Banach Spaces of Analytic Vector-valued Functions

Steven John Barclay

Submitted in accordance with the requirements for the degree of PhD

The University of Leeds

Department of Pure Mathematics

September 2007

The candidate confirms that the work submitted is his own and that appropriate credit has been given where reference has been made to the work of others.

Acknowledgements

Firstly, I would like to thank my supervisor Jonathan R. Partington for his support, invaluable guidance and endless patience. I know that I have not been an easy student to work with, and I have been very lucky to have someone so kind and forgiving as a supervisor. Also, many thanks to all my colleagues in the Mathematics department for their friendliness and interesting mathematical talks and discussions.

I would also like to thank my former housemates Pete, James, Norman, Henry and Marcus for their support and friendship through difficult times, and for taking me out in the evenings and making me laugh. Many thanks to my friend Xhermie in the Philippines for sending messages of support and encouragement. I would like to reserve special thanks to Daniel, my dear friend back home in Mexico now, who made me feel so welcome when I visited last year, and gave me a much needed boost to continue. I wish him well.

I am very grateful to my family, especially my mother, for their love and patience, which has been so valuable in helping me continue my PhD to the end.

Finally, I would like to thank the Engineering and Physical Sciences Research Council for their invaluable financial support.

Abstract

The main theme of the thesis is the study of continuity and approximation problems, involving matrix-valued and vector-valued Hardy spaces on the unit disc $\mathbb{D}$ and its boundary $\mathbb{T}$ in the complex plane. The first part of the thesis looks at the factorization of square matrix-valued boundary functions, beginning with spectral factorization in Chapter 2. Then ideas involving approximations with inner and outer functions are used to solve a matrix analogue of the Douglas-Rudin problem in Chapter 3. In both cases, considerable extra difficulties are created by the noncommutativity of matrix multiplication.

More specifically, we show that the matrix spectral factorization mapping is sequentially continuous from L^p to H^{2p} (where $1 \leq p < \infty$), under the additional assumption of uniform integrability of the logarithms of the spectral densities to be factorized. We show, moreover, that this condition is necessary for continuity, as well as sufficient. Concerning the Douglas-Rudin problem in Chapter 3, we show that any log-integrable essentially bounded square matrix-valued function f can be written in the form h^*g, where h and g lie in H^∞. Extensions to other L^p spaces, with norm bounds on the factors of f, are also provided.

The final part of the thesis takes a somewhat different direction. In Chapter 4, we consider the problem of weighted H^∞ approximation of vector-valued L^∞ functions on the unit circle, subject to a weighted sup-type constraint on the size of the approximant. This involves the development of a suitable theory of vector-valued L^∞ and H^∞ functions on $\mathbb{T}$, taking values in an arbitrary Banach space equipped with a separable predual. We establish existence of a solution under mild assumptions, and characterise some of its properties. We also show that in the scalar case, the unconstrained version of this problem is not well posed in general.

Contents

Chapter 1

Introduction

1.1 Notation and conventions

Let $\mathbb{D}$ denote the open unit disc in the complex plane and let $\mathbb{T}$ denote its boundary, the unit circle. Let μ denote normalized Lebesgue measure on $\mathbb{T}$. We will be working with Hardy spaces of vector-valued and matrix-valued analytic functions on $\mathbb{D}$, together with their measurable boundary functions.

Throughout the thesis, we work exclusively in the disc. However, analoguous results will in most cases hold for the half-plane. This conversion can be achieved in the standard way, by means of a conformal mapping between the disc and the half-plane.

1.1.1 Schatten p-norms and matrix-valued functions

Let $n \in \mathbb{N}$ and let $p \in [1, \infty]$. Throughout the first three chapters of the thesis, we will be working almost exclusively with $n \times n$ matrix-valued functions and there n shall be kept fixed. We shall write $\mathrm{L}^p(\mathcal{L}(\mathbb{C}^n))$ to mean the Lebesgue space of measurable functions from $\mathbb{T}$ to the space of $n \times n$ complex matrices,

$\mathcal{L}(\mathbb{C}^n)$, with finite $\|\cdot\|_p$ norm, given by:

$$\|f\|_p = \begin{cases} \left(\frac{1}{2\pi}\int_0^{2\pi} \|f(e^{i\theta})\|_p^p \, \mathrm{d}\theta\right)^{1/p} & \text{if } 1 \le p < \infty \\ \operatorname*{ess\,sup}_{\theta\in[0,2\pi)} \|f(e^{i\theta})\|_\infty & \text{if } p = \infty, \end{cases}$$

where $f : \mathbb{T} \to \mathcal{L}(\mathbb{C}^n)$ is any measurable function. In this definition, we choose to use the $\|\cdot\|_p$ norm on $\mathcal{L}(\mathbb{C}^n)$, rather than an arbitrary norm. This makes the above norms easier for us to work with, but it makes no difference to the norm topology. For any index $p \in [1,\infty]$ we define the Schatten p-norm on $\mathcal{L}(\mathbb{C}^n)$ by:

$$\|M\|_p = \begin{cases} \operatorname{Tr}(M^*M)^{p/2} & p < \infty \\ \sup_{0\neq v\in\mathbb{C}^n} \|Mv\|_2/\|v\|_2 & p = \infty, \end{cases}$$

for all $n \times n$ matrices $M \in \mathcal{L}(\mathbb{C}^n)$. This norm satisfies the relations:

$$\begin{aligned} \|A^*\|_p &= \|A\|_p \\ |\operatorname{Tr} A| &\le \|A\|_1 \\ \|B^*A\|_p &\le \|A\|_q\|B\|_r, \end{aligned} \tag{1.1}$$

for all $A, B \in \mathcal{L}(\mathbb{C}^n)$ and indices $p, q, r \in [1,\infty]$ satisfying $1/p = 1/q + 1/r$. For a proof of the above three relations, together with the fact that $\|\cdot\|_p$ is a norm as claimed, see Chapter 5 or [4, ch. 11], for example. Alternatively, see [16, section 11.9] for a proof of inequality (1.1) in the well known case $p = 1$.

We shall let L^p denote the usual scalar Lebesgue space of measurable complex-valued functions on $\mathbb{T}$ with finite $\|\cdot\|_p$ norm, given with respect to μ.

For $p, q \in [1,\infty]$ conjugate indices and functions f and g in $\mathrm{L}^p(\mathcal{L}(\mathbb{C}^n))$ and $\mathrm{L}^q(\mathcal{L}(\mathbb{C}^n))$ respectively, we define the inner product:

$$\langle f, g\rangle = \frac{1}{2\pi}\int_0^{2\pi} \operatorname{Tr}(g(e^{i\theta})^* f(e^{i\theta}))\, \mathrm{d}\theta$$

The above integrand is dominated by the mapping $\theta \mapsto \|f(e^{i\theta})\|_p \, \|g(e^{i\theta})\|_q$, which is integrable, and so this inner product is well defined.

By combining the earlier matrix relations with the usual (scalar) formulation of Hölder's inequality, we obtain the following analogue of Hölder's inequality:

$$|\langle f, g\rangle| \leq \|g^* f\|_1 \leq \|f\|_p \|g\|_q,$$

where g^*, the hermitian conjugate of g, is taken pointwise, i.e. $g^*(z) = g(z)^*$ for all $z \in \mathbb{T}$. We will frequently use expressions such as $\det g$, $\operatorname{Tr} g$, fg, g^{-1}, etc. These are always intended to be taken pointwise.

1.1.2 Hardy spaces

For any $p \in [1, \infty]$, let H^p denote the Hardy space of measurable complex-valued Hardy class functions on $\mathbb{T}$ with finite $\|\cdot\|_p$ norm, regarded as a closed subspace of L^p. In other words, H^p consists of those functions in L^p whose harmonic extension to $\mathbb{D}$ is analytic.

For $p \in [1, \infty]$, we write $\mathrm{H}^p(\mathcal{L}(\mathbb{C}^n))$ to mean the Hardy space of matrix-valued functions in $\mathrm{L}^p(\mathcal{L}(\mathbb{C}^n))$, with an analytic extension to $\mathbb{D}$. Thus $\mathrm{H}^p(\mathcal{L}(\mathbb{C}^n))$ is regarded as a closed subspace of $\mathrm{L}^p(\mathcal{L}(\mathbb{C}^n))$. Vector-valued and rectangular matrix-valued Hardy spaces $\mathrm{H}^p(\mathbb{C}^n)$ and $\mathrm{H}^p(\mathcal{L}(\mathbb{C}^m, \mathbb{C}^n))$ can be defined analogously and they will make a brief appearance in the following subsection. We will always be referring to the analytic extension implicitly, when we evaluate H^p functions at points in $\mathbb{D}$.

Let $\mathcal{P}(\mathbb{C}^n)$ denote the vector space of $\mathbb{C}^n$-valued trigonometric polynomials on $\mathbb{T}$. This is dense in $\mathrm{L}^p(\mathbb{C}^n)$ for all $p \in [1, \infty)$. Similarly, we define $\mathcal{P}^+(\mathbb{C}^n)$ to be the space of $\mathbb{C}^n$-valued polynomials over $\mathbb{C}$, considered as a subspace of $\mathcal{P}(\mathbb{C}^n)$. We shall only use these spaces briefly in the following subsection.

Some additional function spaces will also make an occasional appearance: Let C denote the Banach space of complex-valued continuous functions on $\mathbb{T}$, equipped with the uniform norm. Let $\mathrm{A} \subset \mathrm{C}$ denote the *disc algebra*, consisting of those continuous functions on $\mathbb{T}$ which extend continuously to an analytic function on $\mathbb{D}$. The *Wiener algebra* $\mathcal{W}$ (resp. *positive Wiener algebra* $\mathcal{W}^+$) consists of those functions in C (resp. A) with an absolutely convergent Fourier series.

The space $\overline{\mathrm{H}_0^\infty}$, of pointwise complex conjugates $\overline{f}$ of those scalar H^∞ functions f which satisfy $f(0) = 0$, will also make a brief appearance in Chapter 3.

In Chapter 4, we shall encounter more general classes of vector-valued functions on $\mathbb{T}$, and they will be denoted analoguously, such as $\mathrm{L}^1(E)$, $\mathrm{H}^\infty(E)$, $\mathrm{C}(E)$, etc., taking values in a given Banach space E. Such function spaces will be defined later in Chapter 4.

1.1.3 Spectral factorization

Definition 1.1.1 *A matrix-valued function $\rho \in \mathrm{H}^2(\mathcal{L}(\mathbb{C}^n))$ is said to be* outer *if the set of products $\rho\mathcal{P}^+(\mathbb{C}^n)$ is dense in $\mathrm{H}^2(\mathbb{C}^n)$.*

For any $0 \leq m \leq n$, we say that a matrix-valued function $\theta \in \mathrm{H}^\infty(\mathcal{L}(\mathbb{C}^m, \mathbb{C}^n))$ is inner *if it is an isometry almost everywhere on $\mathbb{T}$.*

A matrix-valued function $b \in \mathrm{H}^2(\mathcal{L}(\mathbb{C}^n))$ will be outer if and only if every inner factor of b is constant and unitary almost everywhere on $\mathbb{T}$. To show this, we consider the H^2 closure of the set $b\mathcal{P}^+(\mathbb{C}^n)$. This is a shift-invariant subspace of $\mathrm{H}^2(\mathbb{C}^n)$, so by the vector-valued Beurling-Lax theorem [24, page 14], it is equal to $\theta\mathrm{H}^2(\mathbb{C}^m)$ for some $0 \leq m \leq n$ and an inner function $\theta \in \mathrm{H}^\infty(\mathcal{L}(\mathbb{C}^m, \mathbb{C}^n))$. So if b is not outer then θ cannot be constant and unitary. But θ is clearly an inner factor of b. Conversely, suppose b has an inner factor θ which is nonconstant or nonunitary. Then θ has no H^∞ inverse, which implies that $\theta\mathrm{H}^2(\mathbb{C}^n)$ is not dense. Since this space contains $b\mathcal{P}^+(\mathbb{C}^n)$, b cannot be outer.

Any outer function $\rho \in \mathrm{H}^2(\mathcal{L}(\mathbb{C}^n))$ is almost everywhere invertible. This assertion follows because the set $\rho\mathcal{P}(\mathbb{C}^n)$ is dense in $\mathrm{L}^2(\mathbb{C}^n)$, which implies that $\rho(z)$ has dense range for almost all $z \in \mathbb{T}$.

We will also occasionally need to use the following fact about outer functions. If ρ and ρ' are outer functions in $\mathrm{H}^2(\mathcal{L}(\mathbb{C}^n))$ such that $\rho'\rho^{-1}$ lies in $\mathrm{L}^2(\mathcal{L}(\mathbb{C}^n))$, then $\rho'\rho^{-1}$ lies in $\mathrm{H}^2(\mathcal{L}(\mathbb{C}^n))$ and it is also outer. We prove this as follows:

Let $V = \rho\mathcal{P}^+(\mathbb{C}^n)$. The set $(\rho'\rho^{-1})V$ is clearly a dense subset of $\mathrm{H}^2(\mathbb{C}^n)$. But V is also dense in $\mathrm{H}^2(\mathbb{C}^n)$, so by taking H^2 limits of points in V, we find that

$(\rho'\rho^{-1})\mathrm{H}^2(\mathbb{C}^n)$ is a dense subset of $\mathrm{H}^1(\mathbb{C}^n)$. This shows that $\rho'\rho^{-1} \in \mathrm{H}^2(\mathcal{L}(\mathbb{C}^n))$. We also know that $\mathrm{H}^2(\mathbb{C}^n) \cap (\rho'\rho^{-1})\mathrm{H}^2(\mathbb{C}^n)$ is dense in $\mathrm{H}^2(\mathbb{C}^n)$, since it contains $(\rho'\rho^{-1})V$. This implies that $\rho'\rho^{-1}$ has no nontrivial inner factor, and is therefore outer. (In particular, ρ^{-1} is an outer function in $\mathrm{H}^2(\mathcal{L}(\mathbb{C}^n))$, whenever it lies in $\mathrm{L}^2(\mathcal{L}(\mathbb{C}^n))$.)

Definition 1.1.2 *A positive matrix-valued function $w \in \mathrm{L}^1(\mathcal{L}(\mathbb{C}^n))$ is said to be a* spectral density *if it can be written in the form $\rho^*\rho$, for some matrix-valued outer function $\rho \in \mathrm{H}^2(\mathcal{L}(\mathbb{C}^n))$. We call this a* spectral factorization *of w and the functions ρ and ρ^* are called the* spectral factors *of w.*

If w is a spectral density, then its spectral factors will be unique up to some unitary constant. Indeed, if $w = \rho^*\rho = \rho'^*\rho'$ for outer functions $\rho, \rho' \in \mathrm{H}^2(\mathcal{L}(\mathbb{C}^n))$, then the function $\rho'\rho^{-1}$ is seen to be an isometry almost everywhere. It therefore lies in $\mathrm{L}^2(\mathcal{L}(\mathbb{C}^n))$, proving that it is an outer function in $\mathrm{H}^2(\mathcal{L}(\mathbb{C}^n))$. Hence $\rho'\rho^{-1}$ is both inner and outer. This implies that it is an almost everywhere unitary constant.

There will always be a unique spectral factor $\rho \in \mathrm{H}^2(\mathcal{L}(\mathbb{C}^n))$ for which $\rho(0)$ is a positive matrix. We shall call this the *canonical spectral factor* of w and denote it $\Phi(w)$. This defines a mapping

$$w \mapsto \Phi(w),$$

called the *spectral factorization mapping.*

At the start of Chapter 2, we shall establish the well known fact that a positive matrix-valued function $w \in \mathrm{L}^1(\mathcal{L}(\mathbb{C}^n))$ is a spectral density if and only if it is log-integrable, that is w is almost everywhere nonsingular and $\log\det w$ is an integrable function.

1.2 Continuity of the spectral factorization mapping

Spectral factorization [11] [28, chap. 6] is the process by which a positive matrix-valued function w, on the unit circle in $\mathbb{C}$, is expressed in the form $w(e^{i\theta}) = \rho(e^{i\theta})^* \rho(e^{i\theta})$, for a certain complex matrix-valued function ρ, where $\theta \in [0, 2\pi)$. There are many contexts in which this factorization naturally arises, for example stochastic processes, linear quadratic control design, and H^∞ robust control.

We require the function ρ to have a certain analytic extension to the open unit disc in $\mathbb{C}$, and with a suitable constraint on $\rho(0)$ we can ensure that the spectral factor ρ is unique. In this way, we obtain a mapping from w to this canonical choice of ρ, which we term the *spectral factorization mapping*. (It is also possible to define this mapping for functions on the imaginary line in $\mathbb{C}$, for which the spectral factors are required to have analytic extensions to the right half-plane.)

There are many situations in which it is desirable to know the continuity properties of this mapping. One reason for this is that for applications it common to use approximation methods, when working with spectral factors, and this makes it important to be able to provide error estimates. In Chapter 2, we will address some of the continuity issues.

Our work will improve on some results of earlier authors. Specifically, we show that L^p convergence of a sequence of positive matrix-valued functions is necessary and sufficient for H^{2p} convergence of their spectral factors, where $1 \leq p < \infty$, provided a certain uniform log-integrability condition is placed on this sequence. This condition is fairly weak, and it turns out to be satisfied automatically in many situations.

In the case $p = \infty$, it is already well known that continuity of the spectral factorization mapping fails. (The extra condition above is redundant here.) This is shown by Anderson [1], who provides a counterexample. However, it is also shown in [1] that L^∞ continuity can be recovered by placing a suitable constraint on the derivatives of the functions to be factorized. Additionally, $L^\infty \to H^2$ continuity is established in [1]. This latter result also follows as a corollary of our results, as we shall observe in the penultimate section.

In [21], Jacob et al. work in the half-plane rather than the disc, and consider spectral factorizations of functions on the imaginary line, which extend analytically to some vertical strip about that line. Amongst other things, they establish an $L^p \to H^p$ type continuity result, under the additional assumption that the functions concerned and their inverses are uniformly bounded. However, they deal exclusively with scalar-valued functions.

Other continuity results have been obtained by Jacob and Partington [20], who look at spectral factorization on decomposing Banach algebras. This class includes the Wiener algebra of all absolutely convergent Fourier series. There they establish local Lipschitz continuity of the spectral factorization mapping for such algebras. This provides a rather different class of results to ours.

1.3 The Douglas-Rudin problem

In Chapter 3, we provide an extension to Bourgain's result [6], which characterizes the set of pointwise products $\overline{h}g$ on the unit circle $\mathbb{T}$, for nonzero functions $g, h \in H^\infty$. It is a simple observation that any function $f \in L^\infty$ which takes this form must be *log-integrable*, that is f is almost everywhere nonzero and $\log f$ lies in L^1. The Douglas-Rudin problem asks whether the converse is true. This hypothesis was originally raised in [14], where it was conjectured not to be true. However, Bourgain provides a construction for such a factorization $\overline{h}g$, showing that the log-integrability condition is indeed sufficient.

The problem may be generalized to square matrix-valued functions on $\mathbb{T}$ as follows: One may easily show that a necessary condition for $f \in L^\infty(\mathcal{L}(\mathbb{C}^n))$ to take the form h^*g, for $g, h \in H^\infty(\mathcal{L}(\mathbb{C}^n))$ almost everywhere invertible, is for the function $\log|\det f|$ to be integrable. Is this log-integrability condition also sufficient? In Chapter 3, we provide an affirmative answer to this question.

Our methods follow the steps of Bourgain's constructive proof. Before outlining them, we shall illustrate why obtaining such a factorization is a nontrivial problem, which leads on to the idea behind our construction. For simplicity, let us consider the scalar case: In Bourgain's paper [6], it is observed that any log-integrable function $f \in \mathrm{L}^\infty$ admits a factorization of the form $\phi\rho$, for some outer $\rho \in \mathrm{H}^\infty$ and some a.e. unimodular $\phi \in \mathrm{L}^\infty$. Hence the problem reduces to that of factorizing unimodular functions ϕ.

The two simplest kinds of factorizations we may consider are those of the form $\overline{\theta_2}\theta_1$ for inner functions $\theta_1, \theta_2 \in \mathrm{H}^\infty$, which are manifestly unimodular, and those of the form $\overline{\rho_2}\rho_1$ for outer functions $\rho_1, \rho_2 \in \mathrm{H}^\infty$. However, it turns out that these two classes of products do not suffice to cover every unimodular function. In the last section of Chapter 3, we exhibit a uniformly dense subset of unimodular functions which do not admit a factorization of either form. One such example is the function $f : \mathbb{T} \to \mathbb{C}$ given by:

$$f(z) = \begin{cases} i & \text{if } \mathrm{Im}(z) \geq 0 \\ -i & \text{if } \mathrm{Im}(z) < 0 \,, \end{cases}$$

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

It turns out, however, that each kind of factorization is able to provide a good *approximation* to any given unimodular function. It is shown in [14] and [17, ch. 10], for example, that ratios of inner functions form a uniformly dense subset of the unimodular functions. In the case of outer approximation, observe that for any unimodular ϕ lying in the *Wiener algebra* $\mathcal{W} \subset \mathrm{L}^\infty$, with winding number zero, there exists $f \in \mathcal{W}$ such that $f(z) = \arg\phi(z)$ for all $z \in \mathbb{T}$. But now f has an orthogonal projection $a = \mathrm{P}_{\mathrm{H}^2} f$, onto $\mathcal{W}^+ \subset \mathrm{L}^\infty$, the *positive Wiener algebra* of functions in $\mathcal{W}$ with an analytic extension to $\mathbb{D}$. Setting $\rho = e^{ia/2} \in \mathrm{H}^\infty$ outer, we have:

$$\overline{\rho}^{-1}\rho = e^{i(a+\overline{a})/2} = e^{if} = \phi \,,$$

giving the desired factorization of ϕ with outer functions in H^∞. But such ϕ form an L^p-dense subset of the unimodular functions, for any $p < \infty$. We therefore obtain good uniform approximation of unimodular functions with inner factors, and good L^p approximation with bounded outer factors.

This suggests a possible *iterative* approach to obtaining an exact factorization for a larger class of unimodular functions: Given some $\phi \in L^\infty$ unimodular, we can find a good uniform approximation $\overline{\theta_0}\Theta_0$ to ϕ, for inner functions $\theta_0, \Theta_0 \in H^\infty$. This gives rise to a unimodular error:

$$\theta_0 \phi \Theta_0^{-1} \overset{L^\infty}{\approx} 1 .$$

Now we may find a good L^p approximation $\overline{\tau_1}\rho_1$ to the above function, for outer functions $\rho, \tau \in H^\infty$. We then obtain a unimodular error:

$$\overline{\tau_1}^{-1}(\theta_0\phi\Theta_0^{-1})\rho_1^{-1} \overset{L^p}{\approx} 1 .$$

Now we look for a good uniform approximation $\overline{\Theta_1}\theta_1$ to this new error, for inner functions $\theta_1, \Theta_1 \in H^\infty$, giving rise to the error:

$$\theta_1(\overline{\tau_1}^{-1}\theta_0\phi\Theta_0^{-1}\rho_1^{-1})\Theta_1^{-1} \overset{L^\infty}{\approx} 1 .$$

We proceed iteratively, finding a good L^p approximation $\overline{\tau_k}\rho_k$ to the error, with outer functions $\rho_k, \tau_k \in H^\infty$, then finding a good uniform approximation $\overline{\theta_k}\Theta_k$ to the new error, with inner functions $\theta_k, \Theta_k \in H^\infty$, for each $1 < k \in \mathbb{N}$. By taking successively better and better approximations at each step, we obtain an infinite product,

$$\cdots\ \theta_k\overline{\tau_k}^{-1} \cdots\ \theta_1\overline{\tau_1}^{-1}\theta_0\phi\Theta_0^{-1}\rho_1^{-1}\Theta_1^{-1} \cdots\ \rho_k^{-1}\Theta_k^{-1} \cdots$$

converging in measure. Now by setting:

$$\begin{aligned} g &= \lim_{k\to\infty}(\Theta_k\rho_k \cdots \Theta_1\rho_1)\cdot\Theta_0 ; \\ h &= \lim_{k\to\infty}(\theta_k\tau_k \cdots \theta_1\tau_1)\cdot\theta_0 , \end{aligned}$$

we have the formal identity: $\overline{h}^{-1}\phi g^{-1} = 1$. Thus provided that the above limits converge in H^∞, we obtain the desired factorization $\overline{h}g$ of ϕ.

The point of using the above iterative construction is the fact that by alternating between approximation with outer functions and approximation by inner functions, it is possible to make the limits g and h converge in H^∞ for *any* unimodular function $\phi \in L^\infty$. This is due to the complementary nature of the two different kinds of approximation.

We shall use analogous steps for the matrix case. Spectral factorization is used to reduce the Douglas-Rudin problem to that of factorizing a *unitary-valued* function on $\mathbb{T}$. We then follow the iterative approach above, replacing the scalar-valued functions in the above expressions with matrix-valued functions, and working with unitary errors in place of unimodular errors.

The noncommutativity of matrix multiplication creates considerable extra difficulties in our construction. For example, we cannot simply gather together the inner and outer terms in the above expressions, take their pointwise arguments and then work with sums and differences of real-valued functions, as is done in Bourgain's approach.

1.4 Vector-valued H^∞ approximation

Chapter 4 concerns itself with the task of finding a good Hardy class approximant to a given vector-valued measurable function, defined on the unit circle of the complex plane. Problems of this kind appear in a number of different situations, particularly robust identification in systems theory [2], signal processing [22, 29] and for certain inverse problems for the Laplacian [7]. See the survey article [9] for an overview of these applications.

Our results generalize those of several previous authors, looking at certain H^∞ bounded extremal problems. It is shown, for example, in [2] that for any measurable subset K of the unit circle $\mathbb{T}$, and bounded measurable complex-valued functions f and h on K and $\mathbb{T} \setminus K$ respectively, an H^∞ function g can be found which minimizes $\|f - g|_K\|_\infty$ subject to constraint that $\|h - g|_{\mathbb{T}\setminus K}\|_\infty$ is at most M, for any $M > 0$ sufficiently large. (Such a constraint is necessary to make the problem well posed.) In Chapter 4, we will generalize results such as these to the case of (possibly infinite dimensional) vector or operator-valued functions, with the possibility of using *weighted* L^∞ norms on both K and $\mathbb{T} \setminus K$. Such generality could be of utility in systems theory, where different levels of accuracy may be required on different parts of the frequency domain.

The main idea presented in Chapter 4 is to expess certain H^∞ optimization problems, such as the one just mentioned, in terms of seminorms acting on a Banach space X, where the solutions are to be found in a smaller subspace Y corresponding to the Hardy class functions. More precisely, given two seminorms $\|\cdot\|_A$ and $\|\cdot\|_B$ on X, vectors $x_A, x_B \in X$ and a bound $M > 0$, the problem is to find $y \in Y$ which minimizes $\|x_A - y\|_A$, subject to the constraint that $\|x_b - y\|_B$ is at most M. By replacing X and Y with vector-valued L^∞ and H^∞ spaces respectively, and taking $\|\cdot\|_A$ and $\|\cdot\|_B$ to be weighted L^∞ (semi)norms on X (with possibly vanishing weights), we obtain a very general class of constrained H^∞ approximation problems. We establish sufficient conditions for the existence and (later on) the uniqueness and saturation of the constraints of the solutions to these problems.

A large section of Chapter 4 will be spent developing a suitable theory of vector-valued L^∞ and H^∞ spaces taking values in a Banach space equipped with a separable predual, such as the space $\mathcal{L}(H)$ for a separable Hilbert space H. There does not appear to be an adequate theory of these spaces developed in the literature. In particular, the usual Lesbegue-Bochner definition for vector-valued L^∞ is unsuitable to use in our context, as the notion of measurability in such a space is too strong. This prevents it from having a natural predual, which is vital for our approximation results to hold.

In Chapter 4, we will be concerned with H^p and L^p spaces mainly for the cases $p = \infty$ and $p = 1$. In Section 4.1 at the start of the chapter, the case $p = 2$ is also involved, where we make use of the basic theory of Hankel operators [26]. These are shift-invariant linear maps between the Hilbert spaces H^2 and $(H^2)^\perp = L^2 \ominus H^2$.

Chapter 2

Spectral factorization

In this chapter, we show that the matrix spectral factorization mapping is sequentially continuous from L^p to H^{2p} (where $1 \leq p < \infty$), under the additional assumption of uniform integrability of the logarithms of the spectral densities to be factorized. We shall show, moreover, that this condition is necessary for continuity, as well as sufficient.

The outline of this chapter is as follows: In Section 2.1, we define the spectral factorization mapping and state some of its properties, and it is here that we state the main theorem of this chapter precisely. Section 2.2 will be used to derive an accompanying proposition. In Section 2.3, we will establish some lemmas which will be needed to prove the theorem. In Section 2.4, we will complete the proof and then use it to derive some related results. Finally, we make some concluding remarks in Section 2.5.

2.1 Basic properties of Φ

In this section, we look at some of the basic properties of the matrix spectral factorization mapping.

Not every positive matrix-valued function in $L^1(\mathcal{L}(\mathbb{C}^n))$ is a spectral density. The following theorem, taken from [18], will provide the necessary and sufficient conditions for a spectral factorization to exist.

Theorem 2.1.1 (Helson and Lowdenslager) *Let w be a positive matrix-valued function in $\mathrm{L}^1(\mathcal{L}(\mathbb{C}^n))$. A necessary and sufficient condition for w to have a factorization of the form*

$$w = b^* b,$$

where b is in $\mathrm{H}^2(\mathcal{L}(\mathbb{C}^n))$ with $\det b(0) \neq 0$, is that

$$\int_{\mathbb{T}} \operatorname{Tr} \log w \, \mathrm{d}\mu > -\infty. \tag{2.1}$$

If this condition is satisfied, we can choose b so that

$$\int_{\mathbb{T}} \log |\det b| \, \mathrm{d}\mu = \log |\det b(0)|. \tag{2.2}$$

Equation (2.2) is in fact an extremal condition that is satisfied precisely when b is outer. For arbitrary H^2 functions, (2.2) becomes an inequality with the left hand side greatest. This is just an extension of Jensen's inequality, treated for scalar-valued H^1 functions in [19]. We refer the reader to [18, pages 193–195] for more details relating to the above theorem.

Thus as a corollary of Theorem 2.1.1, we have the following proposition, characterizing spectral densities in $\mathrm{L}^1(\mathcal{L}(\mathbb{C}^n))$.

Proposition 2.1.2 *Let $w \in \mathrm{L}^1(\mathcal{L}(\mathbb{C}^n))$ be a positive matrix-valued function. The following are equivalent:*

1. *w is a spectral density.*
2. *$\log w$ is an integrable function defined almost everywhere on $\mathbb{T}$.*
3. *$\log \det w$ is an integrable function.*
4. *$\langle \log \det w, 1 \rangle > -\infty$.*

Moreover, the above inner product satisfies

$$2 \log |\det \rho(0)| = \langle \log \det w, 1 \rangle = \langle \log w, 1 \rangle, \tag{2.3}$$

whenever ρ is a spectral factor of w.

Proof

1 $\Leftrightarrow$ *4*. As we remarked above, equation (2.2) of Theorem 2.1.1 is equivalent to the H^2 factor of w being outer. Thus w is a spectral density if and only (2.1) is satisified. But since $\operatorname{Tr}\log w$ and $\log\det w$ are equal, the integral in (2.1) is just $\langle \log\det w, 1\rangle$, so this is greater than $-\infty$ if and only if w is spectral density.

4 $\Rightarrow$ *2*. Since $\log\det w > -\infty$ almost everywhere, $w(z)$ is invertible and therefore has a self-adjoint logarithm for almost all $z \in \mathbb{T}$. For all $x \in (0,\infty)$, we have $|\log x| \leq 2x - \log x$. By applying this inequality to the eigenvalues of $w(z)$ and then summing them, we find that

$$\|\log w(z)\|_1 \;\leq\; 2\operatorname{Tr} w(z) - \operatorname{Tr}\log w(z) \;=\; 2\operatorname{Tr} w(z) - \log\det w(z)\,,$$

for almost all $z \in \mathbb{T}$. By integrating over $\mathbb{T}$, we obtain the inequality:

$$\|\log w\|_1 \;\leq\; 2\langle w, 1\rangle - \langle \log\det w, 1\rangle\,,$$

which provides a finite upper bound for $\|\log w\|_1$, since w is integrable.

2 $\Rightarrow$ *3* $\Rightarrow$ *4*. Trivial.

Now from equation (2.2) of Theorem 2.1.1, for any spectral factor ρ we have

$$\log|\det\rho(0)| \;=\; \langle \log|\det\rho|, 1\rangle\,.$$

But $|\det\rho|^2$ is equal to $\det\rho^*\rho$, so the right hand side of the above equation is one half of $\langle \log\det w, 1\rangle = \langle \operatorname{Tr}\log w, 1\rangle = \langle \log w, 1\rangle$, establishing (2.3). □

Equation (2.3) of the above proposition will be fundamental to many of the results we shall later obtain. We will look at some of the continuity properties of the function

$$\Phi : \{\text{spectral densities in } L^1(\mathcal{L}(\mathbb{C}^n))\} \longrightarrow H^2(\mathcal{L}(\mathbb{C}^n))\,,$$

which maps a given spectral density to its canonical spectral factor.

The main result we shall obtain is the following:

Theorem 2.1.3 *Let $w_r, w \in \mathrm{L}^1(\mathcal{L}(\mathbb{C}^n))$ be spectral densities, for all $r \in \mathbb{N}$. The following are equivalent:*

1. *$w_r \to w$ in L^1 as $r \to \infty$ and $\{\log\det w_r : r \in \mathbb{N}\}$ is uniformly integrable.*

2. *$\Phi(w_r) \to \Phi(w)$ in H^2 as $r \to \infty$.*

The above terminology may be unfamiliar. For a collection $\mathcal{G}$ of measurable functions on an arbitrary measure space $(X, \mathcal{E}, \nu)$, we say that $\mathcal{G}$ is *uniformly integrable* if it is bounded in L^1 (i.e. the L^1 norms of the functions in $\mathcal{G}$ form a bounded subset of $\mathbb{R}$) and for all $\varepsilon > 0$, there is some $\delta > 0$ such that

$$\int_E |f|\, \mathrm{d}\nu \; < \; \varepsilon$$

for all $f \in \mathcal{G}$ and $E \in \mathcal{E}$ with $\nu(E) < \delta$. Even though this property is stronger than boundedness in L^1, it is weaker than boundedness in L^p, for any $p > 1$, provided the underlying measure space is finite. There are equivalent definitions of uniform integrability. Consult [27, pages 270–279] for a standard treatment of this topic.

In the next section, we will show that the uniform log-integrability condition for $(\det w_r)_{r\in\mathbb{N}}$ in Theorem 2.1.3, is equivalent to several other conditions on the sequence $(w_r)_{r\in\mathbb{N}}$.

2.2 A preliminary result

First of all, we shall introduce some more notation:

For any matrix $A \in \mathcal{L}(\mathbb{C}^n)$, write $|A|$ to mean the positive matrix $(A^*A)^{1/2}$. Now, for any self-adjoint matrices $A, B \in \mathcal{L}(\mathbb{C}^n)$, we define the following operations:

$$\begin{aligned} A \vee B &= \frac{1}{2}\Big(A + B + |A - B|\Big) \\ A \wedge B &= \frac{1}{2}\Big(A + B - |A - B|\Big) \end{aligned}$$

Note that $\vee$ and $\wedge$ are commutative but not associative (for $n > 1$).

For any self-adjoint matrices $A, B, C \in \mathcal{L}(\mathbb{C}^n)$, the following identities hold:

$$\begin{aligned} -(A \wedge B) &= (-A) \vee (-B) \\ (A+C) \vee (B+C) &= (A \vee B) + C \end{aligned}$$

In addition, the following statements hold:

$$A \wedge B \;\leq\; A, B \;\leq\; A \vee B$$
$$C = A \vee B \quad \text{whenever} \quad A, B \leq C \leq A \vee B$$

In other words, $A \vee B$ is a minimal upper bound for A and B. Minimality follows from the fact that if C is chosen as above, we have

$$0 \;\leq\; (A \vee B) - C \;\leq\; (B - A) \vee 0\,,\; (A - B) \vee 0\,.$$

It follows that

$$\ker((A \vee B) - C) \;\supseteq\; \ker((A - B) \vee 0)\,,\; \ker((B - A) \vee 0)\,.$$

But the ranges of $(A - B) \vee 0$ and $(B - A) \vee 0$ are orthogonal, and so their kernels span $\mathbb{C}^n$. Hence $C = A \vee B$.

Likewise, $A \wedge B$ is a maximal lower bound for A and B. Note that A and B will not have universal upper and lower bounds in general. Indeed, their existence would contradict the non-associativity of $\wedge$ and $\vee$.

For self-adjoint matrix-valued functions $f, g : \mathbb{T} \to \mathcal{L}(\mathbb{C}^n)$, let $f \wedge g$, $f \vee g$ be the functions given by

$$\begin{aligned} (f \wedge g)(z) &= f(z) \wedge g(z) \\ (f \vee g)(z) &= f(z) \vee g(z)\,, \end{aligned}$$

for all $z \in \mathbb{T}$. We will use the binary operators $\wedge$ and $\vee$ in the proof of the main result of this section, and then later in the proof of the main theorem.

We will need the following lemma:

Lemma 2.2.1 *Let $A, B \in \mathcal{L}(\mathbb{C}^n)$ be positive matrices and suppose that $A \leq B$. Then the following inequalities hold:*

1. $\sqrt{A} \leq \sqrt{B}$,

2. $\log A \leq \log B$, *provided* A *and* B *are invertible.*

The first assertion is a special case of the Löwner-Heinz inequality [25], which states that $A^p \leq B^p$ for any positive operators $A \leq B$, whenever $p \in [0,1]$. This fact, along with the second assertion, follow from a general result, Löwner's thoerem, which characterizes the so-called *operator monotone functions*. These are real-valued functions f, defined on an interval $I \subset \mathbb{R}$, such that $f(A) \leq f(B)$ for any self-adjoint operators $A \leq B$ with spectrum in I. See [13] and [23, thm. 2.2.6, p. 47] for more details.

Here we shall provide a separate, elementary proof for the two specific matrix cases given in the lemma statement.

Proof of lemma

Part 1. Let λ be the minimum nonnegative real number such that $\sqrt{B} + \lambda I \geq \sqrt{A}$, or equivalently, such that $\lambda I \geq \sqrt{A} - \sqrt{B}$. Thus we may suppose that λ is (the maximum) eigenvalue of $\sqrt{A} - \sqrt{B}$, since otherwise $\lambda = 0$ and we are done.

Let $v \in \mathbb{C}^n$ be any eigenvector corresponding to λ. Then $(\sqrt{B} + \lambda I)v$ is equal to $\sqrt{A}\,v$. In particular their norms are equal, so that

$$\langle (B + 2\lambda\sqrt{B} + \lambda^2 I)v, v\rangle \;=\; \|(\sqrt{B} + \lambda I)v\|_2^2 \;=\; \|\sqrt{A}\,v\|_2^2 \;=\; \langle Av, v\rangle\,.$$

Rearranging, we have

$$\lambda^2\|v\|_2^2 \;+\; 2\lambda\langle\sqrt{B}v, v\rangle \;+\; \langle (B - A)v, v\rangle \;=\; 0$$

By the positivity of $B - A$ and $\sqrt{B}$, each of the terms of the above expression are nonnegative. This implies that $\lambda = 0$, proving the first assertion.

Part 2. It is well known that the mapping

$$x \mapsto r(x^{1/r} - 1)$$

converges to $x \mapsto \log x$, locally uniformly in $x \in (0, \infty)$, as r tends to ∞. Provided the matrices A and B are invertible, their spectra lie in $(0, \infty)$. Then by the spectral mapping theorem we have

$$\begin{aligned} \|r(A^{1/r}-1)-\log A\| &= \max_{x\in\sigma(A)} |r(x^{1/r}-1)-\log x| \\ &\to 0 \qquad \text{as } r\to\infty\,. \end{aligned}$$

The same happens for B in place of A. In particular, this shows that

$$\begin{aligned} \log A &= \lim_{k\to\infty} 2^k(A^{2^{-k}}-1)\,, \\ \log B &= \lim_{k\to\infty} 2^k(B^{2^{-k}}-1)\,. \end{aligned}$$

We have $A \le B$ by hypothesis. So by induction on k, it follows from the first assertion that

$$A^{2^{-k}} \le B^{2^{-k}} \quad \text{for all } k\in\mathbb{N}\,.$$

Subtracting 1 from both sides, multiplying by 2^k and taking limits, we find that

$$\log A \le \log B\,,$$

as required. □

We remark that there are examples of positive matrices $A, B \in \mathcal{L}(\mathbb{C}^2)$, such that $A \le B$, but A^p and B^p are incomparable for any $p > 1$.

We are now ready to prove the main result of this section, which is stated as follows:

Proposition 2.2.2 *Let $w_r, w \in \mathrm{L}^1(\mathcal{L}(\mathbb{C}^n))$ be spectral densities, for all $r\in\mathbb{N}$. Suppose that $w_r \to w$ in L^1 as $r\to\infty$. Then the following are equivalent:*

1. $\det\Phi(w_r)(0) \to \det\Phi(w)(0)$ *as* $r\to\infty$.
2. $\log w_r \to \log w$ *in* L^1 *as* $r\to\infty$.
3. $\log\det w_r \to \log\det w$ *in* L^1 *as* $r\to\infty$.
4. $\{\log\det w_r : r\in\mathbb{N}\}$ *is uniformly integrable.*

Proof

1 ⇒ 2. For any $r \in \mathbb{N}$, the functions w_r, w and $w_r \vee w$ are a.e. positive invertible matrix-valued, and therefore have self-adjoint logarithms defined almost everywhere on $\mathbb{T}$. Also, since $w_r, w \le w_r \vee w$, we have the relations:

$$\log w_r,\ \log w \ \le\ \log(w_r \vee w) \quad \text{for all } r \in \mathbb{N}\,,$$

as a consequence of Lemma 2.2.1.

We first show that $\log(w_r \vee w) \to \log w$ in L^1 as $r \to \infty$, as follows: For all $x \in (0,\infty)$, the nonnegative real number, $(\log x) \vee 0$, is less than or equal to x. This implies that the positive matrix-valued function

$$\log(w_r \vee w) \vee 0$$

is dominated by $w_r \vee w$. So we obtain the chain of inequalities:

$$\log w \ \le\ \log(w_r \vee w) \ \le\ \log(w_r \vee w) \vee 0 \ \le\ w_r \vee w\,.$$

But since w is a spectral density and $w_r, w \in \mathrm{L}^1(\mathcal{L}(\mathbb{C}^n))$, the terms $\log w$ and $w_r \vee w$ above, are both integrable. Moreover, it is clear from the definition, $2(w_r \vee w) = w_r + w + |w_r - w|$, that $w_r \vee w \to w$ in L^1 as $r \to \infty$. So the collection of functions $\{(w_r \vee w) - \log w : r \in \mathbb{N}\}$ is uniformly integrable. Subtracting $\log w$ from the terms in the above chain of inequalities, we obtain

$$0 \ \le\ \log(w_r \vee w) - \log w \ \le\ (w_r \vee w) - \log w\,,$$

from which we deduce that $\{\log(w_r \vee w) - \log w : r \in \mathbb{N}\}$ is uniformly integrable. But since $\log(w_r \vee w) \to \log w$ in measure as $r \to \infty$, we have

$$\log(w_r \vee w) \to \log w \quad \text{in } \mathrm{L}^1 \quad \text{as } r \to \infty\,.$$

By Minkowski's inequality, we have

$$\begin{aligned}
\|\log w_r - \log w\|_1 &\le \|\log(w_r \vee w) - \log w_r\|_1 + \|\log(w_r \vee w) - \log w\|_1 \\
&= \langle \log(w_r \vee w) - \log w_r,\ 1\rangle + \langle \log(w_r \vee w) - \log w,\ 1\rangle \\
&= \langle \log w - \log w_r,\ 1\rangle + 2\langle \log(w_r \vee w) - \log w,\ 1\rangle \\
&= \langle \log w - \log w_r,\ 1\rangle + 2\underbrace{\|\log(w_r \vee w) - \log w\|_1}_{\to 0 \ \text{ as } r \to \infty}\,.
\end{aligned}$$

So to establish *2*, it is sufficient to prove that $\langle \log w_r, 1\rangle$ tends to $\langle \log w, 1\rangle$ as $r \to \infty$. This is the point where we need to use *1*.

By equation (2.3) of Proposition 2.1.2, we have the identities:

$$\begin{aligned} 2\log\det\Phi(w)(0) &= \langle \log w, 1\rangle \\ 2\log\det\Phi(w_r)(0) &= \langle \log w_r, 1\rangle\,, \end{aligned}$$

for all $r \in \mathbb{N}$. But since $\det\Phi(w_r)(0) \to \det\Phi(w)(0)$ as $r \to \infty$, the continuity of log implies that $\langle \log w_r, 1\rangle$ tends to $\langle \log w, 1\rangle$ as $r \to \infty$, proving *2*.

2 ⇒ *3*. The trace operation is L^1 norm-decreasing, so

$$\begin{aligned} \|\log\det w_r - \log\det w\|_1 &= \|\mathrm{Tr}(\log w_r - \log w)\|_1 \\ &\leq \|\log w_r - \log w\|_1 \\ &\to 0 \qquad \text{as } r \to \infty\,. \end{aligned}$$

3 ⇒ *4*. Trivial. (L^1 convergence always implies uniform integrability.)

4 ⇒ *1*. We have already established the identities:

$$\begin{aligned} 2\log\det\Phi(w)(0) &= \langle \log\det w, 1\rangle &= \int_{\mathbb{T}} \log\det w\,\mathrm{d}\mu \\ 2\log\det\Phi(w_r)(0) &= \langle \log\det w_r, 1\rangle &= \int_{\mathbb{T}} \log\det w_r\,\mathrm{d}\mu \end{aligned}$$

But $\log\det w_r$ converges to $\log\det w$ in measure as r tends to ∞. So by the uniform integrability of $\{\log\det w_r : r \in \mathbb{N}\}$ and the integrability of $\log\det w$, we find that

$$\int_{\mathbb{T}} \log\det w_r\,\mathrm{d}\mu \to \int_{\mathbb{T}} \log\det w\,\mathrm{d}\mu \quad \text{as } r \to \infty\,.$$

In other words, $\log\det\Phi(w_r)(0) \to \log\det\Phi(w)(0)$ as $r \to \infty$. This proves *1*. □

2.3 Some inequalities

In this section we will establish in succession, two inequalities which will lead to a proof of Theorem 2.1.3. The first of these will show that in some appropriate sense, spectral densities which are close to 1 have particular spectral factors which are also close to 1. This will lead to the second result, which deals with two spectral densities which are close to each other, and deduces an inequality concerning their canonical spectral factors.

It will be useful to introduce some more notation here:
Recall that μ denotes normalised Lebesgue measure on $\mathbb{T}$. This makes the triple $(\mathbb{T}, \mathcal{B}(\mathbb{T}), \mu)$ into a probability space, where $\mathcal{B}(\mathbb{T})$ denotes the σ-algebra of Borel subsets of $\mathbb{T}$. So any measurable function

$$X : \mathbb{T} \to \mathcal{L}(\mathbb{C}^n)$$

may be regarded as a matrix-valued random variable. In the case when X is integrable, we have a notion of expectation and variance, given as follows:

$$\begin{aligned} \mathbb{E}X &= \int_{\mathbb{T}} X \, \mathrm{d}\mu \\ \operatorname{var} X &= \tfrac{1}{n} \|X - \mathbb{E}X\|_2^2 \\ &= \tfrac{1}{n} \operatorname{Tr} \mathbb{E}((X - \mathbb{E}X)^*(X - \mathbb{E}X)) \end{aligned} \tag{2.4}$$

We have the identities:

$$\mathbb{E}(X^*(\mathbb{E}X)) = \mathbb{E}((\mathbb{E}X^*)X) = \mathbb{E}((\mathbb{E}X^*)(\mathbb{E}X)) = (\mathbb{E}X)^*(\mathbb{E}X),$$

from which we obtain the alternative formula:

$$\begin{aligned} \operatorname{var} X &= \tfrac{1}{n} \operatorname{Tr}(\mathbb{E}(X^*X) - (\mathbb{E}X)^*(\mathbb{E}X)) \\ &= \tfrac{1}{n} \left(\|X\|_2^2 - \|\mathbb{E}X\|_2^2\right) \end{aligned} \tag{2.5}$$

When X is square integrable, this is just the statement that $\mathbb{E}X$ is orthogonal to $X - \mathbb{E}X$. In fact, $\mathbb{E}X$ is precisely the orthogonal projection of X onto the matrix-valued constant functions.

In the next lemma, we shall be considering the variance of a spectral factor ρ, for a given spectral density $w \in \mathrm{L}^1(\mathcal{L}(\mathbb{C}^n))$. The two key facts we shall use are that:

$$\mathbb{E}\rho = \rho(0),$$

$$\|\rho\|_2^2 = \|\rho^*\rho\|_1 = \|w\|_1 .$$

This shows that the variance of ρ is small provided w is close to 1 and $\rho(0)$ is close to the identity. From this we can deduce that ρ is close to 1, as well. It turns out that the determinant of $\rho(0)$ will provide the constraint we need to show that $\rho(0)$ is close to the identity, so it is involved in the inequality which follows:

Lemma 2.3.1 *Let $w \in \mathrm{L}^1(\mathcal{L}(\mathbb{C}^n))$ be a spectral density and let $\rho \in \mathrm{H}^2(\mathcal{L}(\mathbb{C}^n))$ be a spectral factor of w. Suppose that $\rho(0)$ has positive spectrum. Then*

$$\frac{1}{n}\|\rho - 1\|_2^2 \leq \frac{1}{n}\|w\|_1 - 2(\det \rho(0))^{1/n} + 1 \tag{2.6}$$

$$\leq \frac{1}{n}\|w - 1\|_1 + 2\big(1 - (\det \rho(0))^{1/n}\big) . \tag{2.7}$$

This upper bound tends to 0 whenever $w \to 1$ in L^1 and $\langle \log\det w, 1\rangle \to 0$.

Proof Every eigenvalue of $\rho(0)$ is positive. By applying the AM-GM inequality to these eigenvalues, we find that

$$(\det \rho(0))^{1/n} \leq \frac{1}{n}\mathrm{Tr}\,\rho(0) = \frac{1}{n}\langle \rho(0), 1\rangle .$$

Therefore

$$\begin{aligned} \frac{1}{n}\|\rho(0) - 1\|_2^2 &= \frac{1}{n}\langle \rho(0) - 1, \rho(0) - 1\rangle \\ &= \frac{1}{n}\|\rho(0)\|_2^2 - \frac{2}{n}\langle \rho(0), 1\rangle + 1 \\ &\leq \frac{1}{n}\|\rho(0)\|_2^2 - 2(\det \rho(0))^{1/n} + 1 . \end{aligned} \tag{2.8}$$

From the variance formula (2.5), we have

$$\begin{aligned} \mathrm{var}\,\rho &= \frac{1}{n}\left(\|\rho\|_2^2 - \|\rho(0)\|_2^2\right) , \\ \mathrm{var}(\rho - 1) &= \frac{1}{n}\left(\|\rho - 1\|_2^2 - \|\rho(0) - 1\|_2^2\right) . \end{aligned}$$

But ρ and $\rho - 1$ have the same variance, by the definition given in (2.4). So by adding this variance to both sides of inequality (2.8), we find that

$$\frac{1}{n}\|\rho - 1\|_2^2 \leq \frac{1}{n}\|\rho\|_2^2 - 2(\det \rho(0))^{1/n} + 1 .$$

Substituting $\|w\|_1$ for $\|\rho\|_2^2$, we obtain (2.6). By the triangle inequality,

$$\tfrac{1}{n}\|w\|_1 \;\leq\; \tfrac{1}{n}(\|w-1\|_1 + \|1\|_1) \;=\; \tfrac{1}{n}\|w-1\| + 1\,.$$

So by substituting the right hand side into (2.6), we obtain inequality (2.7).

Recalling identity (2.3) from Proposition 2.1.2,

$$2\log|\det\rho(0)| \;=\; \langle \log\det w, 1\rangle\,,$$

we find that $\det\rho(0) \to 1$ whenever $\langle \log\det w, 1\rangle \to 0$. Therefore the upper bound given in (2.7) tends to 0, whenever $w \to 1$ in L^1 as well. □

We make two remarks about the above lemma. Firstly, a spectral factor ρ, for which $\rho(0)$ has positive spectrum, will not be the canonical spectral factor in general, i.e. $\rho(0)$ need not be positive. In fact, a matrix is diagonalizable and has positive spectrum if and only if it is the product of two positive, invertible matrices. We shall use (and prove) this fact in the next lemma.

Secondly, for functions w and ρ as in the statement of the lemma, we have the following chain of inequalities:

$$(\det\rho(0))^{1/n} \leq \tfrac{1}{n}\langle\rho(0),1\rangle \;\leq\; \tfrac{1}{n}\|\rho(0)\|_1 \;\leq \tfrac{1}{\sqrt{n}}\|\rho(0)\|_2 \;\leq\; \left(\tfrac{1}{n}\|w\|_1\right)^{1/2}$$

When $w \to 1$ in L^1 and $\langle \log\det w, 1\rangle \to 0$, the leftmost and rightmost terms of the above chain tend to unity. Hence all the terms tend to unity, and this forces ρ to tend to 1 in H^2. This summarizes the main idea of the proof.

As a consequence of Lemma 2.3.1, we have following result:

Lemma 2.3.2 *Let $w, w' \in \mathrm{L}^1(\mathcal{L}(\mathbb{C}^n))$ be spectral densities and let $0 < \eta \leq R < \infty$ be given constants.*

Suppose that $w \geq \eta$, and let $P \in \mathrm{L}^\infty(\mathcal{L}(\mathbb{C}^n))$ be the projection-valued function given by

$$P \;=\; \chi_{[0,R]}(w)\,,$$

where $\chi_{[0,R]}$ is the characteristic function of the interval $[0,R]$. Let ρ, ρ' denote the canonical spectral factors $\Phi(w), \Phi(w')$ respectively. Then

$$\|(\rho'-\rho)P\|_2^2 \;\leq\; R\left(\frac{\|w'-w\|_1}{\eta} \;+\; 2n\left[1-\left(\frac{\det\rho'(0)}{\det\rho(0)}\right)^{1/n}\right]\right).$$

Proof Let $\rho'' = \rho'\rho^{-1}$. This function lies in $\mathrm{L}^2(\mathcal{L}(\mathbb{C}^n))$ since ρ^{-1} is bounded. But ρ and ρ' are outer and, as we showed in Section 2.1, this implies that ρ'' is an outer function in $\mathrm{H}^2(\mathcal{L}(\mathbb{C}^n))$. Therefore ρ'' is a spectral factor of the spectral density $w'' \in \mathrm{L}^1(\mathcal{L}(\mathbb{C}^n))$ given by

$$w'' = \rho''^* \rho'' = \rho^{-1^*} w' \rho^{-1} .$$

Now since $\rho''(0) = \rho'(0)\,\rho^{-1}(0)$ is a product of two positive invertible matrices, it has positive spectrum. Indeed, if $P, Q \in \mathcal{L}(\mathbb{C}^n)$ are positive and invertible, then PQ is conjugate to $Q^{1/2} P Q^{1/2}$, which is manifestly positive and invertible and therefore has positive spectrum.

[*Aside:* If $X \in \mathcal{L}(\mathbb{C}^n)$ is diagonalizable with positive spectrum, then $X = L^{-1} D\, L$ for some $D, L \in \mathcal{L}(\mathbb{C}^n)$ invertible with D positive and diagonal. Therefore X is the product of $L^{-1} D\, L^{-1^*}$ and $L^* L$, both of which are positive and invertible.]

So by Lemma 2.3.1, we have the inequality:

$$\begin{aligned} \|\rho'' - 1\|_2^2 &\leq \|w'' - 1\|_1 + 2n\bigl(1 - (\det \rho''(0))^{1/n}\bigr) \\ &= \|w'' - 1\|_1 + 2n\left[1 - \left(\frac{\det \rho'(0)}{\det \rho(0)}\right)^{1/n}\right] . \end{aligned} \tag{2.9}$$

But $w'' - 1$ is equal to $\rho^{-1^*}(w' - w)\rho^{-1}$, and so

$$\|w'' - 1\|_1 \leq \|\rho^{-1^*}\|_\infty \|w' - w\|_1 \|\rho^{-1}\|_\infty \leq \eta^{-1} \|w' - w\|_1 . \tag{2.10}$$

Now, $\rho' - \rho$ is equal to $(\rho'' - 1)\rho$, so we have

$$\begin{aligned} \|(\rho' - \rho)P\|_2^2 &\leq \|\rho'' - 1\|_2^2 \|\rho P\|_\infty^2 \\ &= \|PwP\|_\infty \|\rho'' - 1\|_2^2 \\ &\leq R\,\|\rho'' - 1\|_2^2 . \end{aligned} \tag{2.11}$$

By combining (2.10) and (2.11) with inequality (2.9), we obtain the desired inequality, proving the lemma. □

2.4 The main theorem

We now have all the tools needed to prove Theorem 2.1.3. Recalling the statement of the theorem, we are given a spectral density w, together with a sequence of spectral densities $(w_r)_{r\in\mathbb{N}}$, and we aim to show that the following two statements are equivalent:

1. $w_r \to w$ in L^1 as $r \to \infty$ and $\{\log\det w_r : r \in \mathbb{N}\}$ is uniformly integrable.

2. $\Phi(w_r) \to \Phi(w)$ in H^2 as $r \to \infty$.

We will use Proposition 2.2.2 to reformulate the first statement. This will enable us to prove that 2 implies 1. The other direction is harder, and for this we will make use of Lemma 2.3.2.

Proof of Theorem 2.1.3

2 ⇒ 1. Let ρ, ρ_r denote the canonical spectral factors $\Phi(w), \Phi(w_r)$ respectively, for each $r \in \mathbb{N}$. The function $w_r - w$ can be written in the form:

$$\frac{(\rho_r+\rho)^*(\rho_r-\rho)}{2} + \frac{(\rho_r-\rho)^*(\rho_r+\rho)}{2}$$

So by Hölder's inequality, together with the triangle inequality, we have

$$\begin{aligned}\|w_r - w\|_1 &\leq \tfrac{1}{2}\,\|(\rho_r+\rho)^*(\rho_r-\rho)\|_1 + \tfrac{1}{2}\,\|(\rho_r-\rho)^*(\rho_r+\rho)\|_1 \\ &\leq \tfrac{1}{2}\,\|\rho_r+\rho\|_2\,\|\rho_r-\rho\|_2 + \tfrac{1}{2}\,\|\rho_r-\rho\|_2\,\|\rho_r+\rho\|_2 \\ &\leq (\|\rho_r\|_2 + \|\rho\|_2)\,\|\rho_r-\rho\|_2\,,\end{aligned}$$

for all $r \in \mathbb{N}$. But the sequence of reals

$$(\|\rho_r\|_2 + \|\rho\|_2)_{r\in\mathbb{N}}$$

is bounded since $(\rho_r)_{r\in\mathbb{N}}$ is convergent in H^2. Therefore $\|w_r - w\|_1$ tends to 0 as $r \to \infty$. This establishes the first part of *1.* Now as r tends to ∞,

$$\det\Phi(w_r)(0) \;\to\; \det\Phi(w)(0)\,,$$

since $\Phi(w_r) \to \Phi(w)$ in H^2. So by Proposition 2.2.2, the set $\{\log\det w_r : r \in \mathbb{N}\}$ is uniformly integrable. This establishes the second part of *1.*

1 $\Rightarrow$ *2.* Fix constants $0 < \eta \leq R < \infty$. Let $\tilde{w}$ be the spectral density $w \vee \eta$ and let $\tilde{\rho}$ denote its canonical spectral factor $\Phi(\tilde{w})$. Let $P \in L^\infty(\mathcal{L}(\mathbb{C}^n))$ be the projection-valued function given by

$$P = \chi_{[0,R]}(w) = \chi_{[0,R]}(\tilde{w}).$$

By Lemma 2.3.2, we have the inequalities:

$$\|(\rho - \tilde{\rho})P\|_2^2 \leq \frac{R}{\eta}\|w - \tilde{w}\|_1 + 2nR\left|1 - \left(\frac{\det\rho(0)}{\det\tilde{\rho}(0)}\right)^{1/n}\right|, \tag{2.12}$$

$$\|(\rho_r - \tilde{\rho})P\|_2^2 \leq \frac{R}{\eta}\|w_r - \tilde{w}\|_1 + 2nR\left|1 - \left(\frac{\det\rho_r(0)}{\det\tilde{\rho}(0)}\right)^{1/n}\right|, \tag{2.13}$$

for all $r \in \mathbb{N}$. We shall now obtain an upper bound on $\|\rho_r - \rho\|_2$, in terms of the left hand sides of (2.12) and (2.13).

The functions $(\rho_r - \rho)P$ and $(\rho_r - \rho)(1 - P)$ are orthogonal, since

$$\begin{aligned}\langle(\rho_r - \rho)P, (\rho_r - \rho)1\rangle &= \langle(\rho_r - \rho)P^2, (\rho_r - \rho)1\rangle \\ &= \langle(\rho_r - \rho)P, (\rho_r - \rho)P\rangle.\end{aligned}$$

Therefore by Pythagoras' theorem, we have

$$\begin{aligned}\|\rho_r - \rho\|_2^2 - \|(\rho_r - \rho)P\|_2^2 &= \|(\rho_r - \rho)(1 - P)\|_2^2 \\ &\leq (\|\rho_r(1 - P)\|_2 + \|\rho(1 - P)\|_2)^2 \\ &\leq \left[\|(1 - P)w_r(1 - P)\|_1^{1/2} + \|w(1 - P)\|_1^{1/2}\right]^2 \\ &\leq \Big[\Big(\|(1 - P)(w_r - w)(1 - P)\|_1 \\ &\qquad + \|(1 - P)w(1 - P)\|_1\Big)^{1/2} \\ &\qquad\qquad + \|w(1 - P)\|_1^{1/2}\Big]^2 \\ &\leq \left[(\|w_r - w\|_1 + \|w(1 - P)\|_1)^{1/2} + \right. \\ &\qquad \left. \|w(1 - P)\|_1^{1/2}\right]^2,\end{aligned}$$

which gives us the inequality:

$$\begin{aligned}\|\rho_r - \rho\|_2^2 \;\leq\; & \left(\|(\rho-\tilde{\rho})P\|_2 + \|(\rho_r-\tilde{\rho})P\|_2\right)^2 \;+ \\ & \left[\left(\|w_r - w\|_1 + \|w(1-P)\|_1\right)^{1/2} + \|w(1-P)\|_1^{1/2}\right]^2 \end{aligned} \tag{2.14}$$

Let $\varepsilon > 0$. Recall that the functions P and $\tilde{w}$ depend on R and η respectively. We wish to choose R, η and find $N \in \mathbb{N}$, such that for all $r \geq N$, the following five inequalities are satisfied:

$$\|w(1-P)\|_1 \;\leq\; \frac{\varepsilon^2}{18} \tag{2.15}$$

$$\frac{\|w-\tilde{w}\|_1}{\eta} \;\leq\; \frac{\varepsilon^2}{18R} \tag{2.16}$$

$$\left|1 - \left(\frac{\det\rho(0)}{\det\tilde{\rho}(0)}\right)^{1/n}\right| \;\leq\; \frac{\varepsilon^2}{36nR} \tag{2.17}$$

$$\left|\frac{[\det\rho_r(0)]^{1/n} - [\det\rho(0)]^{1/n}}{[\det\tilde{\rho}(0)]^{1/n}}\right| \;\leq\; \frac{\varepsilon^2}{36nR} \tag{2.18}$$

$$\frac{\|w_r - w\|_1}{\eta} \;\leq\; \frac{\varepsilon^2}{18R} \tag{2.19}$$

We shall establish (2.15) to (2.19) in numerical order. We first choose R to satisfy (2.15), as follows shortly. Next, we look for an η which satisfies (2.16) and (2.17). Finally, we choose N such that (2.18) and (2.19) are satisfied.

The positive scalar-valued function $\operatorname{Tr} w$ is integrable and therefore

$$\int_{E_k} \operatorname{Tr} w \,\mathrm{d}\mu \qquad \left(\text{where } E_k = \{z \in \mathbb{T} : \operatorname{Tr} w(z) > k\}\right)$$

tends to zero whenever $k \to \infty$. Choose R to be any positive k for which the above integral is at most $\varepsilon^2/18$. Since $\|w(z)\|_\infty$ is at most R for any $z \in \mathbb{T} \setminus E_R$, the function $1 - P$ is supported on the set E_R. This implies that

$$\|w(1-P)\|_1 \;=\; \int_{\mathbb{T}} \operatorname{Tr}(w(1-P))\,\mathrm{d}\mu \;\leq\; \int_{E_R} \operatorname{Tr} w \,\mathrm{d}\mu \;\leq\; \frac{\varepsilon^2}{18},$$

establishing (2.15).

For any $x \in (0, \infty)$, we have

$$\frac{(\eta \vee x) - x}{\eta} = \left(1 - \frac{x}{\eta}\right) \vee 0 \leq \left(-\log \frac{x}{\eta}\right) \vee 0 = \log(\eta \vee x) - \log x \,.$$

This implies that

$$\frac{(\eta \vee w) - w}{\eta} \leq \log(\eta \vee w) - \log w \,.$$

So by integrating the trace of both sides, we find that

$$\frac{\|\tilde{w} - w\|_1}{\eta} = \frac{\langle \tilde{w} - w, 1 \rangle}{\eta} \leq \langle \log \tilde{w} - \log w, 1 \rangle \,. \tag{2.20}$$

Now by equation (2.3) from Proposition 2.1.2, the right hand side is just

$$2 \log \det \tilde{\rho}(0) - 2 \log \det \rho(0) = -2 \log \left(\frac{\det \rho(0)}{\det \tilde{\rho}(0)} \right) ,$$

and this is also an upper bound for the expression

$$2n \left[1 - \left(\frac{\det \rho(0)}{\det \tilde{\rho}(0)} \right)^{1/n} \right] \geq 0 \,. \tag{2.21}$$

Now observe that $\tilde{w} \longrightarrow w$ almost everywhere on $\mathbb{T}$ as $\eta \longrightarrow 0$. This is because $w(z)$ is invertible for almost all $z \in \mathbb{T}$. Therefore $\mathrm{Tr}(\log \tilde{w} - \log w) \longrightarrow 0$ almost everywhere as $\eta \longrightarrow 0$. But this sequence of functions is nonnegative and decreasing and therefore dominated in L^1, since $\log w$ and $\log \tilde{w}$ are integrable for all $\eta > 0$. So by the dominated convergence theorem, the integral

$$\langle \log \tilde{w} - \log w, 1 \rangle = \int_{\mathbb{T}} \mathrm{Tr}(\log \tilde{w} - \log w) \, \mathrm{d}\mu$$

tends to zero as $\eta \longrightarrow 0$. Choose $\eta \leq R$ such that this integral is as most $\varepsilon^2/18R$. This is an upper bound for the left hand sides of (2.20) and (2.21), and therefore establishes inequalities (2.16) and (2.17).

Now that η and R are fixed, we shall choose N. By hypothesis, $w_r \longrightarrow w$ in L^1 as $r \longrightarrow \infty$ and by Proposition 2.2.2, $\det \rho_r(0) \longrightarrow \det \rho(0)$ as $r \longrightarrow \infty$. So the left hand sides of (2.18) and (2.19) tend to zero as $r \longrightarrow \infty$. This allows us to choose $N \in \mathbb{N}$ such that inequalities (2.18) and (2.19) are satisfied for all $r \geq N$.

By the triangle inequality, we have

$$\frac{\|w_r - \tilde{w}\|_1}{\eta} \leq \frac{\|w_r - w\|_1}{\eta} + \frac{\|w - \tilde{w}\|_1}{\eta} ,$$

and also

$$\left|1-\left(\frac{\det\rho_r(0)}{\det\tilde{\rho}(0)}\right)^{1/n}\right| \;\le\; \left|1-\left(\frac{\det\rho(0)}{\det\tilde{\rho}(0)}\right)^{1/n}\right| \;+\; \left|\frac{[\det\rho_r(0)]^{1/n}-[\det\rho(0)]^{1/n}}{[\det\tilde{\rho}(0)]^{1/n}}\right| .$$

So from the upper bounds provided by (2.16) to (2.19), we obtain the following inequalities upon substitution into (2.12) and (2.13):

$$\|(\rho-\tilde{\rho})P\|_2^2 \;\le\; \frac{\varepsilon^2}{18}\,(1+1)$$

$$\|(\rho_r-\tilde{\rho})P\|_2^2 \;\le\; \frac{\varepsilon^2}{18}\,(1+1+1+1)\,, \qquad \text{for all } r\ge N$$

Inequality (2.19) implies that $\|w_r-w\|_1\le\varepsilon^2/18$ for all $r\ge N$, since η is chosen to be at most R. So from the above inequalities, together with (2.15), we find upon sustitution into (2.14) that

$$\begin{aligned}\|\rho_r-\rho\|_2^2 \;&\le\; \frac{\varepsilon^2}{18}\times\left((\sqrt{1+1}+\sqrt{1+1+1+1}\,)^2+(\sqrt{1+1}+\sqrt{1}\,)^2\right)\\ &=\; \frac{\varepsilon^2}{18}\times 3\left(1+\sqrt{2}\right)^2,\end{aligned}$$

for all $r\ge N$. Therefore

$$\|\rho_r-\rho\|_2 \;\le\; \left(\frac{1+\sqrt{2}}{\sqrt{6}}\right)\varepsilon \;<\; \varepsilon\,,$$

for all $r\ge N$. This shows that $\rho_r\to\rho$ in H^2 as $r\to\infty$, proving *2.* □

We make two remarks about this theorem. Firstly, unlike Lemmas 2.3.1 and 2.3.2, this result does not give quantitive information about the proximity of two spectral factors in terms of their spectral densities. This is because the values of R and η needed to establish inequalities (2.15) to (2.17) depend upon the distribution of the large values of $\operatorname{Tr} w$ and $-\log\det w$. Even if we constrain the norms of w and $\log\det w$, we can still find spectral densities w, such that R and $1/\eta$ need to be arbitrarily large.

Secondly, for an L^1 convergent sequence of spectral densities, we need only show that their spectral factors converge in measure, in order to establish their H^2 convergence. This is because of their uniform square-integrability. We make use of this idea in the next result, where we extend Theorem 2.1.3 as follows:

Theorem 2.4.1 *Let $p \in [1, \infty)$ and let $w_r, w \in L^p(\mathcal{L}(\mathbb{C}^n))$ be spectral densities, for all $r \in \mathbb{N}$. The following are equivalent:*

1. *$w_r \to w$ in L^p as $r \to \infty$ and $\{\log\det w_r : r \in \mathbb{N}\}$ is uniformly integrable.*
2. *$\Phi(w_r) \to \Phi(w)$ in H^{2p} as $r \to \infty$.*

Proof

2 ⇒ 1. Let ρ, ρ_r denote the canonical spectral factors $\Phi(w), \Phi(w_r)$ respectively, for each $r \in \mathbb{N}$. As in the proof of Theorem 2.1.3, we have the inequality:

$$\|w_r - w\|_p \;\leq\; \tfrac{1}{2}\|(\rho_r + \rho)^*(\rho_r - \rho)\|_p \;+\; \tfrac{1}{2}\|(\rho_r - \rho)^*(\rho_r + \rho)\|_p \,.$$

Let $q \in (1, \infty]$ be the index conjugate to p. For any functions $f, g \in L^{2p}(\mathcal{L}(\mathbb{C}^n))$, we have by Hölder's inequality:

$$\|fg\|_p \;=\; \sup_{\|h\|_q \leq 1} \|fgh\|_1 \;\leq\; \sup_{\|h\|_q \leq 1} \|f\|_{2p}\,\|g\|_{2p}\,\|h\|_q \;=\; \|f\|_{2p}\,\|g\|_{2p}\,,$$

and so we find that

$$\|w_r - w\|_p \;\leq\; \|\rho_r + \rho\|_{2p}\,\|\rho_r - \rho\|_{2p} \;\to\; 0 \quad \text{as } r \to \infty\,.$$

Now we apply Proposition 2.2.2 to find that $\{\log\det w_r : r \in \mathbb{N}\}$ is uniformly integrable, establishing *1*.

1 ⇒ 2. Since $w_r \to w$ in L^1, we find by Theorem 2.1.3 that $\rho_r \to \rho$ in H^2, and in particular, $\rho_r - \rho \to 0$ in measure as $r \to \infty$.

Let $f_r : \mathbb{T} \to [0, \infty)$ be given by

$$f_r(z) \;=\; \|\rho_r(z) - \rho(z)\|_{2p}^{2p}\,, \quad \text{for all } z \in \mathbb{T},\, r \in \mathbb{N}\,.$$

By the triangle inequality,

$$
\begin{aligned}
f_r(z) &\leq (\|\rho_r(z)\|_{2p} + \|\rho(z)\|_{2p})^{2p} \\
&\leq 2^p (\|\rho_r(z)\|_{2p}^2 + \|\rho(z)\|_{2p}^2)^p \\
&\leq 2^p (\|w_r(z) - w(z)\|_p + 2\|w(z)\|_p)^p \\
&\leq 2^p \cdot 3^{p-1}(\|w_r(z) - w(z)\|_p^p + 2\|w(z)\|_p^p), \quad \text{for all } z \in \mathbb{T},
\end{aligned}
$$

and this upper bound is uniformly integrable over $r \in \mathbb{N}$, since $w_r - w \to 0$ in L^p as $r \to \infty$. Now since $(f_r)_{r\in\mathbb{N}}$ is uniformly integrable and convergent in measure to 0, it also tends to 0 in L^1. This implies that $\rho_r \to \rho$ in H^{2p} as $r \to \infty$, establishing *2.* □

Note that the set $\{\log\det w_r : r \in \mathbb{N}\}$ in the first condition of the above theorem, will be uniformly integrable automatically if it is bounded in L^s, for any $s > 1$. Uniform integrability also follows when $\{\log w_r : r \in \mathbb{N}\}$ is bounded in L^s, or when the (convergent) spectral densities have inverses uniformly bounded in L^∞. Thus any of these conditions, together with L^p convergence of the w_r to some spectral density w, will be sufficient for $\Phi(w_r) \to \Phi(w)$ in H^{2p} as $r \to \infty$, for any $p \in [1, \infty)$.

In light of these remarks, the following is immediate from Theorem 2.4.1:

Corollary 2.4.2 *Let $w_r \in \mathrm{L}^1(\mathcal{L}(\mathbb{C}^n))$ be a spectral density for all $r \in \mathbb{N}$. Suppose that $w_r \to w$ in L^∞, where $w, w^{-1} \in \mathrm{L}^\infty(\mathcal{L}(\mathbb{C}^n))$. Then $\Phi(w_r) \to \Phi(w)$ in L^p for all $p \in [1, \infty)$.*

We also note that L^1 boundedness of the set $\{\log w_r : r \in \mathbb{N}\}$ will not guarantee uniform integrability of the functions $(\log\det w_r)_{r\in\mathbb{N}}$, even if the sequence of spectral densities $(w_r)_{r\in\mathbb{N}}$ is uniformly bounded (in L^∞) and convergent to 1 almost everywhere. So by Theorem 2.1.3, their canonical spectral factors will not converge to 1 in H^2. Since they are uniformly bounded, they do not even converge to 1 in measure.

2.5 Concluding remarks

Lemma 2.3.1 is the key ingredient used in the proof of Theorem 2.1.3. This it what establishes the $L^1 \to H^2$ continuity at the constant density 1. Thus the simple variance argument used to prove the lemma is perhaps the most important idea involved in the proof of the theorem. Much of the work needed to complete the proof is simply overcoming the technical difficulties involved in the translation from continuity at 1, to continuity at an arbitary spectral density.

Theorem 2.4.1 and Proposition 2.2.2 complement each other. Together they provide many conditions equivalent to H^{2p} convergence of a sequence of canonical spectral factors. (For example, L^1 convergence of the spectral densities combined with either uniform integrability of their logarithms or the trace of their logarithms, or – via a suitable conformal mapping – convergence of the spectral factors at a given point in the disc $\mathbb{D}$.)

They also suggest that it is the L^1 convergence of the *logarithms* of the spectral densities, and not the densities themselves, which is important for the convergence of their respective spectral factors. Perhaps L^1 log-convergence alone, of any given sequence of spectral densities, is a sufficient condition for their canonical spectral factors to converge in measure.

Chapter 3

The Douglas-Rudin problem

In this chapter we solve the noncommutative Douglas-Rudin problem, showing that any log-integrable essentially bounded square matrix-valued function f can be written in the form h^*g, where h and g lie in H^∞. Extensions to other L^p spaces, with norm bounds on the factors of f, are also provided.

The outline of this chapter is as follows: The next two sections form the bulk of the chapter. In Section 3.1, we obtain good L^2 approximation to unitary-valued functions using outer functions. In Section 3.2, we obtain good L^∞ approximation to unitary-valued functions using inner functions. In order to use these results in the iterative construction outlined in Chapter 1, we provide quantitive information concerning the size of the approximants required.

In Section 3.3, we prove our main results, beginning with the factorization of a general unitary-valued function on $\mathbb{T}$ into the form h^*g, for $g, h \in H^\infty$. We show, moreover, that the uniform norms of g, h, g^{-1} and h^{-1} can simultaneously be made arbitrarily close to 1. This is used to establish the main result of this chapter, Theorem 3.3.2, which solves the Douglas-Rudin problem for bounded matrix-valued functions, as well as a generalization to functions in $L^p(\mathcal{L}(\mathbb{C}^n))$ for all indices $1 \le p \le \infty$. We also obtain bounds on the norms of the factors, which are subsequently shown to be sharp. We then make some concluding remarks in Section 3.4.

3.1 L^2 approximation by outer functions

3.1.1 Continuity of Φ at 1

In order to find the outer factors used for the L^2 approximation of a given unitary-valued function, we will make use of spectral factorization. It will therefore be very helpful for us to have some information about the L^2 continuity of the spectral factorization mapping Φ. For this we have the following result:

Theorem 3.1.1 *Let $w \in \mathrm{L}^\infty(\mathcal{L}(\mathbb{C}^n))$ be a spectral density and let w^{-1} denote its almost everywhere defined inverse. Then $\Phi(w)$ satisfies:*

$$\|\Phi(w) - I\|_2 \;\leq\; \|w\|_\infty(1 + \|w\|_\infty^{1/2})\|I - w^{-1}\|_2 .$$

Proof We may suppose that $\|w^{-1}\|_2 < \infty$, since otherwise there is nothing to prove. Let $\rho = \Phi(w) \in \mathrm{H}^\infty(\mathcal{L}(\mathbb{C}^n))$. Then $\|\rho^{-1}\|_2 < \infty$.

Since ρ is outer, this implies that $\rho^{-1} \in \mathrm{H}^2(\mathcal{L}(\mathbb{C}^n))$. Indeed, for any vector $v \in \mathbb{C}^n$, there exists a sequence $(q_k)_{k=1}^\infty$ of $\mathbb{C}^n$-valued polynomials, such that $\rho q_k \to v$ in H^2, as $k \to \infty$. Multiplying by ρ^{-1}, we find that $q_k = \rho^{-1}\rho q_k \to \rho^{-1}v$ in L^1, as $k \to \infty$. So $\rho^{-1}v \in \mathrm{H}^1(\mathbb{C}^n)$, since $\mathrm{H}^1(\mathbb{C}^n)$ is closed in $\mathrm{L}^1(\mathbb{C}^n)$. But $\mathrm{H}^1(\mathbb{C}^n) \cap \mathrm{L}^2(\mathbb{C}^n) = \mathrm{H}^2(\mathbb{C}^n)$, so $\rho^{-1}v \in \mathrm{H}^2(\mathbb{C}^n)$ for all $v \in \mathbb{C}^n$, which implies that $\rho^{-1} \in \mathrm{H}^2(\mathcal{L}(\mathbb{C}^n))$.

Now we have:

$$\begin{aligned} \rho(0)\cdot\rho\cdot(I - w^{-1}) &= \rho(0)\cdot\rho \;-\; \rho(0)\cdot\rho\rho^{-1}(\rho^{-1})^* \\ &= \rho(0)\cdot\rho \;-\; (\rho^{-1}\cdot\rho(0))^* . \end{aligned}$$

But $\rho^{-1}\cdot\rho(0) \in \mathrm{H}^2(\mathcal{L}(\mathbb{C}^n))$, and since $\|\rho\|_\infty, \|w^{-1}\|_2 < \infty$, the above function lies in $\mathrm{L}^2(\mathcal{L}(\mathbb{C}^n))$. Taking its orthogonal projection onto $\mathrm{H}^2(\mathcal{L}(\mathbb{C}^n))$, we find that:

$$\begin{aligned} \mathrm{P}_{\mathrm{H}^2(\mathcal{L}(\mathbb{C}^n))}\Big(\rho(0)\cdot\rho\cdot(I - w^{-1})\Big) &= \rho(0)\cdot\rho \;-\; (\rho^{-1}(0)\rho(0))^* \\ &= \rho(0)\cdot\rho \;-\; I . \end{aligned}$$

Therefore,

$$\begin{aligned}
\|\rho(0)\cdot\rho - I\|_2 &= \left\|\mathrm{P}_{\mathrm{H}^2(\mathcal{L}(\mathbb{C}^n))}\Big(\rho(0)\cdot\rho\cdot(I-w^{-1})\Big)\right\|_2 \\
&\leq \|\rho(0)\cdot\rho\cdot(I-w^{-1})\|_2 \\
&\leq \|\rho\cdot\rho(0)\|_\infty\|I-w^{-1}\|_2 \\
&\leq \|w\|_\infty\|I-w^{-1}\|_2 .
\end{aligned}$$

Averaging over the unit circle, we find that $\|\rho(0)^2 - I\|_2$ is also bounded by $\|w\|_\infty\|I-w^{-1}\|_2$. Hence

$$\begin{aligned}
\|\rho(0) - I\|_2 &\leq \|\rho(0)^2 - I\|_2\cdot\|(\rho(0)+I)^{-1}\|_\infty \\
&\leq \|\rho(0)^2 - I\|_2 \\
&\leq \|w\|_\infty\|I-w^{-1}\|_2 ,
\end{aligned}$$

since $\rho(0)$ is positive. So we have:

$$\begin{aligned}
\|\rho - I\|_2 &\leq \|\rho\cdot\rho(0) - I\|_2 + \|\rho(\rho(0)-I)\|_2 \\
&\leq \|w\|_\infty\|I-w^{-1}\|_2 + \|\rho\|_\infty\cdot\|w\|_\infty\|I-w^{-1}\|_2 \\
&\leq \|w\|_\infty(1+\|w\|_\infty^{1/2})\|I-w^{-1}\|_2 ,
\end{aligned}$$

since $\|\rho\|_\infty = \|w\|_\infty^{1/2}$. This completes the proof. □

Note that in Chapter 2 (see also [3]), we established the $L^1 \to H^2$ continuity of the spectral factorization mapping under certain mild conditions. However, the above $L^2 \to H^2$ continuity result may be of some interest in its own right, as it provides information about the Lipschitz continuity of Φ. This is more useful in the context of the present chapter.

3.1.2 Some simple lemmas

The following result will be of some utility:

Lemma 3.1.2 *Let $p \in [1, \infty]$ and let $a, b \in \mathrm{L}^p(\mathcal{L}(\mathbb{C}^n))$ be self-adjoint matrix-valued functions. Then we have:*

$$\|e^{ia} - e^{ib}\|_p \leq \|a - b\|_p .$$

Proof Let k be a natural number. Then we have the telescoping sum:

$$e^{ia} - e^{ib} = \sum_{l=1}^{k} e^{ia(k-l)/k}(e^{ia/k} - e^{ib/k})e^{ib(l-1)/k}$$

Taking norms of each side, we find:

$$\begin{aligned} \|e^{ia} - e^{ib}\|_p &\leq \sum_{l=1}^{k} \|e^{ia(k-l)/k}(e^{ia/k} - e^{ib/k})e^{ib(l-1)/k}\|_p \\ &= k\|e^{ia/k} - e^{ib/k}\|_p , \end{aligned} \tag{3.1}$$

since the terms $e^{ia(k-l)/k}$ and $e^{ib(l-1)/k}$ are unitary-valued for all $l \in \{1, \ldots, k\}$.

Let $f : \mathbb{R} \to \mathbb{C}$ be the bounded real analytic function given by:

$$f(x) = \begin{cases} (e^{ix} - 1)/x - i & x \neq 0 \\ 0 & x = 0 , \end{cases}$$

for all $x \in \mathbb{R}$. Now for all natural numbers k, define $f(a/k)$ to be the matrix-valued function on $\mathbb{T}$ given pointwise by the continuous functional calculus, so that $f(a/k)(z) = f(a(z)/k)$ for all $z \in \mathbb{T}$. Then $(f(a/k))_{k=1}^{\infty}$ is a sequence bounded in $\mathrm{L}^\infty(\mathcal{L}(\mathbb{C}^n))$, convergent to zero pointwise on $\mathbb{T}$. Since $\|a(z)f(a/k)(z)\|_p \leq \|a(z)\|_p \, \|f(a/k)(z)\|_\infty$ for all $z \in \mathbb{T}$ and $k \in \mathbb{N}$, we therefore find that:

$$\|k(e^{ia/k} - 1) - ia\|_p = \|af(a/k)\|_p \to 0 ,$$

as $k \to \infty$.

The above argument also works for b in place of a. We therefore find that:

$$\lim_{k \to \infty} \|k(e^{ia/k} - e^{ib/k})\|_p = \|a - b\|_p ,$$

and so by inequality (3.1) the result follows. □

The following result is taken from [8, p. 181]. It gives a slight strengthening of Lemma 1 used in Bourgain's construction [6], which underpins the creation of good scalar outer function approximations to unimodular errors.

Proposition 3.1.3 *Let $f \in \mathrm{L}^\infty$ and let $\eta \in (0, 1/2]$. Then there exist functions $g^+ \in \mathrm{H}^\infty$ and $g^- \in \overline{\mathrm{H}_0^\infty}$ such that:*

1. $\|f - (g^+ + g^-)\|_2 \leq \eta \|f\|_\infty$;

2. $\|g^+\|_\infty + \|g^-\|_\infty \leq K \log(1/\eta) \|f\|_\infty$,

for some absolute constant $K > 0$.

Throughout the rest of this chapter, we shall keep K fixed, with it chosen to satisfy the above proposition statement.

We use the above scalar result to derive a matricial analogue, given as follows:

Lemma 3.1.4 *Let $a \in \mathrm{L}^\infty(\mathcal{L}(\mathbb{C}^n))$ with $a = a^*$ a.e. and let $\varepsilon \in (0, n/2]$. Then there exists $b \in \mathrm{H}^\infty(\mathcal{L}(\mathbb{C}^n))$ with $b(0) + b(0)^* = 0$, such that:*

1. $\|a - (b - b^*)/(2i)\|_2 \leq \varepsilon \|a\|_\infty$;

2. $\|b\|_\infty \leq 2\sqrt{2}\, Kn \log(n/\varepsilon) \|a\|_\infty$.

Proof We have $a = (a_{jk})_{j,k=1}^n$, where $a_{jk} = \overline{a_{kj}} \in \mathrm{L}^\infty$ for all $j, k \in \{1, \ldots, n\}$.

For each $1 \leq j < k \leq n$, we may apply Proposition 3.1.3 with $f = a_{jk}$ and $\eta = \varepsilon/n$, to obtain $g^+, \overline{g^-} \in \mathrm{H}^\infty(\mathcal{L}(\mathbb{C}^n))$ satisfying the proposition statement. Letting $b_{jk} = 2i(g^+ - g^+(0)/2)$ and $b_{kj} = 2i(\overline{g^-} + \overline{g^+(0)}/2)$ for each such j and k, this implies that:

$$\|a_{jk} - (b_{jk} - \overline{b_{kj}})/(2i)\|_2 \leq (\varepsilon/n)\|a_{jk}\|_\infty \tag{3.2}$$

$$\|b_{jk}\|_\infty + \|b_{kj}\|_\infty \leq 4K \log(n/\varepsilon) \|a_{jk}\|_\infty \tag{3.3}$$

$$b_{jk}(0) + \overline{b_{kj}(0)} = 0 \tag{3.4}$$

for all $j, k \in \{1, \ldots, n\}$ with $j \neq k$. For the case $j = k \in \{1, \ldots, n\}$, we may apply Proposition 3.1.3 in the same way as above, but this time let:

$$b_{jj} = i(g^+ - g^+(0)/2) + i(\overline{g^-} + \overline{g^+(0)}/2).$$

Then it is easily seen that (3.2), (3.3) and (3.4) still hold for these choices of b_{jk}.

Now since the scalar functions b_{jk} all lie in H^∞, they form the components of a matrix-valued function $b = (b_{jk})_{j,k=1}^n \in \mathrm{H}^\infty(\mathcal{L}(\mathbb{C}^n))$. The set of equations given by (3.4) then imply that $b(0) + b(0)^* = 0$, as required for b to satisfy the statement of the lemma.

Finally, to show that assertions 1 and 2 hold for this choice of b, we have:

$$\begin{aligned} \|a - (b - b^*)/(2i)\|_2^2 &= \sum_{j=1}^n \sum_{k=1}^n \|a_{jk} - (b_{jk} - \overline{b_{kj}})/(2i)\|_2^2 \\ &\leq (\varepsilon/n)^2 \sum_{j=1}^n \sum_{k=1}^n \|a_{jk}\|_\infty^2 \leq \varepsilon^2 \|a\|_\infty^2, \end{aligned}$$

from the set of inequalities given by (3.2). This establishes assertion 1. We also have the estimate:

$$\begin{aligned} \|b\|_\infty^2 &\leq \sum_{j=1}^n \sum_{k=1}^n \|b_{jk}\|_\infty^2 \\ &\leq \sum_{j=1}^n \sum_{k=1}^n (\|b_{jk}\|_\infty + \|b_{kj}\|_\infty)^2/2 \\ &\leq 8(K \log(n/\varepsilon))^2 \sum_{j=1}^n \sum_{k=1}^n \|a_{jk}\|_\infty^2 \leq 8(Kn \log(n/\varepsilon))^2 \|a\|_\infty^2, \end{aligned}$$

from the set of inequalities given by (3.3). This establishes assertion 2. □

3.1.3 Outer function approximation results

We begin with the following preliminary result, which enables us to overcome the problems posed by the noncommutativity of matrix multiplication:

Lemma 3.1.5 *Let $b \in \mathrm{H}^\infty(\mathcal{L}(\mathbb{C}^n))$ with $b(0) + b(0)^* = 0$ and let $\phi \in \mathrm{L}^\infty(\mathcal{L}(\mathbb{C}^n))$ almost everywhere unitary.*

Then there exist outer functions $\rho, \tau \in \mathrm{H}^\infty(\mathcal{L}(\mathbb{C}^n))$ *such that:*

1. $(\tau^{-1})^* \phi \, \rho^{-1}$ *is almost everywhere unitary;*

2. $\|(\tau^{-1})^* \phi \, \rho^{-1} \; - \; I\|_2 \;\leq\; 319\sqrt{n}\, \|b\|_\infty^2 + 256\, \|\phi - \psi\|_2;$

3. $\|\rho\|_\infty, \|\tau\|_\infty, \|\rho^{-1}\|_\infty, \|\tau^{-1}\|_\infty \;\leq\; e^{\|b\|_\infty/2};$

4. $\|\rho - I\|_2, \|\tau - I\|_2, \|\rho^{-1} - I\|_2, \|\tau^{-1} - I\|_2 \;\leq\; 14\sqrt{n}\, \|b\|_\infty,$

where $\psi = e^{(b-b^*)/2} \in \mathrm{L}^\infty(\mathcal{L}(\mathbb{C}^n))$ *a.e. unitary.*

Proof We may assume without loss of generality that $\|b\|_\infty \leq 1$, for otherwise we may set $\rho = \tau \equiv I$, and then assertions 1, 2, 3 and 4 are satisfied trivially.

We shall begin by establishing assertions 3 and 4 of the lemma, as follows:

Set $\rho = e^{b/2} \in \mathrm{H}^\infty(\mathcal{L}(\mathbb{C}^n))$, with inverse $e^{-b/2} \in \mathrm{H}^\infty(\mathcal{L}(\mathbb{C}^n))$.

Set $\tau = \rho(0)^* \cdot \Phi(\phi\rho^{-1}(\rho^{-1})^*\phi^*)$.

Now $\rho(0)^*\rho(0) = \exp(b(0) + b(0)^*) = I$, and hence $\rho(0)$ and $\rho(0)^*$ are unitary. Therefore ρ^{-1} and τ satisfy:

$$\begin{aligned} \|\tau\|_\infty &= \|\Phi(\phi\rho^{-1}(\rho^{-1})^*\phi^*)\|_\infty \\ &= \|\phi\rho^{-1}(\rho^{-1})^*\phi^*\|_\infty^{1/2} \\ &= \|\rho^{-1}\|_\infty \\ &\leq e^{\|b\|_\infty/2}, \end{aligned}$$

and ρ and τ^{-1} satisfy:

$$\begin{aligned} \|\tau^{-1}\|_\infty &= \|\Phi(\phi\rho^{-1}(\rho^{-1})^*\phi^*)^{-1}\|_\infty \\ &= \|\phi\rho^*\rho\phi^*\|_\infty^{1/2} \\ &= \|\rho\|_\infty \\ &\leq e^{\|b\|_\infty/2}. \end{aligned}$$

This establishes assertion 3 of the lemma.

Now let $w = \phi\rho^{-1}(\rho^{-1})^*\phi^*$. Then $\rho(0)\cdot\tau$ is the canonical spectral factor of w, so by Theorem 3.1.1 we have:

$$\|\rho(0)\tau - I\|_2 \;\leq\; \|w\|_\infty(1+\|w\|_\infty^{1/2})\|I - w^{-1}\|_2\,. \tag{3.5}$$

Since $(\tau^{-1}\cdot\rho(0)^{-1})^{\mathrm{T}}$ is the canonical spectral factor of $(w^{-1})^{\mathrm{T}}$, we also have the following inequality by Theorem 3.1.1:

$$\|\tau^{-1}\rho(0)^{-1} - I\|_2 \;\leq\; \|w^{-1}\|_\infty(1+\|w^{-1}\|_\infty^{1/2})\|I - w\|_2\,. \tag{3.6}$$

The last term of the right hand side of (3.5) satisfies:

$$\begin{aligned} \|I - w^{-1}\|_2 \;&=\; \|I - \phi\rho^*\rho\phi^*\|_2 \;=\; \|I - \rho^*\rho\|_2 \\ &\leq\; \|\rho^*\|_\infty\|I-\rho\|_2 \;+\; \|I-\rho^*\|_2 \\ &=\; (\|\rho^*\|_\infty + 1)\|I-\rho\|_2 \\ &=\; (\|w\|_\infty^{1/2}+1)\|\rho - I\|_2\,. \end{aligned} \tag{3.7}$$

Similarly,

$$\|I - w\|_2 \;\leq\; (\|w^{-1}\|_\infty^{1/2}+1)\|\rho^{-1} - I\|_2\,. \tag{3.8}$$

Therefore, from inequalities (3.5) and (3.7) we obtain:

$$\begin{aligned} \|\tau - I\|_2 \;&\leq\; \|\rho(0)\tau - I\|_2 \;+\; \|\tau\|_\infty\|\rho(0) - I\|_2 \\ &\leq\; \|\rho(0)\tau - I\|_2 \;+\; \|w\|_\infty^{1/2}\,\|\rho - I\|_2 \\ &\leq\; \left(\|w\|_\infty(1+\|w\|_\infty^{1/2})^2 \;+\; \|w\|_\infty^{1/2}\right)\|\rho - I\|_2 \end{aligned} \tag{3.9}$$

Similarly, from inequalities (3.6) and (3.7) we obtain:

$$\|\tau^{-1} - I\|_2 \;\leq\; \left(\|w^{-1}\|_\infty(1+\|w^{-1}\|_\infty^{1/2})^2 \;+\cdot\, \|w^{-1}\|_\infty^{1/2}\right)\|\rho^{-1} - I\|_2\,. \tag{3.10}$$

The last term of the right hand side of (3.9) satisfies:

$$\begin{aligned} n^{-1/2}\,\|\rho - I\|_2 \;&\leq\; \|e^{b/2} - I\|_\infty \\ &=\; \|b/2 \;+\; (b/2)^2/2! \;+\; (b/2)^3/3! \;+\; \cdots\|_\infty \\ &\leq\; (\|b/2\|_\infty \;+\; \|b/2\|_\infty^2/2! \;+\; \|b/2\|_\infty^3/3! \;+\; \cdots) \\ &=\; (e^{\|b\|_\infty/2} - 1) \\ &\leq\; (e^{1/2} - 1)\,\|b\|_\infty\,, \end{aligned} \tag{3.11}$$

and similarly we have:

$$\|\rho^{-1} - I\|_2 \;\leq\; \sqrt{n}\,(e^{1/2} - 1)\,\|b\|_\infty . \tag{3.12}$$

Now since $\|w\|_\infty, \|w^{-1}\|_\infty \leq e^{\|b\|_\infty} \leq e$, inequalities (3.9) to (3.12) imply that:

$$\begin{aligned}\|\tau - I\|_2, \|\rho - I\|_2, \|\tau^{-1} - I\|_2, \|\rho^{-1} - I\|_2 \;&\leq\; (e\cdot(1+e^{1/2})^2 + e^{1/2}) \times \\ &\qquad (e^{1/2}-1)\sqrt{n}\,\|b\|_\infty \\ &\leq\; 14\sqrt{n}\|b\|_\infty ,\end{aligned}$$

establishing assertion 4 of the lemma.

By the unitarity of $\rho(0)$, we have:

$$\begin{aligned}\left((\tau^{-1})^*\phi\,\rho^{-1}\right)^*\left((\tau^{-1})^*\phi\,\rho^{-1}\right) &= (\rho^{-1})^*\,\phi^*(\tau^{-1}\cdot(\tau^{-1})^*)\phi\,\rho^{-1} \\ &= (\rho^{-1})^*\,\phi^*\left(\phi\rho^{-1}(\rho^{-1})^*\phi^*\right)^{-1}\phi\,\rho^{-1} \quad \text{a.e.} \\ &= I \quad \text{a.e.}\end{aligned}$$

So $(\tau^{-1})^*\phi\,\rho^{-1}$ is almost everywhere unitary, establishing assertion 1.

Finally, to establish assertion 2 of the lemma, let $\tilde{\tau} = \tau\rho$ and let $\tilde{w} = \tilde{\tau}^*\tilde{\tau}$. Then we have:

$$\begin{aligned}\|I - \tilde{w}^{-1}\|_2 &= \|I - \rho^{-1}\phi\rho^*\rho\phi^*(\rho^{-1})^*\|_2 \\ &\leq \|\rho^{-1}\|_\infty\|\rho - \phi\rho^*\|_2 + \|\rho^{-1}\phi\rho^*\|_\infty\|\rho^* - \rho\phi^*\|_2\,\|(\rho^{-1})^*\|_\infty \\ &= \left(\|\rho^{-1}\|_\infty + \|\rho^{-1}\phi\rho^*\|_\infty\|(\rho^{-1})^*\|_\infty\right)\cdot\|\phi\rho^* - \rho\|_2 \\ &\leq (e^{\|b\|_\infty/2} + e^{3\|b\|_\infty/2})\cdot\|\phi\rho^* - \rho\|_2 \\ &\leq (e^{1/2} + e^{3/2})\cdot\|\phi\rho^* - \rho\|_2 .\end{aligned}$$

Now $\tilde{\tau}(0) = \tau(0)\rho(0)$ is positive, so $\tilde{\tau}$ is the canonical spectral factor of $\tilde{w}$. Therefore, by another application of Theorem 3.1.1, we obtain:

$$\begin{aligned}\|\tilde{\tau} - I\|_2 \;&\leq\; \|\tilde{w}\|_\infty(1 + \|\tilde{w}\|_\infty^{1/2})\|I - \tilde{w}^{-1}\|_2 \\ &\leq\; e^2(1+e)(e^{1/2} + e^{3/2})\cdot\|\phi\rho^* - \rho\|_2 ,\end{aligned}$$

since $\|\tilde{w}\|_\infty \leq e^{2\|b\|_\infty} \leq e^2$.

Now since $(\tau^{-1})^*\phi\,\rho^{-1}$ is almost everywhere unitary, we have:

$$\begin{aligned}\|(\tau^{-1})^*\phi\,\rho^{-1} - I\|_2 &= \|\tau\phi\rho^* - I\|_2 \\ &\leq \|\tau\|_\infty\,\|\phi\rho^* - \rho\|_2 + \|\tilde{\tau} - I\|_2 \\ &\leq \left(e^{1/2} + e^2(1+e)(e^{1/2}+e^{3/2})\right)\cdot\|\phi\rho^* - \rho\|_2 . \quad (3.13)\end{aligned}$$

We estimate the last term of the above bound as follows:

$$\begin{aligned}\|\phi\rho^* - \rho\|_2 &= \|\phi\cdot e^{b^*/2} - e^{b/2}\|_2 \\ &= \Big\|\phi\cdot(e^{b^*/2} - I - b^*/2) + (\phi-\psi)(I + b^*/2) + \\ &\qquad (\psi - I - (b-b^*)/2)(I + b^*/2) + \\ &\qquad ((b-b^*)/2)(b^*/2) - (e^{b/2} - I - b/2)\Big\|_2 \\ &\leq \|\phi\|_2\,\|e^{b^*/2} - I - b^*/2\|_\infty + \|\phi-\psi\|_2\,\|I + b^*/2\|_\infty + \\ &\qquad \|\psi - I - (b-b^*)/2\|_\infty\|I + b^*/2\|_\infty\|I\|_2 + \\ &\qquad \|(b-b^*)/2\|_\infty\|b^*/2\|_\infty\|I\|_2 + \|e^{b/2} - I - b/2\|_\infty\|I\|_2 \\ &\leq \sqrt{n}\cdot\left(e^{\|b\|_\infty/2} - I - \|b\|_\infty/2\right) + \|\phi-\psi\|_2\cdot(1 + \|b\|_\infty/2) + \\ &\qquad \sqrt{n}\cdot\left(e^{\|b-b^*\|_\infty/2} - 1 - \|b-b^*\|_\infty/2\right)\cdot(1 + \|b\|_\infty/2) + \\ &\qquad \sqrt{n}\cdot\|b-b^*\|_\infty\,\|b\|_\infty/4 + \sqrt{n}(e^{\|b\|_\infty/2} - 1 - \|b\|_\infty/2) .\end{aligned}$$

From the last expression, we obtain the inequality:

$$\begin{aligned}\|\phi\rho^* - \rho\|_2 &\leq \sqrt{n}\cdot\Big[(e^{1/2} - 3/2) + (3/2)(e-2) + 1/2 + (e^{1/2} - 3/2)\Big]\cdot\|b\|_\infty^2 \\ &\qquad + (3/2)\|\phi-\psi\|_2 \\ &= [2(e^{1/2} - 3/2) + (3/2)(e-2) + 1/2]\cdot\sqrt{n}\cdot\|b\|_\infty^2 \\ &\qquad + (3/2)\|\phi-\psi\|_2 .\end{aligned}$$

Substituting this estimate into inequality (3.13), we obtain:

$$\begin{aligned}\|(\tau^{-1})^*\phi\,\rho^{-1} - I\|_2 &\leq k[2(e^{1/2} - 3/2) + (3/2)(e-2) + 1/2]\cdot\sqrt{n}\cdot\|b\|_\infty^2 \\ &\qquad + (3k/2)\|\phi-\psi\|_2 ,\end{aligned}$$

where $k = e^2(1+e)(e^{1/2}+e^{3/2})$. The above absolute constants are less than 319 and 256 respectively, and this establishes assertion 2. □

We are now in a position to apply Lemma 3.1.4, in order to establish the main result of this section, given as follows:

Lemma 3.1.6 *Let* $\phi \in \mathrm{L}^\infty(\mathcal{L}(\mathbb{C}^n))$ *almost everywhere unitary.*

Then there exist outer functions $\rho, \tau \in \mathrm{H}^\infty(\mathcal{L}(\mathbb{C}^n))$ *such that:*

1. $(\tau^{-1})^*\phi\,\rho^{-1}$ *is almost everywhere unitary;*

2. $\|(\tau^{-1})^*\phi\,\rho^{-1} - I\|_2 \leq K_1\|\phi - I\|_\infty^{5/3}$;

3. $\|\rho\|_\infty, \|\tau\|_\infty, \|\rho^{-1}\|_\infty, \|\tau^{-1}\|_\infty \leq \exp(K_2\|\phi - I\|_\infty^{5/6})$;

4. $\|\rho - I\|_2, \|\tau - I\|_2, \|\rho^{-1} - I\|_2, \|\tau^{-1} - I\|_2 \leq K_3\|\phi - I\|_\infty^{5/6}$,

for some constants K_1, K_2 *and* K_3 *which depend only on* n.

Proof We may assume without loss of generality that $\|\phi - I\|_\infty > 0$, for otherwise we may set $\rho = \tau \equiv I$, and then assertions 1,2,3 and 4 are satisfied trivially.

Let $a = \operatorname{Arg}\phi$, given pointwise by the continuous functional calculus, where Arg is the principal branch of the argument, taken to lie in the interval $(-\pi, \pi]$. Thus a is almost everywhere self-adjoint and $\|a\|_\infty \leq (\pi/2)\|\phi - I\|_\infty$.

Now apply Lemma 3.1.4 with this choice of a and with ε equal to $n\exp(-\|\phi - I\|_\infty^{-1/6})$, to obtain some $b \in \mathrm{H}^\infty(\mathcal{L}(\mathbb{C}^n))$ satisfying:

$$\begin{aligned} b(0) + b(0)^* &= 0 \\ \|a - (b-b^*)/(2i)\|_2 &\leq n\exp(-\|\phi - I\|_\infty^{-1/6}) \times (\pi/2)\|\phi - I\|_\infty \qquad (3.14) \\ \|b\|_\infty &\leq 2\sqrt{2}Kn\log\Big(n/\Big(n\exp(-\|\phi - I\|_\infty^{-1/6})\Big)\Big) \times \\ &\qquad (\pi/2)\|\phi - I\|_\infty \end{aligned}$$

Evaluating the last expression above, we find that:

$$\|b\|_\infty \;\leq\; \sqrt{2}\,K\pi n\,\|\phi - I\|_\infty^{5/6}\,. \tag{3.15}$$

The exponential term may be estimated as follows:

$$\begin{aligned}
\exp\left(-\,\|\phi - I\|_\infty^{-1/6}\right) &= \left(\exp(\|\phi - I\|_\infty^{-1/6}/4)\right)^{-4} \\
&\leq \left(1 + \|\phi - I\|_\infty^{-1/6}/4\right)^{-4} \\
&< \left(\|\phi - I\|_\infty^{-1/6}/4\right)^{-4} \\
&= 256\,\|\phi - I\|_\infty^{-2/3}\,.
\end{aligned}$$

Hence,

$$\|a \,-\, (b - b^*)/(2i)\|_2 \;\leq\; 128\pi n\|\phi - I\|^{5/3}\,,$$

by inequality (3.14). Therefore setting $\psi = e^{(b-b^*)/2i}$, as in Lemma 3.1.5, we find upon application of Lemma 3.1.2 that:

$$\|\phi - \psi\|_2 \;\leq\; 128\pi n\|\phi - I\|^{5/3}\,.$$

Now we may apply Lemma 3.1.5 with this choice of b and ϕ, to obtain outer functions $\rho, \tau \in \mathrm{H}^\infty(\mathcal{L}(\mathbb{C}^n))$ satisfying assertion 1. Then from the above estimate, together with (3.15), we have:

$$\begin{aligned}
\|(\tau^{-1})^*\phi\,\rho^{-1} \,-\, I\|_2 &\leq 319\sqrt{n}\,\|b\|_\infty^2 + 32768\pi n\|\phi - I\|_\infty^{5/3} \\
&\leq K_1\|\phi - I\|_\infty^{5/3}\,,
\end{aligned}$$

where $K_1 \;=\; 32768\pi \cdot n \,+\, 638\pi^2 K^2 \cdot n^{5/2}$. Another application of inequality (3.15) gives:

$$\begin{aligned}
\|\rho\|_\infty, \|\tau\|_\infty, \|\rho^{-1}\|_\infty, \|\tau^{-1}\|_\infty &\leq \exp(\|b\|_\infty/2) \\
&\leq \exp(K_2\|\phi - I\|_\infty^{5/6})\,,
\end{aligned}$$

$$\begin{aligned}
\|\rho - I\|_2, \|\tau - I\|_2, \|\rho^{-1} - I\|_2, \|\tau^{-1} - I\|_2 &\leq 14\sqrt{n}\,\|b\|_\infty \\
&\leq K_3\|\phi - I\|_\infty^{5/6}\,,
\end{aligned}$$

where $K_2 = \pi K n/\sqrt{2}$ and $K_3 = 14\sqrt{2}\pi K n^{3/2}$. This establishes assertions 2,3 and 4, completing the proof. □

3.2 L^∞ approximation by inner functions

In order to obtain the main inner function approximation result of this section, we deal first with the scalar case. Then we shall show how a general unitary matrix-valued function in $L^\infty(\mathcal{L}(\mathbb{C}^n))$ can be broken down into unimodular pieces, to which the scalar result can be applied, leading on to the main result of the section.

For the scalar case, we shall make use of essentially the same result as Lemma 2 from Bourgain's construction [6]. A method of proof was outlined there but not explicitly given. In this section, we furnish a complete proof, beginning with the following technical result, based on the proof of [17, p. 430, Lemma 5.5]. (Note that throughout this section, we take the principal branch of the argument to lie in the interval $(-\pi, \pi]$.)

Lemma 3.2.1 *Let $\varepsilon \in (0, \pi)$, let $\delta, \eta > 0$, let $U \subset \mathbb{T}$ relatively open and let $g : U \to (-\pi, \pi]$ be a measurable function. Then there exist finite Blaschke products B_1 and B_2, having simple zeros, such that:*

$$\mu\{z \in U : |g(z) - \operatorname{Arg}(B_1(z)/B_2(z))| > \varepsilon\} < \eta,$$

and such that:

$$-\delta < \operatorname{Arg}(B_1(z)/B_2(z)) < \delta,$$

for all $z \in \mathbb{T} \setminus U$, and also such that $B_1(0)$ and $B_2(0)$ are positive and satisfy:

$$\log(1/B_1(0)), \log(1/B_2(0)) \leq \mu(U) \log(24\pi/\varepsilon).$$

Proof Consider first of all the case:

$$U = \{e^{i\omega} : -\theta < \omega < \theta\}$$

$$g \equiv \alpha,$$

for given $0 < \theta < \pi$ and $\varepsilon - \pi \leq \alpha \leq \pi - \varepsilon$. Then we may assume without loss of generality that $\theta \geq \pi\eta$, for otherwise we may set $B_1 = B_2 \equiv 1$ and then the above inequalities are satisfied trivially.

Fix some $0 < r < 1$, $N \in \mathbb{N}$ and $0 < \phi \leq \theta - \pi/N$, to be determined later. Then the zero sets of B_1 and B_2 are given by:

$$\begin{aligned} Z_1 &= \{r^{1/N}e^{2\pi ik/N} : k \in \mathbb{Z}/N\mathbb{Z}\} \cap \{r^{1/N}e^{i\omega} : -\phi < \omega < \phi\} \\ Z_2 &= e^{i\alpha/N}Z_1 = \{e^{i\alpha/N}z : z \in Z_1\}, \end{aligned}$$

and B_j is given by:

$$B_j(z) = \prod_{w \in \mathbb{Z}_j} \frac{|w|(w-z)}{w-|w|^2 z},$$

for all $z \in \overline{\mathbb{D}}$ and $j \in \{1,2\}$. Thus $B_1(0) = B_2(0) = r^{|Z_1|/N} \geq r^{\mu(U)}$, which implies that:

$$\log(1/B_1(0)), \log(1/B_2(0)) \leq \mu(U)\log(1/r). \tag{3.16}$$

Let $q : \mathbb{T} \to \mathbb{C}$ be the rational function given by:

$$\begin{aligned} q(z) &= \frac{(r-z^N)(e^{i\alpha}-rz^N)}{(1-rz^N)(re^{i\alpha}-z^N)} \\ &= e^{i\alpha} \cdot \left[\frac{(1-rz^{-N})(1-rz^N e^{-i\alpha})}{(1-rz^N)(1-rz^{-N}e^{i\alpha})}\right], \end{aligned}$$

for all $z \in \mathbb{T}$. Then we have:

$$\begin{aligned} |\mathrm{Arg}(e^{-i\alpha}q(z))| &\leq |\mathrm{Arg}(1-rz^{-N})| + |\mathrm{Arg}(1-rz^N e^{-i\alpha})| + \\ &\qquad |\mathrm{Arg}(1-rz^N)| + |\mathrm{Arg}(1-rz^{-N}e^{i\alpha})| \\ &\leq \pi r + \pi r + \pi r + \pi r \\ &= 4\pi r, \quad \text{for all } z \in \mathbb{T}. \end{aligned}$$

Now set $r = \varepsilon/(8\pi)$. Then since $|\alpha| + 4\pi r < \pi$, we find that:

$$|\alpha - \mathrm{Arg}\, q(z)| \leq \varepsilon/2, \quad \text{for all } z \in \mathbb{T}. \tag{3.17}$$

Let $\widetilde{Z}_1$ and $\widetilde{Z}_2$ be the finite sets given by:

$$\begin{aligned} \widetilde{Z}_1 &= \{r^{1/N}e^{2\pi ik/N} : k \in \mathbb{Z}/N\mathbb{Z}\} \setminus Z_1 \\ \widetilde{Z}_2 &= e^{i\alpha}\{r^{1/N}e^{2\pi ik/N} : k \in \mathbb{Z}/N\mathbb{Z}\} \setminus Z_2. \end{aligned}$$

Then we have:

$$B_1(z)/B_2(z) = \prod_{w\in Z_1} \frac{(w-z)(we^{i\alpha/N}-|w|^2z)}{(w-|w|^2z)(we^{i\alpha/N}-z)} \tag{3.18}$$

$$= q(z) \Big/ \prod_{w\in \widetilde{Z}_1} \frac{(w-z)(we^{i\alpha/N}-|w|^2z)}{(w-|w|^2z)(we^{i\alpha/N}-z)}, \tag{3.19}$$

for all $z \in \mathbb{T}$. Now for given $w \in r^{1/N}\mathbb{T}$ and $z \in \mathbb{T}$, each term in the above products may be estimated as follows:

$$\begin{aligned} \left|\frac{(w-z)(we^{i\alpha/N}-|w|^2z)}{(w-|w|^2z)(we^{i\alpha/N}-z)} - 1\right| &= \left|\frac{wz(1-|w|^2)(1-e^{i\alpha/N})}{(w-|w|^2z)(we^{i\alpha/N}-z)}\right| \\ &= \frac{(1-r^{2/N})(1-e^{i\alpha/N})}{|1-\overline{w}z|\,|we^{i\alpha/N}-z|} \\ &\leq \frac{2N^{-1}\log(1/r)\,\cdot\,\alpha N^{-1}}{|1-\overline{w}z|\,|we^{i\alpha/N}-z|} \\ &= \frac{2\alpha N^{-2}\log(1/r)}{|z-w|\,|z-we^{i\alpha/N}|}. \end{aligned}$$

Therefore from equation (3.18) we have:

$$\begin{aligned} \left|\operatorname{Arg}(B_1(z)/B_2(z))\right| &\leq \sum_{w\in Z_1}\left[\pi \times \frac{2\alpha N^{-2}\log(1/r)}{|z-w|\,|z-we^{i\alpha/N}|}\right] \\ &\leq \frac{2\pi\alpha N^{-1}\log(1/r)}{\operatorname{dist}(z,Z_1)\operatorname{dist}(z,Z_2)}, \end{aligned} \tag{3.20}$$

and similarly from (3.19) we have:

$$\left|\operatorname{Arg}(q(z)^{-1}B_1(z)/B_2(z))\right| \leq \frac{2\pi\alpha N^{-1}\log(1/r)}{\operatorname{dist}(z,\widetilde{Z}_1)\operatorname{dist}(z,\widetilde{Z}_2)}, \tag{3.21}$$

for all $z \in \mathbb{T}$. Now set $\phi = \theta - \pi\eta/2$, so that N must be no less than $2/\eta$. Then $\operatorname{dist}(z, Z_1)$ and $\operatorname{dist}(z, Z_2)$ are bounded below over all $N \geq 2/\eta$ and $z \in \mathbb{T}$ with $|\operatorname{Arg} z| \geq \theta$. Similarly, $\operatorname{dist}(z, \widetilde{Z}_1)$ and $\operatorname{dist}(z, \widetilde{Z}_2)$ are bounded below over all $N \geq 2/\eta$ and $z \in \mathbb{T}$ with $|\operatorname{Arg} z| \leq \theta - \pi\eta$.

Therefore, provided N is sufficiently large, the estimates (3.20) and (3.21) are less than δ and $\varepsilon/2$ respectively, for all $z \in \mathbb{T}$ with $|\operatorname{Arg} z| \geq \theta$ and $|\operatorname{Arg} z| \leq \theta - \pi\eta$ respectively. This implies that $-\delta < |\operatorname{Arg}(B_1(z)/B_2(z))| < \delta$ for all $z \in \mathbb{T} \setminus U$, as required. By inequality (3.17), the continuity of estimate (3.21) and the fact that $|\alpha| + \varepsilon \leq \pi$, we also have:

$$\begin{aligned} |\alpha - \operatorname{Arg}(B_1(z)/B_2(z))| &\leq |\alpha - \operatorname{Arg} q(z)| + |\operatorname{Arg}(q(z)^{-1} B_1(0)/B_2(0))| \\ &< \varepsilon/2 + \varepsilon/2 \\ &= \varepsilon, \end{aligned}$$

for all z in some neighbourhood of the arc, $\exp(i[\pi\eta - \theta, \theta - \pi\eta])$, in $\mathbb{T}$. Therefore,

$$\mu\{z \in U : |\alpha - \operatorname{Arg}(B_1(z)/B_2(z))| > \varepsilon\} < \eta,$$

as required. Finally, since $\log(1/r) = \log(8\pi/\varepsilon)$, we have:

$$\log(1/B_1(0)), \log(1/B_2(0)) \leq \mu(U)\log(8\pi/\varepsilon).$$

from inequality (3.16), so that the conclusion of the lemma holds with the constant 8π in place of 24π, in the lemma statement.

To prove the lemma for general U and g, observe that for any nonempty $U \subset \mathbb{T}$ relatively open and any measurable $g : U \to (-\pi, \pi]$, there exists $k \in \mathbb{N}$, real numbers $\alpha_1, \ldots, \alpha_k \in [\varepsilon/3 - \pi, \pi - \varepsilon/3]$ and pairwise disjoint proper arcs $I_1 \ldots, I_k \subset U$ relatively open, such that:

$$\mu\{z \in I_j : |g(z) - \alpha_j| > \varepsilon/3\} < \eta/(3k), \tag{3.22}$$

for all $j \in \{1, \ldots, k\}$, and such that:

$$\sum_{j=1}^{k} \mu(I_j) > \mu(U) - \eta/3. \tag{3.23}$$

Then from the first part of the proof, by replacing ε with $\varepsilon/3$, η with $\eta/(3k)$ and δ with $(\delta \wedge \varepsilon/3)/k$, we therefore obtain (after applying suitable rotations of the domain $\overline{\mathbb{D}}$) Blaschke products $(B_1^{(j)})_{j=1}^k$ and $(B_2^{(j)})_{j=1}^k$, with $B_1^{(j)}(0)$ and $B_2^{(j)}(0)$ positive for all $j \in \{1, \ldots, k\}$, such that:

$$\mu\{w \in I_j : \left|\alpha_j - \operatorname{Arg}\left(B_1^{(j)}(w)/B_2^{(j)}(w)\right)\right| > \varepsilon/3\} < \eta/(3k) \tag{3.24}$$

$$\left|\operatorname{Arg}\left(B_1^{(j)}(z)/B_2^{(j)}(z)\right)\right| < (\delta \wedge \varepsilon/3)/k \tag{3.25}$$

$$\log(1/B_1^{(j)}(0)), \log(1/B_2^{(j)}(0)) \leq \mu(I_j)\log(24\pi/\varepsilon) \tag{3.26}$$

for all $j \in \{1, \ldots, k\}$ and $z \in I_j$. Now let B_1 and B_2 be the Blaschke products given by:

$$B_l = \prod_{j=1}^{k} B_l^{(j)}, \quad \text{for } l = 1, 2.$$

By the construction of each term in the above product, and the fact that the arcs I_j are pairwise disjoint, we find that B_1 and B_2 have simple zeros, as required.

By summing the estimates (3.25) and (3.26) over all indices j, we obtain:

$$|\mathrm{Arg}(B_1(z)/B_2(z))| < \delta$$
$$\log(1/B_1(0)),\ \log(1/B_2(0)) \le \mu(U)\log(24\pi/\varepsilon),$$

with $B_1(0)$ and $B_2(0)$ positive, as required. Finally, by (3.24) and (3.25) we have:

$$\mu\{z \in I_j : |\alpha_j - \mathrm{Arg}(B_1(z)/B_2(z))| > 2\varepsilon/3\} < \eta/(3k),$$

for all $j \in \{1, \ldots, k\}$. Then by applying inequality (3.22), summing the resultant measures over all indices j, and finally applying inequality (3.23), we obtain:

$$\mu\{z \in U : |g(z) - \mathrm{Arg}(B_1(z)/B_2(z))| > \varepsilon\} < \eta,$$

completing the proof. □

The main scalar inner approximation result is given as follows:

Lemma 3.2.2 *Let $\varepsilon \in (0, \pi]$, let $S \subset \mathbb{T}$ measurable and let $f : \mathbb{T} \to (-\pi, \pi]$ be a measurable function vanishing outside of S. Then there exist Blaschke products, B_1 and B_2, such that:*

$$|f(z) - \mathrm{Arg}(B_1(z)/B_2(z))| < \varepsilon,$$

for almost all $z \in \mathbb{T}$, and such that $B_1(0)$ and $B_2(0)$ are positive and satisfy:

$$\log(1/B_1(0)),\ \log(1/B_2(0)) \le \mu(S)\log(100\pi/\varepsilon).$$

Proof We may suppose without loss of generality that $\mu(S) > 0$. Let $(\eta_k)_{k=0}^{\infty}$ and $(\delta_k)_{k=1}^{\infty}$ be any sequences of positive numbers such that:

$$2\sum_{k=0}^{\infty} \eta_k \le \mu(S)\log(25/24); \quad \sum_{k=1}^{\infty} \delta_k \le \varepsilon/4.$$

Let $U_1 \subset \mathbb{T}$ be any relatively open set containing S, such that $\mu(U_1) < \mu(S) + \eta_0$, and let $g_1 : \mathbb{T} \to (-\pi, \pi]$ be given by:

$$g_1(z) = \begin{cases} f(z) - \varepsilon/2 & \text{if } f(z) > 0 \\ f(z) + \varepsilon/2 & \text{if } f(z) \leq 0, \end{cases}$$

for all $z \in \mathbb{T}$. We shall inductively find sequences of sets $(U_k)_{k=2}^{\infty}$, relatively open in $\mathbb{T}$, measurable functions $(g_k)_{k=2}^{\infty}$ from $\mathbb{T}$ to $(-\pi, \pi]$, and Blaschke products $(B_1^{(k)})_{k=1}^{\infty}$ and $(B_2^{(k)})_{k=1}^{\infty}$ such that:

1. $B_1^{(k)}(0)$ and $B_2^{(k)}(0)$ are positive and satisfy:

$$\log(1/B_1^{(k)}(0)),\ \log(1/B_2^{(k)}(0)) \leq \mu(U_k) \log(96\pi/\varepsilon);$$

2. $|\mathrm{Arg}(B_1^{(k)}(z)/B_2^{(k)}(z))| < \delta_k$ for all $z \in \mathbb{T} \setminus U_k$;

3. U_{k+1} contains the set:

$$S_{k+1} = \{z \in U_k : |g_k(z) - \mathrm{Arg}(B_1^{(k)}(z)/B_2^{(k)}(z))| > \varepsilon/4\},$$

and we have:

$$\mu(U_{k+1}) < \mu(S_{k+1}) + \eta_k < 2\eta_k;$$

4. For all $z \in \mathbb{T}$, we have:

$$g_{k+1}(z) = \begin{cases} \mathrm{Arg}\Big(e^{if(z)-\varepsilon/2} / \prod_{j=1}^{k} (B_1^{(j)}(z)/B_2^{(j)}(z))\Big) & \text{if } f(z) > 0 \\ \mathrm{Arg}\Big(e^{if(z)+\varepsilon/2} / \prod_{j=1}^{k} (B_1^{(j)}(z)/B_2^{(j)}(z))\Big) & \text{if } f(z) \leq 0, \end{cases}$$

for all $k \in \mathbb{N}$. Let $l \in \mathbb{N}$ and suppose by induction that we have found U_{k+1}, g_{k+1}, $B_1^{(k)}$ and $B_2^{(k)}$ for every natural number $k < l$, satisfying the above hypotheses for all such k. Then by setting $\eta = \eta_k$, $\delta = \delta_k$, $U = U_k$ and $g = g_k|_U$, and applying Lemma 3.2.1 with $\varepsilon/4$ in place of ε, we obtain Blaschke products B_1 and B_2, with $B_1(0)$ and $B_2(0)$ positive, such that:

$$\begin{aligned} \mu\{w \in U : |g(w) - \mathrm{Arg}(B_1(w)/B_2(w))| > \varepsilon/4\} &< \eta \\ |\mathrm{Arg}(B_1(z)/B_2(z))| &< \delta, \quad \text{for all } z \in \mathbb{T} \setminus U \\ \log(B_1(0)),\ \log(B_2(0)) &\leq \mu(U) \log(96\pi/\varepsilon). \end{aligned}$$

Now by setting $B_1^{(l)} = B_1$ and $B_2^{(l)} = B_2$, and choosing a suitable relatively open set $U_{l+1} \subset \mathbb{T}$, hypotheses 1,2 and 3 above are satisfied for $k = l$. Then hypothesis 4 uniquely defines the function $g_{l+1} : \mathbb{T} \to (-\pi, \pi]$. Hence by induction on l, we obtain the sequences (U_k), (f_k), $(B_1^{(k)})$ and $(B_2^{(k)})$ which satisfy the above hypotheses for all $k \in \mathbb{N}$.

Now by hypothesis 1, we have:

$$\begin{aligned}\sum_{k=1}^{\infty} \log(1/B_1^{(k)}(0)), \sum_{k=1}^{\infty} \log(1/B_2^{(k)}(0)) &\leq \sum_{k=1}^{\infty} \mu(U_k) \log(96\pi/\varepsilon) \\ &\leq \left(\mu(S) + 2\sum_{k=1}^{\infty} \eta_k\right) \log(96\pi/\varepsilon) \\ &\leq \mu(S)(\log(96\pi/\varepsilon) + \log(25/24)) \\ &= \mu(S) \log(100\pi/\varepsilon) \quad < \quad \infty.\end{aligned}$$

This implies that the sequences of zeros of the products:

$$B_1 = \prod_{k=1}^{\infty} B_1^{(k)}; \quad B_2 = \prod_{k=1}^{\infty} B_2^{(k)}$$

both satisfy the Blaschke condition, and hence the above products converge in measure on $\mathbb{T}$ and locally uniformly on $\mathbb{D}$. Therefore $B_1(0)$ and $B_2(0)$ are positive, and satisfy:

$$\log(1/B_1(0)), \log(1/B_2(0)) \leq \mu(S) \log(100\pi/\varepsilon),$$

as required.

Finally, for any $z \in U_1 \setminus \bigcap_{k=2}^{\infty} U_k$ there exists some $l \in \mathbb{N}$ such that:

$$z \in U_l \quad \text{and} \quad z \notin U_k \quad \text{for all } k > l.$$

Therefore, by hypothesis 4 we have:

$$g_l(z) + \operatorname{Arg}\left[\prod_{k=1}^{l-1} (B_1^{(k)}(z)/B_2^{(k)}(z))\right] = \alpha \quad \text{modulo } 2\pi,$$

for some $\alpha \in (\varepsilon/2 - \pi, \pi - \varepsilon/2]$ satisfying $|f(z) - \alpha| = \varepsilon/2$. But $g_k(z)$ lies within $\varepsilon/4$ of $\operatorname{Arg}(B_1^{(l)}(z)/B_2^{(l)}(z))$ by hypothesis 3. Hence,

$$\operatorname{Arg}\left[\prod_{k=1}^{l} (B_1^{(k)}(z)/B_2^{(k)}(z))\right] = \beta,$$

for some $\beta \in (\varepsilon/4 - \pi, \pi - \varepsilon/4]$ satisfying $|f(z) - \beta| \leq 3\varepsilon/4$. Now by hypothesis 2, we have $|\mathrm{Arg}(B_1^{(k)}(z)/B_2^{(k)}(z))| < \delta_k$ for all $k > l$, and these bounds sum to less than $\varepsilon/4$. Hence,

$$\limsup_{m\to\infty} \left| f(z) - \mathrm{Arg}\left[\prod_{k=1}^{m} (B_1^{(k)}(z)/B_2^{(k)}(z)) \right] \right| < \varepsilon ,$$

But the above product is bounded away from $-\pi$ and π and therefore converges to $\mathrm{Arg}(B_1(z)/B_2(z))$ for some sequence of natural numbers m tending to ∞, for almost all $z \in U_1 \setminus \bigcap_{k=2}^{\infty} U_k$. Hence,

$$|f(z) - \mathrm{Arg}(B_1(z)/B_2(z))| < \varepsilon , \tag{3.27}$$

for almost all $z \in U_1 \setminus \bigcap_{k=2}^{\infty} U_k$. Since $\mu(U_k) \to 0$ as k tends to ∞, this intersection has measure zero. Now by a second application of hypothesis 2, we find that:

$$|\mathrm{Arg}(B_1(z)/B_2(z))| \leq \sum_{k=1}^{\infty} \delta_k < \varepsilon ,$$

for all $z \in U_1$. Hence inequality (3.27) holds for almost all $z \in \mathbb{T}$, as required. □

We now turn towards matrix-valued inner function approximation. In order to make use of the previous result, we will need the following simple lemma:

Lemma 3.2.3 *There exist rank one orthogonal projections $P_1, \ldots, P_{n^2} \in \mathcal{L}(\mathbb{C}^n)$ such that any $n \times n$ unitary matrix $U \in \mathcal{L}(\mathbb{C}^n)$ may be expressed in the form:*

$$U = \exp(i\alpha_1 P_1) \exp(i\alpha_2 P_2) \cdots \exp(i\alpha_{n^2} P_{n^2}) , \tag{3.28}$$

for some $\alpha_1, \ldots, \alpha_n \in (-\pi, \pi]$.

Proof Let $e_1, \ldots, e_n$ denote the standard basis of the $n \times 1$ column vectors, and let $P_1, \ldots, P_{n^2} \in \mathcal{L}(\mathbb{C}^n)$ be given by:

$$P_{k^2-l} = \begin{cases} e_k e_k^{\mathrm{T}} & \text{for } l = 0, 2, 4, \ldots, 2(k-1) \\ (e_k + ie_{(l+1)/2})(e_k - ie_{(l+1)/2})^{\mathrm{T}}/2 & \text{for } l = 1, 3, 5, \ldots, 2k-3, \end{cases}$$

for $k = 1, \ldots, n$. Now let $U \in \mathcal{L}(\mathbb{C}^n)$ be any $n \times n$ unitary matrix. We shall show by induction on $k = 1, \ldots, n$ and $l = 0, \ldots, 2(k-1)$, that there are coefficients $\alpha_1, \ldots, \alpha_{n^2}$ such that the unitary matrix:

$$U \exp(-i\alpha_{n^2} P_{n^2}) \cdots \exp(-i\alpha_{k^2-l} P_{k^2-l}) \tag{3.29}$$

takes the form:

$$\left(\begin{array}{ccc|c|ccc} m_{11} & \cdots & m_{1(k-1)} & m_{1k} & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ m_{(k-1)1} & \cdots & m_{(k-1)(k-1)} & m_{(k-1)k} & 0 & \cdots & 0 \\ \hline u_1 & \cdots & u_{k-1} & \lambda & 0 & \cdots & 0 \\ \hline 0 & \cdots & 0 & 0 & 1 & & 0 \\ \vdots & \ddots & \vdots & \vdots & & \ddots & \\ 0 & \cdots & 0 & 0 & 0 & & 1 \end{array}\right) \tag{3.30}$$

for some coefficients $u_1, \ldots, u_{k-1}$, λ and $m_{11}, \ldots, m_{1k}, \ldots, m_{(k-1)1}, \ldots, m_{(k-1)k}$ in $\mathbb{C}$, with the coefficients satisfying $u_j = 0$ for all indices $j \leq (l+1)/2$ and:

$$\begin{cases} \alpha u_{(l+2)/2} = \beta\lambda \quad \text{for some } \alpha, \beta \geq 0 \text{ not both zero} & \text{if } l = 0, 2, \ldots, 2k-4 \\ \lambda = 1 \quad \text{and} \quad m_{1k}, \ldots, m_{(k-1)k} = 0 & \text{if } l = 2k-2. \end{cases}$$

The claim that (3.29) takes this form, for coefficients λ, (u_j) and (m_{ab}) satisfying the above criteria, will form the induction hypothesis.

In the case $k = n$, the matrix U clearly takes the form of (3.30), for some arbitrary complex coefficients λ, (u_j) and (m_{ab}). Now set:

$$\alpha_{n^2} = \begin{cases} \operatorname{Arg}(\lambda/u_1) & \text{if } n > 1 \text{ and } \lambda, u_1 \neq 0 \\ \operatorname{Arg}\lambda & \text{if } n = 1 \\ 0 & \text{otherwise}, \end{cases}$$

By right multiplying the expression (3.30) by $\exp(-i\alpha_{n^2} P_{n^2})$, we find that it takes the same form, with the coefficients $u_1, \ldots, u_{k-1}$ remaining unchanged, but the new value assigned to λ satisfies $\operatorname{Arg}\lambda = \operatorname{Arg} u_1$ in the case that $n > 1$ and λ, u_1 are nonzero, and it satisfies $\lambda = 1$ in the case $n = 1$. Hence the induction hypothesis holds for $k = n$ and $l = 0$.

Now let $k \in \{1, \ldots, n\}$ and $l \in \{0, \ldots, 2(k-1)\}$, and suppose that the induction hypothesis holds for this choice of k and l. To complete the inductive step, we have four cases to consider, as follows:

If l is even and $l < 2(k-1)$, then $\alpha u_{(l+2)/2} = \beta\lambda$ for some $\alpha, \beta \geq 0$ not both zero. Therefore, there is some $\theta \in [0, \pi/2]$ such that $u_{(l+2)/2} \cos\theta = \lambda \sin\theta$. Now by right multiplying expression (3.30) by:

$$\exp(-2i\theta P_{k^2-l-1}) = (I - P_{k^2-l-1}) + e^{-2i\theta} P_{k^2-l-1},$$

we find that it takes the same form, with the coefficients $u_1, \ldots, u_{k-1}$ unchanged, except for $u_{(l+2)/2}$, which changes to:

$$(i/2)(\lambda - iu_{(l+2)/2}) - (i/2)(\lambda + iu_{(l+2)/2})e^{-2i\theta} = 0.$$

So by setting $\alpha_{k^2-l-1} = 2\theta$, we find that the induction hypothesis holds for $l+1$ in place of l.

If l is odd and $l < 2k - 3$, then we shall set:

$$\alpha_{k^2-l-1} = \begin{cases} \operatorname{Arg}(\lambda/u_{(l+1)/2}) & \text{if } \lambda, u_{(l+1)/2} \neq 0 \\ 0 & \text{otherwise}. \end{cases}$$

Now by right mutiplying the expression (3.30) by $\exp(-i\alpha_{k^2-l-1}P_{k^2-l-1})$, we find that it takes the same form, with $u_1, \ldots, u_{k-1}$ unchanged, but the new value assigned to λ satisfies $\operatorname{Arg}\lambda = \operatorname{Arg} u_{(l+1)/2}$, provided λ and $u_{(l+1)/2}$ are nonzero. Hence the induction hypothesis holds for $l+1$ in place of l.

If $l = 2k - 3$ then $u_1, \ldots, u_{k-1}$ are zero, so $|\lambda| = 1$ and $m_{1k}, \ldots, m_{(k-1)k}$ are zero, by the unitarity of (3.30). Now we shall set $\alpha_{k^2-l-1} = \operatorname{Arg}\lambda$. Then by right multiplying (3.30) by $\exp(-i\alpha_{k^2-l-1}P_{k^2-l-1})$, we find that the coefficients (u_j) and (m_{jk}) remain zero, but λ takes the new value 1. So the induction hypothesis again holds for $l+1$ in place of l.

Finally, if $l = 2k - 2$ and $k > 1$, then by induction we have $\lambda = 1$ and $u_1, \ldots, u_{k-1} = 0$ and $m_{1k}, \ldots, m_{(k-1)k} = 0$. Now let us replace k with $k - 1$ and l with 0. Then we find that the product:

$$U \exp(-i\alpha_{n^2} P_{n^2}) \cdots \exp(-i\alpha_{k^2-l+1} P_{k^2-l+1})$$

still takes the form of (3.30), with new arbitrary values assigned to λ, (u_j) and (m_{ab}). Now by setting:

$$\alpha_{k^2-l} = \begin{cases} \operatorname{Arg}(\lambda/u_1) & \text{if } \lambda, u_1 \neq 0 \\ 0 & \text{otherwise}, \end{cases}$$

we find that by right multiplying (3.30) by $\exp(-i\alpha_{k^2-l} P_{k^2-l})$, the coefficients $u_1, \ldots, u_{k-1}$ remain unchanged, but the new value assigned to λ satisfies $\operatorname{Arg}\lambda = \operatorname{Arg} u_1$, provided λ and u_1 are both nonzero. Hence the induction hypothesis holds for the new values of k and l.

This completes the inductive step, and so the induction hypothesis holds for all choices of k and l. In particular, from the case $k = 1$ and $l = 0$, we have:

$$U \exp(-i\alpha_{n^2} P_{n^2}) \cdots \exp(-i\alpha_1 P_1) = I.$$

Hence U takes the form of (3.28), as required. □

We are now in a position to establish the matrix inner approximation result, given as follows:

Lemma 3.2.4 *Let $\varepsilon \in (0, \pi n^2]$, let $S \subset \mathbb{T}$ measurable and let $\phi \in \mathrm{L}^\infty(\mathcal{L}(\mathbb{C}^n))$ be an almost everywhere unitary matrix-valued function, with $\phi(z) = I$ for all $z \in \mathbb{T} \setminus S$. Then there exist matrix-valued inner functions $\Theta, \theta \in \mathrm{H}^\infty(\mathcal{L}(\mathbb{C}^n))$ such that:*

$$\|\theta\phi\Theta^{-1} - I\|_\infty \leq \varepsilon,$$

and such that:

$$\|\Theta - I\|_2,\ \|\theta - I\|_2 \leq \sqrt{2n^5 \mu(S) \log(100\pi n^2/\varepsilon)}.$$

Proof Let $P_1, \ldots, P_{n^2}$ be the rank one orthogonal projections provided by Lemma 3.2.3. Let

$$\theta : (-\pi, \pi]^{n^2} \rightarrow \{U \in \mathcal{L}(\mathbb{C}^n) : U \text{ is unitary}\}$$

be the continuous map given by:

$$\theta(\alpha_1, \ldots, \alpha_{n^2}) = \exp(i\alpha_1 P_1)\exp(i\alpha_2 P_2)\cdots\exp(i\alpha_{n^2} P_{n^2}),$$

for all $\alpha_1, \ldots, \alpha_{n^2}$. By Lemma 3.2.3, this map is onto. It therefore has a measurable right inverse h, which we may construct explicitly as follows:

Let $C \subset [0, 1/2]$ denote the Cantor set, defined by:

$$C = \left\{\sum_{j=1}^{\infty} x_j \cdot 3^{-j} : x_1, x_2, \ldots \in \{0, 1\}\right\}$$

Let $g : C + \mathbb{N} \rightarrow (-\pi, \pi]^{n^2}$ be the continuous surjection given by:

$$g(x + m) = (\pi \ldots, \pi) - (2 - 2^{-m}) \cdot \pi \sum_{k=0}^{\infty} (x_{kn^2+1}, \ldots, x_{kn^2+n^2}) \cdot 2^{-k-1},$$

where $x = \sum_{j=1}^{\infty} x_j \cdot 3^{-j}$, for any sequence $x_1, x_2, \ldots \in \{0, 1\}$ and any $m \in \mathbb{N}$. Then we may define $h : \{U \in \mathcal{L}(\mathbb{C}^n) : U \text{ is unitary}\} \rightarrow (-\pi, \pi]^{n^2}$ by:

$$h(U) = g\Big(\min\{x \in C + \mathbb{N} : \theta(g(x)) = U\}\Big),$$

for all $U \in \mathcal{L}(\mathbb{C}^n)$ unitary. By the continuity and surjectivity of $\theta \circ g$, this is a well defined measurable right inverse to θ, as required.

Now define the measurable functions $f_1, \ldots, f_{n^2} : \mathbb{T} \rightarrow (-\pi, \pi]$ almost everywhere, by:

$$(f_1(z), \ldots, f_{n^2}(z)) = h(\phi(z)),$$

for all $z \in \mathbb{T}$ such that $\phi(z)$ is unitary. So we have the factorization:

$$\phi(z) = \exp(if_1(z)P_1)\exp(if_2(z)P_2)\cdots\exp(if_{n^2}(z)P_{n^2}),$$

for almost all $z \in \mathbb{T}$. Moreover, since $h(I) = (0, \ldots, 0)$, the functions $f_1, \ldots, f_{n^2}$ vanish outside of S. Therefore by Lemma 3.2.2, there exist Blaschke products $B_1, \ldots, B_{n^2} \in \mathrm{H}^\infty$ and Blaschke products $b_1, \ldots, b_{n^2} \in \mathrm{H}^\infty$, such that:

$$|f_k(z) - \mathrm{Arg}(B_k(z)/b_k(z))| < \varepsilon/n^2, \tag{3.31}$$

for almost all $z \in \mathbb{T}$, and such that $B_k(0)$ and $b_k(0)$ are positive and satisfy:

$$\log(1/B_k(0)),\ \log(1/b_k(0)) \ \leq \ \mu(S)\log(100\pi n^2/\varepsilon)\,, \tag{3.32}$$

for all $k \in \{1, \ldots, n^2\}$.

For any orthogonal projection $P \in \mathcal{L}(\mathbb{C}^n)$ and any scalar-valued function $f : \mathbb{T} \to \mathbb{C}$, let f^P denote the matrix-valued function:

$$f^P \ = \ fP + (I-P) \ : \ \mathbb{T} \to \mathcal{L}(\mathbb{C}^n)\,.$$

Clearly f^P is almost everywhere unitary matrix-valued whenever f is a.e. unimodular. Now let $\theta, \Theta \in \mathrm{H}^\infty(\mathcal{L}(\mathbb{C}^n))$ be the inner functions given by:

$$\begin{aligned}\Theta &= b_1^{(I-P_1)} B_1^{P_1} \cdot b_2^{(I-P_2)} B_2^{P_2} \cdots b_{n^2}^{(I-P_{n^2})} B_{n^2}^{P_{n^2}}\,;\\ \theta &= (b_1 \cdot b_2 \cdots b_{n^2}) I\,.\end{aligned}$$

Therefore,

$$\begin{aligned}\theta^{-1}\Theta &= (b_1^{P_1})^{-1} B_1^{P_1} \cdot (b_2^{P_2})^{-1} B_2^{P_2} \cdots (b_{n^2}^{P_{n^2}})^{-1} B_{n^2}^{P_{n^2}}\\ &= (B_1/b_1)^{P_1} \cdot (B_2/b_2)^{P_2} \cdots (B_{n^2}/b_{n^2})^{P_{n^2}}\,.\end{aligned}$$

Now by a simple application of Lemma 3.1.2, we find that:

$$\begin{aligned}\|\theta\phi\Theta^{-1} - I\|_\infty &= \|\phi - \theta^{-1}\Theta\|_\infty\\ &\leq \sum_{k=1}^{n^2} \Big\| e^{if_1P_1} \cdots e^{if_{k-1}P_{k-1}} \times \left(e^{if_kP_k} - (B_k/b_k)^{P_k}\right) \times\\ &\qquad\qquad (B_{k+1}/b_{k+1})^{P_{k+1}} \cdots (B_{n^2}/b_{n^2})^{P_{n^2}} \Big\|_\infty\\ &= \sum_{k=1}^{n^2} \Big\| \exp(if_kP_k) - \exp(i\mathrm{Arg}(B_k/b_k)P_k) \Big\|_\infty\\ &\leq \sum_{k=1}^{n^2} \|f_k - \mathrm{Arg}(B_k/b_k)\|_\infty \ \leq \ \varepsilon\,,\end{aligned}$$

as required, where the last inequality follows from (3.31).

Finally, observe that since B_k and b_k are inner, we have:

$$\|B_k - 1\|_2^2 \ = \ \langle B_k - 1, B_k - 1\rangle \ = \ \|B_k\|_2^2 - 2\mathrm{Re}B_k(0) + 1$$

$$\begin{aligned} &= 2(1 - B_k(0)) \\ &\leq 2\log(1/B_k(0)) \\ &\leq 2\mu(S)\log(100\pi n^2/\varepsilon)\,, \end{aligned}$$

by inequality (3.32), and similarly for b_k in place of B_k. Therefore,

$$\|B_k - 1\|_2,\ \|b_k - 1\|_2 \ \leq\ \sqrt{2\mu(S)\log(100\pi n^2/\varepsilon)}\,,$$

for all $k \in \{1, \ldots, n^2\}$. Hence,

$$\begin{aligned} \|\Theta - I\|_2 &\leq \sum_{k=1}^{n^2} \left\| (b_1^{(I-P_1)} B_1^{P_1} \cdots b_{k-1}^{(I-P_{k-1})} B_{k-1}^{P_{k-1}}) \cdot (b_k^{(I-P_k)} B_k^{P_k} - I) \right\|_2 \\ &= \sum_{k=1}^{n^2} \sqrt{\|b_k^{(I-P_k)} - I\|_2^2 + \|B_k^{P_k} - I\|_2^2} \\ &= \sum_{k=1}^{n^2} \sqrt{(n-1)\|b_k - 1\|_2^2 + \|B_k - 1\|_2^2} \\ &\leq \sum_{k=1}^{n^2} \sqrt{2n\mu(S)\log(100\pi n^2/\varepsilon)}\,. \end{aligned} \tag{3.33}$$

Similarly, we have:

$$\begin{aligned} \|\theta - I\|_2 &\leq \sum_{k=1}^{n^2} \|b_1 \cdots b_{k-1} \cdot (b_k I - I)\|_2 \\ &= \sum_{k=1}^{n^2} \sqrt{n}\,\|b_k - 1\|_2 \\ &\leq \sum_{k=1}^{n^2} \sqrt{2n\mu(S)\log(100\pi n^2/\varepsilon)}\,, \end{aligned} \tag{3.34}$$

and by evaluating the identical sums, (3.33) and (3.34), we obtain the required estimate for $\|\Theta - I\|_2$ and $\|\theta - I\|_2$. □

3.3 The main results

3.3.1 Constructing the factors of unitary-valued functions

We shall now combine the main inner and outer approximation results of the previous two sections, in the iterative construction used to prove the following theorem. First of all, recall the choice of the constants K_1, K_2 and K_3 used in the statement of Lemma 3.1.6.

Theorem 3.3.1 *Let $\varepsilon > 0$ and let $\phi \in \mathrm{L}^\infty(\mathcal{L}(\mathbb{C}^n))$ almost everywhere unitary. Then there exist almost everywhere invertible functions $g, h \in \mathrm{H}^\infty(\mathcal{L}(\mathbb{C}^n))$, with $g^{-1}, h^{-1} \in \mathrm{L}^\infty(\mathcal{L}(\mathbb{C}^n))$, such that:*

$$\phi = h^* g \quad \textit{almost everywhere on } \mathbb{T},$$

and such that:

$$\|g\|_\infty, \|h\|_\infty, \|g^{-1}\|_\infty, \|h^{-1}\|_\infty < 1 + \varepsilon .$$

Proof If $\varepsilon < e^{6K_2} - 1$, then we may apply Lemma 3.2.4 with $S = \mathbb{T}$ and $(\log(1+\varepsilon)/(3K_2))^{6/5}$ in place of ε, to obtain inner functions $\Theta, \theta \in \mathrm{H}^\infty(\mathcal{L}(\mathbb{C}^n))$ such that:

$$\|\theta\phi\Theta^{-1} - I\|_\infty \leq (\log(1+\varepsilon)/(3K_2))^{6/5} . \tag{3.35}$$

Otherwise, we may set $\Theta = \theta \equiv I$ in $\mathrm{H}^\infty(\mathcal{L}(\mathbb{C}^n))$, and then the above inequality is satisfied trivially.

We shall find, by induction, a.e. unitary functions $\phi_k \in \mathrm{L}^\infty(\mathcal{L}(\mathbb{C}^n))$, outer functions $\rho_k, \tau_k \in \mathrm{H}^\infty(\mathcal{L}(\mathbb{C}^n))$ and inner functions $\Theta_k, \theta_k \in \mathrm{H}^\infty(\mathcal{L}(\mathbb{C}^n))$, for every $k \in \mathbb{N}$, with the five sequences of functions satisfying the six hypotheses:

1. $\phi_1 = \theta\phi\Theta^{-1}$;

2. $\phi_{j+1} = ((\theta_j\tau_j)^{-1})^* \phi_j (\Theta_j\rho_j)^{-1}$;

3. $\|\phi_{j+1} - I\|_\infty \leq \frac{1}{2} \|\phi_j - I\|_\infty^{13/12}$;

4. $\|\Theta_j - I\|_2, \|\theta_j - I\|_2 \leq 80\sqrt{2\pi}\, K_1\, n^{7/2} \|\phi_j - I\|_\infty^{1/24}$;

5. $\|\rho_j\|_\infty, \|\tau_j\|_\infty, \|\rho_j^{-1}\|_\infty, \|\tau_j^{-1}\|_\infty \;\leq\; \exp(K_2\|\phi_j - I\|_\infty^{5/6})$;

6. $\|\rho_j - I\|_2, \|\tau_j - I\|_2, \|\rho_j^{-1} - I\|_2, \|\tau_j^{-1} - I\|_2 \;\leq\; K_3\|\phi_j - I\|_\infty^{5/6}$,

where hypotheses 2 to 6 hold for every index $j \in \mathbb{N}$.

Suppose by induction on $k \in \mathbb{N}$, that we have found almost everwhere unitary functions $\phi_1, \ldots, \phi_k \in \mathrm{L}^\infty(\mathcal{L}(\mathbb{C}^n))$, outer functions $\rho_1, \ldots, \rho_{k-1} \in \mathrm{H}^\infty(\mathcal{L}(\mathbb{C}^n))$ and $\tau_1, \ldots, \tau_{k-1} \in \mathrm{H}^\infty(\mathcal{L}(\mathbb{C}^n))$, and inner functions $\Theta_1, \ldots, \Theta_{k-1} \in \mathrm{H}^\infty(\mathcal{L}(\mathbb{C}^n))$ and $\theta_1, \ldots, \theta_{k-1} \in \mathrm{H}^\infty(\mathcal{L}(\mathbb{C}^n))$, satisfying all six hypotheses for all indices $j < k$. Then by applying Lemma 3.1.6 to the function ϕ_k, we obtain outer functions $\rho_k, \tau_k \in \mathrm{H}^\infty(\mathcal{L}(\mathbb{C}^n))$ satisfying hypotheses 5 and 6 for $j = k$, and such that:

$$(\tau_k^{-1})^* \phi_k \, \rho_k^{-1} \text{ is almost everywhere unitary,} \tag{3.36}$$

and also,

$$\|(\tau_k^{-1})^* \phi_k \, \rho_k^{-1} - I\|_2 \;\leq\; K_1\|\phi_k - I\|_\infty^{5/3}. \tag{3.37}$$

Now let $\eta = \frac{1}{4}\|\phi_k - I\|_\infty^{13/12}$ and let $S \subset \mathbb{T}$ be given by:

$$S \;=\; \left\{ z \in \mathbb{T} \,:\, \|((\tau_k^{-1})^* \phi_k \, \rho_k^{-1} - I)(z)\|_\infty \geq \varepsilon \right\}.$$

Then we have:

$$\begin{aligned} \|(\tau_k^{-1})^* \phi_k \, \rho_k^{-1} - I\|_2^2 \;&\geq\; \int_{\mathbb{T}} \|((\tau_k^{-1})^* \phi_k \, \rho_k^{-1} - I)(z)\|_\infty^2 \, \mathrm{d}\mu(z) \\ &\geq\; \eta^2 \mu(S), \end{aligned}$$

and the left hand expression is less than or equal to $K_1^2\|\phi - I\|_\infty^{10/3}$, by inequality (3.37). So provided $\eta > 0$, we have:

$$\begin{aligned} \mu(S)/\eta \;&=\; \|(\tau_k^{-1})^* \phi_k \, \rho_k^{-1} - I\|_2^2/\eta^3 \\ &\leq\; 64K_1^2\|\phi_k - I\|_\infty^{10/3} / \, \|\phi_k - I\|_\infty^{13/4} \\ &\leq\; 64K_1^2\|\phi_k - I\|_\infty^{1/12}. \end{aligned} \tag{3.38}$$

If $\eta = 0$, then $(\tau_k^{-1})^* \phi_k \, \rho_k^{-1} = I$ almost everywhere, so we may set $\Theta_k = \theta_k \equiv I$ and set $\phi_{k+1} = (\tau_k^{-1})^* \phi_k \, \rho_k^{-1}$, to satisfy hypotheses 2, 3 and 4 trivially, and therefore complete the inductive step.

Otherwise, let $\psi \in \mathrm{L}^\infty(\mathcal{L}(\mathbb{C}^n))$ be given by:

$$\psi(z) = \begin{cases} ((\tau_k^{-1})^* \phi_k \rho_k^{-1})(z) & \text{for all } z \in S \\ I & \text{for all } z \in \mathbb{T} \setminus S. \end{cases}$$

By (3.36), this is almost everywhere unitary. So provided $\eta < \pi$, we may apply Lemma 3.2.4 to the function ψ, with the earlier choice of S and with η in place of ε, to obtain inner functions $\Theta_k, \theta_k \in \mathrm{H}^\infty(\mathcal{L}(\mathbb{C}^n))$ such that:

$$\|\theta_k \psi \Theta_k^{-1} - I\|_\infty \le \eta, \tag{3.39}$$

and such that:

$$\begin{aligned} \|\Theta_k - I\|_2, \|\theta_k - I\|_2 &\le \sqrt{200\pi n^7 \mu(S)/\eta} \\ &\le \sqrt{200\pi n^7 \times 64 K_1^2 \|\phi_k - I\|_\infty^{1/12}} \\ &= 80\sqrt{2\pi}\, K_1\, n^{7/2}\, \|\phi_k - I\|_\infty^{1/24}, \end{aligned}$$

where the last inequality comes from estimate (3.38). Hence Θ_k and θ_k satisfy hypothesis 4 with $j = k$. If $\eta \ge \pi$ then by setting $\Theta = \theta \equiv I$, the above inequalities hold trivially, so hypothesis 4 is also satisfied in this case.

Finally, by setting $\phi_{k+1} = \theta_k((\tau_k^{-1})^* \phi_k \rho_k^{-1})\Theta_k^{-1} \in \mathrm{L}^\infty(\mathcal{L}(\mathbb{C}^n))$ a.e. unitary, we find that hypothesis 2 is satisfied for $j = k$, and we obtain the following from estimate (3.39):

$$\begin{aligned} \|\phi_{k+1} - I\|_\infty &\le \|\theta_k((\tau_k^{-1})^* \phi_k \rho_k^{-1} - \psi)\Theta_k^{-1}\|_\infty + \|\theta_k \psi \Theta_k^{-1} - I\|_\infty \\ &\le \eta + \eta, \end{aligned}$$

since Θ_k^{-1} and θ_k are a.e. unitary, and $\|((\tau_k^{-1})^* \phi_k \rho_k^{-1} - I)(z)\|_\infty < \eta$ for every $z \in \mathbb{T} \setminus S$, by the definition of S. Recalling that $\eta = \frac{1}{4}\|\phi_k - I\|_\infty^{13/12}$, we find that ϕ_{k+1} satisfies hypothesis 3 for $j = k$. So all six hypotheses are satisfied for $j = k$, completing the inductive step.

Hence by induction on $k \in \mathbb{N}$, we obtain the desired sequences $(\phi_k)_{k=1}^\infty$, $(\rho_k)_{k=1}^\infty$, $(\tau_k)_{k=1}^\infty$, $(\Theta_k)_{k=1}^\infty$ and $(\theta_k)_{k=1}^\infty$. Now for each natural number k, let f_k be the element of $\mathrm{H}^\infty(\mathcal{L}(\mathbb{C}^n))$ given by the product:

$$f_k = \Theta_k \cdot \rho_k \cdot \Theta_{k-1} \cdot \rho_{k-1} \cdots \Theta_1 \cdot \rho_1 .$$

Then for any natural numbers $j < k$, we have the following from hypotheses 3, 4, 5 and 6:

$$
\begin{aligned}
\|f_k - f_j\|_2 &\le \|\Theta_j\rho_j \cdots \Theta_1\rho_1\|_\infty \sum_{m=j+1}^{k} \|(\Theta_m\rho_m - I)(\Theta_{m-1}\rho_{m-1}\cdots\Theta_{j+1}\rho_{j+1})\|_2 \\
&\le (\|\rho_k\|_\infty \cdots \|\rho_1\|_\infty) \sum_{m=j+1}^{k} \|\Theta_m\rho_m - I\|_2 \,/\, \|\rho_m\|_\infty \\
&= (\|\rho_k\|_\infty \cdots \|\rho_1\|_\infty) \sum_{m=j+1}^{k} \|(\Theta_m - I)\rho_m + (I - \rho_m^{-1})\rho_m\|_2 \,/\, \|\rho_m\|_\infty \\
&\le \exp\Big[\sum_{m=1}^{k} K_2\|\phi_m - I\|_\infty^{5/6}\Big] \times \sum_{m=j+1}^{k} (\|\rho_m^{-1} - I\|_2 + \|\Theta_m - I\|_2) \\
&\le \exp\Big[K_2\sum_{m=1}^{\infty} \|\phi_m - I\|_\infty^{5/6}\Big] \times \sum_{m=j+1}^{k} \Big(K_3\|\phi_m - I\|_\infty^{5/6} + \\
&\qquad\qquad 80\sqrt{2\pi}\, K_1\, n^{7/2}\|\phi_m - I\|_\infty^{1/24}\Big) \\
&\to 0 \quad \text{as } j, k \to 0,
\end{aligned}
$$

since hypothesis 3 implies that $\|\phi_m - I\|_\infty \to 0$ geometrically as $m \to \infty$, so that the above left-hand sum converges, and the above right-hand sum tends to 0 as $j, k \to \infty$. Thus $(f_k)_{k=1}^{\infty}$ is an L^2-Cauchy sequence, with limit:

$$
g\Theta^{-1} = \lim_{k\to\infty} (\Theta_k\rho_k \cdots \Theta_1\rho_1) \in H^2(\mathcal{L}(\mathbb{C}^n)), \tag{3.40}
$$

for some unique $g \in H^2(\mathcal{L}(\mathbb{C}^n))$. Now we may apply hypothesis 5 again to get:

$$
\begin{aligned}
\|g\|_\infty &= \lim_{k\to\infty} \|\Theta_k\rho_k \cdots \Theta_1\rho_1\|_\infty \\
&\le \liminf_{k\to\infty} (\|\rho_k\|_\infty \cdots \|\rho_1\|_\infty) \\
&\le \liminf_{k\to\infty} \exp\Big[\sum_{m=1}^{k} K_2\|\phi_m - I\|_\infty^{5/6}\Big] \\
&= \exp\Big[\sum_{m=1}^{\infty} K_2\|\phi_m - I\|_\infty^{5/6}\Big] < \infty,
\end{aligned}
$$

so $g \in H^\infty(\mathcal{L}(\mathbb{C}^n))$, as required. Similarly, we have:

$$
\begin{aligned}
\|g^{-1}\|_\infty &\le \liminf_{k\to\infty} (\|\rho_1^{-1}\Theta_1^{-1}\|_\infty \cdots \|\rho_k^{-1}\Theta_k^{-1}\|_\infty) \\
&\le \exp\Big[\sum_{m=1}^{\infty} K_2\|\phi_m - I\|_\infty^{5/6}\Big] < \infty,
\end{aligned}
$$

so $g^{-1} \in \mathrm{L}^\infty(\mathcal{L}(\mathbb{C}^n))$, as required.

Now by using the same argument as above, with τ in place of ρ and θ in place of Θ, we find that:

$$\theta_1\tau_1,\ \theta_2\tau_2\theta_1\tau_1,\ \theta_3\tau_3\theta_2\tau_2\theta_1\tau_1,\ \ldots$$

is an L^2-Cauchy sequence in $\mathrm{H}^\infty(\mathcal{L}(\mathbb{C}^n))$, with limit:

$$h\theta^{-1} \;=\; \lim_{k\to\infty}(\theta_k\tau_k\cdots\theta_1\tau_1) \;\in\; \mathrm{H}^2(\mathcal{L}(\mathbb{C}^n)), \tag{3.41}$$

for some unique $h \in \mathrm{H}^2(\mathcal{L}(\mathbb{C}^n))$. Moreover, we find that $h \in \mathrm{H}^\infty(\mathcal{L}(\mathbb{C}^n))$ and $h^{-1} \in \mathrm{L}^\infty(\mathcal{L}(\mathbb{C}^n))$ by the same argument as before, and we also have:

$$\|g\|_\infty,\ \|h\|_\infty,\ \|g^{-1}\|_\infty,\ \|h^{-1}\|_\infty \;\le\; \exp\Big[\sum_{m=1}^{\infty} K_2\|\phi_m - I\|_\infty^{5/6}\Big]. \tag{3.42}$$

Now by repeatedly applying hypothesis 3, we may estimate each term of the sum in the right hand side of (3.42), as follows:

$$\begin{aligned}\|\phi_2 - I\|_\infty &\le 2^{-1}\cdot\|\phi_1 - I\|_\infty^{1/12}\cdot\|\phi_1 - I\|_\infty \\ &\le 2^{-11/12}\cdot\|\phi_1 - I\|_\infty \end{aligned} \tag{3.43}$$

$$\begin{aligned}\|\phi_3 - I\|_\infty &\le 2^{-1}\cdot\|\phi_2 - I\|_\infty^{13/12} \\ &\le 2^{-(11/12\times 13/12)-1}\cdot\|\phi_1 - I\|_\infty^{13/12} \\ &\le 2^{-(143/144)-1+(1/12)}\cdot\|\phi_1 - I\|_\infty \\ &= 2^{-275/144}\cdot\|\phi_1 - I\|_\infty \end{aligned} \tag{3.44}$$

So by induction on k, we have:

$$\begin{aligned}\|\phi_k - I\|_\infty &\le 2^{-1}\cdot\|\phi_{k-1} - I\|_\infty^{13/12} \\ &\le 2^{-1}\cdot(2^{2-(k-1)}\cdot\|\phi_1 - I\|_\infty)^{13/12} \\ &\le 2^{-1+(13/12)(2-(k-1))+(1/12)}\cdot\|\phi_1 - I\|_\infty \\ &= 2^{-2-(13/12)(k-4)}\cdot\|\phi_1 - I\|_\infty \\ &\le 2^{2-k}\cdot\|\phi_1 - I\|_\infty, \end{aligned}$$

for all natural numbers $k > 3$. From the above estimate, together with (3.43) and (3.44), we have:

$$\begin{aligned}\sum_{k=1}^{\infty} K_2\|\phi_m - I\|_{\infty}^{5/6} &\leq K_2\|\phi_1 - I\|_{\infty}^{5/6}\Big[1 + 2^{-11/12\times 5/6} + 2^{-275/144\times 5/6} + \\ &\qquad \sum_{k=4}^{\infty} 2^{(2-k)\times 5/6}\Big] \\ &< 3K_2\|\phi_1 - I\|_{\infty}^{5/6},\end{aligned}$$

in the case that $\|\phi_1 - I\|_\infty > 0$. So from inequality (3.35), together with hypothesis 1, we find that the left hand side is less than $\log(1+\varepsilon)$ in all cases. Substituting this estimate into (3.42), we obtain:

$$\|g\|_\infty, \|h\|_\infty, \|g^{-1}\|_\infty, \|h^{-1}\|_\infty < \exp(\log(1+\varepsilon)) = 1+\varepsilon,$$

as required.

Finally, from the identities (3.40) and (3.41), which define the functions g and h, together with the fact that $\phi_k \to I$ in measure as $k \to \infty$, we have the following:

$$\begin{aligned} h^* g &= \theta^{-1} \cdot \lim_{k\to\infty} (\theta_k \tau_k \cdots \theta_1 \tau_1)^* \cdot \lim_{k\to\infty} \phi_k \cdot \lim_{k\to\infty} (\Theta_k \rho_k \cdots \Theta_1 \rho_1) \cdot \Theta \quad \text{a.e.} \\ &= \theta^{-1} \cdot \lim_{k\to\infty} \Big((\theta_k \tau_k \cdots \theta_1 \tau_1)^* \phi_k (\Theta_k \rho_k \cdots \Theta_1 \rho_1) \Big) \cdot \Theta \quad \text{a.e.} \\ &= \theta^{-1} \cdot \lim_{k\to\infty} \phi_1 \cdot \Theta \\ &= \phi, \end{aligned}$$

with all the above limits converging in measure. This completes the proof. □

3.3.2 Factorization of log-integrable $L^p(\mathcal{L}(\mathbb{C}^n))$ functions

The main result of the chapter is given as follows:

Theorem 3.3.2 *Let $p, q, r \in [1, \infty]$ be indices satisfying $1/p = 1/q + 1/r$, let $\varepsilon > 0$ and let $f \in L^p(\mathcal{L}(\mathbb{C}^n))$. Then the following are equivalent:*

1. $\displaystyle\int_0^{2\pi} \log|\det f(e^{i\theta})|\, d\theta > -\infty$;

2. *There exist almost everywhere invertible functions* $g \in \mathrm{H}^q(\mathcal{L}(\mathbb{C}^n))$ *and* $h \in \mathrm{H}^r(\mathcal{L}(\mathbb{C}^n))$ *such that:*

   (a) $f = h^* g$ *almost everywhere on* $\mathbb{T}$*;*

   (b) $\|f\|_p \;\leq\; \|g\|_q \|h\|_r \;<\; \|f\|_p + \varepsilon$.

Moreover, the left hand inequality holds true for any matrix-valued measurable functions $f, g, h : \mathbb{T} \to \mathcal{L}(\mathbb{C}^n)$ *which satisfy assertion 2(a) above.*

Proof Let $f, g, h : \mathbb{T} \to \mathcal{L}(\mathbb{C}^n)$ be any matrix-valued measurable functions satisfying assertion 2(a), and let $w_f, w_g, w_h : \mathbb{T} \to [0, \infty)$ be the scalar-valued measurable functions given by:

$$w_f(z) \;=\; \|f(z)\|_p; \quad w_g(z) \;=\; \|g(z)\|_q; \quad w_h(z) \;=\; \|h(z)\|_r,$$

for all $z \in \mathbb{T}$. Then by the matrix inequality (1.1) from the start of Chapter 1, we obtain:

$$w_f(z) \;\leq\; w_g(z)\, w_h(z),$$

for almost all $z \in \mathbb{T}$. So in the case $p < \infty$ we have:

$$\begin{aligned} \|f\|_p \;&=\; \|(w_f)^p\|_1^{1/p} \;\leq\; \left(\|(w_g)^p\|_{q/p}\, \|(w_h)^p\|_{r/p}\right)^{1/p} \\ &=\; \|w_g\|_q \|w_h\|_r \\ &=\; \|g\|_q\, \|h\|_r, \end{aligned}$$

by Hölder's inequality for scalar-valued functions. In the case $p = \infty$ we have:

$$\|f\|_\infty \;=\; \|w_f\|_\infty \;\leq\; \|w_g\|_\infty \|w_h\|_\infty \;=\; \|g\|_\infty \|h\|_\infty,$$

by the submultiplicativity of the L^∞ norm. So in both cases, the left hand inequality of assertion 2(b) is satisfied, as required. To complete the proof of the theorem, it remains to show the equivalence of assertions 1 and 2, when $f \in \mathrm{L}^p(\mathcal{L}(\mathbb{C}^n))$.

2 ⇒ 1. By Jensen's inequality, we have:

$$|\det g(z)| \;\leq\; \exp\left(\frac{1}{2\pi}\int_0^{2\pi}\frac{1-|z|^2}{|e^{i\theta}-z|^2}\log|\det g(e^{i\theta})|\,\mathrm{d}\theta\right), \tag{3.45}$$

for all $z \in \mathbb{D}$. (We arrive at the above inequality for $z \neq 0$ by applying a conformal mapping to the domain of g).

Now since g is almost everywhere invertible, and g has almost everywhere convergent radial limits by Fatou's theorem, there is some $\omega \in [0, 2\pi)$ such that $g(e^{i\omega})$ is invertible and $g(re^{i\omega}) \to g(e^{i\omega})$ as $r \to 1^-$. Since the invertibles are open in $\mathcal{L}(\mathbb{C}^n)$, we can therefore find some $r \in [0, 1)$ such that $g(re^{i\omega})$ is invertible. Now $\log|\det g|$ is bounded above pointwise on $\mathbb{T}$ by $\mathrm{Tr}(g^*g)^{1/2}$, which has integral equal to $\|g\|_1 < \infty$. So the integral,

$$J \;=\; \frac{1}{2\pi}\int_0^{2\pi}\log|\det g(e^{i\theta})|\,\mathrm{d}\theta,$$

may be estimated as follows:

$$\begin{aligned}
J \;&=\; \|g\|_1 - \frac{1}{2\pi}\int_0^{2\pi}\left(\mathrm{Tr}(g^*g)^{1/2} - \log|g(e^{i\theta})|\right)\mathrm{d}\theta \\
&\geq\; \|g\|_1 - \frac{1+r}{2\pi(1-r)}\int_0^{2\pi}\frac{1-r^2}{1-2r\cos(\theta-\omega)+r^2}\;\times \\
&\qquad\qquad \left(\mathrm{Tr}(g^*g)^{1/2} - \log|\det g(e^{i\theta})|\right)\mathrm{d}\theta \\
&\geq\; \left[1-\left(\frac{1+r}{1-r}\right)^2\right]\|g\|_1 + \frac{1+r}{2\pi(1-r)}\int_0^{2\pi}\frac{1-r^2}{1-2r\cos(\theta-\omega)+r^2}\;\times \\
&\qquad\qquad \log|\det g(e^{i\theta})|\,\mathrm{d}\theta \\
&=\; \frac{-2r}{(1-r)^2}\|g\|_1 + \frac{1+r}{2\pi(1-r)}\int_0^{2\pi}\frac{1-r^2}{|e^{i\theta}-re^{i\omega}|^2}\log|\det g(e^{i\theta})|\,\mathrm{d}\theta \\
&\geq\; \frac{(1-r^2)\log|\det g(re^{i\omega})| \;-\; 2r\|g\|_1}{(1-r)^2} \;>\; -\infty,
\end{aligned}$$

by inequality (3.45).

Now by the same argument, with h in place of g, we find that:

$$\frac{1}{2\pi}\int_0^{2\pi}\log|\det h(e^{i\theta})|\,\mathrm{d}\theta \;>\; -\infty.$$

Hence,

$$\begin{aligned}\int_0^{2\pi} \log|\det f(e^{i\theta})|\, d\theta &= \int_0^{2\pi} \log\Big(|\det g(e^{i\theta})||\det h(e^{i\theta})|\Big)\, d\theta \\ &= \int_0^{2\pi} \log|\det g(e^{i\theta})|\, d\theta + \int_0^{2\pi} \log|\det h(e^{i\theta})|\, d\theta \\ &> -\infty ,\end{aligned}$$

as required.

1 ⇒ 2. Let $w = (f^*f)^s$ and let $\tilde{w} = (ff^*)^{1-s}$, where $s \in [0,1]$ is chosen so that $1/q = s/p$, or equivalently, $1/r = (1-s)/p$. Then provided $p, q < \infty$, we have:

$$\begin{aligned}\frac{1}{2\pi}\int_0^{2\pi} \operatorname{Tr} w^{q/2}(e^{i\theta})\, d\theta &= \frac{1}{2\pi}\int_0^{2\pi} \operatorname{Tr}(f^*f)^{p/2}(e^{i\theta})\, d\theta \\ &= \|f\|_p^p ,\end{aligned}$$

so that:

$$\|w^{1/2}\|_q = \|f\|_p^s < \infty . \tag{3.46}$$

In the cases $p = \infty$, or $q = \infty$ and $p < \infty$, equation (3.46) holds trivially.

Similarly, we have:

$$\|\tilde{w}^{1/2}\|_r = \|f\|_p^{1-s} < \infty . \tag{3.47}$$

Now we shall construct outer functions $\rho \in \mathrm{H}^q(\mathcal{L}(\mathbb{C}^n))$ and $\tau \in \mathrm{H}^r(\mathcal{L}(\mathbb{C}^n))$ such that $w = \rho^*\rho$ and $\tilde{w} = \tau^*\tau$ almost everywhere. By assertion 1, we have:

$$\begin{aligned}\frac{1}{2\pi}\int_0^{2\pi} \log\det w(e^{i\theta})^{1/2}\, d\theta &= \frac{1}{2\pi}\int_0^{2\pi} \log|\det f(e^{i\theta})|^s\, d\theta \\ &= \frac{s}{2\pi}\int_0^{2\pi} \log|\det f(e^{i\theta})|\, d\theta < \infty ; \\ \frac{1}{2\pi}\int_0^{2\pi} \log\det \tilde{w}(e^{i\theta})^{1/2}\, d\theta &= \frac{1}{2\pi}\int_0^{2\pi} \log|\det f(e^{i\theta})|^{1-s}\, d\theta \\ &= \frac{1-s}{2\pi}\int_0^{2\pi} \log|\det f(e^{i\theta})|\, d\theta < \infty .\end{aligned}$$

So $w^{1/2}$ and $\tilde{w}^{1/2}$ are spectral densities.

Now let $\rho_1, \rho_2, \tau_1, \tau_2 \in \mathrm{H}^2(\mathcal{L}(\mathbb{C}^n))$ be the outer functions given by:

$$\begin{aligned} \rho_1 &= \Phi(w^{1/2})\,; \quad \rho_2 = \Phi(\rho_1\rho_1^*)\,. \\ \tau_1 &= \Phi(\tilde{w}^{1/2})\,; \quad \tau_2 = \Phi(\tau_1\tau_1^*)\,. \end{aligned}$$

And let $\rho, \tau \in \mathrm{H}^1(\mathcal{L}(\mathbb{C}^n))$ be the outer functions given by:

$$\rho = \rho_2\rho_1\,; \quad \tau = \tau_2\tau_1\,.$$

Then we have:

$$\begin{aligned} \rho^*\rho &= \rho_1^*(\rho_2^*\rho_2)\rho_1 = (\rho_1^*\rho_1)(\rho_1^*\rho_1) \quad \text{a.e.} \\ &= w^{1/2}w^{1/2} \quad \text{a.e.} \\ &= w\,; \end{aligned} \tag{3.48}$$

$$\begin{aligned} \tau^*\tau &= \rho_1^*(\tau_2^*\tau_2)\tau_1 = (\tau_1^*\tau_1)(\tau_1^*\tau_1) \quad \text{a.e.} \\ &= \tilde{w}^{1/2}\tilde{w}^{1/2} \quad \text{a.e.} \\ &= \tilde{w}\,. \end{aligned} \tag{3.49}$$

So $\rho \in \mathrm{H}^q(\mathcal{L}(\mathbb{C}^n))$ and $\tau \in \mathrm{H}^r(\mathcal{L}(\mathbb{C}^n))$ with:

$$0 < \|\rho\|_q = \|w^{1/2}\|_q = \|f\|_p^s\ ; \tag{3.50}$$

$$0 < \|\tau\|_r = \|\tilde{w}^{1/2}\|_r = \|f\|_p^{1-s}, \tag{3.51}$$

by equations (3.46) and (3.47).

Now let $\phi : \mathbb{T} \longrightarrow \mathcal{L}(\mathbb{C}^n)$ be the measurable function defined almost everywhere on $\mathbb{T}$ by:

$$\phi = (\tau^{-1})^* f \rho^{-1} \quad \text{a.e.} \tag{3.52}$$

Then from identities (3.48) and (3.49) we have:

$$\begin{aligned} \phi^*\phi = (\rho^{-1})^* f^* \tau^{-1}(\tau^{-1})^* f \rho^{-1} &= (\rho^{-1})^* f^* \tilde{w}^{-1} f \rho^{-1} \\ &= (\rho^{-1})^* f^* (f f^*)^{s-1} f \rho^{-1} \\ &= (\rho^{-1})^* (f^* f)^s \rho^{-1} \\ &= (\rho^{-1})^* \rho^* \rho \rho^{-1} \\ &= I\,, \end{aligned}$$

almost everywhere on $\mathbb{T}$. Hence ϕ is almost everywhere unitary, so by Theorem 3.3.1 there exist functions $\tilde{g}, \tilde{h} \in \mathrm{H}^\infty(\mathcal{L}(\mathbb{C}^n))$ such that:

$$\begin{aligned} \phi &= \tilde{h}^*\tilde{g} \quad \text{a.e.}; \\ \|\tilde{g}\|_\infty, \|\tilde{h}\|_\infty &< (1+\varepsilon/\|f\|_p)^{1/2}. \end{aligned}$$

Now let $g = \tilde{g}\rho$ and let $h = \tilde{h}\tau$. Then we have:

$$\begin{aligned} h^*g &= \tau^*(\tilde{h}^*\tilde{g})\rho = \tau^*\phi\rho \quad \text{a.e.} \\ &= f \quad \text{a.e.}, \end{aligned}$$

by the defining equation (3.52). Thus assertion 2(a) is satisfied.

Now from (3.50) and (3.51) we obtain the estimate:

$$\begin{aligned} \|g\|_q\|h\|_r &\leq \|\rho\|_q \|\tilde{g}\|_\infty \|\tau\|_r \|\tilde{h}\|_\infty \\ &< \|\rho\|_q \cdot (1+\varepsilon/\|f\|_p)^{1/2} \cdot \|\tau\|_r \cdot (1+\varepsilon/\|f\|_p)^{1/2} \\ &\leq (1+\varepsilon/\|f\|_p) \cdot \|f\|_p^s \cdot \|f\|_p^{1-s} \\ &= \|f\|_p + \varepsilon. \end{aligned}$$

Thus g and h satisfy assertion 2(b). This completes the proof. □

In the above theorem, we show that $\|g\|_q\|h\|_r$ can be made arbitrarily close to $\|f\|_p$, for the factorization $f = h^*g$. The next two results show that equality cannot be attained in general. Thus the bounds on the norms of g and h, given by the above theorem, are sharp.

Proposition 3.3.3 *Let $p, q, r \in [1, \infty]$ be indices satisfying $1/p = 1/q + 1/r$, and let $\phi, g, h : \mathbb{T} \to \mathcal{L}(\mathbb{C}^n)$ be measurable matrix-valued functions such that:*

1. *ϕ is almost everywhere unitary;*

2. *$\phi = h^*g$ almost everywhere;*

3. *$\|\phi\|_p = \|g\|_q\|h\|_r$;*

Then λg and $\lambda^{-1}h$ are almost everywhere unitary, for some scalar $\lambda > 0$.

Proof By multiplying g and h by suitable positive scalars, we may assume without loss of generality that $\|g\|_q = n^{1/q}$ and $\|h\|_r = n^{1/r}$. We shall show that in this case, g and h are almost everywhere unitary.

Consider first of all the case $q = \infty$ and $p = r$, and suppose for a contradiction that g is not almost everywhere unitary. Then since $\|g\|_\infty = 1$, there exists a set $S \subset \mathbb{T}$ of positive measure, and some $\varepsilon > 0$, such that:

$$g(z)g(z)^* \;\leq\; (1-\varepsilon)I\,,$$

for all $z \in S$. But now observe that by conditions 1 and 2, the functions g and h are almost everywhere invertible and satisfy:

$$(gg^*)(hh^*) \;=\; I \qquad \text{almost everywhere on } \mathbb{T}. \tag{3.53}$$

Hence,

$$h(z)h(z)^* \;\geq\; \begin{cases} (1-\varepsilon)^{-1}I & \text{for almost all } z \in S; \\ I & \text{for almost all } z \in \mathbb{T} \setminus S. \end{cases}$$

This implies that $\|h\|_r > \|I\|_r = n^{1/r}$, a contradiction, and so g is almost everywhere unitary. Hence h is also a.e. unitary, by conditions 1 and 2.

An analoguous argument holds for the case $r = \infty$ and $p = q$. So now we shall consider the remaining case $p, q, r < \infty$. Since $(x^q - 1)/q, (x^r - 1)/r \geq \log x$ for all $0 < x < \infty$, and since g and h are a.e. invertible, the functions,

$$\begin{aligned} w_1 &= ((gg^*)^{q/2} - I)/q \;-\; \log(gg^*)^{1/2} \\ w_2 &= ((hh^*)^{r/2} - I)/r \;-\; \log(hh^*)^{1/2} \end{aligned}$$

are well defined and positive almost everywhere on $\mathbb{T}$. Taking traces and integrating, we obtain:

$$\begin{aligned} \int_{\mathbb{T}} \operatorname{Tr} w_1 \, d\mu &= (\|g\|_q^q - n)/q - \int_{\mathbb{T}} \operatorname{Tr} \log(gg^*)^{1/2} \, d\mu \\ &= -\int_{\mathbb{T}} \log\det(gg^*)^{1/2} \, d\mu\,; \end{aligned} \tag{3.54}$$

$$\begin{aligned}\int_{\mathbb{T}} \operatorname{Tr} w_2 \, d\mu &= (\|h\|_r^r - n)/r - \int_{\mathbb{T}} \operatorname{Tr} \log(hh^*)^{1/2} \, d\mu \\ &= -\int_{\mathbb{T}} \log\det(hh^*)^{1/2} \, d\mu . \end{aligned} \tag{3.55}$$

By equation (3.53), the integrals (3.54) and (3.55) sum to zero. Hence w_1 and w_2 are almost everywhere zero, which implies that $gg^* = hh^* = I$ almost everywhere, as required. □

Corollary 3.3.4 *Let $\phi \in \mathrm{L}^\infty(\mathcal{L}(\mathbb{C}^n))$ a.e. unitary and suppose that:*

$$0 < \mu\{z \in \mathbb{T} : \phi(z) = I\} < 1 .$$

Let $g, h \in \mathrm{H}^1(\mathcal{L}(\mathbb{C}^n))$ be chosen such that:

$$\phi = h^* g \quad \textit{almost everywhere on } \mathbb{T}.$$

Then we have the strict inequality:

$$\|\phi\|_p < \|g\|_q \|h\|_r ,$$

for all indices $p, q, r \in [1, \infty]$ such that $1/p = 1/q + 1/r$.

Proof As we showed in the proof of Theorem 3.3.2, the inequality:

$$\|\phi\|_p \leq \|g\|_q \|h\|_r$$

holds for all indices $p, q, r \in [1, \infty]$ such that $1/p = 1/q + 1/r$. So suppose for a contradiction that $\|\phi\|_p = \|g\|_q \|h\|_r$, for some such choice of indices p, q and r. Then by Proposition 3.3.3, there is some scalar $\lambda > 0$ such that the functions λg and $\lambda^{-1} h$ are almost everywhere unitary on $\mathbb{T}$, and so they are matrix-valued inner functions.

But now since $h^* = \lambda^2 h^{-1}$ almost everywhere, we have $\lambda^2 g(z) = h(z)$ for almost all $z \in \mathbb{T}$ such that $\phi(z) = I$. Therefore $\lambda^2 g - h \in \mathrm{H}^\infty(\mathcal{L}(\mathbb{C}^n))$ vanishes on a set of positive measure, which implies that $\lambda^2 g = h$ almost everywhere. Hence $\phi = \lambda^2 h^{-1} g = I$ a.e. on $\mathbb{T}$, contradicting the hypotheses satisfied by ϕ. □

3.4 Conclusion

It may be a little surprising that inner and outer functions are sufficent to factorize unitary matrix-valued functions ϕ in combination, but not individually. The essential reason for this ability is that the approximations to ϕ that they provide are of a complementary nature. Outer functions which are close to 1 are very good at correcting a small uniform error in the factorization of ϕ over most, but not all, of the circle. On the other hand, inner functions which are close to 1 are very good at correcting a large uniform error in the factorization of ϕ, provided that it occurs over a small subset of the circle.

It is also notable that nearly all the essential features of the Douglas-Rudin problem and its solution, in the scalar case, carry through to the matrix case without significant alteration, in spite of the difficulties posed by the noncommutativity of matrix multiplication. The main difference between the general outline of our method and Bourgain's method given in [6], is that Bourgain works with sums and differences of the pointwise *arguments* of unimodular functions, whereas we work directly with products of the functions themselves. This is slightly harder, but it is necessary due to the noncommutativity of matrix multiplication. The main outer and inner factorization results from Sections 3.1 and 3.2 are derived from Proposition 3.1.3 and Lemma 3.2.2, respectively. These are almost identical to the two lemmas used in [6].

Finally, we remark that in contrast to spectral factorization, the factorization of a log-integrable function $f \in \mathrm{L}^\infty$ into the form h^*g, for functions $g, h \in \mathrm{H}^\infty$, is in general far from unique. This is clear from the fact that $\|g\|_\infty \|h\|_\infty$ can be made arbitrarily close to $\|f\|_\infty$, but cannot attain it in general. It may be interesting to investigate the class of possible factorizations h^*g of a given fixed function $f \in \mathrm{L}^\infty(\mathcal{L}(\mathbb{C}^n))$. For example, if $f \equiv I$ or f is scalar-valued, then this becomes the problem of determining when an a.e. unitary-valued product $\tau^*\rho$, for outer functions $\rho, \tau \in \mathrm{H}^\infty(\mathcal{L}(\mathbb{C}^n))$, may be expressed as a ratio $\theta\Theta^{-1}$ of inner functions $\Theta, \theta \in \mathrm{H}^\infty(\mathcal{L}(\mathbb{C}^n))$.

3.5 Unimodular functions without a simple factorization

Here we give an example of a uniformly dense subset of the measurable unimodular functions on $\mathbb{T}$, which do not have a simple factorization of the form $\overline{\theta_2}\theta_1$ or $\overline{\rho_2}\rho_1$, for inner $\theta_1, \theta_2 \in \mathrm{H}^\infty$ or outer $\rho_1, \rho_2 \in \mathrm{H}^\infty$:

Proposition 3.5.1 *Let $N \in \mathbb{N}$ and let $f : \mathbb{T} \to \mathbb{C}$ measurable and not a.e. constant, chosen such that $f(z)^N = 1$ for all $z \in \mathbb{T}$. Then f does not have a factorization of the form $\overline{\theta_2}\theta_1$ or $\overline{\rho_2}\rho_1$, for inner $\theta_1, \theta_2 \in \mathrm{H}^\infty$ or outer $\rho_1, \rho_2 \in \mathrm{H}^\infty$.*

Proof *Inner factorization:* Suppose that $f = \overline{\theta_2}\theta_1$ for inner $\theta_1, \theta_2 \in \mathrm{H}^\infty$. Then we have:

$$f \cdot \theta_2 - \theta_1 = 0,$$

everywhere on $\mathbb{T}$. But f has finite range, so there exists some $\lambda \in \mathbb{C}$ and some $K \subset \mathbb{T}$ measurable with $\mu(K) > 0$, such that $f(z) = \lambda$ for all $z \in K$. Hence the linear combination:

$$\lambda\theta_2 - \theta_1$$

is zero everywhere on K. But the above expression lies in H^∞, so we find that it is zero almost everywhere on $\mathbb{T}$. Hence $f = \lambda$ almost everywhere, contradicting the hypothesis that f is not a.e. constant.

Outer factorization: Now suppose that $f = \overline{\rho_2}\rho_1$ for outer $\rho_1, \rho_2 \in \mathrm{H}^\infty$. Since f is everywhere unimodular, we have:

$$|\rho_2| \cdot |\rho_1| = |\rho_2\rho_1| = 1,$$

everywhere on $\mathbb{T}$. Hence the outer function $\rho_2\rho_1$ is almost everywhere constant. This implies that:

$$f = \overline{\rho}^{-1}\rho,$$

for some outer function $\rho \in \mathrm{H}^\infty$ with inverse $\rho^{-1} \in \mathrm{H}^\infty$. Raising both sides to the power of N, we find that:

$$\overline{\rho}^{-N} \rho^N = f^N \equiv 1 .$$

Hence,

$$\rho^N = \overline{\rho^N} \in \mathrm{H}^\infty .$$

This implies that ρ^N has a real-valued analytic extension to $\mathbb{D}$, which must therefore be constant. Hence ρ^N is constant almost everywhere, which implies that ρ and therefore f are also constant almost everywhere. This contradicts the hypotheses on f, as before. □

We remark that if the assumption that the outer functions ρ_1 and ρ_2 are bounded is dropped, then the above result no longer holds. For example, consider the function $f : \mathbb{T} \to \mathbb{C}$ given by:

$$f(z) = \begin{cases} i & \text{if } \mathrm{Im}(z) \geq 0 \\ -i & \text{if } \mathrm{Im}(z) < 0 , \end{cases}$$

for all $z \in \mathbb{T}$. This has a factorization of the form $\overline{\rho}^{-1}\rho$, where:

$$\rho(z) = \sqrt{\frac{1+z}{1-z}} \qquad \text{for all } z \in \overline{\mathbb{D}} \setminus \{1, -1\} ,$$

where the above square root takes values in the right half-plane. This defines ρ and ρ^{-1} almost everywhere on $\mathbb{T}$ as unbounded outer functions, with $\rho, \rho^{-1} \in \mathrm{H}^p$ for all indices $p < 2$.

Chapter 4

Vector-valued H^∞ approximation

In this chapter, we turn to the problem of weighted H^∞ approximation of vector-valued L^∞ functions on the unit circle, subject to a weighted sup-type constraint on the size of the approximant. We shall establish existence of a solution under mild assumptions, and characterise some of its properties.

The layout of this chapter is as follows. In Section 4.1, we look at the problem of unconstrained weighted H^∞ approximation of a scalar-valued continuous function on the unit circle. We show that for weights without a bounded inverse, this problem is not generally well-posed. This provides motivation for the abstract approximation problem studied in Section 4.2, where we replace H^∞ and L^∞ with general Banach spaces equipped with preduals, and impose a constraint to make the problem well-posed. We finish the section with an example of an entirely different approximation problem looked at elsewhere in the literature [10], which also happens to fit into our abstract framework.

In the first half of Section 4.3, we develop a suitable theory of vector-valued Hardy and Lebesgue spaces $H^\infty(E)$ and $L^\infty(E)$, where E is any Banach space with separable predual. In particular, we will show that if Y is a weak* closed subspace of a Banach space X with separable predual X_*, then $H^\infty(Y)$ may be regarded as a weak* closed subspace of $L^\infty(X)$, which has predual $L^1(X_*)$. This will allow these spaces to be used in the abstract framework developed earlier. In the second half of Section 4.3, we define certain sup-type seminorms to use on these spaces, and show that they are in one-to-one correspondence with a class of measurable weights on $\mathbb{T}$, giving a seminorm on X for each point in $\mathbb{T}$.

In Section 4.4, we place the spaces and the seminorms constructed in Section 4.3 into the abstract framework of Section 4.2, giving rise to a general constrained vector-valued H^∞ approximation problem, together with sufficient conditions for the existence of a solution, taken from the abstract setting. With a certain extra condition, we then show that the solutions to this problem saturate the constraints pointwise on $\mathbb{T}$, in a particular sense. This leads on to a uniqueness result under further conditions on the weights used.

Finally, in Section 4.5 we specialize the setup further in a number of examples, including matrix-valued constrained approximation problems with weights given by nonconstant left and right pointwise matrix multiplication, and bounded extremal problems where the two weights are taken to vanish on complementary subsets of the circle.

4.1 Unconstrained scalar H^∞ approximation

In this section, we consider the problem of finding the best H^∞ approximant to a given scalar function $\phi \in C$, where the error of the approximation is to be minimized with respect to a given weighted seminorm $\|\cdot\|_{\infty,w}$ on L^∞. We define such seminorms as follows:

Let $w \in L^\infty$ be any almost everywhere nonnegative, bounded scalar-valued function on $\mathbb{T}$. Then we set

$$\|\psi\|_{\infty,w} = \|w\,\psi\|_\infty ,$$

for any scalar function $\psi \in L^\infty$. We denote the set of all such almost everywhere nonnegative, bounded scalar weights w by L^∞_{pos}.

The approximation problem under consideration is stated as follows:

Problem 4.1.1 (Weighted Nehari) *For given scalar valued functions $w \in$*

$\mathrm{L}^\infty_{\mathrm{pos}}$ *and* $\phi \in \mathrm{C}$, *find a function* $g_0 \in \mathrm{H}^\infty$ *such that*

$$\|\phi - g_0\|_{\infty,w} = \inf\{\|\phi - g\|_{\infty,w} : g \in \mathrm{H}^\infty\} .$$

In the case when w has a bounded inverse, it may be shown that the above problem reduces to the standard Nehari problem of finding the best uniform approximation in H^∞, to a given scalar L^∞ function on the unit circle. This leads to the existence of a unique solution, described in the next result.

In the case when w does not have a bounded inverse, the problem is not generally well posed. That is, the infimum in the above statement of the problem will not be attained, in general. This will be shown later in this section.

First of all, we look at the former case. We omit the proof of the following result, since it is just a straightforward application of the basic theory of Hankel operators (c.f. [26, ch. 1]).

Theorem 4.1.2 *Let* $w \in \mathrm{L}^\infty_{\mathrm{pos}}$ *be a bounded scalar weight with bounded inverse, and let* $\rho \in \mathrm{H}^\infty$ *be the outer function given by the formula:*

$$\rho(z) = \exp\left(\frac{1}{2\pi}\int_0^{2\pi} \frac{e^{i\theta}+z}{e^{i\theta}-z}\log w(e^{i\theta})\,\mathrm{d}\theta\right),$$

for all $z \in \mathbb{D}$, *so that* $|\rho| = w$ *almost everywhere on* $\mathbb{T}$. *Then for any scalar valued function* $\phi \in \mathrm{C}$, *there is a unique* $g_0 \in \mathrm{H}^\infty$ *such that*

$$\|\psi\|_{\infty,w} = \inf\{\|\phi - g\|_{\infty,w} : g \in \mathrm{H}^\infty\} ,$$

where $\psi = \phi - g_0$.

The error satisfies $|\psi(z)| = \|\psi\|_{\infty,w}\cdot w(z)^{-1}$ *for almost all* $z \in \mathbb{T}$, *and* g_0 *is given by the formula:*

$$g_0 = \phi - \frac{\Gamma_{\phi\rho} f}{\rho f}, \tag{4.1}$$

where $f \in \mathrm{H}^2$ *is any maximizing vector of the compact Hankel operator* $\Gamma_{\phi\rho}$. *Furthermore, the norm* $\|\psi\|_{\infty,w}$ *of the error* ψ *is equal to* $\|\Gamma_{\phi\rho}\|$.

We shall now give a result which is in contrast to the above theorem. It concerns the case when the weight does not have a bounded inverse.

Theorem 4.1.3 *Let* $w \in \mathrm{L}^\infty_{\mathrm{pos}}$ *be a bounded scalar weight without a bounded inverse, and let* $\phi \in \mathrm{C}$ *be any scalar valued function such that the infimum,*

$$N \;=\; \inf\left\{\|\phi - g\|_{\infty,w} \,:\, g \in \mathrm{H}^\infty\right\}, \tag{4.2}$$

is greater than 0. *Let* $(g_n)_{n=1}^\infty$ *be any sequence of functions in* H^∞ *such that*

$$\|\phi - g_n\|_{\infty,w} \longrightarrow N\,, \quad \text{as } n \to \infty\,.$$

Then $\|g_n\|_\infty \to \infty$ *as* n *tends to* ∞. *Hence the infimum cannot be attained.*

Proof Suppose, for a contradiction, that $\|g_n\|_\infty$ does not diverge as $n \to \infty$. Then we may choose a subsequence $(g_{n_k})_{k=1}^\infty$ which is uniformly bounded. So let us assume without loss of generality, that $\|g_n\|_\infty \le R$ for all $n \in \mathbb{N}$, for some bound $R > 0$.

Let $\tilde{w} \in \mathrm{L}^\infty_{\mathrm{pos}}$ be the bounded weight with bounded inverse, defined by:

$$\tilde{w}(z) \;=\; w(z) \vee N(R + \|\phi\|_\infty)^{-1}\,,$$

for all $z \in \mathbb{T}$. Now we have:

$$\begin{aligned}
\tilde{w}(z)|\phi(z) - g_n(z)| \;&=\; (w(z) \vee N(R + \|\phi\|_\infty)^{-1})\,|\phi(z) - g_n(z)| \\
&\le\; w(z)|\phi(z) - g_n(z)| \;\vee\; N(R + \|\phi\|_\infty)^{-1}\,\|\phi - g_n\|_\infty \\
&\le\; w(z)|\phi(z) - g_n(z)| \;\vee\; N\,,
\end{aligned}$$

for all $z \in \mathbb{T}$ and $n \in \mathbb{N}$. Hence

$$\begin{aligned}
\|\phi - g_n\|_{\infty,\tilde{w}} \;&\le\; \|\phi - g_n\|_{\infty,w} \vee N \\
&=\; \|\phi - g_n\|_{\infty,w} \\
&\le\; \|\phi - g_n\|_{\infty,\tilde{w}}\,,
\end{aligned}$$

so we have equality throughout, for each $n \in \mathbb{N}$. This shows, in particular, that $\|\phi - g_n\|_{\infty,\tilde{w}} \to N$ as n tends to ∞.

Now by Theorem 4.1.2, the function ϕ has a unique best H^∞ approximant g, with respect to the weight $\tilde{w}$. Letting $\psi = \phi - g$, we have:

$$N \;\le\; \|\psi\|_{\infty,w} \;\le\; \|\psi\|_{\infty,\tilde{w}} \;\le\; \lim_{n\to\infty} \|\phi - g_n\|_{\infty,\tilde{w}} \;=\; N\,,$$

so that $\|\psi\|_{\infty,w} = \|\psi\|_{\infty,\tilde{w}} = N$. This shows that the infimum in equation (4.2) is attained by g. Now, for any $n \in \mathbb{N}$ such that $\|\phi - g_n\|_{\infty,w} = N$, we also have $\|\phi - g_n\|_{\infty,\tilde{w}}$ equal to N. Hence $g_n = g$, by uniqueness of g. This implies that all such g_n are equal.

A further consequence of Theorem 4.1.2 is that $|\phi(z) - g(z)| = N\tilde{w}(z)$, for almost all $z \in \mathbb{T}$. Now observe that since w has no bounded inverse, $\tilde{w}$ varies with the choice of the upper bound R we take. Hence g also varies with R. This implies that the infimum in equation (4.2) is not uniquely attained.

However, this gives a contradiction, since any choice of $(g_n)_{n=1}^{\infty}$ taking only finitely many values and such that $\|\phi - g_n\|_{\infty,w} = N$ for all $n \in \mathbb{N}$, would satisfy the hypotheses of the preceding argument. It would then follow that all the g_n were equal, which implies that the infimum in (4.2) is uniquely attained. □

We remark that the above result also holds when the infimum in equation (4.2) is zero, provided that ϕ does not agree with any H$^\infty$ function on the complement of the zero set of w. Indeed, if $\|g_n\|_\infty$ does not diverge as $n \to \infty$, then the sequence $(g_n)_{n=1}^{\infty}$ has a weak* accumulation point $g \in \mathrm{H}^\infty$, and it can be easily shown that $\|\phi - g\|_{\infty,w}$ is zero.

Finally in this section, we derive a result from Theorem 4.1.2 which will be of use to us much later in the chapter. It concerns the special case where the function $\phi \in \mathrm{C}$, to be approximated in H$^\infty$, is given by $\phi(z) = z^{-1}$ for all $z \in \mathbb{T}$.

Lemma 4.1.4 *Let $w \in \mathrm{L}^\infty_{\mathrm{pos}}$ be a bounded scalar weight with bounded inverse, and let $\rho \in \mathrm{H}^\infty$ be the outer function given by the formula:*

$$\rho(z) \;=\; \exp\left(\frac{1}{2\pi}\int_0^{2\pi} \frac{e^{i\theta}+z}{e^{i\theta}-z}\log w(e^{i\theta})\,\mathrm{d}\theta\right),$$

for all $z \in \mathbb{D}$, as in the statement of Theorem 4.1.2.

Let $\phi \in \mathrm{C}$ be given by $\phi(z) = z^{-1}$ for all $z \in \mathbb{T}$. Then the unique best weighted approximant $g_0 \in \mathrm{H}^\infty$ solving Problem 4.1.1, for this choice of w and ϕ, is given by:

$$g_0 \;=\; \phi\cdot(\rho - \rho(0))\cdot\rho^{-1}.$$

Thus the weighted norm of the error $\psi = \phi - g_0$ is equal to:

$$\|\psi\|_{\infty,w} = \rho(0) = \exp\left(\frac{1}{2\pi}\int_0^{2\pi} \log w(e^{i\theta})\, d\theta\right).$$

Proof It is clear that g_0 is a well defined element of H^∞, with error ψ equal to $\phi\rho(0)\rho^{-1}$, which satisfies $\|\psi\|_{\infty,w} = \|\phi\rho(0)\|_\infty = \rho(0)$. So to show that g_0 is the unique solution to Problem 4.1.1, it is enough by Theorem 4.1.2 to establish equality between $\rho(0)$ and the norm of the Hankel operator $\Gamma_{\phi\rho}$.

Now $\Gamma_{\phi\rho}$ is equal to $\Gamma_\phi \mathrm{T}_\rho$, where T_ρ is the Toeplitz operator with analytic symbol ρ, as in the proof of Theorem 4.1.2. Writing these operators as infinite matrices with respect to the standard orthonormal bases of H^2 and $(\mathrm{H}^2)^\perp$, we have:

$$\Gamma_\phi = \begin{pmatrix} 1 & 0 & 0 & \cdots \\ 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

$$\mathrm{T}_\rho = \begin{pmatrix} \hat{\rho}(0) & 0 & 0 & \cdots \\ \hat{\rho}(1) & \hat{\rho}(0) & 0 & \cdots \\ \hat{\rho}(2) & \hat{\rho}(1) & \hat{\rho}(0) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

where $\hat{\rho}(n)$ is the Fourier coefficient of ρ with index n, for any $n \in \mathbb{Z}$.

From the above matrices, we can easily see that $\Gamma_\phi \mathrm{T}_\rho = \hat{\rho}(0)\Gamma_\phi = \rho(0)\Gamma_\phi$, so that $\|\Gamma_{\phi\rho}\| = \|\Gamma_\phi \mathrm{T}_\rho\| = \rho(0)$, as required. □

4.2 Constrained approximation on Banach spaces

In the previous section, we considered the problem of finding the best H^∞ approximant to a given continuous, scalar L^∞ function on the unit circle, with respect to a bounded weight. It was shown that the problem is not well posed in general, unless the weight has a bounded inverse. More precisely, it was shown that the norm of a suboptimal H^∞ approximant tends to infinity, whenever the weighted norm of its error approaches the infimum over all H^∞ approximants, in the case when the weight has no bounded inverse and the infimum is nonzero.

This suggests that when the weight has no bounded inverse, we should place a constraint on the approximant, to make the problem well posed. In this section, we shall consider an abstract version of such a constrained approximation problem, working with general Banach spaces (with preduals), rather than H^∞ and L^∞. The setup will be as follows:

- Let X be a Banach space with fixed predual X_*. This will take the role of L^∞ in the concrete problem.
- Let $Y \subset X$ be a weak* closed subspace of X. This will take the role of H^∞ in the concrete problem.
- Let $\|\cdot\|_A$ and $\|\cdot\|_B$ be two seminorms on X. The first seminorm takes the role of the weighted norm used in the concrete problem. The second seminorm will be used to form the constraining condition.
- Let $x_A, x_B \in X$. The element x_A will take the role of the L^∞ function to be approximated. The element x_B will be used to form the constraining condition along with $\|\cdot\|_B$.

In addition, we place the following two conditions on $\|\cdot\|_A$ and $\|\cdot\|_B$:

1. There is a constant $\delta > 0$ such that $\|x\|_A \vee \|x\|_B \geq \delta\|x\|$, for all $x \in X$. In other words,

$$\|\cdot\|_A \vee \|\cdot\|_B \quad (\text{or equivalently } \|\cdot\|_A + \|\cdot\|_B)$$

   is bounded below.

2. $\|\cdot\|_A$ and $\|\cdot\|_B$ are weak* lower semicontinuous.

Here we define the *topology of lower semicontinuity* to be the non-Hausdorff topology on $\mathbb{R}$, generated by open intervals of the form (λ, ∞) for $\lambda \in \mathbb{R}$. Thus condition 2 above is equivalent to the assertion that the unit balls of $\|\cdot\|_A$ and $\|\cdot\|_B$ are weak* closed. In particular, the unit balls are norm closed. As the following straightforward result shows, this implies that $\|\cdot\|_A$ and $\|\cdot\|_B$ are bounded. Hence $\|\cdot\|_A \vee \|\cdot\|_B$ is equivalent to the original norm on X.

Proposition 4.2.1 *Let p be any positive homogeneous, nonnegative convex functional on a Banach space E. Then the following are equivalent:*

1. *The convex set B, equal to $\{x \in E : p(x) \leq 1\}$, is norm closed.*
2. *B has nonempty interior.*
3. *p is norm bounded.*
4. *p is norm continuous.*

[*Remark:* Observe that p is the Minkowski functional of the set B.]

Proof *1 ⇒ 2.* Suppose that B has empty interior. Then

$$E \;=\; B \cup 2B \cup 3B \cup \ldots$$

is a countable union of closed, nowhere dense sets and therefore meagre. But E is complete, so by the Baire category theorem this gives a contradiction.

2 ⇒ 3. Let x be an interior point of B. Then for all $\lambda > 0$, the origin 0 is an interior point of the convex set $\lambda(B - x)$. But $-\lambda x \in B$ for λ sufficiently small, so we may choose $\lambda < 1/2$ such that

$$\lambda B - \lambda x \;\subset\; (1-\lambda)B - \lambda x \;\subset\; B\,,$$

by convexity of B. Hence 0 is an interior point of B.

Equivalently, there is some $R \geq 0$ such that for all $x \in E$,

$$\|Rx\| \leq 1 \;\;\Rightarrow\;\; x \in B \;\;\Leftrightarrow\;\; p(x) \leq 1\,.$$

Hence $p(x) \leq R\,\|x\|$ for all $x \in E$, so p is norm bounded.

3 $\Rightarrow$ *4.* By the subadditivity of p, we have for all $x, y \in E$,

$$p(x) - p(-y) \;\leq\; p(x+y) \;\leq\; p(x) + p(y)\,,$$

and therefore

$$p(x) - R\,\|y\| \;\leq\; p(x+y) \;\leq\; p(x) + R\,\|y\|\,.$$

Hence $p(x+y) \longrightarrow p(x)$ whenever $y \longrightarrow 0$ in norm, so p is norm continuous.

4 $\Rightarrow$ *1.* Trivial, since B is the preimage under p of the closed interval $[0, 1]$. □

We remark that there are norms on the Banach space ℓ^∞ (with predual ℓ^1) which are equivalent to the uniform norm, but are not weak* lower semicontinuous. This shows that condition 2 on $\|\cdot\|_A$ and $\|\cdot\|_B$, cannot be derived from the fact that $\|\cdot\|_A \vee \|\cdot\|_B$ is equivalent to the original norm.

With the setup described, the abstract constrained approximation problem is stated as follows:

Problem 4.2.2 (Abstract constrained approximation) *Given the spaces $Y \subset X$, the seminorms $\|\cdot\|_A$, $\|\cdot\|_B$ and the vectors $x_A, x_B \in X$ as defined earlier, and given $M > 0$, find $y_0 \in Y$ which minimizes:*

$$\|x_A - y_0\|_A\,,$$

subject to the constraint that:

$$\|x_B - y_0\|_B \;\leq\; M\,.$$

We shall show that, provided the bound M is sufficiently large, a solution y_0 to this problem always exists. In fact the following result holds, which we shall prove next in this section:

Theorem 4.2.3 *Let X, Y, $\|\cdot\|_A$, $\|\cdot\|_B$ and $x_A, x_B \in X$ be given as earlier, and let $f : [0,\infty] \to [0,\infty]$ be the function defined by:*

$$f(M) \;=\; \inf\{\|x_A - y\|_A \,:\, y \in Y \text{ and } \|x_B - y\|_B \le M\}\,, \qquad (4.3)$$

where this infimum is taken to be ∞ whenever the above set is empty.

Then $f(M)$ is finite for sufficiently large $M < \infty$, and the infimum is attained whenever M and $f(M)$ are both finite. Moreover, f is decreasing and convex, and letting

$$M_0 \;=\; \inf\{M \in [0,\infty) \,:\, f(M) < \infty\}\,,$$

it holds that f is continuous from $[M_0, \infty]$ to $[0,\infty]$, with respect to the standard compact topologies on each of these intervals.

Note that for a given bound M, any $y_0 \in Y$ for which the infimum $f(M)$ in the above theorem is attained, corresponds to a solution to Problem 4.2.2, with respect to this bound. Note also that the case $M = \infty$ is just an unconstrained version of the abstract problem. The above theorem implies that $f(M) \to f(\infty)$ as M tends to ∞.

Proof of Theorem 4.2.3 We shall first show that f is decreasing and convex.

Let $0 \le M_1 \le M_2 \le \infty$ and let $\lambda, \mu \in (0,1)$ with $\lambda + \mu = 1$. We need to show that:

$$f(M_2) \;\le\; f(M_1) \qquad (4.4)$$

$$f(\lambda M_1 + \mu M_2) \;\le\; \lambda f(M_1) + \mu f(M_2) \qquad (4.5)$$

If $f(M_1) = \infty$ then inequalities (4.4) and (4.5) follow trivially. So suppose that $f(M_1)$ is finite. Then for any $\varepsilon > 0$, there is a $y_1 \in Y$ such that:

$$\|x_A - y_1\|_A \;<\; f(M_1) + \varepsilon$$

$$\|x_B - y_1\|_B \;\le\; M_1 \;\le\; M_2$$

Hence $f(M_2) < f(M_1) + \varepsilon$ by the definition of f. Since ε is arbitrary, this establishes inequality (4.4).

Now since $f(M_2)$ is finite, we may choose $y_2 \in Y$ such that:

$$\begin{aligned} \|x_A - y_2\|_A &< f(M_2) + \varepsilon \\ \|x_B - y_2\|_B &\leq M_2 \end{aligned}$$

Let $z = \lambda y_1 + \mu y_2$. By the triangle inequality, we have:

$$\begin{aligned} \|x_A - z\|_A &= \|\lambda(x_A - y_1) + \mu(x_A - y_2)\|_A \\ &\leq \lambda\|x_A - y_1\|_A + \mu\|x_A - y_2\|_A \\ &< \lambda\,(f(M_1) + \varepsilon) + \mu\,(f(M_2) + \varepsilon) \\ &= \lambda f(M_1) + \mu f(M_2) + \varepsilon \end{aligned}$$

$$\begin{aligned} \|x_B - z\|_B &= \|\lambda(x_B - y_1) + \mu(x_B - y_2)\|_B \\ &\leq \lambda\|x_B - y_1\|_B + \mu\|x_B - y_2\|_B \\ &\leq \lambda M_1 + \mu M_2 \end{aligned}$$

Hence

$$f(\lambda M_1 + \mu M_2) \quad \leq \quad \lambda f(M_1) + \mu f(M_2) + \varepsilon\,.$$

Since ε is arbitrary, this establishes inequality (4.5), as required. So f is decreasing and convex.

Now by setting $y = 0$ in the definition of f, given by equation (4.3), we find that $f(\|x_B\|_B) \leq \|x_A\|_A$. Hence M_0 is well defined, and $f(M)$ is finite for all $M \in (M_0, \infty]$. Since f is convex, this implies that f is continuous on (M_0, ∞). To show continuity of f at ∞, observe that since f is a decreasing function, $f(M)$ has some finite limit N as $M \to \infty$, and we have $f(\infty) \leq N$. Now for any $\varepsilon > 0$, there exists $y \in Y$ such that

$$\|x_A - y\|_A < f(\infty) + \varepsilon\,.$$

This gives us the inequalities:

$$N \leq f(\|x_B - y\|_B) \leq \|x_A - y\|_A \leq f(\infty) + \varepsilon\,.$$

Since ε is arbitrary, we find that $f(\infty) = N$, so f is continuous at ∞. We shall establish continuity of f at M_0 in the final part of the proof.

We shall now show that the infimum in (4.3) is attained, whenever M and $f(M)$ are finite. For any such bound M, define the solution set:

$$K(M) \;=\; \{y \in Y \,:\, \|x_A - y\|_A = f(M) \text{ and } \|x_B - y\|_B \leq M\}.$$

Similarly, for any $\varepsilon > 0$, we define:

$$K_\varepsilon(M) \;=\; \{y \in Y \,:\, \|x_A - y\|_A \leq f(M) + \varepsilon \text{ and } \|x_B - y\|_B \leq M\}.$$

Now by condition 1 on the seminorms $\|\cdot\|_A$ and $\|\cdot\|_B$, there is a $\delta > 0$ such that $\|x\|_A \vee \|x\|_B \;>\; \delta\|x\|$ for all $x \in X$. So for any $y \in K_\varepsilon(M)$, we have:

$$\begin{aligned}
\delta\|y\| &\leq \|y\|_A \vee \|y\|_B \\
&\leq (\|x_A - y\|_A + \|x_A\|_A) \;\vee\; (\|x_B - y\|_B + \|x_B\|_B) \\
&\leq (\|x_A - y\|_A \vee \|x_B - y\|_B) \;+\; (\|x_A\|_A \vee \|x_B\|_B) \\
&\leq M + f(M) + \|x_A\|_A + \|x_B\|_B + \varepsilon\,.
\end{aligned}$$

Since this bound is independent of y, we find that $K_\varepsilon(M)$ is a norm bounded set. Now condition 2 (weak* lower semicontinuity) on $\|\cdot\|_A$ and $\|\cdot\|_B$, and the weak* closure of Y, implies that $K_\varepsilon(M)$ is weak* closed in X. Hence $K_\varepsilon(M)$ is weak* compact in X, by Alaoglu's Theorem.

But $K_\varepsilon(M)$ is also nonempty for all $\varepsilon > 0$, by definition of $f(M)$. Hence the solution set,

$$K(M) \;=\; \bigcap_{\varepsilon>0} K_\varepsilon(M)\,,$$

is the intersection of a chain of weak* compact, nonempty sets, and therefore also weak* compact and nonempty. So the infimum in (4.3) is attained, as required.

Finally, to establish continuity of f at M_0, first observe that since f is decreasing, $f(M) \to L$ as M tends to M_0 from above, for some $L \leq f(M_0)$. So it suffices to show that $f(M_0)$ is no greater than L. Assume without loss of generality that L is finite. Then for any $\varepsilon, \eta > 0$, we have $f(M_0 + \eta) \leq L$, so there exists $y \in Y$ such that:

$$\begin{aligned} \|x_A - y\|_A &< f(M_0 + \eta) + \varepsilon \leq L + \varepsilon \\ \|x_B - y\|_B &\leq M_0 + \eta \end{aligned}$$

Letting $\eta \longrightarrow 0$, we find that:

$$\inf \{\|x_B - y\|_B : y \in Y \text{ and } \|x_A - y\|_A \leq L + \varepsilon\} = M_0 .$$

But M_0 and $L + \varepsilon$ are both finite, so by reversing the roles of $(x_A, \|\cdot\|_A)$ and $(x_B, \|\cdot\|_B)$, and by applying the existence of a solution to Problem 4.2.2, as we established above, we find that there is a $y \in Y$ for which the above infimum is attained. Hence

$$f(M_0) \leq L + \varepsilon ,$$

for all $\varepsilon > 0$. This establishes equality of L and $f(M_0)$, as required. □

4.2.1 An example problem

Now we shall look at an example of a constrained approximation problem which can be placed into the abstract framework of this section, under certain conditions. This problem is considered in [10]. It is stated as follows:

Problem 4.2.4 *Let X, X_1 and X_2 be Banach spaces, let $A \in \mathcal{L}(X, X_1)$ and let $B \in \mathcal{L}(X, X_2)$. Now given $x_1 \in X_1$, $x_2 \in X_2$ and $M \in [0, \infty]$, find $x \in X$ which minimizes:*

$$\|Ax - x_1\| ,$$

subject to the constraint that:

$$\|Bx - x_2\| \leq M .$$

In order to place this problem into the earlier framework, we will need to make the following assumptions:

1. The spaces X, X_1 and X_2 are equipped with fixed preduals X_*, X_{1*} and X_{2*} respectively.

2. The operators A and B are weak* continuous, with respect to the above preduals. Equivalently, there exist *preadjoints* $A_* \in \mathcal{L}(X_{1*}, X_*)$ and $B_* \in \mathcal{L}(X_{2*}, X_*)$ to A and B respectively. Thus $A = (A_*)^*$ and $B = (B_*)^*$.

3. There is a constant $\delta > 0$ such that $\|Ax\| + \|Bx\| \geq \delta\|x\|$, for all $x \in X$.

A typical setup for the problem would take X to be reflexive. In this case, the spaces X_1 and X_2 can, without loss of generality, be replaced with their double duals, if necessary. Then the first two assumptions above hold automatically. This is, in fact, the same set of assumptions made in [10], minus their extra assumptions that X_1 and X_2 are smooth and A has dense range.

We will need the following basic result, which is a straightforward corollary of the Banach closed range theorem [5] [15, pp. 487–489]:

Proposition 4.2.5 *Let* $\mathfrak{X} = X_1 \oplus X_2$ *and let* $\mathfrak{Y} \subset \mathfrak{X}$ *be the range of the mapping of* X *into* $\mathfrak{X}$ *given by:*

$$x \mapsto (Ax, Bx),$$

for all $x \in X$. *Then with the three assumptions above, this mapping is an embedding and the subspace* $\mathfrak{Y}$ *is weak* closed, with respect to the natural predual* $\mathfrak{X}_* = X_{1*} \oplus X_{2*}$ *of the space* $\mathfrak{X}$.

Proof It follows clearly from the two assumptions on A and B that the above mapping is a weak* continuous embedding. It therefore has closed range and possesses a preadjoint in $\mathcal{L}(\mathfrak{X}_*, X)$, given explicitly by the mapping:

$$(\alpha, \beta) \mapsto A_*\alpha + B_*\beta \quad \text{for all } (\alpha, \beta) \in \mathfrak{X}_*.$$

So by the closed range theorem, the preadjoint has closed range and hence the embedding has weak* closed range, given by the annihilator of the kernel of its preadjoint. Thus $\mathfrak{Y}$ is weak* closed, as required. □

Now with the spaces $\mathfrak{X}$ and $\mathfrak{Y}$ given by the above proposition, we may define the seminorms $\|\cdot\|_A$ and $\|\cdot\|_B$ on $\mathfrak{X}$ by:

$$\begin{aligned}\|(u,v)\|_A &= \|u\|_{X_1}\\ \|(u,v)\|_B &= \|v\|_{X_2}\end{aligned}$$

for all $(u,v) \in \mathfrak{X}$. These are clearly weak* lower semicontinuous and the norm $\|\cdot\|_A \vee \|\cdot\|_B$ is bounded below.

Fixing $x_1 \in X_1$ and $x_2 \in X_2$, we find that for a given $M \in [0,\infty]$, solving Problem 4.2.4 is equivalent to finding $y_0 \in \mathfrak{Y}$ which minimizes $\|(x_1,x_2) - y_0\|_A$, subject to the constraint that $\|(x_1,x_2) - y_0\|_B \leq M$. For any such y_0, a solution is given by the unique $x \in X$ such that $y_0 = (Ax, Bx)$. Thus Problem 4.2.4 becomes just a special case of Problem 4.2.2, with $\mathfrak{X}$ and $\mathfrak{Y}$ replacing X and Y and x_A and x_B both set equal to (x_1,x_2) in the problem statement. Theorem 4.2.3 may therefore be applied, which establishes the existence of a solution to Problem 4.2.4 for any finite M sufficiently large.

4.3 Vector-valued L^∞ and H^∞ spaces

In this section, we develop much of the theory which will be needed later in Section 4.4. This includes the construction of vector-valued L^∞ and H^∞ spaces on the unit circle $\mathbb{T}$, followed by the study of a certain class of seminorms on these spaces. In Section 4.4, this will be used to formulate a constrained approximation problem adapted from Problem 4.2.2, in which we replace the Banach spaces X and Y with the spaces $\mathrm{L}^\infty(X)$ and $\mathrm{H}^\infty(Y)$, consisting of bounded X-valued and Y-valued measurable functions and Hardy class functions on $\mathbb{T}$, respectively.

First of all, we shall need to define the spaces $\mathrm{L}^\infty(X)$ and $\mathrm{H}^\infty(Y)$. We shall also need to identify their preduals, in order to fit these spaces into the abstract framework of the previous section.

4.3.1 The Banach spaces $\mathrm{L}^\infty(X)$ and $\mathrm{L}^1(X_*)$

Let X be any Banach space. For any given function $f : \mathbb{T} \to X$, we use the following terminology, taken from [12, p. 41]:

1. Say that f is *simple* if it can be written in the form $x_1\chi_{A_1} + \cdots + x_n\chi_{A_n}$, for some choice $x_1, \ldots, x_n \in X$ and Borel sets $A_1, \ldots, A_n \subset \mathbb{T}$.

2. Say that f is *strongly measurable* if it is the almost everywhere, pointwise norm limit of a sequence of simple X-valued functions.

3. Say that f is *weakly* (resp. *weak**) *measurable* if $\alpha \circ f$ is measurable, for all $\alpha \in X^*$ (resp. X_*, provided X is equipped with a predual X_*).

It is easy to see that $1 \Rightarrow 2 \Rightarrow 3$, and that weak measurability implies weak* measurability.

From now on, we work with a fixed Banach space X, which we assume to have a *separable* predual X_*. We shall regard X_* as a closed subspace of the dual space X^*, given by the natural embedding of a Banach space into its double dual. The separability assumption plays an important role in the proceeding arguments. It is also quite a mild assumption since the typical choices for X (such as ℓ^∞, $\mathcal{L}(H)$ for a separable Hilbert space H, or $\mathcal{L}(\mathbb{C}^n, \mathbb{C}^m)$ for $n, m \in \mathbb{N}$), are all equipped with separable preduals.

A rather less typical choice for X might be the space $\mathcal{M}(K)$, consisting of complex Borel measures on a compact metric space K, equipped with the total variation norm. This has a natural predual $\mathrm{C}(K)$, the separable Banach space of continuous scalar-valued functions on K, with the uniform norm.

Definition 4.3.1 *The normed space* $\mathrm{L}^\infty(X)$ *is given by:*

$$\mathrm{L}^\infty(X) = \{f : \mathbb{T} \to X \text{ weak* measurable} : \|f\|_\infty < \infty\},$$

where we define the uniform norm $\|\cdot\|_\infty$ *by:*

$$\|f\|_\infty = \operatorname*{ess\,sup}_{z\in\mathbb{T}} \|f(z)\|,$$

for all $f : \mathbb{T} \to X$ *weak* measurable. Similarly,* $\mathrm{L}^1(X_*)$ *is given by:*

$$\mathrm{L}^1(X_*) \;=\; \{f : \mathbb{T} \to X_* \text{ strongly measurable} : \|f\|_1 < \infty\},$$

where we define the norm $\|\cdot\|_1$ *by:*

$$\|f\|_1 \;=\; \int_{z\in\mathbb{T}} \|f(z)\|\,\mathrm{d}\mu(z),$$

for all $f : \mathbb{T} \to X_*$ *strongly measurable. These two spaces are equipped with the norms* $\|\cdot\|_\infty$ *and* $\|\cdot\|_1$ *respectively, and they are complete.*

The reader may recognize that $\mathrm{L}^1(X_*)$ is just the usual Lebesgue-Bochner space L^1. Note, however, that the definition of $\mathrm{L}^\infty(X)$ differs from that of the Lebesgue-Bochner space L^∞, as given in [12, p. 50]. There, strong measurability is used rather than weak* measurability, and this gives a genuinely smaller space in general. This is seen in the example [12, p. 43, ex. 6], where we take X equal to ℓ^∞ with predual ℓ^1. In fact, the example shows that weak measurability would also give a smaller space. It will turn out, therefore, that weak* measurability is the correct notion to use in our context.

The following lemma will be crucial for the next result:

Lemma 4.3.2 *Let* $\phi : X_* \to \mathrm{L}^\infty$ *be a bounded linear map. Then there is a function* $f \in \mathrm{L}^\infty(X)$*, unique up to equality almost everywhere, such that*

$$\alpha \circ f \;=\; \phi(\alpha) \;\text{ a.e.}, \tag{4.6}$$

for all $\alpha \in X_*$*. Furthermore,* $\|f\|_\infty = \|\phi\|$*. (Here we regard elements of the Banach space* L^∞ *to be equivalence classes of scalar valued functions for which any pair agree on a set of full measure.)*

Proof We shall construct a choice function, $\theta : X_* \times \mathbb{T} \to \mathbb{C}$, for the map ϕ. This will have the properties that $\theta(\cdot, z)$ is linear for all $z \in \mathbb{T}$, and that $\theta(\alpha, \cdot)$ is measurable and equal to $\phi(\alpha)$ almost everywhere, for all $\alpha \in X_*$.

Let $\mathbb{Q}[i]$ denote the number field of Gaussian rationals, $x + iy$, with $x, y \in \mathbb{Q}$. Since X_* is separable, there exists a countable dense set $S \subset X_*$. Now let V be the linear span over $\mathbb{Q}[i]$ of all the elements of S. Then V is a countable vector space over $\mathbb{Q}[i]$, and it is dense in X_*.

For every vector $v \in V$, we may choose a representative $g_v : \mathbb{T} \to \mathbb{C}$ from the equivalence class $[\phi(v)]$. By the linearity and boundedness of ϕ, the following relations hold for almost all $z \in \mathbb{T}$, whenever we fix $u, v \in V$ and $\lambda, \mu \in \mathbb{Q}[i]$:

$$g_{\lambda u+\mu v}(z) = \lambda g_u(z) + \mu g_v(z) \tag{4.7}$$

$$|g_u(z)| \leq \|\phi\| \, \|u\| \tag{4.8}$$

Now since there are only countably many such quadruples (u, v, λ, μ), we may find a subset $K \subset \mathbb{T}$ of full measure, such that relations (4.7) and (4.8) hold for all $z \in K$, $u, v \in V$ and $\lambda, \mu \in \mathbb{Q}[i]$.

For any Cauchy sequence $(v_n)_{n=1}^{\infty}$ in V, and any point $z \in K$, we have

$$\begin{aligned} |g_{v_n}(z) - g_{v_m}(z)| &= |g_{v_n - v_m}(z)| \leq \|\phi\| \, \|v_n - v_m\| \\ &\to 0\,, \end{aligned}$$

whenever $n, m \to \infty$. Hence g_{v_n} converges uniformly on K as $n \to \infty$. Thus we may define $\theta : X_* \times \mathbb{T} \to \mathbb{C}$ as follows:

$$\theta(\alpha, z) = \begin{cases} \lim\limits_{v \in V, v \to \alpha} g_v(z) & \text{if } z \in K \\ 0 & \text{otherwise}\,, \end{cases}$$

for all $\alpha \in X_*$ and $z \in \mathbb{T}$. Since the above limit exists whenever $v \to \alpha$ for elements $v \in V$, it is unique for every $\alpha \in X_*$ and $z \in K$. So θ is well defined.

To show that $\theta(\cdot, z)$ is complex linear, first observe that it is continuous for all $z \in \mathbb{T}$. Indeed, if $\alpha, \beta \in X_*$ and $z \in K$ then there are sequences $(u_n)_{n=1}^{\infty}$ and $(v_n)_{n=1}^{\infty}$ in V, such that $u_n \to \alpha$ and $v_n \to \beta$ in norm as $n \to \infty$. Therefore

$$\begin{aligned} |\theta(\alpha, z) - \theta(\beta, z)| &= \lim_{n\to\infty} |g_{u_n}(z) - g_{v_n}(z)| \\ &= \lim_{n\to\infty} |g_{u_n - v_n}(z)| \\ &\leq \lim_{n\to\infty} \|\phi\| \, \|u_n - v_n\| = \|\phi\| \, \|\alpha - \beta\|\,. \end{aligned}$$

Now since $\theta(\cdot, z)$ is continuous and V is dense in X_* and $\mathbb{Q}[i]$ is dense in $\mathbb{C}$, the $\mathbb{Q}[i]$-linearity of the functional $\theta(\cdot, z)$, on V, implies that it is complex linear on the whole of X_* for all $z \in \mathbb{T}$, as required.

Let $\alpha \in X_*$ and let $(v_n)_{n=1}^{\infty}$ be any sequence in V with limit α as $n \to \infty$. Then $\theta(\alpha, \cdot)$ is the uniform limit, on K, of the functions g_{v_n} as $n \to \infty$. Hence $\theta(\alpha, \cdot)$ is measurable. Moreover,

$$\begin{aligned} \|\theta(\alpha, \cdot) - \phi(\alpha)\|_\infty &\leq \limsup_{n\to\infty} (\|\theta(\alpha, \cdot) - g_{v_n}\|_\infty + \|g_{v_n} - \phi(\alpha)\|_\infty) \\ &\leq 2\,\|\phi\| \lim_{n\to\infty} \|\alpha - v_n\| \\ &= 0\,. \end{aligned}$$

So $\theta(\alpha, \cdot) = \phi(\alpha)$ almost everywhere on $\mathbb{T}$, for all $\alpha \in X_*$, as required.

Finally, for any $z \in \mathbb{T}$, the functional $\theta(\cdot, z)$ is continuous and linear on X_*. Hence there is a unique element $f(z)$ in the dual, X, such that

$$\alpha \circ f(z) \;=\; \theta(\alpha, z)\,,$$

for all $\alpha \in X_*$. This defines a function $f : \mathbb{T} \to X$. Furthermore, $\|f(z)\| = \|\theta(\cdot, z)\| \leq \|\phi\|$ for all $z \in \mathbb{T}$, so f is a bounded function with $\|f\|_\infty \leq \|\phi\|$.

Now for any $\alpha \in X_*$, the scalar function $\alpha \circ f$ satisfies:

$$\alpha \circ f \;=\; \theta(\alpha, \cdot) \;=\; \phi(\alpha) \;\text{ a.e.}\,,$$

This is a measurable function on $\mathbb{T}$. Hence f satisfies equation (4.6), is weak* measurable and therefore lies in $L^\infty(X)$, as required.

To see that f is unique up to equality almost everywhere, suppose that $f' \in L^\infty$ also satisfies equation (4.6). Then $v \circ (f' - f) = 0$ a.e. for all $v \in V$. Since V is countable, this implies that there is a set $L \subset \mathbb{T}$ of full measure on which $v \circ (f' - f)$ is identically zero, for all $v \in V$. But since V is dense in X_*, this implies that $f' - f$ is identically zero on L.

It remains to show that $\|f\|_\infty$ is no smaller than $\|\phi\|$. For any $\varepsilon > 0$, there exists $\alpha \in X_*$ with $\|\alpha\| = 1$ and a set $M \subset \mathbb{T}$ of nonzero measure, such that:

$$\begin{aligned} \|\phi(\alpha)\|_\infty &> \|\phi\| - \varepsilon/2 \\ |\phi(\alpha)(z)| &> \|\phi(\alpha)\|_\infty - \varepsilon/2\,, \end{aligned}$$

for all $z \in M$. Therefore

$$\|f(z)\| \;\geq\; |\alpha(f(z))| \;=\; |\phi(\alpha(z))| \;>\; \|\phi\| - \varepsilon\,,$$

for almost all $z \in M$. This implies that $\|f\|_\infty \geq \|\phi\| - \varepsilon$. But since ε is arbitrary, we have $\|f\|_\infty \geq \|\phi\|$, as required. □

We can now give the main result concerning $\mathrm{L}^\infty(X)$, stated as follows:

Theorem 4.3.3 *There is an isometric linear isomorphism,*

$$\theta \;:\; \mathrm{L}^\infty(X) \;\to\; \mathrm{L}^1(X_*)^*\,,$$

defined by

$$\theta(f)(g) \;=\; \int_{z\in\mathbb{T}} g(z)(f(z))\,\mathrm{d}\mu(z)\,, \tag{4.9}$$

for all functions $f \in \mathrm{L}^\infty(X)$ *and* $g \in \mathrm{L}^1(X_*)$.

Proof We must first show that θ is well defined. Observe that since every function in $\mathrm{L}^1(X_*)$ is an a.e. pointwise norm limit of simple functions, the integrand in (4.9) is measurable. Indeed, for $f \in \mathrm{L}^\infty(X)$ and any simple function $h = \alpha_1\chi_{A_1} + \cdots + \alpha_k\chi_{A_k}$, with vectors $\alpha_1, \ldots, \alpha_k \in X_*$ and Borel sets $A_1, \ldots, A_k \subset \mathbb{T}$, we have:

$$h(z)(f(z)) \;=\; (\alpha_1 \circ f)(z)\,\chi_{A_1}(z) + \ldots + (\alpha_k \circ f)(z)\,\chi_{A_k}(z)\,,$$

which is measurable in $z \in \mathbb{T}$, since $\alpha \circ f$ is measurable for any $\alpha \in X_*$. So if $g_n \to g \in \mathrm{L}^1(X_*)$ almost everywhere as $n \to \infty$, for some simple functions $(g_n)_{n=1}^\infty$, then

$$g_n(z)(f(z)) \;\to\; g(z)(f(z))$$

for almost all $z \in \mathbb{T}$ as $n \to \infty$. The limit is therefore a measurable function.

Furthermore,

$$\begin{aligned}\int_{z\in\mathbb{T}} |g(z)(f(z))|\,\mathrm{d}\mu(z) \;&\leq\; \int_{z\in\mathbb{T}} \|f(z)\|\,\|g(z)\|\,\mathrm{d}\mu(z)\\ &\leq\; \operatorname*{ess\,sup}_{z\in\mathbb{T}} \|f(z)\| \cdot \int_{z\in\mathbb{T}} \|g(z)\|\,\mathrm{d}\mu(z)\\ &=\; \|f\|_\infty \cdot \|g\|_1\,,\end{aligned}$$

so the integral in (4.9) is well defined, with $|\theta(f)(g)| \leq \|f\|_\infty \|g\|_1$. Thus θ is a well defined linear contraction.

To show that θ is onto, let $\rho \in \mathrm{L}^1(X_*)^*$ and let $\alpha \in X_*$. Then the mapping,

$$h \;\mapsto\; \rho(\alpha \cdot h)\,, \tag{4.10}$$

is continuously linear in $h \in \mathrm{L}^1$, with bound at most $\|\rho\|\,\|\alpha\|$. So by the duality between the scalar Lebesgue spaces L^1 and L^∞, there is a unique element $\phi(\alpha)$ of the Banach space L^∞, such that

$$\rho(\alpha \cdot h) \;=\; \int_{z\in\mathbb{T}} \phi(\alpha)(z)\, h(z)\, \mathrm{d}\mu(z)\,,$$

for all $h \in \mathrm{L}^1$. Moreover, $\|\phi(\alpha)\| \leq \|\rho\|\,\|\alpha\|$ for all $\alpha \in X_*$. Thus we obtain a bounded linear functional $\phi : X_* \to \mathrm{L}^\infty$, with $\|\phi\| \leq \|\rho\|$.

Now by Lemma 4.3.2, there is a unique element f in $\mathrm{L}^\infty(X)$, the domain of θ, satisfying equation (4.6). Hence

$$\begin{aligned} \theta(f)(\alpha \cdot h) \;&=\; \int_{z\in\mathbb{T}} \alpha(f(z))\, h(z)\, \mathrm{d}\mu(z) \\ &=\; \int_{z\in\mathbb{T}} \phi(\alpha)(z)\, h(z)\, \mathrm{d}\mu(z) \;=\; \rho(\alpha \cdot h)\,, \end{aligned}$$

for all $\alpha \in X_*$ and $h \in \mathrm{L}^1$. Moreover, equality always occurs between the above two integrals for a unique choice of $f \in \mathrm{L}^\infty(X)$. Since the simple functions are dense in $\mathrm{L}^1(X_*)$ (as shown in [12, ch. 2]), functions of the form $\alpha \cdot h$ span a dense subspace of $\mathrm{L}^1(X_*)$. Hence $\theta(f) = \rho$ with $\|f\|_\infty \leq \|\rho\|$.

This shows, moreover, that θ is an isometry, since the choice $f \in \mathrm{L}^\infty(X)$ for any element $\rho \in \mathrm{L}^1(X_*)^*$ is unique, and $\|\theta(f)\| \leq \|f\|_\infty \leq \|\rho\|$, giving us equality throughout. □

The above theorem shows that we may naturally identify the Banach space $\mathrm{L}^\infty(X)$ with the dual of $\mathrm{L}^1(X_*)$. So from now on, we shall set:

$$\mathrm{L}^\infty(X)_* \;=\; \mathrm{L}^1(X_*)\,,$$

with the predual action of $\mathrm{L}^1(X_*)$ on $\mathrm{L}^\infty(X)$ given by integration, as in equation (4.9) of the statement of Theorem 4.3.3 above.

The argument used to establish the isometry between $\mathrm{L}^\infty(X)$ and $\mathrm{L}^1(X_*)^*$ can be summarized as follows. It is relatively easy to show that $\mathrm{L}^\infty(X)$ embeds into $\mathrm{L}^1(X_*)^*$. The hard part of the argument is to show that the inclusion is onto, and for this we made use of Lemma 4.3.2. This lemma shows that $\mathrm{L}^\infty(X)$ is isometric to $\mathcal{L}(X_*, \mathrm{L}^\infty)$, the space of bounded linear maps from X_* to the scalar-valued functions L^∞. For any element $\rho \in \mathrm{L}^1(X_*)^*$, we obtain a bilinear functional on X_* and L^1, as given by the mapping (4.10) in the above proof. Now we use the fact that $\mathrm{L}^\infty \cong (\mathrm{L}^1)^*$ to obtain from this an element of $\mathcal{L}(X_*, \mathrm{L}^\infty)$. This gives rise to the desired element of $\mathrm{L}^\infty(X)$, mapping to ρ.

In fact we have the following isometries:

$$\begin{aligned} \mathrm{L}^\infty(X) \cong \mathcal{L}(X_*, \mathrm{L}^\infty) \cong \mathcal{L}(\mathrm{L}^1, X) &\cong \mathrm{Bilin}(\mathrm{L}^1, X_*; \mathbb{C}) \\ &\cong \mathcal{L}(\mathrm{L}^1 \widehat{\otimes} X_*, \mathbb{C}) \\ &= (\mathrm{L}^1 \widehat{\otimes} X_*)^* \\ &\cong \mathrm{L}^1(X_*)^*, \end{aligned}$$

where $\mathrm{L}^1 \widehat{\otimes} X_*$ is the projective tensor product of L^1 and X_*, and $\mathrm{Bilin}(\mathrm{L}^1, X_*; \mathbb{C})$ denotes the Banach space of continuous bilinear functionals on L^1 and X_*. The projective tensor product of two Banach spaces is defined in [12, ch. 8]. There it is also shown that $\mathrm{L}^1 \widehat{\otimes} Z$ is isometric to $\mathrm{L}^1(Z)$, for any Banach space Z, thus completing the above chain of isometries between $\mathrm{L}^\infty(X)$ and $\mathrm{L}^1(X_*)^*$.

Note also the above isometry, $\mathrm{L}^\infty(X) \cong \mathcal{L}(\mathrm{L}^1, X)$. This shows that when considered as a Banach space, $\mathrm{L}^\infty(X)$ does not depend on the predual we assign to the space X. However, the definition of $\mathrm{L}^\infty(X)$ as a *function* space does involve the predual of X, since the notion of weak* measurability is used there.

4.3.2 The Banach space $\mathrm{H}^\infty(Y)$

Let Y be any weak* closed subspace of X. Having fixed the predual of $\mathrm{L}^\infty(X)$, we can now define the Banach space $\mathrm{H}^\infty(Y)$. We shall take this to be a weak* closed subspace of $\mathrm{L}^\infty(X)$. First of all, we need another definition:

Definition 4.3.4 *Let* $J : X_* \to \mathrm{L}^1(X_*)$ *be the inclusion map, taking any element* $\alpha \in X_*$ *to the* α*-valued constant function. Define the integration map,*

$$\int_{\mathbb{T}} \cdot \, \mathrm{d}\mu \;:\; \mathrm{L}^\infty(X) \;\to\; X\,,$$

to be the adjoint operator $J^* \in \mathcal{L}(\mathrm{L}^\infty(X), X)$. *It follows that integration is a weak* continuous operation on* $\mathrm{L}^\infty(X)$.

For any function $f \in \mathrm{L}^\infty(X)$ *and any integer* $n \in \mathbb{Z}$, *define the* n^{th} Fourier coefficient, $\hat{f}(n)$, *by:*

$$\hat{f}(n) \;=\; \int_{z\in\mathbb{T}} z^{-n} f(z)\,\mathrm{d}\mu(z)\,.$$

Equivalently, this is seen to be the adjoint of a bounded operator which takes any $\alpha \in X_*$ *to the function, in* $\mathrm{L}^1(X_*)$, *given by* αz^{-n} *over* $z \in \mathbb{T}$. *Thus the mapping* $f \to \hat{f}(n)$ *is also weak* continuous in* $f \in \mathrm{L}^\infty(X)$, *for all* $n \in \mathbb{Z}$.

Using these Fourier coefficients, we can now define $\mathrm{H}^\infty(Y)$ as follows:

Definition 4.3.5 *The Banach space* $\mathrm{H}^\infty(Y)$ *is given by:*

$$\begin{aligned}\mathrm{H}^\infty(Y) \;=\; \{f \in \mathrm{L}^\infty(X) \,:\, &\hat{f}(n) = 0 \text{ for all } n < 0, \\ &\text{and } \hat{f}(n) \in Y \text{ for all } n \geq 0\}\,.\end{aligned}$$

Since Y *is weak* closed in* X, *and the evaluation of each Fourier coefficient is weak* continuous on* $\mathrm{L}^\infty(X)$, *this definition makes* $\mathrm{H}^\infty(Y)$ *into a weak* closed subspace of* $\mathrm{L}^\infty(X)$.

With this definition, it is easily seen that any function $f \in \mathrm{H}^\infty(Y)$ is almost everywhere Y-valued. Indeed, let us choose a countable dense subset S of the preannihilator $Y_\circ \subset X_*$ of the subspace $Y \subset X$, and let $\alpha \in S$. Then

$$(\alpha \circ f)\hat{}(n) \;=\; \alpha(\hat{f}(n)) \;=\; 0\,,$$

for all $n \in \mathbb{Z}$, by the definition of integration on $\mathrm{L}^\infty(X)$. Hence $\alpha \circ f = 0$ almost everywhere, for all $\alpha \in S$. Since S is countable, this implies that $f(z) \in S^\circ$ for almost all $z \in \mathbb{Z}$. But $S^\circ = (Y_\circ)^\circ = Y$, because Y is weak* closed.

For the purposes of this chapter, we take $\mathrm{H}^\infty(Y)$ to be a certain closed subspace of bounded measurable a.e. Y-valued functions on the unit circle $\mathbb{T}$. This definition does not make explicit the connection between $\mathrm{H}^\infty(Y)$ and analyticity. The next result will show how elements of $\mathrm{H}^\infty(Y)$ may be alternatively realised as Y-valued analytic functions on the unit disc $\mathbb{D}$.

Say that a continuous function $f : \mathbb{D} \to Y$ is *analytic* if the derivative,

$$f'(z) \;=\; \lim_{w \to z} \frac{f(w) - f(z)}{w - z},$$

exists for all $z \in \mathbb{D}$, with convergence in norm. It can be shown that this is equivalent to the statement that for every closed disc $K \subset \mathbb{D}$, the function f has a uniformly convergent power series on K about its centre. Thus uniform limits of Y-valued analytic functions on $\mathbb{D}$ are analytic.

The main result describing $\mathrm{H}^\infty(Y)$ is stated as follows:

Theorem 4.3.6 *Let $\mathcal{H}^\infty(Y)$ denote the Banach space of bounded Y-valued analytic functions on $\mathbb{D}$, equipped with the uniform norm. Then the map,*

$$\Omega \;:\; \mathrm{H}^\infty(Y) \;\to\; \mathcal{H}^\infty(Y),$$

given by the power series,

$$\Omega(f)(z) \;=\; \hat{f}(0) + \hat{f}(1)z + \hat{f}(2)z^2 + \dots, \tag{4.11}$$

for any $f \in \mathrm{H}^\infty(Y)$ and $z \in \mathbb{D}$, is a well defined, isometric linear isomorpism.

Proof We shall first show that Ω is a well defined contraction.

Let $f \in \mathrm{H}^\infty(Y)$. From the definition of the Fourier coefficients of f, we see that the mapping $f \mapsto \hat{f}(n)$ is the adjoint of a contraction for all $n \in \mathbb{Z}$. Thus the Fourier coefficient $\hat{f}(n)$ satisfies $\|\hat{f}(n)\| \le \|f\|_\infty$, for all $n \in \mathbb{Z}$. Hence the above power series, in equation (4.11), converges uniformly in z on any closed subset of the unit disc $\mathbb{D}$. This shows that (4.11) defines a Y-valued analytic function $\Omega(f)$ on $\mathbb{D}$. It therefore remains to show that $\Omega(f)$ is a bounded function, with uniform bound $\|f\|_\infty$.

For any number $r \in [0,1)$, we define the *Poisson kernel* $P_r : \mathbb{R} \to (0,\infty)$ by:

$$\begin{aligned} P_r(\theta) &= \frac{1-r^2}{1-2r\cos\theta + r^2} \\ &= 2\,\mathrm{Re}[(1-re^{i\theta})^{-1}] - 1, \quad \text{for all } \theta \in \mathbb{R}. \end{aligned}$$

Thus P_r has Fourier series expansion:

$$P_r(\theta) = \sum_{n\in\mathbb{Z}} r^{|n|}\, e^{in\theta},$$

for all $\theta \in \mathbb{R}$. Hence P_r takes mean value 1 on any interval of length 2π.

Now we may define the *harmonic extension*, $F : \mathbb{D} \to X$, of the boundary function f, by the formula:

$$F(re^{i\phi}) = \frac{1}{2\pi}\int_0^{2\pi} f(e^{i\theta}) P_r(\theta-\phi)\,\mathrm{d}\theta, \tag{4.12}$$

for all $r \in [0,1)$ and $\phi \in [0,2\pi)$. Since P_r takes mean value 1 on any interval of length 2π, we deduce that $\|F(z)\| \le \|f\|_\infty$ for all $z \in \mathbb{D}$.

But from the Fourier series expansion of P_r, we find that:

$$\begin{aligned} F(z) &= \int_{w\in\mathbb{T}} f(w)(1 + z\bar{w} + \bar{z}w + z^2\bar{w}^2 + \bar{z}^2 w^2 + \ldots)\,\mathrm{d}\mu(w) \\ &= \hat{f}(0) + \sum_{n=1}^{\infty}\left[\hat{f}(n)\,z^n + \hat{f}(-n)\,\bar{z}^n\right] \\ &= \Omega(f)(z), \end{aligned}$$

for all $z \in \mathbb{D}$, since $\hat{f}(-n) = 0$ for all $n \in \mathbb{N}$. So $\Omega(f)$ is precisely the harmonic extension of f to $\mathbb{D}$, and hence it is bounded uniformly by $\|f\|_\infty$. This shows that Ω is a well defined contraction.

We shall now show that Ω is onto. Let $F \in \mathcal{H}^\infty(Y)$. Then for any $\alpha \in X_*$, the function $\alpha\circ F$ is analytic, since its complex derivative converges at all points in $\mathbb{D}$, from the earlier definition of Y-valued analyticity. Moreover, $\alpha \circ F$ is uniformly bounded by $\|\alpha\|\,\|F\|_\infty$, and hence there is a boundary function $\phi(\alpha) \in \mathrm{H}^\infty \subset \mathrm{L}^\infty$, unique up to equality almost everywhere, whose harmonic extension to $\mathbb{D}$ is $\alpha\circ F$. This defines a linear map $\phi : X_* \to \mathrm{L}^\infty$ whose norm is at most $\|F\|_\infty$, where we regard the codomain L^∞ to be a Banach space of equivalence classes of functions on $\mathbb{T}$.

By Lemma 4.3.2, there is a function $f \in \mathrm{L}^\infty(X)$ such that $\alpha \circ f = \phi(\alpha)$ almost everywhere, for all $\alpha \in X_*$, and such that $\|f\|_\infty = \|\phi\|$. Hence $\|f\|_\infty \leq \|F\|_\infty$. Now, $\alpha \circ f$ has harmonic extension $\alpha \circ F$ for all $\alpha \in X_*$. This implies that f has harmonic extension F, given by formula (4.12), since

$$\begin{aligned} \alpha\left(\frac{1}{2\pi}\int_0^{2\pi} f(e^{i\theta})P_r(\theta-\phi)\,\mathrm{d}\theta\right) &= \frac{1}{2\pi}\int_0^{2\pi}(\alpha\circ f)(e^{i\theta})P_r(\theta-\phi)\,\mathrm{d}\theta \\ &= \alpha(F(re^{i\phi}))\,, \end{aligned}$$

for all $r \in [0,1)$, $\phi \in [0,2\pi)$ and $\alpha \in X_*$.

Now the formula for F given by

$$F(z) \;=\; \hat{f}(0) + \sum_{n=1}^{\infty}\left[\hat{f}(n)\,z^n + \hat{f}(-n)\,\bar{z}^n\right],$$

for all $z \in \mathbb{D}$, uniquely determines the Fourier coefficients of f. In particular, $\hat{f}(n) = 0$ for all indices $n < 0$, and $\hat{f}(n) \in Y$ for all $n \geq 0$, since F is a Y-valued analytic function. Since the X_*-valued trigonometric polynomials are dense in $\mathrm{L}^1(X_*)$, the Fourier coefficients of the function f determine it uniquely. Hence f is the unique element of $\mathrm{H}^\infty(Y)$ such that $\Omega(f) = F$. This shows that Ω is bijective. Since $\|f\|_\infty \leq \|F\|_\infty$ and Ω is a contraction, we deduce that Ω is an isometric isomorphism, as required. □

The above proof shows that the mapping Ω, given in the statement of the theorem, can be equivalently defined as the harmonic extension of a function in $\mathrm{H}^\infty(Y)$, to the open unit disc $\mathbb{D}$ using the Poisson kernel.

It was shown earlier that the space $\mathrm{L}^\infty(X)$ is independent, up to isometry, of the weak* topology on X, since it is isometric to $\mathcal{L}(\mathrm{L}^1, X)$. Similarly, the above theorem shows that $\mathrm{H}^\infty(Y)$ is independent, up to isometry, of the spaces X and X_*. This is because the space $\mathcal{H}^\infty(Y)$, of bounded analytic Y-valued functions, is defined only in terms of Y.

Note that for a scalar, bounded analytic function G on the disc $\mathbb{D}$, Fatou's theorem states that G converges nontangentially to its H^∞ boundary function, at almost all points on the circle $\mathbb{T}$. However, an example given in [28, p. 92] shows that a vector-valued analytic function, $F \in \mathcal{H}^\infty(Y)$, may in every direction fail to converge radially, and in norm, to its boundary function $\Omega^{-1}(F)$, even in the case that Y is a separable Hilbert space. The following argument shows that by replacing norm convergence with weak* convergence, we recover an analogue of Fatou's theorem:

Let S be any countable dense subset of X_*. Let $F \in \mathcal{H}^\infty(Y)$ and set $f = \Omega^{-1}(F)$. Then for any $\alpha \in S$, the bounded analytic function $\alpha \circ F$ has boundary function $\alpha \circ f$. So by the usual Fatou's theorem, it converges nontangentially to $\alpha \circ f$ at almost all points in $\mathbb{T}$. Since S is countable, this implies that there is a subset $K \subset \mathbb{T}$ of full measure, such that $\alpha \circ F$ converges nontangentially to $\alpha \circ f$ at all points in K, for all $\alpha \in S$. Hence F is weak* nontangentially convergent to f at almost every point in $\mathbb{T}$, as required.

4.3.3 Seminorms on $\mathrm{L}^\infty(X)$

We complete this section by studying a certain class of seminorms on $\mathrm{L}^\infty(X)$. This will include, for example, all the weighted norms $\|\cdot\|_{\infty,w}$ for scalar weights $w \in \mathrm{L}^\infty_{\mathrm{pos}}$, as we defined in Section 4.1. The class of seminorms in which we are interested is given as follows:

Definition 4.3.7 *Let $\|\cdot\|'$ be any seminorm on $\mathrm{L}^\infty(X)$. Then $\|\cdot\|'$ is said to be of* sup-type *if the following two conditions hold:*

1. *$\|\cdot\|'$ is weak* lower semicontinuous.*

2. *For every function $f \in \mathrm{L}^\infty(X)$ and any Lebesgue measurable sets $A, B \subset \mathbb{T}$, we have:*

$$\|f \cdot \chi_{A\cup B}\|' = \|f \cdot \chi_A\|' \vee \|f \cdot \chi_B\|',$$

where χ_A, χ_B and $\chi_{A\cup B}$ are the characteristic functions of the sets A, B and $A \cup B$, respectively.

Before we prove the main results concerning sup-type seminorms, we will need the following theorem, taken from [15, sec. V.5.7]. It concerns convex subsets of any dual Banach space X. So for this result, we may drop the assumption of separability of the predual X_*.

Theorem 4.3.8 (Krein-Šmulian theorem) *Let S be a convex subset of X such that $S \cap RB_X$ is weak* closed for all $R > 0$ (where B_X denotes the closed unit ball of X). Then S is itself weak* closed.*

An immediate corollary of this theorem is the fact that a seminorm $\|\cdot\|'$ on X is weak* lower semicontinuous if and only if the norm, $\|\cdot\|' \vee \varepsilon\|\cdot\|$, which is bounded below, is weak* lower semicontinuous for all $\varepsilon > 0$.

The next result shows that a rather large class of weighted seminorms on $\mathrm{L}^\infty(X)$ are sup-type:

Proposition 4.3.9 *Let $W : \mathbb{T} \times X \to [0,\infty)$ be any given function. Then define the function $\widetilde{W} : \mathbb{T} \times X_* \to [0,\infty]$ by*

$$\widetilde{W}(z,\alpha) = \sup\{|\alpha(x)| : x \in X \text{ with } W(z,x) \leq 1\}, \tag{4.13}$$

for all $z \in \mathbb{T}$ and $\alpha \in X_*$. Now suppose that the following hypotheses hold:

1. *$\widetilde{W}(\cdot,\alpha)$ is a measurable function on $\mathbb{T}$, for all $\alpha \in X_*$.*

2. *$W(\cdot,x)$ is an essentially bounded function on $\mathbb{T}$, for all $x \in X$.*

3. *$W(z,\cdot)$ is a weak* lower semicontinuous seminorm on X, for all $z \in \mathbb{T}$.*

Then $\|\cdot\|_{\infty,W}$ is a sup-type seminorm on $\mathrm{L}^\infty(X)$, where we set:

$$\|f\|_{\infty,W} = \operatorname*{ess\,sup}_{z\in\mathbb{T}} W(z,f(z)),$$

for all $f \in \mathrm{L}^\infty(X)$. Furthermore, $W(z,f(z))$ is a measurable function of $z \in \mathbb{T}$ for any function $f \in \mathrm{L}^\infty(X)$.

Proof We shall first show that $\|f\|_{\infty,W}$ is finite for all $f \in \mathrm{L}^\infty(X)$.

For all $z \in \mathbb{T}$, let B_z be the weak* closed convex set in X, given by:

$$B_z \;=\; \{x \in X \,:\, W(z,x) \le 1\}\,.$$

Now define the essential intersection $B \subset X$ by:

$$B \;=\; \{x \in X \,:\, x \in B_z \text{ for almost all } z \in \mathbb{T}\}\,.$$

Then B is a norm closed convex set, since for any element $x \in X \setminus B$, the mapping:

$$z \;\mapsto\; \mathrm{d}(x, B_z)\,,$$

is not almost everywhere zero over $z \in \mathbb{T}$, and therefore takes values of at least δ, for some $\delta > 0$, on some subset of $\mathbb{T}$ with nonzero measure. But this implies that $y \notin B$ for all $y \in X$ such that $\|x - y\| < \delta$. So B is norm closed.

By hypothesis 2 of the statement of the proposition, the function on X given by the mapping,

$$x \;\mapsto\; \|x \cdot 1\|_{\infty,W}\,,$$

takes finite values for all $x \in X$. It therefore constitutes a seminorm on X, with unit ball equal to B. So by Proposition 4.2.1, this seminorm is bounded. Hence there is a constant $R > 0$ such that

$$W(z,x) \;\le\; R\,\|x\| \quad \text{for almost all } z \in \mathbb{T}\,,$$

for all $x \in X$. Now since X_* is separable, the unit ball B_X of X is weak* metrizable. But B_X is also weak* compact by Alaoglu's theorem and therefore separable with respect to the weak* topology. Hence there is a countable set $V \subset B_X$ which is weak* dense in B_X. Since V is countable, there is a set $K \subset \mathbb{T}$ of full measure such that $W(z,x) \le R$ for all $z \in K$ and $x \in V$. But $W(z,\cdot)$ is weak* lower semicontinuous, which implies that $W(z,x) \le R$ for all $x \in B_X$, whenever $z \in K$. Hence $\|W(z,\cdot)\| \le R$ for almost all $z \in \mathbb{T}$. From the definition of $\|\cdot\|_{\infty,W}$, this implies that

$$\|f\|_{\infty,W} \;\le\; R\,\|f\|_\infty \;<\; \infty\,,$$

for all $f \in \mathrm{L}^\infty(X)$, as required.

Now it is clear from hypothesis 3 that $\|\cdot\|_{\infty,W}$ is subadditive and homogeneous, and thus forms a seminorm on $\mathrm{L}^\infty(X)$. It is also clear that $\|\cdot\|_{\infty,W}$ satisfies condition 2 of the definition of a sup-type seminorm. So it remains to show that it is weak* lower semicontinuous. We do this as follows:

First of all, let us assume that $W(z,\cdot)$ is bounded below for all $z \in \mathbb{T}$. That is, for every $z \in \mathbb{T}$ there is some $\delta > 0$, such that $W(z,x) \geq \delta\|x\|$ for all $x \in X$. It therefore follows from equation (4.13) that $\widetilde{W}(z,\cdot)$ is a bounded norm on X_*, with bound δ^{-1}. So the convex set $\{\alpha \in X_* : \widetilde{W}(z,\alpha) < 1\}$ is an open subset of X_*, for each $z \in \mathbb{T}$.

Let $(\alpha_n)_{n=1}^\infty$ be a fixed sequence for nonzero vectors in X_* with dense range. Now since $W(\cdot,z)$ is weak* lower semicontinuous, it is given by the formula:

$$W(z,x) = \sup\{|\alpha(x)| : \alpha \in X_* \text{ with } \widetilde{W}(z,\alpha) < 1\}, \tag{4.14}$$

for all $x \in X$ and $z \in \mathbb{T}$. But because the set of such α included in this supremum is open, $W(z,x)$ is equal to:

$$\sup\{|\alpha_n(x)| : n \in \mathbb{N} \text{ with } \widetilde{W}(z,\alpha_n) < 1\} = \sup_{n\in\mathbb{N}} |g_n(z)(x)|, \tag{4.15}$$

where $(g_n)_{n=1}^\infty$ is the sequence of measurable, essentially bounded functions in $\mathrm{L}^1(X_*)$ given by

$$g_n(z) = \widetilde{W}(z,\alpha_n)^{-1} \cdot \alpha_n,$$

for all $n \in \mathbb{N}$ and $\alpha \in X_*$. (They are measurable by hypothesis 1 of the proposition statement. They are essentially bounded since otherwise the essential boundedness over $z \in \mathbb{T}$ of $W(z,x)$, given by the supremum (4.15), would be contradicted for some $x \in X$.)

Let $f \in \mathrm{L}^\infty(X)$. From formula (4.15) we have:

$$W(z,f(z)) = \sup_{n\in\mathbb{N}} |g_n(z)(f(z))|,$$

for all $z \in \mathbb{T}$. Since this is a countable supremum of measurable functions, we find that $W(z,f(z))$ varies measurably with $z \in \mathbb{T}$, as required. Furthermore, the essential supremum of this function is given by:

$$\|f\|_{\infty,W} = \sup\Big\{\mu(K)^{-1}\Big|\int_K g_n(z)(f(z))\,\mathrm{d}\mu(z)\Big| \ : \ n \in \mathbb{N} \text{ and } K \subset \mathbb{T} \text{ measurable with } \mu(K) > 0\Big\}.$$

Since this is a supremum of weak* continuous linear functionals applied to $f \in L^\infty(X)$, we find that $\|\cdot\|_{\infty,W}$ is weak* lower semicontinuous.

Finally to complete the proof for general W, let $W_\varepsilon(z,x) = W(z,x) \vee \varepsilon\|x\|$ for all $z \in \mathbb{T}$ and $x \in X$ and any $\varepsilon > 0$. Then we have

$$W(z, f(z)) = \lim_{\varepsilon \to 0} W_\varepsilon(z, f(z)),$$

for all $f \in L^\infty(X)$ and $z \in \mathbb{T}$. Since $W_\varepsilon(z,\cdot)$ is bounded below for all $z \in \mathbb{T}$ and $\varepsilon > 0$, this is a limit of measurable functions of $z \in \mathbb{T}$, and therefore measurable as required. Finally, observe that $\|f\|_{\infty,W_\varepsilon} = \|f\|_{\infty,W} \vee \varepsilon\|f\|_\infty$ for all $f \in L^\infty(X)$. This is weak* lower semicontinuous in f for all $\varepsilon > 0$. So by Theorem 4.3.8, the seminorm $\|\cdot\|_{\infty,W}$ is also weak* lower semicontinuous, completing the proof. □

It turns out that Proposition 4.3.9 has a converse, which we turn to now. Namely, every sup-type seminorm on $L^\infty(X)$ takes the form $\|\cdot\|_{\infty,W}$.

Theorem 4.3.10 *Let* $\|\cdot\|'$ *be a sup-type seminorm on* $L^\infty(X)$*. Then there is a weight* $W : \mathbb{T} \times X \to [0,\infty)$*, unique up to equality on* X *almost everywhere on* $\mathbb{T}$*, for which the hypotheses of Proposition 4.3.9 are satisfied and the seminorms* $\|\cdot\|_{\infty,W}$ *and* $\|\cdot\|'$ *are identical.*

Proof We shall first establish uniqueness of W. Suppose that W and W' are two weights satisfying the hypotheses of Proposition 4.3.9, such that $\|\cdot\|_{\infty,W}$ and $\|\cdot\|_{\infty,W'}$ are identical to $\|\cdot\|'$. Assume to begin with, that $W(z,\cdot)$ and $W'(z,\cdot)$ are bounded below for all $z \in \mathbb{T}$, and let $(\alpha_n)_{n=1}^\infty$ be a sequence of nonzero vectors in X_* with dense range, as in the proof of Proposition 4.3.9.

For any nonzero $\alpha \in X_*$, let $K_\alpha = \{z \in \mathbb{T} : \widetilde{W}(z,\alpha) \le 1\}$, where $\widetilde{W}$ is defined in terms of W, as in the statement of Proposition 4.3.9. Define K'_α and

$\widetilde{W}'$ analogously, in terms of W'. Now observe that:

$$\mu(K_\alpha \cap L) > 0 \quad \Longleftrightarrow \quad \|x \cdot \chi_L\|' \;\geq\; |\alpha(x)| \text{ for all } x \in X\,,$$

for any measurable subset $L \subset \mathbb{T}$. Similarly for K'_α in place of K_α. This shows that the differences $K_\alpha \setminus K'_\alpha$ and $K'_\alpha \setminus K_\alpha$ both have measure zero. Indeed, if not then set L equal to the one with nonzero measure, to arrive at a contradiction. Hence there is a set $M \subset \mathbb{T}$ of full measure, such that the sets $K_{\alpha_n} \cap M$ and $K'_{\alpha_n} \cap M$ agree for all $n \in \mathbb{N}$.

Now since $\widetilde{W}(z,\cdot)$ is a bounded norm on X_*, the set $\{\alpha \in X_* \,:\, \widetilde{W}(z,\alpha) \leq 1\}$ is the closure of:

$$\{\alpha_n \,:\, n \in \mathbb{N} \text{ with } z \in K_{\alpha_n}\}\,,$$

for all $z \in \mathbb{T}$. Similarly for $\widetilde{W}'$ in place of W, and K'_{α_n} in place of K_{α_n} for each $n \in \mathbb{N}$. Therefore the norms $\widetilde{W}(z,\cdot)$ and $\widetilde{W}'(z,\cdot)$ coincide for all $z \in M$. By formula (4.14) in the proof of Proposition 4.3.9, this shows that the seminorms $W(z,\cdot)$ and $W'(z,\cdot)$ also coincide for all $z \in M$, as required.

To establish equality almost everywhere for general W and W', let $\varepsilon > 0$ and let the weight W_ε be given by $W_\varepsilon(z,x) = W(z,x) \vee \varepsilon\|x\|$ for all $z \in \mathbb{T}$ and $x \in X$. Define W'_ε analogously. Since W_ε and W'_ε give rise to the same seminorm, and since $W_\varepsilon(z,\cdot)$ and $W'_\varepsilon(z,\cdot)$ are bounded below, they are equal for almost all $z \in \mathbb{T}$, by the above argument. Letting $\varepsilon \to 0$, we find that $W(z,\cdot)$ and $W'(z,\cdot)$ are equal for almost all $z \in \mathbb{T}$, establishing uniqueness of W.

To prove existence of the weight W, let us first assume that $\|\cdot\|'$ is bounded below. That is, we suppose that there is some $\delta > 0$ such that $\|f\|' \geq \delta\|f\|_\infty$ for all $f \in \mathrm{L}^\infty(X)$. For any $\alpha \in X_*$, define the set function,

$$\nu_\alpha \,:\, \mathcal{B}(\mathbb{T}) \to [0,\infty]\,,$$

(where $\mathcal{B}(\mathbb{T})$ denotes the Borel subsets on $\mathbb{T}$), by:

$$\nu_\alpha(K) \;=\; \sup\left\{ \mathrm{Re}\left(\int_K (\alpha \circ f)\,\mathrm{d}\mu\right) \,:\, f \in \mathrm{L}^\infty(X) \text{ with } \|f\|' \leq 1\right\}, \qquad (4.16)$$

for all $K \in \mathcal{B}(\mathbb{T})$. Let $(A_n)_{n=1}^\infty$ be a sequence of pairwise disjoint Borel subsets of $\mathbb{T}$. Then it is clear from the definition of ν_α that the inequality,

$$\nu_\alpha\Big(\bigcup_{n=1}^{\infty} A_n\Big) \;\leq\; \sum_{n=1}^{\infty} \nu_\alpha(A_n)\,,$$

holds. Now let $A_1, A_2 \subset \mathbb{T}$ be disjoint Borel sets. Then for any $a_1, a_2 \in \mathbb{R}$ with $a_1 < \nu_\alpha(A_1)$ and $a_2 < \nu_\alpha(A_2)$, there are functions $f_1, f_2 \in \mathrm{L}^\infty(X)$ such that $\|f_1\|', \|f_2\|' \leq 1$ and the inequality,

$$\mathrm{Re}\left(\int_{A_j} (\alpha \circ f_j)\,\mathrm{d}\mu\right) \;>\; a_j\,,$$

holds for $j = 1, 2$. So upon setting $a_3 = a_1 + a_2$ and $A_3 = A_1 \cup A_2$ and also $f_3 = f_1 \cdot \chi_A + f_2 \cdot \chi_B$, the above inequality holds for $j = 3$. But since $\|\cdot\|'$ is a sup-type norm, we have

$$\|f_3\|' \;\leq\; \|f_1\|' \vee \|f_2\|' \;\leq\; 1\,.$$

Hence $\nu_\alpha(A_3) > a_3$. This shows that ν_α is finitely additive. Since it is also countably subadditive and $\nu_\alpha(K) = 0$ whenever K is empty, we deduce that ν_α is a measure on $\mathbb{T}$ for all $\alpha \in X_*$.

Now $\|\cdot\|'$ is weak* lower semicontinuous, so by Proposition 4.2.1 it is bounded (and hence equivalent to $\|\cdot\|_\infty$). That is, there is some $R \geq \delta$ such that $\|f\|' \leq R\|f\|_\infty$ for all $f \in \mathrm{L}^\infty(X)$. From formula (4.16), this shows that $\nu_\alpha(K)$ has lower bound:

$$R^{-1} \cdot \sup\left\{ \mathrm{Re}\left(\int_K (\alpha \circ f)\,\mathrm{d}\mu\right) \;:\; f \in \mathrm{L}^\infty(X) \text{ with } \|f\|_\infty \leq 1 \right\}.$$

But this is equal to $R^{-1}\,\|\alpha\|\,\mu(K)$, for all $\alpha \in X_*$ and $K \in \mathcal{B}(\mathbb{T})$ (where μ is normalized Lebesgue measure on $\mathbb{T}$). Similarly, we find that ν_α has upper bound $\delta^{-1}\,\|\alpha\|\,\mu(K)$ for all $K \in \mathcal{B}(\mathbb{T})$. Hence ν_α is absolutely continuous with respect to μ, for any $\alpha \in X_*$. It therefore has a Radon-Nikodym derivative:

$$\frac{\mathrm{d}\nu_\alpha}{\mathrm{d}\mu} \;:\; \mathbb{T} \to [0, \infty)\,,$$

and this function has lower bound $R^{-1}\|\alpha\|$ and upper bound $\delta^{-1}\|\alpha\|$, almost everywhere on $\mathbb{T}$, for any $\alpha \in X_*$.

Now observe that $\nu_{(\lambda\alpha)} = |\lambda|\nu_\alpha$ and that $\nu_{(\alpha+\beta)} \leq \nu_\alpha + \nu_\beta$, for all $\alpha, \beta \in X_*$ and $\lambda \in \mathbb{C}$. Therefore, using a similar argument to that used in the proof of Lemma 4.3.2, we may construct a function $\widetilde{W} : \mathbb{T} \times X_* \to [0,\infty)$ such that $\widetilde{W}(z,\cdot)$ is a norm on X_*, bounded above by δ^{-1} and below by R^{-1} for all $z \in \mathbb{T}$, and also such that:

$$\widetilde{W}(\cdot,\alpha) \;=\; \frac{\mathrm{d}\nu_\alpha}{\mathrm{d}\mu} \quad \text{a.e.},$$

for every $\alpha \in X_*$. Define W by formula (4.14) from the proof of Proposition 4.3.9. That is, $W(z,\cdot)$ is the dual norm of $\widetilde{W}(z,\cdot)$, which is weak* lower semicontinuous, and bounded above by R and below by δ, for all $z \in \mathbb{T}$. Hence $\widetilde{W}$ is given by formula (4.13) of the statement of Proposition 4.3.9, and so W satisfies the hypotheses of the proposition. It remains to show that the sup-type seminorm $\|\cdot\|_{\infty,W}$ is equal to $\|\cdot\|'$.

For any $f \in \mathrm{L}^\infty(X)$ with $\|f\|' \leq 1$, and any $\alpha \in X_*$, we may deduce from formula (4.16) that:

$$\left(\frac{\mathrm{d}\nu_\alpha}{\mathrm{d}\mu}\right)(z) \;\geq\; |(\alpha \circ f)(z)|,$$

for almost all $z \in \mathbb{T}$. Let S be a countable dense subset of X_*. For all $\alpha \in S$, we have $\widetilde{W}(\cdot,\alpha) \leq |\alpha \circ f|$ almost everywhere on $\mathbb{T}$. Since S is countable, there is a subset $K \subset \mathbb{T}$ of full measure such that $\widetilde{W}(z,\alpha) \leq |(\alpha \circ f)(z)|$ for all $z \in K$ and $\alpha \in S$. But S is dense, so by continuity the inequality holds for all $\alpha \in X_*$. This implies that the supremum $W(z, f(z))$ takes values at most 1, for almost all $z \in \mathbb{T}$. Hence $\|f\|_{\infty,W} \leq 1$.

Conversely, let $f \in \mathrm{L}^\infty(X)$ with $\|f\|' > 1$. By the weak* lower semicontinuity of $\|\cdot\|'$, there is a separating linear functional on $\mathrm{L}^\infty(X)$ given by the function $g \in \mathrm{L}^1(X_*)$. That is, $|\theta(f)(g)| > 1$ and $|\theta(h)(g)| < 1$ for all $h \in \mathrm{L}^\infty(X)$ with $\|h\|' \leq 1$, where $\theta : \mathrm{L}^\infty(X) \to \mathrm{L}^1(X_*)^*$ is the natural isometric isomorphism given by equation (4.9) of Theorem 4.3.3. But now since $\|\cdot\|'$ is bounded below, the set of such $g \in \mathrm{L}^1(X_*)$ with these properties is open. So we may assume g to be a simple function, that is $g = \alpha_1\chi_{A_1} + \cdots + \alpha_n\chi_{A_n}$ for some elements $\alpha_1, \ldots, \alpha_n$ of X_* and disjoint Borel sets $A_1, \ldots, A_n \subset \mathbb{T}$.

By condition 2 of the definition of a sup-type seminorm, we have:

$$\nu_{\alpha_1}(A_1) + \cdots + \nu_{\alpha_n}(A_n) \;=\; \sup_{\|h\|' \le 1} |\theta(h)(g)| \;<\; 1,$$

using a similar argument to the one we used to establish additivity of ν_α for any $\alpha \in X_*$. So by the triangle inequality, we find that there is some $1 \le k \le n$ such that $|\theta(f)(\alpha_k \chi_{A_k})| > \nu_{\alpha_k}(A_k)$. In other words, letting $\alpha = \alpha_k$ and $K = A_k$, we have:

$$\left| \int_K (\alpha \circ f)\, \mathrm{d}\mu \right| \;>\; \int_K \frac{\mathrm{d}\nu_\alpha}{\mathrm{d}\mu}\, \mathrm{d}\mu\,.$$

Therefore $|\alpha(f(z))| > \widetilde{W}(z, \alpha)$ for all z in some subset of $\mathbb{T}$ of measure greater than zero. But for any such z, we find that $W(z, f(z)) > 1$. Therefore $\|f\|_{\infty,W} > 1$, thus establishing equality between $\|\cdot\|'$ and $\|\cdot\|_{\infty,W}$.

Now we complete the proof for general sup-type seminorms $\|\cdot\|'$. For each $n \in \mathbb{N}$, let $\|\cdot\|'_n$ be the sup-type seminorm given by $\|f\|'_n = \|f\|' \vee n^{-1}\|f\|_\infty$ for all $f \in \mathrm{L}^\infty(X)$. Since $\|\cdot\|'_n$ is bounded below, it takes the form $\|\cdot\|_{\infty,W_n}$ for some weight W_n. By uniqueness of W_n, we have $W_n(z,x) = W_{n+1}(z,x) \vee n^{-1}\|x\|$ for all $x \in X$, at almost all points $z \in \mathbb{T}$. So there is a subset $K \subset \mathbb{T}$ of full measure, such that $(W_n(z,\cdot))_{n=1}^\infty$ is a decreasing sequence of weak* lower semicontinuous norms, converging uniformly on the unit ball B_X of X, for all $z \in K$.

For each $z \in K$, set $W(z,\cdot)$ to be the limit of the norms $W_n(z,\cdot)$ as $n \to \infty$. By Theorem 4.3.8, this is a weak* lower semicontinous seminorm. For each $z \in \mathbb{T} \setminus K$ and $x \in X$, set $W(z,x) = 0$. Then W forms a weight, satisfying the hypotheses of Proposition 4.3.9. It is easily verified, moreover, that $\|\cdot\|'$ and $\|\cdot\|_{\infty,W}$ are identical, as required. □

Observe that thoughout most of the proofs of the above two results, we were able to assume that all the relevant seminorms on X and $\mathrm{L}^\infty(X)$ are bounded below. The only use of Theorem 4.3.8 in the above proofs is to extend the results to include weights and sup-type seminorms which are not bounded below. (We say that a weight W is bounded below if there is some $\delta > 0$ such that $W(z,x) \ge \delta\|x\|$ for all $x \in X$ and $z \in \mathbb{T}$.)

4.4 Constrained H^∞ approximation

In Section 4.2, we formulated a constrained approximation problem on a general Banach space X, with a predual X_* and a weak* closed subspace $Y \subset X$. In this section, we shall specialize this setup, making use of the theory developed in Section 4.3. We shall replace X with the space $\mathrm{L}^\infty(X)$ of bounded X-valued functions on the unit circle $\mathbb{T}$, and replace Y with the subspace $\mathrm{H}^\infty(Y)$ of bounded Y-valued Hardy class functions on the unit circle, as defined in Section 4.3. This specialization will allow us to derive some further results later in this section, which are specific to the new framework.

The setup for the new problem is as follows:

- Let X be any Banach space with a fixed separable predual X_*.
- Let $Y \subset X$ be any weak* closed subspace of X.
- Let $W_A, W_B : \mathbb{T} \times X$ be any two weights satisfying the hypotheses of Proposition 4.3.9 as well as the following condition: there is some $\delta > 0$ such that $W_A(z,x) \vee W_B(z,x) \geq \delta\|x\|$ for all $z \in \mathbb{T}$ and $x \in X$.
- Let $\phi_A, \phi_B \in \mathrm{L}^\infty(X)$. The function ϕ_A is to be approximated in $\mathrm{H}^\infty(Y)$, while the function ϕ_B is used to form the constraining condition.

From Proposition 4.3.9, the weights W_A and W_B give rise to well defined sup-type seminorms $\|\cdot\|_{\infty,W_A}$ and $\|\cdot\|_{\infty,W_B}$, which we denote by $\|\cdot\|_A$ and $\|\cdot\|_B$, respectively. These clearly satisfy $\|\psi\|_A \vee \|\psi\|_B \geq \delta\|\psi\|_\infty$ for all $\psi \in \mathrm{L}^\infty(X)$. In other words, the seminorm $\|\cdot\|_A \vee \|\cdot\|_B$ is bounded below. Conversely, it is straightforward to show from Theorem 4.3.10 that any pair of sup-type seminorms $\|\cdot\|_A$ and $\|\cdot\|_B$, for which $\|\cdot\|_A \vee \|\cdot\|_B$ is bounded below, arise from some choice of weights W_A and W_B satisfying the conditions of the above setup.

The problem of interest is stated as follows:

Problem 4.4.1 (Constrained H^∞ approximation) *Given the spaces $Y \subset X$, the seminorms $\|\cdot\|_A$ and $\|\cdot\|_B$ and the functions $\phi_A, \phi_B \in \mathrm{L}^\infty(X)$ as defined* *above, and given $M > 0$, find $g \in \mathrm{H}^\infty(Y)$ which minimizes:*

$$\|\phi_A - g\|_A ,$$

subject to the constraint that:

$$\|\phi_B - g\|_B \leq M .$$

Analogously to the definition in the statement of Theorem 4.2.3, we define the function $f : [0, \infty] \to [0, \infty]$ to be the infimum:

$$f(M) = \inf\{\|\phi_A - g\|_A : g \in \mathrm{H}^\infty(Y) \text{ and } \|\phi_B - g\|_B \leq M\},$$

where this is taken to be ∞ whenever the above set is empty. Now we may apply Theorem 4.2.3 to the specialized setup above. We obtain the following result immediately:

Theorem 4.4.2 *Let X,Y,$\|\cdot\|_A$,$\|\cdot\|_B$ and ϕ_A,$\phi_B \in \mathrm{L}^\infty(X)$ and the function f be given as above. Then $f(M)$ is finite for sufficiently large $M \in [0, \infty]$, and there exists a solution $g \in \mathrm{H}^\infty(Y)$ to Problem 4.4.1 with respect to the bound M, whenever M and $f(M)$ are both finite. Moreover, f is decreasing and convex, and letting*

$$M_0 = \inf\{M \in [0, \infty) : f(M) < \infty\} ,$$

it holds that f is continuous from $[M_0, \infty]$ to $[0, \infty]$, with respect to the standard compact topologies on each of these intervals. In particular, $f(M)$ tends to $f(\infty)$ as the bound M tends to ∞.

We now turn to the results which are specific to the new framework, specialized from Section 4.2. We shall establish, under certain conditions, saturation and uniqueness results concerning the solution to Problem 4.4.1. It will be seen in the next section that these provide a generalization of a well known result (c.f. [26, Theorem 1.4]), which establishes uniqueness and constant modulus of the best uniform H^∞ approximant to any continuous function on $\mathbb{T}$.

For any Banach space E, let $\mathrm{C}(E)$ denote the Banach space of continuous E-valued functions on the unit circle $\mathbb{T}$, equipped with the uniform norm. Thus

$\mathrm{C}(X)$ and $\mathrm{C}(Y)$ may be regarded as a closed subspaces of $\mathrm{L}^\infty(X)$. The space $\mathrm{C}(Y)$ is involved in the remaining results of this section.

First of all, we will need the following lemma. This result is closely related to Mergelyan's theorem, which states that for a compact set $S \subset \mathbb{C}$ with connected complement in $\mathbb{C}$, any continuous function from S to $\mathbb{C}$ which is analytic on the interior of S may be uniformly approximated arbitrarily well by polynomials.

Lemma 4.4.3 *Let* $K \subset \mathbb{T}$ *be a measurable set with* $\mu(K) > 0$. *Let* $\phi \in \mathrm{C}(Y)$ *and let* $\varepsilon > 0$. *Then there exists* $g \in \mathrm{H}^\infty(Y)$ *such that* $\|(\phi - g) \cdot \chi_{K^c}\|_\infty < \varepsilon$.

Proof Let $n \in \mathbb{N}$ and let us suppose, first of all, that $\hat{\phi}(k) = 0$ for all $k \leq -n$. In the case $n = 1$, this implies that $\phi \in \mathrm{H}^\infty(Y)$, so we may set $g = \phi$ and then we are done. Otherwise, consider the weight $w = \chi_{K^c} + \lambda\chi_K$, where we set:

$$\lambda = \left((1 + \|\hat{\phi}(1-n)\|)^{-1} \cdot 2^{-n}\varepsilon\right)^{1/\mu(K)} > 0 .$$

By Lemma 4.1.4, there is a unique scalar-valued function $h \in \mathrm{H}^\infty$ such that the weighted norm $\|\zeta - h\|_{\infty,w}$ is minimal, where $\zeta \in \mathrm{C}$ is given by $\zeta(z) = z^{-1}$ for all $z \in \mathbb{T}$. Moreover, $\|\zeta - h\|_{\infty,w} = \lambda^{\mu(K)}$, the geometric mean of w over the unit circle $\mathbb{T}$.

Let $\psi = \phi - \zeta^{n-2}(\zeta - h)\hat{\phi}(1-n)$. Then $\hat{\psi}(k) = 0$ for every $k \leq 1 - n$, so we may assume by induction on n that there exists $g \in \mathrm{H}^\infty(Y)$ such that $\|(\psi - g) \cdot \chi_{K^c}\|_\infty < (1 - 2^{1-n})\varepsilon$. But now we have:

$$\begin{aligned} \|(\phi - g) \cdot \chi_{K^c}\|_\infty &\leq \|(\phi - \psi) \cdot \chi_{K^c}\|_\infty + \|(\psi - g) \cdot \chi_{K^c}\|_\infty \\ &< \|\hat{\phi}(1-n)\| \cdot \|(\zeta - h) \cdot \chi_{K^c}\|_\infty + (1 - 2^{1-n})\varepsilon \\ &\leq \|\hat{\phi}(1-n)\| \cdot \|\zeta - h\|_{\infty,w} + (1 - 2^{1-n})\varepsilon \\ &< 2^{-n}\varepsilon + (1 - 2^{1-n})\varepsilon , \end{aligned}$$

so that $\|(\phi - g) \cdot \chi_{K^c}\|_\infty < (1 - 2^{-n})\varepsilon$, which completes the inductive step. This establishes the desired result, for any $n \in \mathbb{N}$ and $\phi \in \mathrm{C}(Y)$ such that $\hat{\phi}(k) = 0$ for all $k \leq -n$.

To complete the proof for general $\phi \in \mathrm{C}(Y)$, we use the fact that the Y-valued trigonometric polynomials on $\mathbb{T}$ are uniformly dense in $\mathrm{C}(Y)$. This fact may be proved using Fejér summation, to obtain a series converging uniformly to any given element of $\mathrm{C}(Y)$. So there exists $n \in \mathbb{N}$ and some $2n-1$ elements $y_{1-n}, y_{2-n}, \ldots, y_{n-1} \in Y$, such that $\|\phi - \psi\|_\infty < \varepsilon/2$, where

$$\psi(z) = \sum_{k=1-n}^{n-1} y_k \cdot z^k ,$$

for all $z \in \mathbb{T}$. But $\hat{\psi}(k) = 0$ for all $k \leq -n$, so there exists $g \in \mathrm{H}^\infty(Y)$ such that $\|(\psi - g) \cdot \chi_{K^\mathrm{C}}\|_\infty < \varepsilon/2$. But now we have:

$$\begin{aligned} \|(\phi - g) \cdot \chi_{K^\mathrm{C}}\|_\infty &\leq \|\phi - \psi\|_\infty + \|(\psi - g) \cdot \chi_{K^\mathrm{C}}\|_\infty \\ &< \varepsilon/2 + \varepsilon/2 \;=\; \varepsilon , \end{aligned}$$

as required. □

Observe that the conclusions of the above lemma hold equally well when $\phi \in \mathrm{H}^\infty(Y) + \mathrm{C}(Y)$. Now using this lemma, we shall first establish the saturation result concerning the solutions to Problem 4.4.1, which is stated as follows:

Theorem 4.4.4 *Let* $X, Y, \|\cdot\|_A, \|\cdot\|_B$ *and* $\phi_A, \phi_B \in \mathrm{L}^\infty(X)$ *and the function* f *be given as in the statement of Theorem 4.4.2. Let* $M \in [0,\infty)$ *and suppose that* $f(M) < \infty$*, so that there exists a solution* $g \in \mathrm{H}^\infty(Y)$ *to Problem 4.4.1, with respect to this bound* M*. Suppose furthermore that there exists some* Y*-valued function* $h \in \mathrm{H}^\infty(Y) + \mathrm{C}(Y)$ *such that:*

$$\begin{aligned} \|\phi_A - h\|_A &< f(M) \\ \|\phi_B - h\|_B &< M \end{aligned}$$

Then for almost all $z \in \mathbb{T}$*, at least one of the following equations holds:*

$$W_A(z, (\phi_A - g)(z)) = f(M) \tag{4.17}$$

$$W_B(z, (\phi_B - g)(z)) = M \tag{4.18}$$

(In words, at almost all points in $\mathbb{T}$, at least one of the two constraints imposed by the bounds on $\|\phi_A - g\|_A$ and $\|\phi_B - g\|_B$ is saturated.)

Proof Let g be a solution to Problem 4.4.1 for the bound M, and let $N = f(M)$. Suppose, for a contradiction, that equations (4.17) and (4.18) both fail for all z in some subset of measure greater than 0. Then there is some set $K \subset \mathbb{T}$ measurable with $\mu(K) > 0$, such that $\|(\phi_A - g)\cdot\chi_K\|_A < N$ and $\|(\phi_B - g)\cdot\chi_K\|_B < M$.

Let h be given as in the theorem statement. Let $R > 0$ be chosen such that $\|\psi\|_A \vee \|\psi\|_B \le R\,\|\psi\|_\infty$ for all $\psi \in \mathrm{L}^\infty(X)$. Now let $\varepsilon > 0$ be given by:

$$\varepsilon \;=\; R^{-1}((N - \|\phi_A - h\|_A) \wedge (M - \|\phi_B - h\|_B))\,.$$

By Lemma 4.4.3, there exists $\tilde{h} \in \mathrm{H}^\infty(Y)$ such that $\|(\tilde{h} - h)\cdot\chi_{K^{\mathrm{c}}}\|_\infty < \varepsilon$. Thus we have the inequalities:

$$\begin{aligned}\|(\phi_A - \tilde{h})\cdot\chi_{K^{\mathrm{c}}}\|_A \;&\le\; \|(\phi_A - h)\cdot\chi_{K^{\mathrm{c}}}\|_A \;+\; \|(h - \tilde{h})\cdot\chi_{K^{\mathrm{c}}}\|_A \\ &\le\; \|\phi_A - h\|_A \;+\; R\,\|(h - \tilde{h})\cdot\chi_{K^{\mathrm{c}}}\|_\infty \\ &<\; \|\phi_A - h\|_A \;+\; (N - \|\phi_A - h\|_A)\,,\end{aligned}$$

and similarly for ϕ_B in place of ϕ_A and $\|\cdot\|_B$ in place of $\|\cdot\|_A$. Hence,

$$\begin{aligned}\|(\phi_A - \tilde{h})\cdot\chi_{K^{\mathrm{c}}}\|_A \;&<\; N \\ \|(\phi_B - \tilde{h})\cdot\chi_{K^{\mathrm{c}}}\|_B \;&<\; M\,.\end{aligned}$$

Now choose $\lambda \in (0, 1)$ sufficiently small that:

$$\begin{aligned}(1-\lambda)\,\|(\phi_A - g)\cdot\chi_K\|_A \;+\; \lambda\,\|(\phi_A - \tilde{h})\cdot\chi_K\|_A \;&<\; N \\ (1-\lambda)\,\|(\phi_B - g)\cdot\chi_K\|_B \;+\; \lambda\,\|(\phi_B - \tilde{h})\cdot\chi_K\|_B \;&<\; M\,.\end{aligned}$$

Then upon setting $\tilde{g} \;=\; (1-\lambda)g + \lambda\tilde{h} \;\in\; \mathrm{H}^\infty(Y)$, we find that:

$$\begin{aligned}\|\phi_A - \tilde{g}\|_A \;&\le\; \|(\phi_A - \tilde{g})\cdot\chi_K\|_A \;\vee\; \|(\phi_A - \tilde{g})\cdot\chi_{K^{\mathrm{c}}}\|_A \\ &\le\; \Big[(1-\lambda)\,\|(\phi_A - g)\cdot\chi_{K^{\mathrm{c}}}\|_A \;+\; \lambda\,\|(\phi_A - \tilde{h})\cdot\chi_{K^{\mathrm{c}}}\|_A\Big] \\ &\qquad \vee\; \Big[(1-\lambda)\,\|(\phi_A - g)\cdot\chi_K\|_A \;+\; \lambda\,\|(\phi_A - \tilde{h})\cdot\chi_K\|_A\Big] \\ &<\; ((1-\lambda)N \;+\; \lambda N) \;\vee\; N \;=\; N\,.\end{aligned}$$

Similarly, we find that $\|\phi_B - \tilde{g}\|_B < M$. Thus $\tilde{g}$ is a better constrained $\mathrm{H}^\infty(Y)$ approximant to ϕ_A than g, which is a contradiction. □

A special case where all the hypotheses of Theorem 4.4.4 hold is when $\phi_A = \phi_B = \phi$, for some function $\phi \in \mathrm{H}^\infty(Y) + \mathrm{C}(Y)$, and we have $0 < f(M) < \infty$ for a given $0 < M < \infty$. Then all the hypotheses are satisfied when we take $h = \phi$.

In the following corollary to Theorem 4.4.4, we show how pointwise saturation of the constraints in the above way, leads to uniqueness of the solution to Problem 4.4.1, under certain circumstances.

Corollary 4.4.5 *Let* X,Y,$\|\cdot\|_A$,$\|\cdot\|_B$ *and* $\phi_A, \phi_B \in \mathrm{L}^\infty(X)$ *be given as in the statement of Theorem 4.4.2. Let* $M > 0$ *and suppose that all the hypotheses of Theorem 4.4.4 hold for this bound* M. *Suppose, additionally, that the seminorms* $W_A(z,\cdot)$ *and* $W_B(z,\cdot)$ *are strictly convex or zero for almost all* $z \in \mathbb{T}$. *Then there exists a unique solution* $g \in \mathrm{H}^\infty(Y)$ *to Problem 4.4.1, with respect to the bound* M.

Proof Let $N = f(M)$, where the function f is defined as earlier. Suppose, for a contradition, that $g_1, g_2 \in \mathrm{H}^\infty$ are two different solutions to Problem 4.4.1, with respect to the bound M. Then there is some measurable set $K \subset \mathbb{T}$ with $\mu(K) > 0$, such that $g_1(z) \neq g_2(z)$ for all $z \in K$.

Let $g_3 = (g_1 + g_2)/2$. Then g_3 is also a solution to Problem 4.4.1, with respect to the bound M. Now $W_A(z, (\phi_A - g_k)(z)) \leq N$ for all $z \in K$ and $k = 1, 2$. So because $g_1(z) \neq g_2(z)$, $N > 0$ and $W_A(z,\cdot)$ is strictly convex or zero, this implies that $W_A(z, (\phi_A - g_3)(z)) < N$ for all $z \in K$. Similarly, $W_B(z, (\phi_B - g_3)(z)) < M$ for all $z \in K$.

But this is in contradiction with Theorem 4.4.4, applied to the solution g_3, which asserts that the inequalities,

$$\begin{aligned} W_A(z, (\phi_A - g_3)(z)) &< N \\ W_B(z, (\phi_B - g_3)(z)) &< M\,, \end{aligned}$$

may hold only for those z in some set of measure zero. □

4.5 Some applications

In this section, we shall give an example of a more concrete H^∞ approximation problem which fits into the framework of the previous section, so that we can apply some of the earlier results. We shall start with the following setup:

Let $n, m \in \mathbb{N}$ and let $X = \mathcal{L}(\mathbb{C}^n, \mathbb{C}^m)$, the space of $n \times m$ matrices equipped with the Hilbert space operator norm, together with its unique predual X_*. Let $P, R \in \mathrm{L}^\infty(\mathcal{L}(\mathbb{C}^n))$ and let $Q, S \in \mathrm{L}^\infty(\mathcal{L}(\mathbb{C}^m))$. With these matrix-valued functions P, Q, R, S, we define the sup-type seminorms $\|\cdot\|_A$ and $\|\cdot\|_B$ by:

$$\begin{aligned} \|\psi\|_A &= \|Q\psi P\|_\infty \\ \|\psi\|_B &= \|S\psi R\|_\infty, \end{aligned}$$

for all $\psi \in \mathrm{L}^\infty(X)$, where the above matrix products are taken pointwise over $\mathbb{T}$. These correspond to the respective weights $W_A, W_B : \mathbb{T} \times X \to [0, \infty)$, where $W_A(z, T) = \|Q(z)TP(z)\|$ and $W_B(z, T) = \|S(z)TR(z)\|$ for all points $z \in \mathbb{T}$ and matrices $T \in X$. It is easily seen that these weights satisfy the hypotheses of Proposition 4.3.9, and so $\|\cdot\|_A$ and $\|\cdot\|_B$ are well defined sup-type seminorms, as claimed. Additionally, we shall suppose that the functions P, Q, R, S are chosen such that $\|\cdot\|_A \vee \|\cdot\|_B$ is bounded below.

For example, we could take $Q = S = \mathrm{I}_m$, the $m \times m$ identity matrix, and take $P = w_A \mathrm{I}_n$ and $R = w_B \mathrm{I}_n$, where $w_A, w_B \in \mathrm{L}^\infty_{\mathrm{pos}}$ are scalar weights. Then $\|\cdot\|_A$ and $\|\cdot\|_B$ would be equal to $\|\cdot\|_{\infty, w_A}$ and $\|\cdot\|_{\infty, w_B}$ respectively. In this case, the condition that $\|\cdot\|_A \vee \|\cdot\|_B$ is bounded below is equivalent to the condition that the weight $w_A \vee w_B$ has a bounded inverse.

Let us now consider the following specific problem:

Problem 4.5.1 *Let $X = \mathcal{L}(\mathbb{C}^n, \mathbb{C}^m)$. Given $\phi \in \mathrm{L}^\infty(X)$, a bound $M > 0$ and* *the matrix-valued functions P, Q, R, S as above, let*

$$N = \inf\{\|Q(\phi - g)P\|_\infty : g \in \mathrm{H}^\infty(X) \text{ with } \|S(\phi - g)R\|_\infty \le M\},$$

which we take to be ∞ if the above set is empty. Find $g_0 \in \mathrm{H}^\infty(X)$ such that $\|Q(\phi - g_0)P\|_\infty = N$ and $\|S(\phi - g_0)R\|_\infty \le M$.

By setting $Y = X$ and $\phi_A = \phi_B = \phi$, we find that the above problem is just a special case of Problem 4.4.1. So from the results of Section 4.4, it is possible to deduce the following facts concerning solutions to the above problem:

1. By Theorem 4.4.2, a solution g_0 does exist, provided $N < \infty$.

2. Unless $N = \inf\{\|Q(\phi - g)P\|_\infty : g \in \mathrm{H}^\infty(X)\}$, then any solution g_0 will satisfy $\|S(\phi - g_0)T\|_\infty = M$. Indeed, N is a convex and decreasing function of M, and N approaches this infimum as $M \to \infty$, by Theorem 4.4.2. This implies that any decrease in M will cause an increase in N, unless N is equal to the infimum.

Now consider the case that $\phi \in \mathrm{H}^\infty(X) + \mathrm{C}(X)$ with $N > 0$, and that there is a measurable set $K \subset \mathbb{T}$ such that the pointwise Kronecker product $P \otimes Q$ vanishes on K^{C} and similarly the product $R \otimes S$ vanishes on K. Then we have the following additional facts:

3. For any solution g_0, we have $\|Q(z)(\phi - g_0)(z)P(z)\| = N$ for almost all $z \in K$ and $\|S(z)(\phi - g_0)(z)R(z)\| = M$ for almost all $z \in K^{\mathrm{C}}$. This follows from Theorem 4.4.4.

4. If $n = 1$ or $m = 1$ then Corollary 4.4.5 applies, so the solution g_0 is unique.

One special case of Problem 4.5.1 is when $P = \chi_K \mathrm{I}_n$, $R = \chi_{K^{\mathrm{C}}} \mathrm{I}_n$ and $Q = S = \mathrm{I}_m$, for some measurable subset $K \subset \mathbb{T}$. This corresponds to the problem of finding the best uniform matrix-valued H^∞ approximant to a given matrix-valued L^∞ function ϕ, where the approximation is restricted to a subset of the unit circle $\mathbb{T}$, with a bound M on the uniform norm of the error on the complementary subset.

This is a matrix version of the scalar H^∞ approximation problem considered in [2, sec. 1.2]. The above four facts provide a direct analogue to [2, Theorem 2], which describes the solutions to this problem.

In particular, when $K = \mathbb{T}$ and $n = m = 1$, we recover the standard Nehari problem of finding the best uniform scalar-valued H^∞ approximant to $\phi \in \mathrm{L}^\infty$. When $\phi \in \mathrm{H}^\infty + \mathrm{C}$, the last two facts above imply the well known result that the solution is unique and has a.e. constant modulus (c.f. [26, Theorem 1.4]).

Chapter 5

Miscellaneous results

5.1 The matrix inequality: $\|N^*M\|_p \leq \|M\|_q\|N\|_r$

In this section, we establish the above matricial generalization to Hölder's inequality, for all $M, N \in \mathcal{L}(\mathbb{C}^n)$. We shall do this (for $p > 1$ and $q, r < \infty$) by showing that for any given matrices $M, N \in \mathcal{L}(\mathbb{C}^n)$, with $\|M\|_q = 1$ and with M chosen to maximize $\|N^*M\|_p$, the matrix N must be diagonal whenever M is diagonal. The latter statement can always be made to hold by applying suitable unitary transformations to the domain and codomain of M. Then the result becomes a straightforward application of the usual (scalar) Hölder inequality.

Note that this generalization is already known (which I later discovered), as referenced near the start of Chapter 1. In the book [4, ch. 11], a more conventional proof of the result is given using *Horn inequalities*, which are not described here. See the reference for details.

Our method outlined above will depend crucially on the following result, which shows how $\|M\|_p$ changes as we perturb $M \in \mathcal{L}(\mathbb{C}^n)$:

Lemma 5.1.1 *Let $M, N \in \mathcal{L}(\mathbb{C}^n)$ with M nonsingular, and let $p \in [1, \infty)$. Then we have:*

$$\|M + \varepsilon N\|_p^p \quad = \quad \|M\|_p^p \, + \, p\varepsilon \, \mathrm{Re}\,\mathrm{Tr}((M^*M)^{p/2-1}M^*N) \, + \, \mathrm{o}(\varepsilon)\,,$$

as $\varepsilon \to 0$. In the case $p > 1$, the mapping $M \mapsto (M^*M)^{p/2-1}M^*$ extends continuously to all $M \in \mathcal{L}(\mathbb{C}^n)$, and when this continuous limit is substituted into the above formula, it holds for all $M \in \mathcal{L}(\mathbb{C}^n)$.

Proof Suppose without loss of generality that $\|M\|_\infty < 1$. By definition, we have:

$$\begin{aligned} \|M + \varepsilon N\|_p^p &= \operatorname{Tr}((M^* + \varepsilon N^*)(M + \varepsilon N))^{p/2} \\ &= \operatorname{Tr}(M^*M + \varepsilon(M^*N + N^*M) + \varepsilon^2 N^*N)^{p/2} \\ &= \operatorname{Tr}(I + X + \varepsilon Y)^{p/2}, \end{aligned}$$

where $X = M^*M - I$ and $Y = M^*N + N^*M + \varepsilon N^*N$.

Now M^*M is positive and nonsingular with $\|M^*M\|_\infty < 1$, and therefore $\|X\|_\infty < 1$. So we have the series expansion:

$$\begin{aligned} \|M + \varepsilon N\|_p^p &= \operatorname{Tr}(I + (X + \varepsilon Y))^{p/2} \\ &= n + \sum_{k=1}^{\infty} \frac{(p/2)(p/2-1)\cdots(p/2-k+1)}{k!} \operatorname{Tr}(X + \varepsilon Y)^k \quad (5.1) \end{aligned}$$

provided ε is sufficiently small that $\|X\|_\infty + \varepsilon\|Y\|_\infty < 1$. Similarly, we have the series expansions:

$$\|M\|_p^p = n + \sum_{k=1}^{\infty} \frac{(p/2)(p/2-1)\cdots(p/2-k+1)}{k!} \operatorname{Tr} X^k \quad (5.2)$$

$$\begin{aligned} \operatorname{Tr}((M^*M)^{p/2-1}Y) &= \operatorname{Tr} Y + \sum_{l=1}^{\infty} \frac{(p/2-1)(p/2-2)\cdots(p/2-l)}{l!} \operatorname{Tr}(X^l Y) \\ &= \frac{2}{p}\sum_{k=1}^{\infty} \Big[\frac{(p/2)(p/2-1)\cdots(p/2-k+1)}{k!} \times \\ &\qquad k \operatorname{Tr}(X^{k-1}Y)\Big] \quad (5.3) \end{aligned}$$

with all the above series converging absolutely.

For any natural number $k > 2$, we have:

$$(X + \varepsilon Y)^k = X^k + \varepsilon(X^{k-1}Y + X^{k-2}YX + \cdots + YX^{k-1}) + \text{h.o.t.},$$

and so:

$$\begin{aligned}\left|\operatorname{Tr}(X+\varepsilon Y)^k - \operatorname{Tr} X^k - k\varepsilon \operatorname{Tr}(X^{k-1}Y)\right| &\leq n\Big[(\|X\|_\infty + \varepsilon\|Y\|_\infty)^k \\ &\qquad - \|Y\|_\infty^k - k\varepsilon\|X\|_\infty^{k-1}\|Y\|_\infty\Big] \\ &\leq (n\varepsilon^2/\lambda^2)(\|X\|_\infty + \lambda\|Y\|_\infty)^k ,\end{aligned}$$

for any natural number k and any $\lambda > \varepsilon$. So from the series expansions (5.1),(5.2) and (5.3), we have:

$$\begin{aligned}&\Big|\, \|M+\varepsilon N\|_p^p - \|M\|_p^p - (p\varepsilon/2)\operatorname{Tr}((M^*M)^{p/2-1}Y)\Big| \\ &\leq \sum_{k=1}^{\infty} \frac{(p/2)(p/2-1)\cdots(p/2-k+1)}{k!}\Big|\operatorname{Tr}(X+\varepsilon Y)^k - \operatorname{Tr} X^k \\ &\qquad\qquad - k\varepsilon \operatorname{Tr}(X^{k-1}Y)\Big| \\ &\leq \sum_{k=1}^{\infty} \frac{(p/2)(p/2-1)\cdots(p/2-k+1)}{k!} \cdot (n\varepsilon^2/\lambda^2)(\|X\|_\infty + \lambda\|Y\|_\infty)^k \\ &= (n\varepsilon^2/\lambda^2)\Big((1+\|X\|_\infty + \lambda\|Y\|_\infty)^{p/2} - 1\Big),\end{aligned}$$

provided $\|X\| + \lambda\|Y\| < 1$ and $\varepsilon < \lambda$. Choosing such a λ, we find that:

$$\|M+\varepsilon N\|_p^p = \|M\|_p^p + (p\varepsilon/2)\operatorname{Tr}((M^*M)^{p/2-1}Y) + o(\varepsilon),$$

for $0 < \varepsilon < \lambda$. But now:

$$\begin{aligned}\operatorname{Tr}((M^*M)^{p/2-1}Y) &= \operatorname{Tr}((M^*M)^{p/2-1}(M^*N + N^*M)) + O(\varepsilon) \\ &= 2\operatorname{Re}\operatorname{Tr}((M^*M)^{p/2-1}M^*N)) + O(\varepsilon),\end{aligned}$$

and by substituting this into the previous equation, the equation in the statement of the lemma follows.

To show that the mapping $M \mapsto (M^*M)^{p/2-1}M^*$ extends continuously to all $M \in \mathcal{L}(\mathbb{C}^n)$, define the family of continuous functions:

$$\begin{aligned}f_{p,\eta} &: \mathcal{L}(\mathbb{C}^n) \to \mathcal{L}(\mathbb{C}^n) \\ f_{p,\eta} &: M \mapsto (M^*M + \eta I)^{p/2-1}M^*,\end{aligned}$$

for all $\eta > 0$, $p \in (1, \infty)$. Then we have:

$$\begin{aligned}\|f_{p,\eta_1}(M) - f_{p,\eta_2}(M)\|_\infty &= \|[(M^*M + \eta_1 I)^{p-2} - (M^*M + \eta_2 I)^{p-2}]M^*M\|_\infty^{1/2} \\ &= \|g(M^*M)\|_\infty^{1/2},\end{aligned}$$

where $g(x) = x((x+\eta_1)^{p-2} - (x+\eta_2)^{p-2})$ for $x \geq 0$. In the case $0 < \eta_2 \leq \eta_1$, this function is dominated in modulus by:

$$\begin{aligned}&(x+\eta_2) \cdot |(x+\eta_1)^{p-2} - (x+\eta_2)^{p-2}| \\ &\quad \leq \; (x+\eta_1)^{p-1} - (x+\eta_2)^{p-1} + (\eta_1 - \eta_2)(x+\eta_1)^{p-2} \\ &\quad \leq \; (x+\eta_1)^{p-1} - (x+\eta_2)^{p-1} + (\eta_1 - \eta_2)^{(p-1)/p}(x+\eta_1)^{(p-1)^2/p}.\end{aligned}$$

A similar bound holds for $0 < \eta_1 \leq \eta_2$. Hence

$$\begin{aligned}\|f_{p,\eta_1}(M) - f_{p,\eta_2}(M)\|_\infty^2 &\leq \left|(\|M\|_\infty^2 + \eta_1)^{p-1} - (\|M\|_\infty^2 + \eta_2)^{p-1}\right| + |\eta_1 - \eta_2| \\ &\quad + |\eta_1 - \eta_2|^{(p-1)/p} \cdot (\|M\|_\infty^2 + \eta_1 + \eta_2)^{(p-1)^2/p} \\ &\rightarrow 0\end{aligned}$$

locally uniformly in $M \in \mathcal{L}(\mathbb{C}^n)$, as $\eta_1, \eta_2 \rightarrow 0^+$. Therefore we obtain a continuous, local uniform limit:

$$f_p = \lim_{\eta \rightarrow 0^+} f_{p,\eta}. \tag{5.4}$$

Moreover, we find that $f_p(M) = (M^*M)^{p/2-1}M^*$ for all nonsingular $M \in \mathcal{L}(\mathbb{C}^n)$. Hence the function f_p is the unique continuous extension to $\mathcal{L}(\mathbb{C}^n)$ of the mapping $M \mapsto (M^*M)^{p/2-1}M^*$, as required.

Finally, to show that:

$$\|M + \varepsilon N\|_p^p = \|M\|_p^p + p\varepsilon \operatorname{Re} \operatorname{Tr}(f_p(M)N) + o(\varepsilon), \tag{5.5}$$

as $\varepsilon \rightarrow 0$, for all $M, N \in \mathcal{L}(\mathbb{C}^n)$, observe that for M nonsingular and $\lambda, \mu \in \mathbb{R}$ with $\mu \neq 0$, we have:

$$(M + i\mu N) + \lambda N = (\lambda + i\mu)M(M^{-1}N + (\lambda + i\mu)^{-1}I),$$

and this is nonsingular provided $-(\lambda + i\mu)^{-1}$ is not in the spectrum of $M^{-1}N$. Hence there are $\mu \neq 0$ arbitrary close to zero, such that the right hand side is nonsingular for all $\lambda \in \mathbb{R}$. This shows that for any $M, N \in \mathcal{L}(\mathbb{C}^n)$, there exist matrices $\widetilde{M} \in \mathcal{L}(\mathbb{C}^n)$ arbitrarily close to M, such that $\widetilde{M} + \lambda N$ is nonsingular for all $\lambda \in \mathbb{R}$. For any such perturbation $\widetilde{M}$ of M, we have:

$$\|\widetilde{M} + \lambda N\|_p^p \;=\; \|\widetilde{M}\|_p^p + p\int_0^\lambda \operatorname{Re}\operatorname{Tr}(f_p(\widetilde{M} + tN)N)\, dt\,,$$

for all $\lambda \in \mathbb{R}$. Now observing that the functions $\|\cdot\|_p$ and f_p are continuous, we find that both sides of the above equation converge as $\widetilde{M} \to M$ for any fixed λ, and so the above equation also holds for $\widetilde{M} = M$. Hence equation (5.5) holds for all $M, N \in \mathcal{L}(\mathbb{C}^n)$, as required. □

Now we prove the main result of this section, before finally deducing the other crucial properties of the functions $\|\cdot\|_p$.

Theorem 5.1.2 *Let $M, N \in \mathcal{L}(\mathbb{C}^n)$ and let $p, q, r \in [1, \infty]$ be indices satisfying $1/p = 1/q + 1/r$. Then we have:*

$$\|N^*M\|_p \;\leq\; \|M\|_q\|N\|_r\,.$$

Proof *Case $q, r = \infty$.* It is immediate from its definition that $\|\cdot\|_\infty$ is a multiplicative norm. Now by the Cauchy-Schwarz inequality, we have

$$\|N\|_\infty = \sup\{|\langle Nu, v\rangle| \;:\; u, v \in \mathbb{C}^n \text{ with } \|u\|_2, \|v\|_2 \leq 1\}\,,$$

from which it follows that $\|N^*\|_\infty = \|N\|_\infty$. This establishes the case $q, r = \infty$.

Case $p, q, r < \infty$. It suffices to show that:

$$\|N\|_r \;=\; \max\{\|N^*M\|_p \;:\; \|M\|_q = 1\}\,, \tag{5.6}$$

for any matrix $N \in \mathcal{L}(\mathbb{C}^n)$. Observe that since the set of $M \in \mathcal{L}(\mathbb{C}^n)$ with $\|M\|_q = 1$ is compact, the above maximum is always attained. Compactness of this set also implies that the maximum is a continuous function of $p \in [1, q)$, so we may assume without loss of generality that $p > 1$.

Let $M \in \mathcal{L}(\mathbb{C}^n)$ be chosen such that the maximum is attained, with $\|M\|_q = 1$. We may choose unitary matrices $U, V, W \in \mathcal{L}(\mathbb{C}^n)$ such that VMW is positive and diagonal, and $U^*(N^*M)W$ is positive. Now let:

$$C = VNU\,, \quad D = VMW\,.$$

Then $C^*D = U^*(N^*M)V$ is positive with $\|C^*D\|_p = \|N^*M\|_p$. We shall show that C^*D is in fact positive and diagonal, as well as D.

Let $P \in \mathcal{L}(\mathbb{C}^n)$ be any matrix whose diagonal entries are zero. We obtain the following from Lemma 5.1.1:

$$\begin{aligned} \|D + \varepsilon P\|_q^q &= \|D\|_q^q + q\varepsilon \operatorname{Re}\operatorname{Tr}(D^{q-1}P) + o(\varepsilon) \\ &= 1 + o(\varepsilon)\,, \end{aligned}$$

as $\varepsilon \to 0$, since the diagonal entries of $D^{q-1}P$ are zero, and $\|D\|_q = \|M\|_q = 1$.

We find by a similar application of Lemma 5.1.1 that:

$$\|C^*(D + \varepsilon P)\|_p^p = \|C^*D\|_p^p + p\varepsilon \operatorname{Re}\operatorname{Tr}((C^*D)^{p-1}C^*P) + o(\varepsilon)\,,$$

as $\varepsilon \to 0$. Now for any ε sufficiently small, we may let $\widetilde{D} = (D + \varepsilon P)/\,\|D + \varepsilon P\|_q$ and let $\widetilde{M} = V^*\widetilde{D}W^*$. Then $\|\widetilde{M}\|_q = 1$ and we have:

$$\begin{aligned} \|N^*\widetilde{M}\|_p^p &= \|C^*(D + \varepsilon P)\|_p^p \,/\, \|D + \varepsilon P\|_q \\ &= \|C^*D\|_p^p + p\varepsilon \operatorname{Re}\operatorname{Tr}((C^*D)^{p-1}C^*P) + o(\varepsilon) \\ &= \|N^*M\|_p^p + p\varepsilon \operatorname{Re}\operatorname{Tr}((C^*D)^{p-1}C^*P) + o(\varepsilon)\,. \end{aligned}$$

But $\|N^*M\|_p$ is maximal for $M \in \mathcal{L}(\mathbb{C}^n)$ with $\|M\|_q = 1$. This implies that:

$$\operatorname{Re}\operatorname{Tr}((C^*D)^{p-1}C^*P) = 0\,,$$

since otherwise $\|N^*\widetilde{M}\|_p$ would be greater than $\|N^*M\|_p$ for either sufficiently small positive ε or sufficiently small negative ε.

Since this is true for all $P \in \mathcal{L}(\mathbb{C}^n)$ with diagonal entries equal to zero, it shows that $(C^*D)^{p-1}C^*$ is diagonal. Hence $(C^*D)^{p-1}C^*D = (C^*D)^p$ is diagonal, and therefore also C^*D as required. Furthermore, for any $j,k \in \{1,\dots,n\}$ such that $(C^*D)_{jj} \neq 0$ and $j \neq k$, we have $(C^*)_{jk} = 0$, since otherwise the diagonality of $(C^*D)^{p-1}C^*$ is contradicted. Similarly we have $(C^*)_{kj} = 0$, since otherwise the diagonality of C^*D is contradicted. This implies that $((C^*C)^{r/2})_{jk} = 0$ for all such j and k, and hence:

$$\sum_{j\in S} |C_{jj}|^r \;\leq\; \|C\|_r^r\,,$$

where

$$S \;=\; \{j \in \{1,\dots,n\} \,:\, (C^*D)_{jj} \neq 0\}\,.$$

So by the diagonality of (C^*D) we have:

$$\begin{aligned}
\|N^*M\|_p \;=\; \|C^*D\|_p &= \Big(\sum_{j\in S} |C_{jj}|^p |D_{jj}|^p\Big)^{1/p} \\
&\leq \left(\Big(\sum_{j\in S} |C_{jj}|^r\Big)^{p/r}\Big(\sum_{j\in S} |D_{jj}|^q\Big)^{p/q}\right)^{1/p} \\
&\leq \|C\|_r \|D\|_q \\
&= \|N\|_r\,.
\end{aligned}$$

This shows that the maximum (5.6) is at most $\|N\|_r$. To show that the value $\|N\|_r$ is actually attained, set $M = (f_{r/p}(N))^* / \|N\|_r^{r/q}$, where $f_{r/p}$ is defined as in equation (5.4) of the proof of Lemma 5.1.1. Then $(f_{r/p}(N))(f_{r/p}(N))^* = (N^*N)^{r/q}$ and $N^*(f_{r/p}(N))^* = (N^*N)^{(r/p)/2}$, and so $\|M\|_q = 1$ and $\|N^*M\|_p = \|N\|_r$, as required.

Now this implies that equality holds throughout, in the above chain of inequalities. Hence $(C^*)_{jk}$ is nonzero only if $j,k \in S$ with $j = k$. So we may deduce that the three matrices C, D and C^*D are positive and diagonal, and that they share the same range.

Case $q = \infty, r < \infty$. For any matrices $X, Y \in \mathcal{L}(\mathbb{C}^n)$, we may write $X = UDV$

for some matrices U, V unitary and D positive and diagonal, in $\mathcal{L}(\mathbb{C}^n)$. Hence

$$|\mathrm{Tr}\, Y^*X| \;=\; |\mathrm{Tr}\, D(VY^*U)| \;\leq\; \|Y\|_\infty(\mathrm{Tr} D) \;=\; \|X\|_1\,\|Y\|_\infty\,, \tag{5.7}$$

since the diagonal entries of VY^*U have modulus at most $\|Y\|_\infty$. Therefore in the case $r > 1$, it follows from relations (5.6) and (5.7) that:

$$\begin{aligned}
\|N^*M\|_r &= \max\{\|M^*NL\|_1 \,:\, \|L\|_{r/(r-1)} \leq 1\} \\
&= \max\{|\mathrm{Tr}\, M^*(NL)| \,:\, \|L\|_{r/(r-1)} \leq 1\} \\
&\leq \max\{\|NL\|_1\,\|M\|_\infty \,:\, \|L\|_{r/(r-1)} \leq 1\} \\
&= \|M\|_\infty\,\|N\|_r\,.
\end{aligned}$$

In the case $r = 1$, we have from inequality (5.7):

$$\begin{aligned}
\|N^*M\|_1 &= \max\{|\mathrm{Tr}\, N^*MU| \,:\, U \text{ unitary}\} \\
&= \max\{|\mathrm{Tr}\, M^*NU^*| \,:\, U \text{ unitary}\} \\
&\leq \|M\|_\infty\,\|N\|_1\,.
\end{aligned}$$

This completes the proof for all cases of p, q and r. □

Observe that from the above proof of the case $q = \infty, r < \infty$, we obtain the duality relation:

$$\|N\|_p \;=\; \max\{|\mathrm{Tr}\, N^*M| \,:\, \|M\|_{p/(1-p)} \leq 1\}\,,$$

for all $N \in \mathcal{L}(\mathbb{C}^n)$ and $1 \leq p < \infty$. Therefore from the linearity of Tr and the fact that N^*N is unitarily equivalent to NN^*, we conclude that $\|\cdot\|_p$ is a $*$-invariant norm on $\mathcal{L}(\mathbb{C}^n)$, for $1 \leq p \leq \infty$. (The case $p = \infty$ has already been established in the first part of the proof of Theorem 5.1.2).

We remark that, in addition to the fact that $\|\cdot\|_p$ is a $*$-invariant norm for all $p \in [1, \infty]$, one may also deduce the inequality:

$$|\mathrm{Tr} M| \;\leq \|M\|_1 \quad \text{for all } M \in \mathcal{L}(\mathbb{C}^n),$$

from the proof of Theorem 5.1.2. (It is just a special case of inequality (5.7)).

5.2 Smoothness of the predual of H^∞

In Chapter 4, we posed a constrained approximation problem on an arbitrary Banach space X in possession of a predual X_*. After proving the existence of a solution in Section 4.2, the setup was specialized by substituting a vector-valued function space $\mathrm{L}^\infty(X)$ for X, where we derived uniqueness and saturation results for the solution (constrained to lie in a vector-valued H^∞ space), under certain conditions. In particular, the uniqueness result depended on the assumption that the underlying space X is strictly convex. It was originally hoped that a similar strict convexity condition would guarantee uniqueness of the solution in the abstract setup of Section 4.2. However, neither of the spaces L^∞ or H^∞ are strictly convex, so it is unlikely that such a result would be of relevance to constrained H^∞ approximation. This suggests that we should try replacing the condition of strict convexity with the slightly weaker condition of smoothness of the predual, which motivates the result of this section.

We shall prove that the quotient space $\mathrm{L}^1/\,\mathrm{H}^1_0$ is smooth, where H^1_0 is the space of H^1 functions f such that $f(0) = 0$. This quotient space may be naturally identified with the predual of H^∞, since the annihilator of H^1_0 in L^∞ is precisely H^∞. That is, we have:

$$\begin{aligned}(\mathrm{L}^1/\,\mathrm{H}^1_0)^* \;\cong\; (\mathrm{H}^1_0)^\circ &= \left\{ f \in \mathrm{L}^\infty \;:\; \int_{\mathbb{T}} fg \,\mathrm{d}\mu = 0 \;\text{ for all } g \in \mathrm{H}^1_0 \right\} \\ &= \mathrm{H}^\infty .\end{aligned}$$

In order to establish this result, we shall begin with the following proposition, which shows that every function in L^1 has a best H^1_0 approximant. First of all, recall that $\mathrm{A} \subset \mathrm{C}$ is defined to be the algebra of those continuous functions on $\mathbb{T}$ which extend continuously to an analytic function on $\mathbb{D}$.

Proposition 5.2.1 *Let $\phi \in \mathrm{L}^1$ and let $\alpha : \mathrm{A} \to \mathbb{C}$ be the continuous linear functional given by:*

$$\alpha(f) = \int_{\mathbb{T}} f(z)\phi(z)\,\mathrm{d}\mu(z)$$

for all $f \in \mathrm{A}$. *Then we have:*

$$\inf\{\|\phi - g\|_1 \,:\, g \in \mathrm{H}^1_0\} \;=\; \|\alpha\|\,, \tag{5.8}$$

with the infimum attained by some $g_0 \in \mathrm{H}^1_0$.

Proof For any $g \in \mathrm{H}^1_0$, we have:

$$\begin{aligned}\|\phi - g\|_1 &\geq \sup\left\{\left|\int_{\mathbb{T}} f(z)(\phi(z) - g(z))\mathrm{d}\mu(z)\right| \;:\; f \in \mathrm{A} \text{ with } \|f\|_\infty \leq 1\right\} \\ &= \sup\left\{\left|\int_{\mathbb{T}} f(z)\phi(z)\,\mathrm{d}\mu(z)\right| \;:\; f \in \mathrm{A} \text{ with } \|f\|_\infty \leq 1\right\} \\ &= \|\alpha\|\,,\end{aligned}$$

since A lies in the annihilator of H^1_0. Hence the infimum (5.8) is at least $\|\alpha\|$. It remains to show that the infimum is attained.

By the Hahn-Banach theorem, α has a continuous linear extension β to the whole of C, such that $\|\beta\| = \|\alpha\|$. Now by the Riesz representation theorem, there is a unique complex Borel measure ν on $\mathbb{T}$ such that:

$$\beta(f) \;=\; \int_{\mathbb{T}} f\,\mathrm{d}\nu\,,$$

for all $f \in C$. Moreover, we have:

$$\|\nu\| \;=\; |\nu|(\mathbb{T}) \;=\; \|\beta\| \;=\; \|\alpha\|\,.$$

Now observe that:

$$\int_{\mathbb{T}} f\,(\phi\,\mathrm{d}\mu - \mathrm{d}\nu) \;=\; \alpha(f) - \beta(f) \;=\; 0\,,$$

for all $f \in \mathrm{A}$. In particular, for all integers $n \leq 0$, the Fourier coefficients

$$(\phi\,\mathrm{d}\mu - \mathrm{d}\nu)\hat{}\,(n) \;=\; \int_{\mathbb{T}} z^{-n}\,(\phi\,\mathrm{d}\mu - \mathrm{d}\nu)(z)$$

are all zero. Thus the measure $\phi\,\mathrm{d}\mu - \mathrm{d}\nu$ has an analytic extension to $\mathbb{D}$, vanishing at 0. Now the Riesz Brothers' theorem implies that there is a (unique) $g_0 \in \mathrm{H}^1_0$ such that:

$$\phi\,\mathrm{d}\mu - \mathrm{d}\nu \;=\; g_0\,\mathrm{d}\mu\,.$$

Hence $d\nu = (\phi - g_0)d\mu$ and furthermore, $\|\phi - g_0\|_1 = \|\nu\| = \|\alpha\|$, so the infimum (5.8) is attained by g_0, completing the proof. □

We are now ready to prove that L^∞ / H_0^1 is smooth, using the existence of a best H_0^1 approximant g to any given $\phi \in L^1$. This is done by establishing a relationship between g and the unique support functional of $[\phi]$ on the quotient space L^∞ / H_0^1. In addition to this, the following result shows that the best H_0^1 approximant to ϕ is unique.

Theorem 5.2.2 *Let $\phi \in L^1 \setminus H_0^1$. Then there is a function $f \in H^\infty$ of unit norm and a function $g \in H_0^1$, which are both unique up to equality almost everywhere, such that:*

$$\int_{\mathbb{T}} f(z)\phi(z)\,d\mu(z) \;=\; \mathrm{dist}(\phi, H_0^1) \;=\; \|\phi - g\|_1 . \qquad (5.9)$$

Thus every L^1 function has a unique best H_0^1 approximant, and the quotient space L^1 / H_0^1 is smooth. Furthermore, upon letting $\psi = \phi - g$, we find that $f(z) = \overline{\psi(z)} / |\psi(z)|$ for almost all $z \in \mathbb{T}$ such that $\psi(z)$ is nonzero.

Proof By Proposition 5.2.1, ϕ has a best H_0^1 approximant g. Now let $\psi = \phi - g$ and let $f \in H^\infty$ of unit norm be chosen to satisfy equation (5.9). Such a function f is guaranteed to exist by the Hahn-Banach theorem, applied to the quotient space L^1 / H_0^1. Now we have:

$$\int_{\mathbb{T}} f(z)\psi(z)\,d\mu(z) \;=\; \int_{\mathbb{T}} f(z)\phi(z)\,d\mu(z) \;=\; \|\psi\|_1 , \qquad (5.10)$$

since f annihilates g. But $\|f\|_\infty = 1$, so we may deduce that $f(z) = \overline{\psi(z)} / |\psi(z)|$ for almost all $z \in \mathbb{T}$ such that $\psi(z)$ is nonzero, as claimed. Indeed, $|f(z)| \le 1$ for almost all $z \in \mathbb{T}$, and by equation (5.10) we have:

$$\int_{\mathbb{T}} \mathrm{Re}(|\psi(z)| - f(z)\psi(z))\,d\mu(z) \;=\; 0 .$$

But the above integrand is a nonnegative real for almost all $z \in \mathbb{T}$, so it must be zero for almost all $z \in \mathbb{T}$. Therefore $f(z)\psi(z) = |\psi(z)|$ for almost all $z \in \mathbb{T}$, as required.

Now, to show that f is unique up to equality almost everywhere, suppose that $f' \in \mathrm{H}^\infty$ is another function of unit norm, satisfying (5.9) in place of f. Then since $\phi \notin \mathrm{H}_0^1$, there is a subset $K \subset \mathbb{T}$ of nonzero measure such that $\psi(z)$ is nonzero for all $z \in K$. But $f(z) = f'(z) = \overline{\psi(z)}/\,|\psi(z)|$ for almost all $z \in K$, so f and f' agree on a set of nonzero measure. Hence $f = f'$ almost everywhere, as required.

Finally, to show that g is unique up to equality almost everywhere, suppose that g' is another best H_0^1 approximant to ϕ. Then upon setting the new error $\psi' = \phi - g'$, we have:

$$\begin{aligned} f(z)\psi(z) &= |\psi(z)| \\ f(z)\psi'(z) &= |\psi'(z)| \end{aligned}$$

for almost all $z \in \mathbb{T}$. But $\psi - \psi' = g' - g \in \mathrm{H}^\infty$, so the mapping:

$$z \;\mapsto\; f(z)(\psi(z) - \psi'(z)) \;=\; |\psi(z)| - |\psi'(z)| \quad \text{for almost all } z \in \mathbb{T},$$

is both real and H^∞ and therefore a constant, say $\lambda \in \mathbb{R}$, almost everywhere. Now integrating the right hand side over $\mathbb{T}$, we obtain $\|\psi\|_1 - \|\psi'\|_1 = \lambda$. But $\|\psi\|_1 = \|\psi'\|_1 = \mathrm{dist}(\phi, \mathrm{H}_0^1)$ and so $\lambda = 0$. Hence $f(g' - g) = 0$ almost everywhere. So $g = g'$ almost everywhere, as required, since $0 \neq f \in \mathrm{H}^\infty$. □

As a corollary to this uniqueness theorem, we obtain the following decomposition result for L^1:

Corollary 5.2.3 *Let $\phi \in \mathrm{L}^1 \setminus \mathrm{H}_0^1$. Then ϕ has a unique (up to equality almost everywhere) decomposition of the form:*

$$\phi \;=\; g + w\overline{f}\,, \tag{5.11}$$

for some $g \in \mathrm{H}_0^1$, $f \in \mathrm{H}^\infty$ with $\|f\|_\infty = 1$, and $w \in \mathrm{L}^1$ a.e. nonnegative real, with $w(z) = 0$ for almost all $z \in \mathbb{T}$ such that $|f(z)| < 1$. Furthermore, f and g satisfy equation (5.9) of Theorem 5.2.2.

Proof Theorem 5.2.2 guarantees that such a decomposition of ϕ exists. Indeed, upon choosing f and g to satisfy equation (5.9), the theorem implies that $f(z) = \overline{\psi(z)}/\,|\psi(z)|$ for almost all $z \in \mathbb{T}$ such that $\psi(z)$ is nonzero, where $\psi = \phi - g$. Therefore,

$$|\psi(z)|\,\overline{f(z)} \;=\; \psi(z)\,,$$

for almost all $z \in \mathbb{T}$. So there exists $w \in \mathrm{L}^1$ a.e. nonnegative real such that $w\overline{f} = \psi = \phi - g$ and $w(z) = 0$ for almost all $z \in \mathbb{T}$ such that $|f(z)| < 1$, as required.

It remains to show that any decomposition of the form (5.11) gives rise to the choice of f and g which satisfy equation (5.9). Suppose we have such a decomposition. Then since $|f(z)| \leq 1$ for almost all $z \in \mathbb{T}$ and f annihilates H^1_0, we have:

$$\mathrm{Re}\int_{\mathbb{T}} f(z)\phi(z)\,\mathrm{d}\mu(z) \;\leq\; \mathrm{dist}(\phi, \mathrm{H}^1_0) \;\leq\; \|\phi - g\|_1\,. \tag{5.12}$$

But $f\phi \;=\; fg + fw\overline{f} \;=\; fg + w$ and by integrating this we obtain:

$$\int_{\mathbb{T}} f(z)\phi(z)\,\mathrm{d}\mu(z) \;=\; \|w\|_1 \;=\; \|\phi - g\|_1\,. \tag{5.13}$$

Therefore we have equality throughout (5.12), and since equation (5.13) implies that the left hand integral is real, we obtain equation (5.9), as required. $\square$

The $\mathrm{L}^1/\,\mathrm{H}^1_0$ smoothness result appears to be new. However, I was ultimately unable to use it in the constrained H^∞ approximation theory of Chapter 4, as was originally hoped. Instead, the more concrete approach of Section 4.4, to the question of solution uniqueness, was deemed to be more useful.

Bibliography

[1] ANDERSON B. D. O., *Continuity of the spectral factorization operation*, Math. Appl. Comput., 4 (1985), 139–156

[2] BARATCHART L., LEBLOND J. AND PARTINGTON J. R., *Hardy approximation to* L^∞ *functions on subsets of the circle*, Constr. Approx., 12 (1996), 423–435.

[3] BARCLAY S., *Continuity of the spectral factorization mapping*, J. London Math. Soc., 763–779 (70), No. 2, 2004.

[4] BIRMAN M. S. AND SOLOMJAK M. Z., *Spectral theory of self-adjoint operators in Hilbert space*, D. Reidel Publishing Company, Dordrecht, Holland, 1987.

[5] BONSALL F. F., *A general atomic decomposition theorem and Banach's closed range theorem*, Quart. J. Math. Oxford, 42, no. 2 (1991), 9–14.

[6] BOURGAIN J., *A problem of Douglas and Rudin on factorization*, Pacific J. Math., 47–50 (121), No. 1, 1986.

[7] CHAABANE S., JAOUA M. AND LEBLOND J., *Parameter identification for Laplace equation and approximation in Hardy classes*, J. Inverse Ill-Posed Probl., 11 (2003), 33–57.

[8] CHALENDAR I. AND ESTERLE J., L^1*-factorization for* C_{00}*-contractions with isometric functional calculus*, J. Func. Anal., 174–194 (154), 1998.

[9] Chalendar I., Leblond J. and Partington J. R., *Approximation problems in some holomorphic spaces, with applications*, Systems, approximation, singular integral operators, and related topics (Bordeaux, 2000), 143–168, Oper. Theory Adv. Appl., 129, Birkhäuser, Basel, 2001.

[10] Chalendar I., Partington J. R. and Smith M., *Approximation in reflexive Banach spaces and applications to the invariant subspace problem*, Proc. Amer. Math. Soc., 132, no. 4 (2003), 1133–1142.

[11] Curtain R. F. and Zwart H., *An introduction to infinite-dimensional linear systems theory*, Texts in applied mathematics, vol. 21, Springer-Verlag, New York, 1995.

[12] Diestel J. and Uhl J. J., *Vector measures*, Mathematical surveys, no. 15, American Mathematical Society, 1977.

[13] Donoghue W. F., *Monotone matrix functions and analytic continuation*, Springer-Verlag, New York, Heidelberg, Berlin, 1974.

[14] Douglas R. G. and Rudin W., *Approximation by inner functions*, Pacific J. Math., 313–320 (31), No. 2, 1969.

[15] Dunford N. and Schwartz J. T., *Linear operators*, Part 1 : general theory, Wiley classics library edition, Wiley-Interscience, New York, 1988.

[16] Dunford N. and Schwartz J. T., *Linear operators*, Part 2 : spectral theory, Pure and applied mathematics, vol. 7, Interscience, New York, 1963.

[17] Garnett J. B., *Bounded Analytic Functions*, Academic Press, New York, 1981.

[18] Helson H. and Lowdenslager D., *Prediction theory and Fourier series in several variables*, Acta Math., 99 (1958), 165–202.

[19] Hoffman K., *Banach spaces of analytic functions*, Prentice-Hall, Englewood Cliffs, NJ, 1962.

[20] Jacob B. and Partington J. R., *On the boundedness and continuity of the spectral factorization mapping*, SIAM J. Control Optim., 40 (2001), 88–106.

[21] Jacob B., Winkin J. and Zwart H., *Continuity of the spectral factorization on a vertical strip*, Systems Control Lett., 37 (1999), 183–192.

[22] Leblond J. and Olivi M., *Weighted* H^2 *approximation of transfer functions*, Math. Control Signals Systems, 11 (1998), 28–39.

[23] Murphy G. J., *C*-algebras and operator theory*, Academic Press, New York, 1990.

[24] Nikolski N. K., *Operators, functions and systems: an easy reading*, Vol. 1, Mathematical surverys and monographs, vol. 92, American Mathematical Society, 2002.

[25] Pedersen G. K., *Some operator monotone functions*, Proc. Amer. Math. Soc., 36 (1972), 309–310.

[26] Peller V. V., *Hankel operators and their applications*, Springer monographs in mathematics, Springer-Verlag, New York, 2003.

[27] Rana I. K., *An introduction to measure and integration*, 2nd ed., Graduate studies in mathematics, vol. 45, American Mathematical Society, 2002.

[28] Rosenblum M. and Rovnyak J., *Hardy classes and operator theory*, Oxford mathematical monographs, Oxford University Press, 1985.

[29] Slepian, D., *On bandwidth*, Proc. IEEE, 64 (1976), 292–300.