Acknowledgements

I would like to thank my supervisor, Prof. Nicholas Young, for the patient guidance, encouragement and advice he has provided throughout my time as his student. I have been extremely lucky to have a supervisor who cared so much about my work, and who responded to my questions and queries so promptly. I would also like to thank all the members of staff at Newcastle and Lancaster Universities who helped me in my supervisor's absence. In particular I would like to thank Michael White for the suggestions he made in reference to Chapter 4 of this work.

I must express my gratitude to Karina, my wife, for her continued support and encouragement. I was continually amazed by her willingness to proof read countless pages of meaningless mathematics, and by the patience of my mother, father and brother who experienced all of the ups and downs of my research.

Completing this work would have been all the more difficult were it not for the support and friendship provided by the other members of the School of Mathematics and Statistics in Newcastle, and the Department of Mathematics and Statistics in Lancaster. I am indebted to them for their help.

Those people, such as the postgraduate students at Newcastle and *The Blatherers*, who provided a much needed form of escape from my studies, also deserve thanks for helping me keep things in perspective.

Finally, I would like to thank the Engineering and Physical Sciences Research Council, not only for providing the funding which allowed me to undertake this research, but also for giving me the opportunity to attend conferences and meet so many interesting people.

Abstract

We establish necessary conditions, in the form of the positivity of Pick-matrices, for the existence of a solution to the spectral Nevanlinna-Pick problem:

Let k and n be natural numbers. Choose n distinct points z_j in the open unit disc, $\mathbb{D}$, and n matrices W_j in $\mathbb{M}_k(\mathbb{C})$, the space of complex $k \times k$ matrices. Does there exist an analytic function $\phi : \mathbb{D} \to \mathbb{M}_k(\mathbb{C})$ such that

$$\phi(z_j) = W_j$$

for $j = 1, ..., n$ and

$$\sigma(\phi(z)) \subset \overline{\mathbb{D}}$$

for all $z \in \mathbb{D}$?

We approach this problem from an operator theoretic perspective. We restate the problem as an interpolation problem on the symmetrized polydisc Γ_k,

$$\Gamma_k = \{(c_1(z), \ldots, c_k(z)) \mid z \in \overline{\mathbb{D}}\} \subset \mathbb{C}^k$$

where $c_j(z)$ is the j^{th} elementary symmetric polynomial in the components of z. We establish necessary conditions for a k-tuple of commuting operators to have Γ_k as a complete spectral set. We then derive necessary conditions for the existence of a solution ϕ of the spectral Nevanlinna-Pick problem.

The final chapter of this thesis gives an application of our results to complex geometry. We establish an upper bound for the Caratheodory distance on $\operatorname{int} \Gamma_k$.

Chapter 1

Interpolation Problems

This thesis is concerned with establishing necessary conditions for the existence of a solution to the spectral Nevanlinna-Pick problem. In the sections of this chapter which follow, we define a number of interpolation problems beginning with the classical Nevanlinna-Pick problem. After presenting a full solution to this classical mathematical problem, we give a brief summary of some results in linear systems theory. These results demonstrate how the classical Nevanlinna-Pick problem arises as a consequence of robust control theory. We then slightly alter the robust stabilization problem and show that this alteration gives rise to the spectral Nevanlinna-Pick problem. Chapter 1 is completed with the introduction of a new interpolation problem which is closely related to both versions of the Nevanlinna-Pick problem. Although the problems discussed all have relationships with linear systems and control theory, they are interesting mathematical problems in their own right. The engineering motivation presented in Section 1.2 is not essential to the work which follows but allows the reader a brief insight into the applications of our results.

Chapter 2 begins by converting function theoretic interpolation problems into problems concerning the properties of operators on a Hilbert space. Throughout Chapter 2 we are concerned with finding a particular class of polynomials. Although the exact form of the polynomials is unknown to us, we are aware of various properties they must possess. We use these properties to help us define a suitable class of polynomials.

In Chapter 3 we define a class of polynomials based on the results of Chapter 2. We present a number of technical results which allow us to represent this class of polynomials in various forms. This is followed, in Chapter 4, by a proof that certain polynomial pencils which arise as part of a representation in Chapter 3 are non-zero over the polydisc. Although this proof, like the results in the chapter which precedes it, is rather detailed, the resultant simplifications in Chapters 5, 6 and 7 are essential.

The proof of our first necessary condition for the existence of a solution to the spectral Nevanlinna-Pick problem is given in Chapter 5. The actual statement of this necessary condition is presented in operator theoretic terms, but in keeping with the classical motivation for the problem we also present the result in the form of a Pick-matrix. Chapter 5 concludes with a simple example demonstrating the use of the new necessary condition.

Chapter 6 contains the second of our necessary conditions. The results needed to prove this more refined condition are easy extensions of results in the preceding chapters. The main results of Chapter 5 can all be given as special cases of the results in Chapter 6. Although the results of Chapter 6 do bear much similarity to those of Chapter 5, they are presented in full for completeness.

The mathematical part of this thesis is concluded with a new result in complex geometry. In Chapter 7 we prove an upper bound for the Caratheodory distance on a certain domain in $\mathbb{C}^k$. This proof relies heavily on the theory developed in earlier chapters in connection with the necessary conditions for the existence of a solution to the spectral Nevanlinna-Pick problem. It also serves to demonstrate the consequences of our results.

The thesis concludes in Chapter 8 with a brief discussion of possible future avenues of research. We also discuss the connections between our work and other results in the area.

What remains of this chapter is devoted to introducing three interpolation problems. The first of these, the classical Nevanlinna-Pick problem, is to construct an analytic function on the disc subject to a number of interpolation conditions and a condition concerning its supremum. The Nevanlinna-Pick problem is well studied and has an elegant solution (Corollary 1.1.3). We present a full proof of this solution in Section 1.1.

Although the Nevanlinna-Pick problem resides in the realms of function theory and pure mathematics, it has far reaching applications. In Section 1.2 we introduce some of the fundamentals of linear systems theory. The small section of this chapter which is devoted to linear systems is far from a complete study of even the most basic concepts of that subject. The inclusion of the topic is meant only to act as motivation and we hope it will help the reader place the Nevalinna-Pick problem in a wider context. With this aim in mind, we present a simplified demonstration of the importance of the Nevanlinna-Pick problem to a specific control engineering problem. Section 1.2 concludes by investigating how a change in the formulation of the control engineering problem gives rise to a variant of the Nevanlinna-Pick problem.

The variant of the Nevanlinna-Pick problem of interest to us is known as the Spectral Nevanlinna-Pick problem. This is introduced in Section 1.3. As one might suspect from its title, the Spectral Nevanlinna-Pick problem (which we now refer to as the Main Problem) is very similar to the classic Nevanlinna-Pick problem. Unfortunately, as yet, no solution of it is known. The aim of this work is to find necessary conditions for the existence of a solution to the Main Problem.

In Section 1.3 we introduce the Γ_k problem. This is an interpolation problem which is closely related to the Main Problem. Section 1.3 is concluded with the definition of a polynomial which will be of great interest to us throughout our work.

1.1 The Nevanlinna-Pick Problem

The Main Problem studied below is a variant of the classical Nevanlinna-Pick problem, which was first solved by Pick [34] early this century. The classical version of the problem can be stated thus.

Nevanlinna-Pick Problem *Pick $2n$ points $\{z_j\}_1^n, \{\lambda_j\}_1^n$ in $\mathbb{D}$ such that the z_j are distinct. Does there exist an analytic function $\phi : \mathbb{D} \to \mathbb{C}$ such that $\phi(z_j) = \lambda_j$ for $j = 1, \ldots, n$ and $|\phi(z)| \leq 1$ for all $z \in \mathbb{D}$?*

Below we present a solution to this problem in keeping with the methods and philosophies used throughout this work. We hope that the reader will find this solution both illuminating and motivating. First, we require some terminology.

Following convention, we shall let H^2 denote the Hardy space of analytic functions in $\mathbb{D}$ which have square summable Taylor coefficients. That is,

$$H^2 = \left\{ \sum_{n=0}^{\infty} a_n z^n \mid \sum_{n=0}^{\infty} |a_n|^2 < \infty \right\}.$$

For a full account of Hardy spaces see [31]. It is well known that H^2 has a reproducing kernel, the Szegő kernel, which is defined by

$$K(\lambda, z) = \frac{1}{1 - \lambda \bar{z}} \qquad \lambda,\ z \in \mathbb{D}. \tag{1.1}$$

Fixing z, we see that this kernel gives rise to a function in H^2, namely $K_z(\cdot) = K(\cdot, z)$. When the z_j are distinct, the functions K_{z_j} are linearly independent (see, for example, [33]). The kernel is referred to as 'reproducing' because the function K_z has the following property. For all $h \in H^2$ and all $z \in \mathbb{D}$ we have

$$\langle h, K_z \rangle_{H^2} = h(z).$$

For any $z_1, \ldots, z_n \in \mathbb{D}$ and $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ we define the corresponding space $\mathcal{M}$ and model operator $T_{K_{z_1}, \ldots, K_{z_n}; \lambda_1, \ldots, \lambda_n} : \mathcal{M} \to \mathcal{M}$ as follows:

$$\mathcal{M} = \mathrm{Span}\{K_{z_1}, \ldots, K_{z_n}\} \quad \text{and} \quad T_{K_{z_1}, \ldots, K_{z_n}; \lambda_1, \ldots, \lambda_n} K_{z_j} = \bar{\lambda}_j K_{z_j}. \tag{1.2}$$

Thus $T_{K_{z_1},\ldots,K_{z_n};\lambda_1,\ldots,\lambda_n}$ is the operator with matrix

$$\begin{bmatrix} \bar{\lambda}_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \bar{\lambda}_n \end{bmatrix}$$

with respect to the basis $K_{z_1},\ldots,K_{z_n}$ of $\mathcal{M}$. Clearly, two such operators commute.

Definition 1 *The backward shift operator S^* on H^2 is given by*

$$(S^*f)(z) = \frac{1}{z}(f(z) - f(0))$$

for all $z \in \mathbb{D}$.

Lemma 1.1.1 *Let $z_i \in \mathbb{D}$ for $i = 1,\ldots,n$. Define K_{z_i} and $\mathcal{M}$ by (1.2). Then $\mathcal{M}$ is invariant under the action of S^*. Furthermore, $S^*|\mathcal{M}$ commutes with $T_{K_{z_1},\ldots,K_{z_n};\lambda_1,\ldots,\lambda_n}$.*

Proof. Consider K_z, a basis element of $\mathcal{M}$. We have, for all $\lambda \in \mathbb{D}$,

$$S^*K_z(\lambda) = \frac{1}{\lambda}(K_z(\lambda) - K_z(0)) = \frac{1}{\lambda}\left(\frac{1}{1-\lambda\bar{z}} - 1\right) = \frac{\lambda\bar{z}}{\lambda}\left(\frac{1}{1-\lambda\bar{z}}\right) = \bar{z}K_z(\lambda).$$

It follows that $S^*|\mathcal{M} = T_{K_{z_1},\ldots,K_{z_n};z_1,\ldots,z_n}$. The result then holds.

■

The properties of S^* and its relationship to $\mathcal{M}$ described above allow us to prove the following theorem.

Definition 2 *For $\phi \in H^\infty$ define the multiplication operator M_ϕ on H^2 by*

$$M_\phi h(\lambda) = \phi(\lambda)h(\lambda)$$

for all $h \in H^2$ and all $\lambda \in \mathbb{D}$.

Theorem 1.1.2 *Let $z_j \in \mathbb{D}$ and $\lambda_j \in \mathbb{C}$ for $j = 1,\ldots,n$. There exists a bounded function $\phi : \mathbb{D} \to \mathbb{C}$ such that $\|\phi\|_\infty \leq 1$ and $\phi(z_j) = \lambda_j$ for $j = 1,\ldots,n$ if and only if the model operator $T = T_{K_{z_1},\ldots,K_{z_n};\lambda_1,\ldots,\lambda_n}$ is a contraction.*

Proof. For $h \in H^2$ and $\phi \in H^\infty$ consider the inner product

$$\begin{aligned} \langle M_\phi{}^* K_z, h\rangle = \quad & \langle K_z, M_\phi h\rangle = \langle K_z, \phi h\rangle & = \overline{(\phi h)(z)} \\ = \quad & \overline{\phi(z)h(z)} = \overline{\phi(z)}\langle K_z, h\rangle & = \langle \overline{\phi(z)}K_z, h\rangle. \end{aligned} \tag{1.3}$$

That is

$$M_\phi{}^* K_z = \overline{\phi(z)} K_z \tag{1.4}$$

for all $z \in \mathbb{D}$.

Suppose a function $\phi : \mathbb{D} \to \mathbb{C}$ exists such that $\|\phi\| \leq 1$ and $\phi(z_j) = \lambda_j$ for $j = 1, \ldots, k$. Choose a basis element K_{z_j} of $\mathcal{M}$ and consider the operator M_ϕ^*, which has norm at most one. It follows from (1.4) that $M_\phi^* K_{z_j}$ is equal to $\overline{\lambda_j} K_{z_j}$, which in turn is equal to TK_{z_j} by definition. Thus, T and M_ϕ^* coincide on every basis element of $\mathcal{M}$ and are therefore equal on the whole of $\mathcal{M} \subset H^2$. That is, $T = M_\phi^* | \mathcal{M}$ where $\|M_\phi^*\| \leq 1$, and hence $\|T\| \leq 1$.

Conversely, suppose that $T = T_{K_{z_1}, \ldots, K_{z_n}; \lambda_1, \ldots, \lambda_n}$ is a contraction. By Lemma 1.1.1, T commutes with $S^* | \mathcal{M}$ (i.e. the backward shift restricted to $\mathcal{M}$). The minimal co-isometric dilation of $S^* | \mathcal{M}$ is S^* on H^2, and by the Commutant Lifting Theorem (see [26]) it follows that T is the restriction to $\mathcal{M}$ of an operator M, which commutes with the backward shift and is a contraction. It is a well known fact that those operators which commute with the unilateral shift are exactly the multiplication operators M_ϕ where $\phi \in H^\infty$ (see, for example, [29]). It follows that $M = M_\phi^*$ for some $\phi \in H^\infty$. For $j = 1, \ldots, k$ we have

$$\overline{\lambda_j} K_{z_j} = TK_{z_j} = (M_\phi^* | \mathcal{M}) K_{z_j}. \tag{1.5}$$

Hence,

$$\overline{\lambda_j} K_{z_j} = (M_\phi^* | \mathcal{M}) K_{z_j} = M_\phi^* K_{z_j} = \overline{\phi(z_j)} K_{z_j}$$

for $j = 1, \ldots, k$. Moreover $\|M\| \leq 1$, so we have

$$\|\phi\|_\infty = \|M_\phi\| = \|M^*\| \leq 1.$$

The result follows. ∎

We have now shown that the existence of a solution to the classical Nevanlinna-Pick problem is equivalent to a certain operator being a contraction. This result is essentially the result of Pick [34], which is presented in its more familiar form below.

Corollary 1.1.3 *Let $z_j \in \mathbb{D}$ and $\lambda_j \in \mathbb{C}$ for $j = 1, \ldots, n$. There exists a bounded function $\phi : \mathbb{D} \to \mathbb{C}$ such that $\|\phi\|_\infty \leq 1$ and $\phi(z_j) = \lambda_j$ for $j = 1, \ldots, n$ if and only if*

$$\left[\frac{1 - \lambda_i \overline{\lambda_j}}{1 - z_i \overline{z_j}}\right]_{i,j=1}^n \geq 0$$

Proof. Theorem 1.1.2 states that the existence of an interpolating function satisfying the conditions of the result is equivalent to the operator $T = T_{K_{z_1},\ldots,K_{z_n};\lambda_1,\ldots,\lambda_n}$ being a contraction. Clearly, T is a contraction if and only if

$$1 - T^*T \geq 0.$$

That is, T is a contraction if and only if

$$\left[\langle (1 - T^*T)K_{z_i}, K_{z_j}\rangle\right]_{i,j=1}^n \geq 0,$$

which is the same as

$$\left[\langle K_{z_i}, K_{z_j}\rangle - \langle TK_{z_i}, TK_{z_j}\rangle\right]_{i,j=1}^n \geq 0.$$

By the reproducing property of K_{z_i} discussed above, and the definition of T, we see that this holds if and only if

$$\left[K_{z_i}(z_j) - \lambda_j \overline{\lambda_i} K_{z_i}(z_j)\rangle\right]_{i,j=1}^n \geq 0,$$

or, equivalently, if and only if

$$\left[\frac{1}{1 - z_j \overline{z_i}} - \lambda_j \overline{\lambda_i} \frac{1}{1 - z_j \overline{z_i}}\right]_{i,j=1}^n = \left[\frac{1 - \lambda_j \overline{\lambda_i}}{1 - z_j \overline{z_i}}\right]_{i,j=1}^n \geq 0.$$

■

Pick's Theorem is an elegant, self contained piece of pure mathematics. However, the Nevanlinna-Pick problem is much more than that. It arises in certain engineering disciplines as an important tool in the solution of difficult problems. In the next section we present a discussion of one of these applications.

1.2 Linear Systems

In this section we introduce the reader to a small number of simple concepts in linear systems theory. We demonstrate why control engineers may be interested in the Nevanlinna-Pick problem as a tool to

Figure 1.1: A feedback control block diagram

help them solve difficult physical problems. The fields of linear systems theory and control theory are far too large for us to present any more than a cursory introduction. For a more complete study of the kind of control problems related to Nevanlinna-Pick, we recommend the easily readable book by Doyle, Francis and Tannenbaum [23].

Throughout this section, all linear systems are assumed to be finite dimensional. Our attention will centre on *closed loop feedback systems*, that is, systems which can be represented as in Figure 1.1. In Figure 1.1, G represents the *plant* and C represents a *controller*. Essentially, we think of the plant as performing the primary role of the system while the controller ensures that it behaves correctly. From a mathematical viewpoint, in the linear case, the plant and the controller can be seen as multiplication operators (via the Laplace transform, see [23]). Normally, no distinction is drawn between the actual plant/controller and the multiplication operator it induces. In the case where u is a p-dimensional vector input and y is an r-dimensional vector output, the plant G and the controller C will be $r \times p$ and $p \times r$ matrices respectively. Clearly, if u and y are scalar functions, then so are G and C. The case where everything is scalar is described as SISO (single input, single output).

In the simple block diagram Figure 1.2, we see that the input u and output y satisfy $y(s) = G(s)u(s)$. Here u, y and G are the Laplace transforms of the input, output and plant, which in turn are functions of time. Analysis of models of linear systems based on the use of the Laplace transform is described as *frequency domain analysis*. The (possibly matricial) multiplication operators induced by the boxes in the relevant diagram are known as *transfer functions*.

Definition 3 *A system is* stable *if its transfer function is bounded and analytic in the right half-plane.*

Therefore the system given in Figure 1.2 is stable if and only if, for some $M \in \mathbb{R}$, we have $|G(s)| < M$ for all $s \in \mathbb{C}$ with $\operatorname{Re} s \geq 0$.

The system in Figure 1.1 is meant to represent a (simple) physical system and because of this we often ask for it to satisfy more stringent conditions than those of Definition 3. We say a system is *internally stable* if the transfer function between each input and each branch of the system is stable. This stronger notion of stability is necessary because systems which appear to have a stable transfer function can still have internal instabilities.

Clearly the system in Figure 1.2 is internally stable if and only if it is stable. The system in Figure 1.1 is internally stable if and only if each of the transfer functions

$$(I + GC)^{-1}, \qquad (I + GC)^{-1}G, \qquad C(I + GC)^{-1}, \qquad C(I + GC)^{-1}G$$

Figure 1.2: A simple block diagram

are stable.

It would obviously be of great interest to know which controllers C stabilize the system in Figure 1.1 for a given G. Below, we present a parameterization of all such solutions for a wide class of G.

To simplify the problem of parameterizing all controllers which stabilize the system in Figure 1.1, we shall assume that G is rational and therefore has a co-prime factorization. That is, there exist stable matrices M, N, X and Y such that X and Y are proper, real rational, and

$$G = NM^{-1} \qquad \text{and} \qquad YN + XM = I.$$

Youla proved the following result, a full proof of which can be found in [27, Chapter 4]. A simpler proof in the scalar case can be found in [33].

Theorem 1.2.1 *Let G be a rational plant with co-prime factorisation $G = NM^{-1}$ as above. Then C is a rational controller which internally stabilizes the system given in Figure* 1.1 *if and only if*

$$C = (Y + MQ)(X - NQ)^{-1}$$

for some stable, proper, real rational function Q for which $(X - NQ)^{-1}$ exists.

Thus, in the scalar case, if $G = \frac{N}{M}$ then C produces an internally stable system in Figure 1.1 if and only if

$$C = \frac{Y + MQ}{X - NQ}$$

for some $Q \in H^\infty$ with $X - NQ \neq 0$.

Observe that in the case of an internally stable single-input, single-output (SISO) system we have:

$$\begin{aligned}
\frac{C}{1 + GC} &= \frac{Y + MQ}{X - NQ} \frac{1}{I + \frac{N}{M}\frac{Y + MQ}{X - NQ}} \\
&= \frac{Y + MQ}{X - NQ} \frac{M(X - NQ)}{M(X - NQ) + N(Y + MQ)} \\
&= (Y + MQ)\frac{M}{NY + MX} \\
&= M(Y + MQ).
\end{aligned}$$

It was mentioned above that the Nevanlinna-Pick problem arises as a consequence of robust stabilization. The problem of robust stabilization asks if it is possible to construct a controller which not only stabilizes the feedback system of Figure 1.1, but also stabilizes all other such systems whose plants are 'close' to G. We donote the right halfplane by $\mathbb{C}_+$, and donote the system in Figure 1.1 by (G, C). The following result is taken from [33].

Theorem 1.2.2 *Let (G, C) be an internally stable SISO feedback system over $A(\mathbb{C}_+)$ and suppose that*

$$\left\| \frac{C}{I + GC} \right\|_\infty = \varepsilon.$$

Then C stabilizes $G + \Delta$ for all $\Delta \in A(\mathbb{C}_+)$ with

$$\|\Delta\|_\infty < \frac{1}{\varepsilon}.$$

Suppose we seek a controller C which would stabilize the SISO system $(G + \Delta, C)$ whenever $\|\Delta\|_\infty < 1$. Suppose further that G is a real rational function. Clearly, by Theorem 1.2.2, it will suffice to find Q such that

$$\left\| \frac{C}{I + GC} \right\|_\infty = \|M(E + MQ)\|_\infty = \|ME + M^2 Q\|_\infty \stackrel{\text{def}}{=} \|T_1 - T_2 Q\|_\infty \leq 1.$$

By changing variables under the transform $z = (1 - s)/(1 + s)$ we can work with functions over $\mathbb{D}$ rather than $\mathbb{C}_+$. Now if $\phi = T_1 - T_2 Q$ we have $\phi - T_1 = -T_2 Q$. Thus, $\phi(z) = T_1(z)$ for all $z \in \overline{\mathbb{D}}$ with $T_2(z) = 0$.

Conversely, if ϕ does interpolate T_1 at each of the zeros of T_2 then $(T_1 - \phi)/T_2$ is analytic and bounded in $\mathbb{D}$ and as such defines a suitable candidate for Q.

Therefore, our task is to construct a function ϕ on $\mathbb{D}$ such that $\|\phi\|_\infty \leq 1$ and $\phi(z_j) = w_j$ for all z_j satisfying $T_2(z_j) = 0$ where $w_j = T_1(z_j)$. This is clearly the classical Nevanlinna-Pick problem discussed above. It follows that the Nevanlinna-Pick problem is exactly the same as the robust stabilization problem.

Suppose now that we have a slightly different robust stabilization problem. What happens if we know a little about the perturbation Δ? Doyle [21] was the first to consider such *structured robust stabilization problems.* Doyle's approach is based on the introduction of *the structured singular value.* The structured singular value is defined relative to an underlying structure of operators which represent the permissible forms of the perturbation Δ. The definition of μ given here is taken from [25].

Definition 4 *Suppose $\mathcal{H}$ is a Hilbert space and $\mathcal{R}$ is a subalgebra of $\mathcal{L}(\mathcal{H})$ which contains the identity. For $A \in \mathcal{L}(\mathcal{H})$ define*

$$\mu_{\mathcal{R}}(A) = \frac{1}{\inf\{\|T\| \mid T \in \mathcal{R},\ 1 \in \sigma(TA)\}}.$$

Although μ is defined for operators on infinite dimensional Hilbert space, Doyle only defined it in a finite setting which is more in keeping with the name structured singular value. We denote the largest singular value of A by $\bar{\sigma}(A)$.

Figure 1.3: A Robust Stabilization Problem

We consider a closed loop feedback system which is to be stablized, and then remain stable after the addition of a perturbation Δ. Diagramatically, we wish to stabilize the system in Figure 1.3 where $\tilde{G}$ represents the system in Figure 1.1. As before, we shall assume that G is a real rational (matrix) function. We seek a controller C to stabilize $\tilde{G}$ in such a way that it will remain stable under the perturbation Δ.

A result of Doyle's [22, Theorem RSS] states that the system in Figure 1.3 is stable for all Δ of suitable form $\mathcal{R}$ with $\bar{\sigma}(\Delta) < 1$ if and only if

$$\|\tilde{G}\|_\mu \stackrel{\text{def}}{=} \sup_{s\in\mathbb{C}_+} \mu_{\mathcal{R}}(\tilde{G}(s)) \leq 1.$$

If we take the underlying space of matrices $\mathcal{R}$ to be scalar functions (times a suitably sized identity) then $\mu_{\mathcal{R}}(\tilde{G}(s))$ will equal the spectral radius of $\tilde{G(s)}$ which we denote by $r(\tilde{G}(s))$. It follows (with this choice of $\mathcal{R}$) that the system in Figure 1.3 is stable if and only if $r(\tilde{G}(s)) \leq 1$ for all $s \in \mathbb{C}_+$. We obviously require $\tilde{G}$ to be stable under a zero perturbation and we can achieve this by using the Youla parametrization for C. Our task is now to choose the free parameter Q in the Youla parameterization of C in such a way that $r(\tilde{G}(s)) \leq 1$ for all $s \in \mathbb{C}_+$.

Francis [27, Section 4.3, Theorem 1] shows that when C is chosen via the Youla parameterization, there exist matrices T_1, T_2 and T_3 such that $\tilde{G} = T_1 - T_2 Q T_3$. As before, we may choose to work with $\mathbb{D}$ rather than $\mathbb{C}_+$. Thus far we have shown that the system in Figure 1.3 is stable if and only if we can choose a stable, rational, bounded Q such that $r(T_1(z) - T_2(z)Q(z)T_3(z)) \leq 1$ for all $z \in \mathbb{D}$. Clearly, $\tilde{G} - T_1 = -T_2 Q T_3$. Now if x is a vector with the property $T_2(z)Q(z)T_3(z)x = 0$ for some $z \in \mathbb{D}$, then $\tilde{G}(z)x = T_1(z)x$. In other words, the system in Figure 1.3 is stable only if we can construct a stable function $\tilde{G}$ such that $\sup_{z\in\mathbb{D}} \tilde{G}(s) \leq 1$ and $\tilde{G}(z)x = T_1(z)x$ for all z and x such that $T_3(z)x = 0$, with a like condition involving points z and vectors y such that $y^* T_2(z) = 0$. This problem is known as the *tangential spectral Nevanlinna-Pick problem.* Clearly, a special case of the tangential spectral Nevanlinna-Pick problem is the spectral Nevanlinna-Pick problem in which occurs when T_2 and T_3 happen to be scalar functions.

The spectral Nevanlinna-Pick problem is the most difficult case of the tangential spectral Nevanlinna-Pick problem (see [14]). It is also the subject of this work.

1.3 The Spectral Nevanlinna-Pick Problem

We shall study the following interpolation problem and derive a necessary condition for the existence of a solution. Let $\mathbb{M}_k(\mathbb{C})$ denote the space of $k \times k$ matrices with complex entries.

Main Problem *Let k and n be natural numbers. Choose n distinct points z_j in $\mathbb{D}$ and n matrices W_j in $\mathbb{M}_k(\mathbb{C})$. Does there exist an analytic function $\phi : \mathbb{D} \to \mathbb{M}_k(\mathbb{C})$ such that*

$$\phi(z_j) = W_j$$

for $j = 1, ..., n$ and

$$\sigma(\phi(z)) \subset \overline{\mathbb{D}}$$

for all $z \in \mathbb{D}$?

To simplify the statement of the problem we shall introduce the following sets. Let Σ_k denote the set of complex $k \times k$ matrices whose spectra are contained in the closed unit disc. For $z = (z_1, \ldots, z_k) \in \mathbb{C}^k$ let $c_t(z)$ represent the t^{th} elementary symmetric polynomial in the components of z. That is, for each $z \in \mathbb{C}^k$ let

$$c_t(z) = \sum_{1 \le r_1 < \cdots < r_t \le k} z_{r_1} \cdots z_{r_t}.$$

For completeness, define $c_0(z) = 1$ and $c_r(z) = 0$ for $r > k$.

Let Γ_k be the region of $\mathbb{C}^k$ defined as follows:

$$\Gamma_k = \{(c_1(z), ..., c_k(z)) \mid z \in \overline{\mathbb{D}}^k\}.$$

Define the mapping $\pi : \mathbb{C}^k \to \mathbb{C}^k$ by

$$\pi(z) = (c_1(z), \ldots, c_k(z)).$$

Notice that if A is a complex $k \times k$ matrix with eigenvalues $\lambda_1, \ldots, \lambda_k$ (repeated according to multiplicity), then $A \in \Sigma_k$ if and only if $(c_1(\lambda_1, \ldots, \lambda_k), \ldots, c_k(\lambda_1, \ldots, \lambda_k)) \in \Gamma_k$. Motivated by this observation we extend the definition of c_t to enable it to take matricial arguments.

Definition 5 *For a matrix $W \in \mathbb{M}_k(\mathbb{C})$ we define $c_j(W)$ as the coefficient of $(-1)^k \lambda^{k-j}$ in the polynomial $\det(\lambda I_k - W)$. Define $a(W)$ as $(c_1(W), \ldots, c_k(W))$.*

Observe that $c_j(W)$ is a polynomial in the entries of W and is the j^{th} elementary symmetric polynomial in the eigenvalues of W. The function a is a polynomial function in the entries of W_j, and as such is analytic.

We can now show that each target value in Σ_k of the Main Problem gives rise to a corresponding point in Γ_k. If $\phi : \mathbb{D} \to \Sigma_k$ is an analytic function which satisfies the conditions of the Main Problem, then the function

$$a \circ \phi : \mathbb{D} \to \Gamma_k$$

is analytic because it is the composition of two analytic functions, and furthermore maps points in $\mathbb{D}$ to points in Γ_k. Thus, the existence of a solution to the Main Problem implies the existence of an analytic Γ_k-valued function on the disc. In other words, the existence of an analytic interpolating function from the disc into Γ_k is a necessary condition for the existence of an interpolating function from the disc into Σ_k. In the (generic) case, when $W_1, \ldots, W_k$ are non-derogatory (i.e. when their characteristic and minimal polynomials are the same), the Σ_k and Γ_k problems are equivalent (see discussion in Chapter 8). We therefore seek a necessary condition for the existence of a solution to the following problem, which in turn will provide a necessary condition for the existence of a solution to the Main Problem.

Γ_k **Problem** *Given n distinct points z_j in $\mathbb{D}$ and n points γ_j in Γ_k, does there exist an analytic function $\phi : \mathbb{D} \to \Gamma_k$ such that $\phi(z_j) = \gamma_j$ for $j = 1, \ldots, n$?*

Throughout this work, the reader may assume $k > 1$. We show that a necessary condition for the existence of a solution to the Γ_k problem can be expressed in terms of the positivity of a particular operator polynomial. For $k \in \mathbb{N}$ introduce the polynomial P_k given by

$$P_k(x_0, \ldots, x_k; y_0, \ldots, y_k) = \sum_{r,s=0}^{k} (-1)^{r+s}(k - (r + s))y_s x_r. \tag{1.6}$$

Similar polynomials arise throughout this work in subtly different contexts although they can all be represented in terms of P_k. The following result is simply a matter of re-arranging this polynomial.

Lemma 1.3.1 *The following identity holds:*

$$P_k(x_0, \ldots, x_k; y_0, \ldots, y_k) = \frac{1}{k} \left[A_k(y)A_k(x) - B_k(y)B_k(x)\right]$$

where

$$A_k(y) = \sum_{s=0}^{k} (-1)^s (k - s) y_s$$

and

$$B_k(y) = \sum_{s=0}^{k} (-1)^s s y_s.$$

Proof.

$$\begin{aligned}
P_k(x_0,\dots,x_k;y_0,\dots,y_k) &= \sum_{r,s=0}^{k}(-1)^{r+s}(k-(r+s))x_r y_s \\
&= \sum_{r,s=0}^{k}(-1)^{r+s}\frac{1}{k}(k^2-(r+s)k)x_r y_s \\
&= \sum_{r,s=0}^{k}(-1)^{r+s}\frac{1}{k}(k^2-(r+s)k+rs-rs)x_r y_s \\
&= \sum_{r,s=0}^{k}(-1)^{r+s}\frac{1}{k}((k-r)(k-s)-rs)x_r y_s \\
&= \sum_{r,s=0}^{k}(-1)^{r+s}\frac{1}{k}(k-r)(k-s)x_r y_s - \sum_{r,s=0}^{k}(-1)^{r+s}\frac{1}{k}rsx_r y_s \\
&= \frac{1}{k}\left(\sum_{r,s=0}^{k}(-1)^{r+s}(k-r)(k-s)x_r y_s - \sum_{r,s=0}^{k}(-1)^{r+s}rsx_r y_s\right) \\
&= \frac{1}{k}\left(\sum_{s=0}^{k}(-1)^{s}(k-s)y_s\right)\left(\sum_{r=0}^{k}(-1)^{r}(k-r)x_r\right) \\
&\qquad - \frac{1}{k}\left(\sum_{s=0}^{k}(-1)^{s}sy_s\right)\left(\sum_{r=0}^{k}(-1)^{r}rx_r\right)
\end{aligned}$$

■

Thus, for example,

$$P_2(x_0,x_1,x_2;y_0,y_1,y_2) = \frac{1}{2}[(2y_0-y_1)(2x_0-x_1)-(-y_1+2y_2)(-x_1+2x_2)]$$

This polynomial, in a number of different guises, will be of great interest to us while we study the Γ_k problem for given $k \in \mathbb{N}$. Indeed, it is essentially the polynomial which we use to express our necessary condition for the existence of a solution to the Γ_k problem.

Chapter 2

Hereditary Polynomials

2.1 Model Operators and Complete Spectral Sets

We begin this chapter by recalling some definitions from the Introduction. We then prove a result which is analogous to Theorem 1.1.2 in the sense that it interprets an interpolation problem in terms of the properties of a set of operators. In Chapter 1 we demonstrated how such a condition gives rise to a solution to the classical Nevanlinna-Pick Problem. The derivation of this solution relied on the hereditary polynomial $f(x, y) = 1 - yx$. We used the fact that an operator T is a contraction if and only if $f(T, T^*) = 1 - T^*T$ is positive semi-definite. The later part of this chapter contains the derivation of an hereditary polynomial which will be used in a similar way to f. Namely, we derive an hereditary polynomial g with the property that a k-tuple of operators $T_1, \ldots, T_k$ is a Γ_k-contraction only if $g(T_1, \ldots, T_k, T_1^*, \ldots, T_k^*) \geq 0$. In the chapters which follow, we show that the polynomial derived here does indeed possess the desired properties and therefore gives rise to a partial solution to the Γ_k problem.

As in Chapter 1 we denote by H^2 the Hardy space of analytic functions on $\mathbb{D}$ which have square summable Taylor coefficients and by K its reproducing kernel (see equation (1.1)). The reader will recall that for any $(z_1, \ldots, z_n, \lambda_1, \ldots, \lambda_n) \in \mathbb{D}^k \times \mathbb{C}^k$ with $z_i \neq z_j$, we defined the space $\mathcal{M}$ and the operator $T_{K_{z_1},\ldots,K_{z_n};\lambda_1,\ldots,\lambda_n}$ by

$$\mathcal{M} = \operatorname{Span}\{K_{z_1}, \ldots, K_{z_n}\} \quad \text{and} \quad T_{K_{z_1},\ldots,K_{z_n};\lambda_1,\ldots,\lambda_n} K_{z_j} = \bar{\lambda}_j K_{z_j}.$$

We have seen how operators of this type can be used to provide a full solution to the Main Problem when $k = 1$ (Chapter 1). More recently, however, they have been used by Agler and Young [6] to establish a necessary condition for the existence of a solution to the Main Problem when $k = 2$. The rest of this chapter is devoted to showing how the methods of [6] can be extended to give a necessary condition for the existence of a solution to the Main Problem for general k. First we require a definition.

Definition 6 *For any $p \times q$ matricial polynomial h in k variables*

$$h(x_1, \dots, x_k) = \left[\sum_{r_1, \dots, r_k} a_{i,j,r_1 \cdots r_k} {x_1}^{r_1} \cdots {x_k}^{r_k} \right]_{\substack{i=1,\dots,p \\ j=1,\dots,q}}$$

we denote by $h^\vee$ the conjugate polynomial

$$h^\vee(x_1, \dots, x_k) = \left[\sum_{r_1, \dots, r_k} \overline{a_{i,j,r_1 \cdots r_k}} {x_1}^{r_1} \cdots {x_k}^{r_k} \right]_{\substack{i=1,\dots,p \\ j=1,\dots,q}} = \overline{h(\bar{x_1}, \dots, \bar{x_k})}.$$

For any function ϕ analytic in a neighbourhood of each $\overline{\lambda_j}$, $j = 1, \dots, n$ we define

$$\phi(T_{K_{z_1}, \dots, K_{z_n}; \lambda_1, \dots, \lambda_n}) = T_{K_{z_1}, \dots, K_{z_n}; \phi^\vee(\lambda_1), \dots, \phi^\vee(\lambda_n)}.$$

For $j = 1, \dots, n$ let z_j be distinct points in $\mathbb{D}$ and define

$$\mathcal{M} = \operatorname{Span}\{K_{z_1}, \dots, K_{z_n}\}.$$

Pick n points $({c_1}^{(1)}, \dots, {c_k}^{(1)}), \dots, ({c_1}^{(n)}, \dots, {c_k}^{(n)}) \in \Gamma_k$ and let

$$C_i = T_{K_{z_1}, \dots, K_{z_n}; {c_i}^{(1)}, \dots {c_i}^{(n)}} \qquad \text{for } i = 1, \dots, k. \tag{2.1}$$

These operators are diagonal with respect to the basis $K_{z_1}, \dots, K_{z_n}$ and thus they commute. Commutativity can also be proven by considering two such model operators, C_r and C_t, acting on a basis element of $\mathcal{M}$:

$$C_r C_t K_{z_j} = C_r \overline{{c_t}^{(j)}} K_{z_j} = \overline{{c_r}^{(j)} {c_t}^{(j)}} K_{z_j} = \overline{{c_t}^{(j)} {c_r}^{(j)}} K_{z_j} = C_t \overline{{c_r}^{(j)}} K_{z_j} = C_t C_r K_{z_j}.$$

The initial information in the Γ_k problem has now been encoded in a k-tuple of commuting operators on a subspace of H^2.

Next we define the joint spectrum of a k-tuple of operators in keeping with the definition in [6]. This definition is different from that used by Arveson in [9]; it is the simplest of many forms of joint spectrum (see, for example, [24, Chapter 2]), but is sufficient for our purpose.

Definition 7 *Let $X_1, \dots, X_k$ be operators on a Hilbert space and let A be the $*$-algebra generated by these operators. We define $\sigma(X_1, \dots, X_k)$, the joint spectrum of $X_1, \dots, X_k$, by*

$$\sigma(X_1, \dots, X_k) = \{\lambda \in \mathbb{C}^k \mid \exists \text{ a proper ideal } I \subset A \text{ with } \lambda_j - X_j \in I \text{ for } j = 1, \dots, k\}.$$

Definition 8 *A set $E \subset \mathbb{C}^k$ is said to be a* complete spectral set *for a k-tuple of commuting operators $(X_1, \dots, X_k)$ on a Hilbert space H if*

$$\sigma(X_1, \dots, X_k) \subset E$$

and if, for any $q \times p$ matrix-valued function h of k variables which is analytic on $\overline{E}$, we have:

$$\|h(X_1, \ldots, X_k)\|_{\mathcal{L}(H^p, H^q)} \leq \sup_E \|h(z)\|. \tag{2.2}$$

The requirement that h be an analytic function in the above definition can be replaced by a more managable alternative when $E = \Gamma_k$. Below we show that it is sufficient to consider only polynomial functions h. To prove this result we need the following definition and a classical result.

Definition 9 *Let S_k represent the group of permutations on the symbols $1, \ldots, k$ and let $\mathrm{St}_k(1)$ be the subgroup of S_k comprising those permutations which leave the symbol 1 unaltered. Let $\mathrm{St}_k(1)_2$ denote the set of elements of $\mathrm{St}_k(1)$ which have order 2. Suppose $\sigma \in S_k$. For $\lambda \in \mathbb{D}^k$ write $\lambda = (\lambda_1, \ldots, \lambda_k)$ and define λ^σ as*

$$\lambda^\sigma = (\lambda_{\sigma(1)}, \ldots, \lambda_{\sigma(k)}).$$

Thus, for example, if $k = 3$ and (12) denotes the element of S_3 which interchanges the first and second symbols, then $(\lambda_1, \lambda_2, \lambda_3)^{(12)} = (\lambda_2, \lambda_1, \lambda_3)$. A proof of the following classical Lemma may be found in [37].

Lemma 2.1.1 *Let f be a symmetric polynomial in indeterminates $x_1, \ldots, x_k$. Then there exists a polynomial p such that*

$$f(x_1, \ldots, x_k) = p(c_1(x_1, \ldots, x_k), \ldots, c_k(x_1, \ldots, x_k)).$$

With this result we may prove the following Lemma which allows us to replace the analytic matrix functions in (2.2) with polynomial matricial functions.

Lemma 2.1.2 *The space of polynomial functions on Γ_k is dense in the space of analytic functions on Γ_k.*

Proof. If f is an analytic function on Γ_k then $f \circ \pi$ is an analytic function on $\overline{\mathbb{D}}^k$. The set $\mathbb{D}^k$ is a Reinhardt domain (see [28]) and so $f \circ \pi$ can be uniformly approximated by a polynomial function on the closed polydisc. Let $\varepsilon > 0$. Choose a polynomial function p on the polydisc such that

$$\sup_{z \in \mathbb{D}^k} |f(\pi(z)) - p(z)| < \varepsilon.$$

For $z \in \mathbb{D}^k$ let

$$q(z) = \frac{1}{k!} \sum_{\sigma \in S_k} p(z^\sigma).$$

Then, for all $z \in \mathbb{D}^k$ we have,

$$\begin{aligned}
k!|f(\pi(z)) - q(z)| &= \left| k!f(\pi(z)) - \sum_{\sigma \in S_k} p(z^\sigma) \right| \\
&= \left| \sum_{\sigma \in S_k} [f(\pi(z)) - p(z^\sigma)] \right| \\
&\leq \sum_{\sigma \in S_k} |f(\pi(z)) - p(z^\sigma)| \\
&= \sum_{\sigma \in S_k} |f(\pi(z^\sigma)) - p(z^\sigma)| \\
&< \varepsilon k!.
\end{aligned}$$

It follows that q is a symmetric polynomial function on the polydisc which approximates $f \circ \pi$. By Lemma 2.1.1 there exists a polynomial m such that

$$q(z) = m(c_1(z), \ldots, c_k(z)) = m \circ \pi(z)$$

for all $z \in \mathbb{D}^k$. Let $\gamma \in \Gamma_k$. By definition of Γ_k, $\gamma = \pi(z)$ for some z in $\overline{\mathbb{D}}^k$. Therefore,

$$|f(\gamma) - m(\gamma)| = |f(\pi(z)) - m(\pi(z))| = |f(\pi(z)) - q(z)| < \varepsilon.$$

Hence, the polynomial function induced by m on Γ_k uniformly approximates the analytic function f.

■

Theorem 2.1.3 *If there exists a function $\phi : \mathbb{D} \to \Gamma_k$ which is analytic and has the property that $\phi(z_j) = (c_1{}^{(j)}, \ldots, c_k{}^{(j)})$ for $j = 1, \ldots, n$, then Γ_k is a complete spectral set for the commuting k-tuple of operators $(C_1, \ldots, C_k)$ defined by* (2.1).

Proof. Lemma 2.1.2 states that it will suffice to consider only matricial polynomial functions h on Γ_k.

Consider the scalar polynomial case. Let h be a polynomial in k variables given by

$$h(x_1, \ldots, x_k) = \sum_{r_1, \ldots, r_k} a_{i,j,r_1 \cdots r_k} x_1{}^{r_1} \cdots x_k{}^{r_k}.$$

Observe that, for $1 \leq j \leq n$, we have

$$\begin{aligned}
h(C_1, \ldots, C_k)K_{z_j} &= \sum_{r_1, \ldots, r_k} a_{r_1 \cdots r_k} C_1{}^{r_1} \cdots C_k{}^{r_k} K_{z_j} \\
&= \sum_{r_1, \ldots, r_k} a_{r_1 \cdots r_k} \overline{c_1{}^{(j)}}^{r_1} \cdots \overline{c_k{}^{(j)}}^{r_k} K_{z_j} \\
&= \overline{h^\vee \circ \phi(z_j)} K_{z_j} \\
&= h^\circ \phi^\vee (T_{K_{z_1}, \ldots, K_{z_n}; z_1, \ldots, z_n}) K_{z_j}.
\end{aligned}$$

Hence, if $h = [h_{ij}]$ is a $p \times q$ matrix polynomial and $z \in \{z_1, \ldots, z_n\}$ then

$$h(C_1, \ldots, C_k) \begin{bmatrix} 0 \\ \vdots \\ K_z \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} h_{1j}(C_1, \ldots, C_k) K_z \\ \vdots \\ h_{pj}(C_1, \ldots, C_k) K_z \end{bmatrix} = \begin{bmatrix} h_{1j} \circ \phi^\vee(T_{K_{z_1}, \ldots, K_{z_n}; z_1, \ldots, z_n}) K_z \\ \vdots \\ h_{pj} \circ \phi^\vee(T_{K_{z_1}, \ldots, K_{z_n}; z_1, \ldots, z_n}) K_z \end{bmatrix}$$

$$= h \circ \phi^\vee(T_{K_{z_1}, \ldots, K_{z_n}; z_1, \ldots, z_n}) \begin{bmatrix} 0 \\ \vdots \\ K_z \\ \vdots \\ 0 \end{bmatrix}.$$

Thus

$$h(C_1, \ldots, C_k) = h \circ \phi^\vee(T_{K_{z_1}, \ldots, K_{z_n}; z_1, \ldots, z_n}).$$

By von Neumann's inequality [36, Proposition 8.3], since $T_{K_{z_1}, \ldots, K_{z_n}; z_1, \ldots, z_n}$ is the contraction $S^*|\mathcal{M}$,

$$\begin{aligned} \|h(C_1, \ldots, C_k)\| &= \|h \circ \phi^\vee(T_{K_{z_1}, \ldots, K_{z_n}; z_1, \ldots, z_n})\| \\ &\leq \sup_{\mathbb{D}} \|h \circ \phi^\vee(z)\| \\ &= \sup_{\Gamma_k} \|h(\gamma)\|. \end{aligned}$$

That is, Γ_k is a complete spectral set for $(C_1, \ldots, C_k)$ as required.

■

The model operators introduced in (2.1) have now provided a necessary condition for the existence of a solution to the Γ_k problem. Namely, in the notation of that problem, if an interpolating function exists which maps the z_j to the γ_j, then Γ_k is a complete spectral set for the model operators associated with z_j and γ_j. This naturally raises the question as to which k-tuples of commuting operators have Γ_k as a complete spectral set. This is the topic of the next three sections.

2.2 Properties of Polynomials

We shall consider a class of polynomials in $2k$ arguments known as hereditary polynomials. They will play a major role in establishing a necessary condition for k-tuples of commuting contractions to have Γ_k as a complete spectral set. First some definitions.

Definition 10 *Polynomial functions on $\mathbb{C}^k \times \mathbb{C}^k$ of the form*

$$g(\lambda, z) = \sum_{r_1, \ldots, r_{2k}} a_{r_1 \cdots r_{2k}} z_1^{r_1} \cdots {z_k}^{r_k} \lambda_1^{r_{k+1}} \cdots \lambda_k^{r_{2k}},$$

where $\lambda = (\lambda_1, \ldots, \lambda_k), z = (z_1, \ldots, z_k) \in \mathbb{C}^k$, are said to be hereditary polynomials.

If $g(\lambda, z)$ is such a polynomial, then we may define $g(T_1, \ldots, T_k, {T_1}^*, \ldots, {T_k}^*)$ for a k-tuple of commuting operators on a Hilbert space H as

$$g(T_1, \ldots, T_k, {T_1}^*, \ldots, {T_k}^*) = \sum_{r_1, \ldots, r_{2k}} a_{r_1 \cdots r_{2k}} {T_1}^{*r_1} \cdots {T_k}^{*r_k} {T_1}^{r_{k+1}} \cdots {T_k}^{r_{2k}}.$$

For convenience, we shall abbreviate the operator polynomial $g(T_1, \ldots, T_k, {T_1}^*, \ldots, {T_k}^*)$ to $g(T_1, \ldots, T_k)$ and $g(x_1, \ldots, x_k, \bar{x_1}, \ldots, \bar{x_k})$ to $g(x_1, \ldots, x_k)$ whenever suitable.

Note that although the T_j commute with one another, T_j^* need not commute with T_i. The polynomials are said to be hereditary because if $g(T_1, \ldots T_k, {T_1}^*, \ldots {T_k}^*) \geq 0$ on a Hilbert space H, and $\tilde{T_j}$ is the compression of T_j to an invariant subspace of H then $g(\tilde{T_1}, \ldots \tilde{T_k}, \tilde{T_1}^*, \ldots \tilde{T_k}^*) \geq 0$. Next we consider some properties of general polynomials.

Definition 11 *An hereditary polynomial g is said to be* weakly symmetric *if*

$$g(\lambda, z) = g(\lambda^\sigma, z^\sigma)$$

for all $\sigma \in S_k$, $\lambda, z \in \mathbb{C}^k$ and doubly symmetric *if*

$$g(\lambda, z) = g(\lambda^\sigma, z) = g(\lambda, z^\sigma)$$

for all $\sigma \in S_k$, $\lambda, z \in \mathbb{C}^k$.

Note that all doubly symmetric polynomials are weakly symmetric.

Definition 12 *Let X be a set. A function $g : X \times X \to \mathbb{C}$ is said to be* positive semi-definite *if, for any $n \in \mathbb{N}$ and $x_1, \ldots x_n \in X$, we have*

$$[g(x_i, x_j)]_{i,j=1}^n \geq 0.$$

2.3 Hereditary Polynomial Representations

Take h to be a scalar-valued polynomial on $\mathbb{C}^k$ such that $\|(h \circ \pi)(T_1, \ldots, T_k)\| \leq 1$ for all k-tuples of commuting contractions $(T_1, \ldots, T_k)$. We may define an hereditary polynomial $g : \mathbb{D}^k \times \mathbb{D}^k \to \mathbb{C}$ which is positive on all k-tuples of commuting contractions by

$$g(\lambda, \bar{z}) = 1 - \overline{h \circ \pi(z)} h \circ \pi(\lambda). \tag{2.3}$$

It is easy to observe that g is doubly symmetric.

Recall a version of a theorem by Agler [1], which we will refine for our own use.

Theorem 2.3.1 *Let f be a polynomial function defined on $\mathbb{D}^k \times \mathbb{D}^k$. Then $f(T_1, \dots, T_k, {T_1}^*, \dots, {T_k}^*) \geq 0$ for all k-tuples of commuting contractions $(T_1, \dots, T_k)$ if and only if there exist k Hilbert spaces $\mathcal{H}_1, \dots, \mathcal{H}_k$ and k holomorphic functions $f_1, \dots f_k$ such that $f_r : \mathbb{D}^k \to \mathcal{H}_r$ and*

$$f(\lambda, \bar{z}) = \sum_{r=1}^{k} (1 - \lambda_r \bar{z}_r) f_r(z)^* f_r(\lambda)$$

for all $\lambda, z \in \mathbb{D}^k$.

This result holds for all holomorphic functions f on $\mathbb{D}^k \times \mathbb{D}^k$ for which $f(T_1, \dots, T_k)$ is defined, but the stated version is sufficient for our purpose. Since the hereditary polynomials of interest, namely those of the form (2.3), are weakly symmetric, we may extend Agler's theorem in the following way:

Theorem 2.3.2 *Let f be a weakly symmetric hereditary polynomial on $\mathbb{C}^k \times \mathbb{C}^k$. Then f is positive on all k-tuples of commuting contractions if and only if there exists a positive semi-definite function Φ on $\mathbb{D}^k \times \mathbb{D}^k$ such that*

$$f(\lambda, \bar{z}) = \sum_{r=1}^{k} (1 - \lambda_r \bar{z}_r) \Phi(\lambda^{\nu_r}, z^{\nu_r})$$

for all $\lambda, z \in \mathbb{D}^k$ and for any choice of $\nu_1, \dots, \nu_k \in S_k$ such that $\nu_r(1) = r$ for $r = 1, \dots, k$.

Proof. $(\Rightarrow)$ Let f be positive on k-tuples of commuting contractions. By Theorem 2.3.1, there exist k Hilbert spaces $\mathcal{H}_1, \dots, \mathcal{H}_k$ and k $\mathcal{H}_r$-valued functions f_r such that

$$f(\lambda, \bar{z}) = \sum_{r=1}^{k} (1 - \lambda_r \bar{z}_r) f_r(z)^* f_r(\lambda)$$

for all $\lambda, z \in \mathbb{D}^k$. For $r = 1, \dots, k$ let a_r be the positive semidefinite function defined on $\mathbb{D}^k \times \mathbb{D}^k$ by $a_r(\lambda, z) = f_r(z)^* f_r(\lambda)$. Then

$$f(\lambda, \bar{z}) = \sum_{j=1}^{k} (1 - \lambda_r \bar{z}_r) a_r(\lambda, z)$$

for all $\lambda, z \in \mathbb{D}^k$. For $t = 1, \dots, k$ define $b_t : \mathbb{D}^k \times \mathbb{D}^k \to \mathbb{C}$ by

$$b_t(\lambda, z) = \sum_{\sigma \in S_k} \frac{1}{k!} a_{\sigma^{-1}(t)}(\lambda^\sigma, z^\sigma).$$

Clearly each b_t is positive semi-definite. Pick $r \in \{1, \dots, k\}$ and $\tau \in S_k$ such that $\tau(1) = r$. Let $\nu = \tau\sigma$. Consider $b_1(\lambda^\tau, z^\tau)$:

$$\begin{aligned} b_1(\lambda^\tau, z^\tau) &= \sum_{\sigma \in S_k} \frac{1}{k!} a_{\sigma^{-1}(1)}(\lambda^{\tau\sigma}, z^{\tau\sigma}) \\ &= \sum_{\sigma \in S_k} \frac{1}{k!} a_{\sigma^{-1}\tau^{-1}(r)}(\lambda^{\tau\sigma}, z^{\tau\sigma}) \\ &= \sum_{\nu \in S_k} \frac{1}{k!} a_{\nu^{-1}(r)}(\lambda^{\nu}, z^{\nu}) \\ &= b_r(\lambda, z). \end{aligned}$$

Let $\Phi(\lambda, z) = b_1(\lambda, z)$, so that $b_r(\lambda, z) = \Phi(\lambda^\nu, z^\nu)$ whenever $\nu \in S_k$ satisfies $\nu(1) = r$.

Since f is weakly symmetric,

$$f(\lambda, \bar{z}) = \frac{1}{k!} \sum_{\sigma \in S_k} f(\lambda^\sigma, \bar{z}^\sigma) = \frac{1}{k!} \sum_{\sigma \in S_k} \sum_{j=1}^{k} (1 - \lambda_{\sigma(j)} \bar{z}_{\sigma(j)}) a_j(\lambda^\sigma, z^\sigma).$$

Changing variables under the substitution $\sigma(j) = t$ this can be rewritten as

$$\begin{aligned} f(\lambda, \bar{z}) &= \sum_{t=1}^{k} (1 - \lambda_t \bar{z}_t) \sum_{\sigma \in S_k} \frac{1}{k!} a_{\sigma^{-1}(t)}(\lambda^\sigma, z^\sigma) \\ &= \sum_{t=1}^{k} (1 - \lambda_t \bar{z}_t) b_t(\lambda, z). \end{aligned}$$

For any choice of $\nu_1, \ldots, \nu_k \in S_k$ such that $\nu_j(1) = j$, we may substitute Φ defined above to conclude

$$f(\lambda, \bar{z}) = \sum_{r=1}^{k} (1 - \lambda_r \bar{z}_r) \Phi(\lambda^{\nu_r}, z^{\nu_r})$$

as required.

($\Leftarrow$) Suppose there exists a positive semi-definite function Φ such that

$$f(\lambda, \bar{z}) = \sum_{r=1}^{k} (1 - \lambda_r \bar{z}_r) \Phi(\lambda^{\nu_r}, z^{\nu_r})$$

for all $\lambda, \bar{z} \in \mathbb{D}^k$ and for any choice of $\nu_1, \ldots, \nu_k \in S_k$ such that $\nu_r(1) = r$ for $r = 1, \ldots, k$. For $r = 1, \ldots, k$ define $a_r(\lambda, z) = \Phi(\lambda^{\nu_r}, z^{\nu_r})$. Since Φ is positive semi-definite, so is a_r for $r = 1, \ldots, k$. That is, there exist k positive semi-definite functions a_r such that

$$f(\lambda, \bar{z}) = \sum_{r=1}^{k} (1 - \lambda_r \bar{z}_r) a_r(\lambda, z)$$

for all $\lambda, z \in \mathbb{D}^k$. Theorem 2.3.1 then shows that f is positive on k-tuples of commuting contractions.

■

Denote by γ the cycle of order k in S_k which maps k to 1 and t to $t+1$ for $t < k$, i.e. $\gamma = (123 \ldots k)$. Then the above result gives:

Corollary 2.3.3 *Let f be a weakly symmetric hereditary polynomial on $\mathbb{C}^k \times \mathbb{C}^k$. If f is positive on k-tuples of commuting contractions then there exists a positive semi-definite function Φ on $\mathbb{D}^k \times \mathbb{D}^k$ such that*

$$f(\lambda, \bar{z}) = \sum_{t=1}^{k} (1 - \lambda_t \bar{z}_t) \Phi(\lambda^{\gamma^{t-1}}, z^{\gamma^{t-1}})$$

for all $\lambda, \bar{z} \in \mathbb{D}^k$.

Proof. It is clear that $\gamma^{t-1}(1) = t$, and so we may apply Theorem 2.3.2.

$\blacksquare$

Recall a classic theorem of E.H. Moore and N. Aronszajn (for a proof see [7]).

Theorem 2.3.4 *If Ψ is a positive semi-definite function on $\mathbb{D}^k \times \mathbb{D}^k$ then there exists a Hilbert space $\mathcal{E}$ with inner product $\langle \cdot, \cdot \rangle$ and an $\mathcal{E}$-valued function H on $\mathbb{D}^k$ such that*

$$\Psi(\lambda, \bar{z}) = \langle H(\lambda), H(\bar{z}) \rangle$$

for all $\lambda, \bar{z} \in \mathbb{D}^k$.

Denote by Φ the positive definite function formed by applying Corollary 2.3.3 to the function g defined in (2.3) so that,

$$g(\lambda, \bar{z}) = \sum_{t=1}^{k} (1 - \lambda_t \bar{z}_t) \Phi(\lambda^{\gamma^{t-1}}, z^{\gamma^{t-1}}).$$

Let $\mathcal{C}$ be the Hilbert space with inner product $\langle \cdot, \cdot \rangle$ which is formed by applying Theorem 2.3.4 to Φ and let F be the corresponding $\mathcal{C}$-valued function. We may now write g as

$$g(\lambda, \bar{z}) = \sum_{t=1}^{k} (1 - \lambda_t \bar{z}_t) \langle F(\lambda^{\gamma^{t-1}}), F(z^{\gamma^{t-1}}) \rangle \tag{2.4}$$

for all $\lambda, \bar{z} \in \mathbb{D}^k$.

2.4 Finding the Polynomial

In this section we utilise the results of the previous section to study the properties of polynomials of the form (2.3). Our aim is to derive a representation formula for these polynomials.

Lemma 2.4.1 *Let g be the polynomial in $2k$ indeterminates defined in (2.3) and let $F : \mathbb{D}^k \to \mathcal{C}$ be as in (2.4). For every $\sigma \in \mathrm{St}_k(1)$ there exists a corresponding unitary operator $U_\sigma : \mathcal{C} \to \mathcal{C}$ such that*

$$U_\sigma F(\lambda) = F(\lambda^\sigma)$$

for all $\lambda \in \mathbb{D}^k$. Furthermore, the mapping $\sigma \mapsto U_\sigma$ is an anti-representation of $\mathrm{St}_k(1)$.

Proof. Theorem 2.3.2 implies that the function Φ has the property that $\Phi(\lambda^\sigma, z^\sigma) = \Phi(\lambda^\tau, z^\tau)$ whenever $\sigma(1) = \tau(1)$ and in particular, if $\sigma \in \mathrm{St}_k(1)$ we have

$$\langle F(\lambda), F(z) \rangle = \Phi(\lambda, z) = \Phi(\lambda^\sigma, z^\sigma) = \langle F(\lambda^\sigma), F(z^\sigma) \rangle$$

for all $\lambda, z \in \mathbb{D}^k$. It follows that there exists an isometry U_σ mapping $F(\lambda)$ to $F(\lambda^\sigma)$ for all $\lambda \in \mathbb{D}^k$. However, since the linear span of $\{F(\lambda) \mid \lambda \in \mathbb{D}^k\}$ may be assumed dense in $\mathcal{C}$, we see that U_σ is a unitary operator satisfying

$$U_\sigma F(\lambda) = F(\lambda^\sigma)$$

for all $\lambda \in \mathbb{D}^k$. Clearly, every element of $\mathrm{St}_k(1)$ gives rise to a unitary operator in this manner. If σ and τ are elements of $\mathrm{St}_k(1)$ then their product $\sigma\tau$ is also an element of $\mathrm{St}_k(1)$. The three unitaries which are associated with these elements are related as indicated by the following equality:

$$U_{\sigma\tau}F(\lambda) = F(\lambda^{\sigma\tau}) = U_\tau F(\lambda^\sigma) = U_\tau U_\sigma F(\lambda)$$

for all $\lambda \in \mathbb{D}^k$. That is $U_{\sigma\tau} = U_\tau U_\sigma$ and $\sigma \mapsto U_\sigma$ is an anti-representation of $\mathrm{St}_k(1)$.

■

Lemma 2.4.2 *If $\sigma \in \mathrm{St}_k(1)_2$ then the corresponding unitary operator U_σ is self-adjoint. Moreover, if $\mathcal{C}$ is one dimensional, then U_σ is the identity operator.*

Proof. Suppose $\sigma \in \mathrm{St}_k(1)_2$. Then

$$U_\sigma{}^2 F(\lambda) = U_\sigma U_\sigma F(\lambda) = U_\sigma F(\lambda^\sigma) = F(\lambda^{\sigma^2}) = F(\lambda)$$

for all $\lambda \in \mathbb{D}^k$. It follows that U_σ is self-adjoint. This completes the proof of the first statement.

Suppose the space $\mathcal{C}$ corresponding to the hereditary polynomial g is one dimensional. Then $U_\sigma = \pm I$ for all $\sigma \in \mathrm{St}_k(1)_2$. Consider the case where $U_\sigma = -I$ for some σ in $\mathrm{St}_k(1)_2$ and define the diagonal of $\mathbb{D}^k$ by

$$\mathcal{D} = \{\lambda \in \mathbb{D}^k | \lambda_1 = \cdots = \lambda_k\}.$$

By assumption, for all $\lambda \in \mathbb{D}^k$, we have

$$-F(\lambda) = U_\sigma F(\lambda) = F(\lambda^\sigma).$$

Therefore $F(\lambda) = 0$ for all $\lambda \in \mathcal{D}$. By Equation (2.4),

$$g(\lambda, \bar{z}) = \sum_{j=1}^{k} (1 - \lambda_j \bar{z}_j)\langle F(\lambda^{\gamma^{j-1}}), F(z^{\gamma^{j-1}})\rangle.$$

Hence, $g(\lambda, \bar{z}) = 0$ whenever λ or z is in $\mathcal{D}$. Fix $\lambda \in \mathcal{D}$. Then, for all $z \in \mathbb{D}^k$,

$$0 = g(\lambda, \bar{z}) = 1 - \overline{h \circ \pi(z)} h \circ \pi(\lambda).$$

Thus, for any $z \in \mathbb{D}^k$,

$$\overline{h \circ \pi(z)} = (h \circ \pi(\lambda))^{-1}.$$

Since λ is fixed, the right hand side of this equation is constant, thus $h \circ \pi(z)$ is constant on $\mathbb{D}^k$ and h is constant on Γ_k. This contradicts the choice of h as any scalar valued polynomial on Γ_k such that $\|h \circ \pi(T_1, \ldots, T_k)\| \leq 1$ for all k-tuples of commuting contractions $(T_1, \ldots, T_k)$. This contradiction implies $U_\sigma \neq -I$.

Thus, whenever $\mathcal{C}$ is one dimensional U_σ is the identity for all $\sigma \in \mathrm{St}_k(1)_2$.

■

The following corollary extends Lemma 2.4.2. We show that, in the case where $\mathcal{C}$ is one dimensional, $U_\tau = I$ for all $\tau \in \mathrm{St}_k(1)$, not just $\mathrm{St}_k(1)_2$.

Corollary 2.4.3 *Let $g, F, \mathcal{C}$ be as in (2.4) and suppose that $\mathcal{C}$ is one dimensional. Then for every $\tau \in \mathrm{St}_k(1)$, the corresponding unitary U_τ is the identity, and hence,*

$$F(\lambda^\tau) = F(\lambda).$$

Proof. Let $\dim \mathcal{C} = 1$. Lemma 2.4.2 states that $U_\sigma = I$ whenever $\sigma \in \mathrm{St}_k(1)_2$. It is trivial to show that every element of $\mathrm{St}_k(1)$ can be written as a product of elements in $\mathrm{St}_k(1)_2$. Pick an element $\tau \in \mathrm{St}_k(1)$ and suppose that $\tau = \sigma_1 \cdots \sigma_s$ is a factorisation of τ over $\mathrm{St}_k(1)_2$. For every $\lambda \in \mathbb{D}^k$ we have:

$$F(\lambda^\tau) = F(\lambda^{\sigma_1 \cdots \sigma_s}) = U_{\sigma_s} F(\lambda^{\sigma_1 \cdots \sigma_{s-1}}) = \cdots = U_{\sigma_s} \cdots U_{\sigma_1} F(\lambda) = I^s F(\lambda) = F(\lambda).$$

That is, for every element τ in $\mathrm{St}_k(1)$, the associated matrix U_τ is the identity. ∎

Lemma 2.4.4 *Let $\mathcal{C}$ be one dimensional. Choose $t \in \{1, \dots, k\}$ and pick $\sigma \in S_k$ such that $\sigma(1) = t$. Then*

$$F(\lambda^{\gamma^{t-1}}) = F(\lambda^{\sigma}). \tag{2.5}$$

Proof. Let $\tau \in \mathrm{St}_k(1)$. It was shown in Corollary 2.4.3 that $F(\lambda^{\tau}) = F(\lambda)$. Replacing λ with $\lambda^{\gamma^{t-1}}$ in this equation we have,

$$F(\lambda^{\gamma^{t-1}\tau}) = F(\lambda^{\gamma^{t-1}})$$

for all $\lambda \in \mathbb{D}^k$ and all $\tau \in \mathrm{St}_k(1)$. It is easy to show that

$$\{\gamma^{t-1}\tau \mid \tau \in \mathrm{St}_k(1)\} = \{\sigma \mid \sigma \in S_k,\ \sigma(1) = t\}$$

and hence for any $\lambda \in \mathbb{D}^k$ we have

$$F(\lambda^{\gamma^{t-1}}) = F(\lambda^{\sigma})$$

whenever $\sigma(1) = t$.

■

Lemma 2.4.5 *Let $\mathcal{C}$ be one dimensional. There exists an $\alpha \in \mathbb{T}$ such that*

$$(1 - \alpha\lambda_t)F(\lambda^{\gamma^{t-1}}) = (1 - \alpha\lambda_{t+1})F(\lambda^{\gamma^{t}}) \tag{2.6}$$

for all $\lambda \in \mathbb{D}^k$ and $t = 1, \dots, k-1$. Furthermore the function $\mathcal{S} : \mathbb{D}^k \to \mathcal{C}$ defined by

$$\mathcal{S}(\lambda) = (1 - \alpha\lambda_1)F(\lambda)$$

is symmetric on $\mathbb{D}^k$ under the action of S_k.

Proof. Denote by (12) the element of S_k which exchanges the first two symbols and recall that the polynomial g, defined in (2.3), is doubly symmetric. By definition of double symmetry we have

$$g(\lambda, \bar{z}) = g(\lambda^{(12)}, \bar{z}).$$

Using (2.4) and (2.5) we can expand this equality to see that

$$\begin{aligned}
(1 - \lambda_1\bar{z}_1)\langle F(\lambda), F(z)\rangle \quad &+ \quad (1 - \lambda_2\bar{z}_2)\langle F(\lambda^{\gamma}), F(z^{\gamma})\rangle \\
&+ \quad \sum_{j=3}^{k}(1 - \lambda_j\bar{z}_j)\langle F(\lambda^{\gamma^{j-1}}), F(z^{\gamma^{j-1}})\rangle \\
= (1 - \lambda_2\bar{z}_1)\langle F(\lambda^{(12)}), F(z)\rangle \quad &+ \quad (1 - \lambda_1\bar{z}_2)\langle F(\lambda^{(12)\gamma}), F(z^{\gamma})\rangle \\
&+ \quad \sum_{j=3}^{k}(1 - \lambda_j\bar{z}_j)\langle F(\lambda^{\gamma^{j-1}}), F(z^{\gamma^{j-1}})\rangle.
\end{aligned}$$

Cancel common terms and apply (2.5) once more to see

$$\begin{aligned} & (1-\lambda_1\bar{z}_1)\langle F(\lambda), F(z)\rangle + (1-\lambda_2\bar{z}_2)\langle F(\lambda^\gamma), F(z^\gamma)\rangle \\ = \;& (1-\lambda_2\bar{z}_1)\langle F(\lambda^{(12)}), F(z)\rangle + (1-\lambda_1\bar{z}_2)\langle F(\lambda^{(12)\gamma}), F(z^\gamma)\rangle \\ = \;& (1-\lambda_2\bar{z}_1)\langle F(\lambda^\gamma), F(z)\rangle + (1-\lambda_1\bar{z}_2)\langle F(\lambda), F(z^\gamma)\rangle. \end{aligned}$$

Equate the first and last of these expressions and re-factorize as follows:

$$\langle F(\lambda) - F(\lambda^\gamma), F(z) - F(z^\gamma)\rangle = \langle \lambda_1 F(\lambda) - \lambda_2 F(\lambda^\gamma), z_1 F(z) - z_2 F(z^\gamma)\rangle.$$

Consequently, since $\dim \mathcal{C} = 1$, there exists an $\alpha \in \mathbb{T}$ such that

$$\bar{\alpha}(F(\lambda) - F(\lambda^\gamma)) = \lambda_1 F(\lambda) - \lambda_2 F(\lambda^\gamma)$$

for all $\lambda \in \mathbb{D}^k$. Equivalently,

$$(1-\alpha\lambda_1)F(\lambda) = (1-\alpha\lambda_2)F(\lambda^\gamma)$$

for all $\lambda \in \mathbb{D}^k$. Replacing λ by $\lambda^{\gamma^{t-1}}$ gives, for $t = 1, \ldots, k$,

$$(1-\alpha\lambda_t)F(\lambda^{\gamma^{t-1}}) = (1-\alpha\lambda_{t+1})F(\lambda^{\gamma^t})$$

for all $\lambda \in \mathbb{D}^k$. That is (2.6) holds. This completes the proof of the first part of the result. Let $\mathcal{S}$ be as defined in the statement of the result. Pick any $\sigma \in S_k$ and let $t = \sigma(1)$. Then, by virtue of (2.5) and (2.6),

$$\begin{aligned} \mathcal{S}(\lambda^\sigma) &= (1-\alpha\lambda_t)F(\lambda^\sigma) \\ &= (1-\alpha\lambda_t)F(\lambda^{\gamma^{t-1}}) \\ &= (1-\alpha\lambda_{t-1})F(\lambda^{\gamma^{t-2}}) \\ &\;\vdots \\ &= (1-\alpha\lambda_1)F(\lambda) \\ &= \mathcal{S}(\lambda). \end{aligned}$$

Hence $\mathcal{S}$ has the required property.

■

Lemma 2.4.6 *Let $\mathcal{C}$ be one dimensional. For α and $\mathcal{S}$ as in Lemma* 2.4.5*, the hereditary polynomial g defined in* (2.3) *can be expressed in the form*

$$g(\lambda, \bar{z}) = \overline{\psi(z)} p(\lambda, z)\psi(\lambda) \tag{2.7}$$

for all $\lambda, z \in \mathbb{D}^k$ *where*

$$p(\lambda, \bar{z}) = \sum_{j=1}^{k} \left((1 - \lambda_j \bar{z}_j) \prod_{i \neq j} (1 - \bar{\alpha} \bar{z}_i)(1 - \alpha \lambda_i) \right) \tag{2.8}$$

and

$$\psi(\lambda) = \mathcal{S}(\lambda) \prod_{i=1}^{k} (1 - \alpha \lambda_i)^{-1}.$$

Proof. Lemma 2.4.5 allows us to re-write $F(\lambda)$ in terms of the symmetric function $\mathcal{S}$,

$$F(\lambda) = \frac{\mathcal{S}(\lambda)}{1 - \alpha \lambda_1}.$$

Substituting this into (2.4) yields

$$\begin{aligned} g(\lambda, \bar{z}) &= \sum_{j=1}^{k} (1 - \lambda_j \bar{z}_j) \frac{\mathcal{S}(\lambda) \overline{\mathcal{S}(z)}}{(1 - \alpha \lambda_j)(1 - \bar{\alpha} \bar{z}_j)} \\ &= \overline{\mathcal{S}(z)} \left[\sum_{j=1}^{k} \frac{1 - \lambda_j \bar{z}_j}{(1 - \bar{\alpha} \bar{z}_j)(1 - \alpha \lambda_j)} \right] \mathcal{S}(\lambda) \\ &= \overline{\psi(z)} \left[\sum_{j=1}^{k} \left((1 - \lambda_j \bar{z}_j) \prod_{i \neq j} (1 - \bar{\alpha} \bar{z}_i)(1 - \alpha \lambda_i) \right) \right] \psi(\lambda) \\ &= \overline{\psi(z)} p(\lambda, \bar{z}) \psi(\lambda). \end{aligned}$$

■

We have now reached the end of a chain of arguments which will give rise to a necessary condition for the existence of a solution to the Main Problem described in the introduction. We have shown that the existence of an interpolating function which satisfies the constraints of the main problem implies the existence of a solution to the Γ_k problem. We then went on to show that the existence of such a function implies that Γ_k is a complete spectral set for the commuting k-tuple of operators $C_1, \ldots, C_k$. In particular, if h is a polynomial in k indeterminates which is bounded by 1 on Γ_k then $h(C_1, \ldots, C_k)$ is a contraction. We then considered those polynomials, h, which give rise to a contraction for *all* commuting k-tuples of contractions and defined the functions $g = 1 - \overline{h \circ \pi(z)} h \circ \pi(\lambda)$. These functions are positive on contractions. The results of this section show that each g which is "atomic" in the sense that the corresponding Hilbert space $\mathcal{C}$ is one dimensional, has a representation of the form (2.7). Finally, if g is of the form (2.7) and is positive on Γ_k then the polynomial p defined in (2.8) must also be positive on Γ_k.

In Chapter 3 we shall consider a more general class of polynomials. In Chapter 5 we will show that these more general polynomials give rise to a necessary condition for the existence of a solution to the interpolation problems described in the introduction.

Chapter 3

Elementary Symmetric Polynomials

The aim of this chapter is to introduce a class of polynomials motivated by those at the end of Chapter 2. We show that polynomials in this class have a number of possible representations. The results of this chapter are rather technical but they are essential for the work which follows. Both of the representations which are proved in this chapter are used in the proof of the main result of Chapter 5, Theorem 5.1.5. Many of the proofs in Chapters 6 and 7 are also simplified by the results in this chapter.

3.1 Definitions and Preliminaries

In this section we shall generalise the polynomial p, introduced at the end of Chapter 2 (see (2.8)), to define a wider class of doubly symmetric polynomials. We also introduce a differential operator which will be used to help simplify the forms of various polynomial representations.

Definition 13 *For $k \in \mathbb{N}$ and $\alpha \in \mathbb{C}$ we define the polynomial $p_{k,\alpha}$ in $2k$ variables by*

$$p_{k,\alpha}(\lambda, \bar{z}) = \sum_{j=1}^{k} \left((1 - |\alpha|^2 \lambda_j \bar{z}_j) \prod_{i \neq j} (1 - \bar{\alpha} \bar{z}_i)(1 - \alpha \lambda_i) \right). \tag{3.1}$$

When $|\alpha| = 1$ we see that $p_{k,\alpha}$ coincides with the polynomial p in (2.8) .

Definition 14 *For r and s satisfying $1 \leq r, s \leq k$ define the partial differential operator $D_{r,s}$ as*

$$D_{r,s} = \frac{\partial^{r+s}}{\partial \lambda_1 \cdots \partial \lambda_r \partial \bar{z}_1 \cdots \partial \bar{z}_s}. \tag{3.2}$$

The operators $D_{0,s}$ and $D_{r,0}$ are defined as the corresponding differential operators where differentiation is carried out only with respect to the components of either $\bar{z}$ or λ.

By a multi-index m we mean a k-tuple of non-negative integers $(m_1, \ldots, m_k)$ and for such a multi-index, we define

$$\lambda^m = {\lambda_1}^{m_1} \ldots {\lambda_k}^{m_k}$$

for all $\lambda \in \mathbb{C}^k$. We shall refer to a multi-index m as binary if the components of m only take the values zero and one. We define the following sets:

$$\begin{aligned}\mathcal{G} &= \{(m_1, \ldots, m_k) \in \mathbb{N}^k \mid i > j \Rightarrow m_i \leq m_j\} \\ \mathcal{B} &= \{(m_1, \ldots, m_k) \in \mathbb{N}^k \mid m_i \in \{0, 1\},\ i > j \Rightarrow m_i \leq m_j\} \\ &= \{(0, \ldots, 0), (1, 0, \ldots, 0), \ldots, (1, \ldots, 1)\}.\end{aligned}$$

Let $\theta : \{0, 1, \ldots, k\} \to \mathcal{B}$ be the bijection which maps r to the element of $\mathcal{B}$ whose first r terms equal one, and all the rest equal to zero.

Definition 15 *Given a polynomial p in n indeterminates,*

$$p(x_1, \ldots, x_n) = \sum_{i=1}^{N} c_i x_1^{r_{i1}} \cdots x_n^{r_{in}},$$

define the leading power of p to be

$$\max_{\substack{i,j \\ c_i \neq 0}} \{r_{ij}\}.$$

Recall that $c_r(\lambda)$ represents the r^{th} elementary symmetric polynomial in the k co-ordinates of λ. That is, $c_r(\lambda)$ is the sum of all monomials which can be formed by multiplying r distinct co-ordinates of λ together.

Lemma 2.1.1 can be extended to doubly symmetric polynomials in the following way.

Lemma 3.1.1 *Every doubly symmetric polynomial in the indeterminates λ and $\bar{z}$ can be expressed as a polynomial in the elementary symmetric polynomials of the components of λ and $\bar{z}$. That is, if $q(\lambda, \bar{z})$ is doubly symmetric then there exists a polynomial p such that*

$$q(\lambda, \bar{z}) = p(c_1(\lambda), \ldots, c_k(\lambda), c_1(\bar{z}), \ldots, c_k(\bar{z})) \tag{3.3}$$

Proof. Let q be a doubly symmetric polynomial in the indeterminates λ and $\bar{z}$. Then

$$\begin{aligned}q(\lambda, \bar{z}) &= \sum_{m,n \in \mathbb{N}^k} c_{mn} \lambda^n \bar{z}^m. \\ &= \sum_{m \in \mathbb{N}^k} \bar{z}^m \left(\sum_{n \in \mathbb{N}^k} c_{mn} \lambda^n \right).\end{aligned}$$

Since q is invariant under any permutation of $\bar{z}$ it follows that the coefficients of $\bar{z}^m$ and $\bar{z}^{m^\sigma}$ are equal for all $\sigma \in S_k$. Note that every $m \in \mathbb{N}^k$ is of the form w^σ for some $\sigma \in S_k$ and some $w \in \mathcal{G}$. Thus, terms may be grouped as follows

$$\begin{aligned} q(\lambda, \bar{z}) &= \sum_{m \in \mathcal{G}} \left(\sum_{\sigma \in S_k} \bar{z}^{m^\sigma} \right) \left(\sum_{n \in \mathbb{N}^k} c'_{mn} \lambda^n \right) \\ &= \sum_{n \in \mathbb{N}^k} \sum_{m \in \mathcal{G}} c'_{mn} \left(\sum_{\sigma \in S_k} \bar{z}^{m^\sigma} \right) \lambda^n. \end{aligned}$$

The coefficient c'_{mn} may differ from the coefficient c_{mn} since some terms are invariant under the action of S_k. For example $\lambda^{(1,1,1)} = \lambda^{(1,1,1)^\sigma}$ for all $\sigma \in S_k$. The polynomial q is symmetric in λ so using the same argument as above we have

$$q(\lambda, \bar{z}) = \sum_{m,n \in \mathcal{G}} c''_{mn} \left(\sum_{\sigma \in S_k} \bar{z}^{m^\sigma} \right) \left(\sum_{\sigma \in S_k} \lambda^{n^\sigma} \right).$$

Applying Lemma 2.1.1 to the symmetric polynomials on the right hand side of this expression we infer that there exist polynomials p_m and q_n such that

$$q(\lambda, \bar{z}) = \sum_{m,n \in \mathcal{G}} c''_{mn} p_m(c_1(\bar{z}), \ldots, c_k(\bar{z})) q_n(c_1(\lambda), \ldots, c_k(\lambda)).$$

Let $\Lambda = (c_1(\lambda), \ldots, c_k(\lambda))$ and $Z = (c_1(\bar{z}), \ldots c_k(\bar{z}))$. Then the required polynomial p can be taken to be

$$p(\Lambda, Z) = \sum_{m,n \in \mathcal{G}} c''_{mn} p_m(Z) q_n(\Lambda).$$

■

The polynomial defined in (3.1) is doubly symmetric in λ and $\bar{z}$ so we may write it as a polynomial in the elementary symmetric polynomials of λ and $\bar{z}$. However, $p_{k,\alpha}$ is such that it may be expressed in two simpler forms. A number of results are required to prove this fact.

3.2 Doubly Symmetric Polynomials

We shall say that an hereditary polynomial h *contains* $\bar{z}^m \lambda^n$ if

$$h(\lambda, \bar{z}) = \sum_{i,j \in \mathbb{N}^k} c_{ji} \bar{z}^j \lambda^i$$

and $c_{mn} \neq 0$.

Lemma 3.2.1 *Fix $k \in \mathbb{N}$. Let m and n be elements of $\mathcal{B}$ and let λ and $\bar{z}$ be k-tuples of indeterminates. Then a (not unique) doubly symmetric polynomial of elementary symmetric polynomials with fewest terms which contains $\lambda^n \bar{z}^m$ is $c_{\theta^{-1}(n)}(\lambda) c_{\theta^{-1}(m)}(\bar{z})$.*

Proof. Let $\theta^{-1}(n) = r$ and $\theta^{-1}(m) = s$. Then $\lambda^n \bar{z}^m$ is the product of r coefficients of λ with s coefficients of $\bar{z}$. The elementary symmetric polynomial $c_r(\lambda)$ has $\binom{k}{r}$ terms, so the product $c_r(\lambda)c_s(\bar{z})$ has $\binom{k}{r}\binom{k}{s}$ terms. Clearly, this polynomial contains $\lambda^n \bar{z}^m$ since it contains all products of r coefficients of λ with s coefficients of $\bar{z}$.

Now, every doubly symmetric polynomial which contains $\lambda^n \bar{z}^m$ also contains every term of the form $\lambda^{n^\sigma} \bar{z}^{m^\tau}$ where $\sigma, \tau \in S_k$. In other words it must contain *every* product of r coefficients of λ with s coefficients of $\bar{z}$. Since there are $\binom{k}{r}$ ways of choosing r coefficients of λ and $\binom{k}{s}$ ways of choosing s coefficients of $\bar{z}$, it follows that there are $\binom{k}{r}\binom{k}{s}$ such products. Thus any polynomial which is doubly symmetric and contains the term $\lambda^n \bar{z}^m$ must have at least $\binom{k}{r}\binom{k}{s}$ terms. Since $c_r(\lambda)c_s(\bar{z})$ has this many terms, it is clearly a doubly symmetric polynomial containing $\lambda^n \bar{z}^m$ with the fewest possible terms.

∎

Lemma 3.2.2 *Let q be a doubly symmetric polynomial with leading power at most 1. Then q can be written in the following form:*

$$q(\lambda, \bar{z}) = \sum_{r,s=0}^{k} b_{rs} c_r(\lambda) c_s(\bar{z}). \tag{3.4}$$

Furthermore the coefficients b_{rs} can be evaluated as follows:

$$b_{rs} = D_{r,s} q(\lambda, \bar{z})|_{\lambda = \bar{z} = 0}. \tag{3.5}$$

Proof. If the leading power of q is zero, then the result is trivial. More generally the polynomial q is of the form

$$q(\lambda, \bar{z}) = \sum_{n,m \in \mathbb{N}^k} a_{nm} \lambda^n \bar{z}^m$$

for all $\lambda, \bar{z} \in \mathbb{C}^k$. Pick any two multi-indices m and n in $\mathbb{N}^k$ and consider the coefficient of the monomial $\lambda^n \bar{z}^m$ in $q(\lambda, \bar{z})$. Since $q(\lambda, \bar{z})$ is doubly symmetric, the coefficient of $\lambda^n \bar{z}^m$ is equal to that of $\lambda^{n^\sigma} \bar{z}^{m^\tau}$ for every σ and τ in S_k. That is $a_{nm} = a_{n^\sigma m^\tau}$ for all n and m in $\mathbb{N}^k$ and all σ and τ in S_k. It is obvious that every m in $\mathbb{N}^k$ is the image of an element of $\mathcal{G}$ under an element of S_k. Now rewrite the formula for $q(\lambda, \bar{z})$ grouping terms with equal coefficients together, and possibly altering the coeffiecient a_{nm} to a'_{nm} to take into account terms whose multiindices are invariant under certain elements of S_k:

$$q(\lambda, \bar{z}) = \sum_{n,m \in \mathcal{G}} a'_{nm} \left(\sum_{\sigma, \tau \in S_k} \lambda^{n^\sigma} \bar{z}^{m^\tau} \right).$$

Clearly, the polynomial

$$\sum_{\sigma,\tau\in S_k} \lambda^{n^\sigma} \bar{z}^{m^\tau}$$

is doubly symmetric— indeed it is a doubly symmetric polynomial with the fewest possible terms which contains the term $\lambda^n \bar{z}^m$. Thus we may apply Lemma 3.1.1 and write it in the form given in (3.3). We have

$$q(\lambda, \bar{z}) = \sum_{n,m\in\mathcal{G}} a'_{nm} p_{nm}(c_1(\lambda), \ldots, c_k(\lambda), c_1(\bar{z}), \ldots c_k(\bar{z})).$$

Thus, the coefficient of any monomial $\lambda^n \bar{z}^m$ in $q(\lambda, \bar{z})$ is equal to that of the polynomial p_{nm}— a polynomial of the elementary symmetric polynomials in λ and $\bar{z}$ with the fewest terms which contains $\lambda^n \bar{z}^m$.

Now, since q has leading power no greater than one, it follows that the coefficient of $\lambda^n \bar{z}^m$ is zero unless both n and m are binary. Therefore,

$$q(\lambda, \bar{z}) = \sum_{n,m\in\mathcal{B}} a'_{nm} p_{nm}(c_1(\lambda), \ldots, c_k(\lambda), c_1(\bar{z}), \ldots c_k(\bar{z})).$$

By virtue of Lemma 3.2.1, whenever $n, m \in \mathcal{B}$ we have

$$p_{nm} = c_{\theta^{-1}(n)}(\lambda) c_{\theta^{-1}(m)}(\bar{z})$$

so we may make the substitutions $n = \theta(r)$ and $m = \theta(s)$ which give

$$q(\lambda, \bar{z}) = \sum_{r,s=0}^{k} a'_{\theta(r)\theta(s)} c_r(\lambda) c_s(\bar{z}).$$

Finally we may relabel the constants $a_{\theta(r)\theta(s)}$ as b_{rs} to see that the first part of the result holds. Furthermore, we know that the coefficient of $\lambda_1 \cdots \lambda_r \bar{z}_1 \cdots \bar{z}_s$ in $q(\lambda, z)$ is equal to the coefficient of $c_r(\lambda) c_s(\bar{z})$ since this is a doubly symmetric polynomial with the fewest possible terms which contains the given term. That is, b_{rs} is equal to the coefficient of $\lambda_1 \cdots \lambda_r \bar{z}_1 \cdots \bar{z}_s$ in $q(\lambda, z)$. Clearly, this is given by

$$D_{r,s} q(\lambda, \bar{z})|_{\lambda=\bar{z}=0}$$

and hence the result holds.

■

The simplification of the polynomial $p_{k,\alpha}$ defined in (3.1) requires another lemma.

Lemma 3.2.3 *Let r and s be integers such that $0 \leq r, s \leq k$. Then*

$$D_{r,s}\prod_{i=1}^{k}(1-\alpha\lambda_i)\prod_{l=1}^{k}(1-\bar{\alpha}\bar{z}_l) = (-1)^{r+s}\alpha^r\bar{\alpha}^s\prod_{i=r+1}^{k}(1-\alpha\lambda_i)\prod_{l=s+1}^{k}(1-\bar{\alpha}\bar{z}_l).$$

Proof. Pick r and s such that $0 \leq r, s \leq k$, then

$$\begin{aligned}
&D_{r,s}\prod_{i=1}^{k}(1-\alpha\lambda_i)\prod_{l=1}^{k}(1-\bar{\alpha}\bar{z}_l) \\
&\quad=\frac{\partial^{r+s}}{\partial\lambda_1\cdots\partial\lambda_r\partial\bar{z}_1\cdots\partial\bar{z}_s}\prod_{i=1}^{k}(1-\alpha\lambda_i)\prod_{l=1}^{k}(1-\bar{\alpha}\bar{z}_l) \\
&\quad=\prod_{i=1}^{r}\frac{\partial}{\partial\lambda_i}(1-\alpha\lambda_i)\prod_{l=1}^{s}\frac{\partial}{\partial\bar{z}_l}(1-\bar{\alpha}\bar{z}_i)\prod_{i=r+1}^{k}(1-\alpha\lambda_i)\prod_{l=s+1}^{k}(1-\bar{\alpha}\bar{z}_l) \\
&\quad=\prod_{i=1}^{r}(-\alpha)\prod_{l=1}^{s}(-\bar{\alpha})\prod_{i=r+1}^{k}(1-\alpha\lambda_i)\prod_{l=s+1}^{k}(1-\bar{\alpha}\bar{z}_l) \\
&\quad=(-\alpha)^r(-\bar{\alpha})^s\prod_{i=r+1}^{k}(1-\alpha\lambda_i)\prod_{l=s+1}^{k}(1-\bar{\alpha}\bar{z}_l) \\
&\quad=(-1)^{r+s}\alpha^r\bar{\alpha}^s\prod_{i=r+1}^{k}(1-\alpha\lambda_i)\prod_{l=s+1}^{k}(1-\bar{\alpha}\bar{z}_l).
\end{aligned}$$

■

3.3 Two Representations of the Polynomial $p_{k,\alpha}$

In this Section we represent the polynomial $p_{k,\alpha}$ defined in (3.1) in two simple forms. The first representation depends on the results of the previous Section and a heavy dose of basic differentiation. The second representation of $p_{k,\alpha}$ follows from Lemma 1.3.1 and the similarity of the first representation to the polynomial P_k, which was defined in (1.6).

Theorem 3.3.1 *Let $p_{k,\alpha}$ be defined as in* (3.1) *Then*

$$p_{k,\alpha}(\lambda,\bar{z}) = \sum_{r,s=1}^{k}(-1)^{r+s}\alpha^r\bar{\alpha}^s(k-(r+s))c_r(\lambda)c_s(\bar{z}).$$

Proof. Since the leading power of $p_{k,\alpha}(\lambda,\bar{z})$ is one, we may apply Lemma 3.2.2 and write $p_{k,\alpha}$ as

$$p_{k,\alpha}(\lambda, z) = \sum_{r,s=0}^{k} b_{rs}c_r(\lambda)c_s(\bar{z}).$$

The same result shows that it will suffice to prove

$$D_{r,s}p_{k,\alpha}(\lambda,\bar{z})|_{\lambda=\bar{z}=0} = (-1)^{r+s}\alpha^r\bar{\alpha}^s(k-(r+s))$$

for all $r,s \in \{0,\ldots,k\}$. Without loss of generality we may assume that $r < s$.

Consider $D_{r,s}p_{k,\alpha}(\lambda,\bar{z})$,

$$\begin{aligned}
D_{r,s}p_{k,\alpha}(\lambda,\bar{z}) &= D_{r,s}\sum_{j=1}^{k}\left((1-|\alpha|^2\lambda_j\bar{z}_j)\prod_{\substack{i=1\\i\neq j}}^{k}(1-\bar{\alpha}\bar{z}_i)(1-\alpha\lambda_i)\right)\\
&= D_{r,s}\sum_{j=1}^{r}\left((1-|\alpha|^2\lambda_j\bar{z}_j)\prod_{\substack{i=1\\i\neq j}}^{k}(1-\bar{\alpha}\bar{z}_i)(1-\alpha\lambda_i)\right)\\
&+ D_{r,s}\sum_{j=r+1}^{s}\left((1-|\alpha|^2\lambda_j\bar{z}_j)\prod_{\substack{i=1\\i\neq j}}^{k}(1-\bar{\alpha}\bar{z}_i)(1-\alpha\lambda_i)\right)\\
&+ D_{r,s}\sum_{j=s+1}^{k}\left((1-|\alpha|^2\lambda_j\bar{z}_j)\prod_{\substack{i=1\\i\neq j}}^{k}(1-\bar{\alpha}\bar{z}_i)(1-\alpha\lambda_i)\right).
\end{aligned}$$

We now calculate the second component of this expression,

$$\begin{aligned}
&D_{r,s}\sum_{j=r+1}^{s}\left((1-|\alpha|^2\lambda_j\bar{z}_j)\prod_{\substack{i=1\\i\neq j}}^{k}(1-\bar{\alpha}\bar{z}_i)(1-\alpha\lambda_i)\right)\\
&= \sum_{j=r+1}^{s}\left(\frac{\partial}{\partial\bar{z}_j}(1-|\alpha|^2\lambda_j\bar{z}_j)\frac{\partial^{r+s-1}}{\partial\lambda_1\cdots\partial\lambda_r\partial\bar{z}_1\cdots\partial\bar{z}_{j-1}\partial\bar{z}_{j+1}\cdots\partial\bar{z}_s}\prod_{\substack{i=1\\i\neq j}}^{k}(1-\bar{\alpha}\bar{z}_i)(1-\alpha\lambda_i)\right)
\end{aligned}$$

By virtue of Lemma 3.2.3,

$$\begin{aligned}
D_{r,s}&\sum_{j=r+1}^{s}\left((1-|\alpha|^2\lambda_j\bar{z}_j)\prod_{\substack{i=1\\i\neq j}}^{k}(1-\bar{\alpha}\bar{z}_i)(1-\alpha\lambda_i)\right)\\
&= \sum_{j=r+1}^{s}\left(-\alpha\lambda_j\bar{\alpha}(-1)^{r+s-1}\alpha^r\bar{\alpha}^{s-1}\prod_{\substack{i=r+1\\i\neq j}}^{k}(1-\alpha\lambda_i)\prod_{l=s+1}^{k}(1-\bar{\alpha}\bar{z}_l)\right)\\
&= \sum_{j=r+1}^{s}\left(\lambda_j(-1)^{r+s}\alpha^{r+1}\bar{\alpha}^{s}\prod_{\substack{i=r+1\\i\neq j}}^{k}(1-\alpha\lambda_i)\prod_{l=s+1}^{k}(1-\bar{\alpha}\bar{z}_l)\right).
\end{aligned}$$

Using an identical method on each of the remaining components of the sum, we have

$$\begin{aligned}D_{r,s}p_{k,\alpha}(\lambda,\bar{z}) = \sum_{j=1}^{r}&\left(-(-1)^{r+s}\alpha^r\bar{\alpha}^s\prod_{i=r+1}^{k}(1-\alpha\lambda_i)\prod_{l=s+1}^{k}(1-\bar{\alpha}\bar{z}_l)\right)\\ &+\sum_{j=r+1}^{s}\left(\lambda_j(-1)^{r+s}\alpha^{r+1}\bar{\alpha}^s\prod_{\substack{i=r+1\\ i\neq j}}^{k}(1-\alpha\lambda_i)\prod_{l=s+1}^{k}(1-\bar{\alpha}\bar{z}_l)\right)\\ &+\sum_{j=s+1}^{k}\left((1-|\alpha|^2\lambda_j\bar{z}_j)(-1)^{r+s}\alpha^r\bar{\alpha}^s\prod_{\substack{i=r+1\\ i\neq j}}^{k}(1-\alpha\lambda_i)\prod_{\substack{l=s+1\\ l\neq j}}^{k}(1-\bar{\alpha}\bar{z}_l)\right).\end{aligned}$$

Therefore,

$$\begin{aligned}D_{r,s}p_{k,\alpha}(\lambda,\bar{z})|_{\lambda=\bar{z}=0} &= -r(-1)^{r+s}\alpha^r\bar{\alpha}^s+0+(k-s)(-1)^{r+s}\alpha^r\bar{\alpha}^s\\ &= (-1)^{r+s}(k-(r+s))\alpha^r\bar{\alpha}^s.\end{aligned}$$

Hence the result holds.

■

The simplification of $p_{k,\alpha}$ may be carried a little further by virtue of the observation

$$\alpha^r c_r(\lambda) = c_r(\alpha\lambda).$$

Corollary 3.3.2 *Let $p_{k,\alpha}$ be defined as in* (3.1). Then

$$p_{k,\alpha}(\lambda,\bar{z}) = \sum_{r,s=0}^{k}(-1)^{r+s}(k-(r+s))c_r(\alpha\lambda)c_s(\bar{\alpha}\bar{z}).$$

This completes our initial aim of finding an alternative representation of $p_{k,\alpha}$. Notice, in terms of the polynomial P_k defined in (1.6) we have shown,

$$p_{k,\alpha}(\lambda,\bar{z}) = P_k(c_0(\alpha\lambda),\ldots,c_k(\alpha\lambda);c_0(\bar{\alpha}\bar{z}),\ldots,c_k(\bar{\alpha}\bar{z})). \tag{3.6}$$

We now establish a second representation for $p_{k,\alpha}$ using Lemma 1.3.1.

Theorem 3.3.3 *The polynomial $p_{k,\alpha}(\lambda,\bar{z})$ can be expressed as*

$$p_{k,\alpha}(\lambda,\bar{z}) = \frac{1}{k}\left[\overline{A_{k,\alpha}(z)}A_{k,\alpha}(\lambda)-\overline{B_{k,\alpha}(z)}B_{k,\alpha}(\lambda)\right]$$

for all $\lambda,\bar{z}\in\mathbb{C}^k$ where

$$A_{k,\alpha}(\lambda) = \sum_{r=0}^{k}(-1)^r(k-r)c_r(\alpha\lambda)$$

and

$$B_{k,\alpha}(\lambda) = \sum_{r=0}^{k} (-1)^r r c_r(\alpha\lambda).$$

Proof. By Corollary 3.3.2, equation (3.6) and Lemma 1.3.1 we have,

$$\begin{aligned}
p_{k,\alpha}(\lambda, \bar{z}) &= \sum_{r,s=0}^{k} (-1)^{r+s}(k-(r+s))c_r(\alpha\lambda)c_s(\overline{\alpha z}) \\
&= P_k(c_0(\alpha\lambda), \ldots, c_k(\alpha\lambda); c_0(\bar{\alpha}\bar{z}), \ldots, c_k(\bar{\alpha}\bar{z})) \\
&= \frac{1}{k}\left[A_k(c_0(\bar{\alpha}\bar{z}), \ldots, c_k(\bar{\alpha}\bar{z}))A_k(c_0(\alpha\lambda), \ldots, c_k(\alpha\lambda))\right] \\
&\quad - B_k(c_0(\bar{\alpha}\bar{z}), \ldots, c_k(\bar{\alpha}\bar{z}))B_k(c_0(\alpha\lambda), \ldots, c_k(\alpha\lambda)) \\
&= \frac{1}{k}\left[\overline{A_{k,\alpha}(z)}A_{k,\alpha}(\lambda) - \overline{B_{k,\alpha}(z)}B_{k,\alpha}(\lambda)\right]
\end{aligned}$$

■

The polynomials $A_{k,\alpha}$ and $B_{k,\alpha}$ play a vital role in Chapters 5 and 6. In Chapter 4 we study the properties of $A_{k,\alpha}$.

Chapter 4

$\mathbf{A}_{\mathbf{k},\alpha}$ has no Zeros in the Polydisc

In Theorem 3.3.3 we introduced two one-parameter pencils of polynomials of degree k in k variables. These polynomials were denoted by $A_{k,\alpha}$ and $B_{k,\alpha}$, where the parameter α ranges over $\mathbb{D}$. We are particularly interested in the behaviour of these pencils on the k-dimensional polydisc. This chapter contains the proof of the most important of their characteristics, namely that $A_{k,\alpha}$ has no zeros in the polydisc. Formally, we show that for $k \in \mathbb{N}, \alpha \in \mathbb{D}$ we have $A_{k,\alpha}(\lambda) \neq 0$ for all $\lambda \in \mathbb{D}^k$.

The cases $k = 1$ and $k = 2$ are trivial (see Theorem 4.1.1) and although the case $k = 3$ is more complicated (Theorem 4.1.2), it fails to contain all the germs of generality. One must wait until $k = 4$ (Theorem 4.1.6) before the whole picture unfolds and for this reason the proofs of these special cases are presented as a prelude to the full result (Theorem 4.2.2).

The difficult calculations of this chapter are essential to the proofs of the main results in Chapters 5, 6 and 7. A key step in the proofs of the main results of Chapters 5 and 6 will be to show that if a certain hereditary polynomial is applied to a specific k-tuple of commuting operators and the resulting operator is positive semi-definite then the same hereditary polynomial applied to certain compressions of the k-tuple of operators will also yield a positive semi-definite operator. In general this is not true and it only holds because the operators and hereditary polynomial in question are of a special form. The results proved here will enable us to show that the polynomial under investigation is indeed of that special form. We do this by showing that $A_{k,\alpha}$ has no zeros in the polydisc; this will allow us to show that $A_{k,\alpha}(T)$ is invertible for a certain k-tuple of commuting operators T.

4.1 Special Cases

Theorem 4.1.1 *Let $\alpha \in \mathbb{D}$. Then*

$$A_{1,\alpha}(z) \neq 0 \qquad \text{and} \qquad A_{2,\alpha}(\lambda) \neq 0$$

for all $z \in \mathbb{D}$ and all $\lambda \in \mathbb{D}^2$.

Proof. The polynomial $A_{1,\alpha}$ is identically 1. For $k = 2$ pick $\lambda = (\lambda_1, \lambda_2) \in \mathbb{D}^2$. We have

$$\begin{aligned} A_{2,\alpha}(\lambda) &= \sum_{r=0}^{2} (-1)^r (2-r) c_r(\alpha\lambda) \\ &= 2 - c_1(\alpha\lambda) \\ &= 2 - \alpha(\lambda_1 + \lambda_2) \\ &\neq 0 \end{aligned}$$

since $|\alpha(\lambda_1 + \lambda_2)| \leq |\alpha\lambda_1| + |\alpha\lambda_2| < 2$.

∎

The next result demonstrates the increasing complexity as k increases and provides the first indications of the method for a general proof.

Theorem 4.1.2 *Let $\alpha \in \mathbb{D}$. Then*

$$A_{3,\alpha}(z) \neq 0$$

for all $z \in \mathbb{D}^3$.

Proof. We shall argue by contradiction. Suppose there exists $\alpha \in \mathbb{D}$ and $z = (z_1, z_2, z_3) \in \mathbb{D}^3$ such that $A_{3,\alpha}(z) = 0$. Let $\lambda = (\lambda_1, \lambda_2, \lambda_3) = (\alpha z_1, \alpha z_2, \alpha z_3) \in \mathbb{D}^3$. Then

$$\begin{aligned} 0 &= 3 - 2(\lambda_1 + \lambda_2 + \lambda_3) + (\lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1) \\ &= 3 - 2(\lambda_1 + \lambda_2) + \lambda_1\lambda_2 - \lambda_3(2 - (\lambda_1 + \lambda_2)). \end{aligned}$$

Therefore

$$\lambda_3 = \frac{3 - 2(\lambda_1 + \lambda_2) + \lambda_1\lambda_2}{2 - (\lambda_1 + \lambda_2)}. \tag{4.1}$$

Define $F_3 : \mathbb{D} \to \mathbb{C}$ by

$$F_3(z) = \frac{3 - 2\lambda_1 - z(2 - \lambda_1)}{2 - \lambda_1 - z}.$$

Theorem 4.1.1 shows that F_3 is analytic on $\mathbb{D}$. By (4.1) we have $\lambda_3 = F_3(\lambda_2)$. We shall derive a contradiction by showing $F_3(\mathbb{D}) \cap \mathbb{D} = \emptyset$. Clearly F_3 is a linear fractional transformation and as such will map the disc to some other disc in $\mathbb{C}$. Call the centre of this new disc c. The pre-image of a point γ under F_3 will be denoted $F_3^{-1}(\gamma)$. By inspection we have

$$F_3^{-1}(\infty) = 2 - \lambda_1.$$

Since conjugacy is preserved by linear fractional transformations we have

$$F_3^{-1}(c) = \frac{1}{2 - \overline{\lambda_1}}.$$

Therefore c, the centre of the image of $\mathbb{D}$ under F_3 equals

$$\begin{aligned}
F_3\left(\frac{1}{2-\overline{\lambda_1}}\right) &= \frac{3 - 2\lambda_1 - \left(\frac{1}{2-\overline{\lambda_1}}\right)(2-\lambda_1)}{2 - \lambda_1 - \frac{1}{2-\overline{\lambda_1}}} \\
&= \frac{(3-2\lambda_1)(2-\overline{\lambda_1}) - (2-\lambda_1)}{|2-\lambda_1|^2 - 1} \\
&= \frac{[(1-\lambda_1)+(2-\lambda_1)](2-\overline{\lambda_1}) - (2-\lambda_1)}{|2-\lambda_1|^2 - 1} \\
&= \frac{|2-\lambda_1|^2 + (1-\lambda_1)(2-\overline{\lambda_1}) - 1 - (1-\lambda_1)}{|2-\lambda_1|^2 - 1} \\
&= \frac{|2-\lambda_1|^2 + (1-\lambda_1)(2-\overline{\lambda_1} - 1) - 1}{|2-\lambda_1|^2 - 1} \\
&= \frac{|2-\lambda_1|^2 + |1-\lambda_1|^2 - 1}{|2-\lambda_1|^2 - 1} \\
&= 1 + \frac{|1-\lambda_1|^2}{|2-\lambda_1|^2 - 1}.
\end{aligned}$$

We may therefore conclude that $F_3(\mathbb{D})$ is a circle centred at a $c \in \mathbb{R}$ such that $c > 1$. Notice also that $F_3(1) = 1$ so the point 1 lies on the boundary of $F_3(\mathbb{D})$. It follows that $F_3(\mathbb{D}) \cap \mathbb{D} = \emptyset$ which contradicts (4.1).

■

For a proof of the next special case, and indeed the general result, it is convenient to extend the definition of elementary symmetric polynomials given in Chapter 1.

Definition 16 *Define σ_r^n, the r^{th} partial elementary symmetric polynomial on k indeterminates by*

$$\sigma_r^n(\lambda_1, \ldots, \lambda_k) = \begin{cases} c_r(\lambda_1, \ldots, \lambda_n) & \text{if } k \geq n \geq r \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

When no ambiguity can arise, we omit the argument and write σ_r^n.

Notice that $\sigma_r^n(\lambda_1, \ldots \lambda_k) = \sigma_r^n(\lambda_1, \ldots, \lambda_l)$ whenever $k, l \geq n \geq r$. For example

$$\sigma_3^4(\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6) = \lambda_1\lambda_2\lambda_3 + \lambda_1\lambda_2\lambda_4 + \lambda_1\lambda_3\lambda_4 + \lambda_2\lambda_3\lambda_4 = \sigma_3^4(\lambda_1, \lambda_2, \lambda_3, \lambda_4).$$

Also $\sigma_r^k(\lambda_1, \ldots \lambda_k) = c_r(\lambda_1, \ldots \lambda_k)$. With this definition, we can state a recursive formula for partial elementary symmetric polynomials.

Lemma 4.1.3 *Let k, n, r be integers such that $k \geq n \geq r$. Then*

$$\sigma_r^n(\lambda_1, \ldots, \lambda_k) = \sigma_r^{n-1}(\lambda_1, \ldots, \lambda_k) + \lambda_n \sigma_{r-1}^{n-1}(\lambda_1, \ldots, \lambda_k). \tag{4.2}$$

Proof. The result is trivial unless $k \geq n \geq r > 0$ so we shall consider only this case. The polynomial $\sigma_r^{n-1}(\lambda_1, \ldots, \lambda_k)$ contains every product of r indeterminates from $(\lambda_1, \ldots, \lambda_{n-1})$. That is, it contains every possible product of r terms from $(\lambda_1, \ldots, \lambda_n)$ which does not contain λ_n. Similarly, $\sigma_{r-1}^{n-1}(\lambda_1, \ldots, \lambda_k)$ is the sum of every possible product of $r-1$ indetermines from $(\lambda_1, \ldots, \lambda_{n-1})$ which implies that $\lambda_n \sigma_{r-1}^{n-1}(\lambda_1, \ldots, \lambda_k)$ is the sum of every possible product of r of the indeterminates $(\lambda_1, \ldots, \lambda_n)$ which contains λ_n. The RHS of (4.2) is therefore the sum of all possible products of r indeterminates from $(\lambda_1, \ldots, \lambda_n)$ which either contain, or do not contain the term λ_n. Therefore, it is the sum of all possible products of r of the indetermintes $(\lambda_1, \ldots, \lambda_n)$ which by definition is equal to $\sigma_r^n(\lambda_1, \ldots, \lambda_k)$. ■

We also need to extend the definition of the polynomial $A_{k,\alpha}$.

Definition 17 *Let $k, n \in \mathbb{N}$ and $\alpha \in \mathbb{D}$. Define*

$$A_{k,\alpha}^n(\lambda) = \sum_{r=0}^{k} (-1)^r (k-r) \sigma_r^n(\alpha\lambda).$$

Notice that $A_{k,\alpha}(\lambda) = A_{k,\alpha}^k(\lambda)$.

Lemma 4.1.4 *The following identity holds*

$$A_{k,\alpha}^n(\lambda) = A_{k,\alpha}^{n-1}(\lambda) - \alpha\lambda_n A_{k-1,\alpha}^{n-1}(\lambda).$$

Proof. Using Lemma 4.1.3 we have

$$\begin{aligned}
A_{k,\alpha}^n(\lambda) &= \sum_{r=0}^{k} (-1)^r (k-r) \sigma_r^n(\alpha\lambda) \\
&= \sum_{r=0}^{k} (-1)^r (k-r)(\sigma_r^{n-1}(\alpha\lambda) + \alpha\lambda_n \sigma_{r-1}^{n-1}(\alpha\lambda)) \\
&= A_{k,\alpha}^{n-1}(\lambda) - \alpha\lambda_n \sum_{r=0}^{k} (-1)^{r-1}(k-1-(r-1))\sigma_{r-1}^{n-1}(\alpha\lambda) \\
&= A_{k,\alpha}^{n-1}(\lambda) - \alpha\lambda_n \sum_{s=0}^{k-1} (-1)^s (k-1-s))\sigma_s^{n-1}(\alpha\lambda) \\
&= A_{k,\alpha}^{n-1}(\lambda) - \alpha\lambda_n A_{k-1,\alpha}^{n-1}(\lambda).
\end{aligned}$$

$\blacksquare$

Lemma 4.1.5 *Let $\alpha \in \mathbb{D}$, $k \in \mathbb{N}$. The following are equivalent:*

(a) $A_{k,\alpha}(\lambda) \neq 0$ *for all* $\lambda \in \mathbb{D}^k$,

(b) $A_{k,\alpha}^{k-1}(\lambda) - zA_{k-1,\alpha}^{k-1}(\lambda) \neq 0$ *for all* $z \in \mathbb{D}, \lambda \in \mathbb{D}^{k-1}$,

(c) $\left| \dfrac{A_{k,\alpha}^{k-1}(\lambda)}{A_{k-1,\alpha}^{k-1}(\lambda)} \right| \geq 1$ *for all* $\lambda \in \mathbb{D}^{k-1}$,

(d) $|A_{k,\alpha}^{k-1}(\lambda)|^2 - |A_{k-1,\alpha}^{k-1}(\lambda)|^2 \geq 0$ *for all* $\lambda \in \mathbb{D}^{k-1}$.

Proof. (a) $\Leftrightarrow$ (b) follows trivially from Lemma 4.1.4 once we observe that $A_{k-1,\alpha}^{n-1}(\lambda)$ is independent of λ_n.

(b) $\Leftrightarrow$ (c) Suppose (b) does not hold, then there exists a $z \in \mathbb{D}$ and a $\lambda \in \mathbb{D}^{k-1}$ such that

$$A_{k,\alpha}^{k-1}(\lambda) - zA_{k-1,\alpha}^{k-1}(\lambda) = 0.$$

Equivalently, $z \in \mathbb{D}$ may be expressed as

$$z = \frac{A_{k,\alpha}^{k-1}(\lambda)}{A_{k-1,\alpha}^{k-1}(\lambda)}.$$

It follows, since $|z| < 1$ that (c) is false if and only if (b) is false. This completes the proof of (b) $\Leftrightarrow$ (c).

(c) $\Leftrightarrow$ (d) Suppose (c) holds, then for all $\lambda \in \mathbb{D}^{k-1}$,

$$\left| \frac{A_{k,\alpha}^{k-1}(\lambda)}{A_{k-1,\alpha}^{k-1}(\lambda)} \right| \geq 1$$

which is equivalent to,

$$|A_{k,\alpha}^{k-1}(\lambda)| \geq |A_{k-1,\alpha}^{k-1}(\lambda)|$$

or alternatively,

$$|A_{k,\alpha}^{k-1}(\lambda)|^2 \geq |A_{k-1,\alpha}^{k-1}(\lambda)|^2.$$

$\blacksquare$

We may now prove the final special case of the main result of this section. The proof relies heavily on previous results.

Theorem 4.1.6 *For all $\alpha \in \mathbb{D}$, $z \in \mathbb{D}^4$ we have*

$$A_{4,\alpha}(z) \neq 0.$$

Proof. We shall argue by contradiction. Suppose there exists $\alpha \in \mathbb{D}$ and $z \in \mathbb{D}^4$ such that $0 = A_{4,\alpha}(z)$. Let $\lambda = \alpha z$. Then

$$\begin{aligned} 0 &= A_{4,\alpha}(z) \\ &= 4 - 3\sigma_1^4(\lambda) + 2\sigma_2^4(\lambda) - \sigma_3^4(\lambda), \end{aligned}$$

and so, by Lemma 4.1.3,

$$\begin{aligned} 0 &= 4 - 3(\sigma_1^3 + \lambda_4\sigma_0^3) + 2(c_2^3 + \lambda_4\sigma_1^3) - (\sigma_3^3 + \lambda_4\sigma_2^3) \\ &= 4 - 3\sigma_1^3 + 2c_2^3 - \sigma_3^3 - \lambda_4(3 - 2\sigma_1^3 + \sigma_2^3). \end{aligned}$$

Therefore,

$$\begin{aligned} \lambda_4 &= \frac{4 - 3\sigma_1^3 + 2c_2^3 - \sigma_3^3}{3 - 2\sigma_1^3 + \sigma_2^3} && (4.3) \\ &= \frac{4 - 3(\sigma_1^2 + \lambda_3) + 2(\sigma_2^2 + \lambda_3\sigma_1^2) - \sigma_2^2}{3 - 2(\sigma_1^2 + \lambda_3) + (\sigma_2^2 + \lambda_3\sigma_1^3)} \\ &= \frac{4 - 3\sigma_1^2 + 2\sigma_2^2 - \lambda_3(3 - 2\sigma_1^2 + \sigma_2^2)}{3 - 2\sigma_1^2 + \sigma_2^2 - \lambda_3(2 - \sigma_1^2)} \\ &= F_4(\lambda_3) && (4.4) \end{aligned}$$

where $F : \mathbb{D} \to \mathbb{C}$ is defined by

$$F_4(x) = \frac{4 - 3\sigma_1^2 + 2\sigma_2^2 - x(3 - 2\sigma_1^2 + \sigma_2^2)}{3 - 2\sigma_1^2 + \sigma_2^2 - x(2 - \sigma_1^2)}.$$

The denominator of this linear fractional transformation can be written as

$$A_{3,\alpha}^2(z) - xA_{2,\alpha}^2(z)$$

which is non-zero by Theorem 4.1.2 and Lemma 4.1.5. Thus, F_4 is an analytic linear fractional transformation on the disc, and $F_4(\mathbb{D})$ is a disc. Let c represent the centre of $F_4(\mathbb{D})$ and write the pre-image of γ under F_4 as $F_4^{-1}(\gamma)$. By inspection,

$$F_4^{-1}(\infty) = \frac{3 - 2\sigma_1^2 + \sigma_2^2}{2 - \sigma_1^2}.$$

Therefore, since conjugacy is preserved,

$$F_4^{-1}(c) = \frac{2 - \overline{\sigma_1^2}}{3 - 2\overline{\sigma_1^2} + \overline{\sigma_2^2}}.$$

The centre of $F_4(\mathbb{D})$ is equal to

$$
\begin{aligned}
&F_4\left(\frac{2-\overline{\sigma_1^2}}{3-2\overline{\sigma_1^2}+\overline{\sigma_2^2}}\right) \\
&= \frac{4-3\sigma_1^2+2\sigma_2^2-\left(\frac{2-\overline{\sigma_1^2}}{3-2\overline{\sigma_1^2}+\overline{\sigma_2^2}}\right)(3-2\sigma_1^2+\sigma_2^2)}{3-2\sigma_1^2+\sigma_2^2-\left(\frac{2-\overline{\sigma_1^2}}{3-2\overline{\sigma_1^2}+\overline{\sigma_2^2}}\right)(2-\sigma_1^2)} \\
&= \frac{(4-3\sigma_1^2+2\sigma_2^2)(3-2\overline{\sigma_1^2}+\overline{\sigma_2^2})-(2-\overline{\sigma_1^2})(3-2\sigma_1^2+\sigma_2^2)}{|3-2\sigma_1^2+\sigma_2^2|^2-|2-\sigma_1^2|^2} \\
&= \frac{[(3-2\sigma_1^2+\sigma_2^2)+(1-\sigma_1^2+\sigma_2^2)](3-2\overline{\sigma_1^2}+\overline{\sigma_2^2})-(2-\overline{\sigma_1^2})(3-2\sigma_1^2+\sigma_2^2)}{|3-2\sigma_1^2+\sigma_2^2|^2-|2-\sigma_1^2|^2} \\
&= \frac{|3-2\sigma_1^2+\sigma_2^2|^2+(1-\sigma_1^2+\sigma_2^2)(3-2\overline{\sigma_1^2}+\overline{\sigma_2^2})-(2-\overline{\sigma_1^2})(3-2\sigma_1^2+\sigma_2^2)}{|3-2\sigma_1^2+\sigma_2^2|^2-|2-\sigma_1^2|^2} \\
&= \frac{|3-2\sigma_1^2+\sigma_2^2|^2+(1-\sigma_1^2+\sigma_2^2)[(1-\overline{\sigma_1^2}+\overline{\sigma_2^2})+(2-\overline{\sigma_1^2})]-(2-\overline{\sigma_1^2})(3-2\sigma_1^2+\sigma_2^2)}{|3-2\sigma_1^2+\sigma_2^2|^2-|2-\sigma_1^2|^2} \\
&= \frac{|3-2\sigma_1^2+\sigma_2^2|^2+|1-\sigma_1^2+\sigma_2^2|^2+(1-\sigma_1^2+\sigma_2^2)(2-\overline{\sigma_1^2})-(2-\overline{\sigma_1^2})(3-2\sigma_1^2+\sigma_2^2)}{|3-2\sigma_1^2+\sigma_2^2|^2-|2-\sigma_1^2|^2} \\
&= \frac{|3-2\sigma_1^2+\sigma_2^2|^2+|1-\sigma_1^2+\sigma_2^2|^2-(2-\overline{\sigma_1^2})(2-\sigma_1^2)}{|3-2\sigma_1^2+\sigma_2^2|^2-|2-\sigma_1^2|^2} \\
&= \frac{|3-2\sigma_1^2+\sigma_2^2|^2+|1-\sigma_1^2+\sigma_2^2|^2-|2-\sigma_1^2|^2}{|3-2\sigma_1^2+\sigma_2^2|^2-|2-\sigma_1^2|^2} \\
&= 1+\frac{|1-\sigma_1^2+\sigma_2^2|^2}{|3-2\sigma_1^2+\sigma_2^2|^2-|2-\sigma_1^2|^2} \\
&= 1+\frac{|1-\sigma_1^2+\sigma_2^2|^2}{|A_{3,\alpha}^2(\lambda)|^2-|A_{2,\alpha}^2(\lambda)|^2}.
\end{aligned}
$$

But $|A_{3,\alpha}^2(\lambda)|^2-|A_{2,\alpha}^2(\lambda)|^2 \geq 0$ by Theorem 4.1.2 and Lemma 4.1.5. Thus, $c \in \mathbb{R}$ and $c \geq 1$.

Notice also that

$$
\begin{aligned}
F_4(1) &= \frac{4-3\sigma_1^2+2\sigma_2^2-3+2\sigma_1^2-\sigma_2^2}{3-2\sigma_1^2+\sigma_2^2-2+\sigma_1^2} \\
&= \frac{1-\sigma_1^2+\sigma_2^2}{1-\sigma_1^2+\sigma_2^2} \\
&= 1.
\end{aligned}
$$

It follows that 1 lies on the boundary of $F_4(\mathbb{D})$ and therefore that $\mathbb{D} \cap F_4(\mathbb{D}) = \emptyset$. This contradicts (4.3) and the result follows.

■

4.2 The General Case

A recurring technique in the simplification of the expressions in the above proofs has been the subtraction or addition of various alternating polynomials, namely polynomials of the form $1 - \sigma_1^2 + \sigma_2^2$. This observation plays a key role in the general result and gives rise to the following definition and Lemma.

Definition 18 *For $n \leq k$ define,*

$$\eta_\alpha^n(\lambda_1, \ldots, \lambda_k) = \prod_{i=1}^{n}(1 - \alpha\lambda_i) = \sum_{r=0}^{k}(-1)^r \sigma_r^n(\alpha\lambda_1, \ldots \alpha\lambda_k).$$

Polynomials of this type will be described as alternating.

Lemma 4.2.1 *For $\alpha \in \mathbb{D}$, $k > n$, and $\lambda \in \mathbb{D}^k$,*

$$A_{k,\alpha}^n(\lambda) = \eta_\alpha^n(\lambda) + A_{k-1,\alpha}^n(\lambda).$$

Proof. By definition we have

$$\begin{aligned}
A_{k,\alpha}^n(\lambda) &= \sum_{r=0}^{k}(-1)^r(k-r)\sigma_r^n(\alpha\lambda) \\
&= \sum_{r=0}^{k}(-1)^r(k-1-r)\sigma_r^n(\alpha\lambda) + \sum_{r=0}^{k}(-1)^r\sigma_r^n(\alpha\lambda) \\
&= \sum_{r=0}^{k-1}(-1)^r(k-1-r)\sigma_r^n(\alpha\lambda) + \sum_{r=0}^{k}(-1)^r\sigma_r^n(\alpha\lambda) \\
&= A_{k-1,\alpha}^n(\lambda) + \eta_\alpha^n(\lambda).
\end{aligned}$$

∎

We may now prove the main result of this section.

Theorem 4.2.2 *For $k \in \mathbb{N}$, $\alpha \in \mathbb{D}$,*

$$A_{k,\alpha}(\lambda) \neq 0 \tag{4.5}$$

for all $\lambda \in \mathbb{D}^k$.

Proof. We shall argue by induction. Theorems 4.1.1, 4.1.2 and 4.1.6 show that the result holds when $k = 1, 2, 3, 4$. Assume the result holds for $k - 1$, namely

$$A_{k-1,\alpha}(\lambda) \neq 0. \tag{4.6}$$

By Lemma 4.1.5, this induction hypothesis is equivalent to each of the following.

$$A^{k-2}_{k-1,\alpha}(\lambda) - zA^{k-2}_{k-2,\alpha}(\lambda) \neq 0 \text{ for all } z \in \mathbb{D}, \lambda \in \mathbb{D}^{k-2}, \tag{4.7}$$

$$|A^{k-2}_{k-1,\alpha}(\lambda)|^2 - |A^{k-2}_{k-2,\alpha}(\lambda)|^2 \geq 0 \text{ for all } \lambda \in \mathbb{D}^{k-2}. \tag{4.8}$$

We shall prove

$$A_{k,\alpha}(\lambda) \neq 0 \tag{4.9}$$

by assuming the contrary and arriving at a contradiction.

Accordingly, assume there exists $\alpha \in \mathbb{D}$ and $z \in \mathbb{D}^k$ such that

$$A_{k,\alpha}(z) = 0.$$

Then, by Lemma 4.1.4,

$$0 = A^{k-1}_{k,\alpha}(z) - \alpha z_k A^{k-1}_{k-1,\alpha}(z).$$

Therefore, since $A^{k-1}_{k-1,\alpha}(z) \neq 0$ by (4.6),

$$\alpha z_k = \frac{A^{k-1}_{k,\alpha}(z)}{A^{k-1}_{k-1,\alpha}(z)}$$

which implies

$$\left| \frac{A^{k-1}_{k,\alpha}(z)}{A^{k-1}_{k-1,\alpha}(z)} \right| < 1. \tag{4.10}$$

Define $F_k : \mathbb{D} \to \mathbb{C}$ by

$$F_k(x) = \frac{A^{k-2}_{k,\alpha}(z) - xA^{k-2}_{k-1,\alpha}(z)}{A^{k-2}_{k-1,\alpha}(z) - xA^{k-2}_{k-2,\alpha}(z)}.$$

Then

$$F_k(\alpha z_{k-1}) = \frac{A^{k-2}_{k,\alpha}(z) - \alpha z_{k-1} A^{k-2}_{k-1,\alpha}(z)}{A^{k-2}_{k-1,\alpha}(z) - \alpha z_{k-1} A^{k-2}_{k-2,\alpha}(z)}$$

by Lemma 4.1.4,

$$= \frac{A^{k-1}_{k,\alpha}(z)}{A^{k-1}_{k-1,\alpha}(z)}.$$

The inequality (4.10) is therefore equivalent to $|F_k(\alpha z_{k-1})| < 1$. We shall show this is impossible by proving $F_k(\mathbb{D}) \cap \mathbb{D} = \emptyset$. The linear fractional transformation F_k is analytic on $\mathbb{D}$ by the induction hypothesis (4.7), thus $F_k(\mathbb{D})$ is a disc. Let c represent the centre of ths disc and recall that linear fractional transformations preserve conjugacy. We denote the pre-image of γ under F_k by $F_k^{-1}(\gamma)$. By inspection,

$$F_k^{-1}(\infty) = \frac{A^{k-2}_{k-1,\alpha}(z)}{A^{k-2}_{k-2,\alpha}(z)}$$

and so

$$F_k^{-1}(c) = \left(\frac{A_{k-2,\alpha}^{k-2}(z)}{A_{k-1,\alpha}^{k-2}(z)} \right)^{-}.$$

Repeated use of Lemma 4.2.1 shows the centre of the disc $F_k(\mathbb{D})$ is equal to

$$\begin{aligned}
&F_k \left(\frac{A_{k-2,\bar{\alpha}}^{k-2}(\bar{z})}{A_{k-1,\bar{\alpha}}^{k-2}(\bar{z})} \right) \\
&= \frac{A_{k,\alpha}^{k-2}(z) - \left(\frac{A_{k-2,\bar{\alpha}}^{k-2}(\bar{z})}{A_{k-1,\bar{\alpha}}^{k-2}(\bar{z})} \right) A_{k-1,\alpha}^{k-2}(z)}{A_{k-1,\alpha}^{k-2}(z) - \left(\frac{A_{k-2,\bar{\alpha}}^{k-2}(\bar{z})}{A_{k-1,\bar{\alpha}}^{k-2}(\bar{z})} \right) A_{k-2,\alpha}^{k-2}(z)} \\
&= \frac{A_{k,\alpha}^{k-2}(z)A_{k-1,\bar{\alpha}}^{k-2}(\bar{z}) - A_{k-2,\bar{\alpha}}^{k-2}(\bar{z})A_{k-1,\alpha}^{k-2}(z)}{A_{k-1,\alpha}^{k-2}(z)A_{k-1,\bar{\alpha}}^{k-2}(\bar{z}) - A_{k-2,\bar{\alpha}}^{k-2}(\bar{z})A_{k-2,\alpha}^{k-2}(z)} \\
&= \frac{[A_{k-1,\alpha}^{k-2}(z) + \eta_{\alpha}^{k-2}(z)]A_{k-1,\bar{\alpha}}^{k-2}(\bar{z}) - A_{k-2,\bar{\alpha}}^{k-2}(\bar{z})A_{k-1,\alpha}^{k-2}(z)}{|A_{k-1,\alpha}^{k-2}(z)|^2 - |A_{k-2,\alpha}^{k-2}(z)|^2} \\
&= \frac{|A_{k-1,\alpha}^{k-2}(z)|^2 + \eta_{\alpha}^{k-2}(z)A_{k-1,\bar{\alpha}}^{k-2}(\bar{z}) - A_{k-2,\bar{\alpha}}^{k-2}(\bar{z})A_{k-1,\alpha}^{k-2}(z)}{|A_{k-1,\alpha}^{k-2}(z)|^2 - |A_{k-2,\alpha}^{k-2}(z)|^2} \\
&= \frac{|A_{k-1,\alpha}^{k-2}(z)|^2 + \eta_{\alpha}^{k-2}(z)[\eta_{\bar{\alpha}}^{k-2}(\bar{z}) + A_{k-2,\bar{\alpha}}^{k-2}(\bar{z})] - A_{k-2,\bar{\alpha}}^{k-2}(\bar{z})A_{k-1,\alpha}^{k-2}(z)}{|A_{k-1,\alpha}^{k-2}(z)|^2 - |A_{k-2,\alpha}^{k-2}(z)|^2} \\
&= \frac{|A_{k-1,\alpha}^{k-2}(z)|^2 + |\eta_{\alpha}^{k-2}(z)|^2 + A_{k-2,\bar{\alpha}}^{k-2}(\bar{z})[\eta_{\alpha}^{k-2}(z) - A_{k-1,\alpha}^{k-2}(z)]}{|A_{k-1,\alpha}^{k-2}(z)|^2 - |A_{k-2,\alpha}^{k-2}(z)|^2} \\
&= \frac{|A_{k-1,\alpha}^{k-2}(z)|^2 + |\eta_{\alpha}^{k-2}(z)|^2 - |A_{k-2,\alpha}^{k-2}(z)|^2}{|A_{k-1,\alpha}^{k-2}(z)|^2 - |A_{k-2,\alpha}^{k-2}(z)|^2} \\
&= 1 + \frac{|\eta_{\alpha}^{k-2}(z)|^2}{|A_{k-1,\alpha}^{k-2}(z)|^2 - |A_{k-2,\alpha}^{k-2}(z)|^2} \\
&\geq 1
\end{aligned}$$

by (4.8).

Since F_k preserves conjugacy, $F_k(1)$ will lie on the boundary of $F_k(\mathbb{D})$. Now,

$$\begin{aligned}
F_k(1) &= \frac{A_{k,\alpha}^{k-2}(z) - A_{k-1,\alpha}^{k-2}(z)}{A_{k-1,\alpha}^{k-2}(z) - A_{k-2,\alpha}^{k-2}(z)} \\
&= \frac{\eta_{\alpha}^{k-2}(z)}{\eta_{\alpha}^{k-2}(z)} \\
&= \frac{\eta_{\alpha}^{k-2}(z)}{\eta_{\alpha}^{k-2}(z)} \\
&= 1.
\end{aligned}$$

It follows that $F(\mathbb{D}) \cap \mathbb{D} = \emptyset$ which contradicts (4.10); therefore (4.9) holds. The result follows by the principle of induction.

■

Corollary 4.2.3 *Let $k \in \mathbb{N}$. If $\alpha \in \mathbb{D}$ and $\lambda \in \overline{\mathbb{D}}^k$ or $\alpha \in \overline{\mathbb{D}}$ and $\lambda \in \mathbb{D}^k$ then*

$$A_{k,\alpha}(\lambda) \neq 0.$$

Proof. If α and λ satisfy either of the conditions in the statement of the result, then there exists $\beta \in \mathbb{D}$ and $\zeta \in \mathbb{D}^k$ such that $\alpha\lambda = \beta\zeta$. The result then follows by Theorem 4.2.2.

■

After the completion of this Chapter, Dr. Michael White suggested that the results proved above could be demonstrated more simply by noticing that $A_{k,\alpha}$ is similar to a derivative with respect to α. He suggests that one could then employ the Gauss-Lucas Theorem [17, Exercise 4.50] to draw the conclusions given above.

Chapter 5

A Necessary Condition for Spectral Interpolation

5.1 Main Theorem

For each natural number k we shall define a polynomial ρ_k. In terms of this polynomial, we shall give a necessary condition for the solution of the problem of interpolating into Γ_k. Recall the definition of P_k given in (1.6):

$$P_k(x_0, \ldots, x_k; y_0, \ldots, y_k) = \sum_{r,s=0}^{k} (-1)^{r+s}(k-(r+s))y_s x_r.$$

Definition 19 *Let the hereditary polynomial ρ_k be defined as follows:*

$$\rho_k(x_1, \ldots, x_k; y_1, \ldots, y_k) = P_k(1, x_1, \ldots, x_k; 1, y_1, \ldots, y_k) \tag{5.1}$$

By virtue of (3.6), and the fact that $c_0(z) = 1$, we have

$$p_{k,\alpha}(\lambda, \bar{\lambda}) = P_k(c_0(\alpha\lambda), \ldots, c_k(\alpha\lambda); c_0(\overline{\alpha\lambda}), \ldots, c_k(\overline{\alpha\lambda})) = \rho_k(c_1(\alpha\lambda), \ldots, c_k(\alpha\lambda)). \tag{5.2}$$

This observation plays a key role in the proof of our Main Theorem (Theorem 5.1.5).

The proofs of Lemmas 5.1.3 and 5.1.4 are extensions of the proofs in the case $k = 2$, which was dealt with by Agler and Young in [6]. The proofs rely heavily on the following well known theorems. The first theorem is due to Arveson and the second to Stinespring. Proofs of both results can be found in [8] where they are labelled Theorems 1.1.1 and 1.2.9. The version of Arveson's result given below relies on the remarks immediately preceeding Proposition 1.2.11 of [8]. The original proof of Stinespring's result is in [35].

Theorem 5.1.1 (Arveson's Extension Theorem)
Let C be a C^*-algebra with identity and let A be a linear subspace of C containing the identity. If H is a Hilbert space and $\theta : A \to \mathcal{L}(\mathcal{H})$ is a completely contractive linear map then there exists a completely contractive linear map $\Theta : C \to \mathcal{L}(\mathcal{H})$ such that $\theta = \Theta|_A$.

Theorem 5.1.2 (Stinespring's Theorem) *Let C be a C^*-algebra with identity, H be a Hilbert space, and assume $\Theta : C \to \mathcal{L}(\mathcal{H})$ is a completely positive linear map. Then there exists a Hilbert space $\mathcal{K}$, a bounded linear map $V : H \to \mathcal{K}$ and a representation $\pi : C \to \mathcal{L}(\mathcal{K})$ such that $\Theta(x) = V^*\pi(x)V$ for all $x \in C$.*

Having recalled the above results, we are in a position to extend the results of Agler and Young. Recall Definition 7, where we defined the joint spectrum of a k-tuple of commuting operators.

Definition 20 *Define the* distinguished boundary of Γ_k *as*

$$b\Gamma_k = \{\pi(z) \mid z \in \mathbb{T}^k\}.$$

Definition 21 *For $X \subset \mathbb{C}^k$, X compact, $A(X)$ denotes the algebra of continuous functions on X which are analytic on the interior of X.*

Lemma 5.1.3 *Let $(X_1, \dots, X_k)$ be a commuting k-tuple of operators on a Hilbert space $\mathcal{H}$ such that $\sigma(X_1, \dots X_k) \subset \Gamma_k$. If Γ_k is a complete spectral set for $(X_1, \dots, X_k)$ then there exist Hilbert spaces $\mathcal{H}_-, \mathcal{H}_+$ and a k-tuple of commuting normal operators $(\tilde{X}_1, \dots, \tilde{X}_k)$ on $\mathcal{K} \stackrel{def}{=} \mathcal{H}_- \oplus \mathcal{H} \oplus \mathcal{H}_+$ such that $\sigma(\tilde{X}_1, \dots, \tilde{X}_k)$ is contained in the distinguished boundary of Γ_k and each $\tilde{X}_j$ is expressible by an operator matrix of the form*

$$\tilde{X}_j \sim \begin{bmatrix} * & * & * \\ 0 & X_j & * \\ 0 & 0 & * \end{bmatrix} \tag{5.3}$$

with respect to the orthogonal decomposition $\mathcal{K} = \mathcal{H}_- \oplus \mathcal{H} \oplus \mathcal{H}_+$.

Proof. Suppose that Γ_k is a complete spectral set for a commuting k-tuple of operators $(X_1, \dots, X_k)$. That is

$$\|h(X_1, \dots, X_k)\|_{\mathcal{L}(\mathcal{H})} \leq \sup_{z \in \Gamma_k} \|h(z)\| \tag{5.4}$$

for all analytic matrix valued functions h.

Let $\theta : \mathcal{P}_k \to \mathcal{L}(\mathcal{H})$ be the unital representation of the algebra $\mathcal{P}_k$ of polynomials in k variables defined by $\theta(h) = h(X_1, \dots, X_k)$. Inequality (5.4) states that θ is completely contractive and therefore uniformly continuous. Hence θ has a completely contractive extension to $A(\Gamma_k)$. The space $A(\Gamma_k)$ is embedded in the C^*-algebra $C(\mathbb{T}^k)$ of continuous functions on the k-torus by

$$f \in A(\Gamma_k) \mapsto f \circ \pi \in C(\mathbb{T}^k).$$

By Theorem 5.1.1, θ extends to a completely contractive unital linear mapping $\Theta : C(\mathbb{T}^k) \to \mathcal{L}(\mathcal{H})$. By Theorem 5.1.2, there exists a Hilbert space $\mathcal{K}$ and a unital representation $\Phi : C(\mathbb{T}^k) \to \mathcal{L}(\mathcal{K})$ such that $\mathcal{H} \subset \mathcal{K}$ and

$$\Theta(f) = P_{\mathcal{H}} \Phi(f)|_{\mathcal{H}}$$

for all $f \in C(\mathbb{T}^k)$, where $P_{\mathcal{H}}$ is the orthogonal projection from $\mathcal{K}$ to $\mathcal{H}$. Thus, if $f \in A(\Gamma_k)$ we have

$$f(X_1, \dots, X_k) = \theta(f) = \Theta(f) = P_{\mathcal{H}} \Phi(f \circ \pi)|_{\mathcal{H}}.$$

Let f_j represent the j^{th} co-ordinate function, i.e. $f_j(z) = z_j,\ j = 1, \dots, k$. Then for $j = 1, \dots, k$,

$$f_j \circ \pi(\lambda) = c_j(\lambda)$$

and

$$X_j = f_j(X_1, \dots, X_k) = \theta(f_j) = \Theta(f_j \circ \pi) = \Theta(c_j(\lambda)) = P_{\mathcal{H}} \Phi(c_j(\lambda))|_{\mathcal{H}}.$$

For $j = 1, \dots, k$, let

$$\tilde{X}_j = \Phi(c_j(\lambda)).$$

Observe that the operators $\tilde{X}_1, \ldots, \tilde{X}_k$ lie in the commutative *-algebra $\Phi(C(\mathbb{T}^k))$. The $\tilde{X}_j$ are therefore commuting normal operators with $X_1, \ldots, X_k$ as their compressions to $\mathcal{H}$. We may suppose that the smallest subspace of $\mathcal{K}$ which contains $\mathcal{H}$ and reduces each $\tilde{X}_j$ is $\mathcal{K}$. The joint spectrum of the k-tuple of elements $(c_1(\lambda), \ldots, c_k(\lambda))$ in $C(\mathbb{T}^k)$ is the range of this k-tuple of functions in $\mathbb{T}^k$. The range of these functions is exactly the distinguished boundary of Γ_k. Next we may apply the unital representation Φ to deduce that $\sigma(\tilde{X}_1, \ldots, \tilde{X}_k) \subset b\Gamma_k$.

We shall construct spaces $\mathcal{H}_-$ and $\mathcal{N}$ which are invariant for each X_j and satisfy $\mathcal{H}_- \subset \mathcal{N} \subset \mathcal{K}$ and $\mathcal{H} = \mathcal{N} \ominus \mathcal{H}_-$. Let

$$\mathcal{H}_- = \{x \in \mathcal{K} \mid f(\tilde{X}_1, \ldots, \tilde{X}_k)x \perp \mathcal{H} \text{ for every polynomial } f\}.$$

Clearly $\mathcal{H}_-$ is invariant for each X_j, $j = 1, \ldots, k$, and $\mathcal{H}_- \perp \mathcal{H}$. Let $\mathcal{N} = \mathcal{H}_- \oplus \mathcal{H}$. Then $\mathcal{N}$ is also invariant for each X_j. We can prove this by considering any element $x \in \mathcal{H}$ and showing that, for each $j \in \{1, \ldots, k\}$,

$$\tilde{X}_j x - P_{\mathcal{H}} \tilde{X}_j x \in \mathcal{H}_-.$$

Suppose f is any polynomial in k variables; then

$$\begin{aligned} P_{\mathcal{H}}\{f(\tilde{X}_1, \ldots, \tilde{X}_k)(\tilde{X}_j x - P_{\mathcal{H}} \tilde{X}_j x)\} &= P_{\mathcal{H}} f(\tilde{X}_1, \ldots, \tilde{X}_k)\tilde{X}_j x - f(\tilde{X}_1, \ldots, \tilde{X}_k) P_{\mathcal{H}} \tilde{X}_j x \\ &= f(X_1, \ldots, X_k) X_j x - f(X_1, \ldots, X_k) X_j x \\ &= 0. \end{aligned}$$

Thus, $\tilde{X}_j x - P_H \tilde{X}_j x \in \mathcal{H}_-$ and therefore $\mathcal{N}$ is invariant for each X_j as claimed. Clearly, X_j has the form (5.3) with respect to the decomposition $\mathcal{K} = \mathcal{H}_- \oplus \mathcal{H} \oplus \mathcal{N}^\perp$ so we may take $\mathcal{H}_+ = \mathcal{N}^\perp$ to see that the result holds.

■

Lemma 5.1.4 which follows is the key tool in the proof of the Main Theorem of this chapter (Theorem 5.1.5). The results of Chapter 4 are crucial to the proof of Lemma 5.1.4. Without the technical results of that chapter the simplification which Lemma 5.1.4 permits would not be possible and the proof of Theorem 5.1.5 would be unattainable.

Lemma 5.1.4 *Let $(X_1, \ldots, X_k)$ be a commuting k-tuple of operators on a Hilbert space H such that $\sigma(X_1, \ldots X_k) \subset \Gamma_k$. If Γ_k is a complete spectral set for $(X_1, \ldots, X_k)$ and*

$$\rho_k(\alpha c_1(\lambda_1, \ldots, \lambda_k), \ldots, \alpha^k c_k(\lambda_1, \ldots, \lambda_k)) \geq 0$$

for $|\lambda_j| = 1, \alpha \in \overline{\mathbb{D}}$ then

$$\rho_k(\alpha X_1, \ldots, \alpha^k X_k) \geq 0$$

for all $\alpha \in \overline{\mathbb{D}}$.

Proof. Define $\mathcal{K}$ and $\tilde{X}_j$ as in Lemma 5.1.3. Since the $\tilde{X}_j$ are normal and commute it follows that they generate a commutative C^*-subalgebra $\mathcal{A}$ of $\mathcal{L}(\mathcal{K})$. The space $C(\sigma(\tilde{X}_1,\ldots,\tilde{X}_k))$ can be identified with $\mathcal{A}$ by the Gelfand transform. By this transform, $\tilde{X}_j$ can be identified with the j^{th} coordinate function. An operator is positive semi-definite if and only if its Gelfand transform is non-negative. Thus, since $\sigma(\tilde{X}_1,\ldots,\tilde{X}_k) \subset b\Gamma_k$ and

$$\rho_k(\alpha c_1(\lambda_1,\ldots,\lambda_k),\ldots\ ,\alpha^k c_k(\lambda_1,\ldots,\lambda_k)) \geq 0$$

for all $(\alpha,(\lambda_1,\ldots,\lambda_k)) \in \overline{\mathbb{D}} \times \mathbb{T}^k$ it follows that

$$\rho_k(\alpha\tilde{X}_1,\ldots\ ,\alpha^k\tilde{X}_k) \geq 0 \tag{5.5}$$

for all $\alpha \in \overline{\mathbb{D}}$. We wish to show

$$\rho_k(\alpha X_1,\ldots\ ,\alpha^k X_k) \geq 0 \tag{5.6}$$

for all $\alpha \in \overline{\mathbb{D}}$.

To simplify the following calculations we define

$$\tilde{A} = \sum_{r=0}^{k}(-1)^r(k-r)\alpha^r\tilde{X}_r,$$
$$A = \sum_{r=0}^{k}(-1)^r(k-r)\alpha^r X_r,$$
$$\tilde{B} = \sum_{r=0}^{k}(-1)^r r\alpha^r\tilde{X}_r,$$
$$B = \sum_{r=0}^{k}(-1)^r r\alpha^r X_r.$$

Note that

$$\tilde{A} \sim \begin{bmatrix} * & * & * \\ 0 & A & * \\ 0 & 0 & * \end{bmatrix}, \qquad \tilde{B} \sim \begin{bmatrix} * & * & * \\ 0 & B & * \\ 0 & 0 & * \end{bmatrix}. \tag{5.7}$$

with respect to the orthogonal decomposition $\mathcal{K} = \mathcal{H}_- \oplus \mathcal{H} \oplus \mathcal{H}_+$. By Lemma 1.3.1, equation (5.5) is equivalent to,

$$k\rho_k(\alpha\tilde{X}_1,\ldots\ ,\alpha^k\tilde{X}_k) = \tilde{A}^*\tilde{A} - \tilde{B}^*\tilde{B} \geq 0 \tag{5.8}$$

for all $\alpha \in \overline{\mathbb{D}}$.

We next show that $\tilde{A}$ is invertible. Recall $\sigma(\tilde{X}_1, \dots, \tilde{X}_k)$ is contained in the distinguished boundary of Γ_k. Thus, for all $\alpha \in \mathbb{D}$,

$$\begin{aligned}
\sigma(\tilde{A}) &= \sigma\left(\sum_{r=0}^{k}(-1)^r(k-r)\alpha^r\tilde{X}_r\right) \\
&\subset \left\{\sum_{r=0}^{k}(-1)^r(k-r)\alpha^r c_r(\lambda) \mid \lambda \in \mathbb{T}^k\right\} \\
&\subset \left\{\sum_{r=0}^{k}(-1)^r(k-r)c_r(\alpha\lambda) \mid \lambda \in \overline{\mathbb{D}}^k\right\} \\
&= \left\{A_{k,\alpha}(\lambda) \mid \lambda \in \overline{\mathbb{D}}^k\right\}.
\end{aligned}$$

Corollary 4.2.3 states that $A_{k,\alpha}(\lambda) \neq 0$ for all $\alpha \in \mathbb{D}$ and $\lambda \in \overline{\mathbb{D}}^k$. Therefore, for $\alpha \in \mathbb{D}$,

$$0 \notin \sigma(\tilde{A})$$

and $\tilde{A}$ is invertible. The inequality in (5.8) is equivalent to

$$1 - \tilde{A}^{*-1}\tilde{B}^*\tilde{B}\tilde{A}^{-1} \geq 0.$$

That is,

$$\|\tilde{B}\tilde{A}^{-1}\| \leq 1.$$

In view of (5.7), A must be invertible and

$$\tilde{A}^{-1} \sim \begin{bmatrix} * & * & * \\ 0 & A^{-1} & * \\ 0 & 0 & * \end{bmatrix}.$$

Hence

$$\tilde{B}\tilde{A}^{-1} \sim \begin{bmatrix} * & * & * \\ 0 & BA^{-1} & * \\ 0 & 0 & * \end{bmatrix}.$$

Thus, BA^{-1} is the compression to $\mathcal{H}$ of $\tilde{B}\tilde{A}^{-1}$, and so

$$\|BA^{-1}\| \leq 1.$$

It follows in turn that

$$\begin{aligned}
1 - A^{*-1}B^*BA^{-1} &\geq 0, \\
A^*A - B^*B &\geq 0,
\end{aligned}$$

and so

$$\rho_k(\alpha X_1, \dots, \alpha^k X_k) \geq 0$$

for $\alpha \in \mathbb{D}$ and, by continuity, for all $\alpha \in \overline{\mathbb{D}}$. That is, (5.5) implies (5.6).

$\blacksquare$

We may now prove our Main Theorem. It will lead to a necessary condition for a solution to the Main Problem to exist.

Theorem 5.1.5 *Let $(X_1, \dots, X_k)$ be a commuting k-tuple of operators on a Hilbert space H such that $\sigma(X_1, \dots X_k) \subset \Gamma_k$. If Γ_k is a complete spectral set for $(X_1, \dots, X_k)$ then*

$$\rho_k(\alpha X_1, \dots\ , \alpha^k X_k) = \sum_{r,s=0}^{k} (-1)^{r+s}(k-(r+s))\bar{\alpha}^s \alpha^r X_r^* X_s \geq 0$$

for all $\alpha \in \overline{\mathbb{D}}$.

Proof. By Lemma 5.1.4 it will suffice to show that

$$\rho_k(\alpha c_1(\lambda), \dots, \alpha^k c_k(\lambda)) \geq 0$$

for all $\alpha \in \overline{\mathbb{D}}$ and all $\lambda \in \mathbb{T}^k$. However,

$$\rho_k(\alpha c_1(\lambda), \dots, \alpha^k c_k(\lambda)) = \rho_k(c_1(\alpha\lambda), \dots, c_k(\alpha\lambda)),$$

so that

$$\rho_k(\alpha c_1(\lambda), \dots, \alpha^k c_k(\lambda)) = p_{k,\alpha}(\lambda, \bar{\lambda})$$

by (5.2). That is, since $\lambda_j \in \mathbb{T}$,

$$\rho_k(\alpha c_1(\lambda), \dots, \alpha^k c_k(\lambda)) = (1-|\alpha|^2) \sum_{j=1}^{k} \prod_{i \neq j} |1-\alpha\lambda_i|^2$$

This is non-negative for all $\alpha \in \overline{\mathbb{D}}$ and so the result holds.

$\blacksquare$

5.2 Associated Results

The Main Theorem leads directly to a necessary condition for the existence of an interpolating function from $\mathbb{D}$ to Γ_k.

Corollary 5.2.1 *Let $n \in \mathbb{N}$. Choose n distinct points z_j in $\mathbb{D}$ and n points $(c_1^{(j)}, \dots, c_k^{(j)})$ in Γ_k. If there exists an analytic function $\phi : \mathbb{D} \to \Gamma_k$ such that $\phi(z_j) = (c_1^{(j)}, \dots, c_k^{(j)})$ then*

$$\rho_k(\alpha C_1, \dots\ , \alpha^k C_k) \geq 0$$

for all $\alpha \in \overline{\mathbb{D}}$, where $C_1, \ldots, C_k$ are the operators defined by

$$C_i = \operatorname{diag}\{\overline{c_i^{(1)}}, \ldots, \overline{c_i^{(n)}}\}$$

on the Hilbert space

$$\mathcal{M} = \operatorname{Span}\{K_{z_1}, \ldots, K_{z_n}\}.$$

Proof. Suppose such a ϕ exists. Theorem 2.1.3 implies that Γ_k is a complete spectral set for the commuting k-tuple of operators $(C_1, \ldots,\ C_k)$. If this is the case then the given polynomial is positive semi-definite by an application of Theorem 5.1.5.

■

This result may be converted to a partial solution to the Main Problem with the necessary condition in the more familiar form of the positivity of Pick matrices.

Corollary 5.2.2 *Let $n \in \mathbb{N}$. Choose n distinct points z_j in $\mathbb{D}$ and n matrices W_j in $\mathbb{M}_k(\mathbb{C})$. If there exists an analytic function $\phi : \mathbb{D} \to \mathbb{M}_k(\mathbb{C})$ such that $\phi(z_j) = W_j$ and $\sigma(\phi(z)) \subset \overline{\mathbb{D}}$ for all $z \in \mathbb{D}$ then, for every $\alpha \in \overline{\mathbb{D}}$,*

$$\left[\frac{\sum_{r,s=0}^{k}(-1)^{r+s}(k-(r+s))c_r(\alpha W_j)c_s(\overline{\alpha W_i})}{1-\overline{z_i}z_j}\right]_{i,j=1}^{n} \geq 0. \tag{5.9}$$

Proof. Suppose an analytic function $\phi : \mathbb{D} \to \mathbb{M}_k(\mathbb{C})$ is such that $\phi(z_j) = W_j$ for $j = 1, \ldots, n$. The composition of the analytic functions ϕ and a (defined in Definition 5) is also analytic. Let $\Phi : \mathbb{D} \to \Gamma_k$ be defined as

$$\Phi = a \circ \phi.$$

That is, Φ is an analytic function which maps the point z_j in the disk to the point $a(W_j) = (c_1^{(j)}, \ldots c_k^{(j)})$ in Γ_k. It follows from Corollary 5.2.1 that

$$\rho_k(\alpha C_1, \ldots\ , \alpha^k C_k) \geq 0$$

for all $\alpha \in \mathbb{D}$, where C_j is defined as in (2.1). This is the same as

$$\left[\langle \rho_k(\alpha C_1 \ldots\ , \alpha^k C_k)K_{z_i}, K_{z_j}\rangle\right]_{i,j=1}^{n} \geq 0. \tag{5.10}$$

for all $\alpha \in \mathbb{D}$.

Now,

$$
\begin{aligned}
&\langle \rho_k(\alpha C_1, \dots, \alpha^k C_k) K_{z_i}, K_{z_j} \rangle \\
&\qquad = \left\langle \left(\sum_{r,s=0}^{k} (-1)^{r+s} (k-(r+s)) \alpha^r \bar{\alpha}^s C_s^* C_r \right) K_{z_i}, K_{z_j} \right\rangle \\
&\qquad = \sum_{r,s=0}^{k} (-1)^{r+s} (k-(r+s)) \alpha^r \bar{\alpha}^s \langle C_s^* C_r K_{z_i}, K_{z_j} \rangle \\
&\qquad = \sum_{r,s=0}^{k} (-1)^{r+s} (k-(r+s)) \alpha^r \bar{\alpha}^s \langle C_r K_{z_i}, C_s K_{z_j} \rangle \\
&\qquad = \sum_{r,s=0}^{k} (-1)^{r+s} (k-(r+s)) \alpha^r \bar{\alpha}^s \langle \overline{c_r(W_i)} K_{z_i}, \overline{c_s(W_j)} K_{z_j} \rangle \\
&\qquad = \sum_{r,s=0}^{k} (-1)^{r+s} (k-(r+s)) \alpha^r \bar{\alpha}^s \overline{c_r(W_i)} c_s(W_j) \langle K_{z_i}, K_{z_j} \rangle \\
&\qquad = \sum_{r,s=0}^{k} (-1)^{r+s} (k-(r+s)) \overline{c_r(\bar{\alpha} W_i)} c_s(\bar{\alpha} W_j) K_{z_i}(z_j) \\
&\qquad = (1 - z_i \overline{z_j})^{-1} \sum_{r,s=0}^{k} (-1)^{r+s} (k-(r+s)) \overline{c_r(\bar{\alpha} W_i)} c_s(\bar{\alpha} W_j).
\end{aligned}
$$

Therefore (5.10) holds for all $\alpha \in \mathbb{D}$ if and only if (5.9) holds for all $\alpha \in \mathbb{D}$.

■

In the next section we illustrate a use of this result in the simplest case not studied elsewhere. We consider the special case of the Main Problem with 2 interpolation points and 3×3 target matrices.

5.3 An Illustrative Example

Before we present an example demonstrating the use of the results in the previous section we shall simplify their statements by introducing some new notation.

Suppose W_1 and W_2 are 3×3 complex matrices. Let

$$s_j = c_1(W_j), \qquad b_j = c_2(W_j), \qquad p_j = c_3(W_j). \tag{5.11}$$

Using this notation we will specialise Corollary 5.2.2 to the relevant result for a two point interpolation problem whose target values are 3×3 matrices.

Corollary 5.3.1 *Let W_1 and W_2 be 3×3 complex matrices and suppose $z_1, z_2 \in \mathbb{D}$. Define $s_j, b_j, p_j, j = 1, 2$ as in (5.11). If there exists an analytic function $F : \mathbb{D} \to \mathbb{M}_3(\mathbb{C})$ such that $F(z_1) = W_1,\ F(z_2) = W_2$ and $\sigma(F(\mathbb{D})) \subset \overline{\mathbb{D}}$ then*

$$\left[\frac{\begin{array}{c} 3[1 - |\alpha|^6 p_j \bar{p_i}] + 2[\alpha(|\alpha|^4 p_j \bar{b_i} - s_j) + \bar{\alpha}(|\alpha|^4 b_j \bar{p_i} - \bar{s_i})] \\ +[\alpha^2(b_j - |\alpha|^2 s_j \bar{p_i}) + \alpha^2(\bar{b_i} - |\alpha|^2 s_i \bar{p_j}) + |\alpha|^2(s_j \bar{s_i} - b_j \bar{b_i})] \end{array}}{1 - \bar{z}_i z_j} \right]_{i,j=1}^{2} \geq 0$$

for all $\alpha \in \overline{\mathbb{D}}$.

Example 1 *Let*

$$W_1 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \qquad W_2 = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 3/4 \end{bmatrix}.$$

Does there exist an analytic function $F : \mathbb{D} \to \mathbb{M}_3(\mathbb{C})$ such that

$$\sigma(F(\mathbb{D})) \subset \overline{\mathbb{D}} \tag{5.12}$$

and

$$F(0) = W_1, \qquad F(\tfrac{1}{4}) = W_2? \tag{5.13}$$

The eigenvalues of W_1 and W_2 are $(0, 0, 0)$ and $(\frac{3}{4}, 0, 0)$ respectively, so that $\sigma(W_j) \subset \mathbb{D}$. In the notation of (5.11) we have

$$s_1 = b_1 = p_1 = b_2 = p_2 = 0, \quad s_2 = \frac{3}{4}.$$

Consider the Pick-type matrix

$$\begin{aligned}
M_\alpha &= \left[\frac{\begin{array}{c} 3[1-|\alpha|^6 p_j\bar{p_i}] + 2[\alpha(|\alpha|^4 p_j\bar{b_i} - s_j) + \bar{\alpha}(|\alpha|^4 b_j\bar{p_i} - \bar{s_i})] \\ +[\alpha^2(b_j - |\alpha|^2 s_j\bar{p_i}) + \alpha^2(\bar{b_i} - |\alpha|^2 s_i\bar{p_j}) + |\alpha|^2(s_j\bar{s_i} - b_j\bar{b_i})] \end{array}}{1-\bar{z_i}z_j} \right]_{i,j=1}^{2} \\
&= \begin{bmatrix} 3 & 3-2\overline{\alpha s_2} \\ 3-2\alpha s_2 & \dfrac{3-2(\overline{\alpha s_2}+\alpha s_2)+(\overline{\alpha s_2})(\alpha s_2)}{1-|z_2|^2} \end{bmatrix} \\
&= \begin{bmatrix} 3 & 3-\frac{3}{2}\bar{\alpha} \\ 3-\frac{3}{2}\alpha & \dfrac{16(3-3\operatorname{Re}\alpha+\frac{9}{16}|\alpha|^2)}{15} \end{bmatrix}.
\end{aligned}$$

Now,

$$\begin{aligned}
\det M_\alpha &= \frac{3\times 16}{15}(3-3\operatorname{Re}\alpha+\frac{9}{16}|\alpha|^2) - (3-\frac{3}{2}\bar{\alpha})(3-\frac{3}{2}) \\
&= \frac{48}{5} - \frac{48}{5}\operatorname{Re}\alpha + \frac{9}{5}|\alpha|^2 - 9 + 9\operatorname{Re}\alpha - \frac{9}{4}|\alpha|^2 \\
&= \frac{3}{5} - \frac{3}{5}\operatorname{Re}\alpha - \frac{9}{20}|\alpha|^2.
\end{aligned}$$

Therefore $\det M_1 < 0$. It follows that M_1 is not a positive semi-definite matrix. Corollary 5.3.1 states that if there exists an analytic function satisfying (5.12) and (5.13) then $M_\alpha \geq 0$ for all $\alpha \in \overline{\mathbb{D}}$. We may therefore conclude that no such function exists.

Given the results of this chapter, it is natural to wonder whether the necessary conditions established are also sufficient. Unfortunately a number of the links in the chain of implications which are used to prove the above results are only 'one-way'. Agler and Young have succeeded in showing that a number of these are equivalences in the case $k = 2$. Sadly, from the point of view of this work, whenever $k > 2$ it is impossible to use the Commutant Lifting Theorem, which was the main tool of Agler and Young in the proofs of all of their 'backwards' implications. A more detailed discussion of sufficiency and related issues can be found in Chapter 8 along with relevant references to the work of Agler and Young.

Chapter 6

A Refined Necessary Condition for Spectral Interpolation

In Chapter 5 we demonstrated how the technical results of the earlier chapters give rise to a concrete necessary condition for spectral interpolation. In this chapter we show how the same technical results give rise to another necessary condition which is (potentially) stronger and only slightly more difficult to implement. Theorem 5.1.5 can be deduced immediately from the work in this chapter.

The results of Chapter 5 rely on the fact that Γ_k is a complete spectral set for a certain k-tuple of operators if an interpolating function satisfying the required conditions is to exist. The work which follows uses a very similar approach. We show that $\overline{\mathbb{D}} \times \Gamma_k$ must be a complete spectral set for a certain $(k+1)$-tuple of operators if a suitable Γ_k-valued function is to exist and satisfy certain interpolating conditions.

As one might expect, just as the motivation for the two necessary conditions is similar, so are the methods of proof. The reader will notice that all of the proofs in this chapter are extensions of the corresponding results in Chapter 5. For completeness all proofs are given in full.

If $\overline{\mathbb{D}} \times \Gamma_k$ is a complete spectral set for a commuting $(k+1)$-tuple of operators $(A, C_1, \ldots, C_k)$ then $\mathbb{D}$ is a complete spectral set for A and Γ_k is a complete spectral set for $(C_1, \ldots, C_k)$. The converse, however, does not hold. In [20] Crabb and Davie construct a triple of commuting contractions (T_1, T_2, T_3) and a symmetric polynomial f bounded by 1 on $\mathbb{D}^3$ such that

$$\|f(T_1, T_2, T_3)\| > 1.$$

Since the polynomial f is symmetric, there exists a polynomial g which is bounded by 1 on $\overline{\mathbb{D}} \times \Gamma_2$ such that

$$f(x_1, x_2, x_3) = g(x_1, x_2 + x_3, x_2 x_3).$$

Then $(T_2 + T_3, T_2T_3)$ is a complete Γ_2-contraction, but $g(T_1, T_2 + T_3, T_2T_3)$ is not a contraction and so $(T_1, T_2 + T_3, T_2T_3)$ is not a complete $\overline{\mathbb{D}} \times \Gamma_2$-contraction.

6.1 Definitions

For each $k \in \mathbb{N}$ we introduce a polynomial μ_k in $2(k+1)$ variables. This polynomial will play a similar role to that of ρ_k in Chapter 5. Let

$$\begin{aligned} \mu_k(x_0, \ldots, x_k; y_0, \ldots, y_k) = {} & \frac{1}{k}\left(\sum_{r=0}^{k}(-1)^r(k-r)y_0^r y_r\right)\left(\sum_{r=0}^{k}(-1)^r(k-r)x_0^r x_r\right) \\ & - \frac{1}{k}\left(\sum_{r=0}^{k}(-1)^r r y_0^{r-1} y_r\right)\left(\sum_{r=0}^{k}(-1)^r r x_0^{r-1} x_r\right). \end{aligned} \tag{6.1}$$

Let $n, k \in \mathbb{N}$. Suppose we wish to find an analytic function $\phi : \mathbb{D} \to \Gamma_k$ such that $\phi(z_j) = (c_1^{(j)}, \ldots, c_k^{(j)})$ for $j = 1, \ldots, n$. Define the Hilbert space $\mathcal{M}$ with basis $K_{z_1}, \ldots, K_{z_n}$ and the operators $C_1, \ldots, C_k$ as in Section 2.1. In addition, define Λ on $\mathcal{M}$ such that

$$\Lambda \sim \operatorname{diag}\{\bar{z_1}, \ldots, \bar{z_n}\} \tag{6.2}$$

with respect to the basis $K_{z_1}, \ldots, K_{z_n}$. In Chapter 1 it was shown that the $(k+1)$-tuple of operators $\Lambda, C_1, \ldots, C_k$ commute.

6.2 A New Necessary Condition

Theorem 6.2.1 *If there exists a function $\phi : \mathbb{D} \to \Gamma_k$ which is analytic and has the property that $\phi(z_j) = ({c_1}^{(j)}, \ldots, {c_k}^{(j)})$ for $j = 1, \ldots, n$, then $\overline{\mathbb{D}} \times \Gamma_k$ is a complete spectral set for the commuting $(k+1)$-tuple of operators $(\Lambda, C_1, \ldots, C_k)$ as defined in* (2.1) *and* (6.2).

Proof. By Lemma 2.1.2 we need only consider matricial polynomial functions h on $\overline{\mathbb{D}} \times \Gamma_k$. Consider the scalar polynomial case. Let h be a polynomial in $k+1$ variables. Observe that, for $1 \leq j \leq n$, we have

$$\begin{aligned} h(\Lambda, C_1, \ldots, C_k)K_{z_j} &= \sum_{r_0, \ldots, r_k} a_{r_0 \cdots r_k} \Lambda^{r_0} {C_1}^{r_1} \cdots {C_k}^{r_k} K_{z_j} \\ &= \sum_{r_0, \ldots, r_k} a_{r_0 \cdots r_k} \overline{z_j}^{r_0} \overline{{c_1}^{(j)}}^{r_1} \cdots \overline{{c_k}^{(j)}}^{r_k} K_{z_j} \\ &= \overline{h^\vee \circ (\mathrm{id} \times \phi)(z_j)} K_{z_j}. \\ &= h \circ (\mathrm{id} \times \phi)^\vee(\Lambda) K_{z_j}. \end{aligned}$$

Hence, if $h = [h_{ij}]$ is a $p \times q$ matrix polynomial and $z \in \{z_1, \dots, z_n\}$ then

$$h(\Lambda, C_1, \dots, C_k) \begin{bmatrix} 0 \\ \vdots \\ K_z \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} h_{1j}(\Lambda, C_1, \dots, C_k)K_z \\ \vdots \\ h_{pj}(\Lambda, C_1, \dots, C_k)K_z \end{bmatrix} = \begin{bmatrix} h_{1j} \circ (\mathrm{id} \times \phi)^\vee(\Lambda)K_z \\ \vdots \\ h_{pj} \circ (\mathrm{id} \times \phi)^\vee(\Lambda)K_z \end{bmatrix}$$

$$= h^\vee \circ (\mathrm{id} \times \phi)(\Lambda) \begin{bmatrix} 0 \\ \vdots \\ K_z \\ \vdots \\ 0 \end{bmatrix}.$$

Thus

$$h(\Lambda, C_1, \dots, C_k) = h \circ (\mathrm{id} \times \phi)^\vee(\Lambda).$$

By von Neumann's inequality, since Λ is a contraction,

$$\begin{aligned} \|h(\Lambda, C_1, \dots, C_k)\| &= \|h^\vee \circ (\mathrm{id} \times \phi)(\Lambda)\| \\ &\le \sup_{\mathbb{D}} \|h^\vee \circ (\mathrm{id} \times \phi)(z)\| \\ &\le \sup_{\overline{\mathbb{D}} \times \Gamma_k} \|h^\vee(z, \gamma)\| \\ &= \sup_{\overline{\mathbb{D}} \times \Gamma_k} \|h(z, \gamma)\|. \end{aligned}$$

That is, $\overline{\mathbb{D}} \times \Gamma_k$ is a complete spectral set for $(\Lambda, C_1, \dots, C_k)$.

■

Lemma 6.2.2 *Let $(X_0, \dots, X_k)$ be a commuting $(k+1)$-tuple of operators on a Hilbert space $\mathcal{H}$ such that $\sigma(X_0, \dots, X_k) \subset \overline{\mathbb{D}} \times \Gamma_k$. If $\overline{\mathbb{D}} \times \Gamma_k$ is a complete spectral set for $(X_0, \dots, X_k)$ then there exist Hilbert spaces $\mathcal{H}_-, \mathcal{H}_+$ and a $(k+1)$-tuple of commuting normal operators $(\tilde{X}_0, \dots, \tilde{X}_k)$ on $\mathcal{K} \stackrel{def}{=} \mathcal{H}_- \oplus \mathcal{H} \oplus \mathcal{H}_+$ such that $\sigma(\tilde{X}_0, \dots, \tilde{X}_k) \subset \mathbb{T} \times b\Gamma_k$ and each $\tilde{X}_j$ is expressible by an operator matrix of the form*

$$\tilde{X}_j \sim \begin{bmatrix} * & * & * \\ 0 & X_j & * \\ 0 & 0 & * \end{bmatrix} \tag{6.3}$$

with respect to the orthogonal decomposition $\mathcal{K} = \mathcal{H}_- \oplus \mathcal{H} \oplus \mathcal{H}_+$.

Proof. Suppose that $\overline{\mathbb{D}} \times \Gamma_k$ is a complete spectral set for a commuting $(k+1)$-tuple of operators $(X_0, \dots, X_k)$. That is

$$\|h(X_0, \dots, X_k)\|_{\mathcal{L}(\mathcal{H})} \le \sup_{z \in \overline{\mathbb{D}} \times \Gamma_k} \|h(z)\|. \tag{6.4}$$

for all analytic matrix valued functions h.

Let $\theta : \mathcal{P}_{k+1} \to \mathcal{L}(\mathcal{H})$ be the unital representation of the algebra $\mathcal{P}_{k+1}$ of polynomials in $k+1$ variables defined by $\theta(h) = h(X_0, \dots, X_k)$. Inequality (6.4) states that θ is completely contractive and therefore uniformly continuous. It follows that θ has a completely contractive, uniformly continuous extension to $A(\overline{\mathbb{D}} \times \Gamma_k)$. The space $A(\overline{\mathbb{D}} \times \Gamma_k)$ is embedded in the C^*-algebra $C(\mathbb{T}^{k+1})$ of continuous functions on the $(k+1)$-torus by

$$f \in A(\overline{\mathbb{D}} \times \Gamma_k) \mapsto f \circ (\mathrm{id} \times \pi) \in C(\mathbb{T}^{k+1}).$$

By Theorem 5.1.1, θ extends to a completely contractive unital linear mapping $\Theta : C(\mathbb{T}^{k+1}) \to \mathcal{L}(\mathcal{H})$. By Theorem 5.1.2, there exists a Hilbert space $\mathcal{K}$ and a unital representation $\Phi : C(\mathbb{T}^{k+1}) \to \mathcal{L}(\mathcal{K})$ such that $\mathcal{H} \subset \mathcal{K}$ and

$$\Theta(f) = P_{\mathcal{H}} \Phi(f)|_{\mathcal{H}}$$

for all $f \in C(\mathbb{T}^{k+1})$, where $P_{\mathcal{H}}$ is the orthogonal projection from $\mathcal{K}$ to $\mathcal{H}$. Thus, if $f \in A(\overline{\mathbb{D}} \times \Gamma_k)$ we have

$$f(X_0, \ldots, X_k) = \theta(f) = \Theta(f) = P_{\mathcal{H}} \Phi(f \circ \pi)|_{\mathcal{H}}.$$

Let f_j represent the $(j+1)^{th}$ co-ordinate function, i.e. $f_j(z_0, \ldots, z_k) = z_j, j = 0, \ldots, k$. Then for $(\alpha, \lambda_1, \ldots, \lambda_k) \in \mathbb{D} \times \mathbb{D}^k$,

$$f_j \circ (\mathrm{id} \times \pi)(\alpha, \lambda_1, \ldots, \lambda_k) = \begin{cases} \alpha & \text{if } j = 0 \\ c_j(\lambda) & \text{if } j = 1, \ldots, k \end{cases}.$$

For $j = 1, \ldots, k$, let

$$\tilde{X}_j = \Phi(c_j(\lambda)).$$

Set $\tilde{X}_0 = \Phi(\mathrm{id})$. Then

$$X_j = f_j(X_0, \ldots, X_k) = \theta(f_j) = \Theta(f_j \circ (\mathrm{id} \times \pi)) = P_{\mathcal{H}} \tilde{X}_0|_{\mathcal{H}}.$$

Observe that the operators $\tilde{X}_0, \ldots, \tilde{X}_k$ lie in the commutative *-algebra $\Phi(C(\mathbb{T}^{k+1}))$. The $\tilde{X}_j$ are therefore commuting normal operators with $X_0, \ldots, X_k$ as their compressions to $\mathcal{H}$. We may suppose that the smallest subspace of $\mathcal{K}$ which contains $\mathcal{H}$ and reduces each $\tilde{X}_j$ is $\mathcal{K}$. The joint spectrum of the $(k+1)$-tuple of elements $(\alpha, c_1(\lambda), \ldots, c_k(\lambda))$ in $C(\mathbb{T}^{k+1})$ is the range of this $(k+1)$-tuple of functions in $\mathbb{T}^{k+1}$. The range of these functions is exactly $\mathbb{T} \times b\Gamma_k$. Next we may apply the unital representation Φ to deduce that $\sigma(\tilde{X}_0, \ldots, \tilde{X}_k) \subset \mathbb{T} \times b\Gamma_k$.

We shall construct spaces $\mathcal{H}_-$ and $\mathcal{N}$ which are invariant for each X_j and satisfy $\mathcal{H}_- \subset \mathcal{N} \subset \mathcal{K}$ and $\mathcal{H} = \mathcal{N} \ominus \mathcal{H}_-$. Let

$$\mathcal{H}_- = \{x \in \mathcal{K} \mid f(\tilde{X}_0, \ldots, \tilde{X}_k)x \perp \mathcal{H} \text{ for every polynomial } f\}.$$

Clearly $\mathcal{H}_-$ is invariant for each $X_j, j = 0, \ldots, k$, and $\mathcal{H}_- \perp \mathcal{H}$. Let $\mathcal{N} = \mathcal{H}_- \oplus \mathcal{H}$. Then $\mathcal{N}$ is also invariant for each X_j. We can prove this by considering any element $x \in \mathcal{H}$ and showing that, for each

$j \in \{0, \ldots, k\}$,

$$\tilde{X}_j x - P_{\mathcal{H}} \tilde{X}_j x \in \mathcal{H}_-.$$

Suppose f is any polynomial in k variables, then

$$\begin{aligned} P_{\mathcal{H}}\{f(\tilde{X}_0, \ldots, \tilde{X}_k)(\tilde{X}_j x - P_{\mathcal{H}}\tilde{X}_j x)\} &= P_{\mathcal{H}} f(\tilde{X}_0, \ldots, \tilde{X}_k)\tilde{X}_j x - f(\tilde{X}_0, \ldots, \tilde{X}_k) P_{\mathcal{H}} \tilde{X}_j x \\ &= f(X_0, \ldots, X_k) X_j x - f(X_0, \ldots, X_k) X_j x \\ &= 0. \end{aligned}$$

Thus, $\tilde{X}_j x - P_H \tilde{X}_j x \in \mathcal{H}_-$ and therefore $\mathcal{N}$ is invariant for each X_j as claimed. Clearly, X_j has the form (6.3) with respect to the decomposition $\mathcal{K} = \mathcal{H}_- \oplus \mathcal{H} \oplus \mathcal{N}^\perp$ so we may take $\mathcal{H}_+ = \mathcal{N}^\perp$ to see that the result holds.

■

As in the previous chapter, we now present a result which allows us to verify the positivity of an operator polynomial by reducing our calculations to the scalar valued case.

Lemma 6.2.3 *Let $(X_0, \ldots, X_k)$ be a commuting $(k+1)$-tuple of operators on a Hilbert space H such that $\sigma(X_0, \ldots, X_k) \subset \overline{\mathbb{D}} \times \Gamma_k$. Choose $\alpha \in \mathbb{D}$. If $\overline{\mathbb{D}} \times \Gamma_k$ is a complete spectral set for $(X_0, \ldots, X_k)$ and*

$$\mu_k(\beta, \alpha c_1, \ldots, \alpha^k c_k) \geq 0 \tag{6.5}$$

for all $\beta \in \mathbb{T}$ and $(c_1, \ldots, c_k) \in b\Gamma_k$ then

$$\mu_k(X_0, \alpha X_1, \ldots, \alpha^k X_k) \geq 0.$$

Proof. Define $\mathcal{K}$ and $\tilde{X}_j$ as in Lemma 6.2.2. Since the $\tilde{X}_j$ are normal and commute it follows that they generate a commutative C^*-subalgebra $\mathcal{A}$ of $\mathcal{L}(\mathcal{K})$. The space $C(\sigma(\tilde{X}_1, \ldots, \tilde{X}_k))$ can be identified with $\mathcal{A}$ by the Gelfand transform. By this transform, $\tilde{X}_j$ can be identified with the $(j+1)^{th}$ coordinate function. Suppose (6.5) holds. An operator is positive semi-definite if and only if its Gelfand transformation is non-negative, thus on application of an inverse Gelfand transform we have

$$\mu_k(\tilde{X}_0, \alpha\tilde{X}_1, \ldots, \alpha^k \tilde{X}_k) \geq 0. \tag{6.6}$$

To simplify the following calculations we introduce the operators:

$$\tilde{N} = \sum_{r=0}^{k} (-1)^r (k-r) \alpha^r \tilde{X_0}^r \tilde{X}_r,$$
$$N = \sum_{r=0}^{k} (-1)^r (k-r) \alpha^r X_0^r X_r,$$
$$\tilde{M} = \sum_{r=0}^{k} (-1)^r r \alpha^r \tilde{X_0}^{r-1} \tilde{X}_r,$$
$$M = \sum_{r=0}^{k} (-1)^r r \alpha^r X_0^{r-1} X_r.$$

By the definition of $\tilde{X}_j$ we see that

$$\tilde{N} \sim \begin{bmatrix} * & * & * \\ 0 & N & * \\ 0 & 0 & * \end{bmatrix}, \qquad \tilde{M} \sim \begin{bmatrix} * & * & * \\ 0 & M & * \\ 0 & 0 & * \end{bmatrix}. \tag{6.7}$$

with respect to the orthogonal decomposition $\mathcal{K} = \mathcal{H}_- \oplus \mathcal{H} \oplus \mathcal{H}_+$. Using this new notation, (6.6) reads

$$\tilde{N}^* \tilde{N} - \tilde{M}^* \tilde{M} \geq 0. \tag{6.8}$$

Recall Definition 17 and consider $\sigma(\tilde{N})$.

$$\begin{aligned} \sigma(\tilde{N}) &= \sigma\left(\sum_{r=0}^{k} (-1)^r (k-r) \alpha^r \beta^r \tilde{X}_r \right) \\ &\subset \left\{ \sum_{r=0}^{k} (-1)^r (k-r) \alpha^r \beta^r c_r(\lambda) \mid \lambda \in \mathbb{T}^k \right\} \\ &\subset \left\{ \sum_{r=0}^{k} (-1)^r (k-r) c_r(\alpha\beta\lambda) \mid \lambda \in \overline{\mathbb{D}}^k \right\} \\ &= \{ A_{k,\alpha\beta}(\lambda) \mid \lambda \in \overline{\mathbb{D}}^k \}. \end{aligned}$$

Corollary 4.2.3 implies that $0 \notin \sigma(\tilde{N})$. It follows that $\tilde{N}$ is invertible. Rearrange (6.8) to give

$$1 - \tilde{N}^{*-1} \tilde{M}^* \tilde{M} \tilde{N}^{-1} \geq 0.$$

Therefore

$$\|\tilde{M} \tilde{N}^{-1}\| \leq 1.$$

Hence, by (6.7), N must be invertible and

$$\tilde{N}^{-1} \sim \begin{bmatrix} * & * & * \\ 0 & N^{-1} & * \\ 0 & 0 & * \end{bmatrix}.$$

Hence

$$\tilde{M} \tilde{N}^{-1} \sim \begin{bmatrix} * & * & * \\ 0 & MN^{-1} & * \\ 0 & 0 & * \end{bmatrix}.$$

Thus, MN^{-1} is the compression to $\mathcal{H}$ of $\tilde{M}\tilde{N}^{-1}$, and so

$$\|MN^{-1}\| \leq 1.$$

That is

$$\begin{aligned} 1 - N^{*-1}M^*MN^{-1} &\geq 0 \\ N^*N - M^*M &\geq 0 \\ \mu_k(X_0, \alpha X_1, \ldots \alpha^k X_k) &\geq 0 \end{aligned}$$

as required.

■

Lemma 6.2.4 *If $\overline{\mathbb{D}} \times \Gamma_k$ is a complete spectral set for $(X_0, \alpha X_1, \ldots, \alpha^k X_k)$ then it is also a complete spectral set for $(\nu(X_0), \alpha X_1, \ldots, \alpha^k X_k)$ where ν is any automorphism of $\overline{\mathbb{D}}$.*

Proof. If f is an analytic function on $\overline{\mathbb{D}} \times \Gamma_k$ then so is f_ν defined by

$$f_\nu(x_0, \alpha x_1, \ldots, \alpha^k x_k) = f(\nu(x_0), \alpha x_1, \ldots, \alpha^k x_k)$$

Thus

$$\begin{aligned} \|f(\nu(X_0), \alpha X_1, \ldots, \alpha^k X_k)\| &= \|f_\nu(X_0, \alpha X_1, \ldots, \alpha^k X_k)\| \\ &\leq \sup_{\overline{\mathbb{D}} \times \Gamma_k} |f_\nu(z, \gamma)| \\ &= \sup_{\overline{\mathbb{D}} \times \Gamma_k} |f(z, \gamma)| \end{aligned}$$

as required.

■

Theorem 6.2.5 *If $\overline{\mathbb{D}} \times \Gamma_k$ is a complete spectral set for a $(k+1)$-tuple of operators $(X_0, \ldots, X_k)$ and ν is any automorphism of the disc then*

$$\mu_k(\nu(X_0), \alpha X_1, \ldots, \alpha^k X_k) \geq 0 \tag{6.9}$$

for all $\alpha \in \overline{\mathbb{D}}$.

Proof. Suppose $\overline{\mathbb{D}} \times \Gamma_k$ is a complete spectral set for $(X_0, \ldots, X_k)$. Lemma 6.2.4 states that $\overline{\mathbb{D}} \times \Gamma_k$ is a complete spectral set for $(\nu(X_0), \ldots, X_k)$, thus it will suffice to show

$$\mu_k(X_0, \alpha X_1, \ldots, \alpha^k X_k) \geq 0$$

for all $\alpha \in \overline{\mathbb{D}}$. By Lemma 6.2.3, this will follow if

$$\mu_k(\beta, \alpha c_1(\lambda), \ldots, \alpha^k c_k(\lambda)) \geq 0$$

for all $\beta \in \mathbb{T}$, $\lambda \in \mathbb{T}^k$, and $\alpha \in \overline{\mathbb{D}}$. Recall Theorem 3.3.3 and the definition of $p_{k,\alpha}$ given in (3.1). For $\beta \in \mathbb{T}$, $\lambda \in \mathbb{T}^k$, and $\alpha \in \mathbb{D}$ we have

$$\begin{aligned}
\mu_k(\beta, \alpha c_1(\lambda), \ldots, &\alpha^k c_k(\lambda)) \\
&= \frac{1}{k}\left(\sum_{r=0}^{k}(-1)^r(k-r)\alpha^r\beta^r c_r(\lambda)\right)^* \left(\sum_{r=0}^{k}(-1)^r(k-r)\alpha^r\beta^r c_r(\lambda)\right) \\
&\qquad - \frac{1}{k}\left(\sum_{r=0}^{k}(-1)^r r\alpha^r\beta^{r-1} r c_r(\lambda)\right)^* \left(\sum_{r=0}^{k}(-1)^r r\alpha^r\beta^{r-1} r c_r(\lambda)\right) \\
&= \frac{1}{k}\left(\sum_{r=0}^{k}(-1)^r(k-r)\alpha^r\beta^r c_r(\lambda)\right)^* \left(\sum_{r=0}^{k}(-1)^r(k-r)\alpha^r\beta^r c_r(\lambda)\right) \\
&\qquad - \frac{1}{k}\beta\overline{\beta}\left(\sum_{r=0}^{k}(-1)^r r\alpha^r\beta^{r-1} c_r(\lambda)\right)^* \left(\sum_{r=0}^{k}(-1)^r r\alpha^r\beta^{r-1} c_r(\lambda)\right) \\
&= \frac{1}{k}\left(\sum_{r=0}^{k}(-1)^r(k-r)\alpha^r\beta^r c_r(\lambda)\right)^* \left(\sum_{r=0}^{k}(-1)^r(k-r)\alpha^r\beta^r c_r(\lambda)\right) \\
&\qquad - \frac{1}{k}\left(\sum_{r=0}^{k}(-1)^r r\alpha^r\beta^r c_r(\lambda)\right)^* \left(\sum_{r=0}^{k}(-1)^r r\alpha^r\beta^r c_r(\lambda)\right) \\
&= \frac{1}{k}\left[\overline{A_{k,\alpha\beta}}(\lambda)A_{k,\alpha\beta}(\lambda) - \overline{B_{k,\alpha\beta}}(\lambda)B_{k,\alpha\beta}(\lambda)\right] \\
&= p_{k,\alpha\beta}(\lambda, \overline{\lambda}) \\
&= (1-|\alpha\beta|^2)\sum_{j=1}^{k}\prod_{i\neq j}|1-\alpha\beta\lambda_i|^2 \\
&\geq 0.
\end{aligned}$$

as required. Hence (6.9) holds.

■

Corollary 6.2.6 *If there exists an analytic function $\phi : \mathbb{D} \to \Gamma_k$ such that $\phi(z_j) = (c_1^{(j)}, \ldots, c_k^{(j)})$ then*

$$\mu_k(\nu(\Lambda), \alpha C_1, \ldots, \alpha^k C_k) \geq 0$$

for all automorphisms of the disc ν, and all $\alpha \in \overline{\mathbb{D}}$.

Proof. Theorem 6.2.1 states that if there exists an analytic function $\phi : \mathbb{D} \to \Gamma_k$ such that $\phi(z_j) = (c_1^{(j)}, \ldots, c_k^{(j)})$ then $\overline{\mathbb{D}} \times \Gamma_k$ is a complete spectral set for the $(k+1)$-tuple of operators $(\Lambda, C_1, \ldots, C_k)$. Theorem 6.2.5 states that if $\overline{\mathbb{D}} \times \Gamma_k$ is a complete spectral set for the $(k+1)$-tuple of operators $(\Lambda, C_1, \ldots, C_k)$

then

$$\mu_k(\nu(\Lambda), \alpha C_1, \dots, \alpha^k C_k) \geq 0$$

for all automorphisms of the disc ν and all $\alpha \in \mathbb{D}$.

$\blacksquare$

Just as in Chapter 5, this result can be transformed into a necessary condition for the existence of a solution to a certain Nevanlinna-Pick problem in terms of the positivity of certain Pick matrices.

Corollary 6.2.7 *Let $n \in \mathbb{N}$. Choose n distinct points z_j in $\mathbb{D}$ and n matrices W_j in $\mathbb{M}_k(\mathbb{C})$. If there exists an analytic function $\phi : \mathbb{D} \to \mathbb{M}_k(\mathbb{C})$ such that $\phi(z_j) = W_j$ and $\sigma(\phi(z)) \subset \overline{\mathbb{D}}$ for all $z \in \mathbb{D}$ then*

$$\left[\frac{\displaystyle\sum_{r,s=0}^{k} (-1)^{r+s} \overline{c_s(\alpha W_j)\nu(z_j)}^{s-1} [\overline{\nu(z_j)}(k^2 - (r+s)k + rs)\nu(z_i) - rs]\nu(z_i)^{r-1} c_r(\alpha W_i)}{1 - \overline{z_j} z_i} \right]_{i,j=1}^{n}$$

is positive semi-definite for every $\alpha \in \overline{\mathbb{D}}$ and all automorphisms of the disc ν.

Proof. By multiplying out (6.1) we have,

$$\begin{aligned} &\mu_k(x_0, \dots, x_k; \bar{x_0}, \dots, \bar{x_k}) \\ &\qquad = \frac{1}{k} \sum_{r,s=0}^{k} (-1)^{r+s} \bar{x_s} \bar{x_0}^{s-1} [\overline{x_0}(k^2 - (r+s)k + rs)x_0 - rs] x_0^{r-1} x_r. \end{aligned} \tag{6.10}$$

Let $a(W_j) = (c_1^{(j)}, \dots, c_k^{(j)})$ and define C_j by (2.1). Let Λ be given by (6.2). By Corollary 6.2.6, if there exists an analytic function ϕ such that $\phi(z_j) = W_j$ and $\sigma(\phi(z_)) \subset \overline{\mathbb{D}}$ for all $z \in \mathbb{D}$ then for all $\alpha \in \overline{\mathbb{D}}$ and all automorphisms ν of the disc,

$$\mu_k(\nu(\Lambda), \alpha C_1, \dots, \alpha^k C_k) \geq 0. \tag{6.11}$$

For $\alpha \in \mathbb{D}$ and any automorphisms ν of the disc, (6.11) is equivalent to

$$\left[\langle \mu_k(\nu(\Lambda), \alpha C_1, \dots, \alpha^k C_k) K_{z_i}, K_{z_j} \rangle \right]_{i,j=1}^{n} \geq 0. \tag{6.12}$$

Using the expansion given in (6.10) we see that

$$k \langle \mu_k(\nu(\Lambda), \alpha C_1, \dots, \alpha^k C_k) K_{z_i}, K_{z_j} \rangle$$

is equal to

$$\left\langle \sum_{r,s=0}^{k} (-1)^{r+s} \bar{\alpha}^s C_s^* \nu(\Lambda)^{*s-1} [\nu(\Lambda)^*(k^2 - (r+s)k + rs)\nu(\Lambda) - rs]\nu(\Lambda)^{r-1} \alpha^r C_r K_{z_i}, K_{z_j} \right\rangle$$

Using an identical method to Corollary 5.2.2 we see that this is equal to

$$\sum_{r,s=0}^{k}(-1)^{r+s}\overline{\bar{\alpha}c_s(W_j)\nu(z_j)}^{s-1}[\nu(z_j)(k^2-(r+s)k+rs)\nu(z_i)-rs]\nu(z_i)^{r-1}c_r(\bar{\alpha}W_i)K_{z_i}(z_j)$$

which is the same as

$$\frac{\sum_{r,s=0}^{k}(-1)^{r+s}\overline{c_s(\bar{\alpha}W_j)\nu(z_j)}^{s-1}[\nu(z_j)(k^2-(r+s)k+rs)\nu(z_i)-rs]\nu(z_i)^{r-1}c_r(\bar{\alpha}W_i)}{1-\overline{z_j}z_i}. \tag{6.13}$$

Substituting this last expression into (6.12) and observing that

$$\mu_k(\nu(\Lambda),\alpha C_1,\ldots,\alpha^k C_k)\geq 0$$

if and only if

$$k\mu_k(\nu(\Lambda),\alpha C_1,\ldots,\alpha^k C_k)\geq 0$$

yields the result.

■

Corollary 6.2.7 gives us a second necessary condition for the existence of an interpolating function from $\mathbb{D}$ to $\mathbb{M}_k(\mathbb{C})$. This new neccessary condition implies the one proved earlier in Corollary 5.2.2 by choosing ν as the identity. It is not clear however whether the two conditions are equivalent. It is true that $(C_1,\ldots,C_k)$ being a complete Γ_k-contraction is genuinely weaker than $(\Lambda, C_1,\ldots,C_k)$ being a complete $\overline{\mathbb{D}}\times\Gamma_k$-contraction, and this might lead one to suspect that the resultant necessary conditions in terms of polynomials would also be different. Whether this suspicion is true remains an open question.

Chapter 8 contains a discussion of the relative strengths of the two necessary conditions we have presented.

Chapter 7

The Caratheodory Distance on the Symmetrized Polydisc

In this chapter we show how the main theorem of Chapter 5 gives rise to an interesting result concerning the geometry of Γ_k. We derive an upper bound for the Caratheodory distance between two points in Γ_k. Caratheodory introduced his notion of distance in [18] and [19]. Essentially Caratheodory says that we can define a distance between two points in some domain by considering all analytic maps from that domain into the disc and asking how far apart (in the sense of the pseudohyberbolic distance) the images of the points can be. Parallels can clearly be drawn with the distance of Kobayashi [30] in which the roles of the disc and the domain are basically reversed. Kobayashi defines the distance between two points in some domain as the minimum distance between two points in the disc which can be mapped to the points of interest by an analytic function from the disc to the domain.

The method employed in this chapter is inspired by the work of Agler in [2]. Agler and Young used similar methods in [3] to find an explicit formula for the Caratheodory distance on Γ_2. The technical results we need to achieve the goals of this chapter are presented in Section 7.1 while the derivation of an upper bound for the Caratheodory distance is contained in Section 7.2.

The upper bound we present is an infimum of a certain function. It is presented here as an infimum over $\mathbb{T}$. However, without the technical results of Chapter 4 the infimum would have to be taken over the whole disc.

7.1 Spectral Sets

The mathematical objects used in Section 7.2 differ slightly from those used in previous chapters. The main differences occur because we consider spectral sets of operators acting on two dimensional Hilbert spaces rather than *complete* spectral sets of operators acting on Hilbert spaces of arbitrary dimension. We state (without proof) a result of Agler's which, under special circumstances, will allow us to identify complete spectral sets with spectral sets.

Definition 22 *A set $E \subset \mathbb{C}^k$ is said to be a* spectral set *for a commuting k-tuple of operators $(X_1, \dots, X_k)$ on some Hilbert space, if $\sigma(X_1, \dots, X_k) \subset E$ and if, for all scalar-valued functions f of k variables which are analytic on $\overline{E}$, we have*

$$\|f(X_1, \dots, X_k)\| \leq \sup_{(x_1, \dots, x_k) \in E} |f(x_1, \dots, x_k)|.$$

Notice that the concept of a spectral set is tautologically weaker than that of a complete spectral set. Just as we would describe a commuting k-tuple of operators with E as a complete spectral set as a *complete E-contraction* we will describe k-tuples of operators with E as a spectral set as *E-contractions.* An active area of research in Operator Theory is to establish which sets have the property that they are a complete spectral set for every k-tuple of operators which have them as a spectral set. The following result of Agler's ([2, Proposition 3.5]) solves this problem in a special case of interest to us.

Theorem 7.1.1 *Let U be a bounded set in $\mathbb{C}^k$. Assume that $z_1, z_2 \in U$, $z_1 \neq z_2$ and $\sigma(X_1, \dots, X_k) \subset \{z_1, z_2\}$. Then U is a spectral set for $(X_1, \dots, X_k)$ if and only if U is a complete spectral set for $(X_1, \dots, X_k)$.*

Corollary 7.1.2 *If $(X_1, \dots, X_k)$ is a commuting k-tuple of operators on a two dimensional Hilbert space then $(X_1, \dots, X_k)$ is a Γ_k-contraction if and only if it is a complete Γ_k-contraction.*

Proof. The result follows from Theorem 7.1.1 since all reasonable forms of the joint spectrum of a k-tuple of two dimensional operators are equal to the algebraic joint spectrum of those operators, which is clearly a two point set. ∎

We state the following Theorem without proof since the equivalences hold trivially.

Theorem 7.1.3 *Let $(X_1, \dots, X_k)$ be a commuting k-tuple of operators with joint spectrum contained in $\operatorname{int}\Gamma_k$. The following statements are equivalent*

(a) Γ_k *is a spectral set for* $(X_1, \dots, X_k)$,

(b) $\operatorname{int}\Gamma_k$ *is a spectral set for* $(X_1, \dots, X_k)$,

(c) $\|f(X_1, \dots, X_k)\| \leq 1$ *for all analytic* $f : \operatorname{int}\Gamma_k \to \mathbb{D}$.

7.2 Caratheodory Distance

Let $\mathcal{N}$ denote the set of analytic functions $F : \operatorname{int}\Gamma_k \to \mathbb{D}$. The Caratheodory distance D_k on $\operatorname{int}\Gamma_k$ is defined as follows. Let $z_1 = (c_1^{(1)}, \ldots, c_k^{(1)})$ and $z_2 = (c_1^{(2)}, \ldots, c_k^{(2)})$ be distinct points in $\operatorname{int}\Gamma_k$. Then

$$D_k(z_1, z_2) = \sup_{F\in\mathcal{N}} \tanh^{-1} \left| \frac{F(z_1) - F(z_2)}{1 - \overline{F(z_2)}F(z_1)} \right|.$$

To simplify notation, we use $d(\cdot,\cdot)$ to represent the pseudohyperbolic distance on the disc. That is, for $\lambda_1, \lambda_2 \in \mathbb{D}$,

$$d(\lambda_1, \lambda_2) = \left| \frac{\lambda_1 - \lambda_2}{1 - \overline{\lambda_2}\lambda_1} \right|.$$

Thus,

$$D_k(z_1, z_2) = \sup_{F\in\mathcal{N}} \tanh^{-1} d(F(z_1), F(z_2)).$$

We wish to express the Caratheodory distance between $z_1 = (c_1^{(1)}, \ldots, c_k^{(1)})$ and $z_2 = (c_1^{(2)}, \ldots, c_k^{(2)})$ in terms of operators on certain Hilbert spaces. Let H be a two dimensional Hilbert space with inner product $\langle \cdot,\cdot \rangle$ and define

$$\mathcal{U}(H) = \{u = (u_1, u_2) \mid u \text{ is a basis of unit vectors for } H\}.$$

For $u \in \mathcal{U}(H)$ define the commuting k-tuple of operators $(C_{1u}, \ldots, C_{ku})$ on H by

$$C_{ju} \sim \operatorname{diag}\{c_j^{(1)}, c_j^{(2)}\} \tag{7.1}$$

with respect to the basis u. We wish to show

$$\sigma(C_{1u}, \ldots, C_{ku}) = \{z_1, z_2\} \subset \operatorname{int}\Gamma_k. \tag{7.2}$$

We begin by proving $z_1 \in \sigma(C_{1u}, \ldots, C_{ku})$. We have

$$c_j^{(1)} - C_{ju} \sim \operatorname{diag}\{0, c_j^{(1)} - c_j^{(2)}\}$$

for $j = 1, \ldots, k$. Hence,

$$(c_j^{(1)} - C_{ju}) \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 0$$

for $j = 1, \ldots, k$ and therefore

$$\sum_{j=1}^{k} A_j (c_j^{(1)} - C_{ju}) \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 0$$

for all A_j in the *-algebra generated by $(C_{1u}, \ldots, C_{ku})$. It follows that the identity is not contained in the ideal generated by $c_j^{(1)} - C_{ju}$ for $j = 1, \ldots, k$. Hence this ideal is proper and $z_1 \in \sigma(C_{1u}, \ldots, C_{ku})$. Similarly $z_2 \in \sigma(C_{1u}, \ldots, C_{ku})$.

Suppose $z = (c_1, \ldots, c_k) \notin \{z_1, z_2\}$. There exist (not necessarily distinct) integers i and j with $0 < i, j \leq k$ such that

$$c_i \neq c_i^{(1)}, \qquad c_j \neq c_j^{(2)}.$$

Therefore

$$c_i - C_{iu} \sim \begin{bmatrix} * & 0 \\ 0 & c_i - c_i^{(2)} \end{bmatrix}, \qquad c_j - C_{ju} \sim \begin{bmatrix} c_j - c_j^{(1)} & 0 \\ 0 & * \end{bmatrix}.$$

We can therefore find scalars $\alpha,\ \beta$ such that

$$\alpha(c_i - C_{iu}) + \beta(c_j - C_{ju}) = I.$$

The ideal generated by $c_j - C_{ju}$ for $j = 1, \ldots, k$ is therefore not proper, and $z \notin \sigma(C_{1u}, \ldots, C_{ku})$. This completes the proof of (7.2). Finally, we introduce the set $\mathcal{U}'(H)$,

$$\mathcal{U}'(H) = \{u \in \mathcal{U}(H) \mid (C_{1u}, \ldots, C_{ku}) \text{ is a } \Gamma_k \text{ contraction}\}.$$

As a first step to interpreting the Caratheodory distance in terms of the Hilbert space H we have the following Lemma which relates the pseudohyperbolic distance on $\mathbb{D}$ to $\mathcal{U}(H)$.

Lemma 7.2.1 *If $F \in \mathcal{N}, u \in \mathcal{U}(H)$ then*

$$\|F(C_{1u}, \ldots, C_{ku})\| \leq 1 \tag{7.3}$$

if and only if

$$|\langle u_1, u_2\rangle|^2 \leq 1 - d(F(z_1), F(z_2))^2 \tag{7.4}$$

Moreover, (7.3) *holds with equality if and only if* (7.4) *holds with equality.*

Proof. Consider $F(C_{1u}, \ldots, C_{ku})$. Clearly,

$$\|F(C_{1u}, \ldots, C_{ku})\| \leq 1$$

if and only if

$$1 - F(C_{1u}, \ldots, C_{ku})^* F(C_{1u}, \ldots, C_{ku}) \geq 0.$$

Since $u \in \mathcal{U}(H)$ this is equivalent to

$$[\langle (1 - F(C_{1u}, \ldots, C_{ku})^* F(C_{1u}, \ldots, C_{ku}))u_j, u_i\rangle]_{i,j=1}^2 \geq 0. \tag{7.5}$$

With respect to the basis u we have

$$F(C_{1u}, \ldots, C_{ku}) \sim \operatorname{diag}\{F(c_1^{(1)}, \ldots c_k^{(1)}), F(c_1^{(2)}, \ldots c_k^{(2)})\} = \operatorname{diag}\{F(z_1), F(z_2)\}.$$

Thus (7.5) is equivalent to

$$[\langle u_j, u_i\rangle - \langle F(C_{1u},\dots,C_{ku})u_j, F(C_{1u},\dots,C_{ku})u_i\rangle]_{i,j=1}^2 \geq 0,$$

which is

$$\begin{bmatrix} 1-|F(z_1)|^2 & (1-\overline{F(z_1)}F(z_2))\langle u_2,u_1\rangle \\ (1-\overline{F(z_2)}F(z_1))\langle u_1,u_2\rangle & 1-|F(z_2)|^2 \end{bmatrix} \geq 0.$$

Taking determinants, we see this is equivalent to

$$(1-|F(z_1)|^2)(1-|F(z_2)|^2) - (1-\overline{F(z_1)}F(z_2))\langle u_2,u_1\rangle(1-\overline{F(z_2)}F(z_1))\langle u_1,u_2\rangle \geq 0.$$

Thus (7.3) holds if and only if

$$\begin{aligned} |\langle u_1,u_2\rangle|^2 &\leq \frac{(1-|F(z_1)|^2)(1-|F(z_2)|^2)}{|1-\overline{F(z_2)}F(z_1)|} \\ &= 1 - d(F(z_1),F(z_2))^2. \end{aligned}$$

This proves the first part of the result. If (7.3) holds with equality then each of the equivalent statements above will hold with equality. This completes the proof.

■

We can now represent $D_k(z_1,z_2)$ in terms of the Hilbert space H.

Lemma 7.2.2 *Let $z_j = (c_1^{(j)},\dots,c_k^{(j)}) \in \operatorname{int}\Gamma_k$ for $j=1,2$. Suppose H is a two dimensional Hilbert space and let $u \in \mathcal{U}(H)$. If, for $i=1,\dots,k$, C_{iu} is given by (7.1) then Γ_k is a spectral set for $(C_{1u},\dots,C_{ku})$ if and only if*

$$|\langle u_1,u_2\rangle| \leq \operatorname{sech} D_k(z_1,z_2).$$

Furthermore,

$$\operatorname{sech} D_k(z_1,z_2) = \sup_{u\in\mathcal{U}'(H)} |\langle u_1,u_2\rangle|.$$

Proof. Γ_k is a spectral set for $(C_{1u},\dots,C_{ku})$ if and only if

$$\|F(C_{1u},\dots,C_{ku})\| \leq 1$$

for all analytic functions $F : \operatorname{int}\Gamma_k \to \mathbb{D}$. By Lemma 7.2.1 it follows that $(C_{1u},\dots,C_{ku})$ is a Γ_k contraction if and only if

$$\begin{aligned} |\langle u_1,u_2\rangle|^2 &\leq 1 - \sup_{F\in\mathcal{N}} d(F(z_1),F(z_2))^2 \\ &= 1-\tanh^2 D_k(z_1,z_2) \\ &= \operatorname{sech}^2 D_k(z_1,z_2). \end{aligned}$$

This completes the first part of the result.

Since $\langle u_1, u_2\rangle$ may be interpreted as the cosine of the acute angle between the two unit vectors u_1 and u_2 and $0 \le \operatorname{sech} x \le 1$ for all $x \in \mathbb{R}$ it is possible to choose $u' = (u_1', u_2') \in \mathcal{U}(H)$ such that

$$|\langle u_1', u_2'\rangle| = \operatorname{sech}^2 D_k(z_1, z_2).$$

It then follows from above that $(C_{1u}, \dots, C_{ku})$ is a Γ_k-contraction. Therefore

$$\sup_{u \in \mathcal{U}'(H)} |\langle u_1, u_2\rangle| \ge \operatorname{sech}^2 D_k(z_1, z_2).$$

It is clear from (7.3) however that the opposite inequality holds and we may therefore conclude that

$$\sup_{u \in \mathcal{U}'(H)} |\langle u_1, u_2\rangle| = \operatorname{sech}^2 D_k(z_1, z_2)$$

as required.

■

Recall the definition of P_k in (1.6). For $\omega \in \mathbb{T}$ and $x = (x_1, \dots, x_k),\ y = (y_1, \dots, y_k)$ define the hereditary polynomial $v_{k,\omega}$ by

$$v_{k,\omega}(x, y) = P_k(1, \omega x_1, \dots, \omega^k x_k; 1, \overline{\omega} y_1, \dots, \overline{\omega^k} y_k).$$

In Theorem 7.2.5 we prove a specialisation of Theorem 5.1.5. The specialisation relies on the properties of $v_{k,\omega}$ given in the following Lemma.

Lemma 7.2.3 *If $\omega \in \mathbb{T}$ and $x, y \in \operatorname{int}\Gamma_k$ then*

$$v_{k,\omega}(x, \overline{x}) > 0,$$

and

$$v_{k,\omega}(x, \overline{y}) \ne 0.$$

Proof. Since $x \in \operatorname{int}\Gamma_k$ there exists $\lambda = (\lambda_1, \dots, \lambda_k) \in \mathbb{D}$ such that $x = (c_1(\lambda), \dots, c_k(\lambda))$. Thus,

$$\begin{aligned} v_{k,\omega}(x, \overline{x}) &= P_k(1, \omega c_1(\lambda), \dots, \omega^k c_k(\lambda); 1, \overline{\omega c_1(\lambda)}, \dots, \overline{\omega^k c_k(\lambda)}) \\ &= P_k(1, c_1(\omega\lambda), \dots, c_k(\omega\lambda); 1, c_1(\overline{\omega\lambda}), \dots, c_k(\overline{\omega\lambda})). \end{aligned}$$

Then, by Corollary 3.3.2, we have

$$v_{k,\omega}(x, \overline{x}) = p_{k,\omega}(\lambda, \overline{\lambda}).$$

That is, by (3.1),

$$v_{k,\omega}(x, \overline{x}) = \sum_{j=1}^{k} \left((1 - |\lambda_j|^2) \prod_{i \ne j} |1 - \omega\lambda|^2 \right) > 0.$$

This completes the proof of the first statement in the result. If $y \in \operatorname{int}\Gamma_k$ then there exists $\delta \in \mathbb{D}^k$ such that $y = (c_1(\delta), \ldots, c_k(\delta))$. By Theorem 3.3.3,

$$v_{k,\omega}(x, \overline{x}) = \frac{1}{k}|A_{k,\omega}(\lambda)|^2 - \frac{1}{k}|B_{k,\omega}(\lambda)|^2.$$

Furthermore, by Corollary 4.2.3 we have $A_{k,\omega}(\lambda) \neq 0$. Thus $v_{k,\omega}(x, \overline{x}) > 0$ if and only if

$$1 > \left| \frac{B_{k,\omega}(\lambda)}{A_{k,\omega}(\lambda)} \right|. \tag{7.6}$$

Again by Theorem 3.3.3, we have

$$kv_{k,\omega}(x, \overline{y}) = \overline{A_{k,\omega}(\delta)} A_{k,\omega}(\lambda) - \overline{B_{k,\omega}(\delta)} B_{k,\omega}(\lambda).$$

Therefore $v_{k,\omega}(x, \overline{y}) = 0$ if and only if

$$\overline{A_{k,\omega}(\delta)} A_{k,\omega}(\lambda) - \overline{B_{k,\omega}(\delta)} B_{k,\omega}(\lambda) = 0.$$

That is, if and only if

$$1 - \frac{\overline{B_{k,\omega}(\delta)}}{\overline{A_{k,\omega}(\delta)}} \frac{B_{k,\omega}(\lambda)}{A_{k,\omega}(\lambda)} = 0$$

which contradicts (7.6). Thus $v_{k,\omega}(x, \overline{y}) \neq 0$.

■

Lemma 7.2.4 *Let $(X_1, \ldots, X_k)$ be a k-tuple of commuting operators on a two dimensional Hilbert space such that $\sigma(X_1, \ldots, X_k) \subset \operatorname{int}\Gamma_k$. Then*

$$\rho_k(\omega X_1, \ldots, \omega^k X_k) \geq 0 \tag{7.7}$$

for all $\omega \in \mathbb{T}$ if and only if

$$\rho_k(\alpha X_1, \ldots, \alpha^k X_k) \geq 0$$

for all $\alpha \in \mathbb{D}$.

Proof. ($\Leftarrow$) This follows immediately by taking radial limits.
($\Rightarrow$) Suppose (7.7) holds. Then

$$A_\omega^* A_\omega - B_\omega^* B_\omega \geq 0$$

for all $\omega \in \mathbb{T}$, where

$$A_\omega = \sum_{r=0}^{k} (-1)^r (k-r) \omega^r X_r,$$
$$B_\omega = \sum_{r=0}^{k} (-1)^r r \omega^r X_r.$$

Now, A_ω is invertible since by Corollary 4.2.3 its spectrum does not contain the point zero. Therefore (7.7) implies

$$\|B_\omega A_\omega^{-1}\| \leq 1$$

for all $\omega \in \mathbb{T}$. However, Corollary 4.2.3 also tells us that A_α is invertible for all $\alpha \in \overline{\mathbb{D}}$ and hence the map

$$\alpha \mapsto B_\alpha A_\alpha^{-1}$$

is analytic. By the maximum modulus principle,

$$\|B_\alpha A_\alpha^{-1}\| \leq 1 \tag{7.8}$$

for all $\alpha \in \mathbb{D}$. Inequality (7.8) is equivalent to

$$A_\alpha^* A_\alpha - B_\alpha^* B_\alpha \geq 0$$

for all $\alpha \in \mathbb{D}$, and hence to

$$\rho_k(\alpha X_1, \ldots, \alpha^k X_k) \geq 0$$

for all $\alpha \in \mathbb{D}$.

■

Theorem 7.2.5 *Suppose H is a two dimensional Hilbert space. Let $(X_1, \ldots, X_k)$ be a k-tuple of commuting operators on H such that $\sigma(X_1, \ldots, X_k) \subset \operatorname{int} \Gamma_k$. Then Γ_k is a spectral set for $(X_1, \ldots, X_k)$ only if*

$$\rho_k(\omega X_1, \ldots, \omega^k X_k) \geq 0$$

for all $\omega \in \mathbb{T}$.

Proof. Theorem 7.1.2 states that Γ_k is a spectral set for $(X_1, \ldots, X_k)$ if and only if it is a *complete* spectral set for $(X_1, \ldots, X_k)$. Theorem 5.1.5 states that Γ_k is a complete spectral set for $(X_1, \ldots, X_k)$ only if

$$\rho_k(\alpha X_1, \ldots, \alpha^k X_k) \geq 0$$

for all $\alpha \in \mathbb{D}$. The result then follows from Lemma 7.2.4.

■

The above specialisation of Theorem 5.1.5 gives rise to the following inequality, which we use to establish an upper bound for the Caratheodory distance on $\operatorname{int}\Gamma_k$.

Lemma 7.2.6 *Let H be a two dimensional Hilbert space and let $u = (u_1, u_2) \in \mathcal{U}(H)$. Suppose $z_j = (c_1^{(j)}, \ldots, c_k^{(j)}) \in \operatorname{int}\Gamma_k$ for $j = 1, 2$. For $i = 1, \ldots, k$, let C_{iu} be defined by (7.1). If Γ_k is a spectral set for $(C_{1u}, \ldots, C_{ku})$ then*

$$|\langle u_1, u_2\rangle|^2 \leq \inf_{\omega\in\mathbb{T}} \frac{v_{k,\omega}(z_1, \overline{z_1})v_{k,\omega}(z_2, \overline{z_2})}{|v_{k,\omega}(z_1, \overline{z_2})|^2}.$$

Proof. If Γ_k is a spectral set for $(C_{1u}, \ldots, C_{ku})$ then by Theorem 7.2.5,

$$\rho_k(\omega C_{1u}, \ldots, \omega^k C_{ku}) \geq 0$$

for all $\omega \in \mathbb{T}$. This is equivalent to

$$\left[< \rho_k(\omega C_{1u}, \ldots, \omega^k C_{ku})u_j, u_i >\right]_{i,j=1}^2 \geq 0$$

for all $\omega \in \mathbb{T}$. This in turn is equivalent to the matrix

$$\left[< P_k(1, \omega C_{1u}, \ldots, \omega^k C_{ku}; 1, \overline{\omega} C_{1u}^*, \ldots, \overline{\omega^k} C_{ku}^*)u_j, u_i >\right]_{i,j=1}^2$$

being positive semi-definite for all $\omega \in \mathbb{T}$, which is to say,

$$\left[< P_k(1, \omega c_1^{(j)}, \ldots, \omega^k c_k^{(j)}; 1, \overline{\omega c_1(i)}, \ldots, \overline{\omega^k c_k^{(i)}})u_j, u_i >\right]_{i,j=1}^2 \geq 0$$

for all $\omega \in \mathbb{T}$. Equivalently,

$$\begin{bmatrix} v_{k,\omega}(z_1, \overline{z_1}) & v_{k,\omega}(z_2, \overline{z_1})\langle u_2, u_1\rangle \\ v_{k,\omega}(z_1, \overline{z_2})\langle u_1, u_2\rangle & v_{k,\omega}(z_2, \overline{z_2}) \end{bmatrix} \geq 0$$

for all $\omega \in \mathbb{T}$. This last inequality is equivalent to

$$v_{k,\omega}(z_1, \overline{z_1})v_{k,\omega}(z_2, \overline{z_2}) - |v_{k,\omega}(z_2, \overline{z_1})|^2|\langle u_2, u_1\rangle|^2 \geq 0$$

for all $\omega \in \mathbb{T}$. By Lemma 7.2.3, for all $\omega \in \mathbb{T}$,

$$|\langle u_1, u_2\rangle|^2 \leq \frac{v_{k,\omega}(z_1, \overline{z_1})v_{k,\omega}(z_2, \overline{z_2})}{|v_{k,\omega}(z_2, \overline{z_1})|^2}.$$

Therefore

$$|\langle u_1, u_2\rangle|^2 \leq \inf_{\omega\in\mathbb{T}} \frac{v_{k,\omega}(z_1, \overline{z_1})v_{k,\omega}(z_2, \overline{z_2})}{|v_{k,\omega}(z_2, \overline{z_1})|^2}.$$

$\blacksquare$

We are now in a position to establish an upper bound for the Caratheodory distance between two points in $\operatorname{int}\Gamma_k$.

Theorem 7.2.7 *Let $z_1, z_2 \in \operatorname{int}\Gamma_k$. Then*

$$\operatorname{sech}^2 D_k(z_1, z_2) \le \inf_{\omega \in \mathbb{T}} \frac{v_{k,\omega}(z_1, \overline{z_1}) v_{k,\omega}(z_2, \overline{z_2})}{|v_{k,\omega}(z_2, \overline{z_1})|^2}.$$

Proof. Applying Lemmas 7.2.2 and 7.2.6 in turn we have

$$\begin{aligned}
\operatorname{sech}^2 D_k(z_1, z_2) &= \sup\{|\langle u_1, u_2\rangle|^2 \ : \ u \in \mathcal{U}'(H)\} \\
&\le \sup\left\{ |\langle u_1, u_2\rangle|^2 \ : \ |\langle u_1, u_2\rangle|^2 \le \inf_{\omega \in \mathbb{T}} \frac{v_{k,\omega}(z_1, \overline{z_1}) v_{k,\omega}(z_2, \overline{z_2})}{|v_{k,\omega}(z_1, \overline{z_2})|^2} \right\} \\
&\le \inf_{\omega \in \mathbb{T}} \frac{v_{k,\omega}(z_1, \overline{z_1}) v_{k,\omega}(z_2, \overline{z_2})}{|v_{k,\omega}(z_1, \overline{z_2})|^2}.
\end{aligned}$$

■

Chapter 8

Areas of Further Study

This thesis has touched on a number of different mathematical and engineering areas, including interpolation theory, operator theory, Hilbert spaces, complex geometry, linear systems and control engineering. The Main Problem of this thesis is studied in one form or another by specialists in most of these disciplines. It is beyond the scope of this work to place accurately our results alongside those of our colleagues in different fields, or even to speculate on how future approaches to these problems should be made. Instead, we content ourselves by discussing some questions which we believe arise naturally as a consequence of our work.

We draw comparisons mainly with the work of Agler and Young, not because they are the only authors to have dealt with the problem, but because it is their approach which we adopted to derive our results. If the reader is interested in seeing a different approach to spectral interpolation problems, then we recommend the series of papers by Bercovici, Foias and Tannenbaum [10, 11, 12, 14, 15]. The third of these contains a number of illuminating examples.

A more general introduction to interpolation problems (including their applications) can be found in [33], while [27] is an excellent introduction to the wealth of control engineering literature.

It seems that a natural question to ask, when presented with two necessary conditions for the existence of a solution to the spectral Nevanlinna-Pick problem, is whether either, or both, of these conditions is sufficient. I do not know the answer to this question.

In [6] Agler and Young proved that Γ_2 is a complete spectral set for the commuting pair of operators (C_1, C_2) if and only if

$$\rho_2(\alpha C_1, \alpha C_2) \geq 0 \tag{8.1}$$

for all $\alpha \in \overline{\mathbb{D}}$. Furthermore, the same paper contains a realisation formula for all hereditary polynomials which are positive on Γ_2-contractions. This means, when $k = 2$, we can use the realisation formula to verify whether any particular pair of commuting contractions will satisfy (8.1). Unfortunately, for general k, it is impossible to employ the methods of Agler and Young to establish an equivalent result. The realisation formula for hereditary polynomials which are positive on Γ_2-contractions relies on the fact that a certain Hilbert space can be assumed to be one dimensional. Since we were unable to justify such an assumption, we were unable to present a realisation formula for hereditary polynomials which are positive on Γ_k-contractions. Another difference between this work and that of Agler and Young is that the inequality in (8.1) is sufficient (as well as necessary) for (C_1, C_2) to be a Γ_2-contraction. We prove the necessity of (8.1) for general k in Chapter 5. Agler and Young's proof of sufficiency relies on an application of the Commutant Lifting Theorem, which cannot be applied to more than two commuting contractions. It is therefore impossible to apply Agler and Young's methods in the case of arbitrary k. Of course, just because the same method of proof cannot be employed, it does not follow that an equivalent result is not true for arbitrary k. Indeed, I have been unable to produce a counter-example to such a claim.

The fact that (8.1) is equivalent to (C_1, C_2) being a Γ_2-contraction also allowed Agler and Young [3] to give an exact formula for the Caratheodory distance on $\operatorname{int}\Gamma_2$. The formula they present takes exactly the same form as the upper bound for the Caratheodory distance we present in Chapter 7. The additional information given by an exact formula for the Caratheodory distance enabled Agler and Young to produce Caratheodory extremal functions for Γ_2.

Even if one were able to establish that the k-dimensional equivalent of the operator inequality (8.1) is sufficient for a k-tuple of operators to be a Γ_k-contraction, one would still lack a full sufficient condition for spectral interpolation. The argument presented in Chapter 1 to reduce spectral interpolation to Γ_k interpolation fails to prove the equivalence of the two problems. An identical obstacle faced Agler and Young in their treatment of the two dimensional case. In [5] they were able to show that the two types of interpolation problem are equivalent whenever all or none of the target matrices are scalar multiples of the identity. If the target matrices fail to satisfy this condition, then it has been shown that the spectral Nevanlinna-Pick problem is equivalent to a Γ_k interpolation problem with a condition on the derivative of the interpolating function. It would appear that one could begin to approach the question of whether Γ_k interpolation is equivalent to $k \times k$ spectral interpolation in the same way. Namely, one could base an analysis on a suitable generalization of scalar matrices. The work of Agler and Young suggests that the correct generalization would be to consider whether the target matrices are *derogatory.* That is, whether the target matrices have the property that their minimal and characteristic polynomials coincide. In the case of 2×2 matrices, derogatory and scalar are synonymous.

Simplifying the Γ_k problem by considering only two interpolation points yielded success, in terms of a sufficiency result, in the two dimensional case in [4]. Here, Agler and Young were able to prove a Schwarz Lemma for Γ_2. That is, they were able to show that the condition in (8.1) is equivalent to the existence of a solution to the Γ_2 interpolation problem in the case of two interpolation points when one of the target points is $(0, 0)$. The resultant Pick-matrix form of this result is very much in keeping with the classical Schwarz Lemma, and it serves to demonstrate the increasing difficulty one faces in considering higher dimensional versions of the problem. A Schwarz Lemma of this type seems to be an achievable target for $k > 2$. The results of this thesis provide the necessary condition part of such a result. I believe that a Schwarz Lemma for $k > 2$ would present the greatest chance of success if one were to attempt to extend the various results of Agler and Young from the two dimensional case.

Until now we have been concerned with the sufficiency of the necessary conditions for spectral interpolation presented in Chapters 5 and 6. At this point however, it is not even clear whether these different conditions are equivalent to one another. It is easy to see that the necessary condition given in Chapter 5 is a special case of the condition in Chapter 6. Whether the opposite implication holds is an open question. As discussed in Chapter 6, the condition presented there is derived from the fact that $\overline{\mathbb{D}} \times \Gamma_k$ is a complete spectral set for a $(k+1)$-tuple of operators, whereas the necessary condition in Chapter 5 is based on the genuinely weaker fact that Γ_k is a complete spectral set for a certain k-tuple of operators. Although there is a real difference between the two underlying spectral set conditions, it is unclear whether this difference remains when the dust has cleared and we are presented with the two necessary conditions for spectral interpolation in Pick-matrix form. The question of whether the two necessary conditions presented in this thesis are equivalent is open even in the case $k = 2$. Having attempted in vain to construct a number of examples which would show that the conditions are different, I would not like to guess whether they are or not. Agler and Young are currently pursuing research in this area.

A series of papers by Bercovici et al. has examined the spectral Nevanlinna-Pick problem from a different, but still operator theoretic, perspective. In [12] these authors prove a spectral commutant lifting theorem. They then apply this theorem to the spectral Nevanlinna-Pick problem in much the same way as one would use the classical Commutant Lifting Theorem to prove Pick's theorem (Corollary 1.1.3). The solution to the spectral Nevanlinna-Pick theorem which results from this approach is complete in the sense that the condition given is both necessary and sufficient. However, the condition which Bercovici et al. show to be equivalent to the existence of a solution to the spectral Nevanlinna-Pick problem is rather difficult to apply in general. They, like us, construct a Hilbert Space $\mathcal{H}$ and an operator A from the interpolation information. Their theorem states that there exists a solution to the spectral Nevanlinna-

Pick problem if and only if there exists an invertible operator M which commutes with the shift operator (compressed to $\mathcal{H}$) such that $\|MAM^{-1}\| < 1$. The paper concludes with a section consisting of various examples which demonstrate the subtlety of the spectral Nevanlinna-Pick problem. Included in this section is an example in which each of the target matrices is diagonal, there is a solution, but no diagonal interpolating function exists. Another example illustrates the difference between spectral and classical Nevanlinna-Pick interpolation: it shows that a solution to the spectral version of the problem can have arbirarily large norm.

Bercovici, Foias and Tannenbaum extended the results of [12] in [14] where they considered the tangential spectral Nevanlinna-Pick problem. The method employed in this paper is very similar to that of their previous work. The paper [14] includes an algorithm for constructing optimal solutions (in the sense that their spectral radius is as small as possible) to the tangential version of the problem- a theme which is continued in [13], where the authors describe a property of an optimal solution to the spectral Nevanlinna-Pick problem. They summarise their result by saying that optimal solutions to the spectral Nevanlinna-Pick problem are spectral analogues of inner functions which appear as optimal solutions in dilation theory.

The condition which Bercovici et al. prove equivalent to the existence of a solution to the spectral Nevanlinna-Pick problem is closely related to the structured singular value- or more precisely, an upper bound for the structured singular value. The link between these two concepts is examined further in [15] in which the spectral commutant lifting theorm of [12] is used to prove that the structured singular value is equal to its upper bound under certain conditions. The method of dealing with robust stabilization problems via an upper bound for the structured singular value, rather than the quantity itself, is common amongst engineers.

Doyle and Packard published a comprehensive paper [32] in which they discussed various methods of calculating bounds for the sturctured singular value whilst in [16], Braatz et al. demonstrated why so much effort is expended in dealing with bounds for μ when they showed that in many cases the exact calculation of μ is NP-hard. They politely suggest that attempting to calculate μ directly is therefore futile. A more recent, and more operator theoretic study of the structured singular value was undertaken by Feintuch and Markus [25]. These authors again focus on the closeness of upper bounds for μ, but do so in the more general setting of infinite dimensional Hilbert space.

Whether one is interested in robust stablisation, or spectral interpolation there are a wealth of different approaches currently being pursued. I hope the results of this work will make a significant contribution to the field by throwing light on a hard, concrete special case.

Bibliography

[1] J Agler. On the representation of certain holomorphic functions defined on a polydisc. *Operator Theory, Advances and Applications*, 48:47 – 66, 1990.

[2] J Agler. Operator theory and the Caratheodory metric. *Invent. math.*, 101:483–500, 1990.

[3] J Agler and N J Young. Invariant distances on the symmetrized bidisc. Preprint, University of Newcastle upon Tyne, 1999.

[4] J Agler and N J Young. A Schwarz Lemma for the symmetrized bidisc. To Appear in Bull. London Math. Soc.

[5] J Agler and N J Young. The two-point spectral Nevanlinna-Pick problem. Preprint, University of Newcastle upon Tyne, 1999, to appear in Integral equations and Operator Theory.

[6] J Agler and N J Young. A Commutant Lifting Theorem for a domain in $\mathbb{C}^2$. *J. Func. Anal.*, 161:452–477, 1999.

[7] N Aronszajn. Theory of reproducing kernels and applications. *Trans. AMS*, 68:337 – 404, 1950.

[8] W B Arveson. Subalgebras of C^*-algebras. *Acta Math.*, 123:141 – 224, 1969.

[9] W B Arveson. Subalgebras of C^*-algebras II. *Acta Math.*, 128:271–308, 1972.

[10] H Bercovici, C Foias, P. Khargonekar, and A Tannenbaum. On a lifting theorem for the structured singular value. *J. Math. Anal. Appl.*, 187(2):617–627, 1994.

[11] H. Bercovici, C. Foias, and A. Tannenbaum. Spectral variants of the Nevanlinna-Pick interpolation problem. In *Signal processing, scattering and operator theory, and numerical methods (Amsterdam, 1989)*, pages 23–45. Birkhäuser Boston, Boston, MA, 1990.

[12] H Bercovici, C Foias, and A Tannenbaum. A Spectral Commutant Lifting Theorem. *Trans. A.M.S*, 325:741 – 763, 1991.

[13] H Bercovici, C Foias, and A Tannenbaum. On optimal solutions in spectral commutant lifting theory. *J. Func. Anal.*, 101(4):38–49, 1991.

[14] H Bercovici, C Foias, and A Tannenbaum. On spectral tangential Nevanlinna-Pick interpolation. *J. Math. Anal. Appl.*, 155(1):156–176, 1991.

[15] H Bercovici, C Foias, and A Tannenbaum. The structured singular value for linear input/output operators. *SIAM J. Control Optim.*, 34(4):1392–1404, 1996.

[16] R Braatz, P Young, C Doyle, and M Morari. Computational complexity of μ calculation. *IEEE Trans. Automatic Control*, 39(5):1000–1002, May 1994.

[17] R B Burckel. *An Introduction to Classical Complex Analysis: Volume 1.* Pure and Applied Mathematics. Academic Press, 1979.

[18] C Caratheodory. Über die Schwarzsche Lemma bei analytischen Funktionen von zwei komplexen Veränderlichen. *Math. Ann.*, 97:76–98, 1926.

[19] C Caratheodory. Über die Geometrie der analytischen Abbildungen, die durch analytischen Funktionen von zwei Veränderlichen vermittelt werden. *Abh. Math. Sem. Univ. Hamb.*, 6:97–145, 1928.

[20] M J Crabb and A M Davie. von Neumann's inequality for Hilbert space operators. *Bull. London Math. Soc.*, 7:49–50, 1975.

[21] J C Doyle. Structured uncertainty in control system design. *IEEE Conf. on Decision and Control, Ft. Lauderdale*, 1982.

[22] J C Doyle. Structured uncertainty in control systems. IFAC Workshop in Model Error Concepts and Compensation, Boston, June 1985.

[23] J C Doyle, B A Francis, and A R Tannenbaum. *Feedback Control Theory.* Macmillan, New York, NY, 1992.

[24] J Eschmeier and M Putinar. *Spectral Decompositions and Analytic Sheaves*, volume 10 of *LMS Monographs (New Series).* OUP, Oxford, 1996.

[25] A Feintuch and A Markus. The structured norm of a Hilbert space operator with respect to a given algebra of operators. Preprint, Ben-Gurion University of the Negev, Israel, 1997.

[26] C Foias and A E Frazho. *The Commutant Lifting Approach to Interpolation Problems*, volume 44 of *Operator Theory:Advances and Applications.* Birkhäuser Verlag, Basel, 1990.

[27] B A Francis. *A Course in H_∞ Control Theory.* Number 88 in Lecture Notes in Control and Information Sciences. Springer-Verlag, Berlin, 1987.

[28] R C Gunning. *Introduction to Holomorphic Functions of Several Variables*, volume 1. Brooks/Cole, 1990.

[29] P R Halmos. *A Hilbert Space Problem Book.* D.Van Nostrand Company, Inc., Princeton, NJ., 1967.

[30] S Kobayashi. *Holomorphic manifolds and holomorphic mappings.* Marcel Decker, New York, NY, 1970.

[31] P Koosis. *Introduction to H^p Spaces.* Cambridge University Press, 1980.

[32] A Packard and J Doyle. The complex structured singular value. *Automatica*, 29(1):71–109, 1993.

[33] J R Partington. *Interpolation, Identification, and Sampling*, volume 17 of *LMS Monographs, New series.* Oxford University Press, Oxford, 1997.

[34] G Pick. Über die Beschränkungen analytischer Funktionen, welche durch vorgegebene Funktionswerte bewirkt werden. *Math. Ann.*, 77:7–23, 1916.

[35] W F Stinespring. Positive functions on C^*-algebras. *Proc. A.M.S.*, 6:211–216, 1955.

[36] B Sz-Nagy and C Foias. *Harmonic Analysis of Operators on Hilbert Space.* North-Holland Publishing Company, Amsterdam, 1970.

[37] B L van der Waerden. *Modern Algebra.* Frederick Ungar Pub., 1953.