On the Application of Schur Complements to the Spectral Analysis of Hermitian Matrices

Olof Rubin

Master's thesis
2019:E12

Lund University

Faculty of Science
Centre for Mathematical Sciences
Mathematics

CENTRUM SCIENTIARUM MATHEMATICARUM

Contents

Abstract

A second order expansion of the eigenvalues of a parameter perturbation of Hermitian matrices is derived using matrix decomposition methods. The method developed is then applied to give a novel proof of a second order approximation of the eigenvalues of a global perturbation of Hermitian matrices first described by M. Carlsson. The result generalizes an approximation given by G.W Stewart which treats the case of simple eigenvalues. This global second order approximation is then applied to the study of perturbations of singular values for matrices with non-empty kernel.

Acknowledgement

Marcus Carlsson is gratefully acknowledged for sharing his knowledge on matrix perturbations with me. I thank him for our interesting discussions and his help whenever I needed it.

I would also like to thank my high-school teacher Lars Blomgren for helping me develop my interest in mathematics by exposing me to the foundations of mathematical analysis.

1 Introduction

Mathematical spectral analysis is the study of eigenvalues and eigenvectors of operators acting on a linear space. In the finite-dimensional case this coincides with the spectral analysis of matrices which finds applications in virtually any branch of mathematical analysis and its applications.

Determination of eigenvalues is essential for the analysis of the stability of dynamical systems see Chapter 6 of [1]. Via this connection spectral theory finds applications in the theory of feedback systems which is used to control automative processes, see Chapter 4 and Chapter 6 of [2].

The Spectral Theorem is an important result in the spectral analysis of operators. It states that any normal operator has a complete set of eigenvectors [3], meaning that any vector can be written as a linear combination of eigenvectors. In this way the Spectral Theorem is applicable when solving certain differential equations such as the Sturm-Liouville Equations and the Schrodinger Equation see Chapter 5 of [1] and Chapter 3 of [4]. In numerical linear algebra singular values are used to determine properties of matrices such as condition numbers and rank [5].

In this thesis we discuss the spectral properties of perturbations of matrices. Given Hermitian matrices A and E we give approximate formulas for the eigenvalues and eigenvectors of $A + E$ in terms of those of A. We study the perturbation of singular values as an application, and thus remove the condition that the matrices be Hermitian. This approach is also of great interest in numerical applications, where a perturbation represents an uncertainty in measurement. Existence of well-behaved approximative formulas therefore implies stability of singular values, meaning that a small change in the perturbance accounts for a small change of the spectrum which is a favourable property.

1.1 The Perturbation Theory of Hermitian Matrices

We let A and E denote Hermitian operators acting on a finite-dimensional space and consider the perturbation $A + E$. We distinguish the perturbations into two distinct categories:

- The global perturbations of the form $A + E$ where no restriction is placed on E.
- The parameter perturbations of the form $A + E(t)$ where $E(t)$ is an Hermitian operator that depends analytically on the parameter $t \in \mathbb{R}$.

The spectrum of parameter perturbations exhibit more structure than the spectrum of global perturbations, due to its more restrictive form. A theorem by Franz Rellich provided in Chapter I of his book [6] states that the eigenvalues and eigenvectors of parameter perturbations (can be chosen to) depend analytically on the parameter t. This allows for an expansion of the eigenvalues and eigenvectors as a Taylor series in terms of the perturbation parameter t. This expansion is known up to certain finite orders, provided that some restrictions are placed on the perturbation, see Chapter VI of [4].

For global perturbations one can not hope to find an expansion in Taylor series, however, one can consider approximative results. A famous result by Hermann Weyl states that the eigenvalues of the Hermitian perturbation $A + E$ are Lipschitz continuous with respect to E

with Lipschitz constant 1. Some further such results are given by Lidskii's Theorems and Hoffman-Wielandt's Theorem, see Chapter III.4 and Chapter VI.4 of [7]. These are global results, meaning that there is no restriction on the norm of the perturbation, but on the other hand give first-order approximations with respect to the error term. To give higher-order approximations one can restrict the size of the error term to lie in a neighborhood of the origin. This was explored in an article by G.W Stewart, see Lemma 1 of [8], and Chapter V of [9], where a third order approximation is given in the case of simple eigenvalues and real matrices. A generalization of this result was provided by M. Carlsson, see Theorem 4.1 in [10]. These approximations are applicable to operations involving the square root of eigenvalues, such as singular values, or the absolute value of matrices. Based on this generalization of G.W Stewart's result M. Carlsson carries out a study of the matrix square root in [11].

1.2 Parameter Perturbations

David Hilbert and Richard Courant derived a formula for computing the Taylor coefficients for the eigenvalues and eigenvectors of a parameter perturbation under the assumption that these are analytic [12]. In the current study we explored the same problem for the perturbation $A + E(t)$ where $E(t) = tF$ for some fixed Hermitian matrix F. We took a direct approach via matrix decompositions which does not rely on the holomorphic properties of the eigenvalues and eigenvectors proven in [6] as opposed to Hilbert and Courant [12]. In particular we proved the following:

Theorem 1.1 (Hilbert-Courant). *Let* $\operatorname{diag}(\alpha_1, \ldots, \alpha_n) = \Lambda_A$ *be an* $n \times n$ *Hermitian matrix such that the diagonal is ordered non-increasingly. For Hermitian matrices* F *which satisfy*

$$F_{(i,j)} = 0, \quad i \neq j,\ \alpha_i = \alpha_j$$

and

$$F_{(i,i)} > F_{(j,j)}, \quad i \neq j,\ \alpha_i = \alpha_j$$

the following properties on the spectrum are valid:

(i) The eigenvalues of $\Lambda_A + tF$*, denoted* $\xi_1(t), \ldots, \xi_n(t)$ *can be ordered so that they satisfy*

$$\xi_j(t) = \alpha_j + tF_{(j,j)} + t^2 \sum_{i:\alpha_i \neq \alpha_j} \frac{|F_{(i,j)}|^2}{\alpha_i - \alpha_j} + \mathcal{O}(t^3).$$

(ii) Let $\{\boldsymbol{e}_j\}_{j=1}^n$ *denote the canonical basis for* $\mathbb{C}^n$*. There are choices of eigenvectors of* $\Lambda_A + tF$*, denoted* $\boldsymbol{u}_j(t)$*, which form an orthonormal basis of* $\mathbb{C}^n$ *such that*

$$\boldsymbol{u}_j(t) = \boldsymbol{e}_j + t\boldsymbol{v}_j + \mathcal{O}(t^2)$$

where

$$\boldsymbol{v}_j = \sum_{\substack{p:\alpha_p = \alpha_j \\ p \neq j}} \frac{(F^*(\Lambda_\alpha - \alpha_j I)^\dagger F)_{(p,j)}}{F_{(p,p)} - F_{(j,j)}} \boldsymbol{e}_p - \sum_{p:\alpha_p \neq \alpha_j} \frac{F_{(p,j)}}{\alpha_p - \alpha_j} \boldsymbol{e}_p.$$

Here $\dagger$ *denotes the Moore-Penrose inverse.*

Our approach was based on the application of Schur complements in conjuncture with Gershgorin's Circle Theorem. For completeness, we also provided the elegant argument supplied by David Hilbert and Richard Courant in [12], however, omitting Rellich's deep result concerning analyticity [6].

1.3 Global Perturbations and Singular Values

Having studied parameter perturbations, we then applied the method we developed in that context to prove a generalization of G.W. Stewarts approximation due to M. Carlsson (see [8] and [10]).

Theorem 1.2 (Carlsson)**.** *Let Λ_A be an $n \times n$ diagonal Hermitian matrix of the form*

$$\Lambda_A = \begin{pmatrix} \lambda I & 0 \\ 0 & \Lambda_\tau \end{pmatrix}$$

where λI is an $l \times l$ matrix and the diagonal of Λ_τ does not contain λ. For Hermitian $n \times n$ matrices E of the form

$$E = \begin{pmatrix} E_{11} & E_{12} \\ E_{12}^* & E_{22} \end{pmatrix}$$

the eigenvalues of $A + E$ denoted $\{\xi_j\}_{j=1}^n$ can be ordered so that

$$\xi_j = \lambda + \beta_j + \mathcal{O}(\|E\|^3), \quad j \in \{1, \ldots, l\}$$

where $\{\beta_j\}_{j=1}^l$ denotes an ordering of the eigenvalues of $E_{11} - E_{12}(\Lambda_\tau - \lambda I)^{-1}E_{12}^*$.

In the case of real Λ_A and E such that λ is a simple eigenvalue of Λ_A we see that the identification $\alpha = \lambda$, $\tilde{\alpha} = \lambda + E_{11}$, $\tilde{h} = E_{12} = h$, $A = \Lambda_\tau$ and $\tilde{A} = \Lambda_\tau + E_{22}$ shows that Lemma 1 in [8] implies Theorem 1.2 which motivates the identification as a generalization.

In the article [10] the author took advantage of a result by Henri Cartan, see [13], to estimate the roots of the characteristic polynomial of the perturbation. In this presentation we used a modification of our proof of Theorem 1.1 to give a new shorter proof of Theorem 1.2 based on Gershgorin's Circle Theorem.

Having obtained this result, we then considered its applications to singular value approximations in the case of non-empty kernels. A similar result is attained in [8] in the case of simple kernels which also allows for differentiability in the sense of Fréchet. Our result allowed for arbitrary dimensions of the kernels, however, some restrictions needed to be placed on the perturbation. For another approach of a first order expansion of the singular values in the case of simple singular values see [14].

2 Background

In this section we present the background in linear algebra which we will use in the spectral analysis of perturbations. We prefer to argue in the language of abstract operators acting on abstract finite-dimensional vector spaces whenever possible as opposed to matrices acting on $\mathbb{C}^n$. It is our belief that the reasoning is made clearer this way. Specific results which we cover are Gershgorin's Circle Theorem and one of its generalizations together with the Singular Value Decomposition.

2.1 Linear Operators on Finite-Dimensional Hilbert Spaces

Definition 2.1. *A Banach space is a complete normed-linear space. A Hilbert space is a complete normed-linear space whose norm is induced by an inner product.*

We reserve the notation $\mathscr{H}_n$ to denote an n-dimensional Hilbert space over $\mathbb{C}$. Given an orthonormal basis $\{\boldsymbol{e}_j\}_{j=1}^n$ of $\mathscr{H}_n$ there is a natural isometric isomorphism between $\mathscr{H}_n$ and $\mathbb{C}^n$ given by the map

$$\boldsymbol{e}_j \mapsto (\delta_{kj})_{k=1}^n = (0, \ldots, 0, 1, 0, \ldots, 0)$$

δ_{kj} denotes the Kronecker delta function so that the only non-zero entry is at the jth index and the value there is 1. Denoting the space of linear operators on $\mathscr{H}_n$ by $\mathcal{L}(\mathscr{H}_n)$ we find that for $\mathcal{A} \in \mathcal{L}(\mathscr{H}_n)$ there exists $\{A_{(i,j)}\}_{i,j}$ such that

$$\mathcal{A}\boldsymbol{e}_j = \sum_{i=1}^{n} A_{(i,j)}\boldsymbol{e}_i.$$

It is evident that linear operators in $\mathcal{L}(\mathscr{H}_n)$ naturally correspond to matrices of dimension $n \times n$ and that composition of linear operators in turn correspond with matrix multiplication. Due to this property the two concepts can be interchanged whenever suitable. We reserve the notation A to denote the matrix representation of the operator $\mathcal{A}$ with respect to some specified basis. An important function defined on the set $\mathcal{L}(\mathscr{H}_n)$ is the determinant, $\det : \mathcal{L}(\mathscr{H}_n) \to \mathbb{C}$ which contains basic information of the operator.

Definition 2.2. *The determinant is a multilinear function denoted* $\det : \mathcal{M}_{n\times n} \to \mathbb{C}$ *which satisfies*

$$\det A = \sum_{j=1}^{n!} \operatorname{sgn}(\pi_j) \prod_{i=1}^{n} A_{(i,\pi_j(i))}. \tag{1}$$

where $\{\pi_j\}_{j=1}^{n!}$ *is an enumeration of all permutations on* $\{1, \ldots, n\}$.

For specifics concerning the construction of the determinant using permutations we refer the reader to Chapter 5 of [15]. We will, however, require an alternative formula which requires us to define the concept of minors. For an $n \times n$-matrix A the (ij)th minor of A is the $(n-1) \times (n-1)$ matrix $A^{(i,j)}$ formed from A by removing its ith row and jth column. With this definition in hand we can consider the so-called Laplace expansion of the determinant which is given by the recursive formula:

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} A_{(i,j)} A^{(i,j)} = \sum_{j=1}^{n} (-1)^{i+j} A_{(i,j)} A^{(i,j)}. \tag{2}$$

That the values obtained from the calculation in Equation (1) and Equation (2) are equal is proven in [15].

A useful property of the determinant is that it is multiplicative in the sense that

$$\det(AB) = \det A \cdot \det B.$$

This implies that for an invertible matrix S one has the identity $\det(S^{-1}AS) = \det(S^{-1}S)\det A = \det A$. Since changes of basis of $\mathscr{H}_n$ corresponds to changing the matrix representation of an operator to $S^{-1}AS$ the determinant can be defined on $\mathcal{L}(\mathscr{H}_n)$ by $\det \mathcal{A} = \det A$ where A is any matrix representation of $\mathcal{A}$. Among other important properties, the determinant of an operator is 0 if and only if the operator is invertible which directly tie the determinant of an operator to the analysis of its spectrum. The spectrum of an operator $\mathcal{A}$ is defined as the set

$$\{\lambda \in \mathbb{C} \mid \mathcal{A} - \lambda I \text{ is not invertible}\}.$$

For operators acting on finite dimensional spaces the spectrum and the eigenvalues agree as a consequence of the Rank-Nullity Theorem [3]. In this case injectivity is equivalent to surjectivity. It is immediate that the spectrum of a matrix and its transpose are equal since $\det(A - \lambda I) = \det(A^t - \lambda I)$. Furthermore two matrices, A and B are said to be similar if there exists an invertible matrix S such that [16]

$$A = S^{-1}BS.$$

We denote this relation with $A \sim B$. It is clear that two similar matrices share eigenvalues and that their multiplicities are the same. This is something we will often apply as it allows us to rewrite a complex matrix expression A into a simpler one B without changing the spectrum. A natural way to define a norm on $\mathcal{L}(\mathscr{H}_n)$ is the operator norm, denoted $\|\cdot\|$ which is formally defined as

$$\|\mathcal{A}\| = \sup_{\boldsymbol{x} \neq 0} \frac{|\mathcal{A}\boldsymbol{x}|}{|\boldsymbol{x}|}.$$

An important property of this norm is that $\mathcal{L}(\mathscr{H}_n)$ equipped with it becomes a Banach space which allows us to apply topological arguments to the set of operators. Two notions of differentiation on Banach spaces are Gateaux differentiation and Fréchet differentiation. Let f be a function from the Banach space X into the Banach space Y. The function f is said to be Fréchet differentiable at the point x_0 if there exists a linear operator T such that

$$\frac{\|f(x_0 + h) - f(x_0) - Th\|_Y}{\|h\|_X} \to 0$$

as $h \to 0$. The subindices indicate which norm is under consideration. The Gateaux derivative is defined by letting $h = t\xi$ above where t is a scalar and $\xi \in X$ and then letting $t \to 0$.

We will need to give an estimation in norm of the inverse of perturbations of the from $A+E$ where A is invertible. Inspired by the identity

$$\frac{1}{1-z} = \sum_{k=0}^{\infty} z^k$$

for $|z| < 1$ one could apply the same reasoning to any Banach algebra by Cauchy's criterion for convergence. The following generalization was first proven by John Von-Neumann.

Theorem 2.3 (Von-Neumann Series). *Suppose that $\mathcal{A} \in \mathcal{L}(\mathscr{H}_n)$ satisfies $\|\mathcal{A}\| < 1$ then $I - \mathcal{A}$ is invertible with inverse*

$$(I - \mathcal{A})^{-1} = \sum_{k=0}^{\infty} \mathcal{A}^k.$$

Proof. If $N > M$ then

$$\left\| \sum_{k=0}^{N} \mathcal{A}^k - \sum_{k=0}^{M} \mathcal{A}^k \right\| \leq \sum_{k=M+1}^{N} \|\mathcal{A}\|^k \to 0$$

as $M, N \to \infty$. Since $\mathcal{L}(\mathscr{H}_n)$ is a Banach space it follows by the Cauchy criterion that the series converges. To prove that it converges to the inverse of $I - \mathcal{A}$ we consider the limit:

$$(I - \mathcal{A}) \sum_{k=0}^{N} \mathcal{A}^k = I - \mathcal{A}^{N+1} \to I$$

as $N \to \infty$. Here we observed that the sum was telescopic. □

Theorem 2.3 has many applications, one which we will use extensively is that if two operators are close in norm and if one of them is invertible then so is the other. More specifically we have:

Corollary 2.4. *The subset of $\mathcal{L}(\mathscr{H}_n)$ constituting all invertible operators is open.*

Proof. Note that if $\|\mathcal{A} - \mathcal{B}\| < \frac{1}{\|\mathcal{A}^{-1}\|}$ then $I - \mathcal{A}^{-1}(\mathcal{A} - \mathcal{B})$ is invertible, why

$$\mathcal{B} = \mathcal{A} - (\mathcal{A} - \mathcal{B}) = \mathcal{A}(I - \mathcal{A}^{-1}(\mathcal{A} - \mathcal{B}))$$

is invertible as a composition of invertible operators. □

Now if $\mathcal{A}_n$ is a sequence of invertible operators which converges to the invertible operator $\mathcal{A}$ is it then true that $\mathcal{A}_n^{-1} \to \mathcal{A}^{-1}$? The answer is yes based on the following reasoning:

Corollary 2.5. *Consider the operation $\mathcal{A} \mapsto \mathcal{A}^{-1}$ defined on the invertible elements of $\mathcal{L}(\mathscr{H}_n)$. This is a continuous map.*

Proof. Suppose that $\mathcal{B}_n \to \mathcal{A}$ as $n \to \infty$ where $\mathcal{B}_n$ is an invertible operator for every n. Then if $\|\mathcal{B}_n - \mathcal{A}\| < \frac{1}{\|\mathcal{A}^{-1}\|}$ we find that

$$\mathcal{B}_n^{-1} = \left(\sum_{k=0}^{\infty} \left(\mathcal{A}^{-1}(\mathcal{A} - \mathcal{B}_n) \right)^k \right) \mathcal{A}^{-1}$$

why

$$\left\| \mathcal{B}_n^{-1} - \mathcal{A}^{-1} \right\| \leq \sum_{k=1}^{\infty} \left\| \mathcal{A}^{-1}(\mathcal{A} - \mathcal{B}_n) \right\|^k = \frac{\|\mathcal{A}^{-1}(\mathcal{A} - \mathcal{B}_n)\|}{1 - \|\mathcal{A}^{-1}(\mathcal{A} - \mathcal{B}_n)\|} \to 0$$

as $n \to \infty$. □

This implies that if $\mathcal{E} \to 0$ then $(\mathcal{A} + \mathcal{E})^{-1} \to \mathcal{A}^{-1}$ in norm. Thus given $\epsilon > 0$ we can find $\delta > 0$ such that $\|\mathcal{E}\| < \delta \Rightarrow \|(\mathcal{A} + \mathcal{E})^{-1} - \mathcal{A}^{-1}\| < \epsilon$ and in particular $\|\mathcal{A}^{-1}\| - \epsilon \leq \|(\mathcal{A} + \mathcal{E})^{-1}\| \leq \|\mathcal{A}^{-1}\| + \epsilon$. We proceed by stating properties of adjoints of an operator and give some useful properties of the important subclass of Hermitian operators.

Given an operator $\mathcal{A} \in \mathcal{L}(\mathscr{H}_n)$ its adjoint denoted $\mathcal{A}^*$ is the unique operator which satisfies

$$\langle \mathcal{A}\boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{x}, \mathcal{A}^*\boldsymbol{y} \rangle, \quad \forall x, y \in \mathscr{H}_n.$$

Existence of such an operator follows from Riesz-Representation Theorem [3]. An operator, $\mathcal{A} \in \mathcal{L}(\mathscr{H}_n)$, is said to be Hermitian if $\mathcal{A}^* = \mathcal{A}$. Supposing that the matrix representation, A, of the operator $\mathcal{A}$ is given with respect to an orthonormal basis, then $\mathcal{A}$ is Hermitian if and only if A is equal to its conjugate transpose. This follows directly since

$$A_{(i,j)} = \langle \mathcal{A}\boldsymbol{e}_j, \boldsymbol{e}_i \rangle = \langle \boldsymbol{e}_j, \mathcal{A}^*\boldsymbol{e}_i \rangle = \overline{\langle \mathcal{A}^*\boldsymbol{e}_j, \boldsymbol{e}_i \rangle} = \overline{A_{(j,i)}}.$$

Matrices which are equal to their own conjugate transpose are therefore also called Hermitian. For these operators we have the following spectral theorem:

Theorem 2.6 (The Spectral Theorem)**.** *Let $\mathcal{A} \in \mathcal{L}(\mathscr{H}_n)$ denote a Hermitian operator then*

(i) The eigenvalues of $\mathcal{A}$ are real.

(ii) There exists an orthonormal set consisting of eigenvectors of $\mathcal{A}$ which constitute an orthonormal basis of $\mathscr{H}_n$.

For a proof of this we refer the reader to [3]. Since the eigenvalues of a Hermitian operator $\mathcal{A} \in \mathcal{L}(\mathscr{H}_n)$ are real this means that we can impose an ordering of them. In particular we can define the functions $\lambda_i(\mathcal{A})$ for $i \in \{1, \ldots, n\}$ which maps Hermitian operators to the ith eigenvalue of $\mathcal{A}$ ordered non-increasingly where each eigenvalue appears as many times as its algebraic multiplicity. Therefore

$$\lambda_1(\mathcal{A}) \geq \lambda_2(\mathcal{A}) \geq \cdots \geq \lambda_n(\mathcal{A}). \tag{3}$$

An important result in the perturbation theory of Hermitian operators is Weyl's Perturbation Theorem.

Theorem 2.7 (Weyl's Perturbation Theorem)**.** *Let $\mathcal{A}$ and $\mathcal{E}$ be Hermitian operators on $\mathscr{H}_n$ then*

$$|\lambda_j(\mathcal{A}+\mathcal{E})-\lambda_j(\mathcal{A})| \leq \|\mathcal{E}\| \, .$$

For a proof of this we refer to the reader to Chapter III of [7]. Note that this implies that each map λ_i is Lipschitz continuous for a fix $\mathcal{A}$ with respect to Hermitian operators $\mathcal{E}$. An example of Hermitian operators are the orthogonal projections. To define these let $U \subseteq \mathscr{H}_n$ be a linear subspace of $\mathscr{H}_n$. Then $\mathscr{H}_n = U \oplus U^{\perp}$, see [3], where $U^{\perp}$ is the orthogonal complement of U so that any vector in $\boldsymbol{x} \in \mathscr{H}_n$ can be uniquely decomposed as

$$\boldsymbol{x} = x + x^{\perp}$$

where $x \in U$ and $x^{\perp} \in U^{\perp}$.

Definition 2.8. *We say that the operator $\mathcal{P}_U \in \mathcal{L}(\mathscr{H}_n)$ is orthogonal projection onto U if*

$$\mathcal{P}_U(\boldsymbol{x}) = x$$

where $\boldsymbol{x} = x + x^{\perp}$ is the decomposition of $\boldsymbol{x}$ as a direct sum of terms in U and $U^{\perp}$.

From the definition it is immediate that $\mathcal{P}_U^2 = \mathcal{P}_U$. Furthermore note that if the matrix representation of $\mathcal{E}$ with respect to some orthonormal basis $\{\boldsymbol{e}_j\}_{j=1}^n$ is of the block-form

$$E = \begin{pmatrix} E_{11} & E_{12} \\ E_{12}^* & E_{22} \end{pmatrix}$$

where E_{11} is an $l \times l$ matrix then E_{11} is the matrix representation of $\mathcal{P}_{\mathcal{N}}\mathcal{E}\mathcal{P}_{\mathcal{N}}$ where $\mathcal{N} = \operatorname{span}\{\boldsymbol{e}_j\}_{j=1}^l$. Of course similar formulas can be given for the remaining blocks. For this reason it will turn out to be important that orthogonal projections are in fact Hermitian so that the same holds true for the operators $\mathcal{P}_{\mathcal{N}}\mathcal{E}\mathcal{P}_{\mathcal{N}}$.

Lemma 2.9. *An orthogonal projection in $\mathcal{L}(\mathscr{H}_n)$ is an Hermitian operator.*

Proof. Let U be a linear subspace of $\mathscr{H}_n$ and denote $\mathcal{P}_U$ with orthogonal projection onto U. Decompose $\boldsymbol{x}$ and $\boldsymbol{y}$ as an orthogonal sum:

$$\boldsymbol{x} = x + x^{\perp}, \quad \boldsymbol{y} = y + y^{\perp}$$

where $x, y \in U$ and $x^{\perp}, y^{\perp} \in U^{\perp}$. Then

$$\begin{aligned} \langle \mathcal{P}_U(x+x^{\perp}), y+y^{\perp}\rangle &= \langle x, y+y^{\perp}\rangle = \langle x, y\rangle \\ &= \langle x+x^{\perp}, y\rangle = \langle x+x^{\perp}, \mathcal{P}_U(y+y^{\perp})\rangle. \end{aligned}$$

□

In the second part of this thesis we will prove Theorem 1.2. We give the statement in operator form.

Let λ_0 denote a fix eigenvalue of $\mathcal{A}$ and let j_0 be the first occurrence of the eigenvalue in the collection $\{\lambda_j(\mathcal{A})\}_{j=1}^n$. If we further define

$$\mathcal{B} = \mathcal{P}_{\mathcal{N}}\mathcal{E}\mathcal{P}_{\mathcal{N}} - \mathcal{P}_{\mathcal{N}}\mathcal{E}\mathcal{P}_{\mathcal{N}^\perp}\left(\mathcal{P}_{\mathcal{N}^\perp}(\mathcal{A}-\alpha_0 I)^\dagger\mathcal{P}_{\mathcal{N}^\perp}\right)\mathcal{P}_{\mathcal{N}^\perp}\mathcal{E}\mathcal{P}_{\mathcal{N}}$$

then

$$|\lambda_j(\mathcal{A}+\mathcal{E}) - \lambda_j(\mathcal{A}) - \lambda_{j-j_0+1}(\mathcal{B})| \leq c\,\|\mathcal{E}\|^3\,, \quad j \in \{j_0,\dots,j_0+l-1\}$$

for some $c>0$ as $\mathcal{E}\to 0$. Here $\dagger$ indicates the Moore-Penrose inverse.

2.2 The Moore-Penrose Inverse

While our use of the Moore-Penrose inverse is limited we give its precise definition in order to make the presentation complete.

Definition 2.10. *Suppose that $\mathcal{A}$ is a linear operator from $\mathscr{H}_n$ to $\mathscr{H}_m$. An operator $\mathcal{A}^\dagger$ from $\mathscr{H}_m$ to $\mathscr{H}_n$ is a Moore-Penrose inverse of $\mathcal{A}$ if*

(i) $\mathcal{A}\mathcal{A}^\dagger\mathcal{A} = \mathcal{A}$

(ii) $\mathcal{A}^\dagger\mathcal{A}\mathcal{A}^\dagger = \mathcal{A}^\dagger$

(iii) Both $\mathcal{A}\mathcal{A}^\dagger$ and $\mathcal{A}^\dagger\mathcal{A}$ are Hermitian.

The possibility of existence and uniqueness of such an operator is discussed in [16]. We will however only need to consider the special case where $\mathcal{A}$ is an Hermitian operator in $\mathcal{L}(\mathscr{H}_n)$ in this case the Moore-Penrose inverse is unique and has a simple form.

Lemma 2.11. *Suppose that $\mathcal{A}\in\mathcal{L}(\mathscr{H}_n)$ is an Hermitian operator where $\{\boldsymbol{e}_j\}_{j=1}^n$ denotes an orthonormal basis of eigenvectors of $\mathcal{A}$ with corresponding eigenvalues $\lambda_j(\mathcal{A})$. Then $\mathcal{A}^\dagger\in\mathcal{L}(\mathscr{H}_n)$ is given by*

$$\mathcal{A}^\dagger\boldsymbol{e}_j = \begin{cases} 0 & \lambda_j(\mathcal{A}) = 0 \\ \frac{1}{\lambda_j(\mathcal{A})}\boldsymbol{e}_j & \textit{otherwise.}\end{cases}$$

Remark. *We only show that $\mathcal{A}^\dagger$ defined this way is a Moore-Penrose inverse, not that it is unique. Uniqueness is proven in [16].*

Proof. We verify the four conditions given in Definition 2.10. For that purpose it is enough to consider how the operators in (i) and (ii) act on the vectors $\boldsymbol{e}_j$. Suppose that $\boldsymbol{e}_j\notin\ker\mathcal{A}$ then for the operator given in (i) we find that

$$\mathcal{A}\mathcal{A}^\dagger\mathcal{A}\boldsymbol{e}_j = \mathcal{A}\mathcal{A}^\dagger\lambda_j(\mathcal{A})\boldsymbol{e}_j = \mathcal{A}\frac{\lambda_j(\mathcal{A})}{\lambda_j(\mathcal{A})}\boldsymbol{e}_j = \mathcal{A}\boldsymbol{e}_j.$$

For the operator given in (ii) we have that

$$\mathcal{A}^\dagger \mathcal{A}\mathcal{A}^\dagger \boldsymbol{e}_j = \mathcal{A}^\dagger \mathcal{A} \frac{1}{\lambda_j(\mathcal{A})} \boldsymbol{e}_j = \mathcal{A}^\dagger \boldsymbol{e}_j$$

and therefore (i) and (ii) holds. Since $\mathcal{A}\mathcal{A}^\dagger = \mathcal{P}_{\ker \mathcal{A}^\perp} = \mathcal{A}^\dagger \mathcal{A}$ we find that both operators are Hermitian. □

We remark that if Λ_α is a diagonal real matrix of the form

$$\Lambda_{\mathcal{A}} = \begin{pmatrix} 0 & 0 \\ 0 & \Lambda_\tau \end{pmatrix}$$

where Λ_τ is invertible then

$$\Lambda_{\mathcal{A}}^\dagger = \begin{pmatrix} 0 & 0 \\ 0 & \Lambda_\tau^{-1} \end{pmatrix}.$$

2.3 Schur Complement

The Schur complement, while simple, is a useful concept in matrix theory which got its name from its use in a research article by Issai Schur (1875-1941) published in 1917 [17]. We will use it to decompose matrix perturbations so that Gershgorin's Circle Theorem can be applied and as such it is a key tool in our further development.

Definition 2.12 (Schur Complement). *Given a matrix $A \in \mathcal{M}_{n\times n}$ in the block form*

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \tag{4}$$

where A_{11} and A_{22} are square matrices and where A_{22} is invertible, the Schur complement [17] of the block A_{22} is the matrix

$$A/A_{22} = A_{11} - A_{12}A_{22}^{-1}A_{21}.$$

This definition has many interesting applications to matrix theory which are covered in [17]. A useful tool in computing determinants is contained in Lemma 2.13.

Lemma 2.13 (Schur Determinant Formula). *Given a matrix A of the form in equation* (4) *the determinant of A is given by*

$$\det A = \det A_{22} \det A/A_{22}.$$

Proof. By composing with two invertible matrices we find that

$$\begin{pmatrix} I & -A_{12}A_{22}^{-1} \\ 0 & I \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} I & 0 \\ -A_{22}^{-1}A_{21} & I \end{pmatrix} = \begin{pmatrix} I & -A_{12}A_{22}^{-1} \\ 0 & I \end{pmatrix} \begin{pmatrix} A/A_{22} & A_{12} \\ 0 & A_{22}. \end{pmatrix}$$
$$= \begin{pmatrix} A/A_{22} & 0 \\ 0 & A_{22} \end{pmatrix}$$

Since the determinant of a triangular matrix is equal to the product of its diagonal elements and since the determinant is multiplicative we can conclude that

$$\det A = \det \begin{pmatrix} A/A_{22} & 0 \\ 0 & A_{22} \end{pmatrix} = \det A_{22} \det A/A_{22}.$$

□

Another application of Schur complements is that one can transform block-matrices without changing the eigenvalues. It is this crucial property that we will apply in order to derive our main results.

Lemma 2.14. *The matrix*

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

where A_{11} *and* A_{22} *are square matrices and* A_{22} *is invertible is similar to the matrix*

$$\begin{pmatrix} A_{11} - A_{12}A_{22}^{-1}A_{21} & A_{12} \\ A_{22}^{-1}A_{21}(A_{11} - A_{12}A_{22}^{-1}A_{21}) & A_{22}^{-1}A_{21}A_{12} + A_{22} \end{pmatrix} = \begin{pmatrix} A/A_{22} & A_{12} \\ A_{22}^{-1}A_{21}(A/A_{22}) & A_{22}^{-1}A_{21}A_{12} + A_{22} \end{pmatrix}.$$

Proof. The matrix

$$\begin{pmatrix} I & 0 \\ -A_{22}^{-1}A_{21} & I \end{pmatrix}$$

is invertible with inverse

$$\begin{pmatrix} I & 0 \\ A_{22}^{-1}A_{21} & I \end{pmatrix}$$

and since

$$\begin{aligned} &\begin{pmatrix} I & 0 \\ A_{22}^{-1}A_{21} & I \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} I & 0 \\ -A_{22}^{-1}A_{21} & I \end{pmatrix} \\ &= \begin{pmatrix} A_{11} - A_{12}A_{22}^{-1}A_{21} & A_{12} \\ A_{22}^{-1}A_{21}(A_{11} - A_{12}A_{22}^{-1}A_{21}) & A_{22}^{-1}A_{21}A_{12} + A_{22} \end{pmatrix} \end{aligned}$$

the result now follows. □

2.4 Gershgorin's Circle Theorem

A useful result in estimating the eigenvalues of a matrix is due to the Soviet mathematician Semyon Aranovich Gershgorin (1901-1933) and is the content of Theorem 2.15. It roughly translates to the fact that if the diagonal of a matrix dominates its off-diagonal entries, then the eigenvalues will be close to the diagonal entries. We will have particular use of this in the conjuncture with Schur Complements.

Theorem 2.15 (Gershgorin's Circle Theorem)**.** *Let* $A \in \mathcal{M}_{n\times n}$ *and define* $D_i \subseteq \mathbb{C}$ *for* $i \in \{1,\dots,n\}$ *to be the set which consist of all* $z \in \mathbb{C}$ *which satisfy*

$$|z - A_{(i,i)}| \leq \sum_{\substack{j=1 \\ j\neq i}}^{n} |A_{(i,j)}|. \tag{5}$$

Then every eigenvalue of A *belongs to some* D_i*.*

Proof. Let $\boldsymbol{x}'$ be any eigenvector of A and define

$$\boldsymbol{x} = \frac{\boldsymbol{x}'}{\max_{1\leq i\leq n} |x'_i|}$$

so that there exists an index $i \in \{1,\dots,n\}$ which satisfies $|x_j| \leq |x_i| = 1$ for every $j \in \{1,\dots,n\}$. Since $\boldsymbol{x}$ is an eigenvector there exists a $\lambda \in \mathbb{C}$ which corresponds to the eigenvector $\boldsymbol{x}$, and therefore the column vector $A\boldsymbol{x}$ satisfies

$$(A\boldsymbol{x})_{(i,1)} = \lambda x_i \Leftrightarrow \sum_{j=1}^{n} A_{(i,j)} x_j = \lambda x_i \Leftrightarrow \sum_{\substack{j=1 \\ j\neq i}}^{n} A_{(i,j)} x_j = (\lambda - A_{(i,i)}) x_i.$$

Since $|x_j| \leq |x_i| = 1$ we see that

$$|\lambda - A_{(i,i)}| = \left| \sum_{\substack{j=1 \\ j\neq i}}^{n} A_{(i,j)} x_j \right| \leq \sum_{\substack{j=1 \\ j\neq i}}^{n} |A_{(i,j)}|$$

and therefore $\lambda \in D_i$. □

Corollary 2.16. *Let* $A \in \mathcal{M}_{n\times n}$ *and define* $C_i \subseteq \mathbb{C}$ *for* $i \in \{1,\dots,n\}$ *to be the set which consists of all* $z \in \mathbb{C}$ *which satisfy*

$$|z - A_{(i,i)}| \leq \sum_{\substack{j=1 \\ j\neq i}}^{n} |A_{(j,i)}|.$$

Then every eigenvalue of A *belongs to some* C_i*.*

Proof. Since the transpose of A shares eigenvalues with A the result follows from Theorem 2.15. □

The disks D_i and C_i that we defined in Theorem 2.15 and Corollary 2.16 will be referred to as Gershgorin disks. It is a consequence of Theorem 2.15 and Corollary 2.16 that the eigenvalues are in fact contained in the intersection of these disks.

Theorem 2.15 does not imply that every Gershgorin disk contains an eigenvalue. In fact this is not true in general. It is for this reason that we will need to consider a strengthening of this theorem referred to as Gershgorin's Second Circle Theorem. This result relies on continuity of roots of polynomials.

2.5 Continuity of Roots of Polynomials

The eigenvalues of a matrix $A \in \mathcal{M}_{n\times n}$ are the roots of the polynomial equation in ζ defined by

$$\det(A - \zeta I) = 0.$$

In this section we will show that if $\{A_k\}_{k=1}^{\infty}$ denotes a sequence of matrices in $\mathcal{M}_{n\times n}$ which converge to A then the eigenvalues of A_k converge to those of A.

The following theorem is stated without proof in Appendix D of [18].

Theorem 2.17. *Let $p : \mathbb{C} \to \mathbb{C}$ be a polynomial given by*

$$p(\zeta) = \sum_{k=0}^{n} a_k \zeta^k = a_n \prod_{i=1}^{n} (\zeta - \lambda_i)$$

where $a_n \neq 0$. Given $\epsilon > 0$ there exists a $\delta > 0$ such that if $q : \mathbb{C} \to \mathbb{C}$ defined by

$$q(\zeta) = \sum_{k=0}^{n} b_k \zeta^k = b_n \prod_{i=1}^{n} (\zeta - \mu_i), \quad b_n \neq 0$$

satisfies

$$|a_i - b_i| < \delta, \quad i \in \{1, \ldots, n\}$$

then there exists a permutation π of $\{1, \ldots, n\}$ such that

$$|\mu_{\pi(i)} - \lambda_i| < \epsilon, \quad i \in \{1, \ldots, n\}.$$

To prove Theorem 2.17 we need Rouché's Theorem from complex analysis, a proof of this can be found in Chapter 3 of [19]. We state the formulation of the theorem given there.

Theorem 2.18 (Rouché's Theorem)**.** *Let f and g denote analytic functions defined on the open set Ω. If $D \subseteq \Omega$ is a closed disk and if*

$$|f(\zeta)| > |g(\zeta)|, \quad \zeta \in \partial D$$

where ∂D denotes the boundary of the disk, then f and $f + g$ have the same amount of zeros in D counting multiplicity.

Proof of theorem 2.17. Let $\tau : \{1, \ldots, m\} \to \{1, \ldots, n\}$, $1 \leq m \leq n$ be an injective function such that $\lambda_{\tau(1)}, \ldots, \lambda_{\tau(m)}$ is an enumeration of the distinct zeros of p. Choose $\epsilon > 0$ such that the sets

$$D_i = \{\zeta \in \mathbb{C} \mid |\zeta - \lambda_{\tau(i)}| \leq \epsilon\}$$

are disjoint. Then define

$$m_i = \min_{\zeta \in \partial D_i} |p(\zeta)|.$$

Since p is continuous and attains its zero in the interior of D_i and since ∂D_i is compact it follows that $m_i > 0$. Now define

$$M_i = \max_{\zeta \in \partial D_i} \sum_{k=0}^{n} |z|^k$$

and then choose $\delta > 0$ such that

$$\delta \cdot \left(\max_{i \in \{1,\dots,n\}} M_i \right) < \min_{i \in \{1,\dots,n\}} m_i.$$

If $\zeta \in \partial D_j$ then

$$\begin{aligned} |q(\zeta) - p(\zeta)| &= \left| \sum_{k=0}^{n} (b_k - a_k)\zeta^k \right| \\ &\leq \sum_{k=0}^{n} |b_k - a_k||\zeta|^k \\ &\leq \delta \sum_{k=0}^{n} |\zeta^k| \leq \delta \max_{i \in \{1,\dots,n\}} M_i < \min_{i \in \{1,\dots,n\}} m_i \leq |p(\zeta)| \end{aligned}$$

and therefore by Theorem 2.18 p and $p + (q - p) = q$ have the same number of zeros in D_j counting multiplicity. Since this holds for every $j \in \{1, \dots, n\}$ we conclude that every zero of q is contained in a disk D_j since these are disjoint and q has precisely as many zeros as p counting multiplicities. This implies the existence of a permutation of $\{1, \dots, n\}$ denoted π such that

$$|\mu_{\pi(i)} - \lambda_i| < \epsilon, \quad i \in \{1, \dots, n\}.$$

□

Our interest in Theorem 2.17 lies in its applications to the characteristic polynomial of a matrix.

Corollary 2.19. *Let $\{A_k\}_{k=1}^{\infty}$ be a sequence on $\mathcal{M}_{n\times n}$ which converges to A. For every $\epsilon > 0$ there exists a $K \in \mathbb{N}$ such that $k \geq K$ implies that the disks of radius $\epsilon > 0$ centered at the eigenvalues of A each contain as many eigenvalues of A as they contain eigenvalues of A_k.*

Proof. Since $\|A_k - A\| \to 0$ as $k \to 0$ is equivalent to

$$|(A_k)_{(i,j)} - A_{(i,j)}| \to 0, \quad i, j \in \{1, \dots, n\}$$

we see that the coefficients of the characteristic polynomial of A_k given by $q_k(\zeta) = \det(A_k - \zeta I)$ converge to those of $p(\zeta) = \det(A - \zeta I)$. Furthermore note that the leading coefficient of the characteristic polynomial is ± 1. Let $\epsilon > 0$ be given then Theorem 2.17 implies that we can find $\delta > 0$ such that if the coefficients of $p(z)$ and $q_k(z)$ differ by less than δ then the zeros of the polynomials differ by less than ϵ. By increasing k it is clear that the coefficients of p and q_k can be made arbitrarily small. This concludes our proof. □

2.6 Extensions of Gershgorin's Circle Theorem

As previously stated, Theorem 2.15 does not imply that a Gershgorin disk actually contains an eigenvalue however if a particular Gershgorin disk is disjoint from the other Gershgorin disks then it must contain an eigenvalue, see Appendix 7 of [15]. An even stronger result is contained in Theorem 2.20.

Theorem 2.20 (Gershgorin's Second Circle Theorem). *Let $A \in \mathcal{M}_{n\times n}$ and denote $\{C_i\}_{i=1}^n$ to be an enumeration of the Gershgorin discs of A defined in Theorem 2.15. Then if π is a permutation of $\{1,\dots,n\}$ such that*

$$U = \bigcup_{i=1}^{k} C_{\pi(i)} \text{ and } V = \bigcup_{i=k+1}^{n} C_{\pi(i)}$$

are disjoint then U contains exactly k eigenvalues while V contains exactly $n-k$ eigenvalues of A.

The proof of Theorem 2.20 relies on the so called argument principle of complex analysis. We state without proof the version of this theorem given in Chapter VIII of [20], even using the same notation that is used there.

Theorem 2.21 (The Argument Principle). *Let Ω be a bounded open connected set such that $\partial\Omega$ is piecewise smooth. If f denotes a function which is analytic apart from poles contained in the interior of Ω and $f(\zeta) \neq 0$ for $\zeta \in \partial\Omega$ then*

$$\frac{1}{2\pi i}\oint_{\partial\Omega} \frac{f'(\zeta)}{f(\zeta)} d\zeta = N_0 - N_\infty.$$

Here N_0 denotes the number of zeros of f in Ω and N_∞ denotes the number of poles of f in Ω.

For the purpose of the proof we also recall the Hadamard multiplication of matrices denoted $\circ$ which is defined by

$$(A \circ B)_{(i,j)} = A_{(i,j)} \cdot B_{(i,j)}.$$

With this in hand we can express the diagonal of the matrix A as $A \circ I$.

Proof. The proof we present is the one given in Chapter 6 of [18]. Consider the matrix function $B : [0,1] \to \mathcal{M}_{n\times n}$ defined by

$$B(t) = (1-t)I \circ A + tA = I \circ A + t(A - I \circ A).$$

Clearly $B(0)$ is diagonal and thus its eigenvalues are precisely its diagonal elements. Furthermore the Gershgorin disks are in this case reduced to one-point sets of eigenvalues. Since

$$B(t)_{(i,i)} = (1-t)A_{(i,i)} + tA_{(i,i)} = A_{(i,i)}$$

we observe that the Gershgorin disks of $B(t)$, denoted $C_i(t)$, have the same centers as those of A for every t. Furthermore for $t \in (0,1)$ the radius of the disks are scaled by a factor t. To see this note that

$$\sum_{\substack{j=1 \\ j\neq i}}^{n} |B_{(j,i)}(t)| = \sum_{\substack{j=1 \\ j\neq i}}^{n} |tA_{(j,i)}| = t\sum_{\substack{j=1 \\ j\neq i}}^{n} |A_{(j,i)}|.$$

Therefore if we define

$$U(t) = \bigcup_{i=1}^{k} C_{\pi(i)}(t) \text{ and } V(t) = \bigcup_{i=k+1}^{n} C_{\pi(i)}(t)$$

then $U(t) \subseteq U$ and $V(t) \subseteq V$ for $t \in [0,1]$ and therefore these are disjoint. By Corollary 2.16 we know that all eigenvalues of $B(t)$ are contained in $U(t) \cup V(t)$.

Let γ denote a union of piecewise smooth closed curves in $\mathbb{C}$ such that U is contained in the interior of γ and V lies in its exterior, a proof for the existence of such a curve can be found in [21]. Let $p(t,\zeta)$ denote the characteristic polynomial of $B(t)$ then Theorem 2.21 implies that

$$\frac{1}{2\pi i}\oint_{\gamma} \frac{\frac{\partial}{\partial\zeta}p(t,\zeta)}{p(t,\zeta)} d\zeta = N_0(t).$$

Now $N_0(t) \in \mathbb{Z}$ and since Corollary 2.19 implies that the number of zeros of the polynomial $p(t,\zeta)$ and $p(t+\delta,\zeta)$ in the interior of γ are constant for small δ it follows that the integral is continuous which in turn implies that $N_0(t)$ is constant for $t \in [0,1]$ (functions with values in $\mathbb{Z}$ which are continuous on a connected set must be constant). Since $N_0(0) = k$ it follows that $N_0(1) = k$ which concludes our proof. $\square$

2.7 The Theory of the Singular Value Decomposition

For completeness we also include the construction of the Singular Value Decomposition. We again consider arbitrary n-dimensional Hilbert spaces $\mathscr{H}_n$ and recall the definition of $\lambda_i(\mathcal{A})$ for Hermitian operators $\mathcal{A}$ as the ith eigenvalue of $\mathcal{A}$ ordered non-increasingly. In the context of singular values, since these are non-negative, it will be easier to work with the eigenvalues ordered non-decreasingly. We therefore define $\mu_j(\mathcal{A})$ for Hermitian operators $\mathcal{A} \in \mathcal{L}(\mathscr{H}_n)$ by the equation $\mu_j(\mathcal{A}) = \lambda_{n-j+1}(\mathcal{A})$, thus $\mu_j(\mathcal{A})$ is the jth eigenvalue of $\mathcal{A}$ ordered non-decreasingly counting multiplicity.

Definition 2.22. *The singular values of $\mathcal{A}$ are the values attained by the map $\sigma_j : \mathcal{L}(\mathscr{H}_n) \to [0,\infty)$ where $\sigma_j(\mathcal{A}) = \sqrt{\mu_j(\mathcal{A}^*\mathcal{A})}$ counting multiplicity.*

We first argue that this definition is sound. Clearly $\mathcal{A}^*\mathcal{A}$ is Hermitian so for $j \in \{1,\ldots,n\}$ we show that $\mu_j(\mathcal{A}^*\mathcal{A}) \geq 0$ so that we may take its square root. Let $\boldsymbol{v}_j$ denote a normalized eigenvector of $\mathcal{A}^*\mathcal{A}$ corresponding to the eigenvalue $\mu_j(\mathcal{A})$. Then

$$\begin{aligned}\mu_j(\mathcal{A}^*\mathcal{A}) &= \mu_j(\mathcal{A}^*\mathcal{A})\langle \boldsymbol{v}_j, \boldsymbol{v}_j\rangle = \langle \mu_j(\mathcal{A}^*\mathcal{A})\boldsymbol{v}_j, \boldsymbol{v}_j\rangle = \langle \mathcal{A}^*\mathcal{A}\boldsymbol{v}_j, \boldsymbol{v}_j\rangle \\ &= \langle \mathcal{A}\boldsymbol{v}_j, \mathcal{A}\boldsymbol{v}_j\rangle = \|\mathcal{A}\boldsymbol{v}_j\|^2 \geq 0.\end{aligned}$$

The following formulation and proof of the Singular Value Decomposition is from Chapter 7 of [3].

Theorem 2.23 (Singular-Value Decomposition)**.** *For every $\mathcal{A} \in \mathscr{L}(\mathscr{H}_n)$ there exists orthonormal bases of $\mathscr{H}_n$ denoted $\{\boldsymbol{e}_j\}_{j=1}^n$ and $\{\boldsymbol{f}_j\}_{j=1}^n$ respectively such that for every $\boldsymbol{v} \in \mathscr{H}_n$*

$$\mathcal{A}\boldsymbol{v} = \sum_{j=1}^{n} \sigma_j(\mathcal{A})\langle \boldsymbol{v}, \boldsymbol{e}_j\rangle \boldsymbol{f}_j.$$

We first give the definition of the square root of a positive operator.

Definition 2.24. *Let $\mathcal{A} \in \mathcal{L}(\mathscr{H}_n)$ be an Hermitian operator then $\mathcal{A}$ is a positive operator if for every $\boldsymbol{v} \in \mathscr{H}_n$*

$$\langle \mathcal{A}\boldsymbol{v}, \boldsymbol{v}\rangle \geq 0.$$

If we consider a vector space over $\mathbb{C}$ then every positive operator is Hermitian, this is however not true in general.

Lemma 2.25 (Square Root)**.** *For every positive operator $\mathcal{A} \in \mathcal{L}(\mathscr{H}_n)$ there exists a unique positive operator $\mathcal{R}$ such that $\mathcal{R}^2 = \mathcal{A}$. This operator is called the positive square root of $\mathcal{A}$.*

Proof. It is a direct consequence of Definition 2.24 that the eigenvalues of a positive operator are non-negative. Let $\{\boldsymbol{e}_j\}_{j=1}^n$ denote an orthonormal basis of $\mathscr{H}_n$ consisting of eigenvectors of $\mathcal{A}$ such that $\boldsymbol{e}_j$ is the eigenvector of $\mathcal{A}$ corresponding to $\mu_j(\mathcal{A})$. We define $\mathcal{R} \in \mathcal{L}(\mathscr{H}_n)$ by

$$\mathcal{R}\boldsymbol{e}_j = \sqrt{\mu_j(\mathcal{A})}\boldsymbol{e}_j, \quad j \in \{j, \dots, n\}.$$

It is then clear that

$$\mathcal{R}^2\boldsymbol{e}_j = \mathcal{R}\sqrt{\mu_j(\mathcal{A})}\boldsymbol{e}_j = \left(\sqrt{\mu_j(\mathcal{A})}\right)^2 \boldsymbol{e}_j = \mu_j(\mathcal{A})\boldsymbol{e}_j = \mathcal{A}\boldsymbol{e}_j$$

and therefore $\mathcal{R}^2 = \mathcal{A}$. Since $\mathcal{R}$ has positive eigenvalues and a orthonormal set of eigenvectors it is clear that $\mathcal{R}$ defined this way is Hermitian. Suppose that $\mathcal{R}'$ is another positive square root of $\mathcal{A}$ and let $\boldsymbol{v} \in \mathscr{H}_n$ denote an arbitrary eigenvector of $\mathcal{A}$ with corresponding eigenvalue $\mu \geq 0$. Since $\mathcal{R}'$ is Hermitian there is an orthonormal basis of $\mathscr{H}_n$ consisting of eigenvectors of $\mathcal{R}'$ denoted $\{\boldsymbol{f}_j\}_{j=1}^n$ such that $\boldsymbol{f}_j$ corresponds to the non-negative eigenvalue $\mu_j(\mathcal{R}')$. We can therefore express $\boldsymbol{v}$ in this basis:

$$\boldsymbol{v} = \sum_{j=1}^{n}\langle \boldsymbol{v}, \boldsymbol{f}_j\rangle \boldsymbol{f}_j \Rightarrow \mathcal{R}'\boldsymbol{v} = \sum_{j=1}^{n} \mu_j(\mathcal{R}')\langle \boldsymbol{v}, \boldsymbol{f}_j\rangle \boldsymbol{f}_j.$$

Then since

$$\mathcal{R}'(\mathcal{R}'\boldsymbol{v}) = \mathcal{A}\boldsymbol{v} = \mu\boldsymbol{v}$$

we find that

$$\sum_{j=1}^{n} \mu\langle \boldsymbol{v}, \boldsymbol{f}_j\rangle \boldsymbol{f} = \mu\boldsymbol{v} = \sum_{j=1}^{n} \mu_j(\mathcal{R}')^2\langle \boldsymbol{v}, \boldsymbol{f}_j\rangle \boldsymbol{f}_j.$$

This shows that for every index j where $\langle \boldsymbol{v}, \boldsymbol{f}_j\rangle \neq 0$ we must have that $\mu = \mu_j(\mathcal{R}')^2$. Since $\mu_j(\mathcal{R}') \geq 0$ we can thus conclude that $\boldsymbol{v}$ is an eigenvector of $\mathcal{R}'$ with eigenvalue $\sqrt{\mu}$. With the basis $\{\boldsymbol{e}_j\}_{j=1}^n$ we therefore see that $\mathcal{R}'\boldsymbol{e}_j = \sqrt{\mu_j(\mathcal{A})}\boldsymbol{e}_j$. This proves that our definition of the square root is unique. □

Lemma 2.26 (Polar Decomposition)**.** *For every $\mathcal{A} \in \mathcal{L}(\mathscr{H}_n)$ there exists an isometry $\mathcal{U} \in \mathcal{L}(\mathscr{H}_n)$ such that*

$$\mathcal{A} = \mathcal{U}\sqrt{\mathcal{A}^*\mathcal{A}}.$$

Proof. Let $\mathcal{R} = \sqrt{\mathcal{A}^*\mathcal{A}}$. Then for $\boldsymbol{v} \in \mathscr{H}_n$

$$\|\mathcal{A}\boldsymbol{v}\|^2 = \langle \mathcal{A}^*\mathcal{A}\boldsymbol{v}, \boldsymbol{v}\rangle = \langle \mathcal{R}\boldsymbol{v}, \mathcal{R}\boldsymbol{v}\rangle = \|\mathcal{R}\boldsymbol{v}\|^2 \Rightarrow \|\mathcal{A}\boldsymbol{v}\| = \|\mathcal{R}\boldsymbol{v}\|.$$

We define $\mathcal{U}_1 : \text{ran } \mathcal{R} \to \text{ran } \mathcal{A}$ by

$$\mathcal{U}_1(\mathcal{R}\boldsymbol{v}) = \mathcal{A}\boldsymbol{v}.$$

For this definition to make sense we need to show that $\mathcal{R}\boldsymbol{v}_1 = \mathcal{R}\boldsymbol{v}_2$ implies that $\mathcal{A}\boldsymbol{v}_1 = \mathcal{A}\boldsymbol{v}_2$. Suppose to that end that $\mathcal{R}\boldsymbol{v}_1 = \mathcal{R}\boldsymbol{v}_2$ then

$$\begin{aligned} 0 &= \|\mathcal{R}(\boldsymbol{v}_1 - \boldsymbol{v}_2)\|^2 = \langle \mathcal{R}(\boldsymbol{v}_1 - \boldsymbol{v}_2), \mathcal{R}(\boldsymbol{v}_1 - \boldsymbol{v}_2)\rangle \\ &= \langle \mathcal{A}^*\mathcal{A}(\boldsymbol{v}_1 - \boldsymbol{v}_2), \boldsymbol{v}_1 - \boldsymbol{v}_2\rangle = \|\mathcal{A}(\boldsymbol{v}_1 - \boldsymbol{v}_2)\|^2 \Rightarrow \mathcal{A}\boldsymbol{v}_1 = \mathcal{A}\boldsymbol{v}_2. \end{aligned}$$

It is clear that $\mathcal{U}_1$ is surjective however since it is also an isometry onto its range it is bijective. The Rank-Nullity Theorem now implies that

$$\dim \text{ran } \mathcal{R} = \dim \text{ran } \mathcal{A}.$$

Pick orthonormal bases for $(\text{ran } \mathcal{R})^\perp$ and $(\text{ran } \mathcal{A})^\perp$ denoted $\{\boldsymbol{e}_j\}_{j=1}^m$ and $\{\boldsymbol{f}_j\}_{j=1}^m$ respectively. Any vector $\boldsymbol{u} \in \mathscr{H}_n$ can now be uniquely decomposed as

$$\boldsymbol{u} = \boldsymbol{v} + \sum_{j=1}^{m} a_j \boldsymbol{e}_j$$

where $\boldsymbol{v} \in \text{ran } \mathcal{R}$. We define $\mathcal{U} : \mathscr{H}_n \to \mathscr{H}_n$ by

$$\mathcal{U}\boldsymbol{u} = \mathcal{U}\left(\boldsymbol{v} + \sum_{j=1}^{m} a_j \boldsymbol{e}_j\right) = \mathcal{U}_1\boldsymbol{v} + \sum_{j=1}^{m} a_j \boldsymbol{f}_j.$$

Then $\mathcal{U}$ is an isometry since the Pythagorean Theorem implies that

$$\begin{aligned}\|\mathcal{U}(\boldsymbol{u})\|^2 &= \left\|\mathcal{U}_1\boldsymbol{v} + \sum_{j=1}^{m} a_j \boldsymbol{f}_j\right\|^2 = \|\mathcal{U}_1\boldsymbol{v}\|^2 + \sum_{j=1}^{m} |a_j|^2 = \|\boldsymbol{v}\|^2 + \sum_{j=1}^{m} |a_j|^2 \\ &= \left\|\boldsymbol{v} + \sum_{j=1}^{m} a_j \boldsymbol{e}_j\right\|^2 = \|\boldsymbol{u}\|^2 .\end{aligned}$$

We have left to verify that $\mathcal{U}\mathcal{R} = \mathcal{A}$. Let $\boldsymbol{u} \in \mathscr{H}_n$ be arbitrary then

$$\mathcal{U}\mathcal{R}\boldsymbol{u} = \mathcal{U}_1\mathcal{R}\boldsymbol{u} = \mathcal{A}\boldsymbol{u}$$

and thus the result now follows. □

Proof of Theorem 2.23. Let $\{\boldsymbol{e}_j\}_{j=1}^n$ be a basis of eigenvectors of $\sqrt{\mathcal{A}^*\mathcal{A}}$ where $\boldsymbol{e}_j$ corresponds to the eigenvalue $\mu_j(\sqrt{\mathcal{A}^*\mathcal{A}}) = \sigma_j(\mathcal{A})$. Any vector $\boldsymbol{v} \in \mathscr{H}_n$ can be written

$$\boldsymbol{v} = \sum_{j=1}^{n} \langle \boldsymbol{v}, \boldsymbol{e}_j\rangle \boldsymbol{e}_j$$

and applying $\sqrt{\mathcal{A}^*\mathcal{A}}$ we find that

$$\sqrt{\mathcal{A}^*\mathcal{A}}\boldsymbol{v} = \sum_{j=1}^{n} \sigma_j(\mathcal{A})\langle \boldsymbol{v}, \boldsymbol{e}_j\rangle \boldsymbol{e}_j.$$

By the Polar Decomposition we can find a isometry $\mathcal{U}$ such that $\mathcal{U}\sqrt{\mathcal{A}^*\mathcal{A}} = \mathcal{A}$. Note that if we define $\boldsymbol{f}_j = \mathcal{U}\boldsymbol{e}_j$ then $\{\boldsymbol{f}_j\}_{j=1}^n$ is an orthonormal basis of $\mathscr{H}_n$. Indeed

$$\langle \boldsymbol{f}_j, \boldsymbol{f}_i\rangle = \langle \mathcal{U}\boldsymbol{e}_j, \mathcal{U}\boldsymbol{e}_i\rangle = \langle \mathcal{U}^*\mathcal{U}\boldsymbol{e}_j, \boldsymbol{e}_i\rangle = \langle \boldsymbol{e}_j, \boldsymbol{e}_i\rangle.$$

Thus we find that

$$\mathcal{A}\boldsymbol{v} = \mathcal{U}\sqrt{\mathcal{A}^*\mathcal{A}}\boldsymbol{v} = \sum_{j=1}^{n} \sigma_j(\mathcal{A})\langle \boldsymbol{v}, \boldsymbol{e}_j\rangle \boldsymbol{f}_j.$$

□

In matrix form we conclude the following corollary:

Corollary 2.27. *Let $A \in \mathcal{M}_{n\times m}$ then we can find unitary matrices $U \in \mathcal{M}_{n\times n}$ and $V \in \mathcal{M}_{m\times m}$ such that*

$$A = U\Sigma V^*$$

where $\Sigma = \operatorname{diag}\,(\sigma_1(A), \ldots, \sigma_n(A))$.

3 Analytic Theory of the Spectrum of a Parameter Perturbation

We now turn our attention to the spectral analysis of the parameter perturbation $\mathcal{A}+t\mathcal{F}$ where $\mathcal{A}$ and $\mathcal{F}$ are Hermitian operators in $\mathcal{L}(\mathscr{H}_n)$ and $t \in \mathbb{R}$. We study both first and second order Taylor formulas with the aim of proving Theorem 1.1, here reformulated in Theorem 3.5 and Theorem 3.6.

As stated before the results of this section are previously known and can be proven directly if certain properties concerning the analyticity of the spectrum are taken for granted. However we have not found the method which we supply using matrix decompositions in the standard literature on the subject. After having given our proof we will also supply the one presented by David Hilbert and Richard Courant.

Furthermore it turns out that the methods developed in this section will be applicable to global perturbations. This will allow us to give a short proof of Theorem 1.2 which is stated in a recent article [10].

3.1 Decomposing the Perturbation

We begin by introducing some terminology which we will use. Let $\mathcal{A}$ be a given Hermitian operator in $\mathcal{L}(\mathscr{H}_n)$. The Spectral Theorem implies that we can decompose $\mathscr{H}_n$ as a direct sum of eigenspaces of $\mathcal{A}$. These are necessarily orthogonal to each other. Suppose that $\mathcal{A}$ has m different eigenvalues so that $\lambda_{p(j)}(\mathcal{A})$ for $j = 1, \dots, m$ is a distinct enumeration of these with $\lambda_{p(j_0)}(\mathcal{A})$ denoting the first occurrence of $\lambda_{p(j_0)(\mathcal{A})}$ in the collection $\{\lambda_j(\mathcal{A})\}_{j=1}^{n}$. If we denote $\mathcal{N}_j = \ker(\mathcal{A} - \lambda_{p(j)}(\mathcal{A})I)$ we see that

$$\mathscr{H}_n = \bigoplus_{j=1}^{m} \ker(\mathcal{A} - \lambda_{p(j)}(\mathcal{A})I) = \bigoplus_{j=1}^{m} \mathcal{N}_j$$

and thus by choosing a basis for each kernel we find that the matrix representation of $\mathcal{A}$ is given by

$$\begin{pmatrix} \lambda_{p(1)}(\mathcal{A})I & 0 & \cdots & 0 \\ 0 & \lambda_{p(2)}(\mathcal{A})I & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_{p(m)}(\mathcal{A})I \end{pmatrix} \tag{6}$$

where each $\lambda_{p(j)}(\mathcal{A})$ is unique.

Consider now another Hermitian operator $\mathcal{F}$ in $\mathcal{L}(\mathscr{H}_n)$ and let $\mathcal{P}_{\mathcal{N}_j}$ denote orthogonal projection onto $\mathcal{N}_j$. If we write $\mathcal{P}_{\mathcal{N}_j^\perp} = I - \mathcal{P}_{\mathcal{N}_j}$ then

$$\begin{aligned} \mathcal{F} &= (\mathcal{P}_{\mathcal{N}_j} + \mathcal{P}_{\mathcal{N}_j^\perp})\mathcal{F}(\mathcal{P}_{\mathcal{N}_j} + \mathcal{P}_{\mathcal{N}_j^\perp}) \\ &= \mathcal{P}_{\mathcal{N}_j}\mathcal{F}\mathcal{P}_{\mathcal{N}_j} + \mathcal{P}_{\mathcal{N}_j}\mathcal{F}\mathcal{P}_{\mathcal{N}_j^\perp} + \mathcal{P}_{\mathcal{N}_j}^\perp\mathcal{F}(\mathcal{P}_{\mathcal{N}_j} + \mathcal{P}_{\mathcal{N}_j^\perp}). \end{aligned}$$

Since

$$(\mathcal{P}_{\mathcal{N}_j}\mathcal{F}\mathcal{P}_{\mathcal{N}_j})^* = \mathcal{P}_{\mathcal{N}_j}\mathcal{F}\mathcal{P}_{\mathcal{N}_j}$$

we see that $\mathcal{P}_{\mathcal{N}_j}\mathcal{F}\mathcal{P}_{\mathcal{N}_j}$ is an Hermitian operator. Also $\mathcal{P}_{\mathcal{N}_j}\mathcal{F}\mathcal{P}_{\mathcal{N}_j}$ is invariant under $\mathcal{N}_j$ and therefore The Spectral Theorem implies that we can find an orthonormal basis of $\mathcal{N}_j$ consisting of eigenvectors of $\mathcal{P}_{\mathcal{N}_j}\mathcal{F}\mathcal{P}_{\mathcal{N}_j}$. Doing this for every eigenspace we obtain an orthonormal basis for $\mathscr{H}_n$ which we denote with $\{\boldsymbol{e}_i\}_{i=1}^n$. Since $\boldsymbol{e}_i$ is an eigenvector of $\mathcal{A}$ the matrix representation of $\mathcal{A}$ with respect to this basis will be of the form in equation (6). The matrix form of $\mathcal{F}$, denoted F, will be what we choose to call block-wise diagonal in accordance with the same definition given in [10]. What we mean by this is that if $\lambda_i(\mathcal{A}) = \lambda_j(\mathcal{A})$ but $i \neq j$ then $F_{(i,j)} = 0$. Therefore F will be of the form

$$F = \begin{pmatrix} \Lambda_{F_1} & * & \cdots & * \\ * & \Lambda_{F_2} & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & \Lambda_{F_m} \end{pmatrix}$$

where each Λ_{F_j} is the diagonal matrix which represents the operator $\mathcal{P}_{\mathcal{N}_j}\mathcal{F}\mathcal{P}_{\mathcal{N}_j}$ of the form

$$\Lambda_{F_j} = \begin{pmatrix} \lambda_{p(j)}(\mathcal{P}_{\mathcal{N}_j}\mathcal{F}\mathcal{P}_{\mathcal{N}_j}) & 0 & \cdots & 0 \\ 0 & \lambda_{p(j)+1}(\mathcal{P}_{\mathcal{N}_j}\mathcal{F}\mathcal{P}_{\mathcal{N}_j}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_{p(j+1)-1}(\mathcal{P}_{\mathcal{N}_j}\mathcal{F}\mathcal{P}_{\mathcal{N}_j}) \end{pmatrix}. \tag{7}$$

We see that this definition of F_j implies that the diagonal is ordered non-increasingly.

Definition 3.1. *Let $\mathcal{A}$ and $\mathcal{F}$ be given Hermitian operators in $\mathcal{L}(\mathscr{H}_n)$. Choosing an orthonormal basis of eigenvectors of $\mathcal{A}$ such that the matrix representation of $\mathcal{A}$ is of the form in equation* (6) *the matrix representation of $\mathcal{F}$ denoted $F \in \mathcal{M}_{n\times n}$ is said to be block-wise diagonal if*

$$F_{(i,j)} = 0$$

for indices i, j where $\lambda_i(\mathcal{A}) = \lambda_j(\mathcal{A})$ and $i \neq j$. If the additional assumption is placed that $F_{(i,i)} < F_{(j,j)}$ whenever $i < j$ and $\lambda_i(\mathcal{A}) = \lambda_j(\mathcal{A})$ we say that the matrix F has block-wise decreasing diagonal.

We note that the matrix F given in Equation (7) has a block-wise decreasing diagonal precisely when the eigenvalues of $\mathcal{P}_{\mathcal{N}_j}\mathcal{F}\mathcal{P}_{\mathcal{N}_j}$ are distinct.

Definition 3.2. *Given a Hermitian operator $\mathcal{A}$ we say that the Hermitian operator $\mathcal{F}$ is a simple perturbation of $\mathcal{A}$ if the eigenvalues of $\mathcal{P}_{\mathcal{N}_j}\mathcal{F}\mathcal{P}_{\mathcal{N}_j}$ are distinct for every j. We let $\mathscr{F}(\mathcal{A})$ denote the collection of such operators.*

It is worth noting that if $\mathcal{A}$ has simple eigenvalues then $\mathscr{F}(\mathcal{A})$ coincides with the set of Hermitian operators on $\mathscr{H}_n$. We gather the results of this section in the following lemma:

Lemma 3.3. *Given Hermitian operators $\mathcal{A}$ and $\mathcal{F}$ in $\mathcal{L}(\mathscr{H}_n)$ there exists an orthonormal basis $\{\boldsymbol{e}_i\}_{i=1}^n$ consisting of eigenvectors of $\mathcal{A}$ such that the matrix representation of $\mathcal{F}$ with respect to this basis will be block-wise diagonal as represented in equation* (7).

Furthermore if $\mathcal{F} \in \mathscr{F}(\mathcal{A})$ then the matrix representation of $\mathcal{F}$ can be chosen to have a block-wise decreasing diagonal.

3.2 First Order Expansion of the Spectrum

Our first result will concern estimating the first order term in the Taylor expansion of the eigenvalues of $\mathcal{A}+t\mathcal{F}$ where $t \in \mathbb{R}$ and $\mathcal{A}$ and $\mathcal{F}$ are Hermitian operators in $\mathcal{L}(\mathscr{H}_n)$. The reason for considering this specific case before the second order expansion is that it illustrates the method that we will use later for more complicated matrices. Furthermore the result is more general in the sense that it allows for any Hermitian perturbation $\mathcal{F}$ and not just the simple perturbations of $\mathcal{A}$ as will be the case with the second order expansion. Since the argument supplied in the proof will be repeated several times in this text we take extra care to present the proof in a rigorous manner.

Theorem 3.4. *Let $\mathcal{A}$ and $\mathcal{F}$ be Hermitian operators in $\mathcal{L}(\mathscr{H}_n)$ and let j_0 be the first index of occurrence of the eigenvalue $\lambda_{j_0}(\mathcal{A})$ with multiplicity l in the collection $\{\lambda_j(\mathcal{A})\}_{j=1}^n$. The eigenvalues of $\mathcal{A}+t\mathcal{F}$ for real t then satisfy*

$$\lambda_j(\mathcal{A}+t\mathcal{F}) = \lambda_j(\mathcal{A}) + t\lambda_{j-j_0+1}(\mathcal{P}_{\mathcal{N}}\mathcal{F}\mathcal{P}_{\mathcal{N}}) + \mathcal{O}(t^2), \quad j \in \{j_0, \dots, j_0+l-1\}.$$

Proof. Suppose first that the Hermitian operator $\mathcal{A}$ satisfies $\dim\ker\mathcal{A} = l \geq 1$. This is the same as saying that 0 is an eigenvalue of $\mathcal{A}$ with algebraic multiplicity l. Define j_0 as the first index j with $\lambda_j(\mathcal{A}) = 0$. From Lemma 3.3 we gather that there exists an orthonormal basis $\{\boldsymbol{e}_j\}_{j=1}^n$ of eigenvectors of $\mathcal{A}$ such that the matrix representation of $\mathcal{A}+t\mathcal{F}$ is given by

$$\Lambda_{\mathcal{A}} + tF = \begin{pmatrix} 0 & 0 \\ 0 & \Lambda_\tau \end{pmatrix} + \begin{pmatrix} t\Lambda_\rho & tY \\ tY^* & tZ \end{pmatrix}.$$

Here $\{\boldsymbol{e}_j\}_{j=1}^l$ are the eigenvectors which correspond to the eigenvalue 0 while $\{\boldsymbol{e}_j\}_{j=l+1}^n$ are the eigenvectors which correspond to the non-zero eigenvalues of $\mathcal{A}$. The matrix $\Lambda_\rho = \operatorname{diag}(\rho_1, \dots, \rho_l)$ consists of the eigenvalues of $\mathcal{P}_{\ker\mathcal{A}}\mathcal{F}\mathcal{P}_{\ker\mathcal{A}}$ ordered non-increasingly so that $\rho_j = \lambda_j(\mathcal{P}_{\ker\mathcal{A}}\mathcal{F}\mathcal{P}_{\ker\mathcal{A}})$. Since $\Lambda_\tau = \operatorname{diag}(\tau_{l+1}, \dots, \tau_n)$ is a diagonal matrix whose diagonal entries are non-zero it is invertible and therefore Theorem 2.3 implies that $\Lambda_\tau + tZ$ is invertible for small t. By applying the decomposition of Lemma 2.14 we see that another matrix representation of $\mathcal{A}+t\mathcal{F}$ in a basis which is not necessarily orthonormal is given by

$$\begin{aligned} \Psi(t) &= \begin{pmatrix} t\Lambda_\rho - tY(\Lambda_\tau+tZ)^{-1}tY^* & tY \\ (\Lambda_\tau+tZ)^{-1}tY^*\Big(t\Lambda_\rho - tY(\Lambda_\tau+tZ)^{-1}tY^*\Big) & (\Lambda_\tau+tZ)^{-1}tY^*tY + \Lambda_\tau + tZ \end{pmatrix} \\ &= \begin{pmatrix} t\Lambda_\rho + \mathcal{O}(t^2) & \mathcal{O}(t) \\ \mathcal{O}(t^2) & \Lambda_\tau + \mathcal{O}(t) \end{pmatrix}. \end{aligned} \tag{8}$$

Boundedness of $(\Lambda_\tau + tZ)^{-1}$ in a neighborhood of the origin follows from Corollary 2.5 and by Corollary 2.16 we know that the eigenvalues of $\mathcal{A}+t\mathcal{F}$ are contained in Gershgorin disks $C_j(t)$ expanded along the columns of Ψ. These are the disks $\{C_j(t)\}_{j=1}^n$ where

$$\begin{aligned} C_j(t) &= \left\{ \zeta \in \mathbb{C} \mid \left|\zeta - t\rho_j - (tY(\Lambda_\tau+tZ)^{-1}tY^*)_{(j,j)}\right| \leq \sum_{k\neq j} \left|(\Psi(t))_{(k,j)}\right| \right\} \\ &\subseteq \left\{ \zeta \in \mathbb{C} \mid |\zeta - t\rho_j| \leq K_j t^2 \right\} \end{aligned}$$

if $1 \leq j \leq l$ and

$$C_j(t) = \left\{ \zeta \in \mathbb{C} \mid \left| \zeta - \tau_j - t^2 \Big((\Lambda_\tau + tZ)^{-1} Y^* Y \Big)_{(j,j)} - tZ_{(j,j)} \right| \leq \sum_{k \neq j} |(\Psi(t))_{(k,j)}| \right\}$$
$$\subseteq \{ \zeta \in \mathbb{C} \mid |\zeta - \tau_j| \leq K_j |t| \}$$

if $l+1 \leq j \leq n$. Here $\{K_j\}_{j=1}^n$ denotes a collection of positive numbers. As $t \to 0$ it is clear that the l first disks $C_j(t)$ tends to 0 while the remaining $n-l$ disks $C_j(t)$ tends to the non-zero eigenvalues of $\mathcal{A}$ denoted $\{\tau_j\}_{j=l+1}^n$. It is clear from this that we can make the l first disks disjoint from the remaining $n-l$ disks. For a pictorial representation see Figure 1.

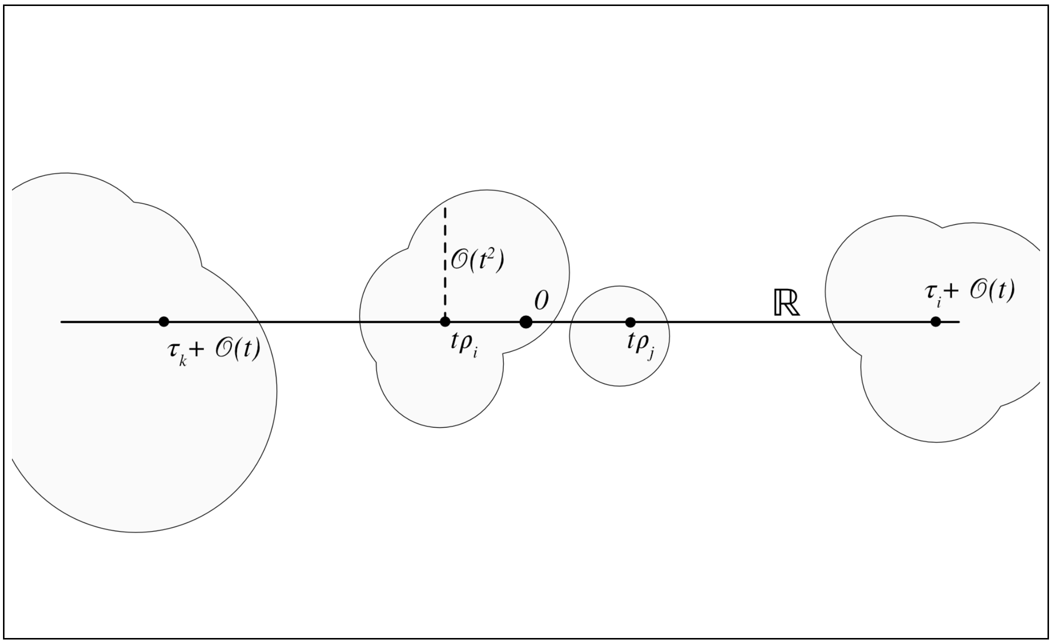

Figure 1: An illustration of the Gershgorin disks for the perturbation $\mathcal{A} + t\mathcal{F}$. In reality we need only consider the restriction of these to the real line since the eigenvalues will be real, however for illustrative reasons we imagine these as subsets of $\mathbb{C}$.

By Theorem 2.20 we know that unions of Gershgorin disks which are disjoint from the remaining Gershgorin disks will contain as many eigenvalues as there are disks in the union. Let $\rho_{j_1} \neq \rho_{j_2}$, we claim that we can choose t such that $C_{j_1}(t)$ and $C_{j_2}(t)$ are disjoint if t is small. Indeed if $\zeta \in C_{j_1}(t) \cap C_{j_2}(t)$ then

$$|\zeta - t\rho_{j_1}| = \mathcal{O}(t^2) \text{ and } |\zeta - t\rho_{j_2}| = \mathcal{O}(t^2)$$

hold simultaneously. If these were to hold for all $t > 0$ then

$$|t\rho_{j_1} - t\rho_{j_2}| \leq |\zeta - t\rho_{j_1}| + |\zeta - t\rho_{j_2}| = \mathcal{O}(t^2)$$
$$\Rightarrow |\rho_{j_1} - \rho_{j_2}| \leq C|t|$$

for some $C > 0$. Now this is a contradiction since $\rho_{j_1} \neq \rho_{j_2}$. Let that ρ_j denote an eigenvalue of Λ_ρ with multiplicity l_j. It follows that we can make the l_j Gersghorin disks corresponding to ρ_j disjoint from the disks we exclude. Therefore each eigenvalue of Λ_ρ can be matched with precisely one eigenvalue of $\mathcal{A}+t\mathcal{F}$ which lies within a distance of $\mathcal{O}(t^2)$. This implies that if t is sufficiently small then

$$\lambda_j(\mathcal{A}+t\mathcal{F}) = t\lambda_{j-j_0+1}(\mathcal{P}_\mathcal{N}\mathcal{F}\mathcal{P}_\mathcal{N}) + \mathcal{O}(t^2), \quad j \in \{j_0, \dots, j_0+l-1\}.$$

The extension to arbitrary eigenvalues is now straight-forward. For a given Hermitian operator $\mathcal{A}$ we let j_0 be the first index of occurrence of $\lambda_{j_0}(\mathcal{A})$ in the collection $\{\lambda_j(\mathcal{A})\}_{j=1}^n$ and suppose that $\lambda_{j_0}(\mathcal{A})$ has multiplicity l. Then the operator $\mathcal{A}' = \mathcal{A} - \lambda_{j_0}(\mathcal{A})I$ satisfies $\lambda_j(\mathcal{A}') = \lambda_j(\mathcal{A}) - \lambda_{j_0}(\mathcal{A})I$ and $\mathcal{N}$ defined by $\mathcal{N} = \ker(\mathcal{A} - \lambda_{j_0}I)$ is the kernel of $\mathcal{A}'$. By applying our previous reasoning we conclude that

$$\lambda_j(\mathcal{A}'+t\mathcal{F}) = t\lambda_{j-j_0+1}(\mathcal{P}_\mathcal{N}\mathcal{F}\mathcal{P}_\mathcal{N}), \quad j \in \{j_0, \dots, j_0+l-1\}.$$

But then since $\lambda_j(\mathcal{A}'+t\mathcal{F}) = \lambda_j(\mathcal{A}-t\mathcal{F}) - \lambda_{j_0}(\mathcal{A})I$ the result now follows. □

3.3 Second Order Expansion of the Spectrum

We again study the perturbation $\mathcal{A}+t\mathcal{F}$ where $\mathcal{A}$ and $\mathcal{F}$ denotes Hermitian operators in $\mathcal{L}(\mathscr{H}_n)$ and $t \in \mathbb{R}$ but this time we add the additional condition that $\mathcal{F} \in \mathscr{F}(\mathcal{A})$ with the purpose of deriving a second order expansion for the eigenvalues.

In order to simplify the presentation we drop the focus on the ordering of the eigenvalues given by the functions λ_j and instead denote the eigenvalues of $\mathcal{A}$ with $\{\alpha_j\}_{j=1}^n$. The eigenvalues of $\mathcal{P}_{\mathcal{N}_j}\mathcal{F}\mathcal{P}_{\mathcal{N}_j}$, previously denoted ρ_i satisfies

$$\rho_i = \langle \mathcal{F}\boldsymbol{e}_i, \boldsymbol{e}_i\rangle$$

where $\{\boldsymbol{e}_i\}_{i=1}^n$ is a basis of eigenvectors of $\mathcal{A}$ chosen such that the matrix representation of $\mathcal{F}$ has a block-wise decreasing diagonal.

Theorem 3.5. *Let $\mathcal{A}$ and $\mathcal{F}$ be Hermitian operators in $\mathcal{L}(\mathscr{H}_n)$ where $\mathcal{F} \in \mathscr{F}(\mathcal{A})$. Let $\{\boldsymbol{e}_j\}_{j=1}^n$ denote an orthonormal basis consisting of eigenvectors of $\mathcal{A}$ with corresponding eigenvalues $\{\alpha_j\}_{j=1}^n$ in which the matrix representation of $\mathcal{F}$ has a block-wise decreasing diagonal. Then the eigenvalues of $\mathcal{A}+t\mathcal{F}$ denoted $\xi_1(t), \dots, \xi_n(t)$ can be ordered so that they satisfy*

$$\xi_j(t) = \alpha_j + t\langle \mathcal{F}\boldsymbol{e}_j, \boldsymbol{e}_j\rangle - t^2 \sum_{p:\alpha_p \neq \alpha_j} \frac{|\langle \mathcal{F}\boldsymbol{e}_j, \boldsymbol{e}_p\rangle|^2}{\alpha_p - \alpha_j} + \mathcal{O}(t^3), \quad j \in \{1, \dots, l\}.$$

Before turning to the proof we will need to construct a basis which will simplify our problem. This decomposition is taken from [10]. Similar to the previous study of a first order formula, we first consider the case where the operator $\mathcal{A}$ has a non-trivial kernel and denote $l = \dim\ker\mathcal{A}$. Since $\mathcal{F} \in \mathscr{F}(\mathcal{A})$ it follows by Lemma 3.3 that we can find an orthonormal basis $\{\boldsymbol{e}_i\}_{i=1}^n$ of $\mathscr{H}_n$ consisting of eigenvectors of $\mathcal{A}$ such that the matrix representation of $\mathcal{F}$ with respect to this basis has a block-wise decreasing diagonal. By arrangement of the basis we can assume that the matrix representation of $\mathcal{A}$ is of the form

$$\begin{pmatrix} 0 & 0 \\ 0 & \Lambda_\tau \end{pmatrix}$$

where Λ_τ is the diagonal matrix which contains the non-zero eigenvalues of $\mathcal{A}$. The matrix representation of F will be of the form

$$\begin{pmatrix} \Lambda_\rho & Y \\ Y^* & Z \end{pmatrix}$$

where $\Lambda_\rho = \text{diag}\ (\rho_1, \dots, \rho_l)$ and $\rho_1 > \rho_2 > \dots > \rho_l$. These are the elements $\langle \mathcal{F}\boldsymbol{e}_j, \boldsymbol{e}_j \rangle$. We define the vectors $\{\boldsymbol{v}_j(t)\}_{j=1}^n$ by

$$\boldsymbol{v}_j(t) = \boldsymbol{e}_j + t \cdot \sum_{\substack{p:\alpha_p=\alpha_j \\ p\neq j}} \frac{\langle (\mathcal{A}-\alpha_j I)^\dagger \mathcal{F}\boldsymbol{e}_j, \mathcal{F}\boldsymbol{e}_p \rangle}{\langle \mathcal{F}\boldsymbol{e}_p, \boldsymbol{e}_p \rangle - \langle \mathcal{F}\boldsymbol{e}_j, \boldsymbol{e}_j \rangle} \boldsymbol{e}_p.$$

Since $\boldsymbol{v}_j(t)$ is a linear combination of eigenvectors of $\mathcal{A}$ corresponding to the same eigenvalue α_j it follows that $\boldsymbol{v}_j(t)$ is an eigenvector of $\mathcal{A}$ corresponding to the eigenvalue α_j. We now aim to express this set of vectors in matrix form to simplify our arguments. To that end we introduce the matrix N defined by

$$N_{(i,j)} = \begin{cases} 0, & \alpha_i \neq \alpha_j \text{ or } i = j \\ \frac{(F^*(\Lambda_\alpha - \alpha_j I)^\dagger F)_{(i,j)}}{F_{(i,i)} - F_{(j,j)}}, & \text{otherwise.} \end{cases} \tag{9}$$

By the definition it is clear that the matrix N will be invariant under the kernels $\ker(\mathcal{A} - \alpha_j I)$ and thus has its representation as

$$N = \begin{bmatrix} \begin{bmatrix} 0 & * \\ * & 0 \end{bmatrix} & 0 & 0 & 0 \\ 0 & \begin{bmatrix} 0 & * \\ * & 0 \end{bmatrix} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \begin{bmatrix} 0 & * \\ * & 0 \end{bmatrix} \end{bmatrix}.$$

Furthermore we see that the matrix whose columns are the coordinates of the vectors $\{\boldsymbol{v}_j(t)\}_{j=1}^n$ with respect to the basis $\{\boldsymbol{e}_j\}_{j=1}^n$ is precisely the matrix denoted $V_{ap}(t)$ which we define as

$$V_{ap}(t) = I + tN. \tag{10}$$

We choose the notation V_{ap} in accordance with [10] where it is claimed that this matrix will form an approximate basis of eigenvectors, something we will prove later. By Theorem 2.3 we see that $V_{ap}(t)$ is invertible if t is small enough which implies that $\{\boldsymbol{v}_j(t)\}_{j=1}^n$ in this case will form a basis of $\mathscr{H}_n$. The inverse is explicitly given by

$$V_{ap}^{-1} = (I + tN)^{-1} = \sum_{k=0}^{\infty} (-tN)^k = I - tN + \mathcal{O}(t^2).$$

Since $\boldsymbol{v}_j(t)$ is an eigenvector of $\mathcal{A}$ the matrix representation of $\mathcal{A}$ with respect to the basis $\{\boldsymbol{v}_j(t)\}_{j=1}^n$ is given by

$$\Lambda_{\mathcal{A}} = \begin{pmatrix} 0 & 0 \\ 0 & \Lambda_\tau \end{pmatrix}.$$

We proceed by studying the effects of this change of basis to the operator $\mathcal{F}$. If we express N in block-form

$$N = \begin{pmatrix} N_{11} & 0 \\ 0 & N_{22} \end{pmatrix}$$

then we see that the matrix representation of $t\mathcal{F}$ with respect to the basis $\{\boldsymbol{v}_j(t)\}_{j=1}^n$ becomes

$$\begin{aligned}
&(I+tN)^{-1}tF(I+tN) \\
&= \left(\begin{pmatrix} I - tN_{11} & 0 \\ 0 & I - tN_{22} \end{pmatrix} + \mathcal{O}(t^2) \right) \begin{pmatrix} t\Lambda_\rho & tY \\ tY^* & tZ \end{pmatrix} \begin{pmatrix} I + tN_{11} & 0 \\ 0 & I + tN_{22} \end{pmatrix} \\
&= \begin{pmatrix} (I - tN_{11})t\Lambda_\rho(I + tN_{11}) & (I - tN_{11})tY(I + tN_{22}) \\ (I - tN_{22})tY^*(I + tN_{11}) & (I - tN_{22})tZ(I + tN_{22}) \end{pmatrix} + \mathcal{O}(t^3) \\
&= \begin{pmatrix} t\Lambda_\rho + t^2(\Lambda_\rho N_{11} - N_{11}\Lambda_\rho) + \mathcal{O}(t^3) & tY + \mathcal{O}(t^2) \\ tY^* + \mathcal{O}(t^2) & tZ + \mathcal{O}(t^2) \end{pmatrix}.
\end{aligned}$$

By definition of N we find that if $1 \leq i, j \leq l = \dim \ker \mathcal{A}$ then

$$N_{(i,j)} = \begin{cases} 0 & i = j \\ \frac{(Y\Lambda_\tau^{-1}Y^*)_{(i,j)}}{\rho_i - \rho_j} & \text{otherwise.} \end{cases}$$

And therefore

$$\begin{aligned}
(\Lambda_\rho N_{11} - N_{11}\Lambda_\rho)_{(i,j)} &= \sum_{k=1}^{l} (\Lambda_\rho)_{(i,k)}(N_{11})_{(k,j)} - (N_{11})_{(i,k)}(\Lambda_\rho)_{k,j} \\
&= \begin{cases} 0 & i = j \\ \frac{(Y\Lambda_\tau^{-1}Y^*)_{(i,j)}}{\rho_i - \rho_j}(\rho_i - \rho_j) & \text{otherwise} \end{cases}
\end{aligned}$$

why

$$\Lambda_\rho N_{11} - N_{11}\Lambda_\rho = (Y\Lambda_\tau^{-1}Y^*) - I \circ (Y\Lambda_\tau^{-1}Y^*).$$

The operation $\circ$ denotes the Hadamard product. The matrix representation of the perturbation $\mathcal{A} + t\mathcal{F}$ with respect to the basis $\{\boldsymbol{v}_j(t)\}$ becomes

$$\begin{pmatrix} t\Lambda_\rho + t^2((Y\Lambda_\tau^{-1}Y^*) - I \circ (Y\Lambda_\tau^{-1}Y^*)) + \mathcal{O}(t^3) & tY + \mathcal{O}(t^2) \\ tY^* + \mathcal{O}(t^2) & \Lambda_\tau + tZ + \mathcal{O}(t^2) \end{pmatrix}. \tag{11}$$

Having introduced this basis we are now in a position to prove Theorem 3.5. While the decomposition is taken from [10] the proof we supply from this point on is novel to the best of our knowledge.

Proof of theorem 3.5. By applying the decomposition of Lemma 2.14 twice in order to the matrix representation of $\mathcal{A}+t\mathcal{F}$ given in Equation (11), while using the fact that $(\Lambda_\tau + tZ + \mathcal{O}(t^2))^{-1} = \Lambda_\tau^{-1} + \mathcal{O}(t)$ we find that

$$\begin{aligned}
&\begin{pmatrix} t\Lambda_\rho + t^2((Y\Lambda_\tau^{-1}Y^*) - I\circ(Y\Lambda_\tau^{-1}Y^*)) + \mathcal{O}(t^3) & tY+\mathcal{O}(t^2) \\ tY^* + \mathcal{O}(t^2) & \Lambda_\tau + tZ + \mathcal{O}(t^2) \end{pmatrix} \\
\sim &\begin{pmatrix} t\Lambda_\rho + t^2((Y\Lambda_\tau^{-1}Y^*) - I\circ(Y\Lambda_\tau^{-1}Y^*)) - t^2 Y\Lambda_\tau^{-1}Y^* + \mathcal{O}(t^3) & tY+\mathcal{O}(t^2) \\ \mathcal{O}(t^2) & \Lambda_\tau + \mathcal{O}(t) \end{pmatrix} \\
= &\begin{pmatrix} t\Lambda_\rho - t^2\Big[I\circ(Y\Lambda_\tau^{-1}Y^*)\Big] + \mathcal{O}(t^3) & tY+\mathcal{O}(t^2) \\ \mathcal{O}(t^2) & \Lambda_\tau + \mathcal{O}(t) \end{pmatrix} \\
\sim &\begin{pmatrix} t\Lambda_\rho - t^2\Big[I\circ(Y\Lambda_\tau^{-1}Y^*)\Big] + \mathcal{O}(t^3) & tY+\mathcal{O}(t^2) \\ \mathcal{O}(t^3) & \Lambda_\tau + \mathcal{O}(t) \end{pmatrix} \\
= &\begin{pmatrix} t\Lambda_\rho - t^2\Big[I\circ(Y\Lambda_\tau^{-1}Y^*)\Big] + \mathcal{O}(t^3) & \mathcal{O}(t) \\ \mathcal{O}(t^3) & \Lambda_\tau + \mathcal{O}(t) \end{pmatrix}.
\end{aligned} \tag{12}$$

Note that the matrix in the upper left position is diagonal up to errors of $\mathcal{O}(t^3)$. We let $\{C_j(t)\}_{j=1}^n$ denote the Gershgorin disks of this matrix where $C_j(t)$ denotes the disk centered at the jth diagonal element and whose radius is given by the sum in absolute value of the off-diagonal elements of the jth column in accordance with Corollary 2.16. Similar to the situation in the proof of Theorem 3.4 it is clear that the l first disks can be made disjoint from the remaining $n-l$ disks by restricting the size of t since the l-first will tend to the origin while the remaining disks will converge to the non-zero eigenvalues of $\mathcal{A}$. Furthermore since each ρ_i is distinct we can make each of the l first Gershgorin disks disjoint from each other. Theorem 2.20 now implies that if we denote the eigenvalues of $\mathcal{A}+t\mathcal{F}$ with $\{\xi_j(t)\}_{i=1}^n$ then we can arrange it so that if $j\in\{1,\dots,l\}$ then

$$\begin{aligned}
\xi_j(t) &= t\rho_j - t^2(Y\Lambda_\tau^{-1}Y^*)_{(j,j)} + \mathcal{O}(t^3) \\
&= t\rho_j - t^2\sum_{p=1}^{n-l}\frac{|Y_{(j,p)}|^2}{\tau_p} = t\langle \mathcal{F}\boldsymbol{e}_j, \boldsymbol{e}_j\rangle - t^2\sum_{p=l}^{n}\frac{|\langle \mathcal{F}\boldsymbol{e}_p, \boldsymbol{e}_j\rangle|^2}{\langle \mathcal{A}\boldsymbol{e}_p, \boldsymbol{e}_p\rangle} + \mathcal{O}(t^3) \\
&= t\langle \mathcal{F}\boldsymbol{e}_j, \boldsymbol{e}_j\rangle - t^2\sum_{p:\alpha_p\neq 0}\frac{|\langle \mathcal{F}\boldsymbol{e}_p, \boldsymbol{e}_j\rangle|^2}{\alpha_p} + \mathcal{O}(t^3).
\end{aligned}$$

This proves that the l eigenvalues of $\mathcal{A}+t\mathcal{F}$ which converge to 0 as $t\to 0$ can be written in the form stipulated by the theorem.

Now consider an arbitrary Hermitian operator $\mathcal{A}$ in $\mathcal{L}(\mathscr{H}_n)$ with eigenvalues $\alpha_1,\dots,\alpha_n$ and an associated basis of eigenvectors $\{\boldsymbol{e}_j\}_{j=1}^n$ chosen such that the matrix representation of $\mathcal{F}$ with respect to this basis has block-wise decreasing diagonal elements. We define $\mathcal{A}' = \mathcal{A} - \alpha_i I$ so that $\mathcal{A}'$ has a non-trivial kernel with $\dim\ker\mathcal{A}' = l > 0$ and denote the eigenvalues of $\mathcal{A}'$ with $\alpha_1',\dots,\alpha_n'$ where $\alpha_j' = \alpha_j - \alpha_i$. Since the eigenspaces of $\mathcal{A}$ and $\mathcal{A}'$ agree we find that $\mathscr{F}(\mathcal{A}) = \mathscr{F}(\mathcal{A}')$. The previous case proves that the eigenvalues of $\mathcal{A}'+t\mathcal{F}$ denoted $\{\xi_j'(t)\}_{j=1}^n$

can be ordered so that if $j \in \{1, \dots, l\}$ then

$$\xi'_j(t) = t\langle \mathcal{F}\boldsymbol{e}_j, \boldsymbol{e}_j\rangle - t^2 \sum_{p:\alpha'_p \neq 0} \frac{|\langle \mathcal{F}\boldsymbol{e}_p, \boldsymbol{e}_j\rangle|^2}{\alpha'_p} + \mathcal{O}(t^3).$$

Since the eigenvectors of $\mathcal{A} + t\mathcal{F}$ are simply $\xi'_j(t) + \alpha_i$ for $j \in \{1, \dots, n\}$ we see that the eigenvalues of $\mathcal{A} + t\mathcal{F}$ denoted $\{\xi_j(t)\}_{j=1}^n$ can be ordered so that if $j \in \{1, \dots, l\}$ then

$$\xi_j(t) = \alpha_i + t\langle \mathcal{F}\boldsymbol{e}_j, \boldsymbol{e}_j\rangle - t^2 \sum_{p:\alpha_p \neq \alpha_i} \frac{|\langle \mathcal{F}\boldsymbol{e}_p, \boldsymbol{e}_j\rangle|^2}{\alpha_p - \alpha_i} + \mathcal{O}(t^3)$$

which was precisely the formula we were supposed to show. □

Since the formula exists even without the restriction that $\mathcal{F} \in \mathscr{F}(\mathcal{A})$ it seems natural to believe that this result should be valid for all perturbations $\mathcal{A} + t\mathcal{F}$. For instance, since the set $\mathscr{F}(\mathcal{A})$ is dense in the space of Hermitian operators in $\mathcal{L}(\mathscr{H}_n)$ we could pick a sequence $\mathcal{F}_k \in \mathscr{F}(\mathcal{A})$ which converges to $\mathcal{F}$ and then conclude that the eigenvalues of $\mathcal{A} + t\mathcal{F}$ as limits of $\mathcal{A} + t\mathcal{F}_n$ would satisfy the formula. However here the reasoning breaks down since the error term incorporates the bad behaviour of the matrix N as the distance between the eigenvalues of Λ_ρ decreases. We have not been able to find another method using matrix decompositions which allows us to lift the restriction that $\mathcal{F} \in \mathscr{F}(\mathcal{A})$, however, using the method described by Hilbert and Courant we will show that the formula extends to every Hermitian $\mathcal{F}$.

3.4 First Order Expansion of the Eigenvectors

Having derived a second order expansion of the eigenvalues of perturbations of the form $\mathcal{A}+t\mathcal{F}$ where $\mathcal{A}$ is Hermitian and $\mathcal{F} \in \mathscr{F}(\mathcal{A})$ we now turn our attention to the eigenvectors. In particular we want to prove the first order expansion given in Theorem 1.1 which formulated in terms of general n-dimensional Hilbert spaces becomes:

Theorem 3.6. *Let $\mathcal{A}$ and $\mathcal{F}$ be Hermitian operators in $\mathcal{L}(\mathscr{H}_n)$ with the additional requirement that $\mathcal{F} \in \mathscr{F}(\mathcal{A})$. If $\{\boldsymbol{e}_j\}_{j=1}^n$ denotes an orthonormal basis of eigenvectors of $\mathcal{A}$ with corresponding eigenvalues $\{\alpha_j\}_{j=1}^n$ then there are orthonormal eigenvectors $\{\boldsymbol{u}_j(t)\}_{j=1}^n$ of $\mathcal{A} + t\mathcal{F}$ which can be ordered so that they satisfy*

$$\boldsymbol{u}_j(t) = \boldsymbol{e}_j + t\left(\sum_{\substack{p:\alpha_p=\alpha_j \\ p\neq j}} \frac{\langle(\mathcal{A}-\alpha_j I)\mathcal{F}\boldsymbol{e}_j, \mathcal{F}\boldsymbol{e}_p\rangle}{\langle \mathcal{F}\boldsymbol{e}_p, \boldsymbol{e}_p\rangle - \langle \mathcal{F}\boldsymbol{e}_j, \boldsymbol{e}_j\rangle}\boldsymbol{e}_p - \sum_{p:\alpha_p\neq\alpha_j} \frac{\langle \mathcal{F}\boldsymbol{e}_j, \boldsymbol{e}_p\rangle}{\alpha_p - \alpha_j}\boldsymbol{e}_p\right) + \mathcal{O}(t^2). \qquad (13)$$

Our method of proof is heavily inspired by a proof of a similar result in [10]. There the result is attained for global perturbations of the form $\mathcal{A} + \mathcal{E}$ where $\mathcal{E}$ varies. We begin by translating the problem into matrix form. Having chosen a basis $\{\boldsymbol{e}_j\}_{j=1}^n$ of eigenvectors of $\mathcal{A}$ such that $\mathcal{F}$ has a block-wise decreasing diagonal in accordance with Definition 3.2 we obtain that the matrix representation of our perturbation $\mathcal{A}+t\mathcal{F}$ becomes $\Lambda_{\mathcal{A}}+tF$ where $\Lambda_{\mathcal{A}} = \text{diag}\,(\alpha_1,\dots,\alpha_n)$. We define the matrix $M \in \mathcal{M}_{n\times n}$ by

$$M_{(i,j)} = \begin{cases} \frac{1}{\alpha_i-\alpha_j}, & \alpha_i \neq \alpha_j \\ 0, & \text{otherwise.} \end{cases}$$

If we recall the definitions of the matrices N and $V_{ap}(t)$ defined in Equation (9) and Equation (13) respectively then we see that the matrix whose jth column correspond to the coordinates of the vector $\boldsymbol{u}_j(t)$ given in Equation (13) with respect to the basis $\{\boldsymbol{e}_j\}_{j=1}^n$ becomes

$$U(t) = V_{ap}(t) - tM \circ F + \mathcal{O}(t^2) = I + t(N - M\circ F) + \mathcal{O}(t^2).$$

We again choose to begin studying the case where $\mathcal{A}$ has a non-trivial kernel of dimension l. The matrix representation of $\mathcal{A}+t\mathcal{F}$ can then be chosen as before to be of the form

$$\begin{pmatrix} 0 & 0 \\ 0 & \Lambda_\tau \end{pmatrix} + t\begin{pmatrix} \Lambda_\rho & Y \\ Y^* & Z \end{pmatrix}.$$

Reusing the notation of Theorem 3.5 we denote the eigenvalues of $\mathcal{A}+t\mathcal{F}$ with $\{\xi_j(t)\}_{j=1}^n$ where the same theorem implies that

$$\xi_j(t) = \alpha_j + t\langle \mathcal{F}\boldsymbol{e}_j, \boldsymbol{e}_j\rangle - t^2 \sum_{p:\alpha_p\neq\alpha_j} \frac{|\langle \mathcal{F}\boldsymbol{e}_j, \boldsymbol{e}_p\rangle|^2}{\alpha_p-\alpha_j} + \mathcal{O}(t^3), \quad j \in \{1,\dots,n\}.$$

Denoting

$$\rho_j = \langle \mathcal{F}\boldsymbol{e}_j, \boldsymbol{e}_j\rangle$$

and

$$\beta_j = \sum_{p:\alpha_p\neq\alpha_j} \frac{|\langle \mathcal{F}\boldsymbol{e}_j, \boldsymbol{e}_p\rangle|^2}{\alpha_p-\alpha_j}$$

we can express the eigenvalues $\xi_j(t)$ as

$$\xi_j(t) = \alpha_j + t\rho_j - t^2\beta_j + \mathcal{O}(t^3).$$

In particular we see that if $j \in \{1,\dots,l\}$ then $\alpha_j = 0$ and therefore $\xi_j(t) = \mathcal{O}(t)$. Furthermore since $\mathcal{F} \in \mathscr{F}(\mathcal{A})$ we see that the ρ_j are all distinct, implying that the eigenvalues of $\mathcal{A}+t\mathcal{F}$ will be simple if $t > 0$ is small enough. Thus for a fix $q \in \{1,\dots,l\}$ we find that

$$\mathcal{A} + t\mathcal{F} - \xi_q(t)I$$

will have a 1-dimensional kernel whose non-zero elements are eigenvectors of $\mathcal{A}+t\mathcal{F}$ corresponding to $\xi_q(t)$. We define the matrix $\Psi(t)$ by

$$\Psi(t) := (V_{ap}(t))^{-1}(A + tF - \xi_q(t)I)V_{ap}(t)$$

It then follows from equation (11) that $\Psi(t)$ is given by

$$\begin{aligned}
&(V_{ap}(t))^{-1}(A + tF - \xi_q(t)I)V_{ap}(t) \\
&= \begin{pmatrix} t\Lambda_\rho + t^2((Y\Lambda_\tau^{-1}Y^*) - I \circ (Y\Lambda_\tau^{-1}Y^*)) + \mathcal{O}(t^3) & tY + \mathcal{O}(t^2) \\ tY^* + \mathcal{O}(t^2) & \Lambda_\tau + tZ + \mathcal{O}(t^2) \end{pmatrix} - \xi_q(t)I \\
&= \begin{pmatrix} t\Lambda_\rho + t^2((Y\Lambda_\tau^{-1}Y^*) - I \circ (Y\Lambda_\tau^{-1}Y^*)) - \xi_q(t)I + \mathcal{O}(t^3) & tY + \mathcal{O}(t^2) \\ tY^* + \mathcal{O}(t^2) & \Lambda_\tau - \xi_q(t)I + tZ + \mathcal{O}(t^2) \end{pmatrix}.
\end{aligned}$$

If we write M in its block form:

$$M = \begin{pmatrix} 0 & M_{12} \\ M_{21} & M_{22} \end{pmatrix}$$

then

$$I - tM \circ F = \begin{pmatrix} I & tM_{12} \circ F_{12} \\ tM_{21} \circ F_{21} & I + tM_{22} \circ F_{22} \end{pmatrix}.$$

Since we assumed that $\alpha_1 = \cdots = \alpha_l = 0$ the blocks M_{12} and M_{21} are given by

$$M_{12} = \begin{bmatrix} \frac{1}{\alpha_{l+1}} & \cdots & \frac{1}{\alpha_n} \\ \vdots & \ddots & \vdots \\ \frac{1}{\alpha_{l+1}} & \cdots & \frac{1}{\alpha_n} \end{bmatrix} = -M_{21}^t.$$

We can therefore write

$$M_{12} \circ F_{12} = M_{12} \circ Y = Y\Lambda_\tau^{-1}$$

and

$$M_{21} \circ F_{21} = -M_{12}^t \circ Y^* = -(Y\Lambda_\tau^{-1})^* = -\Lambda_\tau^{-1}Y^*.$$

In total this implies that

$$I - tM \circ F = \begin{pmatrix} I & tY\Lambda_\tau^{-1} \\ -t\Lambda_\tau^{-1}Y^* & I + tM_{22} \circ Z \end{pmatrix}.$$

Now we again want to apply the theory of Schur complements and in particular Lemma 2.14. To that end we focus our attention on the matrix

$$U_{ap}(t) = \begin{pmatrix} I & 0 \\ -t(\Lambda_\tau + tZ)^{-1}Y^* & I \end{pmatrix} = \begin{pmatrix} I & 0 \\ -t\Lambda_\tau Y^* + \mathcal{O}(t^2) & I \end{pmatrix}$$

noting that the first l columns of $U_{ap}(t)$ differ by $\mathcal{O}(t^2)$ from the first l columns of $I - tM \circ F$. It therefore suffices to prove that the q:th column of $U_{ap}(t)$ given by $U_{ap}(t)e_q$ lies within $\mathcal{O}(t^2)$ from an actual non-zero vector in the kernel. Here the vectors $\{e_j\}_{j=1}^n$ denote the canonical basis for $\mathbb{C}^n$ which should not be confused with the basis $\{\boldsymbol{e}_j\}_{j=1}^n$ of $\mathscr{H}_n$. Applying $U_{ap}(t)$ to the decomposition $\Psi(t)$ leaves us with

$$\begin{aligned}\Psi(t)U_{ap}(t) &= \begin{pmatrix} t\Lambda_\rho - \xi_q(t)I - t^2[I\circ(Y\Lambda_\tau^{-1}Y^*)] + \mathcal{O}(t^3) & tY + \mathcal{O}(t^2) \\ t\xi_q(t)(\Lambda_\tau + tZ)^{-1}Y^* & \Lambda_\tau - \xi_q(t)I + tZ + \mathcal{O}(t^2)\end{pmatrix} \\ &= \begin{pmatrix} t\Lambda_\rho - \xi_q(t)I - t^2[I\circ(Y\Lambda_\tau^{-1}Y^*)] + \mathcal{O}(t^3) & \mathcal{O}(t) \\ \mathcal{O}(t^2) & \Lambda_\tau + \mathcal{O}(t)\end{pmatrix}.\end{aligned} \tag{14}$$

We will use a method described in [10] to find a non-zero vector for the matrix in Equation (14).

Lemma 3.7. *Let $A \in \mathcal{M}_{n\times n}$. If $\det A = 0$ then for a fix $i \in \{1,\dots,n\}$ a vector in $\ker A$ is given by $\boldsymbol{w} = (w_1,\dots,w_n)$ where*

$$w_j = (-1)^{i+j}\det A^{(i,j)}.$$

Remark. *The vector $\boldsymbol{w}$ is obtained by letting its jth component be the determinant of the matrix*

$$\begin{pmatrix} A_{(1,1)} & \cdots & A_{(1,j-1)} & A_{(1,j)} & A_{(1,j+1)} & \cdots & A_{(1,n)} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ A_{(i-1,1)} & \cdots & A_{(i-1,j-1)} & A_{(i-1,j)} & A_{(i-1,j+1)} & \cdots & A_{(i-1,n)} \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ A_{(i+1,1)} & \cdots & A_{(i+1,j-1)} & A_{(i+1,j)} & A_{(i+1,j+1)} & \cdots & A_{(i+1,n)} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ A_{(n,1)} & \cdots & A_{(n,j-1)} & A_{(n,j)} & A_{(n,j+1)} & \cdots & A_{(n,n)} \end{pmatrix}.$$

thus we interchange the ith row of A with the vector e_j.

Proof. Let $\boldsymbol{w} = (w_1,\dots,w_n)$ where $w_j = (-1)^{i+j}\det A^{(i,j)}$. We let A_k denote the kth row of A. If $j \neq i$ the jth row of $A\boldsymbol{w}$ is given by

$$A_j\boldsymbol{w} = \sum_{k=1}^n A_{(j,k)}w_k = \sum_{k=1}^n A_{(j,k)}(-1)^{i+k}\det A^{(i,k)}.$$

We assert that this implies that $A_j\boldsymbol{w} = 0$. Indeed $\sum_{k=1}^n A_{(j,k)}(-1)^{i+k}\det A^{(i,k)}$ is the determinant of the matrix obtained from A by replacing its ith row with its jth row, and since any matrix with two identical rows has determinant value 0 it now follows that $A_j\boldsymbol{w} = 0$. For the remaining case we consider the ith row of $A\boldsymbol{w}$ given by

$$A_i\boldsymbol{w} = \sum_{k=1}^n A_{(i,k)}w_k = \sum_{k=1}^n A_{(i,k)}(-1)^{i+k}\det A^{(i,k)} = \det A = 0.$$

With this we have established that $\boldsymbol{w} \in \ker A$. $\square$

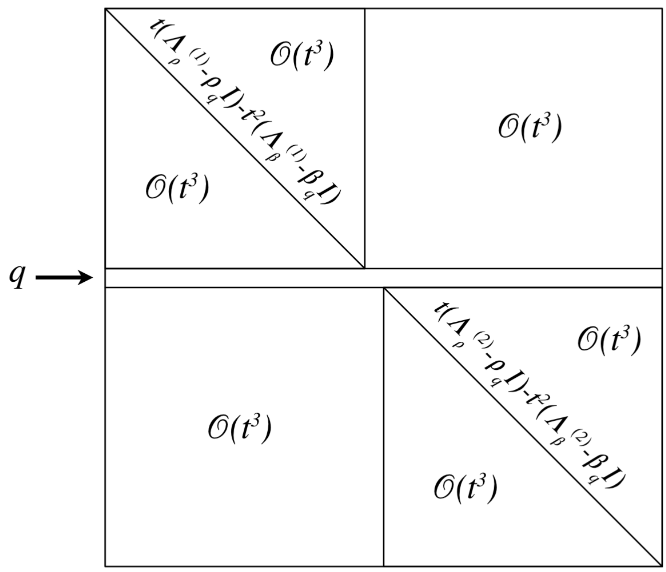

Figure 2: A visual representation of the matrix $\Xi_j(t)$.

Lemma 3.7 does not imply that the vector obtained is non-zero. For instance if the first row of a matrix is zero everywhere but the process is applied to some other row than the first, corresponding to the ith row in the lemma, then the vector $\boldsymbol{w}$ obtained is just the zero vector.

We let $X_j(t)$ denote the matrix obtained from $\Psi(t)U_{ap}(t)$ with the addition that the q:th row equals e_j in accordance with Lemma 3.7. Since $\Lambda_\tau + \mathcal{O}(t)$ is invertible we may consider the Schur complement of the block $\Lambda_\tau + \mathcal{O}(t)$ of the matrix $X_j(t)$ which we denote as $\Xi_j(t)$. To ease with notation we define

$$\Lambda_\rho^{(1)} = \text{diag}\,(\rho_1, \ldots, \rho_{q-1}), \quad \Lambda_\rho^{(2)} = \text{diag}\,(\rho_{q+1}, \ldots, \rho_l)$$
$$\Lambda_\beta^{(1)} = \text{diag}\,(\beta_1, \ldots, \beta_{q-1}), \quad \Lambda_\beta^{(2)} = \text{diag}\,(\beta_{q+1}, \ldots, \beta_l).$$

By definition of the Schur complement it follows that $\Xi_j(t)$ is given by

$$\Xi_j(t) = \begin{pmatrix} t(\Lambda_\rho^{(1)} - \rho_q I) - t^2(\Lambda_\beta^{(1)} - \beta_q I) + \mathcal{O}(t^3) & \mathcal{O}(t^3) \\ * & * \\ \mathcal{O}(t^3) & t(\Lambda_\rho^{(2)} - \rho_q I) - t^2(\Lambda_\beta^{(2)} - \beta_q I) + \mathcal{O}(t^3) \end{pmatrix}.$$

The qth row marked with asterisk is of one of two forms: If $j \in \{1, \ldots, l\}$ then it simply is 0 on all position except for the jth column where it is 1. If $j \in \{l+1, \ldots, n\}$ then it is $\mathcal{O}(t^2)$ at all positions. Lemma 2.13 implies that

$$\det X_j(t) = \det(\Lambda_\tau + \mathcal{O}(t)) \det \Xi_j(t)$$

Now consider the formula for the determinant given in equation (1)

$$\det \Xi_j = \sum_{p=1}^{l!} \text{sgn}\,(\pi_p) \prod_{k=1}^{l} (\Xi_j)_{(k,\pi_p(k))}.$$

We see that if $j \in \{1, \ldots, l\}$ then the permutation in the sum must satisfy $\pi_p(i) = j$ in order for the pth summand to be non-vanishing. This implies that if $j = q$ then the determinant will be of the form

$$\det \Xi_q = t^{l-1} \prod_{\substack{k=1 \\ k \neq q}}^{l} (\rho_k - \rho_q) + \mathcal{O}(t^l).$$

To see this note that the product of all diagonal entries is of the form

$$t^{l-1} \prod_{\substack{k=1 \\ k \neq q}}^{l} (\rho_k - \rho_q) + \mathcal{O}(t^l)$$

and every summand of the determinant where π_k is not the identity will be of the form $\mathcal{O}(t^{l+1})$. This means that we can find δ such that

$$|\det \Xi_q| \geq \frac{|t^{l-1}|}{2} \prod_{\substack{k=1 \\ k \neq q}}^{l} |\rho_k - \rho_q|$$

if $0 < |t| < \delta$.

If $j \in \{1, \ldots, q-1, q+1, \ldots, l\}$ then every summand of the determinant must get contributions from at least two off-diagonal elements which implies that

$$\det \Xi_j = \mathcal{O}(t^{l+3}).$$

Finally if $j \in \{l+1, \ldots, n\}$ then every summand of the determinant must contain at least one element from the qth row which is $\mathcal{O}(t^2)$. If all other choices are from the diagonal then we find that the summand is $\mathcal{O}(t^2 \cdot t^{l-1}) = \mathcal{O}(t^{l+1})$. All other summands of the determinant are of higher or equal degree. We conclude that in this case

$$\det \Xi_j(t) = \mathcal{O}(t^{l+1}).$$

Finally focusing our attention on $\det(\Lambda_\tau + \mathcal{O}(t))$ we use the fact that $A \mapsto \det A$ is continuous to conclude that $\det(\Lambda_\tau + \mathcal{O}(t)) \to \det \Lambda_\tau \neq 0$ as $t \to 0$. In particular $|\det(\Lambda_\tau + \mathcal{O}(t))|$ is bounded below by $\frac{|\det \Lambda_\tau|}{2}$ if $0 < |t| < \delta'$ for some $\delta' \in (0, \delta)$. If we define

$$\boldsymbol{w}(t) = (w_1(t), \ldots, w_n(t))$$

where $w_j(t) = \det X_j(t)$ then $\boldsymbol{w}(t)$ satisfies

$$|w_q(t)| \geq \left(\frac{|\det \Lambda_\tau|}{4} \prod_{\substack{k=1 \\ k \neq q}}^{l} |\rho_k - \rho_q| \right) |t|^{l-1}$$

while

$$w_j(t) = \mathcal{O}(t^{l+1})$$

for $j \in \{1, \ldots, q-1, q+1, \ldots, n\}$. If we define $\boldsymbol{v}(t) = (v_1(t), \ldots, v_n(t))$ in a neighborhood of the origin via

$$\boldsymbol{v}(t) = \frac{\boldsymbol{w}(t)}{w_q(t)}$$

then

$$v_j(t) = \begin{cases} 1, & j = q \\ \mathcal{O}(t^2), & j \neq q \end{cases}$$

and therefore $\boldsymbol{v}(t) = e_q + \mathcal{O}(t^2)$. We have thus shown that a non-zero vector in the kernel of $A + tF - \xi_q I$ is given by

$$\begin{aligned} V_{ap}(t)U_{ap}(t)(e_q + \mathcal{O}(t^2)) &= (I + tN)(I - tM \circ F)e_q + \mathcal{O}(t^2) \\ &= (I + t(N - M \circ F))e_q + \mathcal{O}(t^2). \end{aligned}$$

Normalizing we find that a non-zero vector in the kernel of $A + tF - \xi_q(t)I$ of unit length is given by

$$\frac{(I + t(N - M \circ F))e_q}{\sqrt{1 + \mathcal{O}(t^2)}} + \mathcal{O}(t^2) = (I + t(N - M \circ F))e_q + \mathcal{O}(t^2).$$

This $n \times 1$ matrix corresponds to the vector

$$\boldsymbol{e}_q + t \left(\sum_{\substack{p:\alpha_p = 0 \\ p \neq q}} \frac{\langle (\mathcal{A} - \alpha_q I)\mathcal{F}\boldsymbol{e}_q, \mathcal{F}\boldsymbol{e}_p \rangle}{\langle \mathcal{F}\boldsymbol{e}_p, \boldsymbol{e}_p \rangle - \langle \mathcal{F}\boldsymbol{e}_q, \boldsymbol{e}_q \rangle} \boldsymbol{e}_p - \sum_{p:\alpha_p \neq 0} \frac{\langle \mathcal{F}\boldsymbol{e}_q, \boldsymbol{e}_p \rangle}{\alpha_p} \boldsymbol{e}_p \right) + \mathcal{O}(t^2)$$

in $\mathscr{H}_n$ which we denote with $\boldsymbol{u}_q$. We now complete the proof for the general case.

Proof of Theorem 3.6. We consider an arbitrary Hermitian operator $\mathcal{A}$ in $\mathcal{L}(\mathscr{H}_n)$. The operator $\mathcal{A}'$ defined via $\mathcal{A}' = \mathcal{A} - \alpha_j I$ has the same eigenspaces as $\mathcal{A}$ and thus $\mathscr{F}(\mathcal{A}) = \mathscr{F}(\mathcal{A}')$. We denote the eigenvalues of $\mathcal{A}' + t\mathcal{F}$ with $\{\xi_i'(t)\}_{i=1}^n$ where the l first indices denotes the eigenvalues which tend to 0 as $t \to 0$. By our previous reasoning we gather that a normalized eigenvector of $\mathcal{A}' + t\mathcal{F}$ corresponding to $\xi_q(t)'$ denoted $\boldsymbol{u}_q'(t)$ satisfies

$$\boldsymbol{u}_q'(t) = \boldsymbol{e}_q + t \left(\sum_{\substack{p:\alpha_p' = 0 \\ p \neq q}} \frac{\langle (\mathcal{A}' - \alpha_q' I)\mathcal{F}\boldsymbol{e}_q, \mathcal{F}\boldsymbol{e}_p \rangle}{\langle \mathcal{F}\boldsymbol{e}_p, \boldsymbol{e}_p \rangle - \langle \mathcal{F}\boldsymbol{e}_q, \boldsymbol{e}_q \rangle} \boldsymbol{e}_p - \sum_{p:\alpha_p' \neq 0} \frac{\langle \mathcal{F}\boldsymbol{e}_q, \boldsymbol{e}_p \rangle}{\alpha_p'} \boldsymbol{e}_p \right) + \mathcal{O}(t^2).$$

Now the eigenvectors of $\mathcal{A}' + t\mathcal{F}$ and $\mathcal{A} + t\mathcal{F}$ agree. Since $\alpha_p' = \alpha_p - \alpha_j$ and $\mathcal{A}' = \mathcal{A} + \alpha_j I = \mathcal{A} + \alpha_q I$ we find that a normalized eigenvector of $\mathcal{A} + t\mathcal{F}$ corresponding to $\xi_q(t)$ is given by

$$\boldsymbol{u}_q(t) = \boldsymbol{e}_q + t \left(\sum_{\substack{p:\alpha_p = \alpha_q \\ p \neq q}} \frac{\langle (\mathcal{A} - \alpha_q I)\mathcal{F}\boldsymbol{e}_q, \mathcal{F}\boldsymbol{e}_p \rangle}{\langle \mathcal{F}\boldsymbol{e}_p, \boldsymbol{e}_p \rangle - \langle \mathcal{F}\boldsymbol{e}_q, \boldsymbol{e}_q \rangle} \boldsymbol{e}_p - \sum_{p:\alpha_p \neq \alpha_q} \frac{\langle \mathcal{F}\boldsymbol{e}_q, \boldsymbol{e}_p \rangle}{\alpha_p - \alpha_q} \boldsymbol{e}_p \right) + \mathcal{O}(t^2).$$

Doing this for every q we gather the collection of normalized eigenvectors of $\mathcal{A} + t\mathcal{F}$ denoted $\{\boldsymbol{u}_q\}_{q=1}^n$. Since the eigenvalues of $\mathcal{A} + t\mathcal{F}$ are simple if $0 < |t| < \delta$ for some $\delta > 0$ it follows that these eigenvectors are orthogonal. This completes our proof. ☐

3.5 The Proof by Hilbert and Courant

As we mentioned in our introduction both Theorem 3.6 and Theorem 3.5 were proven by David Hilbert and Richard Courant, see Chapter V in [12]. Their approach however relied on the holomorphic dependence of the eigenvalues and eigenvectors on the perturbation parameter t and not on matrix decompositions like those we have presented. These holomorphic properties were first proven by Franz Rellich (see Chapter 9.4 of [15]) and the proofs rely on the fact that the roots of a polynomial may be expanded in a so called Puiseaux series.

Theorem 3.8 (Rellich)**.** *Let $\mathcal{A} : \mathbb{R} \mapsto \mathcal{L}(\mathscr{H}_n)$ be a real analytic mapping where $\mathcal{A}(t)$ is Hermitian for every t. Then then the eigenvalues are holomorphic functions of the perturbation parameter t. Furthermore there is a complete orthonormal set of eigenvectors of $\mathcal{A}(t)$ which depend analytically on t.*

With this in hand we give the proof of a slight generalization of Theorem 1.1 supplied by Hilbert and Courant in [12].

Theorem 3.9 (Hilbert-Courant)**.** *Let $\operatorname{diag}(\alpha_1, \ldots, \alpha_n) = \Lambda_{\mathcal{A}}$ be an $n \times n$ Hermitian matrix such that the diagonal is ordered non-increasingly. For Hermitian matrices F which satisfy*

$$F_{(i,j)} = 0, \quad i \neq j,\ \alpha_i = \alpha_j$$

the eigenvalues of $\mathcal{A} + t\mathcal{F}$ denoted $\{\xi_j(t)\}_{j=1}^n$ can be ordered so that they satisfy

$$\xi_j(t) = \alpha_j + tF_{(j,j)} + t^2 \sum_{i:\alpha_i \neq \alpha_j} \frac{|F_{(i,j)}|^2}{\alpha_i - \alpha_j} + \mathcal{O}(t^3).$$

If additionally

$$F_{(i,i)} > F_{(j,j)}, \quad i \neq j,\ \alpha_i = \alpha_j$$

holds then there is an ordering of orthonormal eigenvectors of $\Lambda_{\mathcal{A}} + tF$, denoted $\{\boldsymbol{u}_j(t)\}_{j=1}^n$, which satisfy

$$\boldsymbol{u}_j(t) = \boldsymbol{e}_j + t\boldsymbol{v}_j + \mathcal{O}(t^2)$$

where $\{\boldsymbol{e}_j\}_{j=1}^n$ denotes the canonical basis of $\mathbb{C}^n$ and

$$\boldsymbol{v}_j = \sum_{\substack{p:\alpha_p = \alpha_j \\ p \neq j}} \frac{(F^*(\Lambda_\alpha - \alpha_j I)^\dagger F)_{(p,j)}}{F_{(p,p)} - F_{(j,j)}} \boldsymbol{e}_p - \sum_{p:\alpha_p \neq \alpha_j} \frac{F_{(p,j)}}{\alpha_p - \alpha_j} \boldsymbol{e}_p.$$

Here $\dagger$ denotes the Moore-Penrose inverse.

Proof. We consider the perturbation $\Lambda_{\mathcal{A}} + tF$ where $\Lambda_{\mathcal{A}} = \operatorname{diag}(\alpha_1, \ldots, \alpha_n)$ is a diagonal matrix and F is block-wise diagonal. We let $\{\boldsymbol{e}_j\}_{j=1}^n$ denote the canonical basis of $\mathbb{C}^n$. By Theorem 3.8

we can express the eigenvalues and eigenvectors of $\Lambda_{\mathcal{A}} + tF$, denoted $\{\xi_j(t)\}_{j=1}^n$ and $\{\boldsymbol{u}_j(t)\}_{j=1}^n$ respectively as power series:

$$\begin{aligned}\xi_j(t) &= \alpha_j + t\rho_j + t^2\beta_j + \mathcal{O}(t^3)\\ \boldsymbol{u}_j(t) &= \boldsymbol{e}_j + t\boldsymbol{v}_j + t^2\boldsymbol{w}_j + \mathcal{O}(t^2).\end{aligned}$$

Theorem 3.8 further implies that we can assume that $\|\boldsymbol{u}_j(t)\| \equiv 1$. In component form we denote

$$\begin{aligned}\boldsymbol{v}_j &= (v_{j,1}, \dots, v_{j,n}),\\ \boldsymbol{w}_j &= (w_{j,1}, \dots, w_{j,n}).\end{aligned}$$

Applying the perturbation to $\boldsymbol{u}_j(t)$ we obtain

$$\begin{aligned}(\Lambda_{\mathcal{A}} + tF)\boldsymbol{u}_j(t) &= (\Lambda_{\mathcal{A}} + tF)(\boldsymbol{e}_j + t\boldsymbol{v}_j + t^2\boldsymbol{w}_j + \mathcal{O}(t^3))\\ &= \Lambda_{\mathcal{A}}\boldsymbol{e}_j + t(\Lambda_{\mathcal{A}}\boldsymbol{v}_j + F\boldsymbol{e}_j) + t^2(\Lambda_\alpha\boldsymbol{w}_j + F\boldsymbol{v}_j) + \mathcal{O}(t^3).\end{aligned}$$

Since $\boldsymbol{u}_j$ is the eigenvector of $\Lambda_{\mathcal{A}} + tF$ corresponding to $\xi_j(t)$ we also know that

$$\begin{aligned}(\Lambda_{\mathcal{A}} + tF)\boldsymbol{u}_j(t) = \xi_j(t)\boldsymbol{u}_j(t) &= (\alpha_j + t\rho_j + t^2\beta_j + \mathcal{O}(t^3))(\boldsymbol{e}_j + t\boldsymbol{v}_j + t^2\boldsymbol{w}_j + \mathcal{O}(t^2))\\ &= \alpha_j\boldsymbol{e}_j + t(\rho_j\boldsymbol{e}_j + \alpha_j\boldsymbol{v}_j) + t^2(\beta_j\boldsymbol{e}_j + \rho_j\boldsymbol{v}_j + \alpha_j\boldsymbol{w}_j) + \mathcal{O}(t^3).\end{aligned}$$

Equating these two expressions we conclude that

$$\Lambda_{\mathcal{A}}\boldsymbol{e}_j = \alpha_j\boldsymbol{e}_j \tag{15}$$

$$\Lambda_{\mathcal{A}}\boldsymbol{v}_j + F\boldsymbol{e}_j = \rho_j\boldsymbol{e}_j + \alpha_j\boldsymbol{v}_j \tag{16}$$

$$\Lambda_\alpha\boldsymbol{w}_j + F\boldsymbol{v}_j = \beta_j\boldsymbol{e}_j + \rho_j\boldsymbol{v}_j + \alpha_j\boldsymbol{w}_j. \tag{17}$$

Equation (15) is a consequence of the definition of $\Lambda_{\mathcal{A}}$ and does not generate any new information. Consider the ith row of the vector in Equation (16). On the left-hand side we have

$$(\Lambda_{\mathcal{A}}\boldsymbol{v}_j)_i + (F\boldsymbol{e}_j)_i = \alpha_i v_{j,i} + F_{(i,j)}$$

while the right-hand side equates to

$$(\rho_j\boldsymbol{e}_j)_i + \alpha_j(\boldsymbol{v}_j)_i = \delta_{ij}\rho_j + \alpha_j v_{j,i}.$$

Here δ_{ij} denotes the Kronecker delta function. Setting $i = j$ we see that $\rho_j = F_{(j,j)}$ which determines the first order term in the eigenvalue expansion. If $\alpha_i \neq \alpha_j$ then the equation implies that

$$v_{j,i} = \frac{F_{(i,j)}}{\alpha_j - \alpha_i}.$$

The case where $i \neq j$ but $\alpha_i = \alpha_j$ does not present any new information. We now turn our attention to Equation (17). Again studying the ith row we find that

$$\alpha_i w_{j,i} + (F\boldsymbol{v}_j)_i = \beta_j\delta_{ij} + F_{(j,j)}v_{j,i} + \alpha_j w_{j,i}.$$

Setting $i = j$ we get the equation

$$\begin{aligned}\beta_j = (F\boldsymbol{v}_j)_j - F_{(j,j)}v_{j,j} &= \sum_{k:\alpha_k\neq\alpha_j} F_{(j,k)}v_{j,k} \\ &= \sum_{k:\alpha_k\neq\alpha_j} \frac{F_{(j,k)}F_{(k,j)}}{\alpha_j-\alpha_k} = \sum_{k:\alpha_k\neq\alpha_j} \frac{|F_{(j,k)}|^2}{\alpha_j-\alpha_k}.\end{aligned}$$

This determines the coefficients of the eigenvalue expansion.

Assume now that $F_{(i,i)} > F_{(j,j)}$ holds whenever $i \neq j$ and $\alpha_i = \alpha_j$. We consider indices i such that $i \neq j$ but $\alpha_i = \alpha_j$ then Equation (17) implies that

$$\sum_{k=1}^{n} F_{(i,k)}v_{j,k} = F_{(j,j)}v_{j,i}.$$

Since $F_{(i,k)} = 0$ if $i \neq k$ but $\alpha_i = \alpha_k$ we find that the sum equals

$$F_{(i,i)}v_{j,i} + \sum_{k:\alpha_k\neq\alpha_j} F_{(i,k)}v_{j,k} = F_{(i,i)}v_{j,i} + \sum_{k:\alpha_k\neq\alpha_j} \frac{F_{(i,k)}F_{(k,j)}}{\alpha_j-\alpha_k}$$

and therefore we can conclude that

$$v_{j,i} = \frac{1}{F_{(i,i)} - F_{(j,j)}} \sum_{k:\alpha_k\neq\alpha_j} \frac{F_{(i,k)}F_{(k,j)}}{\alpha_j-\alpha_k}.$$

The only component of $\boldsymbol{v}_j$ which is not determined so far is the jth coordinate. We claim that we can choose it to be 0. Since $\boldsymbol{u}_j(t)$ is of constant modulus 1 it follows that

$$\begin{aligned}0 &= \frac{d}{dt}\langle \boldsymbol{u}_j(t), \boldsymbol{u}_j(t)\rangle = \langle \boldsymbol{u}_j'(t), \boldsymbol{u}_j(t)\rangle + \langle \boldsymbol{u}_j(t), \boldsymbol{u}_j'(t)\rangle \\ &= \langle \boldsymbol{v}_j + \mathcal{O}(t), \boldsymbol{e}_j + \mathcal{O}(t)\rangle + \langle \boldsymbol{e}_j + \mathcal{O}(t), \boldsymbol{v}_j + \mathcal{O}(t)\rangle \\ &= v_{j,j} + \overline{v_{j,j}} + \mathcal{O}(t) = 2\text{Re } v_{j,j} + \mathcal{O}(t) \Rightarrow \text{Re } v_{j,j} = 0.\end{aligned}$$

The imaginary component, however, is not well-defined. Any eigenvector may be multiplied by scalars while retaining its properties as an eigenvector. We choose to multiply $\boldsymbol{u}_j(t)$ with a unimodular constant of the form $e^{i\theta t} = 1 + i\theta t + \mathcal{O}(t^2)$, $\theta \in \mathbb{R}$. If we consider the component form of $\boldsymbol{u}_j(t)e^{i\theta t}$ we have

$$\begin{aligned}\boldsymbol{u}_j(t)e^{i\theta t} &= \left(\begin{pmatrix}0\\ \vdots\\ 1\\ \vdots\\ 0\end{pmatrix} + t\begin{pmatrix}v_{j,i}\\ \vdots\\ v_{j,j}\\ \vdots\\ v_{j,n}\end{pmatrix} + \mathcal{O}(t^2)\right)(1 + i\theta t + \mathcal{O}(t^2)) \\ &= \begin{pmatrix}0\\ \vdots\\ 1\\ \vdots\\ 0\end{pmatrix} + t\begin{pmatrix}v_{j,1}\\ \vdots\\ i\theta + v_{j,j}\\ \vdots\\ v_{j,n}\end{pmatrix} + \mathcal{O}(t^2)\end{aligned}$$

Since $v_{j,j}$ is imaginary we may choose $\theta = -\text{Im}\ v_{j,j}$ to obtain a new eigenvector such that the jth component of the first order term is 0. $\square$

It is clear that this proof is less technical than those we presented previously however the proof of Theorem 3.8 which we assumed to be true is a deep result.

4 Global Perturbations and Applications to Singular Values

Having studied perturbations of the form $\mathcal{A} + t\mathcal{F}$ where $\mathcal{A}$ and $\mathcal{F}$ are Hermitian operators in $\mathcal{L}(\mathscr{H}_n)$ we now turn our attention to global perturbations. We recall that these are perturbations of the form $\mathcal{A} + \mathcal{E}$ where $\mathcal{E}$ is a Hermitian operator. However we will now use Schur complements to derive two eigenvalue approximations of global perturbations which are analogous to the results of the previous section. While these results are proven in [10] we present new proofs which rely on the Schur complement and Gershgorin's Circle Theorem. We recall the definition of $\lambda_i(\mathcal{A})$ as the ith eigenvalue of the Hermitian operator $\mathcal{A}$ ordered non-increasingly and counting multiplicities.

4.1 A First Order Approximation of the Spectrum

Our first result gives an approximation of the eigenvalues of $\mathcal{A}+\mathcal{E}$ whose errors are of the form $\mathcal{O}(\|\mathcal{E}\|^2)$.

Theorem 4.1. *Let $\mathcal{A}$ and $\mathcal{E}$ be Hermitian operators in $\mathcal{L}(\mathscr{H}_n)$ and let α_0 denote an eigenvalue of $\mathcal{A}$ of multiplicity $l > 0$. If j_0 is the first occurrence of α_0 in the collection $\{\lambda_j(\mathcal{A})\}_{j=1}^n$ then*

$$\lambda_j(\mathcal{A} + \mathcal{E}) = \lambda_j(\mathcal{A}) + \lambda_{j-j_0+1}(\mathcal{E}_0) + \mathcal{O}(\|\mathcal{E}\|^2), \quad j \in \{j_0, \ldots, j_0 + l - 1\}$$

where $\mathcal{E}_0$ is the restriction of $\mathcal{P}_{\ker(\mathcal{A}-\alpha_0 I)}\mathcal{E}\mathcal{P}_{\ker(\mathcal{A}-\alpha_0 I)}$ to $\ker(\mathcal{A} - \alpha_0 I)$.

Proof. Assume first that $\alpha_0 = 0$, where $\dim\ker\mathcal{A} = l$ and let j_0 be the index of the first occurrence of 0 in $\{\lambda_j(\mathcal{A})\}_{j=1}^n$. If $\mathcal{A}$ and $\mathcal{E}$ are Hermitian operators in $\mathcal{L}(\mathscr{H}_n)$ then it follows from Lemma 3.3 that we can find a basis denoted $\{\boldsymbol{e}_j\}_{j=1}^n$ in which the matrix representation of $\mathcal{E}$, denoted E, is block-wise diagonal and such that the matrix representation of $\mathcal{A}$ is the diagonal matrix $\Lambda_{\mathcal{A}}$ where

$$\Lambda_{\mathcal{A}} = \begin{pmatrix} 0 & 0 \\ 0 & \Lambda_\tau \end{pmatrix}.$$

In this formula the matrix Λ_τ contain the eigenvalues of $\mathcal{A}$ which are non-zero. By expressing E in its corresponding block form we write

$$E = \begin{pmatrix} \Lambda_{\mathcal{E}_0} & E_{12} \\ E_{12}^* & E_{22} \end{pmatrix}$$

where $\Lambda_{\mathcal{E}_0}$ denotes the diagonal matrix whose diagonal consists of the eigenvalues $\{\lambda_j(\mathcal{E}_0)\}_{j=1}^l$ of $\mathcal{E}_0$ which is defined by $\mathcal{E}_0 = \mathcal{P}_{\ker\mathcal{A}}\mathcal{E}\mathcal{P}_{\ker\mathcal{A}}$. If we apply Lemma 2.14 we obtain that

$$\Lambda_{\mathcal{A}} + E = \begin{pmatrix} \Lambda_{\mathcal{E}_0} & E_{12} \\ E_{12}^* & \Lambda_\tau + E_{22} \end{pmatrix} \sim \begin{pmatrix} \Lambda_{\mathcal{E}_0} + \mathcal{O}(\|E\|^2) & E_{12} \\ \mathcal{O}(\|E\|^2) & \Lambda_\tau + \mathcal{O}(\|E\|) \end{pmatrix}.$$

By inspection we find that the first l columns have only terms of the form $\mathcal{O}(\|\mathcal{E}\|^2)$ on the off-diagonal. Let $R > 0$ denote the maximal radius of the l first Gershgorin disks when expanded along the columns. Then $R = \mathcal{O}(\|\mathcal{E}\|^2)$. If we now apply Corollary 2.16 we can see that the l first disks contain l eigenvalues and that these must be the ones which are smallest in modulus. Indeed the $n - l$ other disks converge to the non-zero eigenvalues of $\mathcal{A}$ as $\mathcal{E} \to 0$ and thus becomes separated from the l first disks. Theorem 2.20 now implies that the l first disks indeed contain l eigenvalues of $\mathcal{A}+\mathcal{E}$ counting multiplicites. We claim that we can achieve the ordering

$$\lambda_j(\mathcal{A} + \mathcal{E}) = \lambda_{j-j_0+1}(\mathcal{E}_0) + \mathcal{O}(\|\mathcal{E}\|^2), \quad j \in \{j_0, \dots, j_0 + l - 1\}.$$

Note first that we can find a permutation of $\{j_0, \dots, j_0 + l - 1\}$ denoted π such that

$$\lambda_j(\mathcal{A} + \mathcal{E}) = \lambda_{\pi(j)-j_0+1}(\mathcal{E}_0) + \mathcal{O}(\|\mathcal{E}\|^2).$$

This follows since any cluster of the l first Gershgorin disks will have a diameter bounded by $lR = \mathcal{O}(\|\mathcal{E}\|^2)$ and contain as many eigenvalues as disks in the cluster. Now to achieve the ordering we note that if $\pi(j) > \pi(j + 1)$ then since

$$0 \leq \lambda_{\pi(j+1)-j_0+1}(\mathcal{E}_0) - \lambda_{\pi(j)-j_0+1}(\mathcal{E}_0)$$

and

$$\lambda_{\pi(j)-j_0+1} + \mathcal{O}(\|\mathcal{E}\|^2) \geq \lambda_{\pi(j+1)-j_0+1}(\mathcal{E}_0) + \mathcal{O}(\|\mathcal{E}\|^2)$$

we find that $\lambda_{\pi(j+1)-j_0+1}(\mathcal{E}_0) - \lambda_{\pi(j)-j_0+1}(\mathcal{E}_0) = \mathcal{O}(\|\mathcal{E}\|^2)$ and thus we can exchange the ordering of the permutation so that $\pi(j) < \pi(j + 1)$ but keeping the form of the error term. Doing this for every j we obtain

$$\lambda_j(\mathcal{A} + \mathcal{E}) = \lambda_{j-j_0+1}(\mathcal{E}_0) + \mathcal{O}(\|\mathcal{E}\|^2), \quad j \in \{j_0, \dots, j_0 + l - 1\}.$$

Assume now that α_0 is an arbitrary eigenvalue of $\mathcal{A}$ with multiplicity l and such that j_0 is the first index j where $\lambda_j(\mathcal{A}) = \alpha_0$. Consider then the operator $\mathcal{A}'$ defined by $\mathcal{A}' = \mathcal{A} - \alpha_0 I$. Since the eigenspaces of $\mathcal{A}'$ and $\mathcal{A}$ agree we find that $\mathcal{E}_0$ as defined above actually satisfies $\mathcal{E}_0 = \mathcal{P}_{\ker(\mathcal{A}-\alpha_0 I)}\mathcal{E}\mathcal{P}_{\ker(\mathcal{A}-\alpha_0 I)}$ on $\ker(\mathcal{A} - \alpha_0 I)$. If we apply our previous result we conclude that

$$\lambda_j(\mathcal{A}' + \mathcal{E}) = \Lambda_{j-j_0+1}(\mathcal{E}_0) + \mathcal{O}(\|\mathcal{E}\|^2).$$

Since $\lambda_j(\mathcal{A}' + \mathcal{E}) = \lambda_j(\mathcal{A} + \mathcal{E}) - \alpha_0 = \lambda_j(\mathcal{A} + \mathcal{E}) - \lambda_{j_0}(\mathcal{A})$ the result now follows. □

The proof carried out here differs from the one presented in [10] since it does not require the matrix M defined in Section 3. In the case of simple eigenvalues Theorem 4.1 implies differentiability. To see this we consider the perturbation of the form

$$A + E$$

then we can diagonalize A via the unitary matrix U so that

$$U^*(A + E)U = \Lambda_{\mathcal{A}} + \hat{E}$$

where $\hat{E} = U^* E U$. If we assume that $\ker(A - \lambda_{j_0}(A)I)$ is one dimensional then we can express the perturbation as

$$\Lambda_{\mathcal{A}} + \hat{E} = \begin{pmatrix} \alpha_0 & 0 \\ 0 & \Lambda_\tau \end{pmatrix} + \begin{pmatrix} \xi & \hat{E}_{12} \\ \hat{E}_{12}^* & \hat{E}_{22}. \end{pmatrix}$$

The matrix $\hat{E}_{12}$ is then of the form $1 \times (n-1)$. In this case Theorem 4.1 implies that the j_0th eigenvalue of $A + E$ satisfies

$$\lambda_{j_0}(A + E) = \lambda_{j_0}(A) + \lambda_1(\xi) + \mathcal{O}(\|E\|^2) = \lambda_{j_0}(A) + \xi + \mathcal{O}(\|E\|^2).$$

Therefore in this case $E \mapsto \lambda_{j_0}(A + E)$ is Fréchet differentiable with derivative $(U^* E U)_{(1,1)}$.

4.2 On a Second Order Perturbation result due to Marcus Carlsson

We now turn our attention to deriving an $\mathcal{O}(\|\mathcal{E}\|^3)$ approximation of the eigenvalues of $\mathcal{A} + \mathcal{E}$ where $\mathcal{A}$ and $\mathcal{E}$ are Hermitian operators in $\mathcal{L}(\mathscr{H}_n)$. The result we are interested in is first proven in [10], however, our proof is novel to our best knowledge and does not rely on estimating the roots of polynomials but instead uses the trick of applying Lemma 2.14 twice in succession.

Theorem 4.2. *Let $\mathcal{A}$ and $\mathcal{E}$ be Hermitian operators in $\mathcal{L}(\mathscr{H}_n)$ where α_0 is an eigenvalue of $\mathcal{A}$ with multiplicity $l > 0$ and set j_0 to denote the first index where $\lambda_{j_0}(\mathcal{A}) = \alpha_0$. Let $\mathcal{N} = \ker(\mathcal{A} - \alpha_0 I)$ and define the operator $\mathcal{B} : \mathcal{N} \to \mathcal{N}$ by*

$$\mathcal{B} = \mathcal{P}_{\mathcal{N}} \mathcal{E} \mathcal{P}_{\mathcal{N}} - \mathcal{P}_{\mathcal{N}} \mathcal{E} \mathcal{P}_{\mathcal{N}^\perp} \left(\mathcal{P}_{\mathcal{N}^\perp} (\mathcal{A} - \alpha_0 I)^\dagger \mathcal{P}_{\mathcal{N}^\perp} \right) \mathcal{P}_{\mathcal{N}^\perp} \mathcal{E} \mathcal{P}_{\mathcal{N}}$$

then

$$\lambda_j(\mathcal{A} + \mathcal{E}) = \lambda_j(\mathcal{A}) + \lambda_{j - j_0 + 1}(\mathcal{B}) + \mathcal{O}(\|\mathcal{E}\|^3).$$

Remark. *It should be noted that Theorem 4.1 follows from Theorem 4.2 together with Theorem 2.7.*

Proof. Again we consider first the situation where $\alpha_0 = 0$ and choose an orthonormal basis of $\mathscr{H}_n$ consisting of eigenvectors of $\mathcal{A}$ denoted $\{\boldsymbol{e}_i\}_{i=1}^n$ where the l first vectors form a basis for $\mathcal{N} = \ker \mathcal{A}$. Since $\mathscr{H}_n = \mathcal{N} \oplus \mathcal{N}^\perp$ we see that we can simultaneously choose the basis such that they are eigenvectors of $\mathcal{B}$ since $\mathcal{B}$ is Hermitian. We define $\Lambda_{\mathcal{B}} = \text{diag}\ (\lambda_1(\mathcal{B}), \ldots, \lambda_l(\mathcal{B}))$ and set Λ_τ to denote the $(n-l) \times (n-l)$ diagonal matrix consisting of the non-zero eigenvalues of $\mathcal{A}$.

In the basis $\{\boldsymbol{e}_i\}_{i=1}^n$ the matrix representation of $\mathcal{A} + \mathcal{E}$ becomes

$$\Lambda_{\mathcal{A}} + E = \begin{pmatrix} 0 & 0 \\ 0 & \Lambda_\tau \end{pmatrix} + \begin{pmatrix} \Lambda_{\mathcal{B}} + C \Lambda_\tau^{-1} C^* & C \\ C^* & D \end{pmatrix}.$$

Applying Lemma 2.14 twice we obtain that

$$\begin{aligned}\Lambda_{\mathcal{A}} + E &\sim \begin{pmatrix} \Lambda_{\mathcal{B}} + \mathcal{O}(\|E\|^3) & C \\ \mathcal{O}(\|E\|^2) & \Lambda_\tau + \mathcal{O}(\|E\|) \end{pmatrix} \\ &\sim \begin{pmatrix} \Lambda_{\mathcal{B}} + \mathcal{O}(\|E\|^3) & C \\ \mathcal{O}(\|E\|^3) & \Lambda_\tau + \mathcal{O}(\|E\|) \end{pmatrix}.\end{aligned} \tag{18}$$

From here on the argument is exactly the same as in the proof of Theorem 4.1. By Corollary 2.16 we find that the eigenvalues of $\Lambda_{\mathcal{A}} + E$ are contained in the Gershgorin disks centered at the diagonal entries of the last matrix in equation (18). As $\|E\| \to 0$ we see that we can make the l Gershgorin disks centered at the diagonal of $\Lambda_{\mathcal{B}} + \mathcal{O}(\|E\|^3)$ disjoint from the remaining $n - l$ Gershgorin disks. Since any cluster of disks centered at $\Lambda_{\mathcal{B}} + \mathcal{O}(\|E\|^3)$ would have a diameter of the form $\mathcal{O}(\|E\|^3)$ Theorem 2.20 now implies that

$$\lambda_j(\mathcal{A} + \mathcal{E}) = \lambda_{j-j_0+1}(\mathcal{B}), \quad j \in \{j_0, \ldots, j_0 + l - 1\}.$$

Now if $\alpha_0 \neq 0$ we again consider $\mathcal{A}' = \mathcal{A} - \alpha_0 I$ and then the result follows. □

If we instead consider a perturbation in matrix form: $A + E$ then we can find a unitary matrix U which diagonalize A so that

$$U^* A U = \Lambda_{\mathcal{A}} = \begin{pmatrix} \alpha_0 I & 0 \\ 0 & \Lambda_\tau \end{pmatrix}.$$

Now the corresponding block form of E can be written

$$U^* E U = \hat{E} = \begin{pmatrix} \hat{E}_{11} & \hat{E}_{12} \\ \hat{E}_{12}^* & \hat{E}_{22} \end{pmatrix}.$$

We write $\Lambda_{\tau-\alpha_0} = \Lambda_\tau - \alpha_0 I$ and define

$$B = \hat{E}_{11} - \hat{E}_{12} \Lambda_{\tau-\alpha_0}^{-1} \hat{E}_{12}^*$$

in accordance with the hypothesis of Theorem 4.2. If we then denote the eigenvalues of B with $\{\beta_i\}_{i=1}^l$ then Theorem 4.2 implies that there is an ordering of the eigenvalues of $A + E$ denoted $\{\xi_i\}_{i=1}^n$ such that

$$\xi_i = \alpha_0 + \beta_i + \mathcal{O}(\|E\|^3), \quad i \in \{1, \ldots, l\}.$$

An application of this result to the spectral analysis of the matrix square root and absolute value can be found in the recent article [11]. We make a final remark that one could drop the dependence of $\hat{E}_{22}$ in the error term if one instead considered B of the form

$$B = \hat{E}_{11} - \hat{E}_{12}(\Lambda_{\tau-\alpha_0} + \hat{E}_{22})^{-1} \hat{E}_{12}^*.$$

For the purpose of this text however the approximation presented in Theorem 4.2 will suffice.

4.3 Applications to the Perturbation of Singular Values

Recall the definition of $\mu_j(\mathcal{A})$ as the jth eigenvalue of $\mathcal{A}$ ordered non-decreasingly counting multiplicities or equivalently $\mu_j(\mathcal{A}) = \lambda_{n-j+1}(\mathcal{A})$. Let $\mathcal{A}$ and $\mathcal{E}$ be arbitrary operators in $\mathcal{L}(\mathscr{H}_n)$, here we do not require that these are Hermitian. By picking a basis we get the corresponding matrix representations A and E in $\mathcal{M}_{n\times n}$. Corollary 2.27 implies the existence of unitary matrices U and V such that

$$A = U\Sigma V^*$$

but then

$$A + E = U(\Sigma + U^* E V)V^*.$$

Note further that

$$(A + E)^*(A + E) = V(\Sigma + U^* E V)^*(\Sigma + U^* E V)V^*$$

which implies that $A + E$ and $\Sigma + U^*EV$ share singular values. We have thus reduced the problem of studying arbitrary matrix perturbations $A + E$ to the study of the simple case of finding singular values of matrices of the form $\Sigma + U^*EV$. We consider first a generalization of Theorem 1 in [8].

Theorem 4.3. *Let Σ and Ξ be $n \times n$ matrices where Σ is a diagonal matrix of the form*

$$\begin{pmatrix} \varsigma I & 0 \\ 0 & \Lambda_s \end{pmatrix}$$

such that $\Lambda_s - \varsigma I$ is invertible and all diagonal entries are non-negative. Let j_0 denote the index of the first occurrence of ς in $\{\sigma_j(\Sigma)\}_{j=1}^n$ and denote its multiplicity with l. Assume further that Ξ is of the form

$$\begin{pmatrix} P & Q \\ R & S \end{pmatrix}$$

and finally denote $C = \varsigma Q + R^*\Lambda_s$ and $B = \varsigma(P + P^*) + P^*P + R^*R - C(\Lambda_s^2 - \varsigma^2 I)^{-1}C^*$. Then

$$\sigma_j(\Sigma + \Xi)^2 = \sigma_j(\Sigma)^2 + \mu_{j-j_0+1}(B) + \mathcal{O}(\|\Xi\|^3), \quad j \in \{j_0, \dots, j_0 + l - 1\}.$$

Proof. By definition the singular-values of $\Sigma + \Xi$ are the square roots of the eigenvalues of $(\Sigma + \Xi)^*(\Sigma + \Xi)$. This expression equals

$$(\Sigma + \Xi)^*(\Sigma + \Xi) = \Sigma^*\Sigma + \Sigma^*\Xi + \Xi^*\Sigma + \Xi^*\Xi.$$

We proceed by computing each of the summands independently in block-form.

$$\Sigma^2 = \begin{pmatrix} \varsigma^2 I & 0 \\ 0 & \Lambda_{s^2} \end{pmatrix}$$

For the perturbation we find that

$$\Sigma^*\Xi = \begin{pmatrix} \varsigma I & 0 \\ 0 & \Lambda_s \end{pmatrix}\begin{pmatrix} P & Q \\ R & S \end{pmatrix} = \begin{pmatrix} \varsigma P & \varsigma Q \\ \Lambda_s R & \Lambda_s S \end{pmatrix},$$
$$\Xi^*\Sigma = \begin{pmatrix} P^* R^* \\ Q^* S^* \end{pmatrix}\begin{pmatrix} \varsigma I & 0 \\ 0 & \Lambda_s \end{pmatrix} = \begin{pmatrix} \varsigma P^* & R^*\Lambda_s \\ \varsigma Q^* & S^*\Lambda_s \end{pmatrix},$$
$$\Xi^*\Xi = \begin{pmatrix} P^* & R^* \\ Q^* & S^* \end{pmatrix}\begin{pmatrix} P & Q \\ R & S \end{pmatrix} = \begin{pmatrix} P^*P + R^*R & P^*Q + R^*S \\ Q^*P + S^*R & Q^*Q + S^*S \end{pmatrix}.$$

By denoting

$$C = \varsigma Q + R^*\Lambda_s$$

and

$$B = \varsigma(P + P^*) + P^*P + R^*R - C(\Lambda_s^2 - \varsigma^2 I)^{-1}C^*$$

Theorem 4.2 together with Theorem 2.7 now implies that the singular values satisfy

$$\sigma_j(\Sigma + \Xi)^2 = \sigma_j(\Sigma)^2 + \mu_{j-j_0+1}(B) + \mathcal{O}(\|\Xi\|^3), \quad j \in \{j_0, \ldots, j_0 + l - 1\}.$$

□

Now the form of B may seem unwieldily however in the case of $\varsigma = 0$ it actually simplifies as we show in our next theorem where we consider the case of a non-empty kernel.

Theorem 4.4. *Let Σ and Ξ be elements of $\mathcal{M}_{n\times n}$ where Σ is a diagonal matrix of the form*

$$\begin{pmatrix} 0 & 0 \\ 0 & \Lambda_s \end{pmatrix},$$

Λ_s is invertible and all diagonal entires are non-negative. Assume further that Ξ is of the form

$$\Xi = \begin{pmatrix} P & Q \\ R & S \end{pmatrix}$$

and that $\sigma_j(P) \geq c\,\|\Xi\|$ for some fix $c > 0$. Then

$$\sigma_j(\Sigma + \Xi) = \sigma_j(P) + \mathcal{O}(\|\Xi\|^2), \quad j \in \{1, \ldots, l\}$$

where $l = \dim\ker\Sigma$.

Proof. By Theorem 4.3 we find that the singular values satisfy

$$\sigma_j(\Sigma + \Xi)^2 = \sigma_j(\Sigma)^2 + \mu_j(B) + \mathcal{O}(\|\Xi\|^3), \quad j \in \{1, \ldots, l\}.$$

Here $\sigma_j(\Sigma) = 0$ if $j \in \{1, \ldots, l\}$ by assumption. The matrix B is in this case given by

$$B = P^*P + R^*R - R^*\Lambda_s(\Lambda_s^2)^{-1}\Lambda_s R = P^*P.$$

Since $\sigma_j(P) = \mu_j(P^*P)$ we conclude that

$$\sigma_j(\Sigma + \Xi)^2 = \sigma_j(P)^2 + \mathcal{O}(\|\Xi\|^3), \quad j \in \{1, \dots, l\}.$$

By Taylor's Theorem $\sqrt{1+x} = 1 + \mathcal{O}(x)$ which implies that

$$\sigma_j(\Sigma + \Xi) = \sigma_j(P)\sqrt{1 + \frac{\mathcal{O}(\|\Xi\|^3)}{\sigma_j(P)^2}} = \sigma_j(P)(1 + \mathcal{O}(\|\Xi\|)) = \sigma_j(P) + \mathcal{O}(\|\Xi\|^2)$$

for $j \in \{1, \dots, l\}$. Here we used the requirement that $\sigma_j(P) \geq c\,\|\Xi\|$.

□

We now use Theorem 4.1 to prove an analogous result for the non-zero singular values. If ς is non-zero in Theorem 4.3 then if we disregard the terms of the form $\mathcal{O}(\|\Xi\|^2)$ we find that $B = \varsigma(P + P^*) + \mathcal{O}(\|\Xi\|^2)$. By Theorem 2.7 the eigenvalues of B lie within $\mathcal{O}(\|\Xi\|^2)$ from those of $\varsigma(P + P^*)$. This gives us the following result.

Theorem 4.5. *Let Σ be an $n \times n$ diagonal matrix with non-negative diagonal of the form*

$$\Sigma = \begin{pmatrix} \varsigma I & 0 \\ 0 & \Lambda_s \end{pmatrix}.$$

Here Λ_s is a diagonal matrix whose diagonal is different from σ_0. Let $\Xi \in \mathcal{M}_{n\times n}$ be of the form

$$\Xi = \begin{pmatrix} P & Q \\ R & S \end{pmatrix}.$$

If j_0 is the first occurrence of the singular value ς in the collection $\{\sigma_j(\Sigma)\}_{j=1}^n$ and if its multiplicity is $l > 0$ then

$$\sigma_j(\Sigma + \Xi) = \sigma_j(\Xi) + \frac{1}{2}\mu_{j-j_0+1}(P + P^*) + \mathcal{O}(\|\Xi\|^2), \quad j \in \{j_0, \dots, j_0 + l - 1\}.$$

Proof. Let j_0 be the first index of ς in the collection $\{\sigma_j(\Sigma)\}_{j=1}^n$ and denote its multiplicity with l as in the statement of the theorem. Theorem 4.3 and Theorem 2.7 now implies that

$$\sigma_j(\Sigma + \Xi)^2 = \sigma_j(\Sigma)^2 + \sigma_j(\Sigma)\mu_{j-j_0+1}(P + P^*) + \mathcal{O}(\|\Xi\|^2)$$

for $j \in \{j_0, \dots, j_0 + l - 1\}$. By applying Taylor's Theorem we conclude that

$$\begin{aligned}\sigma_j(\Sigma + \Xi) &= \sqrt{\sigma_j(\Sigma)^2 + \sigma_j(\Sigma)\mu_{j-j_0+1}(P + P^*) + \mathcal{O}(\|\Xi\|^2)} \\ &= \sigma_j(\Sigma)\left(1 + \frac{1}{2}\frac{\mu_{j-j_0+1}(P + P^*)}{\sigma_j(\Sigma)} + \mathcal{O}(\|\Xi\|^2)\right) \\ &= \sigma_j(\Sigma) + \frac{1}{2}\mu_{j-j_0+1}(P + P^*) + \mathcal{O}(\|\Xi\|^2), \quad j \in \{j_0, \dots, j_0 + l - 1\}\end{aligned}$$

□

In the case that $l = 1$ then $P + P^*$ are simply scalars and as such $\mu_{j-j_0+1}(P + P^*)$ is linear with respect to Ξ if we consider $\mathcal{M}_{n\times n}$ as a vector space over $\mathbb{R}$. This implies that we may consider the Fréchet derivative of $\sigma_{j_0} : \mathcal{M}_{n\times n} \to [0, \infty)$.

Corollary 4.6. *Let A and E be elements of $\mathcal{M}_{n\times n}$ where $\mathcal{M}_{n\times n}$ is considered a vector space over $\mathbb{R}$ and let $U\Sigma V^*$ be the singular value decomposition of A such that*

$$\Sigma = \begin{pmatrix} \varsigma & 0 \\ 0 & \Lambda_s \end{pmatrix}$$

where $\Lambda_s - \varsigma I$ is non-singular. Suppose that $\sigma_{j_0}(A) = \varsigma$ then $E \mapsto \sigma_{j_0}(A + E)$ is Fréchet differentiable at the origin with derivative

$$E \mapsto \frac{1}{2}(UEV^* + VE^*U^*)_{(1,1)}.$$

Proof. By Theorem 4.5 we find that

$$\begin{aligned}\sigma_{j_0}(A + E) &= \sigma_{j_0}(A) + \frac{1}{2}\mu_1((UEV^* + VE^*U^*)_{(1,1)}) + \mathcal{O}(\|E\|^2) \\ &= \sigma_{j_0}(A) + \frac{1}{2}(UEV^* + VE^*U^*)_{(1,1)} + \mathcal{O}(\|E\|^2).\end{aligned}$$

Therefore

$$\lim_{E\to 0} \frac{\left\|\sigma_{j_0}(A + E) - \sigma_{j_0}(A) - \frac{1}{2}(UEV^* + VE^*U^*)_{(1,1)}\right\|}{\|E\|} = 0.$$

□

Remark. *While $E \mapsto UEV^* + VE^*U^*$ is only linear when we consider $\mathcal{M}_{n\times n}$ as a vector space over $\mathbb{R}$ the limit exists in either case.*

This result is previously known, see for instance Section 4 of [22] for the case where the matrices are real so that

$$\frac{1}{2}(UEV^* + VE^*U^*)_{(1,1)} = (UEV^t)_{(1,1)}$$

in accordance with that article.

5 Conclusion and the Possibility of Further Study

In this text we showed how Schur complements can be applied to the spectral analysis of perturbations of matrices to give eigenvalue approximations. We derived a second order expansion of the eigenvalues together with a first order expansion of the eigenvectors of a particular parameter perturbation and compared our method with the one given by David Hilbert and Richard Courant. The application of Schur complements together with Gershgorin's Circle Theorem allowed us to give a new short proof of the eigenvalue approximation presented in [10]. We were unable to find a third order expansion by iterating this method further. We applied the second order approximation to give a generalization of a singular value perturbation result found in [8].

Since both the Schur complement and Gershgorin's Circle Theorem generalize to certain infinite-dimensional linear spaces (see [23] and [24]) one could expect that Theorem 4.2 also could be generalized to those situations. Gershgorin's Circle Theorem, however, requires some additional summability conditions and therefore one would need to restrict the theorem to certain operators which satisfy these conditions.

References

[1] G. Teschl. *Ordinary Differential Equations and Dynamical Systems*. Graduate studies in mathematics. American Mathematical Society, 2012.

[2] K.J. Aström and R.M. Murray. *Feedback Systems: An Introduction for Scientists and Engineers*. Princeton University Press, 2008.

[3] S. Axler. *Linear Algebra Done Right*. Springer-Verlag New York Berlin Heidelberg, 2015.

[4] L. D. Landau and E. M. Lifshitz. *Course of Theoretical Physics - Volume 3*. Reed Educational and Professional Publishing Ltd, 1958.

[5] W. Ford. Chapter 15 - the singular value decomposition. In W. Ford, editor, *Numerical Linear Algebra with Applications*, pages 299 – 320. Academic Press, Boston, 2015.

[6] F. Rellich. *Perturbation theory of eigenvalue problems: lectures delivered fall 1953*. Institute of Mathematical Sciences, New York University, 1953.

[7] R. Bhatia. *Matrix Analysis*. Springer-Verlag New York Berlin Heidelberg, 1997.

[8] G.W. Stewart. A second order perturbation expansion for small singular values. *Linear Algebra and its Applications*, 56:231 – 235, 1984.

[9] G.W. Stewart and Ji guang Sun. *Matrix Perturbation Theory*. Computer science and scientific computing. Academic Press, 1990.

[10] M. Carlsson. Perturbation theory for the spectral decomposition of hermitian matrices. `https://arxiv.org/abs/1809.09480v3`. Accessed: 2018-12-30.

[11] M. Carlsson. Perturbation theory for the matrix square root and matrix modulus. `https://arxiv.org/pdf/1810.01464.pdf`. Accessed: 2019-03-01.

[12] R. Courant, C.I.M.S.R. Courant, D. Hilbert, and D. Hilbert. *Methods of Mathematical Physics*. Number v. 1 in Methods of Mathematical Physics. Wiley, 1953.

[13] B.I.A. Levin. *Distribution of Zeros of Entire Functions*. Translations of mathematical monographs. American Mathematical Society, 1964.

[14] S. Gratton and J. Tshimanga-Ilunga. On a second-order expansion of the truncated singular subspace decomposition. *Numerical Linear Algebra with Applications*, 23(3):519–534, 2016.

[15] P.D. Lax. *Linear Algebra and Its Applications*. Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts. Wiley, 2013.

[16] A. Holst and V. Ufnarovski. *Matrix Theory*. Studentlitteratur, Lund, 2014.

[17] F. Zhang. *The Schur Complement and Its Applications*. Numerical Methods and Algorithms. Springer US, 2010.

[18] R.A. Horn and C.R. Johnson. *Matrix Analysis*. Matrix Analysis. Cambridge University Press, 2013.

[19] E.M. Stein and R. Shakarchi. *Complex Analysis*. Princeton lectures in analysis. Princeton University Press, 2010.

[20] T.W. Gamelin. *Complex Analysis*. Undergraduate Texts in Mathematics. Springer New York, 2013.

[21] J.R. Munkres. *Topology*. Prentice-Hall, 2000.

[22] G.W Stewart. Perturbation theory for the singular value decomposition. *SVD and Signal Processing II, Algorithms, Analysis and Applications*, pages 99–109, 1991.

[23] C. Bacuta. Schur complements on hilbert spaces and saddle point systems. *Journal of Computational and Applied Mathematics*, 225(2):581 – 593, 2009.

[24] P.N. Shivakumar, J.J. Williams, and N. Rudraiah. Eigenvalues for infinite matrices. *Linear Algebra and its Applications*, 96:35 – 63, 1987.

Master's Theses in Mathematical Sciences 2019:E12
ISSN 1404-6342

LUNFMA-3103-2019

Mathematics
Centre for Mathematical Sciences
Lund University
Box 118, SE-221 00 Lund, Sweden

http://www.maths.lth.se/