



# Faculteit Bètawetenschappen

# Weaver's conjecture

BACHELOR THESIS

Artuur Oerlemans Wiskunde

> Dr. M.P.T. Caspers Supervisor Mathematical Institute 13-06-2017

# Contents

1	Introduction	<b>1</b>						
	1.1 Structure of thesis	1						
<b>2</b>	Preliminaries							
	2.1 General notations	2						
	2.2 Matrices	2						
	2.2.1 Hermitian Matrices	3						
	2.2.2 Characteristic Polynomials	3						
	2.2.3 Positive Definite Matrices	3						
	2.3 Norm	4						
	2.4 Probability	5						
3	Polynomial Roots	6						
	3.1 Hurwitz' Theorem	6						
	3.2 Interlacing Families	6						
	3.3 Stable Polynomials	9						
	3.4 Determinants	11						
4	Mixed Characteristic Polynomials 14							
<b>5</b>	Largest Root 1							
6	Proof Weaver's Conjecture							
7	Paving Conjecture	27						
•	7.1 Theorems for proving the Paving Conjecture	27						
	7.2 Proving the Paving Conjecture	28						
8	3 Conclusion							
R	References							

# 1 Introduction

This thesis will mainly focus on proving Weaver's Conjecture  $KS_2$  in a way that is easier to understand for 3 year bachelor student. Without further ado Weaver's Conjecture  $KS_2$ .

**Conjecture 1.0.1** (Weaver's Conjecture KS<sub>2</sub> [13]). There exist universal constants  $\eta \geq 2$  and  $\theta > 0$  such that the following holds. Let  $w_1, ..., w_m \in \mathbb{C}^d$  be vectors such that  $||w_i|| \leq 1$  for all i and that

$$\sum_{i=1}^{m} |\langle u, w_i \rangle|^2 = \eta \tag{1}$$

for every unit vector  $u \in \mathbb{C}^d$ . Then there exists a partition  $S_0, S_1$  of [m] so that

$$\sum_{i \in S_j} |\langle u, w_i \rangle|^2 \le \eta - \theta \tag{2}$$

for every unit vector  $u \in \mathbb{C}^d$  and every j.

The reason why Nik Weaver created this conjecture, is that a positive answer implies the Kadison-Singer problem.

**Question 1.0.2** (Kadison-Singer problem, Question 1.1 of [9]). Does every pure state on the C<sup>\*</sup>-algebra  $\mathbb{D}$  of bounded diagonal operators on  $\ell_2$  have a unique extension to a pure state on  $B(\ell_2)$ , the C<sup>\*</sup>-algebra of all bounded operators on  $\ell_2$ ?

The Kadison-Singer problem finds its roots in quantum mechanics in the 1940s and was first formulated in 1959 by Richard Kadison and Isadore Singer [6]. It then went on to become one of the most important problems in functional analysis, due to it being equivalent with many other open mathematical problems in a variety fields and of course the implications it has for quantum mechanics.

#### 1.1 Structure of thesis

Overall, the order in which things are proven is rather like that of the main source material [9]. The difference mainly lies in how much is proven. With this thesis, we aim to create more of a complete image of the proof. So, we will go into more detailed proofs and will also work out some of the references used in the main source material.

This thesis will start with the preliminaries, in which we will explain all necessary notations for this thesis, including those needed for understanding Weaver's Conjecture  $KS_2$ . In the preliminaries, we won't give any information about the terms used in the Kadison-Singer problem, because this thesis will not directly prove the positive solution to the Kadison-Singer Problem.

After that we will introduce interlacing families and stable polynomials, which will be our main tools for solving Weaver's Conjecture KS<sub>2</sub>. We will then go on to apply such functions in mixed characteristic polynomials, which we will then show have a limit on their largest roots. This limit we will then later be able to put in relation with  $\eta - \theta$  from the Weaver's Conjecture KS<sub>2</sub>.

All that has been proven and introduced will then come together in chapter 6, where we will finally prove Weaver's Conjecture, the main aim of this thesis. This will be done by formulating and proving a theorem which has similar properties to Weaver's Conjecture. To prove that theorem we will use the connection between the operator norm and eigenvalues. This is where we will get our use out of all that we have previously proven about mixed characteristic polynomials. After having proven this theorem, we will prove a generalization of Weaver's Conjecture, which we can then use to prove Weaver's Conjecture.

In chapter 7 we will try to get a bit closer to fully proving the positive result of the Kadison-Singer problem. We will do this by proving the Paving Conjecture which is equivalent to the Kadison-Singer Problem and is also what Nik Weaver used to show that Weaver's Conjecture implies the result of the Kadison-Singer problem. This proof will be based on the generalization of Weaver's Conjecture in chapter 6.

# 2 Preliminaries

## 2.1 General notations

The natural numbers  $\mathbb{N}$  will be defined as  $\{1, 2, ...\}$  in this thesis. The collection [m] is equal to  $\{1, ..., m\}$ .

#### 2.2 Matrices

In this thesis, we are going to be working a lot with Hermitian and positive definite matrices. So here is some pre-required knowledge about them.

A matrix  $X \in \mathbb{C}^{n \times n}$  is equal to

$$\begin{pmatrix} X_{1,1} & X_{1,2} & \cdots & X_{1,n} \\ X_{2,1} & X_{2,2} & & \vdots \\ \vdots & & \ddots & \vdots \\ X_{n,1} & \cdots & \cdots & X_{n,n} \end{pmatrix}.$$

A transpose is noted down as

$$X^{t} = \begin{pmatrix} X_{1,1} & X_{2,1} & \cdots & X_{n,1} \\ X_{1,2} & X_{2,2} & & \vdots \\ \vdots & & \ddots & \vdots \\ X_{1,n} & \cdots & \cdots & X_{n,n} \end{pmatrix}$$

The determinant of a matrix is denoted as

$$det(X) = \begin{vmatrix} X_{1,1} & X_{1,2} & \cdots & X_{1,n} \\ X_{2,1} & X_{2,2} & & \vdots \\ \vdots & & \ddots & \vdots \\ X_{n,1} & \cdots & \cdots & X_{n,n} \end{vmatrix}$$

1.( 17)

The adjugate of a matrix is denoted as

$$\begin{aligned} \text{adj}(X) &= \\ \left( \begin{array}{c} + \begin{vmatrix} X_{2,2} & X_{2,3} & \cdots & X_{2,n} \\ X_{3,2} & X_{3,3} & \vdots \\ \vdots & \ddots & \vdots \\ X_{n,2} & \cdots & X_{n,n} \\ X_{2,1} & X_{2,3} & \cdots & X_{2,n} \\ X_{2,1} & X_{2,3} & \cdots & X_{2,n} \\ \vdots & \ddots & \vdots \\ X_{n,1} & \cdots & \cdots & X_{n,n} \end{vmatrix} \right) & - \begin{vmatrix} X_{1,2} & X_{1,3} & \cdots & X_{1,n} \\ X_{3,2} & X_{3,3} & \vdots \\ \vdots & \ddots & \vdots \\ X_{n,2} & \cdots & \cdots & X_{n,n} \\ X_{3,1} & X_{3,3} & \vdots \\ \vdots & \ddots & \vdots \\ X_{n,1} & \cdots & \cdots & X_{n,n} \end{vmatrix} \\ & & \vdots & & \ddots & \vdots \\ \left( (-1)^{n+1} \begin{vmatrix} X_{2,1} & X_{2,2} & \cdots & X_{2,n-1} \\ X_{3,1} & X_{3,2} & \vdots \\ \vdots & \ddots & \vdots \\ X_{n,1} & \cdots & \cdots & X_{n,n-1} \end{vmatrix} \right) & & & & \cdots & + \begin{vmatrix} X_{1,1} & X_{1,2} & \cdots & X_{1,n} \\ X_{1,1} & X_{1,2} & \cdots & X_{n-1,n} \end{vmatrix}$$

Important to remember here is that the value of column i and row j of the adjugate is equal to the determinant of the matrix that misses column j and row i multiplied by  $(-1)^{i+j}$ . The adjugate can be used for calculating the inverse, since

$$X^{-1} = \frac{\operatorname{adj}(X)}{\det(X)}.$$

Some basic matrices are:

- $I_n$  is the unit matrix that is  $n \times n$ .
- $\mathbf{1}_{n \times n}$  is the matrix filled with only ones and  $\mathbf{1}_n$  is the vector filled with only ones.
- $\mathbf{0}_{n \times n}$  is the matrix filled with only zeros and  $\mathbf{0}_n$  is the vector filled with only zeroes.

A coordinate projection is a matrix which is  $\mathbf{0}_{n \times n}$  except for certain values on the diagonal, which are then equal to 1. Example

$$\begin{pmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 0
\end{pmatrix}$$

A unitary matrix U is a matrix such that  $U = U^* = U^{-1}$ .

#### 2.2.1 Hermitian Matrices

Hermitian matrices are  $n \times n$  matrices that stay the same during a Hermitian transpose which is defined as  $M^* = \overline{M}^t$ . An example of a Hermitian transpose

$$\begin{pmatrix} 1 & 1+i \\ 0 & 2-i \end{pmatrix}^* = \begin{pmatrix} 1 & 0 \\ 1-i & 2+i \end{pmatrix}$$

Here are some important properties of Hermitian matrices

Lemma 2.2.1. All its eigenvalues are real.

Lemma 2.2.2. It is possible to diagonalize Hermitian matrices using unitary matrices.

These lemmas have all been proven in the undergraduate course linear algebra.

#### 2.2.2 Characteristic Polynomials

For a matrix  $M \in \mathbb{C}^{n \times n}$  there is a characteristic polynomial of the form

$$\chi[M](x) = \det(xI - M)$$

where I is the unit matrix. The roots of this polynomial are M's eigenvalues.

#### 2.2.3 Positive Definite Matrices

A  $n \times n$  Hermitian matrix H is positive definite when all values of  $v^*Hv$  are positive for every non-zero vector v. Similar to this, a positive semidefinite matrix is a matrix H such that  $v^*Hv$  is non-negative. Another way of telling if a Hermitian matrix is positive definite or semidefinite is by looking at the eigenvalues. If all eigenvalues are positive, then the matrix is positive definite and when they are non-negative the matrix is positive semidefinite.

Important properties of the positive definite matrices are

Lemma 2.2.3. Every positive definite matrix is invertible and its inverse is also positive definite.

**Lemma 2.2.4.** When you have matrices M and N that are positive semidefinite, then also M + N is positive semidefinite. If M or N were also to be positive definite, then M + N is also positive definite.

*Proof.* For this simply look at the way we first defined positive definite and semidefinite matrices. It should then be obvious from the fact that  $v^*(M+N)v = v^*Mv + v^*Nv$ .

**Lemma 2.2.5.** If M is positive semidefinite and Q has the same size, then also  $Q^*MQ$  is positive semidefinite. If moreover Q is invertible and M is positive definite then  $Q^*MQ$  is positive definite.

**Lemma 2.2.6.** Positive definite matrices have a unique square root matrix, that is positive definite. The same is true for positive semidefinite matrices, except their square roots are positive semidefinite.

Lemma 2.2.7. A Gram matrix is always positive semidefinite.

*Proof.* Put the vectors needed for the Gram matrix G in a matrix M such that  $M^*M = G$ . Then again by using the first definition

$$v^*Gv = v^*M^*Mv = (Mv)^*(Mv) \ge 0.$$

So, G must be positive semidefinite.

Lemma 2.2.8. If A and B are positive definite matrices of the same size, then

 $\operatorname{Tr}(AB) \ge 0$ 

*Proof.* Use lemma 2.2.6 to get that A = XX and B = YY. Using the fact that X and Y are Hermitian we get  $A = XX^*$  and  $B = YY^*$ .

$$Tr(AB) = Tr(XX^*YY^*) = Tr((Y^*X)(XY^*)) = Tr((XY^*)^*(XY^*))$$

Because a vector multiplied by its Hermitian transpose always gives a non-negative value, all the values on the diagonal are also non-negative and the same is true for the trace.  $\Box$ 

#### 2.3 Norm

A norm ||...|| on a vector a space V is mapped  $||...||: V \to R_{\geq 0}$  such that

- 1. ||x|| = 0 iff x = 0
- 2.  $||\lambda x|| = |\lambda| * ||x||$  where  $\lambda \in \mathbb{R}$
- 3.  $||x + y|| \le ||x|| + ||y||$

In this thesis when  $x \in \mathbb{R}^n$  we will use the Euclidean norm. For matrices, we will use the norm:

$$||M|| = \max_{||x||=1} ||Mx|| = \sqrt{\max\{\text{eigenvalues } \{M^*M\}\}}.$$

**Lemma 2.3.1.** When M is positive semidefinite, then ||M|| is equal to the largest eigenvalue of M.

*Proof.* Because M is positive semidefinite, it follows from our definition that it is also Hermitian.

 $||M|| = \sqrt{\max\{\text{eigenvalues } \{M^*M\}\}}$  $= \sqrt{\max\{\text{eigenvalues } \{MM\}\}}$ 

 $= \sqrt{\max\{\text{eigenvalues } \{M\}\}^2} \text{ (when you raise a matrix to a certain value, the same will happen for it's eigenvalues)}$  $= \max\{\text{eigenvalues } \{M\}\}$ 

**Definition 2.3.2.** We will define the inner product of  $v, w \in \mathbb{C}^n$  as

$$\langle v, w \rangle = v^{\mathrm{t}} \overline{w}.$$

**Lemma 2.3.3.** Let  $A \in \mathbb{C}^{m \times n}$  and  $B \in \mathbb{C}^{n \times m}$ , then  $||AB|| \leq ||A||||B||$ .

*Proof.* This follows from the fact that  $||AB|| = \max_{||x||=1} ||ABx||$ , so splitting it into two parts means that there more possible values added to take the maximum of.

**Lemma 2.3.4.** Let  $M \in \mathbb{C}^{m \times n}$ , then  $||MM^*|| = ||M^*M||$ .

*Proof.* This proof is based on Proposition 2.7 of [3]. Set ||x|| = 1, then

$$\begin{split} ||Mx||^2 &= \langle Mx, Mx \rangle \\ &= \langle M^*Mx, x \rangle \\ &\leq ||M^*Mx||||x|| \\ &= ||M^*Mx|| \\ &\leq ||M^*M|| \text{ (follows from the definition of the matrix norm we use)} \\ &\leq ||M^*||||M|| \text{ (using Lemma 2.3.3).} \end{split}$$

Using the definition ||...|| for matrices we get that

$$||Mx||^{2} \leq ||M||^{2} \\ \leq ||M^{*}M|| \\ \leq ||M^{*}||||M||.$$

By now dividing all the previous values by ||M||, we get that  $||M|| \le ||M^*||$ . Since we didn't define any specific properties for M, this property also works the other way around;  $||M|| = ||M^*||$ . By now reversing the division by ||M||, we get that

$$\begin{split} ||M||^2 &= ||M^*M|| \\ &= ||M^*||||M|| \\ &= ||M||||M^*|| \\ &= ||MM^*|| \end{split}$$

#### 2.4 Probability

The notation  $\mathbb{E}$  means expected value,  $\mathbb{P}$  means the probability of something occurring. When we state independent random vectors, we don't mean that the vectors are linearly independent, but that the probability of getting a certain vector value is independent of the other vectors.

# **3** Polynomial Roots

This chapter shall particularly focus on the roots of polynomials. Using interlacing families to group polynomials together and stability to avoid complex roots.

### 3.1 Hurwitz' Theorem

Here is a bit of an abstract theorem about the roots of polynomials, which shall be rather useful in the following proofs.

**Theorem 3.1.1** (Hurwitz' theorem, Theorem 2.3 of [1]). Let D be a domain (open connected set) in  $\mathbb{C}^n$ and suppose that  $\{f_k\}_{k=1}^{\infty}$  is a sequence of nonvanishing (no roots) analytic (a value can locally be given by a convergent power series) functions on D that converge to f uniformly on compact subsets of D (on a Euclidean space this is regular convergence). Then f is either nonvanishing on D or else identical to zero.

To simplify the theorem a bit, here is the following lemma.

Lemma 3.1.2. A polynomial is always an analytic function.

*Proof.* This comes from the fact that every polynomial can be written as a finite power series.

### 3.2 Interlacing Families

This part will be about interlacing families existing out of real rooted polynomials. Real rooted means that all its roots and **coefficients** are real.

**Definition 3.2.1.** Let g and f be real rooted polynomials of the form  $g(x) = \alpha_0 \prod_{i=1}^{n-1} (x - \alpha_i)$  and  $f(x) = \beta_0 \prod_{i=1}^n (x - \beta_i)$ . We say that g interlaces f when  $\beta_1 \leq \alpha_1 \leq \beta_2 \leq \alpha_2 \leq \ldots \leq \alpha_{n-1} \leq \beta_n$ .

**Definition 3.2.2.** When there are multiple polynomials interlaced by the same polynomial we speak of a common interlacing. Since we can construct polynomials with the roots where ever we want, we can create a definition of this that doesn't contain a polynomial that interlaces them both, only the concept that such a polynomial can be created. This means when we have polynomials  $f_1, ..., f_k$  which are real rooted of the form  $f_j(x) = \beta_{j,0} \prod_{i=1}^n (x - \beta_{j,i})$  we can say that they have a common interlacing when  $\max(\beta_{1,i}, ..., \beta_{k,i}) \leq \min(\beta_{1,i+1}, ..., \beta_{k,i+1})$  for  $i \in [n]$ .

To make the following text a bit less complicated, we will introduce the following notation. When f is a polynomial, the largest root of f shall be denoted by  $\xi(f)$ 

**Lemma 3.2.3.** Let  $f_1, \ldots, f_k$  be real rooted, have a positive leading coefficient and have the same degree. Define

$$f_{\emptyset} = \sum_{i=1}^{k} f_i$$

If  $f_1, \ldots, f_k$  have a common interlacing, then there exists an i such that  $\xi(f_i)$  is at most  $\xi(f_0)$ .

*Proof.* Because every  $f_j$  with  $j \in [k]$  has a positive leading coefficient, the function  $f_j(x)$  will always converge to  $\infty$  when x goes to  $\infty$ . We now know that when x is greater than  $\xi(f_j)$ , that  $f_j(x) > 0$ . Since we are working with a common interlacing, setting  $c = \inf\{\xi(f_j) | j \in [k]\}$  we have  $f_j(c) \leq 0$  and  $f_{\emptyset}(c) \leq 0$ . So,  $\xi(f_{\emptyset})$  must be greater or equal to c.

**Definition 3.2.4.** Let  $S_1, ..., S_m$  be a finite sets and for every  $s_1, ..., s_m \in S_1 \times ... \times S_m$  let  $f_{s_1,...,s_m}$  be a real rooted polynomial of degree n with a positive leading coefficient. When k < m we define that

$$f_{s_1,...,s_k} = \sum_{s_{k+1} \in S_{k+1},...,s_m \in S_m} f_{s_1,...,s_k,s_{k+1},...,s_m}$$

also, that

$$f_{\emptyset} = \sum_{s_1 \in S_1, \dots, s_m \in S_m} f_{s_1, \dots, s_n}$$

We say that the polynomials  $\{f_{s_1,...,s_m}\}$  are an interlacing family when for each k = 0, ..., m-1 and  $s_1, ..., s_k \in S_1 \times ... \times S_k$ , the collection  $\{f_{s_1,...,s_k,t}\}_{t \in S_{k+1}}$  has a common interlacing.

**Theorem 3.2.5.** Let  $\{f_{s_1,...,s_m}\}$  be an interlacing family of polynomials, then there exists an  $s_1,...,s_m$  such that  $\xi(f_{s_1,...,s_m})$  is at most  $\xi(f_{\emptyset})$ .

Proof. We know that  $\{f_t\}_{t\in s_1}$  has a common interlacing and since the sum of them is equal to  $f_{\emptyset}$ , we can use Lemma 3.2.3 to show that there exist an  $i_1 \in s_1$  such that  $\xi(f_{i_1})$  is at most  $\xi(f_{\emptyset})$ . Now we can start using Lemma 3.2.3 inductively, since  $f_{i_1,\ldots,i_k}$  is the sum of functions  $\{f_{i_1,\ldots,i_k,t}\}_{t\in s_{k+1}}$ , which leads to that there exists a  $i_{k+1}$  such that  $\xi(f_{i_1,\ldots,i_k+1})$  is at most  $\xi(f_{i_1,\ldots,i_k})$  which is at most  $\xi(f_{\emptyset})$ . So, for function  $f_{i_1,\ldots,i_m}$  goes that  $\xi(f_{i_1,\ldots,i_m})$  is at most  $\xi(f_{\emptyset})$ .

Later in this thesis we will need a proper way of determining when there is a common interlacing, for this the next theorem can be used.

**Theorem 3.2.6.** Let  $f_1, ..., f_m$  be polynomials with a positive leading coefficient (it can easily be extended to work without this property, but it in the proof for this we will use a positive leading coefficient) and all have degree n. Then  $f_1, f_2, ..., f_m$  have a common interlacing if and only if all convex combinations  $\sum_{i=1}^m \lambda_i f_i$  where  $\lambda_i \geq 0$  and  $\sum_{i=1}^m \lambda_i = 1$  are real rooted.

To prove the aforementioned theorem, we are going first prove a lemma (Theorem 2.1 of [4]) that can then be used inductively.

Lemma 3.2.7. Let f and g be polynomials with a positive leading coefficient of degree n. Then

1. f and g have a common interlacing

if and only if

2. all convex combinations  $h_{\lambda} = \lambda f + (1 - \lambda)g \ \lambda \in [0, 1]$  are real rooted.

Proof of Lemma 3.2.7  $\{(1) \Rightarrow (2)\}$ . Let f and g have roots  $l_1, ..., l_n$  and  $r_1, ..., r_n$  respectively with  $l_i \leq l_{i+1}$  and  $r_i \leq r_{i+1}$ .

Here we only need to look at the case where  $\lambda \in ]0, 1[$ , because for there to be a common interlacing f and g must be real rooted.

First, the case where f and g are simple rooted (meaning  $i \neq j$ , then  $l_i \neq l_j$ ). Using the fact that the we only have simple roots and that they are part of a common interlacing, we can set  $\max(l_i, r_i) < b_i < \min(l_{i+1}, r_{i+1})$ with  $b_0 = -\infty$  and  $b_n = \infty$ . Since f and g both have a positive leading coefficient

$$\operatorname{sign}(f(b_i)) = \operatorname{sign}(g(b_i)) = \operatorname{sign}(h_{\lambda}(b_i)) = -(-1)^i.$$

This means that  $h_{\lambda}$  will switch *n* times between positive and negative for real values. So, because  $h_{\lambda}$  is a continuous function it must be real rooted.

Second, the case that f and g aren't simple rooted. This is done by creating function  $f_{\epsilon}$  and  $g_{\epsilon}$ , where if  $\epsilon \to 0$ , that  $f_{\epsilon} \to f$  and  $g_{\epsilon} \to g$ . Important here is for which  $\epsilon$  they here have a common interlacing. As to demonstrate the proper approach for creating a good  $f_{\epsilon} \to f$  and  $g_{\epsilon} \to g$  here an example. Let f(x) = (x-5)(x-5)(x-5) and g(x) = (x-6)(x-5)(x-3), then  $f_{\epsilon}(x) = (x-5-\epsilon)(x-5)(x-5+\epsilon)$  and  $g_{\epsilon}(x) = (x-6)(x-5)(x-3)$ . If  $\epsilon \in ]-\infty, 0[\cup]0, \infty[$  there will exist a common interlacing for these functions as we proved earlier. Define

$$h_{\lambda,\epsilon} = \lambda f_{\epsilon} + (1 - \lambda)g_{\epsilon}$$

When  $\epsilon \to 0$ , then  $h_{\lambda,\epsilon} \to h_{\lambda}$ . By using the Hurwitz' theorem (3.1.1) we can now prove that  $h_{\lambda}$  doesn't have any complex roots. For this we will use the Hurwitz' theorem (3.1.1) twice. We define sequence  $\{h_{\lambda,3^{-k}}\}_{k=1}^{\infty}$ , these polynomials are analytic. Since those polynomials are all real rooted the sequence is nonvanishing in domains  $D_1 = \{x \in \mathbb{C} | \operatorname{Im}(x) < 0\}$  and  $D_2 = \{x \in \mathbb{C} | \operatorname{Im}(x) > 0\}$  (separated, since otherwise the domains wouldn't be connected domains). Now according Hurwitz' theorem (3.1.1)  $h_{\lambda}$  is either nonvanishing for values with an imaginary part or always equal to zero, we know the latter is not true due to  $h_{\lambda}$  being the sum of two polynomials with a positive leading coefficient. So, by process of elimination the roots of  $h_{\lambda}$  have to be real. Proof of Lemma 3.2.7  $\{(2) \Rightarrow (1)\}$ . First the proof for the case where all roots are distinct from each other. We will do this by a proof by contradiction. Assume that there isn't a common interlacing. Set k to be the lowest value such that  $\max(l_k, r_k) > \min(l_{k+1}, r_{k+1})$  and since f and g are interchangeable we can say  $r_k > l_{k+1}$  without losing generality. Set  $\max(l_i, r_i) < b_i < \min(l_{i+1}, r_{i+1})$  for  $i \in [k-1]$  and  $b_0 = -\infty$ . Define

$$q_{\lambda}(x) = \frac{h_{\lambda}(x)}{(1-\lambda)f} = \frac{\lambda}{1-\lambda} + \frac{g(x)}{f(x)}$$

 $q_{\lambda}$ . It has the same roots as  $h_{\lambda}$ , because f and g have no common roots (Meaning that  $q_{\lambda}$  is defined for all roots of  $h_{\lambda}$ ). Hence,  $(1 - \lambda)f$  also doesn't have any common roots with  $h_{\lambda}$ . Now we will make a small table to show where the contradiction comes from.

	$b_{k-1}$	$l_k$		$l_{k+1}$	$r_k$
$\operatorname{sign}(f)$	$-(-1)^{k-1}$	0	$-(-1)^{k}$	0	$-(-1)^{k+1}$
$\operatorname{sign}(g)$	$-(-1)^{k-1}$	$-(-1)^{k-1}$	$-(-1)^{k-1}$	$-(-1)^{k-1}$	0
$\operatorname{sign}(q_{\lambda})$	+		-		-

Here it shows that in domain  $]l_k, l_{k+1}[$  the function  $q_{\lambda}$  is always negative, but when  $\lambda \uparrow 1$ , then  $\frac{\lambda}{1-\lambda} \to \infty$ . So for every  $x \in ]l_k, l_{k+1}[$ , there should exist a  $\lambda$  such that  $q_{\lambda}(x)$  is positive; a contradiction.

Proof for the case that f and/or g have a multiple root (the opposite of a simple root. As an example,  $l_k = l_{k+1}$ , this would break the previous part of the proof), but still don't have any roots in common. Define

$$f_{\epsilon} = \epsilon f + (1 - \epsilon)g = h_{\epsilon}$$
$$g_{\epsilon} = (1 - \epsilon)f + \epsilon g = h_{1 - \epsilon}$$

where  $\epsilon > 0$  and low enough to keep common roots between the two functions from happening. Since these functions are real rooted, if we manage to prove that they are also simple rooted, we can then use the fact that when they have a common interlacing the limit of when  $\epsilon \to 0$  should have a common interlacing. Since the limit of  $\max(l_{\epsilon,i}, r_{\epsilon,i}) \leq \min(l_{\epsilon,i+1}, r_{\epsilon,i+1})$  should also hold this inequality.

To prove that  $f_{\epsilon}$  and  $g_{\epsilon}$  are simple rooted, we can prove that  $h_{\epsilon}$  is simple rooted and that can be done by showing that  $q_{\epsilon}$  is simple rooted. Let's assume there is a multiple root in  $q_{\epsilon}$ , because  $\epsilon$  is an element of an open interval it can always be increased or decreased in value. If  $q_{\epsilon}$  is positive before and after the multiple root, then you can create a complex root by increasing the value of  $\lambda$  so that it doesn't touch the x-axis there any more, the other way around when the function is negative before and after the multiple root. In the case that it goes from positive to negative or the other way around it should be clearly visible that a complex root comes into being when the value of  $\lambda$  is changed. So,  $f_{\epsilon}$  and  $g_{\epsilon}$  are simple rooted.

Finally, the proof for the case for when f and g have a common root. If it is just one, we can rewrite the functions as  $f(x) = (x - a)^b f_a(x)$  and  $g(x) = (x - a)^b g_a(x)$ , where a is the common root and b is in case it is also a multiple root for both functions. Since  $f_a$  and  $g_a$  don't have any common roots they will have a common interlacing, giving us  $\max(l_{a,i}, r_{a,i}) \leq \min(l_{a,i+1}, r_{a,i+1})$  for  $i \in [n - b]$ . Now there are two possibilities:

- 1. there exists a k such that  $\max(l_{a,k}, r_{a,k}) \leq a \leq \min(l_{a,k+1}, r_{a,k+1})$  in which case  $\max(l_k, r_k) \leq \min(l_{k+1}, r_{k+1}) = a$  and  $a = (l_{k+b}, r_{k+b}) \leq \min(l_{k+b+1}, r_{k+b+1})$ .
- 2. there exists a k such that  $\min(l_{a,k}, r_{a,k}) \le a \le \max(l_{a,k}, r_{a,k})$  in which case  $a = \max(\min(l_{a,k}, r_{a,k}), a) = \max(l_k, r_k) \le \min(l_{k+1}, r_{k+1}) = \min(\max(l_{a,k}, r_{a,k}), a) = a$  and  $a = \max(\min(l_{a,k}, r_{a,k}), a) = \max(l_{k+b}, r_{k+b}) \le \min(l_{k+b+1}, r_{k+b+1}) = \min(\max(l_{a,k}, r_{a,k}), a) = a$ .

Either way, f and g have a common interlacing. This proof can be done similarly for the case that there is more than one common root.

Proof of Theorem 3.2.6. First the proof for when  $f_1, ..., f_m$  have a common interlacing, then every convex combination is real rooted. The easiest way this can be done is by taking the first half of the proof of Lemma 3.2.7 and extending it to work with m functions. This can be done by redefining  $b_i$  as  $\max(l_{1,i}, ..., l_{m,i}) \leq 1$ 

 $b_i \leq \min(l_{1,i+1}, \dots, l_{m,i+1})$  where  $l_{j,i}$  is the *i*th root of  $f_j$ .

Now for the proof for when every convex combination of  $f_1, ..., f_m$  is real rooted, then there is also a common interlacing. For this we will use Lemma 3.2.7 inductively. The base case can be derived directly from Lemma 3.2.7 when there are only 2 functions. Now to prove that when  $f_1, ..., f_{m-1}$  has a common interlacing and all the convex combinations of  $f_1, ..., f_m$  are real rooted, then also  $f_1, ..., f_m$  has a common interlacing. Here we need to use the fact that every convex combination of  $f_m$  with  $f_i$  is real rooted, so  $f_m$  has a common interlacing with every  $f_1, ..., f_{m-1}$ . Let's assume that  $f_1, ..., f_m$  doesn't have a common interlacing, then there exist an  $k \in [n-1]$  and a  $p \in [m-1]$  such that  $l_{m,k} > \min(l_{1,k+1}, ..., l_{m-1,k+1}) = l_{p,k+1}$  or  $l_{p,k} = \max(l_{1,k}, ..., l_{m-1,k}) < l_{m,k+1}$ . This means there is no common interlacing between  $f_m$  and  $f_p$ ; a contradiction.

#### 3.3 Stable Polynomials

Here we will focus on the creation of stable polynomials. What these are, will be explained in the following definition.

**Definition 3.3.1.** Let  $f(z_1, ..., z_m)$  be a polynomial, we say f is stable if it either is the zero polynomial or for f when  $\text{Im}(z_i) > 0$  for all  $i \ f(z_1, ..., z_m) \neq 0$ . We call f real stable if in addition also its coefficients are real.

While this definition is easy to understand, it can be changed a bit such that the stability of a multivariate function can be proven with only univariate functions.

**Lemma 3.3.2.**  $f(z_1, ..., z_m)$  is stable if and only if for all  $\alpha \in \mathbb{R}^n$  and  $v \in \{x > 0 | \mathbb{R}\}^n$  univariate function  $f(\alpha + vt)$  with  $t \in \mathbb{C}$  is stable.

*Proof.* Since if a function is stable, is only determined by the values for which the function can't be zero. We only need to show that Im(t) > 0 if and only if  $\text{Im}(z_i) > 0$  for all *i*.

When Im(t) > 0, also  $\text{Im}(v_i t) > 0$  for all *i* and since  $\text{Im}(\alpha_i) = 0$  for all *i*, it goes without saying that also  $\text{Im}(z_i) > 0$  for all *i*.

When  $\text{Im}(z_i) > 0$  for all *i*, it can only be that Im(t) > 0, because *v* values can't be negative or equal to zero and *t* is the only part that can take a complex value.

Knowing this, we know that there is a bijection between the two (this is partially possible because of  $\text{Im}(z_i) > 0$  for all *i*, since with  $(\alpha + vt)$  the imaginary part is either negative, positive or non existent).

Since we are busy with univariate polynomials, it's time to prove an important property of them.

**Lemma 3.3.3.** A univariate polynomial f is real stable if and only if it is real rooted.

*Proof.* It's a given that when f is real rooted it is also real stable, because real rooted is simply real stable with added constraint that there can't be any roots with a negative imaginary part.

Now for when we already now f is real stable. Let's assume we have found a function that is real stable, but not real rooted.  $f(z) = \sum_{j=0}^{n} c_j z^j$  with  $c_j \in \mathbb{R}$  and root a - bi where b is positive.

$$0 = f(a - bi) = \overline{f(a - bi)}$$
$$= \sum_{j=0}^{n} \overline{c_j (a - bi)^j}$$
$$= \sum_{j=0}^{n} c_j \overline{(a - bi)^j}$$
$$= \sum_{j=0}^{n} c_j (a + bi)^j$$
$$= f(a + bi)$$

Now f has a root with a positive imaginary part, this is in contradiction with f being real stable. So when f is real stable, it has to also be real rooted.

Now to create some real stable polynomials of our own. For this we will use Proposition 2.4 of [1].

**Proposition 3.3.4** (Proposition 2.4 of [1]). When  $A_1, \ldots, A_m$  are positive semidefinite matrices, then

$$f(z_1, ..., z_m) = \det\left(\sum_{i=1}^m z_i A_i\right)$$

is real stable.

Proof. We start by proving that if f is stable, it is real stable. This is because an alternative way of calculating the determinant of a matrix is by multiplying all its eigenvalues. Since  $\sum_{i=1}^{n} z_i A_i$  is a Hermitian matrix when  $z \in \mathbb{R}^m$ , we can now make use of that Hermitian matrices only have real eigenvalues, that means f also only gives real values when  $(z_1, ..., z_m) \in \mathbb{R}^m$ . Hence, we can conclude that all coefficients of f also must be real. Now only the stability is left to be proven, for this we start with the case that all  $A_i$  are positive definite. By using Lemma 3.3.2 we can simplify  $f(z_1, ..., z_m)$  to a univariate polynomial  $f(\alpha + vt)$ , where  $\alpha \in \mathbb{R}^m$ ,  $v \in \{x > 0 | \mathbb{R}\}^m$  and  $t \in \mathbb{C}$ . Now we can start rewriting  $f(\alpha + vt)$  in a form that is more useful to us. For this we will use Lemma 2.2.3 which says there is a positive definite inverse for positive definite matrices, Lemma 2.2.4 tell us that summations of positive definite matrices are also positive semidefinite and 2.2.6 tells us that there are exist unique positive definite square roots for all positive definite matrices.

$$\begin{aligned} f(\alpha + vt) &= \det\left(\sum_{i=1}^{n} (\alpha_i + v_i t) A_i\right) \\ &= \det\left(t\sum_{i=1}^{n} v_i A_i + \sum_{i=1}^{n} \alpha_i A_i\right) \\ &= \det\left(\left(\sum_{i=1}^{n} v_i A_i\right)^{\frac{1}{2}}\right) \det\left(\left(t\sum_{i=1}^{n} v_i A_i\right)^{\frac{1}{2}} + \left(\sum_{i=1}^{n} \alpha_i A_i\right) \left(\sum_{i=1}^{n} v_i A_i\right)^{-\frac{1}{2}}\right) \\ &= \det\left(\sum_{i=1}^{n} v_i A_i\right) \det\left(tI + \left(\sum_{i=1}^{n} v_i A_i\right)^{-\frac{1}{2}} \left(\sum_{i=1}^{n} \alpha_i A_i\right) \left(\sum_{i=1}^{n} v_i A_i\right)^{-\frac{1}{2}}\right) \\ &= \det\left(P\right) \det\left(tI + P^{-\frac{1}{2}} H P^{-\frac{1}{2}}\right) \\ &= \det\left(P\right) \det\left(tI + P^{-\frac{1}{2}} H \left(P^{-\frac{1}{2}}\right)^*\right) \end{aligned}$$

The function, which we are now left with is a constant multiplied by the characteristic polynomial of  $P^{-\frac{1}{2}}H\left(P^{-\frac{1}{2}}\right)^*$ , so the eigenvalues of that matrix are its roots. Since  $P^{-\frac{1}{2}}H\left(P^{-\frac{1}{2}}\right)^*$  is positive definite, it is also Hermitian. This means that  $f(\alpha + vt)$  only has real roots and that  $f(z_1, ..., z_m)$  is real stable according to Lemma 3.3.2.

To prove it is also true when  $A_i$  is only positive semidefinite, we need to use the Hurwitz' theorem (3.1.1). Let  $D = \{(z_1, ..., z_m) \in \mathbb{C}^m | \text{Im}(x_i) > 0 \text{ for all } i\}$  and  $\{f_k\}_{k=1}^{\infty}$  be a sequence of functions that are created the same way as f, but with only positive definite matrices, so  $f_k$  is nonvanishing and analytic on D. Now since  $\{f_k\}_{k=1}^{\infty}$  can converge to f, it means that f is nonvanishing on D or identical to zero in other words stable.

To further increase our possibilities with real stable polynomials, we will need to prove that they stay real stable after having been multiplied by  $(1 - \partial_{z_i})$ . To prove this, we first need the following lemma.

**Lemma 3.3.5.** Let  $q \in \mathbb{C}[z]$  be stable, then  $q(z) - \partial_z q(z)$  is also stable.

*Proof.* Define  $q(z) = a_0(z-a_1)...(z-a_n)$  this means that  $\partial_z q(z) = a_0 \sum_{i=1}^n ((z-a_1)...(z-a_{i-1})(z-a_{i+1})...(z-a_n))$ . In this proof, we can ignore cases where  $a_i$  is a root of both q and  $q(z) - \partial_z q(z)$ , because we then already know that  $a_i$  is a stable root. In order to ignore them, we will simply remove them from the domain

of the functions.

We may then assume that  $\frac{q-\partial_z q}{q}$  has the same roots as  $q-\partial_z q$ .

$$\frac{q(z) - \partial_z q(z)}{q(z)} = 1 - \frac{\partial_z q(z)}{q(z)}$$
$$= 1 - \sum_{i=1}^n \frac{1}{z - a_i}$$
$$= 1 - \sum_{i=1}^n \frac{1}{z - a_i} \frac{\overline{z} - \overline{a_i}}{\overline{z} - \overline{a_i}}$$
$$= 1 - \sum_{i=1}^n \frac{\overline{z} - \overline{a_i}}{|z - a_i|^2}$$

When  $(\partial_z q - q)(z)$  is zero, we know that this equation has to be upheld.

$$\begin{split} 0 &= 1 - \sum_{i=1}^{n} \frac{\overline{z} - \overline{a_i}}{|z - a_i|^2} \\ \overline{z} \sum_{i=1}^{n} \frac{1}{|z - a_i|^2} &= 1 + \sum_{i=1}^{n} \frac{\overline{a_i}}{|z - a_i|^2} \\ z \sum_{i=1}^{n} \frac{1}{|z - a_i|^2} &= 1 + \sum_{i=1}^{n} \frac{a_i}{|z - a_i|^2} \end{split}$$

Since  $\sum_{i=1}^{n} \frac{1}{|z-a_i|^2}$  has a real positive value and  $\operatorname{Im}\left(1+\sum_{i=1}^{n} \frac{a_i}{|z-a_i|^2}\right) \leq 0$ , hence all roots of  $q - \partial_z q$  also have to be stable.

**Theorem 3.3.6.** Let  $p \in \mathbb{R}[z_1, ..., z_m]$  be real stable, then  $(1 - \partial_{z_i})p(z_1, ..., z_m)$  is also real stable.

Proof. Define

$$q(z_i) = p(z_1, \dots, z_m)|_{z_1 = x_1, \dots, z_{i-1} = x_{i-1}, z_{i+1} = x_{i+1}, \dots, z_m = x_m}$$

where  $x_1, ..., x_{i-1}, x_{i+1}, ..., x_m$  have a positive imaginary part. According to the previous lemma (3.3.5) since  $q(z_i)$  is stable,  $(1 - \partial_{z_i})q(z_i)$  is also stable. This implies that  $(1 - \partial_{z_i})p(z_1, ..., z_m)$  has no roots where all values have a positive imaginary part, hence is stable. It is also real stable, because you can't create complex values by taking the derivative of a polynomial.

**Theorem 3.3.7** (Lemma 2.4(d) of [12]). Let  $f \in \mathbb{R}[z_1, ..., z_m]$  be real stable and  $a \in \mathbb{R}$ , then  $f|_{z_i=a} \in \mathbb{R}[z_1, ..., z_{i-1}, z_{i+1}, ..., z_m]$  with  $i \in [m]$  is also real stable.

*Proof.* For this we will again need to use Hurwitz' Theorem (3.1.1). Let  $D = (z_1, ..., z_m) \in \{\mathbb{C}^m | \operatorname{Im}(x_i) > 0 \text{ for all } i\}$  and  $\{f_k\}_{k=1}^{\infty} = \{f|_{z_i=a+i2^{-k}}\}_{k=1}^{\infty}$ . Every  $f_k$  is real stable since  $z_i$  gets a value with a positive imaginary part. Since this sequence converges to  $f|_{z_i=a}$ , so  $f|_{z_i=a}$  must be stable. Also since f's coefficients are real and a is real,  $f|_{z_i=a}$  is real stable.

#### 3.4 Determinants

Here we will focus on some properties of determinants.

**Lemma 3.4.1** (Lemma 1.1 of [5]). When A is an invertible  $n \times n$  matrix and u, v are vectors, then

$$\det(A + uv^*) = \det(A)(1 + v^*A^{-1}u)$$

*Proof.* The first step is to separate the determinant of A, which gets us  $\det(A) \det(I + (A^{-1}uv^*))$ . Now the only thing left to prove is that  $\det(I + A^{-1}uv^*) = (1 + v^*A^{-1}u)$ . The first step is to create a larger  $(n+1) \times (n+1)$  matrix with the same determinant.

$$\det(I + A^{-1}uv^*) = \begin{pmatrix} I + A^{-1}uv^* & A^{-1}u \\ 0 & 1 \end{pmatrix}$$

This matrix should be read as the matrix  $I + A^{-1}uv^*$  with column vector  $A^{-1}u$  pressed against it and with a bottom row added which is filled with zeros except the last value. Since the bottom row is only zero except the last value, all values in the vector  $A^{-1}u$  are ignored in the determinant. Now we can go and add some more matrices with a determinant value of 1.

$$\det \begin{pmatrix} I + A^{-1}uv^* & A^{-1}u \\ 0 & 1 \end{pmatrix} = \det \begin{pmatrix} I & 0 \\ v^* & 1 \end{pmatrix} \det \begin{pmatrix} I + A^{-1}uv^* & A^{-1}u \\ 0 & 1 \end{pmatrix} \det \begin{pmatrix} I & 0 \\ -v^* & 1 \end{pmatrix}$$
$$= \det \begin{pmatrix} I & 0 \\ v^* & 1 \end{pmatrix} \det \begin{pmatrix} I & A^{-1}u \\ -v^* & 1 \end{pmatrix}$$
$$= \det \begin{pmatrix} I & A^{-1}u \\ 0 & v^*A^{-1}u + 1 \end{pmatrix}$$
$$= 1 + v^*A^{-1}u$$

There you have it.

Let A(t) be a differentiable function from  $\mathbb{C} \to \mathbb{C}^{n \times n}$ . We could turn this back into a value in  $\mathbb{C}$  by taking the determinant of it. Now one might wonder what the derivative of this might be, Jacobi's formula was created to simplify this process.

**Theorem 3.4.2** (Jacobi's formula, theorem 8.3 of [8]). When  $A(t) : \mathbb{C} \to \mathbb{C}^{n \times n}$ , then  $\partial_t \det(A) = \operatorname{Tr}(\operatorname{adj}(A)(\partial_t A))$ 

*Proof.* We will first define the determinant as a function, so that we can use the chain rule on it.

$$det(X) = det \begin{pmatrix} X_{1,1} & X_{1,2} & \cdots & x_{1,n} \\ X_{2,1} & X_{2,2} & \vdots \\ \vdots & \ddots & \vdots \\ X_{n,1} & \cdots & \cdots & X_{n,n} \end{pmatrix}$$
$$= det(X_{1,1}, X_{1,2}, X_{1,3}, \dots, X_{2,1}, X_{2,2}, X_{2,3}, \dots, X_{n,n})$$
$$= \sum_{j=1}^{n} X_{i,j} adj^{t}(X)_{i,j}$$

where  $X \in \mathbb{C}^{n \times n}$ . Here the determinant is written using the Laplace expansion, where *i* is an arbitrary row. Now we can apply the chain rule.

$$\partial_t \det(A) = \sum_{i=1}^n \sum_{j=1}^n (\partial_{X_{i,j}} \det(A)) (\partial_t A)_{i,j}$$

Look at the derivative of the determinant, important here is that  $adj^{t}(A)_{i,j}$  doesn't use the value of  $A_{i,j}$ .

n

$$\partial_{X_{i,j}} \det(A) = \partial_{X_{i,j}} \sum_{j=k}^{n} A_{ik} \operatorname{adj^{t}}(A)_{ik}$$
$$= \sum_{j=k}^{n} \partial_{X_{i,j}} A_{i,k} \operatorname{adj^{t}}(A)_{i,k}$$
$$= \partial_{X_{i,j}} A_{ij} \operatorname{adj^{t}}(A)_{i,j} \text{ (if } j \neq k, \text{ then that part doesn't use } X_{i,j})$$
$$= \operatorname{adj^{t}}(A)_{i,j}$$

Now by only having used the chain rule we get

$$\partial_t \det(A) = \sum_{i=1}^n \sum_{j=1}^n \operatorname{adj}^{\mathsf{t}}(A)_{i,j} (\partial_t A)_{i,j}$$

By using how the trace is taken of two matrices that are multiplied by each other, we get

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \operatorname{adj}^{t}(A)_{i,j}(\partial_{t}A)_{i,j} = \operatorname{Tr}(\operatorname{adj}(A)(\partial_{t}A))$$

**Corollary 3.4.3.** Let A(t) = B + Ct, where B is an invertible square matrix and C a matrix of the same size. If the value of t is equal to 0 then

$$\partial_t \det(B + Ct)|_{t=0} = \det(B) \operatorname{Tr}(B^{-1}C).$$

Similarly, when we only know that B + Ct is invertible

$$\partial_t \det(B + Ct) = \det(B + Ct) \operatorname{Tr}((B + Ct)^{-1}C).$$

*Proof.* This proof is near trivial due the previous theorem (3.4.2).

$$\begin{aligned} \partial_t \det(B + Ct)|_{t=0} &= \operatorname{Tr}(\operatorname{adj}(B + Ct)(\partial_t(B + Ct)))|_{t=0} \\ &= \operatorname{Tr}(\operatorname{adj}(B + Ct)C)|_{t=0} \\ &= \operatorname{Tr}(\operatorname{adj}(B)C) \\ &= \operatorname{Tr}(\det(B)B^{-1}C) \text{ (making use of that } B^{-1} = \frac{\operatorname{adj}(B)}{\det(B)}) \\ &= \det(B)\operatorname{Tr}(B^{-1}C) \end{aligned}$$

The proof for when we only know that B + Ct is invertible is nearly the same and is left as an exercise to the reader.

# 4 Mixed Characteristic Polynomials

In this chapter, we will explore mixed characteristic polynomials. This is a polynomial particularly created for solving the Weaver conjecture.

**Definition 4.0.1.** Let  $A_1, ..., A_m$  be  $n \times n$  matrices. We then call  $\mu[A_1, ..., A_m](x)$  the mixed characteristic polynomial of  $A_1, ..., A_m$ , which is the function given by

$$\mu[A_1, ..., A_m](x) = \left(\prod_{i=1}^m 1 - \partial_{z_i}\right) \det\left(xI + \sum_{i=1}^m z_i A_i\right) \Big|_{z_1 = \dots = z_m = 0}$$

We are particularly interested in mixed characteristic polynomials when all  $A_i$  are covariance matrices, what covariance matrices are we will define now.

**Definition 4.0.2.** Let v be a random column vector taken from a distribution, then the covariance of its possible values is shown by matrix C also known as the covariance matrix. Define

$$C = \mathbb{E} v v^*$$

**Definition 4.0.3.** For distributions in thesis we will only use finite support, meaning that there is a finite set of possible values the vector v can take. Each of those values has its own probability of occurring.

**Theorem 4.0.4.** Let  $v_1, ..., v_m \in \mathbb{C}^n$  be independent random column vectors with finite support and set  $A_i = \mathbb{E} v_i v_i^*$ . Then

$$\mathbb{E}\chi\left[\sum_{i=1}^{m} v_{i}v_{i}^{*}\right](x) = \left(\prod_{i=1}^{m} 1 - \partial_{z_{i}}\right)\det\left(xI + \sum_{i=1}^{m} z_{i}A_{i}\right)\Big|_{z_{1}=\ldots=z_{m}=0} = \mu[A_{1},\ldots,A_{m}](x)$$
(3)

To prove the above Theorem 4.0.4, we will need an operator which can show the relation between expected values and partial derivatives.

**Lemma 4.0.5.** For every square matrix  $M \in \mathbb{C}^{n \times n}$  and random column vector  $v \in \mathbb{C}^n$ , we have

$$\mathbb{E}\det\left(M - vv^*\right) = (1 - \partial_t)\det\left(M + t\,\mathbb{E}\,vv^*\right)|_{t=0}.$$
(4)

*Proof.* We will start with the case that M is invertible.

$$\mathbb{E} \det (M - vv^*) = \mathbb{E} \det(A)(1 - v^*M^{-1}v) \text{ (Using Lemma 3.4.1)}$$

$$= \mathbb{E} \det(M)(1 - v^*(M^{-1}v))$$

$$= \mathbb{E} \det(M)(1 - \operatorname{Tr}((M^{-1}v)v^*))$$

$$= \det(M) \mathbb{E}(1 - \operatorname{Tr}(M^{-1}vv^*))$$

$$= \det(M) - \det(M) \mathbb{E} \operatorname{Tr}(M^{-1}vv^*)$$

$$= \det(M) - \det(M) \operatorname{Tr}(M^{-1} \mathbb{E} vv^*) \text{ (Possible, because trace is linear)}$$

We will now be adding the t that is set to zero.

$$\det(M) - \det(M)\operatorname{Tr}(M^{-1} \mathbb{E} vv^*) = \det(M + t \mathbb{E} vv^*)|_{t=0} - \det(M)\operatorname{Tr}(M^{-1} \mathbb{E} vv^*)$$
$$= \det(M + t \mathbb{E} vv^*)|_{t=0} - \partial_t \det(M + t \mathbb{E} vv^*)|_{t=0} \text{ (Using Corollary 3.4.3)}$$
$$= (1 - \partial_t) \det(M + t \mathbb{E} vv^*)|_{t=0}$$

Now we have proven that the lemma holds for the case that M is invertible. So now it is time for the case where M isn't invertible.

For this we will create function N(t) = M - tI. When  $t \to 0$ , then  $N(t) \to M$ . If t isn't equal to an eigenvalue of M, we know that N(t) is invertible. If we now put N(t) in both sides of Equation (4), we can use the fact that both sides are then polynomials. So, continuity then implies that it also works for non invertible matrices.

*Proof of Theorem 4.0.4.* In order to prove Equation (3) inductively we will need to create a more generic version of it. This is done by rewriting it to this

$$\mathbb{E}\det\left(M-\sum_{i=1}^{m}v_{i}v_{i}^{*}\right)=\left(\prod_{i=1}^{m}1-\partial_{z_{i}}\right)\det\left(M+\sum_{i=1}^{m}z_{i}A_{i}\right)\Big|_{z_{1}=\ldots=z_{m}=0},$$

where M is an arbitrary matrix in  $\mathbb{C}^{d \times d}$ . Start with the base case that m = 0. This case is rather trivial due to all the summation and products being gone.

$$\mathbb{E}\det(M) = \det(M)$$

Now the proof for when m > 0 and it works for m - 1. (Notation note: here we will use  $\mathbb{E}$  with an under script, to show that the expected value is taken at a different time)

$$\mathbb{E} \det \left( M - \sum_{i=1}^{m} v_i v_i^* \right) = \mathbb{E}_{v_1, \dots, v_{m-1}} \mathbb{E}_{v_m} \det \left( M - \sum_{i=1}^{m-1} v_i v_i^* - v_m v_m^* \right) \text{ (possible due to independence)}$$

$$= \mathbb{E}_{v_1, \dots, v_{m-1}} (1 - \partial_{z_m}) \det \left( M + z_m A_m - \sum_{i=1}^{m-1} v_i v_i^* \right) \Big|_{z_m = 0} \text{ (Lemma 4.0.5)}$$

$$= (1 - \partial_{z_m}) \mathbb{E}_{v_1, \dots, v_{m-1}} \det \left( (M + z_m A_m) - \sum_{i=1}^{m-1} v_i v_i^* \right) \Big|_{z_m = 0}$$

$$= (1 - \partial_{z_m}) \left( \prod_{i=1}^{m-1} 1 - \partial_{z_i} \right) \det \left( M + z_m A_m + \sum_{i=1}^{m-1} z_i A_i \right) \Big|_{z_1 = \dots = z_{m-1} = 0} \Big|_{z_m = 0}$$

$$= \left( \prod_{i=1}^{m} 1 - \partial_{z_i} \right) \det \left( M + \sum_{i=1}^{m-1} z_i A_i \right) \Big|_{z_1 = \dots = z_m = 0}$$

Now by replacing M with xI, we get that

$$\left(\prod_{i=1}^{m} 1 - \partial_{z_i}\right) \det\left(xI + \sum_{i=1}^{m-1} z_i A_i\right) \Big|_{z_1 = \dots = z_m = 0} = \mathbb{E} \det\left(xI - \sum_{i=1}^{m} v_i v_i^*\right) = \mathbb{E} \chi\left[\sum_{i=1}^{m} v_i v_i^*\right](x).$$

**Corollary 4.0.6.** When  $A_1, ..., A_m$  are positive semidefinite, then  $\mu[A_1, ..., A_m]$  is real rooted.

Proof. We start by using Theorem 4.0.4

$$\mu[A_1, ..., A_m](x) = \left(\prod_{i=1}^m 1 - \partial_{z_i}\right) \det\left(xI + \sum_{i=1}^{m-1} z_i A_i\right) \Big|_{z_1, ..., z_m = 0}$$

Since *I* is positive definite, we can use Proposition 3.3.4 to show that det  $\left(xI + \sum_{i=1}^{m-1} z_i A_i\right)$  is real stable, then Theorem 3.3.6 to show, that  $\left(\prod_{i=1}^{m} 1 - \partial_{z_i}\right) \det \left(xI + \sum_{i=1}^{m-1} z_i A_i\right)$  maintains the real stability and then by inductively using Theorem 3.3.7 we get that  $\left(\prod_{i=1}^{m} 1 - \partial_{z_i}\right) \det \left(xI + \sum_{i=1}^{m-1} z_i A_i\right) \Big|_{z_1,\dots,z_m=0}$  is also real stable. Now to finish this off we will use Lemma 3.3.3, which tells us that being real stable and real rooted is the same thing for univariate functions.

Now to be able to use the previous Corollary 4.0.6, we need to prove it also works for  $\mathbb{E} vv^*$  matrices.

**Lemma 4.0.7.** Let A be equal to  $\mathbb{E}vv^*$ , where  $v \in \mathbb{C}^n$  is a random column vector with finite support, then A is positive semidefinite.

*Proof.* Let  $u \in \mathbb{C}^n$  be one of the finite possibilities for v. Since  $uu^*$  can be seen as the Gram matrix of a collection of vectors in  $\mathbb{C}^1$ , it follows from Lemma 2.2.7 that  $uu^*$  is positive semidefinite. Since there is only a finite amount of possibilities for u, the value of  $\mathbb{E}vv^*$  can be calculated by a finite summation. So, then according to Lemma 2.2.4 A is also positive semidefinite.

Now we can start using the real rooted property of mixed characteristic polynomials to prove that they can form interlacing families using the values of its vectors. For this we will need to set some values related to finite support. We will set  $l_i$  equal to the amount of possible values for  $v_i$ , with  $w_{i,1}, ..., w_{i,l_i}$  being the possible values. Each of those values has a probability of  $p_{i,1}, ..., p_{i,l_i}$  to occur. Let  $j_i \in [l_i]$ , define

$$q_{j_1,\dots,j_m} = \left(\prod_{i=1}^m p_{i,j_i}\right) \chi \left[\sum_{i=1}^m w_{i,j_i} w_{i,j_i}^*\right]$$

**Theorem 4.0.8.** The polynomials  $q_{j_1,...,j_m}$  form an interlacing family.

*Proof.* For  $k \in [m-1]$ , define

$$q_{j_1,\dots,j_k} = \left(\prod_{i=1}^k p_{i,j_i}\right) \underset{v_{k+1},\dots,v_m}{\mathbb{E}} \chi \left[\sum_{i=1}^k w_{i,j_i} w_{i,j_i}^* + \sum_{i=k+1}^m v_i v_i^*\right].$$

Also, let

$$q_{\emptyset} = \mathop{\mathbb{E}}_{v_1, \dots, v_m} \chi \left[ \sum_{i=1}^m v_i v_i^* \right].$$

We now need to prove for all possible  $j_1, ..., j_k \in [l_1] \times ... \times [l_k]$  that the collection

$$\{q_{j_1,...,j_k,t}\}$$
 with  $t \in [l_{k+1}]$ 

has a common interlacing. Due to Theorem 3.2.6 we can instead choose to prove that the convex combination are real rooted. In order to show all convex combinations, let  $\lambda_1, ..., \lambda_{l_{k+1}}$  be nonnegative numbers of which the sum is equal to 1. This gives us the polynomial

$$\sum_{t=1}^{l_{k+1}} \lambda_i q_{j_1,\dots,j_k,t}(x).$$

Now all that is left, is to show that it is real rooted. To prove this, let  $u_{k+1}$  be a random vector that is equal to  $w_{k+1,t}$  with a probability of  $\lambda_t$ . We then get that  $\sum_{t=1}^{l_{k+1}} \lambda_i q_{j_1,\ldots,j_k,t}(x)$  equals

$$\left(\prod_{i=1}^{k} p_{i,j_i}\right) \underset{u_{k+1}, v_{k+1}, \dots, v_m}{\mathbb{E}} \left[\sum_{i=1}^{k} w_{i,j_i} w_{i,j_i}^* + u_{k+1} u_{k+1}^* + \sum_{i=k+2}^{m} v_i v_i^*\right] (x).$$

Since we can treat  $w_{i,j_i}$  as a random vector from a collection that only has 1 value, we can treat the above polynomial as mixed characteristic polynomial multiplied by a real constant. According to Corollary 4.0.6 combined with Lemma 4.0.7, it is real rooted. So, every convex combination is real rooted, hence  $q_{j_1,\ldots,j_m}$  is an interlacing family.

# 5 Largest Root

This chapter will focus on proving the existence of an upper bound on the roots of mixed characteristic polynomials with certain properties.

**Theorem 5.0.1.** Suppose we have Hermitian positive semidefinite matrices  $A_1, ..., A_m$ , which satisfy  $\sum_{i=1}^m A_i = I$  and  $\operatorname{Tr}(A_i) < \epsilon$  for all *i*. Then the largest root of  $\mu[A_1, ..., A_m](x)$  is at most  $(1 + \sqrt{\epsilon})^2$ .

First, we will prove how the  $\sum_{i=1}^{m} A_i = I$  property can be applied in our favour.

**Lemma 5.0.2.** Let  $A_1, ..., A_m$  be Hermitian positive semidefinite matrices. If  $\sum_{i=1}^m A_i = I$ , then

$$\mu[A_1, \dots A_m](x) = \left(\prod_{i=1}^m 1 - \partial_{y_i}\right) \det\left(\sum_{i=1}^m y_i A_i\right) \Big|_{y_1 = \dots = y_m = x}.$$
(5)

Proof.

$$\mu[A_1, \dots A_m](x) = \left(\prod_{i=1}^m 1 - \partial_{z_i}\right) \det \left(xI + \sum_{i=1}^m z_i A_i\right) \Big|_{z_1 = \dots = z_m = 0}$$
$$= \left(\prod_{i=1}^m 1 - \partial_{z_i}\right) \det \left(x\sum_{i=1}^m A_i + \sum_{i=1}^m z_i A_i\right) \Big|_{z_1 = \dots = z_m = 0}$$
$$= \left(\prod_{i=1}^m 1 - \partial_{z_i}\right) \det \left(\sum_{i=1}^m (x + z_i) A_i\right) \Big|_{z_1 = \dots = z_m = 0}$$
$$= \left(\prod_{i=1}^m 1 - \partial_{y_i}\right) \det \left(\sum_{i=1}^m y_i A_i\right) \Big|_{y_1 = \dots = y_m = x}$$

The last step is possible because x can be treated as a constant to the partial derivatives that are being taken.

**Definition 5.0.3.** Let  $p(z_1, ..., z_m)$  be a multivariate polynomial, then  $z \in \mathbb{R}^m$  is above the roots of p when for all  $t = (t_1, ..., t_m) \in \mathbb{R}^m$  where  $t_i \ge 0$ 

$$p(z+t) \neq 0.$$

The collection of all the points that fulfil the previous definition are denoted as  $Ab_p$ .

To prove Theorem 5.0.1, we will need to show that  $(1 + \sqrt{\epsilon})^2 \mathbf{1}_m \in Ab_Q$ . For this, we will use induction, by slowly building up to the polynomial Q and proving that for each step the property  $(1 + \sqrt{\epsilon})^2 \mathbf{1}_m \in Ab_Q$  is true.

To prove Theorem 5.0.1, we will start working with the multivariate polynomial occurring in the right side of Equation (5). We will call this part

$$Q(y_1, ..., y_m) = \left(\prod_{i=1}^m 1 - \partial_{y_i}\right) \det\left(\sum_{i=1}^m y_i A_i\right).$$

The reason why this is useful, is that when  $(u_1, ..., u_m)$  is above the roots of Q, then  $\max_{i=1}^m (u_i)$  is also above the roots of  $\mu[A_1, ..., A_m]$ .

**Definition 5.0.4.** Let  $p(z_1, ..., z_m)$  be a real stable polynomial and  $z = (z_1, ..., z_m)$  an element of Ab<sub>p</sub>. Then the barrier function of p in direction i at z is

$$\Phi_p^i(z) = \frac{(\partial_{z_i} p(z))}{p(z)}$$
$$= \partial_{z_i} \log \pm p(z) \text{ (reverse chain rule)}$$

with the  $\pm$  always being such that  $\pm p(z) > 0$ . The value of  $\pm$  also won't need to change when the value of z changes, because z is above the roots of p.

**Lemma 5.0.5.** Analogous to  $\Phi_p^i(z)$  there is the univariate version, for in the case we are only interested in one variable.

$$\Phi_{p,z}^i(z_i) = \sum_{j=1}^r \frac{1}{z_i - \lambda_j}$$

where  $\lambda_i, ..., \lambda_r$  are the roots of univariate polynomial

$$q_{z,i}(t) = p(z_1, ..., z_{i-1}, t, z_{i+1}, ..., z_m).$$
(6)

Those roots are real, as a result of Theorem 3.3.7.

*Proof.* The polynomial  $q_{z,i}$  can be written as  $q_{z,i}(t) = \lambda_0 \prod_{j=1}^r (t - \lambda_j)$ , so  $\partial q_{z,i} = \lambda_0 \sum_{k=1}^r (t - \lambda_1) \dots (t - \lambda_{k-1})(t - \lambda_{k+1}) \dots (t - \lambda_r)$  according to the product rule. Now we can conclude that

$$\Phi_{p,z}^{i}(z_{i}) = \frac{(\partial q_{z,i}(z_{i}))}{q_{z,i}(z_{i})} = \sum_{j=1}^{r} \frac{1}{z_{i} - \lambda_{j}}.$$

Now some important properties of the barrier function, which can later be used inductively.

**Theorem 5.0.6.** Let  $p \in \mathbb{R}[z_1, ..., z_m]$  be real stable and  $z \in Ab_p$ , then for all  $i, j \in [m]$  and  $\delta \geq 0$ 

 $\Phi_p^i(z+\delta e_j) \le \Phi_p^i(z)$  (non-increasing)

and

$$\Phi_p^i(z+\delta e_j) \le \Phi_p^i(z) + \delta \partial_{z_j} \Phi_p^i(z+\delta e_j) \quad (convexity).$$
(7)

Also, that  $\Phi_p^i(z) > 0$ .

The following proof is based on Lemma 17 of [10].

*Proof.* We will start with the case that i = j. Since we are only interested in one variable of p, we can use Lemma 5.0.5  $\Phi_p^i(z) = \sum_{k=1}^r \frac{1}{z_i - \lambda_k}$  where  $\lambda_k$  is real. Since  $z \in Ab_p$ , we know that  $z_i > \lambda_k$  for all k, we can apply this further giving us  $0 < \frac{1}{z_i - \lambda_k} > \frac{1}{z_i + \delta - \lambda_k}$  from this the non-increasing property follows and that  $\Phi_p^i(z) > 0$ .

Now the convexity property. We can do this by first generalizing taking the derivative of  $\Phi_p^i(z)$ .

$$\partial_{z_i}^l \Phi_p^i(z) = (-1)^l l! \sum_{k=1}^r \frac{1}{(z_i - \lambda_k)^{l+1}}$$
(8)

for integer l > 0. If it is positive or not completely depends on the value of l. Since we already know that  $\partial z_i \Phi_p^i(z)$  is non-increasing and from Equation (8). We can deduce that the convexity property also must be true.

The proof for the case when  $i \neq j$ . Since we are now only interested in two variables, we can turn p into a bivariate function without losing generality.

$$q_{z,i,j}(z_i, z_j) = p(z_1, ..., z_m)$$

To achieve our goal, we will first have to prove a generalization, that

$$(-1)^{l} \partial_{z_{j}}^{l} \Phi_{q_{z,i,j}}^{i}(z_{i}, z_{j}) \ge 0$$
(9)

where l > 0. To make the problem a bit more understandable

$$(-1)^{l}\partial_{z_{j}}\Phi_{q_{z,i,j}}(z_{i},z_{j})^{i} = (-1)^{l}\partial_{z_{j}}^{l}\partial_{z_{i}}\log\pm q_{z,i,j}(z_{i},z_{j}) \text{ (definition 5.0.4)}$$
$$=\partial_{z_{i}}\left((-1)^{l}\partial_{z_{j}}^{l}\log\pm q_{z,i,j}(z_{i},z_{j})\right) \text{ (possible, because } z \text{ is above the roots of } q_{z,i,j})$$

So, we need to proof that  $(-1)^l \partial_{z_i}^l \log \pm q_{z,i,j}(z_i, z_j)$  is non-decreasing in  $z_i$ .

For this we will make use of the fact that polynomial  $q_{z,j}$  from Equation (6) is real rooted and that the roots of a polynomial are a continuous function of its coefficients. Now we can set  $y_1(x_i), ..., y_d(x_i)$  to be the roots of  $q_{z,j}$ , where d is constant. We then get,

$$(-1)^{l} \partial_{z_{j}}^{l} \log \pm q_{z,i,j}(z_{i}, z_{j}) = -(l-1)! \sum_{k=1}^{d} \frac{1}{(x_{j} - y_{k}(x_{i}))^{l}}.$$

We can now prove Inequality (9) by showing that the value of  $\frac{1}{(z_j - y_k(z_i))^l}$  is non-increasing for all k in  $z_i$ . It makes sense that when the value of  $z_i$  increases, then the value of  $y_k(z_i)$  decreases, because we are getting further away from the roots, causing  $\frac{1}{(x_j - y_k(x_i))^l}$  to decrease in value. Let's assume this isn't the case and that  $y_k(x_i)$  increases, meaning that  $y_k(x_i)$  has a positive derivative for all  $x_1$  in a neighbourhood. Then there exists a point  $x^0 \in \mathbb{R}^m$  such that the derivative of  $y_k(x^0)$  is positive and that for all  $x_i$  sufficiently close enough to  $x^0, y_k(x_i)$  has a positive derivative and there exists a local constant multiplicity m (meaning that there are m roots for  $q_{z,j}$  in the point  $y_k(x_i)$  for all  $x_i$  that are close enough).

Now we have 2 different cases. We will start with the case that m is 1, because of this we know that polynomial  $t \to q_{z,i,j}(x^0, t)$  has to be increasing or decreasing for value  $t = y_k(x^0)$ , otherwise the function would be constant. Since  $q_{z,i,j}(x^0, y_k(x^0)) = 0$  and because  $q_{z,i,j}(x^0, x_j)$  is analytic (all polynomials are analytic), we can use the implicit function theorem (Theorem 8.6 of [7]) extended for complex values.

The implicit function theorem extended to complex values tells us that when f(x, y) is an analytic function,  $x^0, y^0$  in  $\mathbb{C}^2$  have a neighbourhood, where  $x^0, y^0$  are chosen such that  $f(x^0, y^0) = 0$  and

$$\partial_y f|_{(x^0, y^0)} \neq 0.$$

Then there exist an analytic function  $g :\subseteq \mathbb{C} \to \mathbb{C}$  such that f(x, g(x)) = 0. Applying this to what we just have proven, we get that  $y_k$  is an analytic function. Now we can use the Cauchy-Riemann Equation (Opmerking 4.31 of [11]), which states when  $f : \mathbb{U} \to \mathbb{C}$  is an analytic function ( $\mathbb{U}$  is an open subset of  $\mathbb{C}$ ), it can then be rewritten as

$$f(x+yi) = u(x,y) + iv(x,y)$$

where  $u, v : \mathbb{R}^2 \to \mathbb{R}$ , with following properties. That  $\partial_x u = \partial_y v$  and  $\partial_y u = -\partial_x v$ .

From this we learn that because  $y_k$  has positive derivative for real values in  $x^0$ , it also must have a positive derivative for imaginary values in  $x^0$ . So, there has to be a t > 0 such that  $\text{Im}(y_k(x^0 + ti)) > 0$ , meaning that  $q_{z,i,j}$  has a root for which both  $x_i$  and  $x_j$  have a positive imaginary value. This contradicts that  $q_{z,i,j}$  is real stable, so  $y_k$  has to be non-increasing when it has multiplicity 1.

For the case that m isn't 1, we can do almost the exact same thing. Only now we need to use  $\partial_{z_j}^{m-1}q_{z,i,j}$  instead of  $q_{z,i,j}$  in the implicit function theorem. This works, because  $\partial_{z_j}^{m-1}q_{z,i,j}(x_i, y_k(x_i)) = 0$  and the multiplicity of  $y_k(x_i)$  for this polynomial is also 1.

So, now we have successfully proven Equation (9), from that immediately follows the non-increasing and convexity property of  $\Phi_p^i$ .

**Lemma 5.0.7.** Let  $p(z_1, ..., z_m)$  be a real stable polynomial. When  $z \in Ab_p$  and  $\Phi_p^i(z) < 1$ , then  $z \in Ab_{p-\partial_{z_i}p}$ .

*Proof.* Using the previous Theorem 5.0.6, we know that  $0 < \Phi_p^i(z) < 1$ . From this we learn that for every nonnegative vector t, when p(z) is positive also p(z+t) and  $\partial_{z_i}p(z+t)$  are positive. The same is true the other way around when p(z) is negative. Since the proof for both cases is near identical, we will only do the case where p(z) is positive. Using the non-increasing property of  $\Phi_p^i$  we get that  $\Phi_p^i(z+t) < 1$ . This leads to

$$\Phi_p^i(z+t) = \frac{\partial_{z_i} p(z+t)}{p(z+t)} < 1$$

which is equivalent to

$$\partial_{z_i} p(z+t) < p(z+t).$$

This gives us the desired result of  $(p - \partial_{z_i} p)(z+t) > 0$ , which means that z is above all roots of  $p - \partial_{z_i} p$ .  $\Box$ 

**Lemma 5.0.8.** Let  $p(z_1, ..., z_m)$  be real stable and  $z \in Ab_p$ . When  $\Phi_p^j(z) \leq 1 - \frac{1}{\delta}$  for  $\delta > 0$ , then for all i

$$\Phi^{i}_{p-\partial_{z_j}p}(z+\delta e_j) \le \Phi^{i}_p(z).$$
(10)

*Proof.* To start we will express  $\Phi^i_{p-\partial_{z_i}p}$  in terms of  $\Phi_p$ .

$$\begin{split} \Phi_{p-\partial_{z_j}p}^i &= \frac{\partial_{z_i}(p - \partial_{z_j}p)}{p - \partial_{z_j}p} \\ &= \frac{\partial_{z_i}((1 - \Phi_p^j)p)}{(1 - \Phi_p^j)p} \\ &= \frac{(1 - \Phi_p^j)(\partial_{z_i}p)}{(1 - \Phi_p^j)p} + \frac{\partial_{z_i}(1 - \Phi_p^j)p}{(1 - \Phi_p^j)p} \text{ (product rule)} \\ &= \frac{\partial_{z_i}p}{p} + \frac{\partial_{z_i}(1 - \Phi_p^j)}{(1 - \Phi_p^j)} \\ &= \Phi_p^i - \frac{\partial_{z_i}\Phi_p^j}{1 - \Phi_p^j} \end{split}$$

To make it so that it only contains  $\Phi_p^i$ , we will have to rewrite  $\partial_{z_i} \Phi_p^j$  using  $\Phi_p^i$ .

$$\begin{split} \partial_{z_i} \Phi_p^j &= \partial_{z_i} \frac{\partial_{z_j} p}{p} \\ &= \frac{(\partial_{z_j} \partial_{z_i} p) p - (\partial_{z_i} p) (\partial_{z_j} p)}{p} \\ &= \frac{(\partial_{z_i} \partial_{z_j} p) p - (\partial_{z_j} p) (\partial_{z_i} p)}{p} \\ &= \partial_{z_j} \Phi_p^i. \end{split}$$

So,

$$\Phi^i_{p-\partial_{z_j}p} = \Phi^i_p - \frac{\partial_{z_j}\Phi^i_p}{1-\Phi^j_p}.$$
(11)

We now get back to proving Equation (10). By applying Equation (11) to it, we get the equivalent statement

$$\Phi_p^i(z+\delta e_j) - \frac{\partial_{z_j} \Phi_p^i(z+\delta e_j)}{1 - \Phi_p^j(z+\delta e_j)} \le \Phi_p^i(z)$$

The convexity property that we got from Theorem 5.0.6 tell us that

$$\Phi_p^i(z+\delta e_j) - \delta \partial_{z_j} \Phi_p^i(z+\delta e_j) \le \Phi_p^i(z).$$

By combining this with what we already had, we now know that in order to prove this Lemma, we only need to show that

$$-\frac{\partial_{z_j}\Phi_p^i(z+\delta e_j)}{1-\Phi_p^j(z+\delta e_j)} \le -\delta\partial_{z_j}\Phi_p^i(z+\delta e_j)$$
(12)

In the proof of Theorem 5.0.6 we have shown that  $-\partial_{z_j}\Phi_p^i(z+\delta e_j) \ge 0$ , so Equation (12), can be implied by

$$\frac{1}{1 - \Phi_p^j(z + \delta e_j)} \le -\delta$$

Now by using the non-increasing property of  $\Phi_p^i$  we get that

$$\frac{1}{1 - \Phi_p^j(z + \delta e_j)} \le \frac{1}{1 - \Phi_p^j(z)} \le \delta$$

This can be changed into

$$\Phi_p^j(z) \le 1 - \frac{1}{\delta}$$

which is the condition for this lemma.

Proof of Theorem 5.0.1. We start by defining  $P(y_1, ..., y_m) = \det(\sum_{i=1}^m y_i A_i)$ . Let  $t > \epsilon$  (the exact value for t will be set later on), then

$$P(t,...,t) = \det(t\sum_{i=1}^{m} A_i) = \det(tI_n) = tn > 0.$$

To prove that this means that  $t\mathbf{1}_m \in Ab_P$ , we need to show that it is still true when the vector  $v = (v_1, ..., v_m)$ (where  $v_k \ge 0$  for all  $k \in [m]$ ) is added to  $t\mathbf{1}_m$ . This can be done by using Lemma 2.2.4, which says that because  $tI_n$  is positive definite, also  $tI_n + \sum_{k=1}^m v_k A_i$  is positive definite. Using Corollary 3.4.3 and assuming that P(y) isn't zero.

$$\begin{split} \Phi_p^i(y_1, ..., y_m) &= \frac{\partial_{z_i} \det(\sum_{k=1}^m y_k A_k)}{\det(\sum_{k=1}^m y_k A_k)} \\ &= \frac{\det(\sum_{k=1}^m y_k A_k) \operatorname{Tr}((\sum_{k=1}^m y_k A_k)^{-1} A_i)}{\det(\sum_{k=1}^m y_k A_k)} \\ &= \operatorname{Tr}((\sum_{k=1}^m y_k A_k)^{-1} A_i). \end{split}$$

So,

$$\Phi_p^i(t,...,t) = \text{Tr}((tI)^{-1}A_i) = \frac{\text{Tr}(A_i)}{t} \le \frac{\epsilon}{t} < 1.$$

Since we want to use Lemma 5.0.8, we need to turn this into something of form  $1 - \frac{1}{\delta}$ .

$$1 - \frac{1}{\delta} = \frac{\epsilon}{t}$$

is equivalent to

$$\delta = \frac{1}{1 - \frac{\epsilon}{t}}.$$

So that is how we will define  $\delta$ . For  $k \in [m]$ , define

$$P_k(y_1, ..., y_m) = \left(\prod_{j=1}^k (1 - \partial_{z_j})\right) P(y_1, ..., y_m)$$

This will be used to build up toward  $Q = P_m$ .

To prove that  $\Phi_{P_m}^i((t+\delta)\mathbf{1}_m) \leq 1-\frac{1}{\delta}$ , we will use a proof by induction. To do this we will increase the value of k, while also increasing  $t\mathbf{1}_m$  step by step into  $(t+\delta)\mathbf{1}_m$ . In order to increase the value, we put in  $\Phi_{P_k}^i$ , we will set

$$x^k = t\mathbf{1}_m + \sum_{j=1}^k \delta e_j.$$

For the inductive proof, we already have the base case;  $\Phi_{P_0}^i(x^0) \leq 1 - \frac{1}{\delta}$  and  $x^0 \in Ab_{P_0}$ . Now the proof for when  $\Phi_{P_k}^i(x^k) \leq 1 - \frac{1}{\delta}$  and  $x^k \in Ab_{P_k}$ , then also  $\Phi_{P_{k+1}}^i(x^{k+1}) \leq 1 - \frac{1}{\delta}$  and  $x^{k+1} \in Ab_{P_{k+1}}$ . The order in which these two properties is proven doesn't matter. So, we will start with  $x^{k+1} \in Ab_{P_{k+1}}$ . Use Lemma 5.0.7, which gives us that  $x^k \in Ab_{P_{k+1}}$  and since  $x^{k+1}$  is above  $x^k$  also  $x^{k+1} \in Ab_{P_{k+1}}$ . By using Lemma 5.0.8, we get that  $\Phi_{P_{k+1}}^i(x^{k+1}) \leq \Phi_{P_k}^i(x^k) \leq 1 - \frac{1}{\delta}$ .

Now we have that the largest root of Q is at most  $x^m$ , this means that the largest root of  $\mu[A_1, ..., A_m]$  is at most  $t + \delta$ . By now finally setting the value of t to  $\sqrt{\epsilon} + \epsilon$ , we get that

$$t + \delta = t + \frac{1}{1 - \frac{\epsilon}{t}}$$
$$= \sqrt{\epsilon} + \epsilon + \frac{1}{1 - \frac{\epsilon}{\sqrt{\epsilon} + \epsilon}}$$
$$= \sqrt{\epsilon} + \epsilon + \frac{\sqrt{\epsilon} + \epsilon}{\sqrt{\epsilon}}$$
$$= \sqrt{\epsilon} + \epsilon + 1 + \sqrt{\epsilon}$$
$$= (1 + \sqrt{\epsilon})^2$$

So, the largest root of  $\mu[A_1, ..., A_m]$  is at most  $(1 + \sqrt{\epsilon})^2$ .

# 6 Proof Weaver's Conjecture

The goal of this chapter is, as you might have read in the title, to prove Weaver's conjecture. We will do this by first proving a theorem which is an accumulation of all that we have proven so far. Then we will use that to prove a generalisation of Weaver's conjecture.

**Theorem 6.0.1.** Let  $\epsilon > 0$  and  $u_1, \ldots, u_m$  are independent random vectors in  $\mathbb{C}^n$  with finite support. If

$$\sum_{i=1}^{m} \mathbb{E} u_i u_i^* = I$$

 $\mathbb{E} ||u_i||^2 < \epsilon \text{ for all } i,$ 

and

then

$$\mathbb{P}\left[\left|\left|\sum_{i=1}^{m} u_{i} u_{i}^{*}\right|\right| \le (1+\sqrt{\epsilon})^{2}\right] > 0$$
(13)

*Proof.* To simplify notation a bit, set  $M = \sum_{i=1}^{m} u_i u_i^*$ . We start with looking at the value of ||M||. Since M is a summation of positive definite matrices, it is also positive definite. Lemma 2.3.1 tells us that then ||M|| is equal to the largest eigenvalue of M. This means Equation (13) is equivalent to saying that there is a chance that the largest root of  $\chi[M]$  is smaller or equal to  $(1 + \sqrt{\epsilon})^2$ .

To prove this, we will need to have a way to separately handle the possible values of  $u_i$ . For this we will use the same notation as we did at the end of Chapter 4. We set  $l_i$  equal to the number of possible values for  $u_i$ , with  $w_{i,1}, ..., w_{i,l_i}$  being the possible values. Each of those values has a probability of  $p_{i,1}, ..., p_{i,l_i}$  to occur. Let  $j_i \in [l_i]$ , define

$$q_{j_1,\dots,j_m} = \left(\prod_{i=1}^m p_{i,j_i}\right) \chi\left[\sum_{i=1}^m w_{i,j_i} w_{i,j_i}^*\right].$$

All possible  $\chi[M]$  can now be linked to a  $q_{j_1,...,j_m}$  (this is a one on one relation). By using Theorem 4.0.8, we get that these polynomials form an interlacing family. Then according to 3.2.5, there is a  $q_{j_1,...,j_m}$  of which the largest root is smaller than the largest root of  $q_{\emptyset}$  (sum of all possible  $q_{j_1,...,j_m}$ ). Define  $A_i = \mathbb{E} u_i u_i^*$ . Theorem 4.0.4 tells us that we are now working with the characteristic polynomial of  $A_i, ..., A_m$ .

$$q_{\emptyset} = \sum_{j_1 \in [l_1], \dots, j_m \in [l_m]} q_{j_1, \dots, j_m} = \mathbb{E} \chi \left[ \sum_{i=1}^m u_i u_i^* \right] = \mu[A_1, \dots, A_m]$$

To finish this proof, we will use Theorem 5.0.1. For that we already have the condition that  $\mathbb{E} M = I$ , we only need to prove that  $\operatorname{Tr}(A_i) \leq \epsilon$ . This follows from

$$\operatorname{Tr}(A_i) = \mathbb{E}\operatorname{Tr}(u_i u_i^*) = \mathbb{E}u_i^* u_i = \mathbb{E}||u_i||^2 \le \epsilon.$$

From Theorem 5.0.1, we get that because the largest root of  $\tilde{q}_{\emptyset}$  is smaller than  $(1 + \sqrt{\epsilon})^2$ , also Equation (13) is true.

We shall move on to proving a generalization of Weaver's Conjecture 1.0.1.

**Corollary 6.0.2.** Let  $v_1, ..., v_m \in \mathbb{C}^d$  be column vectors such that

$$\sum_{i=1}^{m} v_i v_i^* = I \tag{14}$$

and  $||v_i||^2 \leq \delta$  for all *i*. Then for every positive integer *r* there exists a partition  $\{S_1, ..., S_r\}$  of [m] such that

$$\left\| \sum_{i \in S_j} v_i v_i^* \right\| \le \left( \frac{1}{\sqrt{r}} + \sqrt{\delta} \right)^2 \tag{15}$$

*Proof.* This proof will mainly focus on how this corollary is a weaker or equally strong expression as Theorem 6.0.1.

Define  $w_{i,k} \in \mathbb{C}^{rd}$  where  $i \in [m]$  and  $k \in [r]$ . The value of  $w_{i,k}$  is that of r vectors combined together, where the kth is  $v_i$  and the others are  $\mathbf{0}_d$ . For example,

$$w_{i,1} = \begin{pmatrix} v_i \\ \mathbf{0}_d \\ \vdots \\ \mathbf{0}_d \end{pmatrix}.$$

For the independent random vector  $u_i$  from Theorem 6.0.1, we will use vectors of the size n = rd. The values will be set as follows  $u_i$  is equal to  $\{\sqrt{r}w_{i,k}\}_{k=1}^r$  with probability  $\frac{1}{r}$ . Now to guarantee these values match the conditions.

$$\mathbb{E} u_i u_i^* = \begin{pmatrix} \frac{1}{r} \sqrt{r} v_i \sqrt{r} v_i^* & \mathbf{0}_{d \times d} & \cdots & \mathbf{0}_{d \times d} \\ \mathbf{0}_{d \times d} & \frac{1}{r} \sqrt{r} v_i \sqrt{r} v_i^* & \vdots \\ \vdots & & \ddots & \vdots \\ \mathbf{0}_{d \times d} & \cdots & \mathbf{0}_{d \times d} \\ \end{bmatrix}$$
$$= \begin{pmatrix} v_i v_i^* & \mathbf{0}_{d \times d} & \cdots & \mathbf{0}_{d \times d} \\ \mathbf{0}_{d \times d} & v_i v_i^* & \vdots \\ \vdots & \ddots & \vdots \\ \mathbf{0}_{d \times d} & \cdots & v_i v_i^* \end{pmatrix}$$

and

$$\mathbb{E} ||u_i||^2 = ||\sqrt{r}v_i||^2 = r||v_i||^2 \le r\delta.$$

So,  $\sum_{i=1}^{m} \mathbb{E} u_i u_i^* = I_{rd}$  and  $\epsilon = r\delta$ . Now we can apply Theorem 6.0.1, which tells us that there exists an assignment of  $u_i' \in \{\sqrt{r}w_{i,k}\}_{k=1}^r$  such that

$$(1+\sqrt{r\delta})^2 \ge \left| \left| \sum_{i=1}^m u_i' u_i'^* \right| \right|$$

To go back to the values we started with, we will use  $S_k = \{i | u'_i = \sqrt{r}w_{i,k}\}$ .

$$\left\| \sum_{i=1}^{m} u_{i}' u_{i}'^{*} \right\| = \left\| \sum_{k=1}^{m} \sum_{S_{k}} (\sqrt{r} w_{i,k}) (\sqrt{r} w_{i,k})^{*} \right\|.$$

In conclusion for all k,

$$\begin{split} \left\| \sum_{i \in S_k} v_i v_i^* \right\| &= \left\| \sum_{S_k} w_{i,k} w_{i,k}^* \right\| \\ &\leq \left\| \left\| \sum_{k=1}^m \sum_{S_k} w_{i,k} w_{i,k}^* \right\| \right\| \\ &= \frac{1}{r} \left\| \left\| \sum_{k=1}^m \sum_{S_k} (\sqrt{r} w_{i,k}) (\sqrt{r} w_{i,k})^* \right\| \right\| \\ &\leq \frac{1}{r} (1 + \sqrt{r\delta})^2 \\ &= \left( \frac{1}{\sqrt{r}} + \sqrt{\delta} \right)^2. \end{split}$$

*Proof of Weaver's Conjecture 1.0.1.* This proof will mainly show how Corollary 6.0.2 is a generalization of Weaver's Conjecture.

First, we bring back Corollary 6.0.2 to something simpler by setting r to 2, since we are only interested in the case where a partition into 2 parts is made. Set  $v_i = \frac{w_i}{\sqrt{\eta}}$ . This means that

$$\sum_{i=1}^{m} |\langle u, v_i \rangle|^2 = 1.$$

To prove that this means that Equation (14) is true, we will rewrite Condition (1).

$$1 = \sum_{i=1}^{m} |\langle u, v_i \rangle|^2$$
$$= \sum_{i=1}^{m} |u^{t}\overline{v_i}|^2$$
$$= \sum_{i=1}^{m} (u^{t}\overline{v_i})\overline{(u^{t}\overline{v_i})}$$
$$= \sum_{i=1}^{m} (u^{t}\overline{v_i})(u^{*}v_i)$$
$$= \sum_{i=1}^{m} (u^{t}\overline{v_i})(v_i^{t}\overline{u})$$
$$= \sum_{i=1}^{m} u^{t}(\overline{v_i}v_i^{t})\overline{u}$$
$$= u^{t} \left(\sum_{i=1}^{m} \overline{v_i}v_i^{t}\right)\overline{u}$$

Making use of the fact that u is a unit vector;  $u^{t}\overline{u} = 1$ , so

$$\sum_{i=1}^{m} \overline{v_i} v_i^{t} = I$$
$$= \overline{I}$$
$$= \sum_{i=1}^{m} v_i v_i^{*}$$

From  $||w_i|| \leq 1$  for all *i*, we learn that  $||v_i||^2 \leq \frac{1}{\eta} = \delta$ . We now have all the conditions for Corollary 6.0.2.

$$\left\|\sum_{i\in S_j} v_i v_i^*\right\| \le \left(\frac{1}{\sqrt{r}} + \sqrt{\delta}\right)^2$$

is equivalent to

$$\left|\sum_{i\in S_j} w_i w_i^*\right| \le \eta \left(\frac{1}{\sqrt{2}} + \sqrt{\frac{1}{\eta}}\right)^2 = \frac{\eta}{2} + \sqrt{2\eta} + 1.$$

So, if  $\eta$  reaches a certain value, then  $\frac{\eta}{2} + \sqrt{2\eta} + 1 < \eta$  (Here it is important to note that not for every  $\eta$  there has to exist a  $\theta$ , only that there is a combination of the two that exists).

Now to finish this prove, we need to prove that this is also true for  $\max_u \left( \sum_{i \in S_j} |\langle u, w_i \rangle|^2 \right)$ .

$$\max_{u} \left( \sum_{i \in S_j} |\langle u, w_i \rangle|^2 \right) = \max_{u} \left( u^* \left( \sum_{i \in S_j} w_i w_i^* \right) u \right)$$
$$= \max_{u} \left\langle \left( \sum_{i \in S_j} w_i w_i^* \right) u, u \right\rangle$$

because u is a unit vector, we are now calculating the largest eigenvalue of  $\sum_{i \in S_j} w_i w_i^*$ . This means that

$$\max_{u} \left( \sum_{i \in S_j} |\langle u, w_i \rangle|^2 \right) = \max \left\{ \text{eigenvalues} \left( \sum_{i \in S_j} w_i w_i^* \right) \right\}$$
$$= \left| \left| \sum_{i \in S_j} w_i w_i^* \right| \right| \text{ (Using Lemma 2.3.1)}$$

Now to give some examples of possible values for Weaver's Conjecture KS<sub>2</sub>. When  $\frac{\eta}{2} - \sqrt{2\eta} - 1 > 0$ , then  $\eta$  and  $\theta = \frac{\eta}{2} - \sqrt{2\eta} - 1$  are valid solutions for Weaver's Conjecture KS<sub>2</sub>. To give a natural number example there is  $\eta = 18$  and  $\theta = 2$ .

# 7 Paving Conjecture

Using Corollary 6.0.2 we have already shown that we can prove Weaver's Conjecture  $KS_2$  which is known for implying the Kadison-Singer problem. Another thing we can do with it is prove the Paving Conjecture. The reason why the Paving Conjecture is interesting, is because Weaver's conjecture  $KS_2$  doesn't directly imply the positive solution to Kadison-Singer problem, but instead it implies the Paving Conjecture which is known to be equivalent to the positive solution of the Kadison-Singer problem. So, by proving the Paving Conjecture we are one step closer to fully proving the positive solution to the Kadison-Singer problem. Before we can explain the Paving Conjecture we first need to define what pavings are.

**Definition 7.0.1.** Matrix  $M \in \mathbb{C}^{n \times n}$  can be  $(r, \epsilon)$ -paved where  $\epsilon > 0$  and  $r \in \mathbb{N}$ , if there exist coordinate projections  $P_1, ..., P_r$  with  $\sum_{i=1}^m P_i = I_n$  such that

$$||P_i M P_i|| \le \epsilon ||M||$$

for all  $i \in [m]$ .

Since we now understand what pavings are, we can go onto the Paving Conjecture.

**Conjecture 7.0.2** (Paving Conjecture). For every  $\epsilon > 0$  there is a  $r \in \mathbb{N}$  such that every Hermitian matrix with a zero diagonal can be  $(r, \epsilon)$ -paved.

#### 7.1 Theorems for proving the Paving Conjecture

In this chapter, we will work with a variety of matrix types. To not define the same types of matrices over and over again, we will define some collections.

Definition 7.1.1. We set:

- $\mathcal{H}$  is the collection of all Hermitian matrices with zero diagonal.
- $\mathcal{U}$  are all unitary matrices in  $\mathcal{H}$ .
- $\mathcal{Q}$  is the collection of all Hermitian projection matrices where the values on the diagonal are  $\frac{1}{2}$ .

In order to prove the Paving Conjecture, we wish to stop working with matrices in  $\mathcal{H}$  and step by step convert it into a problem for matrices in  $\mathcal{Q}$ .

**Lemma 7.1.2** (Theorem 3 {(5)  $\rightarrow$  (4)}of [2]). If there is a function  $r : \mathbb{R}^{>0} \rightarrow \mathbb{N}$  such that every  $2n \times 2n$  matrix  $U \in \mathcal{U}$  can be  $(r(\epsilon), \epsilon)$ -paved, then every  $n \times n$  matrix  $H \in \mathcal{H}$  can be  $(r(\epsilon), \epsilon)$ -paved.

*Proof.* We use H to create matrix

$$U = \begin{pmatrix} H & \sqrt{I - H^2} \\ \sqrt{I - H^2} & -H \end{pmatrix} \in \mathcal{U}.$$

Since U can be  $(r(\epsilon), \epsilon)$ -paved, we can now create the paving for H by simplifying the diagonal projections for U leaving only the top left quarter behind. So, H can be  $(r(\epsilon), \epsilon)$ -paved.

Now we can prove the Paving Conjecture without using  $\mathcal{H}$ , but instead using  $\mathcal{U}$ . To push this even one step further, we will first need to prove some connections between  $\mathcal{Q}$  and  $\mathcal{U}$ .

**Lemma 7.1.3.** The function  $f: \mathcal{U} \to \mathcal{Q}$  defined by  $f(U) = \frac{I+U}{2}$  is bijective.

*Proof.* To do this, we simply need to show that there exists a one on one relation between  $\mathcal{U}$  and  $\mathcal{Q}$  through f. Since f only uses multiplication and summations to constant variables, we know that f creates a link from each U to a unique element in  $f(\mathcal{U})$  and the other way around. So, we only need to prove that  $f(\mathcal{U}) = \mathcal{Q}$ .

$$f(U) = \left(\frac{I+U}{2}\right)^2 = \frac{I+2U+U^2}{4} = \frac{I+U}{2} \in Q$$

and

$$f^{-1}(Q) = (2Q - I)^2 = 4Q - 4Q + I = I.$$

So,  $f(\mathcal{U}) \in \mathcal{Q}$  and  $Q \in f(\mathcal{U})$ .

**Lemma 7.1.4** (Theorem 3 {(6)  $\rightarrow$  (5)}of [2]). If there is a function  $r : \mathbb{R}^{>0} \rightarrow \mathbb{N}$  such that every  $n \times n$  matrix  $Q \in \mathcal{Q}$  can be  $(r(\epsilon), \frac{1+\epsilon}{2})$ -paved, then every  $n \times n$  matrix  $U \in \mathcal{U}$  can be  $(r(\epsilon)^2, \epsilon)$ -paved.

*Proof.* Here it is important to first notice that ||U|| = || - U||. We start by using Lemma 7.1.3. For U we can create matrix  $Q_1 = f(U) \in \mathcal{Q}$  with coordinate projections  $P_{1,1}, ..., P_{1,r(\epsilon)}$ . Now we know that  $||P_{1,i}Q_1P_{1,i}|| \leq \frac{1+\epsilon}{2}$ . Sadly, we cannot directly rely this back to ||U||, this is because for some cases

$$||P_{1,i}Q_1P_{1,i}|| \le ||P_{1,i}f(-U)P_{1,i}||$$

(due to that in f the value I is added, which can both increase or decrease the value of ||...||, but if adding I increase it, then it means that substracting I would decrease the value, also the other way around). In order to solve this problem, we create  $Q_2 = f(-U) \in \mathcal{Q}$  with diagonal projections  $P_{2,1}, ..., P_{2,r(\epsilon)}$ . We now can go and create a collection that covers both cases by creating  $P_{3,1}, ..., P_{3,r(\epsilon)^2}$  from all possible products of  $P_{1,1}, ..., P_{1,r(\epsilon)}$  and  $P_{2,1}, ..., P_{2,r(\epsilon)}$ . If  $P_{3,k} = P_{1,i}P_{2,j}$ , then

$$||P_{3,k}UP_{3,k}||$$

is smaller or equal to either

$$2||P_{1,i}Q_1P_{1,i}|| - 1 \le \epsilon$$

or

 $2||P_{2,j}Q_2P_{2,j}|| - 1 \le \epsilon.$ 

So, for all  $U \in \mathcal{U}$  can be  $(r(\epsilon)^2, \epsilon)$ -paved.

Now by combining Lemma 7.1.2 and Lemma 7.1.4, we can rewrite the Paving Conjecture to something that is easier for us to solve.

#### 7.2 Proving the Paving Conjecture

Proof of the Paving Conjecture 7.0.2. By using Lemma 7.1.2 and 7.1.4, we can simplify the conjecture so that we only need to prove that for every  $\epsilon > 0$  there is an r such that every  $2n \times 2n$  matrix  $Q \in \mathcal{Q}$  can be  $(r, \frac{1+\epsilon}{2})$ -paved.

Our first step will be to rewrite Q into a Gram matrix of vectors which fit the  $\sum_{2n}^{i=1} v_i v_i^* = I_n$  condition of Corollary 6.0.2 (important to remember is that we are really aiming at creating a unit matrix of size  $n \times n$ , not  $2n \times 2n$ ). For this we will first diagonalize Q using unitary matrix U (possible, because Lemma 2.2.2, which states it is possible for Hermitian matrices).

$$\widetilde{Q} = UQU^*$$

Now we are going to deduce the eigenvalues of  $\tilde{Q}$ . Q has the same eigenvalues as  $\tilde{Q}$ , so we can start by looking at the trace of Q. Since, all diagonal entries of Q are equal to  $\frac{1}{2}$ , we know that the sum of all eigenvalues is n. To get even more information about the eigenvalues of  $\tilde{Q}$  we will use that Q is a projection matrix, so  $Q^2 = Q$ .

$$\begin{split} \tilde{Q}^2 &= UQU^*UQU^* \\ &= UQQU^* \\ &= UQU^* \\ &= \tilde{Q} \end{split}$$

This means that all eigenvalues of  $\widetilde{Q}$  are either 1 or 0. By combining this with the fact that the sum of the eigenvalues is n, we get that the half the eigenvalues is 1 and the other 0. Knowing this, we can create a collection of vectors  $\tilde{v}_1, ..., \tilde{v}_{2n} \in \mathbb{C}^n$  which are either unit vectors or equal to  $0_n$ , such that the Gram matrix of them is  $\widetilde{Q} = (\tilde{v}_1, ..., \tilde{v}_n)(\tilde{v}_1, ..., \tilde{v}_n)^*$  and  $\sum_{2n}^{i=1} \tilde{v}_i \tilde{v}_i^* = I_n$ . To ease the notation, we will write

$$\widetilde{V} = \left(\widetilde{v}_1, \dots, \widetilde{v}_{2n}\right)$$

So,  $\widetilde{Q} = \widetilde{V}^*\widetilde{V}$ .

$$Q = U^* \widetilde{Q} U$$
$$= U^* \widetilde{V}^* \widetilde{V} U$$
$$= (\widetilde{V} U)^* \widetilde{V} U$$

This means Q is also a Gram matrix of vectors of length n. Define

$$(v_1, ..., v_{2n}) = V = VU.$$

We can now get back to the condition of 6.0.2.

$$\sum_{i=1}^{2n} v_i v_i^* = VV^*$$
$$= \widetilde{V}U(\widetilde{V}U)^*$$
$$= \widetilde{V}UU^*\widetilde{V}$$
$$= \widetilde{V}\widetilde{V}^*$$
$$= I_n$$

This means we fulfil the first condition of Corollary 6.0.2. To obtain the other condition for Corollary 6.0.2;  $||v_i||^2 \leq \delta$  we use the fact that all values on the diagonal of Q are  $\frac{1}{2}$ , meaning  $||v_i||^2 = \frac{1}{2} = \delta$ . Now we can use Corollary 6.0.2, giving us partition  $S_1, ..., S_r$  of [2n] such that

$$\left\| \sum_{i \in S_j} v_i v_i^* \right\| \le \left( \frac{1}{\sqrt{r}} + \sqrt{\delta} \right)^2 = \left( \frac{1}{\sqrt{r}} + \frac{1}{\sqrt{2}} \right)^2.$$

Define diagonal projection  $P_k$  which is based on  $S_k$ , so when  $i \in S_k$  then the *i*th value of the diagonal is equal to 1 else it is 0. This gives us  $||P_kQP_k||$ , to deduce its value we will turn it into a Gram matrix. Define  $V_k \in \mathbb{C}^{2n \times n}$ , for which all columns which number is in  $S_k$  have the same value as their respective column in V, the rest is equal to zero. Now  $P_kQP_k = V_k^*V_k$ , a Gram matrix of the columns of  $V_k$ . Now we can conclude that

$$\begin{split} ||P_k Q P_k|| &= ||V_k^* V_k|| \\ &= ||V_k V_k^*|| \text{ (using Lemma 2.3.4)} \\ &= \left| \left| \sum_{i \in S_k} v_i v_i^* \right| \right| \\ &= \left( \frac{1}{\sqrt{r}} + \frac{1}{\sqrt{2}} \right)^2 \end{split}$$

By picking  $r = \frac{36}{\epsilon^2}$ , we get that

$$||P_k Q P_k|| \le \left(\frac{1}{\sqrt{\frac{36}{\epsilon^2}}} + \frac{1}{\sqrt{2}}\right)^2 = \epsilon.$$

Now we have proven the Paving Conjecture.

# 8 Conclusion

We have done what we have set out to do. We have proven Weaver's Conjecture  $KS_2$  and got a bit closer to fully proving the Kadison-Singer Problem by proving the Paving Conjecture. To give a more complete version of the proof, we also included the proofs that were only referenced in the source material. This is most noticeable in chapter 3, 5 and 7. The only things that have been left open are Hurwitz' Theorem 3.1.1, the implicit function theorem extended to complex values and the Cauchy-Riemann Equation each with their own reference. These are all things that the main source material for this thesis never mentioned, because it left all the work of dealing with these to its references. They are all far more general knowledge than most things mentioned in this thesis, that is why I'm okay with not proving them in this thesis. It especially felt odd that Hurwitz' Theorem hadn't been mentioned in the source material, even though, it is extremely useful in the context of Chapter 3. I even incorporated it into the proof for Lemma 3.2.7.

But before I get further off-topic. Overall, I'm fairly satisfied with the level of detail in this proof of Weaver's Conjecture  $KS_2$  and the Paving Conjecture. I hope that anyone who reads this finds it easier to understand and more complete than the source material.

# References

- J. Borcea and P. Branden. Applications of stable polynomials to mixed determinants: Johnson's conjectures, unimodality, and symmetrized Fischer products. *Duke Math J.*, 143(2):205–223, 2008.
- [2] P. Casazza, D. Edidin, D. Kalra, and V. I. Paulsen. Projections and the Kadison-Singer problem. oper. matrices, 1:391–408, 2007.
- [3] J. B. Conway. A Course in Functional Analysis. Graduate Texts in Mathematics. Springer New York, 1994. ISBN 9780387972459.
- [4] J. P. Dedieu. Obreschkoff's theorem revisited: What convex sets are contained in the set of hyperbolic polynomials? *Journal of Pure and Applied Algebra*, 81:269–278, 1992.
- [5] J. Ding and A. Zhou. Eigenvalues of rank-one updated matrices with some applications. Applied Mathematics Letters, pages 1223–1226, 2007.
- [6] R. V. Kadison and I. M. Singer. Extensions of pure states. Amer. J. Math., 81:383–400, 1959.
- [7] L. Kaup, B. Kaup, and G. Barthel. Holomorphic Functions of Several Variables: An Introduction to the Fundamental Theory. De Gruyter studies in mathematics. W. de Gruyter, 1983. ISBN 9783110041507.
- [8] J. Magnus and H. Neudecker. Matrix Differential Calculus with Applications in Statistics and Econometrics. Wiley Series in Probability and Statistics: Texts and References Section. Wiley, 1999. ISBN 9780471986331.
- [9] A. W. Marcus, D. A. Spielman, and N. Srivastava. Interlacing families II: Mixed characteristic polynomials and the Kadison-Singer problem. Annals of Mathematics, 182:327–350, 2014.
- [10] T. Tao. Real stable polynomials and the Kadison-Singer problem. Available at https://terrytao. wordpress.com/2013/11/04/real-stable-polynomials-and-the-kadison-singer-problem/, 2013.
- [11] E. van den Ban. Dictaat functies en reeksen. Lecture notes, 7 2016.
- [12] D. G. Wagner. Multivariate stable polynomials: theory and applications. Bull. Amer. Math. Soc., 28: 53–84, 2011.
- [13] N. Weaver. The Kadison-Singer problem in discrepancy theory. Discrete Mathematics, 278:227 239, 2004.