which yields a precise control on the multiplicity of the second-order spectral points. The argument is reminiscent of the method of proof of Goerisch theorem (see Plum 1990).

Theorem 3.4 also determines the precise manner in which second-order spectra encode information about the multiplicity of points in the spectrum of $A$. When an approximating space is ‘sufficiently close’ to the eigenspace corresponding to an eigenvalue of finite multiplicity, a finite set of conjugate pairs in the second-order spectrum becomes ‘isolated’ and clusters near the eigenvalue. It turns out that the total multiplicity of these conjugate pairs exactly matches the multiplicity of the eigenvalue. This indicates that second-order spectra capture, in a reliable manner, points to the discrete spectrum and their multiplicities, even under the above variational collapse condition. In §4 we examine the practical validity of this statement on benchmark differential operators for subspaces generated by a basis of finite elements.

A final appendix is aimed at finding the minimal region in the complex plane where the limit of second-order spectra is allowed to accumulate. It turns out that, modulo a subset of topological dimension zero, this minimal region is completely determined by the essential spectrum of $A$. This gives an insight on the difficulties involving the problem of finding conditions on the test subspaces to guarantee convergence to the essential spectrum.

(a) Notation

Below we will denote by $D(A)$ the domain of $A$ and by $\text{Spec}(A)$ its spectrum. We decompose $\text{Spec}(A)$ in the standard manner as the union of essential and discrete spectrum (see Kato 1980), and denote $\text{Spec}(A) = \text{Spec}_{\text{ess}}(A) \cup \text{Spec}_{\text{disc}}(A)$. Recall that the discrete spectrum is the set of isolated eigenvalues of finite multiplicity of $A$.

Let $(a, b) \subset \mathbb{R}$. Below $D(a, b)$ will be the open disc in the complex plane with centre $(a + b)/2$ and radius $(b - a)/2$, and $D[a, b] = D(a, b)$. We allow $a = -\infty$ or $b = +\infty$ in the obvious way to denote half-planes or the whole of $\mathbb{C}$. For $b = a$, $D[a, a] = \{a\}$ and $D(a, a) = \emptyset$. 

Proc. R. Soc. A
Second-order spectra

Let $\mathcal{K}$ be a Hilbert space. Let $\mathcal{Q} \subset \mathcal{K}$ be an arbitrary subset and let $\mathcal{B} \subset \mathcal{K}$ be a finite subset. We will often write

$$\text{dist}_\mathcal{K}(\mathcal{B}, \mathcal{Q}) := \max_{u \in \mathcal{B}} \inf_{v \in \mathcal{Q}} \|u - v\|_\mathcal{K} \quad \text{and} \quad \text{dist}_\mathcal{K}(u, \mathcal{Q}) = \text{dist}_\mathcal{K}(\{u\}, \mathcal{Q}).$$

Here, we include the possibility of $\mathcal{K} \equiv \mathbb{C}$ and write $\text{dist}(u, \mathcal{Q}) = \text{dist}_\mathbb{C}(u, \mathcal{Q})$.

Let $p \in \mathbb{N} \cup \{0\}$. We endow $D(A^p)$ with the graph inner product defined for all $u, v \in D(A^p)$ as $\langle u, v \rangle_p^p := \sum_{q=0}^{p} (A^q u, A^q v)$. By construction, $(D(A^p), \langle \cdot, \cdot \rangle_p)$ is a Hilbert space with the associated norm denoted by $\| \cdot \|_p$. Below we will consider sequences of finite-dimensional subspaces $(\mathcal{L}_n) = (\mathcal{L}_n)_{n \in \mathbb{N}}$ growing towards $D(A^p)$ with a given density property determined as follows:

$$A_p = \{(\mathcal{L}_n) \subset D(A^p) : \forall f \in D(A^p), \ \text{dist}_{D(A^p)}(f, \mathcal{L}_n) \to 0\}.$$

Note that if $\|A\| < \infty$, then $A_0 = A_p$ for any $p \in \mathbb{N}$. If $(\mathcal{L}_n) \subset A_p$ and $\mathcal{L}_n \subset \mathcal{L}_{n+1}$, then $\bigcup_{n=1}^\infty \mathcal{L}_n$ is dense norm $\| \cdot \|_p$.

For a family of closed subsets $\mathcal{Q}_n \subset \mathbb{C}$, the limit set of this family is defined as

$$\lim_{n \to \infty} \mathcal{Q}_n = \{z \in \mathbb{C} : \exists z_n \in \mathcal{Q}_n, \ z_n \to z\} = \{z \in \mathbb{C} : \text{dist}(z, \mathcal{Q}_n) \to 0\}.$$

The limit set of a family of closed subsets is always closed.

Let $\{b_j\}_{j=1}^n$ be a basis for a subspace $\mathcal{L} \subset \mathcal{H}$. Without further mention, we will identify $u \in \mathcal{L}$ with a corresponding $u \in \mathbb{C}^n$ via: $u = \sum_{j=1}^n \langle u, b_j^* \rangle b_j$ and $u = (\langle u, b_1^* \rangle, \ldots, \langle u, b_n^* \rangle)$, where $\{b_j^*\}$ is the basis conjugate to $\{b_j\}$. If the vectors $b_j$ are mutually orthogonal and $\|b_j\| = 1$, then $b_j^* = b_j$. Below the orthogonal projection onto $\mathcal{L}$ is denoted by $P : \mathcal{H} \rightarrow \mathcal{L}$. When referring to sequences of subspaces $(\mathcal{L}_n)$ we will use $P_n$ instead. Clearly $(\mathcal{L}_n) \in A_0$ iff $P_n \rightarrow I$ strongly.

(b) Toy models of spectral pollution in the Galerkin method

We now present a series of examples which will later illustrate our main results. We follow theorem 2.1 of Levitin & Shargorodsky (2004) and remark 2.5 of Lewin & Séré (2010). Here and below we denote $|v\rangle \langle w| = \langle u, w \rangle v$.

Example 1.1. Let $\mathcal{H} = \overline{\text{Span}\{e_{n}^\pm\}_{n=1}^\infty}$. Let $e_{n}^\pm$ be an orthonormal set of vectors and let $\mathcal{L}_n = \text{Span}\{e_{1}^\pm, \ldots, e_{n-1}^\pm, e_{n}^-\}$. Let $A = \sum_{n \in \mathbb{N}} n^2 |f_n^+\rangle \langle f_n^+| - \sum_{n \in \mathbb{N}} n |f_n^-\rangle \langle f_n^-|$ defined in its maximal domain, where $f_n^\pm = 1/\sqrt{2}(e_{n}^+ \pm e_{n}^-)$. Then, $\text{Spec}(P_n A \upharpoonright \mathcal{L}_n) = \{0, \pm 1, \ldots, \pm (n-1)\}$ and so $\lim_{n \to \infty} \text{Spec}(P_n A \upharpoonright \mathcal{L}_n) \setminus \text{Spec}(A) = \{0\} \neq \emptyset$. Note that the resolvent of $A$ is compact.

Example 1.2. If $A$ is strongly indefinite and it has a compact resolvent, then there exists a sequence $(\mathcal{L}_n) \in A_p$ for all $p \in \mathbb{N}$, such that $\lim_{n \to \infty} \text{Spec}(P_n A \upharpoonright \mathcal{L}_n) = \mathbb{R}$. Indeed, let $\text{Spec}(A) = \{\lambda_m^\pm\}_{m \in \mathbb{N}}$ where the eigenvalues are repeated according to multiplicity with $\lambda_{m+1}^- \leq \lambda_m^- < 0$ and $0 \leq \lambda_m^+ < \lambda_{m+1}^+$ for $1 \leq j \leq n$. Let $e_{m}^\pm$ be eigenvectors associated to $\lambda_m^\pm$ and assume that $\{e_{m}^\pm\}$ is an orthonormal set. Let $\{\gamma_j\}_{j \in \mathbb{N}}$ be an ordering of $\mathbb{Q}$. For each $n \in \mathbb{N}$, choose $n < k \in \mathbb{N}$ with $\lambda_k^- < \gamma_j < \lambda_k^+$ for $1 \leq j \leq n$. Let $\theta_j \in (-\pi, \pi]$ be such that $\gamma_j = \cos^2(\theta_j)\lambda_{k+j}^- + \sin^2(\theta_j)\lambda_{k+j}^+$ for $1 \leq j \leq n$ and define $f_{n} = \cos(\theta_j) e_{k+j}^- + \sin(\theta_j) e_{k+j}^+$. The desired conclusion holds by taking $\mathcal{L}_n = \text{Span}\{e_{1}^\pm, \ldots, e_{n}^\pm, f_{n1}, \ldots, f_{nn}\} \subset D(A^p)$. 

Proc. R. Soc. A
Example 1.3. Let $\mathcal{H}$ and $\mathcal{L}_n$ be as in example 1.1. Let $f_n^\pm = \sin(1/n) e_n^\pm \pm \cos(1/n) e_n^\pm$. For $r > 0$, define $A = \sum_{n \in \mathbb{N}} r^n |f_n^+\rangle\langle f_n^+| - \sum_{n \in \mathbb{N}} r^n |f_n^-\rangle\langle f_n^-|$, and $\sum_{n \in \mathbb{N}} r^n |f_n^+\rangle\langle f_n^+| - \sum_{n \in \mathbb{N}} r^n |f_n^-\rangle\langle f_n^-|$ in its maximal domain. Then $\text{Spec}_\text{ess}(A) = \{-1\}$ and $\text{Spec}_\text{dis}(A) = \{r^n\}_{n \in \mathbb{N}}$. It is not difficult to see that now $\text{Spec}(P_n A | \mathcal{L}_n) = \{-1, n^r \sin(1/n)^2 - \cos(1/n)^2, 1, \ldots, (n - 1)^r\}$, where $-1$ is an eigenvalue of multiplicity $n$. Then, if $r = 2$, the origin is an accumulation point of $\text{Spec}(P_n A | \mathcal{L}_n)$. Note that the resolvent of $A$ is not compact and $A$ is now semi-bounded below.

2. The second-order spectra of a self-adjoint operator

Definition 2.1 (see Levitin & Shargorodsky 2004). Given a non-trivial subspace $\mathcal{L} \subset D(A)$, the second-order spectrum of $A$ relative to $\mathcal{L}$ is the set

$$\text{Spec}_2(A, \mathcal{L}) := \{z \in \mathbb{C} : \exists u \in \mathcal{L}\setminus\{0\}, \langle (A - zI)u, (A - \bar{z}I)v \rangle = 0 \ \forall v \in \mathcal{L}\}.$$ (2.1)

If $\mathcal{L} \subset D(A^2)$, then

$$\text{Spec}_2(A, \mathcal{L}) = \{z \in \mathbb{C} : \exists u \in \mathcal{L}\setminus\{0\}, \langle P(A - z)^2 u, v \rangle = 0 \ \forall v \in \mathcal{L}\}.$$ (2.2)

Typically, $\text{Spec}_2(A, \mathcal{L})$ contains non-real points. From the definition, it follows that $z \in \text{Spec}_2(A, \mathcal{L})$ if and only if $\bar{z} \in \text{Spec}_2(A, \mathcal{L})$.

(a) **Algebraic and geometric multiplicity**

Let $\mathcal{L} = \text{Span}\{b_j\}_{j=1}^n \subset D(A)$ where the $b_j$ are linearly independent. Let $B, L, M \in \mathbb{C}^{n \times n}$ be matrices with entries given by

$$B_{jk} = \langle Ab_k, Ab_j \rangle, \quad L_{jk} = \langle Ab_k, b_j \rangle \quad \text{and} \quad M_{jk} = \langle b_k, b_j \rangle,$$ (2.1)

and $Q(z) = B - 2zL + z^2M$. It is readily seen that $z \in \text{Spec}_2(A, \mathcal{L})$ if and only if

$$\exists u \in \mathbb{C}^n \setminus\{0\} \text{ with } Q(z)u = 0.$$ (2.2)

By definition 2.1, $\text{Spec}_2(A, \mathcal{L})$ is independent of the basis chosen. Since $\det M \neq 0$, it follows that $\text{Spec}_2(A, \mathcal{L})$ consists of at most $2n$ points.

The quadratic eigenvalue problem (2.2) may be solved via suitable linearizations (see Gohberg et al. 2005). One such linearization is given by the matrices

$$T = \begin{pmatrix} 0 & I \\ -B & 2L \end{pmatrix} \quad \text{and} \quad S = \begin{pmatrix} I & 0 \\ 0 & M \end{pmatrix}.$$ (2.3)

Clearly, $z \in \text{Spec}_2(A, \mathcal{L})$ if and only if $(T - zS)u = 0$ for some $u \neq 0$. Equivalently, one can consider the matrix $T$ where

$$T := S^{-1}T = \begin{pmatrix} 0 & I \\ -M^{-1}B & 2M^{-1}L \end{pmatrix},$$ (2.3)

so that $\text{Spec}(T) = \text{Spec}_2(A, \mathcal{L})$. Let

$$E(z) := \begin{pmatrix}zM - 2L & M \\ I & 0 \end{pmatrix} \quad \text{and} \quad F(z) := \begin{pmatrix} I & 0 \\ zI & I \end{pmatrix}.$$ (2.4)

A straightforward calculation yields $Q(z) \oplus (-I) = E(z)(zI - T)F(z)$, so $Q(z) \oplus (-I)$ and $(zI - T)$ are equivalent in the sense of matrix polynomials.
Every matrix polynomial is known to be equivalent to a diagonal matrix polynomial \(\text{diag}[i_1(z), i_2(z), \ldots, i_p(z), 0, \ldots, 0]\), where the diagonal entries \(i_j(z)\) are polynomials of the form \(\Pi_k(z - z_{jk})^{\beta_k}\) with the property that \(i_j(z)\) is divisible by \(i_{j-1}(z)\). The factors \((z - z_{jk})^{\beta_k}\) are called elementary divisors and the \(z_{jk}\) are the eigenvalues. The degrees of the elementary divisors associated to a particular eigenvalue \(z_0\) are called the partial multiplicities at \(z_0\) (see Gohberg et al. 2005). Therefore, \(\text{Spec}_2(A)\) is defined to be the geometric and algebraic multiplicity of \(z\) as an eigenvalue of \(A\). For a linear matrix polynomial, such as \((T - zI)\), the degrees of the elementary divisors associated to a particular eigenvalue \(z\) follow from the fact that two matrix polynomials are equivalent if and only if they have the same collection of elementary divisors (see Gohberg et al. 2005, theorem A.6.2).

**Definition 2.2.** The geometric and algebraic multiplicities of \(z \in \text{Spec}_2(A, \mathcal{L})\) are defined to be the geometric and algebraic multiplicity of \(z\) as an eigenvalue of \(T\).

Note that definition 2.2 is independent of the basis used to assemble the matrices \(B, L, M\) and \(T\). Indeed, let \(B, L, M\) be assembled with respect to a basis \(\{b_j\}_{j=1}^n\) and \(\tilde{B}, \tilde{L}, \tilde{M}\) with respect to a basis \(\{c_j\}_{j=1}^n\), where \(\mathcal{L} = \text{Span}\{b_j\}_{j=1}^n = \text{Span}\{c_j\}_{j=1}^n\). Then,

\[
\tilde{T} := \begin{pmatrix}
0 & I \\
-M^{-1}\tilde{B} & 2M^{-1}\tilde{L}
\end{pmatrix} = \begin{pmatrix}
N^{-1} & 0 & I \\
0 & N^{-1} & -M^{-1}B & 2M^{-1}L
\end{pmatrix} \begin{pmatrix}
N & 0 \\
0 & N
\end{pmatrix},
\]

where \(N_{ij} = \langle c_j, b_i^\dagger \rangle\), and therefore \(\tilde{T}\) and \(T\) are equivalent.

(b) **The approximate spectral distance**

Suppose that the given basis \(\{b_j\}_{j=1}^n\) of \(\mathcal{L}\) is orthonormal so that \(M = I\) in (2.1) and \(Q(z) = B - 2zL + z^2I\). For \(z \in \mathbb{C}\), we define \(\sigma_{\mathcal{A},\mathcal{L}}(z) = \min_{x \in \mathbb{C}^n} \|Q(z)x\|/\|x\|\). The right-hand side of this expression is independent of the orthonormal basis chosen for \(\mathcal{L}\). If \(\mathcal{L} \subset D(A^2)\), then \(\sigma_{\mathcal{A},\mathcal{L}}(z) = \min_{x \in \mathcal{L}} \|P(A - z)^2v\|/\|v\|\), and for \(z \notin \text{Spec}_2(A, \mathcal{L})\) we have \(\sigma_{\mathcal{A},\mathcal{L}}(z) = \|Q(z)^{-1}\|^{-1}\). Furthermore, \(z \in \text{Spec}_2(A, \mathcal{L})\) if and only if \(\sigma_{\mathcal{A},\mathcal{L}}(z) = 0\). The map \(\sigma_{\mathcal{A},\mathcal{L}} : \mathbb{C} \rightarrow [0, \infty)\) is a Lipschitz subharmonic function (Davies 1998; Boulton 2007). Therefore, \(\text{Spec}_2(A, \mathcal{L})\) is completely characterized via the maximum principle.

(c) **The spectrum and the second-order spectra**

Let \(\lambda_{\min}(A) = \inf[\text{Spec}(A)]\) and \(\lambda_{\max}(A) = \sup[\text{Spec}(A)]\). For all \(\mathcal{L} \subset D(A)\),

\[
\text{Spec}_2(A, \mathcal{L}) \subset D[\lambda_{\min}(A), \lambda_{\max}(A)]; \quad (2.4)
\]

(see Shargorodsky 2000; Strauss in press). Thus, \(\lim_{n \to \infty} \text{Spec}_2(A, \mathcal{L}_n)\) is a subset of \(D[\lambda_{\min}(A), \lambda_{\max}(A)]\) for any sequence \((\mathcal{L}_n)\).

A growing interest in the second-order relative spectrum and corresponding limit sets has been stimulated by the following property: if \((a, b) \cap \text{Spec}(A) = \emptyset\), then

\[
\text{Spec}_2(A, \mathcal{L}) \cap D(a, b) = \emptyset \quad \forall \mathcal{L} \subset D(A); \quad (2.5)
\]
(see Levitin & Shargorodsky 2004; also lemma 2.3 below). Thus,

$$\text{Spec}(A) \cap [\text{Re } z - |\text{Im } z|, \text{Re } z + |\text{Im } z|] \neq \emptyset \quad \text{for } z \in \text{Spec}_2(A, \mathcal{L}).$$

(2.6)

Therefore, intervals of inclusion for points in the spectrum of $A$ can be obtained from $\text{Re } z$ with a two-sided explicit residual given by $|\text{Im } z|$.

In fact, the order of magnitude of the residue in the approximation of $\text{Spec}(A)$ by projecting $\text{Spec}_2(A, \mathcal{L})$ into $\mathbb{R}$ can be improved to $|\text{Im } z|$, if some information on the localization of $\text{Spec}(A)$ is at hand. Indeed, if $(a, b) \cap \text{Spec}(A) = \{\lambda\}$ and $z \in \text{Spec}_2(A, \mathcal{L})$ with $z \in \mathbb{D}(a, b)$, then

$$\left[\text{Re } z - \frac{|\text{Im } z|^2}{b - \text{Re } z}, \text{Re } z + \frac{|\text{Im } z|^2}{\text{Re } z - a}\right] \cap \text{Spec}(A) = \{\lambda\}$$

(2.7)

(see Boulton & Levitin 2007; Strauss in press).

The following lemma is an improvement upon theorem 5.2 of Shargorodsky (2000). It immediately implies equation (2.5).

**Lemma 2.3.** Let $a, b \in \mathbb{R}$ be such that $(a, b) \cap \text{Spec}(A) = \emptyset$. If $z \in \mathbb{D}(a, b)$, then $\sigma_{A, \mathcal{L}}(z) \geq \alpha(z)$ where

$$\alpha(z) = \alpha_{(a, b)}(z) = \frac{(b - a)^2 - |z - a|^2 - |z - b|^2}{2|z - b||z - a|}\text{dist}(z, \{a, b\})^2 > 0.$$  

(2.8)

**Proof.** Without loss of generality assume that $\text{Im } z \geq 0$, and consider the triangle $\Delta azb$. Let $\theta$ be the angle of $\Delta azb$ at $z$. Since $z \in \mathbb{D}(a, b)$, then $\pi/2 < \theta \leq \pi$. Let $\theta_1$ and $\theta_2$ be the angles at $b$ and $a$. Then $\theta + \theta_1 + \theta_2 = \pi$. The region $F_z := \{\mu - z : \mu \in \text{Spec}(A)\}$ is contained in two sectors, one between the real line and the ray $re^{i\theta_1}$ ($r \geq 0$), the other between the real line and $re^{i(\pi + \theta_2)}$, where $0 \leq \theta_1 + \theta_2 < \pi/2$. We have $\cos(\theta_1 + \theta_2) > 0$, and by the cosine rule, $\cos(\theta_1 + \theta_2) = ((b - a)^2 - |z - b|^2 - |z - a|^2)/(2|z - b||z - a|)$. Now $F_z^2 = \{w^2 : w \in F_z\}$ is contained in a sector between the rays $re^{-i\theta_1}$ and $re^{i\theta_2}$. Therefore

$$\min\{\text{Re } y : y \in e^{i(\theta_1 - \theta_2)}F_z^2\} = \min\{|a - z|^2, |b - z|^2\} \cos(\theta_1 + \theta_2) = \alpha(z).$$

By virtue of the spectral theorem,

$$\text{Re } e^{i(\theta_1 - \theta_2)}(Q(z)\mu, \nu)_{\mathcal{C}^n} = \text{Re } e^{i(\theta_1 - \theta_2)}(A - z)v, (A - \bar{z})v)$$

$$= \text{Re } e^{i(\theta_1 - \theta_2)}\int_{\mathbb{R}} (\mu - z)^2 d\langle E_\mu v, v \rangle \geq \alpha(z)||v||^2.$$  

The result follows from the Cauchy–Schwarz inequality. 

If $(a, b) \cap \text{Spec}(A) = \emptyset$, it follows from equation (2.5) that $\mathbb{D}(a, b)$ does not intersect $\lim_{n \to \infty} \text{Spec}_2(A, \mathcal{L}_n)$. Therefore, $\mathbb{R} \cap \lim_{n \to \infty} \text{Spec}_2(A, \mathcal{L}_n) \subseteq \text{Spec}(A)$.

We now consider two examples where $\lim_{n \to \infty} \text{Spec}_2(A, \mathcal{L}_n) = \text{Spec}(A)$. The first one has a compact resolvent while the second one has $-1$ in the essential spectrum.

**Example 2.4.** Suppose that $\mathcal{H}$, $\mathcal{L}_n$ and $A$ are as in example 1.1. It is readily seen that $\text{Spec}_2(A, \mathcal{L}_n) = \{\pm in, \pm 1, \ldots, \pm (n - 1)\}$.

**Example 2.5.** Suppose that $\mathcal{H}$, $\mathcal{L}_n$ and $A$ be as in example 1.3. It is readily seen that $\text{Spec}_2(A, \mathcal{L}_n) = \{\alpha_n \pm i\gamma_n, -1, 1, \ldots, (n - 1)'\}$, where $\alpha_n = n^r \sin(1/n)^2 - \cos(1/n)^2$ and $\gamma_n = (n^r + 1)\sin(1/n)\cos(1/n)$. The geometric multiplicity of the
second-order spectral point $-1$ is $n - 1$ and its algebraic multiplicity is $2(n - 1)$. As $n \to \infty$, the non-real point $\alpha_n \pm i \gamma_n \to \beta$, where $\beta = -1$ for $0 < r < 1$, $\beta = -1 \pm i$ for $r = 1$ and $\beta = \infty$ for $r > 1$. Hence, $\lim_{n \to \infty} \text{Spec}_2(A, L_n) = \text{Spec}(A)$ for $r \neq 1$.

(d) Mapping of second-order spectra

Lemma 2.6. For $a, b, c, d \in \mathbb{R}$, let $f(w) = aw + b$, $g(w) = cw + d$ and $F(w) = f(w)g(w)^{-1}$. If $ad \neq cb$ and $g(A)^{-1} \in \mathcal{B}(\mathcal{H})$, then

$$z \in \text{Spec}_2(A, L) \iff F(z) \in \text{Spec}_2(F(A), g(A)L).$$

Moreover, the algebraic and geometric multiplicities of $z$ and $F(z)$ coincide.

Proof. Since $g(A)$ preserves linear independence, a set of vectors $\{b_j\}_{j=1}^n \subset L$ is a basis for $L$ if and only if the corresponding set $\{g(A)b_j\}_{j=1}^n$ is a basis for $g(A)L$. Let $\tilde{b}_j = g(A)b_j$. It is readily seen that $(\tilde{b}_j)^* = g(A)^{-1}b_j^*$. Let $u, v \in L$, $\tilde{u} = g(A)u$ and $\tilde{v} = g(A)v$. Then, $\tilde{u} = u \in \mathbb{C}^n$ for any $u \in L$, obtaining the left side from $\{\tilde{b}_j\}$ and the right side from $\{b_j\}$. Let $B, L, M$ be as in (2.1) for $A$ and $\{b_j\}$. Let $\tilde{B}, \tilde{L}, \tilde{M}$ be as in (2.1) for $F(A)$ and $\{b_j\}$. Let $z \in \text{Spec}_2(A, L)$. Then

$$\langle (B - 2zL + z^2M)u, v \rangle = \langle (A - z)u, (A - \tilde{z})v \rangle$$

$$= (ad - cb)^{-2}g(z)^2((F(A) - F(z))\tilde{u}, (F(A) - F(\tilde{z}))\tilde{v})$$

$$= (ad - cb)^{-2}g(z)^2(\tilde{B} - 2F(z)\tilde{L} + F(z)^2\tilde{M})u, v \rangle.$$ 

Here $u, v \in \mathbb{C}^n$ are arbitrary, so we deduce that

$$B - 2zL + z^2M = (ad - cb)^{-2}g(z)^2(\tilde{B} - 2F(z)\tilde{L} + F(z)^2\tilde{M}).$$

Thus, $F(z) \in \text{Spec}_2(F(A), g(A)L)$, and moreover, since the function $(ad - cb)^{-2}g(w)^2$ is analytical and non-zero at $z$, the algebraic and geometric multiplicities of $z$ and $F(z)$ are the same (see Gohberg et al. 2005, theorem A.6.6).

3. Multiple points of the second-order spectra

(a) Neighbourhoods of the discrete spectrum

Throughout this section we assume that

$$(a, b) \cap \text{Spec}(A) = \{\lambda_1 < \cdots < \lambda_s\} \subseteq \text{Spec}_{\text{dis}}(A).$$

Let $1_j = 1_{(\lambda_j - \epsilon, \lambda_j + \epsilon)}(A)$, where $\epsilon > 0$ is small enough so that $m_j := \text{Rank}(1_j)$ is equal to the multiplicity of $\lambda_j$. Let $m = \sum_{j=1}^s m_j$ be the total multiplicity of the group of eigenvalues $\lambda_1, \ldots, \lambda_s$. We will denote by $\mathcal{E}$ the eigenspace associated to the group of eigenvalues in $(a, b)$, that is, $\mathcal{E} = \text{Eig} \text{sp}(\text{Rank} 1_{(a,b)}(A))$. We fix orthonormal bases $\mathcal{B}_j = \{u_{kj}^j\}_{k=1}^{m_j}$ of $\text{Rank} 1_j$ and let $\mathcal{B} = \bigcup_{j=1}^s \mathcal{B}_j$ be a corresponding basis of $\mathcal{E}$.
Lemma 3.1. Let \( z \in \mathbb{D}(a, b) \). Let
\[
\gamma = 2\sqrt{5}(b - a)^2((1 + \max(|a|, |b|))^2 + (b - a)^2(2\sqrt{5}s + 8)).
\]

If the subspace \( \mathcal{L} \subset D(A^2) \) is such that
\[
\text{dist}_{D(A^2)}(\mathcal{B}, \mathcal{L}) < \delta < \alpha_{(a, b)}(z)[\text{dist}(z, \text{Spec } A)^2\gamma^{-1},
\]
then \( \sigma_{A, \mathcal{L}}(z) \geq \frac{\delta}{b - a} - 2\alpha_{(a, b)}(z)[\text{dist}(z, \text{Spec } A)^2 - \delta\gamma] > 0 \).

Proof. Let \( c \in \mathbb{R}\setminus(a, b) \). Define \( \tilde{A} = A + \sum_{j=1}^s (c - \lambda_j)1_j \) where \( D(\tilde{A}) = D(A) \). Then \( \tilde{A} \) is self-adjoint and \( \text{Spec}(\tilde{A}) = \{c\} \cup \text{Spec}(A) \setminus \{\lambda_1, \ldots, \lambda_s\} \). Let
\[
K(z) = -\sum_{j=1}^s (c - \lambda_j)(\lambda_j + c - 2z)1_j,
\]
then \( K(z) \) is a finite rank operator with range \( \mathcal{E} \) and \( (\tilde{A} - z)^2 = (A - zI)^2 - K(z) \). Evidently, \( (\tilde{A} - z)^2 \) is invertible and for all \( v \in \mathcal{H} \) we have
\[
\|v + (\tilde{A} - z)^{-2}K(z)v\|^2 = \|\mathbb{1}_{\mathbb{R}\setminus(a, b)}v\|^2 + \|\mathbb{1}_{(a, b)}v + (\tilde{A} - z)^{-2}(A - z)^2\mathbb{1}_{(a, b)}v\|^2
\]
\[
= \|\mathbb{1}_{\mathbb{R}\setminus(a, b)}v\|^2 + \|\mathbb{1}_{(a, b)}v\|^2 \geq \beta_1(z)^2\|v\|^2,
\]
where \( \beta_1(z) = \min\{1, \text{dist}(z, \text{Spec } A)\}/|c - z| \).

Let \( Q(z) = P(\tilde{A} - z)^2 \mathbb{1}_{\mathcal{L}} \). For each eigenvector \( u_j^k \) we assign a \( v_j^k \in \mathcal{L} \) such that \( \|Ap(u_j^k - v_j^k)\| \leq \delta \) for \( p = 0, 1, 2 \). In general \( Pu_j^k \neq v_j^k \); however, \( \|Pv_j^k - v_j^k\| \leq \delta \). The hypothesis on \( \delta \) ensures that
\[
\|Q(z)v_j^k - (\tilde{A} - z)^2u_j^k\| \leq \|P(\tilde{A} - z)^2(u_j^k - v_j^k)\| + |c - z|^2\|Pu_j^k - u_j^k\|
\]
\[
\leq \|(\tilde{A} - z)^2(u_j^k - v_j^k)\| + |c - z|^2\|Pu_j^k - u_j^k\|
\]
\[
\leq \left\| \left( (A - zI)^2 + \sum_{i=1}^s (c - \lambda_i)(\lambda_i + c - 2z)1_i \right) (u_j^k - v_j^k) \right\|
\]
\[
+ \delta|c - z|^2 \leq \delta\beta_2(z),
\]
where \( \beta_2(z) = (1 + |z|)^2 + s\mu(z) + |c - z|^2 \) and \( \mu(z) = \max_{1 \leq j \leq s} |(c - \lambda_j)(\lambda_j + c - 2z)| \).
Note that $\tilde{A}$ satisfies the hypothesis of lemma 2.3. Let $\alpha(z)$ be given by equation (2.8). Then
\[
\|((\tilde{A} - z)^{-2}u_j^k - Q(z)^{-1}Pu_j^k)\|
\leq\|((\tilde{A} - z)^{-2}u_j^k - Q(z)^{-1}v_j^k)\| + \|Q(z)^{-1}Pu_j^k - Q(z)^{-1}v_j^k)\|
\leq\|(c - z)^{-2}v_j^k - Q(z)^{-1}v_j^k\| + (\alpha(z)^{-1} + |c - z|^{-2})\delta
\leq\alpha(z)^{-1}\|(c - z)^{-2}Q(z)v_j^k - u_j^k\| + (2\alpha(z)^{-1} + |c - z|^{-2})\delta
\leq\alpha(z)^{-1}|c - z|^{-2}\|Q(z)v_j^k - (\tilde{A} - z)^{-2}u_j^k\| + (2\alpha(z)^{-1} + |c - z|^{-2})\delta \leq \beta_3(z)\delta
\]
where $\beta_3(z) = \alpha(z)^{-1}|c - z|^{-2}[\beta_2(z) + 2|c - z|^2 + \alpha(z)]$. Thus
\[
\|[(\tilde{A} - z)^{-2} - Q(z)^{-1}P]K(z)v\|
= \left\|[((\tilde{A} - z)^{-2} - Q(z)^{-1}P)\sum_{j=1}^{s} (c - \lambda_j)(2z - c - \lambda_j)1_jv\right]\|
\leq \sum_{j=1}^{s} |c - \lambda_j||2z - c - \lambda_j|\|[(\tilde{A} - z)^{-2} - Q(z)^{-1}P]1_jv\|
\leq \mu(z)||v|| \sum_{j=1}^{s} \sum_{k=1}^{m_j} \|[(\tilde{A} - z)^{-2} - Q(z)^{-1}P]u_k\| \leq \delta m\mu(z)\beta_3(z)||v||.
\]
Hence, for any $v \in \mathcal{L}$,
\[
\|P(A - zI)^2v\|
= \|Q(z)v + PK(z)v\| \geq \|Q(z)^{-1}\|^{-1}||v + Q(z)^{-1}PK(z)v||
\geq \alpha(z)||v + (\tilde{A} - z)^{-2}K(z)v|| - \|((\tilde{A} - z)^{-1}K(z)v - Q(z)^{-1}PK(z)v)||
\geq \alpha(z)[\beta_1(z)^2 - \delta m\mu(z)\beta_3(z)]||v||.
\]
Let
\[
c = \begin{cases} 
  b + (b - a) = 2b - a & \text{if } \text{Re } z \geq \frac{(a + b)}{2} \\
  a - (b - a) = 2a - b & \text{if } \text{Re } z \leq \frac{(a + b)}{2},
\end{cases}
\]
then $b - a \leq |c - z| \leq (\sqrt{10}/2)(b - a)$. Thus
\[
\beta_1(z) = \frac{\text{dist}(z, \text{Spec}(A))}{|c - z|}, \quad \mu(z) \leq 2\sqrt{5}(b - a)^2,
\]
\[
\beta_2(z) \leq (1 + \max(|a|, |b|))^2 + (b - a)^2 \left(2s\sqrt{5} + \frac{10}{4}\right) \quad \text{and } \alpha(z) \leq (b - a)^2/2.
\]
Therefore
\[
\sigma_{A,L}(z) \geq \alpha(z)[\beta_1(z)^2 - \delta m \mu(z) \beta_3(z)]
\]
\[
\geq \frac{\alpha(z)[\text{dist}(z, \text{Spec}(A))]^2}{|c - z|^2} - \frac{\delta m 2\sqrt{5}(b - a)^2[\beta_2(z) + 2|c - z|^2 + \alpha(z)]}{|c - z|^2}
\]
\[
\geq \frac{\alpha(z)[\text{dist}(z, \text{Spec}(A))]^2 - \delta \gamma}{|c - z|^2} \geq \frac{2}{5}(b - a)^2(\alpha(z)[\text{dist}(z, \text{Spec}(A))]^2 - \delta \gamma).
\]

\[\square\]

**Lemma 3.2.** Let \( \lambda \in \text{Spec}_{\text{dis}}(A) \) be an eigenvalue of multiplicity \( m \). Denote by \( d_\lambda = \text{dist}(\lambda, \text{Spec}(A) \setminus \{\lambda\}) \). If \( \text{Rank } \mathbb{1}_{(a,b)}(A) = E \subset L \), then
\[
\text{Spec}_2(A, L) \cap \mathbb{D}(\lambda - d_\lambda, \lambda + d_\lambda) = \{\lambda\},
\]
and \( \lambda \) as member of \( \text{Spec}_2(A, L) \) has geometric multiplicity \( m \) and algebraic multiplicity \( 2m \).

**Proof.** Assume that \( z \in \text{Spec}_2(A, L) \cap \mathbb{D}(\lambda - d_\lambda, \lambda + d_\lambda) \) and \( z \neq \lambda \). Then, there exists a normalized \( u \in L \setminus \{0\} \), such that \( \langle (A - zI)u, (A - \bar{z}I)u \rangle = 0 \) for all \( w \in L \). In particular \( \langle u, u^k \rangle = 0 \) for each \( u^k \in E \). If we take \( w = u \) above, then
\[
\|(A - \text{Re } z)u\|^2 - |\text{Im } z|^2\|u\|^2 - 2i|\text{Im } z|\langle (A - \text{Re } z)u, u \rangle = 0.
\]
Thus, \( |\text{Im } z| = \|(A - \text{Re } z)u\| \neq 0 \) and \( \langle (A - \text{Re } z)u, u \rangle = 0 \), so that
\[
\|(A - \lambda)u\|^2 = |\lambda - \text{Re } z|^2\|u\|^2 + \|(A - \text{Re } z)u\|^2 + 2(\lambda - \text{Re } z)\langle (A - \text{Re } z)u, u \rangle
\]
\[
= |\lambda - \text{Re } z|^2 + |\text{Im } z|^2 = |\lambda - z|^2.
\]
Since \( u \perp E \),
\[
d_\lambda = \text{dist}(\lambda, \text{Spec}(A) \setminus \{\lambda\}) \leq \frac{\|(A - \lambda)(I - \mathbb{1}(\lambda - d_\lambda, \lambda + d_\lambda))u\|}{\|(I - \mathbb{1}(\lambda - d_\lambda, \lambda + d_\lambda))u\|} = \|(A - \lambda)u\| = |\lambda - z|,
\]
however, the right-hand side is strictly less than \( d_\lambda \), and equation (3.3) follows from the contradiction.

Let \( \{b_j\}_{j=1}^n \) be a basis for \( L \) and let \( B, L \) and \( M \) be the matrices (2.1) associated to this basis. Let \( v \) be an arbitrary member of \( L \). Then
\[
0 = \langle (A - \lambda I)u^k, (A - \lambda I)v \rangle_{\mathcal{H}} = \langle (B - 2\lambda L + \lambda^2 M)u_k, v \rangle_{\mathcal{C}^n}
\]
so that \( (B - 2\lambda L + \lambda^2 M)u_k = 0 \). Hence,
\[
[T - \lambda S]\begin{pmatrix} u_k^k \\ \lambda u_k \end{pmatrix} = \begin{pmatrix} 0 & I \\ -B & 2L \end{pmatrix} - \lambda \begin{pmatrix} I & 0 \\ 0 & M \end{pmatrix} \begin{pmatrix} u_k^k \\ \lambda u_k \end{pmatrix} = 0.
\]

Since \( \det(S) \neq 0 \), \( \{\begin{pmatrix} u_k^1 \\ \lambda_{u_k}^1 \end{pmatrix}, \ldots, \begin{pmatrix} u_k^m \\ \lambda_{u_k}^m \end{pmatrix}\} \) are eigenvectors of \( T \) corresponding to the eigenvalue \( \lambda \). For \( k \neq j \) we have \( \langle Mu_k, u^j \rangle_{\mathcal{C}^n} = \langle u_k, u^j \rangle_{\mathcal{H}} = 0 \), so the vectors \( \{u_k^j, \ldots, u_k^m\} \) are linearly independent. There is a one-to-one correspondence between the eigenvectors of \( T \) associated to \( \lambda \) and the eigenvectors of \( A \) associated to \( \lambda \). Thus, the geometric multiplicity of \( \lambda \in \text{Spec}(A, L) \) is equal to \( m \).
We now compute the algebraic multiplicity. Let \( v, w \in \mathbb{C}^n \) and \( u \in \text{Span} \{ u_k \}_{k=1}^m \).

\[
\begin{bmatrix}
v \\
\lambda w
\end{bmatrix} = (\mathcal{T} - \lambda I) \begin{bmatrix} v \\
w \end{bmatrix} = \left[ \left( \begin{array}{cc} 0 & B \\ -M^{-1} & 2M^{-1}L \end{array} \right) - \lambda \left( \begin{array}{cc} I & 0 \\ 0 & I \end{array} \right) \right] \begin{bmatrix} v \\
w \end{bmatrix},
\]

then \( w = \lambda v + u \) and \( -M^{-1}Bv + 2M^{-1}Lw - \lambda w = \lambda u \). Therefore,

\[
-Bv + 2\lambda L - \lambda^2 Mv = 2\lambda Mu - 2Lu = 0
\]

so

\[
\begin{bmatrix} v \\
w \end{bmatrix} = \begin{bmatrix} v \\
\lambda v \end{bmatrix} + \begin{bmatrix} 0 \\
u \end{bmatrix}
\]

where \( v \in \text{Span} \{ u_k \}_{k=1}^m \). If

\[
\begin{bmatrix} 0 \\
u \end{bmatrix} = (\mathcal{T} - \lambda I) \begin{bmatrix} v \\
w \end{bmatrix},
\]

then \( w = \lambda v \) and \( -M^{-1}Bv + 2M^{-1}Lw - \lambda w = u \). Therefore, \( -Bv + 2\lambda L - \lambda^2 Mv = Mu \). Thus,

\[
\langle Mu, u \rangle = -\langle Bv - 2\lambda Lu + \lambda^2 Mv, u \rangle = -\langle v, Bu - 2\lambda Lu + \lambda^2 Mu \rangle = 0
\]

so that \( u = 0 \). This ensures that the spectral subspace associated with \( \lambda \) as eigenvalue of \( \mathcal{T} \) is given by

\[
\text{Span} \left\{ \begin{bmatrix} u_1 \\
\lambda u_1 \end{bmatrix}, \begin{bmatrix} 0 \\
u_1 \end{bmatrix}, \ldots, \begin{bmatrix} u_m \\
\lambda u_m \end{bmatrix}, \begin{bmatrix} 0 \\
u_m \end{bmatrix} \right\}.
\]

\[\blacksquare\]

**Lemma 3.3.** Let \( \mathcal{K} \) be a Hilbert space and \( \mathcal{A} = \{ u_1, \ldots, u_m \} \subset \mathcal{K} \) be a set of orthogonal vectors with \( \| u_j \|_{\mathcal{K}} \geq 1 \) for \( 1 \leq j \leq m \). Let \( \mathcal{L} \) be an \( n \)-dimensional subspace of \( \mathcal{K} \) where \( n \geq m \). Let \( w_j = \tilde{P}u_j \), where \( \tilde{P} : \mathcal{K} \to \mathcal{L} \) is the orthogonal projection associated to \( \mathcal{L} \). Let \( w_j(t) = tw_j + (1-t)u_j \) for \( t \in [0,1] \). If \( \| w_j - u_j \| < 1/\sqrt{m} \) for each \( 1 \leq j \leq m \), then there exist vectors \( \{ w_{m+1}, \ldots, w_n \} \), such that \( \{ w_1(t), \ldots, w_m(t), w_{m+1}, \ldots, w_n \} \) is linearly independent for all \( t \in [0,1] \).

**Proof.** If \( a_1 w_1 + \cdots + a_m w_m = 0 \), where not all of the \( a_j = 0 \), we would have

\[
\sum_{j=1}^m |a_j|^2 = \| a_1(u_1 - w_1) + \cdots + a_m(u_m - w_m) \|^2
\]

\[
\leq \left( \sum_{j=1}^m |a_j| \| u_j - w_j \| \right)^2 < \frac{(\sum_{j=1}^m |a_j|^2)^2}{m},
\]

which contradicts Hölder’s inequality. Hence, necessarily \( \{ w_1, \ldots, w_m \} \) is linearly independent. Let \( \{ w_{m+1}, \ldots, w_n \} \) be any completion of \( \{ w_1, \ldots, w_m \} \) to a basis of \( \mathcal{L} \).
Let $t \in [0, 1]$. Suppose now that $a_1 w_1(t) + \cdots + a_m w_m(t) + a_{m+1} w_{m+1} + \cdots + a_n w_n$ vanishes. Then

$$0 = \sum_{j=1}^m a_j w_j(t) + \sum_{j=m+1}^n a_j w_j = \sum_{j=1}^n a_j w_j + (I - \tilde{P}) \sum_{j=1}^m (1-t)a_j u_j.$$ 

The two terms on the right-hand side are orthogonal and therefore each must vanish. As the set $\{w_1, \ldots, w_n\}$ is linearly independent, all $a_j = 0$. 

(b) The main result

**Theorem 3.4.** Let $a, b \not\in \text{Spec}(A)$. Assume that condition (3.1) holds for a group of eigenvalues $\{\lambda_1 < \ldots < \lambda_s\}$ with corresponding multiplicities $m_j$ and their total multiplicity is $m = \sum_{j=1}^s m_j$. Let $d = \text{dist}(\{a, b\}, \text{Spec}(A) \setminus \{\lambda_j\}_{j=1}^s)$. Let $\kappa = d^2/\gamma$ where $\gamma$ is defined by equation (3.2). Let

$$0 < \varepsilon < \min \left\{ \frac{1}{m^{1/4}K^{1/2}}, \min_{0 \leq j \leq s} \frac{|\lambda_j - \lambda_{j+1}|}{2} \right\},$$

where $\lambda_0 = a$ and $\lambda_{s+1} = b$.

If $\mathcal{L} \subset D(A^2)$ is such that $\text{dist}_{D(A^2)}(\mathcal{B}, \mathcal{L}) < \kappa \varepsilon^2$ for an orthonormal set of eigenfunctions $\mathcal{B}$ associated to $\{\lambda_j\}_{j=1}^s$, then

$$\text{Spec}_2(A, \mathcal{L}) \cap D(a, b) \subset \bigcup_{j=1}^s D(\lambda_j - \varepsilon, \lambda_j + \varepsilon),$$

and the total algebraic multiplicity of $\text{Spec}_2(A, \mathcal{L}) \cap D(\lambda_j - \varepsilon, \lambda_j + \varepsilon)$ is $2m_j$.

**Proof.** Let $\tilde{a} = a - d$ and $\tilde{b} = b + d$, then

$$\inf_{z \in D(a, b)} \alpha_{(\tilde{a}, \tilde{b})}(z) = \min_{\theta \in (-\pi, \pi)} \left\{ \alpha_{(\tilde{a}, \tilde{b})}(z) : z = \frac{b-a}{2} e^{i\theta} + \frac{a+b}{2} \right\}.$$ 

In order to find the right-hand side, we may assume without loss of generality that $a = -r$ and $b = r$. If $z = x + iy$ where $x^2 + y^2 = r$, then

$$\alpha_{(\tilde{a}, \tilde{b})}(z) = \frac{(2r + 2d)^2 - |x + iy + r + d|^2 - |x + iy - r - d|^2}{2|x + iy + r + d||x + iy - r - d|} \times \text{dist}(x + iy, \{-r - d, r + d\})^2.$$ 

We also assume that $x \geq 0$ as the case where $x < 0$ can be treated analogously. A straightforward calculation yields

$$\min_{0 \leq x \leq r, y^2 + x^2 = r} \alpha_{(\tilde{a}, \tilde{b})}(x + iy) = d(d + 2r) \left[ \min_{0 \leq x \leq r, y^2 + x^2 = r} \frac{|x + iy - r - d|}{|x + iy + r + d|} \right] = d^2,$$

hence lemma 3.1 applied to the interval $(\tilde{a}, \tilde{b})$ ensures (3.4).

We now prove the second part of the theorem. Let $A = \mathcal{B}$, $\mathcal{K} = D(A^2)$ and define subspaces $\mathcal{L}_t = \text{Span}\{w_1^t(t), \ldots, w_m^t(t), w_{m+1}, \ldots, w_n\}$ for $t \in [0, 1]$. 

Proc. R. Soc. A
(see lemma 3.3). Note that $\mathcal{B}$ is also an orthogonal set in $D(A^2)$ and $\|u^k\|_{D(A^2)} \geq 1$. Since $\kappa e^2 < m^{-1/2}$, we get $\dim \mathcal{L}_t = n$ for all $t \in [0, 1]$. Now,
\[
\|u^k_j - w^k_j(t)\|_{D(A^2)} = t\|u^k_j - w^k_j\|_{D(A^2)} = t\|(1 - \tilde{P})u^k_j\|_{D(A^2)} \leq \text{dist}_{D(A^2)}(u^k_j, \mathcal{L}) \quad \text{for all} \quad t \in [0, 1],
\]
thus $\text{dist}_{D(A^2)}(\mathcal{B}, \mathcal{L}_t) < \kappa e^2$, and by virtue of lemma 3.1 we obtain
\[
\text{Spec}_2(A, \mathcal{L}_t) \cap \mathbb{D}(a, b) \subset \bigcup_{j=1}^{s} \mathbb{D}(\lambda_j - \varepsilon, \lambda_j + \varepsilon) \quad \text{for all} \quad t \in [0, 1].
\]
Denote by $\mathcal{T}(t)$ the linearization matrix defined as in equation (2.3) for $\mathcal{L}_t$. According to lemma 3.2, the algebraic multiplicities of $\lambda_j \in \text{Spec} \mathcal{T}(1)$ is $2m_j$. If we let $\pi_j(t)$ be the spectral projection associated to $\text{Spec} \mathcal{T}(t) \cap \mathbb{D}(\lambda_j - \varepsilon, \lambda_j + \varepsilon)$, then
\[
\pi_j(t_1) - \pi_j(t_2) = -\frac{1}{2i\pi} \oint_{|\lambda_j - z| = \varepsilon} (\mathcal{T}(t_1) - \zeta)^{-1} - (\mathcal{T}(t_2) - \zeta)^{-1} \, d\zeta.
\]
Thus, $\pi_j(t) \to \pi_j(1)$ as $t \to 1$. By virtue of lemma 4.10 of §1.4.6 in Kato (1980), it is guaranteed that $\text{Rank} \pi_j(0) = \text{Rank} \pi_j(1)$. □

An immediate consequence of theorem 3.4 is the fact that, if a sequence of test subspaces $\mathcal{L}_n$ approximates a normalized eigenfunction $u$ associated to a simple eigenvalue $\lambda$ at a given rate in the graph norm of $A^2$, $\|u - u_n\|_{D(A^2)} = \delta(n) \to 0$ for $u_n \in \mathcal{L}_n$, then there exist a conjugate pair $z_n, \bar{z}_n \in \text{Spec}_2(A, \mathcal{L}_n)$ such that $|\lambda - z_n| = O(\delta^{1/2})$. Simple examples show that this estimate is sub-optimal.

**Example 3.5.** Let $\mathcal{H} = \text{Span}\{e_n\}_{n=0}^{\infty}$. Let $\mathcal{L}_n = \text{Span}\{e_1, \ldots, e_{n-1}, \alpha_n e_0 + \beta_n e_n\}$, where $\alpha_n^2 + \beta_n^2 = 1$ and $\beta_n \to 0$. Then $\text{Spec}_2(A, \mathcal{L}_n) = \{1, \ldots, n - 1, \gamma_n, \overline{\gamma}_n\}$ where $\gamma_n \sim in\beta_n$ for $A = \sum n(|e_n|^2)\langle e_n^* |$. On the other hand, $\|A^p((\alpha_n - 1)e_0 + \beta_n e_n)\| \sim n^p \beta_n$ for $p = 0, 1, 2$. Thus, $|\gamma_n| \sim n^2 \beta_n$ while $\text{dist}_{D(A^2)}(\mathcal{B}, \mathcal{L}_n) \sim n^{3/2} \beta_n$. Moreover, we observe that in general $\text{dist}_{D(A^2)}(\mathcal{B}, \mathcal{L}_n)$ might diverge as $n \to \infty$ and still we might be able to recover $|\gamma_n| \to 0$; take for instance $\beta_n = 1/n^{3/2}$.

An application of lemma 2.6 yields convergence to eigenvalues under the weaker assumption $\mathcal{L}_n \in A_1$.

**Corollary 3.6.** Let $-\infty \leq a < b \leq \infty$ be such that $(a, b) \cap \text{Spec}_{\text{ess}}(A) = \emptyset$. If $(\mathcal{L}_n) \in A_1$, then $\lim_{n \to \infty} \text{Spec}_2(A, \mathcal{L}_n) \cap \mathbb{D}(a, b) = \text{Spec}_{\text{dis}}(A) \cap (a, b)$.

**Proof.** Pick $\varepsilon > 0$ such that $a + \varepsilon < \min(\lambda \in \text{Spec}(A) \cap (a, b))$. Let $\mu \in (a, a + \varepsilon)$. Let $g(w) = w - \mu$ and $h(z) = g(z)^{-1}$. The operator $h(A)$ is bounded and $g(w)$ satisfies the hypothesis of lemma 2.6. Let $\mathcal{M}_n = g(A)\mathcal{L}_n$. Then $(\mathcal{M}_n) \in A_0$. Indeed, any $u \in \mathcal{H}$ can be expressed as $u = g(A)v$ for $v \in D(A)$, and since $(\mathcal{L}_n) \in A_1$, we can find $v_n \in \mathcal{L}_n$ such that $g(A)v_n \to g(A)v$. By virtue of theorem 3.4,
\[
\lim_{n \to \infty} \text{Spec}_2(h(A), \mathcal{M}_n) \cap \mathbb{D}(h(b), h(a + \varepsilon)) = \text{Spec}_{\text{dis}}(h(A)) \cap (h(b), h(a + \varepsilon)).
\]
The fact that $g(w)$ is a conformal mapping, lemma 2.6 and the spectral mapping theorem ensure the desired conclusion. □
Example 3.7. Let $A$ be the operator of examples 1.3 and 2.5. By virtue of (2.5) and corollary 3.6, $\lim_{n \to \infty} \text{Spec}_2(A, \mathcal{L}_n) \subset \text{Spec}(A) \cup (-1 + i\mathbb{R})$ for any $(\mathcal{L}_n) \in \Lambda_1$.

Corollary 3.8. Let $A$ be an operator with a compact resolvent. For all $(\mathcal{L}_n) \in \Lambda_0$, $\lim_{n \to \infty} \text{Spec}_2(A, \mathcal{L}_n) = \text{Spec}(A)$.

When $A$ is not semi-bounded, this statement is in stark contrast to Example 1.2.

4. Numerical applications

The identity (2.6) leads to guaranteed intervals of enclosures for $\lambda \in \text{Spec}_{\text{disc}}(A)$. Theorem 3.4 yields a priori upper bounds for the length of these enclosure in terms of bounds on the distance from the test space, $\mathcal{L}$, to an orthonormal basis of eigenfunctions $\mathcal{B}$. In practice, these upper bounds are computed from interpolation estimates for bases of $\mathcal{L}$. We now examine various aspects of the applicability of these results.

(a) On multiplicity and approximation

Suppose that $\lambda$ is of multiplicity $m$, so that (3.4) ensures an upper bound on the estimation of $\lambda$ by $m$ conjugate pairs of $\text{Spec}_2(A, \mathcal{L})$. If $\mathcal{L}$ captures well a basis of only $l < m$ eigenfunctions and rather poorly the remaining $m - l$ elements of $\mathcal{B}$, then it would be expected that only $l$ conjugate pairs of $\text{Spec}_2(A, \mathcal{L})$ will be close to $\lambda$ while the remaining $m - l$ will lie far from this eigenvalue. We can illustrate this locking effect on the multiplicity in a simple numerical experiment.

Let $A = -\partial_x^2 - \partial_y^2$ subject to Dirichlet boundary conditions on $\Omega = [0, \pi]^2 \subset \mathbb{R}^2$. Then $\text{Spec}(A) = \{ j^2 + k^2 : j, k \in \mathbb{N} \}$ and a family of eigenfunctions is given by $u_{jk}(x, y) = \sin(jx) \sin(ky)$. Some eigenvalues of $A$ are simple and some are multiple. In particular, the ground eigenvalue $2 = 1 + 1$ is simple and $1 + k^2$ for $k = 2, 3, 4, 5$ are double. Contrast between the two eigenfunctions $u_{11}$ and $u_{1k}$ will increase as $k$ increases. The former will be highly oscillatory in the $x$-direction while the latter will be so in the $y$-direction. If $\mathcal{L}$ captures well oscillations in only one direction, we should expect one conjugate pair of $\text{Spec}_2(A, \mathcal{L})$ to be close to $1 + k^2$ and the other to be not so close to this eigenvalue.

In order to implement a finite element scheme for the computation of the second-order spectra of the Dirichlet Laplacian, the condition $\mathcal{L} \subset \text{D}(A)$ prescribes the corresponding basis to be at least $C^1$-conforming. We let $\mathcal{L} = \mathcal{L}(s)$ be generated by a basis of Argyris elements on given triangulations of $\Omega$. Contrast between the residues in the interpolation of $u_{1k}$ and $u_{11}$ is achieved by considering triangulation that are stretched either in the $x$ or in the $y$ direction (figure 1b).

Below, we generate triangulations with a fixed total number of 240 elements, resulting from diagonally bisecting a decomposition of $\Omega$ as the union of 120 rectangles of equal ratio. We consider mesh with $s = 3, 4, 5, 6, 8, 10, 12, 15, 20, 24, 30$ and 40 elements of equal size on the lower edge $[0, \pi] \times \{0\} \subset \partial \Omega$.

We then compute the pairs $z_{1k}(s), z_{1k}(s), z_{k1}(s), z_{k1}(s) \in \text{Spec}_2(A, \mathcal{L}(s))$, which are closer to the eigenvalues $1 + k^2$ than to any other point in $\text{Spec}(A)$. We know that $\text{Re} z_{1k}(s) \pm |\text{Im} z_{1k}(s)|$ and $\text{Re} z_{k1}(s) \pm |\text{Im} z_{k1}(s)|$ are guaranteed bounds for this eigenvalue.
Second-order spectra

15

Figure 1. (a) Change in the length of the enclosure predicted by (2.6) for the eigenvalues $\lambda = 1 + k^2$ of the two-dimensional Dirichlet Laplacian on $[0, \pi]^2$. (b) Corresponding mesh for the first six iterates of (a) graph.

Even though we have the same number of elements in each of the above mesh, in general they do not have the same amount of elements intersecting $\partial \Omega$. Then typically $\dim \mathcal{L}(s) \neq \dim \mathcal{L}(t)$ for $s \neq t$, although these numbers do not differ substantially. The precise dimension of the test spaces are: $\dim \mathcal{L}(3) = 2774$, $\dim \mathcal{L}(4) = 2648$, $\dim \mathcal{L}(5) = 2578$, $\dim \mathcal{L}(6) = 2536$, $\dim \mathcal{L}(8) = 2494$ and $\dim \mathcal{L}(10) = 2480$.

In figure 1 we have depicted the residues $2|\text{Im } z_{1k}(s)|$ and $2|\text{Im } z_{k1}(s)|$ for each one of the eigenvalues $1 + k^2$ in the vertical axis, versus $s$ in the horizontal axis on a semi-log scale. The graph suggests that the order of approximation for all the eigenvalues changes by at least two orders of magnitude as $s$ varies. The minimal residue in the approximation of the ground eigenvalue is achieved when $s = 10$ and $s = 12$. This corresponds to low contrast in the basis of $\mathcal{L}(s)$. When the eigenvalue is multiple, however, the minimal residue is achieved by increasing the contrast in the basis. As this contrast increases, one conjugate pair will get closer to the real axis while the other will move away from it. The greater the $k$ is, the greater contrast is needed to achieve a minimal residue and the further away the conjugate pairs travel from each other.

This experiment suggests a natural extension for theorem 3.4. If only $l < m$ members of $\mathcal{B}$ are close to $\mathcal{L}$, then only $l$ conjugate pairs on $\text{Spec}_2(A, \mathcal{L})$ will be close to the corresponding eigenvalue. At present it is not clear whether a more precise statement in this respect can be formulated along the lines of theorem 3.4.

(b) Optimality of convergence to eigenvalues

We saw in example 3.5 that the upper bound established in theorem 3.4 is sub-optimal. We now examine this assertion from a computational perspective.
Let $A = -\partial_x^2 + V$ acting on $\mathcal{H} = L^2[0, \pi]$, where $V$ is a smooth real-valued bounded potential. Let $D(A) = \{ u \in H^2[0, \pi] : u(0) = u(\pi) = 0 \}$, so that $A$ defines a self-adjoint semi-bounded operator. Note that $\lambda \in \text{Spec}(A) = \text{Spec}_{\text{disc}}(A)$ if and only if $\lambda$ solves the Sturm–Liouville eigenvalue problem $Au = \lambda u$ subject to homogeneous Dirichlet boundary conditions at 0 and $\pi$.

Let $\Xi$ be an equidistant partition of $[0, \pi]$ into $n$ sub-intervals $I_l = [x_{l-1}, x_l]$ of length $h = \pi/n = x_l - x_{l-1}$. Let

$$L(h, k, r) = V_h(k, r, \Xi) = \{ v \in C^k(0, \pi) : v \mid_{I_l} \in P_r(I_l), \; 1 \leq l \leq n, \; v(0) = v(\pi) = 0 \}$$

be the finite element space generated by $C^k$-conforming elements of order $r$ subject to Dirichlet boundary conditions at 0 and $\pi$. An implementation of standard interpolation error estimates for finite elements combined with theorem 3.4 ensures the following.

**Lemma 4.1.** Let $A$, $\mathcal{H}$ and $L(h, k, r)$ be as above. Let $\text{Spec}(A) = \{ \lambda_1 < \lambda_2 < \ldots \}$. Let $a_j = (1/4)\lambda_j + (3/4)\lambda_{j-1}$ and $b_j = (1/4)\lambda_j + (3/4)\lambda_{j+1}$, where $\lambda_0 = -\infty$. For all $r > k \geq 3$, there exist a constant $c > 0$, dependent on $j$, $k$ and $r$, but independent of $h$, such that $S_{\text{disc}}(A, L(h, k, r)) \cap \mathbb{D}(a_j, b_j) = \{ z_{hkr}, \bar{z}_{hkr} \} \subset \mathbb{D}(\lambda_j - ch^{(r-3)/2}, \lambda_j + ch^{(r-3)/2})$ for all $h > 0$ sufficiently small.

**Proof.** Use the well-known estimate $\| v - v_h \|_{H^p[0, \pi]} \leq ch^{r+1-p}$, where $v_h \in L(h, k, r)$ is the finite element interpolant of $v \in C^k \cap H^{r+1}_0(0, \pi)$ and note that all eigenvalues of $A$ are simple while their associated eigenfunctions are $C^\infty$.

Therefore, each individual eigenvalue $\lambda_j$ is approximated by second-order spectral points at a rate $O(h^{(r-3)/2})$ for test subspaces generated by a basis of $C^3$-conforming finite elements of order $r > 4$. Because of the high regularity required on the approximating basis, this result is only of limited practical use. In fact, only $k \geq 1$ (and $r \geq 3$) is required for $L(h, k, r) \subset D(A)$. The following simple numerical experiment confirms that the exponent predicted by lemma 4.1 can be improved.

**Example 4.2.** Suppose that $V = 0$ so that $A = -\partial_x^2$. In figure 2 we have fixed an equidistant partition $\Xi$ and let $\mathcal{L} = L(h, 1, r)$ be the space of Hermite elements of order $r = 3, 4, 5$ satisfying Dirichlet boundary conditions in 0 and $\pi$. We then find the conjugate pairs $\{ z_{h1r}(j), \bar{z}_{h1r}(j) \} \in \text{Spec}_{\text{disc}}(A, \mathcal{L})$, which are close to $\lambda_j = j^2$ in $\text{Spec}(A)$ for $j = 1, 2, 3, 4, 5$.

**Example 4.3.** The previous example suggests an order of approximation of the eigenvalue $\lambda$ from its closest $z \in \text{Spec}_{\text{disc}}(A, \mathcal{L})$ of $| z - \lambda | \sim \text{dist}_{D(A)}(B, \mathcal{L})$. Indeed, interpolation by Hermite elements of order $r$ has an $H^2$-error proportional to $h^{r-1}$. The same convergence rate is confirmed by example 3.5. Let $V(x) = 2 \cos(2x)$ be the Mathieu potential. Let $\mathcal{L}$ be as in example 4.2. The exponents $p$ reported in table 1 further confirm this conjecture.

It is well established that the error in the estimation of the eigenvalues of a one-dimensional elliptic problem of order $2p$ by the Galerkin method using Hermite elements of order $r$ is proportional to $h^{2(r+1-p)}$. If $A$ is a second-order differential operator, the quadratic eigenvalue problem (2.2) gives rise to a
Second-order spectra

![Figure 2](link-to-image)

Figure 2. See example 4.2. (a) Log-log plot of $2|\text{Im} z_{h1r}(j)|$ versus element sizes. (b) Value of $p$ such that $|\text{Im} z_{h1r}(j)| \sim h^p$ found via linear fittings. Black, order 3; dark grey, order 4; light grey, order 5; line with circles, $\lambda = 1$; line with crosses, $\lambda = 4$; line with triangles, $\lambda = 9$; line with squares, $\lambda = 16$; line with crosses, $\lambda = 25$.

Table 1. Eigenvalue enclosures for the Mathieu Schrödinger operator. See examples 4.3 and 4.4. We have computed the enclosures of $\lambda_j$ by direct application of (2.6) and by the improved bound (2.7). Here, $\mathcal{L} = \mathcal{L}(\pi/n, 1, r)$ has been chosen.

| $r$ | $j$ | $\lambda$ | $r=3$ | $r=4$ | $r=5$ |
|-----|-----|-----------|-------|-------|-------|
| 3   | 1   | $-0.11087_{117}$ | $3.9212_{127}$ | $9.0619_{334}$ | $16.07_{599}$ | $25.12_{599}$ |
|     |     | $-0.11024_{906}$ | $3.91702_{122}$ | $9.0477_{100}$ | $16.03_{277}$ | $25.02_{199}$ |
|     | $n=10:5:50$ | $|\text{Im} z_{h13}(j)| \sim h^p$ | $p \approx 1.9915$ | $p \approx 1.9847$ | $p \approx 1.9790$ | $p \approx 1.9682$ | $p \approx 1.9532$ |
| 4   | 2   | $-0.1106_{37}$ | $3.91767_{637}$ | $9.04_{555}$ | $16.0_{253}$ | $25.0_{499}$ |
|     |     | $-0.11024_{817}^{170}$ | $3.91702_{282}$ | $9.0477_{385}$ | $16.0329_{748}$ | $25.0208_{795}$ |
|     | $n=9:2:19$ | $|\text{Im} z_{h14}(j)| \sim h^p$ | $p \approx 2.9816$ | $p \approx 2.9680$ | $p \approx 2.9629$ | $p \approx 2.9486$ | $p \approx 2.9238$ |
| 5   | 3   | $-0.1101_{51}$ | $3.91_{636}$ | $9.0488_{661}$ | $16.0_{286}$ | $25.0_{356}$ |
|     |     | $-0.11024_{932}^{697}$ | $3.91702_{506}$ | $9.0477_{385}$ | $16.0329_{277}$ | $25.0208_{210}$ |
|     | $n=9:1:12$ | $|\text{Im} z_{h15}(j)| \sim h^p$ | $p \approx 3.9768$ | $p \approx 4.0214$ | $p \approx 3.9557$ | $p \approx 3.9432$ | $p \approx 3.9183$ |

non-self-adjoint fourth-order problem, which is to be solved by a projection-type method. Thus, example 4.3 is consistent with this estimate if we consider the improved enclosure (2.7).

Proc. R. Soc. A
Table 2. Enclosures for eigenvalues of the perturbed periodic Schrödinger operator. Enclosures for the first three eigenvalues of the operator in example 4.5. The numerical values of $a$ and $b$ are found in Abramowitz & Stegun (1964).

| eigenvalue | upper bounds (2.6) | lower bounds (2.6) | $a$ | $b$ | $l$ |
|------------|-------------------|-------------------|-----|-----|-----|
| $\lambda_1$ | $-0.409535$ | $-0.378490$ | $-\infty$ | $-0.378409$ | 25 |
| $\lambda_2$ | $0.377791$ | $0.347670$ | $0.594800$ | 50 |
| $\lambda_3$ | $1.18219$ | $0.918058$ | $1.29317$ | 100 |

(c) Improved accuracy

Guaranteed enclosures for $\lambda \in \text{Spec}_{\text{disc}}(A)$ can be improved by combining (2.7) with (2.6). For this we should find an upper bound $a > v = \sup\{\tilde{\lambda} \in \text{Spec}(A) : \tilde{\lambda} < \lambda\}$, a lower bound $b < \nu = \inf\{\tilde{\lambda} \in \text{Spec}(A) : \tilde{\lambda} > \lambda\}$ and $z \in \text{Spec}_2(A, \mathcal{L})$ such that $(\Re z - |\Im z|, \Re z + |\Im z|) \cap \text{Spec}(A) = \{\lambda\}$. Here $\mu$ and $\nu$ can be elements of the discrete or the essential spectrum of $A$. The preliminary bounds $a$ and $b$ can be found from $\text{Spec}_2(A, \mathcal{L})$ or by analytical means. If $b - a$ is sufficiently large and $|\Im z|$ is sufficiently small, then (2.7) improves upon (2.6). This technique can be implemented as follows.

Example 4.4. Let $A = -\partial_x^2 + 2 \cos(2x)$, $\mathcal{H} = L^2[0, \pi]$ and $\mathcal{L} = \mathcal{L}(h, 1, r)$ for $r = 3, 4, 5$ be as in examples 4.2 and 4.3. In table 1 we have computed inclusions for the first five eigenvalues of $A$ by directly employing (2.6) and by the technique described in the previous paragraph. In the latter case, we have found $a_j = a$ from the computed upper bound for $\lambda_{j-1} < \lambda_j$ ($a_1 = -\infty$) using (2.6). Similarly for $b_j = b$. Compare these calculations with those of Arbenz (1983).

We can also consider an example from solid-state physics to illustrate the case where $\mu$ and $\nu$ are not in the discrete spectrum.

Example 4.5. Let $A = -\partial_x^2 + \cos(x) - e^{-x^2}$ on $L^2(-\infty, \infty)$, where $D(A) = H^2(-\infty, \infty)$. Then $A$ is a semi-bounded operator but now $\text{Spec}_{\text{ess}}(A)$ consists of an infinite number of non-intersecting bands, separated by gaps, determined by the periodic part of the potential. The endpoints of these bands can be found analytically. They correspond to the so called Mathieu characteristic values. The addition of a fast decaying perturbation gives rise to a non-trivial discrete spectrum. Non-degenerate isolated eigenvalues can appear below the bottom of the essential spectrum or in the gaps between bands.

In table 2 we report on computation of inclusions for the first three eigenvalues in $\text{Spec}_{\text{disc}}(A)$: $\lambda_1 < \min \text{Spec}_{\text{ess}}(A)$, $\lambda_2$ in the first gap of the essential spectrum and $\lambda_3$ in the second gap. No other eigenvalue is to be found in any of these gaps. Here $\mathcal{L}$ is the space of Hermite elements of order 3 on a mesh of 20 segments of
equal size $h = 0.1$ in $[-l, l]$ subject to Dirichlet boundary conditions at $\pm l$. For the improved enclosure, $a$ and $b$ are approximations of the endpoints of the gaps where the $\lambda_j$ lie.

Since the eigenvectors of $A$ decay exponentially fast as $x \rightarrow \pm \infty$, the members of $\text{Spec}_{\text{disc}}(A)$ are close to eigenvalues of the regular Sturm–Liouville problem $Au = \lambda u$, subject to $u(-l) = u(l) = 0$, for sufficiently large $l < \infty$. The numerical method considered in this example does not distinguish between the $l = \infty$ and the large $l < \infty$ eigenvalue problem. For instance, the inclusion found for $\lambda_1$ is also an inclusion for the Dirichlet ground eigenvalue of $A \mid L^2(-25, 25)$. In the case of $\lambda_2$ and $\lambda_3$, which lie in gaps of $\text{Spec}_{\text{ess}}(A)$, they should be close to high-energy eigenvalues of the finite interval problem. Indeed, for the parameters considered in table 2, the inclusion for $\lambda_2$ is also an inclusion for the 17th eigenvalue of $A \mid L^2(-50, 50)$. Similarly that for $\lambda_3$ is an inclusion for the 65th eigenvalue of $A \mid L^2(-100, 100)$.

We are grateful to M. Langer for fruitful discussions. L.B. wishes to acknowledge the hospitality of Ceremade where part of this research was carried out. M.S. gratefully acknowledges the support of EPSRC grant no. EP/E037844/1.

Appendix A. Accumulation points outside the real line

Since $\mathbb{R}$ is second countable and the essential spectrum of $A$ is closed, there always exists a family of open intervals $(a_j, b_j) \subset \mathbb{R}$ such that $\mathbb{R} \setminus \text{Spec}_{\text{ess}}(A) = \bigcup_{j=1}^{\infty} (a_j, b_j)$. Throughout this section we will be repeatedly referring to the following two $A$-dependent regions of the complex plane: $A = \mathbb{A} = \mathbb{D}[\inf \text{Spec}_{\text{ess}} A, \sup \text{Spec}_{\text{ess}} A] \setminus \mathbb{B}$, where $\mathbb{B} = \mathbb{B}_A = \bigcup_{j=1}^{\infty} \mathbb{D}(a_j, b_j)$. For $\Omega \subset \mathbb{C}$ and $\epsilon > 0$ we will denote the open $\epsilon$-neighbourhood of $\Omega$ by $[\Omega]_\epsilon = \{z \in \mathbb{C} : \text{dist}(z, \Omega) < \epsilon\}$.

**Lemma A.1.** Let $\mathcal{N} = \mathcal{N} \subset \mathbb{B}$ be compact and such that $\mathcal{N} \cap \text{Spec}(A) = \emptyset$. Let $(\mathcal{L}_n) \in A_2$. There exists $s_{\mathcal{N}} > 0$ and $N > 0$ such that $\sigma_{A, \mathcal{L}_n}(z) \geq s_\mathcal{N}$ for all $z \in \mathcal{N}$ and $n \geq N$.

**Proof.** It follows from lemmas 2.3 and 3.1 applied in $(a_j + \epsilon, b_j - \epsilon)$, where $\bigcup \mathbb{D}(a_j, b_j)$ is a finite covering of $\mathcal{N}$, by analogous arguments as in the proof of theorem 3.4. ■

Combining this lemma with an argument as in the proof of corollary 3.6, we immediately achieve a generalization of corollary 3.8:

$$
\text{Spec}_{\text{disc}}(A) \subseteq \lim_{n \rightarrow \infty} \text{Spec}_{\text{ess}}(A, \mathcal{L}_n) \subseteq A \cup \text{Spec}_{\text{disc}}(A), \quad (\mathcal{L}_n) \in A_1.
$$

We now examine accumulation of $\text{Spec}_{\text{ess}}(A, \mathcal{L}_n)$ in compact subsets of $A$.

**Lemma A.2.** Let $c \in \text{Spec}_{\text{ess}}(A)$. For any given $m \in \mathbb{N}$ and $\delta > 0$ there exists $\{x_l\}_{l=1}^{m} \subset 1_{(c-\delta, c+\delta)}(A)$ such that

(i) $\|x_l\| = 1$ for all $l = 1, \ldots, m$

(ii) $\langle A^p x_l, A^q x_l \rangle = 0$ for all $l \neq \tilde{l}$ and $p, q = 0, 1$

(iii) $Ax_l = cx_l + \hat{x}_l$ where $\|\hat{x}_l\| < \delta$ for any $l = 1, \ldots, m$.
Proof. If $c$ is isolated from the other points in the spectrum, the proof is elementary. Otherwise, there exists $c_1 \in \text{Spec}(A) \setminus \{c\}$. By substituting $c_l$ by a sub-sequence if necessary, we can assume that $|c_l - c| < \delta/2$ for $0 < \delta(l + 1) < \delta(l)$, such that $(c_l - (\delta(l)/2), c_l + (\delta(l)/2)) \cap \{c_l\}_{l=1}^\infty = \{c_l\}$. If we pick $x_l \in \mathbb{1}_{(c_l-\delta(l)/2, c_l+(\delta(l)/2))}(A)$ for $l = 1, \ldots, m$ with $\|x_l\| = 1$, the desired conclusion is guaranteed.

For $c_1 \leq c_2 \leq c_3$ denote $A(c_1, c_2, c_3) = \mathbb{D}(c_1, c_3) \setminus (\mathbb{D}(c_1, c_2) \cup \mathbb{D}(c_2, c_3))$.

**Lemma A.3.** Let $z \in A(c_1, c_2, c_3)$. For $\delta > 0$ let $x_1, x_2, x_3 \in \mathbb{D}(A)$ be such that (i)–(iii) of Lemma A.2 hold for $m = 3$. There exist $\alpha_k \in \mathbb{R}$ independent of $\delta$ such that if $y = \sum_{k=1}^3 \alpha_k x_k$, the polynomial $\langle Ay, Ay \rangle - 2\lambda \langle Ay, y \rangle + \lambda^2 \langle y, y \rangle$ has roots $\lambda_+ = z + O(\delta)$ and $\lambda_- = z + O(\delta)$.

Proof. The cases where $c_k = c_{\tilde{k}}$ for $k \neq \tilde{k}$ are easy, so we assume that $c_1 < c_2 < c_3$. Moreover, without loss of generality we can fix $c_1 = -1$, $c_2 = c$, where $|c| < 1$ and $c_3 = 1$. Then $z$ and $\tilde{z}$ are roots of the polynomial $a^2 - 2\lambda a \cos \theta + \lambda^2$ for $z = ae^{i\theta}$. If $y = \sum_{k=1}^3 \alpha_k x_k$ we get

$$\langle y, y \rangle = \alpha_1^2 + \alpha_2^2 + \alpha_3^2,$$

$$\langle Ay, Ay \rangle = -\alpha_1^2 + c\alpha_2^2 + \alpha_3^2 + O(\delta)$$

and

$$\langle Ay, Ay \rangle = \alpha_1^2 + c^2\alpha_2^2 + \alpha_3^2 + O(\delta).$$

The solution of the system

$$\alpha_1^2 + \alpha_2^2 + \alpha_3^2 = 1,$$

$$-\alpha_1^2 + c\alpha_2^2 + \alpha_3^2 = a \cos \theta$$

$$\alpha_1^2 + c^2\alpha_2^2 + \alpha_3^2 = a^2$$

is $\alpha_1 = \sqrt{\beta_+}$, $\alpha_2 = \sqrt{(1 - a^2)/(1 - c^2)}$ and $\alpha_3 = \sqrt{\beta_-}$, where $\beta_{\pm} = 1/2((a^2 \pm c)/(1 \pm c) \mp a \cos \theta)$. The proof is completed by an elementary trigonometric argument, showing that $\beta_{\pm} \geq 0$ for any $z \in A(-1, c, 1)$.

**Theorem A.4.** Let $\mathcal{N} = \tilde{\mathcal{N}} \subseteq \mathbb{A}$ be compact. For all $\varepsilon > 0$ there exists $\mathcal{L}_\varepsilon \subset \mathbb{D}(A)$ such that $\mathcal{N} \subset [\text{Spec}_2(A, \mathcal{L}_\varepsilon)]_\varepsilon \subset [\mathcal{N}]_{2\varepsilon}$.

Proof. Let $\mathcal{N}$ and $\varepsilon > 0$ be fixed. Since $\mathcal{N}$ is compact, there exists $\{z_j\}_{j=1}^n \subset \mathcal{N}$, such that $\mathcal{N} \subset \{(z_j)_{j=1}^n \}_{\varepsilon/4} \subset \mathcal{N}_{\varepsilon/2}$. The proof will be completed if $\mathcal{L}_\varepsilon \subset \mathbb{D}(A)$ is such that $\{(z_j)_{j=1}^n \}_{\varepsilon/4} \subset \{\text{Spec}_2(A, \mathcal{L}_\varepsilon)\}_\varepsilon \subset \{(z_j)_{j=1}^n \}_{\varepsilon/2}$. Below we will choose $\delta > 0$ small enough.

Let $\{(c_k)_{k=1,j=1}^n \subset \text{Spec}_\varepsilon(A)\}$ be such that $z_j \in \mathbb{A}(c_{1j}, c_{2j}, c_{3j})$. By virtue of lemma A2 there is a family of vectors $\{x_{kj}\}_{k=1,j=1}^n \subset \mathbb{D}(A)$, such that: $\|x_{kj}\| = 1$ for all $k = 1, 2, 3$ and $j = 1, \ldots, n$; $\langle A^px_{kj}, A^qx_{kj} \rangle = 0$ for all $(k, j) \neq (\tilde{k}, \tilde{j})$ and $p, q = 0, 1$; $Ax_{kj} = c_{kj}x_{kj} + \hat{x}_{kj}$, where $\|\hat{x}_{kj}\| < \delta$ for any $k = 1, 2, 3$ and $j = 1, \ldots, n$. Let $\alpha_{kj} = \alpha_{k}$ be as in lemma A3 for $z = z_j$. Choosing $\delta > 0$ small enough and defining $y_j = \sum_{k=1}^3 \alpha_{kj}x_{kj}$ and $\mathcal{L}_\varepsilon = \text{Span}\{y_1, \ldots, y_n\}$ completes the proof.
This result has two straightforward consequences. Given any compact subset \( \mathcal{N} \subseteq \mathbb{A} \), there exists \( \{ \mathcal{L}_n \} \subset D(A) \), such that \( \mathcal{N} = \lim_{n \to \infty} \text{Spec}_2(A, \mathcal{L}_n) \). Evidently, \( \{ \mathcal{L}_n \} \) might fall outside \( \mathcal{A}_0 \) in general. On the other hand, we can always find \( \{ \mathcal{L}_n \} \in \mathcal{A}_1 \), such that \( \lim_{n \to \infty} \text{Spec}(A, \mathcal{L}_n) = \mathbb{A} \cup \text{Spec}_{\text{disc}}(A) \).

References

Abramowitz, M. & Stegun, I. A. 1964 *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. Washington, DC: US Government Printing Office.

Arbenz, P. 1983 Finite element interpolation error bounds with applications to eigenvalue problems. *Z. Angew. Math. Phys.* 34, 180–191. (doi:10.1007/BF00944591)

Boulton, L. 2007 Non-variational approximation of discrete eigenvalues of self-adjoint operators. *IMA J. Numer. Anal.* 27, 102–121. (doi:10.1093/imanum/drl015)

Boulton, L. & Boussaid, N. 2010 Non-variational computation of the eigenstates of Dirac operators with radially symmetric potentials. *LMS J. Comput. Math.* 13, 10–32. (doi:10.1112/S1461157008000429)

Boulton, L. & Levitin, M. 2007 On approximation of the eigenvalues of perturbed periodic Schrödinger operators. *J. Phys. A* 40, 9319–9329. (doi:10.1088/1751-8113/40/31/010)

Boulton, L. & Strauss, M. 2007 Stability of quadratic projection methods. *Oper. Matrices* 1, 217–233.

Dauge, M. & Suri, M. 2002 Numerical approximation of the spectra of non-compact operators arising in buckling problems. *J. Numer. Math.* 10, 193–219.

Davies, E. B. 1998 Spectral enclosures and complex resonances for general self-adjoint operators. *LMS J. Comput. Math.* 1, 42–74.

Gohberg, I., Lancaster, P. & Rodman, L. 2005 *Indefinite linear algebra and applications*. Basel, Switzerland: Birkhäuser.

Kato, T. 1980 *Perturbation theory for linear operators*. Berlin, Germany: Springer.

Levitin, M. & Shargorodsky, E. 2004 Spectral pollution and second-order relative spectra for self-adjoint operators. *IMA J. Numer. Anal.* 24, 393–416. (doi:10.1093/imananum/24.3.393)

Lewin, M. & Séré, E. 2010 Spectral pollution and how to avoid it (with applications to Dirac and periodic Schrödinger operators). *Proc. Lond. Math. Soc.* 100, 864–900. (doi:10.1112/plms/pdp046)

Plum, M. 1990 Eigenvalue inclusions for second-order ordinary differential operators by a numerical homotopy method. *Z. Angew. Math. Phys.* 41, 205–226. (doi:10.1007/BF00945108)

Rappaz, J., Sanchez Hubert, J., Sanchez Palencia, E. & Vassiliev, D. 1997 On spectral pollution in the finite element approximation of thin elastic ‘membrane’ shells. *Numer. Math.* 75, 473–500. (doi:10.1007/s002110050249)

Shargorodsky, E. 2000 Geometry of higher order relative spectra and projection methods. *J. Operator Theory* 44, 43–62.

Strauss, M. In press. Quadratic projection methods for approximating the spectrum of self-adjoint operators. *IMA J. Numer. Anal.* (doi:10.1093/imanum/drp013)