PERTURBATION AND SUPERCONVERGENCE FOR EIGENVALUES IN GAPS

MICHAEL STRAUSS†

Abstract. We consider the problem of how to compute the spectrum of a self-adjoint operator when a direct application of the Galerkin (finite-section) method is unreliable. Typically, it is eigenvalues in gaps of the essential spectrum which are hard to approximate. A new perturbation method for identifying these eigenvalues has recently emerged. The idea being to perturb them off the real line and, consequently, away from regions where the Galerkin method fails. We propose a much simpler approach to this idea which results in an extremely accurate technique. We present new perturbation results for a non-self-adjoint perturbation of a self-adjoint operator. These enable us to control, very precisely, how eigenvalues are perturbed from the real line. We then show that our method is superconvergent. The main results are demonstrated with examples including magnetohydrodynamics, Schrödinger and Dirac operators.

Keywords: Eigenvalue problem, perturbation of eigenvalues, spectral pollution, Galerkin method, finite-section method, superconvergence.

2010 Mathematics Subject Classification: 47A55, 47A58.

† School of Mathematics, Cardiff University, Cardiff CF24 4AG, Wales, UK. Email: michaelstrauss27@hotmail.com, Tel: 0044(0)7580574971.

1. Introduction

Computational spectral theory for operators which act on infinite dimensional Hilbert spaces has advanced significantly in recent years. For self-adjoint operators, the introduction of quadratic methods has enabled the approximation of those eigenvalues which are not reliably located by direct application of the Galerkin method. The latter is due to spectral pollution; see Section 2 and [2, 3, 7, 10, 13, 19, 22]. Eigenvalues affected by pollution are typically (though not exclusively, see examples [4, 5, & 5.1]) located within gaps in the essential spectrum. Notable amongst the quadratic techniques are the Davies & Plum method [30], the Zimmermann & Mertins method [13], and the second order relative spectra [4, 5, 6, 7, 11, 19, 21, 25, 26, 27]. Although not hindered by spectral pollution, these methods do have many drawbacks and we shall review some of these in Section 2. Spectral approximation with the second order relative spectra has also been studied for normal operators; see [27]. For recent advances in spectral approximation for arbitrary operators see, for example, [13, 15, 16] and references therein.

The present manuscript is concerned with a technique for self-adjoint operators which is Galerkin based, pollution-free, and non-quadratic. The idea is to perturb eigenvalues, away from \( \mathbb{R} \) and into \( \mathbb{C}^+ \), then approximate them with the Galerkin method. This idea was initially proposed for a particular class of differential operators; see [20, 21]. An abstract version of this approach, for arbitrary bounded
self-adjoint operators, was formulated in [28]. In this general setting, the eigenvalues are perturbed using a spectral projection which is obtained from the Galerkin method; this requires a priori information about the location of gaps in the essential spectrum. We shall remove the requirement of this spectral projection and, consequently, also of the a priori information. Unless stated otherwise, $A$ will denote a semi-bounded (from below) self-adjoint operator acting on a Hilbert space $\mathcal{H}$. The quadratic form, spectrum, resolvent set, discrete spectrum, essential spectrum and spectral measure we denote by $a$, $\sigma(A)$, $\rho(A)$, $\sigma_{\text{dis}}(A)$, $\sigma_{\text{ess}}(A)$ and $E$, respectively.

We now briefly describe our main results.

Section 3 is concerned with the spectrum of operators of the form $A + iP$ where $P$ is self-adjoint with $0 \leq P \leq I$. The case of a non-self-adjoint perturbation, of a self-adjoint operator, has been largely neglected; see for example [12] and references therein. We show that the region into which new spectra may be introduced, by the non-self-adjoint perturbation $iP$, is far smaller and more nuanced than shown in [28, Section 3]. Our approach, which is based on a spectral enclosure result due to Kato, appears to lend itself to more general situations than those considered here; this will be pursued in a subsequent paper. The main results are Theorem 3.6 and Corollary 3.9. For any gap in $\sigma(A)$, Theorem 3.6 provides a region in $\rho(A + iP)$ and bounds on the resolvent. For any gap in $\sigma_{\text{ess}}(A)$, Corollary 3.9 imposes a condition on $P$ which ensures a region in $\rho(A + iP)$ and bounds on the resolvent. Section 4.2 concerns the spectra of $A + iP_n$ where $(P_n)_{n \in \mathbb{N}}$ is a sequence of orthogonal projections. By Theorem 4.6, for an eigenvalue $\lambda \in \sigma_{\text{dis}}(A)$ we have

$$\text{dist}(\lambda + i, \sigma(A + iP_n)) \to 0$$

where the convergence rate is typically faster than superconvergence; even an order of magnitude faster as demonstrated by Example 4.7. This rapid convergence is an unexpected and extremely attractive property of the perturbation method. Section 4.3 is concerned with convergence properties of the Galerkin method when applied to $A + iP_n$ for a fixed $n \in \mathbb{N}$. Theorem 4.10 shows that we achieve superconvergence to a neighbourhood of $\lambda + i$. By considering the convergence to a neighbourhood of $\lambda + i$, rather than eigenvalues of $A + iP_n$, we avoid the issue of non-semi-simple eigenvalues which would compromise convergence rates. The superconvergence of the perturbation method is demonstrated in Example 4.11. In Section 5 we apply the method to several more unbounded operators arising in magnetohydrodynamics, non-relativistic and relativistic quantum mechanics. Our theoretical results are, for the most part, focused on the perturbation and approximation of $\sigma_{\text{dis}}(A)$, however, the examples indicate that our new perturbation method may also capture $\sigma_{\text{ess}}(A)$.

2. Galerkin and Quadratic methods

The Galerkin eigenvalues of $A$ with respect to a finite-dimensional trial space $\mathcal{L} \subset \text{Dom}(a)$, denoted $\sigma(A, \mathcal{L})$, consists of those $\mu \in \mathbb{C}$ for which $\exists u \in \mathcal{L} \setminus \{0\}$ with

$$a(u, v) = \mu(u, v) \quad \forall v \in \mathcal{L}.$$ 

Unless stated otherwise, $(\mathcal{L}_n)_{n \in \mathbb{N}} \subset \text{Dom}(a)$ will be a sequence of finite-dimensional trial spaces with corresponding sequence of orthogonal projections $(P_n)$. We shall always assume that:

$$\forall u \in \text{Dom}(a) \quad \exists u_n \in \mathcal{L}_n : \|u - u_n\|_a \to 0$$
where \( \| \cdot \|_a \) is the norm associated to the Hilbert space \( \mathcal{H}_a \) with inner-product

\[
\langle u, v \rangle_a := a(u, v) - (m - 1)(u, v) \quad \forall u, v \in \text{Dom}(a) \quad \text{where} \quad m = \min \sigma(A).
\]

The distance from a subspace \( M \subset \mathcal{H} \) to another subspace \( N \subset \mathcal{H} \) is defined as

\[
\delta(M, N) = \sup_{u \in M, \|u\| = 1} \text{dist}(u, N),
\]

the gap between the two subspaces is

\[
\hat{\delta}(M, N) = \max \{ \delta(M, N), \delta(N, M) \};
\]

see [18, Section IV.2.1] for further details. We shall write \( \delta_a \) and \( \hat{\delta}_a \) to indicate the distance and gap between subspaces of \( \mathcal{H}_a \). For trial spaces satisfying (2.1) the Galerkin method is an extremely powerful tool for approximating those eigenvalues which lie below the essential spectrum; see for example [9]. It is well-known that

\[
\lim_{n \to \infty} \sigma(A, \mathcal{L}_n) \cap (-\infty, \min \sigma_{\text{ess}}(A)) = \sigma(A) \cap (-\infty, \min \sigma_{\text{ess}}(A)).
\]

Furthermore, for an eigenvalue \( \lambda < \min \sigma_{\text{ess}}(A) \) with eigenspace \( \mathcal{L}(\{\lambda\}) \), we have

the superconvergence property

\[
\text{dist}(\lambda, \sigma(A, \mathcal{L}_n)) = \mathcal{O}(\delta_a(\mathcal{L}(\{\lambda\}), \mathcal{L}_n)^2).
\]

In general, the Galerkin method cannot be relied upon for approximating eigenvalues above \( \min \sigma_{\text{ess}}(A) \). This is due to a phenomenon known as spectral pollution which is the presence of sequences of Galerkin eigenvalues which converge to points in \( \rho(A) \). A typical situation is \( \min \sigma_{\text{ess}}(A) \leq \alpha < \beta \), \( (\alpha, \beta) \cap \sigma_{\text{ess}}(A) = \emptyset \), and

\[
\lim_{n \to \infty} \sigma(A, \mathcal{L}_n) \cap (\alpha, \beta) = (\alpha, \beta).
\]

Hence, any approximation of \( \sigma_{\text{dis}} \cap (\alpha, \beta) \) is lost within an increasingly dense fog of spurious Galerkin eigenvalues; see examples [14] & [15] and [2, 3, 7, 10, 13, 19, 22]. Although this means that a direct application of the Galerkin method often fails to identify eigenvalues, in view of (2.2) and (2.3), there is every reason to suppose that eigenvalues above \( \min \sigma_{\text{ess}}(A) \) could, in principle, be approximated with a superconvergent technique using trial spaces satisfying only (2.1). The absence of such a technique has resulted in the development of quadratic methods.

The quadratic methods are so-called because of their reliance on truncations of the square of the operator in question; the Galerkin method relies only on the quadratic form. They have been studied and applied extensively over the last two decades. The quadratic method which has received the most attention is the second order relative spectrum. This is because it can be applied without \( \alpha \) priori information and it was widely thought to approximate the whole spectrum of an arbitrary self-adjoint operator. The latter has recently been shown to be false; see [26]. However, it is known that the method will reliably approximate the discrete spectrum of a self-adjoint operator and part of the discrete spectrum of a normal operator; see [4] and [24], respectively. The appeal of quadratic methods is that they can approximate eigenvalues without interference from spectral pollution, in fact, they can even provide enclosures for eigenvalues. The latter is often regarded as a major selling point of these methods. In practice though, we are more likely to be interested in accuracy and convergence rather than enclosures.
A drawback of quadratic methods is that they require trial spaces to belong to the operator domain. From a computational perspective this can be highly awkward as typically FEM software will not support the operator domain. Particularly inconvenient, is that for a second order differential operator we cannot use the standard FEM space of piecewise linear trial functions. Furthermore, it is straightforward to show that (2.1), with the added condition $L_n \subset \text{Dom}(A)$ $\forall n \in \mathbb{N}$, is not sufficient to ensure approximation of $\sigma_{\text{dis}}(A)$. A sufficient condition is

$$\forall u \in \text{Dom}(A) \exists u_n \in L_n : \|u - u_n\|_A \to 0;$$

see [7, Corollary 3.6]. With (2.4) satisfied, we have for each $\lambda \in \sigma_{\text{dis}}(A)$ an element $z_n$ belonging to the second order spectrum of $A$ relative to $L_n$ with

$$|\lambda - z_n| = O(\delta_A(L(\{\lambda\}), L_n)) \quad \text{and} \quad |\lambda - \text{Re } z_n| = O(\delta_A(L(\{\lambda\}), L_n)^2)$$

where $\delta_A(L(\{\lambda\}), L_n)$ is the distance from the eigenspace $L(\{\lambda\})$ to the trial space $L_n$ with respect to the graph norm; see [27, Section 6]. That the convergence rate is measured in terms of the graph norm means that the convergence, and therefore the accuracy, of this method can be poor when compared to the superconvergence of the Galerkin method; see Example 5.4 and [7, Example 3.5 & 4.3]. Convergence rates analogous to the right hand side of (2.5) are also known for the Davies & Plum and Zimmermann & Mertins methods; see [8, Lemma 2].

3. The spectrum of $A + iP$

In this section we are concerned with the spectrum of operators of the form $A + iP$ where $P$ is self-adjoint with $0 \leq P \leq I$. Evidently, we have

$$\sigma(A + iP) \subseteq W(A + iP) \subseteq \{z \in \mathbb{C} : \min \sigma(A) \leq \text{Re } z \text{ and } 0 \leq \text{Im } z \leq 1\}$$

where $W(\cdot)$ is the numerical range. We assume throughout that $\alpha, \beta \in \mathbb{R}$ with $\alpha < \beta$, and we denote $\Delta := (\alpha, \beta)$. The spectral projection associated to $A$ and the interval $\Delta$ is denoted by $E(\Delta)$ and the range of $E(\Delta)$ is denote $L(\Delta)$.

**Definition 3.1.** The function $\gamma^{\pm}(t) : [0, 1] \to \mathbb{C}$ we define as:

1. if $\beta - \alpha < 1$ we set
   $$\text{Re } \gamma^{\pm}(t) = \alpha(1 - t) + \beta t,$$
   $$\text{Im } \gamma^{\pm}(t) = \frac{1}{2} \pm \sqrt{\frac{1}{4} + (\text{Re } \gamma^{\pm}(t) - \alpha)(\text{Re } \gamma^{\pm}(t) - \beta)},$$

2. if $\beta - \alpha \geq 1$ we set
   $$\text{Re } \gamma^{\pm}(t) = \frac{\alpha + \beta}{2} \pm \sqrt{\left(\frac{\beta - \alpha}{2}\right)^2 - t(1 - t)},$$
   $$\text{Im } \gamma^{\pm}(t) = t.$$
Lemma 3.2. If $\beta - \alpha < 1$, then
\[
0 \leq \text{Im } \gamma^-(t) \leq \frac{1}{2} - \sqrt{\frac{1 - (\beta - \alpha)^2}{4}} < \frac{1}{2} \quad \text{and}
\]
\[
1 \geq \text{Im } \gamma^+(t) \geq \frac{1}{2} + \sqrt{\frac{1 - (\beta - \alpha)^2}{4}} > \frac{1}{2} \quad \forall t \in [0, 1].
\]
If $\beta - \alpha \geq 1$, then
\[
\alpha \leq \text{Re } \gamma^-(t) \leq \frac{\alpha + \beta}{2} - \sqrt{\frac{(\beta - \alpha)^2 - 1}{2}} \leq \frac{\alpha + \beta}{2} \quad \text{and}
\]
\[
\beta \geq \text{Re } \gamma^+(t) \geq \frac{\alpha + \beta}{2} + \sqrt{\frac{(\beta - \alpha)^2 - 1}{2}} \geq \frac{\alpha + \beta}{2} \quad \forall t \in [0, 1].
\]

Definition 3.3. The set $\mathcal{U}_\Delta$ we define as:
(1) if $\beta - \alpha < 1$ set
\[
\mathcal{U}_\Delta := \{ z \in \mathbb{C} : \alpha < \text{Re } z < \beta, \text{Im } z < \text{Im } \gamma^-(\text{Re } z) \text{ or } \text{Im } z > \text{Im } \gamma^+(\text{Re } z), \text{ and } z \notin \sigma(A + i) \},
\]
(2) if $\beta - \alpha \geq 1$ set
\[
\mathcal{U}_\Delta := \{ z \in \mathbb{C} : 0 \leq \text{Im } z \leq 1, \text{Re } \gamma^- (\text{Im } z) < \text{Re } z < \text{Re } \gamma^+(\text{Im } z), \text{ and } z \notin \sigma(A + i) \}.
\]

Figure 1 shows $\gamma^\pm$ and the region $\mathcal{U}_\Delta$ for differing values of $\beta - \alpha$. We shall make extensive use of the following spectral enclosure result: for any $u \in \text{Dom}(A)$ with $\|u\| = 1$, $\langle Au, u \rangle = \eta$ and $\| (A - \eta)u \| = \zeta$, we have
\[
(3.2) \quad \alpha < \eta \Rightarrow \left( \alpha, \eta + \frac{\zeta^2}{\eta - \alpha} \right) \cap \sigma(A) \neq \emptyset;
\]
see [17, Lemma 1]. Property (3.2) also follows easily from [24, theorems 2.6 & 3.1].

Lemma 3.4. If $z \in \gamma^\pm$, $z \neq \alpha$ and $z \neq \alpha + i$, then
\[
\text{Re } z + \frac{\text{Im } z(1 - \text{Im } z)}{\text{Re } z - \alpha} = \beta.
\]
If $z \in \mathcal{U}_\Delta$, then
\[
\text{Re } z + \frac{\text{Im } z(1 - \text{Im } z)}{\text{Re } z - \alpha} < \beta.
\]

Proof. We assume that $\beta - \alpha \geq 1$, the case where $\beta - \alpha < 1$ being treated similarly. Let $x, y \in \mathbb{R}$ with $x > \alpha$ and $0 \leq y \leq 1$, then
\[
x + \frac{y(1 - y)}{x - \alpha} = \beta \quad \Leftrightarrow \quad x^2 - (\alpha + \beta)x + \alpha\beta + y(1 - y) = 0
\]
\[
\Leftrightarrow \quad x = \frac{\alpha + \beta}{2} \pm \sqrt{\left( \frac{\beta - \alpha}{2} \right)^2 - y(1 - y)}
\]
where the discriminant is non-negative for every $0 \leq y \leq 1$. The first assertion follows. Now,
\[
x + \frac{y(1 - y)}{x - \alpha} < \beta
\]
Figure 1. With $\alpha = 0$ and $\beta = 2$, 1.5, 1.25, 1.1, 1.01, 1, 0.99, 0.9 and 0.5 the curves $\gamma^\pm(\cdot)$ are shown. The shaded regions are $\mathcal{U}_\Delta$.

iff

$$\frac{\alpha + \beta}{2} - \sqrt{\left(\frac{\beta - \alpha}{2}\right)^2 - y(1 - y)} < x < \frac{\alpha + \beta}{2} + \sqrt{\left(\frac{\beta - \alpha}{2}\right)^2 - y(1 - y)}$$

from which the second assertion follows. $\square$

Lemma 3.5. If $z \in \mathcal{U}_\Delta$, then

$$\beta - \text{Re } z \geq \frac{\text{dist}(z, \gamma^\pm)^2}{\text{Re } z - \alpha} + \frac{\text{Im } z(1 - \text{Im } z)}{\text{Re } z - \alpha}.$$  

Proof. First we consider the case where $\beta - \alpha < 1$. Let $z \in \mathcal{U}_\Delta$, then by Lemma 3.2 we have $\text{Im } z \neq 1/2$. We assume that $\text{Im } z < 1/2$, the case where $\text{Im } z > 1/2$ can be treated similarly. For some $0 < \text{dist}(z, \gamma^\pm) \leq d$ and $t \in [0, 1]$, we have

$$\text{Re } z + (\text{Im } z + d)i = \gamma^-(t).$$
Then using Lemma 3.4,\[ \frac{\text{Re } z + (\text{Im } z + d)(1 - \text{Im } z - d)}{\text{Re } z - \alpha} = \beta. \]

From which we obtain
\[ \beta - \text{Re } z - \frac{\text{Im } z(1 - \text{Im } z)}{\text{Re } z - \alpha} = \frac{d(1 - 2\text{Im } z - d)}{\text{Re } z - \alpha} > \frac{d^2}{\text{Re } z - \alpha} \geq \frac{\text{dist}(z, \gamma^\pm)^2}{\text{Re } z - \alpha}. \]

Now we consider the case where $\beta - \alpha \geq 1$. Let $z \in \mathbb{U}_\Delta$ with $2\text{Re } z \leq \alpha + \beta$, the case where $2\text{Re } z > \alpha + \beta$ can be treated similarly. For some $0 < \text{dist}(z, \gamma^\pm) \leq d$ we have
\[ \text{Re } z - d + \frac{\text{Im } z(1 - \text{Im } z)}{\text{Re } z - d - \alpha} = \beta, \]
hence
\[ \beta - \text{Re } z - \frac{\text{Im } z(1 - \text{Im } z)}{\text{Re } z - \alpha} = d \left( \frac{d + \alpha + \beta - 2\text{Re } z}{\text{Re } z - \alpha} \right) > \frac{d^2}{\text{Re } z - \alpha} \geq \frac{\text{dist}(z, \gamma^\pm)^2}{\text{Re } z - \alpha}. \]

**Theorem 3.6.** If $\Delta \subset \rho(A)$, then $\mathbb{U}_\Delta \subset \rho(A + iP)$ and there exists a constant $K > 0$, dependent only on $\beta - \alpha$, such that
\[ \|(A + iP - z)^{-1}\| \leq \frac{K}{\text{dist}(z, \gamma^\pm)} \forall z \in \mathbb{U}_\Delta. \]

**Proof.** Let $z \in \mathbb{U}_\Delta$, $u \in \text{Dom}(A)$ with $\|u\| = 1$ and $\|(A + iP - z)u\| = \varepsilon$. Let us assume that
\[ \varepsilon < \min \left\{ 1, \frac{\text{dist}(z, \gamma^\pm)}{2} \right\}. \]
For some $v \in \mathcal{H}$ with $\|v\| = 1$ we have $(A + iP - z)v = \varepsilon v$. Then
\[ \langle (A - \text{Re } z)u, u \rangle + i\langle (P - \text{Im } z)u, u \rangle = \varepsilon \langle v, u \rangle, \]
hence
\[ \langle (A - \text{Re } z)u, u \rangle = \varepsilon \text{Re } \langle v, u \rangle \quad \text{and} \quad \langle (P - \text{Im } z)u, u \rangle = \varepsilon \text{Im } \langle v, u \rangle. \]
Furthermore,
\[ \|(A - \text{Re } z)u\| \leq \varepsilon + \|(P - \text{Im } z)u\| \]
where (recalling that $0 \leq P \leq I$)
\[ \|(P - \text{Im } z)u\|^2 = \|Pu\|^2 - 2\text{Im } z\langle Pu, u \rangle + (\text{Im } z)^2 \]
\[ \leq \langle Pu, u \rangle - 2\text{Im } z\langle Bu, u \rangle + (\text{Im } z)^2 \]
\[ \leq \text{Im } z + \varepsilon - 2\text{Im } z(\text{Im } z - \varepsilon) + (\text{Im } z)^2 \]
\[ = \text{Im } z(1 - \text{Im } z) + \varepsilon(1 + 2\text{Im } z) \]
\[ \leq \text{Im } z(1 - \text{Im } z) + 3\varepsilon. \]
Combining this estimate with (3.3), then using (3.2), (3.3) and (3.4), yields
\[ \left( \alpha, \text{Re } z + \varepsilon + \frac{\varepsilon + \sqrt{\text{Im } z(1 - \text{Im } z) + 3\varepsilon}}{\text{Re } z - \varepsilon - \alpha} \right) \cap \sigma(A) \neq \emptyset. \]
Furthmore, since \( \Delta \cap \sigma(A) = \emptyset \),

\[
\beta - \Re z \leq \varepsilon + \frac{(\varepsilon + \sqrt{\Im z(1 - \Im z) + 3\varepsilon})^2}{\Re z - \varepsilon - \alpha}
\]

\[
= \varepsilon + \frac{\varepsilon^2 + \Im z(1 - \Im z) + 3\varepsilon + 2\varepsilon \sqrt{\Im z(1 - \Im z) + 3\varepsilon}}{\Re z - \varepsilon - \alpha}
\]

\[
\leq \varepsilon + \frac{\varepsilon^2 + \Im z(1 - \Im z) + 2\varepsilon \sqrt{\Im z(1 - \Im z) + 2\varepsilon \sqrt{3\varepsilon}}}{\Re z - \varepsilon - \alpha}
\]

\[
\leq \varepsilon + \frac{\varepsilon + \Im z(1 - \Im z) + 3\varepsilon + \varepsilon + 2\varepsilon \sqrt{3}}{\Re z - \varepsilon - \alpha}
\]

\[
= \varepsilon + \frac{\Im z(1 - \Im z) + \varepsilon(5 + 2\sqrt{3})}{\Re z - \varepsilon - \alpha}.
\]

Combining this estimate with Lemma 3.3 gives

\[
\frac{\dist(z, \gamma^\pm)^2}{\Re z - \alpha} \leq \varepsilon + \frac{\Im z(1 - \Im z)}{\Re z - \varepsilon - \alpha} - \frac{\Im z(1 - \Im z)}{\Re z - \alpha} + \frac{\varepsilon(5 + 2\sqrt{3})}{\Re z - \varepsilon - \alpha}
\]

\[
= \varepsilon + \frac{\varepsilon \Im z(1 - \Im z)}{(\Re z - \varepsilon - \alpha)(\Re z - \alpha)} + \frac{\varepsilon(5 + 2\sqrt{3})}{\Re z - \varepsilon - \alpha}.
\]

It follows from (3.3) that

\[
\Re z - \varepsilon - \alpha \geq \dist(z, \gamma^\pm) - \varepsilon > \frac{\dist(z, \gamma^\pm)}{2},
\]

hence

\[
\frac{\dist(z, \gamma^\pm)^2}{\Re z - \alpha} \leq \left(1 + \frac{1}{\dist(z, \gamma^\pm)^2} + \frac{2(5 + 2\sqrt{3})}{\dist(z, \gamma^\pm)}\right) \varepsilon,
\]

and therefore

\[
\frac{\dist(z, \gamma^\pm)^4}{\beta - \alpha} < \frac{\dist(z, \gamma^\pm)^4}{\Re z - \alpha} \leq \left(\frac{(\beta - \alpha)^2}{4} + 1 + (5 + 2\sqrt{3})(\beta - \alpha)\right) \varepsilon.
\]

The result follows from this estimate combined with the assumption (3.3). □

The consequences of this theorem are quite surprising. It is obvious that any \( z \in \sigma(A + iP) \) with \( z \notin \sigma(A) \) has the property that \( 0 < \Im z \leq 1 \). However, if \( (\alpha, \beta) \subset \rho(A) \) and \( \alpha < \Re z < \beta \), then the distance that \( \Re z \) can be from \( \{\alpha, \beta\} \) decreases as \( \beta - \alpha \) increases. For example, this means that if \( \alpha \) is an eigenvalue of \( A \) which is moved by the perturbation \( iP \), then how far the real part can move to the right depends on the distance of \( \alpha \) to \( \sigma(A) \backslash \{-\infty, \alpha\} \) and the greater this distance the less the eigenvalue can move to the right. One explanation of this phenomenon would be \( z \in \sigma(A + iP) \Rightarrow \Re z \in \sigma(A) \), however, the following example demonstrates that this is false and furthermore that whenever \( \beta - \alpha > 1 \) we can have \( \gamma^\pm(1/2) \in \sigma(A + iP) \).

**Example 3.7.** Let \( \beta - \alpha > 1 \) and consider the matrices

\[
A = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} \quad \text{and} \quad P = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}.
\]
The matrix $P$ is an orthogonal projection and is therefore self-adjoint with $\|P\| = 1$. Direct calculation yields

$$
\sigma(A + iP) = \left\{ \frac{\alpha + \beta}{2} + \frac{i}{2} \pm \sqrt{\left(\frac{\beta - \alpha}{4}\right)^2 - \frac{1}{4}} \right\} = \left\{ \gamma^-(1/2), \gamma^+(1/2) \right\}.
$$

**Definition 3.8.** For $z \in U_\Delta$ we set

$$
d(z) = \min \left\{ \frac{\dist(z, \gamma^\pm)^4}{K}, \dist(z, \sigma(A) + i) \right\} \quad \text{with } K \text{ as in Theorem 3.6}
$$

For $\varepsilon > 0$ we set

$$
X_\varepsilon := \left\{ z \in U_\Delta : d(z) > 3\varepsilon \right\}.
$$

**Corollary 3.9.** Let $P$ be an orthogonal projection with $\|(I - P)E(\Delta)\| = \varepsilon$, then

$$
(3.6) \quad X_\varepsilon \subset \rho(A + iB) \quad \text{with} \quad \|(A + iP - z)^{-1}\| \leq \frac{1}{d(z) - 3\varepsilon} \quad \text{for all } z \in X_\varepsilon.
$$

**Proof.** For simplicity let us denote $E := E(\Delta)$. We readily deduce that

$$
(3.7) \quad \|(I - E)PE\| \leq \varepsilon \quad \text{and} \quad \|EP(I - E)\| \leq \varepsilon.
$$

Let $z \in X_\varepsilon$ and $u \in \Dom(A)$. Using (3.7) and the identity

$$
P = EPE + (I - E)PE + EP(I - E) + (I - E)P(I - E),
$$

we obtain

$$
\|(A + iP - z)u\| = \|(A - z)(I - E)u + (A - z)Eu + iPu\|
$$

$$
= \|(A - z)(I - E)u + (A - z)Eu
$$

$$
\quad + i(EP + (I - E)P \quad
$$

$$
\quad + EP(I - E) + (I - E)P(I - E))u\|
$$

$$
\geq \|(A - z)(I - E)u + i(I - E)P(I - E)u
$$

$$
\quad + (A - z)Eu + iEPEu\|
$$

$$
\quad - \|(I - E)PE + EP(I - E)\|
$$

$$
\geq \|(A - z)(I - E)u + i(I - E)P(I - E)u
$$

$$
\quad + (A - z)Eu + iEPEu\| - 2\varepsilon\|u\|.
$$

The vector $(A - z)Eu + iEPEu$ satisfies the estimate

$$
\|(A - z)Eu + iEPEu\| = \|(A - z + i)Eu + iE(P - I)Eu\|
$$

$$
\geq (\dist(z, \sigma(A) + i) - \varepsilon)\|Eu\|
$$

$$
\geq (d(z) - \varepsilon)\|Eu\|.
$$

Next consider the vector $(A - z)(I - E)u + i(I - E)P(I - E)u$. Evidently, we have

$$
(A - z)(I - E) + i(I - E)P(I - E) : \mathcal{H} \oplus \mathcal{L}(\Delta) \to \mathcal{H} \oplus \mathcal{L}(\Delta).
$$

The restriction of $A$ to $\mathcal{H} \oplus \mathcal{L}(\Delta)$ is a self-adjoint operator with no spectrum in the interval $\Delta$. The restriction of $(I - E)P$ to $\mathcal{H} \oplus \mathcal{L}(\Delta)$ is a self-adjoint operator with $0 \leq (I - E)P \leq 1$. Therefore, by Theorem 3.6

$$
\|(A - z)(I - E)u + i(I - E)P(I - E)u\| \geq \frac{\dist(z, \gamma^\pm)^4}{K}\|(I - E)u\|
$$

$$
\geq d(z)\|(I - E)u\|.
$$
Combining these three estimates yields the result.

Note that for any compact set $X \subset \mathcal{U}_\Delta$ with $X \cap \{\sigma(A) + i\} = \emptyset$ it follows, from Corollary $3.9$ that $X \subset \rho(A + iP)$ for all $P$ with $\|(I - P)E(\Delta)\|$ sufficiently small.

4. The Perturbation Method

The perturbation method aims to perturb eigenvalues from $\sigma_{\text{dis}}(A)$, off the real line, by adding a perturbation $iP$ where $P$ is a finite-rank orthogonal projection. By [28, Theorem 2.5], all non-real eigenvalues of $A + iP$ are approximated by the Galerkin method; there is no spectral pollution away from the real line. Furthermore, by [28, Theorem 2.9], the Galerkin method will also capture the multiplicity of the non-real eigenvalues.

We assume throughout this section that $P$ is a finite-rank orthogonal projection and that $\Delta \cap \sigma(A) = \{\lambda_1, \ldots, \lambda_d\} \subset \sigma_{\text{dis}}(A)$ where $d < \infty$. If $\|(I - P)E(\Delta)\|$ is sufficiently small, by Corollary $3.9$ it follows that $\sigma(A + iP) \cap \mathcal{U}_\Delta$ can consist only of eigenvalues in a small neighbourhood of the $\lambda_j + i$ or in a small neighbourhood of the curves $\gamma^\pm$. The non-real eigenvalues of $A + iP$ can then, of course, be approximated without incurring spectral pollution. Naturally, we are interested in approximate eigenvalues which are contained in $\mathcal{U}_\Delta$, have imaginary part close to one, and which are not close to $\gamma^\pm$. We fix an $r > 0$ and assume throughout that, for each $1 \leq j \leq d$, we have

$$0 < r < \max \left\{ \frac{\text{dist}(\lambda_j, \sigma(A) \backslash \{\lambda_j\})}{2}, \frac{1}{2} \right\}$$

and $\mathbb{D}(\lambda_j + i, r) \cap \gamma^\pm = \emptyset$

where $\mathbb{D}(\lambda_j + i, r)$ the closed disc with center $\lambda_j + i$ and radius $r$. Associated to the restriction of the form $a$ to the trial space $\mathcal{L}_n$ is a self-adjoint operator acting in the Hilbert space $\mathcal{L}_n$: denote this operator and the corresponding spectral measure by $A_n$ and $E_n$, respectively.

In Subsection 4.1, aided by Corollary $3.9$, we establish that $\mathbb{D}(\lambda_j + i, r)$ will intersect $\sigma(A + iP)$ whenever $\|(I - P)E(\Delta)\|$ is sufficiently small. Subsection 4.2 considers the convergence of elements from $\sigma(A + iP_n)$ to $\lambda_j + i$, Theorem 4.3 establishes a remarkably fast convergence rate. In Subsection 4.3, we prove, for a fixed $n \in \mathbb{N}$, the superconvergence of $\sigma(A + iP_n, \mathcal{L}_n) \cap \mathcal{U}_\Delta$ to a neighbourhood of $\lambda_j + i$.

4.1. Preliminary properties of $\sigma(A + iP)$

**Corollary 4.1.** There exist constants $c_r, \varepsilon_r > 0$, independent of $P$, such that whenever $\|(I - P)E(\Delta)\| \leq \varepsilon_r$ and $|\lambda_j + i - z| = r$ for some $1 \leq j \leq d$, we have

$$z \in \rho(A + iP) \quad \text{with} \quad \|(A + iP - z)^{-1}\| \leq \frac{1}{c_r}.$$  

**Proof.** We may choose any $\varepsilon_r > 0$ such that, for each $1 \leq j \leq d,$

$$d(z) - 3\varepsilon_r > 0 \quad \text{for all} \quad |\lambda_j + i - z| = r.$$  

It then follows, from Corollary $3.9$ that $z \in \rho(A + iP)$ and $\|(A + iP - z)^{-1}\|$ is uniformly bounded for all $z \in \mathcal{U}_\Delta$ with $|\lambda_j + i - z| = r$. It will therefore suffice to show that $A + iP - z$ is also invertible with uniformly bounded inverse whenever $|\lambda_j + i - z| = r$ and $\text{Im} z > 1$. That $z \in \rho(A + iP)$ follows immediately from $3.1$. Suppose that $(z_n^\pm)$ is sequence converging to $\lambda_j \pm r + i$ with $\text{Im} z_n^\pm > 1$ for every $n \in \mathbb{N}$. Let $(S_n)$ be a sequence of orthogonal projections with $\|(I - S_n)E(\Delta)\| \leq$
eigenspace for $A$ which is a contradiction since $\|A\|$ corresponding spectral subspace contains $I$. The left hand side is real from which it follows that $(u, v)$.

Let $\varepsilon_r > 0$ satisfying (4.1) now follows from the estimate

$$\|(A + iP - z)^{-1}\| \leq \frac{1}{\text{dist}(z, W(A + iP))} \quad \forall z \notin W(A + iP).$$

**Corollary 4.2.** Let $||(I - P)E(\Delta)|| = 0$, then $\mathcal{U}_\Delta \subset \rho(A + iP)$, and for each $1 \leq j \leq d$ we have $\lambda_j + i \in \sigma(A + iP)$ with spectral subspace $\mathcal{L}(\{\lambda_j\})$.

**Proof.** The first assertion follows immediately from Corollary 3.9. For any $u \in \mathcal{L}(\{\lambda_j\})$ we have $(A + iP)u = (\lambda_j + i)u$, so that $\lambda_j + i \in \sigma(A + iP)$ and the corresponding spectral subspace contains $\mathcal{L}(\{\lambda_j\})$. Now let $(A + iP)v = (\lambda_j + i)v$. Then $(A - \lambda_j)v = i(I - P)v$, and therefore

$$\langle (A - \lambda_j)v, v \rangle = i\|(I - P)v\|^2.$$

The left hand side is real from which it follows that $(I - P)v = 0$, therefore $(A - \lambda_j)v = 0$ and hence $v \in \mathcal{L}(\{\lambda_j\})$. We deduce that $\mathcal{L}(\{\lambda_j\})$ is the geometric eigenspace for $A + iP$ and the eigenvalue $\lambda_j + i$. Suppose that $\lambda_j + i$ is not semi-simple. Then there exists a non-zero vector $w \perp \mathcal{L}(\{\lambda_j\})$ with $(A + iP - \lambda_j - i)w = u \in \mathcal{L}(\{\lambda_j\})$. Hence,

$$(A - \lambda_j - i)w \perp \mathcal{L}(\{\lambda_j\}) \quad \text{with} \quad \|(A - \lambda_j - i)w\| > \|w\|,$$

and

$$iPw = u - (A - \lambda_j - i)w \quad \text{where} \quad u \perp (A - \lambda_j - i)w.$$

It follows that

$$\|Pw\|^2 = \|u\|^2 + \|(A - \lambda - i)w\|^2 > \|w\|^2,$$

which is a contradiction since $\|P\| = 1$. □

**Corollary 4.3.** Let $||(I - P)E(\Delta)|| \leq \varepsilon_r$, then

$$\mathcal{D}(\lambda_j + i, r) \cap \sigma(A + iP) \neq \emptyset \quad \text{for each} \quad 1 \leq j \leq d.$$

The dimension of the corresponding spectral subspace is equal $\dim(\mathcal{L}(\{\lambda_j\}))$ for each $1 \leq j \leq d$.

**Proof.** Let $u_1, \ldots, u_k$ be an orthonormal basis for $\mathcal{L}(\Delta)$. Set $v_j = Pu_j$ and let $v_{k+1}, \ldots, v_p$ be such that

$$\text{Range}(P) = \text{span}\{v_1, \ldots, v_p\}.$$

For $t \in [0, 1]$ set $w_j(t) = tu_j + (1 - t)v_j$ and let $P_t$ be the orthogonal projection onto $\text{span}\{w_1(t), \ldots, w_k(t), v_{k+1}, \ldots, v_p\}$. For any normalised $u \in \mathcal{L}(\Delta)$ we have $u = c_1u_1 + \cdots + c_ku_k$ and

$$||(I - P_t)u|| \leq \|c_1u_1 + \cdots + c_ku_k - c_1w_1(t) - \cdots - c_kw_k(t)|| = (1 - t)||u||.$$

Therefore $||(I - P_t)E(\Delta)|| \leq \varepsilon_r$ for all $t \in [0, 1]$ and $||(I - P_t)E(\Delta)|| = 0$. By Corollary 4.2, we have $\lambda_j + i \in \sigma(A + iP_1)$ with spectral subspace $\mathcal{L}(\{\lambda_j\})$. By
Corollary 4.1, the operator $A + iP_t - zI$ is invertible with uniformly bounded inverse for $|z - \lambda_j - i| = r$ and $t \in [0, 1]$. Hence, we may define the family of projections

$$Q_t := \int_{|\lambda_j + i - z| = r} (A + P_t - \zeta)^{-1} \, d\zeta.$$ 

Evidently, $Q(t)$ is a continuous family and therefore

$$\dim \left( \mathcal{L}(\{\lambda_j\}) \right) = \text{Rank}(Q_1) = \text{Rank}(Q_t) \quad \forall t \in [0, 1].$$

□

Example 4.4. With $\mathfrak{H} = \left[ L^2((0,1)) \right]^2$ we consider the following block-operator matrix

$$A_0 = \begin{pmatrix}
-d^2/dx^2 & -d/dx \\
-d/dx & 2I
\end{pmatrix}$$

with $\text{Dom}(A_0) = H^2((0,1)) \cap H^1_0((0,1))$. $A_0$ is essentially self-adjoint with closure $A$. We have $\sigma_{\text{ess}}(A) = \{1\}$ (see for example [29, Example 2.4.11]) while $\sigma_{\text{dis}}(A)$ consists of the simple eigenvalue 2 with eigenvector $(0, 1)^T$, and the two sequences of simple eigenvalues

$$\lambda_{\pm} := \frac{2 + k^2 \pi^2 \pm \sqrt{(k^2 \pi^2 + 2)^2 - 4k^2 \pi^2}}{2}.$$ 

The sequence $\lambda_{\pm}$ lies below, and accumulates at the essential spectrum. The sequence $\lambda_{\pm}$ lies above the eigenvalue 2 and accumulates at $\infty$.

Denote by $\mathcal{L}_h^0$ the FEM space of piecewise linear functions on $[0, 1]$ with a uniform mesh of size $h$ and which satisfy homogeneous Dirichlet boundary conditions. Denote by $\mathcal{L}_h$ the space without boundary conditions. Applying the Galerkin method directly to $A$, with the trial spaces $L_h = \mathcal{L}_h^0 \oplus \mathcal{L}_h$, we find that the only region which appears to incur spectral pollution is the interval $(1, 2) \subset \rho(A)$. Figure 2 shows the spurious Galerkin eigenvalues in the interval $(1, 2)$ for the trial space $L_{1/16}$. This situation appears to deteriorate further as $h$ is reduced. The approximation of the eigenvalue 2 is obscured.

![Figure 2. Spurious Galerkin eigenvalues in the interval (1, 2).](image)

Let $P_{1/16}$ be the orthogonal projection onto the trial space $L_{1/16}$. With $\Delta = (1, \lambda_1^+) = \sigma(A_{L_{1/16}})$ we note that $\Delta \cap \sigma(A)$ consists only of the simple eigenvalue 2 and that the
corresponding eigenvector $(0,1)^T \in L_h$ for all $h \in (0,1]$. From Corollary 4.2, we deduce that

$$U_\Delta \cap \sigma(A + iP_{1/16}) = \emptyset \quad \text{and} \quad 2 + i \in \sigma(A + iP_{1/16}).$$

Furthermore, by [28] Theorem 2.5, we can approximate the eigenvalue $2 + i \in \sigma(A + iP_{1/16})$ without incurring any spectral pollution in the region $U_\Delta \setminus \mathbb{R}$. Figures 3 & 4 show the Galerkin approximation of $\sigma(A + iP_{1/16})$ with the trial space $L_{1/1024}$. We see that $2 + i \in \sigma(A + iP_{1/16}, L_{1/1014})$ and the only elements from $\sigma(A + iP_{1/16}, L_{1/1014})$ contained in $U_\Delta$ are very close to the real line where spectral pollution is still permitted. Hence, the perturbation method has demonstrated that the Galerkin eigenvalues in the interval $(1,2)$ are all spurious. Furthermore, the genuine eigenvalue, $2$, is approximated by the perturbation method without being obscured by pollution.

**Figure 3.** Galerkin eigenvalues for the operator $A + iP_{1/16}$ and the curves $\gamma^\pm$.

**Figure 4.** The Galerkin eigenvalues near $1 + i$ all lie to the left of the curve $\gamma^-$ and are, therefore, not contained in $U_\Delta$. 
4.2. Superconvergence of $\sigma(A+iP_n)$. For some $N \in \mathbb{N}$, the hypothesis of Corollary [4.3] with $P = P_n$ (with $P_n$ as in Section 2), is satisfied for all $n \geq N$. For the remainder this section we assume that $n \geq N$. We denote by $\mathcal{M}_n(\{\lambda_j+i\})$ the spectral subspace corresponding to the eigenvalues of $A+iP_n$ which are enclosed by the circle $|\lambda_j+i-z|=r$. We also denote $\varepsilon_n = \delta(\mathcal{L}(\Delta),\mathcal{L}_n)$.

**Lemma 4.5.** $\hat{\delta}_n(\mathcal{L}(\{\lambda_j\}), \mathcal{M}_n(\{\lambda_j+i\})) = O(\varepsilon_n)$.

**Proof.** For simplicity, let us denote $E := E(\Delta)$. Note that

$$\sigma(A+iE) = (\sigma(A) \setminus \{\lambda_1+i, \ldots, \lambda_d+i\},$$

the spectral subspace associated to $\lambda_j+i$ is $\mathcal{L}(\{\lambda_j\})$, and for any $|\lambda_j+i-z|=r$

$$z \in \rho(A+iE) \text{ with } \|(A+iE-z)^{-1}\| = \frac{1}{r}.$$ 

By Corollary [4.1] we have

$$z \in \rho(A+iP_n) \text{ with } \|(A+iP_n-zI)^{-1}\| \leq \frac{1}{cr} \text{ for all } |\lambda_j+i-z|=r.$$ 

Let $u \in \mathcal{H}$ with $\|u\| = 1$ then, using the identity

$$(A+iP_n-z)^{-1} = (A+iE-z)^{-1} + (A+iE-z)^{-1}(iE-iP_n)(A+iP_n-z)^{-1}$$

and recalling that $m = \min \sigma(A)$, we obtain

$$\|(A-m+1)^{\frac{1}{2}}(A+iP_n-z)^{-1}u\|$$

$$\leq \|(A-m+1)^{\frac{1}{2}}(A+iE-z)^{-1}u\|$$

$$+ \|(A-m+1)^{\frac{1}{2}}(A+iE-z)^{-1}(iE-iP_n)(A+iP_n-z)^{-1}u\|$$

$$\leq \|(A-m+1)^{\frac{1}{2}}(A+iE-z)^{-1}\|$$

$$+ 2\|(A-m+1)^{\frac{1}{2}}(A+iE-z)^{-1}\|\|\|(A+iP_n-z)^{-1}\|\leq \max_{|\lambda_j+i-z|=r} \left\{ \frac{(2+c_r)^{\frac{1}{2}}(A-m+1)^{\frac{1}{2}}(A+iE-z)^{-1}}{c_r} \right\} =: M.$$ 

Now let $u \in \mathcal{L}(\{\lambda_j\})$ with $\|u\| = 1$ and let $|\lambda_j+i-z|=r$. The above estimate gives

$$\|(A+iE-z)^{-1}u - (A+iP_n-z)^{-1}u\|$$

$$= \|(A+iP_n-z)^{-1}(P_n-E)(A+iE-z)^{-1}u\|$$

$$= \frac{\|(A-m+1)^{\frac{1}{2}}(A+iP_n-z)^{-1}(P_n-I)u\|}{r}$$

$$\leq \frac{M\|(I-P_n)E\|}{r}$$

$$= \frac{M\delta(\mathcal{L}(\Delta),\mathcal{L}_n)}{r}.$$ 

(4.2) Set

$$u_n := -\frac{1}{2i\pi} \int_{|\lambda_j+i-z|=r} (A+iP_n-\zeta)^{-1}u \, d\zeta,$$
then \( u_n \in \mathcal{M}_n(\{\lambda_j + i\}) \). Using estimate (12),

\[
\begin{align*}
\left\| \frac{u}{\|u\|} - \frac{u_n}{\|u_n\|} \right\|_a &= \frac{1}{2\pi \|u\|_a} \int_{|\lambda_j + i| = r} |(A + iE - \zeta)^{-1}u - (A + iP_n - \zeta)^{-1}u| \, d\zeta \\
&\leq \frac{1}{2\pi} \int_{|\lambda_j + i| = r} \left\| (A + iE - \zeta)^{-1}u - (A + iP_n - \zeta)^{-1}u \right\|_a \, d\zeta \\
&= O(\delta(\mathcal{L}(\Delta), \mathcal{L}_n)).
\end{align*}
\]

We deduce that

\[
\delta_a(\mathcal{L}(\{\lambda_j\}), \mathcal{M}_n(\{\lambda_j + i\})) = O(\delta(\mathcal{L}(\Delta), \mathcal{L}_n)).
\]

Moreover, since

\[\dim \mathcal{M}_n(\{\lambda_j + i\}) = \dim \mathcal{L}(\{\lambda_j\}) < \infty,\]

the following formula holds

\[
\delta_a(\mathcal{M}_n(\{\lambda_j + i\}), \mathcal{L}(\{\lambda_j\})) \leq \frac{\delta_a(\mathcal{L}(\{\lambda_j\}), \mathcal{M}_n(\{\lambda_j + i\}))}{1 - \delta_a(\mathcal{L}(\{\lambda_j\}), \mathcal{M}_n(\{\lambda_j + i\}))}.
\]

\[\Box\]

**Theorem 4.6.** There exists a constant \( c_1 \geq 0 \) such that

\[
\sigma(A + iP_n) \cap \mathbb{D}(\lambda_j + i, c_1 \varepsilon_n^2) \neq \emptyset
\]

and the spectral subspace corresponding to the intersection is \( \mathcal{M}_n(\{\lambda_j + i\}) \).

**Proof.** Let \( \lambda_j \) have multiplicity \( \kappa \) and let \( \mu_{n,1}, \ldots, \mu_{n,\kappa} \) be the repeated eigenvalues of \( A + iP_n \) which are enclosed by the circle \(|\lambda + i - z| = r\). It will suffice to show

\[
(4.3) \quad \max_{1 \leq i \leq \kappa} |\mu_{n,i} - \lambda_j - i| = O(\varepsilon_n).
\]

Let \( u_1, \ldots, u_\kappa \) be an orthonormal basis for \( \mathcal{L}(\{\lambda_j\}) \). Let \( Q_n \) be the orthogonal projection from \( \mathcal{H}_a \) onto \( \mathcal{M}_n(\{\lambda_j + i\}) \) and set \( u_{n,l} = Q_n u_l \) for each \( 1 \leq l \leq \kappa \). By Lemma 4.5

\[
\|u_l - u_{n,l}\|_a = \| (I - Q_n) u_l \|_a = \text{dist}_a(u_l, \mathcal{M}_n(\{\lambda_j + i\})) = O(\varepsilon_n),
\]

and we may assume that \( Q_n \) maps \( \mathcal{L}(\{\lambda_j\}) \) one-to-one onto \( \mathcal{M}_n(\{\lambda_j + i\}) \). Consider the \( \kappa \times \kappa \) matrices

\[
[L_n]_{p,q} = (A + iP_n) u_{n,q}, u_{n,p} \quad \text{and} \quad [M_n]_{p,q} = (u_{n,q}, u_{n,p}).
\]

Evidently, \( M_n \) converges to the \( \kappa \times \kappa \) identity matrix and \( \sigma(L_n M_n^{-1}) \) is precisely the set \( \{\mu_{n,1}, \ldots, \mu_{n,\kappa}\} \). We have

\[
[L_n]_{p,q} = a(u_{n,q}, u_{n,p}) + i(P_n u_{n,q}, u_{n,p}).
\]

Consider the first term on the right hand side,

\[
a(u_{n,q}, u_{n,p}) = a((Q_n - I) u_q, u_p) + a((Q_n - I) u_q, (Q_n - I) u_p) + a(u_q, (Q_n - I) u_p) + a(u_q, u_p)
\]

\[
= \lambda_j (u_q, u_p) + a((Q_n - I) u_q, (Q_n - I) u_p) + \lambda_j (u_q, (Q_n - I) u_p) + \lambda_j \delta_{pq}
\]
where

\[
|\langle \lambda_j - m + 1 \rangle a, (Q_n - I)u_p \rangle| = |\langle a, (Q_n - I)u_p \rangle + (1 - m)\langle u_q, (Q_n - I)u_p \rangle|
\]

\[
= |\langle a, (Q_n - I)u_p \rangle_a|
\]

\[
= \left| \langle (Q_n - I)u_q, (Q_n - I)u_p \rangle \right|
\]

\[
\leq \| (Q_n - I)u_q \|_{a}\| (Q_n - I)u_p \|_{a},
\]

hence \(a(u_{n,q}, u_{n,p}) = \lambda_j \delta_{pq} + O(\varepsilon_n^2)\). Similarly,

\[
\langle P_n a_{n,q}, u_{n,p} \rangle = \langle (Q_n - I)u_q, (Q_n - I)u_p \rangle + \langle (Q_n - I)u_q, (Q_n - I)u_p \rangle
\]

\[
+ \langle (Q_n - I)u_q, (Q_n - I)u_p \rangle + \langle (P_n - I)u_q, (Q_n - I)u_p \rangle
\]

\[
+ \langle u_q, (Q_n - I)u_p \rangle + \langle (P_n - I)u_q, u_p \rangle + \delta_{pq}
\]

and

\[
|M_n|_{p,q} = \langle (Q_n - I)u_q, (Q_n - I)u_p \rangle + \langle (Q_n - I)u_q, (Q_n - I)u_p \rangle + \langle u_q, (Q_n - I)u_p \rangle + \delta_{pq}.
\]

Hence

\[
i\langle P_n a_{n,q}, u_{n,p} \rangle = i\delta_{pq} + O(\varepsilon_n^2) \quad \text{and} \quad |M_n|_{p,q} = \delta_{pq} + O(\varepsilon_n^2).
\]

Then

\[
|L_n|_{p,q} = (\lambda_j + i) \delta_{pq} + O(\varepsilon_n^2) \quad \text{and} \quad |M_n|^{-1} = \delta_{pq} + O(\varepsilon_n^2),
\]

and we deduce that \([L_n M_n^{-1}]_{p,q} = (\lambda_j + i) \delta_{pq} + O(\varepsilon_n^2)\). Property \((4.3)\) now follows from the Gersgorin circle theorem.

\[\square\]

**Example 4.7.** Let \(A\) and \(L_h\) be as in Example 4.4. We consider the spectrum of the operator \(A + P_h\) where \(P_h\) is the orthogonal projection onto \(L_h\). Applying the Galerkin method directly to \(A\) we do not incur spectral pollution near the eigenvalue \(\lambda_1^+\) and consequently we have the standard superconvergence result

\[(4.4)\]

\[
\text{dist}(\lambda_1^+, \sigma(A, L_h)) = O(h^2).
\]

By Theorem 4.6 we have the convergence rate

\[(4.5)\]

\[
\text{dist}(\lambda_1^+ + i, \sigma(A + iP_h)) = O(h^4).
\]

The second column in Table 1 shows the distance of \(\lambda_1^+\) to \(\sigma(A, L_h)\). The third column in Table 1 shows the distance of \(\lambda_1^+ + i\) to a Galerkin approximation (with very refined mesh) of the eigenvalue of \(A + iP_h\) which is close to \(\lambda_1^+ + i\). Figure 5, which displays a loglog plot of the data in Table 1, verifies \((4.4)\) and \((4.5)\).

| \(h\) | \(\text{dist}(\lambda_1^+, \sigma(A, L_h))\) | \(\text{dist}(\lambda_1^+ + i, \sigma(A + iP_h, L_{h \times 2 - r}))\) |
|---|---|---|
| 1/2 | 1.861045647858232 | 0.014440864705963 |
| 1/4 | 0.458746253205135 | 0.000609676693732 |
| 1/8 | 0.113442149493080 | 0.000034835584324 |
| 1/16 | 0.028273751580725 | 0.00002688221958 |

**Table 1.** Approximation of \(\lambda_1^+\) from \(\sigma(A, L_h)\) and from an approximation of \(\sigma(A + iP_h)\).
4.3. Superconvergence of $\sigma(A + iP_n, L_k)$. In this section, we are concerned with the Galerkin method approximation of certain eigenvalues of the operator $A + iP_n$. Unless stated otherwise we assume that $n \geq N$ is fixed. Then, by [28, Theorem 2.3], for all sufficiently large $k$,

$$z \in \rho(A + iP_n, L_k) \quad \text{for every} \quad |\lambda_j + i - z| = r.$$  

Furthermore, by [28, Theorem 2.5], the Galerkin method approximates all non-real eigenvalues of $A + iP_n$ without incurring spectral pollution and, by [28, Theorem 2.9], the Galerkin method will capture the multiplicity of the non-real eigenvalues.

Then, for all sufficiently large $k$,

$$\dim M_{n,k}(|\lambda_j + i|) = \dim M_n(|\lambda_j + i|) = \dim \mathcal{L}(|\lambda_j|).$$

**Lemma 4.8.** There exists a constant $c_2 > 0$, independent of $n \geq N$, such that

$$\max_{|\lambda_j + i - z| = r} \| (A_k + iP_kP_n - z)^{-1} \| \leq c_2 \quad \text{for all sufficiently large } k.$$  

**Proof.** We assume that the assertion is false. Then there exist sequences $(n_p)$ and $(\gamma_p)$ with $\gamma_p \to \infty$, such that, for each fixed $p$ there is a subsequence $k_q$ with

$$\max_{|\lambda_j + i - z| = r} \| (A_{k_q} + iP_{k_q}P_{n_p} - z)^{-1} \| > \gamma_p \quad \text{for all sufficiently large } q.$$  

Let us fix a $p$. We may assume without loss of generality that there exists a $|\lambda_j + i - z| = r$, such that

$$\| (A_k + iP_kP_{n_p} - z)^{-1} \| > \gamma_p \quad \text{for all sufficiently large } k.$$  

Hence, there exists a normalised sequence $u_k \in \mathcal{L}_k$ for which

$$\max_{v \in \mathcal{L}_k} \frac{|a(u_k, v) + i\langle P_{n_p}u_k, v \rangle - z(u_k, v)|}{\|v\|} < \frac{1}{\gamma_p}.$$
The sequence $P_{n_k} u_k$ has a convergent subsequence. We assume without loss of
generality that $iP_{n_k} u_k \to w$. Therefore

$$\max_{v \in \mathcal{H}_a} |a(u_k, v) + \langle w, v \rangle - z\langle u_k, v \rangle| < \frac{1}{\gamma_p} + \alpha_k \quad \text{for some} \quad 0 \leq \alpha_k \to 0.$$ 

Denote by $\hat{P}_k$ the orthogonal projection from $\mathcal{H}_a$ onto $\mathcal{L}_k$. Let $x = -(A - z)^{-1}w$
and set $x_k = \hat{P}_k x$, then for any $v \in \mathcal{L}_k$

$$a(x_k, v) - z\langle x_k, v \rangle = a(x, v) - z\langle x, v \rangle - a((I - \hat{P}_k)x, v) + z\langle (I - \hat{P}_k)x, v \rangle$$

$$= a(x, v) - z\langle x, v \rangle + (z - m + 1)\langle (I - \hat{P}_k)x, v \rangle$$

$$= -\langle w, v \rangle + (z - m + 1)\langle (I - \hat{P}_k)x, v \rangle.$$ 

We deduce that

$$\max_{v \in \mathcal{L}_k} |a(u_k - x_k, v) - z\langle u_k - x_k, v \rangle| < \frac{1}{\gamma_p} + \beta_k \quad \text{for some} \quad 0 \leq \beta_k \to 0,$$

hence

$$\|u_k - x_k\| < \left(\frac{1}{\gamma_p} + \beta_k\right)/\text{Im} \ z \leq \frac{1}{\gamma_p(1 - r)} + \frac{\beta_k}{(1 - r)}$$

and therefore

$$(4.7) \quad \|x\| - \|x_k\| > 1 - \frac{1}{\gamma_p(1 - r)} - \frac{\beta_k}{(1 - r)} \to 1 - \frac{1}{\gamma_p(1 - r)}.$$ 

Let $y = (A + iP_{n_k} - z)x = -w - iP_{n_k}(A - z)^{-1}w$. Since $iP_{n_k} u_k \to w$ implies that
$w \in \mathcal{L}_{n_k} \subset \mathcal{H}_a$, we deduce that $y \in \mathcal{H}_a$ and we set $y_k = \hat{P}_k y$. Using Corollary 4.1
and (4.7),

$$|a(x_k, y_k) + i\langle P_{n_k} x_k, y_k \rangle - z\langle x_k, y_k \rangle| \to \|(A + iP_{n_k} - z)x\|^2 \geq c_r^2 \left(\frac{1}{\gamma_p(1 - r)}\right)^2.$$ 

Furthermore, using the estimates above we have

$$|a(x_k, y_k) + i\langle P_{n_k} x_k, y_k \rangle - z\langle x_k, y_k \rangle|$$

$$= |a(x_k - u_k, y_k) + i\langle P_{n_k} (x_k - u_k), y_k \rangle - z\langle x_k - u_k, y_k \rangle|$$

$$+ |a(u_k, y_k) + i\langle P_{n_k} u_k, y_k \rangle - z\langle u_k, y_k \rangle|$$

$$\leq \|a(x_k - u_k, y_k) - z\langle x_k - u_k, y_k \rangle\| + \|\langle P_{n_k} (x_k - u_k), y_k \rangle\|$$

$$+ \|a(u_k, y_k) + i\langle P_{n_k} u_k, y_k \rangle - z\langle u_k, y_k \rangle\|$$

$$< \left(\frac{1}{\gamma_p} + \beta_k\right)\|y_k\|^2 + \left(\frac{1}{\gamma_p(1 - r)} + \frac{\beta_k}{(1 - r)}\right)\|y_k\|^2 + \frac{1}{\gamma_p}\|y_k\|^2.$$ 

Since $y = (A + iP_{n_k} - z)x = -w - iP_{n_k}(A - z)^{-1}w$ where $\|w\| \leq 1$,

$$\|y_k\| \to \|y\| = \|w - iP_{n_k}(A - z)^{-1}w\| \leq \|w\| + \|iP_{n_k}(A - z)^{-1}w\| \leq 1 + \frac{1}{1 - r},$$
Therefore, we have
\[ \left( \frac{1}{\gamma_p} + \beta_k \right) \| y_k \| + \left( \frac{1}{\gamma_p (1 - r)} + \frac{\beta_k}{1 - r} \right) \| y_k \| + \frac{1}{\gamma_p} \| y_k \| \]
\[ \to \left( \frac{2}{\gamma_p} + \frac{1}{\gamma_p (1 - r)} \right) \| y \| \]
\[ \leq \left( \frac{2}{\gamma_p} + \frac{1}{\gamma_p (1 - r)} \right) \left( 1 + \frac{1}{1 - r} \right). \]

Therefore, we have
\[ c_r^2 \left( 1 - \frac{1}{\gamma_p (1 - r)} \right)^2 \leq \| (A + iP_n - z)x \|^2 \]
\[ \leftarrow |a(x_k, y_k) + i(P_n x_k, y_k) - z(x_k, y_k)| \]
\[ \leq \left( \frac{1}{\gamma_p} + \beta_k \right) \| y_k \| + \left( \frac{1}{\gamma_p (1 - r)} + \frac{\beta_k}{1 - r} \right) \| y_k \| + \frac{1}{\gamma_p} \| y_k \| \]
\[ \to \left( \frac{2}{\gamma_p} + \frac{1}{\gamma_p (1 - r)} \right) \| y \| \]
\[ \leq \left( \frac{2}{\gamma_p} + \frac{1}{\gamma_p (1 - r)} \right) \left( 1 + \frac{1}{1 - r} \right). \]

Evidently, the left hand side is larger than the right hand side for all sufficiently large \( p \). The result follows from the contradiction. \( \square \)

**Lemma 4.9.** There exists a constant \( c_3 > 0 \), independent of \( n \geq N \), such that
\[ \delta_n(\mathbb{M}_n(\{\lambda_j + i\}), \mathbb{M}_{n,k}(\{\lambda_j + i\})) \leq c_3 \delta_n(\mathbb{M}_n(\{\lambda_j + i\}), \mathcal{L}_k) \]
for all sufficiently large \( k \).

**Proof.** Let \( u \in \mathbb{M}_n(\{\lambda_j + i\}) \) with \( \|u\| = 1 \). For \( |\lambda_j + i - z| = r \), we denote
\[ A_k(z) = A_k + iP_k P_n - z \quad \text{and} \quad x(z) = (A + iP_n - z)^{-1} u \in \mathbb{M}_n(\{\lambda_j + i\}). \]

Then, using Corollary 4.4, we have \( \| x(z) \| \leq c_r^{-1} \) and therefore
\[ \| x(z) \|_a = a(x(z)) - (m - 1) \| x(z) \|^2 \]
\[ = \langle Ax(z), x(z) \rangle - (m - 1) \| x(z) \|^2 \]
\[ = (A + iP_n - z)^{-1} u, x(z) - (m - 1) \| x(z) \|^2 \]
\[ = \langle u, x(z) \rangle - ((iP_n - z)x(z), x(z)) - (m - 1) \| x(z) \|^2 \]
\[ \leq \| x(z) \| + (2 + m + |z|) \| x(z) \|^2 \]
\[ \leq \frac{1}{c_r} + \frac{2 + m + |z|}{c_r^2}. \]

We deduce that
\[ \|(A + iP_n - z)^{-1} u\|_a = \| x(z) \|_a \leq K_1 \]
for some constant \( K_1 \geq 0 \) which is independent of \( n \geq N \) and \( |\lambda_j + i - z| = r \). Let \( v \in \mathcal{L}_k \) with \( \|v\| = 1 \), then
\[ \langle A_k(z) \hat{P} x(z) - u, v \rangle = a(\hat{P} x(z), v) + i\langle P_n \hat{P} x(z), v \rangle - z \langle \hat{P} x(z), v \rangle - \langle u, v \rangle \]
\[ = i\langle P_n (\hat{P} - I) x(z), v \rangle - (z - m + 1) \langle (\hat{P} - I) x(z), v \rangle. \]
Hence
\[ \|A_k(z)\tilde{P}_k x(z) - P_k u\| \leq (1 + |(z - m + 1)|)\|(\tilde{P}_k - I)x(z)\| \]
then, using Lemma 4.8,
\[ \|A_k(z)^{-1}P_k u - \tilde{P}_k x(z)\| \leq c_2\|A_k(z)\tilde{P}_k x(z) - P_k u\| \]
\[ \leq c_2(1 + |(z - m + 1)|)\|(\tilde{P}_k - I)x(z)\|. \]
Furthermore,
\[ \|A_k(z)^{-1}P_k u - x(z)\|_a \leq \|A_k(z)^{-1}P_k u - \tilde{P}_k x(z)\|_a + \|(\tilde{P}_k - I)x(z)\|_a \]
where
\[ \|A_k(z)^{-1}P_k u - \tilde{P}_k x(z)\|_a^2 \]
\[ = (a - m)\|A_k(z)^{-1}P_k u - \tilde{P}_k x(z)\| + \|A_k(z)^{-1}P_k u - \tilde{P}_k x(z)\|^2 \]
\[ = \langle P_k u - A_k(z)\tilde{P}_k x(z), A_k(z)^{-1}P_k u - \tilde{P}_k x(z) \rangle \]
\[ - \langle iP_k P_n - z(A_k(z)^{-1}P_k u - \tilde{P}_k x(z)), A_k(z)^{-1}P_k u - \tilde{P}_k x(z) \rangle \]
\[ + (1 - m)\|A_k(z)^{-1}P_k u - \tilde{P}_k x(z)\|^2 \]
\[ \leq \|P_k u - A_k(z)\tilde{P}_k x(z)\|\|A_k(z)^{-1}P_k u - \tilde{P}_k x(z)\| \]
\[ + |iP_k P_n - z|\|A_k(z)^{-1}P_k u - \tilde{P}_k x(z)\|^2 \]
\[ + |1 - m|\|A_k(z)^{-1}P_k u - \tilde{P}_k x(z)\|^2 \]
\[ \leq c_2(1 + |(z - m + 1)|)^2\|(\tilde{P}_k - I)x(z)\|^2 \]
\[ + (1 + |z|)c_2^2(1 + |(z - m + 1)|)^2\|(\tilde{P}_k - I)x(z)\|^2 \]
\[ + |1 - m|c_2(1 + |(z - m + 1)|)^2\|(\tilde{P}_k - I)x(z)\|^2. \]
Therefore,
\[ \|A_k(z)^{-1}P_k u - (A + iP_n - z)^{-1}u\|_a \leq K_2\|(\tilde{P}_k - I)x(z)\|_a \]
\[ \leq K_2\|x(z)\|_a\delta_a(M_n(\{\lambda_j + i\}), \mathcal{L}_k) \]
\[ \leq K_1K_2\delta_a(M_n(\{\lambda_j + i\}), \mathcal{L}_k) \]
for some constant \(K_2 \geq 0\) which is independent of \(n \geq N\) and \(|\lambda_j + i - z| = r\). Set
\[ u_k := -\frac{1}{2\pi} \int_{|\lambda_j + i - z| = r} A_k(\zeta)^{-1}P_k u \, d\zeta, \]
then \(u_k \in M_n,k(\{\lambda_j + i\})\) and
\[ \left\| \frac{u}{\|u\|_a} - \frac{u_k}{\|u\|_a} \right\|_a \leq \frac{1}{2\pi\|u\|_a} \left\| \int_{|\lambda_j + i - z| = r} A_k(\zeta)^{-1}P_k u - (A + iP_n - \zeta)^{-1}u \, d\zeta \right\|_a \]
\[ \leq \frac{1}{2\pi\|u\|_a} \int_{|\lambda_j + i - z| = r} \|A_k(\zeta)^{-1}P_k u - (A + iP_n - \zeta)^{-1}u\|_a \, d\zeta. \]
Combining this estimate with (4.9), we deduce that for some constant \(K_4 \geq 0\) which is independent of \(n \geq N\), we have
\[ \delta_a(M_n(\{\lambda_j + i\}), M_n,k(\{\lambda_j + i\})) \leq K_4\delta_a(M_n(\{\lambda_j + i\}), \mathcal{L}_k). \]
Then, by virtue of (4.6), the following formula holds for all sufficiently large $k$,
\[
\delta_a(M_{n,k}(\{\lambda_j + i\}), M_n(\{\lambda_j + i\})) \leq \frac{\delta_a(M_n(\{\lambda_j + i\}), M_{n,k}(\{\lambda_j + i\}))}{1 - \delta_a(M_n(\{\lambda_j + i\}), M_{n,k}(\{\lambda_j + i\}))}.
\]

Recall that $\varepsilon_n = \delta(\mathcal{L}(\Delta), \mathcal{L}_n(\Delta))$ and set $\varepsilon_{n,k} = \delta_a(M_n(\{\lambda_j + i\}), \mathcal{L}_k)$.

**Theorem 4.10.** There exist a constant $c_4 \geq 0$, independent of $n \geq N$, such that for all sufficiently large $k$,
\[
\sigma(A + iP_n, \mathcal{L}_k) \cap \mathbb{D}(\lambda_j + i, c_4(\varepsilon_n + \varepsilon_{n,k})^2) \neq \emptyset
\]
and spectral subspace corresponding to the intersection is $M_{n,k}(\{\lambda_j + i\})$.

**Proof.** Let $u_1, \ldots, u_\kappa$ be an orthonormal basis for $\mathcal{L}(\{\lambda_j\})$, and let $\mu_{n,k,1}, \ldots, \mu_{n,k,\kappa}$ be the repeated eigenvalues of $A_k + iP_kP_n$ which, for all sufficiently large $k$, are enclosed by the circle $|\lambda_j + i - z| = r$. It will suffice to show that
\[
\max_{1 \leq l \leq \kappa} |\mu_{n,k,l} - \lambda_j - i| \leq c_4(\varepsilon_n + \varepsilon_{n,k})^2.
\]
Let $R_k$ be the orthogonal projection from $\mathcal{H}_\kappa$ onto $M_{n,k}(\{\lambda_j + i\})$ and set $u_{j,k} = R_ku_j$. By Lemma 4.3 there exists a $c_0 > 0$ such that
\[
\hat{\delta}_a(\mathcal{L}(\{\lambda_j\}), M_n(\{\lambda_j + i\})) \leq c_0 \varepsilon_n.
\]
Using this estimate and Lemma 4.9
\[
\|u_j - u_{j,k}\|_a = \|(I - R_k)u_j\|_a = \text{dist}_a(u_j, M_{n,k}(\{\lambda_j + i\})) \\
\leq \|u_j\|_a \hat{\delta}_a(\mathcal{L}(\{\lambda_j\}), M_{n,k}(\{\lambda_j + i\})) \\
\leq \|u_j\|_a \left(\hat{\delta}_a(\mathcal{L}(\{\lambda_j\}), M_n(\{\lambda_j + i\})) + \hat{\delta}_a(M_n(\{\lambda_j + i\}), M_{n,k}(\{\lambda_j + i\}))\right) \\
\leq \|u_j\|_a (c_0 \varepsilon_n + c_3 \varepsilon_{n,k}) \\
\leq K_5(\varepsilon_n + \varepsilon_{n,k})
\]
where $K_5 = \max\{c_0, c_3\}\sqrt{b - m + 1}$ and is independent of $n \geq N$. Consider the matrices
\[
[L_{n,k}]_{p,q} = \langle (A + iP_n)u_{q,k}, u_{p,k}\rangle \quad \text{and} \quad [M_{n,k}]_{p,q} = \langle u_{q,k}, u_{p,k}\rangle.
\]
Evidently, $\sigma(L_{n,k}M_{n,k}^{-1})$ is precisely the set $\{\mu_{n,k,1}, \ldots, \mu_{n,k,\kappa}\}$. We have
\[
[L_{n,k}]_{p,q} = a(u_{q,k}, u_{p,k}) + i\langle P_nu_{q,k}, u_{p,k}\rangle.
\]
Consider the first term on the right hand side,
\[
a(u_{q,k}, u_{p,k}) = a(R_k - I)u_q, u_p + a(R_k - I)u_q, (R_k - I)u_p + a(u_q, (R_k - I)u_p) + a(u_q, u_p) \\
= \lambda_j \langle (R_k - I)u_q, u_p \rangle + a((R_k - I)u_q, (R_k - I)u_p) + \lambda_j \langle u_q, (R_k - I)u_p \rangle + \lambda_j \delta_{qp}
\]
hence

\[ |a(u_{q,k}, u_{p,k}) - \lambda_j \delta_{q,p}| \leq K_6(\varepsilon_n + \varepsilon_{n,k})^2, \]

for some \( K_6 \geq 0 \) which is independent of \( n \geq N \). Similarly,

\[
\langle P_n u_{q,k}, u_{p,k} \rangle = \langle P_n (R_k - I)u_q, (R_k - I)u_p \rangle + \langle (R_k - I)u_q, (P_n - I)u_p \rangle + (u_q, (R_k - I)u_p) + \langle (P_n - I)u_q, (R_k - I)u_p \rangle + (u_q, (R_k - I)u_p) + \langle (P_n - I)u_q, (R_k - I)u_p \rangle + (u_q, (R_k - I)u_p),
\]

hence

\[ |i\langle P_n u_{q,k}, u_{p,k} \rangle - i\delta_{q,p}| \leq K_7(\varepsilon_n + \varepsilon_{n,k})^2, \]

for some \( K_7 \geq 0 \) which is independent of \( n \geq N \). Furthermore,

\[ [M_{n,k}]_{p,q} = \langle (R_k - I)u_q, (R_k - I)u_p \rangle + \langle (R_k - I)u_q, (P_n - I)u_p \rangle + (u_q, (R_k - I)u_p), \]

for some \( K_8, K_9 \geq 0 \) both independent of \( n \geq N \), we have

\[ |[M_{n,k}]_{pq} - \delta_{pq}| \leq K_8(\varepsilon_n + \varepsilon_{n,k})^2 \quad \Rightarrow \quad |[M_{n,k}]^{-1}_{pq} - \delta_{pq}| \leq K_9(\varepsilon_n + \varepsilon_{n,k})^2. \]

Therefore,

\[ |[L_{n,k}^{-1}M_{n,k}]_{p,q} - (\lambda_j + i)\delta_{p,q}| \leq K_{10}(\varepsilon_n + \varepsilon_{n,k})^2 \]

for some \( K_{10} \geq 0 \) which is independent of \( n \geq N \). The result follows from the Gershgorin circle theorem. \( \square \)

**Example 4.11.** Let \( A \) and \( L_h \) be as in Example 4.3. We consider the spectrum of \( A + P_h \) where \( P_h \) is the orthogonal projection onto \( L_h \). We approximate \( \lambda_1^+ + i \) with \( \sigma(A + iP_{1/2}, L_h), \sigma(A + iP_{1/2}, L_h) \) and \( \sigma(A + iP_{1/2}, L_h) \) for decreasing values of \( h \). We also approximate \( \lambda_1^+ \) using \( \sigma(A, L_h) \) and recall that there is no spectral pollution obscuring the approximation of \( \lambda_1^+ \). The results are displayed in Figure 6; we see that the approximation and convergence achieved by the perturbation method are essentially the same as those achieved by the Galerkin method. It is clear, and consistent with Theorem 4.11 that we need not be overly concerned with locking-in poor accuracy with a relatively low dimensional projection \( P_h \). In fact, it is quite remarkable that the approximation with \( \sigma(A + iP_{1/2}, L_{1/32 \times 2^7}) \) is essentially the same as \( \sigma(A, L_{1/32 \times 2^7}) \).

5. **Further Examples**

**Example 5.1.** With \( \Omega = [L^2((0,1), \rho_0 dx)]^3 \) we consider the magneto-hydrodynamics operator

\[
A = \begin{pmatrix}
-\frac{d}{dx}(v_a^2 + v_b^2) + k^2 v_a^2 - i\left(\frac{d}{dx}(v_a^2 + v_b^2) - 1\right)k_\perp & -i\left(\frac{d}{dx}v_a^2 - 1\right)k_\parallel \\
-i k_\perp (v_a^2 + v_b^2) + k_\perp^2 v_a^2 + k_\perp k_\parallel v_a^2 & k_\perp k_\parallel v_a^2 \\
-i k_\parallel (v_a^2 + v_b^2) + k_\parallel^2 v_a^2 & k_\parallel^2 v_b^2
\end{pmatrix},
\]

With \( \rho_0 = k_\perp = k_\parallel = g = 1 \), \( v_a(x) = \sqrt{7/8 - x/2} \) and \( v_b(x) = \sqrt{1/8 + x/2} \), we have

\[ \sigma_{ess}(A) = [7/64, 1/4] \cup [3/8, 7/8]. \]
The discrete spectrum contains a sequence of simple eigenvalues which accumulate only at $\infty$. These eigenvalues are above, and not close to, the essential spectrum. They are approximated by the Galerkin method, with trial spaces $L_h = L^0_h \oplus L_h$, without incurring spectral pollution. It was shown, using the second order relative spectrum, that there is also an eigenvalue $\lambda_1 \approx 0.279$ in the gap in the essential spectrum; see [26, Example 2.7]. Figure 7 shows many Galerkin eigenvalues in the gap in the essential spectrum; we suspect that all but one of these is spectral pollution. Figure 9 shows many Galerkin eigenvalues lying just above the essential spectrum; again we suspect that most of these are spectral pollution. Indeed, Figure 9 shows a genuine approximation of $\sigma(A + iP_{1/64}L_{1/1024})$ off the real line. The two bands of essential spectrum are clearly approximated along with an approximate eigenvalue $\lambda_0$ in the gap, and an approximate eigenvalue $\lambda_2$ above the essential spectrum. The latter two eigenvalues are obscured by a direct application of the Galerkin method; see figure 7 & 8.

Example 5.2. With $\mathcal{H} = L^2(\mathbb{R})$ we consider the Schrödinger operator

$$Au = -u'' + \left( \cos x - e^{-x^2} \right) u.$$

The essential spectrum of $A$ has a band structure. The first three intervals of essential spectrum are approximately

$$[-0.37849, -0.34767], \quad [0.5948, 0.918058] \quad \text{and} \quad [1.29317, 2.28516].$$

The second order relative spectrum has been applied to this operator, see [6], where the following approximate eigenvalues were identified

$$\lambda_1 \approx -0.40961, \quad \lambda_2 \approx 0.37763, \quad \text{and} \quad \lambda_3 \approx 1.18216.$$
Figure 7. Spurious Galerkin eigenvalues in the gap in the essential spectrum which obscure the approximation of $\lambda_1$.

Figure 8. Spurious Galerkin eigenvalues above the essential spectrum which obscure the approximation of $\lambda_2$.

has identified the first two bands of essential spectrum plus the eigenvalues $\lambda_1$ below the essential spectrum, $\lambda_2$ in the first gap, and $\lambda_3$ in the second gap.

Example 5.3. With $\mathcal{H} = L^2((0, \infty))$ we consider the Schrödinger operator

$$Au = -u'' + \left(\sin x - \frac{40}{1 + x^2}\right)u, \quad u(0) = 0.$$ 

This example has been also been considered in [21]. The first three bands of essential spectrum are the same as in the previous example. However, this time there are infinitely many eigenvalues in the gaps which accumulate at the lower end point of the bands with their spacing becoming exponentially small; see [22]. We apply the perturbation method with the trial spaces $\mathcal{L}_{(X,Y)}$ which is a $Y$-dimensional space of piecewise linear trial functions on the interval $[0, X]$ which vanish at the boundary. The operator $P_{(X,Y)}$ is the orthogonal projection onto trial space $\mathcal{L}_{(X,Y)}$. Figure 11 shows that the perturbation method has approximated four eigenvalues in the first gap of the essential spectrum.
Our final example is outside much of the theory presented here, this is because the operator concerned is indefinite. However, the numerical results suggest that the perturbation method can be extended to the indefinite case. The second order relative spectra has been applied to this example and the code made available online; see [5] and [1], respectively. We use this code to apply the perturbation method and also to provide a direct comparison with the second order relative spectrum, the Davies & Plum method and the Zimmermann & Mertins method.

Example 5.4. With \( \mathcal{H} = [L^2((0, \infty))]^2 \) we consider the Dirac operator

\[
A = \begin{pmatrix}
I - \frac{1}{2x} & -\frac{d}{dx} - \frac{1}{x} \\
\frac{d}{dx} - \frac{1}{x} & -I - \frac{1}{2x}
\end{pmatrix}.
\]
Figure 11. Approximation of four eigenvalues in the interval \((b, c)\)
\((a = -0.37849, b = -0.34767, c = 0.5948)\), also shown is the curve
\(\gamma^+\) which corresponds to the interval \((b, c)\).

We have \(\sigma_{\text{ess}}(A) = (-\infty, -1] \cup [1, \infty)\) and the interval \((-1, 1)\) contains the eigenvalues
\[
\sigma_{\text{dis}}(A) = \left(1 + \frac{1}{4(j + \sqrt{3/4})^2}\right)^{-1/2}, \quad j = 0, 1, 2, \ldots.
\]

There is no spectral pollution incurred by the Galerkin method in this example, therefore we can also compare the perturbation method with the Galerkin method. We denote by \(\text{Spec}_2(A, L_n)\) the second order spectrum of \(A\) relative to trial space \(L_n\). We have a sequence \(z_n \in \text{Spec}_2(A, L_n)\) with \(z_n \rightarrow \lambda_1\). The sequence of real parts \((\text{Re} z_n)\) will converge to \(\lambda_1\) at least an order of magnitude faster; see (2.5).

The Davies & Plum and Zimmermann & Mertins methods are equivalent, given an interval containing only \(\lambda_1\) they both provide a tighter enclosure for \(\lambda_1\). For our approximation of \(\lambda_1\) we take the mid-point of this encloser which we denote by \(w_n\). Figure 12 shows a loglog plot of the dimension of the trial space \(n\) against the approximation of \(\lambda_1\) offered by each method. The results indicate that
\[
\text{dist}(\lambda_1 + i, \sigma(A + iP_n/2, P_n)) = O(n^{-0.9}), \quad \text{dist}(\lambda_1, \sigma(A, P_n)) = O(n^{-0.9}),
\]
\[
|\lambda_1 - z_n| = O(n^{-0.2}), \quad |\lambda_1 - \text{Re} z_n| = O(n^{-0.7}) \quad \text{and} \quad |\lambda_1 - w_n| = O(n^{-0.4}).
\]
Again we see the performance of the perturbation method is essentially the same as the Galerkin method. The quadratic methods converges slowly by comparison.

6. Conclusions and further research

For computing eigenvalues, which are not reliably located by a direct application of the Galerkin method, the perturbation method is extremely effective. The rapid convergence assured by theorems 4.6 & 4.10 (demonstrated in examples 4.7 & 4.11) mean that, in terms of accuracy and convergence, we can expect the perturbation method to significantly outperform the quadratic techniques. Unlike some quadratic techniques the perturbation method requires no \(\alpha\) priori information. Furthermore, it can be applied with trial spaces satisfying condition (2.1) rather than the awkward condition (2.4). Consequently, it is much easier to apply than quadratic techniques. In particular, we are able to apply the method, to second
order differential operators, using the FEM space of piecewise linear trial functions and this is a huge advantage.

In our examples we have chosen to perturb $A$ with projections from the same sequence with which we apply the Galerkin method. While this is convenient, it is not a requirement, and it will be interesting to apply the method by perturbing the operator with one sequence of projections whilst applying the Galerkin method with an unrelated sequence of trial spaces.

Our examples have suggested that as well as approximating the discrete spectrum, the perturbation method actually captures the whole spectrum, that is,

$$\left( \lim_{n \to \infty} \sigma(A + iP_n) \right) \cap \{ z \in \mathbb{C} : \text{Im } z = 1 \} = \sigma(A);$$

if so, then we have found a simple, easy to apply, pollution-free, superconvergent method for locating the whole spectrum.

7. ACKNOWLEDGEMENTS

The author is grateful to Marco Marletta for many useful discussions. This research was conducted with the support of the Wales Institute of Mathematical and Computational Sciences and the Leverhulme Trust grant: RPG-167.

REFERENCES

[1] T. Betcke, N. J. Higham, V. Mehrmann, C. Schröder, and F. Tisseur, NLEVP: A Collection of Nonlinear Eigenvalue Problems, MIMS EPrint 2011.116, December 2011.
[2] D. Boffi, F. Brezzi, L. Gastaldi, On the problem of spurious eigenvalues in the approximation of linear elliptic problems in mixed form. Math. Comp., 69 (229) (2000) 121–140.
[3] D. Boffi, R. G. Duran, L. Gastaldi, A remark on spurious eigenvalues in a square. Appl. Math. Lett., 12 (3) (1999) 107–114.
[4] L. Boulton, Limiting set of second order spectrum. Math. Comput., 75 (2006) 1367–1382.
[5] L. Boulton, N. Boussaïd, Non-variational computation of the eigenstates of Dirac operators with radially symmetric potentials. LMS J. Comput. Math. 13 (2010) 10–32.
[6] L. Boulton, M. Levitin, On Approximation of the Eigenvalues of Perturbed Periodic Schrödinger Operators. J. Phys. A: Math. Theor. 40 (2007) 9319–9329.
[7] L. Boulton, M. Strauss, On the convergence of second-order spectra and multiplicity. Proc. R. Soc. A 467 (2011) 264–275.
[8] L. Boulton, M. Strauss, Eigenvalues enclosures and convergence for the linearized MHD operator. Bit Numer. Math. 52 (2012) 801–825.
[9] F. Chatelin, Spectral Approximation of Linear Operators. Academic Press (1983).
[10] M. Dauge, M. Suri, Numerical approximation of the spectra of non-compact operators arising in buckling problems. J. Numer. Math. 10 (2002) 193–219.
[11] E. B. Davies, Spectral enclosures and complex resonances for general self-adjoint operators. LMS J. Comput. Math. 1 (1998) 42–74.
[12] E. B. Davies, Sectorial perturbations of self-adjoint matrices and operators. arXiv:1206.1703.
[13] E. B. Davies, M. Plum, Spectral Pollution. IMA J. Numer. Anal. 24 (2004) 417–438.
[14] A. C. Hansen, On the approximation of spectra of linear operators on Hilbert spaces. J. Funct. Anal. 254 (8) (2008) 2092–2126.
[15] A. C. Hansen, Infinite dimensional numerical linear algebra; theory and applications, Proc. R. Soc. Lond. Ser. A. 466 (2124) (2010) 3539–3559.
[16] A. C. Hansen, On the Solvability Complexity Index, the n-Pseudospectrum and Approximations of Spectra of Operators, J. Amer. Math. Soc. 24 (1) (2011) 81–124.
[17] T. Kato, On the upper and lower bounds of eigenvalues. J. Phys. Soc. Jpn. 4 (1949) 334–339.
[18] T. Kato, Perturbation Theory for Linear Operators. Springer-Verlag (1995).
[19] M. Levitin, E. Shargorodsky, Spectral pollution and second order relative spectra for self-adjoint operators. IMA J. Numer. Anal. 24 (2004) 393–416.
[20] M. Marletta, Neumann-Dirichlet maps and analysis of spectral pollution for non-self-adjoint elliptic PDEs with real essential spectrum. IMA J. Numer. Analysis 30 (2010) 917–939.
[21] M. Marletta, R. Scheichl, Eigenvalues in Spectral Gaps of Differential Operators. J. Spectral Theory 2 (3) (2012) 293–320.
[22] J. Rappaz, J. Sanchez Hubert, E. Sanchez Palencia & D. Vassiliev, On spectral pollution in the finite element approximation of thin elastic membrane shells. Numer. Math. 75 (1997) 473–500.
[23] K.M. Schmidt, Critical coupling constants and eigenvalue asymptotics of perturbed periodic Sturm-Liouville operators. Comm. Math. Phys. 211 (2000) 465–485.
[24] E. Shargorodsky, Geometry of higher order relative spectra and projection methods. J. Oper. Theory, 44 (2000) 43–62.
[25] E. Shargorodsky, On the limit behaviour of second order relative spectra of self-adjoint operators. J. Spectral Theory 3 (4) (2013) 535–552.
[26] M. Strauss, Quadratic Projection Methods for Approximating the Spectrum of Self-Adjoint Operators, IMA J. Numer. Anal. 31 (2011) 40–60.
[27] M. Strauss, The second order spectrum and optimal convergence. Math. Comp. 82 (2013) 2305–2325.
[28] M. Strauss, The Galerkin Method for Perturbed Self-Adjoint Operators and Applications. J. Spectral Theory (to appear).
[29] C. Tretter, Spectral Theory Of Block Operator Matrices And Applications. Imperial College Press (2007).
[30] S. Zimmermann, U. Mertins, Variational bounds to eigenvalues of self-adjoint eigenvalue problems with arbitrary spectrum. Z. Anal. Anwend. 14 (1995) 327–345.