INTERPOLATION IN HARDY SPACES, WITH APPLICATIONS

JONATHAN R. PARTINGTON

Abstract. This tutorial paper discusses a number of classical interpolation results from the theory of Hardy spaces, and stresses the role played by reproducing kernels. Next, the reproducing kernel thesis is presented in several different contexts. Some basics of the theory of linear systems are described, leading into two applications of the abstract results: the first is in the recovery of functions from boundary values, and the second is in the theory of semigroups, admissibility, and controllability.

1. Introduction

This paper will present some classical results from the theory of interpolation in the Hardy spaces of analytic functions, together with some more recent extensions and applications with which I have myself been involved. To some extent the account given here may be seen as complementary to [8], which surveyed some related results, but considered them more from an approximation point of view. The material of this paper was presented as a mini-course in Le Touquet in September 2010, and I am grateful to the organizers, and in particular Catalin Badea, for inviting me to participate in this way.

2. Hardy spaces, classical interpolation results

We begin with some standard results on Hardy spaces, which may be found in many places, for example the books [11, 13, 17, 30]. For $1 \leq p < \infty$ the Hardy space $H^p$ on the unit disc $\mathbb{D}$ is the Banach space of analytic functions in $\mathbb{D}$ such that the norm

$$\|f\|_p = \left( \sup_{0 < r < 1} \frac{1}{2\pi} \int_0^{2\pi} |f(re^{it})|^p \, dt \right)^{1/p} < \infty.$$ 

Likewise, $H^\infty$ is the space of bounded analytic functions in $\mathbb{D}$, and

$$\|f\|_\infty = \sup_{z \in \mathbb{D}} |f(z)|.$$ 

Now $H^p$ functions can be given boundary values on the circle $\mathbb{T}$ (almost everywhere) by setting

$$\tilde{f}(e^{it}) = \lim_{r \to 1^-} f(re^{it}),$$
and indeed
\[ \| \tilde{f} \|_{L^p(T)} = \| f \|_{H^p}. \]
We shall usually write \( f \) instead of \( \tilde{f} \). Note that, if
\[
f(z) = \sum_{n=0}^{\infty} a_n z^n, \quad \text{then} \quad \| f \|_2 = \left( \sum_{n=0}^{\infty} |a_n|^2 \right)^{1/2}.
\]
Indeed \( H^2 \) is a Hilbert space and \( (e_n)_{n=0}^{\infty} \) forms an orthonormal basis, where \( e_n(z) = z^n \).

In order to discuss interpolation, one needs to consider the sets on which \( H^p \) functions vanish. The zeroes of a function \( f \neq 0 \) form a finite or countable set \( (z_n) \) in \( \mathbb{D} \), satisfying the Blaschke condition
\[
\sum_{n} (1 - |z_n|) < \infty.
\]
Given a Blaschke sequence (i.e., satisfying the Blaschke condition), we can find a function in \( H^\infty \) (and hence all \( H^p \)) that has precisely these zeroes (repeated zeroes are allowed). Let
\[
B(z) = z^p \prod_{n: z_n \neq 0} \frac{|z_n|}{z_n} \frac{z_n - z}{1 - z_n z}.
\]
This is called a Blaschke product, and converges pointwise (locally uniformly) to an analytic function. Note that \( |B(e^{it})| = 1 \) a.e., i.e., \( B \) is an inner function. There are other inner functions without zeroes (singular inner functions), for example
\[
\Theta(z) = \exp \left( \frac{z - 1}{z + 1} \right).
\]
The discussion of zeroes on the boundary is more delicate as the functions are now only defined a.e. In fact, if \( f \) is in an \( H^p \) space, then
\[
\int_0^{2\pi} \log |f(e^{it})| \, dt > -\infty,
\]
so \( f \neq 0 \) a.e. on \( T \).

Inner functions arise in another way. Let \( S: H^2 \to H^2 \) be the shift operator, \( (Sf)(z) = zf(z) \). Its closed invariant subspaces, i.e., subspaces \( \mathcal{M} \) such that \( SM \subseteq \mathcal{M} \), are classified as follows.

**Theorem 2.1** (Beurling). Apart from \( \{0\} \) and \( H^2 \), the nontrivial shift-subspaces of \( H^2 \) have the form \( \mathcal{M} = \Theta H^2 = \{ \Theta f : f \in H^2 \} \), where \( \Theta \) is inner.

For example, the set of functions vanishing at a sequence \( (z_n) \) is \( BH^2 \) where \( B \) is the Blaschke product with those points as zeroes.

The cyclic functions, i.e., those that lie in no proper \( S \)-invariant subspace, are called outer functions: for example, the functions \( z \mapsto z - 1 \) and \( z - 2 \)
are outer. One consequence is that every $H^p$ function is a product of an inner function and an outer function.

It is well known how $B(z) = z^n$ maps $T$ to itself. In fact, finite Blaschke products show a similar behaviour. If

$$B(z) = \prod_{k=1}^{n} \frac{z - z_k}{1 - \overline{z_k}z},$$

is a Blaschke product with $n$ zeroes including multiplicity, then it maps $T \to T$ as $n$-to-1. Moreover, if $B(e^{i\theta}) = e^{i\phi}$, then $\frac{d\phi}{d\theta} > 0$. (This is easy algebra.)

In fact, the zeroes of $B'$ lie in the convex hull of 0 and the zeroes of $B$ (this was proved by Walsh [38], for finite products, and Cassier–Chalendar [6] for infinite products).

**Nevanlinna–Pick interpolation**

Suppose $z_1, \ldots, z_n$ are distinct points in $\mathbb{D}$ and $w_1, \ldots, w_n$ are also points in $\mathbb{D}$.

- Can we find an interpolating function $\phi \in H^\infty$ of norm at most 1 with $\phi(z_k) = w_k$ for each $k$?
- What is the minimal $H^\infty$ norm of an interpolating function $\phi$?

The following results (see for example [32]) gives an answer to the first question.

**Theorem 2.2** (Pick). Consider the $n \times n$ matrix $Q$ such that

$$Q_{j,k} = \frac{1 - w_j \overline{w_k}}{1 - z_j \overline{z_k}} \quad (j, k = 1, \ldots, n).$$

Then we can find such an interpolating function $\phi$ if and only if $Q \succeq 0$ (i.e., positive semi-definite). When $Q \succeq 0$, there is an interpolating $\phi$, which is a Blaschke product of degree at most $n$.

To see how $Q$ arises, we introduce the Cauchy–Szegő (reproducing) kernels, $k_a$ for $a \in \mathbb{D}$, where

$$k_a(z) = 1/(1 - \overline{a}z),$$

satisfying

$$f(a) = \langle f, k_a \rangle, \quad (f \in H^2).$$

The functions $k_{z_1}, \ldots, k_{z_n}$ form a basis for a finite-dimensional space $K_B$, the orthogonal complement of $BH^2$, where $B$ is the Blaschke product with zeroes $z_1, \ldots, z_n$.

Now look at multiplication operators $M_\phi : H^2 \to H^2$, given by $M_\phi f = \phi f$ for $f \in H^2$. Note that $\langle f, M_\phi^* k_{z_j} \rangle = \langle M_\phi f, k_{z_j} \rangle = \phi(z_j)f(z_j)$, and so $M_\phi^* k_{z_j} = \overline{\phi(z_j)}k_{z_j}$ for each $j$. 
It can be shown that $\phi$ interpolates, and $M_\phi$ is a contraction, if and only if $I - TT^* \geq 0$, where $T = M_\phi^*(K_B)$, which satisfies $Tk_j = \overline{w}_j z_j$ for each $j$. But then $\langle (I - TT^*)k_j, k_z \rangle$ is just the $Q_{j,k}$ in the Pick matrix.

We now consider the minimal-norm interpolant. Let $p$ be any function in $H^\infty$ such that $p(z_j) = w_j$ for all $j$; for example, a polynomial. Then the set of all interpolating $\phi$ is $p + BH^\infty$, where $B$ is the Blaschke product with zeroes $z_1, \ldots, z_n$.

Now $\|p + Bg\|_\infty = \|B^{-1}p + g\|_{L^\infty(T)}$.

We refer to [31] for more details.

**Boundary interpolation**

We now look at interpolation from boundary values.

**Theorem 2.3** (Rudin–Carleson). If $K \subset T$ is closed, with measure 0, then any $f_0 \in C(K)$ can be extended to a function $f_1 \in A(D) = C(\overline{D}) \cap H^\infty$, with $\|f_1\|_{H^\infty} = \|f_0\|_{C(K)}$.

Here, $A(D)$ is the disc algebra, the closure of the polynomials in $H^\infty$.

Indeed, the zero sets (sets of uniqueness) for $A(D)$ are the closed subsets $Z \subset \overline{D}$ such that (i) $Z \cap T$ has measure 0; and (ii) $Z \cap D$ is at most a countable sequence $(z_n)$, satisfying the Blaschke condition $\sum (1 - |z_n|) < \infty$.

**Carleman formulae**

A good source for this material is the book [1].

For $f \in H^p$, we know that $f$ is uniquely determined by its values on any set $S$ of positive measure; that is, we have a restriction mapping $R : H^p \to L^p(S)$, which is injective, but not surjective, although it has dense range (for $1 \leq p < \infty$). The problem is: how do we invert $R$?

If we know $f$ on all of $T$ we can find it in $D$ by harmonic extension (the Dirichlet problem whose solution involves the Poisson kernel) or by complex analysis (Cauchy’s formula):

$$f(w) = \frac{1}{2\pi i} \int_T \frac{f(z) \, dz}{z - w}.$$  

Let us rewrite this as

$$f(w) = \frac{1}{2\pi} \int_0^{2\pi} f(e^{it}) \, dt \frac{1}{1 - we^{-it}}.$$  

We explain the Goluzin–Krylov technique for a closed subset $K$ of positive measure – an arc is the most important example.
We begin by constructing a function $H \in H^\infty$ (invertible) such that
\[
|H(e^{it})| = \exp \chi_K(e^{it})
\]
a.e. on $T$. Here $\chi$ denotes characteristic function (1 on $K$, 0 off $K$). Indeed, we may take
\[
H(z) = \exp \left\{ \frac{1}{2\pi} \int_0^{2\pi} \frac{e^{it} + z}{e^{it} - z} \chi_K(e^{it}) \, dt \right\}.
\]
Now $|H| > 1$ in the open disc $D$, and thus, for fixed $w \in D$, \( (H(e^{it})/H(w))^n \) becomes very small on $T \setminus K$ as $n$ increases. $H$ is called a quenching function.

Given the values on $K$ of the unknown $f$ define
\[
f_n(w) = \frac{1}{2\pi} \int_K f(e^{it}) \left( \frac{H(e^{it})}{H(w)} \right)^n \frac{dt}{1 - we^{-it}}.
\]
If we took this integral on the whole circle we would get $f(w)$ exactly (but we do not know $f$ except on $K$).

The quenching property can be used to prove easily that $f_n(w) \to f(w)$ pointwise as $n \to \infty$ for any $f \in H^1$ (so for $f$ in any $H^p$ space for $p \geq 1$).

But more is true as shown in [34].

**Theorem 2.4** (Patil). For $f \in H^p$, and $1 < p < \infty$, we have $\|f_n - f\|_p \to 0$.

The proof involves Toeplitz operators. Indeed
\[
f_n = \frac{e^{2n}}{e^{2n} - 1} (f - T_n f),
\]
where $T_n$ are operators (inverses of Toeplitz operators) such that $\|T_n f\| \to 0$ for all $f \in H^p$. Later we shall look at what can be done if we only know $f$ approximately on $K$ (as in many applications).

For recovery from values in $D$, one can recover $f \in H^p$ from $(f(z_n))_n$, provided that $\sum (1 - |z_n|) = \infty$, i.e., $(z_n)$ is a non-Blaschke sequence. For example, Totik [37] provided such a formula.

**Carleson interpolation**

Suppose that $(z_j)_j$ is a sequence of distinct points in $D$. Then we have a contractive mapping
\[
R : H^\infty \to \ell^\infty, \quad f \mapsto (f(z_j))_j.
\]
If $R$ is surjective, then we call $(z_j)_j$ a Carleson sequence.

By the open mapping theorem, if $R$ is surjective then, for some $M > 0$ it holds that all $(a_j) \in \ell^\infty$ there is an $f \in H^\infty$ with
\[
f(z_j) = a_j \quad \text{for all } j
\]
and
\[
\|f\|_\infty \leq M\|(a_j)\|_\infty.
\]
One easy observation is that a Carleson sequence $(z_j)$ must satisfy the Blaschke condition (1). For there will be an $f \in H^\infty$ such that $f(z_1) = 1$.
and \( f(z_j) = 0 \) for all \( j > 1 \). So there is a non-zero \( H^\infty \) function vanishing at \((z_j)_{j \geq 2}\), and hence the Blaschke condition holds.

We can take this further. For each \( k = 1, 2, \ldots \) we can find \( f_k \in H^\infty \) such that \( \|f_k\|_\infty \leq M \), with \( f_k(z_k) = 1 \) and \( f_k(z_j) = 0 \) for all \( j \neq k \). Then write \( f_k = B_k g_k \), where \( B_k \) is a Blaschke product with zeroes \((z_j)_{j \neq k}\). Now

\[
\prod_{j \neq k} \frac{z_j - z_k}{1 - \overline{z_k} z_j} = |B_k(z_k)| \geq \frac{1}{M} > 0.
\]

Our conclusion is that the condition that

\[
\delta := \inf_{k \geq 1} \prod_{j \neq k} \left| \frac{z_j - z_k}{1 - \overline{z_k} z_j} \right| > 0
\]

is necessary for an Carleson sequence.

This is much stronger than just being a Blaschke sequence. It implies that the points are well-separated from each other (e.g. there is a minimum hyperbolic distance between any pair). Note that finite unions of Carleson sequences do not need to be Carleson, whereas finite unions of Blaschke sequences are still Blaschke. The big theorem here was first published in [5]:

**Theorem 2.5 (Carleson).** Condition (C) is necessary and sufficient for \((z_j)\) to be an Carleson sequence.

Thus in this case \( R : H^\infty \to \ell^\infty \) is surjective. There are several proofs. Some (e.g. that of P. Jones [24]) construct an explicit interpolating function.

The \( H^2 \) version is also important, but now we have to weight the sequence. Shapiro and Shields [36] show that Condition (5) is necessary and sufficient for the following problem: given \((v_j) \in \ell^2\), find \( f \in H^2 \) such that

\[
f(z_j)(1 - |z_j|^2)^{1/2} = v_j
\]

for all \( j \).

This is more subtle, since for an arbitrary sequence \((z_j)_j\) there might not even be a mapping from \( H^2 \) to \( \ell^2 \) given by \( T f = (f(z_j)(1 - |z_j|^2)^{1/2})_j \). To understand this better, we look at Carleson measures.

A **Carleson measure** on \( \mathbb{D} \) is a Borel measure \( \mu \) such that the natural mapping \( H^2 \to L^2(\mathbb{D}, \mu) \) is well-defined and bounded. That is, for some \( C > 0 \),

\[
\int_{\mathbb{D}} |f(z)|^2 \, d\mu(z) \leq C\|f\|^2_2
\]

for all \( f \in H^2 \).

Thus a sequence \((z_j)\) has the property that there is a bounded operator \( T : H^2 \to \ell^2 \), defined by \( T f = (f(z_j)(1 - |z_j|^2)^{1/2})_j \), if and only if the
measure $\mu := \sum_j (1 - |z_j|^2) \delta_{z_j}$ is a Carleson measure. Here $\delta_j$ is a Dirac point mass at $z_j$.

How do we recognise a Carleson measure? The answer is that $\mu$ is a Carleson measure if and only if there is a constant $M > 0$ such that $\mu(Q_{h,\alpha}) \leq Mh$ for all $0 < h < 1$ and $\alpha \in [0, 2\pi]$, where

$$Q_{h,\alpha} = \{ z = re^{i\theta} : 1 - h < r < 1, |\theta - \alpha| < h \},$$

a curvilinear rectangle. We shall see another way of testing these later. In fact, if $(z_j)_j$ is a Carleson sequence then $\mu := \sum_j (1 - |z_j|^2) \delta_{z_j}$ is a Carleson measure. Conversely, if $\mu$ is a Carleson measure, then $(z_j)_j$ is a finite union of Carleson sequences.

**Interpolation by inner functions**

Recall that an inner function $\Theta \in H^\infty$, satisfies $|\Theta(e^{it})| = 1$ a.e., i.e., $\Theta$ maps $\mathbb{D}$ to $\mathbb{D}$ and $T$ to $T$. We now look at values on the circle.

Recall that for a finite Blaschke product, $B$ is an $n$-to-1 mapping from $T$ to $T$. Suppose now that $\Theta$ is inner, and extends analytically across $T$ except at $z = 1$. Suppose also that we know the sequence $(t_n)$ on $T$ where $\Theta(t_n) = 1$. This must accumulate at 1: let us assume first that it does so on both sides. What can we say about $\Theta$?

Let us transform to the upper half-plane by $\psi(z) = i(1 + z)/(1 - z)$. Note that $\psi(1) = \infty$. We are now interested in $F := \psi \circ \Theta \circ \psi^{-1}$. Then $F$ is meromorphic on $\mathbb{C}$ with real poles $(b_n)$ accumulating at $\pm\infty$. It maps $\mathbb{C}^+ \to \mathbb{C}^+$ and $\mathbb{C}^- \to \mathbb{C}^-$. Such functions are called strongly real.

To simplify explanations we shall suppose $F(0) \neq 0, \infty$. The Hermite–Biehler (Krein) theorem [27] says that the zeroes $(a_n)$ and poles $(b_n)$ are interlaced, i.e., $b_n < a_n < b_{n+1}$ for all $n$ and

$$F(z) = c \prod_{n \in \mathbb{Z}} \frac{1 - z/a_n}{1 - z/b_n},$$

where $c > 0$ unless some $a_nb_n < 0$, when $c < 0$. We can choose the $(a_n)$ how we like, if interlaced with the $(b_n)$. As given in the literature, there are other cases not always considered, e.g. accumulation at $-\infty$ but not $+\infty$, but they can be handled.

Thus, if there is one limit point on $T$, the set $\Theta^{-1}(1)$ does not determine $\Theta$. However, the sets $\Theta^{-1}(1)$ and $\Theta^{-1}(-1)$ together do tell us $\Theta$ (to within composition by a Möbius map fixing 1 and $-1$).

Recent extensions have been given in [7]. We understand the case of finitely-many singularities (so $F$ has some finite essential singularities). Certain unexpected non-uniqueness cases appear. For example, if $a_n \searrow 1$ at $-\infty$ and $a_n \nearrow +\infty$ at $+\infty$, then the function can be

$$F(z) = c \prod_{n \in \mathbb{Z}} \frac{1 - z/a_n}{1 - z/b_n},$$
but it can also be
\[ F(z) = c(z - 1) \prod_{n \in \mathbb{Z}} \frac{1 - z/a_n}{1 - z/b_n}, \]
with \( c > 0 \) (and that’s all).

3. Reproducing kernels.

We have seen the Cauchy–Szegö (reproducing) kernels, \( k_a \) for \( a \in \mathbb{D} \), defined in (3) and satisfying (4).

Clearly, they play a key role in interpolation. Let us write \( K \) for the closed linear span of a sequence \( (k_{z_n})_n \). Then \( K^\perp \) is the space of all functions \( f \) such that
\[
f(z_n) = \langle f, k_{z_n} \rangle = 0 \quad \text{for all } n.\]

Suppose we take a Blaschke sequence \( (z_n)_n \). Then \( K = (BH^2)^\perp \), where \( B \) is the Blaschke product with zeroes \( (z_n)_n \). If \( (z_n)_n \) is non-Blaschke, then \( K = H^2 \), since \( K^\perp = \{0\} \).

Now reproducing kernels give another way of characterising Carleson sequences. A Riesz basic sequence in a Hilbert space is a sequence \( (g_n)_n \) such that for some \( A, B > 0 \),
\[
A \sum |a_n|^2 \leq \left\| \sum a_n g_n \right\|^2 \leq B \sum |a_n|^2
\]
for all scalar sequences \( (a_n)_n \). If the closed linear span is \( H \) then it is a Riesz basis. Equivalently, \( (g_n)_n \) is a Riesz basis if \( (Tg_n)_n \) is an orthonormal basis (i.e., \( A = B = 1 \) in (6)) for some linear isomorphism \( T : H \to H \).

In fact the sequence \( (z_n)_n \) is a Carleson sequence if and only if the normalized kernels
\[
e_n := (1 - |z_n|^2)^{1/2}k_{z_n}
\]
form a Riesz basis for \( K \). This gives yet another approach to constructing interpolating functions.

The Reproducing Kernel Thesis (RKT)

The RKT refers to a body of powerful results and some examples of it are discussed in [30, 41]. Consider the complete collection of normalized reproducing kernels \( \{h_a : a \in \mathbb{D}\} \), where
\[
h_a(z) = \frac{(1 - |a|^2)^{1/2}}{1 - az}, \quad (z \in \mathbb{D}),
\]
for \( a \in \mathbb{D} \). Clearly, if \( T : H^2 \to X \) is any operator, we have
\[
\sup_{a \in \mathbb{D}} \|Th_a\| \leq \|T\|.
\]

We have no right to expect a converse. After all, the closed absolute convex hull of the reproducing kernels does not contain \( \{f \in H^2 : \|f\| \leq \epsilon\} \) for any \( \epsilon > 0 \). For example, if \( \psi_n(f) = \langle f, e_n \rangle \), where \( e_n(z) = z^n \), then \( \|\psi_n\| = 1 \), but
\[
\sup_{a \in \mathbb{D}} |\psi_n(h_a)| = O(n^{-1/2}).
\]
That is, estimating the norm by testing with reproducing kernels can be very inefficient. However for certain classes of operator, we do have a converse.

**RKT for Toeplitz operators on** $H^2$.

Take $\phi \in L^\infty(\mathbb{T})$ and define the Toeplitz operator $T_\phi : H^2 \to H^2$ by

$$T_\phi f = P_{H^2}(\phi.f) \quad (f \in H^2),$$

where $P_{H^2} : L^2(\mathbb{T}) \to H^2(\mathbb{D})$ is the orthogonal projection. It is well known that $\|T_\phi\| = \|\phi\|_\infty$ (clearly it is $\leq$). Now

$$\langle Th_a, h_a \rangle = \langle T_\phi k_a, k_a \rangle = \frac{\langle \phi, |k_a|^2 \rangle}{\|k_a\|^2} = \langle \phi, P_a \rangle,$$

where $P_a$ is the Poisson kernel. By the general theory of harmonic extensions, we get

$$\sup_{a \in \mathbb{D}} \|T_\phi h_a\| = \|\phi\|_\infty = \|T_\phi\|.$$

**Carleson–Vinogradov embedding theorem**

Recall that $\mu$ is a Carleson measure when the embedding $J : H^2 \to L^2(\mathbb{D}, \mu)$ is bounded; i.e., there is a constant $C > 0$ such that

$$\int_\mathbb{D} |f(z)|^2 d\mu(z) \leq C\|f\|_2^2$$

for all $f \in H^2$.

**Theorem 3.1.** The embedding $J$ is bounded if and only if

$$\sup_{a \in \mathbb{D}} \|J h_a\|_{L^2(\mu)} < \infty.$$

That is, if

$$\sup_{a \in \mathbb{D}} \int_\mathbb{D} \frac{1 - |a|^2}{|1 - az|^2} d\mu(z) < \infty.$$

Indeed in [35] it is shown that $\|J\| \leq \sqrt{2e} \sup_{a \in \mathbb{D}} \|J h_a\|$.

**Bonsall’s theorem for Hankel operators**

Write $L^2(\mathbb{T}) = H^2 \oplus \overline{H_0^2}$, where $H^2$ is spanned by $\{e^{i n t} : n \geq 0\}$, and its complement $\overline{H_0^2}$ is spanned by $\{e^{-i n t} : n \geq 1\}$. For $\phi \in L^\infty$, define the Hankel operator $\Gamma_\phi : H^2 \to \overline{H_0^2}$ by

$$\Gamma_\phi f = P_{\overline{H_0^2}}(\phi.f) \quad (f \in H^2).$$

Then $\|\Gamma_\phi\| = \text{dist}(\phi, H^\infty) = \inf_{g \in H^\infty} \|\phi - g\|_\infty$. (Nehari’s theorem.)

The following result is given in [4].

**Theorem 3.2 (Bonsall).** The RKT holds for Hankel operators; i.e., for some $C > 0$, $\|\Gamma_\phi\| \leq C \sup_{a \in \mathbb{D}} \|\Gamma_\phi h_a\|$.

Indeed, we may take $C = 4\sqrt{2e}$ [22], and a very recent preprint of Treil improves on this estimate.

Many instances of the RKT can be derived from the Carleson–Vinogradov result. For example, we mention weighted composition operators on $H^2$. 

Let \( \phi : \mathbb{D} \to \mathbb{D} \) be analytic, and define \( C_\phi : H^2 \to H^2 \) by \( C_\phi f = f \circ \phi \). By Littlewood’s subordination theorem [28], \( C_\phi \) is always bounded on \( H^2 \).

Now let \( g \in H^2 \) and define \( W_{g,\phi} \) by \( W_{g,\phi} f = g(f \circ \phi) \) for \( f \in H^2 \). Then a necessary and sufficient condition for \( W_{g,\phi} \) to map boundedly into \( H^2 \) (see [15]) is that \( \sup_{a \in \mathbb{D}} \| W_{g,\phi} h_a \| < \infty \), i.e.,
\[
\sup_{a \in \mathbb{D}} \left\| \frac{(1 - |a|^2)^{1/2}g}{1 - a\phi} \right\|_{H^2} < \infty.
\]

4. Linear systems and Hardy spaces

Discrete-time linear systems

Informally, linear systems have inputs \( u(0), u(1), u(2), \ldots \), which are often vector-valued, and outputs \( y(0), y(1), y(2), \ldots \), also vector-valued. More formally, we look at operators \( T \) defined on an input space \( U \), such as \( \ell^2(\mathbb{Z}^+, \mathbb{C}^m) \), and mapping into an output space \( Y \), such as \( \ell^2(\mathbb{Z}^+, \mathbb{C}^p) \). Here \( H \) and \( K \) are Hilbert spaces, usually finite-dimensional in practice, say \( H = \mathbb{C}^m \) and \( K = \mathbb{C}^p \).

Sometimes we work with SISO (single-input, single-output) systems, that is, \( m = p = 1 \). Physically we would expect inputs and outputs to be real, i.e., expect \( \ell^2(\mathbb{Z}^+, \mathbb{R}^m) \) to map into \( \ell^2(\mathbb{Z}^+, \mathbb{R}^p) \). Our operators may also be unbounded, and defined on a domain \( \mathcal{D}(T) \), a proper subspace of \( \ell^2(\mathbb{Z}^+, \mathbb{C}^m) \).

**Example.** Let \( y(t) = \sum_{k=0}^t u(k) \), a discrete integrator or ‘summer’. Clearly even \((1, 0, 0, \ldots) \notin \mathcal{D}(T)\).

**Causality.** We say a system is causal if whenever \( u \in \mathcal{D}(T) \) and \( u(t) = 0 \) for \( t \leq n \), then \( y(t) = 0 \) for \( t \leq n \). That is, the past cannot depend on the future. Algebraically, \( P_n TP_n u = P_n Tu \), where \( P_n u = (u(0), \ldots, u(n), 0, 0, \ldots) \).

The ‘summer’ example above is causal, and has dense domain. Causality corresponds to a lower triangular (block) matrix representation using the standard orthonormal basis of \( \ell^2 \).

**Shift invariance.** Let \( S \) be the right shift on \( \mathcal{U} = \ell^2(\mathbb{Z}^+, \mathbb{C}^m) \), so that
\[
S(u_0, u_1, u_2, \ldots) = (0, u_0, u_1, \ldots).
\]

We also use \( S \) for the analogous operator on \( \mathcal{Y} = \ell^2(\mathbb{Z}^+, \mathbb{C}^p) \). The system is shift-invariant if whenever \( u \in \mathcal{D}(T) \) and \( y = Tu \), then \( Su \in \mathcal{D}(T) \), and \( Sy = T(Su) \).

There is an automatic continuity result here: if \( T \) is shift-invariant and \( \mathcal{D}(T) = \mathcal{U} \), then \( T \) is a bounded operator, at least for \( \mathcal{U} = \ell^2(\mathbb{Z}^+, \mathbb{C}^m) \). Shift-invariant operators with \( \mathcal{D}(T) = \mathcal{U} \) will also be causal (easy).

**Transfer functions.** Shift-invariant operators have a representation as multiplication operators [12] using the theory of Hardy spaces. We’ll work
with $H^2(\mathbb{D}, \mathbb{C}^m)$, i.e., analytic vector-valued functions
\[ U(z) = \sum_{k=0}^{\infty} u(k) z^k, \]
with
\[ \|U\|_2^2 = \sum_{k=0}^{\infty} \|u(k)\|^2 < \infty. \]
These be regarded as power series in the disc $\mathbb{D}$, extending to give $L^2$ vector-valued functions on the circle $\mathbb{T}$. Likewise, $H^\infty(\mathbb{D}, L(\mathbb{C}^m, \mathbb{C}^p))$ consists of the bounded analytic matrix-valued functions in $\mathbb{D}$, extending also to $L^\infty$ functions on $\mathbb{T}$. Here
\[ \|G\|_\infty = \sup_{|z|<1} \|G(z)\|. \]
Using the obvious unitary equivalence between $\ell^2(\mathbb{Z}_+)$ and $H^2$, the shift-invariant operators $T$ become multiplications
\[ Y(z) = G(z)U(z) \quad \text{and} \quad \|T\| = \|G\|_\infty. \]
On $\ell^2(\mathbb{Z}_+)$ they look like convolutions
\[ (Tu)(t) = \sum_{k=0}^{t} h(k)u(t-k), \]
where $h(0), h(1), \ldots$ are the Fourier coefficients of an $H^\infty$ transfer function.

**Finite-dimensional systems**

Finite-dimensional systems correspond to rational (matrix-valued) functions, and are convenient from a computational point of view. They can be realized using finite state matrices:
\[
\begin{align*}
x(t+1) &= Ax(t) + Bu(t), \\
y(t) &=Cx(t) + Du(t),
\end{align*}
\]
where $x(t) \in \mathbb{C}^n$ denotes the state of the system. If $x(0) = 0$, then the associated transfer function is
\[ D + Cz(I - zA)^{-1}B. \]
Many infinite-dimensional systems can be realized using operators $A, B, C, D$, rather than matrices, as we shall see later.

**Continuous-time systems**

We now work with operators
\[ T : L^2(0, \infty; \mathbb{C}^m) \to L^2(0, \infty; \mathbb{C}^p). \]
Again, notions such as causality and shift-invariance make sense.
For shift-invariance (i.e., time-invariance) we suppose that $T$ commutes with all right shifts $S_{\tau}$. To translate this into function theory, we use the Laplace transform

$$L: L^2(0, \infty; \mathbb{C}^m) \to H^2(\mathbb{C}_+; \mathbb{C}^m),$$

$$(Lu)(s) = \int_0^\infty e^{-st}u(t) \, dt,$$

giving an isometry (up to a constant) between $L^2(0, \infty; \mathbb{C}^m)$ and a Hardy space of analytic vector-valued functions on the right half-plane $\mathbb{C}_+$ (the Paley–Wiener theorem). The norm we take is

$$\|f\|^2 = \sup_{x>0} \int_{-\infty}^{\infty} \|f(x+iy)\|^2 \, dy < \infty.$$

Note that $H^2(\mathbb{C}_+, \mathbb{C}^m)$ can be seen as a closed subspace of $L^2(i\mathbb{R}; \mathbb{C}^m)$.

Again the causal, bounded, everywhere-defined, shift-invariant operators correspond to multiplication by \textit{transfer functions}, that is, functions in the space $H^\infty(\mathbb{C}_+, \mathcal{L}(\mathbb{C}^m, \mathbb{C}^p))$. These are bounded analytic matrix-valued functions in $\mathbb{C}_+$, extending also to $L^\infty$ functions on $i\mathbb{R}$. Note that a shift by $T > 0$ in $L^2(0, \infty)$ (the time domain) corresponds to a multiplication by $e^{-st}$ on $H^2(\mathbb{C}_+)$ (the frequency domain).

We may define a continuous-time linear system in state form by the equations

$$\begin{align*}
\frac{dx(t)}{dt} &= Ax(t) + Bu(t), \\
y(t) &= Cx(t) + Du(t).
\end{align*}$$

In the finite-dimensional case, these are matrices; more generally, they are operators (more details later). The associated transfer function is

$$C(sI - A)^{-1}B + D,$$

which we suppose to be matrix-valued and analytic in some right half-plane.

Here are two examples (with zero initial conditions).

$$\frac{dy(t)}{dt} + ay(t) = u(t), \quad G(s) = 1/(s + a);$$

this is $H^\infty(\mathbb{C}_+)$ stable only if $a > 0$.

$$\frac{dy(t)}{dt} + ay(t - 1) = u(t), \quad G(s) = 1/(s + ae^{-s});$$

this is a delay system, and is $H^\infty(\mathbb{C}_+)$ stable only if $0 < a < \pi/2$.

**Graphs and invariant subspaces.**

We now deal with operators

$$T : \mathcal{D}(T) \to H^2(\mathbb{C}^p)$$
that have closed shift-invariant graphs. Closed, because in fact systems stabilizable by feedback (i.e., useful ones) will be closable. We will not specify whether we are working in discrete or continuous time (i.e., $\mathbb{D}$ or $\mathbb{C}_+$), unless it makes a difference.

Note that the graph $G(T)$ is defined to be
\[
\left\{ \begin{pmatrix} u \\ Tu \end{pmatrix} : u \in D(T) \right\} \subset H^2(\mathbb{C}^m) \times H^2(\mathbb{C}^p) = H^2(\mathbb{C}^{m+p}).
\]

We can now use the Beurling–Lax theorems on shift-invariant subspaces of $H^2(\mathbb{C}^N)$ to classify the closed shift-invariant operators, by means of their graphs.

An application to systems theory was given in [14].

**Theorem 4.1** (Georgiou-Smith). Let $T : D(T) \to H^2(\mathbb{C}^p)$ be closed, shift-invariant, with $D(T) \subseteq H^2(\mathbb{C}^m)$. Then there exist $r \leq m$, a nonsingular $M \in H^\infty(\mathcal{L}(\mathbb{C}^r, \mathbb{C}^m))$, and $N \in H^\infty(\mathcal{L}(\mathbb{C}^r, \mathbb{C}^p))$ such that
\[
G(T) = \begin{pmatrix} M \\ N \end{pmatrix} H^2(\mathbb{C}^r) = \Theta H^2(\mathbb{C}^r),
\]
where $\Theta$ is inner in the sense that $\|\Theta u\| = \|u\|$ for all $u \in H^2(\mathbb{C}^r)$.

If $M$ is allowed to be singular, then we do not obtain a graph. Now take $m = p = 1$. Then
\[
G(T) = \begin{pmatrix} M \\ N \end{pmatrix} H^2,
\]
with $M, N \in H^\infty$ and $|M(z)|^2 + |N(z)|^2 = 1$ a.e. on $T$ or $i\mathbb{R}$ (as appropriate). This means that $T$ acts as multiplication by $N/M$. The domain $\mathcal{D}(T)$ is $MH^2$, which is dense provided that $M$ is outer. Causality can be characterized in terms of inner divisors of $M$ and $N$.

**Example.** For an unstable delay system
\[
\frac{dy(t)}{dt} - y(t) = u(t-1),
\]
we have
\[
G(s) = \frac{e^{-s}}{s-1}.
\]
We take
\[
N(s) = \frac{e^{-s}}{s + \sqrt{2}}, \quad M(s) = \frac{s-1}{s + \sqrt{2}},
\]
obtaining a normalized coprime factorization and
\[
G(T) = \begin{pmatrix} M \\ N \end{pmatrix} H^2.
\]

Let us now link these ideas with some interpolation problems.
5. Recovery of functions from boundary values

Take $I \subset \mathbb{T}$ of positive measure, e.g. an arc. Now, if $f \in H^2(\mathbb{D})$ and $f = 0$ on $I$, then $f = 0$ everywhere by (2). That is, $R : H^2(\mathbb{D}) \to L^2(I)$, the restriction mapping, is injective. It also has dense range.

**Problem:** Given data $g \in L^2(I)$ can we find $f$ in $H^2(\mathbb{D})$ such that $f|_I \approx g$?

This was considered by Krein and Nudel’man [25]. In fact, for all $g \in L^2(I)$, there is a sequence $(f_n)$ in $H^2$ with $f_n|_I \to g$, but unless $g \in H^2|_I$, necessarily $\|f_n\| \to \infty$.

In applications (signal processing, control theory and inverse problems for PDEs), this is no use. It motivates the *Bounded Extremal Problem (BEP)*.

In its simplest form:

*Given* $g \in L^2(I)$ *and* $M > 0$, *find* $f \in H^2(\mathbb{D})$ *such that* $\|f\|_2 \leq M$ *and* $\|f|_I - g\|_{L^2(I)}$ *is minimized.*

The solution involves Toeplitz operators (as we shall see later) [2].

The $H^\infty$ version of the problem is useful in control theory. One difference may be noticed here: the closure of $H^\infty|_I$ is no longer $L^\infty(I)$. It is an open question to give a ‘useful’ description of it.

We may formulate the analogous Bounded Extremal Problem (BEP):

*Given* $g \in L^\infty(I)$ *and* $M > 0$, *find* $f \in H^\infty(\mathbb{D})$ *such that* $\|f\|_\infty \leq M$ *and* $\|f|_I - g\|_{L^\infty(I)}$ *is minimized.*

The solution to the $L^\infty$ version involves Hankel operators [3].

Returning to the $L^2$ problem, it is known that the solution is unique if $M > \|g\|$, and it saturates the constraint, i.e., $\|f\| = M$. Since it is an approximation procedure it can be used to recover functions in the presence of (small) noise and errors in measurements.

**Applications:** Krein–Nudel’man [25] were originally concerned with signal processing, and band-limited signals. Alpay–Baratchart–Leblond [2] provided applications in control theory and system identification. More recent work has focused on inverse problems for PDEs (e.g. heat flux).

**Example:** if $\phi$ real harmonic on $\mathbb{D}$ and we know $\phi$ and $\frac{\partial \phi}{\partial r}$ on $I$, then $\phi = \text{Re} \ f$, with $f$ analytic in $\mathbb{D}$, and by Cauchy–Riemann, $f \approx \phi + i \int \frac{\partial \phi}{\partial r}$. 
It is convenient to formulate a more general constrained approximation problem [9] (a Banach space version appears in [10]):

Let $H$, $K$, $L$ be Hilbert spaces and $A : H \to K$ and $B : H \to L$ linear operators. Given $k \in K$ and $\ell \in L$, find $h \in H$ to minimize $\|Ah - k\|$ subject to the condition $\|Bh - \ell\| \leq M$.

The solution (subject to conditions on $A$ and $B$ to make it well-posed) is as follows: there is a constant $\gamma > 0$ such that

$$(A^*A + \gamma B^*B)h = A^*k + \gamma B^*\ell$$

and $\|Bh - \ell\| = M$.

In our application, the spaces are $H = H^2(\D)$, $K = L^2(I)$, $L = H^2(\D)$, and the operators are given by $Af = f|_I$, and $Bf = f$. Thus, $A^* : L^2(I) \to H^2$, satisfies $A^*g = P_{H^2}(g \uplus 0)$, where $g \uplus 0$ is defined to be equal to $g$ on $I$ and to 0 on its complement. Hence $A^*A$ is a Toeplitz operator on $H^2$ with symbol $\chi_I$. This has links to the Carleman interpolation discussed above.

The annulus

Take the domain $A = \D \setminus \overline{r_0\D}$, with circular boundaries $r_0T$ and $T$, where $0 < r_0 < 1$. One application is to fault detection in pipelines [23, 26]. Functions in $H^2(A)$ have a Laurent series

$$f(z) = \sum_{n=-\infty}^{\infty} a_n z^n, \quad \text{with} \quad \sum_{n=-\infty}^{\infty} |a_n|^2(1 + r_0^{2n}) < \infty.$$  

Unfortunately harmonic functions in $A$ are not always the real part of analytic functions: consider $\log|z|$. But this is the only problem, and in physical situations can be dealt with.

Problem. Given (noisy) data $g \in L^2(K)$ of an unknown $H^2(A)$ function $f$ on $K \subset \partial A$, find the function. Similarly for a harmonic function and its normal derivative. Again we have a BEP: find $\min ||f|_K - g||_{L^2(K)}$ for $f \in H^2(A)$ with $||f|| \leq M$.

Once more, the solution involves Toeplitz operators, and in the most important case $K = T$ these are diagonal with respect to the orthonormal basis

$$\left(\frac{z^n}{(1 + r_0^{2n})^{1/2}}\right)_{n \in \Z}.$$  


Let $H$ be a complex Hilbert space, $(T_t)_{t \geq 0}$ a strongly continuous semigroup of bounded operators; i.e., $T(0) = I$, $T_{t+u} = T_t T_u$ for all $t, u \geq 0$, and $t \mapsto T_t x$ is continuous for each $x \in H$. We write $A$ for the infinitesimal generator, defined on a domain $\mathcal{D}(A)$, and given by

$$Ax = \lim_{t \to 0} \frac{1}{t}(T_t - I)x.$$
Note that the equation
\[ \dot{x}(t) = Ax(t), \quad x(0) = x_0, \]
has ‘mild’ solution \( x(t) = T_t x_0 \), for \( x_0 \in \mathcal{D}(A) \).

A continuous-time linear system in state form satisfies the equations (7), with \( x(0) = x_0 \), say. Often we take \( D = 0 \). Sometimes \( B \) and \( C \) (the control and observation operators) are bounded. If \( B \) is unbounded, we may allow it to map into a larger Hilbert space \( \tilde{H} \), and extend the semigroup to act on \( \tilde{H} \).

We consider a PDE example (the equation of an undamped beam).

**Example 6.1.**
\[ \frac{\partial^2 y}{\partial t^2} = \frac{\partial^4 y}{\partial x^4}, \quad 0 \leq x \leq 1, \quad t \geq 0, \]
with initial conditions on the position and velocity,
\[ y(x, 0) = y_1(x) \quad \text{and} \quad y_t(x, 0) = y_2(x), \]
given, and boundary conditions
\[ y(0, t) = y(1, t) = y_{xx}(0, t) = y_{xx}(1, t) = 0, \]
i.e., the beam is fixed at the endpoints. Let \( E = -\frac{d^2}{dx^2} \) with domain
\[ \mathcal{D}(E) = \{ z \in L^2(0, 1) : z, \frac{dz}{dx}, \frac{d^2 z}{dx^2} \in L^2(0, 1), z(0) = z(1) = 0 \}. \]

We can rewrite the equation as
\[ \frac{dz}{dt} = Az, \quad \text{with} \quad z = \begin{pmatrix} y \\ \frac{dy}{dt} \end{pmatrix}, \quad \text{and} \quad A = \begin{pmatrix} 0 & I \\ -E^2 & 0 \end{pmatrix}, \]
where \( z \) lies in \( \mathcal{D}(A) \), a subspace of the Hilbert space \( H = \mathcal{D}(E) \oplus L^2(0, 1) \), equipped with the norm
\[ \|(z_1, z_2)\|^2 = \|Ez_1\|^2 + \|z_2\|^2. \]

**(Infinite-time) admissibility**
There is a duality here between control and observation. Observation is simpler to explain.

**Admissibility of control operators.** Consider
\[ \frac{dx(t)}{dt} = Ax(t) + Bu(t), \quad x(0) = x_0, \quad t \geq 0, \]
where \( u(t) \in \mathcal{U} \) is the input at time \( t \), \( \mathcal{U} \) is a separable Hilbert space, and \( B : \mathcal{D}(B) \to \tilde{H} \) with \( \mathcal{D}(B) \subseteq \mathcal{U} \). How can we ensure that the state \( x(t) \) lies in \( H \)?
It is sufficient that $B \in \mathcal{L}(\mathcal{U}, \mathcal{D}(A^*)')$ and $\exists m_0 > 0$ such that
\[
\left\| \int_0^\infty T_t B u(t) \, dt \right\|_H \leq m_0 \| u \|_{L^2(0, \infty; \mathcal{U})}
\]
(the admissibility condition for $B$).

**Admissibility of observation operators.**
Consider the system
\[
\frac{dx(t)}{dt} = Ax(t), \quad y(t) = Cx(t),
\]
with $x(0) = x_0$, say. Let $C : \mathcal{D}(A) \to \mathcal{Y}$, Hilbert, be an $A$-bounded `observation operator', i.e., for some $m_1, m_2 > 0$,
\[
\|Cz\| \leq m_1 \|z\| + m_2 \|Az\|.
\]
$C$ is admissible, if $\exists m_0 > 0$ such that $y \in L^2(0, \infty; \mathcal{Y})$ and $\|y\|_2 \leq m_0 \|x_0\|$. Note that $y(t) = CT_t x_0$.

The duality here is that $B$ is an admissible control operator for $(T_t)_{t \geq 0}$ if and only if $B^*$ is an admissible observation operator for the dual semigroup $(T_t^*)_{t \geq 0}$.

**The Weiss conjecture**
Suppose $C$ is admissible, and take Laplace transforms: $y(t) = CT_t x_0$, so
\[
\hat{y}(s) = \int_0^\infty e^{-st} y(t) \, dt = C(sI - A)^{-1} x_0.
\]
Now if $y \in L^2(0, \infty; \mathcal{Y})$, then $\hat{y} \in H^2(\mathbb{C}_+, \mathcal{Y})$, the Hardy space on the right half-plane, and
\[
\|\hat{y}(s)\| = \left\| \int_0^\infty e^{-st} y(t) \, dt \right\| \leq \frac{\|y\|_2}{\sqrt{2 \text{Re} s}},
\]
by Cauchy–Schwarz. Thus admissibility, i.e.,
\[
\|CT_t x_0\|_{L^2(0, \infty; \mathcal{Y})} \leq m_0 \|x_0\|,
\]
implies the resolvent condition: there exists an $m_1 > 0$ such that
\[
\|C(sI - A)^{-1}\| \leq \frac{m_1}{\sqrt{2 \text{Re} s}}, \quad \forall s \in \mathbb{C}_+.
\]
Georges Weiss [40] conjectured that the two conditions are equivalent, and proved this in the case that the semigroup consists of normal operators. Interestingly, the known cases of the general conjecture imply several big theorems in function theory in an elementary way.

In fact the conjecture holds for contraction semigroups if $\dim \mathcal{Y} < \infty$ [18], but not in the general case [20, 42].
Example 1 Consider the case
\[ H = L^2(\mathbb{C}_+, \mu), \]
\[ (T_t(x))(\lambda) = e^{-\lambda t} x(\lambda), \]
\[ (Ax)(\lambda) = -\lambda x(\lambda). \]

For which Borel measures \( \mu \) on \( \mathbb{C}_+ \) does \( C \) defined by
\[ Cf = \int_{\mathbb{C}_+} f(\lambda) \, d\mu(\lambda) \]
satisfy the resolvent condition? Answer: if and only if
\[ \int_{\mathbb{C}_+} \frac{d\mu(\lambda)}{|s + \lambda|^2} \leq \frac{M}{\text{Re } s} \quad \forall s \in \mathbb{C}_+. \]
This actually means that \( \mu \) is a Carleson measure for \( \mathbb{C}_+ \), i.e., that the \( \mu \)-measure of square \([0, 2h] \times [a-h, a+h] \) is always at most \( O(h) \).

So when is \( C \) admissible? Precisely when there is a continuous embedding \( H^2(\mathbb{C}_+) \to L^2(\mathbb{C}_+, \mu) \).
This equivalence is the Carleson–Vinogradov embedding theorem for the half-plane, which we discussed in Section 3. Thus the Weiss conjecture for the above semigroup is equivalent to the above embedding theorem, an example of the reproducing kernel thesis. The link between Carleson measures and admissibility for normal semigroups goes back to Ho and Russell [16] and Weiss [39].

Example 2 Take the right shift semigroup on \( H = L^2(0, \infty) \), defined by
\[ (T_t x)(\tau) = \begin{cases} 0 & \text{if } \tau < t, \\ x(\tau - t) & \text{if } \tau \geq t. \end{cases} \]
Equivalently,
\[ H = H^2(\mathbb{C}_+), \]
\[ (T_t(x))(\lambda) = e^{-\lambda t} x(\lambda), \]
\[ (Ax)(\lambda) = -\lambda x(\lambda). \]

Now \( C : \mathcal{D}(A) \to \mathbb{C} \) is \( A \)-bounded if and only if it has the form
\[ Cx = \int_{-\infty}^{\infty} \frac{c(i\omega)x(i\omega)}{c(z)/(1 + z)} \, d\omega, \]
where \( c(z)/(1 + z) \in H^2(\mathbb{C}_+) \) (easy).
Then \( C \) is admissible if and only if the following Hankel operator is bounded:
\[ \Gamma_c : H^2(\mathbb{C}_-) \to H^2(\mathbb{C}_+), \quad \Gamma_c u = \Pi_+(c. u), \]
where \( \Pi_+ \) is the orthogonal projection from \( L^2(i\mathbb{R}) = H^2(\mathbb{C}_+) \oplus H^2(\mathbb{C}_-) \) onto \( H^2(\mathbb{C}_+) \).
The resolvent condition is equivalent to
\[
\left\| \Gamma_e \left( \frac{1}{s-a} \right) \right\| \leq m' \left\| \frac{1}{s-a} \right\| \quad \forall a \in \mathbb{C}_+.
\]

Hence the Weiss conjecture for the shift semigroup is equivalent to Bonsall’s theorem for the half-plane: a Hankel operator \( \Gamma_e \) bounded if and only if bounded on normalized reproducing kernels. This was shown in [33].

The proof of the Weiss conjecture for contraction semigroups uses results for unitary semigroups, a stronger form of the result for the shift, and some auxiliary results. Combining the ‘smaller’ theorems gives the big theorem, which in turn implies the smaller theorems.

**Controllability**

Assume we have an exponentially stable semigroup \((T_t)_{t \geq 0}\), i.e.,
\[
\|T_t\| \leq Me^{-\lambda t}, \quad (t \geq 0),
\]
for some \( M > 0 \) and \( \lambda > 0 \). Again we consider the equation
\[
\frac{dx(t)}{dt} = Ax(t) + Bu(t),
\]
with solution
\[
x(t) = T_tx_0 + \int_0^t T_{t-s}Bu(s) \, ds,
\]
suitably interpreted. Suppose that \( B \) is admissible (an easier case to describe). Then we have a bounded operator \( B_\infty : L^2(0, \infty; \mathcal{U}) \to H \), defined by
\[
B_\infty u = \int_0^\infty T_tBu(t) \, dt.
\]
The system is **exactly controllable**, if \( \text{Im} B_\infty = H \), i.e., we can steer the system where we like, using the input \( u \). Alternatively, it is **approximately controllable**, if \( \text{Im} B_\infty \) is dense. There are dual notions of exact and approximate observability, which we omit.

Controllability involves more links with the theory of interpolation, as follows.

**Diagonal semigroups.** We now describe an important special case where most things are known (includes some heat equations, vibrating structures, etc.) Suppose that
\[
A\phi_n = \lambda_n \phi_n,
\]
with \((\phi_n)\) normalized eigenvectors forming a Riesz basis. Let \((\psi_n)\) be the dual basis. So every \( x \in H \) can be written
\[
x = \sum_{n=1}^{\infty} \langle x, \psi_n \rangle \phi_n
\]
with
\[ K_1 \sum |\langle x, \psi_n \rangle|^2 \leq \|x\|^2 \leq K_2 \sum |\langle x, \psi_n \rangle|^2. \]

Note that
\[ T_t \sum_{n=1}^{\infty} c_n \phi_n = \sum_{n=1}^{\infty} c_n e^{\lambda_{nt}} \phi_n. \]
By exponential stability \( \sup_n \Re \lambda_n < 0 \).

We shall concentrate on the case \( \dim \mathcal{U} = 1 \), which is simpler to explain. However, finite-dimensional \( \mathcal{U} \) can be handled similarly. If \( \dim \mathcal{U} = 1 \) then \( B : \mathbb{C} \rightarrow \mathcal{H} \) and is represented by a vector,
\[ b = \sum_{n=1}^{\infty} b_n \phi_n. \]

Let us calculate \( B_{\infty} \):
\[ B_{\infty} u = \int_{0}^{\infty} T_t Bu(t) \, dt = \sum_{n=1}^{\infty} b_n \int_{0}^{\infty} e^{\lambda_{nt}} u(t) \, dt \, \phi_n. \]
This is just
\[ \sum_{n=1}^{\infty} b_n \hat{u}(-\lambda_n) \phi_n \]
(the hat denoting a Laplace Transform).

Since \( \mathcal{L} : L^2(0, \infty) \rightarrow H^2(\mathbb{C}_+) \) is an isomorphism, exact controllability is equivalent to the following:

for every \((c_n) \in \ell^2\) there is a function \( g \in H^2(\mathbb{C}_+) \) such that
\[ b_n g(-\lambda_n) = c_n. \]
This brings us back directly to Carleson interpolation problems in the \( H^2 \) (Shapiro–Shields) version. The exact result needed was given by McPhail [29], although once again we have to rewrite things for the half-plane. The scalar case was given in [19], and the vectorial case \( \dim \mathcal{U} > 1 \) in [21]. The latter required some new interpolation theory which is beyond the scope of this paper.

The necessary and sufficient condition for exact controllability is that
\[ \nu = \sum_{n=1}^{\infty} \frac{|\Re \lambda_n|^2}{|b_n|^2} \prod_{k \neq n} \frac{|\lambda_n + \lambda_k|^2}{|\lambda_k - \lambda_n|^2} \delta_{-\lambda_n} \]
is a Carleson measure on \( \mathbb{C}_+ \), i.e., that the \( \nu \)-measure of a square \([0, 2h] \times [a-h, a+h]\) is \( O(h) \). Here \( \delta_\lambda \) denotes a Dirac (point mass) at \( \lambda \).

We don’t need \((-\lambda_n)\) to be a Carleson sequence, but it does need to be a Blaschke sequence. However \( \lambda_n = -n^\beta \) with \( 0 < \beta \leq 1 \) is never exactly
controllable (if $\dim U < \infty$).

Approximate controllability (dense range) is much easier and just requires $b_n \neq 0$ for all $n$ and distinct eigenvalues $(\lambda_n)$. For finite-dimensional systems (working with matrices, not just operators) it is well known that exact and approximate controllability coincide.

Finally, an intermediate concept is null controllability, which corresponds to being able to steer a system to the ‘zero’ state. We just require $\text{Im} \mathcal{B}_\infty$ to contain $T_{t_1}H$ for some $t_1 \geq 0$. The necessary and sufficient condition now becomes that

$$\nu = \sum_{n=1}^{\infty} \frac{|\text{Re}\lambda_n|^2 e^{2t_1 \text{Re}\lambda_n}}{|b_n|^2} \prod_{k \neq n} \frac{|\lambda_n + \lambda_k|^2 |b_k - \lambda_n|^2 \delta_{-\lambda_n}}{|\lambda_k - \lambda_n|^2}$$

should be a Carleson measure on $\mathbb{C}_+$.

REFERENCES


School of Mathematics, University of Leeds, Leeds LS2 9JT, U.K.

E-mail address: J.R_Partington@leeds.ac.uk