Operators and Matrices
Volume 1, Number 2 (2007), 217–233

STABILITY OF QUADRATIC PROJECTION METHODS

LYONELL BOULTON AND MICHAEL STRAUSS

(*communicated by Leiba Rodman*)

Abstract. We discuss stability properties of the method studied recently in [7] and [2], for computing eigenvalues in gaps of the essential spectrum of self-adjoint operators.

1. Introduction

1.1. Spectral Pollution in the Galerkin method

Let A be a self-adjoint operator acting on an infinite dimensional Hilbert space $\mathcal{H}$, with a dense domain $\mathrm{Dom}(A)$. The spectrum of A, $\mathrm{Spec}(A)$, may be expressed as the union of the discrete spectrum consisting of all isolated eigenvalues of finite multiplicity, $\mathrm{Spec}_{\mathrm{disc}}(A)$, and the essential spectrum, where $\mathrm{Spec}_{\mathrm{ess}}(A) := \mathrm{Spec}(A) \backslash \mathrm{Spec}_{\mathrm{disc}}(A)$. In most standard situations the essential spectrum can be found analytically, but points in $\mathrm{Spec}_{\mathrm{disc}}(A)$ are usually estimated by numerical procedures.

The estimation of $\mathrm{Spec}_{\mathrm{disc}}(A)$ is often performed through subspaces $\mathcal{L} \subset \mathrm{Dom}(A)$ and corresponding truncations of A. Standard numerical techniques, such as the finite element method, aim at solving Galerkin approximate problems posed in weak form:

$$\text{(P)} \qquad \begin{array}{l} \text{find } 0 \neq u \in \mathcal{L} \text{ and } \lambda \in \mathbb{R} \text{ such that} \\ \langle Au, v \rangle = \lambda \langle u, v \rangle \qquad \forall v \in \mathcal{L} \end{array}$$

where $\mathcal{L}$ is finite dimensional.

Backed by the Rayleigh-Ritz variational principle, when applicable, the Galerkin method represents a powerful tool in the analysis of spectral properties of linear operators. However, the Galerkin method is not foolproof, in general, the solutions of (P) might fail to provide reliable information about the location of $\mathrm{Spec}(A)$ (see [3], [4], [7], [8], [9]).

The drawbacks in the Galerkin method are due in part to the so called spectral pollution phenomenon which we now describe. Let $\mathcal{L}_n \subset \mathrm{Dom}(A)$ be a sequence of subspaces approaching $\mathcal{H}$, as $n \to \infty$ (e.g. satisfying (3) below with $p = 0, 1$ only). Suppose we found $0 \neq u_n \in \mathcal{L}_n$ and $\lambda_n \in \mathbb{R}$ solutions of (P) with $\mathcal{L} = \mathcal{L}_n$, satisfying

Mathematics subject classification (2000): 47B36, 47B39, 81-08.

Key words and phrases: Non-variational projection methods, spectral pollution, numerical approximation of the spectrum.

Paper No. 01-15

$\lambda_n \to \mu$ and $\|u_n\|^{-1}u_n \to w \in \mathrm{Dom}(A)$ in the weak topology. By the approximating property of $\mathcal{L}_n$, we may obtain

$$\langle Aw - \mu w, v\rangle = 0 \qquad \forall v \in \mathrm{Dom}(A),$$

which appears to suggest that μ is in $\mathrm{Spec}(A)$. Unfortunately, the latter conclusion is not ensured in general. Without further information about the structure of A (e.g. compactness properties), w might be $0 \in \mathrm{Dom}(A)$, so convergent solutions of the approximate problem might produce "polluted" sequences $\lambda_n \to \mu \notin \mathrm{Spec}(A)$.

The emergence of spurious eigenvalues in gaps of $\mathrm{Spec}_{\mathrm{ess}}(A)$ represents a serious difficulty in applications such as elasticity theory and solid state physics (see [3] and [9]), as there is no universal recipe to detect or prevent them for a given operator A and sequence of approximate subspaces $\mathcal{L}_n$.

1.2. Pollution-free strategies and quadratic methods

Spectral pollution is a consequence of the fact that in (P) we are truncating simultaneously both u and v. Indeed, let Π be the orthogonal projection onto $\mathcal{L}$ and

$$\hat{F}_{\mathcal{L}}(x) := \min_{0\neq v\in\mathcal{L}} \frac{\|\Pi(x-A)v\|}{\|v\|}.$$

Then $\hat{\lambda} \in \mathbb{R}$ satisfies (P) if, and only if, $\hat{F}_{\mathcal{L}}(\hat{\lambda}) = 0$. That is to say, there exists $\hat{u} \in \mathcal{L}$ such that $(\hat{\lambda} - A)\hat{u} \perp \mathcal{L}$. As $\|(\hat{\lambda} - A)\hat{u}\|/\|\hat{u}\|$ is not guaranteed to be small, we have no indication whether $\hat{\lambda}$ is close to $\mathrm{Spec}(A)$ or not.

This argument suggests that the correct quantity to look at is

$$F_{\mathcal{L}}(x) := \min_{0\neq v\in\mathcal{L}} \frac{\|(x-A)v\|}{\|v\|}.$$

As

$$F_{\mathcal{L}}(x) \geqslant \inf_{u\in\mathrm{Dom}(A)} \frac{\|(x-A)u\|}{\|u\|} = \|(x-A)^{-1}\|^{-1} = \mathrm{dist}\,[x, \mathrm{Spec}(A)],$$

$F_{\mathcal{L}}(x)$ can be close to 0 only when x is close to a point in the spectrum of A.

In [8], Davies and Plum considered a pollution-free strategy for finding $\mathrm{Spec}(A)$ based on computing the profile of $F_{\mathcal{L}}(x)$ for $x \in \mathbb{R}$. If $\mathcal{L} \subset \mathrm{Dom}(A^2)$,

$$\begin{aligned} F_{\mathcal{L}}(x)^2 &= \min_{0\neq v\in\mathcal{L}} \frac{\langle \Pi(x-A)^2 v, v\rangle}{\|v\|^2} = \|[\Pi(x-A)^2 \restriction \mathcal{L}]^{-1}\|^{-1} \\ &= \min_{0\neq v\in\mathcal{L}} \frac{\|\Pi(x-A)^2 v\|}{\|v\|} =: G_{\mathcal{L}}(x). \end{aligned} \tag{1}$$

Therefore estimating $F_{\mathcal{L}}(x)$ reduces to computing eigenvalues of self-adjoint matrices depending on the parameter $x \in \mathbb{R}$.

The approach developed in [8] relies heavily on being able to find accurately a matrix representation for $\Pi(x-A)^2 \restriction \mathcal{L}$ in terms of an orthonormal basis of $\mathcal{L}$. This is a drawback, for instance, if $\mathcal{L}$ is given by the finite element method, where an orthonormalisation of the basis will be numerically expensive.

An alternative pollution-free method which is independent of the matrix representation of $\Pi(x-A)^2 \restriction \mathcal{L}$ is also available and it may be obtained by considering the zeros of the function $G_{\mathcal{L}}(z)$ for $z \in \mathbb{C}$. Typically $G_{\mathcal{L}}(z)$ and $F^2_{\mathcal{L}}(z)$ only coincide at $z \in \mathbb{R}$. The $(2 \dim \mathcal{L})$ zeros of the polynomial $\det\big(\Pi(z-A)^2 \restriction \mathcal{L}\big)$ are the zeros of $G_{\mathcal{L}}(z)$ and, on the other hand, $F_{\mathcal{L}}(z) \neq 0$ unless z is an eigenvalue of A, with corresponding eigenvector $u \in \mathcal{L}$, a very unlikely situation. It is remarkable, however, that the non-real zeros of $G_{\mathcal{L}}(z)$ also provide reliable information about the location of $\mathrm{Spec}(A)$.

This alternative procedure has been recently discussed in [7], [1] and [2], and it can be traced back to [4] and [10]. A central role is played by the problem

$$\text{(Q)} \qquad \begin{array}{l} \text{find } \zeta \in \mathbb{C} \text{ such that } \exists u \in \mathcal{L} \text{ with} \\ \langle Au, Av\rangle - 2\zeta\langle Au, v\rangle + \zeta^2\langle u, v\rangle = 0, \quad \forall v \in \mathcal{L}. \end{array}$$

It is readily seen that $G_{\mathcal{L}}(\zeta) = 0$ if, and only if, ζ is a solution of (Q). The philosophy of the method is to regard (Q), in place of (P), as an approximate spectral problem for operator A.

The following universal non-pollution result justifies favouring (Q) over (P) (see [7, Theorem 2.6] or Theorem 3 below): if ζ is a solution of (Q), then

$$\mathrm{dist}[\mathrm{Re}\ \zeta, \mathrm{Spec}(A)] \leqslant |\mathrm{Im}\ \zeta|. \tag{2}$$

That is to say, ζ can be close to $\mathbb{R}$, *only when* it is also close to the spectrum of A.

Problem (Q) gives rise to a matrix spectral problem quadratic in the spectral parameter. This added complication balances out with the reliability of the method expressed in the above result.

Now, will a solution of (Q) *ever* be close to $\mathbb{R}$? As for the Galerkin method, in general, additional conditions on a sequence of subspaces $\mathcal{L}_n$ are required for convergence. A precise statement reads as follows, see [2] or Theorem 5 below. Let $\lambda \in \mathrm{Spec}_{\mathrm{disc}}(A)$ and Π_n be the orthogonal projection onto $\mathcal{L}_n \subset \mathrm{Dom}(A^2)$. If

$$\|\Pi_n A^p \Pi_n u - \lambda^p u\| \to 0, \qquad \begin{array}{l} \forall p = 0, 1, 2, \\ \forall u \in \mathrm{Dom}(A) : Au = \lambda u, \end{array} \tag{3}$$

then there exists $\zeta_n \in \mathbb{C}$ satisfying (Q) with $\zeta = \zeta_n$ and $\mathcal{L} = \mathcal{L}_n$, such that $|\zeta_n - \lambda| \to 0$. The above hypothesis is fulfilled immediately, for instance, if A is bounded and $\Pi_n v \to v$ for all $v \in \mathcal{H}$.

The combination of these two results appears to provide a general pollution-free procedure for finding discrete eigenvalues of self-adjoint operators. Although this might seem too optimistic at the present moment, one of the advantages of this method lies in the fact that it is applicable without any special restriction upon the structure of $\mathrm{Spec}(A)$. Moreover, the requirements on $\mathcal{L}_n$ are analogous to those needed in the Galerkin method.

1.3. Stability of Quadratic Projection Methods

On the downside, here we are confronted with a more difficult problem to solve. In general, the finite-dimensional eigenvalue problem associated to (Q) is non-Hermitian. Accuracy, as well as stability of the method becomes a delicate matter. The main goal of the present note is to discuss how non-pollution and convergence of the method are affected, when the coefficients of problem (Q) are known only approximately.

In Section 2. we will show that the non-pollution property remains stable in a sense which will be specified below. In Section 3. we will discuss stability of approximation. Note that a consistent formulation of (Q) only requires $\mathcal{L}_n \subset \mathrm{Dom}(A)$, see Remark 4. Under a suitable hypothesis on the subspaces $\mathcal{L}_n$, our Theorem 5 extends the analogous result of [2] by allowing $\mathcal{L}_n \cap [\mathrm{Dom}(A) \setminus \mathrm{Dom}(A^2)] \neq \emptyset$. In the final section we report on various numerical experiments performed on a simple example.

2. Pollution-free Stability

We devote this section to showing that, given error bounds in the computation of the coefficients of problem (Q), it is possible to control errors in the pollution-free estimation of $\mathrm{Spec}(A)$ by the quadratic method described in Section 1.2.

Let us begin by fixing some notation. Below $\mathcal{L}_n$ is an n-dimensional subspace of $\mathrm{Dom}(A)$ with basis $\{e_1, \ldots, e_n\}$. This basis will always be normalised, $\|e_j\| = 1$ for all $j = 1, \ldots, n$. When sufficiently clear from the context, we will suppress the sub-index and write $\mathcal{L} \equiv \mathcal{L}_n$.

For any $u \in \mathcal{L}$, $u = \overline{u_1} e_1 + \cdots + \overline{u_n} e_n$, from which we define the following norm on $\mathcal{L}$,

$$\|u\|_0 := \left(|u_1|^2 + \cdots + |u_n|^2\right)^{\frac{1}{2}}.$$

Since $\mathcal{L}$ is a finite dimensional space, there exists $\beta > 0$, such that

$$\|u\| = \langle u, u \rangle^{\frac{1}{2}} \geqslant \beta \|u\|_0 \quad \forall u \in \mathcal{L}. \tag{4}$$

If $\{e_1, \ldots, e_n\}$ is an orthonormal basis, then $\|\cdot\| = \|\cdot\|_0$. However when the basis is far from being orthonormal, β will be small. We will occasionally write $\mathbf{u} = (u_1, \ldots, u_n) \in \mathbb{C}^n$.

Let matrices A_0, A_1 and A_2 in $\mathbb{C}^{n \times n}$ be given entrywise by

$$[A_0]_{jk} = \langle Ae_j, Ae_k \rangle, \quad [A_1]_{jk} = \langle Ae_j, e_k \rangle, \quad [A_2]_{jk} = \langle e_j, e_k \rangle. \tag{5}$$

Define the matrix polynomial $M(z) \in \mathbb{C}^{n \times n}$ as

$$M(z) := A_0 - 2zA_1 + z^2 A_2, \qquad z \in \mathbb{C}. \tag{6}$$

Then $\zeta \in \mathbb{C}$ is a solution of (Q) if, and only if, $\det[M(\zeta)] = 0$.

The stability results we establish below give a positive answer to the following question. Suppose we are only able to estimate the matrices A_p by $\tilde{A}_p$ and the norm of the error $\|A_p - \tilde{A}_p\| \leqslant \varepsilon_p$, $p = 0, 1, 2$. Can we recover information about the spectrum of A from the approximate problem

$$(\tilde{\mathrm{Q}}) \qquad \text{find } \zeta \in \mathbb{C} : \det[\tilde{A}_0 - 2\zeta\tilde{A}_1 + \zeta^2\tilde{A}_2] = 0$$

with accuracy possibly depending upon ε_p?

The following preliminary result will be needed.

LEMMA 1. *For $z \in \mathbb{C}$ and $\delta > 0$, let*

$$\begin{aligned} J &= [\operatorname{Re} z - |\operatorname{Im} z| - \delta, \operatorname{Re} z + |\operatorname{Im} z| + \delta], \\ \Omega &= \{(x-z)^2 : x \in \mathbb{R}\backslash J\}. \end{aligned}$$

Then,

$$\inf_{\upsilon \in \Omega} \operatorname{Re} \upsilon = 2\delta|\operatorname{Im} z| + \delta^2. \tag{7}$$

Proof. Let $z = a + ib$, $a, b \in \mathbb{R}$, then for any $\upsilon \in \Omega$ we have

$$\upsilon = (x-a)^2 - b^2 - 2b(x-a)i$$

for some $x \in \mathbb{R}\backslash J$. It is clear that $\operatorname{Re} \upsilon > 0$ and moreover

$$\begin{aligned} \inf_{\upsilon \in \Omega} \operatorname{Re} \upsilon &= \inf_{x \in \mathbb{R}\backslash J} (x-a)^2 - b^2 \\ &= (|b| + \delta)^2 - b^2 \\ &= 2\delta|b| + \delta^2 \end{aligned}$$

verifying (7). □

THEOREM 2. *Let A be a self-adjoint operator acting on a Hilbert space $\mathcal{H}$, and $\mathcal{L}$ be an n-dimensional subspace of* $\operatorname{Dom}(A)$. *Let B be a singular $n \times n$ matrix. For any $z \in \mathbb{C}$, let $M(z)$ and β be as in (6) and (4). Let $\alpha_z \in \mathbb{R}$ with $\alpha_z \geqslant \|M(z) - B\|_{\mathbb{C}^n}$. If $\delta > 0$ is such that $(\delta^2 + 2\delta|\operatorname{Im} z|)\beta^2 > \alpha_z$, then*

$$\operatorname{Spec}(A) \cap [\operatorname{Re} z - |\operatorname{Im} z| - \delta, \operatorname{Re} z + |\operatorname{Im} z| + \delta] \neq \emptyset. \tag{8}$$

Proof. Let $\delta > 0$ be as in the hypothesis and suppose the intersection (8) is empty. Using the spectral theorem and (7), we have for all $u \in \mathcal{L}$

$$\begin{aligned} \operatorname{Re}(\overline{\mathbf{u}}^{\mathrm{T}} M(z)\mathbf{u}) &= \operatorname{Re} \sum_{jk=1}^{n} \langle (A-z)e_j, (A-\overline{z})e_k \rangle \, u_k \overline{u}_j \\ &= \operatorname{Re} \sum_{jk=1}^{n} \langle (A-z)\overline{u}_j e_j, (A-\overline{z})\overline{u}_k e_k \rangle = \operatorname{Re} \langle (A-z)u, (A-\overline{z})u \rangle \\ &= \int_{\mathbb{R}} \operatorname{Re}(\lambda - z)^2 d\langle E_\lambda u, u \rangle \geqslant (2\delta|\operatorname{Im} z| + \delta^2)\|u\|^2 \\ &\geqslant (2\delta|\operatorname{Im} z| + \delta^2)\beta^2 \|u\|_0^2 = (2\delta|\operatorname{Im} z| + \delta^2)\beta^2 \|\mathbf{u}\|_{\mathbb{C}^n}^2, \end{aligned}$$

where E_λ is the spectral measure associated to A. It then follows from the Schwarz inequality that for any $\mathbf{u} \in \mathbb{C}^n$

$$\|M(z)\mathbf{u}\|_{\mathbb{C}^n} \geqslant (2\delta|\operatorname{Im} z| + \delta^2)\beta^2 \|\mathbf{u}\|_{\mathbb{C}^n},$$

so that the operator $M(z) : \mathbb{C}^n \to \mathbb{C}^n$ is invertible and

$$\|M(z)^{-1}\|_{\mathbb{C}^n} \leqslant \left((2\delta|\operatorname{Im} z| + \delta^2)\beta^2\right)^{-1} < \alpha_z^{-1}.$$

In particular $\|M(z)^{-1}\|_{\mathbb{C}^n}^{-1} > \|M(z)-B\|_{\mathbb{C}^n}$, from which it follows that B is not singular. The result follows from the obtained contradiction. □

The next theorem is the main result of this section and it is an improvement on [7, Theorem 2.6].

THEOREM 3. *Let A be a self-adjoint operator acting on a Hilbert space $\mathcal{H}$, and $\mathcal{L}$ be an n-dimensional subspace of* $\mathrm{Dom}(A)$. *Define A_0, A_1 and A_2 as in (5). Let $\tilde{A}_p$ be $n \times n$ matrices, such that for $\varepsilon_p \geqslant 0$*

$$\|A_p - \tilde{A}_p\|_{\mathbb{C}^n} \leqslant \varepsilon_p, \qquad p = 0, 1, 2.$$

If the matrix $\tilde{A}_0 - 2\zeta\tilde{A}_1 + \zeta^2\tilde{A}_2$ is singular for some $\zeta \in \mathbb{C}$, then

$$\mathrm{Spec}(A) \cap [\mathrm{Re}\ \zeta - \tilde{\delta}, \mathrm{Re}\ \zeta + \tilde{\delta}] \neq \emptyset \tag{9}$$

for

$$\tilde{\delta} = \sqrt{|\mathrm{Im}\ \zeta|^2 + \beta^{-2}(|\zeta|^2\varepsilon_2 + 2|\zeta|\varepsilon_1 + \varepsilon_0)}.$$

Proof. With the notation of Theorem 2, take $B = \tilde{A}_0 - 2\zeta\tilde{A}_1 + \zeta^2\tilde{A}_2$. Since

$$\|M(\zeta) - B\|_{\mathbb{C}^n} \leqslant (|\zeta|^2\varepsilon_2 + 2|\zeta|\varepsilon_1 + \varepsilon_0),$$

(9) follows from Theorem 2. □

In particular, under the hypothesis above,

$$\mathrm{dist}[\mathrm{Re}\ \zeta, \mathrm{Spec}(A)] \leqslant |\mathrm{Im}\ \zeta| + \beta^{-1}\epsilon \tag{10}$$

with $\epsilon = \sqrt{|\zeta|^2\varepsilon_2 + 2|\zeta|\varepsilon_1 + \varepsilon_0}$. If the basis $\{e_1, \ldots, e_n\}$ is orthonormal, then (9) and (10) hold with $\beta = 1$. Note that the case $\varepsilon_0 = \varepsilon_1 = \varepsilon_2 = 0$ corresponds to [7, Theorem 2.6], see (2).

3. Stability of Convergence in the Quadratic Method

A consistent formulation of (Q) only requires $\mathcal{L} \subset \mathrm{Dom}(A)$. However, the available approximation results for the quadratic method (cf. [1] and [2]) impose the hypothesis $\mathcal{L} \subset \mathrm{Dom}(A^2)$. In this section we show that, if $\mathcal{L}_n \subset \mathrm{Dom}(A)$ approach reasonably well the eigenspace associated to an eigenvalue $\lambda \in \mathrm{Spec}_{\mathrm{disc}}(A)$, then solutions of (Q) will converge to λ in the large n limit, and the process remains stable under perturbation of the matrix coefficients of the polynomial $M(z)$.

REMARK 4. Allowing the possibility of test spaces $\mathcal{L} \not\subseteq Dom(A^2)$ is only relevant when A is unbounded. If A is a differential operator of order $2m$ and the trial spaces are constructed using the finite element method, $\mathcal{L} \subset \mathrm{Dom}(A^2)$ requires C^{4m-1} conforming elements, while $\mathcal{L} \subset \mathrm{Dom}(A)$ only requires C^{2m-1} conforming elements. The performance of the interpolation algorithm in the finite element method is usually compromised as m increases.

Below we highlight explicitly the dependency on n of approximate subspaces and operators, so we denote matrices $M(z)$ and A_p, corresponding to $\mathcal{L}_n$, by $M^{(n)}(z)$ and $A_p^{(n)}$, respectively. We also assume throughout this section that the basis $\{e_1, \dots, e_n\}$ of $\mathcal{L}_n$ is orthonormal. In general we do not assume that $\mathcal{L}_n \subseteq \mathcal{L}_m$ whenever $n < m$. Strictly speaking we should denote the basis functions of $\mathcal{L}_n$ by $\{e_j^{(n)}\}$. However we suppress this notation as no confusion shall arise.

For $u \in \mathrm{Dom}(A)$, the projection of u onto $\mathcal{L}_n$ is then given by

$$\Pi_n u = \sum_{k=1}^{n} \overline{u}_k e_k.$$

Since $\{e_1, \dots, e_n\}$ is orthonormal, we can isometrically identify $\mathcal{L}_n$ with $\mathbb{C}^n$.

Our key result assumes the following hypothesis on the sequence $\mathcal{L}_n$:

$$\text{(H)} \qquad \begin{array}{l} \forall p, q = 0, 1, \text{ and } \forall u \in \mathrm{Dom}(A) : Au = \lambda u, \\ \| \sum_{j=1}^{n} \langle A^p \Pi_n u, A^q e_j \rangle \, e_j - \lambda^{p+q} u \| \to 0, \text{ as } n \to \infty. \end{array}$$

Whenever $\mathcal{L}_n \subset \mathrm{Dom}(A^2)$, (H) reduces to (3). Furthermore, if A is bounded and Π_n converges strongly to the identity, then (H) holds true for all $\lambda \in \mathrm{Spec}_{\mathrm{disc}}(A)$.

The following result is an improvement upon [2, Theorem 2.2].

THEOREM 5. *Let A be a self-adjoint operator on a Hilbert space. Suppose that the sequence of approximate subspaces $\mathcal{L}_n \subset \mathrm{Dom}(A)$ satisfy (H). Let $\lambda \in \mathrm{Spec}_{\mathrm{disc}}(A)$ and let $d := \mathrm{dist}[\lambda, \mathrm{Spec}(A) \setminus \{\lambda\}]$. Given $0 < \delta < d/4$, there always exist N, $\varepsilon_0, \varepsilon_1, \varepsilon_2 > 0$ ensuring the following. If $n > N$ and the matrices $\tilde{A}_p \in \mathbb{C}^{n \times n}$ satisfy*

$$\|\tilde{A}_p - A_p^{(n)}\| < \varepsilon_p, \qquad p = 0, 1, 2,$$

then

(a) we can always find $\zeta \in \mathbb{C}$ with $\det[\tilde{A}_0 - 2\zeta\tilde{A}_1 + \zeta^2\tilde{A}_2] = 0$ and $|\zeta - \lambda| < \delta$,

(b) the set $\{\mu \in \mathbb{C} : \det[\tilde{A}_0 - 2\mu\tilde{A}_1 + \mu^2\tilde{A}_2] = 0\}$ does not intersect the annulus $\{w \in \mathbb{C} : \delta < |w - \lambda| \leqslant d/4\}$.

The proof of this result will be given at the end of this section. It will be a consequence of various technical lemmas, in particular, suitable extensions of [2, Lemmas 5.1 and 5.3] and various regularity properties of $G_{\mathcal{L}}(z)$.

We begin with the rigorous definition of the right hand side of (1) in the case $\mathcal{L} \subset \mathrm{Dom}(A)$. For $z \in \mathbb{C}$, let

$$G_{\mathcal{L}}(z) := \min_{0 \neq \mathbf{v} \in \mathbb{C}^n} \frac{\|M(z)\mathbf{v}\|_{\mathbb{C}^n}}{\|\mathbf{v}\|_{\mathbb{C}^n}}.$$

If $\mathcal{L} \subset \mathrm{Dom}(A^2)$, then $G_{\mathcal{L}}(z)$ coincides with the right hand side of (1). Below we will write $G_n(z) \equiv G_{\mathcal{L}_n}(z)$.

Clearly $G_{\mathcal{L}}(\zeta) = 0$ if, and only if, $\det M(\zeta) = 0$; so the solutions of problem (Q) are completely characterised as the zeros of $G_{\mathcal{L}}(z)$. It is readily seen that:

$$G_{\mathcal{L}}(z) = \|[M(z)]^{-1}\|^{-1} = \text{least singular value of } M(z). \tag{11}$$

In fact $G_{\mathcal{L}}(z)^{-1}$ is a continuous subharmonic function in the region $\{z \in \mathbb{C} : \det M(z) \neq 0\}$, with singularities at the zeros of $\det[M(z)]$ (see e.g. [4] or [2, Lemma 4.1]). This property will play a central role below.

The statement of Theorem 5 will be obtained as a consequence of the fact that $G_{\mathcal{L}}(z)$ is small if, and only if, for small enough $\varepsilon_p > 0$, $z = \zeta$ is a solution of an approximate problem ($\tilde{\mathrm{Q}}$). The following notion, which has recently become standard, will simplify considerably most of our arguments. Let

$$\Lambda^{\mathcal{L}}(\varepsilon_0, \varepsilon_1, \varepsilon_2) := \{z \in \mathbb{C} : G_{\mathcal{L}}(z) - (\varepsilon_0 + 2\varepsilon_1|z| + \varepsilon_2|z|^2) \leqslant 0\}.$$

This set is called the structured pseudospectrum of the matrix polynomial $M(z)$, see [5].

The proof of the following fundamental property of the pseudospectrum is a direct consequence of (11) and [5, Lemma 2.1]. It clearly suggests how to verify the validity of (a) and (b) of Theorem 5.

LEMMA 6. *The complex number* $\zeta \in \Lambda^{\mathcal{L}}(\varepsilon_0, \varepsilon_1, \varepsilon_2)$ *if, and only if,* $\det[\tilde{A}_0 - 2\zeta\tilde{A}_1 + \zeta^2\tilde{A}_2] = 0$ *for some* $\tilde{A}_p \in \mathbb{C}^{n\times n}$ *satisfying* $\|\tilde{A}_p - A_p\| \leqslant \varepsilon_p$, $p = 0, 1, 2$.

Furthermore, cf. [6, Theorem 2.3],

LEMMA 7. *Let* Ω *be a connected component of* $\Lambda^{\mathcal{L}}(\varepsilon_0, \varepsilon_1, \varepsilon_2)$, *such that* $\det M(\mu) = 0$ *for some* $\mu \in \Omega$. *If* $\|\tilde{A}_p - A_p\| \leqslant \varepsilon_p$, *there always exist* $\zeta \in \Omega$ *such that* $\det[\tilde{A}_0 - 2\zeta\tilde{A}_1 + \zeta^2\tilde{A}_2] = 0$.

We now establish two key relations between the large n limit of $G_n(z)$ and $\mathrm{dist}[z, \mathrm{Spec}(A)]^2$ in a neighbourhood of the discrete spectrum of A.

LEMMA 8. *Let* $\lambda \in \mathrm{Spec}_{\mathrm{disc}}A$. *If the sequence of approximate subspaces* $\mathcal{L}_n \subset \mathrm{Dom}(A)$ *satisfy (H), then*

$$\lim_{n\to\infty} G_n(\lambda) = 0.$$

Proof. Let $u \in \mathrm{Dom}(A)\backslash\{0\}$ be such that $Au = \lambda u$. Consider the vector $\mathbf{u} = (\langle e_1, u\rangle, \ldots, \langle e_n, u\rangle)$. We have

$$\begin{aligned}
[M^{(n)}(\lambda)\mathbf{u}]_i &= \sum_{j=1}^{n} [M(\lambda)]_{ij}\langle e_j, u\rangle \\
&= \sum_{j=1}^{n} \langle Ae_i, A\langle u, e_j\rangle e_j\rangle - 2\lambda\langle Ae_i, \langle u, e_j\rangle e_j\rangle + \lambda^2\langle e_i, \langle u, e_j\rangle e_j\rangle \\
&= \langle Ae_i, A\Pi_n u\rangle - 2\lambda\langle Ae_i, \Pi_n u\rangle + \lambda^2\langle e_i, \Pi_n u\rangle,
\end{aligned}$$

so that

$$\begin{aligned}
\|M^{(n)}(\lambda)\mathbf{u}\|^2_{\mathbb{C}^n} &= \sum_{j=1}^{n} |\langle A\Pi_n u, Ae_j\rangle - 2\lambda\langle \Pi_n u, Ae_j\rangle + \lambda^2\langle \Pi_n u, e_j\rangle|^2 \\
&= \|\sum_{j=1}^{n} \langle A\Pi_n u, Ae_j\rangle e_j - 2\lambda\langle \Pi_n u, Ae_j\rangle e_j + \lambda^2\langle \Pi_n u, e_j\rangle e_j\|^2.
\end{aligned}$$

The right hand side converges to zero by virtue of (H). Also $\|\mathbf{u}\|_{\mathbb{C}^n} \to \|u\| \neq 0$ as $n \to \infty$. Now, fix $\varepsilon > 0$. Then, for all n large enough,

$$G_n(\lambda) \leqslant \frac{\|M^{(n)}(\lambda)\mathbf{u}\|_{\mathbb{C}^n}}{\|\mathbf{u}\|_{\mathbb{C}^n}} \leqslant \varepsilon.$$

As $G_n(z)$ is non-negative and ε is arbitrary the lemma follows. □

In general it is possible to construct examples where $\lim_{n\to\infty} G_n(z) = 0$ for certain $z \notin \mathbb{R}$, [1]. However, this is not possible for z in the vicinity of discrete eigenvalues of A.

LEMMA 9. *Let $\lambda \in \mathrm{Spec}_{\mathrm{disc}}(A)$ and let $d > 0$ be as in Theorem 5. Assume that the sequence of approximate subspaces $\mathcal{L}_n \subset \mathrm{Dom}(A)$ satisfy (H). For all $0 < \delta < d/4$, there exist a constant $0 < s \leqslant 1$ such that*

$$\liminf_{n\to\infty} G_n(z) \geqslant s\delta^2 \qquad \text{for all} \qquad \delta \leqslant |z-\lambda| \leqslant d/4. \tag{12}$$

Proof. If $\mathcal{L}_n \subset \mathrm{Dom}(A^2)$, the result has been established in [2, Lemma 5.3]. We treat the more general case by considering approximate subspace $\tilde{\mathcal{L}}_n \subset \mathrm{Dom}(A^2)$ with orthonormal bases sufficiently close to $\mathcal{L}_n$ in the sense specified by (i)-(iii) below.

As a first step, we recall the following standard result. For any $v \in \mathrm{Dom}(A)$, there exists a sequence $v_n \in \mathrm{Dom}(A^2)$ such that $v_n \to v$ and $Av_n \to Av$. That is to say, $\mathrm{Dom}(A^2)$ is a core (in the operator sense) for A.

Let

$$c_n = \max\big\{1, \max_{j,k=1\ldots n;\, p,q=0,1} \{|\langle A^p e_j, A^q e_k\rangle|\}\big\}.$$

Then it is always possible to find a set $\{\tilde{e}_1, \ldots, \tilde{e}_n\} \subset \mathrm{Dom}(A^2)$ such that

(i) $\{\tilde{e}_1, \ldots, \tilde{e}_n\}$ is orthonormal,
(ii) $\|e_j - \tilde{e}_j\| \leqslant c_n^{-1} e^{-n}$,
(iii) $|\langle A^p e_j, A^q e_k\rangle - \langle A^p \tilde{e}_j, A^q \tilde{e}_k\rangle| \leqslant e^{-n}$,

for $j,k = 1,\ldots,n$ and $p,q = 0,1$. We may find $\tilde{e}_j$ by applying the Gram-Schmidt orthogonalisation procedure to vectors of $\mathrm{Dom}(A^2)$ sufficiently close to the e_j.

Let $\tilde{\mathcal{L}}_n := \mathrm{Span}\{\tilde{e}_1, \ldots, \tilde{e}_n\} \subset \mathrm{Dom}(A^2)$. In this proof, the symbol $\sim$ on top of matrices and operators denotes that they are constructed using $\mathcal{L} = \tilde{\mathcal{L}}_n$. Note that (iii) ensures the existence of complex numbers w_{jk}^{pq} such that $|w_{jk}^{pq}| \leqslant 1$ and

$$\langle A^p \tilde{e}_j, A^q \tilde{e}_k\rangle = \langle A^p e_j, A^q e_k\rangle + w_{jk}^{pq} e^{-n}.$$

Property (iii) yields

$$\begin{aligned} |[\tilde{M}^{(n)}(z) - M^{(n)}(z)]_{jk}| &= |2z[\tilde{A}_1^{(n)} - A_1^{(n)}]_{jk} + [\tilde{A}_0^{(n)} - A_0^{(n)}]_{jk}| \\ &\leqslant 2|z|\,|\langle A\tilde{e}_j, \tilde{e}_k\rangle - \langle Ae_j, e_k\rangle| + |\langle A\tilde{e}_j, A\tilde{e}_k\rangle - \langle Ae_j, Ae_k\rangle| \\ &\leqslant (2|z|+1)e^{-n}, \end{aligned}$$

Thus,

$$\|\tilde{M}^{(n)}(z) - M^{(n)}(z)\| \leqslant (2|z|+1)ne^{-n}. \tag{13}$$

Let $u \in \mathrm{Dom}(A)$ be such that $Au = \lambda u$. We next show that (ii) and the fact that (H) holds for $\mathcal{L}_n$, ensures that (H) also holds for $\tilde{\mathcal{L}}_n$. Indeed,

$$\begin{aligned}
&\| \sum_{k=1}^{n} \langle A^p \tilde{\Pi}_n u, A^q \tilde{e}_k \rangle \, \tilde{e}_k - \lambda^{p+q} u \| \\
&\quad \leqslant \| \sum_{k=1}^{n} \langle A^p \tilde{\Pi}_n u, A^q \tilde{e}_k \rangle \, \tilde{e}_k - \langle A^p \Pi_n u, A^q e_k \rangle \, e_k \| \\
&\qquad + \| \sum_{k=1}^{n} \langle A^p \Pi_n u, A^q e_k \rangle \, e_k - \lambda^{p+q} u \| \\
&\quad \leqslant \| \sum_{k=1}^{n} \langle A^p \tilde{\Pi}_n u, A^q \tilde{e}_k \rangle \, \tilde{e}_k - \langle A^p \tilde{\Pi}_n u, A^q \tilde{e}_k \rangle \, e_k \| \\
&\qquad + \| \sum_{k=1}^{n} \langle A^p \tilde{\Pi}_n u, A^q \tilde{e}_k \rangle \, e_k - \langle A^p \Pi_n u, A^q e_k \rangle \, e_k \| \\
&\qquad + \| \sum_{k=1}^{n} \langle A^p \Pi_n u, A^q e_k \rangle \, e_k - \lambda^{p+q} u \| = T_1 + T_2 + T_3.
\end{aligned}$$

Since $\mathcal{L}_n$ satisfies condition (H), $T_3 \to 0$. We must show that the first two terms also converge to zero. Consider the first term,

$$\begin{aligned}
T_1 &= \| \sum_{jk=1}^{n} \langle u, \tilde{e}_j \rangle \, \langle A^p \tilde{e}_j, A^q \tilde{e}_k \rangle \, (\tilde{e}_k - e_k) \| \\
&\leqslant \|u\| \sum_{jk=1}^{n} |\langle A^p \tilde{e}_j, A^q \tilde{e}_k \rangle| \, \|\tilde{e}_k - e_k\| \\
&= \|u\| \sum_{jk=1}^{n} |(\langle A^p e_j, A^q e_k \rangle + w_{jk}^{pq} e^{-n})| \, \|\tilde{e}_k - e_k\| \\
&\leqslant \|u\| n e^{-n} \sum_{k=1}^{n} \|\tilde{e}_k - e_k\| + \|u\| \sum_{jk=1}^{n} |\langle A^p e_j, A^q e_k \rangle| \, \|\tilde{e}_k - e_k\|.
\end{aligned}$$

Using (ii) it is clear that $T_1 \to 0$. For the second term we have,

$$\begin{aligned}
T_2 &= \| \sum_{jk=1}^{n} \langle u, \tilde{e}_j \rangle \, \langle A^p \tilde{e}_j, A^q \tilde{e}_k \rangle \, e_k - \langle u, e_j \rangle \, \langle A^p e_j, A^q e_k \rangle \, e_k \| \\
&= \| \sum_{jk=1}^{n} \langle u, \overline{\langle A^p \tilde{e}_j, A^q \tilde{e}_k \rangle} \, \tilde{e}_j - \overline{\langle A^p e_j, A^q e_k \rangle} \, e_j \rangle \, e_k \| \\
&\leqslant \|u\| \sum_{jk=1}^{n} \| \langle A^p \tilde{e}_j, A^q \tilde{e}_k \rangle \, \tilde{e}_j - \langle A^p e_j, A^q e_k \rangle \, e_j \|
\end{aligned}$$

$$= \|u\| \sum_{jk=1}^{n} \|(\langle A^p e_j, A^q e_k\rangle + w_{jk}^{pq} e^{-n})\tilde{e}_j - \langle A^p e_j, A^q e_k\rangle e_j\|$$

$$\leqslant \|u\| n^2 e^{-n} + \|u\| \sum_{jk=1}^{n} |\langle A^p e_j, A^q e_k\rangle| \|\tilde{e}_j - e_j\|.$$

Again, using (ii) it is clear that $T_2 \to 0$. This ensures that $\tilde{\mathcal{L}}_n$ also satisfies (H) so (12) is valid for $\tilde{G}_n(z)$.

The proof of (12) follows. Fix $\varepsilon > 0$. Let $\mathbf{v}_n \in \mathbb{C}^n$ such that $\|\mathbf{v}_n\| = 1$ and

$$\|M^{(n)}(z)\mathbf{v}_n\| \leqslant G_n(z) + \varepsilon.$$

Then, by virtue of (13),

$$\begin{aligned} \tilde{G}_n(z) &\leqslant \|\tilde{M}^{(n)}(z)\mathbf{v}_n\| \\ &\leqslant \|[\tilde{M}^{(n)}(z) - M^{(n)}(z)]\mathbf{v}_n\| + \|M^{(n)}(z)\mathbf{v}_n\| \\ &\leqslant (2|z|+1)ne^{-n} + G_n(z) + \varepsilon. \end{aligned}$$

As this happens for all $\varepsilon > 0$, the fact that $\tilde{G}_n(z)$ satisfies (12) implies that also $G_n(z)$ satisfies this inequality. □

Proof of Theorem 5. Let $0 < s \leqslant 1$ be as in Lemma 9 and d be as in the hypothesis of the theorem. By virtue of Lemmas 8 and 9, there exists $N > 0$ such that, $G_n(\lambda) \leqslant s\delta^2$ and

$$G_n(z) > s\delta^2 \quad \text{whenever } \delta < |z - \lambda| \leqslant d/4, \tag{14}$$

for all $n > N$. The subharmonicity of $G_n(z)^{-1}$ ensures that the only local minima of $G_n(z)$ are those points where the function vanishes. Thus, for all $n > N$, there always exists $\zeta_n \in \mathbb{C}$ satisfying

$$|\zeta_n - \lambda| < \delta \qquad \text{and} \qquad G_n(\zeta_n) = 0. \tag{15}$$

Let $\varepsilon_0, \varepsilon_1, \varepsilon_2 > 0$ be small enough such that

$$\varepsilon_0 + 2\varepsilon_1|z| + \varepsilon_2|z|^2 < s\delta^2, \qquad |z - \lambda| < d/4.$$

Suppose that $n > N$. Then, by (14),

$$G_n(z) - (\varepsilon_0 + 2\varepsilon_1|z| + \varepsilon_2|z|^2) > 0$$

for all $\delta < |z - \lambda| \leqslant d/4$, so $\Lambda^{\mathcal{L}_n}(\varepsilon_0, \varepsilon_1, \varepsilon_2) \cap \{\delta < |z - \lambda| \leqslant d/4\} = \emptyset$. This, along with Lemma 6, ensures (b). On the other hand, by virtue of (15), $\Lambda^{\mathcal{L}_n}(\varepsilon_0, \varepsilon_1, \varepsilon_2) \cap \{|z - \lambda| < \delta\} \neq \emptyset$. Thus Lemma 7 yield (a). □

4. Case Study

Finite rank perturbations of multiplication operators have been considered previously in connection with spectral pollution (see [8], [7] and [2]) due to their simple structure. In this final section we report on various numerical experiments we have performed on a model operator of this type.

Let $e_k(x) = (2\pi)^{-\frac{1}{2}} e^{ikx}$ and

$$a(x) = \begin{cases} 0 & \text{for } -\pi \leqslant x < 0, \\ 1 & \text{for } 0 \leqslant x < \pi. \end{cases}$$

In this section we assume that

$$\begin{aligned} \mathcal{H} &= L^2[-\pi, \pi], \\ \mathcal{L}_{2n+1} &= \operatorname{span}\{e_{-n}(x), \ldots, e_n(x)\} \qquad \text{and} \\ A\phi(x) &= a(x)\phi(x) + \langle \phi, e_0 \rangle e_0(x), \qquad \phi \in \mathcal{H}. \end{aligned}$$

Clearly A is bounded and self-adjoint in $\mathcal{H}$. Moreover, the spectrum of A is found explicitly. Since A is a rank one perturbation of the multiplication operator by the symbol a, Weyl's Theorem ensures that $\operatorname{Spec}_{\text{ess}}(A) = \operatorname{Range}(a) = \{0, 1\}$. On the other hand, the isolated eigenvalues of finite multiplicity of A are the solutions of $\langle (\lambda - a)^{-1} e_0, e_0 \rangle = 1$, [4]. A straightforward calculation reveals the two solutions $\lambda^{\pm} = 1 \pm \sqrt{2}/2$, which comprise the discrete spectrum of A. The eigenvalue λ^- is inside the gap $(0, 1)$ of the essential spectrum.

As the symbol $a(x)$ is discontinuous, the Fourier basis $\{e_k\}$ is not a good choice for approximating λ^- using the Galerkin method. Indeed, the solutions of (P) pollute the whole interval $[0, 1]$ as the dimension of $\mathcal{L}_{2n+1}$ increases. Let us test the quadratic method described in the preceding sections in this very simple model.

Since A is bounded and $\Pi_{2n+1}\phi \to \phi$ for all $\phi \in \mathcal{H}$, condition (H) of Section 3. is satisfied. Thus, by virtue of Theorem 5, both discrete eigenvalues are approached by solutions of (Q) as $n \to \infty$, free from spectral pollution according to Theorem 3.

All the calculation described in this section were carried out using the computer package MATLAB. Fully functional m-codes are available at the web page [11].

We compute the exact solutions of (Q), by finding the $\zeta \in \mathbb{C}$ such that $\det M^{(2n+1)}(\zeta) = 0$. The matrix coefficients $A_p^{(2n+1)}$ may be found explicitly using (5). They are sparse and Hermitian with entries either purely real or purely imaginary. The errors in solving (Q) are negligible for n of reasonable size (< 1000).

In order to test the results established in the previous sections, we force large errors in the matrix entries, and compute the corresponding "perturbed" solution of the problem ($\tilde{\text{Q}}$). For simplicity, we fix $\varepsilon_0 = \varepsilon_1 = \varepsilon_2 = \varepsilon$.

Let

$$[\tilde{A}_p]_{jk} = [A_p]_{jk} + \frac{\varepsilon}{2n+1} \alpha_{jk}^{(p)} \tag{16}$$

where $\alpha_{jk}^{(p)}$ are random variables sampled from the unit disk $\{|z| \leqslant 1\}$ with additional constraints specified below. Then

$$\begin{aligned}\|(A_p-\tilde{A}_p)\mathbf{u}\|^2 &= \sum_{j=1}^{2n+1}\left|\sum_{k=1}^{2n+1}\frac{\varepsilon}{2n+1}\alpha_{jk}^{(p)}u_k\right|^2\\ &\leqslant \frac{\varepsilon^2}{(2n+1)^2}\sum_{k=1}^{2n+1}|u_k|^2\sum_{jk=1}^{2n+1}|\alpha_{jk}^{(p)}|^2\\ &\leqslant \frac{\varepsilon^2}{(2n+1)^2}\sum_{k=1}^{2n+1}|u_k|^2\sum_{jk=1}^{2n+1}1\leqslant\varepsilon^2\sum_{k=1}^{2n+1}|u_k|^2,\end{aligned}$$

so $\|A_p-\tilde{A}_p\|\leqslant\varepsilon$. Moreover this bound is sharp. Indeed, the matrix T such that $[T]_{jk}=\frac{\varepsilon}{2n+1}$ for all $1\leqslant j,k\leqslant 2n+1$, satisfies

$$\|T\|^2=\|T^*T\|=\varepsilon\|T\|.$$

We consider two types of restrictions on the random variable $\alpha_{jk}^{(p)}$. On the one hand, Theorem 3 covers the general situation of moving all entries of A_p along randomly chosen directions in the complex plane. Thus, we perform *unstructured perturbations* by allowing all $\alpha_{jk}^{(p)}\neq 0$. On the other hand, however, in order to reproduce the effect made by rounding errors in the estimation of the entries, we perform *non-zero-Hermitian perturbations* by imposing the condition:

$$\alpha_{jk}^{(p)}=\begin{cases}0 & \text{if } [A_p]_{jk}=0,\\ \overline{\alpha_{kj}^{(p)}}\neq 0 & \text{if } [A_p]_{jk}\neq 0.\end{cases}$$

In Figure 1. we depict the exact solutions of (Q) for $n=50$. According to (2), the points which are in the vicinity of the real axis are necessarily close to the spectrum.

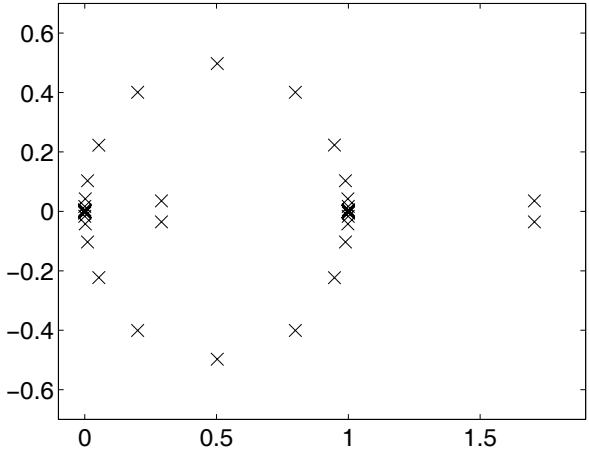

Figure 1. Exact solutions to (Q) *for* $n=50$.

Figure 2., on the other hand, depicts the solutions of $(\tilde{\mathrm{Q}})$ corresponding to 100 different random perturbations. Each of the graphs were constructed by prescribing a different constraint on the random variables. Here $n = 50$ and $\varepsilon = 10^{-1}$. From the perturbed solutions one can identify $\mathrm{Spec}(A)$ less accurately but, once again, without pollution by virtue of Theorem 3. The correction δ of Theorem 3, will depend on ζ and ε, but notably not on n. Furthermore, Theorem 5 ensures that the clouds observed in Figure 2. will cluster near to each of the exact solutions of (Q) as $\varepsilon \to 0$.

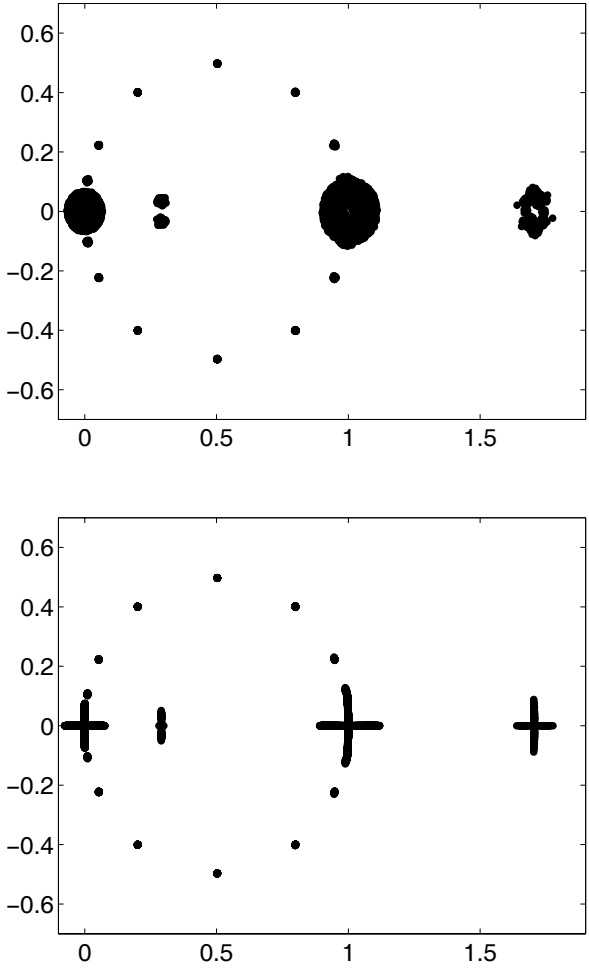

Figure 2. Top: solutions to $(\tilde{\mathrm{Q}})$ *for 100 unstructured random perturbations. Bottom: solutions to* $(\tilde{\mathrm{Q}})$ *for 100 non-zero-Hermitian random perturbations. Here* $\varepsilon = 10^{-1}$.

Figures 3.-4. show the outcome of running Monte Carlo simulations in this model. We fix again $\varepsilon = 10^{-1}$. These pictures have been constructed in the following manner. For each fixed n, we have found ζ_n^-, the closest point to the eigenvalue λ^- such that $\det M^{(2n+1)}(\zeta_n^-) = 0$. Then we have performed 20 constrained perturbations and averaged the solutions of $(\tilde{\mathrm{Q}})$ which are closest to ζ_n^-. We know that solutions of the approximate problems close to ζ_n^- always exist, as a consequence of Theorem 5. Denote these averages by $\zeta_n^{\mathrm{u},-}$ and $\zeta_n^{\mathrm{s},-}$ for the unstructured and non-zero-Hermitian cases respectively. In Figure 3. we depict $|\mathrm{Im}\ \zeta_n^-|$, $|\mathrm{Im}\ \zeta_n^{\mathrm{u},-}|$ and $|\mathrm{Im}\ \zeta_n^{\mathrm{s},-}|$ for $n = 5 : 10 : 100$. Similarly in Figure 4. we depict $|\lambda^- - \mathrm{Re}\ \zeta_n^-|$, $|\lambda^- - \mathrm{Re}\ \zeta_n^{\mathrm{u},-}|$ and $|\lambda^- - \mathrm{Re}\ \zeta_n^{\mathrm{s},-}|$.

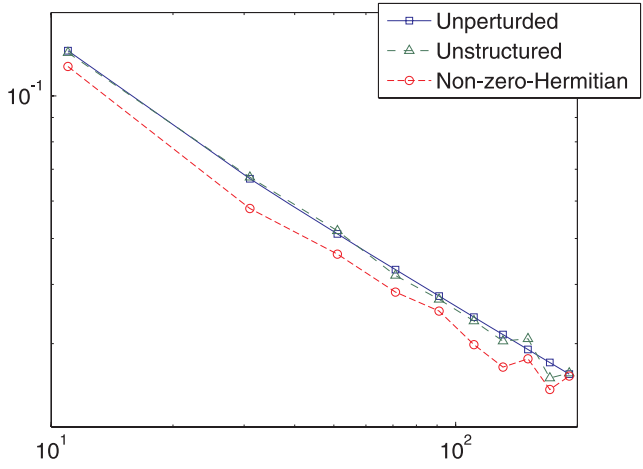

Figure 3. Error predicted by Theorem 3 in the approximation of λ^-. Here we depict $|\mathrm{Im}\ \zeta_n^-|$ (unperturbed), $|\mathrm{Im}\ \zeta_n^{\mathrm{u},-}|$ and $|\mathrm{Im}\ \zeta_n^{\mathrm{s},-}|$ for $n = 5 : 10 : 100$. We average the two perturbed solutions of $(\tilde{\mathrm{Q}})$ over a sample of 20 problems with $\varepsilon = 10^{-1}$. The scaling is log-log and the horizontal axis shows $2n + 1$.

Figure 3. provides clear numerical evidence that the convergence of the quadratic method applied to this simple model is not lost even when the perturbations are large in modulus. Figures 4. suggests that structured perturbations are considerably superior to the unstructured ones, in the test $|\lambda^- - \mathrm{Re}\ \zeta_n^-|$.

By combining Figure 3. and Theorem 3, we immediately predict a rate of convergence of $|\lambda^- - \mathrm{Re}\ \zeta_n^-| = o(n^{-r})$ for $r \approx 1/2$. It is remarkable, however, that Figure 4. strongly suggests an actual exponent of $r \approx 1$ for this rate of convergence. An explanation of this phenomenon is linked to the fact that λ^- is an isolated point of the spectrum, see [10, Section 2]. We will be reporting on this issue elsewhere.

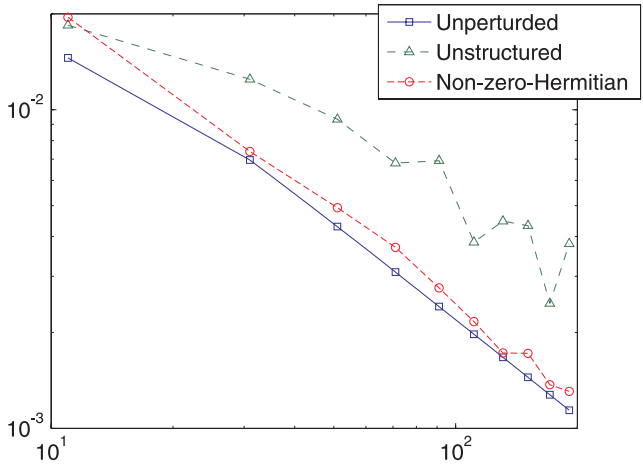

Figure 4. Actual error in the approximation of λ^- *using the real part of solutions of* (Q) *and* $(\tilde{Q})$. *Here we depict* $|\lambda^- - \mathrm{Re}\ \zeta_n^-|$ *(unperturbed),* $|\lambda^- - \mathrm{Re}\ \zeta_n^{u,-}|$ *and* $|\lambda^- - \mathrm{Re}\ \zeta_n^{s,-}|$ *for* $n = 5 : 10 : 100$. *We average the two perturbed solutions of* $(\tilde{Q})$ *over a sample of* 20 *problems with* $\varepsilon = 10^{-1}$. *The scaling is log-log and the horizontal axis shows* $2n + 1$.

5. Acknowledgements

We kindly thank Eugene Shargorodsky and Michael Levitin for encouraging us to write this manuscript in a first place and for their valuable comments during the various stages of its preparation.

REFERENCES

[1] L. Boulton, *Limiting set of second order spectrum*, to appear Math. Comp., **75** (2006) 1367–1382.

[2] L. Boulton, *Non-variational approximation of discrete eigenvalues of self-adjoint operators*, IMA J. Numer. Anal., **27** (2007) 102–121.

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

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

[5] N. J. Higham, F. Tisseur, *Structured pseudospectra for polynomial eigenvalue problems with applications*, SIAM J. Matrix Anal. Appl., **23** (2001) 187–208.

[6] P. Lancaster, P. Psarrakos, *On the pseudospectra of matrix polynomial*, SIAM J. Matrix Anal. Appl., **27** (2005) 115–129.

[7] M. Levitin, E. Shargorodsky, *Spectral pollution and second order relative spectra for self-adjoint operators*, IMA J. Numer. Anal., **24** (2004) 393–416.

[8] E.B. Davies, M. Plum, *Spectral pollution*, IMA J. Numer. Anal., **24** (2004) 417–438.

[9] J. Rapaz, J. Sanchez Hubert, J. Sanchez Palencia, D. Vasiliev, *On spectral pollution in the finite element approximation of thin elastic 'membrane' shells*, Numer. Math., **75** (1997) 473–500.
[10] E. Shargorodsky, *Geometry of higher order relative spectra and projection methods*, J. Oper. Theo., **44** (2000) 43–62.
[11] Web page http://www.ma.hw.ac.uk/∼lyonell/stable

(Received June 16, 2006)

Lyonell Boulton
Department of Mathematics and
Maxwell Institute for Mathematical Sciences
Heriot-Watt University, Edinburgh EH14 2AS, Scotland
e-mail: L.Boulton@hw.ac.uk

Michael Strauss
Department of Mathematics, Kings College London
Strand, London WC2R 2LS, England
e-mail: michael.strauss@kcl.ac.uk