Generalisations of Pick's Theorem to Reproducing Kernel Hilbert Spaces

Ph.D. Thesis submitted to Lancaster University
by Peter Philip Quiggin M.A.
Department of Mathematics and Statistics
Lancaster University, United Kingdom

August 1994

Contents

Generalisations of Pick's Theorem to Reproducing Kernel Hilbert Spaces

Ph.D. Thesis submitted to Lancaster University

by Peter Philip Quiggin M.A.

Department of Mathematics and Statistics

Lancaster University, United Kingdom.

August 1994

Abstract

Pick's theorem states that there exists a function in H^∞, which is bounded by 1 and takes given values at given points, if and only if a certain matrix is positive. H^∞ is the space of multipliers of H^2 and this theorem has a natural generalisation when H^∞ is replaced by the space of multipliers of a general reproducing kernel Hilbert space $H(K)$ (where K is the reproducing kernel). J. Agler showed that this generalised theorem is true when $H(K)$ is a certain Sobolev space or the Dirichlet space. This thesis widens Agler's approach to cover reproducing kernel Hilbert spaces in general and derives sufficient (and usable) conditions on the kernel K, for the generalised Pick's theorem to be true for $H(K)$. These conditions are then used to prove Pick's theorem for certain weighted Hardy and Sobolev spaces and for a functional Hilbert space introduced by Saitoh. The reproducing kernel approach is then used to derived results for several related problems. These include the uniqueness of the optimal interpolating multiplier, the case of operator-valued functions and a proof of the Adamyan-Arov-Kreĭn theorem.

MSC 1991 : Primary 30E05

Secondary 46E22, 47A20.

Acknowledgments

This research draws heavily on ideas developed by Prof. J. Agler in an (as yet) unpublished draft paper in which he proves Pick's theorem for the Dirichlet space. The overall approach stems from Agler's work, and in particular lemma 1.4.2 and the main idea for the proof of lemma 1.3.3 are derived from that paper. I would like to thank Prof. Agler for the stimulus his ideas have given. I would also like to thank Scott McCullough, who showed the way forward on the completely NP kernels described in chapter 5.

Most importantly, however, I would like to thank my supervisor, Prof. N. Young, for many useful and enjoyable discussions, much inspiration and several key ideas, in particular for the proofs of lemma 1.4.3 and theorem 2.2.3.

This work was supported by a grant from the Science and Engineering Research Council, for which I am very grateful.

Declaration

The contents of chapters 1 and 2 of this thesis have previously been published as a paper in the journal Integral Equations and Operator Theory [Qui93].

Introduction and Overview

Pick's theorem is a result about interpolation for complex-valued functions. Suppose we are asked to find an analytic function $\phi : \mathbb{D} \to \mathbb{C}$ on the unit disk $\mathbb{D}$ whose supremum norm $\|\phi\|_\infty = \sup_{z\in\mathbb{D}} |\phi(z)|$ is as small as possible and yet ϕ satisfies the interpolation requirement that $\phi(x_i) = z_i$ $(i = 1, \ldots, n)$. Here $x_1, \ldots, x_n \in \mathbb{D}$ are some given points and $z_1, \ldots, z_n \in \mathbb{C}$ are given values that the function must take at those points. Pick's theorem addresses the question: how small can we make $\|\phi\|_\infty$ and yet still satisfy the interpolation requirement? It states:

> There exists an analytic function $\phi : \mathbb{D} \to \mathbb{C}$ for which $\|\phi\|_\infty \leq 1$ and $\phi(x_i) = z_i$ $(i = 1, \ldots, n)$ if and only if the matrix
>
> $$\left(\frac{1 - z_i z_j^*}{1 - x_i x_j^*}\right)_{i,j=1,\ldots,n}$$
>
> is positive.

Because scaling the data values z_i simply scales the possible solutions ϕ and their norms $\|\phi\|_\infty$, this result effectively answers our question. For

$$\begin{aligned}
& \text{there exists } \phi \text{ such that } \|\phi\|_\infty \leq r \text{ and } \phi(x_i) = z_i \\
\Leftrightarrow\ & \text{there exists } \phi \text{ such that } \|\phi\|_\infty \leq 1 \text{ and } \phi(x_i) = z_i/r \\
\Leftrightarrow\ & \left(\frac{1 - z_i z_j^*/r^2}{1 - x_i x_j^*}\right)_{i,j=1,\ldots,n} \text{ is positive} \\
\Leftrightarrow\ & \left(\frac{r^2 - z_i z_j^*}{1 - x_i x_j^*}\right)_{i,j=1,\ldots,n} \text{ is positive}
\end{aligned}$$

so the best we can achieve for $\|\phi\|_\infty$ is

$$\inf_{r\geq 0}\left(r : \left(\frac{r^2 - z_i z_j^*}{1 - x_i x_j^*}\right)_{i,j=1,\ldots,n} \text{ is positive } \right).$$

Pick proved his theorem in 1916 [Pic16] using function-theoretic methods, but because of its importance it continued to be studied by many authors and in 1967 Donald Sarason [Sar67] gave a radically new, and particularly natural, operator-theoretic interpretation and proof of the result. In Sarason's approach the function ϕ is represented by the operator on the Hardy space H^2 of multiplication by ϕ, the norm $\|\phi\|_\infty$ equals the norm of this operator and Pick's theorem is then a result about multiplication operators on H^2. This is one example of an interplay between function theory and operator theory that has proved fruitful in recent decades—some aspects of Pick's theorem are easier to understand from the operator-theoretic viewpoint, whereas others are easier from a function-theoretic viewpoint.

H^2 is a reproducing kernel Hilbert space, i.e. it is a Hilbert space of complex-valued functions on some set X ($= \mathbb{D}$ in this case) in which all of the point-evaluation functionals $f \mapsto f(x)$, $x \in X$, are continuous. Each such space has a unique associated function $K : X \times X \to \mathbb{C}$, called its reproducing kernel, which completely characterises the space and can be used to study it. I denote the space corresponding to kernel K by $H(K)$.

In the late 1980s, J. Agler shone further light on Pick's theorem using H^2's reproducing kernel. He showed that the operator-theoretic form of Pick's theorem still holds if H^2 is replaced by certain other reproducing kernel Hilbert spaces. The research in this thesis began when I read an early draft paper by Agler in which he proved this variation of Pick's theorem for the Dirichlet space. It appeared that Agler's approach could be generalised to any reproducing kernel Hilbert space whose kernel satisfied suitable properties and this lead me to ask the question 'For which reproducing kernel Hilbert spaces is Pick's theorem true?'. Chapters 1 and 2 are the result of this enquiry and are essentially as published in my paper [Qui93]. Chapter 1 derives sufficient conditions on the kernel K for Pick's theorem to hold for $H(K)$ and chapter 2 applies this result to several specific cases.

The following two chapters examine whether the reproducing kernel approach can also be used to prove and generalise some other results related to Pick's theorem. Chapter 3 studies the known fact that for H^2 there is a *unique* analytic interpolating function that has the smallest possible norm $\|\phi\|_\infty$ and shows that this uniqueness result also holds for a class of spaces closely related to H^2. In chapter 4 the reproducing kernel approach is generalised to interpolation of operator-valued functions.

The sufficient conditions derived in chapter 1 frustrated me. Numerous computer calculations with randomly chosen kernels that failed to satisfy the condition all found that the generalised Pick's theorem also failed to hold with those kernels. This lead me into a lengthy, but failed, attempt to prove that the conditions on the kernel K are also necessary. Only later did it become apparent, from studying related work by Scott McCullough [McC92], that this attempt was doomed to fail since the conditions are *not* necessary—they in fact characterise a slightly stronger property of the kernel. Chapter 5 describes this stronger property and derives an explicit counter-example to necessity.

The final two chapters cover another very important derivative of Pick's theorem—the Adamyan-Arov-Kreĭn theorem. This difficult result, proved in 1971 [AAK71], has important applications in control theory. Chapter 7 addresses the question 'Can the reproducing kernel approach shed light on the Adamyan-Arov-Kreĭn theorem and perhaps even generalise it further?'. My attempts to answer this question lead me to some much more general results about reproducing kernels, which are described in chapter 6.

Chapter 1

The Scalar-Valued Case

This chapter studies Pick's theorem for reproducing kernel Hilbert spaces of scalar-valued functions.

1.1 Reproducing Kernel Hilbert Spaces

I will start by defining notation and stating the main facts which we will need from the theory of reproducing kernel Hilbert spaces. For details and proofs of these facts, see [Aro50] and [Sai88].

Given a set X, a *kernel* K on X is a complex-valued function on $X \times X$. When the set X is a finite set I will call K a *finite* kernel. Finite kernels can be viewed as matrices; in fact kernels are effectively generalisations of matrices, and in line with this analogy I will call the functions $K(x, \cdot)$ and $K(\cdot, x)$ the x-row and x-column of K, respectively. The *restriction* of K to the set $E \times E$, where E is any non-empty subset of X, will be denoted K_E, and kernel K is called *positive* or *positive-definite*, denoted $K \geq 0$ and $K > 0$, whenever its finite restrictions are all positive or all positive-definite matrices, respectively.

Given any Hilbert space H of complex-valued functions on a set X, for which the point-evaluation linear functionals are all continuous, it can be shown that there exists a unique positive kernel K which has the following 'reproducing' property for H:

$$\text{for all } f \in H \text{ and all } x \in X,\ K(\cdot, x) \in H \text{ and } f(x) = \langle f, K(\cdot, x)\rangle$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product in H.

Conversely, it is also true that to every positive kernel K on a set X there corresponds a unique Hilbert space of complex-valued functions on X, for which K is the reproducing kernel. I indicate this association between the positive kernel K and the corresponding reproducing kernel Hilbert space by denoting the latter as $H(K)$. The space $H(K)$ is easily described—it is the completion of the span of the columns of K; i.e. the functions in $H(K)$ are all limits of finite linear combinations of the columns of K. And the inner product in $H(K)$ is the extension to the whole space of the inner product defined on the columns of K by

$$\langle K(\cdot, y), K(\cdot, x)\rangle = K(x, y).$$

The functions in $H(K)$, and their norms, can also be characterised by a positivity condition—they are exactly the functions f on X for which $r^2K - f \otimes f^* \geq 0$ for some real $r \geq 0$ (where $r^2K - f \otimes f^*$ denotes the kernel $r^2K(x, y) - f(x)f(y)^*$). The set of such r (for which $r^2K - f \otimes f^* \geq 0$) is either empty if f is not in $H(K)$, or else is the half open interval $[\|f\|, \infty)$, where $\|f\|$ is f's norm in $H(K)$.

Given a positive kernel K, a function $\phi : X \to \mathbb{C}$ is called a *multiplier* of $H(K)$ if $H(K)$ is closed under multiplication by ϕ. When this happens multiplication by ϕ always turns out to be a bounded linear operator on $H(K)$, which I denote by M_ϕ, and the operator norm of M_ϕ gives a 'multiplier' norm for ϕ, denoted $\|\phi\|_{M(K)}$. I will denote the space of multipliers of $H(K)$, with this multiplier norm, by $M(K)$. The functions in $M(K)$, and their multiplier norms, are again characterised by a positivity condition—they are the functions ϕ on X for which $(r^2 - \phi \otimes \phi^*)K \geq 0$

for some non-negative r. The set of such r (for which $(r^2 - \phi \otimes \phi^*)K \geq 0$) is either empty if ϕ is not in $M(K)$, or else is the half open interval $[\|\phi\|_{M(K)}, \infty)$.

If K is positive-definite, then when we restrict K to a subset E of X the resulting spaces $H(K_E)$ and $M(K_E)$ relate to $H(K)$ and $M(K)$ in a simple way. Let $H(K, E)$ denote the closed subspace of $H(K)$ spanned by the columns of K corresponding to the subset E. Then the mapping

$$U : H(K_E) \to H(K, E) \qquad K_E(\cdot, x) \mapsto K(\cdot, x) \quad \text{for all } x \in E$$

is unitary, since the E columns of K_E and of K have the same Gram matrix, namely K_E. Therefore U is a unitary embedding of $H(K_E)$ in $H(K)$.

Given a function $f : E \to \mathbb{C}$ that is in $M(K_E)$, its corresponding multiplication operator M_f is an operator on $H(K_E)$. But U is effectively an identification of $H(K_E)$ as $H(K, E)$ and so allows us to work instead with the unitarily equivalent operator UM_fU^{-1} on $H(K, E)$, which I will denote as $\tilde{M}_f$. We will sometimes work with $\tilde{M}_f$ rather than M_f, the advantage being that $\tilde{M}_f$ is an operator on a subspace of $H(K)$.

If K is a positive kernel and $\phi \in M(K)$ is a multiplier of $H(K)$, let ϕ_E be its restriction to the subset E. Then from the above characterisation of multiplier norms we have $(\|\phi\|^2_{M(K)} - \phi \otimes \phi^*)K \geq 0$. But $(\|\phi\|^2_{M(K)} - \phi_E \otimes \phi_E^*)K_E$ is a restriction of this kernel, and so must also be positive; so ϕ_E is a multiplier of $H(K_E)$ and $\|\phi_E\|_{M(K_E)} \leq \|\phi\|_{M(K)}$. In words, the restriction to E of a multiplier of $H(K)$ is a multiplier of $H(K_E)$ with equal or smaller norm.

1.2 Generalising Pick's Theorem

Nevanlinna [Nev19] and Pick [Pic16] studied the problem of interpolation by bounded analytic functions early this century, and much more recently Sarason [Sar67] set their results in an operator-theoretic context.

Pick's original theorem states:

> There exists a function $\phi \in H^\infty$, with $\|\phi\|_\infty \leq 1$, that takes the n given data values $z_i \in \mathbb{C}$ at the n given data points $x_i \in \mathbb{D}$ if and only if
>
> $$\left(\frac{1 - z_i z_j^*}{1 - x_i x_j^*} \right)_{i,j=1,\ldots,n} \geq 0.$$

We can now set this into the context of reproducing kernel Hilbert spaces and the notation defined above. Let

$$\begin{aligned} X &= \text{the open unit disc } \mathbb{D} \\ K &= \text{the Szegö kernel } K(x,y) = 1/(1 - xy^*) \\ E &= \text{the set of data points } \{x_1, \ldots, x_n\} \\ K_E &= \text{the restriction of } K \text{ to } E \\ \text{and } f &= \text{the function on } E \text{ defined by } f(x_i) = z_i. \end{aligned}$$

Then $H(K) = H^2$, $M(K) = H^\infty$, $\|\cdot\|_{M(K)} = \|\cdot\|_\infty$, and the matrix in Pick's theorem is simply the finite kernel $(1 - f \otimes f^*)K_E$, positivity of which is, from above, equivalent to $\|f\|_{M(K_E)} \leq 1$. We can therefore restate Pick's theorem as

> **(1)** There exists a multiplier $\phi \in M(K)$, with $\|\phi\|_{M(K)} \leq 1$, that is an extension of the multiplier $f \in M(K_E)$
>
> if and only if
>
> **(2)** $\|f\|_{M(K_E)} \leq 1$.

It is now clear why (1) $\Rightarrow$ (2); simply because f is then the restriction of ϕ to E and so (see section 1.1) must have multiplier norm no greater than that of ϕ. So the 'meat' of Pick's theorem is the sufficiency of (2), i.e. that every multiplier of $H(K_E)$ has an extension to a multiplier of $H(K)$ with no greater (and therefore equal) norm.

With this in mind I define the following terminology. If E and F are non-empty subsets of X and F contains E, then I will say that *all multipliers on E extend isometrically to F* if all multipliers in $M(K_E)$ have extensions in $M(K_F)$ with equal norm.

If multipliers on E extend isometrically to F, then they clearly also extend isometrically to any subset F_1 of F that contains E, since we can extend any multiplier in $M(K_E)$ to F and then restrict back to F_1, all without increasing the multiplier norm.

It is clear from the above discussion that the following, which I will refer to as the *full Pick theorem*, is true for $H(K)$ if and only if all multipliers on all non-empty subsets of X extend isometrically to X:

> Given a function f on any non-empty subset E of X, there exists an extension $\phi \in M(K)$ of f, with $\|\phi\|_{M(K)} \leq 1$ if and only if the kernel $(1 - f \otimes f^*)K_E$ is positive.

I will call the weakening that results if only finite subsets E are allowed, the *finite Pick theorem*:

> Given a function f on any *finite* non-empty subset E of X, there exists an extension $\phi \in M(K)$ of f, with $\|\phi\|_{M(K)} \leq 1$ if and only if the kernel $(1 - f \otimes f^*)K_E$ is positive.

Clearly it is true if and only if all multipliers on non-empty *finite* subsets of X extend isometrically to X.

In this terminology, Pick's original theorem is the finite Pick theorem for the case where $H(K)$ is the Hardy space H^2. It is also known that the finite Pick theorem is true for several other reproducing kernel Hilbert spaces, including a

Sobolev space on [0,1] (see [Agl90]), and the Dirichlet space on $\mathbb{D}$ [Agl88]. And more restricted forms of Pick's theorem, for example with conditions placed on the subset E, have been proved for some classes of reproducing kernel Hilbert spaces [Sai88] [BB84] [Sza86]. However neither the finite nor the full Pick theorem are true for all reproducing kernel Hilbert spaces—Beatrous and Burbea [BB84] show this, and in section 2.3 I give other examples.

The question that we would like to answer is: for which positive kernels K are these two theorems true for $H(K)$? Hopefully it is now clear that this question comes down to asking which extensions of multipliers can be done isometrically.

1.3 One-Point Extensions Are Sufficient

The property of being able to extend all multipliers isometrically, from one subset E of X to a larger one F, is clearly transitive. I.e. if we can extend all multipliers isometrically from E to F, and also from F to G, then we can do so from E to G. Because of this, the extensions where exactly one extra point is added (i.e. from E to $E \cup \{t\}$ where $t \in X \setminus E$) are elementary extensions from which larger extensions can be built, and I will call these *one-point extensions*. In this section I show it is sufficient to know that all one-point extensions can be done isometrically.

Firstly, we will need the following characterisation of multiplication operators on reproducing kernel Hilbert spaces.

LEMMA 1.3.1 *Given a positive kernel K on a set X, an operator M on $H(K)$ is a multiplication operator (i.e. M is the same as multiplication by some multiplier in $M(K)$) if and only if every column of K is an eigenvector of M^*. Further, if M is a multiplication operator then M^*'s eigenvalue corresponding to the y-column of K is $\phi(y)^*$ where ϕ is the corresponding multiplier.*

Proof: $\Rightarrow$ If M is the operator of multiplication by multiplier ϕ, then for all $y \in X$

$$\begin{aligned} M^*(K(\cdot,y))(x) &= \langle M^*(K(\cdot,y)), K(\cdot,x)\rangle \\ &= \langle K(\cdot,y), M(K(\cdot,x))\rangle \\ &= \langle K(\cdot,y), \phi(\cdot)K(\cdot,x)\rangle \\ &= \langle \phi(\cdot)K(\cdot,x), K(\cdot,y)\rangle^* \\ &= \phi(y)^*K(y,x)^* = \phi(y)^*K(x,y) \\ &= (\phi(y)^*K(\cdot,y))(x). \end{aligned}$$

Therefore $K(\cdot,y)$ is an eigenvector of M^*, with eigenvalue $\phi(y)^*$.

$\Leftarrow$ Given an operator M such that all columns of K are eigenvectors of M^*, define the function $\phi : X \to \mathbb{C}$ by $\phi(y)^* = M^*$'s eigenvalue corresponding to the y-column of K. Then for all $y \in X$

$$\begin{aligned} M(K(\cdot,y))(x) &= \langle M(K(\cdot,y)), K(\cdot,x)\rangle \\ &= \langle K(\cdot,y), M^*(K(\cdot,x))\rangle \\ &= \langle K(\cdot,y), \phi(x)^*K(\cdot,x)\rangle \\ &= (\phi(\cdot)K(\cdot,y))(x). \end{aligned}$$

So M and multiplication by ϕ agree on all columns of K, and by linearity they must also agree on the linear span of K's columns. But for any $f \in H(K)$, f is the limit of some sequence of functions $(f_n)_{n\in\mathbb{N}}$ in the linear span of K's columns, so

$$\begin{aligned} M(f)(x) &= \lim(M(f_n))(x) \\ &= \lim(M(f_n)(x)) \\ &= \lim(\phi(x)f_n(x)) \\ &= \phi(x)\lim(f_n(x)) \\ &= \phi(x)f(x) \end{aligned}$$

Therefore M and multiplication by ϕ agree on the whole of $H(K)$.

■

We can now prove that one-point isometric extensions are sufficient. Following Agler, I will call a positive kernel for which all one-point extensions can be done isometrically a Nevanlinna-Pick kernel, abbreviated for convenience to NP kernel.

DEFINITION 1.3.2 *A positive kernel K on a set X will be called an NP kernel provided for every subset E of X and every $t \in X \setminus E$, all multipliers in $M(K_E)$ can be isometrically extended to multipliers in $M(K_{E \cup \{t\}})$.*

This new terminology is justified by the following lemma.

LEMMA 1.3.3 *Given a positive-definite kernel K on a set X, the full Pick theorem is true for $H(K)$ if and only if K is an NP kernel.*

Proof: The forward implication is clear, since given a subset E and a point $t \in X \setminus E$, all multipliers in $M(K_E)$ can be isometrically extended to X and then restricted back to $E \cup \{t\}$, all without increasing the norm. So given a multiplier f defined on subset E of X, and assuming all one-point extensions can be done isometrically, we need to show there exists an isometric extension of f to X.

Consider the set A of isometric extensions of f, i.e. of isometric extensions of f to multipliers on subsets of X that contain E. Then A is non-empty since it contains f itself, and all the multipliers in A have the same norm as f. Now define a partial ordering of A by defining that $g \leq h$ whenever h is an extension of g, i.e. whenever h's domain contains g's domain and h agrees with g on g's domain. We now show that if G is any totally ordered subset of A, then G is bounded by some element of A.

SubProof: For each $g \in G$ let

F_g = domain of g

$F = \cup_{g \in G} F_g$

K_{F_g} = the restriction of K to F_g

$H_g = H(K_F, F_g)$, the closed subspace of $H(K_F)$ spanned by the columns of K_F corresponding to F_g

$\tilde{M}_g$ = the operator on H_g that corresponds to multiplication of $H(K_{F_g})$ by g. Recall that $H(K_{F_g})$ embeds naturally as H_g in $H(K_F)$, so operators on $H(K_{F_g})$ induce unitarily equivalent operators on H_g.

Then the operators $\tilde{M}_g$ all have norm $\|f\|_{M(K_E)}$, and their adjoints $\tilde{M}_g^*$ form a 'nest' in the following sense. Let $\tilde{M}_{g_1}^*$ and $\tilde{M}_{g_2}^*$ be any two such operators; since G is totally ordered we can assume (without loss of generality) that $g_1 \leq g_2$. Then H_{g_1} is a $\tilde{M}_{g_2}^*$-invariant subspace of H_{g_2}, since it is generated by the F_{g_1}-columns of K, which are all eigenvectors of $\tilde{M}_{g_2}^*$. And $\tilde{M}_{g_1}^*$ is simply $\tilde{M}_{g_2}^*$ restricted to H_{g_1}.

Now define the linear transformation M^*, on the union of the H_g's, by $M^*(h) = \tilde{M}_g^*(h)$ where g is any element of G such that $h \in H_g$. Because we have a 'nest' of operators this is well-defined, and because each $\tilde{M}_g^*$ has norm $\|f\|_{M(K_E)}$ then so does M^*.

We can now extend M^* by continuity, without increasing its norm, to an operator on the norm-closure of the union of the H_g's in $H(K_F)$, which is the whole of $H(K_F)$. This gives a well-defined operator M^* on $H(K_F)$ whose restrictions to the subspaces H_g (which are M^*-invariant) are the operators $\tilde{M}_g^*$.

Since $\tilde{M}_g^*$'s eigenvectors include all the F_g-columns of K_F for all $g \in G$, all columns of K_F are eigenvectors of M^*, so (by lemma 1.3.1) M is a multiplication operator on $H(K_F)$. Let the corresponding multiplier be

ϕ; then from the way M was constructed, ϕ is an extension of each g in G, and has the same norm $\|f\|_{M(K_E)}$. Therefore ϕ is in A, and $g \leq \phi$ for all $g \in G$, so G is bounded in A, as required. ▪

Since all totally ordered subsets of A are bounded (and assuming the Axiom of Choice), A has a maximal element ϕ say, by Zorn's lemma. But then ϕ's domain must be the whole of X, since otherwise (by assumption) there would exist a one-point extension of ϕ which would be in A and greater than ϕ, contradicting ϕ's maximality. ϕ is therefore an isometric extension of f to the whole of X, as required. ■

1.4 Minimal Norm Extension

Throughout this section, let X be any set, K be a *positive-definite* kernel on X, t be any point of X, and $E = X \setminus \{t\}$. We will consider the problem of extending a given multiplier f on E to a multiplier ϕ on X which has the smallest possible norm. Agler [Agl88] has given an explicit formula for this smallest possible norm, and shown that it is achievable. Here we will derive Agler's result from Parrott's theorem, but first we need to move over to an operator-theoretic way of viewing the problem. Initially, let us restrict attention to X being a *finite* set; at the end of this section we will see that the main result also holds for infinite X.

Corresponding to the given multiplier f is its multiplication operator adjoint M_f^* on $H(K_E)$, which in turn induces an operator $\tilde{M}_f^*$ on $H(K, E)$ (the closure of the span of the E-columns, which is the subspace of $H(K)$ corresponding to $H(K_E)$). $\tilde{M}_f^*$ is simply the operator that has the E-columns of K as its eigenvectors with $f(x)^*$ as its eigenvalues; in other words it is a diagonal operator with respect to the basis given by the E-columns of K. Because K is positive-definite and E is finite, $K(\cdot, t)$ is not in $H(K, E)$ so the process of one-point extending f to the extra point t is therefore equivalent to extending $\tilde{M}_f^*$ to an operator on $H(K)$ by choosing a new eigenvalue $f(t)^*$ for the new (given) eigenvector $K(\cdot, t)$. That is, using the decomposition $H(K) = \text{span}(E\text{-columns of } K) + \text{span}(t\text{-column of } K)$, it is the process of completing the partially defined block matrix

$$\begin{pmatrix} \tilde{M}_f^* & 0 \\ 0 & ? \end{pmatrix}$$

by choosing a new eigenvalue to go in the bottom right-hand corner. This block matrix completion problem is unusual, however, since the blocks are defined with respect to a (in general) non-orthogonal decomposition of $H(K)$. To handle this we need the following minor generalisation of Parrott's theorem (which in its usual form applies only to orthogonal decompositions).

LEMMA 1.4.1 *Let $H = H_1 + H_2$ and $J = J_1 + J_2$ be (not necessarily orthogonal) decompositions of Hilbert spaces H and J into linearly independent non-trivial subspaces H_1, H_2 and J_1, J_2, and let $\mathcal{L}(H_i, J_j)$ denote the Banach space of bounded linear operators from H_i to J_j. Given any operators $A \in \mathcal{L}(H_1, J_1)$, $B \in \mathcal{L}(H_2, J_1)$, and $C \in \mathcal{L}(H_1, J_2)$, let T_D denote the linear transformation from H to J given by*

$$T_D = \begin{pmatrix} A & B \\ C & D \end{pmatrix}.$$

Then

$$\min_{D \in \mathcal{L}(H_2, J_2)} \|T_D\| = \max(\|R\|, \|S\|)$$

where

$$\begin{aligned} R &= \textit{common restriction of all completions } T_D \textit{ to } H_1 \\ \textit{and } S &= \textit{common projection of all completions } T_D \textit{ onto } J_2^{\perp}. \end{aligned}$$

(Note that R and S are independent of the variable operator D.)

Proof: Re-expressing T_D with respect to the *orthogonal* decompositions $H = H_1 \oplus H_1^{\perp}$ and $J = J_2^{\perp} \oplus J_2$ gives

$$T_D = \begin{pmatrix} A' & B' \\ C' & D' \end{pmatrix}$$

$$
\begin{aligned}
\text{where } A' &= P_{J_2^\perp} A \\
B' &= P_{J_2^\perp}(AQ_1 + BQ_2) \\
C' &= P_{J_2} A + C \\
D' &= P_{J_2}(AQ_1 + BQ_2) + (CQ_1 + DQ_2).
\end{aligned}
$$

Here $P_{J_2^\perp}$ and P_{J_2} are the orthogonal projections onto those subspaces, and $Q_1 \in \mathcal{L}(H_1^\perp, H_1)$ and $Q_2 \in \mathcal{L}(H_1^\perp, H_2)$ are the operators that map a vector $v \in H_1^\perp$ to its unique representation as a sum of vectors in H_1 and H_2, respectively.

Now Q_2 is invertible—its inverse is $P_{H_1^\perp}$—so as D ranges over all of $\mathcal{L}(H_2, J_2)$, D' ranges over all of $\mathcal{L}(H_1^\perp, J_2)$, and Parrott's theorem [Par78] therefore tells us the smallest possible value for $\|T_D\|$ is

$$
\max(\| \begin{pmatrix} A' \\ C' \end{pmatrix} \|, \| \begin{pmatrix} A' & B' \end{pmatrix} \|)
$$

and that it is achievable. Since

$$
\begin{pmatrix} A' \\ C' \end{pmatrix} = R \text{ and } \begin{pmatrix} A' & B' \end{pmatrix} = S
$$

the result follows. ■

We can now obtain Agler's explicit minimal achievable norm for a one-point extension:

Lemma 1.4.2 *(J. Agler [Agl88])* *Let K be a positive-definite kernel on a finite set X, $t \in X$, $E = X \setminus \{t\}$ and f be a multiplier of $H(K_E)$. Also, let $K^{(t)}$ denote the kernel on X given by*

$$
K^{(t)}(x, y) = K(x, y) - \frac{K(x, t)K(t, y)}{K(t, t)}.
$$

(The kernel $K^{(t)}$ is the Schur complement of the diagonal term $K(t, t)$ in K.)

Then $\|f\|_{M(K_E)}$ and $\|f\|_{M(K_E^{(t)})}$ are both lower bounds for the norm of any one-point extension of f to X, and the larger of the two, if finite, is achievable, i.e. there exists a one-point extension of f with that norm.

Proof: In the notation of lemma 1.4.1, let

$$\begin{aligned}
H_1 = J_1 &= H(K,E) = \text{ isomorphic to } H(K_E) \\
H_2 = J_2 &= H(K,\{t\}) = \text{ one-dimensional} \\
H = J &= H(K) \\
A &= \tilde{M}_f^* = \text{ operator on } H_1 \text{ corresponding to } M_f^* \text{ on } H(K_E) \\
B = C &= 0.
\end{aligned}$$

Then the variable operator D is just a single complex number, and the completions T_D (if any bounded ones exist) are exactly the multiplication operator adjoints of one-point extensions of f to X (if any exist). Therefore, by lemma 1.4.1, the smallest possible norm of a one-point extension of f is $\max(\|R\|, \|S\|)$ and if finite this norm is achievable. But

$$\begin{aligned}
\|R\| &= \|A\| && \text{since } C = 0 \\
&= \|M_f^*\| \\
&= \|f\|_{M(K_E)}
\end{aligned}$$

$$\begin{aligned}
\text{and } \|S\| &= \|\text{common projection of all completions } T_D \text{ onto } J_2^\perp\| \\
&= \|\text{compression, W say, of } T_D \text{ onto } H_2^\perp\| && \text{since } B = 0.
\end{aligned}$$

Now $H_2^\perp$ is itself a reproducing kernel Hilbert space and we have

$$\begin{aligned}
\langle K^{(t)}(\cdot,y), K(\cdot,t)\rangle &= \langle K(\cdot,y) - \frac{K(\cdot,t)K(t,y)}{K(t,t)}, K(\cdot,t)\rangle \\
&= K(t,y) - \frac{K(t,t)K(t,y)}{K(t,t)} \\
&\quad \text{(using } K\text{'s reproducing property)} \\
&= 0
\end{aligned}$$

so all the columns of $K^{(t)}$ are in $K(\cdot,t)^\perp = H_2^\perp$. Indeed its columns $K_y^{(t)}$ are, by direct calculation, simply the projections of K's columns orthogonal to $K(\cdot,t)$. Further, $K^{(t)}$ also has the reproducing property for $K(\cdot,t)^\perp$ since for all $f \in K(\cdot,t)^\perp$ we have

$$\begin{aligned}\langle f, K^{(t)}(\cdot,y)\rangle &= \langle f, K(\cdot,y) - \frac{K(\cdot,t)K(t,y)}{K(t,t)}\rangle \\ &= f(y)\end{aligned}$$

(using the reproducing property of K, and noting that $\langle f, K(\cdot,t)\rangle = 0$). Therefore $K^{(t)}$ is the reproducing kernel of $K(\cdot,t)^\perp = H_2^\perp$.

Now for $y \in E$

$$\begin{aligned}W(K^{(t)}(\cdot,y)) &= P_{H_2^\perp} T_D \left(K(\cdot,y) - \frac{K(\cdot,t)K(t,y)}{K(t,t)} \right) \\ &= P_{H_2^\perp}(A(K(\cdot,y)) - \frac{K(t,y)}{K(t,t)} D K(\cdot,t)) \\ &= P_{H_2^\perp}(f(y)^* K(\cdot,y)) \qquad \text{since } P_{H_2}^\perp D = 0 \\ &= f(y)^* K^{(t)}(\cdot,y).\end{aligned}$$

Hence the compression W is just the operator with the E-columns of $K^{(t)}$ as eigenvectors and $f(y)^*$ as eigenvalues, so it corresponds to the adjoint of multiplication by f on $H(K_E^{(t)})$. So $\|S\| = \|f\|_{M(K_E^{(t)})}$ and the proof is complete. ■

So the smallest norm of a one-point extension of f is $\max\{\|f\|_{M(K_E)}, \|f\|_{M(K_E^{(t)})}\}$ and if this is infinite (because f is not a multiplier of $H(K_E^{(t)})$) then no extension of f exists in $M(K)$. To illustrate the operator-theoretic view used above to show this, consider the case of a finite domain set, say $X = \{x_1, x_2, x_3\}$. Then $H(K)$ is isomorphic to $\mathbb{C}^3$ and we can think of K as being the Gram matrix of the vectors, a_1, a_2, a_3 say, in $\mathbb{C}^3$ corresponding to its columns. If $E = \{x_1, x_2\}$ and $t = x_3$, then $H(K_E)$ corresponds to span$\{a_1, a_2\}$ and $H(K_E^{(t)})$ corresponds to span$\{w_1, w_2\}$, where w_i is the projection of a_i onto $a_3^\perp$. With notation consistent with the previous lemma, $\|f\|_{M(K_E)} = \|A\|$ and $\|f\|_{M(K_E^{(t)})} = \|W\|$, where

$$\begin{aligned} A &= \text{operator on span}\{a_1, a_2\} \text{ with eigenvectors } a_1, a_2 \\ &\quad \text{and eigenvalues } f(x_1)^*, f(x_2)^* \\ \text{and } W &= \text{operator on span}\{w_1, w_2\} \text{ with eigenvectors } w_1, w_2 \\ &\quad \text{and eigenvalues } f(x_1)^*, f(x_2)^*. \end{aligned}$$

And the lemmas show that the smallest possible norm of a one-point extension of f is $\max(\|A\|, \|W\|)$, and hence f can be extended isometrically from E to X if and only if $\|W\| \leq \|A\|$.

This allows us to explicitly construct a non-NP finite kernel. If we choose a_1 and a_2 to be unit vectors and such that $\langle a_1, a_2 \rangle$ is real, then the vectors $a_1 + a_2$ and $a_1 - a_2$ are orthogonal—they are simply the diagonals of the parallelogram formed by a_1 and a_2. And if we let $f(x_1) = 1$ and $f(x_2) = -1$, then A maps these two diagonals to each other—its matrix with respect to the basis $[a_1 + a_2, a_1 - a_2]$ is

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

It is then easy to see that the shorter of the two diagonals is A's maximising vector, and so $\|A\|$ is simply the ratio of the lengths of the two diagonals of A's unit-eigenvector parallelogram. Exactly the same applies to W, except that its eigenvectors are w_1, w_2. Amongst these operators with eigenvalues $+1$ and -1, the smallest norm ones are those with their eigenvectors orthogonal, giving norm 1, and the large norm ones are those with highly non-orthogonal eigenvectors, giving arbitrarily large norm. So if we fix a_1 and a_2 to be orthogonal, making $\|A\| = 1$, and choose a_3 such that the projections w_1, w_2 of a_1, a_2 onto $a_3^\perp$ are not orthogonal, then W will have larger norm than A. The Gram matrix of a_1, a_2, a_3 then cannot be an NP kernel, by lemma 1.4.2. The vectors $\begin{bmatrix} 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \end{bmatrix}$

give an example of this, so their Gram matrix

$$\begin{pmatrix} 1 & 0 & 1/\sqrt{3} \\ 0 & 1 & 1/\sqrt{3} \\ 1/\sqrt{3} & 1/\sqrt{3} & 1 \end{pmatrix}$$

is not an NP kernel.

Finally in this section, the following lemma shows that to test if a kernel is an NP kernel, it is sufficient to check its finite restrictions.

LEMMA 1.4.3 *Let K be a positive-definite kernel on a set X, t be any point of X, $E = X \setminus \{t\}$, and f be a multiplier in $M(K_E)$. Then if N is a positive real such that all restrictions of f to finite subsets G have one-point extensions to $G \cup \{t\}$ with norms $\leq N$, there exists an extension of f to X with norm $\leq N$.*

Proof: The result is trivial for sets X of finite cardinality, and we will prove the result by transfinite induction on the cardinality of X. Assume X is infinite, let $\aleph_\mu = \text{card}(X) = \text{card}(E)$ (where $\text{card}(X)$ denotes the cardinality of X), and assume that the result is true for all sets X of cardinality less than $\aleph_\mu$.

Well order E minimally, so that $E = \{e_\alpha : \alpha < \omega_\mu\}$ and each e_α has fewer than $\aleph_\mu$ predecessors. Note that this also minimally well-orders the E-columns of K. Now for each $\alpha < \omega_\mu$ define an operator, M_α, as follows:

1. restrict the multiplier f to $\{e_\beta : \beta < \alpha\}$, i.e. to the E-columns of K that precede the α-column.
2. one-point extend that multiplier to a multiplier on $\{e_\beta : \beta < \alpha\} \cup \{t\}$ with norm $\leq N$. This is possible, by the induction hypothesis, since we well-ordered E minimally and so $\{e_\beta : \beta < \alpha\}$ has lower cardinality than $\aleph_\mu$.
3. take the operator on $H(K)$ that corresponds to this multiplier on the closed span of the columns of K corresponding to $\{e_\beta : \beta < \alpha\} \cup \{t\}$, and is zero on the orthogonal complement of this subspace.

Then the adjoint operator M_α^* has the following properties:

- it is an operator on $H(K)$
- it has norm $\leq N$
- it has the t-column of K as an eigenvector
- for all $\beta < \alpha$, it has the β-column of K as an eigenvector with eigenvalue $f(e_\beta)^*$.

The operators M_α^* form a net in the ball B of radius N in $\mathcal{L}(H(K))$, and under the weak operator topology B is compact, so this net has some cluster point, M^* say. Since all the operators M_α^* have the t-column of K as an eigenvector, it is easily shown that so does M^*. And for each $\alpha < \omega_\mu$, since M^* is also a cluster point of the net $\{M_\beta^* : \beta > \alpha\}$, all of whose members have the α-column of K as an eigenvector with eigenvalue $f(e_\alpha)^*$, then again so does M^*.

M^* therefore has all the columns of K as eigenvectors, and is therefore a multiplication operator adjoint (on $H(K)$) for some multiplier ϕ. Since M^* has the correct eigenvalues for each E-column of K, ϕ is an extension of the multiplier f. Finally M^* is a weak operator topology limit of operators of norm $\leq N$, and so must itself have norm $\leq N$ (see for example Halmos [Hal74] problem 109). ϕ is therefore an extension of the multiplier f with multiplier norm $\leq N$. ■

COROLLARY 1.4.4 *A positive-definite kernel K is an NP kernel if and only if all finite restrictions of K are NP kernels.*

Proof: All finite restrictions of an NP kernel are clearly NP kernels themselves, so only the reverse implication needs to be proved. But this is shown by lemma 1.4.3 with N set to $\|f\|_{M(K_E)}$. ■

Corollary 1.4.5 *If K is a positive-definite kernel on a set X, then for $H(K)$ the finite Pick theorem and the full Pick theorem are equivalent.*

Proof: The full Pick theorem clearly implies the finite Pick theorem. But the finite Pick theorem implies that all finite restrictions of K are NP kernels, which by corollary 1.4.4 implies K is an NP kernel and so, by lemma 1.3.3, the full Pick theorem is true for $H(K)$. ∎

Corollary 1.4.6 *Lemma 1.4.2 also holds when X is infinite.*

Proof: Since restricting a function to a smaller domain never increases its multiplier norm, there cannot be an extension of f with multiplier norm less than

$$N = \sup_{G \subseteq E, G \text{ finite}} \inf(\|g : G \cup \{t\} \to \mathbb{C}\|_{M(K_{G\cup\{t\}})} : g|G = f|G).$$

But by lemma 1.4.2 the infimum here equals $\max\left(\|f|G\|_{M(K_G)}, \|f|G\|_{M(K_G^{(t)})}\right)$ and is achievable, so by lemma 1.4.3 an extension with multiplier norm $\leq N$ exists. Therefore (interpreting the multiplier norm as ∞ for functions that are not multipliers) we have

$$\begin{aligned}
& \min(\|\phi : X \to \mathbb{C}\|_{M(K)} : \phi|E = f) \\
= & \sup_{G \subseteq E, G \text{ finite}} \max\left(\|f|G\|_{M(K_G)}, \|f|G\|_{M(K_G^{(t)})}\right) \\
= & \max\left(\sup_{G \subseteq E, G \text{ finite}} \|f|G\|_{M(K_G)}, \sup_{G \subseteq E, G \text{ finite}} \|f|G\|_{M(K_G^{(t)})}\right) \\
= & \max\left(\|f\|_{M(K_E)}, \|f\|_{M(K_E^{(t)})}\right).
\end{aligned}$$

This last equality holds since the multiplier norm of any function is the supremum of the multiplier norms of its finite restrictions, a fact that follows from the kernel positivity characterisation of the multiplier norm. ∎

1.5 Sufficient Conditions for Pick's Theorem

Corollary 1.4.6 shows that the one-point extension of f from E to X can be done isometrically if and only if $\|f\|_{M(K_E^{(t)})} \leq \|f\|_{M(K_E)}$. In this section I use the positivity formulation of these norms to derive sufficient conditions for this to occur.

Lemma 1.5.1 *Let K be a positive-definite kernel on $X \times X$, t be any point of X, and $E = X \setminus \{t\}$. Then for $\|f\|_{M(K_E^{(t)})} \leq \|f\|_{M(K_E)}$ to hold for all multipliers in $M(K_E)$ it is sufficient that K is completely non-zero (i.e. $K(x,y)$ is non-zero for all $x, y \in X$) and $(1/K)_E^{(t)} \leq 0$.*

Proof:

$$\begin{aligned}
\frac{K_E^{(t)}}{K_E}(x,y) &= \frac{K(x,y) - K(x,t)K(t,y)K(t,t)^{-1}}{K(x,y)} \\
&= -\left(\frac{K(x,t)K(t,y)}{K(t,t)}\right)\left(K(x,y)^{-1} - \frac{K(x,t)^{-1}K(t,y)^{-1}}{K(t,t)^{-1}}\right) \\
&= -\left(\frac{K(x,t)K(t,y)}{K(t,t)}\right)(1/K)_E^{(t)}(x,y)
\end{aligned}$$

But the first term is a rank one positive and has a rank one positive reciprocal, so $K_E^{(t)}/K_E \geq 0$ if and only if $(1/K)_E^{(t)} \leq 0$. So whenever $r \geq \|f\|_{M(K_E)}$ we have

$$\begin{aligned}
(r^2 - f \otimes f^*)K_E^{(t)} &= (r^2 - f \otimes f^*)K_E . K_E^{(t)}/K_E \\
&= \text{pointwise product of two positive kernels} \\
&\geq 0 \qquad \text{by the Schur product theorem.}
\end{aligned}$$

Therefore $r \geq \|f\|_{M(K_E^{(t)})}$. ∎

The next lemma makes these sufficient conditions more tractable by characterising those kernels K for which $(1/K)^{(t)} \leq 0$. I am grateful to Dr. G. Naevdal and a journal referee for pointing out the proof given here, which greatly shortens my original proof.

LEMMA 1.5.2 *For $m > 1$ let A be an $m \times m$ Hermitian matrix with real and strictly positive diagonal entries. Then the Schur complement of any diagonal term is negative if and only if A has exactly one positive eigenvalue.*

Proof: We can assume, without loss of generality, that the diagonal term is the bottom right term $A_{m,m}$ since we can always re-order the rows and column of A without affecting its number of negative eigenvalues. So partition A by separating off the last row and column, and let the result be

$$A = \begin{pmatrix} B & v \\ v^* & a \end{pmatrix}$$

where B is an $(m-1) \times (m-1)$ array. Then a is real and positive (by assumption), all the matrices involved are Hermitian, and

$$A^{(m)} = \begin{pmatrix} B - vv^*/a & 0 \\ 0 & 0 \end{pmatrix}.$$

The matrix

$$C = \begin{pmatrix} I_{m-1,m-1} & -v/a \\ 0 & 1 \end{pmatrix}$$

is non-singular and direct calculation shows that

$$CAC^* = A^{(m)} + \begin{pmatrix} 0 & 0 \\ 0 & a \end{pmatrix}.$$

Now let $\kappa(A)$ denote the inertia of A, i.e. the triplet $(\kappa^-(A), \kappa^\circ(A), \kappa^+(A))$ formed from the numbers of negative, zero and positive eigenvalues of A, counting multiplicities. Then, by Sylvester's law of inertia (see, for example, reference [HJ85]), $\kappa(A) = \kappa(CAC^*)$, so

$$\kappa(A) = \kappa(CAC^*) = \kappa(A^{(m)}) + \kappa(a) = \kappa(A^{(m)}) + (0, 0, 1)$$

from which the result follows. ■

Corollary 1.5.3 *If K is a completely non-zero, positive-definite, finite kernel on a set X, then any of the following are sufficient for K to be an NP kernel.*

EITHER (i) *the matrix $[1/K]$ has exactly one positive eigenvalue.*

OR (ii) $\text{sign}(\det([1/K]_n^1)) = (-1)^{n-1}$ *for each $n = 1, \ldots, m$, where m is the order of K and $[1/K]_n^1$ denotes the principal submatrix formed from the first n rows and columns.*

OR (iii) *there exists a function $f : X \to \mathbb{C}$ such that $f \otimes f^* - 1/K \geq 0$.*

Proof: **(i)** The interlacing theorem for bordered Hermitian matrices [HJ85] states that when a Hermitian $n \times n$ matrix A is extended to a new Hermitian matrix B by adding an extra row and column, the new eigenvalues interlace the old eigenvalues in the sense that

$$\lambda_1(B) \leq \lambda_1(A) \leq \lambda_2(B) \leq \ldots \leq \lambda_n(B) \leq \lambda_n(A) \leq \lambda_{n+1}(B)$$

where λ_i denotes the i^{th} eigenvalue when they are arranged in increasing order. Hence if $[1/K]$ has exactly one positive eigenvalue, then so do all its principal submatrices. By lemmas 1.5.1 and 1.5.2, all one-point extensions are therefore achievable isometrically.

(ii) Since the principal minors, $\det([1/K]_n^1)$, are all non-zero and alternate in sign, the matrix $[1/K]$ must be non-singular and, by the interlacing theorem for bordered Hermitian matrices, must have exactly one positive eigenvalue. So (ii) $\Rightarrow$(i).

(iii) Weyl's theorem for Hermitian matrices [HJ85] states that if A and B are Hermitian $n \times n$ matrices, then for each $k = 1, \ldots, n$ the k^{th} eigenvalue of $A + B$ is contained in the closed interval $[\lambda_k(A) + \lambda_1(B), \lambda_k(A) + \lambda_n(B)]$. But $1/K = f \otimes f^* -$ (something positive) and $f \otimes f^*$ has at most one positive eigenvalue, so $1/K$ has at most one positive eigenvalue. Since $1/K$ has at least one positive eigenvalue, (iii)$\Rightarrow$(i). ∎

We can now combine lemma 1.3.3 and corollary 1.5.3 to give the main result of this chapter—sufficient conditions for the full Pick theorem to be true for $H(K)$.

THEOREM 1.5.4 *Let K be a completely non-zero, positive definite kernel on a domain set X. If*

either *all finite restrictions of $1/K$ have exactly one positive eigenvalue.*

or *all finite restrictions of $1/K$ have non-zero determinant with sign $(-1)^{n-1}$, where n is the size of the restriction.*

or *there exists a function $f : X \to \mathbb{C}$ such that $f \otimes f^* - 1/K \geq 0$.*

then the full Pick theorem is true for $H(K)$.

Chapter 2

Scalar-Valued Applications

This chapter applies the sufficient conditions for Pick's theorem that were derived in chapter 1 to prove Pick's theorem for various reproducing kernel Hilbert spaces of scalar-valued functions.

2.1 Weighted Hardy Spaces

THEOREM 2.1.1 *(Pick's theorem for certain weighted Hardy spaces) Let* $(w_n)_0^\infty$ *be a real, strictly positive weight sequence for which* $\sum_0^\infty z^n/w_n$ *is convergent and non-zero on* $\mathbb{D}$*. Further, let* $H^2(w_n)$ *denote the weighted Hardy space with weight sequence* $(w_n)_0^\infty$*, i.e. the Hilbert space of complex-valued analytic functions on the unit disc that are bounded with respect to the inner product* $\langle f, g\rangle = \sum_0^\infty w_n \hat{f}(n)\hat{g}(n)^*$*, where* $\hat{f}(n)$ *and* $\hat{g}(n)$ *are the coefficients of the power series representations of* f *and* g*. Then if the sequence* $(w_n/w_{n+1})_0^\infty$ *is non-decreasing, the full Pick theorem is true for* $H^2(w_n)$*.*

Proof: Consider the kernel $K(x,y) = \sum_0^\infty (1/w_n)(xy^*)^n$. The y-column of K satisfies

$$\langle K(\cdot,y), K(\cdot,y)\rangle = \sum_0^\infty w_n \left(\frac{(y^*)^n}{w_n}\right)\left(\frac{y^n}{w_n}\right) < \infty$$

since $\sum_0^\infty z^n/w_n$ is convergent on $\mathbb{D}$. But since $\langle x^n, K(\cdot,y)\rangle = y^n$, K has the reproducing property for the polynomials, which are dense in $H^2(w_n)$, so K has the reproducing property for all of $H^2(w_n)$. Therefore K is the reproducing kernel for $H^2(w_n)$.

Now let $\sum_0^\infty a_n(xy^*)^n$ be the expansion of $1/K(x,y)$ as a power series in xy^*—it has such a representation since $K(x,y)$ is completely non-zero. It is easy to verify that the coefficients a_n are related to the weights w_n by the recurrence relations

$$a_0 = w_0, \qquad 0 = \frac{a_n}{w_0} + \frac{a_{n-1}}{w_1} + \frac{a_{n-2}}{w_2} + \ldots + \frac{a_0}{w_n} \qquad \text{for } n \geq 1.$$

So $a_0 > 0$, and the condition that $(w_n/w_{n+1})_0^\infty$ is non-decreasing implies that $a_n \leq 0$ for $n \geq 1$. We will verify this using proof by induction, the result being true for n=1 since $a_1 = -w_0a_0/w_1 < 0$. Assume $a_i \leq 0$ for $1 \leq i \leq n$. Then for $n \geq 1$

$$0 = \frac{a_n}{w_0} + \frac{a_{n-1}}{w_1} + \ldots + \frac{a_0}{w_n} \leq \frac{a_n}{w_1} + \frac{a_{n-1}}{w_2} + \ldots + \frac{a_0}{w_{n+1}}$$

since $(w_n/w_{n+1})_0^\infty$ is non-decreasing, and therefore the change made involves scaling the non-positive terms (all except the last term) by no more than the factor applied to the last term (which is the only positive one). Hence

$$0 = \frac{a_{n+1}}{w_0} + \frac{a_n}{w_1} + \frac{a_{n-1}}{w_2} + \ldots + \frac{a_0}{w_{n+1}} = \frac{a_{n+1}}{w_0} + \text{ something non-negative}$$

and therefore $a_{n+1} \leq 0$. Therefore, by induction, $a_n \leq 0$ for $n \geq 1$.

Now, with f denoting the constant function $\sqrt{w_0}$, $1/K$ satisfies

$$f \otimes f^* - 1/K = \sum_1^\infty (-a_n)(xy^*)^n \geq 0$$

since the right hand side is the limit of a sum of positive rank-1 kernels. K therefore satisfies the conditions of theorem 1.5.4, so the full Pick theorem is true for $H(K)$. ■

Corollary 2.1.2 *The full Pick theorem is true for the Hardy space H^2 on $\mathbb{D}$ and for the Dirichlet space on $\mathbb{D}$.*

Proof: The case of H^2 is the original Pick theorem, and it is now a direct corollary of theorem 2.1.1, since this is the weighted Hardy space with $w_n = 1$ for all n. Hence w_n/w_{n+1} is non-decreasing and $\sum_0^\infty z^n/w_n = 1/(1-z)$ is non-zero on $\mathbb{D}$, so the conditions of theorem 2.1.1 are satisfied.

Agler proves the finite Pick theorem for the Dirichlet space in [Agl88], in which he develops many of the ideas expounded above. It is now a direct corollary of theorem 2.1.1, since the Dirichlet space is the weighted Hardy space with $w_n = n+1$, for which w_n/w_{n+1} is non-decreasing, and the function

$$\sum_0^\infty z^n/(n+1) = \left(\frac{1}{z}\right)\log\left(\frac{1}{1-z}\right)$$

is non-zero on $\mathbb{D}$. ■

2.2 Weighted Sobolev Spaces

In order to obtain the second application of theorem 1.5.4, we first need the following technical lemma and its corollary.

Lemma 2.2.1 *If A is an $n \times n$ symmetric matrix such that $A_{ij} = f_i g_j$ for all $j \geq i$ (and hence $A_{ij} = A_{ji} = f_j g_i$ for all $i \leq j$) and $g_i \neq 0$ for all i, then*

$$\det([A]) = f_1 g_n \prod_{i=2}^{n} (f_i g_{i-1} - f_{i-1} g_i)$$

Proof: The proof is by induction, the result being clear for $n = 1$. We therefore assume the result for $n-1$ and show this implies the result for n. Let A_n denote the matrix for size n, and $D_n = \det(A_n)$. Since the matrix is of the form

$$A_n = \begin{pmatrix} f_1 g_1 & \cdots & f_1 g_{n-1} & f_1 g_n \\ \cdots & \cdots & \cdots & \cdots \\ f_1 g_{n-1} & \cdots & f_{n-1} g_{n-1} & f_{n-1} g_n \\ f_1 g_n & \cdots & f_{n-1} g_n & f_n g_n \end{pmatrix}$$

then using the standard determinant formula, going backwards along the last row, gives

$$\begin{aligned} D_n &= f_n g_n \det(A_{n-1}) \\ &- f_{n-1} g_n \det(A_{n-1} \text{with last column multiplied by } g_n/g_{n-1}) \\ &+ \sum_{i=1}^{n-2} \det(\text{matrix with last two columns linearly dependent}) \\ &= f_n g_n D_{n-1} - \frac{f_{n-1} g_n^2}{g_{n-1}} D_{n-1} + 0 \\ &= \left(f_n g_n - \frac{f_{n-1} g_n^2}{g_{n-1}} \right) f_1 g_{n-1} \prod_{i=2}^{n-1} (f_i g_{i-1} - f_{i-1} g_i) \\ &= f_1 g_n \prod_{i=2}^{n} (f_i g_{i-1} - f_{i-1} g_i) \end{aligned}$$

The induction step is therefore proved. ■

Corollary 2.2.2 *Let $H(K)$ be a reproducing kernel Hilbert space on a totally ordered domain X, with kernel of the form*

$$K(x,y) = \begin{cases} f(x)g(y) & \text{if } x \le y \\ f(y)g(x) & \text{if } y \le x \end{cases}$$

If f and g are strictly positive real-valued and f/g is strictly increasing, then the restriction of the kernel $1/K$ to the n points $\{x_1, \ldots, x_n\}$ satisfies

$$\text{sign}(\det([1/K(x_i, x_j)])) = (-1)^{n-1}.$$

Proof: First order the x_i's so that they are in increasing order. Swapping two x's corresponds to swapping their rows and columns simultaneously, so this does not alter the determinant. Then using the above formula,

$$\begin{aligned} \det([1/K]) &= \left(\frac{1}{f(x_1)g(x_n)} \right) \prod_{i=2}^{n} \left(\frac{1}{f(x_i)g(x_{i-1})} - \frac{1}{f(x_{i-1})g(x_i)} \right) \\ &= (+ve) \times \prod_{i=2}^{n} (-ve) \end{aligned}$$

since

$$\begin{aligned}
\operatorname{sign}\left(\frac{1}{f(x_i)g(x_{i-1})} - \frac{1}{f(x_{i-1})g(x_i)}\right) &= \operatorname{sign}\left(\frac{g(x_i)}{f(x_i)} - \frac{g(x_{i-1})}{f(x_{i-1})}\right) \\
&= \text{negative}
\end{aligned}$$

(as f and g are strictly positive, and f/g is strictly increasing).

Therefore $\operatorname{sign}(\det([1/K])) = (-1)^{n-1}$. ■

We can now give the second application of theorem 1.5.4.

THEOREM 2.2.3 *(Pick's theorem for a weighted Sobolev space) Let $w_0(x)$ and $w_1(x)$ be real, strictly positive, continuous functions on $[0,1]$, and let $w_1(x)$ be continuously differentiable. Further, let $H(K)$ denote the weighted Sobolev space of complex-valued, absolutely continuous functions on $[0,1]$, that have derivatives in $L^2[0,1]$ and which are bounded with respect to the inner product*

$$\langle f, g\rangle = \int_0^1 w_0(x) f(x) g(x)^* dx + \int_0^1 w_1(x) f'(x) g'(x)^* dx.$$

Then the full Pick theorem is true for the weighted Sobolev space $H(K)$.

Proof: Note that in [Agl90] Agler proves the finite Pick theorem for the special case where $w_0(x) = w_1(x) = 1$, but by a different method.

Under the conditions placed on the weight functions, the space defined is indeed a reproducing kernel Hilbert space. Let K be its kernel and consider the functions $f \in H(K)$ for which $f'(0) = f'(1) = 0$. Then

$$\begin{aligned}
f(y) &= \langle f, K_y\rangle \\
&= \int_0^1 w_0 f K_y^* + \int_0^1 w_1 f' K_y'^* \\
&= \int_0^1 w_0 f K_y^* + [w_1 f' K_y^*]_0^1 - \int_0^1 K_y^* (w_1 f')' \text{ (integration by parts)} \\
&= \int_0^1 (-K_y^*) g \qquad \text{since } f'(0) = f'(1) = 0
\end{aligned}$$

where $g = (w_1 f')' - w_0 f$.

This suggests that $-K(x,y)^*$ is the Green's function for the boundary-value problem

$$Lf = g \ , \ f'(0) = f'(1) = 0$$

where L is the linear differential operator $Lf = (w_1 f')' - w_0 f$. This is a regular Sturm-Liouville problem and the extensive theory of such problems (see for example [You88]) gives detailed information about its associated Green's function $G(x,y)$. In particular G must be of the form

$$G(x,y) = \begin{cases} u(x)v(y) & \text{if } x \leq y \\ u(y)v(x) & \text{if } y \leq x \end{cases}$$

where u and v are real and differentiable, and $w_1(uv' - vu') = 1$.

From this it can be verified that the kernel $-G$ does indeed satisfy the requirements for being the reproducing kernel for $H(K)$. That is, its columns $-G_y$ are all members of $H(K)$, and it has the reproducing property for $H(K)$. Therefore $K(x,y) = -G(x,y)$. Although this does not give an explicit expression for $K(x,y)$, we can now use the information we have about G to show that K satisfies the conditions of theorem 1.5.4.

Because $H(K)$ contains the constant function 1, which is everywhere non-zero, no column of K can be completely zero. But K is a positive kernel, so $K(x,x)$ must be non-zero on $[0,1]$, and therefore u and v must also be non-zero on $[0,1]$. Since u and v are differentiable, they must both have constant sign, and since K is positive they must have opposite signs.

Now let $f = |u|$ and $g = |v|$. Then

$$K(x,y) = \begin{cases} f(x)g(y) & \text{if } x \leq y \\ f(y)g(x) & \text{if } y \leq x \end{cases}$$

and $(f/g)' = -(u/v)' = (uv' - vu')/v^2 = 1/(w_1 v^2) > 0$, so f/g is strictly increasing. Therefore, by corollary 2.2.2, K satisfies the condition of 1.5.4, so the full Pick theorem is true for $H(K)$. ■

2.3 Other Cases

Lastly, the third application of theorem 1.5.4.

Theorem 2.3.1 *Let ρ be any real, positive, continuous and integrable function on the interval (a, b). Then the full Pick theorem is true for the reproducing kernel Hilbert space of absolutely continuous, complex-valued functions on (a, b) that satisfy $\lim_{x\to a} f(x) = 0$ and which are bounded with respect to the inner product $\langle f, g\rangle = \int_a^b f'(x)g'^*(x)/\rho(x)dx$.*

Proof: This space is considered by Saitoh [Sai88, theorem 5.3], who gives the reproducing kernel. The conditions of theorem 1.5.4 are easily checked:

- the reproducing kernel for this space is $K(x, y) = \int_a^{\min(x,y)} \rho(t)dt$ which is completely non-zero on (a, b).
- since
$$K(x, y) = \begin{cases} \int_a^x \rho(t)dt.1 & \text{if } x \le y \\ 1.\int_a^y \rho(t)dt & \text{if } y \le x \end{cases}$$
and $\int_a^x \rho(t)dt$ and the constant function 1 are strictly positive, and $\int_a^x \rho(t)dt/1$ is strictly increasing on $[0, 1]$, corollary 2.2.2 to lemma 2.2.1 shows that every finite restriction of $1/K$ has a non-zero determinant with sign $(-1)^{n-1}$ (where n is the restriction's size). ■

Finally, some comments on the conditions (i) and (ii) in theorem 1.5.4; (i) is simply a restriction of the approach—there *do* exist positive definite kernels with zero entries for which Pick's theorem is true, for example the identity kernel on $\{1, 2, 3\}$. But amongst positive definite kernels satisfying (i), it is not clear from our analysis so far whether condition (ii) is necessary as well as sufficient. In extensive computer experiments with kernels that satisfy (i) but fail (ii) I found that all kernels tested also failed Pick's theorem, counter examples being easy to find. The following result gives examples of this. However, we will see in chapter 5 that condition (ii) is in fact *not* necessary.

RESULT 2.3.2 *The finite Pick theorem (and hence also the full Pick theorem) does NOT hold for the following reproducing kernel Hilbert spaces:*

1. *The Bergman space of complex-valued analytic functions on the unit disc that are square-integrable with respect to normalised area measure dA, with inner product given by $\langle f, g\rangle = \int f g^* dA$.*

2. *The space $H^2(\mathbb{D}^2)$ of functions that are analytic and square-integrable on the bi-disc.*

3. *the Sobolev space of complex-valued functions on $[0, 1]$ with inner product given by*

$$\langle f, g\rangle = \int_0^1 f(x)g(x)^* dx + \int_0^1 f'(x)g'(x)^* dx + \int_0^1 f''(x)g''(x)^* dx$$

Proof: 1. In the terminology of theorem 2.1.1 this Bergman space is the weighted Hardy space with weights $w_n = 1/(n+1)$, for which w_n/w_{n+1} is decreasing. Its kernel is $K(x, y) = 1/(1 - xy^*)^2$ [Axl88]. By lemma 1.4.2, it is sufficient to show that there exists a finite subset E of $\mathbb{D}$, a point $t \in \mathbb{D} \setminus E$, and a multiplier ϕ such that $\|\phi\|_{M(K_E^{(t)})} > \|\phi\|_{M(K_E)}$, where $X = E \cup \{t\}$. Computer trials quickly found the following example of this:

$$\begin{aligned} E &= \{x_1 = 0.0, x_2 = 0.5\} & t &= 0.3 \\ \phi(x_1) &= +1.0 & \phi(x_2) &= -1.0 \\ \|\phi\|_{M(K_E)} &= 2.65 & \|\phi\|_{M(K_E^{(t)})} &= 3.05. \end{aligned}$$

2. The kernel of this space is

$$K((x_1, x_2), (y_1, y_2)) = \frac{1}{(1 - x_1 y_1^*)(1 - x_2 y_2^*)}$$

and its restriction to the diagonal set in $\mathbb{D}^2$ is

$$K((x,x),(y,y)) = 1/(1-xy^*)^2$$

which is the same as the kernel of the Bergman space considered in (1). Since Pick's theorem fails for this Bergman space, this restriction is NOT an NP kernel, and so therefore neither is the kernel of $H^2(\mathbb{D}^2)$. Note that Agler's unpublished paper [Agl88] goes on to derive the correct variant of Pick's theorem, with the matrix condition modified, that is valid for this space.

3. I have calculated the kernel for this space—it is a sum of exponentials that is rather too complex to document here. Again, computer calculations with small restrictions of this kernel easily find examples of multipliers ϕ for which $\|\phi\|_{M(K_E^{(t)})} > \|\phi\|_{M(K_E)}$. ∎

Chapter 3

Uniqueness of the Optimal Interpolant

For the original Nevanlinna-Pick problem (i.e. when K is the Szegö kernel, $H(K)$ is H^2 and $M(K)$ is H^∞) there have been previous studies of when the minimal norm interpolating multiplier is unique [Den29, Wal56, Sar67]. The answer is that uniqueness always holds when the set of data points E is finite, but may or may not hold when E is infinite, depending on the data values. In fact when E is infinite there are always some sets of data values for which uniqueness holds and some for which uniqueness fails. This section examines the uniqueness question for general reproducing kernel Hilbert spaces $H(K)$ for which Pick's theorem holds.

The fact that uniqueness holds if K is the Szegö kernel and E is finite is fairly easy to prove, and the method does give some information for other NP kernels, so I will outline the argument for the general case. We know that the smallest possible norm of an interpolating multiplier $\phi \in M(K)$ (i.e. such that $\phi|E = f$) is $\|\tilde{M}_f\|$, where $\tilde{M}_f$ is the operator on $H(K, E)$ given by $\tilde{M}_f^* : K(\cdot, y) \to f(y)^* K(\cdot, y)$, $y \in E$. The reason for this is that

$$\phi|E = f \Leftrightarrow M_\phi^* \text{ is an extension of } \tilde{M}_f^* \Leftrightarrow M_\phi \text{ is a lifting of } \tilde{M}_f$$

and we know that, since K is NP, $\tilde{M}_f^*$ can be isometrically extended to a multiplication operator adjoint.

Let ϕ be a minimal norm solution, i.e. $\|M_\phi\| = \|\tilde{M}_f\|$, and suppose $\tilde{M}_f$ has a maximising vector g. Then, since M_ϕ is an isometric lifting of $\tilde{M}_f$, g must also be maximising for M_ϕ and the action of $\tilde{M}_f$ and M_ϕ on g must be the same, i.e. $\phi g = M_\phi g = \tilde{M}_f g$. Therefore ϕ is determined by $\tilde{M}_f$ on g's support—it must equal $\tilde{M}_f g/g$—and so ϕ is uniquely determined by $\tilde{M}_f$ on the union of the supports of $\tilde{M}_f$'s maximising vectors. In particular if $\tilde{M}_f$ has a completely non-zero maximising vector then the minimal norm interpolating multiplier ϕ must be unique.

Now let K be the Szegö kernel and E be finite. Then $\tilde{M}_f$ has finite rank and so must have a maximising vector, g. If $g = 0$ then $f = \phi = 0$ and ϕ is unique. Otherwise, since the support of a non-zero H^2 function is dense in $\mathbb{D}$, ϕ is determined uniquely on a dense subset and so, being analytic, on all of $\mathbb{D}$. There is therefore a unique minimal norm solution when K is the Szegö kernel and E is finite, and it is given, on a dense subset of $\mathbb{D}$, by $\tilde{M}_f g/g$ where g is any maximising vector of $\tilde{M}_f$.

For the case where E is infinite $\tilde{M}_f$ may not have any maximising vectors, so the above method breaks down. It also breaks down for a general NP kernel since then $H(K)$ does not have the special properties of H^2 such as non-zero functions having dense support.

3.1 Uniqueness for One Point Extension

For NP kernels, we can study uniqueness of the minimal norm interpolating multiplier by examining when uniqueness holds for one-step extension problems. For if there were two different isometric extensions of f to X, ϕ_1 and ϕ_2, which differed at t say, then $\phi_1|E \cup \{t\}$ and $\phi_2|E \cup \{t\}$ would be two different isometric extensions of f from E to $E \cup \{t\}$. Conversely, if f_1 and f_2 were two different isometric extensions of f from E to $E \cup \{t\}$ for some $t \in X \setminus E$, then since K is NP we could isometrically extend f_1 and f_2 to all of X and so obtain two different isometric extensions to X.

We must therefore ask 'when is the isometric one-point extension of f to $E \cup \{t\}$ unique?'. From section 1.4 we know that a general one-point extension problem is effectively a non-orthogonal Parrott completion problem and for *orthogonal* Parrott completion problems there is a well-known parameterisation of the contractive solutions. The following lemma uses this to characterise when uniqueness holds in finite-dimensional (not necessarily orthogonal) completion problems where the completing operator has one-dimensional domain and target spaces. Note that $\mathcal{D}_W$ denotes the defect operator $(I - W^*W)^{1/2}$ of an operator W, ran(W) denotes its range and max(W) its space of maximising vectors.

LEMMA 3.1.1 *As in lemma 1.4.1, let $H = H_1 + H_2$ and $J = J_1 + J_2$ be (not necessarily orthogonal) decompositions of Hilbert spaces H and J into linearly independent non-trivial subspaces H_1, H_2 and J_1, J_2, and let $A \in \mathcal{L}(H_1, J_1)$, $B \in \mathcal{L}(H_2, J_1)$, and $C \in \mathcal{L}(H_1, J_2)$ be any given operators. Further, assume that H and J are finite-dimensional and H_2 and J_2 are 1-dimensional. Then there is a unique operator $D \in \mathcal{L}(H_2, J_2)$ for which $T_D = \begin{pmatrix} A & B \\ C & D \end{pmatrix} : H \to J$ has smallest possible norm if and only if* $\max(R) \neq \max(S)$ *where*

$$\begin{aligned} R &= \textit{common restriction of all completions } T_D \textit{ to } H_1 \\ \textit{and } S &= \textit{common projection of all completions } T_D \textit{ onto } J_2^{\perp} \end{aligned}$$

are as in lemma 1.4.1.

Proof: By lemma 1.4.1 the minimum norm of any completion is $\max(\|R\|, \|S\|)$. The result holds when $\max(\|R\|, \|S\|) = 0$, since then the zero operator is the unique completion of minimum norm, and $\max(R) = H_1 \neq H = \max(S)$, so assume (by scaling the problem if necessary) that $\max(\|R\|, \|S\|) = 1$. Then the minimum norm is 1 and the question is whether there is a unique contractive completion.

As in lemma 1.4.1, and using the same notation, the given completion problem is equivalent to the orthogonal problem of completing

$$T_D = \begin{pmatrix} A' & B' \\ C' & D' \end{pmatrix} : H_1 \oplus H_1^{\perp} \to J_2^{\perp} \oplus J_2$$

by choice of D'. The well-known parameterisation of the solutions [You88, page 152] tells us that the set of contractive completions are given by

$$D' = \mathcal{D}_{Z^*} V \mathcal{D}_Y - Z A'^* Y$$

where $Y : H_1^{\perp} \to J_2^{\perp}$ and $Z : H_1 \to J_2$ are given by

$$\begin{aligned} B'^* &= Y^* \mathcal{D}_{A'^*} \text{ and } Y^* = 0 \text{ on } \operatorname{ran}(\mathcal{D}_{A'^*})^{\perp} \\ C' &= Z \mathcal{D}_{A'} \text{ and } Z = 0 \text{ on } \operatorname{ran}(\mathcal{D}_{A'})^{\perp} \end{aligned}$$

and $V : H_1^{\perp} \to J_2$ is an arbitrary contraction.

The result is now proved by the following sequence of implications, the last few of which are justified afterwards. Uniqueness fails

$$\begin{aligned}
&\Leftrightarrow \quad \mathcal{D}_{Z^*} \neq 0 \text{ and } \mathcal{D}_Y \neq 0 \\
&\Leftrightarrow \quad \text{neither } Z^* \text{ nor } Y \text{ is an isometry} \\
&\Leftrightarrow \quad \|Z\| < 1 \text{ and } \|Y^*\| < 1 \\
&\qquad\; \text{(since } Z^* \text{ and } Y \text{ have 1-dimensional domain spaces)} \\
&\Leftrightarrow \qquad \|\mathcal{D}_{A'} u\|^2 > \|C' u\|^2 \text{ for all } u \in H_1 \setminus \ker(C') \\
&\qquad\; \text{and } \|\mathcal{D}_{A'^*} v\|^2 > \|B'^* v\|^2 \text{ for all } v \in J_2^{\perp} \setminus \ker(B'^*) \\
&\Leftrightarrow \qquad \|u\|^2 > \|A' u\|^2 + \|C' u\|^2 \text{ for all } u \in H_1 \setminus \ker(C') \\
&\qquad\; \text{and } \|v\|^2 > \|A'^* v\|^2 + \|B'^* v\|^2 \text{ for all } v \in J_2^{\perp} \setminus \ker(B'^*) \qquad (1) \\
&\Leftrightarrow \qquad \|A'\| = 1 \\
&\qquad\; \text{and } \max \begin{pmatrix} A' \\ C' \end{pmatrix} = \max(A') \\
&\qquad\; \text{and } \max \begin{pmatrix} A'^* \\ B'^* \end{pmatrix} = \max(A'^*) \qquad (2)
\end{aligned}$$

$$\Leftrightarrow \quad \max\begin{pmatrix} A' \\ C' \end{pmatrix} = \max\begin{pmatrix} A' & B' \end{pmatrix} \tag{3}$$

$$\Leftrightarrow \quad \max(R) = \max(S).$$

This completes the proof, except that some of the later implications here need explanation.

$(1) \Rightarrow (2)$: Since $\max(\|\begin{pmatrix} A' \\ C' \end{pmatrix}\|, \|\begin{pmatrix} A'^* \\ B'^* \end{pmatrix}\|) = \max(\|R\|, \|S\|) = 1$ there exists either a non-zero vector $u \in H_1$ such that $\|u\|^2 = \|A'u\|^2 + \|C'u\|^2$ or else a non-zero vector $v \in J_2^\perp$ such that $\|v\|^2 = \|A'^*v\|^2 + \|B'^*v\|^2$. In the former case $(1) \Rightarrow u \in \ker(C') \Rightarrow \|A'\| = 1$, and in the latter case $(1) \Rightarrow v \in \ker(B'^*) \Rightarrow \|A'^*\| = 1$, so either way $\|A'\| = \|\begin{pmatrix} A' \\ C' \end{pmatrix}\| = \|\begin{pmatrix} A'^* \\ B'^* \end{pmatrix}\| = 1$. But then any vector that maximises A' also maximises $\begin{pmatrix} A' \\ C' \end{pmatrix}$ and $(1) \Rightarrow$ the converse is also true, so $\max\begin{pmatrix} A' \\ C' \end{pmatrix} = \max(A')$. Similarly $(1) \Rightarrow \max\begin{pmatrix} A'^* \\ B'^* \end{pmatrix} = \max(A'^*)$.

$(2) \Rightarrow (1)$: If $\|A'\| = 1$ then $\|\begin{pmatrix} A' \\ C' \end{pmatrix}\| = 1$ and so $\max\begin{pmatrix} A' \\ C' \end{pmatrix} = \max(A')$ implies that any vector that maximises $\begin{pmatrix} A' \\ C' \end{pmatrix}$ also maximises A' and so must be annihilated by C'. Similarly, (2) implies that any vector that maximises $\begin{pmatrix} A'^* \\ B'^* \end{pmatrix}$ must be annihilated by B'^*. These two together give condition (1).

$(2) \Rightarrow (3)$: (2) implies

$$\begin{aligned} \max\begin{pmatrix} A' & B' \end{pmatrix} &= \begin{pmatrix} A'^* \\ B'^* \end{pmatrix} \max\begin{pmatrix} A'^* \\ B'^* \end{pmatrix} \\ &= A'^* \max(A'^*) \qquad \text{since } \begin{pmatrix} A'^* \\ B'^* \end{pmatrix} = A'^* \text{ on } \max(A'^*) \\ &= \max(A') \\ &= \max\begin{pmatrix} A' \\ C' \end{pmatrix}. \end{aligned}$$

$(3) \Rightarrow (2)$:

$$(3) \quad \Rightarrow \quad \max\begin{pmatrix} A' & B' \end{pmatrix} \subseteq H_1$$

$$\Rightarrow \begin{pmatrix} A' & B' \end{pmatrix} = A' \text{ on } \max\begin{pmatrix} A' & B' \end{pmatrix}$$
$$\Rightarrow \max\begin{pmatrix} A' \\ C' \end{pmatrix} = \max\begin{pmatrix} A' & B' \end{pmatrix} = \max(A')$$
$$\text{and } \max\begin{pmatrix} A'^* \\ B'^* \end{pmatrix} = A'\max(A') = \max(A'^*).$$

But then $\|A'\| = \|\begin{pmatrix} A' \\ C' \end{pmatrix}\| = \|\begin{pmatrix} A' & B' \end{pmatrix}\|$ so all three norms must equal 1. ■

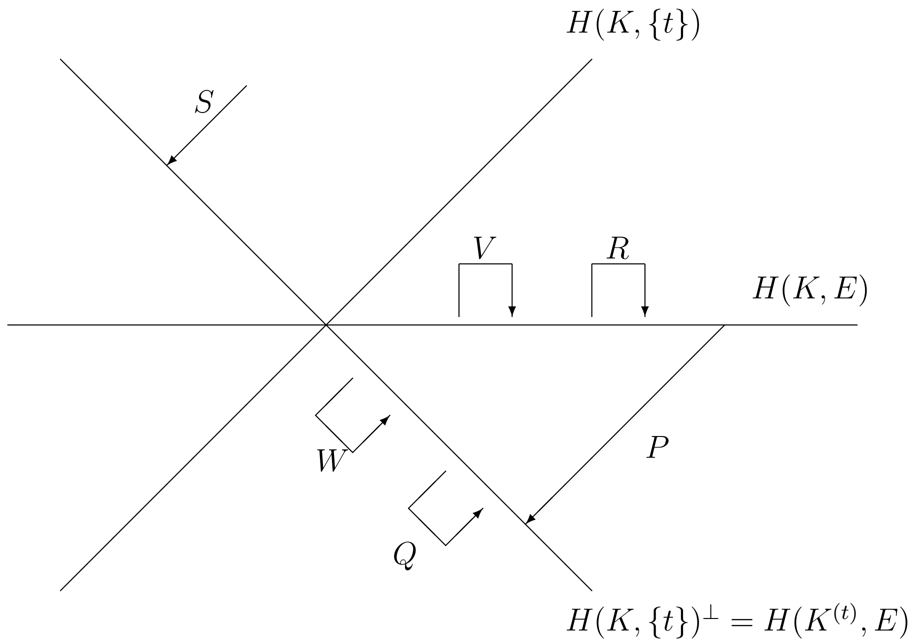
We can now give a proof that uniqueness holds for the Szegö kernel, using reproducing kernel methods.

LEMMA 3.1.2 *Let $K : X \times X \to \mathbb{C}$ be the restriction of the Szegö kernel to a finite subset X of $\mathbb{D}$. Then for any $t \in X$ and any function $f : E \to \mathbb{C}$, where $E = X \setminus \{t\}$, there is a unique minimal norm extension of f from E to X.*

Proof: As in lemma 1.4.2 we know this multiplier extension problem is equivalent to the problem of finding a minimal norm completion

$$M_\phi^* = \begin{pmatrix} \tilde{M}_f^* & 0 \\ 0 & \phi(t)^* \end{pmatrix} : H(K,E) + H(K,\{t\}) \to H(K,E) + H(K,\{t\})$$

by choice of $\phi(t)^*$. By lemma 3.1.1, to show uniqueness we need to show that $\max(R) \neq \max(S)$, where $R = M_\phi^*|H(K,E)$, $S = P_{H(K,\{t\})^\perp} M_\phi^*$ and ϕ is any extension of f to X. We can assume $f \neq 0$ since otherwise the zero function on X is the unique minimal norm extension of f.



Now consider the various operators that are shown schematically in the above diagram of $H(K)$, where P is the orthogonal projection of $H(K,E)$ onto $H(K,\{t\})^{\perp}$ and $V \in \mathcal{L}(H(K,E))$, $W \in \mathcal{L}(H(K,\{t\})^{\perp})$ are given by

$$\begin{aligned} V &= P_{H(K,E)}R = \text{ compression of } M_{\phi}^* \text{ onto } H(K,E) \\ \text{and } W &= S|H(K,\{t\})^{\perp} = \text{ compression of } M_{\phi}^* \text{ onto } H(K,\{t\})^{\perp} \end{aligned}$$

where ϕ is any extension of f. Then $\max(R) = \max(V)$, since R leaves $H(K,E)$ invariant, and $\max(S) = \max(W)$, since S annihilates $H(K,\{t\})^{\perp}$. We therefore need to show $\max(V) \neq \max(W)$.

As in lemma 1.4.2, $K^{(t)}$'s columns are the orthogonal projections of K's columns onto $H(K,\{t\})^{\perp}$ and the E-columns of $K^{(t)}$ span $H(K,\{t\})^{\perp}$, i.e. $K_y^{(t)} = PK_y$ and $H(K,\{t\})^{\perp} = H(K^{(t)},E)$.

Now by direct calculation with the Szegö kernel we find that

$$\begin{aligned} K^{(t)}(x,y) &= (1-xy^*)^{-1} - \frac{(1-xt^*)^{-1}(1-ty^*)^{-1}}{(1-tt^*)^{-1}} \\ &= \frac{(1-xt^*)(1-ty^*) - (1-xy*)(1-tt^*)}{(1-xt^*)(1-ty^*)}K(x,y) \\ &= \frac{xy^* + tt^* - xt^* - ty^*}{(1-xt^*)(1-ty^*)}K(x,y) \end{aligned}$$

$$
\begin{aligned}
&= \left(\frac{x-t}{1-xt^*}\right)\left(\frac{y-t}{1-yt^*}\right)^* K(x,y) \\
&= b_t(x)b_t(y)^*K(x,y)
\end{aligned}
$$

where $b_t(x) = (x-t)/(1-xt^*)$ denotes the Blaschke factor associated with the point t.

This shows that $K^{(t)}$ is a rank-1 positive Schur multiple of K, so the vectors $\{K_y^{(t)} : y \in E\}$, which are non-zero since K is completely non-zero, can be re-scaled so that they have Gram matrix equal to K_E. Let Q be the operator on $H(K,\{t\})^\perp$ that does this rescaling. Then the vector sets $\{QK_y^{(t)} : y \in E\}$ and $\{K_y : y \in E\}$ have the same Gram matrices and therefore the operator

$$
U : H(K,E) \to H(K,\{t\})^\perp \qquad U = QP = K_y \mapsto QK_y^{(t)}
$$

is unitary.

But we have

$$
V \in \mathcal{L}(H(K,E)) \qquad V = K_y \mapsto f(y)^* K_y \quad y \in E
$$

and

$$
\begin{aligned}
W \in \mathcal{L}(H(K^{(t)},E)) \qquad W &= K_y^{(t)} \mapsto f(y)^* K_y^{(t)} \quad y \in E \\
&= UK_y \mapsto f(y)^* UK_y \quad y \in E.
\end{aligned}
$$

Therefore, since V and W have the same eigenvalues and U maps V's eigenvector to W's eigenvectors, we have $V = U^{-1}WU$ and so V and W are unitarily equivalent via the unitary operator $U = QP$.

Now assume the opposite of what we want to prove, i.e. that $\max(V) = \max(W)$, and let M denote this common maximising space. Then U must leave M invariant (since it must map V's maximising space to W's maximising space) and so U must have an eigenvector $v \in M$, with corresponding unimodular eigenvalue λ say. But since $M = \max(W) \subseteq H(K,\{t\})^\perp$, P

has no effect on M, so v must also be an eigenvector of Q with the same unimodular eigenvalue λ.

But Q already has a spanning set of eigenvectors, namely $\{K_y^{(t)} : y \in E\}$, and the corresponding eigenvalues are all greater than 1 in modulus since

$$\begin{aligned} \|QK_y^{(t)}\| &= \|K_y\| && (\text{since } \{QK_y^{(t)} : y \in E\} \text{ has Gram matrix } K) \\ &> \|PK_y\| && (K(t,y) \neq 0, \text{ so } K_y \text{ is not orthogonal to } K_t) \\ &= \|K_y^{(t)}\|. \end{aligned}$$

Therefore Q cannot also have λ as an eigenvalue. This contradiction proves that the assumption that $\max(V) = \max(W)$ must be wrong, so completing the proof. ∎

Since the Szegö kernel is NP, uniqueness of one-point extensions implies uniqueness of extensions to all of $\mathbb{D}$, so we have the following corollary.

Corollary 3.1.3 *Let $K : \mathbb{D} \times \mathbb{D} \to \mathbb{C}$ be the Szegö kernel and $f : E \to \mathbb{C}$ be a given function on a non-empty, finite subset of $\mathbb{D}$. Then there is a unique minimal norm extension of f from E to $\mathbb{D}$.*

3.2 Blaschke Kernels

In our proof that uniqueness holds for the Szegö kernel, lemma 3.1.2, the only special properties of the Szegö kernel that we used, other than it being NP, were that it is completely non-zero and that $K^{(t)}$ is a rank 1 positive Schur multiple of K. It is this latter property that is most important—a lot of the rich structure of H^2 stems from this property of its kernel—so we now consider kernels of this type. We saw in the case of the Szegö kernel that the rank 1 Schur multiplier is formed from the Blaschke factor b_t, so I will call these the Blaschke kernels.

Definition 3.2.1 *A kernel* $K : X \times X \to \mathbb{C}$ *is a Blaschke kernel if and only if for each* $t \in X$ *there exists a function* $b_t : X \to \mathbb{C}$ *such that* $K^{(t)} = (b_t \otimes b_t^*)K$*. The functions* b_t *are then called the generalised Blaschke factors associated with the points of* X*.*

The following two lemmas show what form the Blaschke kernels take.

Lemma 3.2.2 *Let* $K : X \times X \to \mathbb{C}$ *be a positive-definite kernel. Then there exist kernels* $L[t]$ *such that* $K^{(t)} = L[t]K$ *for each* $t \in X$ *if and only if, for some ordering of* X*,* K *is block diagonal with completely non-zero blocks.*

Proof: If such kernels $L[t]$ exist then consider the graph $G(K)$ that has the points of X as vertices and has an edge connecting x to y whenever $K(x, y) \neq 0$. From the equation

$$K^{(t)} = K(x, y) - K(x, t)K(y, t)^*/K(t, t) = L[t](x, y)K(x, y)$$

we see that if $K(x, y) = 0$ then at least one of $K(x, t)$ and $K(y, t)$ must also be zero, so $G(K)$ has the property that any pair of vertices connected via a third are also connected directly. $G(K)$ is therefore a union of disjoint cliques, a clique being a sub-graph with an edge joining every pair of vertices. Translating this back into the location of zeros in K, this shows that for some ordering of X, K is block diagonal with completely non-zero blocks, as claimed.

The reverse implication is clear, since if K is block diagonal with completely non-zero blocks, then we can define $L[t]$ to be $K^{(t)}/K$ on t's block and $L[t](x, y) = 1$ elsewhere. ∎

LEMMA 3.2.3 *A positive-definite kernel $K : X \times X \to \mathbb{C}$ is a Blaschke kernel if and only if, for a suitable ordering of X, K is block diagonal with completely non-zero blocks, each of which is of the form $1/(p \otimes p^* - q \otimes q^*)$ where p and q are functions on that block's domain.*

Proof: First assume K is a Blaschke kernel and is therefore, by the previous lemma, block diagonal with completely non-zero blocks. Consider one of those blocks, K_E say. For $x, y \in E$ we can rearrange $K^{(t)} = (b_t \otimes b_t^*)K$ as

$$K(x,y) = \frac{K(x,t)K(y,t)^*/K(t,t)}{1 - b_t(x)b_t(y)^*} = \left(\frac{1}{p \otimes p^* - q \otimes q^*}\right)(x,y)$$

where

$$\begin{aligned} p(x) &= \sqrt{K(t,t)}/K(x,t) \\ q(x) &= b_t(x)p(x). \end{aligned}$$

K therefore has the claimed form.

Now assume that K is block diagonal and that each block K_E has the form $K_E = 1/(p \otimes p^* - q \otimes q^*)$. By lemma 3.2.2, $K^{(t)} = L[t]K$ for some kernel $L[t]$ and on the block containing t we have

$$\begin{aligned} & L[t](x,y) = K^{(t)}(x,y)/K(x,y) \\ = \; & 1 - \frac{(p(x)p(y)^* - q(x)q(y)^*)(p(t)p(t)^* - q(t)q(t)^*)}{(p(x)p(t)^* - q(x)q(t)^*)(p(t)p(y)^* - q(t)q(y)^*)} \\ = \; & \frac{p(x)p(y)^*q(t)q(t)^* + q(x)q(y)^*p(t)p(t)^* - p(x)q(y)^*p(t)^*q(t) - p(y)^*q(x)p(t)q(t)^*}{(p(x)p(t)^* - q(x)q(t)^*)(p(t)p(y)^* - q(t)q(y)^*)} \\ = \; & \left(\frac{p(x)q(t) - q(x)p(t)}{p(x)p(t)^* - q(x)q(t)^*}\right)\left(\frac{p(y)q(t) - q(y)p(t)}{p(y)p(t)^* - q(y)q(t)^*}\right)^* \\ = \; & b_t(x)b_t(y)^* \end{aligned}$$

where

$$b_t(x) = \frac{p(x)q(t) - q(x)p(t)}{p(x)p(t)^* - q(x)q(t)^*}.$$

If either of x or y is not in the same block as t then $L[t](x,y) = 1$, so if we extend b_t from E to X by setting $b_t(x) = 1$ outside E then $K^{(t)} = (b_t \otimes b_t^*)K$ on all of X. Therefore K is a Blaschke kernel, as claimed. ■

Note that the generalised Blaschke factor b_t is not fully determined by K. On the block containing the point t it is determined up to a unitary scalar factor. Outside that block it may take any unitary values, though it is natural to choose the value 1, as we did above.

Finally, let us return to the uniqueness question for Pick's theorem. We can now extend our uniqueness result to the completely non-zero Blaschke kernels.

LEMMA 3.2.4 *Let $K : X \times X$ be a completely non-zero, positive definite, Blaschke kernel, i.e. a positive definite kernel of the form $K = 1/(p \otimes p^* - q \otimes q^*)$, where p and q are any complex-valued functions on X. Then Pick's theorem is true for $H(K)$ (i.e. K is NP) and when the data set E is finite the minimal norm interpolating multiplier is unique.*

Proof: K is NP, by theorem 1.5.4, since $p \otimes p^* - 1/K = q \otimes q^* \geq 0$. Therefore K satisfies all the conditions that we assumed in our argument in proving lemma 3.1.2, i.e. it is positive-definite, NP, completely non-zero and satisfies $K^{(t)} = (b_t \otimes b_t^*)K$ for some functions b_t. Uniqueness therefore holds when E is finite. ∎

Note that the limitation that K be completely non-zero is certainly necessary. For otherwise K is block diagonal with more than one block and is effectively the direct sum of two or more sub-kernels that are completely independent of each other. Correspondingly, any interpolation problem for K is then equivalent to two or more simultaneous sub-problems that are independent except that the overall multiplier norm is the maximum of the norms arising in the sub-problems. Clearly, in this situation, the overall minimal norm extension will only be uniquely determined on the critical blocks, i.e. those on which the multiplier achieves its overall norm.

The simplest example of this is the identity kernel on a finite set X, i.e. $K(x, y) = 1$ if x=y, $K(x, y) = 0$ otherwise. The multiplier norm is then simply the supremum norm and we can extend a function f isometrically to a new point t by choosing any new value $\leq \|f\|_{M(K)}$.

Chapter 4

The Operator-Valued Case

It is well known [You86, theorem 1] [BGR90] that Pick's original theorem generalises to operator-valued H^∞ spaces. The result can be expressed in the following form:

> Let H be any Hilbert space, and $H^\infty(\mathcal{L}(H))$ denote the Banach space of $\mathcal{L}(H)$-valued analytic functions on $\mathbb{D}$ with the supremum norm
>
> $$\|\Phi\|_\infty = \sup_{x \in \mathbb{D}} \|\Phi(x)\|_{\mathcal{L}(H)}.$$
>
> Then there exists a function $\Phi \in H^\infty(\mathcal{L}(H))$, with $\|\Phi\|_\infty \leq 1$, that takes the n given operator values $Z_i \in \mathcal{L}(H)$ at the n given data points $x_i \in \mathbb{D}$ if and only if the operator matrix
>
> $$\left(\frac{1 - Z_i Z_j^*}{1 - x_i x_j^*} \right)_{i,j=1,\ldots,n}$$
>
> is positive.

As in the scalar-valued case covered so far, $H^\infty(\mathcal{L}(H))$ is the space of multipliers of $H^2(H)$, the space of square-integrable analytic H-valued functions on $\mathbb{D}$, and the above operator-valued Pick theorem again has a natural possible generalisation to reproducing kernel Hilbert spaces, this time to vector-valued spaces. In this chapter I examine how far the method used in chapter 1 generalises to vector-valued reproducing kernel Hilbert spaces.

Sections 1.1 to 1.4 generalise fairly directly, so to ease comparison the corresponding sections here use the same numbering of lemmas, definitions etc. as in chapter 1. To avoid repetition I do not go through the arguments step by step, but instead describe how the results generalise, only going into detail where significant differences arise.

Section 1.5 does not generalise to the non-commutative world of the operator-valued case; section 4.5 shows that its results can only really be applied to special cases where commutativity holds.

4.1 Vector-Valued Reproducing Kernel Spaces

In the generalisation of reproducing kernel Hilbert spaces the kernel becomes operator-valued and the space becomes one of vector-valued functions. This section outlines the theory, generally without proof. For further details see [BM84].

Before proceeding, we must define positivity for operator-valued kernels. For my preferred definition, developed in the beautiful generalisation work of Laurent Schwartz [Sch64], see chapter 6. However an equivalent, more elementary definition is as follows. A $\mathcal{L}(H)$-valued kernel K on a finite set X is positive (positive definite) if and only if the operator on $\sum_{x \in X} \oplus H$ with operator matrix K is a positive (positive definite) operator. This is then extended to infinite sets X by defining a kernel K to be positive (positive-definite) if and only if all finite restrictions of K are positive (positive-definite).

Now let X be any set, H be a Hilbert space and K be a positive $\mathcal{L}(H)$-valued kernel on $X \times X$. Then for any $y \in X$ and any vector $h \in H$, the mapping

$$K(\cdot, y)h : X \to H \qquad x \mapsto K(x, y)h$$

defines an H-valued function on X, and as h varies over H this gives a space of functions from X to H

$$H_y = \{x \mapsto K(x, y)h : h \in H\}$$

which is associated with the y-column of K. The reproducing kernel Hilbert space generated by K, $H(K)$, is the space of H-valued functions generated by these 'seed' spaces H_y. That is

$$H(K) = \overline{\text{span}}\{K(\cdot, y)h : y \in X, h \in H\}$$

with inner product defined to be the conjugate-linear extension of

$$\langle K(\cdot, y)h_y, K(\cdot, x)h_x\rangle_{H(K)} = \langle K(x, y)h_y, h_x\rangle_H.$$

The kernel K is called the 'reproducing' kernel for $H(K)$ since it has the following properties with respect to $H(K)$:

- for each $y \in X$ and each $h \in H$ the function $K(\cdot, y)h$, obtained by applying the y-column of K to h, is a member of $H(K)$
- the functions $K(\cdot, y)h$ $(y \in X, h \in H)$ 'reproduce' all functions $f \in H(K)$ in the sense that

  $$\langle f(y), h\rangle_H = \langle f, K(\cdot, y)h\rangle_{H(K)}$$

  which can be read as

  $$h \text{ component of } f(y) \text{ in } H = K(\cdot, y)h \text{ component of } f \text{ in } H(K).$$

Furthermore, $H(K)$ is the only Hilbert space of H-valued functions on X for which K has these reproducing properties. Conversely, given any Hilbert space of H-valued functions on a set X, for which all the 'point-evaluation' linear operators from $H(K)$ to H are bounded, there exists a unique positive kernel K on $X \times X$ with these reproducing properties.

To be the kernel of a reproducing kernel Hilbert space the kernel K need only be positive, not necessarily positive-definite. However, to avoid the unhelpful complication of degeneracy we will assume from now on that the kernel K is positive-definite.

Since members of $H(K)$ are H-valued functions, a multiplier of $H(K)$ must be an $\mathcal{L}(H)$-valued function on X. The multiplier space $M(K)$ is defined to be the space of $\mathcal{L}(H)$-valued functions on X for which multiplication of $H(K)$ is a bounded operator. As in the scalar case, both membership and the norms in $H(K)$ and $M(K)$ can be characterised in terms of positivity, this time positivity of *operator-valued* kernels:

- $\|f\|_{H(K)} = \inf_{r\geq 0}\{r : r^2K - f \otimes f^* \geq 0\}$
  and $f \in H(K)$ if and only if this infimum is finite. Here $f \otimes f^*$ is the kernel whose values are the rank-1 operators given by
  $$(f \otimes f^*)(x, y) = f(x)f(y)^* = h \mapsto \langle h, f(y)\rangle_H f(x).$$
- $\|\Phi\|_{M(K)} = \inf_{r\geq 0}\{r : r^2K - \Phi \otimes K \otimes \Phi^* \geq 0\}$
  and $\Phi \in M(K)$ if and only if this infimum is finite. Here $\Phi \otimes K \otimes \Phi^*$ is the kernel given by $(\Phi \otimes K \otimes \Phi^*)(x, y) = \Phi(x)K(x, y)\Phi(y)^*$.

Since restrictions of positive operator-valued kernels are positive, it follows from the above positivity characterisation of $\|\cdot\|_{M(K)}$ that the restriction of a multiplier in $M(K)$ to a subset E of X is a multiplier in $M(K_E)$ with equal or smaller norm.

In the scalar-valued case the columns of K played the role of the 'reproducing' functions in $H(K)$—evaluation of a function in $H(K)$ corresponded to taking its inner product with the columns of K. They were also used to characterise the multiplication operator adjoints on $H(K)$ as those that have all these reproducing functions as eigenvectors. In the operator-valued case, these two roles can be played by the subspaces H_y, by identifying H_y with H via various mappings.

Firstly note that each subspace H_y is unitarily equivalent to H. For, given any $y \in X$, $K(y,y) > 0$ so the mapping $K(\cdot,y)h \mapsto K(y,y)^{1/2}h$ defines a linear transformation of H_y onto H. But

$$\langle K(\cdot,y)h, K(\cdot,y)k\rangle_{H(K)} = \langle K(y,y)h,k\rangle_H = \langle K(y,y)^{1/2}h, K(y,y)^{1/2}k\rangle_H$$

so this mapping is an isometry, and therefore defines a unitary equivalence of H_y onto H.

Now let E_y denote the operator from H_y to H of evaluation at y. Since

$$\begin{aligned}\langle K(\cdot,y)h, E_y^*h'\rangle_{H(K)} &= \langle E_yK(\cdot,y)h,h'\rangle_H \\ &= \langle K(y,y)h,h'\rangle_H \\ &= \langle K(\cdot,y)h, K(\cdot,y)h'\rangle_{H(K)}\end{aligned}$$

then E_y^* is simply multiplication by $K(\cdot,y)$, i.e. $E_y^* = M_{K_y} = h \mapsto K(\cdot,y)h$. Letting P_y denote the orthogonal projection of $H(K)$ onto H_y, we have

$$\begin{aligned}\langle f(y),h\rangle_H &= \langle f, K(\cdot,y)h\rangle_{H(K)} \\ &= \langle P_yf, K(\cdot,y)h\rangle_{H(K)} \\ &= \langle P_yf, E_y^*h\rangle_{H(K)} \\ &= \langle E_yP_yf,h\rangle_H\end{aligned}$$

so $f(y) = E_yP_yf$. In other words, under the (generally non-unitary) identifications $E_y : H_y \to H$ of the subspaces H_y with H, the values of a function in $H(K)$ are its projections onto the subspaces H_y. This roughly justifies saying that the subspaces H_y play the role of 'reproducing' subspaces of $H(K)$—projection onto them evaluates functions in $H(K)$ (under the identifications E_y).

4.2 Generalising the Operator Pick Theorem

We now have enough facts and notation to restate a Pick theorem for the operator-valued case in terms of vector-valued reproducing kernel Hilbert spaces. Let

$$
\begin{aligned}
X &= \mathbb{D} \\
K &= \text{the operator-valued Szegö kernel } K(x,y) = I_H/(1-xy^*) \\
E &= \text{the set of data points } \{x_1, \ldots, x_n\} \\
K_E &= \text{the restriction of } K \text{ to } E \\
\text{and } F &= \text{the function on } E \text{ defined by } F(x_i) = Z_i.
\end{aligned}
$$

Then $H(K) = H^2(H)$ and $M(K) = H^\infty(\mathcal{L}(H))$ and Pick's theorem for $H^\infty(\mathcal{L}(H))$ can be restated as:

(1) There exists a multiplier $\Phi \in M(K)$ such that $\Phi|E = F$ and $\|\Phi\|_{M(K)} \leq 1$

if and only if

(2) the $\mathcal{L}(H)$-valued kernel $K_E - F \otimes K_E \otimes F^*$ is positive.

For general sets X, $E \subseteq X$, Hilbert space H and positive kernel K on X, this theorem will be called the generalised operator-valued Pick theorem for $H(K)$ and we will now investigate for which kernels K it is true.

As in the scalar-valued case, (2) is equivalent to $\|\Phi|E\|_{M(K_E)} \leq 1$, so (1) $\Rightarrow$ (2) is always true, because restricting Φ to E cannot increase its multiplier norm. The truth of the generalised operator-valued Pick's theorem therefore depends on whether (2) $\Rightarrow$ (1), which is equivalent to always being able to extend multipliers from subsets of X to the whole of X without increasing multiplier norm.

4.3 One-Point Extensions Are Sufficient

The results of section 1.3 generalise fairly directly. Lemma 1.3.1 becomes:

LEMMA 4.3.1 *Let H be a Hilbert space and K be a positive $\mathcal{L}(H)$-valued kernel on a set X. Then an operator M on $H(K)$ is a multiplication operator if and only if the reproducing subspaces H_y are eigenspaces of M^* for all $y \in X$. Furthermore, if M is a multiplication operator then the corresponding multiplier Φ is given by $M^*|H_y = M_{K_y}\Phi(y)^* M_{K_y}^{-1} = K(\cdot, y)h \mapsto K(\cdot, y)\Phi(y)^*h$.*

Proof: Given a multiplier $\Phi \in M(K)$

$$\begin{aligned}
\langle M_\Phi^*(K(\cdot,y)h_y), K(\cdot,x)h_x\rangle_{H(K)} &= \langle K(\cdot,y)h_y, \Phi(\cdot)K(\cdot,x)h_x\rangle_{H(K)} \\
&= \langle h_y, \Phi(y)K(y,x)h_x\rangle_H \\
&= \langle K(x,y)\Phi(y)^*h_y, h_x\rangle_H \\
&= \langle K(\cdot,y)\Phi(y)^*h_y, K(\cdot,x)h_x\rangle_{H(K)}
\end{aligned}$$

and since this is for all $x, y \in X$ and all $h_x, h_y \in H$, we have

$$M_\Phi^*(K(\cdot,y)h_y) = K(\cdot,y)\Phi(y)^*h_y.$$

Therefore M_Φ^* leaves H_y invariant and $M_\Phi^*|H_y = M_{K_y}\Phi(y)^*M_{K_y}^{-1}$, as claimed. Conversely, if M^* is a bounded operator on $H(K)$ that leaves H_y invariant for each $y \in X$, then M^* and the adjoint of multiplication by the $\mathcal{L}(H)$-valued function $\Phi(y)^* = M_{K_y}^{-1}(M^*|H_y)M_{K_y}$ clearly agree on the reproducing subspaces H_y, so by linearity and continuity they must also agree on their closed span, i.e. $H(K)$. ■

The definition of NP kernels needs no change from that given in chapter 1:

DEFINITION 4.3.2 *A positive, $\mathcal{L}(H)$-valued kernel K on a set X will be called an NP kernel provided for every subset E of X and every $t \in X \setminus E$, all multipliers in $M(K_E)$ can be isometrically extended to multipliers in $M(K_{E \cup \{t\}})$.*

LEMMA 4.3.3 *Given a Hilbert space H and a positive-definite $\mathcal{L}(H)$-valued kernel K on a set X, then the full Pick theorem is true for $H(K)$ if and only if K is an NP kernel.*

Proof: The proof given in chapter 1 applies equally well in the operator-valued case, the only places where generalisation are needed being:

- references to columns of a kernel need to be changed throughout to refer instead to the reproducing subspaces associated with the columns. For example H_g's definition becomes 'the closed subspace of $H(K_F)$ spanned by the reproducing subspaces of $H(K_F)$ corresponding to F_g'.
- references to eigenvectors need to be changed to eigenspaces
- the above lemma 4.3.1 replaces the use of lemma 1.3.1. ■

4.4 Minimal Norm Extension

The non-orthogonal generalisation of Parrott's lemma (1.4.1) developed in chapter 1 is already general enough for the operator-valued case and so stands without change. However the other results of section 1.4 do need some generalisation.

LEMMA 4.4.2 *Let K be a positive-definite, $\mathcal{L}(H)$-valued kernel on a set X, $t \in X$, $E = X \setminus \{t\}$ and F be a multiplier of $H(K_E)$. Further, let $K^{(t)}$ denote the Schur complement of K(t,t) in K, given by*

$$K^{(t)}(x, y) = K(x, y) - K(x, t)K(t, t)^{-1}K(t, y).$$

Then $\|F\|_{M(K_E)}$ *and* $\|F\|_{M(K_E^{(t)})}$ *are both lower bounds for the norm of any one-point extension of* F *to* X*, and the larger of the two, if finite, is achievable, i.e. there exists a one-point extension of* F *with that norm.*

Proof: The proof given for the corresponding lemma (1.4.2) in chapter 1 generalises fairly directly. To cover this operator-valued case we need to show that:

- $K^{(t)}$ is the reproducing kernel of the subspace $H(K,\{t\})^{\perp}$. To verify this, note that

$$\begin{aligned}
& \langle K^{(t)}(\cdot,y)h_y, K(\cdot,t)h_t\rangle_{H(K)} \\
= \; & \langle K(\cdot,y)h_y - K(\cdot,t)K(t,t)^{-1}K(t,y)h_y, K(\cdot,t)h_t\rangle_{H(K)} \\
= \; & \langle K(t,y)h_y - K(t,t)K(t,t)^{-1}K(t,y)h_y, h_t\rangle_{H} \\
& \text{(using } K\text{'s reproducing property)} \\
= \; & 0.
\end{aligned}$$

  Therefore $K^{(t)}(\cdot,y)h_y \in H(K,\{t\})^{\perp}$ for all $y \in X$. Also

$$\begin{aligned}
\langle f, K^{(t)}(\cdot,y)h_y\rangle &= \langle f, K(\cdot,y)h_y - K(\cdot,t)K(t,t)^{-1}K(t,y)h_y\rangle_{H(K)} \\
&= \langle f(y), h_y\rangle_H
\end{aligned}$$

  for all $f \in H(K,\{t\})^{\perp}$ (using K's reproducing property and noting that $\langle f, K(\cdot,t)K(t,t)^{-1}K(t,y)h_y\rangle_{H(K)} = 0$) so $K^{(t)}$ also has the reproducing property for $H(K,\{t\})^{\perp}$.

- the compression W, of T_D onto $H(K,\{t\})^{\perp}$, leaves the subspaces

$$\{K^{(t)}(\cdot,y)h : h \in H\} \qquad (y \in X)$$

  invariant and satisfies

$$W|\{K^{(t)}(\cdot,y)h : h \in H\} = M_{K_y^{(t)}}F(y)^* M_{K_y^{(t)}}^{-1}.$$

This is verified by the following calculation.

$$
\begin{aligned}
W(K^{(t)}(\cdot,y)h) &= P_{H_2^\perp}T_D\left(K(\cdot,y)h - K(\cdot,t)K(t,t)^{-1}K(t,y)h\right) \\
&= P_{H_2^\perp}(A(K(\cdot,y)h) - DK(\cdot,t)K(t,t)^{-1}K(t,y)h) \\
&= P_{H_2^\perp}(K(\cdot,y)F(y)^*h) \qquad \text{since } P_{H_2}^\perp D = 0 \\
&= K^{(t)}(\cdot,y)F(y)^*h.
\end{aligned}
$$

■

LEMMA 4.4.3 *Let K be a positive-definite, $\mathcal{L}(H)$-valued kernel on X, t be any point of X, $E = X \setminus \{t\}$, and F be a multiplier in $M(K_E)$. Then if N is a positive real such that all restrictions of F to finite subsets G have one-point extensions to $G \cup \{t\}$ with norms $\leq N$, there exists an extension of F to X with norm $\leq N$.*

Proof: The induction proof of the corresponding lemma (1.4.3) in chapter 1 still works in this operator-valued case, provided we again generalise eigenvectors to invariant subspaces and eigenvalues to the restrictions of the operator to those invariant subspaces.

In a little more detail, the operators M_α^* are defined as before, again they form a net in the ball B of radius N in $\mathcal{L}(H(K))$ and since B is compact there is a cluster point M^*. The new, slightly stronger, fact we now need is that given a net of operators all of which leave a given subspace invariant and which all agree on that subspace, then any cluster point of the net must also share these properties. Using this we can conclude that M^* must leave all the reproducing subspaces of $H(K)$ invariant and so, by lemma 4.3.1, be the adjoint of multiplication by some multiplier Φ. Further, for $y \in E$, $M^*|H_y$ must equal $M_{K_y}F(y)^*M_{K_y}^{-1}$ so $\Phi|E = F$ and therefore Φ is an extension of F. That Φ has multiplier norm $\leq N$ follows as before. ■

As in chapter 1, it follows from 4.4.3 that for operator-valued kernels we have 'finitely-NP $\equiv$ NP' and 'finite Pick theorem $\equiv$ full Pick theorem'. More precisely:

Corollary 4.4.4 *A positive-definite, $\mathcal{L}(H)$-valued kernel K is an NP kernel if and only if all finite restrictions of K are NP kernels.*

Corollary 4.4.5 *If K is a positive-definite, $\mathcal{L}(H)$-valued kernel on a set X, then for $H(K)$ the finite Pick theorem and the full Pick theorem are equivalent.*

4.5 Sufficient Conditions for Pick's Theorem

We have now reached the point where the arguments we used in chapter 1 no longer generalise fully to the operator-valued case, the barrier basically being non-commutativity. For example in section 1.5 a key tool used is the Schur product theorem, i.e. that the pointwise product of two scalar-valued positive kernels is positive. However this does not hold for two operator-valued positive kernels; for instance a positive kernel must have positive diagonal entries but the operator product of two positive operators is not even necessarily Hermitian.

It turns out that the methods from chapter 1 can only be applied if fairly strong conditions are placed on the operator-valued kernel K, so that commutativity is restored. Before showing this it will be useful to develop a new view of the one-point extension method, using the terminology of linear maps on C^*-algebras; this view will also be needed in the next chapter on completely NP kernels.

For the operator-valued case, consider the problem of isometrically extending a multiplier $F \in M(K_E)$ from $E = X \setminus \{t\}$ to X. By corollary 4.4.4 we can restrict attention to the case of X being finite. By lemma 4.4.2 isometric extension will be possible if and only if

$$\|\text{multiplication of } H(K_E) \text{ by } F\| \leq \|\text{multiplication of } H(K_E^{(t)}) \text{ by } F\|$$

so isometric extension will be possible for all multipliers $F \in M(K_E)$ if and only if the map

$$\beta : M^*_{F,K_E} \to M^*_{F,K_E^{(t)}}$$

is contractive. Here we have extended the notation M_F to indicate the space being multiplied.

Let $\mathcal{M}(K_E)$ and $\mathcal{M}(K_E^{(t)})$ denote the Banach spaces of multiplication operators on $H(K_E)$ and $H(K_E^{(t)})$ respectively—$\mathcal{M}(K_E)$ and $\mathcal{M}(K_E^{(t)})$ are effectively $M(K_E)$ and $M(K_E^{(t)})$ except that their members are the multiplication operators rather than the multipliers themselves. These spaces, and the corresponding spaces of adjoints $\mathcal{M}(K_E)^*$ and $\mathcal{M}(K_E^{(t)})^*$, are subspaces of the C^*-algebras $\mathcal{L}(H(K_E))$, $\mathcal{L}(H(K_E^{(t)}))$ and as such are known as *operator spaces.* The property that interests us, i.e. the ability to isometrically extend any multiplier, has therefore been characterised in terms of the contractivity of a linear map β from one operator space onto another. A lot is known about such maps [Pau86] and in the next chapter this C^*-algebra way of viewing our problem will prove useful. However, for the moment it is simply a useful shorthand.

Now consider another way of characterising our property, this time in terms of positivity of a map operating on kernels. Let γ denote the map

$$K_E - F \otimes K_E \otimes F^* \mapsto K_E^{(t)} - F \otimes K_E^{(t)} \otimes F^*$$

defined on the kernels of the form $K_E - F \otimes K_E \otimes F^*$, where F ranges over all multipliers $F \in M(K_E)$. We know that positivity of the kernels $K_E - F \otimes K_E \otimes F^*$ and $K_E^{(t)} - F \otimes K_E^{(t)} \otimes F^*$ corresponds exactly to contractivity of F as a multiplier of $H(K_E)$ and $H(K_E^{(t)})$, respectively, so β is contractive if and only if γ is positive, i.e. maps positive kernels to positive kernels. In the scalar-valued case (where everything commutes), and if K is completely non-zero, then γ is simply Schur multiplication by $K_E^{(t)}/K_E$ and this allows us to use the Schur product theorem.

With this new perspective, we can now summarise the proof of Pick's theorem given in chapter 1 as having the following steps.

1. Pick's theorem is equivalent to always being able to isometrically extend any multiplier from any given subset E of X to any given extra point $t \in X \setminus E$. (Shown in sections 1.2 and 1.3.)

2. It is sufficient that isometric extension is always possible when E is finite. (Shown in section 1.4.)

3. Isometric extension is always possible from E to $E \cup \{t\}$ if and only if the associated map β is contractive. (This was shown, though not in this terminology, in section 1.4.)

4. β is contractive if and only if

$$\gamma : K_E - F \otimes K_E \otimes F^* \to K_E^{(t)} - F \otimes K_E^{(t)} \otimes F^*$$

is positive. (This is essentially the fact that $\|A\| \leq 1 \Leftrightarrow I - A^*A \geq 0$.)

5. When K is completely non-zero, γ is Schur multiplication by $K_E^{(t)}/K_E$ and so γ is a positive map if this kernel is positive. (By the Schur product theorem.)

6. $K_E^{(t)}/K_E$ is positive for each $E \subset X$ and $t \in X \setminus E$ if and only if $\kappa^+(1/K) = 1$ (shown in section 1.5). Here

$$\kappa^+(1/K) = \sup_{G \subseteq X, G \text{ finite}} (\text{number of positive eigenvalues of } (1/K)|G)$$

is the 'positivity' of the kernel $1/K$.

From this bird's eye view of chapter 1's method we can see how far the method extends to the operator-valued case now being considered. Steps 1 to 4 generalise fairly directly, with no special constraints on the kernel K, as has been shown in sections 4.2 to 4.4. However in step 5 the map

$$\gamma : K_E - F \otimes K_E \otimes F^* \to K_E^{(t)} - F \otimes K_E^{(t)} \otimes F^*$$

will not be a Schur product map unless we can commute the factors K_E and $K_E^{(t)}$ out from the centres of the second terms on each side. This will only be possible for all multipliers F if K is of the form kI_H for some scalar-valued kernel k, and to guarantee γ is a Schur product map k will need to be completely non-zero. We also need K to be of this form to use the Schur product theorem in step 5, since this theorem holds for the product of a positive operator-valued kernel by a positive scalar-valued kernel, but not in general (as noted above) when both kernels are operator-valued. If K *is* of the form kI_H, with k a completely non-zero scalar-valued kernel, then γ is Schur multiplication by the scalar-valued kernel k and step 6 for the operator-valued case is exactly as it was for the scalar-valued case.

From this review of chapter 1's method, and its generalisation to the operator-valued case, we can see that the method is limited to cases where the operator-valued kernel K is a scalar multiple of the identity. However, for this case we obtain the following sufficient conditions for the full Pick theorem to be true for $H(K)$.

THEOREM 4.5.1 *Let H be a Hilbert space, X be any set and k be a completely non-zero scalar-valued kernel on $X \times X$ such that $\kappa^+(1/k) = 1$. Then the full Pick theorem is true for $H(kI_H)$.*

Although this result is somewhat limited, it does give as corollaries:

- the original case of Pick's theorem for $H^2(H)$
- Pick's theorem for vector-valued Dirichlet spaces.

Chapter 5

Completely NP Kernels

Chapter 1 showed that for a completely non-zero, scalar-valued kernel K to be NP it is sufficient that $\kappa^+(1/K) = 1$. It is tempting to hope that this condition might also be necessary and I expended substantial time and effort hoping to prove this, but only succeeded in proving its necessity for completely non-zero kernels of size up to 3×3.

Work by Scott McCullough [McC] then showed that it is better to study a stronger property than NP-ness, called complete NP-ness. McCullough shows that for completely non-zero finite kernels the property $\kappa^+(1/K) = 1$ in fact characterises complete NP-ness rather than the weaker property of NP-ness. In this chapter we

- analyse exactly which are the 3×3 NP kernels; they are all completely NP

- extend McCullough's work to a complete characterisation of the completely NP kernels on any set X

- construct an explicit example of a 4×4 kernel that is NP but is not completely NP.

5.1 Characterisation of 3 by 3 NP Kernels

When working with a finite, scalar-valued, positive-definite kernel K it is convenient to normalise it by Schur multiplying by $a \otimes a^*$ where

$$a(x) = 1/\sqrt{K(x,x)}.$$

This renders all the diagonal terms of K equal to 1 and is equivalent to altering the norm in $H(K)$ to rescale the reproducing functions to be unit vectors. Because a multiplication operator adjoint M_f^* is the operator with these vectors as eigenvectors and $f(x)^*$ as eigenvalues, rescaling the vectors has no effect on the multiplier norm and so no effect on interpolation questions. We will therefore work with normalised kernels throughout this section.

When the domain set X is small enough NP-ness of a kernel K can be analysed by direct hand calculation. The property is meaningless if X contains only one point since at least two points are needed in order to do a one-point extension. For the next size up, i.e. $\operatorname{card}(X) = 2$, isometric one point extension of a multiplier f from $\{x_1\}$ to $\{x_1, x_2\}$ is always possible, since we can simply set $f(x_2) = f(x_1)$. Therefore all 2×2 kernels are NP and it is not until $\operatorname{card}(X) = 3$ that interesting things start to happen.

Take a 3×3 normalised positive-definite kernel

$$K = \begin{pmatrix} 1 & a & b \\ a^* & 1 & c \\ b^* & c^* & 1 \end{pmatrix}$$

and consider isometrically extending a given multiplier $f : \{x_1, x_2\} \to \{f_1, f_2\}$ by choosing a value for f_3. Let $E = \{x_1, x_2\}$ and $t = x_3$. By lemma 1.4.2 this can be done if and only if

$$\|f\|_{M(K_E^{(t)})} \leq \|f\|_{M(K_E)} \qquad (\star)$$

and in this case these are multiplier norms associated with 2×2 kernels so we can calculate them directly.

Let $L = \begin{pmatrix} 1 & a \\ a^* & 1 \end{pmatrix}$ be any positive-definite normalised 2×2 kernel on $\{x_1, x_2\}$. Note that L is positive definite $\Leftrightarrow \det(L) > 0 \Leftrightarrow |a| < 1$. Then

$$\begin{aligned}
\|f\|^2_{M(L)} &= \inf(r \geq 0 : (r - f \otimes f^*)L \geq 0) \\
&= \text{largest root of } \det \begin{pmatrix} r - f_1 f_1^* & a(r - f_1 f_2^*) \\ a^*(r - f_2 f_1^*) & r - f_2 f_2^* \end{pmatrix} = 0 \\
&= \text{largest root of } (r - f_1 f_1^*)(r - f_2 f_2^*) - aa^*(r - f_2 f_1^*)(r - f_1 f_2^*) = 0 \\
&= \text{largest root of } r^2(1 - aa^*) \\
&\quad -r((f_1 - f_2)(f_1^* - f_2^*) + (1 - aa^*)(f_1 f_2^* + f_2 f_1^*)) \\
&\quad +(1 - aa^*)|f_1|^2|f_2|^2 = 0 \\
&= \text{largest root of } r^2 - Br + C = 0 \\
&= \frac{B + \sqrt{B^2 - 4C}}{2C} \\
&\quad \text{where } B = |f_1 - f_2|^2/(1 - |a|^2) + 2\Re(f_1 f_2^*) \text{ and } C = |f_1|^2|f_2|^2.
\end{aligned}$$

Although this formula is fairly complicated we can derive from it some useful information about how $\|f\|_{M(L)}$ varies with a. Since C is independent of a the variation only occurs through B. But $\|f\|_{M(L)}$ is an increasing function of B since

$$\frac{\partial(2C\|f\|^2_{M(L)})}{\partial B} = 1 + \frac{B}{\sqrt{B^2 - 4C}} = \frac{2C\|f\|^2_{M(L)}}{\sqrt{B^2 - 4C}} \geq 0$$

and B is an increasing function of $|a|$, so $\|f\|_{M(L)}$ is an increasing function of $|a|$. This tells us how to compare multiplier norms—the larger the off-diagonal term in the normalised kernel then the larger the multiplier norm.

Applying this to the inequality $(\star)$ we see that isometric extension of f will be possible

$$\begin{aligned}
&\Leftrightarrow \quad |\text{off-diagonal term in normalised } K_E^{(t)}|^2 \leq |a|^2 \\
&\Leftrightarrow \quad \frac{(a - bc^*)(a^* - b^*c)}{(1 - bb^*)(1 - cc^*)} \leq aa^* \\
&\qquad \text{since } K_E^{(t)} = \begin{pmatrix} 1 - bb^* & a - bc^* \\ a^* - b^*c & 1 - cc^* \end{pmatrix} \\
&\Leftrightarrow \quad (a - bc^*)(a^* - b^*c) \leq aa^*(1 - bb^*)(1 - cc^*) \\
&\Leftrightarrow \quad aa^*bb^*cc^* - aa^*bb^* - aa^*cc^* - bb^*cc^* + ab^*c + a^*bc^* \geq 0.
\end{aligned}$$

Because this condition is symmetrical in a, b and c the ability to isometrically extend all multipliers from 2 points to the third does not depend on which two points are chosen. Therefore these criteria actually characterise NP-ness of K. Also, we can see from the form of the inequalities that:

- if exactly one of the off-diagonal terms is zero then K is not NP. For example if $a = 0$ then

$$aa^*bb^*cc^* - aa^*bb^* - aa^*cc^* - bb^*cc^* + ab^*c + a^*bc^* = -bb^*cc^* \not\geq 0$$

- if more than one of the off-diagonal terms is zero then K is NP.

- if K is completely non-zero then it is NP if and only if $\kappa^+(1/K) = 1$. To see this, note that $1/K$ must have at least 1 negative and 1 positive eigenvalue since each of its principal 2 by 2 submatrices have negative determinants. Therefore the only possible values for $\kappa^+(1/K)$ are 1 and 2 and the latter value gives

$$0 > aa^*bb^*cc^* \det(1/K) = aa^*bb^*cc^* - aa^*bb^* - aa^*cc^* - bb^*cc^* + ab^*c + a^*bc^*.$$

We have therefore characterised the 3×3 NP kernels—they are the completely non-zero kernels for which $\kappa^+(1/K) = 1$ plus the kernels having zeros but which are block-diagonal. Later in this chapter it will be shown that these are exactly the completely NP kernels, so for 3×3 kernels the adverb 'completely' adds nothing new.

Although the above approach can in principle also be applied to 4×4 kernels, in practice the algebra becomes totally unmanageable—the kernel then has 6 complex degrees of freedom and a cubic equation must be solved. This is a pity since, as will be seen in section 5.3, it turns out that in the 4×4 case the adverb 'completely' does indeed make a difference.

5.2 Characterisation of Completely NP kernels

In section 4.5 we introduced a way of viewing NP-ness of a finite kernel K in terms of a family of maps β defined on operator spaces. Let $\beta_{K,E,t}$ denote the map associated with extension from E to t with kernel K, i.e.

$$\beta_{K,E,t} : \mathcal{M}(K_E)^* \to \mathcal{M}(K_E^{(t)})^* \qquad M^*_{f,K_E} \mapsto M^*_{f,K_E^{(t)}}.$$

Then we know that

$$K \text{ is NP} \Leftrightarrow \beta_{K,E,t} \text{ is contractive for each } \emptyset \subset E \subset X \text{ and } t \in X \setminus E.$$

From the C^*-algebra viewpoint, the map $\beta_{K,E,t}$ has some nice properties:

- its domain and range operator spaces $\mathcal{M}(K_E)^*$ and $\mathcal{M}(K_E^{(t)})^*$ are unital (they contain the identity, the adjoint of multiplication by the constant function 1) and are in fact closed non-self-adjoint subalgebras of their respective containing C^*-algebras.
- $\beta_{K,E,t}$ is unital, i.e. it maps the unit in $\mathcal{M}(K_E)^*$ to the unit in $\mathcal{M}(K_E^{(t)})^*$
- $\beta_{K,E,t}$ is spectrum-preserving and so is a positive map (i.e. maps positive elements to positive elements). To see this, note that for *any* positive-definite kernel L on E, the spectrum of M_f^* on $H(L)$ is simply $f(E)^*$.

This enables us to use the theory of contractivity and positivity of linear maps between operator spaces in C^*-algebra, which has been made easily accessible by Vern Paulsen in his very useful book 'Completely Bounded Maps and Dilations' [Pau86]. An important role is played in this theory by rather stronger, and better behaved, properties of C^*-algebra maps called complete contractivity and complete positivity. Given two C^*-algebras A and B, let $M_n(A)$ and $M_n(B)$ ($n \in \mathbb{N}$) denote the induced C^*-algebras of $n \times n$ matrices over A and B respectively. Then a linear map $\phi : A \to B$ induces maps $\phi_n : M_n(A) \to M_n(B)$ defined by applying ϕ elementwise:

$$\phi_n((a_{ij})_1^n) = (\phi(a_{ij}))_1^n$$

and ϕ is called *completely* contractive if all the induced maps ϕ_n, $n \in \mathbb{N}$, are contractive. In general, appending the adverb 'completely' to a given property of ϕ means that ϕ_n has that property for each $n \in \mathbb{N}$.

This process can equally well be viewed as one of tensor product by the identity of $M_n(\mathbb{C})$, by making the following identifications:

$$\begin{aligned} M_n(A) &\simeq M_n(\mathbb{C}) \otimes A \\ (a_{ij})_1^n &\simeq \sum_{i,j=1}^{n} E_{ij} \otimes a_{ij} \\ \phi_n &\simeq I_n \otimes \phi \end{aligned}$$

where $\{E_{ij} : i, j = 1, \ldots, n\}$ is the standard basis for $M_n(\mathbb{C})$ and I_n is the identity in $M_n(\mathbb{C})$.

In our case the C^*-algebra involved is $\mathcal{L}(H(K))$ and we have

$$M_n(\mathbb{C}) \otimes \mathcal{L}(H(K)) = \mathcal{L}(M_n(\mathbb{C}) \otimes H(K)) = \mathcal{L}(H(I_n \otimes K))$$

where $I_n \otimes K$ is the operator-valued kernel on $X \times X$ given by

$$(K \otimes I_n)(x, y) = K(x, y)I_n.$$

We are therefore led to consider the reproducing kernel Hilbert spaces with kernels of the form $I_n \otimes K$. Burbea and Masani [BM84] refer to these kernels as the *inflations* of K, terminology that I will extend to refer to this general process of generating larger from smaller, by referring to the spaces $M_n(\mathbb{C}) \otimes A$ and maps ϕ_n as the inflations of A and ϕ, respectively.

We can now define complete NP-ness: a kernel K is completely NP if and only if all of its inflations are NP. The following lemma shows that this is a reasonable use of the adverb.

LEMMA 5.2.1 *Let K be a positive-definite, scalar-valued kernel on a finite set X. Then K is completely NP if and only if the maps $\beta_{K,E,t}$ are completely contractive for each $\emptyset \subset E \subset X$ and $t \in X \setminus E$.*

Proof: Consider what happens when K is inflated to $I_n \otimes K$. The inflated Hilbert space $H(I_n \otimes K)$ is $\overline{\text{span}}\{K(\cdot, y)h : y \in X, h \in \mathbb{C}^n\}$ and can be identified as the direct sum of n copies of $H(K)$, via the identification

$$K(\cdot, y)e_i \simeq 0 \oplus \ldots \oplus 0 \oplus K(\cdot, y) \oplus 0 \oplus \ldots \oplus 0$$

where $\{e_i : i = 1, \ldots, n\}$ is the standard orthonormal basis of $\mathbb{C}^n$ and the term $K(\cdot, y)$ appears on the right in the i'th position.

The proof of our lemma is now essentially just the (easy, though complicated) job of following this identification through to the various objects involved in our study. Under this identification $H(I_n \otimes K) \simeq \sum_1^n \oplus H(K)$ we find that:

- the algebra of all bounded operators on $H(I_n \otimes K)$ is identified as the algebra of $n \times n$ matrices with entries taken from $\mathcal{L}(H(K))$. That is

$$\mathcal{L}(H(I_n \otimes K)) \simeq M_n \otimes \mathcal{L}(H(K) = \text{ the inflation of } \mathcal{L}(H(K)).$$

- the $y-$reproducing subspace of $H(I_n \otimes K)$ is identified as the subspace $K(\cdot, y)\mathbb{C}^n$ of $\sum_1^n \oplus H(K)$.
- the multiplication operator adjoints on $H(I_n \otimes K)$ are therefore identified as the operators on $\sum_1^n \oplus H(K)$ that leave each of the subspaces $K(\cdot, y)\mathbb{C}^n$ $(y \in X)$ invariant, so they are exactly the $n \times n$ matrices with entries that are operators on $H(K)$ that leave each $K(\cdot, y)$ invariant. That is

$$\mathcal{M}(I_n \otimes K)^* \simeq M_n(\mathbb{C}) \otimes \mathcal{M}(K)^* = \text{ the inflation of } \mathcal{M}(K)^*.$$

- the process of inflation commutes with restriction to a subset $E \subset X$. That is

$$H((I_n \otimes K)_E) = H(I_n \otimes (K_E)).$$

- inflation also commutes with Schur complementation. That is

$$H((I_n \otimes K)_E^{(t)}) = \sum_1^n \oplus H(K_E^{(t)}) = H(I_n \otimes K_E^{(t)}).$$

- $\beta_{I_n \otimes K,E,t}$ acts on an operator in $\mathcal{M}((I_n \otimes K)_E)^*$, which we have seen can be viewed as an $n \times n$ matrix of operators from $\mathcal{M}(K_E)^*$, by simply applying $\beta_{K,E,t}$ elementwise. That is

$$\beta_{I_n \otimes K,E,t} \simeq (\beta_{K,E,t})_n.$$

Having identified $\beta_{I_n \otimes K,E,t}$ with $(\beta_{K,E,t})_n$ our result now follows directly since K is completely NP

$$\begin{aligned}
&\Leftrightarrow \quad I_n \otimes K \text{ is NP for each } n \in \mathbb{N} \\
&\Leftrightarrow \quad \beta_{I_n \otimes K,E,t} \text{ is contractive for all } \emptyset \subset E \subset X, t \in X \setminus E, n \in \mathbb{N} \\
&\Leftrightarrow \quad (\beta_{K,E,t})_n \text{ is contractive for all } \emptyset \subset E \subset X, t \in X \setminus E, n \in \mathbb{N} \\
&\Leftrightarrow \quad \beta_{K,E,t} \text{ is completely contractive for all } \emptyset \subset E \subset X, t \in X \setminus E.
\end{aligned}$$

■

We can now start to characterise the completely NP kernels by proving, using our viewpoint and terminology, McCullough's result for completely non-zero, finite kernels. The key step in the next proof—the induction step—is derived from McCullough's work [McC].

Lemma 5.2.2 *(S. McCullough)* *Let K be a completely non-zero positive definite kernel on a finite set X. Then the following are equivalent:*

1. *K is completely NP.*
2. *The maps $\beta_{K,E,t} : \mathcal{M}(K_E)^* \to \mathcal{M}(K_E^{(t)})^*$ are completely contractive for each non-empty proper subset E of X and each $t \in X \setminus E$.*
3. *$K_E^{(t)}/K_E \geq 0$ for each non-empty proper subset E of X and each $t \in X \setminus E$.*
4. *$\kappa^+(1/K) = 1$.*

Proof: Lemma 5.2.1 shows the equivalence of (1) and (2), and in chapter 1 we showed (lemmas 1.5.1 and 1.5.2) that (3) and (4) are equivalent. It therefore remains to show that (2) $\Leftrightarrow$ (3).

To prove this we will use a matricial representation of the operators involved. Choose any $E \subset X$, of cardinality n say, and any $t \in X \setminus E$. We can consider K_E as a matrix over $\mathbb{C}^n$, in which case its positive square root $K_E^{1/2}$ is such that its columns are vectors in $\mathbb{C}^n$ whose Gram matrix is $(K_E^{1/2})^*(K_E^{1/2}) = K_E$, i.e. the same as the reproducing functions in $H(K_E)$. The identification

$$K(\cdot, y) \simeq y\text{-column of } K_E^{1/2}$$

is therefore a Hilbert space isomorphism of $H(K_E)$ onto $\mathbb{C}^n$ which allows us to represent $H(K_E)$ as $\mathbb{C}^n$.

Using this representation we can form the matrices of operators on $H(K_E)$ with respect to the standard orthonormal basis of $\mathbb{C}^n$. The operator of interest is the multiplication operator adjoint M_f^*, which under our representation is the operator with the columns of $K_E^{1/2}$ as eigenvectors and values $f(\cdot)^*$ as eigenvalues. With respect to the standard orthonormal basis of $\mathbb{C}^n$, the matrix of M_f^* is therefore

$$M_f^* \simeq K_E^{1/2} \mathrm{diag}(f^*) K_E^{-1/2}$$

since the matrix $K_E^{-1/2}$ maps the columns of $K_E^{1/2}$ to the standard basis vectors, $\mathrm{diag}(f^*)$ scales them correctly and then $K_E^{1/2}$ maps them back again.

We can represent $H(K_E^{(t)})$ and its operators as $\mathbb{C}^n$ and $M_n(\mathbb{C})$ in exactly the same way. The map $\beta_{K,E,t}$ is then represented by a linear map on a subspace of $M_n(\mathbb{C})$. We now have vectorial or matricial representations of all the elements of our problem:

$$\begin{aligned}
H(K_E) &\simeq \mathbb{C}^n \\
M^*_{f,K_E} &\simeq K_E^{1/2}\mathrm{diag}(f^*)K_E^{-1/2} \\
\mathcal{M}(K_E)^* &\simeq \text{matrices of the form } K_E^{1/2}\mathrm{diag}(f^*)K_E^{-1/2} \\
H(K_E^{(t)}) &\simeq \mathbb{C}^n \\
M^*_{f,K_E^{(t)}} &\simeq K_E^{(t)\,1/2}\mathrm{diag}(f^*)K_E^{(t)\,-1/2} \\
\mathcal{M}(K_E^{(t)})^* &\simeq \text{matrices of the form } K_E^{(t)\,1/2}\mathrm{diag}(f^*)K_E^{(t)\,-1/2} \\
\beta_{K,E,t} &\simeq K_E^{1/2}\mathrm{diag}(f^*)K_E^{-1/2} \mapsto K_E^{(t)\,1/2}\mathrm{diag}(f^*)K_E^{(t)\,-1/2}
\end{aligned}$$

For brevity, we will now consider the spaces and transformations of the left to equal those on the right, rather than simply being represented by them. This is valid since the representations involved preserve all properties of interest. Also, since K, E and t are now fixed we will now simplify the notation by omitting the subscripts and simply writing β.

Now consider the following diagram:

$$\begin{array}{ccc}
\mathcal{M}(K_E)^* \subset M_n(\mathbb{C}) & \xrightarrow{\ \mu\ } & M_n(\mathbb{C}) \\
\Big\downarrow \beta & & \Big\downarrow S \\
\mathcal{M}(K_E^{(t)})^* \subset M_n(\mathbb{C}) & \xrightarrow[\ \nu\ ]{} & M_n(\mathbb{C})
\end{array}$$

$$\text{where } \mu \;=\; \text{congruence by } K_E^{1/2} = A \mapsto K_E^{1/2} A K_E^{1/2}$$

$$\begin{aligned} \nu &= \text{congruence by } {K_E^{(t)}}^{1/2} = A \mapsto {K_E^{(t)}}^{1/2} A {K_E^{(t)}}^{1/2} \\ \text{and } S &= \text{Schur multiplication by } K_E^{(t)}/K_E. \end{aligned}$$

For any function $f : E \to \mathbb{C}$, M^*_{f,K_E} is represented by $K_E^{1/2}\mathrm{diag}(f^*)K_E^{-1/2}$ so

$$\begin{aligned} (\nu^{-1}S\mu)(M^*_{f,K_E}) &= (\nu^{-1}S)(K_E^{1/2}K_E^{1/2}\mathrm{diag}(f^*)K_E^{-1/2}K_E^{1/2}) \\ &= (\nu^{-1}S)(K_E\mathrm{diag}(f^*)) \\ &= \nu^{-1}(K_E^{(t)}\mathrm{diag}(f^*)) \\ &= {K_E^{(t)}}^{1/2}\mathrm{diag}(f^*){K_E^{(t)}}^{-1/2} \\ &= \beta_{K,E,t}(M^*_{f,K_E}). \end{aligned}$$

Therefore $\beta = \nu^{-1}S\mu|\mathcal{M}(K_E)^*$; in words, β is congruent to a restriction of Schur multiplication by $K_E^{(t)}/K_E$. We can therefore extend β to a map $\hat{\beta}$ on all of $M_n(\mathbb{C})$ by taking $\hat{\beta} = \nu^{-1}S\mu$. $\hat{\beta}$ and S are then congruent C^*-algebra maps via the congruences μ and ν, and because congruences are completely positive maps (see Paulsen [Pau86, page 28]) the complete positivity of $\hat{\beta}$ and S are equivalent.

There is one more map that we need. Following Paulsen, we extend β to a map $\tilde{\beta}$ on the self-adjoint span $\mathcal{M}(K_E)^* + \mathcal{M}(K_E)$ by defining

$$\tilde{\beta}(M^*_{f,K_E} + M_{g,K_E}) = \beta(M^*_{f,K_E}) + \beta(M^*_{g,K_E})^*.$$

The importance of this extension is that $\tilde{\beta}$'s domain, being a self-adjoint subspace, is an *operator system*, allowing us to use more of Paulsen's results on complete positivity and contractivity.

Using the congruence between $\hat{\beta}$ and S and the fact that $K_E^{(t)}/K_E$ is Hermitian, it is easily verified that $\tilde{\beta}$ is simply $\hat{\beta}|\mathcal{M}(K_E)^* + \mathcal{M}(K_E)$. We therefore have the following enlarged commuting diagram.

$$
\begin{array}{ccccccc}
\mathcal{M}(K_E)^* & \subset \mathcal{M}(K_E)^* + \mathcal{M}(K_E) & \subset M_n(\mathbb{C}) & \xrightarrow{\ \mu\ } & M_n(\mathbb{C}) \\
\downarrow \beta & \downarrow \tilde{\beta} & \downarrow \hat{\beta} & & \downarrow S \\
\mathcal{M}(K_E^{(t)})^* & \subset \mathcal{M}(K_E^{(t)})^* + \mathcal{M}(K_E^{(t)}) & \subset M_n(\mathbb{C}) & \xleftarrow[\ \nu^{-1}\ ]{} & M_n(\mathbb{C})
\end{array}
$$

Operator Spaces   Operator Systems

We finally have enough machinery to prove (2) $\Leftrightarrow$ (3):

(3)$\Rightarrow$(2):

$K_E^{(t)}/K_E \geq 0 \;\Rightarrow\;$ S is completely positive, since Schur multiplication by a positive matrix is completely positive (see [Pau86, page 31])

$\Rightarrow$ $\hat{\beta} = \nu^{-1} S \mu$ is completely positive, since μ and ν^{-1} are both completely positive

$\Rightarrow$ $\tilde{\beta}$ is completely positive and so also completely contractive, since it is unital [Pau86, prop. 3.5].

$\Rightarrow$ β is completely contractive.

(2)$\Rightarrow$(3): We prove this by induction on n, both (2) and (3) being trivially true for $n = 1$, so assume the result for size $n-1$. Since β is completely contractive and unital then $\tilde{\beta}$ is completely positive, by Paulsen [Pau86, proposition 3.4]. Now consider the effect of $\hat{\beta}$ on the positive elements of $\{I_n - \sum_1^n A_i^* A_i : A_i \in \mathcal{M}(K_E)^*\}$. We have

$$I_n - \sum_1^n A_i^* A_i \geq 0 \;\Leftrightarrow\; \begin{pmatrix} A_1 & 0 & \dots & 0 \\ A_2 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A_n & 0 & \dots & 0 \end{pmatrix} \text{ is contractive}$$

$$\Leftrightarrow \left(\begin{array}{cccc|cccc} 1 & 0 & \dots & 0 & A_1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & A_2 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & A_n & 0 & \dots & 0 \\ \hline A_1^* & A_2^* & \dots & A_n^* & 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 1 \end{array}\right) \geq 0$$

by Paulsen [Pau86, lemma 3.1]

$$\Rightarrow \left(\begin{array}{ccc|ccc} 1 & \dots & 0 & \tilde{\beta}(A_1) & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \dots & 0 \\ 0 & \dots & 1 & \tilde{\beta}(A_n) & \dots & 0 \\ \hline \tilde{\beta}(A_1^*) & \dots & \tilde{\beta}(A_n^*) & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & \dots & 0 & 0 & \dots & 1 \end{array}\right) \geq 0$$

since $\tilde{\beta}$ is completely positive and unital

$$\Leftrightarrow \quad I_n - \sum_1^n \beta(A_i)^* \beta(A_i) \geq 0$$

again using [Pau86, lemma 3.1]

$$\Leftrightarrow \quad \hat{\beta}(I_n - \sum_1^n A_i^* A_i) \geq 0$$

since direct calculation shows that

$\hat{\beta}(A^*A) = \beta(A)^*\beta(A)$ whenever $A \in \mathcal{M}(K_E)^*$.

This shows that $\hat{\beta}$ is positive on $\{I_n - \sum_1^n A_i^* A_i : A_i \in \mathcal{M}(K_E)^*\}$ and so, moving across to S, that S is positive on

$$\mu(\{I_n - \sum_1^n A_i^* A_i : A_i \in \mathcal{M}(K_E)^*\})$$

which is simply the set of matrices

$$\{(J_n - P) \circ K_E : P \in M_n(\mathbb{C}), P \geq 0\}$$

where $\circ$ denotes Schur multiplication and $J_n \in M_n(\mathbb{C})$ is the matrix with all entries equal to 1. Since Schur multiplication by any completely non-zero, rank-1, positive matrix is a positive map with positive inverse, S must also be positive on

$$\{(P_1 - P) \circ K_E : P_1 \text{ completely non-zero, rank-1 and } \geq 0, P \geq 0\}.$$

Now, since K is completely NP so is K_E. Hence, by the induction hypothesis $\kappa^+(1/K_E) = 1$, so

$$1/K_E = P_1 - P$$

where $P_1 \geq 0$, $P \geq 0$ and $\text{rank}(P_1) = 1$. Moreover P_1 must be completely non-zero, since otherwise it would have a zero diagonal entry and the corresponding diagonal entry of K_E would be ≤ 0. Hence

$$J_n = (P_1 - P) \circ K_E$$

is a positive matrix contained in the set on which S is positive, so $K_E^{(t)}/K_E = S(J) \geq 0$. Since this is true for all proper subsets E of X the induction step is complete. ■

We can now characterise the completely NP kernels.

THEOREM 5.2.3 *Let K be a scalar-valued positive definite kernel on a set X. Then K is completely NP if and only if for some ordering of the points of X, K is block-diagonal with each block, K_α say, being completely non-zero and such that $\kappa^+(1/K_\alpha) = 1$.*

Proof: $\Leftarrow$ Let K be a block diagonal with completely non-zero blocks

$$K_\alpha : X_\alpha \to \mathbb{C}$$

satisfying $\kappa^+(1/K_\alpha) = 1$. Then for each $n \in \mathbb{N}$, $H(I_n \otimes K)$ is simply the orthogonal direct sum of the spaces $H(I_n \otimes K_\alpha)$. Consider attempting to isometrically extend a given multiplier $F : E \to \mathbb{C}$ of $H(I_n \otimes K)$ to a new point t that is in block α_t say. M_F is simply the direct sum of the operators M_{F_α}, where $F_\alpha = F|E \cap X_\alpha$, and F's multiplier norm is therefore simply the largest of the multiplier norms of the functions F_α. The problem of isometrically extending F to t therefore reduces to that of isometrically extending F_{α_t}. This is indeed possible if X is finite, since then K_{α_t} is completely NP, by the previous lemma, so $I_n \otimes K$ is therefore NP. If X is infinite, the argument only shows that all finite restrictions of $I_n \otimes K$ are NP, but then $I_n \otimes K$ is itself NP, by corollary 4.4.4. We have therefore shown that all the inflations of K are NP, i.e. that K is completely NP.

$\Rightarrow$ Form the graph $G(K)$, associated with K, with the points of X as vertices and having vertices x and y joined by an edge if and only if $K(x, y) \neq 0$. (Since K is Hermitian $K(x, y) \neq 0 \Leftrightarrow K(y, x) \neq 0$.)

Now consider any two vertices x and y that are joined via a third vertex a, i.e. $K(a, x) \neq 0$ and $K(a, y) \neq 0$. Then $K_{\{a,x,y\}}$ is a completely NP kernel of the form

$$\begin{pmatrix} K(a,a) \neq 0 & K(a,x) \neq 0 & K(a,y) \neq 0 \\ K(x,a) \neq 0 & K(x,x) \neq 0 & K(x,y) \\ K(y,a) \neq 0 & K(y,x) & K(y,y) \neq 0 \end{pmatrix}$$

so by section 5.1 $K(x, y)$ must also be non-zero. It follows that in $G(K)$ any two vertices that are joined via any path along edges are also joined directly by an edge. In other words $G(K)$ is a union of disjoint cliques (a clique being a sub-graph with an edge joining every pair of vertices). Translating this back into the location of zeros in K, this shows that K is block diagonal, with completely non-zero blocks K_α say, for some ordering of the points of X, as claimed.

Finally, since all restrictions of completely NP kernels are completely NP, every finite restriction, L say, of each block K_α is completely NP and so satisfies $\kappa^+(1/L) = 1$, by lemma 5.2.2. Hence $\kappa^+(1/K_\alpha) = 1$ for each block K_α, as claimed. ■

Finally, note that all 1×1 and 2×2 completely non-zero positive-definite kernels L satisfy $\kappa^+(1/L) = 1$. Therefore the 3×3 completely NP kernels comprise all the completely non-zero kernels K for which $\kappa^+(1/K) = 1$, together with all the block diagonal kernels with zeros. These are exactly the NP kernels, characterised in section 5.1, so all NP 3×3 kernels are also completely NP.

5.3 An NP kernel that is not Completely NP

We now know that 4×4 is the smallest possible size of any NP kernel that is not completely NP. We also know, by lemma 5.2.1, that to build such a kernel we must arrange that each of the associated maps $\beta_{K,E,t}$ is contractive but at least one of them is not completely contractive. The prime example of a C^*-algebra map that is contractive but not completely contractive is matrix transposition [Pau86, page 5]. We will therefore attempt to build a kernel

$$K = \begin{pmatrix} 1 & a & b & d \\ a^* & 1 & c & e \\ b^* & c^* & 1 & f \\ d^* & e^* & f^* & 1 \end{pmatrix}$$

on $\{1, \ldots, 4\} \times \{1, \ldots, 4\}$ in such a way that the map $\beta_{K,E,t}$, with $E = \{1, 2, 3\}$ and $t = 4$, is equivalent to matrix transposition.

Using the matricial representation developed in section 5.2, $\beta_{K,E,t}$ is (equivalent to) the map on $M_3(\mathbb{C})$ whose domain is all matrices of the form $K_E^{1/2} D K_E^{-1/2}$, where D is any diagonal 3×3 matrix, and whose action is given by

$$\beta_{K,E,t} : K_E^{1/2} D K_E^{-1/2} \to {K_E^{(t)}}^{1/2} D {K_E^{(t)}}^{-1/2}.$$

One route, therefore, is to arrange that $K_E^{(t)} = K_E^{-T}$ since then $\beta_{K,E,t}$ will be matrix transposition and so will be contractive but, hopefully, not completely contractive.

(We cannot guarantee that $\beta_{K,E,t}$ will not be completely contractive, since $\beta_{K,E,t}$'s domain is not all of $M_3(\mathbb{C})$.)

We have

$$\det(K_E)K_E^{-T} = \begin{pmatrix} 1-cc^* & b^*c-a^* & a^*c^*-b^* \\ bc^*-a & 1-bb^* & b^*a-c^* \\ ac-b & ba^*-c & 1-aa^* \end{pmatrix}$$

and

$$K_E^{(t)} = \begin{pmatrix} 1-dd^* & a-de^* & b-df^* \\ a^*-d^*e & 1-ee^* & c-ef^* \\ b^*-d^*f & c^*-e^*f & 1-ff^* \end{pmatrix}.$$

Therefore if we choose $d = c^*$, $e = b$, $f = a$ and a and b real then

$$\det(K_E)K_E^{-T} = \begin{pmatrix} 1-cc^* & bc-a & c^*a-b \\ bc^*-a & 1-b^2 & ab-c^* \\ ca-b & ab-c & 1-a^2 \end{pmatrix}$$

and

$$K_E^{(t)} = \begin{pmatrix} 1-cc^* & -(bc-a)^* & -(c^*a-b) \\ -(bc^*-a)^* & 1-b^2 & -(ab-c^*)^* \\ -(ca-b) & -(ab-c)^* & 1-a^2 \end{pmatrix}$$

so that all the terms have the correct magnitude. All that remains is to align the phases of the off-diagonal terms and to do this we can Schur multiply $K_E^{(t)}$ by a rank-1 positive of the form $(u_i u_j^*)$ where the u_i, $i = 1, \ldots, 3$, are unitary scalars. The effect of this is to add

$$\begin{aligned} &\phi(u_1)-\phi(u_2) \quad \text{to} \quad (K_E^{(t)})_{12} \\ &\phi(u_2)-\phi(u_3) \quad \text{to} \quad (K_E^{(t)})_{23} \\ \text{and} \qquad &\phi(u_3)-\phi(u_1) \quad \text{to} \quad (K_E^{(t)})_{31} \end{aligned}$$

where $\phi(\cdot)$ denotes the phase of a complex number. We cannot adjust the phases of the 3 above-diagonal terms in K_E independently, since the phase shifts achieved must sum to zero. In fact we can choose suitable u_i to achieve the desired phases

$$\begin{aligned} \Leftrightarrow \quad &\phi(bc-a)-\phi(-(bc-a)^*) \\ &+\phi(ab-c^*)-\phi(-(ab-c^*)^*) \\ &+\phi(c^*a-b)-\phi(-(c^*a-b)) = 0 \pmod{2\pi} \end{aligned}$$

$$\Leftrightarrow \quad \pi - 2\phi(-(bc-a)^*) + \pi - 2\phi(-(ab-c^*)^*) + \pi = 0 \pmod{2\pi}$$

$$\Leftrightarrow \quad \phi(a-bc^*) + \phi(c-ab) = 3\pi/2 \pmod{2\pi}$$

$$\Leftrightarrow \quad \phi((a-bc^*)(c-ab)) = 3\pi/2 \pmod{2\pi}$$

$$\Leftrightarrow \quad \phi((ax - a^2b - bx^2 - by^2 + ab^2x) + i(ay - ab^2y)) = 3\pi/2 \pmod{2\pi}$$

where $c = x + iy$.

Finally, to arrange this last condition we can choose

$$0 < a, 0 < b < 1 \text{ and } y = -\sqrt{(ax - a^2b - bx^2 + ab^2x)/b}$$

since this gives

$$(a - bc^*)(c - ab) = -ia(1-b^2)\sqrt{(ax - a^2b - bx^2 + ab^2x)/b}$$

and one simple way to achieve this is to take

$$0 < a = b = x < 1 \qquad y = -\sqrt{x - 2x^2 + x^3} = -(1-x)\sqrt{x}.$$

These values, together with the values $d = c^*$, $e = b$ and $f = a$ chosen earlier, give

$$K = \begin{pmatrix} 1 & x & x & x + i(1-x)\sqrt{x} \\ x & 1 & x - i(1-x)\sqrt{x} & x \\ x & x + i(1-x)\sqrt{x} & 1 & x \\ x - i(1-x)\sqrt{x} & x & x & 1 \end{pmatrix}$$

so this kernel is such that $K_E^{(t)}$ is a completely non-zero, rank-1 Schur multiple of K_E^{-T}. As mentioned earlier, the rank-1 Schur multiple has no effect on the multiplier norm, so in terms of contractivity $\beta_{K,E,t}$ is now equivalent to a restriction of matrix transposition and so is contractive (but we hope not completely so).

Therefore isometric extension of any multiplier from the first three points to the fourth is always possible. Further, because of K's strong symmetry it is easy to see by row and column swapping that the same must be true for extension from any three points to the remaining point. If we now take $x = \frac{1}{4}$ as a particular example, numerical calculations show that

$$\kappa^+(1/K_{\{1,2,3\}}) = \kappa^+(1/K_{\{2,3,4\}}) = \kappa^+(1/K_{\{3,4,1\}}) = \kappa^+(1/K_{\{4,1,2\}}) = 1$$

so each of the 3×3 principle subkernels are completely NP. Therefore all one-point extensions of multipliers can be done isometrically and so this does indeed give an NP kernel.

However we find by direct calculation that

$$\kappa^{+}(1/K) = 2$$

so K is *not* completely NP. Therefore

$$K = \begin{pmatrix} 1 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} + \frac{3i}{8} \\ \frac{1}{4} & 1 & \frac{1}{4} - \frac{3i}{8} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} + \frac{3i}{8} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} - \frac{3i}{8} & \frac{1}{4} & \frac{1}{4} & 1 \end{pmatrix}$$

is an explicit example of an NP kernel that is *not* completely NP.

Chapter 6

Generalised Kernels

So far, our use of reproducing kernels has been based on Aronszajn's paper[Aro50] linking the positive kernels on a set X with the Hilbert spaces of functions on X. The positivity of the inner product in such a space stems directly from the positivity of the kernel, so it is natural to hope that this link between kernels and spaces can be extended to one between the Hermitian kernels on X and the indefinite inner product spaces of functions on X. One would then expect the Hermitian kernels of finite negativity to give rise to reproducing kernel Pontrjagin spaces and those of infinite negativity to give reproducing kernel Kreĭn spaces.

In 1964 Laurent Schwartz showed that such a link does indeed exist[Sch64], provided we restrict ourselves to Hermitian kernels that are the difference of two positive kernels. Schwartz's brilliant paper in fact does much more than this—it explores and generalises the whole area of kernels and associated spaces—but unfortunately seems to have been missed by the literature on reproducing kernels. As a result, the extension of Aronszajn's work to reproducing kernel Pontrjagin spaces was later independently proved by Sorjonen[Sor75] and even later, and again independently, by Alpay[AD]. Alpay later discovered Schwartz's work and brought it to my attention.

This chapter is a short interlude from interpolation questions, with a twofold purpose:

- to briefly describe Schwartz's approach and results, since they give great insight into reproducing kernels.
- to use his approach to prove a result on the inertia of kernels, which we will need in the next chapter.

6.1 Schwartz Kernels

Schwartz's kernels are more general beasts than Aronszajn's—they are operators on topological vector spaces—so we must start with some definitions. Let E be any locally convex, quasi-complete, Hausdorff topological vector space, where quasi-complete means that all closed bounded subsets are complete. We denote by $\bar{E}'$ the conjugate dual of E, i.e. the space of all continuous linear functionals on E endowed with pointwise addition and the conjugated scalar multiplication given by

$$(\lambda f)(e) = \lambda^* f(e) \qquad (e \in E, f \in \bar{E}', \lambda \in \mathbb{C}).$$

Although E starts with a given topology we will always work with the weak topology induced on E by $\bar{E}'$; similarly on $\bar{E}'$ we will use the weak topology induced by E, i.e. the weakest topology that makes all the point evaluation functionals continuous.

A Schwartz kernel relative to E is any weakly continuous linear operator from $\bar{E}'$ to E. We can add, subtract and scalar multiply Schwartz kernels in the usual pointwise manner.

Because we are using the conjugate dual of E, the natural scalar product of elements from E and $\bar{E}'$ given by $(e, f) = f(e)$ ($e \in E$, $f \in \bar{E}'$) is a conjugate-linear product. The interaction between this scalar product and the weakly continuous operators from $\bar{E}'$ to E closely follows that between the inner product on a Hilbert space and the operators on that space. For example we can define the adjoint of a kernel K to be the kernel K^* given by

$$(K^* f_1, f_2) = (K f_2, f_1)^* \text{ whenever } f_1, f_2 \in \bar{E}'$$

and call K Hermitian if and only if $K^* = K$ and positive, denoted $K \geq 0$, if and only if $(Kf, f) \geq 0$ for all $f \in \bar{E}'$. As with operators on Hilbert spaces, the positive kernels are all Hermitian and form a convex cone (using the partial order $L \geq K \Leftrightarrow L - K \geq 0$) in the space of all kernels relative to E.

Given a Hermitian kernel K we can define an inner product $\langle \cdot, \cdot \rangle$ on $\text{ran}(K)$ by $\langle Kf, Kg \rangle = (Kf, g) = g(Kf)$. Because K is Hermitian this is well-defined, since if $Kf = Kf_1$ and $Kg = Kg_1$ we have

$$\begin{aligned}
\langle Kf_1, Kg_1 \rangle &= (Kf_1, g_1) \qquad \text{by definition} \\
&= (Kf, g_1) \\
&= (Kg_1, f)^* \qquad \text{since } K^* = K \\
&= (Kg, f)^* \\
&= (Kf, g) \qquad \text{again since } K^* = K \\
&= \langle Kf, Kg_2 \rangle.
\end{aligned}$$

However, although this *is* an inner product space naturally associated with the Hermitian kernel K, we wish to associate a *complete* space with K and as yet we have no topology defined on $\text{ran}(K)$.

Now suppose K is positive rather than just Hermitian; then this inner product on $\text{ran}(K)$ is positive-definite and so induces a norm and a topology. Schwartz shows that, with this topology, $\text{ran}(K)$ has a unique completion that is continuously embedded in E. This completion is called the reproducing kernel Hilbert space associated with K, denoted $H(K)$, since for it the kernel K satisfies the following reproducing property:

$$f(e) = \langle e, Kf \rangle \text{ whenever } e \in H(K) \text{ and } f \in \bar{E}'.$$

Schwartz shows that the spaces $H(K)$ that arise from the positive kernels are exactly those Hilbert spaces that are continuously embedded in E (from the norm topology on $H(K)$ to the weak topology on E). Indeed, on the set of such spaces, denoted Hilb(E), he defines the following partial order, addition, multiplication by non-negative reals

- $H_1 \leq H_2 \Leftrightarrow H_1$ is contractively embedded in H_2
- $H_1 + H_2 = \text{span}(H_1, H_2)$ with norm $\|h\|^2 = \inf_{h=h_1+h_2} \|h_1\|^2 + \|h_2\|^2$
- $\lambda H = H$ with the inner product scaled by $1/\lambda$ ($\lambda H = \{0\}$ if $\lambda = 0$)

and shows that Hilb(E) is then a convex cone which is isomorphic to the cone of positive kernels via the identification $K \leftrightarrow H(K)$. This isomorphism allows us to translate between an operator-theoretic view and a spatial view; for example it tells us that if K and L are two positive kernels relative to E then $K \leq L$ if and only if $H(K)$ is contractively embedded in $H(L)$.

Before we consider non-positive Hermitian kernels, it is worth noting that under this approach any Hilbert space H can be considered as a reproducing kernel Hilbert space, simply by taking E to be any locally convex, quasi-complete, Hausdorff topological vector space in which H is continuously embedded. One important example is to take E to be a Hilbert space containing H. E is then self-dual and if we use the natural identification of $\bar{E}'$ with E then H's reproducing kernel relative to E is simply TT^*, where T is the embedding of H in E. There are therefore many possible choices for E, and H has many reproducing kernels, exactly one for each possible choice for E.

6.2 Reproducing Kernel Kreĭn Spaces

We have seen above that there is a one-to-one correspondence between the positive kernels relative to E and the Hilbert spaces continuously embedded in E. Things do not work out quite so neatly when this is extended to non-positive Hermitian kernels. Here is a summary of Schwartz's results:

- Each Kreĭn space H that is continuously embedded in E (from the strong topology on H to the weak topology on E) has a unique corresponding reproducing Hermitian kernel K, which can be expressed as the difference of two positive kernels. I will call a kernel that can be so expressed a *splittable* kernel.

- Depending on E there may exist Hermitian kernels which are *not* splittable, in which case they do not have any corresponding reproducing kernel Kreĭn spaces.

- Each splittable Hermitian kernel K is the reproducing kernel for at least one Kreĭn space continuously embedded in E. In general there may be many such spaces with the same reproducing kernel, in which case Schwartz calls K a kernel of multiplicity. If there is only one he calls K a kernel of uniqueness and only in this case can we refer to *the* reproducing kernel space with kernel K and use the notation $H(K)$ unambiguously.

- Two positive kernels K and L are called independent if $H(K) \cap H(L) = \{0\}$. A splittable Hermitian kernel K can always be expressed as the difference of two independent positive kernels.

- If a Hermitian kernel K can be split into the difference of two positive kernels such that one of them is of finite rank then K is a kernel of uniqueness. In other words, distinct Pontrjagin spaces continuously embedded in E have distinct reproducing kernels, which are kernels of uniqueness.

For a splittable Hermitian kernel K these results allow us to define the positivity $\kappa^+(K)$ and negativity $\kappa^-(K)$ to equal the dimensions of any maximal positive-definite and negative-definite subspaces, respectively, of any of its corresponding reproducing kernel Kreĭn spaces. There is no ambiguity here since if there is a choice here for the Kreĭn space they must all have infinite positivity and infinite negativity.

In the case described earlier where E is a Hilbert space and $K = TT^*$, T being the embedding of $H(K)$ in E, these definitions of κ^- and κ^+ coincide with the usual definitions of the negativity and positivity of the Hermitian operator TT^*. That is $\kappa^-(TT^*)$ and $\kappa^+(TT^*)$ equal the dimensions of the spectral subspaces corresponding to $(-\infty, 0)$ and $(0, \infty)$ respectively. The definitions also agree with the more elementary definitions given earlier for the negativity and positivity of a scalar-valued Aronszajn kernel. That is

$$\kappa^{\pm}(K) = \sup_{F \text{ finite}} \text{ number of +ve/-ve eigenvalues of } K_F.$$

6.3 Relationship to Aronszajn's Kernels

To see how Schwartz's kernels generalise Aronszajn's we must choose E to be the vector space $\mathbb{C}^X$ of all complex-valued functions on the domain set X, with the weak topology, i.e. that of pointwise convergence. $\bar{E}'$ is then simply the finite linear span of the point evaluation functionals and can therefore be identified with the space of all finitely-supported functions on X by taking a finitely-supported function f to correspond to the linear functional on E given by

$$e \mapsto \sum_{x \in \text{support}(f)} f(x)e(x).$$

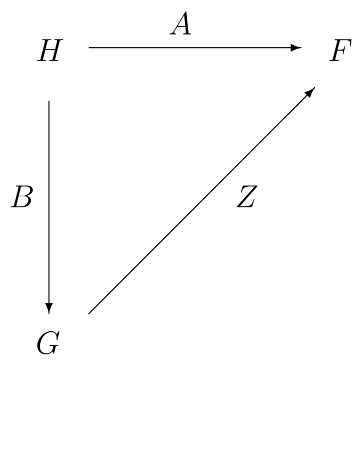
Using this identification the Dirac delta functions $\{\delta_x : x \in X\}$ correspond to the point evaluation functionals and $\bar{E}'$ is simply $\text{span}\{\delta_x : x \in X\}$, a dense subspace of E.

The delta functions form a natural basis for both $\bar{E}'$, which they span, and also for E, of which they span a dense subspace, and with respect to these bases the weakly continuous linear operator from $\bar{E}'$ to E correspond exactly to the set of *all* $X \times X$ matrices. The Aronszajn kernels on X are hence simply the matrices, with respect to these bases, of the Schwartz kernels relative to $\mathbb{C}^X$.

6.4 Negativity of Kernel Differences

If K and L are two positive Schwartz kernels relative to E, we know from Schwartz's results that $L-K \geq 0$ if and only if $H(K)$ is contractively embedded in $H(L)$, but what happens if $L-K$ is *not* positive? In this section we use Schwartz's results to obtain a result that we will need in the next chapter, namely a characterisation of $\kappa^-(L-K)$ in terms of the relationship between $H(K)$ and $H(L)$. Our approach is to first prove the corresponding result for Hilbert space operators, by generalising Douglas's lemma, and then extend this to Schwartz kernels.

Lemma 6.4.1 *In the following diagram let A and B be given bounded linear operators between Hilbert spaces F, G and H.*



Then

$$\kappa^-(B^*B - A^*A) = \min_{\|Z\|\le 1} \operatorname{rank}(A - ZB).$$

Note that ∞ is a possible value for each side.

Proof: We will first show that

$$\kappa^-(B^*B - A^*A) \le \min_{\|Z\|\le 1} \operatorname{rank}(A - ZB).$$

Let $Z : G \to F$ be any contraction, $J = \ker(A - ZB)$, $k = \operatorname{rank}(A - ZB) = \operatorname{codim}(J)$ and assume $k < \infty$. Then for any $h \in J$ we have

$$\begin{aligned}
\langle (B^*B - A^*A)h, h\rangle &= \|Bh\|^2 - \|Ah\|^2 \\
&= \|Bh\|^2 - \|ZBh\|^2 \text{ since } h \in \ker(A - ZB) \\
&\ge \|Bh\|^2 - \|Bh\|^2 \text{ since } Z \text{ is a contraction} \\
&= 0
\end{aligned}$$

so $(B^*B - A^*A)|J$ is positive. Hence J must have trivial intersection with the spectral subspace of $B^*B - A^*A$ corresponding to $(-\infty, 0)$, on which $B^*B - A^*A$ is negative definite, so that spectral subspace must have dimension $\le k$. Therefore $\kappa^-(B^*B - A^*A) \le k$, as claimed.

Conversely, to show

$$\kappa^-(B^*B - A^*A) \ge \min_{\|Z\|\le 1} \operatorname{rank}(A - ZB)$$

let $\kappa^-(B^*B - A^*A) = k < \infty$ and J be the spectral subspace of $B^*B - A^*A$ corresponding to $[0, \infty)$. Then $(B^*B - A^*A)|J$ is positive, so define Z on $B(J)$ by $Zx = Ay$ where y is any vector in J for which $x = By$. This defines Zx unambiguously since if $y' \in J$ is another such vector then

$$\begin{aligned}
0 \le \|Ay - Ay'\|^2 &= \langle A^*A(y - y'), y - y'\rangle \\
&\le \langle B^*B(y - y'), y - y'\rangle \\
&= \|B(y - y')\|^2 \\
&= 0.
\end{aligned}$$

Also, Z is contractive on $B(J)$ since for $x \in B(J)$

$$\|Zx\|^2 = \|Ay\|^2 = \langle A^*Ay, y\rangle \leq \langle B^*By, y\rangle = \|By\|^2 = \|x\|^2.$$

We can now extend Z to a contraction on the whole of G, by extending by continuity to the closure of $B(J)$ and then by zero extension to the orthogonal complement that remains. For the resulting contraction Z we have that $J \subseteq \ker(A - ZB)$, so

$$\operatorname{rank}(A - ZB) = \operatorname{codim}(\ker(A - ZB)) \leq \operatorname{codim}(J) = \kappa^-(B^*B - A^*A) = k.$$

Therefore $\kappa^-(B^*B - A^*A) \geq \min_{\|Z\|\leq 1} \operatorname{rank}(A - ZB)$ as claimed and the proof is complete. ∎

We can now extend this to differences of Schwartz kernels.

THEOREM 6.4.2 *Let K and L be positive Schwartz kernels relative to a locally convex, quasi-complete, Hausdorff topological vector space E. Then*

$$\kappa^-(L - K) = \min(\operatorname{codim}(J))$$

where the minimum is taken over all closed subspaces J of $H(K)$ that are contractively embedded in $H(L)$.

Proof: Apply the above generalised Douglas's lemma with

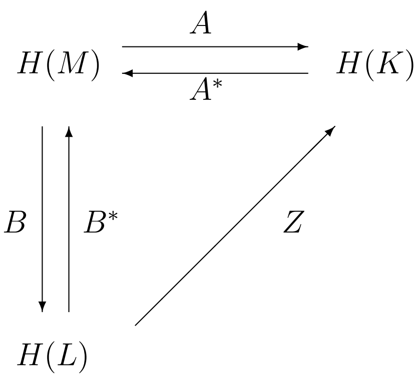
H = any member $H(M)$ of $\operatorname{Hilb}(E)$ that contains $H(K)$ and $H(L)$.
For example taking $M = K + L$ will suffice.

F = $H(K)$

G = $H(L)$

A^* = embedding of $H(K)$ in $H(M)$

B^* = embedding of $H(L)$ in $H(M)$.



Because $H(K)$ and $H(L)$ are continuously embedded in $H(M)$ we can consider their kernels relative to $H(M)$ as well as relative to E. They are A^*A and B^*B, respectively, and by Schwartz's results [Sch64, Proposition 21] we also have that $K = A^*AM$ and $L = B^*BM$.

Now let H be a Kreĭn space continuously embedded in $H(M)$ whose kernel relative to $H(M)$ is $B^*B - A^*A$; there is at least one such. Since $H(M)$ is continuously embedded in E so is H. Moreover, again by Schwartz [Sch64, Proposition 21], its kernel relative to E is $B^*BM - A^*AM = L - K$. In other words $B^*B - A^*A$ and $L - K$ are the kernels of the same Kreĭn space H relative to two different containing topological vector spaces. Hence they must have the same negativity, i.e. $\kappa^-(L - K) = \kappa^-(B^*B - A^*A)$.

Having changed our underlying topological space from E to the Hilbert space $H(M)$ we can now apply lemma 6.4.1, giving

$$\begin{aligned}
\kappa^-(L-K) = \kappa^-(B^*B - A^*A) &= \min_{\|Z\|\leq 1} \operatorname{rank}(A - ZB) \\
&= \min_{\|Z\|\leq 1} \operatorname{codim}(\ker(A - ZB)) \\
&= \min_{\|Z^*\|\leq 1} \operatorname{codim}(\ker(A^* - B^*Z^*)).
\end{aligned}$$

But if $Z^* : H(K) \to H(L)$ is a contraction then for all $v \in \ker(A^* - B^*Z^*)$

$$v = A^*v \qquad \text{since } A^* \text{ is the embedding of } H(K) \text{ into } H(M)$$

$$= B^*Z^*v \qquad \text{since } v \in \ker(A^* - B^*Z^*)$$

$$= Z^*v \qquad \text{since } B^* \text{ is the embedding of } H(L) \text{ into } H(M)$$

so Z^* contractively embeds $\ker(A^* - B^*Z^*)$ in $H(L)$. This argument is clearly reversible, so the contractive embeddings of closed subspaces of $H(K)$ into $H(L)$ correspond exactly to contractions $Z^* : H(K) \to H(L)$ and the associated spaces $\ker(A^* - B^*Z^*)$. Therefore, as claimed, $\kappa^-(L - K)$ equals the smallest possible codimension of any closed subspace of $H(K)$ that is contractively embedded in $H(L)$. ∎

Our detour into Schwartz's work has enabled us to prove the above theorem, which we will need in the next chapter. However, we will not otherwise use the huge generality of Schwartz's kernels—in all that follows the kernels considered are Aronszajn kernels, i.e. matrices over a set X.

Chapter 7

The Adamyan-Arov-Kreĭn Theorem

From our reproducing kernel viewpoint, Pick's theorem is that $\|M_f\|$ is the smallest possible norm of any bounded analytic function on $\mathbb{D}$ that interpolates $f : E \to \mathbb{C}$, where M_f is the operator on $H(K_E)$ of multiplication by f and K is the Szegö kernel on $\mathbb{D}$. $\|M_f\|$ is the first singular value of M_f and in 1971 V.M. Adamyan, D.Z. Arov and M.G. Kreĭn [AAK71] proved a result which showed that further information about 'smallest' interpolating functions can be extracted from the operator M_f by looking at its other singular values. This chapter develops a reproducing kernel Hilbert space approach to proving their result and studies whether it can be generalised to multipliers of other reproducing kernel Hilbert spaces.

7.1 Introduction

The Adamyan-Arov-Kreĭn theorem (AAK theorem for short) is not, in its original form, an interpolation result. However it can be reformulated as one and it is this form that we will state and work on throughout. To state the AAK theorem accurately we need to make some definitions. Firstly we define notation for singular values. Let $s_0(A), s_1(A), \ldots$ denote the singular values of an operator A on a Hilbert space H, defined by

$$s_k(A) = \inf \|A|J\|$$

where the infimum is over all closed subspaces J of H of codimension $\leq k$. We will be working almost exclusively with multiplication operators on reproducing kernel Hilbert spaces, so for convenience we will extend the concept of singular values to apply to functions, as follows. Relative to a given positive-definite kernel K on a set X, we define the singular values of any function f on a non-empty subset E of X by

$$s_k(f, K) = \inf \|M_f : J \to H(K_E)\|$$

where J ranges over all closed subspaces of $H(K_E)$ of co-dimension $\leq k$. Of course when E is infinite, multiplication by f may not map J into $H(K_E)$ or else may do so unboundedly; in these cases we take the norm in the definition to be ∞. The singular values of a function relative to a kernel are hence a decreasing sequence of non-negative real or infinite values. Since the underlying kernel K we use will always be fixed, we will usually abbreviate $s_k(f, K)$ to $s_k(f)$.

Secondly, let $b_t = (z - t)/(1 - zt^*)$ denote the Blaschke factor associated with the point $t \in \mathbb{D}$ and, for each $k \in \mathbb{N}$, let B_k denote the set of Blaschke products with up to k factors, i.e. the functions of the form $b_{t_1} b_{t_2} \ldots b_{t_i}$ for some set of (not necessarily distinct) points $t_1, \ldots, t_i$ $(i \leq k)$ in the disc. Note that, for convenience, we will take the constant function 1 to be the (unique) Blaschke product with no factors, so $B_0 = \{1\}$.

We can now state the AAK theorem in the form that we will prove:

> Let K be the Szegö kernel and $f : E \to \mathbb{C}$ be any given function on a non-empty subset E of $\mathbb{D}$. Then for each $k = 0, 1, \ldots$
>
> $$s_k(f) = \min \|\phi\|_\infty$$

where the minimum is taken over all $\phi \in H^\infty$ such that $\phi|E = fb|E$ for some $b \in B_k$ or is taken to be ∞ if there is no such ϕ.

I have expressed the theorem here in a very compact form, but at the cost of masking what it tells us about functions that interpolate f. To obtain information about interpolating f we must divide the equation $\phi|E = fb|E$ through by b, so let us examine what happens when we do this.

Assume first that b has no zeros in E. Then $(\phi/b)|E = f$ and so ϕ/b is a meromorphic function, with at most k poles, which interpolates f and has supremum norm $s_k(f)$ around the disc edge (since b is unimodular there). What happens if b has a zero in E, at α say? Then $\phi(\alpha) = b(\alpha)f(\alpha) = 0$ so when we divide through b's zero at α is cancelled by ϕ's zero there so ϕ/b has one less pole. However, to compensate for ϕ/b being, as it were, one step nearer to analyticity, we no longer know that $(\phi/b)(\alpha) = f(\alpha)$ so we are possibly one step away from interpolating f. This pattern continues when b has multiple zeros in E; for each zero in E, ϕ/b has one less pole but has one more possible interpolation failure.

These arguments all turn out to be reversible, so when the AAK theorem is interpreted as a statement about interpolating f with a meromorphic function it says

> $s_k(f)$ is the smallest possible supremum norm (around the disc edge) of any meromorphic function on $\mathbb{D}$ that has a total of at most k failures of analyticity or points where it fails to interpolate f. If there is no such function then $s_k(f) = \infty$.

When E is finite, the zeros of b are often all outside E and then an interpolating meromorphic function is obtained. Because of this the AAK theorem is sometimes thought of as saying that f can be interpolated by a meromorphic function that is bounded by $s_k(f)$ around the disc edge and has at most k poles, but this is not always true. A simple counter-example is obtained by taking E to have k or fewer points. Then multiplication of $H(K_E)$ by f has rank $\leq k$, so $s_k(f) = 0$ and therefore ϕ must be zero around the disc edge. Being meromorphic, ϕ must be the zero function, so in this case ϕ is analytic but fails to interpolate f wherever it is non-zero.

It is easy to overlook the possibility that b may have zeros in E. This is why I have preferred to state the AAK theorem in terms of an analytic function ϕ that interpolates $fb|E$. In this form the k 'failures' of analyticity or interpolation are treated equally—they are simply zeros of b.

7.2 Kernel Characterisation of s_k

An important tool in our proof of Pick's theorem was the characterisation of the multiplier norm in terms of kernel positivity:

$$s_0(f) = \|f\|_{M(K)} = \inf_{r \geq 0}(r : (r^2 - f \otimes f^*)K \geq 0).$$

Before we can tackle the AAK theorem we will need the corresponding tool for the case $k > 0$, given by the following lemma.

Lemma 7.2.1 *Let X be any non-empty set and $K : X \times X \to \mathbb{C}$ be a positive definite kernel on X. Then the singular values of a function $f : X \to \mathbb{C}$ relative to K are given by*

$$s_k(f) = \inf_{r \geq 0}(r : \kappa^-((r^2 - f \otimes f^*)K) \leq k) \qquad k = 0, 1, \ldots$$

and the infimum, if finite, is attained.

Proof: Consider the space $M_f(H(K))$, the image of $H(K)$ under multiplication by f, endowed with the inner product defined by

$$\langle fh_1, fh_2\rangle = \langle h_1, h_2\rangle_{H(K)} \text{ whenever } h_1, h_2 \in H(K) \ominus \ker(M_f).$$

Direct calculation confirms that $(f \otimes f^*)K$ has all its columns in $M_f(H(K))$ and satisfies the reproducing property for the inner product defined, so this space is $H((f \otimes f^*)K)$. In other words, $H((f \otimes f^*)K)$ is simply the image of $H(K)$ under multiplication by f, given the inner product mapped across from $H(K) \ominus \ker(M_f)$ unitarily by M_f.

For any subspace J of $H(K)$ we have $\|M_f|J + \ker(M_f)\| = \|M_f|J\|$ and $\text{codim}(J+\ker(M_f)) \leq \text{codim}(J)$, so in the definition of $s_k(f)$ we may restrict attention to the closed subspaces J of $H(K)$ that contain $\ker(M_f)$. But if M_f contractively maps such a subspace J back into $H(K)$ then $H = M_f(J)$ is a closed subspace of $H((f \otimes f^*)K)$ that is contractively embedded in $H(K)$ and

$$\text{codim}(J \text{ in } H(K)) = \text{codim}(H \text{ in } H((f \otimes f^*)K)).$$

Conversely, if H is a closed subspace of $H((f \otimes f^*)K)$ that is contractively embedded in $H(K)$, then $J = M_f^{-1}H$ is a closed subspace of $H(K)$ that contains $\ker(M_f)$ and is contractively mapped back into $H(K)$ by multiplication by f, and again

$$\text{codim}(J \text{ in } H(K)) = \text{codim}(H \text{ in } H((f \otimes f^*)K)).$$

Therefore, by simple scaling, the closed subspaces J of $H(K)$ that contain $\ker(M_f)$ and are mapped by multiplication by f back into $H(K)$ with norm r correspond exactly to the closed subspaces $H = M_fJ$ of $H((f \otimes f^*)K)$ that are embedded in $H(K)$ with norm r. Hence

$$\begin{aligned} & s_k(f) \\ = \; & \inf(\|M_f|J\| : J \text{ a closed subspace of } H(K), \ker(M_f) \subseteq J, \text{codim}(J) = k) \end{aligned}$$

$$
\begin{aligned}
&= \inf(\|\text{embedding } H \to H(K)\|) \\
&\quad \text{where the infimum is over all closed, } k\text{-codimensional subspaces} \\
&\quad H \text{ of } H((f \otimes f^*)K) \text{ that are embedded in } H(K) \\
&= \text{infimum of } r \text{ for which a } k\text{-codimensional subspace of } H((f \otimes f^*)K) \\
&\quad \text{is contractively embedded in } H(r^2 K) \\
&\quad (\text{since } H(r^2K) \text{ is } H(K) \text{ with the norm scaled by } 1/r) \\
&= \inf_{r \geq 0}(r : \kappa^-((r^2 - f \otimes f^*)K) \leq k) \text{ by lemma 6.4.2.}
\end{aligned}
$$

This last infimum is attained, since

$$
\begin{aligned}
& \kappa^-((r^2 - f \otimes f^*)K) \leq k \text{ for all } r > s_k(f) \\
\Rightarrow\ & \kappa^-((r^2 - (f|F) \otimes (f|F)^*)K_F) \leq k \text{ for all finite } F \subseteq X \text{ and } r > s_k(f) \\
\Rightarrow\ & \kappa^-((s_k(f)^2 - (f|F) \otimes (f|F)^*)K_F) \leq k \text{ for all finite subsets } F \subseteq X \\
& \text{since going to the limit of a weakly convergent sequence} \\
& \text{of matrices cannot increase negativity} \\
\Rightarrow\ & \kappa^-((s_k(f)^2 - f \otimes f^*)K) \leq k.
\end{aligned}
$$

All the other infima used must therefore also be attained, since they have been shown to be over exactly corresponding sets of values. ■

This characterisation of s_k is a valuable tool. For instance, because the negativity of a kernel is the supremum of the negativities of its finite restrictions, it gives us the following corollary.

Corollary 7.2.2 *Under the same assumptions as lemma 7.2.1 we have*

$$s_k(f) = \sup(s_k(f|F) : F \textit{ is a finite subset of } X)$$

and therefore extending a function cannot reduce its singular values.

Finally in this section, we will use our characterisation of s_k to obtain the following interlacing result for the Blaschke kernels that we defined in section 3.2.

LEMMA 7.2.3 *Let $K : X \times X$ be a Blaschke kernel and b_t denote its generalised Blaschke factors. Then for any function $f : E \to \mathbb{C}$ on a non-empty subset E of X and any $t \in X$*

$$s_{k+1}(f) \leq s_k(fb_t|E) \leq s_k(f).$$

Proof: The kernel whose negativity characterises $s_k(fb_t|E)$ is given by:

$$\begin{aligned}
& (r^2 - (fb_t|E) \otimes (fb_t|E)^*)K_E \\
= \; & r^2 K_E - (f \otimes f^*)K_E^{(t)} \text{ since } K \text{ is a Blaschke kernel} \\
= \; & (r^2 - f \otimes f^*)K_E + (f \otimes f^*)(K(\cdot,t) \otimes K(\cdot,t)^*)/K(t,t) \\
= \; & (\text{kernel that characterises } s_k(f)) + (\text{a rank 1 positive kernel}).
\end{aligned}$$

The result therefore follows from lemma 7.2.1 and the fact that adding a rank 1 positive kernel to a given kernel can only either leave the negativity unchanged or else reduce it by 1. ■

7.3 s_k for Functions on the Disc

For the special case of $E = \mathbb{D}$ the AAK theorem effectively states that

$$\begin{aligned}
s_k(f) < \infty \quad \Leftrightarrow \quad & \text{there exists a Blaschke product } b \in B_k \text{ such that } fb \in H^\infty \\
& \text{and then } s_k(f) = \|fb\|_\infty.
\end{aligned}$$

As a first step to proving the theorem we will now prove this. To do so we first need the following general result.

LEMMA 7.3.1 *Let $K : X \times X \to \mathbb{C}$ be a positive definite kernel on a set X. Then any function ϕ that is a bounded multiplier of a dense subspace S of $H(K)$ back into $H(K)$, is also a bounded multiplier on the whole of $H(K)$, with the same bound.*

Proof: Assume, without loss of generality, that $\|M_\phi|S\| = 1$ and let f be any function in $H(K)$, with norm r say. Since S is a dense subspace of $H(K)$ we can find functions $f_i \in S$, $i \in \mathbb{N}$, with the same norm r, such that $f_i \to f$ in norm and hence also weakly (i.e. pointwise). Then

$$\begin{aligned} \|M_\phi|S\| = 1 \quad &\Rightarrow \quad \|\phi f_i\| \leq r \text{ for all } i \\ &\Rightarrow \quad r^2K - (\phi f_i)(\phi f_i)^* \geq 0 \text{ for all } i \\ &\Rightarrow \quad r^2K - (\phi f)(\phi f)^* \geq 0 \end{aligned}$$

where the last implication holds because $r^2K-(\phi f_i)(\phi f_i)^* \to r^2K-(\phi f)(\phi f)^*$ pointwise and the pointwise limit of positive kernels must also be positive. Therefore $\phi f \in H(K)$ and $\|\phi f\| \leq r = \|f\|$, so ϕ indeed multiplies all functions in $H(K)$ contractively back into $H(K)$. ∎

We can now give function-theoretic meaning to the singular values of functions on the disc.

Lemma 7.3.2 *Let $f : \mathbb{D} \to \mathbb{C}$ be any given function and K be the Szegö kernel. If there exists a Blaschke product $b \in B_k$ such that $fb \in H^\infty$ then*

$$s_k(f) = \|fb\|_\infty = \lim_{r\to 1-} \sup_{z\in\mathbb{T}} |f(rz)|.$$

Otherwise $s_k(f) = \infty$.

Proof: We first prove that if $b \in B_k$ and $fb \in H^\infty$ then $s_k(f) \leq \|fb\|_\infty$:

If such a Blaschke product b exists, then M_{fb} maps H^2 into H^2 with norm $\|fb\|_\infty$, so M_f maps bH^2 into H^2 with the same norm, since multiplication by any Blaschke product is an isometry on H^2. But bH^2 has codimension in H^2 equal to the number of Blaschke factors in b, i.e. no more than k, so taking $J = bH^2$ in the definition of $s_k(f)$ gives the desired conclusion.

Secondly, we show that if $s_k(f) < \infty$ then there exists $b \in B_k$ such that $fb \in H^\infty$ and $\|fb\|_\infty \leq s_k(f)$:

Because scaling a function simply scales its singular values, we may assume that $s_k(f) = 1$. Consider the linear manifold $H = \{h \in H^2 : fh \in H^2\}$ of functions in H^2 that M_f maps back into H^2. H is clearly invariant under multiplication by z and moreover so is its closure in H^2, since if $h_i \in H$ and $h_i \to h \in H^2$ in norm, then $zh_i \to zh$ in norm and $zh_i \in H$, so $zh \in \text{clos}(H)$. Therefore, by Beurling's theorem [Beu49], $\text{clos}(H) = bH^2$ for some inner function b. Since $s_k(f) < \infty$ then, by the definition of $s_k(f)$, H must have codimension $\leq k$ in H^2, so b must be a finite Blaschke product with at most k factors.

Now consider the unitary equivalence $M_b : H^2 \to bH^2$.

$$
\begin{array}{ccc}
H & \subseteq & bH^2 = \text{clos}(H) \\
\uparrow M_b & & \uparrow M_b \\
M_b^{-1}H & \subseteq & H^2
\end{array}
$$

Since H is dense in bH^2 then, by this equivalence, $M_b^{-1}H$ is dense in H^2, so multiplication by fb contractively maps a dense subspace of H^2, i.e. $M_b^{-1}H$, back into H^2. By lemma 7.3.1, fb is therefore a contractive multiplier on the whole of H^2, so $fb \in H^\infty$ and $\|fb\|_\infty \leq 1$, as claimed.

Finally note that, because Blaschke products are unimodular around the disc edge, $\|fb\|_\infty = \lim_{r \to 1-} \sup_{z \in \mathbb{T}} |f(rz)|$ for any Blaschke product b such that $fb \in H^\infty$. The lemma now follows from the fact that this value is independent of the Blaschke product b chosen, together with the above two implications.

■

Corollary 7.3.3 *Relative to the Szegö kernel, the singular values of any function* $f : \mathbb{D} \to \mathbb{C}$ *are given by*

$$(s_0(f), \ldots, s_{k-1}(f), s_k(f), s_{k+1}(f), \ldots) = (\infty, \ldots, \infty, s, s, \ldots)$$

where $s = \lim_{r \to 1-} \sup_{z \in \mathbb{T}} |f(rz)|$ *and* k *is the smallest integer such that* $fb \in H^\infty$ *for some* $b \in B_k$*.*

Taking $\phi = fb$ in this corollary completes our proof that the AAK theorem holds when $E = \mathbb{D}$. However, the purpose of proving this has been mainly to enable us to see that the AAK theorem can be viewed as a function extension problem.

Starting from the function f with k'th singular value $s_k(f)$, suppose we could extend f from E to $\mathbb{D}$ without increasing the k'th singular value. The resulting function $\psi : \mathbb{D} \to \mathbb{C}$ would satisfy $s_k(\psi) = s_k(f)$ and so by lemma 7.3.2 there would exist a Blaschke product $b \in B_k$ such that $\psi b \in H^\infty$ and $\|\psi b\|_\infty = s_k(f)$. The function $\phi = \psi b$ would therefore satisfy $\phi \in H^\infty$, $\|\phi\|_\infty = s_k(f)$ and $\phi|E = \psi b|E = fb|E$ and so satisfy the requirements of the theorem.

It appears, therefore, that proving the AAK theorem is now a matter of showing that a function can always be extended from E to $\mathbb{D}$ without increasing its k'th singular value. Our approach will therefore be as with Pick's theorem in chapter 1. The main issue involved is whether it is always possible to one-point extend a function without increasing s_k. The next section tackles this problem.

7.4 One Point Extension

We now consider the problem of one-point extending a given function $f : E \to \mathbb{C}$, where E is any proper subset of X, to a new point t without increasing k'th singular value. Our approach is just as in the proof of Pick's theorem in section 1.4, except that now the k'th singular value replaces the norm, i.e. k is now non-zero. There we translated the function extension problem into a corresponding multiplication operator completion problem and then used Parrott's theorem to obtain the lower bound on the norm of any completion. I will not repeat the argument step by step, but will instead concentrate on the changes that must be made for the argument to work for non-zero k.

The first difficulty is that multiplication by f might now *not* be a bounded operator on $H(K_E)$, since we only know that $s_k(f)$ is bounded. Note, however, that this difficulty does not arise if E is finite, since then $H(K_E)$ is finite dimensional and contains all functions on E and we know that all operators on it are bounded. Let us therefore assume that E is finite; we will see later, in lemma 7.4.1, that we can handle an infinite E by working with its net of finite subsets.

The second change is that we need to replace our use of Parrott's theorem with a generalisation that tells us how small the k'th singular value of a completion can be. Such a generalisation has been developed, by Arsene, Constantinescu and Gheondea [ACG87, CG89, CG92]. Their result [CG92, theorem 1.1] shows that

$$\inf_D s_k \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \max(s_k(R), s_k(S))$$

where R and S are as in lemma 1.4.1, closely analogous to the Parrott's theorem result.

As in chapter 1, the operators R and S are unitarily equivalent to multiplication of $H(K_E)$ by f and multiplication of $H(K_E^{(t)})$ by f, respectively, so again the lower limit on the k'th singular value of any extension of f is $\max(s_k(f, K), s_k(f, K^{(t)}))$. Hence preservation of k'th singular value will only be possible if

$$s_k(f, K^{(t)}) \le s_k(f, K)$$

which by lemma 7.2.1 is equivalent to

$$\inf_{r \ge 0}(r : \kappa^-((r^2 - f \otimes f^*)K_E^{(t)}) \le k) \le \inf_{r \ge 0}(r : \kappa^-((r^2 - f \otimes f^*)K_E) \le k).$$

This is the point in the argument where we need conditions on the kernel K for this inequality to hold. In chapter 1 we assumed that K is completely non-zero and so the kernel on the left is obtained from that on the right by Schur multiplication by $K_E^{(t)}/K_E$. We then argued that the inequality holds if $K_E^{(t)}/K_E \geq 0$, since the Schur product of two positive kernels is positive. However now we have kernels with non-zero negativity and it is easily checked that Schur multiplication by a positive kernel *can* increase negativity.

We therefore need to place stronger conditions on the kernel K for the inequality to hold. The simplest extra condition that we can apply is that $K^{(t)}$ is not only a positive Schur multiple of K but is a *rank 1* positive Schur multiple of K. Schur multiplication by a rank 1 positive cannot increase negativity, so with this extra condition the inequality will hold. We know from section 3.2 that the Blaschke kernels satisfy this stronger condition, so we will assume that K is a Blaschke kernel.

There is one final difficulty that arises: in Arsene, Constantinescu and Gheondea's generalisation of Parrott's theorem, the infimum on the left hand side is *not* always attainable when $k > 0$. This means that the most we can say is that, if E is finite and K is a Blaschke kernel, then there exist extensions of f that are arbitrarily close to preserving s_k. The possibility of not being able to extend f without increasing s_k really does happen in the AAK theorem when $k > 0$, as shown by the following example. Take K to be the Szegö kernel, $k = 1$, $E = \{-\frac{1}{2}, \frac{1}{2}\}$, $f(-\frac{1}{2}) = -2$, $f(\frac{1}{2}) = 2$ and $t = 0$. Direct calculation for this case shows that $s_k(f) = 1$ but that $s_k(f \text{ extended by } 0 \mapsto z) > 1$ for all $z \in \mathbb{C}$. In fact as $z \to \infty$

$$s_k(f \text{ extended by } 0 \mapsto z) \to 1 \text{ from above}$$

so by taking $|z|$ large enough we can come arbitrarily close to preserving s_k but we cannot actually attain the value 1. The reason here is that the 'hidden' function we are trying to build, $1/z$, has a pole at t.

In a sense the new value wanted in our example is $z = \infty$ but allowing infinite values would cause many difficulties in our analysis. The breakdown of attainability when $k > 0$ therefore puts a barrier in the way of our one-point extension approach to proving the AAK theorem. The following lemma provides our escape route when this barrier arises, by showing that we can instead 'factor out' the bad point t by multiplication by the Blaschke factor b_t. At the same time it removes the first difficulty we met—that of E having to be finite.

LEMMA 7.4.1 *Let $K : X \times X \to \mathbb{C}$ be a Blaschke kernel, $f : E \to \mathbb{C}$ be a given function on a proper subset E of X and t be any point in $X \setminus E$. Then either there exists an extension of f to $E \cup \{t\}$ that preserves s_k or else $s_{k-1}(fb_t|E) = s_k(f)$, or possibly both.*

Proof: Assume, without loss of generality, that $s_k(f) = 1$. Then for any finite subset F of E we have $s_k(f|F) \leq s_k(f) = 1$, so by the arguments above there exists an extension of $f|F$ to the new point t having $s_k < r$, where r can be made arbitrarily close to 1.

Let us now choose a net (F_α) of finite subsets of E such that $F_\alpha \to E$ (i.e. any given $e \in E$ is in F_α for all sufficiently large α) and a corresponding net of bounds (r_α) such that $r_\alpha \to 1$ from above. (If E is infinite then we can take (F_α) to be the net of finite subsets of E; otherwise we can use $\mathbb{N}$ as the index set and take $F_\alpha = E$ for all α.) Then for each α we can find an extension $f_\alpha : F_\alpha \cup \{t\} \to \mathbb{C}$ of $f|F_\alpha$ such that $s_k(f_\alpha) < r_\alpha$ and so, by lemma 7.2.1, the kernel $L_\alpha = (r_\alpha^2 - f_\alpha \otimes f_\alpha^*)K_{F_\alpha \cup \{t\}}$ satisfies $\kappa^-(L_\alpha) \leq k$.

The net $(f_\alpha(t))$ of function values at t must have a cluster point $z \in \mathbb{C} \cup \{\infty\}$ and, by taking a subnet if necessary, we can assume $f_\alpha(t) \to z$. The two cases now arise from whether $z < \infty$ or $z = \infty$:

$z < \infty$ As α increases, the zero extension of L_α to $E \cup \{t\}$ converges weakly, i.e. pointwise, to $(1-\phi\otimes\phi^*)K_{E\cup\{t\}}$, where ϕ is the extension of f to t obtained by assigning $\phi(t)=z$. Therefore $\kappa^-((1-\phi\otimes\phi^*)K_{E\cup\{t\}})\le k$, since neither zero-extending a kernel nor passing to a weak limit of a net of kernels can increase negativity. By lemma 7.2.1 this shows that ϕ is an extension of f that preserves s_k.

$z=\infty$ $L_\alpha(t,t) = (r_\alpha^2 - |f_\alpha(t)|^2)K(t,t)$ is negative for large enough α, so (by the same arguments as used in lemma 1.5.2) the Schur complement $L_\alpha^{(t)}$ then has negativity $\le k-1$. Now direct calculation shows that

$$\begin{aligned}
& L_\alpha^{(t)}(x,y) \\
= \;& (r_\alpha^2 - f(x)f(y)^*)K(x,y) \\
& -\frac{(r_\alpha^2 - f(x)f_\alpha(t)^*)K(x,t)(r_\alpha^2 - f_\alpha(t)f(y)^*)K(t,y)}{(r_\alpha^2 - f_\alpha(t)f_\alpha(t)^*)K(t,t)} \\
\to \;& (1-f(x)f(y)^*)K(x,y) + f(x)K(x,t)K(t,y)f(y)^*/K(t,t) \\
& \text{since } r_\alpha\to 1 \text{ and } f_\alpha(t)\to\infty \\
= \;& K(x,y) - f(x)\left(K(x,y) - \frac{K(x,t)K(t,y)}{K(t,t)}\right)f(y)^* \\
= \;& K(x,y) - f(x)K^{(t)}(x,y)f(y)^* \\
= \;& K(x,y) - (f(x)b_t(x))K(x,y)(f(y)b_t(y))^*
\end{aligned}$$

since K is a Blaschke kernel.

Therefore $(1-(fb_t|E)\otimes(fb_t|E)^*)K_E$ is the weak limit of a net of kernels that each have negativity $\le k-1$, so it also has negativity $\le k-1$. Hence $s_{k-1}(fb_t|E)\le s_k(f)$ (by lemma 7.2.1) and since we also have $s_{k-1}(fb_t|E)\ge s_k(f)$ (by lemma 7.2.3) the two must be equal, as claimed.

■

7.5 Proof of the AAK Theorem

We finally have all the results we need to prove the AAK theorem in the form we stated in section 7.1 and all that remains is to thread them together. Most of the results we need apply to Blaschke kernels, so we can in fact prove the following small generalisation of the AAK theorem.

THEOREM 7.5.1 *Let $K : X \times X$ be a Blaschke kernel, $f : E \to \mathbb{C}$ be any given function on a non-empty subset E of X and k be a non-negative integer. Then there exists a function $\psi : X \to \mathbb{C}$ and a generalised Blaschke product b with at most k factors such that $\psi|E = fb|E$ and*

$$s_{k-j}(\psi) = s_k(f)$$

where j is the number of factors in b.

Proof: Well order $X \setminus E$ and then construct ψ by simply trying to extend f from E to X, one point at a time in order, without increasing s_k.

Only two issues arise. The first is that we may reach a point t to which extension is not possible without increasing s_k. By lemma 7.4.1 we can then extract the generalised Blaschke factor b_t and continue the process with $fb_t|E$ replacing f and $k-1$ replacing k. The need for this can only arise at most k times since once k reaches zero then one point extension is always possible, because K is an NP kernel.

The second issue is that we must show that s_k does not increase when we reach a limit ordinal in the ordering of $X \setminus E$. Suppose at some point in the process we have succeeded in extending (without increasing s_k) up to and including each of the points preceding some point t. For each point $a < t$ let ψ_a be the extension to $\{x : x \leq a\}$. To one-point extend to t we need to start from a function ψ on $\{x : x < t\}$. If t has an immediate predecessor, s say, in the ordering then ψ_s is such a function. But if t is a limit point in the ordering then it has no immediate predecessor, so none of the functions built so far can act as our starting function for the next step. Our functions

$\{\psi_x | x < t\}$ do define a function $\psi : \{x : x < t\} \to \mathbb{C}$, since any two agree on the intersection of their domains, *but* we need to show that $s_k(\psi) = s_k(f)$. However, this follows directly from our characterisation of s_k, lemma 7.2.1, in particular from its corollary 7.2.2, since

$$\begin{aligned} s_k(\psi) &= \sup_{F \text{ finite}} (s_k(\psi|F)) \\ &= \sup_{F \text{ finite}, x<t} (s_k(\psi_x|F)) \\ &= s_k(f). \qquad \blacksquare \end{aligned}$$

Note that we could have proved this result using Zorn's lemma, as we did with the corresponding result in chapter 1, giving a proof that would perhaps be more clearly rigorous, but probably less illuminating. The AAK theorem for the Szegö kernel now follows easily.

THEOREM 7.5.2 *Let K be the Szegö kernel and $f : E \to \mathbb{C}$ be any given function on a non-empty subset E of $\mathbb{D}$. Then for each $k = 0, 1, \ldots$*

$$s_k(f) = \min \|\phi\|_\infty$$

where the minimum is taken over all $\phi \in H^\infty$ such that $\phi|E = fb|E$ for some $b \in B_k$ or is taken to be ∞ if there is no such ϕ.

Proof: By the previous theorem there exist functions $\psi : \mathbb{D} \to \mathbb{C}$ and $b \in B_j$ such that $\psi|E = fb|E$ and $s_{k-j}(\psi) = s_k(f)$. But now, by lemma 7.3.2, there exists another Blaschke product $b' \in B_{k-j}$ such that $b'\psi \in H^\infty$ and $\|b'\psi\|_\infty = s_k(f)$. The function $\phi = b'\psi$ clearly satisfies the requirements. $\blacksquare$

Our proof of Pick's theorem in chapter 1 worked for a fairly wide range of kernels, which of course included the Szegö kernel. Can our proof of the AAK theorem be similarly widened to give a 'generalised AAK theorem' for some class of kernels? We already have a generalisation, i.e. theorem 7.5.1, but it is limited in two ways.

Firstly, it only applies to Blaschke kernels, which do not include any important kernels other than the Szegö kernel. This is perhaps a necessary restriction, given that it is a result about multiplication by Blaschke factors. This possibility is in keeping with computer calculations that I have done with many small randomly chosen kernels that are NP but not Blaschke kernels. These showed that it is in general easy to find interpolation data for which $s_k(S) > s_k(R)$ and so preservation of s_k cannot even be approached, let alone achieved. I included cases where the underlying reproducing kernel Hilbert space was the Dirichlet space and the Sobolev space considered in section 2.2, with $w_0 = w_1 = 1$. Consequently the analogue of the AAK theorem does *not* hold for either of these spaces.

The second limitation of theorem 7.5.1 is that it does not give a bounded multiplier of $H(K)$, i.e. a function ϕ for which $s_0(\phi) = s_k(f)$, since it does not include the final step of reducing k to zero by multiplying by more Blaschke factors. In theorem 7.5.2 we used lemma 7.3.2, and hence Beurling's theorem, to make this step, but for Blaschke kernels in general we do not have any analogue of Beurling's theorem. Indeed it is clear that this last step is only possible with the Szegö kernel because its domain $\mathbb{D}$ is 'complete' in the sense that it contains all the points needed to provide the required extra Blaschke factors. Even if we simply took K to be a restriction of the Szegö kernel we would have removed points, and hence Blaschke factors, that in general can be needed for this last step.

It appears therefore that theorem 7.5.1 is about as general as we can reasonably expect.

Conclusions

Reproducing kernels are a powerful and often economical tool for analysing functional Hilbert spaces and their operators. Summarising, they have enabled us to:

- develop an approach to understanding Pick's theorem which applies to any reproducing kernel Hilbert space, including vector-valued spaces
- obtain sufficient conditions for Pick's theorem to hold, i.e. for the kernel to be NP, and use these conditions to prove Pick's theorem for several specific spaces
- recognise that several of the special properties of H^2, such as those of Blaschke factors, derive from the special form of the kernel and are shared by a class of kernels, the Blaschke kernels.
- show that the optimal interpolating multiplier is unique if the kernel is a completely non-zero Blaschke kernel
- characterise the completely NP kernels, i.e. those for which Pick's theorem holds for all inflations of the space
- give an alternative proof of the Adamyan-Arov-Kreĭn theorem and show that an analogue of this result holds for all Blaschke kernels.

What further work is there to be done in this area? One obvious 'loose end' is that we have not succeeded in characterising the NP kernels. However, from the

C^*-algebra perspective developed in chapter 5 it seems quite possible that there is no neat characterisation, just as there is no neat characterisation, as far as I know, of the positive $\mathbb{C}^*$-algebra maps, only of the completely positive maps.

More promising, perhaps, is the possibility of finding a much simpler and shorter proof of the Adamyan-Arov-Kreĭn theorem. It is apparent, from our work in chapter 7, that the AAK theorem implies that the following result must hold:

> Let K be the Szegö kernel on $\mathbb{D}$, $f : E \to \mathbb{C}$ be any given function on a non-empty subset E of $\mathbb{D}$ and k be a positive integer. Then there exists a point $t \in \mathbb{D}$ such that $s_k(fb_t|E) = s_{k-1}(f)$.

Conversely, repeated application of this result quickly reduces the AAK theorem to Pick's theorem. It seems hopeful that such a simple result could be proved directly using reproducing kernel techniques, but this possible approach only became apparent at the end of my research, and I have not yet been able to study to it.

Finally, there is no doubt that reproducing kernels are very promising for other types of operator problems on functional Hilbert spaces. We have worked only on multiplication operators, but composition operators on functional spaces are also neatly characterised by their action on the reproducing functions; see [Cow92] for example. Indeed, de Branges used this in his proof of the Bieberbach conjecture. It seems likely that reproducing kernels have much more to tell us yet.

Nomenclature

b_t	Blaschke factor associated with point t
$\mathbb{C}$	complex plane
$\mathrm{card}(X)$	cardinality of set X
$\mathbb{D}$	open unit disc in $\mathbb{C}$
$\mathcal{D}_W$	$(I - W^*W)^{1/2}$, the defect operator of W
E	domain set of a multiplier;
	in chapter 6 a separable, quasi-complete topological vector space
$\bar{E}'$	conjugate dual of topological vector space E
$H(K)$	reproducing kernel Hilbert or Kreĭn space with kernel K
$H(K, E)$	closed span of E-columns of K in $H(K)$
H^∞	Banach space of bounded analytic functions on $\mathbb{D}$; supremum norm
H^2	Hilbert space of square-integrable analytic functions on $\mathbb{D}$, with norm
	$\|f\|^2 = \sup_{r\leq 1} \int_{r\mathbb{T}} \|f\|^2 d\mu$ (μ = normalised Lebesgue measure on $r\mathbb{T}$)
I_H	identity operator on H
κ^-, κ^+	negativity and positivity of an operator or kernel
K	positive (usually positive definite) kernel : $X \times X \to \mathbb{C}$ or $\mathcal{L}(H)$, or a
	Schwartz kernel relative to a topological vector space
$K^{(t)}$	Schur complement of $K(t,t)$ in K
K_E	restriction of kernel K to a subset E
K_y	y-column of kernel K
$\mathcal{L}(H)$	Banach space of continuous linear operator on Hilbert space H
M_f	operator of multiplication of $H(K_E)$ by f, E being f's domain
$M_{f,L}$	operator of multiplication of $H(L)$ by f
$\tilde{M}_f$	operator unitarily equivalent to M_f
$M(K)$	Banach space of multipliers of $H(K)$
$\mathcal{M}(K)$	as $M(K)$ but the multiplication operators themselves
$\|\phi\|_{M(K)}$	operator norm of multiplication of $H(K)$ by ϕ
$\max(W)$	space of maximising vectors of operator W
P_N	orthogonal projection onto subspace N
$\Re(z)$	real part of z
$\mathrm{ran}(W)$	range of operator W
$s_k(\cdot)$	k'th singular value of an operator or function
$\overline{\mathrm{span}}(S)$	closure of the span of the vectors in S
$\mathbb{T}$	unit circle in $\mathbb{C}$
$\geq 0, > 0$	positive, positive-definite, respectively
$\circ$	Schur product of two matrices

Bibliography

[AAK71] V.M. Adamyan, D.Z. Arov, and M.G. Kreĭn. Analytic properties of Schmidt pairs for a Hankel operator and the generalized Schur-Takagi problem. *Math. USSR Sbornik*, 15:31–73, 1971.

[ACG87] G. Arsene, T. Constantinescu, and A. Gheondea. Lifting of operators and prescribed numbers of negative squares. *Mich. Math. J.*, 34:201–216, 1987.

[AD] D. Alpay and H. Dym. On reproducing kernel spaces, the Schur algorithm, and interpolation in a general class of domains. Preprint. To appear in Sapporo workshop OT series.

[Agl88] J. Agler. Interpolation. Lecture at AMS Summer School, Durham, New Hampshire, 1988.

[Agl90] J. Agler. Nevanlinna-Pick interpolation on Sobolev space. *Proc. Amer. Math. Soc.*, 108:341–351, 1990.

[Aro50] N Aronszajn. Theory of reproducing kernels. *Trans. Amer. Math. Soc.*, 68:337–404, 1950.

[Axl88] S. Axler. Bergman spaces and their operators. In J.B. Conway and B.B. Morrel, editors, *Surveys of some recent results in operator theory, Volume 1*. Longman, 1988.

[BB84] F. Beatrous and J. Burbea. Positive-definiteness and its applications to interpolation problems for holomorphic functions. *Trans. Amer. Math. Soc.*, 284:247–270, 1984.

[Beu49] A. Beurling. On two problems concerning linear transformations in Hilbert space. *Acta Math.*, 81:239–255, 1949.

[BGR90] J.A. Ball, I. Gohberg, and L. Rodman. *Interpolation of rational matrix functions*, volume 45 of *Operator Theory : Advances and Applications*. Birkhauser-Verlag, 1990.

[BM84] J. Burbea and P. Masani. *Banach and Hilbert spaces of vector-valued functions*, volume 90 of *Pitman Research Notes in Mathematics*. Pitman, 1984.

[CG89] T. Constantinescu and A. Gheondea. Minimal signature in lifting of operators I. *J. Operator Theory*, 22:345–367, 1989.

[CG92] T. Constantinescu and A. Gheondea. Minimal signature in lifting of operators II. *J. Functional Analysis*, 103:317–351, 1992.

[Cow92] C. Cowen. Transferring subnormality of adjoint compositions operators. *Integral Equations and Operator Theory*, 15:167–171, 1992.

[Den29] A. Denjoy. Sur une classe de fonctions analytiques. *C. R. Acad. Sci. Paris*, 188:140–142, 1084–1086, 1929.

[Hal74] P.R. Halmos. *A Hilbert space problem book*. Springer-Verlag, 1974.

[HJ85] R.A. Horn and C.A. Johnson. *Matrix analysis*. Cambridge University Press, 1985.

[McC] S. McCullough. The local De Branges-Rovnyak construction and complete Nevanlinna-Pick kernels. Preprint. University of Florida.

[McC92] S. McCullough. Caratheodory interpolation kernels. *Integral Equations and Operator Theory*, 15:43–71, 1992.

[Nev19] R. Nevanlinna. Uber beschränkte Funktionen, die in gegeben Punkten vorgeschreibene Werte annehmen. *Ann. Acad. Sci. Fenn. Ser. B*, 13, 1919.

[Par78] S. Parrott. On the quotient norm and the Sz.Nagy-Foias lifting theorem. *J. Functional Analysis*, 30:311–328, 1978.

[Pau86] V.I. Paulsen. *Completely bounded maps and dilations.* Longman, 1986.

[Pic16] G. Pick. Uber der Beschränkungen analytische Funktionen, welche durch vorgegebene Funktionswerte bewirkte werden. *Math. Ann.*, 77:7–23, 1916.

[Qui93] P.P. Quiggin. For which reproducing kernel Hilbert spaces is Pick's theorem true? *Integral Equations and Operator Theory*, 16:244–266, 1993.

[Sai88] S. Saitoh. *Theory of reproducing kernels and its applications*, volume 189 of *Pitman Research Notes in Mathematics.* Longman, 1988.

[Sar67] D. Sarason. Generalized interpolation in H^∞. *Trans. Amer. Math. Soc.*, 127:179–203, 1967.

[Sch64] L. Schwartz. Sous-espaces Hilbertiens d'espaces vectoriels topologiques et noyaux associés (noyaux reproduisants). *J. D'Analyse Mathematique*, 13:115–256, 1964.

[Sor75] P. Sorjonen. Pontryagin spaces with a reproducing kernel. *Ann. Acad. Sci. Fennicae, A.I.*, 594:1–30, 1975.

[Sza86] S.H. Szafraniec. On bounded holomorphic interpolation in several variables. *Monatsh. Math.*, 101:59–66, 1986.

[Wal56] J.L. Walsh. *Interpolation and approximation by rational functions in the complex domain*, volume 20 of *Amer. Math. Soc. Colloq. Publ.* American Math. Soc., 1956.

[You86] N. Young. The Nevanlinna-Pick problem for matrix-valued functions. *J. Operator Theory*, 15:239–265, 1986.

[You88] N. Young. *An introduction to Hilbert space*. Cambridge University Press, 1988.