Trends in Mathematics

Alexandru Aleman
Haakan Hedenmalm
Dmitry Khavinson
Mihai Putinar
Editors

Analysis of Operators on Function Spaces

The Serguei Shimorin Memorial Volume

Birkhäuser

Birkhäuser

Trends in Mathematics

Trends in Mathematics is a series devoted to the publication of volumes arising from conferences and lecture series focusing on a particular topic from any area of mathematics. Its aim is to make current developments available to the community as rapidly as possible without compromise to quality and to archive these for reference.

Proposals for volumes can be submitted using the Online Book Project Submission Form at our website www.birkhauser-science.com.

Material submitted for publication must be screened and prepared as follows:

All contributions should undergo a reviewing process similar to that carried out by journals and be checked for correct use of language which, as a rule, is English. Articles without proofs, or which do not contain any significantly new results, should be rejected. High quality survey papers, however, are welcome.

We expect the organizers to deliver manuscripts in a form that is essentially ready for direct reproduction. Any version of TEX is acceptable, but the entire collection of files must be in one particular dialect of TEX and unified according to simple instructions available from Birkhäuser.

Furthermore, in order to guarantee the timely appearance of the proceedings it is essential that the final version of the entire material be submitted no later than one year after the conference.

More information about this series at http://www.springer.com/series/4961

Alexandru Aleman • Haakan Hedenmalm •
Dmitry Khavinson • Mihai Putinar
Editors

Analysis of Operators on Function Spaces

The Serguei Shimorin Memorial Volume

Birkhäuser

Editors
Alexandru Aleman
Mathematics Department
Lund University
Lund, Sweden

Haakan Hedenmalm
Department of Mathematics
The Royal Institute of Technology
Stockholm, Sweden

Dmitry Khavinson
Department of Mathematics
University of South California
Tampa, FL, USA

Mihai Putinar
Mathematics Department
University of California
Santa Barbara, CA, USA

ISSN 2297-0215 ISSN 2297-024X (electronic)
Trends in Mathematics
ISBN 978-3-030-14639-9 ISBN 978-3-030-14640-5 (eBook)
https://doi.org/10.1007/978-3-030-14640-5

Mathematics Subject Classification (2010): 30-XX, 32-XX, 46-XX, 47-XX, 12-XX, 31-XX

This book is published under the imprint Birkhäuser, www.birkhauser-science.com by the registered company Springer Nature Switzerland AG.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This book is dedicated to Serguei Shimorin, eminent mathematician and virtuoso lensman.

Serguei's mathematics was forged at one of the preeminent hubs of mathematical analysis of the twentieth century: St. Petersburg University and the Steklov Institute of Mathematics at St. Petersburg. He carried over countries and generations the precious legacy of Havin and Nikolski's legendary seminar. The choices of research topics he made, well represented by the articles contained in the present volume, are the product of years of accumulated skills and quiet introspection. Serguei's vast mathematical culture, especially in Function Theory and Functional Analysis, is held in high esteem by his colleagues and collaborators.

Shimorin was a man of few words, written or spoken. He sought beauty in simplicity, uncontaminated by the tumult of modern life. For those close to him, including probably his many undergraduate students at the Royal Institute of Technology at Stockholm, he emanated pure light—a lasting, invaluable heritage.

Serguei's mathematical gift was complemented by artistic dispositions. He was a devout classical music concertgoer (especially on the lavish St. Petersburg scene) and an accomplished pianist. His religious admiration for nature took many forms and in particular he distilled the serenity and mystery of landscape into photos. They speak by themselves about the man we celebrate and mourn in this volume:

http://www.photosight.ru/users/119352/

Some reminiscences from the editors:

Alexandru Aleman: Serguei was a dear friend. I admired his mathematical work before knowing him and am truly grateful for the beautiful subtle ideas he exchanged with me.

Dmitry Khavinson: For me personally he has always represented (by now almost non-existent) legendary Russian, St. Petersburg 'intelligentsia', always firmly guided in life by formidable, non-bending moral humanistic principles.

Mihai Putinar: On the occasion of his visit to Singapore, where we had ample professional, social and cultural activities, he was obstinate in declining our company to discover the city. He wanted to experience alone the wonders of the luxuriant vegetation, unique architecture and human variety. For Serguei the freshness of discovery was primordial.

Serguei is survived by his wife Olga Muzhdaba (geographer), daughter Anastasia Shimorina (linguist) and son Mikhail Shimorin (musician).

We thank Dorothy Mazlum and the team at Birkhäuser Verlag for making possible this book project.

Lund, Sweden — Alexandru Aleman
Stockholm, Sweden — Haakan Hedenmalm
Tampa, FL, USA — Dmitry Khavinson
Santa Barbara, CA, USA — Mihai Putinar

Serguei Shimorin (1965–2016)

Contents

My Recollections of Serguei Shimorin

Haakan Hedenmalm

Abstract The author shares his personal reminiscences of Serguei Shimorin.

Keyword Shimorin theorem

MSC Codes: 01A70

Serguei Shimorin died July 18, 2016, as the result of a hiking accident in the mountains of Abkhazia, a disputed part of Georgia in the Caucasus. He was hiking with two friends, Andrei and Roman (who were brothers), and he was about the to pioneer the crossing of the creek Dzhampal. Unfortunately the attempt ended really tragically. He was a young man, born in 1965 in Leningrad. His death is a major loss for Mathematics and Swedish Mathematics in particular. I would now like to share some of my recollections concerning Serguei and his scientific achievements in Mathematics, at Lund University and later at KTH.

In the fall semester of 1990 I visited Leningrad through an academic exchange program, involving KVA (Royal Swedish Academy of Sciences) and the Academy of Sciences of the USSR. I obtained a nominal stipend in rapidly devaluing roubles, but, more importantly, I was supplied with a free hotel room during my visit. While in Leningrad, I encountered several prominent participants of the Analysis Seminar at LOMI (nowadays POMI), the Steklov Institute located at Fontanka 27. Among these were Nikolai Nikolski, Nikolai Makarov, Vladimir Peller, Alexei Alexandrov, to mention a few. This was a difficult time in the USSR, and the country fell apart just a year later. But there was nothing wrong with the hospitality, and I remember being invited home to both Peller and Nikolski. At around this time the factorization methods involving extremal functions in Bergman spaces had already been developed by myself as well as others (I was influenced by Boris Korenblum from SUNY Albany) and I made a couple of presentations on this topic during the fall semester. I did not notice it then, but later I understood that among the attentive

H. Hedenmalm (✉)
The Royal Institute of Technology, Department of Mathematics, Stockholm, Sweden
e-mail: haakanh@math.kth.se; https://people.kth.se/~haakanh/engindex.html

A. Aleman et al. (eds.), *Analysis of Operators on Function Spaces*,
Trends in Mathematics, https://doi.org/10.1007/978-3-030-14640-5_1

listeners was a shy young man by the name of Serguei Shimorin. He was studying with Stanislav A. Vinogradov, who himself had studied with Victor P. Havin in 1968 in Leningrad. Apparently my presentations made an impression on Serguei, as a bit later he sent me a preprint entitled "Factorization of analytic functions in weighted Bergman spaces" which subsequently appeared in *Algebra i Analiz* and in English translation in *St. Petersburg Math. J.* in 1994. This work (presumably a part of Serguei's 1993 thesis) was highly original, especially as he invented a kind of pseudodifferential operators Δ_α such that the analog of Green's formula

$$\int_{\mathbb{D}} (h_2 \Delta_\alpha h_1 - h_1 \Delta_\alpha h_2) dA_\alpha = \int_{\partial\mathbb{D}} (h_2 \partial_n h_1 - h_1 \partial_n h_2) ds$$

would hold for an interval in α, where $dA_\alpha(z) = (1 - |z|^2)^\alpha dA(z)$ is weighted area measure. Later on he developed the theory of these Δ_α further and computed a related weighted biharmonic Green function and obtained its positivity. The issue whether biharmonic Green functions are positive is a delicate matter going back to work of Boggio and Hadamard from around 1900. Another fascinating contribution is "Single-point extremal functions in weighted Bergman spaces", where Serguei developed a *new idea* connecting between univalence and divisor properties of one-point divisors (analogs of Blaschke factors).

At some point around 1994–1995 I was invited to the home of Serguei and his family (wife Olga and children Anastasia and Mikhail) and I recall that I asked whether he had taken part in any kind of Math Olympiad. That kind of almost sporty activity was encouraged in the USSR and apparently Serguei had done well in some local such olympiad, and showed me a diploma. He also mentioned assuming programming work right after the 1987 diploma before entering the PhD program.

In 1996 I was taking part in a conference in Trondheim organized by Kristian Seip. There I spoke with Alexander Borichev, with whom I had collaborated successfully when he was a "forskarassistent" in Uppsala in the 1990s. He had then left Sweden for France and now that I had moved to Lund from Uppsala, I suggested he might come to Lund. He declined, but suggested I could be interested in Serguei whom he found an excellent mathematician. With this strong recommendation Shimorin was hired as "forskarassistent" in Lund around 1998 with an NFR grant, after spending a year at Université de Bordeaux. The lectureship which I mentioned to attract Borichev went instead to Alexandru Aleman.

I got Serguei interested in the project to show that the biharmonic Green function was positive for a general weight which was reproducing for a point and also logarithmically subharmonic. This was conjectured, but proved difficult to obtain. In the end we succeeded, and Serguei had fundamental insight toward the solution. He derived a property of the corresponding Bergman kernel which together with a twice applied Hele–Shaw flow led to the conjectured positivity. Also in the work on Hele–Shaw flow on hyperbolic surfaces Serguei supplied key insight. He was always meticulously careful and sought elegant arguments whenever possible. As a spin-off he produced the impressive paper "Wold-type decompositions and wandering subspaces for operators close to isometries" published by *Crelle* in 2001. Another work, “Approximate spectral synthesis in the Bergman space” (*Duke Math. J.*, 2000) appeared in this productive period. After 2002, Serguei and I moved to KTH and we continued collaborating on what is known as “Brennan’s conjecture”. Serguei had an initial insight developed first in *IMRN* in 2003, and later jointly in *Duke Math. J.* in 2005. In 2004, Serguei was honored with the prestigeous Wallenberg prize of the Swedish Mathematical Society (shared with Julius Borcea).

Serguei was very interested in problems of significance in Operator Theory, such as related to “Commutant lifting” and “Complete Nevanlinna–Pick kernels”. As I recall, I heard from US colleagues that one of Serguei’s works was presented at a seminar in Berkeley, and that supposedly at the end, when the gist of the argument was put forth, Donald Sarason exclaimed “That was smart!”

As a scientist, Serguei was original with technical ability. But as a person he was very private and rather shy and humble. He was not career oriented, but rather an “artist within mathematics”, who from time to time would find a beautiful flower and wanted to show it to the world. In our Swedish university system such an individual does not get the appropriate appreciation, I believe: we tend to have a hierarchic and career oriented perspective, and you need to grab all the opportunities that come your way. Serguei invested a lot of effort in his work assignments, especially his lectures, and the students appreciated him very much. His scientific talent exceeded that of several full professors, but his shyness made him less visible. It is my opinion that we should give more room for original individuals like Serguei.

In the later years Serguei took an interest in photography (see photosight.ru, under the pseudonym “Serge de la Mer”) (Figs. 1, 2 and 3).

Fig. 1 Le soleil levant

Fig. 2 Summer evening on Khadat

Fig. 3 Morning at artists' lake

Localization of Zeros in Cauchy–de Branges Spaces

Evgeny Abakumov, Anton Baranov, and Yurii Belov

Dedicated to the memory of Serguei Shimorin, a brilliant mathematician and a wonderful person.

Abstract We study the class of discrete measures in the complex plane with the following property: up to a finite number, all zeros of any Cauchy transform of the measure (with ℓ^2-data) are localized near the support of the measure. We find several equivalent forms of this property and prove that the parts of the support attracting zeros of Cauchy transforms are ordered by inclusion modulo finite sets.

Keywords Cauchy transforms · de Branges spaces · Distribution of zeros of entire functions · Polynomial approximation

1991 Mathematics Subject Classification 30D10, 30D15, 46E22, 41A30, 34B20

The work was supported by the joint grant of Russian Foundation for Basic Research (project 17-51-150005-NCNI-a) and CNRS, France (project PRC CNRS/RFBR 2017-2019 “Noyaux reproduisants dans des espaces de Hilbert de fonctions analytiques”).

E. Abakumov (✉)
University Paris-Est, LAMA (UMR 8050), UPEM, UPEC, CNRS, Marne-la-Vallée, France
e-mail: evgueni.abakoumov@u-pem.fr

A. Baranov
Department of Mathematics and Mechanics, St. Petersburg State University, St. Petersburg, Russia

National Research University Higher School of Economics, St. Petersburg, Russia

Y. Belov
St. Petersburg State University, St. Petersburg, Russia

A. Aleman et al. (eds.), *Analysis of Operators on Function Spaces*,
Trends in Mathematics, https://doi.org/10.1007/978-3-030-14640-5_2

1 Introduction and Main Results

Representations of analytic functions via Cauchy transforms of planar measures is a classical theme in function theory. Of special interest are expansions of meromorphic functions as Cauchy transforms of discrete (atomic) measures. A substantial number of papers deal with the distribution of zeros of such Cauchy transforms. Note that zeros of Cauchy transforms are equilibrium points of logarithmic potentials for the corresponding discrete measures.

Let $T = \{t_n\}_{n\in\mathbb{N}}$ be a sequence of distinct complex numbers with $|t_n| \to \infty$, $n \to \infty$. Then, for any sequence $a = \{a_n\}$ such that $\sum_n |t_n|^{-1}|a_n| < \infty$ one can consider the Cauchy transform

$$\mathcal{C}_a(z) = \sum_{n=1}^{\infty} \frac{a_n}{z - t_n}.$$

Clunie et al. [11] conjectured that the Cauchy transform $\mathcal{C}_a$ has infinitely many zeros if all a_n are positive. In the general case this conjecture remains open, for related results see [11, 13, 16]. Clearly, if the coefficients a_n are not positive, the corresponding sum can be the inverse to an entire function and, thus, can have no zeros.

In [1] we studied some phenomena connected with the following heuristic principle: *if the coefficients a_n are extremely small, then all* (*except a finite number of*) *zeros of $\mathcal{C}_a$ are located near the set T or near its part*. We called this *localization property*. In [1] only the case $T \subset \mathbb{R}$ was considered. This enabled us to relate the problem with the theory of de Branges spaces and the structure of Hamiltonians for canonical systems. In the present paper we consider the case of general (complex) t_n and extend many of the results from [1] to this setting. One of the motivations for this study is the role of discrete Cauchy transforms in the functional model for rank one perturbations of compact normal operators (see [4, 6]).

1.1 The Spaces of Cauchy Transforms and Related Spaces of Entire Functions

Let $\mu := \sum_n \mu_n \delta_{t_n}$ be a positive measure on $\mathbb{C}$ such that $\sum_n \frac{\mu_n}{|t_n|^2+1} < \infty$. With any such μ we associate the Hilbert space $\mathcal{H}(T, \mu)$ of the Cauchy transforms

$$\mathcal{H}(T, \mu) := \left\{ f : f(z) = \sum_n \frac{a_n \mu_n^{1/2}}{z - t_n}, \quad a = \{a_n\} \in \ell^2 \right\}$$

equipped with the norm $\|f\|_{\mathcal{H}(T,\mu)} := \|a\|_{\ell^2}$. Note that the series in the definition of $\mathcal{H}(T, \mu)$ converges absolutely and uniformly on compact sets separated from T.

The spaces $\mathcal{H}(T, \mu)$ consist of meromorphic functions which are analytic in $\mathbb{C} \setminus T$. To get rid of the poles, we will usually consider isometrically isomorphic Hilbert spaces of entire functions. Let A be an entire function which has only simple zeros and whose zero set $\mathcal{Z}_A$ coincides with T. With any T, A and μ as above we associate the space $\mathcal{H}(T, A, \mu)$ of entire functions,

$$\mathcal{H}(T, A, \mu) := \left\{ F : F(z) = A(z) \sum_n \frac{a_n \mu_n^{1/2}}{z - t_n}, \quad a = \{a_n\} \in \ell^2 \right\},$$

where again the norm is given by $\|F\|_{\mathcal{H}(T,A,\mu)} := \|a\|_{\ell^2}$. Clearly, the mapping $f \mapsto Af$ is a unitary operator from $\mathcal{H}(T, \mu)$ to $\mathcal{H}(T, A, \mu)$. We will use the term *Cauchy–de Branges spaces* for the spaces $\mathcal{H}(T, A, \mu)$. This is related to the fact that the class of the spaces $\mathcal{H}(T, A, \mu)$ with $T \subset \mathbb{R}$ coincides with the class of all de Branges spaces (see [12]).

The spaces $\mathcal{H}(T, A, \mu)$ were introduced in full generality by Belov et al. [8]. They can also be described axiomatically. It is clear that the reproducing kernels of $\mathcal{H}(T, A, \mu)$ at the points t_n (which are of the form $\overline{A'(t_n)\mu_n} \cdot \frac{A(z)}{z-t_n}$) form an orthogonal basis in $\mathcal{H}(T, A, \mu)$. Conversely, if $\mathcal{H}$ is a reproducing kernel Hilbert space of entire functions such that

1. $\mathcal{H}$ has the *division property*, that is, $\frac{f(z)}{z-w} \in \mathcal{H}$ whenever $f \in \mathcal{H}$ and $f(w) = 0$,
2. there exists an orthogonal basis of reproducing kernels in $\mathcal{H}$,

then $\mathcal{H} = \mathcal{H}(T, A, \mu)$ for some choice of the parameters T, A and μ. One can replace existence of an orthogonal basis by existence of a Riesz basis of normalized reproducing kernels. In this case $\mathcal{H}$ coincides with some space $\mathcal{H}(T, A, \mu)$ as sets with equivalence of norms.

1.2 Localization and Strong Localization

To simplify certain formulas, we will always assume in what follows that

$$|t_n| \geq 2, \qquad t_n \in T.$$

Also we will always assume that T is *a power separated sequence*: there exist numbers $C > 0$ and $N > 0$ such that, for any n,

$$\operatorname{dist}(t_n, \{t_m\}_{m \neq n}) \geq C|t_n|^{-N}. \tag{1.1}$$

Note that condition (1.1) implies that for some $c, \rho > 0$ and for sufficiently large n we have $|t_n| \geq cn^\rho$. We always will choose A to be an entire function of finite order with zeros at T. Without loss of generality we may always fix N in (1.1) so large that $\sum_n |t_n|^{-N} < \infty$.

For an entire function f we denote by $\mathcal{Z}_f$ the set of all zeros of f. Let $D(z, r)$ stand for the open disc centered at z of radius r.

Now we introduce the notion of *zeros localization*.

Definition 1.1 We say that the space $\mathcal{H}(T, A, \mu)$ with a power separated sequence T has the localization property if there exists a sequence of disjoint disks $\{D(t_n, r_n)\}$ with $r_n \to 0$ such that for any nonzero $f \in \mathcal{H}(T, A, \mu)$ the set $\mathcal{Z}_f \setminus \cup_n D(t_n, r_n)$ is finite and each disk $D(t_n, r_n)$ contains at most one point of $\mathcal{Z}_f$ for any n except, possibly, a finite number.

Since the space $\mathcal{H}(T, A, \mu)$ has the division property, one can construct a function from $\mathcal{H}(T, A, \mu)$ with zeros at any given finite set. Therefore the notion of localization of the zeros near T makes sense only up to finite-dimensional sets.

Our first result shows that the localization property in $\mathcal{H}(T, A, \mu)$ can be expressed in several natural ways. For a set E, we denote by $\#E$ the number of elements in E.

Theorem 1.1 *Let $\mathcal{H}(T, A, \mu)$ be a Cauchy–de Branges space with a power separated T. The following statements are equivalent:*

(i) $\mathcal{H}(T, A, \mu)$ *has the localization property*;
(ii) *There exists an unbounded set $S \subset \mathbb{C}$ such that the set $\mathcal{Z}_f \cap S$ is finite for any nonzero $f \in \mathcal{H}(T, A, \mu)$*;
(iii) *For any $f \in \mathcal{H}(T, A, \mu) \setminus \{0\}$ and $M > 0$ we have $\#\big(\mathcal{Z}_f \setminus \cup_n D(t_n, |t_n|^{-M})\big) < \infty$*;
(iv) *There is no nonzero $f \in \mathcal{H}(T, A, \mu)$ with infinite number of multiple zeros.*

Similarly to [1] one can introduce the notion of *strong localization* where the zeros are localized only near the whole set T. We say that the space $\mathcal{H}(T, A, \mu)$ with a power separated sequence T has the *strong localization property* if there exists a sequence of disjoint disks $\{D(t_n, r_n)\}_{t_n \in T}$ with $r_n \to 0$ such that for any nonzero $f \in \mathcal{H}(T, A, \mu)$ the set $\mathcal{Z}_f \setminus \cup_n D(t_n, r_n)$ is finite and each disk $D(t_n, r_n)$ contains exactly one point of $\mathcal{Z}_f$ for any n except, possibly, a finite number.

As in [1], one can show that the strong localization property is equivalent to the approximation by polynomials.

Theorem 1.2 *The space $\mathcal{H}(T, A, \mu)$ has the strong localization property if and only if the polynomials belong to $L^2(\mu)$ and are dense there.*

Note that polynomials belong to $L^2(\mu)$ whenever $\mathcal{H}(T, A, \mu)$ has the localization property (see Proposition 3.1). Density of polynomials in weighted L^p spaces is a classical problem in analysis which was studied extensively (see, e.g., [3, 9, 10, 15, 17]). All these works treat the case when the measure in question is supported by the real line. For measures in $\mathbb{C}$, the problem seems to be largely open.

1.3 Attraction Sets

Let $\mathcal{H}(T, A, \mu)$ have the localization property. By the property (iii) from Theorem 1.1, with any nonzero function $f \in \mathcal{H}(T, A, \mu)$ we may associate a set $T_f \subset T$ such that for some disjoint disks $D(t_n, r_n)$ all zeros of f except, may be, a finite number are contained in $\cup_{t_n \in T} D(t_n, r_n)$ and there exists exactly one point of $\mathcal{Z}_f$ in each disk $D(t_n, r_n)$, $t_n \in T_f$, except, may be, a finite number of indices n. Thus, the set T_f is uniquely defined by f up to finite sets. Let us also note that we can always take $r_n = |t_n|^{-M}$ for any $M > 0$.

Definition 1.2 Let $\mathcal{H}(T, A, \mu)$ have the localization property. We will say that $S \subset T$ is an attraction set if there exists $f \in \mathcal{H}(T, A, \mu)$ such that $T_f = S$ up to a finite set.

Note that $f(z) = \frac{A(z)}{z-t_0} \in \mathcal{H}(T, A, \mu)$ for any $t_0 \in T$, and so T is always an attraction set.

It turns out that the localization property implies the following ordering theorem for the attraction sets of $\mathcal{H}(T, A, \mu)$.

Theorem 1.3 *Let $\mathcal{H}(T, A, \mu)$ be a Cauchy–de Branges space with the localization property. Then for any two attraction sets S_1, S_2 either $S_1 \subset S_2$ or $S_2 \subset S_1$ up to finite sets.*

This ordering rule has some analogy with the de Branges Ordering Theorem for the chains of de Branges subspaces. For the case of de Branges spaces (i.e., $T \subset \mathbb{R}$) the ordering structure of attraction sets was proved in [1, Theorem 1.8]. Two different proofs were given: one of them used the de Branges Ordering Theorem, while the other used only a variant of Phragmén–Lindelöf principle due to de Branges [12, Lemma 7], a deep result which is one of the main steps for the Ordering Theorem.

These methods are no longer available in the case of nonreal t_n. However, it turned out that one can give a completely elementary proof of Theorem 1.3, independent of de Branges' Lemma. Thus, we can essentially simplify the proof of [1, Theorem 1.8].

1.4 Localization of Type N

We say that the space $\mathcal{H}(T, A, \mu)$ has *the localization property of type N* if there exist N subsets $T_1, T_2, \ldots, T_N$ of T such that $T_j \subset T_{j+1}$, $1 \le j \le N-1$, $\#(T_{j+1} \setminus T_j) = \infty$ and for any nonzero $f \in \mathcal{H}(T, A, \mu)$ we have $T_f = T_j$ for some j, $1 \le j \le N$, up to finite sets, moreover, N is the smallest integer with this property. Clearly, in this case $T_N = T$ up to a finite set. The strong localization is the localization of type 1.

In what follows we say that an entire function F of finite order is in the *generalized Hamburger–Krein class* if F has simple zeros $\{z_n\}$,

$$\lim_{n\to\infty} |F'(z_n)|^{-1}|z_n|^M = 0 \qquad \text{for any} \qquad M > 0, \tag{1.2}$$

and

$$\frac{1}{F(z)} = \sum_n \frac{1}{F'(z_n)(z - z_n)}. \tag{1.3}$$

Note that when $\{z_n\}$ is power separated and (1.2) is satisfied, one can replace (1.3) by the condition that, for any $K > 0$, $|F(z)| \gtrsim 1$ when $z \notin \cup_n D(z_n, (|z_n| + 1)^{-K})$.

Now we state the description of spaces with localization property of type 2; localization of type N can be described similarly (see [1, Theorem 6.1]).

Theorem 1.4 *The space $\mathcal{H}(T, A, \mu)$ has the localization property of type 2 if and only if there exists a partition $T = T_1 \cup T_2$, $T_1 \cap T_2 = \emptyset$, such that the following three conditions hold:*

(i) *There exists an entire function A_2 in the generalized Hamburger–Krein class such that $\mathcal{Z}_{A_2} = T_2$;*
(ii) *The polynomials belong to the space $L^2(T_2, \mu|_{T_2})$ and are not dense there, but their closure is of finite codimension in $L^2(T_2, \mu|_{T_2})$.*
(iii) *The polynomials belong to the space $L^2(T_1, \tilde{\mu})$ and are dense there, where $\tilde{\mu} = \sum_{t_n \in T_1} \mu_n |A_2(t_n)|^2 \delta_{t_n}$.*

Moreover, T_1 and T are the attraction sets for $\mathcal{H}(T, A, \mu)$.

In Sect. 6 we will give a number of examples of Cauchy–de Branges spaces having localization property of type 2.

2 Equivalent Forms of Zeros Localization

In this section we will prove Theorem 1.1.

A sequence $\{z_k\} \subset \mathbb{C}$ will be said to be *lacunary* if $\inf_k |z_{k+1}|/|z_k| > 1$. A zero genus canonical product over a lacunary sequence will be said to be a *lacunary canonical product*.

For $\Omega \subset \mathbb{C}$, we define its *upper area density* by

$$D^+(\Omega) = \limsup_{R\to\infty} \frac{m_2(\Omega \cap D(0, R))}{\pi R^2},$$

where m_2 denotes the area Lebesgue measure in $\mathbb{C}$. If $D^+(\Omega) = 0$ we say that Ω is a set of zero area density. We say that a set $E \subset \mathbb{R}$ has zero linear density if $|E \cap (0, R)| = o(R)$, $R \to \infty$, where $|e|$ denotes one-dimensional Lebesgue measure of e.

The following result ([2, Theorem 2.6]) will play an important role in what follows. We will often need to verify that a certain entire function belongs to the space $\mathcal{H}(T, A, \mu)$. In the de Branges space setting a much stronger statement is given in [12, Theorem 26].

Theorem 2.1 *Let $\mathcal{H}(T, A, \mu)$ be a Cauchy–de Branges space and let A be of finite order. Then an entire function f is in $\mathcal{H}(T, A, \mu)$ if and only if the following three conditions hold:*

(i) $\sum_n \dfrac{|f(t_n)|^2}{|A'(t_n)|^2\mu_n} < \infty$;
(ii) *there exist a set $E \subset (0, \infty)$ of zero linear density and $N > 0$ such that $|f(z)| \le |z|^N |A(z)|$, $|z| \notin E$;*
(iii) *there exists a set Ω of positive upper area density such that $|f(z)| = o(|A(z)|)$, $|z| \to \infty$, $z \in \Omega$.*

Condition (iii) can be replaced by a stronger conditions that $|f(z)| = o(|A(z)|)$ as $|z| \to \infty$ outside a set of zero area density. It should be mentioned that (iii) is a consequence of the following standard fact about planar Cauchy transforms (see, e.g., [7, Proof of Lemma 4.3]). If ν is a finite complex Borel measure in $\mathbb{C}$, then, for any $\varepsilon > 0$, there exists a set Ω of zero area density such that

$$\left| \int_{\mathbb{C}} \frac{d\nu(\xi)}{z-\xi} - \frac{\nu(\mathbb{C})}{z} \right| < \frac{\varepsilon}{|z|}, \qquad z \in \mathbb{C} \setminus \Omega.$$

We will frequently use the following corollary of this fact: if ν is orthogonal to all polynomials (meaning that $\int |\xi|^k d|\nu|(\xi) < \infty$ and $\int \xi^k d\nu(\xi) = 0$, $k \in \mathbb{Z}_+$), then for any $M > 0$ we have

$$\left| \int_{\mathbb{C}} \frac{d\nu(\xi)}{z-\xi} \right| = o(|z|^{-M}) \tag{2.1}$$

as $|z| \to \infty$ outside some set of zero area density.

Proof of Theorem 1.1 It is obvious that the localization property implies each of the conditions (ii) and (iv). We will show that (ii)$\Longrightarrow$(iii), (iii)$\Longrightarrow$(iv), and (iii) & (iv) $\Longrightarrow$(i).

(ii) $\Longrightarrow$ (iii). Assume that (iii) is not true. Then for some $M > 0$ there exists a nonzero function $F \in \mathcal{H}$ for which there exists an infinite number of zeros $z \in \mathcal{Z}_F$ with $\operatorname{dist}(z, T) \ge |z|^{-M}$.

Let S be an unbounded set which satisfies (ii). Then we can choose two sequences $s_k \in S$ and $z_k \in \mathcal{Z}_F$ such that $2|z_k| \le |s_k| \le |z_{k+1}|/2$ and $\operatorname{dist}(z_k, T) \ge |z_k|^{-M}$. Now put

$$H(z) = F(z) \prod_k \frac{1 - z/s_k}{1 - z/z_k}.$$

A simple estimate of the lacunary infinite products implies that $|H(z)| \lesssim |z|^{M+1}|F(z)|$ for $|z| \geq 1$ and $\operatorname{dist}(z, \{z_k\}) \geq |z_k|^{-M}/2$, in particular, for $z \in T$. Now dividing H by some polynomial P of degree $M+1$ with $\mathcal{Z}_F \subset \mathcal{Z}_F \setminus \{z_k\}$, we conclude by Theorem 2.1 that $\tilde{H} = H/P$ is in $\mathcal{H}$. This contradicts (ii) since $\mathcal{Z}_{\tilde{H}} \cap S$ is an infinite set.

(iii) $\Longrightarrow$(iv) Assume that (iv) is not true. Then there exist a nonzero function $F \in \mathcal{H}(T, A, \mu)$ and a sequence $\{z_n\}$ of its multiple zeros. Without loss of generality we may assume that the sequence z_k is lacunary. By (iii) there exists a sequence n_k such that $|z_k - t_{n_k}| = o(|t_{n_k}|^{-K})$ for any $K > 0$ as $k \to \infty$. Now put

$$\tilde{F}(z) = F(z) \prod_k \frac{(z - t_{n_k} - |t_{n_k}|^{-M})(z - t_{n_k})}{(z - z_k)^2}$$

for some sufficiently large $M > N$, where N is the constant in (1.1). It is easy to see that $|\tilde{F}(z)| \lesssim |F(z)|$ when $z \notin \cup_n D(t_n, C|t_n|^{-N}/2)$ and so $\tilde{F}$ is in $\mathcal{H}(T, A, \mu)$ by Theorem 2.1.

(iii) & (iv) $\Longrightarrow$ (i). Let F be a nonzero function in $\mathcal{H}(T, A, \mu)$. By (iii), all zeros of F, except a finite number, are localized in the disks $D(t_n, |t_n|^{-M})$ for any fixed M. Assume that an infinite subsequence of disks $D(t_{n_k}, |t_{n_k}|^{-M})$ (where t_{n_k} is a lacunary sequence) contains two zeros $z_k, \tilde{z}_k$ of F. Then the function

$$\tilde{F}(z) = F(z) \prod_k \frac{(z - t_{n_k})^2}{(z - z_k)(z - \tilde{z}_k)}$$

is in $\mathcal{H}(T, A, \mu)$ by Theorem 2.1, a contradiction with (iv). □

3 Localization and Polynomial Density

This section is devoted to the proof of Theorem 1.2. In Sect. 3.1 we show that the polynomial density implies the strong localization property. In Sect. 3.2 we will prove the converse statement.

First of all we prove that the localization property implies that μ_n decrease superpolynomially.

Proposition 3.1 *Let $\mathcal{H}(T, A, \mu)$ have the localization property. Then for any $M > 0$ we have $\mu_n \lesssim |t_n|^{-M}$.*

Proof Assume the converse. Then there exist $M > 0$ and an infinite subsequence $\{n_k\}$ such that $\mu_{n_k} \geq |t_{n_k}|^{-M}$. Without loss of generality we can assume that $\{t_{n_k}\}$ is lacunary. Let U be the lacunary product with zeros $t_{n_{10k}}$. Put

$$f(z) = A(z)U^3(z) \prod_k \left(1 - \frac{z}{t_{n_k}}\right)^{-1}.$$

Then, by simple estimates of lacunary canonical products, we have $|f(t_{n_k})| \lesssim |t_{n_k}|^{-K}|A'(t_{n_k})|$ for any fixed $K > 0$, whence f satisfies condition (i) of Theorem 2.1. Since $\frac{A(z)}{z-t_{n_1}} \in \mathcal{H}(T, A, \mu)$ and $|U(z)|^3 \lesssim |z|^{-K} \prod_k |1 - z/t_{n_k}|$ for any $K > 0$ and $z \notin \cup_k D(t_{n_k}, 1)$, we conclude that f satisfies conditions (ii) and (iii) of Theorem 2.1 and so $f \in \mathcal{H}(T, A, \mu)$. This contradicts the property (iv) from Theorem 1.1. □

3.1 Polynomial Density ⟹ Strong Localization Property

Let $f \in \mathcal{H}(T, A, \mu) \setminus \{0\}$. If the polynomials are dense in $L^2(\mu)$, then it is not difficult to show that for any $M > 0$ there exist $L > 0$ and $R > 0$ such that

$$\inf\{|z|^L |f(z)| : \operatorname{dist}(z, T) \geq |z|^{-M}, |z| > R\} > 0. \tag{3.1}$$

A simple proof of this fact is given in detail in [1, Section 3.1] and we omit it.

In particular, it follows from (3.1) that for any $M > 0$ all zeros of $f \in \mathcal{H}(T, A, \mu) \setminus \{0\}$ except, may be, a finite number, are in $\cup_n D(t_n, |t_n|^{-M})$. Therefore by Theorem 1.1, the space $\mathcal{H}(T, A, \mu)$ has the localization property, and so any disc $D(t_n, |t_n|^{-M})$ except a finite number contains at most one zero of f.

Now we show that the disk $D(t_k, |t_k|^{-M})$ contains exactly one point of $\mathcal{Z}_f$ if $|k|$ is sufficiently large. Let

$$f(z) = \sum_n \frac{d_n \mu_n^{1/2}}{z - t_n}, \qquad g(z) = \sum_{n \neq k} \frac{d_n \mu_n^{1/2}}{z - t_n}.$$

Recall that $|f(z)| \geq c|z|^{-L}$ for $|z - t_k| = |t_k|^{-M}$ and sufficiently large k, where L is the number from (3.1). Since $\mu_k = o(|t_k|^{-\tilde{L}})$, $k \to \infty$, for any $\tilde{L} > 0$, we conclude that $|f(z) - g(z)| < c|z|^{-L}/2$ for $|z - t_k| = |t_k|^{-M}$, $k \geq k_0$.

Put $F = Af$, $G = Ag$. Then F, G are entire and $|F - G| < |G|$ on $|z - t_k| = |t_k|^{-M}$, $k \geq k_0$. By the Rouché theorem, F and G have the same number of zeros in $D(t_k, |t_k|^{-M})$, $k \geq k_0$. Since $G(t_k) = 0$, we conclude that $F = Af$ has a zero in $D(t_k, |t_k|^{-M})$, $|k| \geq k_0$. The strong localization property is proved.

3.2 Strong Localization ⟹ Polynomial Density

This implication is almost trivial. Let $\{u_n\} \in \ell^2$ be a nonzero sequence such that $\sum_n u_n t_n^k \mu_n^{1/2} = 0$ for any $k \in \mathbb{N}_0$. Consider the function

$$F(z) = A(z) \sum_n \frac{u_n \mu_n^{1/2}}{z - t_n}.$$

Then F belongs to the Cauchy–de Branges space $\mathcal{H}(T, A, \mu)$ and since all the moments of u_n are zero, it is easy to see that for any $K > 0$, $|F(z)/A(z)| = o(|z|^{-K})$ as $|z| \to \infty$ and $z \notin \cup_n D(t_n, C|t_n|^{-N}/2)$, where C, N are parameters from (1.1). On the other hand, since we have the strong localization property, for any $M > 0$ all but a finite number of zeros of f lie in $\cup_n D(t_n, r_n)$, where $r_n = |t_n|^{-M}$ and $\#\big(\mathcal{Z}_f \cap D(t_n, r_n)\big) \le 1$ for all indices n except, possibly, a finite number.

Let T_1 be the set of those t_n for which the corresponding disk $D(t_n, r_n)$ contains exactly one zero of F (denoted by z_n with the same index n) and let $A = A_1 A_2$ be the corresponding factorization of A, where A_2 is a polynomial with finite zero set $T \setminus T_1$. Put

$$F_1(z) = A_1(z) \prod_{t_n \in T_1} \frac{z - z_n}{z - t_n}.$$

We can choose M to be so large that the above product converges, and, moreover, $|F_1(z)| \asymp |A_1(z)|$ when $\operatorname{dist}(z, T_1) \ge C|z|^{-N}/2$. Then we can write $F = F_1 F_2$, and it is easy to see that in this case F_2 is at most a polynomial. Thus, for some $L > 0$, we have $|F(z)|/|A(z)| \gtrsim |z|^{-L}$, as $|z| \to \infty$ and $z \notin \cup_n D(t_n, C|t_n|^{-N}/2)$, a contradiction. □

4 Ordering Theorem for the Zeros of Cauchy Transforms

First we show that in the proof of ordering for attraction sets one can consider only functions with zeros in T.

Lemma 4.1 *Let $f \in \mathcal{H}(T, A, \mu)$, $f \ne 0$, and let T_f be defined as in Sect. 1.3. Then there exists a function $A_f \in \mathcal{H}(T, A, \mu)$ which vanishes exactly on T_f up to a finite set.*

Proof Let z_n be a zero of f closest to the point $t_n \in T_f$. Since T_f is defined up to finite sets, we may assume without loss of generality that this is a one-to-one correspondence between $\mathcal{Z}_f$ and T_f. Put

$$A_f(z) = f(z) \prod_{t_n \in T_f} \frac{z - t_n}{z - z_n}.$$

Since we have $|z_n - t_n| \le |t_n|^{-M}$ with M much larger than N from the power separation condition (1.1), it is easy to see that $|A_f(z)| \asymp |f(z)|$, $\operatorname{dist}(z, T) \ge C|t_n|^{-N}/2$, and $|A_f(t_n)| \asymp |f(t_n)|$, $t_n \in T \setminus T_f$. Hence, $A_f \in \mathcal{H}(T, A, \mu)$ by Theorem 2.1. □

Corollary 4.1 *Let $\mathcal{H}(T, A, \mu)$ have the localization property and assume that the zeros of a function $f \in \mathcal{H}(T, A, \mu)$ are localized near the whole set T up to a* *finite set. Then for any* $K > 0$ *there exist* $c, M > 0$ *such that for the discs* $D_k = D(t_k, |t_k|^{-K})$, $t_k \in T$, *we have*

$$|f(z)| \geq c|z|^{-M}|A(z)|, \qquad z \notin \cup_k D_k. \tag{4.1}$$

Proof Let A_f be a function constructed from f as in Lemma 4.1. Since the zero set T_f of A_f differs from T by a finite set and all functions in $\mathcal{H}(T, A, \mu)$ are of finite order, we can write $A_f = APQ^{-1}e^R$, where P, Q, R are some polynomials. Let us show that R is a constant. Indeed, since $A/(z - t_n) \in \mathcal{H}(T, A, \mu)$ for any $t_n \in T$, we have

$$A_f - \frac{A}{z - t_n} = A\left(\frac{P}{Q}e^R - \frac{1}{z - t_n}\right) \in \mathcal{H}(T, A, \mu).$$

If $R \neq const$, then the function in brackets will have infinitely many zeros, a contradiction to localization.

As mentioned above, $|A_f(z)| \asymp |f(z)|$, dist $(z, T) \geq C|t_n|^{-N}/2$, where C, N are constants from power separation condition (1.1). This implies (4.1). □

Now we pass to the proof of Theorem 1.3. By Lemma 4.1 we may assume, in what follows, that $f = A_1$, $g = \tilde{A}_1$, where $\mathcal{Z}_{A_1}, \mathcal{Z}_{\tilde{A}_1} \subset T$. Thus, we may write $A = A_1A_2 = \tilde{A}_1\tilde{A}_2$ for some entire functions A_2 and $\tilde{A}_2$.

Let $A_1 = BA_0$, $\tilde{A}_1 = \tilde{B}A_0$, where B and $\tilde{B}$ have no common zeros. To prove Theorem 1.3, we need to show that either B or $\tilde{B}$ has finite number of zeros.

Note that $A_1 - \alpha\tilde{A}_1$ is in $\mathcal{H}(T, A, \mu)$ for any $\alpha \in \mathbb{C}$. Therefore, the zeros of $B - \alpha\tilde{B}$ are localized near T. As we will see, this is a very strong restriction which cannot hold unless one of the functions B or $\tilde{B}$ has finite number of zeros.

4.1 Key Proposition

The following proposition is the crucial step of the argument. In [1] a similar statement was proved using a deep result of de Branges [12, Lemma 7]; it was valid even without localization assumption. This argument is no longer applicable in non-de Brangean case when t_n are nonreal. However, taking into account the localization property, one can give an elementary proof in the general case.

Proposition 4.1 *If the functions* B *and* $\tilde{B}$ *defined above have infinitely many zeros, then there exists* $M > 0$ *such that at least one of the following two statements holds*:

(i) *there exists a subsequence* $\{t_{n_k}\} \subset \mathcal{Z}_B$ *such that* $|B'(t_{n_k})| \leq 8|t_{n_k}|^M|\tilde{B}(t_{n_k})|$;
(ii) *there exists a subsequence* $\{t_{n_k}\} \subset \mathcal{Z}_{\tilde{B}}$ *such that* $|\tilde{B}'(t_{n_k})| \leq 8|t_{n_k}|^M|B(t_{n_k})|$.

Proof We will often use the following obvious observation: if f is a function of finite order, $|z_0| > 2$ and $\mathrm{dist}\,(z_0, \mathcal{Z}_f) \ge c|z_0|^{-K}$, then there exists $L > 0$ (depending on K and c) such that

$$|f(z_0)|/2 \le |f(z)| \le 2|f(z_0)|, \qquad z \in D(z_0, |z_0|^{-L}). \tag{4.2}$$

This statement follows by standard estimates of canonical products.

Step 1. Assume that neither of the conclusions of the proposition holds. Consider the case where $|B'(t_n)| > 8|t_n|^M |\tilde{B}(t_n)|$ for any M and all n except a finite number. It follows from (4.2) (applied to $B(z)/(z - t_n)$) that there exists L such that $|B(z)| \ge |t_n|^{-L}|B'(t_n)|/2$ and also $|\tilde{B}(z)| \le 2|\tilde{B}(t_n)|$ for $z \in C_n = \{|z - t_n| = |t_n|^{-L}\}$. Now if $M > L$, we have $2|\tilde{B}(z)| < |B(z)|$. By the Rouché theorem, we conclude that for any α with $1 \le |\alpha| \le 2$, the function $B - \alpha\tilde{B}$ has exactly one zero in the disc $D(t_n, |t_n|^{-L})$, $t_n \in \mathcal{Z}_B$. Similarly, $B - \alpha\tilde{B}$ has exactly one zero in the disc $D(t_n, |t_n|^{-L})$, $t_n \in \mathcal{Z}_{\tilde{B}}$. We conclude that

for any sufficiently large $L > 0$ and any α with $1 \le |\alpha| \le 2$, the function $B - \alpha\tilde{B}$ has exactly one zero in each disc $D(t_n, |t_n|^{-L})$, $t_n \in \mathcal{Z}_B \cup \mathcal{Z}_{\tilde{B}}$, except a finite number.

Step 2. Next we prove the following: *there exists an infinite set $\mathcal{A} \subset \{1 \le |z| \le 2\}$ such that for any $\alpha \in \mathcal{A}$ and any $L > 0$ the function $B - \alpha\tilde{B}$ has at most finite number of zeros outside the union of the discs $D_n = D(t_n, |t_n|^{-L})$, $t_n \in \mathcal{Z}_B \cup \mathcal{Z}_{\tilde{B}}$.* In view of localization we know that all zeros of $B - \alpha\tilde{B}$ are localized near T. Thus, we only need to show that for many values of α the function $B - \alpha\tilde{B}$ has no zeros in a neighborhood of $t_n \in T \setminus (\mathcal{Z}_B \cup \mathcal{Z}_{\tilde{B}})$. Put

$$w_n = \frac{B(t_n)}{\tilde{B}(t_n)}, \qquad t_n \in T \setminus (\mathcal{Z}_B \cup \mathcal{Z}_{\tilde{B}}).$$

Since B, $\tilde{B}$ are entire functions of finite order whose zeros are power separated from $T \setminus (\mathcal{Z}_B \cup \mathcal{Z}_{\tilde{B}})$, it follows that for any $M > 0$ there exists sufficiently large $L > 0$ such that for $z \in D_n = D(t_n, |t_n|^{-L})$ with $t_n \in T \setminus (\mathcal{Z}_B \cup \mathcal{Z}_{\tilde{B}})$,

$$\left| \frac{B(z)}{\tilde{B}(z)} - w_n \right| < |t_n|^{-M}, \qquad z \in D_n. \tag{4.3}$$

Obviously, the discs $D(w_n, |t_n|^{-M})$, $t_n \in T \setminus (\mathcal{Z}_B \cup \mathcal{Z}_{\tilde{B}})$ do not cover the annulus $\{1 \le |z| \le 2\}$ if M is sufficiently large. Therefore, (4.3) implies that we have a continuum of α with $1 \le |\alpha| \le 2$ such that $\alpha \ne B(z)/\tilde{B}(z)$ for $z \in \cup_{t_n \in T \setminus (\mathcal{Z}_B \cup \mathcal{Z}_{\tilde{B}})} D_n$, Thus, for such α, all zeros of $B - \alpha\tilde{B}$ up to a finite number belong to $\cup_{t_n \in \mathcal{Z}_B \cup \mathcal{Z}_{\tilde{B}}} D_n$ as required.

Step 3. By Steps 1 and 2, if $\alpha \in \mathcal{A}$ all zeros of the function $B - \alpha\tilde{B}$ are located near $\mathcal{Z}_B \cup \mathcal{Z}_{\tilde{B}}$. Then we can write

$$B - \alpha\tilde{B} = B\tilde{B}R_\alpha e^{Q_\alpha}\Pi_\alpha, \qquad \Pi_\alpha(z) = \prod_{t_n \in \mathcal{Z}_B \cup \mathcal{Z}_{\tilde{B}}} \frac{z - s_n}{z - t_n},$$

where $s_n \in D_n$ are the zeros of $B - \alpha\tilde{B}$, R_α is some rational function and Q_α is a polynomial.

Step 4. Assume that there exist $\alpha \neq \beta$ such that $Q_\alpha = Q_\beta = Q$. Then for $B_1 = e^Q B$ and $\tilde{B}_1 = e^Q\tilde{B}$ we have

$$B_1 - \alpha\tilde{B}_1 = B_1\tilde{B}_1R_\alpha\Pi_\alpha, \qquad B_1 - \beta\tilde{B}_1 = B_1\tilde{B}_1R_\beta\Pi_\beta.$$

It follows that

$$\beta - \alpha = B_1(R_\alpha\Pi_\alpha - R_\beta\Pi_\beta), \qquad \alpha^{-1} - \beta^{-1} = \tilde{B}_1(\alpha^{-1}R_\alpha\Pi_\alpha - \beta^{-1}R_\beta\Pi_\beta).$$

Since for any fixed α the zeros s_n of $B - \alpha\tilde{B}$ satisfy $|s_n - t_n| < |t_n|^{-K}$ for any $K > 0$, it is easy to see that Π_α admits the expansion

$$\Pi_\alpha = 1 + \sum_{k=1}^{K} \frac{c_k}{z^k} + O\left(\frac{1}{z^{K+1}}\right)$$

as $|z| \to \infty$, $z \notin \cup_{t_n \in \mathcal{Z}_B \cup \mathcal{Z}_{\tilde{B}}} D_n$ for any fixed L. Therefore, the function $R_\alpha\Pi_\alpha - R_\beta\Pi_\beta$ is either equivalent to cz^{-K} for some $K \in \mathbb{Z}$ or decays faster than any power when $|z| \to \infty$, $z \notin \cup_{t_n \in \mathcal{Z}_B \cup \mathcal{Z}_{\tilde{B}}} D_n$.
Assume that B_1 is not a polynomial. Then $|B_1|$ tends to infinity faster than any power along some sequence of points outside $\cup_{t_n \in \mathcal{Z}_B \cup \mathcal{Z}_{\tilde{B}}} D_n$. In view of the form of Π_α and Π_β this implies that for any $K > 0$ we have

$$|R_\alpha\Pi_\alpha - R_\beta\Pi_\beta| = o(|z|^{-K}), \qquad |z| \to \infty, \ \ z \notin \cup_{t_n \in \mathcal{Z}_B \cup \mathcal{Z}_{\tilde{B}}} D_n.$$

Similarly, if $\tilde{B}_1$ is not a polynomial, then for any $K > 0$,

$$|\alpha^{-1}R_\alpha\Pi_\alpha - \beta^{-1}R_\beta\Pi_\beta| = o(|z|^{-K}), \qquad |z| \to \infty, \ \ z \notin \cup_{t_n \in \mathcal{Z}_B \cup \mathcal{Z}_{\tilde{B}}} D_n.$$

Since $\alpha \neq \beta$, it follows that $R_\alpha\Pi_\alpha$ decays faster than any power, a contradiction. Thus, either B_1 or $\tilde{B}_1$ is a polynomial.

Step 5. It remains to consider the case when $Q_\alpha \neq Q_\beta$ for any $\alpha, \beta \in \mathcal{A}$, $\alpha \neq \beta$. Without loss of generality we may assume that there exist $\alpha_0, \alpha_1, \alpha_2, \alpha_3 \in \mathcal{A}$ with Q_{α_j} of the same degree m such that the coefficients c_{α_j} at z^m are different.

Dividing by $e^{Q_{\alpha_0}}$ we obtain new functions B_1 and $\tilde{B}_1$ satisfying

$$\alpha_0 - \alpha_j = B_1(e^{\tilde{Q}_j} R_{\alpha_j} \Pi_{\alpha_j} - R_{\alpha_0} \Pi_{\alpha_0}), \qquad j = 1, 2, 3,$$

where R_α, Π_α are defined as above and $\tilde{Q}_j = Q_{\alpha_j} - Q_{\alpha_0}$. Thus,

$$|B_1(z)| \asymp |R_{\alpha_j}(z)|^{-1} e^{-\operatorname{Re} Q_j(z)}$$

as $|z| \to \infty$ along each ray $\{z = re^{i\theta}\}$ on which $\lim_{r\to\infty} \operatorname{Re} Q_j(re^{i\theta}) = \infty$ and outside the set $\cup_{t_n \in \mathcal{Z}_B \cup \mathcal{Z}_{\tilde{B}}} D_n$. Since the real part of a polynomial tends to infinity approximately on the half of the rays, there exists an angle Γ of positive size such that two of the expressions $|R_{\alpha_j}(z)|e^{\operatorname{Re} Q_j(z)}$ have the same asymptotics inside the angle, say, $|R_{\alpha_1}(z)|e^{\operatorname{Re} Q_1(z)} \asymp |R_{\alpha_2}(z)|e^{\operatorname{Re} Q_2(z)}$ as $|z| \to \infty$, $z \in \Gamma$. This is obviously impossible if the leading coefficients of Q_1 and Q_2 are different. This contradiction completes the proof of the proposition. □

4.2 End of the Proof of Theorem 1.3

The rest of the proof is similar to the proof of [1, Theorem 1.8]. Recall that $A = A_1 A_2 = \tilde{A}_1 \tilde{A}_2$. Since A_1 and $\tilde{A}_1$ belong to $\mathcal{H}(T, A, \mu)$, we have

$$\sum_{t_n \in T} \frac{|A_1(t_n)|^2}{\mu_n |A'(t_n)|^2} = \sum_{t_n \in \mathcal{Z}_{A_2}} \frac{1}{\mu_n |A_2'(t_n)|^2} < \infty, \tag{4.4}$$

and, analogously,

$$\sum_{t_n \in \mathcal{Z}_{\tilde{A}_2}} \frac{1}{\mu_n |\tilde{A}_2'(t_n)|^2} < \infty, \tag{4.5}$$

Assume that (i) in Proposition 4.1 holds. Dividing if necessary B by a polynomial we may assume that $|B'(t_{n_l})| \le |\tilde{B}(t_{n_l})|$. Hence, we may construct a lacunary canonical product U_1 such that $\mathcal{Z}_{U_1} \subset \mathcal{Z}_B$ and

$$|B'(t_n)| \le |\tilde{B}(t_n)|, \qquad t_n \in \mathcal{Z}_{U_1}.$$

Let U_2 be another lacunary product with zeros in $\mathbb{C} \setminus \cup_{t_n \in T} D(t_n, C|t_n|^{-N})$ such that

$$|U_2(t_n)| = o(|U_1(t_n)|), \qquad n \to \infty, \qquad t_n \in T \setminus \mathcal{Z}_{U_1}, \tag{4.6}$$

$$|U_2(t_n)| = o(|U_1'(t_n)|), \qquad t_n \in T. \tag{4.7}$$

This may be achieved if we choose zeros of U_2 to be much sparser than the zeros of U_1. Let us show that in this case

$$f := A_1 \cdot \frac{U_2}{U_1} \in \mathcal{H}(T, A, \mu),$$

which contradicts the localization. Since A_1 is in $\mathcal{H}(T, A, \mu)$, while U_1 and U_2 are lacunary products, it is clear that conditions (ii) and (iii) hold for f. It remains to show that

$$\sum_{t_n \in T} \frac{|f(t_n)|^2}{|A'(t_n)|^2 \mu_n} < \infty.$$

Since f vanishes on $\mathcal{Z}_{A_1} \setminus \mathcal{Z}_{U_1}$, we need to estimate the sums over $\mathcal{Z}_{A_2}$ and $\mathcal{Z}_{U_1}$. By (4.6) and (4.4), we have

$$\sum_{t_n \in \mathcal{Z}_{A_2}} \frac{|f(t_n)|^2}{|A'(t_n)|^2 \mu_n} = \sum_{t_n \in \mathcal{Z}_{A_2}} \frac{|U_2(t_n)|^2}{|U_1(t_n)|^2} \cdot \frac{1}{|A_2'(t_n)|^2 \mu_n} < \infty$$

To estimate the sum over $\mathcal{Z}_{U_1}$ note first that BA_0 divides $A = \tilde{B} A_0 \tilde{A}_2$, whence B divides $\tilde{A}_2$. Thus $\mathcal{Z}_{U_1} \subset \mathcal{Z}_{\tilde{A}_2}$. Also, for $t_n \in \mathcal{Z}_{U_1}$,

$$|A_1'(t_n)| = |B'(t_n)| \cdot |A_0(t_n)| \le |A_0(t_n)| \cdot |\tilde{B}(t_n)| = |\tilde{A}_1(t_n)|. \tag{4.8}$$

Now by (4.7), (4.8), and (4.5) we have

$$\sum_{t_n \in \mathcal{Z}_{U_1}} \frac{|f(t_n)|^2}{|A'(t_n)|^2 \mu_n} = \sum_{t_n \in \mathcal{Z}_{U_1}} \frac{|U_2(t_n)|^2}{|U_1'(t_n)|^2} \cdot \frac{|A_1'(t_n)|^2}{|\tilde{A}_1(t_n)|^2 |\tilde{A}_2'(t_n)|^2 \mu_n} \lesssim \sum_{t_n \in \mathcal{Z}_{U_1}} \frac{1}{|\tilde{A}_2'(t_n)|^2 \mu_n} < \infty.$$

Thus, $f \in \mathcal{H}(T, A, \mu)$ and this contradiction completes the proof of Theorem 1.3. □

5 Localization of Type 2

In this section we prove Theorem 1.4. In what follows we will need the following property of functions in the generalized Hamburger–Krein class: since $1/F$ is a Cauchy transform and $|F'(z_n)|$ decays faster than any power, we have

$$\sum_n \frac{z_n^k}{F'(z_n)} = 0, \qquad k \in \mathbb{Z}_+. \tag{5.1}$$

Otherwise, by the arguments from Sect. 3.1, $1/F$ decays at most polynomially away from the zeros whence F itself is a polynomial. Furthermore, it follows from (5.1) that for any $K, M > 0$

$$|z|^M = o(|F(z)|), \qquad |z| \to \infty, \ \ \mathrm{dist}\,(z, \{z_n\}) \ge |z|^{-K}. \tag{5.2}$$

5.1 Proof of Sufficiency in Theorem 1.4

Assume that $\mathcal{H}(T, A, \mu)$ satisfies the conditions (i)–(iii). We will show that in this case $\mathcal{H}(T, A, \mu)$ has localization of type 2. Let $T = T_1 \cup T_2$ and let $A = A_1 A_2$, where A_2 is the Hamburger–Krein class function from (i). Let $\mathcal{H}_2$ be the Cauchy–de Branges space constructed from T_2 and $\mu|_{T_2}$, i.e., $\mathcal{H}_2 = \mathcal{H}(T_2, A_2, \mu|_{T_2})$.

By the hypothesis, the orthogonal complement $\mathcal{L}$ to the polynomials in $L^2(T_2, \mu|_{T_2})$ is finite-dimensional. If $\{d_n\} \in L^2(T_2, \mu|_{T_2}) \setminus \mathcal{L}$, then there exists a nonzero moment for the sequence $\{d_n\}$, that is, $\sum_{t_n \in T_2} \mu_n d_n t_n^K \ne 0$ for some $K \in \mathbb{N}_0$. If $f(z) = A_2(z) \sum_{t_n \in T_2} \frac{\mu_n d_n}{z - t_n}$ is the corresponding function from $\mathcal{H}_2$, then, for any $M > 0$, the function f has a zero in $D(t_n, |t_n|^{-M})$, $t_n \in T_2$, when n is sufficiently large (see Sect. 3.1). Thus, for any function in $\mathcal{H}_2$ except some finite-dimensional subspace, its zeros are localized near the whole set T_2.

Now let $\mathcal{G}$ be the subspace of the Cauchy–de Branges space $\mathcal{H}_2$ defined by

$$\mathcal{G} = \left\{ A_2 \sum_{t_n \in T_2} \frac{\mu_n d_n}{z - t_n} : \{d_n\} \in \mathcal{L} \right\}.$$

This is a finite-dimensional subspace of $\mathcal{H}_2$ and it is easy to see that $F \in \mathcal{G}$ if and only if $F \in \mathcal{H}_2$ and, for any $M > 0$, $|F(z)/A_2(z)| = o(|z|^{-M})$, as $|z| \to \infty$ outside a set of zero density (see (2.1) and (3.1)). Thus, $\mathcal{G}$ is a finite-dimensional space of entire functions with the division property and so it consists of the functions of the form SP where S is some fixed zero-free function and P is any polynomial of degree less than some fixed number L. Note that if $SP \in \mathcal{G}$ and so SP/A_2 decays faster than any power away from zeros of A_2, then we may conclude that A_2/S also is a function in Hamburger–Krein class. Replacing A_2 by A_2/S we may assume that $\mathcal{G}$ consists of polynomials.

We conclude that $\mathcal{H}_2$ has the localization property and for any $F \in \mathcal{H}_2$ we either have $T_F = \emptyset$ (i.e., F is a polynomial) or $T_F = T_2$.

Let $f \in \mathcal{H}(T, A, \mu)$, $f(z) = A(z) \sum_{t_n \in T} \frac{c_n \mu_n^{1/2}}{z - t_n}$. Since by (iii), $|A_2(t_n)| \mu_n^{1/2}$ tends to zero faster than any power of $t_n \in T_1$ when $|t_n| \to \infty$, we have

$$A_2(z) \sum_{t_n \in T_1} \frac{c_n \mu_n^{1/2}}{z - t_n} = \sum_{t_n \in T_1} \frac{A_2(t_n) c_n \mu_n^{1/2}}{z - t_n} + H(z) \tag{5.3}$$

for some entire function H (note that the residues on the left and the right coincide). Let us show using Theorem 2.1 that $H \in \mathcal{H}_2$. Indeed, the Cauchy transform on the left-hand side of (5.3) is bounded on T_2 and so

$$\sum_{t_n \in T_2} \frac{|H(t_n)|^2}{|A_2'(t_n)|^2 \mu_n} < \infty.$$

Conditions (ii) and (iii) of Theorem 2.1 are fulfilled since $1/A_2$ is a Cauchy transform whence the same is true for H/A_2.

Note also that $F(z) := A_2(z) \sum_{t_n \in T_2} \frac{c_n \mu_n^{1/2}}{z - t_n}$ is by definition in $\mathcal{H}_2$. Thus,

$$f = A_1(g + H + F)$$

where $g(z) = \sum_{t_n \in T_1} \frac{c_n A_2(t_n) \mu_n^{1/2}}{z - t_n}$ and $H + F \in \mathcal{H}_2$.

Assume that $H + F \not\equiv 0$. Then, either $H + F$ is a polynomial or the zeros of $H + F$ are localized near T_2 up to a finite set. In both cases, there exists $K > 0$ such that the discs $D(t_k, r_k)$, $t_k \in T$, $r_k = |t_k|^{-K}$, are pairwise disjoint and, for sufficiently large k we have

$$|H(z) + F(z)| > 1, \qquad |z - t_k| = r_k.$$

In the case when the zeros of $H + F$ are localized near the whole set T_2 up to a finite set, we use Corollary 4.1 and (5.2) applied to A_2. Since $|g(z)| \to 0$ whenever $|z - t_k| = r_k$ and $k \to \infty$, we conclude by the Rouché theorem that $A_1(g + H + F)$ has exactly one zero in each $D(t_k, r_k)$, $t_k \in T_1$, except possibly a finite number. Also, if $H + F$ is not a polynomial, then f has zeros near the whole set T_2 up to a finite subset (again apply the Rouché theorem to small disks $D(t_k, r_k)$, $t_k \in T_2$, $r_k = |t_k|^{-K}$, and use the fact that $|H + F| \gtrsim 1$, $|z - t_k| = r_k$).

It remains to consider the case $H + F = 0$, i.e., $f = A_1 g$. Since the polynomials are dense in $L^2(T, \tilde{\mu})$, the space $\mathcal{H}(T_1, A_1, \tilde{\mu})$ has the strong localization property, and so $T_f = T_1$ up to a finite set.

5.2 *Proof of Necessity in Theorem 1.4*

Assume that $\mathcal{H}(T, A, \mu)$ has the localization property of type 2. Let f be a function from $\mathcal{H}(T, A, \mu)$ such that $\#(T \setminus T_f) = \infty$. Then, by Lemma 4.1 there exist T_1 ($T_1 = T_f$ up to a finite set) and a function A_1 with simple zeros in T_1 such that $A_1 \in \mathcal{H}_2$. We now may write $A = A_1 A_2$ for some entire A_2 with $\mathcal{Z}_{A_2} = T_2$.

Proof of (i). Since $A_1 \in \mathcal{H}(T, A, \mu)$, we have

$$\frac{1}{A_2(z)} = \frac{A_1(z)}{A(z)} = \sum_{t_n \in T} \frac{c_n \mu_n^{1/2}}{z - t_n}$$

for some $\{c_n\} \in \ell^2$. It is immediate that $c_n = 1/A_2'(t_n)$, $t_n \in T_2$, and $c_n = 0$ otherwise. Also we have

$$\sum_{t_n \in T_2} \frac{1}{|A_2'(t_n)|^2 \mu_n} < \infty.$$

Since localization property implies that μ_n decay faster than any power, we conclude that A_2 belongs to the generalized Hamburger–Krein class.

Proof of (ii). Note that by (5.1) (applied to A_2) we have

$$\sum_{t_n \in T_2} \frac{t_n^k}{A_2'(t_n)} = 0, \qquad k \in \mathbb{Z}_+,$$

whence the sequence $\{(\mu_n A_2'(t_n))^{-1}\}_{t_n \in T_2}$ is orthogonal to all polynomials in $L^2(T_2, \mu|_{T_2})$.

Now assume that $\{c_n\} \in \ell^2$ and $\{c_n \mu_n^{-1/2}\}$ is orthogonal to all polynomials in $L^2(T_2, \mu|_{T_2})$. Consider the function $f(z) = A_2(z) \sum_{t_n \in T_2} \frac{c_n \mu_n^{1/2}}{z - t_n}$ which belongs to $\mathcal{H}_2 = \mathcal{H}(T_2, A_2, \mu|_{T_2})$. Since $A_1 f \in \mathcal{H}(T, A, \mu)$ and $\mathcal{H}(T, A, \mu)$ has localization property of type 2, the zeros $\{z_n\}$ of f either form a finite set or are localized near T_2. However, in the latter case f satisfies (4.1), a contradiction to the fact that, by (2.1), f/A_2 decays faster than any power outside some set of zero area density.

Thus, any function f constructed above is of the form PS where P is a polynomial and S is some zero-free entire function. It is clear from the localization property that the function S must be the same for all such f-s (up to multiplication by a constant), and that A_2/S also is a Hamburger–Krein class function. Replacing A_2 by A_2/S we may assume that f is a polynomial.

To summarize, for any $\{c_n \mu_n^{-1/2}\}$ which is orthogonal to all polynomials in $L^2(T_2, \mu|_{T_2})$, the function f is a polynomial. Since $A_1 f \in \mathcal{H}(T, A, \mu)$, it remains to show that the degrees of polynomials P such that $PA_1 \in \mathcal{H}(T, A, \mu)$ are uniformly bounded. Let us show that the property that $PA_1 \in \mathcal{H}(T, A, \mu)$ for any polynomial P contradicts the localization property of type 2. If $PA_1 \in \mathcal{H}(T, A, \mu)$ for any polynomial P, then the function $A_1(z) \sum_{k \ge 0} a_k z^k$ is in $\mathcal{H}(T, A, \mu)$ for any sequence $\{a_k\}$ such that

$$\sum_{k \ge 0} |a_k| \cdot \|z^k A_1\|_{\mathcal{H}(T, A, \mu)} < \infty.$$

This contradicts the localization property since for nonzero a_k decaying sufficiently rapidly the function $A_1(z)\sum_{k\ge 0} a_k z_k$ cannot have all but finite number of zeros localized near T_1 or near T.

Proof of (iii). Assume that (iii) is not satisfied, that is, the polynomials are not dense in $\mathcal{H}(T_1, \tilde{\mu})$, and so this space does not have the strong localization property. Then there exists $G(z) = A_1(z)\sum_{t_n\in T_1} \frac{c_n A_2(t_n)\mu_n^{1/2}}{z-t_n} \in \mathcal{H}(T_1, A_1, \tilde{\mu})$, $\{c_n\}_{t_n\in T_1} \in \ell^2$, with the property that there exists an infinite sequence of disks $D(t_{n_j}, |t_{n_j}|^{-M})$, $t_{n_j} \in T_1$, such that $\#D(t_{n_j}, |t_{n_j}|^{-M}) \cap \mathcal{Z}_G = 0$. Now put

$$H(z) = A_2(z)\sum_{t_n\in T_1} \frac{c_n\mu_n^{1/2}}{z-t_n} - \sum_{t_n\in T_1} \frac{c_n A_2(t_n)\mu_n^{1/2}}{z-t_n}.$$

The function H is entire and, as in the proof of sufficiency, $H \in \mathcal{H}_2$. This means that H can be written as

$$H(z) = -A_2(z)\sum_{t_n\in T_2} \frac{c_n\mu_n^{1/2}}{z-t_n} \qquad \text{for some } \{c_n\}_{t_n\in T_2} \in \ell^2.$$

Now put $f(z) = \sum_{t_n\in T} \frac{c_n\mu_n^{1/2}}{z-t_n}$. Then $f \in \mathcal{H}(T, A, \mu)$ and, by the construction,

$$f(z) = -A_1(z)H(z) + A_1(z)\left(\sum_{t_n\in T_1} \frac{c_n A_2(t_n)\mu_n^{1/2}}{z-t_n} + H(z)\right) = G(z).$$

However, the zeros of g are not localized near the whole T_1, a contradiction.

6 Examples of Localization of Type 2

Here we give a series of examples of spaces $\mathcal{H}(T, A, \mu)$ with localization of type 2. Clearly, the most subtle part is to satisfy condition (ii) of Theorem 1.4. However, there exists a standard way to avoid completeness of polynomials with finite defect. For similar constructions, see [10].

Let A be an entire function with power separated zero set $T = \{t_n\}$ and of the Hamburger–Krein class. Then, in particular, for any $K, M > 0$, we have $|A(z)| \gtrsim |z|^M$ when $z \notin \cup_n D(t_n, (|t_n|+1)^{-K})$. Fix some $N \in \mathbb{N}$ such that $\sum_n |t_n|^{-N} < \infty$, and put

$$\mu_n = |t_n|^{2N}|A'(t_n)|^{-2}. \tag{6.1}$$

Then the polynomials belong to the space $L^2(\mu)$, $\mu = \sum_n \mu_n \delta_{t_n}$, but are not dense there. Indeed, for any $k \in \mathbb{N}_0$, we have

$$\frac{z^{k+1}}{A(z)} = \sum_n \frac{t_n^{k+1}}{A'(t_n)(z - t_n)},$$

whence $\sum_n c_n \mu_n t_n^k = 0$ (take $z = 0$). Hence, for $c_n = (A'(t_n)\mu_n)^{-1}$ we have $\{c_n\} \in L^2(T, \mu)$. It remains to show that the polynomials have finite codimension in $L^2(T, \mu)$. The following proposition shows that this is often true.

Recall that for a positive increasing function w on $\mathbb{R}_+$, its Legendre transform $w^\#$ is defined as $w^\#(x) = \sup_{t \in \mathbb{R}_+} \big(xt - w(t)\big)$. If, moreover, w is convex, then $(w^\#)^\# = w$. In what follows we will use the following technical condition on w:

$$w^\#(x+t) - w^\#(x) \lesssim x^{-1} w(x), \qquad x > 1,\ 0 \le t \le 1. \tag{6.2}$$

Also recall that a positive increasing function M on $\mathbb{R}_+$ is said to be a *normal weight* if $w(t) = \log M(e^t)$ is a convex function of t.

Proposition 6.1 *Let A be a Hamburger–Krein class function and let μ be defined by* (6.1)*. Assume that there exist*

- *a finite set of rays $L_j = \{e^{i\theta_j}\}$, $j = 1, \dots, J$, which divide the plane into a union of angles of size less than π/ρ, where ρ is the order of A*;
- *a finite set of positive increasing normal weights M_j on $\mathbb{R}_+$ such that the Legendre transforms of the functions $w_j(t) = \log M_j(e^t)$ satisfy* (6.2),

such that, for some $K > 0$,

$$|A(z)| \le (|z| + 1)^K M_j(|z|), \qquad z \in L_j, \tag{6.3}$$

and

$$|A'(t_n)| \gtrsim |t_n|^{-K} \max_j M_j(|t_n|), \qquad t_n \in T. \tag{6.4}$$

Then the polynomials have finite (*and nonzero*) *codimension in $L^2(T, \mu)$.*

Example 6.1 The following functions A satisfy the conditions of Proposition 6.1 (if not specified, A is assumed to be a zero genus canonical product with zero set T):

- $t_n = 2^n$, $n \in \mathbb{N}$;
- $t_n = n^\alpha$, $n \in \mathbb{N}$, $\alpha > 2$;
- $t_n = |n|^\alpha \operatorname{sign} n$, $n \in \mathbb{Z}$, $\alpha > 1$;
- $A(z) = z^{-1} \sin(\pi z) \sin(\pi i z)$, $T = \mathbb{Z} \cup i\mathbb{Z}$;
- $A(z) = \sigma(z)$, the Weierstrass σ-function, $T = \mathbb{Z} + i\mathbb{Z}$.

In all the above examples except the first one, one should take $w_j(t) = e^{\beta t}$ for some $\beta > 0$, whence $w^\#(x) = \frac{x}{\beta}\big(\log\frac{x}{\beta} - 1\big)$. In the first example w and $w^\#$ are quadratic functions. Condition (6.2) is satisfied in all these cases.

Proof of Proposition 6.1 Assume that $\{c_n\}$ is orthogonal to the polynomials in $L^2(T, \mu)$ and consider the function

$$f(z) = A(z)\sum_n \frac{c_n\mu_n}{z - t_n}.$$

We will show that any such function f is a polynomial whose degrees are uniformly bounded above. This will prove the proposition.

Note that $f \in \mathcal{H}(T, A, \mu)$, since $\{c_n\mu_n^{1/2}\} \in \ell^2$. It is clear that, for any $k \in \mathbb{N}_0$, one has

$$z^k f(z) - A(z)\sum_n \frac{c_n\mu_n t_n^k}{z - t_n} \equiv 0.$$

Since T is power separated, there exists a constant $N_0 \in \mathbb{N}$ such that the discs $D_n = D(t_n, |t_n|^{-N_0})$ are pairwise disjoint and, from (6.3), $|A(z)| \le |z|^{-K} M_j(|z|)$ for any j when $\operatorname{dist}(z, L_j) \le 2|z|^{-N_0}$ and $|z|$ is sufficiently large.

Now fix some ray $L_j = L$ and let $M = M_j$, $w = w_j$. For $z \notin \cup_n D_n$ and for any $k \in \mathbb{N}_0$ we have, using (6.1) and (6.4),

$$\begin{aligned}|f(z)| &\le \frac{|A(z)|}{|z|^k}\sum_n \frac{|t_n|^k|c_n|\mu_n}{|z - t_n|} \le \frac{|A(z)|}{|z|^k}\sum_n \frac{|c_n|\mu_n^{1/2}|t_n|^{k+N}}{|A'(t_n)|\cdot|z - t_n|}\\ &\lesssim \frac{|A(z)|}{|z|^k}\sum_n \frac{|t_n|^{k+2N+N_0+K}}{|t_n|^N M(|t_n|)} \lesssim \frac{|A(z)|}{|z|^k}\sup_n \frac{|t_n|^{k+2N+N_0+K}}{M(|t_n|)}.\end{aligned}$$

Put $m = 2N + N_0 + K$. Then

$$\log\sup_n \frac{|t_n|^{k+m}}{M(|t_n|)} = \sup_n\Big((k+m)\log|t_n| - w(\log|t_n|)\Big) \le w^\#(k+m).$$

Since the estimate for f holds for all k, we now have for $z \notin \cup_n D_n$ and $\operatorname{dist}(z, L_j) \le 2|z|^{-N_0}$,

$$\log|f(z)| \le \log M(r) + K\log r + \inf_{k\in\mathbb{N}_0}\Big(w^\#(k+m) - k\log r\Big) + O(1),$$

when $r = |z|$ is sufficiently large. It follows from (6.2) that $x\log r - w^\#(x) = y\log r - w^\#(y) + O(\log r)$ whenever $|x - y| \le 1$ and $w^\#(x) \le x\log r$. Then, obviously,

$$\inf_{k\in\mathbb{N}_0}\Big(w^{\#}(k+m)-k\log r\Big)=m\log r-\sup_{k\in\mathbb{N}_0}\Big((k+m)\log r-w^{\#}(k+m)\Big)$$
$$\leq -\sup_{x\geq 0}(x\log r-w^{\#}(x))+O(\log r)=-w(\log r)+O(\log r),$$

where the constants involved in $O(\log r)$ depend only on m and the constants from (6.2). We used that w is convex and so $(w^{\#})^{\#}=w$. Since $w(\log r)=M(r)$, we conclude that

$$|f(z)|\lesssim(|z|+1)^{N_1},\qquad z\notin\cup_n D_n,\ \ \mathrm{dist}\,(z,L_j)\leq 2(|z|+1)^{-N_0},$$

where N_1 admits a uniform bound. It follows that $|f(z)|\lesssim(|z|+1)^{N_1}$ on the ray L_j for any j. Now the standard Phragmén–Lindelöf principle shows that f is a polynomial of degree at most N_1. □

Remark 6.1 Note that using the same argument one can show that in conditions of Proposition 6.1 the polynomials are dense in $L^2(T,\mu)$ when $\mu_n=|t_n|^{-N}|A'(t_n)|^{-2}$ and N is sufficiently large.

Example 6.1 One can easily give examples of localization of type 2 choosing the measure μ_1 on T_1 to be sufficiently small. The space $\mathcal{H}(T,A,\mu)$ has localization of type 2 in all cases given below

(i) Let $T_2=\{|n|^{\alpha}\,\mathrm{sign}\,n\}_{n\in\mathbb{Z}}$, $\alpha>1$, let $A_2(z)=\prod_{t_n\in T_2}(1-z/t_n)$, and let $T_1=\{ik^{\beta}\}_{k\geq 1}$, $0<\beta<\alpha$. Put

$$\mu_n=\begin{cases}|t_n|^N|A_2'(t_n)|^{-2}, & t_n\in T_2,\\ e^{-|t_n|^{\gamma}}, & t_n\in T_1,\end{cases}$$

where $N>0$ and $\gamma>1/2$ for $0<\beta\leq 1/2$, while for $\beta>2$ we assume $\gamma>1/\beta$. Density of polynomials in $L^2(T_1,\tilde{\mu}_1)$, $\tilde{\mu}_1=\sum_{t_n\in T_1}|A_2'(t_n)|^2\mu_n\delta_{t_n}$, follows from [10, Appendix 2].

(ii) Let $T_2=\mathbb{Z}\cup i\mathbb{Z}$, $A(z)=z^{-1}\sin(\pi z)\sin(\pi i z)$, and let $T_1=e^{i\pi/4}(\mathbb{Z}\cup i\mathbb{Z})\setminus\{0\}$. For $M,N\in\mathbb{N}$, put

$$\mu_n=\begin{cases}|t_n|^M e^{-2\pi|t_n|}, & t_n\in T_2,\\ |t_n|^{-N}e^{-(2\sqrt{2}+2)\pi|t_n|}, & t_n\in T_1.\end{cases}$$

(iii) Let $T_2=\mathbb{Z}+i\mathbb{Z}$, $A(z)=\sigma(z)$, and let $T_1=\mathbb{Z}+i\mathbb{Z}+1/2$. For $N\in\mathbb{N}$ and $\gamma>2$, put

$$\mu_n=\begin{cases}|t_n|^N|\sigma'(t_n)|^{-2}, & t_n\in T_2,\\ e^{-|t_n|^{\gamma}}, & t_n\in T_1.\end{cases}$$

References

1. E. Abakumov, A. Baranov, Y. Belov, Localization of zeros for Cauchy transforms. Int. Math. Res. Notices **2015**(15), 6699–6733 (2015)
2. E. Abakumov, A. Baranov, Y. Belov, Krein-type theorems and ordered structure of Cauchy–de Branges spaces. J. Funct. Anal. (2019, to appear). arXiv:1802.03385
3. N.I. Akhiezer, On the weighted approximation of continuous functions by polynomials on the real axis. Uspekhi Mat. Nauk **11** (1956), 3–43; AMS Transl. (Ser. 2) **22**, 95–137 (1962)
4. A. Baranov, Spectral theory of rank one perturbations of normal compact operators. Algebra Anal. **30**(5), 1–56 (2018)
5. A.D. Baranov, D.V. Yakubovich, Completeness and spectral synthesis of nonselfadjoint one-dimensional perturbations of selfadjoint operators. Adv. Math. **302**, 740–798 (2016)
6. A.D. Baranov, D.V. Yakubovich, Completeness of rank one perturbations of normal operators with lacunary spectrum. J. Spectral Theory **8**(1), 1–32 (2018)
7. A. Baranov, Y. Belov, A. Borichev, Summability properties of Gabor expansions. J. Funct. Anal. **274**(9), 2532–2552 (2018)
8. Y. Belov, T. Mengestie, K. Seip, Discrete Hilbert transforms on sparse sequences. Proc. Lond. Math. Soc. **103**(3), 73–105 (2011)
9. A. Borichev, M. Sodin, Weighted polynomial approximation and the Hamburger moment problem, in *Complex Analysis and Differential Equations, Proceedings of the Marcus Wallenberg Symposium in Honor of Matts Essén* (Uppsala University, Uppsala, 1998)
10. A. Borichev, M. Sodin, The Hamburger moment problem and weighted polynomial approximation on discrete subsets of the real line. J. Anal. Math. **76**, 219–264 (1998)
11. J. Clunie, A. Eremenko, J. Rossi, On equilibrium points of logarithmic and Newtonian potentials. J. Lond. Math. Soc. **47**(2), 309–320 (1993)
12. L. de Branges, *Hilbert Spaces of Entire Functions* (Prentice–Hall, Englewood Cliffs, 1968)
13. A. Eremenko, J. Langley, J. Rossi, On the zeros of meromorphic functions of the form $f(z) = \sum_{k=1}^{\infty} a_k/(z - z_k)$. J. Anal. Math. **62**, 271–286 (1994)
14. V. Havin, B. Jöricke, *The Uncertainty Principle in Harmonic Analysis* (Springer, Berlin, 1994)
15. P. Koosis, *The Logarithmic Integral. I* (Cambridge University Press, Cambridge, 1988)
16. J.K. Langley, J. Rossi, Meromorphic functions of the form $f(z) = \sum_{n=1}^{\infty} a_n/(z - z_n)$. Rev. Mat. Iberoamericana **20**(1), 285–314 (2004)
17. S. Mergelyan, Weighted approximation by polynomials. Uspekhi Mat. Nauk **11**, 107–152 (1956); AMS Transl. **10**, 59–106 (1958)

Radially Weighted Besov Spaces and the Pick Property

Alexandru Aleman, Michael Hartz, John E. McCarthy, and Stefan Richter

Abstract For $s \in \mathbb{R}$ the weighted Besov space on the unit ball $\mathbb{B}_d$ of $\mathbb{C}^d$ is defined by

$$B^s_\omega = \{f \in \mathrm{Hol}(\mathbb{B}_d) : \int_{\mathbb{B}_d} |R^s f|^2 \omega dV < \infty\}.$$

Here R^s is a power of the radial derivative operator $R = \sum_{i=1}^d z_i \frac{\partial}{\partial z_i}$, V denotes Lebesgue measure, and ω is a radial weight function not supported on any ball of radius < 1.

Our results imply that for all such weights ω and ν, every bounded column multiplication operator $B^s_\omega \to B^t_\nu \otimes \ell^2$ induces a bounded row multiplier $B^s_\omega \otimes \ell^2 \to B^t_\nu$. Furthermore we show that if a weight ω satisfies that for some $\alpha > -1$ the ratio $\omega(z)/(1 - |z|^2)^\alpha$ is nondecreasing for $t_0 < |z| < 1$, then B^s_ω is a complete Pick space, whenever $s \geq (\alpha + d)/2$.

M.H. was partially supported by a Feodor Lynen Fellowship.
J.M. was partially supported by National Science Foundation Grant DMS 1565243.

A. Aleman
Lund University, Mathematics, Faculty of Science, Lund, Sweden
e-mail: alexandru.aleman@math.lu.se

M. Hartz
Department of Mathematics, Washington University in St. Louis, St. Louis, MO, USA

Fakultät für Mathematik und Informatik, Fern Universität in Hagen, Hagen, Germany
e-mail: mphartz@wustl.edu; michael.hartz@fernuni-hagen.de

J. E. McCarthy
Department of Mathematics, Washington University in St. Louis, St. Louis, MO, USA
e-mail: mccarthy@wustl.edu

S. Richter (✉)
Department of Mathematics, University of Tennessee, Knoxville, TN, USA
e-mail: srichter@utk.edu; Richter@math.utk.edu

A. Aleman et al. (eds.), *Analysis of Operators on Function Spaces*,
Trends in Mathematics, https://doi.org/10.1007/978-3-030-14640-5_3

Keywords Complete Pick space · Besov space · Multiplier

2010 Mathematics Subject Classification Primary 46E22; Secondary 30H15, 30H25

1 Introduction

Let $d \in \mathbb{N}$. In this paper we will address certain questions about functions and multipliers in weighted Besov Hilbert spaces of analytic functions in the unit ball $\mathbb{B}_d = \{z \in \mathbb{C}^d : |z| < 1\}$. In particular, we will show that results about multipliers in standard and Bekollé weighted Besov spaces of [17] and [9] extend to hold for all radial weights, and we will provide simple, but general conditions on radial weight functions ω that imply that all results of [4] can be applied to such a weighted Besov space.

We will use V to denote Lebesgue measure on $\mathbb{C}^d$ restricted to $\mathbb{B}_d$, normalized so that $V(\mathbb{B}_d) = 1$. A non-negative integrable function ω on $\mathbb{B}_d$ is called a radial weight, if for each $0 < r < 1$ the value $\omega(rz)$ is independent of $z \in \partial\mathbb{B}_d$ and the non-degeneracy condition

$$\int_{|z|>r} \omega dV > 0 \quad \text{for each } 0 < r < 1 \tag{1.1}$$

holds. It is easily checked that for radial weights the weighted Bergman space $L^2_a(\omega) = L^2(\omega dV) \cap \mathrm{Hol}(\mathbb{B}_d)$ is closed in $L^2(\omega dV)$, and that point evaluations $f \to f(z)$ are bounded on $L^2_a(\omega)$ for each $z \in \mathbb{B}_d$.

We now fix a radial weight ω. Then we have

$$\|f\|^2_{L^2_a(\omega)} = \int_{\mathbb{B}_d} |f|^2 \omega dV = \sum_{n\geq 0} \|f_n\|^2_{L^2_a(\omega)},$$

where $f = \sum_{n\geq 0} f_n$ is the decomposition of the analytic function f into a sum of homogeneous polynomials f_n of degree n. We associate a one-parameter family of weighted Besov spaces $\{B^s_\omega\}_{s\in\mathbb{R}}$ with ω as follows:

$$\|f\|^2_{B^s_\omega} = \|\omega\|_{L^1(V)} |f(0)|^2 + \sum_{n=1}^{\infty} n^{2s} \|f_n\|^2_{L^2_a(\omega)} \tag{1.2}$$

$$B^s_\omega = \{f \in \mathrm{Hol}(\mathbb{B}_d) : \|f\|^2_{B^s_\omega} < \infty\}.$$

Let $R = \sum_{i=1}^d z_i \frac{\partial}{\partial z_i}$ denote the radial derivative operator, then $Rf = \sum_{n\geq 1} n f_n$. More generally, for each nonzero $s \in \mathbb{R}$ we may consider the "fractional" transformation $R^s : \sum_{n\geq 0} f_n \to \sum_{n\geq 1} n^s f_n$. It is thus clear that

$$B^s_\omega = \{f \in \mathrm{Hol}(\mathbb{B}_d) : R^s f \in L^2_a(\omega)\},$$
$$\|f\|^2_{B^s_\omega} = \|\omega\|_{L^1(V)}|f(0)|^2 + \int_{\mathbb{B}_d} |R^s f|^2 \omega dV.$$

One checks that (1.1) implies that each B^s_ω is a Hilbert space, and point evaluations for all points in $\mathbb{B}_d$ are bounded. A space $\mathcal{H}$ of analytic functions that occurs as one of the spaces B^s_ω for a radial weight ω and some $s \in \mathbb{R}$ will be called a weighted Besov space.

If $\omega(z) = 1$, $s \in \mathbb{R}$, and $f \in \mathrm{Hol}(\mathbb{B}_d)$, then $f \in B^s_\omega$ if and only if $R^s f \in L^2_a$, the unweighted Bergman space. Thus, in this case the collection B^s_ω consists of standard weighted Bergman or Besov spaces. We have $B^{d/2}_{\mathbf{1}} = H^2_d$, the Drury–Arveson space, $B^{1/2}_{\mathbf{1}} = H^2(\partial\mathbb{B}_d)$, the Hardy space of the Ball, and for $s < 1/2$ we obtain the weighted Bergman spaces $B^s_{\mathbf{1}} = L^2_a((1-|z|^2)^{-2s}dV)$, where all equalities are understood to mean equality of spaces with equivalence of norms. These spaces have been extensively studied in the literature. We refer the reader to [25], where the L^p-analogues of these spaces were considered as well. If $d = 1$ and $s = 1$, then $B^1_{\mathbf{1}} = D$, the classical Dirichlet space of the unit disc. More generally, if $d = 1$ and $s > 1/2$, then these spaces are referred to as Dirichlet-type spaces, see [8].

If $\omega(z) = (1-|z|^2)^\alpha$ for some $\alpha > -1$, then ω is called a standard weight, and we obtain the same spaces as for $\omega_0 = 1$, but with a shift in indices: $B^s_\omega = B^{s-\frac{\alpha}{2}}_{\mathbf{1}}$. This can be verified by using polar coordinates and the asymptotics $\int_0^1 t^n(1-t)^\alpha dt = \frac{\Gamma(n+1)\Gamma(\alpha+1)}{\Gamma(n+\alpha+2)} \approx n^{-\alpha-1}$, which follows, e.g., from Stirling's formula. We refer the reader to Sect. 2 of the current paper for more detail on further calculations of this type.

Observe that for standard weights ω the spaces B^s_ω are weighted Bergman spaces for all $s \leq 0$. More generally, the following will be Theorem 2.4.

Theorem 1.1 *Let ω be a radial weight, let $s > 0$, and for $z \in \mathbb{B}_d$ define*

$$\omega_s(z) = \frac{1}{d}|z|^{2-2d} \int_{|w|\geq|z|} \frac{(|w|^2 - |z|^2)^{2s-1}}{\Gamma(2s)} \omega(w) dV(w).$$

Then ω_s is a weight and $B^t_\omega = B^{t+s}_{\omega_s}$ with equivalence of norms for all $t \in \mathbb{R}$. In particular, $L^2_a(\omega_s) = B^{-s}_\omega$ with equivalence of norms.

One checks that for all $s \leq 0$ and all radial weights ω, we have $\mathrm{Mult}(B^s_\omega) = H^\infty$. Here H^∞ denotes the bounded analytic functions on $\mathbb{B}_d$, and

$$\mathrm{Mult}(\mathcal{B}) = \{\varphi \in \mathrm{Hol}(\mathbb{B}_d) : \varphi f \in \mathcal{B} \text{ for all } f \in \mathcal{B}\}$$

denotes the multiplier algebra of $\mathcal{B}$.

In this paper we are interested in $\mathrm{Mult}(B^s_\omega)$ for $s > 0$. In general in those cases it turns out that $\mathrm{Mult}(B^s_\omega)$ is a proper subset of H^∞, but it is worthwhile to note that there are radial weights ω such that $B^s_\omega = L^2_a(\mu_s)$ for each $s \in \mathbb{R}$ for some weight μ_s and hence $\mathrm{Mult}(B^s_\omega) = H^\infty$ holds for all $s \in \mathbb{R}$. The weight $\omega(z) = e^{\frac{-1}{1-|z|^2}}$ is an example of a weight where this happens. Indeed, in this case for each positive integer N the function $(1-|z|^2)^{-4N}\omega(z)$ is also integrable, and in Example 4.9 we will show that $R^N f \in L^2_a(\omega)$ if and only if

$$\int_{\mathbb{B}_d} |f|^2(1-|z|^2)^{-4N}\omega dV < \infty.$$

By Theorem 1.1 this implies that B^s_ω is a weighted Bergman space for each $s \le N$, and since N was arbitrary it follows that the same is true for all $s \in \mathbb{R}$.

If $\mathcal{H} \subseteq \mathrm{Hol}(\mathbb{B}_d)$ is a Hilbert function space and if $\mathcal{E}$ is an auxiliary Hilbert space, then the identification of elementary tensors of the type $f \otimes x$, $f \in \mathcal{H}, x \in \mathcal{E}$ with $\mathcal{E}$-valued functions $f(\cdot)x$ extends to define a Hilbert space $\mathcal{H}(\mathcal{E})$ of $\mathcal{E}$-valued analytic functions on $\mathbb{B}_d$ that is isomorphic to $\mathcal{H} \otimes \mathcal{E}$. If $\mathcal{H}$ and $\mathcal{K}$ are two Hilbert spaces of analytic functions on $\mathbb{B}_d$ and if $\mathcal{E}$ and $\mathcal{F}$ are auxiliary Hilbert spaces, then $\mathrm{Mult}(\mathcal{H}(\mathcal{E}), \mathcal{K}(\mathcal{F}))$ will denote the multipliers from $\mathcal{H}(\mathcal{E})$ to $\mathcal{K}(\mathcal{F})$, i.e., those functions $\Phi : \mathbb{B}_d \to \mathcal{B}(\mathcal{E}, \mathcal{F})$ such that $F \to M_\Phi F, (M_\Phi F)(z) = \Phi(z)F(z)$ defines a bounded linear transformation from $\mathcal{H}(\mathcal{E})$ to $\mathcal{K}(\mathcal{F})$. We will write $\mathrm{Mult}(\mathcal{H}, \mathcal{K}) = \mathrm{Mult}(\mathcal{H}(\mathbb{C}), \mathcal{K}(\mathbb{C}))$ for the scalar-valued multipliers.

In the paper [4], an important role was played by the multiplier inclusion condition. For a weighted Besov space B^N_ω, where $N \in \mathbb{N}$, this condition means that

$$\mathrm{Mult}(B^N_\omega, B^N_\omega(\ell^2)) \subseteq \mathrm{Mult}(B^{N-1}_\omega, B^{N-1}_\omega(\ell^2)) \subseteq \cdots \subseteq \mathrm{Mult}(B^0_\omega, B^0_\omega(\ell^2))$$

with continuous inclusions. We established this condition for the Drury–Arveson space and a few other standard weighted Besov spaces using an elementary method. It is also possible to use the complex method of interpolation to establish inclusions of multiplier spaces. Indeed, if $s, t, \alpha \in \mathbb{R}$ with $s \le t$ and $\alpha \ge 0$, then it is shown in [9] that for Bekollé–Bonami weights ω one has

$$\mathrm{Mult}(B^{t+\alpha}_\omega, B^{s+\alpha}_\omega) \subseteq \mathrm{Mult}(B^t_\omega, B^s_\omega).$$

Note that for Bekollé–Bonami weights that are not necessarily radial the following definition is used for the weighted Besov space:

$$B^s_\omega = \{f \in \mathrm{Hol}(\mathbb{B}_d) : R^N f \in L^2_a((1-|z|^2)^{2(N-s)}\omega(z))\},$$

where N is any non-negative integer $\ge s$. For radial weights satisfying a Bekollé–Bonami condition this coincides with the definition used here since in that case $(1-|z|^2)^{2(N-s)}\omega \approx \omega_{N-s}$, see, e.g., Lemmas 4.2 and 4.7.

In this paper, we use a third method to establish a general result about inclusions of multiplier spaces of unitarily invariant Hilbert function spaces on $\mathbb{B}_d$, using the fact that multiplication operators are triangular with respect to the common orthogonal basis of monomials. In particular, we obtain the following theorem, which shows that the multiplier inclusion condition holds whenever ω is a radial weight. It is proved in Corollary 3.8 (also see Corollary 3.4).

Theorem 1.2 *Let ω and ν be radial weights in $\mathbb{B}_d$ and let $s, t, s', t' \in \mathbb{R}$ with $t \leq s$ and $t' - s' \leq t - s$. Then for any pair $\mathcal{E}, \mathcal{F}$ of separable Hilbert spaces,*

$$\mathrm{Mult}(B^s_\omega(\mathcal{E}), B^{s'}_\nu(\mathcal{F})) \subseteq \mathrm{Mult}(B^t_\omega(\mathcal{E}), B^{t'}_\nu(\mathcal{F}))$$

and the inclusion is contractive.

Given a sequence $\Phi = \{\varphi_1, \varphi_2, \ldots\} \subseteq \mathrm{Mult}(\mathcal{H}, \mathcal{K})$ of multipliers, we can consider the column operator $\Phi^C : h \to (\varphi_1 h, \varphi_2 h, \ldots)^T$ and the row operator $\Phi^R : (h_1, h_2, \ldots)^T \to \sum_{i \geq 1} \varphi_i h_i$. Here for ease of writing we have used $(h_1, \ldots)^T$ to denote the transpose of a row vector. We write $M^C(\mathcal{H}, \mathcal{K})$ for the set of those sequences Φ whose column operator Φ^C is bounded, that is, $\Phi^C \in \mathrm{Mult}(\mathcal{H}, \mathcal{K}(\ell_2))$. Similarly, let $M^R(\mathcal{H}, \mathcal{K})$ denote all sequences Φ for which the row operator is bounded, i.e., $\Phi^R \in \mathrm{Mult}(\mathcal{H}(\ell_2), \mathcal{K})$. We will abbreviate the notations to $M^R(\mathcal{H})$ and $M^C(\mathcal{H})$, if $\mathcal{H} = \mathcal{K}$.

Trent showed that for the Dirichlet space D of the unit disc $\mathbb{D} \subseteq \mathbb{C}$ one has the continuous inclusion $M^C(D) \subseteq M^R(D)$ and the norm of the inclusion is at most $\sqrt{18}$, see Lemma 1 of [24]. The results in [4] establish that $M^C(\mathcal{H}) \subseteq M^R(\mathcal{H})$ for certain standard weighted Besov spaces $\mathcal{H}$ including the Drury–Arveson space. Using Theorem 1.2, we now obtain a more general result, which is Theorem 3.9.

Theorem 1.3 *Let ω and ν be radial weights in $\mathbb{B}_d$, and let $s, t \in \mathbb{R}$. Then*

$$M^C(B^s_\omega, B^t_\nu) \subseteq M^R(B^s_\omega, B^t_\nu)$$

and the inclusion is continuous.

It is known and easy to verify that $M^C(L^2_a(\omega)) = M^R(L^2_a(\omega)) = H^\infty(\ell_2)$, where

$$H^\infty(\ell_2) = \{(\varphi_1, \varphi_2, \ldots) : \varphi_j \in H^\infty \text{ and } \sup_{z \in \mathbb{B}_d} \sum_j |\varphi_j(z)|^2 < \infty\}.$$

One application of Theorem 1.3 is to provide another proof of the characterization of interpolating sequences established in [3] in the case of radially weighted Besov spaces with the complete Pick property. The proof in [3] uses the Marcus–Spielman–Srivastava theorem [15], but as explained in Remark 3.7 in [3], this theorem can be avoided for spaces $\mathcal{H}$ with the property that $M^C(\mathcal{H}) \subseteq M^R(\mathcal{H})$.

A Hilbert function space $\mathcal{H}$ is a Hilbert space of complex-valued functions on a set X such that point evaluations for points in X define continuous linear functionals on $\mathcal{H}$. Every Hilbert function space $\mathcal{H}$ has a reproducing kernel, i.e., a function $k : X \times X \to \mathbb{C}$ such that $f(w) = \langle f, k_w \rangle$ for all $w \in X$, where $k_w(z) = k(z, w)$. We say that k is normalized, if there is a $z_0 \in X$ such that $k_{z_0} = 1$.

By a normalized complete Pick kernel we mean a normalized reproducing kernel of the form $k_w(z) = \frac{1}{1-u_w(z)}$, where $u_w(z)$ is positive definite, i.e., whenever $n \in \mathbb{N}$, $z_1, \dots, z_n \in X$, and $a_1, \dots, a_n \in \mathbb{C}$ we have $\sum_{i,j} a_i \overline{a}_j u_{z_j}(z_i) \geq 0$. (Normally complete Pick kernels are defined intrinsically, but by the McCullough–Quiggin theorem they are precisely of this form. See [1, 16, 20].)

An important example of such a complete Pick kernel is the Szegő kernel $k_w(z) = (1 - \overline{w}z)^{-1}$. It is the reproducing kernel for the Hardy space H^2 of the unit disc $\mathbb{D}$. Many properties of the Hardy space carry over to other spaces with complete Pick kernels—see [5] for some examples. We will say that a Hilbert function space $\mathcal{H}$ is a complete Pick space, if there is an equivalent norm on the space such that the reproducing kernel for that norm is a normalized complete Pick kernel. In [5] it is proven that complete Pick spaces $\mathcal{H}$ are contained in the Smirnov class $N^+(\mathcal{H})$ associated with $\mathcal{H}$, where

$$N^+(\mathcal{H}) = \{f = \frac{\varphi}{\psi} : \varphi, \psi \in \text{Mult}(\mathcal{H}), \psi \text{ cyclic in } \mathcal{H}\}$$

and a multiplier ψ is called cyclic if $\psi\mathcal{H}$ is dense in $\mathcal{H}$.

It is known that for all $s \geq d/2$ the spaces $B^s_{\mathbf{1}}$ are complete Pick spaces (this can be seen as in Corollary 7.41 of [2]). In particular, for $d/2 \leq s < (d+1)/2$ the space $B^s_{\mathbf{1}}$ has reproducing kernel $\frac{1}{(1-\langle z,w\rangle)^{d+1-2s}}$ (up to equivalence of norms), which can be seen to be a complete Pick kernel by consideration of the binomial series coefficients. On the other hand, if $s < d/2$, then $B^s_{\mathbf{1}}$ is not a complete Pick space, because $B^s_{\mathbf{1}} \not\subseteq N^+(B^s_{\mathbf{1}})$. Indeed, in this case $B^s_{\mathbf{1}}$ has a reproducing kernel of the type $\frac{1}{(1-\langle z,w\rangle)^{\gamma}}$ for some $\gamma > 1$. If $d = 1$, then $B^s_{\mathbf{1}}$ is a weighted Bergman space, which will contain functions that are not in the Nevanlinna class, and hence cannot be ratios of multipliers. The same is true if $d > 1$. In that case the $d = 1$ result implies that there are functions of the form $f(z_1, 0, \dots, 0)$ in $B^s_{\mathbf{1}}$ that are not the ratio of two bounded functions.

An observation that was shared years ago with us by Serguey Shimorin is that if the Cauchy dual of a space of functions in the unit disc is a weighted Bergman space, then the original space is a complete Pick space. An analogue of this observation holds for functions in $\mathbb{B}_d$; for radially symmetric spaces, this is Lemma 5.1. For many radially symmetric weighted Besov spaces, this result leads to a condition which is easy to check.

Theorem 1.4 *Let $\alpha > -1$, $0 \leq r_0 < 1$, and let ω be a radial weight such that $\frac{\omega(z)}{(1-|z|^2)^{\alpha}}$ is nondecreasing in $|z|$ for $r_0 < |z| < 1$. Then B^s_{ω} is a complete Pick space for all $s \geq \frac{\alpha+d}{2}$.*

This will follow from Theorem 5.2, which holds for weights that satisfy a related, but weaker condition.

If $\mathcal{H}$ and $\mathcal{K}$ are two Hilbert function spaces on the same set, then we will write $\mathcal{H} = \mathcal{K}$ to mean that $\mathcal{H}$ and $\mathcal{K}$ agree as vector spaces and their norms are equivalent, but not necessarily equal. If $\| \cdot \|_1$ and $\| \cdot \|_2$ are two norms, then we will write $\|f\|_1 \approx \|f\|_2$ to denote that the norms are equivalent. Similarly, if $a_n, b_n \geq 0$, then $a_n \approx b_n$ will mean that there are constants $c, C > 0$ such that $ca_n \leq b_n \leq Ca_n$ holds for all $n \in \mathbb{N}$.

The remainder of this paper is organized as follows. In Sect. 2, we collect basic facts about radially weighted Besov spaces and then prove Theorem 1.1. In Sect. 3, we prove several results about inclusions of multiplier algebras and of multiplier spaces. In particular, we show Theorems 1.2 and 1.3. Section 4 is devoted to the study of several finer properties of weights. In particular, we introduce weakly normal weights, which will be important in the proof of Theorem 1.4. Section 5 then contains the proof of Theorem 1.4. In the final Sect. 6, we use the methods developed in this paper to establish some additional properties of multipliers of weighted Besov spaces.

2 Radially Weighted Besov Spaces and Index Shifts

2.1 Basics About Radially Weighted Besov Spaces

Let ω be a radial weight on $\mathbb{B}_d$. We will temporarily write $u_\omega(r) = \omega(r, 0, \dots, 0)$ if $r \in (0, 1)$. Let σ be Lebesgue measure on $\partial\mathbb{B}_d$, normalized so that $\sigma(\partial\mathbb{B}_d) = 1$. Then for any non-negative measurable function h on $\mathbb{B}_d$ we have the change of variables

$$\int_{\mathbb{B}_d} h\omega dV = \int_0^1 \left(\int_{\partial\mathbb{B}_d} h(rw) d\sigma(w) \right) u_\omega(r) 2d r^{2d-1} dr.$$

In particular, if $f \in \mathrm{Hol}(\mathbb{B}_d)$ with homogeneous expansion $f = \sum_{n=0}^\infty f_n$, then

$$\int_{\mathbb{B}_d} |f|^2 \omega dV = \sum_{n=0}^\infty a_n(\omega) \|f_n\|^2_{H^2(\partial\mathbb{B}_d)}, \tag{2.1}$$

where

$$a_n(\omega) = 2d \int_0^1 r^{2n+2d-1} u_\omega(r) dr = \int_0^1 t^n v(t) dt.$$

Here we used $v(t)$ is the product $d \cdot t^{d-1} \cdot u_\omega(\sqrt{t})$ and note that $v \in L^1[0, 1]$. It is clear that this process can be reversed and any positive $L^1[0, 1]$-function v can be used as above to associate a function ω on $\mathbb{B}_d$. The non-degeneracy condition (1.1) is equivalent to

$$\int_t^1 v(x)dx > 0 \qquad \text{for all } t < 1. \tag{2.2}$$

We will say that a non-negative function $v \in L^1[0, 1]$ is a weight if (2.2) holds.

The following elementary lemma about moments of weights will be useful in several places.

Lemma 2.1 *Let* $v, w \in L^1[0, 1]$ *be two non-negative weights such that* $\lim_{t \nearrow 1} \frac{v(t)}{w(t)} = 1$ *(with the convention* $0/0 = 1$*). Then*

$$\lim_{n\to\infty} \frac{\int_0^1 t^n v(t)\, dt}{\int_0^1 t^n w(t)\, dt} = 1.$$

Proof By symmetry, it suffices to show that

$$\limsup_{n\to\infty} \frac{\int_0^1 t^n v(t)\, dt}{\int_0^1 t^n w(t)\, dt} \le 1.$$

To this end, let $r \in (0, 1)$ be such that $\frac{v(t)}{w(t)}$ is finite for $t \in [r, 1]$. Then

$$\int_0^1 t^n w(t)\, dt \ge \int_r^1 t^n w(t)\, dt \ge r^{n/2} \int_{\sqrt{r}}^1 w(t)\, dt,$$

where the last quantity is strictly positive by (2.2). Moreover,

$$\begin{aligned} \int_0^1 t^n v(t)\, dt &= \int_0^r t^n v(t)\, dt + \int_r^1 t^n v(t)\, dt \\ &\le r^n \int_0^1 v(t)\, dt + \sup_{x\in[r,1]} \frac{v(x)}{w(x)} \int_r^1 t^n w(t)\, dt. \end{aligned}$$

Therefore,

$$\frac{\int_0^1 t^n v(t)\, dt}{\int_0^1 t^n w(t)\, dt} \le \sup_{x\in[r,1]} \frac{v(x)}{w(x)} + r^{n/2} \frac{\int_0^1 v(t)\, dt}{\int_{\sqrt{r}}^1 w(t)\, dt},$$

so that

$$\limsup_{n\to\infty} \frac{\int_0^1 t^n v(t)\, dt}{\int_0^1 t^n w(t)\, dt} \leq \sup_{x\in[r,1]} \frac{v(x)}{w(x)}.$$

This is true for all r sufficiently close to 1. The result now follows by taking the limit $r \nearrow 1$. ■

Let now ω be a radial weight in $\mathbb{B}_d$. We will use the moments $a_n(\omega) = \int_0^1 t^n v(t)dt$ to express the norm of B^s_ω. For $f \in \text{Hol}(\mathbb{B}_d)$ we will continue to write $f = \sum f_n$ for its expansion into a sum of homogeneous polynomials.

Let $s \in \mathbb{R}$. Comparison of (1.2) and (2.1) shows that

$$\|f\|^2_{B^s_\omega} = a_0(\omega)|f(0)|^2 + \sum_{n=1}^{\infty} n^{2s} a_n(\omega)\|f_n\|^2_{H^2(\partial\mathbb{B}_d)}. \tag{2.3}$$

Since $R^N f = \sum_{n\geq 0} n^N f_n$ it is clear that for each $s \in \mathbb{R}$ we have $f \in B^s_\omega$, if and only if $R^N f \in B^{s-N}_\omega$.

We also remark that the reproducing kernel of B^s_ω is of the form

$$k_w(z) = \sum_{n=0}^{\infty} b_n \langle z, w\rangle^n,$$

where for $n \geq 1$

$$b_n = ||z_1^n||^{-2}_{B^s_\omega} = n^{-2s} a_n(\omega)^{-1} \|z_1^n\|^{-2}_{H^2(\partial\mathbb{B}_d)} \approx n^{-2s+d-1} a_n(\omega)^{-1}.$$

It follows from Lemma 2.1 that $\frac{\int_0^1 t^{n+1} v(t)dt}{\int_0^1 t^n v(t)dt} \to 1$ as $n \to \infty$ for any weight $v \in L^1[0, 1]$. Hence, $\lim_{n\to 1} b_n/b_{n+1} = 1$. This condition is frequently useful in operator theoretic contexts. For instance, it implies that the tuple $(M_{z_1}, \ldots, M_{z_d})$ of multiplication operators by the coordinate functions is essentially normal and has essential Taylor spectrum $\partial\mathbb{B}_d$, see Theorem 4.5 of [12].

2.2 Index Shift

Recall from the Introduction that $B^s_{\mathbf{1}} = B^{s+\frac{\alpha}{2}}_{\omega_\alpha}$ for all $s \in \mathbb{R}$ and $\alpha > -1$, where $\omega_\alpha(z) = (1 - |z|^2)^\alpha$ is a standard weight. We now introduce a generalization of this procedure which will allow us to shift the index s of the space B^s_ω for more general radial weights ω.

We saw in Sect. 2.1 that by a change to polar coordinates any radial weight ω on $\mathbb{B}_d$ is associated with a non-negative function $v \in L^1[0,1]$. More generally, let μ be a finite Borel measure on $[0,1]$. For $x > 0$ consider

$$\int_0^1 \int_{[t,1]} (s-t)^{x-1} d\mu(s)dt = \int_{[0,1]} \int_0^s (s-t)^{x-1} dt d\mu(s) = \int_{[0,1]} \frac{s^x}{x} d\mu(s) < \infty.$$

Thus, for all $x > 0$ we can define a non-negative $L^1[0,1]$-function v_x by

$$v_x(t) = \int_{[t,1]} \frac{(s-t)^{x-1}}{\Gamma(x)} d\mu(s), t \in [0,1).$$

Here $\Gamma(x)$ denotes the Gamma function. It is easy to check that the functions v_x obey the semigroup law $(v_x)_y = v_{x+y}$ for all $x, y > 0$. We also remark that if $v_1(t) > 0$ for all $r \in (0,1)$, then v_x satisfies (2.2) for all $x > 0$.

The following lemma will be used repeatedly. It will allow us to perform the desired index shift for B^s_ω (see Theorem 2.4 below).

Lemma 2.2 *Let* μ *be a finite positive Borel measure on* $[0,1]$, *for* $x > 0$ *let* v_x *be the function associated with* μ *as above, and assume that* $v_1(t) > 0$ *for all* $0 \le t < 1$.

Then for each $x > 0$ *we have*

$$\lim_{n\to\infty} \frac{n^x \int_0^1 t^n v_x(t)dt}{\int_{[0,1]} t^n d\mu} = 1.$$

Proof We start with the observation that for any integer $n > 0$ we have $\int_0^1 t^{n-1} (\log(1/t))^{x-1} dt = n^{-x}\Gamma(x)$. This can easily be verified with the substitution $t = e^{-\frac{u}{n}}$ (see [14], p. 56). Next we define the auxiliary function

$$v_x^*(t) = \int_{[t,1]} \frac{\left(\log \frac{s}{t}\right)^{x-1}}{\Gamma(x)} d\mu(s).$$

An application of Fubini's theorem and the earlier observation shows that

$$n^x \int_0^1 t^{n-1} v_x^*(t)dt = \int_{[0,1]} t^n d\mu(t), \quad n = 1, 2, \ldots$$

So in order to prove the Lemma, it suffices to show that

$$\lim_{n\to\infty} \frac{\int_0^1 t^{n-1} v_x^*(t)dt}{\int_0^1 t^n v_x(t)dt} = 1.$$

Since $v_1(t) > 0$ for all $t \in (0, 1)$, the weights v_x and v_x^* satisfy (2.2), so the last statement follows from Lemma 2.1 and the observation that $\lim_{t \nearrow 1} \frac{v_x(t)}{v_x^*(t)} = 1$ by elementary properties of the natural logarithm. ■

We will now again restrict attention to absolutely continuous measures $d\mu = v(t)dt$. In this case, it makes sense to define $v_0(t) = v(t)$. We also write

$$\widehat{v}(t) = v_1(t) = \int_t^1 v(x)dx.$$

Note that in this case $v_{x+1}(t) = \int_t^1 v_x(s)ds = \hat{v}_x(t)$ is valid for all $x \geq 0$, and thus the functions v_x get smoother as x increases. They also decay faster near 1. The estimate in the following lemma is obvious.

Lemma 2.3 *If $v \in L^1[0, 1]$ is positive, and v_x is as above, then for all $x, \alpha > 0$ we have $v_{x+\alpha}(t) \leq \frac{\Gamma(x)}{\Gamma(x+\alpha)}(1-t)^\alpha v_x(t)$ for all $t \in [0, 1)$.*

We now investigate this procedure on the level of radial weights in the ball. Let ω be a radial weight in $\mathbb{B}_d$. For each $x \geq 0$ we define a radial weight ω_x by

$$\omega_x(z) = \frac{1}{d}|z|^{2-2d} \int_{|w| \geq |z|} \frac{(|w|^2 - |z|^2)^{2x-1}}{\Gamma(2x)} \omega(w)dV(w).$$

Then ω_x is the radial weight that corresponds to the $L^1[0, 1]$-function v_{2x} that is associated with v as in Lemma 2.2.

Theorem 2.4 *Let ω be a radial weight and let $x \geq 0$.*

Then ω_x is a weight,

$$\|f\|^2_{B_\omega^{-x}} \approx \int_{\mathbb{B}_d} |f|^2 \omega_x dV,$$

and for each $s \in \mathbb{R}$ we have $B^s_\omega = B^{s+x}_{\omega_x}$ with equivalence of norms.

Proof Since ω is a radial weight, so is ω_x. Lemma 2.2 implies that $n^{2x} a_n(\omega_x) \approx a_n(\omega)$ as $n \to \infty$. Now the Theorem follows from (2.3). ■

For later reference we note that Lemma 2.3 applies and we conclude that for all $x > 0$ and $\alpha \geq 0$

$$\frac{\omega_{x+\alpha}(z)}{(1-|z|^2)^{2\alpha}} \leq \frac{\Gamma(2x)}{\Gamma(2x+2\alpha)} \omega_x(z) \text{ for all } z \in \mathbb{B}_d. \tag{2.4}$$

3 Multiplier Inclusions

3.1 Inclusion of Multiplier Algebras

Let ω be a radial weight in $\mathbb{B}_d$ and let $N \in \mathbb{N}$. A crucial condition in [4] is the multiplier inclusion condition for B^N_ω, which demands that

$$\mathrm{Mult}(B^N_\omega, B^N_\omega(\ell^2)) \subseteq \mathrm{Mult}(B^{N-1}_\omega, B^{N-1}_\omega(\ell^2)) \subseteq \cdots \subseteq \mathrm{Mult}(B^0_\omega, B^0_\omega(\ell^2)) \tag{3.1}$$

with continuous inclusions. In this section we will show that all weighted Besov spaces defined by radial weights satisfy this multiplier inclusion condition. In fact, we will prove a more general result about inclusion of the multipliers between spaces of analytic functions on the unit ball with unitarily invariant kernels.

We first recall a few notions from the theory of operator spaces. Let $\mathcal{H}$ be a Hilbert space and let $\mathcal{M} \subseteq \mathcal{B}(\mathcal{H})$ be a subspace. For $n \in \mathbb{N}$, let $M_n(\mathcal{M})$ denote the space of all $n \times n$ matrices with entries in $\mathcal{M}$. The natural identification of $M_n(\mathcal{B}(\mathcal{H}))$ with $\mathcal{B}(\mathcal{H}^n)$ allows us to endow each space $M_n(\mathcal{M})$ with a norm. Suppose now that $\mathcal{K}$ is another Hilbert space and that $\Phi : \mathcal{M} \to \mathcal{B}(\mathcal{K})$ is a linear map. Then for each $n \in \mathbb{N}$, we obtain an induced linear map

$$\Phi^{(n)} : M_n(\mathcal{M}) \to M_n(\mathcal{B}(\mathcal{K})), \quad [m_{ij}] \mapsto [\Phi(m_{ij})].$$

In this setting, we say that Φ is completely contractive if each map $\Phi^{(n)}$ is contractive.

In Sect. 3.2 we will see that this notion has a natural analogue for operators between possibly different Hilbert spaces, and then we will mostly be interested in the case when $\mathcal{M} = \mathrm{Mult}(\mathcal{H}, \mathcal{K})$ for Hilbert function spaces $\mathcal{H}$ and $\mathcal{K}$. In this case, $M_n(\mathrm{Mult}(\mathcal{H}, \mathcal{K}))$ can be identified with $\mathrm{Mult}(\mathcal{H}(\mathbb{C}^n), \mathcal{K}(\mathbb{C}^n))$, so this approach allows us to deal with operator-valued multipliers.

We begin with the following result, which is essentially due to Kacnelson [13], see also [11, Theorem 2.1]. For completeness, we provide a proof. If $\mathcal{H}$ is a Hilbert space with an orthogonal basis (e_n), let $\mathcal{T}(\mathcal{K})$ denote the algebra of all bounded lower triangular operators on $\mathcal{K}$ with respect to (e_n).

Lemma 3.1 (Kacnelson) *Let $\mathcal{H}$ be a Hilbert space with orthonormal basis (e_n), let (d_n) be a nonincreasing sequence of strictly positive numbers and let D denote the diagonal operator on $\mathcal{H}$ with diagonal (d_n), and let D^{-1} be its possibly unbounded inverse. Then for every $T \in \mathcal{T}(\mathcal{H})$, the densely defined operator DTD^{-1} is bounded and the homomorphism*

$$\mathcal{T}(\mathcal{H}) \to \mathcal{T}(\mathcal{H}), \quad T \mapsto DTD^{-1},$$

is completely contractive.

Proof If P_n denotes the orthogonal projection onto the linear span of $e_0, \dots, e_n$, then P_n commutes with every diagonal operator. Thus, a straightforward approximation argument shows that it suffices to prove the following assertion: For every $n \in \mathbb{N}$ and every nonincreasing sequence of strictly positive numbers $d_0, \dots, d_n$, the map

$$\Phi : \mathcal{T}_{n+1} \mapsto \mathcal{T}_{n+1}, \quad T \mapsto \operatorname{diag}(d_0, \dots, d_n) T \operatorname{diag}(d_0, \dots, d_n)^{-1},$$

is completely contractive. Here, $\mathcal{T}_{n+1}$ denotes the algebra of all lower triangular $(n+1) \times (n+1)$ matrices, and $\operatorname{diag}(d_0, \dots, d_n)$ is the diagonal matrix with diagonal $d_0, \dots, d_n$.

To this end, let $d_0, \dots, d_n$ be nonincreasing strictly positive numbers. By multiplying the sequence $d_0, \dots, d_n$ with d_0^{-1}, we may assume that $d_0 = 1$. For $j \geq 1$, let $\alpha_j = d_j / d_{j-1}$ and $\alpha = (\alpha_1, \dots, \alpha_n)$. Then $d_j = \alpha_1 \dots \alpha_j$ for $j \geq 1$ and $\alpha_j \in (0, 1]$ by assumption.

We will use the maximum modulus principle to show that the map Φ is completely contractive. For $z = (z_1, \dots, z_n) \in (\mathbb{C} \setminus \{0\})^n$, define

$$D(z) = \operatorname{diag}(1, z_1, z_1 z_2, \dots, z_1 z_2 \dots z_n).$$

In particular, $D(\alpha) = \operatorname{diag}(d_0, \dots, d_n)$. If $T = [t_{ij}] \in \mathcal{T}_{n+1}$ and $i \geq j$, then the (i, j)-entry of $D(z) T D(z)^{-1}$ is given by

$$z_1 z_2 \dots z_i t_{ij} z_1^{-1} z_2^{-1} \dots z_j^{-1} = t_{ij} z_{j+1} \dots z_i.$$

Since T is lower triangular, we therefore conclude that the map $z \mapsto D(z) T D(z)^{-1}$ extends to an analytic M_{n+1}-valued map on $\mathbb{C}^n$.

Let $[T_{ij}] \in M_r(\mathcal{T}_{n+1})$. By the maximum modulus principle,

$$||[\Phi(T_{ij})]|| = ||[D(\alpha) T_{ij} D(\alpha)^{-1}]|| \leq \sup_{z \in \mathbb{T}^n} ||[D(z) T_{ij} D(z)^{-1}]||.$$

But if $z \in \mathbb{T}^n$, then $D(z)$ is unitary, hence

$$||[D(z) T_{ij} D(z)^{-1}]|| = ||(D(z) \otimes I_r)[T_{ij}](D(z) \otimes I_r)^{-1}|| = ||[T_{ij}||,$$

which finishes the proof. ■

The following corollary is merely a reformulation of Lemma 3.1.

Corollary 3.2 *Let $\mathcal{K}$ be a Hilbert space with an orthonormal basis (e_n). Suppose that $\mathcal{H}$ is another Hilbert space such that $\mathcal{H} \subseteq \mathcal{K}$ as vector spaces, such that (e_n) is an orthogonal basis for $\mathcal{H}$ and such that the sequence $(||e_n||_{\mathcal{H}})$ is nondecreasing. Then $\mathcal{T}(\mathcal{H}) \subseteq \mathcal{T}(\mathcal{K})$, and the inclusion is a complete contraction.*

Proof Observe that every operator in $\mathcal{T}(\mathcal{H})$ is at least densely defined on $\mathcal{K}$. Let D be the diagonal operator on $\mathcal{H}$ with diagonal $(||e_n||_{\mathcal{H}}^{-1})$. Then D extends to a unitary operator $\mathcal{K} \to \mathcal{H}$. Thus, if $[T_{ij}] \in M_r(\mathcal{T}(\mathcal{H}))$, then by Lemma 3.1,

$$||[T_{ij}]||_{\mathcal{B}(\mathcal{K}^r)} = ||[DT_{ij}D^{-1}]||_{\mathcal{B}(\mathcal{H}^r)} \leq ||[T_{ij}]||_{\mathcal{B}(\mathcal{H}^r)}.$$

This shows that $\mathcal{T}(\mathcal{H}) \subseteq \mathcal{T}(\mathcal{K})$ completely contractively. ∎

Let $\mathcal{H}$ be a reproducing kernel Hilbert space on $\mathbb{D}$ with a reproducing kernel of the form

$$k_w(z) = \sum_{n=0}^{\infty} a_n z \overline{w}^n,$$

where $a_n > 0$ for all $n \in \mathbb{N}_0$. Then

$$||z||^2_{\mathrm{Mult}(\mathcal{H})} = \sup_{n \in \mathbb{N}_0} \frac{a_n}{a_{n+1}}.$$

This motivates the condition in the following result.

Proposition 3.3 *Let $\mathcal{H}$ and $\mathcal{K}$ be two reproducing kernel Hilbert spaces on $\mathbb{B}_d$, $d \in \mathbb{N}$, with reproducing kernels $k_w(z) = \sum_{n=0}^{\infty} a_n \langle z, w\rangle^n$ and $\ell_w(z) = \sum_{n=0}^{\infty} b_n \langle z, w\rangle^n$, respectively. Assume that $a_n, b_n > 0$ for all $n \in \mathbb{N}_0$. If*

$$\frac{b_n}{b_{n+1}} \leq \frac{a_n}{a_{n+1}} \quad \textit{for all } n \in \mathbb{N}_0,$$

then $\mathrm{Mult}(\mathcal{H}) \subseteq \mathrm{Mult}(\mathcal{K})$*, and the inclusion is a complete contraction.*

Proof Observe that $\mathcal{H}$ and $\mathcal{K}$ each have orthonormal bases consisting of monomials. If we order the monomials such that their degrees are nondecreasing, then every multiplication operator on $\mathcal{H}$ is lower triangular with respect to such an orthonormal basis. Moreover, if p is a monomial of degree n with $||p||_{\mathcal{K}} = 1$, then

$$||p||_{\mathcal{H}} = \sqrt{\frac{b_n}{a_n}}.$$

The assumption implies that the sequence $\sqrt{b_n/a_n}$ is nondecreasing. In particular, there exists a constant $C > 0$ such that $a_n \leq C b_n$, so that $\mathcal{H}$ is densely contained in $\mathcal{K}$ and every multiplication operator on $\mathcal{H}$ is at least densely defined on $\mathcal{K}$. An application of Corollary 3.2 now shows that every multiplication operator on $\mathcal{H}$ is bounded on $\mathcal{K}$, and hence a bounded multiplication operator, and that the inclusion $\mathrm{Mult}(\mathcal{H}) \subseteq \mathrm{Mult}(\mathcal{K})$ is a complete contraction. ∎

We obtain the following consequence for multiplier algebras of weighted Besov spaces.

Corollary 3.4 *Let ω be a radial weight in $\mathbb{B}_d$ and let $s, t \in \mathbb{R}$ with $t \le s$. Then*

$$\mathrm{Mult}(B^s_\omega) \subseteq \mathrm{Mult}(B^t_\omega)$$

and the inclusion is a complete contraction. In particular,

$$\mathrm{Mult}(B^s_\omega, B^s_\omega(\ell_2)) \subseteq \mathrm{Mult}(B^t_\omega, B^t_\omega(\ell_2))$$

and the inclusion is a contraction.

In particular, by taking $s = n$ and $t = n - 1$ for $n = 1, 2.., N$ we see that any weighted Besov space $\mathcal{H} = B^N_\omega$ associated with a radial weight satisfies the multiplier inclusion condition (3.1).

Proof We saw in Sect. 2 that B^s_ω and B^t_ω have reproducing kernels of the form

$$k_w(z) = \sum_{n=0}^{\infty} a_n \langle z, w\rangle^n \quad \text{and} \quad \ell_w(z) = \sum_{n=0}^{\infty} b_n \langle z, w\rangle^n,$$

respectively, where $a_n = ||z_1^n||^{-2}_{B^s_\omega}$ and $b_n = ||z_1^n||^{-2}_{B^t_\omega}$. From Eq. (2.3), we deduce that for $n \ge 1$,

$$\frac{a_n}{b_n} = n^{2(t-s)}$$

and $a_0/b_0 = 1$. Since $t \le s$, the sequence (a_n/b_n) is nonincreasing, so that the result is a special case of Proposition 3.3. ∎

It was shown in [4, Theorem 1.5] that the multiplier inclusion condition (3.1) for B^N_ω implies that every bounded column multiplication operator on B^N_ω is also a bounded row multiplication operator. Moreover, by Theorem 2.4, each Besov space B^s_ω can also be regarded as a space of the form $B^N_{\widetilde{\omega}}$ for a suitable radial weight $\widetilde{\omega}$ and $N \in \mathbb{N}$. Thus, we obtain the following consequence.

Corollary 3.5 *Let ω be a radial weight in $\mathbb{B}_d$ and let $s \in \mathbb{R}$. Then*

$$M^C(B^s_\omega) \subseteq M^R(B^s_\omega)$$

and the inclusion is continuous.

We do not know if the inclusion in the preceding corollary is contractive, even in the case of the Drury–Arveson space. Even though Corollary 3.4 shows that the multiplier inclusion condition (3.1) holds with contractive inclusions, [4, Theorem 1.5] only yields boundedness of the inclusion $M^C(B^N_\omega) \subseteq M^R(B^N_\omega)$.

3.2 Inclusion of Multiplier Spaces

We also require a version of the preceding result for multipliers between different spaces. Thus, we seek conditions that imply inclusions of the form $\mathrm{Mult}(\mathcal{H}, \mathcal{H}') \subseteq \mathrm{Mult}(\mathcal{K}, \mathcal{K}')$. The proofs based on Kacnelson's lemma (Lemma 3.1) generalize to this setting. The results in this subsection contain the results of the preceding subsection as a special case. But for the sake of readability, we chose to treat inclusions of multiplier algebras first.

We begin with a version of Corollary 3.2 for four Hilbert spaces. First of all, observe that if $\mathcal{H}$ and $\mathcal{H}'$ are Hilbert space, then $\mathcal{B}(\mathcal{H}, \mathcal{H}')$ can be identified with a subspace of $\mathcal{B}(\mathcal{H} \oplus \mathcal{H}')$, hence the notion of a completely contractive map applies in this setting as well. Equivalently, $M_r(\mathcal{B}(\mathcal{H}, \mathcal{H}'))$ is normed by means of the identification with $\mathcal{B}(\mathcal{H}^r, (\mathcal{H}')^r)$. If $\mathcal{H}$ and $\mathcal{H}'$ are Hilbert spaces with orthogonal bases (e_n) and (e_n'), respectively, let $\mathcal{T}(\mathcal{H}, \mathcal{H}') \subseteq \mathcal{B}(\mathcal{H}, \mathcal{H}')$ denote the space of all operators that are lower triangular with respect to (e_n) and (e_n'). Thus, an operator $T \in \mathcal{B}(\mathcal{H}, \mathcal{H}')$ belongs to $\mathcal{T}(\mathcal{H}, \mathcal{H}')$ if and only if

$$\langle T e_i, e_j' \rangle = 0 \quad \text{whenever } j > i.$$

Corollary 3.6 *Let $\mathcal{K}$ and $\mathcal{K}'$ be Hilbert spaces with orthonormal bases (e_n) and (e_n'), respectively. Let $\mathcal{H}$ and $\mathcal{H}'$ be another pair of Hilbert spaces such that*

- *$\mathcal{H} \subseteq \mathcal{K}$ and $\mathcal{H}' \subseteq \mathcal{K}'$ as vector spaces,*
- *(e_n) is an orthogonal basis for $\mathcal{H}$ and (e_n') is an orthogonal basis for $\mathcal{H}'$,*
- *the sequence $(||e_n||_{\mathcal{H}})$ is nondecreasing, and*
- *$||e_n||_{\mathcal{H}} \le ||e_n'||_{\mathcal{H}'}$ for all $n \in \mathbb{N}$.*

Then $\mathcal{T}(\mathcal{H}, \mathcal{H}') \subseteq \mathcal{T}(\mathcal{K}, \mathcal{K}')$ and the inclusion is completely contractive.

Proof Every operator in $\mathcal{T}(\mathcal{H}, \mathcal{H}')$ is at least a densely defined operator from $\mathcal{K}$ to $\mathcal{K}'$. Our goal is to show that these operators are bounded.

In the proof, we will require the following diagonal operators. Let D be the diagonal operator on $\mathcal{H}$ with diagonal $(||e_n||_{\mathcal{H}}^{-1})$. Similarly, let D' be the diagonal operator on $\mathcal{H}'$ with diagonal $(||e_n'||_{\mathcal{H}'}^{-1})$. Observe that D extends to a unitary operator from $\mathcal{K}$ to $\mathcal{H}$ and D' extends to a unitary operator from $\mathcal{K}'$ to $\mathcal{H}'$. Moreover, let $U \in \mathcal{B}(\mathcal{H}', \mathcal{H})$ be the unique unitary operator with

$$U e_n' = \frac{||e_n'||_{\mathcal{H}'}}{||e_n||_{\mathcal{H}}} e_n \quad (n \in \mathbb{N}).$$

Suppose now that $[T_{ij}] \in M_r(\mathcal{T}(\mathcal{H}, \mathcal{H}'))$. Then by Lemma 3.1, we find that

$$\begin{aligned} \|[T_{ij}]\|_{\mathcal{B}(\mathcal{K}^r, (\mathcal{K}')^r)} = \|[U D' T_{ij} D^{-1}]\|_{\mathcal{B}(\mathcal{H}^r)} &= \|[D D^{-1} U D' T_{ij} D^{-1}]\|_{\mathcal{B}(\mathcal{H}^r)} \\ &\le \|[D^{-1} U D' T_{ij}]\|_{\mathcal{B}(\mathcal{H}^r)}. \end{aligned}$$

Observe that

$$D^{-1}UD'e'_n = e_n = USe'_n \quad (n \in \mathbb{N}),$$

where S is the diagonal operator on $\mathcal{H}'$ with diagonal $(\frac{||e_n||_{\mathcal{H}}}{||e'_n||_{\mathcal{H}'}})$. By assumption, this operator is a contraction. From the estimate above and the identity $D^{-1}UD' = US$, we infer that

$$\begin{aligned}\|[T_{ij}]\|_{\mathcal{B}(\mathcal{K}^r,(\mathcal{K}')^r)} \le \|[UST_{ij}]\|_{\mathcal{B}(\mathcal{H}^r)} &= \|[ST_{ij}]\|_{\mathcal{B}(\mathcal{H}^r,(\mathcal{H}')^r)} \\ &\le ||[T_{ij}]||_{\mathcal{B}(\mathcal{H}^r,(\mathcal{H}')^r)},\end{aligned}$$

which finishes the proof. ∎

The following result is a generalization of Proposition 3.3.

Proposition 3.7 *Let $d \in \mathbb{N}$ and let $\mathcal{H}, \mathcal{H}', \mathcal{K}, \mathcal{K}'$ be reproducing kernel Hilbert spaces on $\mathbb{B}_d$ with respective reproducing kernels $k_w(z) = \sum_{n=0}^{\infty} a_n \langle z, w\rangle^n$, $k'_w(z) = \sum_{n=0}^{\infty} a'_n \langle z, w\rangle^n$, $\ell_w(z) = \sum_{n=0}^{\infty} b_n \langle z, w\rangle^n$, and $\ell'_w(z) = \sum_{n=0}^{\infty} b'_n \langle z, w\rangle^n$. Suppose that for all $n \in \mathbb{N}_0$, the inequalities $a_n, a'_n, b_n, b'_n > 0$ and*

$$\frac{b_n}{b_{n+1}} \le \frac{a_n}{a_{n+1}}$$

and

$$\frac{b_n}{a_n} \le \frac{b'_n}{a'_n}$$

hold. Then

$$\mathrm{Mult}(\mathcal{H}, \mathcal{H}') \subseteq \mathrm{Mult}(\mathcal{K}, \mathcal{K}'),$$

and the inclusion is completely contractive.

Proof This follows as in the proof of Proposition 3.3 from an application of Corollary 3.6. Indeed, all four spaces have an orthogonal basis of monomials and if we order the monomials such that their degrees are nondecreasing, then every operator in $\mathrm{Mult}(\mathcal{H}, \mathcal{H}')$ is lower triangular. Moreover, if p is a monomial of degree n, then

$$||p||^2_{\mathcal{H}} = \frac{b_n}{a_n}||p||^2_{\mathcal{K}} \quad \text{and } ||p||^2_{\mathcal{H}'} = \frac{b'_n}{a'_n}||p||^2_{\mathcal{K}'},$$

from which it readily follows that the last two conditions in Corollary 3.6 hold. Finally, the assumptions imply that both sequences $(\frac{a_n}{b_n})$ and $(\frac{a_n'}{b_n'})$ are bounded above, so that $\mathcal{H} \subseteq \mathcal{K}$ and $\mathcal{H}' \subseteq \mathcal{K}'$. ■

The last result applies in particular to the spaces B^s_ω.

Corollary 3.8 *Let ω and ν be radial weights in $\mathbb{B}_d$ and let $s, t, s', t' \in \mathbb{R}$ with $t \leq s$ and $t' - s' \leq t - s$. Then*

$$\operatorname{Mult}(B^s_\omega, B^{s'}_\nu) \subseteq \operatorname{Mult}(B^t_\omega, B^{t'}_\nu)$$

and the inclusion is completely contractive. In particular,

$$\operatorname{Mult}(B^s_\omega, B^s_\nu(\ell_2)) \subseteq \operatorname{Mult}(B^t_\omega, B^t_\nu(\ell_2))$$

and the inclusion is contractive.

Proof We apply Proposition 3.7 with $\mathcal{H} = B^s_\omega$, $\mathcal{H}' = B^{s'}_\nu$, $\mathcal{K} = B^t_\omega$, and $\mathcal{K}' = B^{t'}_\nu$. With notation as in that proposition, the argument in the proof of Corollary 3.4 shows that

$$\frac{a_n}{b_n} = n^{2(t-s)}$$

for $n \geq 1$ and $a_0/b_0 = 1$, so the sequence (a_n/b_n) is nonincreasing as $t \leq s$. Similarly,

$$\frac{a_n'}{b_n'} = n^{2(t'-s')}$$

for $n \geq 1$ and $a_0'/b_0' = 1$. Since $t' - s' \leq t - s$, we conclude that $a_n'/b_n' \leq a_n/b_n$ for all $n \in \mathbb{N}_0$, so the result is a special case of Proposition 3.7. ■

We also obtain a multiplier space version of Corollary 3.5.

Theorem 3.9 *Let ω and ν be radial weights in $\mathbb{B}_d$, and let $s, t \in \mathbb{R}$. Then*

$$M^C(B^s_\omega, B^t_\nu) \subseteq M^R(B^s_\omega, B^t_\nu)$$

and the inclusion is continuous.

Proof Let $\mathcal{H} = B^s_\omega$ and $\mathcal{K} = B^t_\nu$. We will use [4, Theorem 4.2], according to which the result follows from the multiplier inclusion condition for the pair $(\mathcal{H}, \mathcal{K})$. To establish this property, by definition, we have to show that there are weights $\widetilde{\omega}$ and $\widetilde{\nu}$ and $N \in \mathbb{N}$ such that $\mathcal{H} = B^N_{\widetilde{\omega}}$, $\mathcal{K} = B^N_{\widetilde{\nu}}$ (with equivalent norms) and

$$\operatorname{Mult}(B^N_{\widetilde{\omega}}, B^N_{\widetilde{\nu}}(\ell^2)) \subseteq \operatorname{Mult}(B^{N-1}_{\widetilde{\omega}}, B^{N-1}_{\widetilde{\nu}}(\ell^2)) \subseteq \cdots \subseteq \operatorname{Mult}(B^0_{\widetilde{\omega}}, B^0_{\widetilde{\nu}}(\ell^2))$$

with continuous inclusions.

To this end, let $x, y \geq 0$ be real numbers such that $s + x = t + y \in \mathbb{N}$ and let $N = s + x = t + y$ be this common value. Moreover, let $\widetilde{\omega} = \omega_x$ and $\widetilde{\nu} = \nu_y$. Then by Theorem 2.4, we have $\mathcal{H} = B^N_{\widetilde{\omega}}$ and $\mathcal{K} = B^N_{\widetilde{\nu}}$. The continuity of the inclusions above now follows from Corollary 3.8, which concludes the proof. ■

4 Weakly Normal Weights

In this section, we will study several finer properties of $L^1[0, 1]$ weights that will translate to Hilbert space properties of the associated radially weighted Besov spaces.

4.1 A Doubling Condition

Recall from Sect. 2.2 that if $v \in L^1[0, 1]$ is non-negative, then we defined for $x > 0$ a weight $v_x \in L^1[0, 1]$ by

$$v_x(t) = \int_{[t,1]} \frac{(s-t)^{x-1}}{\Gamma(x)} v(s)ds, t \in [0, 1)$$

and we also write $\widehat{v} = v_1$. We will now discuss a class of weights on [0, 1], where one has an asymptotics of the type $v_{x+\alpha}(t) \approx (1-t)^\alpha v_x(t)$ for all $t \in [0, 1]$ at least when $x \geq 1$. As in [19] we define the class $\widehat{\mathcal{D}}$ by saying that a non-negative integrable function v is in $\widehat{\mathcal{D}}$ if $\hat{v}$ is doubling near 1, i.e., if there is a constant $c > 0$ such that $\hat{v}(t) \leq c\hat{v}(\frac{1+t}{2})$ for all $t \in [0, 1)$. It is clear that if $v \in \widehat{\mathcal{D}}$ is not identically equal to 0, then it is a weight. For later reference we record the following elementary lemma.

Lemma 4.1 *If $v \in L^1[0, 1]$ is a weight, then $v \in \widehat{\mathcal{D}}$ if and only if there is $M > 1$ such that*

$$\int_t^1 \hat{v}(s)ds \leq (1-t)\hat{v}(t) \leq M \int_t^1 \hat{v}(s)ds.$$

Proof The inequality on the left is true for all $v \geq 0$ since $\hat{v}$ is nonincreasing. First suppose that $\hat{v}$ is doubling. Then there is a $C > 0$ such that $\hat{v}(t) \leq C\hat{v}(\frac{1+t}{2})$ for all $t \in [0, 1)$. Now fix $t \in [0, 1)$, then

$$\int_t^1 \hat{v}(s)ds \geq \int_t^{(1+t)/2} \hat{v}(s)ds$$
$$\geq \hat{v}(\frac{1+t}{2})\frac{1-t}{2}$$

$$\geq \hat{v}(t)\frac{1-t}{2C}.$$

Next suppose that there is $M > 1$ such that $(1-t)\hat{v}(t) \leq M\int_t^1 \hat{v}(s)ds$ for all $t \in [0,1)$. Then one checks by taking a derivative that $\frac{\int_t^1 \hat{v}(s)ds}{(1-t)^M}$ is nondecreasing and thus

$$\begin{aligned}\frac{\hat{v}(t)}{(1-t)^{M-1}} &\leq M\frac{\int_t^1 \hat{v}(s)ds}{(1-t)^M}\\ &\leq M\frac{\int_{\frac{1+t}{2}}^1 \hat{v}(s)ds}{(1-\frac{1+t}{2})^M}\\ &\leq M\hat{v}(\frac{1+t}{2})\frac{2^{M-1}}{(1-t)^{M-1}}.\end{aligned}$$

It follows that $\hat{v}$ is doubling. ■

Notice that for $x > 0$ one has $\hat{v_x}(t) = v_{x+1}(t) = \int_t^1 \frac{(s-t)^{x-1}}{\Gamma(x)}\hat{v}(s)ds$ and with that it is easy to show that if $v \in \widehat{\mathcal{D}}$, then $v_x \in \widehat{\mathcal{D}}$ for each $x > 0$. Then the previous lemma along with Lemma 2.3 implies that for every $v \in \widehat{\mathcal{D}}$ we have $v_{x+1}(t) \approx (1-t)v_x(t)$ for every $x \geq 1$. Since $\hat{v} = v_1$ we inductively obtain $v_{n+1}(t) \approx (1-t)^n\hat{v}(t)$ for each $n \in \mathbb{N}$. But then we have for $0 \leq x \leq n$ that

$$1 \geq \frac{\int_t^1 \left(\frac{s-t}{1-t}\right)^x v(s)ds}{\hat{v}(t)} \geq \frac{\int_t^1 \left(\frac{s-t}{1-t}\right)^n v(s)ds}{\hat{v}(t)} = \frac{\int_t^1 (s-t)^n v(s)ds}{(1-t)^n\hat{v}(t)} \geq C_n.$$

Thus we have proved the following lemma.

Lemma 4.2 *If $v \in \widehat{\mathcal{D}}$, then for all $x \geq 0$ we have $v_{x+1}(t) \approx (1-t)^x\hat{v}(t) = (1-t)^x v_1(t)$.*

4.2 Weakly Normal Weights

The following definition goes back to S.N. Bernstein, [7].

Definition 4.3 Let $a < b$. A function $f : [a,b) \to [0,\infty)$ is called almost decreasing if there is some $C > 0$ such that $f(t) \leq Cf(s)$, whenever $a \leq s \leq t < b$. Almost increasing is defined similarly.

One reason this definition is useful for weights is the following lemma.

Lemma 4.4 *$f : [a, b) \to [0, \infty)$ is almost decreasing, if and only if there is a nonincreasing function $g : [a, b) \to \mathbb{R}$ and $c, C > 0$ such that*

$$cg(t) \le f(t) \le Cg(t)$$

for all $t \in [a, b)$. If f is continuous, then g can be chosen to be continuous as well.

Proof Suppose that g is nonincreasing such that $cg(t) \le f(t) \le Cg(t)$ for all $t \in [a, b)$. Then for $a \le s \le t < b$ we have

$$f(t) \le Cg(t) \le Cg(s) \le \frac{C}{c} f(s) = C' f(s).$$

Conversely, suppose that f is almost decreasing, then for $t \in [a, b)$ set

$$g(t) = \inf\{f(s) : s \le t\}.$$

Note that if f is continuous, then g is continuous. Clearly g is nonincreasing and $g(t) \le f(t)$ for all $t \in [a, b)$. Furthermore, the hypothesis on f implies the existence of $C > 0$ such that $f(t) \le Cf(s)$, whenever $a \le s \le t < b$. This implies $f(t) \le Cg(t)$ for all $t \in [a, b)$. ∎

Lemma 4.5 *Let $v \in L^1[0, 1]$ be a weight. If $t_0 \in [0, 1)$, $\alpha \in \mathbb{R}$, and $x \ge 0$ such that $\frac{(1-t)^\alpha}{v_x(t)}$ is almost decreasing in $[t_0, 1)$, then so is $\frac{(1-t)^{\alpha+y}}{v_{x+y}(t)}$ for every $y \ge 0$.*

Proof We consider reciprocals and thus prove a statement about almost increasing functions. One verifies that $v_{x+y} = (v_x)_y$ for all $x, y \ge 0$. Let $t, t' \in [t_0, 1)$ with $t < t'$, and define $\lambda = \frac{1-t'}{1-t}$. Then $(1 - \lambda) + \lambda t = t'$ and

$$\begin{aligned} v_{x+y}(t) &= \int_t^1 \frac{(s-t)^{y-1}}{\Gamma(y)} \frac{v_x(s)}{(1-s)^\alpha} (1-s)^\alpha ds \\ &\le C \int_t^1 \frac{(s-t)^{y-1}}{\Gamma(y)} \frac{v_x((1-\lambda)+\lambda s)}{\lambda^\alpha (1-s)^\alpha} (1-s)^\alpha ds \\ &= C\lambda^{-(\alpha+y)} \int_{t'}^1 \frac{(u-t')^{y-1}}{\Gamma(y)} v_x(u) du \\ &= C\lambda^{-(\alpha+y)} v_{x+y}(t'). \end{aligned}$$

The Lemma follows. ∎

Now recall from [23] that a weight function v is called normal, if there are $\alpha > \beta \in \mathbb{R}$ such that $\frac{(1-t)^\beta}{v(t)}$ is almost increasing in $[t_0, 1)$ and $\frac{(1-t)^\alpha}{v(t)}$ is almost decreasing in $[t_0, 1)$ for some $0 \le t_0 < 1$. Actually, Shields and Williams required $\beta > 0$ for their results, and they wanted the limits to be ∞ and 0. Furthermore, in the paper [22] the ratios were assumed to be nondecreasing (resp. nonincreasing), but this definition was modified in the later paper [23]. That is a convention that has been used by many authors since then.

Definition 4.6 Let $\alpha \in \mathbb{R}$. We call a weight v *weakly normal of order* α, if there is $x \geq 0$ such that $\frac{(1-t)^{\alpha+x}}{v_x(t)}$ is almost decreasing in $[t_0, 1)$ for some $0 \leq t_0 < 1$. The weight v is called weakly normal, if it is weakly normal of order α for some $\alpha \in \mathbb{R}$.

Since v_x is nonincreasing for all $x \geq 1$, we do not require an assumption corresponding to the parameter β above.

If a weight is weakly normal of order α, then $\alpha > -1$. Indeed, if v is weakly normal of order α, then by Lemma 4.5 we may assume that $\frac{(1-t)^{\alpha+x}}{v_x(t)}$ is almost decreasing in $[t_0, 1)$ for some $x \geq 1$ and $0 \leq t_0 < 1$. Then for $t \in [t_0, 1)$ we have

$$(1-t)^{\alpha+x} \leq C v_x(t) \leq \frac{C}{\Gamma(x)}(1-t)^{x-1}\hat{v}(t).$$

But we have $\hat{v}(t) \to 0$ as $t \to 1$. We see that this is only possible if $\alpha > -1$.

Obviously $v(t) = (1-t)^{\alpha}$ is weakly normal of order α, whenever $\alpha > -1$. It is also clear from the identity $v_{x+y} = (v_x)_y$ and Lemma 4.5 that v is weakly normal, if and only if v_x is weakly normal for each $x \geq 0$, and this happens if and only if v_x is weakly normal for some $x \geq 0$. In the following Lemma we have summarized the relationship of the weakly normal weights with the class $\widehat{\mathcal{D}}$ and with another class of weights that has been considered in the literature. For $\eta > -1$ the Bekollé–Bonami class $B_2(\eta)$ is defined by

$$\frac{v(t)}{(1-t)^{\eta}} \in B_2(\eta) \iff \int_t^1 v(s)ds \int_t^1 \frac{(1-s)^{2\eta}}{v(s)}ds \approx (1-t)^{2\eta+2}.$$

This is the radial weight version of a more general definition that characterizes the weights ω on $\mathbb{B}_d$ such that a corresponding Bergman projection is bounded on $L^2(\omega)$, see, e.g., [6].

Lemma 4.7 *Let* $v \in L^1[0, 1]$ *be a weight.*

(a) If $\eta > -1$ *and* $\frac{v(t)}{(1-t)^{\eta}} \in B_2(\eta)$*, then* v *is weakly normal of order* $2\eta + 1$.
(b) Let $v \in L^1[0, 1]$ *be non-negative. Then the following are equivalent:*

(i) v *is weakly normal,*
(ii) *there are* $x \geq 0$ *and* $\eta > -1$ *such that* $\frac{v_x(t)}{(1-t)^{\eta}} \in B_2(\eta)$*,*
(iii) *there is* $x \geq 0$ *such that* $v_x \in \widehat{\mathcal{D}}$.

Proof

(a) Let $\eta > -1$ and suppose $\frac{v(t)}{(1-t)^{\eta}} \in B_2(\eta)$, then $g(t) = \int_t^1 \frac{(1-s)^{2\eta}}{v(s)}ds$ is nonincreasing and the hypothesis implies that

$$g(t) \approx \frac{(1-t)^{2\eta+2}}{v_1(t)}.$$

Thus Lemma 4.4 implies that $\frac{(1-t)^{2\eta+2}}{v_1(t)}$ is almost decreasing, i.e., v satisfies the definition of weakly normal of order $2\eta + 1$ with $x = 1$.

(b) $(ii) \Rightarrow (i)$ follows from (a) and the earlier observation that v is weakly normal if and only if v_x is weakly normal for some $x \geq 0$.

$(iii) \Rightarrow (ii)$ By Lemma 4.1 $v_x \in \widehat{\mathcal{D}}$ if and only if there is a $C > 1$ such that $(1-t)v_{x+1}(t) \leq C v_{x+2}(t)$. By use of a first derivative one sees that this is equivalent to $\frac{(1-t)^C}{v_{x+2}(t)}$ being nonincreasing. Hence

$$\int_t^1 v_{x+2}(s)ds \int_t^1 \frac{(1-s)^C}{v_{x+2}(s)} ds \leq v_{x+2}(t)(1-t)\frac{(1-t)^C}{v_{x+2}(t)}(1-t) = (1-t)^{C+2}.$$

Thus $\frac{v_{x+2}(t)}{(1-t)^{C/2}} \in B_2(C/2)$.

$(i) \Rightarrow (iii)$ If v is weakly normal, then there are $x \geq 1$ and $\alpha > -1$ such that $\frac{(1-t)^{\alpha+x}}{v_x(t)}$ is almost decreasing. Then one easily checks directly that $v_x \in \widehat{\mathcal{D}}$. ■

Weights of the type $(1-t)^{\alpha}\left(\frac{1}{t}\log\frac{1}{1-t}\right)^{\beta}$ for $\alpha > -1, \beta \geq 0$ are weakly normal of order α. If $\beta < 0$, then such a weight would be weakly normal of order γ for each $\gamma > \alpha$. This also holds when $\alpha = -1$, although for $\beta < -1$ the weight $\frac{\left(\frac{1}{t}\log\frac{1}{1-t}\right)^{\beta}}{1-t}$ is not a Bekollé weight.

Part (b) of the previous lemma could be paraphrased by saying that the weakly normal weights could also have been called "weakly doubling" or "weak Bekollé weights." For us the viewpoint of weakly normal is important, because the order of a weakly normal weight determines the cut-off for a weighted Besov space to have the Pick property, see Theorem 5.2. The following theorem is instrumental for the proof.

Theorem 4.8 *Let $v \in L^1[0,1]$ be a weight. If v is weakly normal of order $\alpha > -1$, then there is a positive Borel measure μ on $[0,1]$ such that*

$$\int_0^1 t^n v(t)dt \int_{[0,1]} t^n d\mu(t) \approx n^{-\alpha-1} \ \textit{as } n \to \infty.$$

Proof By Lemma 2.2 it will suffice to show that for some $x \geq 0$ there is a measure μ with

$$\int_0^1 t^n v_x(t)dt \int_{[0,1]} t^n d\mu(t) \approx n^{-\alpha-x-1} \ \text{ as } n \to \infty.$$

Note that v_x is continuous for all $x \geq 1$. Thus, by the hypothesis and Lemmas 4.5 and 4.4 there is $x \geq 1$ and a nonincreasing continuous function g on $[0,1)$ such that $g(t) \approx \frac{(1-t)^{\alpha+x}}{v_x(t)}$ for $t \in [t_0, 1)$.

By Lemmas 4.5 and 4.7 we may assume that $x \geq 1$, $v_{x-1} \in \widehat{\mathcal{D}}$, and hence that v_x is nonincreasing.

We set $g(1) = \lim_{t\to 1} g(t)$. Then there is a Borel measure μ on $[0, 1]$ such that $g(t) = \mu([t, 1])$. Note that $g \in L^1[0, 1]$,

$$\hat{g}(t) = \int_t^1 g(s)ds \approx \int_t^1 \frac{(1-s)^{\alpha+x}}{v_x(s)}ds \leq \int_t^1 \frac{(1-s)^{\alpha+x}}{v_x(\frac{1+s}{2})}ds, \ t \in [t_0, 1).$$

Hence

$$\hat{g}(t) \lesssim 2^{\alpha+x+1} \int_{\frac{1+t}{2}}^1 \frac{(1-u)^{\alpha+x}}{v_x(u)}du \approx \hat{g}(\frac{1+t}{2}).$$

This implies that $g \in \widehat{\mathcal{D}}$.

Since $v_{x-1} \in \widehat{\mathcal{D}}$ we also have $v_x \in \widehat{\mathcal{D}}$. Thus Lemma A of [19] with $v_x, g \in \widehat{\mathcal{D}}$ implies that $\int_0^1 t^n g(t)dt \approx \int_{1-\frac{1}{n}}^1 g(t)dt$ and $\int_0^1 t^n v_x(t)dt \approx \int_{1-\frac{1}{n}}^1 v_x(t)dt$. Furthermore, by Lemma 4.1 applied with v_{x-1} we have $\int_t^1 v_x(s)ds \approx (1-t)v_x(t)$. Also noting that Lemma 2.2 implies that $\int_0^1 t^n d\mu \approx n \int_0^1 t^n g(t)dt$ we obtain

$$\int_0^1 t^n v_x(t)dt \int_{[0,1]} t^n d\mu(t) \approx v_x(1-\frac{1}{n}) \int_{1-\frac{1}{n}}^1 \frac{(1-s)^{\alpha+x}}{v_x(s)}ds. \tag{4.1}$$

Since v_x is nonincreasing we immediately obtain

$$\int_0^1 t^n v_x(t)dt \int_{[0,1]} t^n d\mu(t) \geq c \int_{1-\frac{1}{n}}^1 (1-s)^{\alpha+x}ds \approx n^{-\alpha-x-1}.$$

By the hypothesis the ratio $\frac{v_x(t)}{(1-t)^{\alpha+x}}$ is almost increasing, hence there is $C > 0$ such that for $s \geq 1 - \frac{1}{n}$

$$v_x(1-\frac{1}{n}) \leq Cn^{-\alpha-x} \frac{v_x(s)}{(1-s)^{\alpha+x}}.$$

Thus we may substitute this inequality into (4.1), and this concludes the proof of the theorem. ∎

We will say that a radial weight ω on $\mathbb{B}_d$ is weakly normal (of order $\alpha > -1$), if the associated $L^1[0, 1]$-function v (see Sect. 2) is weakly normal (of order $\alpha > -1$). For weakly normal radial weights Lemmas 4.2 and 4.7 imply that there is $x_0 \geq 0$ such that $\omega_{x+x_0} \approx (1-|z|^2)^{2x}\omega_{x_0}$ for all $x \geq 0$.

Example 4.9 Examples of weights that are not weakly normal are $\omega(z) = (1 - |z|^2)^\beta e^{\frac{-1}{1-|z|^2}}$, $\beta \in \mathbb{R}$. One checks with Lemma 4.4 that such a weight would be weakly normal, if and only if $v(t) = (1-t)^\beta e^{\frac{-1}{1-t}}$ is weakly normal. We calculate

$$\begin{aligned}(1-t)^2 v(t) &= -\int_t^1 \frac{d}{ds}(1-s)^{\beta+2} e^{\frac{-1}{1-s}} ds \\ &= \int_t^1 (1+(\beta+2)(1-s))v(s)ds \\ &\approx \int_t^1 v(s)ds = \hat{v}(t)\end{aligned}$$

Iteration of this shows that for each positive integer N we have $(1-t)^{2N} v(t) \approx v_N(t)$ and hence $\omega_N \approx (1-|z|^2)^{4N}\omega$. For more on such weights, see [18].

5 Radial Weights and Complete Pick Spaces

Our result on radially weighted Besov spaces that are complete Pick spaces is based on the following lemma.

Lemma 5.1 *Let μ be a probability measure on* $[0, 1]$. *Then there are $c_n \geq 0$ such that*

$$\int_{[0,1]} \frac{1}{1-tz} d\mu(t) = \frac{1}{1-\sum_{n=1}^\infty c_n z^n}$$

for all $|z| < 1$.

It follows that $k_w(z) = \int_{[0,1]} \frac{1}{1-t\langle z,w\rangle} d\mu(t)$ defines a normalized Pick kernel in $\mathbb{B}_d$.

Proof Let $F(s) = \int_{[0,1]} t^s d\mu(t)$ be the moment generating function for this setup. It is well known that $\log F(s)$ defines a convex function on $[0, \infty)$. In fact, it easily follows from Hölder's inequality that $F(\lambda s_1 + (1-\lambda)s_2) \leq F(s_1)^\lambda F(s_2)^{1-\lambda}$ for all $s_1, s_2 \in [0, \infty)$ and $0 < \lambda < 1$. The logarithmic convexity of F follows from this. Thus for each $n \geq 0$ we have

$$\log F(n+1) - \log F(n) \leq \log F(n+2) - \log F(n+1),$$

which is equivalent to $\frac{F(n+1)}{F(n)}$ being nondecreasing in n. Now the conclusion of the lemma follows from Kaluza's lemma (see, e.g., [2], Lemma 7.38) since $\int_{[0,1]} \frac{1}{1-tz} d\mu(t) = \sum_{n=0}^\infty F(n) z^n$. ∎

Theorem 5.2 *If ω is a weakly normal radial weight of order $\alpha > -1$, then B^s_ω is a complete Pick space for all $s \geq \frac{\alpha+d}{2}$.*

Proof Let v be the $L^1[0,1]$-function associated with ω and set $\alpha' = 2s - d \geq \alpha$. Then v is a weakly normal weight of order α'.

For $n \in \mathbb{N}_0$ let $a_n = a_n(\omega) = \int_0^1 t^n v(t)dt$, and choose a probability measure μ such that $b_n = \int_{[0,1]} t^n d\mu \approx \left(n^{\alpha'+1} a_n\right)^{-1} = \left(n^{2s-d+1} a_n\right)^{-1}$. This can be done by Theorem 4.8. Define

$$k_w(z) = \int_{[0,1]} \frac{1}{1 - t\langle z, w\rangle} d\mu(t).$$

Then $k_w(z) = \sum_{n=0}^{\infty} b_n \langle z, w\rangle^n$ is a normalized complete Pick kernel by Lemma 5.1. Let $\mathcal{H}$ be the reproducing kernel Hilbert space with kernel k, and let $\|f\|_{H^2_d}$ denote the Drury–Arveson norm of a function $f = \sum_n f_n$. It is easy to check and well known that $\|f\|^2_{\mathcal{H}} = \sum_{n=0}^{\infty} \frac{1}{b_n} \|f_n\|^2_{H^2_d}$. Recall that for homogeneous polynomials f_n of degree n we have

$$\|f_n\|^2_{H^2_d} = c_n \|f_n\|^2_{H^2(\partial\mathbb{B}_d)}, \text{ where } c_n \approx (n+1)^{d-1},$$

see, for example, formula (2.2) of [21]. We now apply the above and the definition of the B^s_ω-norm to obtain

$$\begin{aligned}
\|f\|^2_{B^s_\omega} &= |f(0)|^2 + \sum_{n=1}^{\infty} n^{2s} a_n \|f_n\|^2_{H^2(\partial\mathbb{B}_d)} \\
&\approx |f(0)|^2 + \sum_{n=1}^{\infty} n^{2s-d+1} a_n \|f_n\|^2_{H^2_d} \\
&\approx |f(0)|^2 + \sum_{n=1}^{\infty} \frac{1}{b_n} \|f_n\|^2_{H^2_d} \\
&= \|f\|^2_{\mathcal{H}}.
\end{aligned}$$

■

Corollary 5.3 *If ω is a weakly normal radial weight of order $\alpha > -1$, then for every $s_0 \geq (\alpha + d)/2$, there is a positive nonincreasing continuous function $g \in \widehat{\mathcal{D}}$ such that for every $x, y \geq 0$*

$$k_w(z) = \int_0^1 \frac{(1-t)^x}{(1 - t\langle z, w\rangle)^{x+3+2y}} \hat{g}(t)dt$$

is a reproducing kernel for $B^{s_0 - y}_\omega$.

By this we mean that there is an alternate norm on $B_\omega^{s_0-y}$ which is equivalent to the natural norm and such that $k_w(z)$ is the reproducing kernel for the space under the alternate norm.

Proof Since $s_0 \geq (\alpha+d)/2$ Theorem 5.2 implies that the space $B_\omega^{s_0}$ is a complete Pick space. Furthermore, the proof of Theorem 5.2 shows that

$$k_w^{s_0}(z) = \int_0^1 \frac{1}{1-t\langle z,w\rangle} d\mu(t) = \sum_{n=0} \langle z,w\rangle^n \int_{[0,1]} t^n d\mu(t)$$

is a reproducing kernel for $B_\omega^{s_0}$. The existence of the measure μ was established by means of Theorem 4.8, whose proof shows that μ can be chosen so that $g(t) = \mu([t,1])$ is continuous and satisfies $g \in \widehat{\mathcal{D}}$. For $x \geq 0$ let w_x be the $L^1[0,1]$-function associated with μ as in Lemma 2.2, then $w_1 = g$, $w_2 = \hat{g}$, and Lemma 4.2 implies that $w_{x+2}(t) \approx (1-t)^x \hat{g}(t)$. Now consider the power series

$$k_w(z) = \int_0^1 \frac{(1-t)^x}{(1-t\langle z,w\rangle)^{x+3+2y}} \hat{g}(t)dt = \sum_{n=0}^{\infty} a_n \langle z,w\rangle^n,$$

where

$$\begin{aligned} a_n &\approx (n+1)^{x+2+2y} \int_0^1 t^n (1-t)^x \hat{g}(t)dt \\ &\approx (n+1)^{x+2+2y} \int_0^1 t^n w_{x+2}(t)dt \\ &\approx (n+1)^{2y} \int_{[0,1]} t^n d\mu(t) \end{aligned}$$

by Lemma 2.2. It is easy to see that if $k_w^{s_0}(z)$ is a reproducing kernel for $B_\omega^{s_0}$, then $k_w(z)$ is a reproducing kernel for $B_\omega^{s_0-y}$. ■

Corollary 5.4 *Let ω be a weakly normal radial weight on $\mathbb{B}_d$. For $s \in \mathbb{R}$ let $k_w^s(z)$ be the reproducing kernel for B_ω^s.*

Then for each $s \leq t$ there is $c > 0$ such that $k_z^s(z) \leq c\frac{k_z^t(z)}{(1-|z|^2)^{2(t-s)}}$ for all $z \in \mathbb{B}_d$.

Proof If v is weakly normal of order $\alpha > -1$, then choose $s_0 \geq \max(t, (\alpha+d)/2)$. Then by the previous corollary with $x = 0$ we have

$$k_z^s(z) \approx \int_0^1 \frac{\hat{g}(u)}{(1-u|z|^2)^{3+2(s_0-s)}} du$$

and

$$k_z^t(z) \approx \int_0^1 \frac{\hat{g}(u)}{(1-u|z|^2)^{3+2(s_0-t)}} du.$$

The corollary follows from this. ■

6 Further Results About Multipliers of B_ω^s

Some of the main results from the previous sections are about bounded column operators on weighted Besov spaces with radial weights. In this section we collect more facts about such operators.

Let $\alpha \geq 0$ be a real parameter. We will need to use the growth space $A^{-\alpha}(\ell_2)$ defined by

$$A^{-\alpha}(\ell_2) = \{\Phi = (\varphi_1, \varphi_2, \ldots), \varphi_i \in \mathrm{Hol}(\mathbb{B}_d), \|\Phi\|_{A^{-\alpha}(\ell_2)} < \infty\},$$

where

$$\|\Phi\|_{A^{-\alpha}(\ell_2)}^2 = \sup_{z \in \mathbb{B}_d} (1-|z|^2)^{2\alpha} \sum_{i=1}^{\infty} |\varphi_i(z)|^2.$$

If $\alpha = 0$, then we just obtain the bounded analytic functions and we observe $H^\infty(\mathbb{C}, \ell_2) = A^0(\ell_2)$ and $\|\Phi\|_\infty = \|\Phi\|_{A^0(\ell_2)}$.

The following lemma is well known.

Lemma 6.1 *Let $\gamma > 0$, $n \in \mathbb{N}$. Then there is a $c > 0$ such that for all sequences of analytic functions $\Phi = (\varphi_1, \varphi_2, \ldots)$ on $\mathbb{B}_d$ we have*

$$\frac{1}{c}\|\Phi\|_{A^{-\gamma}(\ell_2)} \leq \|\Phi(0)\|_{\ell_2} + \|R^n\Phi\|_{A^{-\gamma-n}(\ell_2)} \leq c\|\Phi\|_{A^{-\gamma}(\ell_2)}.$$

and hence $\Phi \in A^{-\gamma}(\ell_2)$ if and only if $R^n\Phi \in A^{-\gamma-n}(\ell_2)$ and $\Phi(0) \in \ell_2$.

Furthermore, if $\Phi \in H^\infty(\mathbb{C}, \ell_2)$, then $R^n\Phi \in A^{-n}(\ell_2)$ and

$$\|R^n\Phi\|_{A^{-n}(\ell_2)} \leq c\|\Phi\|_{H^\infty}.$$

Proof By induction it follows that it suffices to prove the case where $n = 1$. Furthermore, that case follows easily from the formulas $\varphi(z) = \varphi(0) + \int_0^1 R\varphi(tz)\frac{dt}{t}$ and $R\varphi(z) = \frac{1}{2\pi i}\int_{|\lambda-1|=r} \frac{\varphi(\lambda z)}{(\lambda-1)^2} d\lambda$, $r = (1-|z|)/2$. ■

Theorem 6.2 *Let ω be a radial weight, let $s, t \in \mathbb{R}$ with $t \leq s$, and let $\Phi \in A^{-(s-t)}(\ell_2)$.*

Then the following are equivalent:

(a) $\Phi \in \mathrm{Mult}(B^s_\omega, B^t_\omega(\ell_2))$,
(b) there exists $n \in \mathbb{N}_0$ *such that* $R^n\Phi \in \mathrm{Mult}(B^s_\omega, B^t_{\omega_n}(\ell_2))$,
(c) for all $n \in \mathbb{N}_0$ *we have* $R^n\Phi \in \mathrm{Mult}(B^s_\omega, B^t_{\omega_n}(\ell_2))$.

In fact, for each $n \in \mathbb{N}$ *we have*

$$\|\Phi\|_{A^{-(s-t)}(\ell_2)} + \|\Phi\|_{\mathrm{Mult}(B^s_\omega, B^t_\omega(\ell_2))} \approx \|\Phi\|_{A^{-(s-t)}(\ell_2)} + \|R^n\Phi\|_{\mathrm{Mult}(B^s_\omega, B^t_{\omega_n}(\ell_2))}.$$

Proof Let $n \in \mathbb{N}_0$. The equivalence of the three conditions and the equivalence of norms will follow from an obvious inductive argument once we show the two inequalities

$$\|R^{n+1}\Phi\|_{\mathrm{Mult}(B^s_\omega, B^t_{\omega_{n+1}}(\ell_2))} \lesssim \|R^n\Phi\|_{\mathrm{Mult}(B^s_\omega, B^t_{\omega_n}(\ell_2))} \tag{6.1}$$

$$\|R^n\Phi\|_{\mathrm{Mult}(B^s_\omega, B^t_{\omega_n}(\ell_2))} \lesssim \|\Phi\|_{A^{-(s-t)}(\ell_2)} + \|R^{n+1}\Phi\|_{\mathrm{Mult}(B^s_\omega, B^t_{\omega_{n+1}}(\ell_2))}. \tag{6.2}$$

Let $R^n\Phi \in \mathrm{Mult}(B^s_\omega, B^t_{\omega_n}(\ell_2))$. It follows from Corollary 3.8 that

$$\|R^n\Phi\|_{\mathrm{Mult}(B^{s-1}_\omega, B^{t-1}_{\omega_n}(\ell_2))} \le \|R^n\Phi\|_{\mathrm{Mult}(B^s_\omega, B^t_{\omega_n}(\ell_2))}.$$

Since $B^{s-1}_\omega = B^s_{\omega_1}$ and $B^{t-1}_{\omega_n} = B^t_{\omega_{n+1}}$ with equivalence of norms by Theorem 2.4, we conclude that for $h \in B^s_\omega$

$$\begin{aligned}
&\|(R^{n+1}\Phi)h\|_{B^t_{\omega_{n+1}}(\ell_2)}\\
\le&\|R((R^n\Phi)h)\|_{B^t_{\omega_{n+1}}(\ell_2)} + \|(R^n\Phi)Rh\|_{B^t_{\omega_{n+1}}(\ell_2)}\\
\lesssim&\|(R^n\Phi)h\|_{B^t_{\omega_n}(\ell_2)} + \|R^n\Phi\|_{\mathrm{Mult}(B^s_{\omega_1}, B^t_{\omega_{n+1}}(\ell_2))}\|Rh\|_{B^s_{\omega_1}}\\
\lesssim&2\|R^n\Phi\|_{\mathrm{Mult}(B^s_\omega, B^t_{\omega_n}(\ell_2))}\|h\|_{B^s_\omega}.
\end{aligned}$$

Thus (6.1) holds and $R^{n+1}\Phi \in \mathrm{Mult}(B^s_\omega, B^t_{\omega_{n+1}}(\ell_2))$.

Next we assume that $R^{n+1}\Phi \in \mathrm{Mult}(B^s_\omega, B^t_{\omega_{n+1}}(\ell_2))$, we write

$$M_{n+1}(\Phi) = \|R^{n+1}\Phi\|_{\mathrm{Mult}(B^s_\omega, B^t_{\omega_{n+1}}(\ell_2))} + \|(R^n\Phi)(0)\|_{\ell_2}$$

and we choose an integer $N \ge s$. Let k be an integer with $0 \le k \le N$. Since $B^s_{\omega_k} = B^{s-k}_\omega$ and $B^t_{\omega_{n+1+k}} = B^{t-k}_{\omega_{n+1}}$ with equivalence of norms, Corollary 3.8 applied to the function $R^{n+1}\Phi$ implies that

$$\|R^{n+1}\Phi\|_{\mathrm{Mult}(B^s_{\omega_k}, B^t_{\omega_{n+1+k}}(\ell_2))} \le M_{n+1}(\Phi).$$

Then for all $h \in B^s_\omega$ we have

$$
\begin{aligned}
&\|(R^n\Phi)h\|_{B^t_{\omega_n}(\ell_2)}\\
\lesssim&\|(R^n\Phi)(0)h(0)\|_{\ell_2} + \|R((R^n\Phi)h)\|_{B^t_{\omega_{n+1}}(\ell_2)}\\
\lesssim&M_{n+1}(\Phi)\|h\|_{B^s_\omega} + \|(R^{n+1}\Phi)h\|_{B^t_{\omega_{n+1}}(\ell_2)} + \|(R^n\Phi)Rh\|_{B^t_{\omega_{n+1}}(\ell_2)}\\
\lesssim&2M_{n+1}(\Phi)\|h\|_{B^s_\omega} + \|R((R^n\Phi)Rh)\|_{B^t_{\omega_{n+2}}(\ell_2)}\\
\lesssim&2M_{n+1}(\Phi)\|h\|_{B^s_\omega} + \|(R^{n+1}\Phi)Rh\|_{B^t_{\omega_{n+2}}(\ell_2)} + \|(R^n\Phi)R^2h\|_{B^t_{\omega_{n+2}}(\ell_2)}\\
\lesssim&2M_{n+1}(\Phi)\|h\|_{B^s_\omega} + M_{n+1}(\Phi)\|Rh\|_{B^s_{\omega_1}} + \|(R^n\Phi)R^2h\|_{B^t_{\omega_{n+2}}(\ell_2)}\\
\lesssim&3M_{n+1}(\Phi)\|h\|_{B^s_\omega} + \|(R^n\Phi)R^2h\|_{B^t_{\omega_{n+2}}(\ell_2)}.
\end{aligned}
$$

Thus iteration of this argument shows that

$$
\|(R^n\Phi)h\|_{B^t_{\omega_n}(\ell_2)} \lesssim (N+1)M_{n+1}(\Phi)\|h\|_{B^s_\omega} + \|(R^n\Phi)R^Nh\|_{B^t_{\omega_{n+N}}(\ell_2)}.
$$

Since $M_{n+1}(\Phi)$ is dominated by the right-hand side of (6.2), it remains to estimate the second summand. Note that as $n+N \geq t$ we have $B^t_{\omega_{n+N}}(\ell_2) = L^2_a(\omega_{n+N-t}, \ell_2)$ with equivalence of norms. The growth hypothesis on Φ and Lemma 6.1 imply that $R^n\Phi \in A^{-(s-t+n)}(\ell_2)$ with $\|R^n\Phi\|_{A^{-(s-t+n)}} \lesssim \|\Phi\|_{A^{-(s-t)}}$, so using (2.4), we see that

$$
\begin{aligned}
\|(R^n\Phi)R^Nh\|^2_{B^t_{\omega_{n+N}}(\ell_2)} &\approx \int_{\mathbb{B}_d} \|(R^n\Phi)(z)R^Nh(z)\|^2_{\ell_2}\,\omega_{n+N-t}dV\\
&\lesssim \|R^n\Phi\|^2_{A^{-(s-t+n)}} \int_{\mathbb{B}_d} |R^Nh(z)|^2 \frac{\omega_{n+N-t}}{(1-|z|^2)^{2(s-t+n)}}dV\\
&\lesssim \|R^n\Phi\|^2_{A^{-(s-t+n)}} \int_{\mathbb{B}_d} |R^Nh(z)|^2\omega_{N-s}dV\\
&\lesssim \|\Phi\|^2_{A^{-(s-t)}(\ell_2)}\|R^Nh\|^2_{L^2_a(\omega_{N-s})}\\
&\lesssim \|\Phi\|^2_{A^{-(s-t)}(\ell_2)}\|h\|^2_{B^s_\omega}.
\end{aligned}
$$

Thus (6.2) holds and this concludes the proof. ■

Since the multipliers of a space into itself are always bounded we obtain an immediate consequence.

Theorem 6.3 *Let ω be a radial weight in $\mathbb{B}_d$, and let $s \in \mathbb{R}$, $N \in \mathbb{N}_0$. Then*

$$
\mathrm{Mult}(B^s_\omega, B^s_\omega(\ell_2)) = \{\Phi \in H^\infty(\mathbb{C}, \ell_2) : R^N\Phi \in \mathrm{Mult}(B^s_\omega, B^{s-N}_\omega(\ell_2))
$$

and $\|\Phi\|_{\text{Mult}(B^s_\omega, B^s_\omega(\ell_2))} \approx \|R^N\Phi\|_{\text{Mult}(B^s_\omega, B^{s-N}_\omega(\ell_2))} + \|\Phi\|_\infty$.

Note that if $N \geq s$, then $B^{s-N}_\omega = L^2_a(\omega_{N-s})$ is a weighted Bergman space and the condition in the Corollary says that the higher order derivatives of multipliers satisfy a Carleson measure condition that is appropriate for the space B^s_ω (see [10]): There exists a $c > 0$ such that

$$\int_{\mathbb{B}_d} |f(z)|^2 \|R^N\Phi(z)\|^2_{\ell_2} \omega_{N-s}(z) dV \leq c\|f\|^2_{B^s_\omega}$$

for all $f \in B^s_\omega$.

For the standard weights $\omega(z) = (1 - |z|^2)^\eta$, $\eta > -1$ the scalar case of this theorem is due to Fabrega and Ortega and [17], also see [10]. For general Bekollé–Bonami weights (not necessarily radial) it is in [9]. We note that in those contexts the L^p-case was treated as well.

We don't know whether the equivalence of (b) and (c) of Theorem 6.2 for $n \geq 1$ remains true without the hypothesis that $\Phi \in A^{-(s-t)}(\ell_2)$. For general Bekollé–Bonami weights that is the case, see [9]. We will now see that it also holds for all weakly normal radial weights.

Lemma 6.4 *Let ω be a weakly normal radial weight, and let $s, \alpha \in \mathbb{R}$ with $\alpha \geq 0$. Then*

$$\text{Mult}(B^s_\omega, B^{s-\alpha}_\omega(\ell^2)) \subseteq A^{-\alpha}(\ell_2).$$

Proof Let $\Phi = \{\varphi_1, \dots\} \in \text{Mult}(B^s_\omega, B^{s-\alpha}_\omega(\ell^2))$. Then for each $z \in \mathbb{B}_d$ we have

$$\begin{aligned}
\sum_{n\geq 1} |\varphi_n(z)|^2 &= \sum_{n\geq 1} \frac{|\langle \varphi_n k^s_z, k^{s-\alpha}_z\rangle|^2}{|k^s_z(z)|^2} \\
&\leq \frac{\|\Phi\|^2_{\text{Mult}(B^s_\omega, B^{s-\alpha}_\omega(\ell^2))} \|k^s_z\|^2 \|k^{s-\alpha}_z\|^2}{|k^s_z(z)|^2} \\
&= \|\Phi\|^2_{\text{Mult}(B^s_\omega, B^{s-\alpha}_\omega(\ell^2))} \frac{\|k^{s-\alpha}_z\|^2}{\|k^s_z\|^2} \\
&\leq c\|\Phi\|^2_{\text{Mult}(B^s_\omega, B^{s-\alpha}_\omega(\ell^2))} (1 - |z|^2)^{-2\alpha} \quad \text{by Corollary 5.4.}
\end{aligned}$$

■

Corollary 6.5 *Let ω be a weakly normal radial weight, and let $s, t \in \mathbb{R}$ with $t < s$. Then the following are equivalent:*

(a) $\Phi \in \text{Mult}(B^s_\omega, B^t_\omega(\ell_2))$,
(b) there exists $n \in \mathbb{N}_0$ such that $R^n\Phi \in \text{Mult}(B^s_\omega, B^t_{\omega_n}(\ell_2))$,
(c) for all $n \in \mathbb{N}_0$ we have $R^n\Phi \in \text{Mult}(B^s_\omega, B^t_{\omega_n}(\ell_2))$.

Proof This follows from Theorem 6.2, because by Lemmas 6.4 and 6.1 each of the cases (a), (b), or (c) automatically implies the required growth hypothesis. ■

In particular, applying this to $R\Phi$ and $t = s - 1$, we conclude that for weakly normal radial weights we have $R\Phi \in \text{Mult}(B^s_\omega, B^{s-1}_\omega(\ell_2))$ if and only if there is $n \in \mathbb{N}$ such that $R^n\Phi \in \text{Mult}(B^s_\omega, B^{s-n}_\omega(\ell_2))$.

References

1. J. Agler, J.E. McCarthy, Complete Nevanlinna-Pick kernels. J. Funct. Anal. **175**(1), 111–124 (2000)
2. J. Agler, J.E. McCarthy, Pick interpolation and Hilbert function spaces, vol 44, in *Graduate Studies in Mathematics* (American Mathematical Society, Providence, 2002)
3. A. Aleman, M. Hartz, J. McCarthy, S. Richter, Interpolating sequences in spaces with the complete Pick property. Int. Math. Res. Not. (to appear). https://doi.org/10.1093/imrn/rnx237
4. A. Aleman, M. Hartz, J. McCarthy, S. Richter, Weak products of complete Pick spaces. Indiana U. Math. J. (to appear). arXiv:1804.10693
5. A. Aleman, M. Hartz, J. McCarthy, S. Richter, The Smirnov class for spaces with the complete Pick property. J. Lond. Math. Soc. (2) **96**(1), 228–242 (2017)
6. D. Bekollé, Inégalité à poids pour le projecteur de Bergman dans la boule unité de $\mathbf{C}^n$. Studia Math. **71**(3), 305–323 (1981/1982)
7. S.N. Bernstein, On majorants of finite or quasi-finite growth. Dokl. Akad. Nauk SSSR **65**, 117–120 (1949)
8. L. Brown, A.L. Shields, Cyclic vectors in the Dirichlet space. Trans. Am. Math. Soc. **285**(1), 269–303 (1984)
9. C. Cascante, J. Fàbrega, Bilinear forms on weighted Besov spaces. J. Math. Soc. Japan **68**(1), 383–403 (2016)
10. C. Cascante, J. Fàbrega, J.M. Ortega, On weighted Toeplitz, big Hankel operators and Carleson measures. Integr. Equ. Oper. Theory **66**(4), 495–528 (2010)
11. I. Chalendar, J.R. Partington, Norm estimates for weighted composition operators on spaces of holomorphic functions. Complex Anal. Oper. Theory **8**(5), 1087–1095 (2014)
12. K. Guo, J. Hu, X. Xu, Toeplitz algebras, subnormal tuples and rigidity on reproducing $\mathbf{C}[z_1, \dots, z_d]$-modules. J. Funct. Anal. **210**(1), 214–247 (2004)
13. V.È. Kacnel′son, A remark on canonical factorization in certain spaces of analytic functions. Zap. Naucn. Sem. Leningrad. Otdel. Mat. Inst. Steklov (LOMI) **30**, 163–164 (1972). Investigations on linear operators and the theory of functions, III
14. X. Li, *Moment Sequences and Their Applications*, PhD thesis. Virginia Polytechnic Institute and State University, Virginia, 1994
15. A.W. Marcus, D.A. Spielman, N. Srivastava, Interlacing families II: mixed characteristic polynomials and the Kadison-Singer problem. Ann. Math. (2) **182**(1), 327–350 (2015)
16. S. McCullough, Carathéodory interpolation kernels. Integr. Equ. Oper. Theory **15**(1), 43–71 (1992)
17. J.M. Ortega, J. Fàbrega, Pointwise multipliers and decomposition theorems in analytic Besov spaces. Math. Z. **235**(1), 53–81 (2000)
18. M. Pavlović, J.Á. Peláez, An equivalence for weighted integrals of an analytic function and its derivative. Math. Nachr. **281**(11), 1612–1623 (2008)
19. J.Á. Peláez, J. Rättyä, Two weight inequality for Bergman projection. J. Math. Pures Appl. (9) **105**(1), 102–130 (2016)
20. P. Quiggin, For which reproducing kernel Hilbert spaces is Pick's theorem true? Integr. Equ. Oper. Theory **16**(2), 244–266 (1993)

21. S. Richter, J. Sunkes, Hankel operators, invariant subspaces, and cyclic vectors in the Drury-Arveson space. Proc. Am. Math. Soc. **144**(6), 2575–2586 (2016)
22. A.L. Shields, D.L. Williams, Bonded projections, duality, and multipliers in spaces of analytic functions. Trans. Am. Math. Soc. **162**, 287–302 (1971)
23. A.L. Shields, D.L. Williams, Bounded projections, duality, and multipliers in spaces of harmonic functions. J. Reine Angew. Math. **299/300**, 256–279 (1978)
24. T.T. Trent, A corona theorem for multipliers on Dirichlet space. Integr. Equ. Oper. Theory **49**(1), 123–139 (2004)
25. R. Zhao, K. Zhu, Theory of Bergman spaces in the unit ball of $\mathbb{C}^n$. Mém. Soc. Math. Fr. (N.S.) (115), vi+103 (2009) 2008

Interpolation Between Hilbert Spaces

Yacin Ameur

Abstract This note comprises a synthesis of certain results in the theory of exact interpolation between Hilbert spaces. In particular, we examine various characterizations of interpolation spaces and their relations to a number of results in operator theory and in function theory.

Keywords Interpolation · Hilbert space · Calderón pair · Pick function · Matrix monotonicity

Mathematics Subject Classification (2010) 46B70, 47A57, 47A63

1 Interpolation Theoretic Notions

1.1 Interpolation Norms

When X, Y are normed spaces, we use the symbol $\mathcal{L}(X;Y)$ to denote the totality of bounded linear maps $T : X \to Y$ with the operator norm

$$\| T \|_{\mathcal{L}(X;Y)} = \sup \{ \| Tx \|_Y \, ; \, \| x \|_X \le 1 \} .$$

When $X = Y$ we simply write $\mathcal{L}(X)$.

Consider a pair of Hilbert spaces $\overline{\mathcal{H}} = (\mathcal{H}_0, \mathcal{H}_1)$ which is *regular* in the sense that $\mathcal{H}_0 \cap \mathcal{H}_1$ is dense in $\mathcal{H}_0$ as well as in $\mathcal{H}_1$. We assume that the pair is *compatible*, i.e., both $\mathcal{H}_i$ are embedded in some common Hausdorff topological vector space $\mathcal{M}$.

Y. Ameur (✉)
Department of Mathematics, Faculty of Science, Lund University, Lund, Sweden
e-mail: Yacin.Ameur@math.lu.se; yacin@maths.lth.se

A. Aleman et al. (eds.), *Analysis of Operators on Function Spaces*,
Trends in Mathematics, https://doi.org/10.1007/978-3-030-14640-5_4

We define the *K-functional*[1] for the couple $\overline{\mathcal{H}}$ by

$$K(t,x) = K\left(t,x;\ \overline{\mathcal{H}}\right) = \inf_{x=x_0+x_1}\{\|x_0\|_0^2 + t\|x_1\|_1^2\}, \qquad t>0,\ x\in\mathcal{M}.$$

The *sum* of the spaces $\mathcal{H}_0$ and $\mathcal{H}_1$ is defined to be the space consisting of all $x\in\mathcal{M}$ such that the quantity $\|x\|_\Sigma^2 := K(1,x)$ is finite; we denote this space by the symbols

$$\Sigma = \Sigma(\overline{\mathcal{H}}) = \mathcal{H}_0 + \mathcal{H}_1.$$

We shall soon see that Σ is a Hilbert space (see Lemma 1.1). The *intersection*

$$\Delta = \Delta(\overline{\mathcal{H}}) = \mathcal{H}_0 \cap \mathcal{H}_1$$

is a Hilbert space under the norm $\|x\|_\Delta^2 := \|x\|_0^2 + \|x\|_1^2$.

A map $T:\Sigma(\overline{\mathcal{H}})\to\Sigma(\overline{\mathcal{K}})$ is called a *couple map* from $\overline{\mathcal{H}}$ to $\overline{\mathcal{K}}$ if the restriction of T to $\mathcal{H}_i$ maps $\mathcal{H}_i$ boundedly into $\mathcal{K}_i$ for $i=0,1$. We use the notations $T\in\mathcal{L}(\overline{\mathcal{H}};\ \overline{\mathcal{K}})$ or $T:\overline{\mathcal{H}}\to\overline{\mathcal{K}}$ to denote that T is a couple map. It is easy to check that $\mathcal{L}(\overline{\mathcal{H}};\ \overline{\mathcal{K}})$ as a Banach space, when equipped with the norm

$$\|T\|_{\mathcal{L}(\overline{\mathcal{H}};\overline{\mathcal{K}})} := \max_{j=0,1}\{\|T\|_{\mathcal{L}(\mathcal{H}_j;\mathcal{K}_j)}\}. \tag{1.1}$$

If $\|T\|_{\mathcal{L}(\overline{\mathcal{H}};\overline{\mathcal{K}})}\le 1$ we speak of a *contraction* from $\overline{\mathcal{H}}$ to $\overline{\mathcal{K}}$.

A Banach space X such that $\Delta\subset X\subset\Sigma$ (continuous inclusions) is called *intermediate* with respect to the pair $\overline{\mathcal{H}}$.

Let X, Y be intermediate spaces with respect to couples $\overline{\mathcal{H}}$, $\overline{\mathcal{K}}$, respectively. We say that X, Y are (relative) *interpolation spaces* if there is a constant C such that $T:\overline{\mathcal{H}}\to\overline{\mathcal{K}}$ implies that $T:X\to Y$ and

$$\|T\|_{\mathcal{L}(X;Y)} \le C\,\|T\|_{\mathcal{L}(\overline{\mathcal{H}};\overline{\mathcal{K}})}\,. \tag{1.2}$$

In the case when $C=1$ we speak about *exact interpolation*. When $\overline{\mathcal{H}}=\overline{\mathcal{K}}$ and $X=Y$ we simply say that X is an (exact) interpolation space with respect to $\overline{\mathcal{H}}$.

Let H be a suitable function of two positive variables and X, Y spaces intermediate to the couples $\overline{\mathcal{H}}$, $\overline{\mathcal{K}}$, respectively. We say that the spaces X, Y are of *type* H (relative to $\overline{\mathcal{H}}$, $\overline{\mathcal{K}}$) if for any positive numbers M_0, M_1 we have

$$\|T\|_{\mathcal{L}(\mathcal{H}_i;\mathcal{K}_i)}\le M_i,\quad i=0,1 \qquad \text{implies} \qquad \|T\|_{\mathcal{L}(X;Y)}\le H(M_0,M_1). \tag{1.3}$$

[1] More precisely, this is the *quadratic version* of the classical Peetre K-functional.

The case $H(x, y) = \max\{x, y\}$ corresponds to exact interpolation, while $H(x, y) = x^{1-\theta} y^{\theta}$ corresponds to the convexity estimate

$$\| T \|_{\mathcal{L}(X;Y)} \leq \| T \|_{\mathcal{L}(\mathcal{H}_0;\mathcal{K}_0)}^{1-\theta} \| T \|_{\mathcal{L}(\mathcal{H}_1;\mathcal{K}_1)}^{\theta} . \tag{1.4}$$

In the situation of (1.4), one says that the interpolation spaces X, Y are of *exponent* θ with respect to the pairs $\overline{\mathcal{H}}$, $\overline{\mathcal{K}}$.

1.2 *K*-Spaces

Given a regular Hilbert couple $\overline{\mathcal{H}}$ and a positive Radon measure ϱ on the compactified half-line $[0, \infty]$ we define an intermediate quadratic norm by

$$\| x \|_*^2 = \| x \|_\varrho^2 = \int_{[0,\infty]} \left(1 + t^{-1}\right) K\left(t, x; \overline{\mathcal{H}}\right) d\varrho(t). \tag{1.5}$$

Here the integrand $k(t) = \left(1 + t^{-1}\right) K(t, x)$ is defined at the points 0 and ∞ by $k(0) = \| x \|_1^2$ and $k(\infty) = \| x \|_0^2$; we shall write $\mathcal{H}_*$ or $\mathcal{H}_\varrho$ for the Hilbert space defined by the norm (1.5).

Let $T \in \mathcal{L}\left(\overline{\mathcal{H}}; \overline{\mathcal{K}}\right)$ and suppose that $\| T \|_{\mathcal{L}(\mathcal{H}_i;\mathcal{K}_i)} \leq M_i$, then

$$K\left(t, Tx; \overline{\mathcal{K}}\right) \leq M_0^2 K\left(M_1^2 t / M_0^2, x; \overline{\mathcal{H}}\right), \quad x \in \Sigma. \tag{1.6}$$

In particular, $M_i \leq 1$ for $i = 0, 1$ implies $\| Tx \|_{\mathcal{K}_\varrho} \leq \| x \|_{\mathcal{H}_\varrho}$ for all $x \in \mathcal{H}_\varrho$. It follows that the spaces $\mathcal{H}_\varrho$, $\mathcal{K}_\varrho$ are exact interpolation spaces with respect to $\overline{\mathcal{H}}$, $\overline{\mathcal{K}}$.

1.2.1 Geometric Interpolation

When the measure ϱ is given by

$$d\varrho(t) = c_\theta \frac{t^{-\theta}}{1+t} dt, \qquad c_\theta = \frac{\pi}{\sin \theta\pi}, \quad 0 < \theta < 1,$$

we denote the norm (1.5) by

$$\| x \|_\theta^2 := c_\theta \int_0^\infty t^{-\theta} K(t, x) \frac{dt}{t}. \tag{1.7}$$

The corresponding space $\mathcal{H}_\theta$ is easily seen to be of exponent θ with respect to $\overline{\mathcal{H}}$. In Sect. 3.1, we will recognize $\mathcal{H}_\theta$ as the geometric interpolation space which has been studied independently by several authors, see [25, 27, 40].

1.3 Pick Functions

Let $\overline{\mathcal{H}}$ be a regular Hilbert couple. The squared norm $\| x \|_1^2$ is a densely defined quadratic form in $\mathcal{H}_0$, which we represent as

$$\| x \|_1^2 = \langle Ax , x \rangle_0 = \| A^{1/2} x \|_0^2$$

where A is a densely defined, positive, injective (perhaps unbounded) operator in $\mathcal{H}_0$. The domain of the positive square-root $A^{1/2}$ is Δ.

Lemma 1.1 *We have in terms of the functional calculus in* $\mathcal{H}_0$

$$K(t, x) = \left\langle \frac{tA}{1+tA} x , x \right\rangle_0 , \quad t > 0. \tag{1.8}$$

In the formula (1.8), we have identified the bounded operator $\frac{tA}{1+tA}$ with its extension to $\mathcal{H}_0$.

Proof Fix $x \in \Delta$. By a straightforward convexity argument, there is a unique decomposition $x = x_{0,t} + x_{1,t}$ which is *optimal* in the sense that

$$K(t, x) = \| x_{0,t} \|_0^2 + t \| x_{1,t} \|_1^2 . \tag{1.9}$$

It follows that $x_{i,t} \in \Delta$ for $i = 0, 1$. Moreover, for all $y \in \Delta$ we have

$$\frac{d}{d\epsilon}\{ \| x_{0,t} + \epsilon y \|_0^2 + t \| x_{1,t} - \epsilon y \|_1^2 \}|_{\epsilon=0} = 0,$$

i.e.,

$$\langle A^{-1/2} x_{0,t} - tA^{1/2} x_{1,t} , A^{1/2} y \rangle_0 = 0, \qquad y \in \Delta .$$

By regularity, we conclude that $A^{-1/2} x_{0,t} = tA^{1/2} x_{1,t}$, whence

$$x_{0,t} = \frac{tA}{1+tA} x \qquad \text{and} \qquad x_{1,t} = \frac{1}{1+tA} x. \tag{1.10}$$

(Note that the operators in (1.10) extend to bounded operators on $\mathcal{H}_0$.) Inserting the relations (1.10) into (1.9), one finishes the proof of the lemma. □

Now fix a positive Radon measure ϱ on $[0, \infty]$. The norm in the space $\mathcal{H}_\varrho$ (see (1.5)) can be written

$$\| x \|_\varrho^2 = \langle h(A) x , x \rangle_0 , \tag{1.11}$$

where

$$h(\lambda) = \int_{[0,\infty]} \frac{(1+t)\lambda}{1+t\lambda}\, d\varrho(t). \tag{1.12}$$

The class of functions representable in this form for some positive Radon measure ϱ is the class P' of *Pick functions, positive and regular on* $\mathbf{R}_+$.

Notice that for the definition (1.11) to make sense, we just need h to be defined on $\sigma(A) \setminus \{0\}$, where $\sigma(A)$ is the spectrum of A. (The value $h(0)$ is irrelevant since A is injective).

A calculus exercise shows that for the space $\mathcal{H}_\theta$ (see (1.7)) we have

$$\| x \|_\theta^2 = \langle A^\theta x \, , \, x \rangle_0. \tag{1.13}$$

1.4 Quadratic Interpolation Norms

Let $\mathcal{H}_*$ be any *quadratic* intermediate space relative to $\overline{\mathcal{H}}$. We write

$$\| x \|_*^2 = \langle Bx \, , \, x \rangle_0$$

where B is a positive injective operator in $\mathcal{H}_0$ (the domain of $B^{1/2}$ is Δ).

For a map $T \in \mathcal{L}(\overline{\mathcal{H}})$ we shall often use the simplified notations

$$\| T \| = \| T \|_{\mathcal{L}(\mathcal{H}_0)} \quad , \quad \| T \|_A = \| T \|_{\mathcal{L}(\mathcal{H}_1)} \quad , \quad \| T \|_B = \| T \|_{\mathcal{L}(\mathcal{H}_*)}.$$

The reader can check the identities

$$\| T \|_A = \| A^{1/2} T A^{-1/2} \| \quad \text{and} \quad \| T \|_B = \| B^{1/2} T B^{-1/2} \|.$$

We shall refer to the following lemma as *Donoghue's lemma*, cf. [13, Lemma 1].

Lemma 1.2 *If $\mathcal{H}_*$ is exact interpolation with respect to $\overline{\mathcal{H}}$, then B commutes with every projection which commutes with A and $B = h(A)$ where h is some positive Borel function on $\sigma(A)$.*

Proof For an orthogonal projection E on $\mathcal{H}_0$, the condition $\| E \|_A \le 1$ is equivalent to that $EAE \le A$, i.e., that E commutes with A. The hypothesis that $\mathcal{H}_*$ be exact interpolation thus implies that every spectral projection of A commutes with B. It now follows from von Neumann's bicommutator theorem that $B = h(A)$ for some positive Borel function h on $\sigma(A)$. □

In view of the lemma, the characterization of the exact quadratic interpolation norms of a given type H reduces to the characterization of functions $h : \sigma(A) \to \mathbf{R}_+$ such that for all $T \in \mathcal{L}\left(\overline{\mathcal{H}}\right)$

$$\| T \| \le M_0 \quad \text{and} \quad \| T \|_A \le M_1 \quad \Rightarrow \quad \| T \|_{h(A)} \le H(M_0, M_1), \tag{1.14}$$

or alternatively,

$$T^*T \le M_0^2 \quad \text{and} \quad T^*AT \le M_1^2 A \quad \Rightarrow \quad T^*h(A)T \le H(M_0, M_1)^2 h(A). \tag{1.15}$$

The set of functions $h : \sigma(A) \to \mathbf{R}_+$ satisfying these equivalent conditions forms a convex cone $C_{H,A}$; its elements are called *interpolation functions of type H relative to A*. In the case when $H(x, y) = \max\{x, y\}$ we simply write C_A for $C_{H,A}$ and speak of *exact interpolation functions relative to A*.

1.5 Exact Calderón Pairs and the K-Property

Given two intermediate normed spaces Y, X relative to $\overline{\mathcal{H}}$, $\overline{\mathcal{K}}$, we say that they are (relatively) *exact K-monotonic* if the conditions

$$x^0 \in X \quad \text{and} \quad K\left(t, y^0; \overline{\mathcal{H}}\right) \le K\left(t, x^0; \overline{\mathcal{K}}\right), \qquad t > 0$$

imply that

$$y^0 \in Y \quad \text{and} \quad \| y^0 \|_Y \le \| x^0 \|_X .$$

It is easy to see that *exact K-monotonicity implies exact interpolation.*

Proof of this. If $\| T \|_{\mathcal{L}(\overline{\mathcal{K}};\overline{\mathcal{H}})} \le 1$, then $\forall x, t$: $K\left(t, Tx; \overline{\mathcal{H}}\right) \le K\left(t, x; \overline{\mathcal{K}}\right)$ whence $\| Tx \|_Y \le \| x \|_X$, by exact K-monotonicity. Hence $\| T \|_{\mathcal{L}(X;Y)} \le 1$. □

Two pairs $\overline{\mathcal{H}}$, $\overline{\mathcal{K}}$ are called *exact relative Calderón pairs* if any two exact interpolation (Banach-) spaces Y, X are exact K-monotonic. Thus, with respect to exact Calderón pairs, exact interpolation is equivalent to exact K-monotonicity. The term "Calderón pair" was coined after thorough investigation of Calderón's study of the pair (L_1, L_∞), see [10, 11].

In our present discussion, it is not convenient to work directly with the definition of exact Calderón pairs. Instead, we shall use the following, closely related notion.

We say that a pair of couples $\overline{\mathcal{H}}$, $\overline{\mathcal{K}}$ has the *relative (exact) K-property* if for all $x^0 \in \Sigma(\overline{\mathcal{K}})$ and $y^0 \in \Sigma(\overline{\mathcal{H}})$ such that

$$K\left(t, y^0; \overline{\mathcal{H}}\right) \le K\left(t, x^0; \overline{\mathcal{K}}\right), \quad t > 0, \tag{1.16}$$

there exists a map $T \in \mathcal{L}(\overline{\mathcal{K}}; \overline{\mathcal{H}})$ such that $Tx^0 = y^0$ and $\| T \|_{\mathcal{L}(\overline{\mathcal{K}};\overline{\mathcal{H}})} \le 1$.

Lemma 1.3 *If $\overline{\mathcal{H}}$, $\overline{\mathcal{K}}$ have the relative K-property, then they are exact relative Calderón pairs.*

Proof Let Y, X be exact interpolation spaces relative to $\overline{\mathcal{H}}$, $\overline{\mathcal{K}}$ and take $x^0 \in X$ and $y^0 \in \Sigma(\overline{\mathcal{H}})$ such that (1.16) holds. By the K-property there is $T : \overline{\mathcal{K}} \to \overline{\mathcal{H}}$ such that $Tx^0 = y^0$ and $\| T \| \le 1$. Then $\| T \|_{\mathcal{L}(X;Y)} \le 1$, and so $\| y^0 \|_Y = \| Tx^0 \|_Y \le \| x^0 \|_X$. We have shown that Y, X are exact K-monotonic. □

In the diagonal case $\overline{\mathcal{H}} = \overline{\mathcal{K}}$, we simply say that $\overline{\mathcal{H}}$ is an *exact Calderón couple* if for intermediate spaces Y, X, the property of being exact interpolation is equivalent to being exact K-monotonic. Likewise, we say that $\overline{\mathcal{H}}$ has the *K-property* if the pair of couples $\overline{\mathcal{H}}$, $\overline{\mathcal{H}}$ has that property.

Remark 1.4 For an operator $T : \overline{\mathcal{K}} \to \overline{\mathcal{H}}$ to be a contraction, it is necessary and sufficient that

$$K\left(t, Tx;\, \overline{\mathcal{H}}\right) \le K\left(t, x;\, \overline{\mathcal{K}}\right), \qquad x \in \Sigma(\overline{\mathcal{K}}),\ t > 0. \tag{1.17}$$

Indeed, the necessity is immediate. To prove the sufficiency it suffices to observe that letting $t \to \infty$ in (1.17) gives $\| Tx \|_0 \le \| x \|_0$, and dividing (1.17) by t, and then letting $t \to 0$, gives that $\| Tx \|_1 \le \| x \|_1$.

2 Mapping Properties of Hilbert Couples

2.1 Main Results

We shall elaborate on the following main result from [2].

Theorem I *Any pair of regular Hilbert couples $\overline{\mathcal{H}}$, $\overline{\mathcal{K}}$ has the relative K-property.*

Before we come to the proof of Theorem I, we note some consequences of it. We first have the following corollary, which shows that a strong form of the K-property is true.

Corollary 2.1 *Let $\overline{\mathcal{H}}$ be a regular Hilbert couple and $x^0, y^0 \in \Sigma$ elements such that*

$$K\left(t, y^0\right) \le M_0^2\, K\left(M_1^2 t / M_0^2\,,\, x^0\right), \qquad t > 0. \tag{2.1}$$

Then

(i) *There exists a map $T \in \mathcal{L}\left(\overline{\mathcal{H}}\right)$ such that $Tx^0 = y^0$ and $\| T \|_{\mathcal{L}(\mathcal{H}_i)} \le M_i$, $i = 0, 1$.*
(ii) *If $x^0 \in X$ where X is an interpolation space of type H, then*

$$\| y^0 \|_X \le H\,(M_0, M_1)\, \| x^0 \|_X.$$

Proof (i) Introduce a new couple $\overline{\mathcal{K}}$ by letting $\| x \|_{\mathcal{K}_i} = M_i \| x \|_{\mathcal{H}_i}$. The relation (2.1) then says that

$$K\left(t, y^0; \overline{\mathcal{H}}\right) \leq K\left(t, x^0; \overline{\mathcal{K}}\right), \qquad t > 0.$$

By Theorem I there is a contraction $T : \overline{\mathcal{K}} \to \overline{\mathcal{H}}$ such that $Tx^0 = y^0$. It now suffices to note that $\| T \|_{\mathcal{L}(\mathcal{H}_i)} = M_i \| T \|_{\mathcal{L}(\mathcal{K}_i;\mathcal{H}_i)}$; (ii) then follows from Lemma 1.3. □

We next mention some equivalent versions of Theorem I, which uses the families of functionals K_p and E_p defined (for $p \geq 1$ and $t, s > 0$) via

$$\begin{aligned} K_p(t) &= K_p(t, x) = K_p\left(t, x; \overline{\mathcal{H}}\right) = \inf_{x = x_0 + x_1} \left\{ \| x_0 \|_0^p + t \| x_1 \|_1^p \right\} \\ E_p(s) &= E_p(s, x) = E_p\left(s, x; \overline{\mathcal{H}}\right) = \inf_{\| x_0 \|_0^p \leq s} \left\{ \| x - x_0 \|_1^p \right\}. \end{aligned} \tag{2.2}$$

Note that $K = K_2$ and that $E_p(s) = E_1\left(s^{1/p}\right)^p$; the E-functionals are used in approximation theory. One has that E_p is decreasing and convex on $\mathbf{R}_+$ and that

$$K_p(t) = \inf_{s>0} \left\{ s + tE_p(s) \right\},$$

which means that K_p is a kind of *Legendre transform* of E_p. The inverse Legendre transformation takes the form

$$E_p(s) = \sup_{t>0} \left\{ \frac{K_p(t)}{t} - \frac{s}{t} \right\}.$$

It is now immediate that, for all $x \in \Sigma\left(\overline{\mathcal{K}}\right)$ and $y \in \Sigma\left(\overline{\mathcal{H}}\right)$, we have

$$K_p(t, y) \leq K_p(t, x), \quad t > 0 \qquad \Leftrightarrow \qquad E_p(s, y) \leq E_p(s, x), \quad s > 0. \tag{2.3}$$

Since moreover $E_p(s) = E_2\left(s^{2/p}\right)^{p/2}$, the conditions in (2.3) are equivalent to that $K(t, y) \leq K(t, x)$ for all $t > 0$. We have shown the following result.

Corollary 2.2 *In Theorem I, one can substitute the K-functional for any of the functionals K_p or E_p.*

Define an exact interpolation norm $\| \cdot \|_{\varrho,p}$ relative to $\overline{\mathcal{H}}$ by

$$\| x \|_{\varrho,p}^p = \int_{[0,\infty]} \left(1 + t^{-1}\right) K_p(t, x)\, d\varrho(t)$$

where ϱ is a positive Radon measure on $[0, \infty]$. This norm is non-quadratic when $p \neq 2$, but is of course equivalent to the quadratic norm corresponding to $p = 2$.

2.2 Reduction to the Diagonal Case

It is not hard to reduce the discussion of Theorem I to a diagonal situation.

Lemma 2.3 *If the K-property holds for regular Hilbert couples in the diagonal case $\overline{\mathcal{H}} = \overline{\mathcal{K}}$, then it holds in general.*

Proof Fix elements $y^0 \in \Sigma(\overline{\mathcal{H}})$ and $x^0 \in \Sigma(\overline{\mathcal{K}})$ such that the inequality (1.16) holds. We must construct a map $T : \overline{\mathcal{K}} \to \overline{\mathcal{H}}$ such that $Tx^0 = y^0$ and $\| T \| \le 1$.

To do this, we form the direct sum $\overline{\mathcal{S}} = (\mathcal{H}_0 \oplus \mathcal{K}_0, \mathcal{H}_1 \oplus \mathcal{K}_1)$. It is clear that $\mathcal{S}_0 + \mathcal{S}_1 = (\mathcal{H}_0 + \mathcal{H}_1) \oplus (\mathcal{K}_0 + \mathcal{K}_1)$, and that

$$K\left(t, x \oplus y; \overline{\mathcal{S}}\right) = K\left(t, x; \overline{\mathcal{H}}\right) + K\left(t, y; \overline{\mathcal{K}}\right).$$

Then

$$K\left(t, 0 \oplus y^0; \overline{\mathcal{S}}\right) \le K\left(t, x^0 \oplus 0; \overline{\mathcal{S}}\right).$$

Hence assuming that the couple $\overline{\mathcal{S}}$ has the K-property, we can assert the existence of a map $S \in \mathcal{L}(\overline{\mathcal{S}})$ such that $S(x^0 \oplus 0) = 0 \oplus y^0$ and $\| S \| \le 1$. Letting $P : \mathcal{S}_0 + \mathcal{S}_1 \to \mathcal{K}_0 + \mathcal{K}_1$ be the orthogonal projection, the assignment $Tx = PS(x \oplus 0)$ now defines a map such that $Tx^0 = y^0$ and $\| T \|_{\mathcal{L}(\overline{\mathcal{H}};\overline{\mathcal{K}})} \le 1$. □

2.3 The Principal Case

The core content of Theorem I is contained in the following statement.

Theorem 2.4 *Suppose that a regular Hilbert couple $\overline{\mathcal{H}}$ is finite dimensional and that all eigenvalues of the corresponding operator A are of unit multiplicity. Then $\overline{\mathcal{H}}$ has the K-property.*

We shall settle for proving Lemma 2.4 in this section, postponing to Sect. 5 the general case of Theorem I.

To prepare for the proof, we write the eigenvalues λ_i of A in the increasing order,

$$\sigma(A) = \{\lambda_i\}_1^n \quad \text{where} \quad 0 < \lambda_1 < \cdots < \lambda_n.$$

Let e_i be corresponding eigenvectors of unit length for the norm of $\mathcal{H}_0$. Then for a vector $x = \sum x_i e_i$ we have

$$\| x \|_0^2 = \sum_1^n |x_i|^2 \qquad , \qquad \| x \|_1^2 = \sum_1^n \lambda_i |x_i|^2.$$

Working in the coordinate system (e_i), the couple $\overline{\mathcal{H}}$ becomes identified with the

n-dimensional weighted ℓ_2 couple

$$\overline{\ell_2^n}(\lambda) := \left(\ell_2^n, \ell_2^n(\lambda)\right),$$

where we write λ for the sequence $(\lambda_i)_1^n$.

We will henceforth identify a vector $x = \sum x_i e_i$ with the point $x = (x_i)_1^n$ in $\mathbf{C}^n$; accordingly, the space $\mathcal{L}\left(\ell_2^n\right)$ is identified with the C^*-algebra $M_n(\mathbf{C})$ of complex $n \times n$ matrices.

It will be convenient to reparametrize the K-functional for the couple $\overline{\ell_2^n}(\lambda)$ and write

$$k_\lambda(t, x) := K\left(1/t, x;\ \overline{\ell_2^n}(\lambda)\right). \tag{2.4}$$

By Lemma 1.1 we have

$$k_\lambda(t, x) = \sum_{i=1}^n \frac{\lambda_i}{t + \lambda_i} |x_i|^2, \qquad x \in \mathbf{C}^n. \tag{2.5}$$

2.4 Basic Reductions

To prove that the couple $\overline{\ell_2^n}(\lambda)$ has the K-property, we introduce an auxiliary parameter $\rho > 1$. The exact value of ρ will change meaning during the course of the argument, the main point being that it can be chosen arbitrarily close to 1.

Initially, we pick any $\rho > 1$ such that $\rho\lambda_i < \lambda_{i+1}$ for all i; we assume also that we are given two elements $x^0, y^0 \in \mathbf{C}^n$ such that

$$k_\lambda\left(t, y^0\right) < \frac{1}{\rho} k_\lambda\left(t, x^0\right), \qquad t \geq 0. \tag{2.6}$$

We must construct a matrix $T \in M_n(\mathbf{C})$ such that

$$Tx^0 = y^0 \quad \text{and} \quad k_\lambda(t, Tx) \leq k_\lambda(t, x), \quad x \in \mathbf{C}^n,\ t > 0. \tag{2.7}$$

Define $\tilde{x}^0 = (|x_i^0|)_1^n$ and $\tilde{y}^0 = (|y_i^0|)_1^n$ and suppose that

$$k_\lambda(t, \tilde{y}^0) < \frac{1}{\rho} k_\lambda(t, \tilde{x}^0), \qquad t \geq 0.$$

Suppose that we can find an operator $T_0 \in M_n(\mathbf{C})$ such that $T_0\tilde{x}^0 = \tilde{y}^0$ and $k_\lambda(t, T_0 x) < k_\lambda(t, x)$ for all $x \in \mathbf{C}^n$ and $t > 0$. Writing $x_k^0 = e^{i\theta_k}\tilde{x}_k^0$ and $y_k^0 = e^{i\varphi_k}\tilde{y}_k^0$ where $\theta_k, \varphi_k \in \mathbf{R}$, we then have $Tx^0 = y^0$ and $k_\lambda(t, Tx) < k_\lambda(t, x)$ where

$$T = \operatorname{diag}(e^{i\varphi_k}) T_0 \operatorname{diag}(e^{-i\theta_k}).$$

Replacing x^0, y^0 by $\tilde{x}^0$, $\tilde{y}^0$ we can thus assume that the coordinates x_i^0 and y_i^0 are non-negative; replacing them by small perturbations if necessary, we can assume that they are strictly positive, at the expense of slightly diminishing the number ρ.

Now put $\beta_i = \lambda_i$ and $\alpha_i = \rho\lambda_i$. Our assumption on ρ means that

$$0 < \beta_1 < \alpha_1 < \cdots < \beta_n < \alpha_n.$$

Using the explicit expression for the K-functional, it is plain to check that

$$k_\beta(t, x) \le k_\alpha(t, x) \le \rho k_\beta(t, x), \qquad x \in \mathbf{C}^n,\ t \ge 0.$$

Our assumption (2.6) therefore implies that

$$k_\alpha(t, y^0) < k_\beta(t, x^0), \quad t \ge 0. \tag{2.8}$$

We shall verify the existence of a matrix $T = T_\rho = T_{\rho,x^0,y^0}$ such that

$$Tx^0 = y^0 \quad \text{and} \quad k_\alpha(t, Tx) \le k_\beta(t, x), \qquad x \in \mathbf{C}^n,\ t > 0. \tag{2.9}$$

It is clear by compactness that, as $\rho \downarrow 1$, the corresponding matrices T_ρ will cluster at some point T satisfying $Tx^0 = y^0$ and $\| T \|_{\mathcal{L}(\overline{\mathcal{H}})} \le 1$. (See Remark 1.4.)

In conclusion, the proof of Theorem 2.4 will be complete when we can construct a matrix T satisfying (2.9) with ρ arbitrarily close to 1.

2.5 Construction of *T*

Let $\mathcal{P}_k$ denote the linear space of complex polynomials of degree at most k. We shall use the polynomials

$$L_\alpha(t) = \prod_1^n (t + \alpha_i) \quad , \quad L_\beta(t) = \prod_1^n (t + \beta_i),$$

and the product $L = L_\alpha L_\beta$. Notice that

$$L'(-\alpha_i) < 0 \quad , \quad L'(-\beta_i) > 0. \tag{2.10}$$

Recalling the formula (2.5), it is clear that we can define a real polynomial $P \in \mathcal{P}_{2n-1}$ by

$$\frac{P(t)}{L(t)} = k_\beta\left(t, x^0\right) - k_\alpha\left(t, y^0\right). \tag{2.11}$$

Clearly $P(t) > 0$ when $t \ge 0$. Moreover, a consideration of the residues at the poles of the right-hand member shows that P is uniquely defined by the values

$$P\left(-\beta_i\right) = (x_i^0)^2 \beta_i L'\left(-\beta_i\right) \quad , \quad P\left(-\alpha_i\right) = -(y_i^0)^2 \alpha_i L'\left(-\alpha_i\right). \tag{2.12}$$

Combining with (2.10), we conclude that

$$P\left(-\alpha_i\right) > 0 \quad \text{and} \quad P\left(-\beta_i\right) > 0. \tag{2.13}$$

Perturbing the problem slightly, it is clear that we can assume that P has exact degree $2n - 1$, and that all zeros of P have multiplicity 1. (We here diminish the value of $\rho > 1$ somewhat, if necessary.)

Now, P has $2n - 1$ simple zeros, which we split according to

$$P^{-1}\left(\{0\}\right) = \{-r_i\}_{i=1}^{2m-1} \cup \{-c_i, -\bar{c}_i\}_{i=1}^{n-m},$$

where the r_i are positive and the c_i are non-real, and chosen to have positive imaginary parts. The following is the key observation.

Lemma 2.5 *We have that*

$$L'\left(-\beta_i\right) P\left(-\beta_i\right) > 0 \quad , \quad L'\left(-\alpha_i\right) P\left(-\alpha_i\right) < 0 \tag{2.14}$$

and there is a splitting $\{r_i\}_{i=1}^{2m-1} = \{\delta_i\}_{i=1}^{m} \cup \{\gamma_i\}_{i=1}^{m-1}$ *such that*

$$L\left(-\delta_j\right) P'\left(-\delta_j\right) > 0 \quad , \quad L\left(-\gamma_k\right) P'\left(-\gamma_k\right) < 0. \tag{2.15}$$

Proof The inequalities (2.14) follow immediately from (2.13) and (2.10). It remains to prove (2.15).

Let $-h$ denote the leftmost real zero of the polynomial LP (of degree $4n - 1$). We claim that $P(-h) = 0$. If this were not the case, we would have $h = \alpha_n$. Since the degree of P is odd, $P(-t)$ is negative for large values of t, and so $P(-\alpha_n) < 0$ contradicting (2.13). We have shown that $P(-h) = 0$. Since all zeros of LP have multiplicity 1, we have $(LP)'(-h) \neq 0$, whence

$$L(-h)P'(-h) = (LP)'(-h) > 0.$$

We write $\delta_m = h$ and put $P_*(t) = P(t)/(t + \delta_m)$. Since $t + \delta_m > 0$ for $t \in \{-\alpha_i, -\beta_i\}_1^n$, we have by (2.13) that for all i

$$P_*(-\alpha_i) > 0 \quad \text{and} \quad P_*(-\beta_i) > 0.$$

Denote by $\{-r_j{}^*\}_{j=1}^{2m-2}$ the real zeros of P_*. Since the degree of LP_* is even and the polynomial $(LP_*)'$ has alternating signs in the set $\{-\alpha_i, -\beta_i\}_{i=1}^n \cup \{-r_i{}^*\}_{i=1}^{2m-2}$, we can split the zeros of P_* as $\{-\delta_i, -\gamma_i\}_{i=1}^{m-1}$, where

$$L(-\delta_i)P_*'(-\delta_i) > 0 \quad , \quad L(-\gamma_i)P_*'(-\gamma_i) < 0. \tag{2.16}$$

Since $P'(-r_j{}^*) = (\delta_m - r_j{}^*)P'_*(-r_j{}^*)$ and $\delta_m > r_j{}^*$, the signs of $P'(-r_j{}^*)$ and $P'_*(-r_j{}^*)$ are equal, proving (2.15). □

Recall that $\{-c_i\}_1^{n-m}$ denote the zeros of P such that $\operatorname{Im} c_i > 0$. We put (with the convention that an empty product equals 1)

$$L_\delta(t) = \prod_{i=1}^{m}(t+\delta_i) \quad , \quad L_\gamma(t) = \prod_{i=1}^{m-1}(t+\gamma_i) \quad , \quad L_c(t) = \prod_{i=1}^{n-m}(t+c_i).$$

We define a linear map $F : \mathbf{C}^{n+m} \to \mathbf{C}^{n+m-1}$ in the following way. First define a subspace $U \subset \mathcal{P}_{2n-1}$ by

$$U = \{\, L_c q \,;\, q \in \mathcal{P}_{n+m-1} \,\}.$$

Notice that U has dimension $n+m-1$ and that $P \in U$; in fact $P = aL_cL_c^*L_\delta L_\gamma$ where a is the leading coefficient and the $*$-operation is defined by $L^*(z) = \overline{L(\bar{z})}$.

For a polynomial $Q \in U$ we have

$$\begin{aligned} \frac{|\,Q(t)\,|^2}{L(t)P(t)} &= \sum_{i=1}^{n}|x_i|^2\frac{\beta_i}{t+\beta_i} + \sum_{i=1}^{n}|x_i'|^2\frac{\delta_i}{t+\delta_i} \\ &\quad - \sum_{i=1}^{n}|y_i|^2\frac{\alpha_i}{t+\alpha_i} - \sum_{i=1}^{m-1}|y_i'|^2\frac{\gamma_i}{t+\gamma_i}, \end{aligned} \tag{2.17}$$

where, for definiteness,

$$x_i = \frac{Q(-\beta_i)}{\sqrt{\beta_i L'(-\beta_i)P(-\beta_i)}} \quad ; \quad x_j' = \frac{Q(-\delta_j)}{\sqrt{\delta_j L'(-\delta_j)P(-\delta_j)}} \tag{2.18}$$

$$y_i = \frac{Q(-\alpha_i)}{\sqrt{-\alpha_i L'(-\alpha_i)P(-\alpha_i)}} \quad ; \quad y_j' = \frac{Q(-\gamma_j)}{\sqrt{-\gamma_j L'(-\gamma_j)P(-\gamma_j)}}. \tag{2.19}$$

The identities in (2.18) give rise to a linear map

$$M : \mathbf{C}^n \oplus \mathbf{C}^m \to U \quad ; \quad \left[x; x'\right] \mapsto Q. \tag{2.20}$$

We can similarly regard (2.19) as a linear map

$$N : U \to \mathbf{C}^n \oplus \mathbf{C}^{m-1} \quad ; \quad Q \mapsto \left[y; y'\right]. \tag{2.21}$$

Our desired map F is defined as the composite

$$F = NM : \mathbf{C}^n \oplus \mathbf{C}^m \to \mathbf{C}^n \oplus \mathbf{C}^{m-1} \quad ; \quad [x; x'] \mapsto [y; y'].$$

Notice that if $Q = M\left[x; x'\right]$ and $\left[y; y'\right] = F\left[x; x'\right]$ then (2.17) means that

$$k_{\beta\oplus\delta}\left(t, \left[x; x'\right]\right) - k_{\alpha\oplus\gamma}\left(t, F\left[x; x'\right]\right) = \frac{|\,Q(t)\,|^2}{L(t)P(t)} \geq 0, \qquad t \geq 0.$$

This implies that F is a contraction from $\overline{\ell_2^{n+m}}(\beta \oplus \delta)$ to $\overline{\ell_2^{n+m-1}}(\alpha \oplus \gamma)$.

We now define T as a "compression" of F. Namely, let $E : \mathbf{C}^n \oplus \mathbf{C}^{m-1} \to \mathbf{C}^n$ be the projection onto the first n coordinates, and define an operator T on $\mathbf{C}^n$ by

$$Tx = EF\,[x; 0]\,, \qquad x \in \mathbf{C}^n.$$

Taking $Q = P$ in (2.17) we see that $Tx^0 = y^0$. Moreover,

$$\begin{aligned} k_\beta\,(t, x) - k_\alpha\,(t, Tx) &= \sum_{i=1}^n |x_i|^2 \frac{\beta_i}{t + \beta_i} - \sum_{i=1}^n |y_i|^2 \frac{\alpha_i}{t + \alpha_i} \\ &\geq \sum_{i=1}^n |x_i|^2 \frac{\beta_i}{t + \beta_i} - \sum_{i=1}^n |y_i|^2 \frac{\alpha_i}{t + \alpha_i} - \sum_{j=1}^{m-1} |y_i'|^2 \frac{\gamma_i}{t + \gamma_i} \\ &= k_{\beta\oplus\delta}\,(t, [x; 0]) - k_{\alpha\oplus\gamma}\,(t, F\,[x; 0]) = \frac{|\,Q(t)\,|^2}{L(t)P(t)}. \end{aligned}$$

Since the right-hand side is non-negative, we have shown that

$$k_\alpha\,(t, Tx) \leq k_\beta(t, x), \quad t > 0,\ x \in \mathbf{C}^n,$$

as desired. The proof of Theorem 2.4 is finished. □

2.6 Real Scalars

Theorem 2.4 holds also in the case of Euclidean spaces over the real scalar field. To see this, assume without loss of generality that the vectors $x^0, y^0 \in \mathbf{C}^n$ have *real entries* (still satisfying $k_\lambda\left(t, y^0\right) \leq k_\lambda\left(t, x^0\right)$ for all $t > 0$).

By Theorem 2.4 we can find a (complex) contraction T of $\overline{\ell_2^n}(\lambda)$ such that $Tx^0 = y^0$. It is clear that the operator T^* defined by $T^*x = \overline{T\,(\bar{x})}$ satisfies those same conditions. Replacing T by $\frac{1}{2}\,(T + T^*)$ we obtain a real matrix $T \in M_n(\mathbf{R})$, which is a contraction of $\overline{\ell_2^n}(\lambda)$ and maps x^0 to y^0. □

2.7 Explicit Representations

We here deduce an explicit representation for the operator T constructed above.

Let x^0 and y^0 be two non-negative vectors such that

$$k_\lambda\left(t, y^0\right) \le k_\lambda\left(t, x^0\right), \qquad t > 0.$$

For small $\rho > 0$ we perturb x^0, y^0 slightly to vectors $\tilde{x}^0$, $\tilde{y}^0$ which satisfy the conditions imposed on the previous subsections. We can then construct a matrix $T = T_\rho$ such that

$$T\tilde{x}^0 = \tilde{y}^0 \quad \text{and} \quad k_\alpha(t, Tx) \le k_\beta(t, x), \qquad t > 0,\ x \in \mathbf{C}^n, \tag{2.22}$$

where $\beta = \lambda$ and $\alpha = \rho\lambda$. As ρ, $\tilde{x}^0$, $\tilde{y}^0$ approaches 1, x^0, respectively y^0, it is clear that any cluster point T of the set of contractions T_ρ will satisfy

$$Tx^0 = y^0 \quad \text{and} \quad k_\lambda(t, Tx) \le k_\lambda(t, x), \quad t > 0,\ x \in \mathbf{C}^n.$$

Theorem 2.6 *The matrix* $T = T_\varrho = (\tau_{ik})_{i,k=1}^n$ *where*

$$\tau_{ik} = \operatorname{Re}\left[\frac{1}{\alpha_i - \beta_k}\frac{\tilde{x}_k^0}{\tilde{y}_i^0}\frac{\beta_k L_\delta(-\alpha_i)L_c(-\alpha_i)L_\alpha(-\beta_k)}{\alpha_i L_\delta(-\beta_k)L_c(-\beta_k)L'_\alpha(-\alpha_i)}\right] \tag{2.23}$$

satisfies (2.22).

Proof The range of the map $\mathbf{C}^n \to U$, $x \mapsto M[x; 0]$ (see 2.20) is precisely the n-dimensional subspace

$$V := L_\delta L_c \cdot \mathcal{P}_{n-1} = \{L_\delta L_c R;\ R \in \mathcal{P}_{n-1}\} \subset U. \tag{2.24}$$

We introduce a basis $(Q_k)_{k=1}^n$ for V by

$$Q_k(t) = \frac{L_\delta(t)L_c(t)L_\beta(t)}{t + \beta_k}\frac{\sqrt{\beta_k L'(-\beta_k)P(-\beta_k)}}{L_\delta(-\beta_k)L_c(-\beta_k)L'_\beta(-\beta_k)}.$$

Then

$$\frac{Q_k(-\beta_i)}{\sqrt{\beta_i L'(-\beta_i)P(-\beta_i)}} = \begin{cases} 1 & i = k, \\ 0 & i \ne k. \end{cases}$$

Denoting by (e_i) the canonical basis in $\mathbf{C}^n$ and using (2.18), (2.19) we get

$$\begin{aligned}\tau_{ik} = (Te_k)_i &= \frac{Q_k(-\alpha_i)}{\sqrt{\alpha_i L'(-\alpha_i)P(-\alpha_i)}}\\ &= \frac{1}{\beta_k - \alpha_i}\frac{L_\delta(-\alpha_i)L_c(-\alpha_i)L_\beta(-\alpha_i)}{L_\delta(-\beta_k)L_c(-\beta_k)L'_\beta(-\beta_k)}\left(\frac{\beta_k L'(-\beta_k)P(-\beta_k)}{-\alpha_i L'(-\alpha_i)P(-\alpha_i)}\right)^{1/2}.\end{aligned}$$

Inserting the expressions (2.12) for $P(-\alpha_i)$ and $P(-\beta_k)$ and taking real parts (see the remarks in Sect. 2.6), we obtain the formula (2.23). □

Remark 2.7 It is easy to see that, if we pick all matrix-elements real, some elements τ_{ik} of the matrix T in (2.23) will be negative, even while the numbers x_i^0 and y_k^0 are positive. It was proved in [2], Theorem 2.3, that this is necessarily so. Indeed, one there constructs an example of a five-dimensional couple $\overline{\ell_2^5}(\lambda)$ and two vectors $x^0, y^0 \in \mathbf{R}^5$ having non-negative entries such that *no* contraction $T = (\tau_{ik})_{i,k=1}^5$ on $\overline{\ell_2^5}(\lambda)$ having all matrix entries $\tau_{ik} \geq 0$ can satisfy $Tx^0 = y^0$. On the other hand, if one settles for using a matrix with $\| T \| \leq \sqrt{2}$, then it is possible to find one with only non-negative matrix entries. Indeed, such a matrix was used by Sedaev [35], see also [38].

2.8 On Sharpness of the Norm-Bounds

We shall show that if $m < n$ (i.e., if the polynomial P has at least one non-real zero), then the norm $\| T \|_{\mathcal{L}(\mathcal{H}_i)}$ of the contraction T constructed above is very close to 1 for $i = 0, 1$.

We first claim that $\| T \|_{\mathcal{L}(\mathcal{H}_0)} = 1$. To see this, we notice that if $m < n$, then there is a non-trivial polynomial $Q^{(1)}$ in the space V (see (2.24)) which vanishes at the points $0, \gamma_1, \ldots, \gamma_{m-1}$. If $x_i^{(1)}$ and $y_i^{(1)}$ are defined by the formulas (2.18) and (2.19) (while $(x_j^{(1)})' = (y_k^{(1)})' = 0$), we then have $Tx^{(1)} = y^{(1)}$ and

$$k_\beta(t, x^{(1)}) - k_\alpha(t, y^{(1)}) = \frac{|\, Q^{(1)}(t)\,|^2}{L(t)P(t)}, \qquad t > 0.$$

Choosing $t = 0$ we conclude that $\| x^{(1)} \|_{\ell_2^n}^2 - \| Tx^{(1)} \|_{\ell_2^n}^2 = 0$, whence $\| T \|_{\mathcal{L}(\mathcal{H}_0)} \geq 1$, proving our claim.

Similarly, the condition $m < n$ implies the existence of a polynomial $Q^{(2)} \in V$ of degree at most $n + m - 2$ vanishing at the points $\gamma_1, \ldots, \gamma_{m-1}$. Constructing vectors $x^{(2)}, y^{(2)}$ via (2.18) and (2.19) we will have $Tx^{(2)} = y^{(2)}$ and

$$k_\beta(t, x^{(2)}) - k_\alpha(t, y^{(2)}) = \frac{|\, Q^{(2)}(t)\,|^2}{L(t)P(t)}, \qquad t > 0.$$

Multiplying this relation by t and then sending $t \to \infty$, we find that $\| x^{(2)} \|^2_{\ell_2^n(\beta)} - \| Tx^{(2)} \|^2_{\ell_2^n(\alpha)} = 0$, which implies $\| T \|_{\mathcal{L}(\mathcal{H}_1)} \geq \rho^{-1/2}$.

2.9 A Remark on Weighted ℓ_p-Couples

As far as we are aware, if $1 < p < \infty$ and $p \neq 2$, it is still an open question whether the couple $\overline{\ell_p^n}(\lambda) = \left(\ell_p^n, \ell_p^n(\lambda)\right)$ is an exact Calderón couple or not. (When $p = 1$ or $p = \infty$ it is exact Calderón; see [36] for the case $p = 1$; the case $p = \infty$ is essentially just the Hahn–Banach theorem.)

It is well known, and easy to prove, that the K_p-functional (see (2.2)) corresponding to the couple $\overline{\ell_p^n}(\lambda)$ is given by the explicit formula

$$K_p\left(t, x; \overline{\ell_p^n}(\lambda)\right) = \sum_{i=1}^n |x_i|^p \frac{t\lambda_i}{(1 + (t\lambda_i)^{\frac{1}{p-1}})^{p-1}}.$$

It was proved by Sedaev [35] (cf. [38]) that if $K_p\left(t, y^0; \overline{\ell_p^n}(\lambda)\right) \leq K_p\left(t, x^0; \overline{\ell_p^n}(\lambda)\right)$ for all $t > 0$ then there is $T : \overline{\ell_p^n}(\lambda) \to \overline{\ell_p^n}(\lambda)$ of norm at most $2^{1/p'}$ such that $Tx^0 = y^0$. (Here p' is the exponent conjugate to p.)

Although our present estimates are particular for the case $p = 2$, our construction still shows that, if we re-define $P(t)$ to be the polynomial

$$\frac{P(t)}{L(t)} = \sum_1^n (\tilde{x}_i^0)^p \frac{\beta_i}{t + \beta_i} - \sum_1^n (\tilde{y}_i^0)^p \frac{\alpha_i}{t + \alpha_i}, \tag{2.25}$$

then the matrix T defined by

$$\tau_{ik} = \operatorname{Re}\left[\frac{1}{\alpha_i - \beta_k} \frac{(\tilde{x}_k^0)^{p-1}}{(\tilde{y}_i^0)^{p-1}} \frac{\beta_k L_\delta(-\alpha_i) L_c(-\alpha_i) L_\alpha(-\beta_k)}{\alpha_i L_\delta(-\beta_k) L_c(-\beta_k) L'_\alpha(-\alpha_i)}\right] \tag{2.26}$$

will satisfy $T\tilde{x}^0 = \tilde{y}^0$, at least, provided that $P(t) > 0$ when $t \geq 0$. (Here L_δ and L_c are constructed from the zeros of P as in the case $p = 2$.)

The matrix (2.26) differs from those used by Sedaev [35] and Sparr [38]. Indeed the matrices from [35, 38] have *non-negative entries*, while this is not so for the matrices (2.26). It seems to be an interesting problem to estimate the norm $\| T \|_{\mathcal{L}(\overline{\ell_p^n}(\lambda))}$ for the matrix (2.26), when $p \neq 2$. The motivation for this type of question is somewhat elaborated in Sect. 6.7, but we shall not discuss it further here.

2.10 A Comparison with Löwner's Matrix

In this subsection, we briefly explain how our matrix T is related to the matrix used by Löwner [26] in his original work on monotone matrix functions.[2]

We shall presently display four kinds of partial isometries; Löwner's matrix will be recognized as one of them. In all cases, operators with the required properties can alternatively be found using the more general construction in Theorem 2.4.

The following discussion was inspired by the earlier work of Sparr [39], who seems to have been the first to note that Löwner's matrix could be constructed in a similar way.

In this subsection, scalars are assumed to be real. In particular, when we write "ℓ_2^n" we mean the (real) Euclidean n-dimensional space.

Suppose that two vectors $x^0, y^0 \in \mathbf{R}^n$ satisfy

$$k_\lambda\left(t, y^0\right) \leq k_\lambda\left(t, x^0\right), \qquad t > 0.$$

Let

$$L_\lambda(t) = \prod_1^n (t + \lambda_i),$$

and let $P \in \mathcal{P}_{n-1}$ be the polynomial fulfilling

$$\frac{P(t)}{L_\lambda(t)} = k_\lambda\left(t, x^0\right) - k_\lambda\left(t, y^0\right) = \sum_{i=1}^n \frac{\lambda_i}{t + \lambda_i}\left[(x_i^0)^2 - (y_i^0)^2\right].$$

By assumption, $P(t) \geq 0$ for $t \geq 0$. Moreover, P is uniquely determined by the n conditions

$$P(-\lambda_i) = \frac{(x_i^0)^2 - (y_i^0)^2}{\lambda_i L_\lambda'(-\lambda_i)}.$$

Let $u_1, v_1, u_2, v_2, \ldots$ denote the canonical basis of ℓ_2^n and let

$$\ell_2^n = O \oplus E$$

be the corresponding splitting, i.e.,

$$O = \text{span}\{u_i\} \quad , \quad E = \text{span}\{v_i\}.$$

[2] By "Löwner's matrix," we mean the unitary matrix denoted "V" in Donoghue's book [12], on p. 71. A more explicit construction of this matrix is found in [26], where it is called "T."

Notice that

$$\dim O = \lfloor (n-1)/2 \rfloor + 1 \quad , \quad \dim E = \lfloor (n-2)/2 \rfloor + 1,$$

where $\lfloor x \rfloor$ is the integer part of a real number x.

We shall construct matrices $T \in M_n(\mathbf{R})$ such that

$$Tx^0 = y^0 \quad \text{and} \quad k_\lambda(t, Tx) \le k_\lambda(t, x), \quad t > 0,\ x \in \mathbf{R}^n, \tag{2.27}$$

in the following special cases:

(1) $P(t) = q(t)^2$ where $q \in \mathcal{P}_{(n-1)/2}(\mathbf{R})$, $x^0 \in O$, and $y^0 \in E$,
(2) $P(t) = tq(t)^2$ where $q \in \mathcal{P}_{(n-2)/2}(\mathbf{R})$, $x^0 \in E$, and $y^0 \in O$.

Here $\mathcal{P}_x$ should be interpreted as $\mathcal{P}_{\lfloor x \rfloor}$.

Remark 2.8 In this connection, it is interesting to recall the well-known fact that any polynomial P which is non-negative on $\mathbf{R}_+$ can be written $P(t) = q_0(t)^2 + tq_1(t)^2$ for some real polynomials q_0 and q_1.

To proceed with the solution, we rename the λ_i as $\lambda_i = \xi_i$ when i is odd and $\lambda_i = \eta_i$ when i is even. We also write

$$L_\xi(t) = \prod_{i \text{ odd}} (t + \xi_i) \quad , \quad L_\eta(t) = \prod_{i \text{ even}} (t + \eta_i),$$

and write $L = L_\xi L_\eta$. Notice that $L'_\lambda(-\xi_i) > 0$ and $L'_\lambda(-\eta_i) < 0$.

2.10.1 Case 1

Suppose that $P(t) = q(t)^2$, $q \in \mathcal{P}_{(n-1)/2}(\mathbf{R})$, $x^0 \in O$, and $y^0 \in E$. Then

$$\frac{q(t)^2}{L_\lambda(t)} = \sum_{k \text{ odd}} \frac{\xi_k}{t + \xi_k} (x_k^0)^2 - \sum_{i \text{ even}} \frac{\eta_i}{t + \eta_i} (y_i^0)^2,$$

where

$$x_k^0 = \frac{\varepsilon_k q(-\xi_k)}{\sqrt{\xi_k L'_\lambda(-\xi_k)}} \quad , \quad y_i^0 = \frac{\zeta_i q(-\eta_i)}{\sqrt{-\eta_i L'_\lambda(-\eta_i)}} \tag{2.28}$$

for some choice of signs $\varepsilon_k, \zeta_i \in \{\pm 1\}$.

By (2.28) are defined linear maps

$$O \to \mathcal{P}_{(n-1)/2}(\mathbf{R}) \quad : \quad x \mapsto Q \quad ; \quad \mathcal{P}_{(n-1)/2}(\mathbf{R}) \to E \quad : \quad Q \mapsto y.$$

The composition is a linear map

$$T_0 : O \to E \quad : \quad x \mapsto y.$$

We now define $T \in M_n(\mathbf{R})$ by

$$T : O \oplus E \to O \oplus E \quad : \quad [x; v] \mapsto [0; T_0 x].$$

Then clearly $Tx^0 = y^0$ and

$$\begin{aligned} k_\lambda(t, [x; v]) - k_\lambda(t, T[x; v]) \\ &\geq k_\xi(t, x) - k_\eta(t, T_0 x) \\ &= \frac{Q(t)^2}{L_\lambda(t)} \geq 0, \quad t > 0,\ x \in O,\ v \in E. \end{aligned} \tag{2.29}$$

We have verified (2.27) in case 1. A computation similar to the one in the proof of Theorem 2.6 shows that, with respect to the bases u_k and v_i,

$$(T_0)_{ik} = \frac{\varepsilon_k \zeta_i}{\xi_k - \eta_i} \frac{L_\xi(-\eta_i)}{L'_\xi(-\xi_k)} \left(\frac{\xi_k L'_\xi(-\xi_k) L_\eta(-\xi_k)}{-\eta_i L_\xi(-\eta_i) L'_\eta(-\eta_i)} \right)^{1/2}.$$

Notice that, multiplying (2.29) by t, then letting $t \to \infty$ implies that

$$\sum_{k \text{ odd}} x_k^2 \xi_k - \sum_{i \text{ even}} (T_0 x)_i^2 \eta_i = 0.$$

This means that T *is a partial isometry from* O *to* E *with respect to the norm of* $\ell_2^n(\lambda)$.

2.10.2 Case 2

Now assume that $P(t) = tq(t)^2$, $q \in \mathcal{P}_{(n-2)/2}(\mathbf{R})$, $x^0 \in E$, and $y^0 \in O$. Then

$$\frac{tq(t)^2}{L_\lambda(t)} = -\sum_{i \text{ odd}} (y_i^0)^2 \frac{\xi_i}{t + \xi_i} + \sum_{k \text{ even}} \frac{\eta_k}{t + \eta_k} (x_k^0)^2,$$

where

$$y_i^0 = \frac{\varepsilon'_i q(-\xi_i)}{\sqrt{L'_\lambda(-\xi_i)}} \quad , \quad x_k^0 = \frac{-\zeta'_k q(-\eta_k)}{\sqrt{-L'_\lambda(-\eta_k)}} \tag{2.30}$$

for some $\varepsilon'_i, \zeta'_k \in \{\pm 1\}$.

By (2.30) are defined linear maps

$$E \to \mathcal{P}_{(n-2)/2}(\mathbf{R}) \quad : \quad x \mapsto Q \quad ; \quad \mathcal{P}_{(n-2)/2}(\mathbf{R}) \to O \quad : \quad Q \mapsto y.$$

We denote their composite by

$$T_1 : E \to O \quad : \quad x \mapsto y.$$

Define $T \in M_n(\mathbf{R})$ by

$$T : O \oplus E \to O \oplus E \quad : \quad [u; x] \mapsto [T_1 x; 0].$$

We then have

$$\begin{aligned} -k_\lambda\,(t, T[u; x]) + k_\lambda\,(t, [u; x]) & \\ &\geq -k_\xi\,(t, T_1 x) + k_\eta(t, x) \\ &= \frac{t Q(t)^2}{L_\lambda(t)} \geq 0, \quad t > 0,\ u \in O,\ x \in E, \end{aligned} \tag{2.31}$$

and (2.27) is verified also in case 2.

A computation shows that, with respect to the bases v_k and u_i,

$$(T_1)_{ik} = \frac{\varepsilon_i' \zeta_k'}{\eta_k - \xi_i} \frac{L_\eta(-\xi_i)}{L_\eta'(-\eta_k)} \left(\frac{-L_\xi(-\eta_k) L_\eta'(-\eta_k)}{L_\xi'(-\xi_i) L_\eta(-\xi_i)} \right)^{1/2}.$$

Inserting $t = 0$ in (2.31) we find that

$$-\sum_{i \text{ odd}} (T_1 x)_i^2 + \sum_{k \text{ even}} (x_k)^2 = 0,$$

i.e., *T is a partial isometry form E to O with respect to the norm of ℓ_2^n.*

In the case of even n, the matrix T_1 coincides with Löwner's matrix.

3 Quadratic Interpolation Spaces

3.1 A Classification of Quadratic Interpolation Spaces

Recall that an intermediate space X with respect to $\overline{\mathcal{H}}$ is said to be of *type* H if $\| T \|_{\mathcal{L}(\mathcal{H}_i)} \leq M_i$ for $i = 0, 1$ implies that $\| T \|_{\mathcal{L}(X)} \leq H(M_0, M_1)$. We shall henceforth make a mild restriction, and assume that H be homogeneous of degree one. This means that we can write

$$H(s,t)^2 = s^2\,\mathbf{H}(t^2/s^2) \tag{3.1}$$

for some function $\mathbf{H}$ of one positive variable. In this situation, we will say that X is of type $\mathbf{H}$. The definition is chosen so that the estimates $\| T \|^2_{\mathcal{L}(\mathcal{H}_i)} \le M_i$ for $i = 0, 1$ imply $\| T \|^2_{\mathcal{L}(X)} \le M_0\,\mathbf{H}\,(M_1/M_0)$.

In the following we will make the *standing assumptions*: $\mathbf{H}$ is an increasing, continuous, and positive function on $\mathbf{R}_+$ with $\mathbf{H}(1) = 1$ and $\mathbf{H}(t) \le \max\{1, t\}$.

Notice that our assumptions imply that all spaces of type $\mathbf{H}$ are exact interpolation. Note also that $\mathbf{H}(t) = t^\theta$ corresponds to geometric interpolation of exponent θ.

Suppose now that $\overline{\mathcal{H}}$ is a regular Hilbert couple and that $\mathcal{H}_*$ is an exact interpolation space with corresponding operator B. By Donoghue's lemma, we have that $B = h(A)$ for some positive Borel function h on $\sigma(A)$.

The statement that $\mathcal{H}_*$ is intermediate relative to $\overline{\mathcal{H}}$ is equivalent to that

$$c_1 \frac{A}{1+A} \le B \le c_2(1+A) \tag{3.2}$$

for some positive numbers c_1 and c_2.

Let us momentarily assume that $\mathcal{H}_0$ be *separable*. (This restriction is removed in Remark 3.1.) We can then define the *scalar-valued spectral measure* of A,

$$\nu_A(\omega) = \sum 2^{-k}\,\langle E(\omega)e_k, e_k\rangle_0$$

where E is the spectral measure of A, $\{e_k;\ k = 1, 2, \ldots\}$ is an orthonormal basis for $\mathcal{H}_0$, and ω is a Borel set. Then, for Borel functions h_0, h_1 on $\sigma(A)$, one has that $h_1 = h_2$ almost everywhere with respect to ν_A if and only if $h_1(A) = h_2(A)$.

Note that the regularity of $\overline{\mathcal{H}}$ means that $\nu_A(\{0\}) = 0$.

Theorem II *If $\mathcal{H}_*$ is of type $\mathbf{H}$ with respect to $\overline{\mathcal{H}}$, then $B = h(A)$ where the function h can be modified on a null-set with respect to ν_A so that*

$$h(\lambda)/h(\mu) \le \mathbf{H}\,(\lambda/\mu)\,, \quad \lambda, \mu \in \sigma(A) \setminus \{0\}. \tag{3.3}$$

Proof Fix a (large) compact subset $K \subset \sigma(A) \cap \mathbf{R}_+$ and put $\mathcal{H}_0' = \mathcal{H}_1' = E_K(\mathcal{H}_0)$ where E is the spectral measure of A, and the norms are defined by restriction,

$$\| x \|_{\mathcal{H}_i'} = \| x \|_{\mathcal{H}_i} \quad , \quad \| x \|_{\mathcal{H}_*'} = \| x \|_{\mathcal{H}_*} \quad , \quad x \in E_K\,(\mathcal{H}_0)\,.$$

It is clear that the operator A' corresponding to $\overline{\mathcal{H}'}$ is the compression of A to $\mathcal{H}_0'$ and likewise the operator B' corresponding to $\mathcal{H}_*'$ is the compression of B to $\mathcal{H}_0'$. Moreover, $\mathcal{H}_*'$ is of interpolation type $\mathbf{H}$ with respect to $\overline{\mathcal{H}'}$ and the operator $B' = (h|_K)\,(A')$. For this reason, and since the compact set K is arbitrary, it clearly suffices to prove the statement with $\overline{\mathcal{H}}$ replaced by $\overline{\mathcal{H}'}$. Then A is bounded above and below. Moreover, by (3.2), also B is bounded above and below.

Let $c < 1$ be a positive number such that $\sigma(A) \subset \left(c, c^{-1}\right)$. For a fixed $\varepsilon > 0$ with $\varepsilon < c/2$ we set

$$E_\lambda = \sigma(A) \cap (\lambda - \varepsilon, \lambda + \varepsilon)$$

and consider the functions

$$m_\varepsilon(\lambda) = \operatorname{ess\,inf}_{E_\lambda} h, \qquad M_\varepsilon(\lambda) = \operatorname{ess\,sup}_{E_\lambda} h,$$

the essential inf and sup being taken with respect to ν_A.

Now fix a small positive number ε' and two unit vectors e_λ, e_μ supported by E_λ, E_μ, respectively, such that

$$\| e_\lambda \|_*^2 \geq M_\varepsilon(\lambda) - \varepsilon', \qquad \left\| e_\mu \right\|_*^2 \leq m_\varepsilon(\mu) + \varepsilon'.$$

Now fix $\lambda, \mu \in \sigma(A)$ and let $Tx = \left\langle x, e_\mu \right\rangle_0 e_\lambda$. Then

$$\begin{aligned} \| Tx \|_1^2 = \left| \left\langle x, e_\mu \right\rangle_0 \right|^2 \| e_\lambda \|_1^2 &\leq \frac{1}{(\mu - \varepsilon)^2} \left| \left\langle x, e_\mu \right\rangle_1 \right|^2 (\lambda + \varepsilon) \\ &\leq \frac{(\mu + \varepsilon)(\lambda + \varepsilon)}{(\mu - \varepsilon)^2} \| x \|_1^2 . \end{aligned}$$

Likewise,

$$\| Tx \|_0^2 \leq \left| \left\langle x, e_\mu \right\rangle_0 \right|^2 \leq \| x \|_0^2 ,$$

so $\| T \| \leq 1$ and $\| T \|_A^2 \leq \alpha_{\mu,\lambda,\varepsilon}$ where $\alpha_{\mu,\lambda,\varepsilon} = \frac{(\mu+\varepsilon)(\lambda+\varepsilon)}{(\mu-\varepsilon)^2}$.

Since $\mathcal{H}_*$ is of type **H**, we conclude that

$$\| T \|_B^2 \leq \mathbf{H}\left(\alpha_{\mu,\lambda,\varepsilon}\right),$$

whence

$$\begin{aligned} M_\varepsilon(\lambda) - \varepsilon' \leq \| e_\lambda \|_*^2 = \left\| Te_\mu \right\|_*^2 &\leq \mathbf{H}\left(\alpha_{\mu,\lambda,\varepsilon}\right) \left\| e_\mu \right\|_*^2 \\ &\leq \mathbf{H}\left(\alpha_{\mu,\lambda,\varepsilon}\right) \left(m_\varepsilon(\mu) + \varepsilon'\right). \end{aligned} \tag{3.4}$$

In particular, since ε' was arbitrary, and $m_\varepsilon(\lambda) \leq \| e_\lambda \|_*^2 \leq \| B \|$, we find that

$$M_\varepsilon(\lambda) - m_\varepsilon(\lambda) \leq \left[\mathbf{H}\left(\alpha_{\mu,\lambda,\varepsilon}\right) - 1\right] \| B \| .$$

By assumption, $\mathbf{H}$ is continuous and $\mathbf{H}(1) = 1$. Hence, as $\varepsilon \downarrow 0$, the functions $M_\varepsilon(\lambda)$ diminish monotonically, converging uniformly to a function $h_*(\lambda)$ which is also the uniform limit of the family $m_\varepsilon(\lambda)$. It is clear that h_* is continuous, and since $m_\varepsilon \le h_* \le M_\varepsilon$, we have $h_* = h$ almost everywhere with respect to ν_A. The relation (3.3) now follows for $h = h_*$ by letting ε and ε' tend to zero in (3.4). □

A partial converse to Theorem II is found below, see Theorem 6.3.

Remark 3.1 (The Non-Separable Case) Now consider the case when $\mathcal{H}_0$ is non-separable. (By regularity this means that also $\mathcal{H}_1$ and $\mathcal{H}_*$ are non-separable.)

First assume that the operator A is bounded. Let $\mathcal{H}_0'$ be a separable reducing subspace for A such that the restriction A' of A to $\mathcal{H}_0'$ has the same spectrum as A. The space $\mathcal{H}_0'$ reduces B by Donoghue's lemma; by Theorem II the restriction B' of B to $\mathcal{H}_0'$ satisfies $B' = h'(A')$ for some continuous function h' satisfying (3.3) on $\sigma(A)$. Let $\mathcal{H}_0''$ be any other separable reducing subspace, where (as before) $B'' = h''(A'')$. Then $\mathcal{H}_0' \oplus \mathcal{H}_0''$ is a separable reducing subspace on which $B = h(A)$ for some third continuous function h on $\sigma(A)$. Then $h(A') \oplus h(A'') = h'(A') \oplus h''(A'')$ and by continuity we must have $h = h' = h''$ on $\sigma(A)$. The function h thus satisfies $B = h(A)$ as well as the estimate (3.3).

If A is unbounded, we replace A by its compression to $P_n\mathcal{H}_0$ where P_n is the spectral projection of A corresponding to the spectral set $[0, n] \cap \sigma(A)$, $n = 1, 2, \ldots$. The same reasoning as above shows that B appears as a continuous function of A on $\sigma(A) \cap [0, n]$. Since n is arbitrary, we find that $B = h(A)$ for a function h satisfying (3.3).

3.2 Geometric Interpolation

Now consider the particular case when $\mathcal{H}_*$ is of exponent θ, viz. of type $\mathbf{H}(t) = t^\theta$ with respect to $\overline{\mathcal{H}}$. We write $B = h(A)$ where h is the continuous function provided by Theorem II (and Remark 3.1 in the non-separable case).

Fix a point $\lambda_0 \in \sigma(A)$ and let $C = h(\lambda_0)\lambda_0^{-\theta}$. The estimate (3.3) then implies that $h(\lambda) \le C\lambda^\theta$ and $h(\mu) \ge C\mu^\theta$ for all $\lambda, \mu \in \sigma(A)$. We have proved the following theorem.

Theorem 3.2 ([27, 40]) *If $\mathcal{H}_*$ is an exact interpolation Hilbert space of exponent θ relative to $\overline{\mathcal{H}}$, then $B = h(A)$ where $h(\lambda) = C\lambda^\theta$ for some positive constant C.*

Theorem 3.2 says that $\mathcal{H}_* = \mathcal{H}_\theta$ up to a constant multiple of the norm, where $\mathcal{H}_\theta$ is the space defined in (1.7). In the guise of operator inequalities: for any fixed positive operators A and B, the condition

$$T^*T \le M_0 \quad , \quad T^*AT \le M_1 A \quad \Rightarrow \quad T^*BT \le M_0^{1-\theta} M_1^{\theta} B$$

is equivalent to that $B = A^\theta$.

It was observed in [27] that $\mathcal{H}_\theta$ also equals to the complex interpolation space $C_\theta(\overline{\mathcal{H}})$. For the sake of completeness, we supply a short proof of this fact in the appendix.

Remark 3.3 An exact quadratic interpolation method, the *geometric mean* was introduced earlier by Pusz and Woronowicz [33] (it corresponds to the $C_{1/2}$-method). In [40], Uhlmann generalized that method to a method (the *quadratic mean*) denoted QI_t where $0 < t < 1$; this method is quadratic and of exponent t.

In view of Theorem 3.2 and the preceding remarks we can conclude that $\mathrm{QI}_\theta(\overline{\mathcal{H}}) = C_\theta(\overline{\mathcal{H}}) = \mathcal{H}_\theta$ for any regular Hilbert couple $\overline{\mathcal{H}}$. We refer to [40] for several physically relevant applications of this type of interpolation.

Finally, we want to mention that in [32] Peetre introduces the "Riesz method of interpolation"; in Section 5 he also defines a related method "QM" which comes close to the complex $C_{1/2}$-method.

3.3 Donoghue's Theorem

The exact quadratic interpolation spaces relative to a Hilbert couple were characterized by Donoghue in the paper [14]. We shall here prove the following equivalent version of Donoghue's result (see [2, 3]).

Theorem III *An intermediate Hilbert space $\mathcal{H}_*$ relative to $\overline{\mathcal{H}}$ is an exact interpolation space if and only if there is a positive radon measure ϱ on* $[0,\infty]$ *such that*

$$\| x \|_*^2 = \int_{[0,\infty]} \left(1 + t^{-1}\right) K\left(t, x; \overline{\mathcal{H}}\right) d\varrho(t).$$

Equivalently, $\mathcal{H}_*$ is exact interpolation relative to $\overline{\mathcal{H}}$ if and only if the corresponding operator B can be represented as $B = h(A)$ for some function $h \in P'$.

The statements that all norms of the given form are exact quadratic interpolation norms have already been shown (see Sect. 1.2). There remains to prove that there are no others.

Donoghue's original formulation of the result, as well as other equivalent forms of the theorem, is found in Sect. 6 below. Our present approach follows [2] and is based on K-monotonicity.

Remark 3.4 The condition that $\mathcal{H}_*$ be exact interpolation with respect to $\overline{\mathcal{H}}$ means that $\mathcal{H}_*$ is of type $\mathbf{H}$ where $\mathbf{H}(t) = \max\{1, t\}$. In view of Theorem II (and Remark 3.1), this means that we can represent $B = h(A)$ where h is *quasi-concave* on $\sigma(A) \setminus \{0\}$,

$$h(\lambda) \le h(\mu) \max\{1, \lambda/\mu\}, \qquad \lambda, \mu \in \sigma(A) \setminus \{0\}. \tag{3.5}$$

In particular, h is locally Lipschitzian on $\sigma(A) \cap \mathbf{R}_+$.

Remark 3.5 A related result concerning *non-exact* quadratic interpolation was proved by Ovchinnikov [30] using Donoghue's theorem. Cf. also [4].

3.4 The Proof for Simple Finite-Dimensional Couples

Similar to our approach to Calderón's problem, our strategy is to reduce Theorem III to a case of "simple couples."

Theorem 3.6 *Assume that $\mathcal{H}_0 = \mathcal{H}_0 = \mathbf{C}^n$ as sets and that all eigenvalues $(\lambda_i)_1^n$ of the corresponding operator A are of unit multiplicity. Consider a third Hermitian norm $\|x\|_*^2 = \langle Bx, x\rangle_0$ on $\mathbf{C}^n$. Then $\mathcal{H}_*$ is exact interpolation with respect to $\overline{\mathcal{H}}$ if and only if $B = h(A)$ where $h \in P'$.*

Remark 3.7 The lemma says that the class of functions h on $\sigma(A)$ satisfying

$$T^*T \leq 1 \quad , \quad T^*AT \leq A \quad \Rightarrow \quad T^*h(A)T \leq h(A), \qquad (T \in M_n(\mathbf{C})) \tag{3.6}$$

is precisely the set $P'|\sigma(A)$ of restrictions of P'-functions to $\sigma(A)$. In this way, the condition (3.6) provides an operator-theoretic solution to the interpolation problem by positive Pick functions on a finite subset of $\mathbf{R}_+$.

Proof of Theorem 3.6 We already know that the spaces $\mathcal{H}_*$ of the asserted form are exact interpolation relative to $\overline{\mathcal{H}}$ (see Sects. 1.2 and 1.3).

Now let $\mathcal{H}_*$ be any exact quadratic interpolation space. By Donoghue's lemma and the argument in Sect. 2.3, we can for an appropriate positive sequence $\lambda = (\lambda_i)_1^n$ identify $\overline{\mathcal{H}} = \overline{\ell_2^n}(\lambda)$, $A = \text{diag}(\lambda_i)$, and $B = h(A)$ where h is some positive function defined on $\sigma(A) = \{\lambda_i\}_1^n$.

Our assumption is that $\ell_2^n\,(h(\lambda))$ is exact interpolation relative to $\overline{\ell_2^n}(\lambda)$. We must prove that $h \in P'|\sigma(A)$. To this end, write

$$k_{\lambda_i}(t) = \frac{(1+t)\lambda_i}{1+t\lambda_i},$$

and recall that (see Lemma 1.1)

$$K\left(t, x; \overline{\ell_2^n}(\lambda)\right) = \left(1+t^{-1}\right)^{-1} \sum_1^n |x_i|^2\, k_{\lambda_i}(t).$$

Let us denote by C the algebra of continuous complex functions on $[0, \infty]$ with the supremum norm $\|\, u \,\|_\infty = \sup_{t>0} |\, u(t)\,|$. Let $V \subset C$ be the linear span of the k_{λ_i}

for $i = 1, \ldots, n$. We define a positive functional ϕ on V by

$$\phi(\sum_1^n a_i k_{\lambda_i}) = \sum_1^n a_i \, h(\lambda_i).$$

We claim that ϕ is a *positive functional*, i.e., if $u \in V$ and $u(t) \geq 0$ for all $t > 0$, then $\phi(u) \geq 0$.

To prove this let $u = \sum_1^n a_i k_{\lambda_i}$ be non-negative on $\mathbf{R}_+$ and write $a_i = |x_i|^2 - |y_i|^2$ for some $x, y \in \mathbf{C}^n$. The condition that $u \geq 0$ means that

$$\begin{aligned}\left(1 + t^{-1}\right) K\left(t, x; \overline{\ell_2^n}(\lambda)\right) &= \sum_{i=1}^n |x_i|^2 \, k_{\lambda_i}(t) \\ &\geq \sum_{i=1}^n |y_i|^2 \, k_{\lambda_i}(t) \\ &= \left(1 + t^{-1}\right) K\left(t, y; \overline{\ell_2^n}(\lambda)\right), \qquad t > 0.\end{aligned} \tag{3.7}$$

Since $\overline{\ell_2^n}(\lambda)$ is an exact Calderón couple (by Theorem 2.4), the space $\ell_2^n(h(\lambda))$ is exact K-monotonic. In other words, (3.7) implies that

$$\| \, x \, \|_{\ell_2^n(h(\lambda))} \geq \| \, y \, \|_{\ell_2^n(h(\lambda))} \, ,$$

i.e.,

$$\phi(u) = \sum_1^n \left(|x_i|^2 - |y_i|^2\right) \, h(\lambda_i) \geq 0.$$

The asserted positivity of ϕ is thereby proved.

Replacing λ_i by $c\lambda_i$ for a suitable positive constant c we can without losing generality assume that $1 \in \sigma(A)$, i.e., that the unit $\mathbf{1}(x) \equiv 1$ of the C^*-algebra C belongs to V. The positivity of ϕ then ensures that

$$\| \, \phi \, \| = \sup_{u \in V; \; \|u\|_\infty \leq 1} |\phi(u)| = \phi(\mathbf{1}).$$

Let Φ be a Hahn–Banach extension of ϕ to C and note that

$$\| \, \Phi \, \| = \| \, \phi \, \| = \phi(\mathbf{1}) = \Phi(\mathbf{1}).$$

This means that Φ is a positive functional on C (cf. [29], §3.3). By the Riesz representation theorem there is thus a positive Radon measure ϱ on $[0, \infty]$ such that

$$\Phi(u) = \int_{[0,\infty]} u(t)\, d\varrho(t), \qquad u \in C.$$

In particular

$$h(\lambda_i) = \phi\left(k_{\lambda_i}\right) = \Phi\left(k_{\lambda_i}\right) = \int_{[0,\infty]} \frac{(1+t)\lambda_i}{1+t\lambda_i}\, d\varrho(t), \quad i = 1, \ldots, n.$$

We have shown that h is the restriction to $\sigma(A)$ of a function of class P'. □

3.5 *The Proof of Donoghue's Theorem*

We here prove Theorem III in full generality.

We remind the reader that if $S \subset \mathbf{R}_+$ is a subset, we write $P'|S$ for the convex cone of restrictions of P'-functions to S. We first collect some simple facts about this cone.

Lemma 3.8

(i) *The class $P'|S$ is closed under pointwise convergence.*
(ii) *If S is finite and if $\lambda = (\lambda_i)_{i=1}^n$ is an enumeration of the points of S, then h belongs to $P'|S$ if and only if $\ell_2^n(h(\lambda))$ is exact interpolation with respect to the pair $\overline{\ell_2^n}(\lambda)$.*
(iii) *If S is infinite, then a continuous function h on S belongs to $P'|S$ if and only if $h \in P'|\Lambda$ for every finite subset $\Lambda \subset S$.*

Proof

(i) Let h_n be a sequence in P' converging pointwise on S and fix $\lambda \in S$. It is clear that the boundedness of the numbers $h_n(\lambda)$ is equivalent to boundedness of the total masses of the corresponding measures ϱ_n on the compact set $[0, \infty]$. It now suffices to apply Helly's selection theorem.
(ii) This is Theorem 3.6.
(iii) Let Λ_n be an increasing sequence of finite subsets of S whose union is dense. Let $h_n = h|\Lambda_n$ where h is continuous on S. If $h_n \in P'|\Lambda_n$ for all n, then the sequence h_n converges pointwise on $\cup\Lambda_n$ to h. By part (i) we then have $h \in P'|\sigma(A)$.

□

We can now finish the proof of Donoghue's theorem (Theorem III).

Let $\mathcal{H}_*$ be exact interpolation with respect to $\overline{\mathcal{H}}$ and represent the corresponding operator as $B = h(A)$ where h satisfies (3.3). By the remarks after Theorem III, the function h is locally Lipschitzian.

In view of Lemma 3.8 we shall be done when we have proved that $\ell_2^n(h(\lambda))$ is exact interpolation with respect to $\overline{\ell_2^n}(\lambda)$ for all sequences $\lambda = (\lambda_i)_1^n \subset \sigma(A)$ of distinct points. Let us arrange the sequences in the increasing order: $0 < \lambda_1 < \cdots < \lambda_n$.

Fix $\varepsilon > 0$, $\varepsilon < \min\{c, \lambda_1, 1/\lambda_n\}$ and let $E_i = [\lambda_i - \varepsilon, \lambda_i + \varepsilon] \cap \sigma(A)$; we assume that ε is sufficiently small that the E_i be disjoint. Let $M = \cup_1^n E_i$. We can assume that h has Lipschitz constant at most 1 on M.

Let $\mathcal{M}$ be the reducing subspace of $\mathcal{H}_0$ corresponding to the spectral set M, and let $\tilde{A}$ be the compression of A to $\mathcal{M}$. We define a function g on M by $g(\lambda) = \lambda_i$ on E_i. Then $|g(\lambda) - \lambda| < \varepsilon$ on $\sigma(\tilde{A})$, so

$$\| \tilde{A} - g(\tilde{A}) \| \leq \varepsilon \quad , \quad \| h(\tilde{A}) - h(g(\tilde{A})) \| \leq \varepsilon. \tag{3.8}$$

Lemma 3.9 *Suppose that* $A', A'' \in \mathcal{L}(\mathcal{M})$ *satisfy* $A', A'' \geq \delta > 0$ *and* $\left\| A' - A'' \right\| \leq \varepsilon$. *Then* $\| T \|_{A''} \leq \sqrt{1 + 2\varepsilon/\delta} \max\{\| T \|, \| T \|_{A'}\}$ *for all* $T \in \mathcal{L}(\mathcal{M})$.

Proof By definition, $\| T \|_{A'}$ is the smallest number $C \geq 0$ such that $T^* A' T \leq C^2 A'$. Thus

$$\begin{aligned} T^* A'' T &= T^*(A'' - A')T + T^* A' T \\ &\leq \| T \|^2 \varepsilon + \| T \|_{A'}^2 \left(A'' + \left(A' - A''\right)\right) \\ &\leq 2\varepsilon \max\{\| T \|^2, \| T \|_{A'}^2\} + \| T \|_{A'}^2 A'' \\ &\leq \max\{\| T \|^2, \| T \|_{A'}^2\} (1 + 2\varepsilon/\delta) A''. \end{aligned}$$

□

We can find $\delta > 0$ such that the operators $\tilde{A}$, $g(\tilde{A})$, $h(\tilde{A})$, and $h(g(\tilde{A}))$ are $\geq \delta$. Then by repeated use of Lemma 3.9,

$$\begin{aligned} \| T \|_{h\left(g\left(\tilde{A}\right)\right)} &\leq \sqrt{1 + 2\varepsilon/\delta} \max\{\| T \|, \| T \|_{h\left(\tilde{A}\right)}\} \\ &\leq \sqrt{1 + 2\varepsilon/\delta} \max\left\{\| T \|, \| T \|_{\tilde{A}}\right\} \\ &\leq (1 + 2\varepsilon/\delta) \max\{\| T \|, \| T \|_{g\left(\tilde{A}\right)}\}, \quad T \in \mathcal{L}(\mathcal{M}). \end{aligned}$$

Let e_i be a unit vector supported by the spectral set E_i and define a space $\mathcal{V} \subset \mathcal{M}$ to be the n-dimensional space spanned by the e_i. Let A_0 be the compression of $g(\tilde{A})$ to $\mathcal{V}$, then

$$\| T \|_{h(A_0)} \leq (1 + 2\varepsilon/\delta) \max\left\{\| T \|, \| T \|_{A_0}\right\}, \quad T \in \mathcal{L}(\mathcal{V}). \tag{3.9}$$

Identifying $\mathcal{V}$ with ℓ_2^n and A_0 with the matrix $\operatorname{diag}(\lambda_i)$, we see that (3.9) is independent of ε. Letting ε diminish to 0 in (3.9) now gives that $\ell_2^n(h(\lambda))$ is exact interpolation with respect to $\overline{\ell_2^n}(\lambda)$. In view of Lemma 3.8, this finishes the proof of Theorem III. □

4 Classes of Matrix Functions

In this section, we discuss the basic properties of interpolation functions: in particular, the relation to the well-known classes of monotone matrix functions. We refer to the books [12] and [34] for further reading on the latter classes.

4.1 Interpolation and Matrix Monotone Functions

Let A_1 and A_2 be positive operators in ℓ_2^n ($n = \infty$ is admitted). Suppose that $A_1 \leq A_2$ and form the following operators on $\ell_2^n \oplus \ell_2^n$:

$$T_0 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad , \quad A = \begin{pmatrix} A_2 & 0 \\ 0 & A_1 \end{pmatrix}.$$

It is then easy to see that ${T_0}^* T_0 \leq 1$ and that ${T_0}^* A T_0 = \begin{pmatrix} A_1 & 0 \\ 0 & 0 \end{pmatrix} \leq A$.

Now assume that a function h on $\sigma(A)$ belongs to the class C_A defined in Sect. 1.4, i.e., that h satisfies

$$T^*T \leq 1 \quad , \quad T^*AT \leq A \quad \Rightarrow \quad T^*h(A)T \leq h(A), \tag{4.1}$$

where T denotes an operator on ℓ_2^{2n}.

We then have ${T_0}^* h(A) T_0 \leq h(A)$, or

$$\begin{pmatrix} h(A_1) & 0 \\ 0 & 0 \end{pmatrix} \leq \begin{pmatrix} h(A_2) & 0 \\ 0 & h(A_1) \end{pmatrix}.$$

In particular, we find that $h(A_1) \leq h(A_2)$. We have shown that (under the assumptions above)

$$A_1 \leq A_2 \qquad \Rightarrow \qquad h(A_1) \leq h(A_2). \tag{4.2}$$

We now change our point of view slightly. Given a positive integer n, we let C_n denote the convex of positive functions h on $\mathbf{R}_+$ such that (4.1) holds for *all* positive operators A on ℓ_2^n and all $T \in \mathcal{L}(\ell_2^n)$.

Similarly, we let P'_n denote the class of all positive functions h on $\mathbf{R}_+$ such that $h(A_1) \leq h(A_2)$ whenever A_1, A_2 are positive operators on ℓ_2^n such that $A_1 \leq A_2$. We refer to P'_n as the cone of positive functions *monotone of order* n on $\mathbf{R}_+$.

We have shown above that $C_{2n} \subset P'_n$.

In the other direction, assume that $h \in P'_{2n}$. Let A, T be bounded operators on ℓ_2^n with $A > 0$, $T^*T \leq 1$ and $T^*AT \leq A$. Assume also that h be continuous. We will use the following lemma due to Hansen [19]. We recall the proof for completeness.

Lemma 4.1 ([19]) $T^*h(A)T \leq h(T^*AT)$.

Proof Put $S = (1 - TT^*)^{1/2}$ and $R = (1 - T^*T)^{1/2}$ and consider the $2n \times 2n$ matrix

$$U = \begin{pmatrix} T & S \\ R & -T^* \end{pmatrix} \quad , \quad X = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}.$$

It is well known, and easy to check, that U is unitary and that

$$U^*XU = \begin{pmatrix} T^*AT & T^*AS \\ SAT & SAS \end{pmatrix}.$$

Next fix a number $\varepsilon > 0$, a constant $\lambda > 0$ (to be fixed), and form the matrix

$$Y = \begin{pmatrix} T^*AT + \varepsilon & 0 \\ 0 & 2\lambda \end{pmatrix}$$

which, provided that we choose $\lambda \geq \|SAS\|$, satisfies

$$Y - U^*XU = \begin{pmatrix} \varepsilon & -T^*AS \\ -SAT & 2\lambda - SAS \end{pmatrix} \geq \begin{pmatrix} \varepsilon & D \\ D^* & \lambda \end{pmatrix},$$

where we have written $D = -T^*AS$.

If we now also choose λ so that $\lambda \geq \|D\|^2/\varepsilon$, then we obtain for all $\xi, \eta \in \mathbf{C}^n$ that

$$\left\langle \begin{pmatrix} \varepsilon & D \\ D^* & \lambda \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix}, \begin{pmatrix} \xi \\ \eta \end{pmatrix} \right\rangle = \varepsilon\|\xi\|^2 + \langle D\eta, \xi\rangle + \langle D^*\xi, \eta\rangle + \lambda\|\eta\|^2$$

$$\geq \varepsilon\|\xi\|^2 - 2\|D\|\|\xi\|\|\eta\| + \lambda\|\eta\|^2 \geq 0.$$

Hence $U^*XU \leq Y$ and as a consequence $U^*h(X)U = h(U^*XU) \leq h(Y)$, since h is matrix monotone of order $2n$. The last inequality means that

$$\begin{pmatrix} T^*h(A)T & T^*h(A)S \\ Sh(A)T & Sh(A)S \end{pmatrix} \leq \begin{pmatrix} h(T^*AT + \varepsilon) & 0 \\ 0 & h(2\lambda) \end{pmatrix},$$

so in particular $T^*h(A)T \leq h(T^*AT + \varepsilon)$. Since $\varepsilon > 0$ was arbitrary, and since h is assumed to be continuous, we conclude the lemma. □

We now continue our discussion. Assuming that $T^*T \le 1$ and $T^*AT \le A$, and that $h \in P'_{2n}$ is continuous, we have $h(T^*AT) \le h(A)$ [since $h \in P'_n$], so $T^*h(A)T \le h(A)$ by Lemma 4.1. We conclude that $h \in C_n$.

To prove that $P'_{2n} \subset C_n$, we need to remove the continuity assumption on h made above. This is completely standard: let φ be a smooth positive function on $\mathbf{R}_+$ such that $\int_0^\infty \varphi(t)\,dt/t = 1$, and define a sequence h_k by $h_k(\lambda) = k^{-1}\int_0^\infty \varphi\left(\lambda^k/t^k\right) h(t)\,dt/t$. The class P'_{2n} is a convex cone, closed under pointwise convergence [12], so the functions $h_1, h_2, \ldots$ are of class P'_{2n}. They are furthermore continuous, so by the argument above, they are of class C_n. By Lemma 3.8, the cone C_n is also closed under pointwise convergence, so $h = \lim h_n \in C_n$.

To summarize, we have the inclusions $C_{2n} \subset P'_n$, $P'_{2n} \subset C_n$, and also $C_{n+1} \subset C_n$, $P'_{n+1} \subset P'_n$. In view of Theorem III, we have the identity $\cap_1^\infty C_n = P'$. The inclusions above now imply the following result, sometimes known as "Löwner's theorem on matrix monotone functions."

Theorem 4.2 *We have* $\cap_1^\infty P'_n = \cap_1^\infty C_n = P'$.

The identity $\cap_1^\infty P'_n = P'$ says that a positive function h is monotone of all finite orders if and only if it is of class P'. The somewhat less precise fact that $P'_\infty = P'$ is interpreted as that the class of *operator monotone functions* coincides with P'.

The identity $C_\infty = P'$ is, except for notation, contained in the work of Foiaş and Lions, from [16]. See Sect. 6.4.

Note that the inclusion $P'_{2n} \subset C_n$ shows that a matrix monotone functions of order $2n$ can be interpolated by a positive Pick function at n points. Results of a similar nature, where it is shown, in addition, that an interpolating Pick function can be taken rational of a certain degree, are discussed, for example, in Donoghue's book [14, Chapter XIII] or (more relevant in the present connection) in the paper [13].

It seems somewhat inaccurate to refer to the identity $\cap_1^\infty P'_n = P'$ as "Löwner's theorem," since Löwner discusses more subtle results concerning matrix monotone functions *of a given finite order* n. In spite of this, it is common nowadays to let "Löwner's theorem" refer to this identity.

4.2 More on the Cone C_A

We can now give an short proof of the following result due to Donoghue [13].

Theorem 4.3 *For a positive function h on* $\sigma(A)$ *we define two positive functions* $\tilde{h}$ *and* h^* *on* $\sigma\left(A^{-1}\right)$ *by* $\tilde{h}(\lambda) = \lambda h(1/\lambda)$ *and* $h^*(\lambda) = 1/h(1/\lambda)$. *Then the following conditions are equivalent:*

(i) $h \in C_A$,
(ii) $\tilde{h} \in C_{A^{-1}}$,
(iii) $h^* \in C_{A^{-1}}$.

Proof Let $\mathcal{H}_*$ be a quadratic intermediate space relative to a regular Hilbert couple $\overline{\mathcal{H}}$; let $B = h(A)$ be the corresponding operator. It is clear that $\mathcal{H}_*$ is exact interpolation relative to $\overline{\mathcal{H}}$ if and only if $\mathcal{H}_*$ is exact interpolation relative to the *reverse couple* $\overline{\mathcal{H}^{(r)}} = (\mathcal{H}_1, \mathcal{H}_0)$. The latter couple has corresponding operator A^{-1} and it is clear that the identity $\| x \|_*^2 = \langle h(A)x, x\rangle_0$ is equivalent to that $\| x \|_*^2 = \left\langle A^{-1}\tilde{h}\left(A^{-1}\right)x, x\right\rangle_1$. We have shown the equivalence of (i) and (ii).

Next let $\overline{\mathcal{H}^*} = \left(\mathcal{H}_0{}^*, \mathcal{H}_1{}^*\right)$ be the *dual couple*, where we identify $\mathcal{H}_0{}^* = \mathcal{H}_0$. With this identification, $\mathcal{H}_1{}^*$ becomes associated with the norm $\| x \|^2_{\mathcal{H}_1^*} = \left\langle A^{-1}x, x\right\rangle_0$, and $\mathcal{H}_*{}^*$ is associated with $\| x \|^2_{\mathcal{H}_*^*} = \left\langle B^{-1}x, x\right\rangle_0$. It remains to note that $\mathcal{H}_*$ is exact interpolation relative to $\overline{\mathcal{H}}$ if and only if $\mathcal{H}_*{}^*$ is exact interpolation relative to $\overline{\mathcal{H}^*}$, proving the equivalence of (i) and (iii). □

Combining with Theorem III, one obtains alternative proofs of the interpolation theorems for P'-functions discussed by Donoghue in the paper [13].

Remark 4.4 The exact quadratic interpolation spaces which are fixed by the duality, i.e., which satisfy $\mathcal{H}_*{}^* = \mathcal{H}_*$, correspond precisely to the class of P'-functions which are *self-dual*: $h^* = h$. This class was characterized by Hansen in the paper [20].

4.3 Matrix Concavity

A function h on $\mathbf{R}_+$ is called *matrix concave of order n* if we have Jensen's inequality

$$\lambda h\left(A_1\right) + (1-\lambda)h\left(A_2\right) \leq h\left(\lambda A_1 + (1-\lambda)A_2\right)$$

for all positive $n \times n$ matrices A_1, A_2, and all numbers $\lambda \in [0, 1]$. Let us denote by Γ_n the convex cone of positive concave functions of order n on $\mathbf{R}_+$. The fact that $\cap_n \Gamma_n = P'$ follows from the theorem of Kraus [23]. Following [2] we now give an alternative proof of this fact.

Proposition 4.5 *For all n we have the inclusion $C_{3n} \subset \Gamma_n \subset P'_n$. In particular $\cap_1^\infty \Gamma_n = P'$.*

Proof Assume first that $h \in C_{3n}$ and pick two positive matrices A_1 and A_2. Define $A_3 = (1-\lambda)A_1 + \lambda A_2$ where $\lambda \in [0, 1]$ is given, and define matrices A and T of order $3n$ by

$$A = \begin{pmatrix} A_3 & 0 & 0 \\ 0 & A_1 & 0 \\ 0 & 0 & A_2 \end{pmatrix} \quad , \quad T = \begin{pmatrix} 0 & 0 & 0 \\ \sqrt{1-\lambda} & 0 & 0 \\ \sqrt{\lambda} & 0 & 0 \end{pmatrix}.$$

It is clear that $T^*T \leq 1$ and

$$T^*AT = \begin{pmatrix} A_3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \leq A,$$

so, since $h \in C_{3n}$, we have $T^*h(A)T \leq h(A)$, or

$$\begin{pmatrix} (1-\lambda)h(A_1) + \lambda h(A_2) & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \leq \begin{pmatrix} h(A_3) & 0 & 0 \\ 0 & h(A_1) & 0 \\ 0 & 0 & h(A_2) \end{pmatrix}.$$

Comparing the matrices in the upper left corners, we find that $h \in \Gamma_n$.

Assume now that $h \in \Gamma_n$, and take positive definite matrices A_1, A_2 of order n with $A_1 \leq A_2$. Also pick $\lambda \in (0,1)$. Then $\lambda A_2 = \lambda A_1 + (1-\lambda)A_3$ where $A_3 = \lambda(1-\lambda)^{-1}(A_2 - A_1)$. By matrix concavity, we then have

$$h(\lambda A_2) \geq \lambda h(A_1) + (1-\lambda)h(A_3) \geq \lambda h(A_1),$$

where we used non-negativity to deduce the last inequality. Being concave, h is certainly continuous. Letting $\lambda \uparrow 1$ one thus finds that $h(A_1) \leq h(A_2)$. We have shown that $h \in P'_n$. □

For a further discussion of classes of convex matrix functions and their relations to monotonicity, we refer to the paper [21].

4.4 Interpolation Functions of Two Variables

In this section, we briefly discuss a class of interpolation functions of two matrix variables. We shall not completely characterize the class of such generalized interpolation functions here, but we hope that the following discussion will be of some use for a future investigation.

Let H_1 and H_2 be Hilbert spaces. We turn $H_1 \otimes H_2$ into a Hilbert space by defining the inner product on elementary tensors via $\langle x_1 \otimes x_2, x'_1 \otimes x'_2 \rangle := \langle x_1, x_1{}' \rangle_1 \cdot \langle x_2, x_2{}' \rangle_2$ (then extend via sesqui-linearity). Similarly, if T_i are operators on H_i, the tensor product $T_1 \otimes T_2$ is defined on elementary tensors via $(T_1 \otimes T_2)(x_1 \otimes x_2) = T_1x_1 \otimes T_2x_2$. It is then easy to see that if A_i are positive operators on H_i for $i = 1, 2$, then $A_1 \otimes A_2 \geq 0$ as an operator on the tensor product. Furthermore, we have $A_1 \otimes A_2 \leq A'_1 \otimes A'_2$ if $A_i \leq A'_i$ for $i = 1, 2$.

Given two positive definite matrices A_i of orders n_i and a function h on $\sigma(A_1) \times \sigma(A_2)$, we define a matrix $h(A_1, A_2)$ by

$$h(A_1, A_2) = \sum_{(\lambda_1,\lambda_2)\in\sigma(A_1)\times\sigma(A_2)} h(\lambda_1, \lambda_2)\, E^1_{\lambda_1} \otimes E^2_{\lambda_2}$$

where E^j is the spectral resolution of the matrix A_j.

We shall say that h *gives rise to exact interpolation* relative to (A_1, A_2), and write $h \in C_{A_1,A_2}$, if the condition

$$T_j{}^*T_j \le 1 \quad , \quad T_j{}^*A_jT_j \le A_j, \quad j = 1, 2 \tag{4.3}$$

implies

$$\begin{aligned} &h(A_1, A_2) + (T_1 \otimes T_2)^* h(A_1, A_2)(T_1 \otimes T_2) \\ &\quad - (T_1 \otimes 1)^* h(A_1, A_2)(T_1 \otimes 1) - (1 \otimes T_2)^* h(A_1, A_2)(1 \otimes T_2) \ge 0. \end{aligned} \tag{4.4}$$

Taking $T_1 = T_2 = 0$ we see that $h \ge 0$ for all $h \in C_{A_1,A_2}$. It is also clear that C_{A_1,A_2} is a convex cone closed under pointwise convergence on the finite set $\sigma(A_1) \times \sigma(A_2)$.

If $h = h_1 \otimes h_2$ is an elementary tensor where $h_j \in C_{A_j}$ is a function of one variable, then (4.3) implies $T_j{}^*h_j(A_j)T_j \le h_j(A_j)$, whence $(h_1(A_1) - T_1{}^*h_1(A_1)T_1) \otimes (h_2(A_2) - T_2{}^*h_2(A_2)T_2) \ge 0$, which implies (4.4). We have shown that $C_{A_1} \otimes C_{A_2} \subset C_{A_1,A_2}$.

Since for each $t \ge 0$ the P'-function $\lambda \mapsto \frac{(1+t)\lambda}{1+t\lambda}$ is of class C_{A_j}, we infer that every function representable in the form

$$h(\lambda_1, \lambda_2) = \iint_{[0,\infty]^2} \frac{(1 + t_1)\lambda_1}{1 + t_1\lambda_1} \frac{(1 + t_2)\lambda_2}{1 + t_2\lambda_2}\, d\varrho(t_1, t_2) \tag{4.5}$$

with some positive Radon measure ϱ on $[0, \infty]^2$ is in the class C_{A_1,A_2}.

We shall say that a function h on $\sigma(A_1) \times \sigma(A_2)$ has the *separate interpolation-property* if for each fixed $b \in \sigma(A_2)$ the function $\lambda_1 \mapsto h(\lambda_1, b)$ is of class C_{A_1}, and a similar statement holds for all functions $\lambda_2 \mapsto h(a, \lambda_2)$.

Lemma 4.6 *Each function of class C_{A_1,A_2} has the separate interpolation-property.*

Proof Let $T_2 = 0$ and take an arbitrary T_1 with $T_1{}^*T_1 \le 1$ and $T_1{}^*A_1T_1 \le A_1$. By hypothesis,

$$(T_1 \otimes 1)^* h(A_1, A_2)(T_1 \otimes 1) \le h(A_1, A_2).$$

Fix an eigenvalue b of A_2 and let y be a corresponding normalized eigenvector. Then for all $x \in H_1$ we have $\langle h(A_1, A_2)x \otimes y, x \otimes y\rangle = \langle h(A_1, b)x, x\rangle_{H_1}$ and $\langle (T_1 \otimes 1)^* h(A_1, A_2)(T_1 \otimes 1)x \otimes y, x \otimes y\rangle = \left\langle T_1{}^* h(A_1, b)T_1 x, x\right\rangle_{H_1}$ so

$$\left\langle T_1{}^* h(A_1, b)T_1 x, x\right\rangle_{H_1} \le \langle h(A_1, b)x, x\rangle_{H_1} .$$

The functions $h(a, \lambda_2)$ can be treated similarly. □

Example The function $h(\lambda_1, \lambda_2) = (\lambda_1 + \lambda_2)^{1/2}$ clearly has the separate interpolation-property for all A_1, A_2. However, it is not representable in the form (4.5). Indeed, $\mathrm{Re}\{h(i, i) - h(-i, i)\} = 1$ while it is easy to check that $\mathrm{Re}\{h(\lambda_1, \lambda_2) - h(\bar{\lambda}_1, \lambda_2)\} \le 0$ whenever $\mathrm{Im}\,\lambda_1, \mathrm{Im}\,\lambda_2 > 0$ and h is of the form (4.5).

Let us say that a function $h(\lambda_1, \lambda_2)$ defined on $\mathbf{R}_+ \times \mathbf{R}_+$ is an interpolation function (of two variables) if $h \in C_{A_1, A_2}$ for all A_1, A_2. Lemma 4.6 implies that interpolation functions are separately real-analytic in $\mathbf{R}_+ \times \mathbf{R}_+$ and that the functions $h(a, \cdot)$ and $h(\cdot, b)$ are of class P' (cf. Theorem III).

The above notion of interpolation function is close to Korányi's definition of monotone matrix function of two variables: $f(\lambda_1, \lambda_2)$ is *matrix monotone* in a rectangle $I = I_1 \times I_2$ (I_1, I_2 intervals in $\mathbf{R}$) if $A_1 \le A_1'$ (with spectra in I_1) and $A_2 \le A_2'$ (with spectra in I_2) implies

$$f(A_1', A_2') - f(A_1', A_2) - f(A_1, A_2') + f(A_1, A_2) \ge 0.$$

Lemma 4.7 *Each interpolation function is matrix monotone in* $\mathbf{R}_+ \times \mathbf{R}_+$.

Proof Let $0 < A_i \le A_i'$ and put $\tilde{A}_i = \begin{pmatrix} A_i' & 0 \\ 0 & A_i \end{pmatrix}$, $T_i = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$. Since $T_i{}^* \tilde{A}_i T_i \le \tilde{A}_i$, an interpolation function h will satisfy the interpolation inequality (4.4) with A_i replaced by $\tilde{A}_i$. Applying this inequality to vectors of the form $\begin{pmatrix} x_1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} x_2 \\ 0 \end{pmatrix}$ we readily obtain

$$\begin{aligned} &\left\langle h(A_1', A_2')x_1 \otimes x_2, x_1 \otimes x_2\right\rangle - \left\langle h(A_1, A_2')x_1 \otimes x_2, x_1 \otimes x_2\right\rangle \\ &- \left\langle h(A_1', A_2)x_1 \otimes x_2, x_1 \otimes x_2\right\rangle + \langle h(A_1, A_2)x_1 \otimes x_2, x_1 \otimes x_2\rangle \ge 0. \end{aligned}$$

The same result obtains with $x_1 \otimes x_2$ replaced by a sum $x_1 \otimes x_2 + x_1' \otimes x_2' + \ldots$, i.e., h is matrix monotone. □

Remark 4.8 Assume that f is of the form $f(\lambda_1, \lambda_2) = g_1(\lambda_1) + g_2(\lambda_2)$. Then f is matrix monotone for all g_1, g_2 and f is an interpolation function if and only if $g_1, g_2 \in P'$. In order to disregard "trivial" monotone functions of the above type, Korányi [22] imposed the normalizing assumption (a) $f(\lambda_1, 0) = f(0, \lambda_2) = 0$ for all λ_1, λ_2.

It follows from Lemma 4.7 and the proof of [22, Theorem 4] that if h is a C^2-smooth interpolation function, then the function

$$k(x_1, x_2; y_1, y_2) = \frac{h(x_1, x_2) - h(x_1, y_2) - h(y_1, x_2) + h(y_1, y_2)}{(x_1 - y_1)(x_2 - y_2)}$$

is *positive definite* in the sense that $\sum_m \sum_n k(x_m, y_m; x_n, y_n)\alpha_m\bar{\alpha}_n \geq 0$ for all finite sequences of positive numbers x_j, y_k and all complex numbers α_l. (The proof uses Löwner's matrix.) Korányi uses essentially this positive definiteness condition (and condition (a) in the remark above) to deduce an integral representation formula for h as an integral of products of Pick functions. See Theorem 3 in [22]. However, in contrast to our situation, Korányi considers functions monotone on the rectangle $(-1, 1) \times (-1, 1)$, so this last result cannot be immediately applied. (It easily implies local representation formulas, valid in finite rectangles, but these representations do not appear to be very natural from our point of view.)

This is not the right place to attempt to extend Korányi's methods to functions on $\mathbf{R}_+ \times \mathbf{R}_+$; it would seem more appropriate to give a more direct characterization of the classes C_{A_1,A_2} or of the class of interpolation functions. At present, we do not know if there is an interpolation function which is not representable in the form (4.5).

5 Proof of the K-Property

In this section we extend the result of Theorem 2.4 to obtain the full proof of Theorem I. The discussion is in principle not hard, but it does require some care to keep track of both norms when reducing to a finite-dimensional case.

Recall first that, by Lemma 2.3, it suffices to consider the diagonal case $\overline{\mathcal{H}} = \overline{\mathcal{K}}$.

To prove Theorem I we fix a regular Hilbert couple $\overline{\mathcal{H}}$; we must prove that it has the K-property (see Sect. 1.5). By Theorem 2.4, we know that this is true if $\overline{\mathcal{H}}$ is finite dimensional and the associated operator only has eigenvalues of unit multiplicity.

We shall use a weak* type compactness result ([2]). To formulate it, let $\mathcal{L}_1(\overline{\mathcal{H}})$ be the unit ball in the space $\mathcal{L}(\overline{\mathcal{H}})$. Moreover, let Σ_t be the sum $\mathcal{H}_0 + \mathcal{H}_1$ normed by $\| x \|^2_{\Sigma_t} := K(t, x)$. Note that $\| \cdot \|_{\Sigma_t}$ is an equivalent norm on Σ and that $\Sigma_1 = \Sigma$ isometrically. We denote by $\mathcal{L}_1(\Sigma_t)$ the unit ball in the space $\mathcal{L}(\Sigma_t)$.

In view of Remark 1.4, one has the identity

$$\mathcal{L}_1(\overline{\mathcal{H}}) = \bigcap_{t \in \mathbf{R}_+} \mathcal{L}_1(\Sigma_t). \tag{5.1}$$

We shall use this to define a compact topology on $\mathcal{L}_1(\overline{\mathcal{H}})$.

Lemma 5.1 *The subset* $\mathcal{L}_1\left(\overline{\mathcal{H}}\right) \subset \mathcal{L}_1(\Sigma)$ *is compact relative to the weak operator topology inherited from* $\mathcal{L}(\Sigma)$.

Recall that the *weak operator topology on* $\mathcal{L}(H)$ is the weakest topology such that a net T_i converges to the limit T if the inner product $\langle T_i x, y\rangle_H$ converges to $\langle Tx, y\rangle_H$ for all $x, y \in H$.

Proof of Lemma 5.1 The weak operator topology coincides on the unit ball $\mathcal{L}_1(\Sigma)$ with the weak*-topology, which is compact, due to Alaoglu's theorem (see [29], Chap. 4 for details). It is clear that for a fixed $t > 0$, the subset $\mathcal{L}_1(\Sigma) \cap \mathcal{L}_1(\Sigma_t)$ is weak operator closed in $\mathcal{L}_1(\Sigma)$; hence it is also compact. In view of (5.1), the set $\mathcal{L}_1\left(\overline{\mathcal{H}}\right)$ is an intersection of compact sets. Hence the set $\mathcal{L}_1\left(\overline{\mathcal{H}}\right)$ is itself compact, provided that we endow it with the subspace topology inherited from $\mathcal{L}_1(\Sigma)$. □

Denote by P_n the projections $P_n = E_{\sigma(A)\cap[n^{-1},n]}$ on $\mathcal{H}_0$ where E is the spectral resolution of A and $n = 1, 2, 3, \ldots$. Consider the couple

$$\overline{\mathcal{H}^{(n)}} = (P_n(\mathcal{H}_0), P_n(\mathcal{H}_1)),$$

the associated operator of which is the compression A_n of A to the subspace $P_n(\mathcal{H}_0)$. Note that the norms in the couple $\overline{\mathcal{H}^{(n)}}$ are equivalent, i.e., the associated operator A_n is bounded above and below.

We shall need two lemmas.

Lemma 5.2 *If* $\overline{\mathcal{H}^{(n)}}$ *has the* K*-property for all* n*, then so does* $\overline{\mathcal{H}}$.

Proof Note that $\| P_n \|_{\mathcal{L}(\overline{\mathcal{H}})} = 1$ for all n, and that $P_n \to 1$ as $n \to \infty$ relative to the strong operator topology on $\mathcal{L}(\Sigma)$. Suppose that $x^0, y^0 \in \Sigma$ are elements such that, for some $\rho > 1$,

$$K\left(t, y^0\right) < \frac{1}{\rho} K\left(t, x^0\right), \qquad t > 0. \tag{5.2}$$

Then $K\left(t, P_n y^0\right) \leq K\left(t, y^0\right) < \rho^{-1} K\left(t, x^0\right)$. Moreover, the identity $K\left(t, P_n y^0\right) = \left\langle \frac{tA_n}{1+tA_n} P_n y^0, P_n y^0\right\rangle_0$ shows that we have an estimate of the form $K(t, P_n y^0) \leq C_n \min\{1, t\}$ for $t > 0$ and large enough C_n (this follows since A_n is bounded above and below).

The functions $K\left(t, P_m x^0\right)$ increase monotonically, converging uniformly on compact subsets of $\mathbf{R}_+$ to $K\left(t, x^0\right)$ when $m \to \infty$. By concavity of the function $t \mapsto K\left(t, P_m x^0\right)$ we will then have

$$K\left(t, P_n y^0\right) < \frac{1}{\tilde{\rho}} K\left(t, P_m x^0\right), \qquad t \in \mathbf{R}_+, \tag{5.3}$$

provided that m is sufficiently large, where $\tilde{\rho}$ is any number in the interval $1 < \tilde{\rho} < \rho$.

Indeed, let $A = \lim_{t\to\infty} K\left(t, P_n y^0\right)$ and $B = \lim_{t\to 0} K\left(t, P_n y^0\right)/t$. Take points $t_0 < t_1$ such that $K(t, P_n y^0) \geq A/\rho'$ when $t \geq t_1$ and $K(t, P_n y^0)/t \leq B\rho'$ when $t \leq t_0$. Here ρ' is some number in the interval $1 < \rho' < \rho$.

Next use (5.2) to choose m large enough that $K(t, P_m x^0) > \rho K(t, P_n y^0)$ for all $t \in [t_0, t_1]$. Then $K(t, P_m x^0) > (\rho/\rho')K(t, P_n y^0)$ for $t = t_1$, hence for all $t \geq t_1$, and $K(t, P_m x^0)/t > (\rho/\rho')K(t, P_n y^0)/t$ for $t = t_0$ and hence also when $t \leq t_0$. Choosing $\rho' = \rho/\tilde{\rho}$ now establishes (5.3).

Put $N = \max\{m, n\}$. If $\overline{\mathcal{H}^{(N)}}$ has the K-property, we can find a map $T_{nm} \in \mathcal{L}_1(\overline{\mathcal{H}})$ such that $T_{nm} P_m x^0 = P_n y^0$. (Define $T_{mn} = 0$ on the orthogonal complement of $P_N\left(\mathcal{H}_0\right)$ in Σ.) In view of Lemma 5.1, the maps T_{nm} must cluster at some point $T \in \mathcal{L}_1(\overline{\mathcal{H}})$. It is clear that $Tx^0 = y^0$. Since $\rho > 1$ was arbitrary, we have shown that $\overline{\mathcal{H}}$ has the K-property. □

Lemma 5.3 *Given $x^0, y^0 \in \mathcal{H}_0^{(n)}$ and a number $\epsilon > 0$ there exists a positive integer n and a finite-dimensional couple $\overline{\mathcal{V}} \subset \overline{\mathcal{H}^{(n)}}$ such that $x^0, y^0 \in \mathcal{V}_0 + \mathcal{V}_1$ and*

$$(1-\epsilon)K\left(t, x; \overline{\mathcal{H}}\right) \leq K\left(t, x; \overline{\mathcal{V}}\right) \leq (1+\epsilon)K\left(t, x; \overline{\mathcal{H}}\right), \quad t > 0,\ x \in \mathcal{V}_0 + \mathcal{V}_1. \tag{5.4}$$

Moreover, $\overline{\mathcal{V}}$ can be chosen so that all eigenvalues of the associated operator $A_{\overline{\mathcal{V}}}$ are of unit multiplicity.

Proof Let A_n be the operator associated with the couple $\overline{\mathcal{H}^{(n)}}$; thus $1/n \leq A_n \leq n$.

Take $\eta > 0$ and let $\{\lambda_i\}_1^N$ be a finite subset of $\sigma\left(A_n\right)$ such that $\sigma\left(A_n\right) \subset \cup_1^N E_i$ where $E_i = (\lambda_i - \eta/2, \lambda_i + \eta/2)$. We define a Borel function $w : \sigma\left(A_n\right) \to \sigma\left(A_n\right)$ by $w(\lambda) = \lambda_i$ on $E_i \cap \sigma\left(A_n\right)$, then $\| w\left(A_n\right) - A_n \|_{\mathcal{L}(\mathcal{H}_0)} \leq \eta$.

Let $k_t(\lambda) = \frac{t\lambda}{1+t\lambda}$. It is easy to check that the Lipschitz constant of the restriction $k_t \,\big|\, \sigma\left(A_n\right)$ is bounded above by $C_1 \min\{1, t\}$ where $C_1 = C_1(n)$ is independent of t. Hence

$$\| k_t\left(w\left(A_n\right)\right) - k_t\left(A_n\right) \|_{\mathcal{L}(\mathcal{H}_0)} \leq C_1 \eta \min\{1, t\}.$$

It follows readily that

$$\left|\langle (k_t\left(w\left(A_n\right)\right) - k_t\left(A_n\right)) x, x\rangle_0\right| \leq C_1 \eta \min\{1, t\} \| x \|_0^2, \quad x \in P_n\left(\mathcal{H}_0\right).$$

Now let $c > 0$ be such that $A \geq c$. The elementary inequality $k_t(c) \geq (1/2)\min\{1, ct\}$ shows that

$$\langle k_t\left(A_n\right) x, x\rangle_0 \geq C_2 \min\{1, t\} \| x \|_0^2, \quad x \in P_n\left(\mathcal{H}_0\right),$$

where $C_2 = (1/2)\min\{1, c\}$. Combining these estimates, we deduce that

$$\left|\langle k_t\left(w\left(A_n\right)\right) x, x\rangle_0 - \langle k_t\left(A_n\right) x, x\rangle_0\right| \leq C_3 \eta \langle k_t\left(A_n\right) x, x\rangle_0, \quad x \in P_n\left(\mathcal{H}_0\right) \tag{5.5}$$

for some suitable constant $C_3 = C_3(n)$.

Now pick unit vectors e_i, f_i supported by the spectral sets $E_i \cap \sigma(A_n)$ such that x^0 and y^0 belong to the space $\mathcal{W}$ spanned by $\{e_i, f_i\}_1^N$. Put $\mathcal{W}_0 = \mathcal{W}_1 = \mathcal{W}$ and define norms on those spaces by

$$\| x \|_{\mathcal{W}_0} = \| x \|_{\mathcal{H}_0} \quad , \quad \| x \|_{\mathcal{W}_1}^2 = \langle w(A)\, x, x\rangle_{\mathcal{H}_0} .$$

The operator associated with $\overline{\mathcal{W}}$ is then the compression of $w(A_n)$ to $\mathcal{W}_0$, i.e.,

$$\| x \|_{\mathcal{W}_1}^2 = \left\langle A_{\overline{\mathcal{W}}} x, x\right\rangle_{\mathcal{W}_0} = \langle w(A_n)x, x\rangle_{\mathcal{H}_0} , \qquad x \in \mathcal{W}.$$

Let $\epsilon = 2C_3\eta$ and observe that, by (5.5)

$$\left| K\left(t, x; \overline{\mathcal{W}}\right) - K\left(t, x; \overline{\mathcal{H}}\right) \right| \leq (\epsilon/2) K\left(t, x; \overline{\mathcal{H}}\right), \quad f \in \mathcal{W}. \tag{5.6}$$

The eigenvalues of $A_{\overline{\mathcal{W}}}$ typically have multiplicity 2. To obtain unit multiplicity, we perturb $A_{\overline{\mathcal{W}}}$ slightly to a positive matrix $A_{\overline{\mathcal{V}}}$ such that $\left\| A_{\overline{\mathcal{W}}} - A_{\overline{\mathcal{V}}} \right\|_{\mathcal{L}(\mathcal{H}_0)} < \epsilon/2C_3$. Let $\overline{\mathcal{V}}$ be the couple associated with $A_{\overline{\mathcal{V}}}$, i.e., put $\mathcal{V}_i = \mathcal{W}$ for $i = 0, 1$ and

$$\| x \|_{\mathcal{V}_0} = \| x \|_{\mathcal{W}_0} \quad \text{and} \quad \| x \|_{\mathcal{W}_1}^2 = \left\langle A_{\overline{\mathcal{V}}} x, x\right\rangle_{\mathcal{V}_0} .$$

It is then straightforward to check that

$$\left| K\left(t, f; \overline{\mathcal{W}}\right) - K\left(t, f; \overline{\mathcal{V}}\right) \right| \leq (\epsilon/2) K\left(t, f; \overline{\mathcal{H}}\right), \quad f \in \mathcal{W}.$$

Combining this with the estimate (5.6), one finishes the proof of the lemma. □

Proof of Theorem 1 Given two elements $x^0, y^0 \in \Sigma$ as in (5.2) we write $x^n = P_n\left(x^0\right)$ and $y^n = P_n\left(y^0\right)$. By the proof of Lemma 5.2 we then have $K\left(t, y^n\right) \leq \tilde{\rho}^{-1} K\left(t, x^n\right)$ for large enough n, where $\tilde{\rho}$ is any given number in the interval $(1, \rho)$.

We then use Lemma 5.3 to choose a finite-dimensional sub-couple $\overline{\mathcal{V}} \subset \overline{\mathcal{H}^{(n)}}$ such that

$$\begin{aligned} K\left(t, y^n; \overline{\mathcal{V}}\right) &\leq (1+\epsilon) K\left(t, y^n; \overline{\mathcal{H}}\right) \\ &< \tilde{\rho}^{-1} K\left(t, x^n; \overline{\mathcal{V}}\right) + \epsilon\left(K\left(t, x^n; \overline{\mathcal{H}}\right) + K\left(t, y^n; \overline{\mathcal{H}}\right)\right). \end{aligned}$$

Here $\epsilon > 0$ is at our disposal.

Choosing ϵ sufficiently small, we can arrange that

$$K\left(t, y^n; \overline{\mathcal{V}}\right) \leq K(t, x^n; \overline{\mathcal{V}}), \quad t > 0. \tag{5.7}$$

By Theorem 2.4, the condition (5.7) implies the existence of an operator $T' \in \mathcal{L}_1\left(\overline{\mathcal{V}}\right)$ such that $T'x^n = y^n$. Considering the canonical inclusion and projection

$$I : \Sigma\left(\mathcal{V}\right) \to \Sigma\left(\mathcal{H}\right) \quad \text{and} \quad \Pi : \Sigma\left(\mathcal{H}\right) \to \Sigma\left(\mathcal{V}\right),$$

we have, by virtue of Lemma 5.3,

$$\| I \|^2_{\mathcal{L}(\overline{\mathcal{V}};\overline{\mathcal{H}})} \leq (1-\epsilon)^{-1} \quad \text{and} \quad \| \Pi \|^2_{\mathcal{L}(\overline{\mathcal{H}};\overline{\mathcal{V}})} \leq 1+\epsilon.$$

Now let $T = T_\varepsilon := IT'\Pi \in \mathcal{L}(\overline{\mathcal{H}^{(n)}})$. Then $\| T \|^2 \leq \frac{1+\epsilon}{1-\epsilon}$ and $Tx^n = y^n$. As $\epsilon \downarrow 0$ the operators T_ϵ will cluster at some point $T \in \mathcal{L}_1(\overline{\mathcal{H}^{(n)}})$ such that $Tx^n = y^n$ (cf. Lemma 5.1).

We have shown that $\overline{\mathcal{H}^{(n)}}$ has the K-property . In view of Lemma 5.2, this implies that $\overline{\mathcal{H}}$ has the same property. The proof of Theorem I is therefore complete. □

6 Representations of Interpolation Functions

6.1 Quadratic Interpolation Methods

Let us say that an interpolation method defined on regular Hilbert couples taking values in Hilbert spaces is a *quadratic interpolation method*. (Donoghue [14] used the same phrase in a somewhat wider sense, allowing the methods to be defined on non-regular Hilbert couples as well.)

If F is an exact quadratic interpolation method, and $\overline{\mathcal{H}}$ a Hilbert couple, then by Donoghue's theorem III there exists a positive Radon measure ϱ on $[0,\infty]$ such that $F\left(\overline{\mathcal{H}}\right) = \mathcal{H}_\varrho$, where the latter space is defined by the familiar norm $\|x\|^2_\varrho = \int_{[0,\infty]} \left(1+t^{-1}\right) K(t,x)\, d\varrho(t)$.

A priori, the measure ϱ could depend not only on F but also on the particular $\overline{\mathcal{H}}$. That ϱ is independent of $\overline{\mathcal{H}}$ can be realized in the following way. Let $\overline{\mathcal{H}'}$ be a regular Hilbert couple such that every positive rational number is an eigenvalue of the associated operator. Let B' be the operator associated with the exact quadratic interpolation space $F\left(\overline{\mathcal{H}'}\right)$. There is then clearly a unique P'-function h on $\sigma\left(A'\right)$ such that $B' = h\left(A'\right)$, viz. there is a unique positive Radon measure ϱ on $[0,\infty]$ such that $F\left(\overline{\mathcal{H}'}\right) = \mathcal{H}'_\rho$ (see Sect. 1.2 for the notation).

If $\overline{\mathcal{H}}$ is any regular Hilbert couple, we can form the direct sum $\overline{\mathcal{S}} = \overline{\mathcal{H}'} \oplus \overline{\mathcal{H}}$. Denote by $\tilde{A}$ the corresponding operator and let $\tilde{B} = \tilde{h}(\tilde{A})$ be the operator corresponding to the exact quadratic interpolation space $F\left(\overline{\mathcal{S}}\right)$. Then $\tilde{h}(\tilde{A}) = \tilde{h}(A') \oplus \tilde{h}\,(A) = h\left(A'\right) \oplus \tilde{h}(A)$. This means that $\tilde{h}\left(A'\right) = h\left(A'\right)$, i.e., $\tilde{h} = h$. In particular, the operator B corresponding to the exact interpolation space $F(\overline{\mathcal{H}})$ is equal to $h\,(A)$. We have shown that $F(\overline{\mathcal{H}}) = \mathcal{H}_\varrho$. We emphasize our conclusion with the following theorem.

Theorem 6.1 *There is a one-to-one correspondence* $\varrho \mapsto F$ *between positive Radon measures and exact quadratic interpolation methods.*

We will shortly see that Theorem 6.1 is equivalent to the theorem of Foiaş and Lions [16]. As we remarked above, a more general version of the theorem, admitting for non-regular Hilbert couples, is found in Donoghue's paper [14].

6.2 Interpolation Type and Reiteration

In this subsection, we prove some general facts concerning quadratic interpolation methods; we shall mostly follow Fan [15].

Fix a function $h \in P'$ of the form

$$h(\lambda) = \int_{[0,\infty]} \frac{(1+t)\lambda}{1+t\lambda}\, d\varrho(t).$$

It will be convenient to write $\overline{\mathcal{H}}_h$ for the corresponding exact interpolation space $\mathcal{H}_\varrho$. Thus, we shall denote

$$\|x\|_h^2 = \langle h(A)x, x\rangle_0 = \int_{[0,\infty]} \left(1 + t^{-1}\right) K\,(t, x)\; d\varrho(t).$$

More generally, we shall use the same notation when h is any quasi-concave function on $\mathbf{R}_+$, then $\overline{\mathcal{H}}_h$ is a quadratic interpolation space, but not necessarily exact.

Recall that, given a function $\mathbf{H}$ of one variable, we say that $\mathcal{H}_*$ is of type $\mathbf{H}$ with respect to $\overline{\mathcal{H}}$ if $\| T \|^2_{\mathcal{L}(\mathcal{H}_i)} \le M_i$ implies $\| T \|^2_{\mathcal{L}(\mathcal{H}_*)} \le M_0\, \mathbf{H}\,(M_1/M_0)$.

We shall say that a quasi-concave function h on $\mathbf{R}_+$ is *of type* $\mathbf{H}$ if $\overline{\mathcal{H}}_h$ is of type $\mathbf{H}$ relative to any regular Hilbert couple $\overline{\mathcal{H}}$. The following result somewhat generalizes Theorem 3.2. The class of functions of type $\mathbf{H}$ clearly forms a convex cone.

Theorem 6.2 *Let h be of type $\mathbf{H}$, where (i) $\mathbf{H}(1) = 1$ and $\mathbf{H}(t) \le \max\{1, t\}$, and (ii) $\mathbf{H}$ has left and right derivatives $\theta_\pm = H'(1\pm)$ at the point 1, where $\theta_- \le \theta_+$. Then for any positive constant c,*

$$\min\left\{\lambda^{\theta_-}, \lambda^{\theta_+}\right\} \le \frac{h(c\lambda)}{h(c)} \le \max\left\{\lambda^{\theta_-}, \lambda^{\theta_+}\right\}, \qquad \lambda \in \mathbf{R}_+. \tag{6.1}$$

In particular, if $\mathbf{H}(t)$ is differentiable at $t = 1$ and $\mathbf{H}'(1) = \theta$, then $h(\lambda) = \lambda^\theta$, $\lambda \in \mathbf{R}_+$.

Proof Replacing A by cA, it is easy to see that if h is of type $\mathbf{H}$, then so is $h_c(t) = h(ct)/h(c)$. Fix $\mu > 0$ and consider the function $h_0(t) = h_c(\mu t)/h_c(\mu)$. By Theorem II, we have $h_0(t) \le \mathbf{H}(t)$ for all t. Furthermore $h_0(1) = \mathbf{H}(1) = 1$ by (i). Since h_0 is differentiable, the assumption (ii) now gives $\theta_- \le h_0'(1) \le \theta_+$, or

$$\theta_- \le \frac{\mu h_c'(\mu)}{h_c(\mu)} \le \theta_+.$$

Dividing through by μ and integrating over the interval $[1, \lambda]$, one now verifies the inequalities in (6.1). □

The following result provides a partial converse to Theorem II.

Theorem 6.3 ([15]) *Let $h \in P'$ and set $\mathbf{H}(t) = \sup_{s>0} h(st)/h(s)$. Then h is of type $\mathbf{H}$.*

Proof Let $T \in \mathcal{L}(\overline{\mathcal{H}})$ be a non-zero operator; put $M_j = \| T \|^2_{\mathcal{L}(\mathcal{H}_j)}$ and $M = M_1/M_0$. We then have (by Lemma 1.1)

$$\begin{aligned}
\| Tx \|_h^2 &= \int_{[0,\infty]} \left(1 + t^{-1}\right) K\,(t, Tx)\; d\varrho(t) \\
&\le M_0 \int_{[0,\infty]} \left(1 + t^{-1}\right) K\,(tM, x)\; d\varrho(t) \\
&= M_0 \int_{[0,\infty]} \left\langle \frac{(1+t)MA}{1+tMA} x, x \right\rangle_0 d\varrho(t) \\
&= M_0 \left\langle h\,(MA)\, x, x \right\rangle_0 .
\end{aligned}$$

Letting E be the spectral resolution of A, we have

$$\langle h\,(MA)\, x, x\rangle_0 = \int_0^\infty h\,(M\lambda)\; d\,\langle E_\lambda x, x\rangle_0 .$$

Since $h\,(M\lambda)\,/\,h(\lambda) \le \mathbf{H}\,(M)$, we conclude that

$$\| Tx \|_h^2 \le M_0 \mathbf{H}\,(M) \int_0^\infty h\,(\lambda)\; d\,\langle E_\lambda x, x\rangle_0 = M_0 \mathbf{H}\,(M)\, \| x \|_h^2 ,$$

which finishes the proof. □

Given a function h of a positive variable, we define a new function $\tilde{h}$ by

$$\tilde{h}(s, t) = s\, h\,(t/s)\,.$$

The following reiteration theorem is due to Fan.

Theorem 6.4 ([15]) *Let $h, h_0, h_1 \in P'$, and $\varphi(\lambda) = \tilde{h}\,(h_0(\lambda), h_1(\lambda))$. Then $\overline{\mathcal{H}}_\varphi = (\overline{\mathcal{H}}_{h_0}, \overline{\mathcal{H}}_{h_1})_h$ with equal norms. Moreover, $\overline{\mathcal{H}}_\varphi$ is an exact interpolation space relative to $\overline{\mathcal{H}}$.*

Proof Let $\overline{\mathcal{H}'}$ denote the couple $(\overline{\mathcal{H}}_{h_0}, \overline{\mathcal{H}}_{h_1})$. The corresponding operator A' then obeys

$$\| x \|_{\overline{\mathcal{H}}_{h_1}} = \| (A')^{1/2} x \|_{\mathcal{H}'_0} = \| \varphi_0(A)^{1/2} (A')^{1/2} x \|_0, \quad x \in \Delta(\overline{\mathcal{H}'}).$$

On the other hand, $\| x \|_{\overline{\mathcal{H}}_{h_1}} = \| \varphi_1(A)^{1/2} x \|_0$, so

$$(A')^{1/2}x = \varphi_0(A)^{-1/2}\varphi_1(A)^{1/2}x, \quad x \in \Delta\left(\overline{\mathcal{H}'}\right).$$

We have shown that $A' = \varphi_0(A)^{-1}\varphi_1(A)$, whence (by Lemma 1.1)

$$\begin{aligned} K\left(t, x; \overline{\mathcal{H}'}\right) &= \left\langle \frac{t\varphi_0(A)^{-1}\varphi_1(A)}{1 + t\varphi_0(A)^{-1}\varphi_1(A)} x, x \right\rangle_{\mathcal{H}'_0} \\ &= \left\langle \frac{t\varphi_1(A)}{1 + t\varphi_0(A)^{-1}\varphi_1(A)} x, x \right\rangle_{\mathcal{H}'_0}. \end{aligned} \tag{6.2}$$

Now let the function $h \in P'$ be given by

$$h(\lambda) = \int_{[0,\infty]} \frac{(1+t)\lambda}{1+t\lambda}\, d\varrho(t),$$

and note that the function $\varphi = \tilde{h}\,(h_0, h_1)$ is given by

$$\varphi(\lambda) = \int_{[0,\infty]} \frac{(1+t)h_1(\lambda)}{1 + th_1(\lambda)/h_0(\lambda)}\, d\varrho(t).$$

Combining with (6.2), we find that

$$\begin{aligned} \| x \|^2_{\mathcal{H}'_h} &= \int_{[0,\infty]} \left(1 + t^{-1}\right) K\left(t, x; \overline{\mathcal{H}'}\right) d\varrho(t) \\ &= \int_0^\infty \left[\int_{[0,\infty]} \frac{(1+t)h_1(\lambda)}{1 + th_1(\lambda)/h_0(\lambda)}\, d\varrho(t) \right] d\,\langle E_\lambda x, x\rangle_0 = \| x \|^2_{\mathcal{H}_\varphi}. \end{aligned}$$

This finishes the proof of the theorem. □

Combining with Donoghue's theorem III, one obtains the following, purely function-theoretic corollary. Curiously, we are not aware of a proof which does not use interpolation theory.

Corollary 6.5 ([15]) *Suppose that $h \in P'$ and that $h_0, h_1 \in P'|F$, where F is some closed subset of $\mathbf{R}_+$. Then the function $\varphi = \tilde{h}(h_0, h_1)$ is also of class $P'|F$.*

6.3 Donoghue's Representation

Let $\overline{\mathcal{H}}$ be a regular Hilbert couple. In Donoghue's setting, the principal object is the space $\Delta = \mathcal{H}_0 \cap \mathcal{H}_1$ normed by $\| x \|^2_\Delta = \| x \|^2_0 + \| x \|^2_1$. In the following, all involutions are understood to be taken with respect to the norm of Δ.

We express the norms in the spaces $\mathcal{H}_i$ as

$$\| x \|_0^2 = \langle Hx, x\rangle_\Delta \qquad \text{and} \qquad \| x \|_1^2 = \langle (1-H)x, x\rangle_\Delta\,,$$

where H is a bounded positive operator on Δ, $0 \le H \le 1$. The regularity of $\overline{\mathcal{H}}$ means that neither 0 nor 1 is an eigenvalue of H.

To an arbitrary quadratic intermediate space $\mathcal{H}_*$ there corresponds a bounded positive injective operator K on Δ such that

$$\| x \|_*^2 = \langle Kx, x\rangle_\Delta\,.$$

It is then easy to see that $\mathcal{H}_*$ is exact interpolation if and only if, for bounded operators T on Δ, the conditions $T^*HT \le H$ and $T^*(1-H)T \le 1-H$ imply $T^*KT \le K$. It is straightforward to check that the relations between H, K and the operators A, B used in the previous sections are

$$H = \frac{1}{1+A} \quad , \quad A = \frac{1-H}{H} \quad , \quad K = \frac{B}{1+A} \quad , \quad B = \frac{K}{H}. \tag{6.3}$$

(It follows from the proof of Lemma 1.2 that H and K commute.)

By Theorem III we know that $\mathcal{H}_*$ is an exact interpolation space if and only if $B = h(A)$ for some $h \in P'$. By (6.3), this is equivalent to that $K = k(H)$ where

$$k(H) = \frac{h(A)}{1+A} = H\,h\left(\frac{1-H}{H}\right).$$

In its turn, this means that

$$\begin{aligned} k(\lambda) &= \lambda \int_{[0,\infty]} \frac{(1+t)(1-\lambda)/\lambda}{1+t(1-\lambda)/\lambda}\, d\varrho(t) \\ &= \int_{[0,\infty]} \frac{(1+t)\lambda(1-\lambda)}{\lambda + t(1-\lambda)}\, d\varrho(t), \qquad \lambda \in \sigma(H), \end{aligned}$$

where ϱ is a suitable Radon measure. Applying the change of variables $s = 1/(1+t)$ and defining a positive Radon measure μ on $[0, 1]$ by $d\mu(s) = d\varrho(t)$, we arrive at the expression

$$k(\lambda) = \int_0^1 \frac{\lambda(1-\lambda)}{(1-s)(1-\lambda) + s\lambda}\, d\mu(s), \qquad \lambda \in \sigma(H), \tag{6.4}$$

which gives the representation exact quadratic interpolation spaces originally used by Donoghue in [14].

6.4 J-Methods and the Foiaş-Lions Theorem

We define the (quadratic) J-functional relative to a regular Hilbert couple $\overline{\mathcal{H}}$ by

$$J(t,x) = J\left(t,x;\overline{\mathcal{H}}\right) = \| x \|_0^2 + t \| x \|_1^2, \qquad t>0,\ x \in \Delta(\overline{\mathcal{H}}).$$

Note that $J(t,x)^{1/2}$ is an equivalent norm on Δ and that $J(1,x) = \| x \|_\Delta^2$.

Given a positive Radon measure ν on $[0,\infty]$, we define a Hilbert space $J_\nu(\overline{\mathcal{H}})$ as the set of all elements $x \in \Sigma(\overline{\mathcal{H}})$ such that there exists a measurable function $u : [0,\infty] \to \Delta$ such that

$$x = \int_{[0,\infty]} u(t)\, d\nu(t) \quad \text{(convergence in } \Sigma) \tag{6.5}$$

and

$$\int_{[0,\infty]} \frac{J(t,u(t))}{1+t}\, d\nu(t) < \infty. \tag{6.6}$$

The norm in the space $J_\nu(\overline{\mathcal{H}})$ is defined by

$$\| x \|_{J_\nu}^2 = \inf_u \int_{[0,\infty]} \frac{J(t,u(t))}{1+t}\, d\nu(t) \tag{6.7}$$

over all u satisfying (6.5) and (6.6).

The space (6.7) was (with different notation) introduced by Foiaş and Lions in the paper [16], where it was shown that there is a unique minimizer $u(t)$ of the problem (6.7), namely

$$u(t) = \varphi_t(A)x \qquad \text{where} \qquad \varphi_t(\lambda) = \frac{1+t}{1+t\lambda}\left(\int_{[0,\infty]} \frac{1+s}{1+s\lambda}\, d\nu(s)\right)^{-1}. \tag{6.8}$$

Inserting this expression for u into (6.7), one finds that

$$\| x \|_{J_\nu}^2 = \langle h(A)x, x\rangle_0$$

where

$$h(\lambda)^{-1} = \int_{[0,\infty]} \frac{1+t}{1+t\lambda}\, d\nu(t). \tag{6.9}$$

It is easy to verify that the class of functions representable in the form (6.9) for some positive Radon measure ν coincides with the class P'. We have thus arrived at the following result.

Theorem 6.6 *Every exact quadratic interpolation space $\mathcal{H}_*$ can be represented isometrically in the form $\mathcal{H}_* = J_\nu(\overline{\mathcal{H}})$ for some positive Radon measure ν on* $[0, \infty]$. *Conversely, any space of this form is an exact quadratic interpolation space.*

In the original paper [16], Foiaş and Lions proved the less precise statement that each exact quadratic interpolation method F can be represented as $F = J_\nu$ for some positive Radon measure ν.

6.5 The Relation Between the *K*- and *J*-Representations

The assignment $K_\varrho = J_\nu$ gives rise to a non-trivial bijection $\varrho \mapsto \nu$ of the set of positive Radon measures on $[0, \infty]$. In this bijection, ϱ and ν are in correspondence if and only if

$$\int_{[0,\infty]} \frac{(1+t)\lambda}{1+t\lambda}\, d\varrho(t) = \left(\int_{[0,\infty]} \frac{1+t}{1+t\lambda}\, d\nu(t)\right)^{-1}.$$

As an example, let us consider the geometric interpolation space (where $c_\theta = \pi / \sin(\pi\theta)$)

$$\| x \|_\theta^2 = \left\langle A^\theta x, x\right\rangle_0 = c_\theta \int_0^\infty t^{-\theta} K(t, x)\, \frac{dt}{t}.$$

The measure ϱ corresponding to this method is $d\varrho_\theta(t) = \frac{c_\theta t^{-\theta}}{1+t}\, dt$. On the other hand, it is easy to check that

$$\lambda^\theta = \left(\int_0^\infty \frac{1+t}{1+t\lambda}\, d\nu_\theta(t)\right)^{-1} \quad \text{where} \quad d\nu_\theta(t) = \frac{c_\theta t^\theta}{1+t}\frac{dt}{t}.$$

We leave it to the reader to check that the norm in $\mathcal{H}_\theta$ is the infimum of the expression

$$c_\theta \int_0^\infty t^\theta J(t, u(t))\, \frac{dt}{t}$$

over all representations

$$x = \int_0^\infty u(t)\, \frac{dt}{t}.$$

We have arrived at the Hilbert space version of Peetre's J-method of exponent θ. The identity $J_{\nu_\theta} = K_{\varrho_\theta}$ can now be recognized as a sharp (isometric) Hilbert space version of the equivalence theorem of Peetre, which says that the standard K_θ and J_θ-methods give rise to equivalent norms on the category of Banach couples (see [7]).

The problem of determining the pairs ϱ, ν having the property that the K_ϱ and J_ν methods give equivalent norms was studied by Fan in [15, Section 3].

6.6 Other Representations

As we have seen in the preceding subsections, using the space $\mathcal{H}_0$ to express all involutions and inner products leads to a description of the exact quadratic interpolation spaces in terms of the class P'. If we instead use the space Δ as the basic object, we get Donoghue's representation for interpolation functions. Similarly, one can proceed from any fixed interpolation space $\mathcal{H}_*$ to obtain a different representation of interpolation functions.

6.7 On Interpolation Methods of Power p

Fix a number p, $1 < p < \infty$. We shall write $L_p = L_p(X, \mathcal{A}, \mu)$ for the usual L_p-space associated with an arbitrary but fixed (σ-finite) measure μ on a measure space $(X, \mathcal{A})$. Given a positive measurable weight function w, we write $L_p(w)$ for the space normed by

$$\| f \|_{L_p(w)}^p = \int_X |f(x)|^p \, w(x) \, d\mu(x).$$

We shall write $\overline{L}_p(w) = \big(L_p, L_p(w)\big)$ for the corresponding weighted L_p couple. Note that the conditions imposed mean precisely that $\overline{L}_p(w)$ be separable and regular.

Let us say that an exact interpolation functor F defined on the totality of separable, regular weighted L_p-couples and taking values in the class of weighted L_p-spaces is *of power p*.

Define, for a positive Radon measure ϱ on $[0, \infty]$, an exact interpolation functor $F = K_\varrho(p)$ by the definition

$$\| f \|_{F(\overline{L}_p(w))}^p := \int_{[0,\infty]} (1 + t^{-\frac{1}{p-1}})^{p-1} K_p\left(t, f; \overline{L}_p(w)\right) d\varrho(t).$$

We contend that F is of power p.

Indeed, it is easy to verify that

$$K_p\left(t, f; \overline{L}_p(w)\right) = \int_X |f(x)|^p \, \frac{t w(x)}{(1 + (t w(x))^{\frac{1}{p-1}})^{p-1}} \, d\mu(x),$$

so Fubini's theorem gives that

$$\| f \|_{F(\overline{L}_p(w))}^p = \int_X |f(x)|^p \, h(w(x)) \, d\mu(x),$$

where

$$h(\lambda) = \int_{[0,\infty]} \frac{(1 + t^{\frac{1}{p-1}})^{p-1} \lambda}{(1 + (t\lambda)^{\frac{1}{p-1}})^{p-1}} \, d\varrho(t), \qquad \lambda \in w(X). \tag{6.10}$$

We have shown that $F(\overline{L}_p(w)) = L_p(h(w))$, so F is indeed of power p.

Let us denote by $\mathcal{K}(p)$ the totality of positive functions h on $\mathbf{R}_+$ representable in the form (6.10) for some positive Radon measure ϱ on $[0, \infty]$.

Further, let $\mathcal{I}(p)$ denote the class of all (exact) *interpolation functions of power* p, i.e., those positive functions h on $\mathbf{R}_+$ having the property that for each weighted L_p couple $\overline{L}_p(w)$ and each bounded operator T on $\overline{L}_p(w)$, it holds that T is bounded on $L_p(h(w))$ and

$$\| T \|_{\mathcal{L}(L_p(h(w)))} \le \| T \|_{\mathcal{L}(\overline{L}_p(w))} .$$

The class $\mathcal{I}(p)$ is in a sense the natural candidate for the class of "operator monotone functions on L_p-spaces." The class $\mathcal{I}(p)$ clearly forms a convex cone; it was shown by Peetre [31] that this cone is contained in the class of concave positive functions on $\mathbf{R}_+$ (with equality if $p = 1$).

We have shown that $\mathcal{K}(p) \subset \mathcal{I}(p)$. By Theorem 6.1, we know that equality holds when $p = 2$. For other values of p it does not seem to be known whether the class $\mathcal{K}(p)$ exhausts the class $\mathcal{I}(p)$, but one can show that we would have $\mathcal{K}(p) = \mathcal{I}(p)$ provided that each finite-dimensional L_p-couple $\overline{\ell_p^n}(\lambda)$ has the K_p-property (or equivalently, the K-property, see (2.2)). Naturally, the latter problem (about the K_p-property) also seems to be open, but some comments on it are found in Remark 2.9.

Let ν be a positive Radon measure on $[0, \infty]$. In [16], Foiaş and Lions introduced a method, which we will denote by $F = J_\nu(p)$ in the following way. Define the J_p-functional by

$$J_p\left(t, f; \overline{L}_p(\lambda)\right) = \| f \|_0^p + t \, \| f \|_1^p , \quad f \in \Delta, \, t > 0.$$

We then define an intermediate norm by

$$\| f \|_{F(\overline{L}_p(\lambda))}^p := \inf \int_{[0,\infty]} (1 + t)^{-\frac{1}{p-1}} J_p\left(t, u(t); \overline{L}_p(\lambda)\right) d\nu(t),$$

where the infimum is taken over all representations

$$f = \int_{[0,\infty]} u(t)\, d\nu(t)$$

with convergence in Σ. It is straightforward to see that the method F so defined is exact; in [16] it is moreover shown that it is of power p. More precisely, it is there proved that

$$\| f \|^p_{F(\overline{L}_p(\lambda))} = \int_X |f(x)|^p\, h(w(x))\, d\mu(x),$$

where

$$h(\lambda)^{-\frac{1}{p-1}} = \int_{[0,\infty]} \frac{(1+t)^{\frac{1}{p-1}}}{(1+t\lambda)^{\frac{1}{p-1}}}\, d\nu(t), \qquad \lambda \in w(X). \tag{6.11}$$

Let us denote by $\mathcal{J}(p)$ the totality of functions h representable in the form (6.11). We thus have that $\mathcal{J}(p) \subset \mathcal{I}(p)$. In view of our preceding remarks, we conclude that if all weighted L_p-couples have the K_p property, then necessarily $\mathcal{J}(p) \subset \mathcal{K}(p)$. Note that $\mathcal{J}(2) = \mathcal{K}(2)$ by Theorem 6.6.

Appendix: The Complex Method is Quadratic

Let $S = \{z \in \mathbf{C};\ 0 \le \operatorname{Re} z \le 1\}$. Fix a Hilbert couple $\overline{\mathcal{H}}$ and let $\mathcal{F}$ be the set of functions $S \to \Sigma$ which are bounded and continuous in S, analytic in the interior of S, and which maps the line $j + i\mathbf{R}$ into $\mathcal{H}_j$ for $j = 0, 1$. Fix $0 < \theta < 1$. The norm in the complex interpolation space $C_\theta\left(\overline{\mathcal{H}}\right)$ is defined by

$$\| x \|_{C_\theta\left(\overline{\mathcal{H}}\right)} = \inf\left\{\| f \|_{\mathcal{F}};\ f(\theta) = x\right\}. \tag{*}$$

Let $\mathcal{P}$ denote the set of polynomials $f = \sum_1^N a_i z^i$ where $a_i \in \Delta$. We endow $\mathcal{P}$ with the inner product

$$\langle f, g\rangle_{M_\theta} = \sum_{j=0,1} \int_{\mathbf{R}} \langle f(j+it), g(j+it)\rangle_j\, P_j(\theta, t)\, dt,$$

where $\{P_0, P_1\}$ is the Poisson kernel for S,

$$P_j(\theta, t) = \frac{e^{-\pi t} \sin\theta\pi}{\sin^2\theta\pi + (\cos\theta\pi - (-1)^j e^{-\pi t})^2}.$$

Let M_θ be the completion of $\mathcal{P}$ with this inner product. It is easy to see that the elements of M_θ are analytic in the interior of S, and that evaluation map $f \mapsto f(\theta)$ is continuous on M_θ. Let N_θ be the kernel of this functional and define a Hilbert space $\mathcal{H}_\theta$ by

$$\mathcal{H}_\theta = M_\theta / N_\theta.$$

We denote the norm in $\mathcal{H}_\theta$ by $\| \cdot \|_\theta$.

Proposition A.1 *$C_\theta\left(\overline{\mathcal{H}}\right) = \mathcal{H}_\theta$ with equality of norms.*

Proof Let $f \in \mathcal{F}$. By the Calderón lemma in [7, Lemma 4.3.2], we have the estimate

$$\log \| f(\theta) \|_{C_\theta(\overline{\mathcal{H}})} \leq \sum_{j=0,1} \int_{\mathbf{R}} \log \| f(j+it) \|_j P_j(\theta,t)\, dt.$$

Applying Jensen's inequality, this gives that

$$\| f(\theta) \|_{C_\theta(\overline{\mathcal{H}})} \leq \Big(\sum_{j=0,1} \int_{\mathbf{R}} \| f(j+it) \|_j^2 \, P_j(\theta,t)\, dt\Big)^{1/2} = \| f \|_{M_\theta} .$$

Hence $\mathcal{H}_\theta \subset C_\theta(\overline{\mathcal{H}})$ and $\|\cdot\|_{C_\theta(\overline{\mathcal{H}})} \leq \| \cdot \|_\theta$. On the other hand, for $f \in \mathcal{P}$ one has the estimates

$$\| f(\theta) \|_\theta \leq \| f \|_{M_\theta} \leq \sup\{ \| f(j+it) \|_j \,;\, t \in \mathbf{R},\ j = 0, 1\} = \| f \|_{\mathcal{F}} ,$$

whence $C_\theta(\overline{\mathcal{H}}) \subset \mathcal{H}_\theta$ and $\|\cdot\|_{C_\theta(\overline{\mathcal{H}})} \geq \|\cdot\|_\theta$. □

It is well known that the method C_θ is of exponent θ (see, e.g., [7]). We have shown that C_θ is an exact quadratic interpolation method of exponent θ.

Complex Interpolation with Derivatives

In [15, pp. 421–422], Fan considers the more general complex interpolation method $C_{\theta(n)}$ for the n:th derivative. This means that in (*), one consider representations $x = \frac{1}{n!} f^{(n)}(\theta)$ where $f \in \mathcal{F}$; the complex method C_θ is thus the special case $C_{\theta(0)}$. It is shown in [15] that, for $n \geq 1$, the $C_{\theta(n)}$-method is represented, up to equivalence of norms, by the quasi-power function $h(\lambda) = \lambda^\theta / (1 + \frac{\theta(1-\theta)}{n} \,|\log \lambda\,|)^n$. The complex method with derivatives was introduced by Schechter [37]; for more details on that method, we refer to the list of references in [15].

.

References

1. J. Agler, J.E. McCarthy, N. Young, Operator monotone functions and Löwner functions of several variables. Ann. Math. **176**, 1783–1826 (2012)
2. Y. Ameur, The Calderón problem for Hilbert couples. Ark. Mat. **41**, 203–231 (2003)
3. Y. Ameur, A new proof of Donoghue's interpolation theorem. J. Funct. Space Appl. **3**, 253–265 (2004)
4. Y. Ameur, A note on a theorem of Sparr. Math. Scand. **94**, 155–160 (2004)
5. Y. Ameur, M. Cwikel, On the K-divisibility constant for some special, finite-dimensional Banach couples. J. Math. Anal. Appl. **360**, 130–155 (2009)
6. N. Aronszajn, W. Donoghue, On exponential representations of functions. J. Anal. Math. **5**, 321–388 (1956–57)
7. J. Bergh, J. Löfström, *Interpolation Spaces, an Introduction* (Springer, Berlin, 1976)
8. Y.A. Brudnyi, N.Y. Krugljak, *Interpolation Functors and Interpolation Spaces* (North Holland, Amsterdam, 1991)
9. A.P. Calderón, Intermediate spaces and interpolation, the complex method. Studia Math. **24**, 113–190 (1964)
10. A.P. Calderón, Spaces between L_1 and L_∞ and the theorem of Marcinkiewicz. Studia Math. **26**, 273–299 (1966)
11. M. Cwikel, Monotonicity properties of interpolation spaces. Ark. Mat. **14**, 213–236 (1976)
12. W. Donoghue, *Monotone Matrix Functions and Analytic Continuation* (Springer, Berlin, 1974)
13. W. Donoghue, The theorems of Loewner and Pick. Israel J. Math. **4**, 153–170 (1966)
14. W. Donoghue, The interpolation of quadratic norms. Acta Math. **118**, 251–270 (1967)
15. M. Fan, Quadratic interpolation and some operator inequalities. J. Math. Inequal. **5**, 413–427 (2011)
16. C. Foiaş, J.L. Lions, Sur certains théorèmes d'interpolation. Acta Sci. Math. **22**, 269–282 (1961)
17. C. Foiaş, S.C. Ong, P. Rosenthal, An interpolation theorem and operator ranges. Integr. Equ. Oper. Theory **10**, 802–811 (1987)
18. P.R. Halmos, Quadratic interpolation. J. Oper. Theory **7**, 303–305 (1982)
19. F. Hansen, An operator inequality. Math. Ann. **246**, 249–250 (1980)
20. F. Hansen, Selfadjoint means and operator monotone functions. Math. Ann. **256**, 29–35 (1981)
21. O. Heinävaara, Local characterizations for the matrix monotonicity and convexity of fixed order. Proc. Am. math. Soc. **146**, 3791–3799 (2018)
22. A. Korányi, On some classes of analytic functions of several variables. Trans. Am. Math. Soc. **101**, 520–554 (1961)
23. F. Kraus, Über konvexe Matrixfunktionen. Math. Z. **41**, 18–42 (1936)
24. J.L. Lions, Espaces intermédiaires entre espaces Hilbertiens et applications. Bull. Math. de la Soc. Sci. Math. Phys. de la R. P. R. **2**, 419–432 (1958)
25. J.L. Lions, E. Magenes, *Non-Homogeneous Boundary Value Problems and Applications*, vol. 1 (Springer, New York, 1972)
26. K. Löwner, Über monotone Matrixfunktionen. Math. Z. **38**, 177–216 (1934)
27. J.E. McCarthy, Geometric interpolation of Hilbert spaces. Ark. Mat. **30**, 321–330 (1992)
28. B. Mityagin, An interpolation theorem for modular spaces. Mat. Sbornik **66**, 473–482 (1965)
29. G.J. Murphy, *C^*-Algebras and Operator Theory* (Academic, New York, 1991)
30. V.I. Ovchinnikov, The method of orbits in interpolation theory. Math. Rep. **1**, 349–515 (1984)
31. J. Peetre, On interpolation functions III. Acta Sci. Math. (Szeged) **30**, 235–239 (1969)
32. J. Peetre, Two new interpolation methods based on the duality map. Acta Math. **173**, 73–91 (1979)
33. W. Pusz, S.L. Woronowicz, Functional calculus for sesquilinear forms and the purification map. Rep. Math. Phys. **8**, 159–170 (1975)

34. M. Rosenblum, J. Rovnyak, *Hardy Classes and Operator Theory* (Dover, Mineola, 1997)
35. A. Sedaev, Description of interpolation spaces for the pair $\left(L_p(a_0), L_p(a_1)\right)$. Soviet Math. Dokl. **14**, 538–541 (1973)
36. A. Sedaev, E.M. Semenov, On the possibility of describing interpolation spaces in terms of the K-functional of Peetre. Optimizacja **4**, 98–114 (1971)
37. M. Shechter, Complex interpolation. Compos. Math. **18**, 117–147 (1967)
38. G. Sparr, Interpolation of weighted L^p spaces. Studia Math. **62**, 229–271 (1978)
39. G. Sparr, A new proof of Löwner's theorem. Math. Scand **47**, 266–274 (1980)
40. A. Uhlmann, Relative entropy and the Wigner–Yanase–Dyson–Lieb concavity in an interpolation theory. Commun. Math. Phys. **54**, 21–32 (1977)
41. H. Vasudeva, On monotone matrix functions of two variables. Trans. Am. Math. Soc. **176**, 305–318 (1973)

A Panorama of Positivity. I: Dimension Free

Alexander Belton, Dominique Guillot, Apoorva Khare, and Mihai Putinar

Serguei Shimorin, in memoriam

Abstract This survey contains a selection of topics unified by the concept of positive semidefiniteness (of matrices or kernels), reflecting natural constraints imposed on discrete data (graphs or networks) or continuous objects (probability or mass distributions). We put emphasis on entrywise operations which preserve positivity, in a variety of guises. Techniques from harmonic analysis, function theory, operator theory, statistics, combinatorics, and group representations are invoked. Some partially forgotten classical roots in metric geometry and distance transforms are

D.G. is partially supported by a University of Delaware Research Foundation grant, by a Simons Foundation collaboration grant for mathematicians, and by a University of Delaware Research Foundation Strategic Initiative grant. A.K. is partially supported by Ramanujan Fellowship SB/S2/RJN-121/2017 and MATRICS grant MTR/2017/000295 from SERB (Govt. of India), by grant F.510/25/CAS-II/2018(SAP-I) from UGC (Govt. of India), and by a Young Investigator Award from the Infosys Foundation.

A. Belton
Department of Mathematics and Statistics, Lancaster University, Lancaster, UK
e-mail: a.belton@lancaster.ac.uk

D. Guillot
University of Delaware, Newark, DE, USA
e-mail: dguillot@udel.edu

A. Khare (✉)
Indian Institute of Science, Analysis and Probability Research Group, Bengaluru, India
e-mail: khare@iisc.ac.in

M. Putinar
University of California at Santa Barbara, Santa Barbara, CA, USA

Newcastle University, Newcastle upon Tyne, UK
e-mail: mputinar@math.ucsb.edu; mihai.putinar@ncl.ac.uk

A. Aleman et al. (eds.), *Analysis of Operators on Function Spaces*,
Trends in Mathematics, https://doi.org/10.1007/978-3-030-14640-5_5

presented with comments and full bibliographical references. Modern applications to high-dimensional covariance estimation and regularization are included.

Keywords Metric geometry · Positive semidefinite matrix · Toeplitz matrix · Hankel matrix · Positive definite function · Completely monotone functions · Absolutely monotonic functions · Entrywise calculus · Generalized Vandermonde matrix · Schur polynomials · Symmetric function identities · Totally positive matrices · Totally non-negative matrices · Totally positive completion problem · Sample covariance · Covariance estimation · Hard/soft thresholding · Sparsity pattern · Critical exponent of a graph · Chordal graph · Loewner monotonicity · Convexity · Super-additivity

2010 Mathematics Subject Classification 15-02, 26-02, 15B48, 51F99, 15B05, 05E05, 44A60, 15A24, 15A15, 15A45, 15A83, 47B35, 05C50, 30E05, 62J10

This is the first part of a two-part survey; we include on p. 165 the table of contents for the second part [10]. The survey in its unified form may be found online; see [9]. The abstract, keywords, MSC codes, and introduction are the same for both parts.

1 Introduction

Matrix positivity, or positive semidefiniteness, is one of the most wide-reaching concepts in mathematics, old and new. Positivity of a matrix is as natural as positivity of mass in statics or positivity of a probability distribution. It is a notion which has attracted the attention of many great minds. Yet, after at least two centuries of research, positive matrices still hide enigmas and raise challenges for the working mathematician.

The vitality of matrix positivity comes from its breadth, having many theoretical facets and also deep links to mathematical modelling. It is not our aim here to pay homage to matrix positivity in the large. Rather, the present survey, split for technical reasons into two parts, has a limited but carefully chosen scope.

Our panorama focuses on entrywise transforms of matrices which preserve their positive character. In itself, this is a rather bold departure from the dogma that canonical transformations of matrices are not those that operate entry by entry. Still, this apparently esoteric topic reveals a fascinating history, abundant characteristic phenomena, and numerous open problems. Each class of positive matrices or kernels (regarding the latter as continuous matrices) carries a specific toolbox of internal transforms. Positive Hankel forms or Toeplitz kernels, totally positive matrices, and group-invariant positive definite functions all possess specific *positivity preservers*. As we see below, these have been thoroughly studied for at least a century.

One conclusion of our survey is that the classification of positivity preservers is accessible in the dimension-free setting, that is, when the sizes of matrices are unconstrained. In stark contrast, precise descriptions of positivity preservers in fixed dimension are elusive, if not unattainable with the techniques of modern mathematics. Furthermore, the world of applications cares much more about matrices of fixed size than in the free case. The accessibility of the latter was by no means a sequence of isolated, simple observations. Rather, it grew organically out of distance geometry, and spread rapidly through harmonic analysis on groups, special functions, and probability theory. The more recent and highly challenging path through fixed dimensions requires novel methods of algebraic combinatorics and symmetric functions, group representations, and function theory.

As well as its beautiful theoretical aspects, our interest in these topics is also motivated by the statistics of big data. In this setting, functions are often applied entrywise to covariance matrices, in order to induce sparsity and improve the quality of statistical estimators (see [45, 46, 75]). Entrywise techniques have recently increased in popularity in this area, largely because of their low computational complexity, which makes them ideal to handle the ultra high-dimensional datasets arising in modern applications. In this context, the dimensions of the matrices are fixed, and correspond to the number of underlying random variables. Ensuring that positivity is preserved by these entrywise methods is critical, as covariance matrices must be positive semidefinite. Thus, there is a clear need to produce characterizations of entrywise preservers, so that these techniques are widely applicable and mathematically justified. We elaborate further on this in the second part of the survey [10].

We conclude by remarking that, while we have tried to be comprehensive in our coverage of the field of matrix positivity and the entrywise calculus, our panorama is far from being complete. We apologize for any omissions.

2 From Metric Geometry to Matrix Positivity

2.1 Distance Geometry

During the first decade of the twentieth century, the concept of a metric space emerged from the works of Fréchet and Hausdorff, each having different and well-anchored roots, in function spaces and in set theory and measure theory. We cannot think today of modern mathematics and physics without referring to metric spaces, which touch areas as diverse as economics, statistics, and computer science. Distance geometry is one of the early and ever-lasting by-products of metric-space theory. One of the key figures of the Vienna Circle, Karl Menger, started a systematic study in the 1920s of the geometric and topological features of spaces that are intrinsic solely to the distance they carry. Menger published his findings in a series of articles having the generic name "Untersuchungen über allgemeine Metrik," the first one being [63]; see also his synthesis [64]. His work was very influential in the decades to come [16], and by a surprising and fortunate stroke not often encountered in mathematics, Menger's distance geometry has been resurrected in recent times by practitioners of convex optimization and network analysis [26, 60].

Let (X, ρ) be a metric space. One of the naive, yet unavoidable, questions arising from the very beginning concerns the nature of operations $\phi(\rho)$ which may be performed on the metric and which enhance various properties of the topological space X. We all know that $\rho/(\rho + 1)$ and ρ^γ, if $\gamma \in (0, 1)$, also satisfy the axioms of a metric, with the former making it bounded. Less well known is an observation due to Blumenthal, that the new metric space (X, ρ^γ) has the four-point property if $\gamma \in (0, 1/2]$: every four-point subset of X can be embedded isometrically into Euclidean space [16, Section 49].

Metric spaces which can be embedded isometrically into Euclidean space, or into infinite-dimensional Hilbert space, are, of course, distinguished and desirable for many reasons. We owe to Menger a definitive characterization of this class of metric spaces. The core of Menger's theorem, stated in terms of certain matrices built from the distance function (known as Cayley–Menger matrices) was slightly reformulated by Fréchet and cast in the following simple form by Schoenberg.

Theorem 2.1 (Schoenberg [81]) *Let $d \geq 1$ be an integer and let (X, ρ) be a metric space. An $(n+1)$-tuple of points $x_0, x_1, \ldots, x_n$ in X can be isometrically embedded into Euclidean space $\mathbb{R}^d$, but not into $\mathbb{R}^{d-1}$, if and only if the matrix*

$$[\rho(x_0, x_j)^2 + \rho(x_0, x_k)^2 - \rho(x_j, x_k)^2]_{j,k=1}^n,$$

is positive semidefinite with rank equal to d.

Proof This is surprisingly simple. Necessity is immediate, since the Euclidean norm and scalar product in $\mathbb{R}^d$ give that

$$\begin{aligned}
&\rho(x_0, x_j)^2 + \rho(x_0, x_k)^2 - \rho(x_j, x_k)^2 \\
&= \|x_0 - x_j\|^2 + \|x_0 - x_k\|^2 - \|(x_0 - x_j) - (x_0 - x_k)\|^2 \\
&= 2\langle x_0 - x_j, x_0 - x_k\rangle,
\end{aligned}$$

and the latter are the entries of a positive semidefinite Gram matrix of rank less than or equal to d.

For the other implication, we consider first a full-rank $d \times d$ matrix associated with a $(d+1)$-tuple. The corresponding quadratic form

$$Q(\lambda) = \frac{1}{2} \sum_{j,k=1}^{d} (\rho(x_0, x_j)^2 + \rho(x_0, x_k)^2 - \rho(x_j, x_k)^2)\lambda_j \lambda_k$$

is positive definite. Hence there exists a linear change of variables

$$\lambda_k = \sum_{j=1}^{d} a_{jk}\mu_j \qquad (1 \leq j \leq d)$$

such that

$$Q(\lambda) = \mu_1^2 + \mu_2^2 + \cdots + \mu_d^2.$$

Interpreting $(\mu_1, \mu_2, \ldots, \mu_d)$ as coordinates in $\mathbb{R}^d$, the standard simplex with vertices

$$e_0 = (0, \ldots, 0), \quad e_1 = (1, 0, \ldots, 0), \quad \ldots, \quad e_d = (0, \ldots, 0, 1)$$

has the corresponding quadratic form (of distances) equal to $\mu_1^2 + \mu_2^2 + \cdots + \mu_d^2$. Now we perform the coordinate change $\mu_j \mapsto \lambda_j$. Specifically, set $P_0 = 0$ and let $P_j \in \mathbb{R}^d$ be the point with coordinates $\lambda_j = 1$ and $\lambda_k = 0$ if $k \neq j$. Then one identifies distances:

$$\begin{aligned} &\|P_0 - P_j\| = \rho(x_0, x_j) \qquad (0 \le j \le d) \\ \text{and} \quad &\|P_j - P_k\| = \rho(x_j, x_k) \qquad (1 \le j, k \le d). \end{aligned}$$

The remaining case with $n > d$ can be analyzed in a similar way, after taking an appropriate projection. □

In the conditions of the theorem, fixing a "frame" of d points and letting the $(d+1)$-th point float, one obtains an embedding of the full metric space (X, ρ) into $\mathbb{R}^d$. This idea goes back to Menger, and it led, with Schoenberg's touch, to the following definitive statement. Here and below, all Hilbert spaces are assumed to be separable.

Corollary 2.2 (Schoenberg [81], following Menger) *A separable metric space (X, ρ) can be isometrically embedded into Hilbert space if and only if, for every $(n+1)$-tuple of points $(x_0, x_1, \ldots, x_n)$ in X, where $n \ge 2$, the matrix*

$$[\rho(x_0, x_j)^2 + \rho(x_0, x_k)^2 - \rho(x_j, x_k)^2]_{j,k=1}^n$$

is positive semidefinite.

The notable aspect of the two previous results is the interplay between purely geometric concepts and matrix positivity. This will be a recurrent theme of our survey.

2.2 Spherical Distance Geometry

One can specialize the embedding question discussed in the previous section to submanifolds of Euclidean space. A natural choice is the sphere.

For two points x and y on the unit sphere $S^{d-1} \subset \mathbb{R}^d$, the rotationally invariant distance between them is

$$\rho(x, y) = \sphericalangle(x, y) = \arccos\langle x, y\rangle,$$

where the angle between the two vectors is measured on a great circle and is always less than or equal to π.

A straightforward application of the simple, but central, Theorem 2.1 yields the following result.

Theorem 2.3 (Schoenberg [81]) *Let (X, ρ) be a metric space and let $(x_1, \ldots, x_n)$ be an n-tuple of points in X. For any integer $d \geq 2$, there exists an isometric embedding of $(x_1, \ldots, x_n)$ into S^{d-1} endowed with the geodesic distance but not S^{d-2} if and only if*

$$\rho(x_j, x_k) \leq \pi \qquad (1 \leq j, k \leq n)$$

and the matrix $\left[\cos \rho(x_j, x_k)\right]_{j,k=1}^n$ is positive semidefinite of rank d.

Indeed, the necessity is assured by choosing x_0 to be the origin in $\mathbb{R}^d$. In this case,

$$\begin{aligned}\rho(x_0, x_j)^2 + \rho(x_0, x_k)^2 - \rho(x_j, x_k)^2 &= \|x_j\|^2 + \|x_k\|^2 - \|x_j - x_k\|^2 \\ &= 2\langle x_j, x_k\rangle \\ &= 2\cos \rho(x_j, x_k).\end{aligned}$$

The condition is also sufficient, by possibly adding an external point x_0 to the metric space, subject to the constraints that $\rho(x_0, x_j) = 1$ for all j. The details can be found in [81].[1]

2.3 Distance Transforms

A notable step forward in the study of the existence of isometric embeddings of a metric space into Euclidean or Hilbert space was made by Schoenberg. In a series of articles [82, 84, 85, 94], he changed the set-theoretic lens of Menger, by initiating a harmonic-analysis interpretation of this embedding problem. This was a major turning point, with long-lasting, unifying, and unexpected consequences.

[1] An alternate proof of sufficiency is to note that $A := [\cos \rho(x_j, x_k)]_{j,k=1}^n$ is a Gram matrix of rank r, hence equal to $B^T B$ for some $r \times n$ matrix B with unit columns. Denoting these columns by $\mathbf{b}_1, \ldots, \mathbf{b}_n \in S^{r-1}$, the map $x_j \mapsto b_j$ is an isometry since $\rho(x_j, x_k)$ and $\sphericalangle(y_j, y_k) \in [0, \pi]$. Moreover, since A has rank r, the $\mathbf{b}_j$ cannot all lie in a smaller-dimensional sphere.

We return to a separable metric space (X, ρ) and seek distance-function transforms $\rho \mapsto \phi(\rho)$ which enhance the geometry of X, to the extent that the new metric space $\big(X, \phi(\rho)\big)$ is isometrically equivalent to a subspace of Hilbert space. Schoenberg launched this whole new chapter from the observation that the Euclidean norm is such that the matrix

$$[\exp\big(-\|x_j - x_k\|^2\big)]_{j,k=1}^N$$

is positive semidefinite for any choice of points $x_1, \ldots, x_N$ in the ambient space. Once again, we see the presence of matrix positivity. While this claim may not be obvious at first sight, it is accessible once we recall a key property of Fourier transforms.

An even function $f : \mathbb{R}^d \to \mathbb{C}$ is said to be *positive definite* if the complex matrix $[f(x_j - x_k)]_{j,k=1}^N$ is positive semidefinite for any $N \geq 1$ and any choice of points $x_1, \ldots, x_N \in \mathbb{R}^d$. We will call $f(x - y)$ a *positive semidefinite kernel* on $\mathbb{R}^d \times \mathbb{R}^d$ in this case.

Bochner's theorem [18] characterizes positive definite functions on $\mathbb{R}^d$ as Fourier transforms of even positive measures of finite mass:

$$f(\xi) = \int e^{-\mathrm{i}x\cdot\xi}\, \mathrm{d}\mu(x).$$

Indeed,

$$f(\xi - \eta) = \int e^{-\mathrm{i}x\cdot\xi} e^{\mathrm{i}x\cdot\eta}\, \mathrm{d}\mu(x)$$

is a positive semidefinite kernel because it is the average over μ of the positive kernel $(\xi, \eta) \mapsto e^{-\mathrm{i}x\cdot\xi} e^{\mathrm{i}x\cdot\eta}$. Since the Gaussian e^{-x^2} is the Fourier transform of itself (modulo constants), it turns out that it is a positive definite function on $\mathbb{R}$, whence $\exp(-\|x\|^2)$ has the same property as a function on $\mathbb{R}^d$. Taking one step further, the function $x \mapsto \exp(-\|x\|^2)$ is positive definite on any Hilbert space.

With this preparation we are ready for a second characterization of metric subspaces of Hilbert space.

Theorem 2.4 (Schoenberg [84]) *A separable metric space (X, ρ) can be embedded isometrically into Hilbert space if and only if the kernel*

$$X \times X \to (0, \infty); \ (x, y) \mapsto \exp(-\lambda^2 \rho(x, y)^2)$$

is positive semidefinite for all $\lambda \in \mathbb{R}$.

Proof Necessity follows from the positive definiteness of the Gaussian discussed above. (We also provide an elementary proof below; see Lemma 4.7 and the subsequent discussion). To prove sufficiency, we recall the Menger–Schoenberg characterization of isometric subspaces of Hilbert space. We have to derive, from the positivity assumption, the positivity of the matrix

$$[\rho(x_0, x_j)^2 + \rho(x_0, x_k)^2 - \rho(x_j, x_k)^2]_{j,k=1}^n.$$

Elementary algebra transforms this constraint into the requirement that

$$\sum_{j,k=0}^n \rho(x_j, x_k)^2 c_j c_k \le 0 \qquad \text{whenever } \sum_{j=0}^n c_j = 0.$$

By expanding $\exp(-\lambda^2 \rho(x_j, x_k)^2)$ as a power series in λ^2, and invoking the positivity of the exponential kernel, we see that

$$0 \le -\lambda^2 \sum_{j,k=0}^n \rho(x_j, x_k)^2 c_j c_k + \frac{\lambda^4}{2} \sum_{j,k=0}^n \rho(x_j, x_k)^4 c_j c_k - \cdots$$

for all $\lambda > 0$. Hence the coefficient of $-\lambda^2$ is non-positive. □

The flexibility of the Fourier-transform approach is illustrated by the following application, also due to Schoenberg [84].

Corollary 2.5 *Let H be a Hilbert space with norm $\|\cdot\|$. For every $\delta \in (0, 1)$, the metric space $(H, \|\cdot\|^\delta)$ is isometric to a subspace of a Hilbert space.*

Proof Note first the identity

$$\xi^\alpha = c_\alpha \int_0^\infty (1 - e^{-s^2\xi^2}) s^{-1-\alpha} \, \mathrm{d}s \qquad (\xi > 0,\ 0 < \alpha < 2),$$

where c_α is a normalization constant. Consequently,

$$\|x - y\|^\alpha = c_\alpha \int_0^\infty (1 - e^{-s^2\|x-y\|^2}) s^{-1-\alpha} \, \mathrm{d}s.$$

Let $\delta = \alpha/2$. For points $x_0, x_1, \ldots, x_n$ in H and weights $c_0, c_1, \ldots, c_n$ satisfying

$$c_0 + c_1 + \cdots + c_n = 0,$$

it holds that

$$\sum_{j,k=0}^n \|x_j - x_k\|^{2\delta} c_j c_k = -c_\alpha \int_0^\infty \sum_{j,k=0}^n c_j c_k \, e^{-s^2\|x_j - x_k\|^2} s^{-1-\alpha} \, \mathrm{d}s \le 0,$$

and the proof is complete. □

Several similar consequences of the Fourier-transform approach are within reach. For instance, Schoenberg observed in the same article that if the L^p norm is raised to the power γ, where $0 < \gamma \leq p/2$ and $1 \leq p \leq 2$, then $L^p(0, 1)$ is isometrically embeddable into Hilbert space.

2.4 Altering Euclidean Distance

By specializing the theme of the previous section to Euclidean space, Schoenberg and von Neumann discovered an arsenal of powerful tools from harmonic analysis that were able to settle the question of whether Euclidean space equipped with the altered distance $\phi(\|x - y\|)$ may be isometrically embedded into Hilbert space [83, 94]. The key ingredients are characterizations of Laplace and Fourier transforms of positive measures, that is, Bernstein's completely monotone functions [14] and Bochner's positive definite functions [18].

Here we present some highlights of the Schoenberg–von Neumann framework. First, we focus on an auxiliary class of distance transforms. A real continuous function ϕ is called *positive definite in Euclidean space* $\mathbb{R}^d$ if the kernel

$$(x, y) \mapsto \phi(\|x - y\|)$$

is positive semidefinite. Bochner's theorem and the rotation-invariance of this kernel prove that such a function ϕ is characterized by the representation

$$\phi(t) = \int_0^\infty \Omega_d(tu) \,\mathrm{d}\mu(u),$$

where μ is a positive measure and

$$\Omega_d(\|x\|) = \int_{\|\xi\|=1} e^{\mathrm{i}x\cdot\xi} \,\mathrm{d}\sigma(\xi),$$

with σ the normalized area measure on the unit sphere in $\mathbb{R}^d$; see [83, Theorem 1]. By letting d tend to infinity, one finds that positive definite functions on infinite-dimensional Hilbert space are precisely of the form

$$\phi(t) = \int_0^\infty e^{-t^2u^2} \,\mathrm{d}\mu(u),$$

with μ a positive measure on the semi-axis. Notice that positive definite functions in $\mathbb{R}^d$ are not necessarily differentiable more than $(d - 1)/2$ times, while those which are positive definite in Hilbert space are smooth and even complex analytic in the sector $|\arg t| < \pi/4$.

The class of functions f which are continuous on $\mathbb{R}_+ := [0, \infty)$, smooth on the open semi-axis $(0, \infty)$, and such that

$$(-1)^n f^{(n)}(t) \geq 0 \qquad \text{for all } t > 0$$

was studied by Bernstein, who proved that they coincide with Laplace transforms of positive measures on $\mathbb{R}_+$:

$$f(t) = \int_0^\infty e^{-tu} \, \mathrm{d}\mu(u). \tag{2.1}$$

Such functions are called *completely monotonic* and have proved highly relevant for probability theory and approximation theory; see [14] for the foundational reference. Thus we have obtained a valuable equivalence.

Theorem 2.6 (Schoenberg) *A function f is completely monotone if and only if $t \mapsto f(t^2)$ is positive definite on Hilbert space.*

The direct consequences of this apparently innocent observation are quite deep. For example, the isometric-embedding question for altered Euclidean distances is completely answered via this route. The following results are from [83] and [94].

Theorem 2.7 (Schoenberg–von Neumann) *Let H be a separable Hilbert space with norm $\|\cdot\|$.*

(1) *For any integers $n \geq d > 1$, the metric space $(\mathbb{R}^d, \phi(\|\cdot\|))$ may be isometrically embedded into $(\mathbb{R}^n, \|\cdot\|)$ if and only if $\phi(t) = ct$ for some $c > 0$.*
(2) *The metric space $(\mathbb{R}^d, \phi(\|\cdot\|))$ may be isometrically embedded into H if and only if*

$$\phi(t)^2 = \int_0^\infty \frac{1 - \Omega_d(tu)}{u^2} \, \mathrm{d}\mu(u),$$

where μ is a positive measure on the semi-axis such that

$$\int_1^\infty \frac{1}{u^2} \, \mathrm{d}\mu(u) < \infty.$$

(3) *The metric space $(H, \phi(\|\cdot\|))$ may be isometrically embedded into H if and only if*

$$\phi(t)^2 = \int_0^\infty \frac{1 - e^{-t^2 u}}{u} \, \mathrm{d}\mu(u),$$

where μ is a positive measure on the semi-axis such that

$$\int_1^\infty \frac{1}{u} \, \mathrm{d}\mu(u) < \infty.$$

In von Neumann and Schoenberg's article [94], special attention is paid to the case of embedding a modified distance on the line into Hilbert space. This amounts to characterizing all *screw lines* in a Hilbert space H: the continuous functions

$$f : \mathbb{R} \to H;\ t \mapsto f_t$$

with the translation-invariance property

$$\|f_s - f_t\| = \|f_{s+r} - f_{t+r}\| \qquad \text{for all } s, r, t \in \mathbb{R}.$$

In this case, the gauge function ϕ is such that $\phi(t-s) = \|f_s - f_t\|$ and $t \mapsto f_t$ provides the isometric embedding of $(\mathbb{R}, \phi(|\cdot|))$ into H. Von Neumann seized the opportunity to use Stone's theorem on one-parameter unitary groups, together with the spectral decomposition of their unbounded self-adjoint generators, to produce a purely operator-theoretic proof of the following result.

Corollary 2.8 *The metric space* $(\mathbb{R}, \phi(|\cdot|))$ *isometrically embeds into Hilbert space if and only if*

$$\phi(t)^2 = \int_0^\infty \frac{\sin^2(tu)}{u^2}\,\mathrm{d}\mu(u) \qquad (t \in \mathbb{R}),$$

where μ *is a positive measure on* $\mathbb{R}_+$ *satisfying*

$$\int_1^\infty \frac{1}{u^2}\,\mathrm{d}\mu(u) < \infty.$$

Moreover, in the conditions of the corollary, the space $(\mathbb{R}, \phi(|\cdot|))$ embeds isometrically into $\mathbb{R}^d$ if and only if the measure μ consists of finitely many point masses, whose number is roughly $d/2$; see [94, Theorem 2] for the precise statement. To give a simple example, consider the function

$$\phi : \mathbb{R} \to \mathbb{R}_+;\ t \mapsto \sqrt{t^2 + \sin^2 t}.$$

This is indeed a screw function, because

$$\begin{aligned}\phi(t-s)^2 &= (t-s)^2 + \sin^2(t-s)\\ &= (t-s)^2 + \frac{1}{4}\big(\cos(2t) - \cos(2s)\big)^2 + \frac{1}{4}\big(\sin(2t) - \cos(2s)\big)^2.\end{aligned}$$

Note that a screw line is periodic if and only if it is not injective. Furthermore, one may identify screw lines with period $\tau > 0$ by the geometry of the support of the representing measure: this support must be contained in the lattice $(\pi/\tau)\mathbb{Z}_+$, where $\mathbb{Z}_+ := \mathbb{Z} \cap \mathbb{R}_+ = \{0, 1, 2, \dots\}$. Consequently, all periodic screw lines in Hilbert space have a gauge function ϕ such that

$$\phi(t)^2 = \sum_{k=1}^{\infty} c_k \sin^2(k\pi t/\tau), \tag{2.2}$$

where $c_k \geq 0$ and $\sum_{k=1}^{\infty} c_k < \infty$; see [94, Theorem 5].

2.5 *Positive Definite Functions on Homogeneous Spaces*

Having resolved the question of isometrically embedding Euclidean space into Hilbert space, a natural desire was to extend the analysis to other special manifolds with symmetry. This was done almost simultaneously by Schoenberg on spheres [86] and by Bochner on compact homogeneous spaces [19].

Let X be a compact space endowed with a transitive action of a group G and an invariant measure. We seek G-invariant distance functions, and particularly those which identify X with a subspace of a Hilbert space. To simplify terminology, we call the latter *Hilbert distances*.

The first observation of Bochner is that a G-invariant symmetric kernel $f : X \times X \to \mathbb{R}$ satisfies the Hilbert-space embeddability condition,

$$\sum_{k=0}^{n} c_k = 0 \quad \Longrightarrow \quad \sum_{j,k=0}^{n} f(x_j, x_k) c_j c_k \geq 0,$$

for all choices of weights c_j and points $x_j \in X$, if and only if f is of the form

$$f(x, y) = h(x, y) - h(x_0, x_0) \qquad (x, y \in X),$$

where h is a G-invariant positive definite kernel and x_0 is a point of X. One implication is clear. For the other, we start with a G-invariant function f subject to the above constraint and prove, using G-invariance and integration over X, the existence of a constant c such that $h(x, y) = f(x, y) + c$ is a positive semidefinite kernel. This gives the following result.

Theorem 2.9 (Bochner [19]) *Let X be a compact homogeneous space. A continuous invariant function ρ on $X \times X$ is a Hilbert distance if and only if there exists a continuous, real-valued, invariant, positive definite kernel h on X and a point $x_0 \in X$, such that*

$$\rho(x, y) = \sqrt{h(x_0, x_0) - h(x, y)} \qquad (x, y \in X).$$

Privileged orthonormal bases of G-invariant functions, in the L^2 space associated with the invariant measure, provide a canonical decompositions of positive definite kernels. These generalized spherical harmonics were already studied by Cartan, Weyl, and von Neumann; see, for instance [96]. We elaborate on two important particular cases.

Let $X = \mathbb{T} = \{e^{i\theta} : \theta \in \mathbb{R}\}$ be the unit torus, endowed with the invariant arc-length measure. A continuous positive definite function $h : \mathbb{T} \times \mathbb{T} \to \mathbb{R}$ admits a Fourier decomposition

$$h(e^{ix}, e^{iy}) = \sum_{j,k\in\mathbb{Z}} a_{jk} e^{ijx} e^{-iky}.$$

If h is further required to be rotation invariant, we find that

$$h(e^{ix}, e^{iy}) = \sum_{k\in\mathbb{Z}} a_k e^{ik(x-y)},$$

where $a_k \geq 0$ for all $k \in \mathbb{Z}$ and $a_k = a_{-k}$ because h takes real values. Moreover, the series is Abel summable: $\sum_{k=0}^{\infty} a_k = h(1,1) < \infty$. Therefore, a rotation-invariant Hilbert distance ρ on the torus has the expression (after taking its square):

$$\begin{aligned}\rho(e^{ix}, e^{iy})^2 = h(1,1) - h(e^{ix}, e^{iy}) &= \sum_{k=1}^{\infty} a_k (2 - e^{ik(x-y)} - e^{-ik(x-y)}) \\ &= 2\sum_{k=1}^{\infty} a_k (1 - \cos k(x-y)) \\ &= 4\sum_{k=1}^{\infty} a_k \sin^2(k(x-y)/2).\end{aligned}$$

These are the periodic screw lines (2.2) already investigated by von Neumann and Schoenberg.

As a second example, we follow Bochner in examining a separable, compact group G. A real-valued, continuous, positive definite, and G-invariant kernel h admits the decomposition

$$h(x, y) = \sum_{k\in\mathbb{Z}} c_k \chi_k(yx^{-1}),$$

where $c_k \geq 0$ for all $k \in \mathbb{Z}$, $\sum_{k\in Z} c_k < \infty$ and χ_k denote the characters of irreducible representations of G. In conclusion, an invariant Hilbert distance ρ on G

is characterized by the formula

$$\rho(x, y)^2 = \sum_{k\in\mathbb{Z}} c_k\Big(1 - \frac{\chi_k(yx^{-1}) + \chi_k(xy^{-1})}{2\chi_k(1)}\Big),$$

where $c_k \geq 0$ and $\sum_{k\in\mathbb{Z}} c_k < \infty$.

For details and an analysis of similar decompositions on more general homogeneous spaces, we refer the reader to [19].

The above analysis of positive definite functions on homogeneous spaces was carried out separately by Schoenberg in [86]. First, he remarks that a continuous, real-valued, rotationally invariant, and positive definite kernel f on the sphere S^{d-1} has a distinguished Fourier-series decomposition with non-negative coefficients. Specifically,

$$f(\cos\theta) = \sum_{k=0}^{\infty} c_k P_k^{(\lambda)}(\cos\theta) \tag{2.3}$$

where $\lambda = (d-2)/2$, $P_k^{(\lambda)}$ are the ultraspherical orthogonal polynomials, $c_k \geq 0$ for all $k \geq 0$ and $\sum_{k=0}^{\infty} c_k < \infty$. This decomposition is in accord with Bochner's general framework, with the difference lying in Schoenberg's elementary proof, based on induction on dimension. As with all our formulas concerning the sphere, θ represents the geodesic distance (arc length along a great circle) between two points.

To convince the reader that expressions in the cosine of the geodesic distance are positive definite, let us consider points $x_1, \ldots, x_n \in S^{d-1}$. The Gram matrix with entries

$$\langle x_j, x_k\rangle = \cos\theta(x_j, x_k)$$

is obviously positive semidefinite, with constant diagonal elements equal to 1. According to the Schur product theorem [90], all functions of the form $\cos^k\theta$, where k is a non-negative integer, are therefore positive definite on the sphere.

At this stage, Schoenberg makes a leap forward and studies invariant positive definite kernels on S^∞, that is, functions $f(\cos\theta)$ which admit representations as above for all $d \geq 2$. His conclusion is remarkable in its simplicity.

Theorem 2.10 (Schoenberg [86]) *A real-valued function $f(\cos\theta)$ is positive definite on all spheres, independent of their dimension, if and only if*

$$f(\cos\theta) = \sum_{k=0}^{\infty} c_k \cos^k\theta, \tag{2.4}$$

where $c_k \geq 0$ for all $k \geq 0$ and $\sum_{k=0}^{\infty} c_k < \infty$.

This provides a return to the dominant theme, of isometric embedding into Hilbert space.

Corollary 2.11 *The function $\rho(\theta)$ is a Hilbert distance on S^∞ if and only if*

$$\rho(\theta)^2 = \sum_{k=0}^{\infty} c_k (1 - \cos^k \theta),$$

where $c_k \geq 0$ for all $k \geq 0$ and $\sum_{k=0}^{\infty} c_k < \infty$.

However, there is much more to derive from Schoenberg's theorem, once it is freed from the spherical context.

Theorem 2.12 (Schoenberg [86]) *Let $f : [-1, 1] \to \mathbb{R}$ be a continuous function. If the matrix $[f(a_{jk})]_{j,k=1}^n$ is positive semidefinite for all $n \geq 1$ and all positive semidefinite matrices $[a_{jk}]_{j,k=1}^n$ with entries in $[-1, 1]$, then, and only then,*

$$f(x) = \sum_{k=0}^{\infty} c_k x^k \qquad (x \in [-1, 1]),$$

where $c_k \geq 0$ for all $k \geq 0$ and $\sum_{k=0}^{\infty} c_k < \infty$.

Proof One implication follows from the Schur product theorem [90], which says that if the $n \times n$ matrices A and B are positive semidefinite, then so is their entrywise product $A \circ B := [a_{jk} b_{jk}]_{j,k=1}^n$. Indeed, inductively setting $B = A^{\circ k} = A \circ \cdots \circ A$, the k-fold entrywise power shows that every monomial x^k preserves positivity when applied entrywise. That the same property holds for functions $f(x) = \sum_{k \geq 0} c_k x^k$, with all $c_k \geq 0$, now follows from the fact that the set of positive semidefinite $n \times n$ matrices forms a closed convex cone, for all $n \geq 1$.

For the non-trivial, reverse implication we restrict the test matrices to those with leading diagonal terms all equal to 1. By interpreting such a matrix A as a Gram matrix, we identify n points on the sphere $x_1, \ldots, x_n \in S^{n-1}$ satisfying

$$a_{jk} = \langle x_j, x_k \rangle = \cos \theta(x_j, x_k) \quad (1 \leq j, k \leq n).$$

Then we infer from Schoenberg's theorem that f admits a uniformly convergent Taylor series with non-negative coefficients. □

We conclude this section by mentioning some recent avenues of research that start from Bochner's theorem (and its generalization in 1940, by Weil, Povzner, and Raikov, to all locally compact abelian groups) and Schoenberg's classification of positive definite functions on spheres. On the theoretical side, there has been a profusion of recent mathematical activity on classifying positive definite functions (and strictly positive definite functions) in numerous settings, mostly related to spheres [4, 5, 23, 98–100], two-point homogeneous spaces[2] [2, 3, 21], locally compact abelian groups and homogeneous spaces [28, 41], and products of these [11, 13, 40, 42–44].

Moreover, this line of work directly impacts applied fields. For instance, in climate science and geospatial statistics, one uses positive definite kernels and Schoenberg's results (and their sequels) to study trends in climate behavior on the Earth, since it can be modelled by a sphere, and positive definite functions on $S^2 \times \mathbb{R}$ characterize space-time covariance functions on it. See [39, 65, 71] for more details on these applications. Other applied fields include genomics and finance, through high-dimensional covariance estimation. We elaborate on this in the second part of the survey: see [10] or the full version [9, Chapter 7].

There are several other applications of Schoenberg's work on positive definite functions on spheres (his paper [86] has more than 160 citations) and we mention here just a few of them. Schoenberg's results were used by Musin [66] to compute the kissing number in four dimensions, by an extension of Delsarte's linear-programming method. Moreover, the results also apply to obtain new bounds on spherical codes [67], with further applications to sphere packing [25]. There are also applications to approximating functions and interpolating data on spheres, pseudodifferential equations with radial basis functions, and Gaussian random fields.

Remark 2.13 Another modern-day use of Schoenberg's results in [86] is in Machine Learning; see [91, 92], for example. Given a real inner-product space H and a function $f : \mathbb{R} \to \mathbb{R}$, an alternative notion of f being *positive definite* is as follows: for any finite set of vectors $x_1, \ldots, x_n \in H$, the matrix

$$[f(\langle x_j, x_k \rangle)]_{j,k=1}^n$$

is positive semidefinite. This is in contrast to the notion promoted by Bochner, Weil, Schoenberg, Pólya, and others, which concerns positivity of the matrix with entries $f(\langle x_j - x_k, x_j - x_k \rangle^{1/2})$. It turns out that every positive definite kernel on H, given by

$$(x, y) \mapsto f(\langle x, y \rangle)$$

for a function f which is positive definite in this alternate sense, gives rise to a reproducing-kernel Hilbert space, which is a central concept in Machine Learning. We restrict ourselves here to mentioning that, in this setting, it is desirable for the kernel to be strictly positive definite; see [68] for further clarification and theoretical results along these lines.

[2] Recall [95] that a metric space (X, ρ) is *n-point homogeneous* if, given finite sets $X_1, X_2 \subset X$ of equal size no more than n, every isometry from X_1 to X_2 extends to a self-isometry of X. This property was first considered by Birkhoff [15], and of course differs from the more common usage of the terminology of a homogeneous space G/H, whose study by Bochner was mentioned above.

2.6 Connections to Harmonic Analysis

Positivity and sharp continuity bounds for linear transformations between specific normed function spaces go hand in hand, especially when focusing on the kernels of integral transforms. The end of 1950s marked a fortunate condensation of observations, leading to a quasi-complete classification of preservers of positive or bounded convolution transforms acting on spaces of functions on locally compact abelian groups. In particular, these results can be interpreted as Schoenberg-type theorems for Toeplitz matrices or Toeplitz kernels. We briefly recount the main developments.

A groundbreaking theorem of the 1930s attributed to Wiener and Levy asserts that the pointwise inverse of a non-vanishing Fourier series with coefficients in L^1 exhibits the same summability behavior of the coefficient sequence. To be more precise, if ϕ is never zero and has the representation

$$\phi(\theta) = \sum_{n=-\infty}^{\infty} c_n e^{\mathrm{i}n\theta}, \quad \text{where} \sum_{n=-\infty}^{\infty} |c_n| < \infty,$$

then its reciprocal has a representation of the same form:

$$(1/\phi)(\theta) = \sum_{n=-\infty}^{\infty} d_n e^{\mathrm{i}n\theta}, \quad \text{where} \sum_{n=-\infty}^{\infty} |d_n| < \infty.$$

It was Gelfand [38] who in 1941 cast this permanence phenomenon in the general framework of commutative Banach algebras. Gelfand's theory applied to the Wiener algebra $W := \widehat{L^1(\mathbb{Z})}$ of Fourier transforms of L^1 functions on the dual of the unit torus proves the following theorem.

Theorem 2.14 (Gelfand [38]) *Let $\phi \in W$ and let $f(z)$ be an analytic function defined in a neighborhood of $\phi(\mathbb{T})$. Then $f(\phi) \in W$.*

The natural inverse question of deriving smoothness properties of inner transformations of Lebesgue spaces of Fourier transforms was tackled almost simultaneously by several analysts. For example, Rudin proved in 1956 [76] that a coefficient-wise transformation $c_n \mapsto f(c_n)$ mapping the space $\widehat{L^1(\mathbb{T})}$ into itself implies the analyticity of f in a neighborhood of zero. In a similar vein, Rudin and Kahane proved in 1958 [53] that a coefficient-wise transformation $c_n \mapsto f(c_n)$ which preserves the space of Fourier transforms $\widehat{M(\mathbb{T})}$ of finite measures on the torus implies that f is an entire function. In the same year, Kahane [52] showed that no quasi-analytic function (in the sense of Denjoy–Carleman) preserves the space $\widehat{L^1(\mathbb{Z})}$ and Katznelson [56] refined an inverse to Gelfand's theorem above, by showing the semi-local analyticity of transformers of elements of $\widehat{L^1(\mathbb{Z})}$ subject to some support conditions.

Soon after, the complete picture emerged in full clarity. It was unveiled by Helson, Kahane, Katznelson, and Rudin in an Acta Mathematica article [48]. Given a function f defined on a subset E of the complex plane, we say that f operates on the function algebra A, if $f(\phi) \in A$ for every $\phi \in A$ with range contained in E. The following metatheorem is proved in the cited article.

Theorem 2.15 (Helson–Kahane–Katznelson–Rudin [48]) *Let G be a locally compact abelian group and let Γ denote its dual, and suppose both are endowed with their respective Haar measures. Let $f : [-1, 1] \to \mathbb{C}$ be a function satisfying $f(0) = 0$.*

(1) *If Γ is discrete and f operates on $\widehat{L^1(G)}$, then f is analytic in some neighborhood of the origin.*
(2) *If Γ is not discrete and f operates on $\widehat{L^1(G)}$, then f is analytic in $[-1, 1]$.*
(3) *If Γ is not compact and f operates on $\widehat{M(G)}$, then f can be extended to an entire function.*

Rudin refined the above results to apply in the case of various L^p norms [78, 79], by stressing the lack of continuity assumption for the transformer f in all results (similar in nature to the statements in the above theorem). From Rudin's work we extract a highly relevant observation, *à la* Schoenberg's theorem, aligned to the spirit of the present survey.

Theorem 2.16 (Rudin [77]) *Suppose $f : (-1, 1) \to \mathbb{R}$ maps every positive semidefinite Toeplitz kernel with elements in $(-1, 1)$ into a positive semidefinite kernel:*

$$[a_{j-k}]_{j,k=-\infty}^{\infty} \geq 0 \qquad \Longrightarrow \qquad [f(a_{j-k})]_{j,k=-\infty}^{\infty} \geq 0.$$

Then f is absolutely monotonic, that is analytic on $(-1, 1)$ with a Taylor series having non-negative coefficients:

$$f(x) = \sum_{n=0}^{\infty} c_k x^k, \qquad \text{where } c_k \geq 0 \text{ for all } k \geq 0.$$

The converse is obviously true by the Schur product theorem. The elementary proof, quite independent of the derivation of the metatheorem stated above, is contained in [77]. Notice again the lack of a continuity assumption in the hypotheses.

In fact, Rudin proves more, by restricting the test domain of positive semidefinite Toeplitz kernels to the two-parameter family

$$a_n = \alpha + \beta \cos(n\theta) \qquad (n \in \mathbb{Z}) \tag{2.5}$$

with θ fixed so that θ/π is irrational and $\alpha, \beta \geq 0$ such that $\alpha + \beta < 1$. Rudin's proof commences with a mollifier argument to deduce the continuity of the transformer, then uses a development in spherical harmonics very similar to the original argument of Schoenberg. We will resume this topic in Sect. 3.3, setting it in a wider context.

With the advances in abstract duality theory for locally convex spaces, it is not surprising that proofs of Schoenberg-type theorems should be accessible with the aid of such versatile tools. We will confine ourselves here to mentioning one pertinent convexity-theoretic proof of Schoenberg's theorem, due to Christensen and Ressel [24].

Skipping freely over the details, the main observation of these two authors is that the multiplicatively closed convex cone of positivity preservers of positive semidefinite matrices of any size, with entries in $[-1, 1]$, is closed in the product topology of $\mathbb{R}^{[-1,1]}$, with a compact base K defined by the normalization $f(1) = 1$. The set of extreme points of K is readily seen to be closed, and an elementary argument identifies it as the set of all monomials x^n, where $n \geq 0$, plus the characteristic functions $\chi_1 \pm \chi_{-1}$. An application of Choquet's representation theorem now provides a proof of a generalization of Schoenberg's theorem, by removing the continuity assumption in the statement.

3 Entrywise Functions Preserving Positivity in All Dimensions

3.1 History

With the above history to place the present survey in context, we move to its dominant theme: entrywise positivity preservers. In analysis and in applications in the broader mathematical sciences, one is familiar with applying functions to the spectrum of diagonalizable matrices: $A = UDU^*$ then $f(A) = Uf(D)U^*$. More formally, one uses the Riesz–Dunford holomorphic functional calculus to define $f(A)$ for classes of matrices A and functions f.

Our focus in this survey will be on the parallel philosophy of *entrywise calculus*. To differentiate this from the functional calculus, we use the notation $f[A]$.

Definition 3.1 Fix a domain $I \subset \mathbb{C}$ and integers $m, n \geq 1$. Let $\mathcal{P}_n(I)$ denote the set of $n \times n$ Hermitian positive semidefinite matrices with all entries in I.

A function $f : I \to \mathbb{C}$ acts *entrywise* on a matrix

$$A = [a_{jk}]_{1 \leq j \leq m,\ 1 \leq k \leq n} \in I^{m \times n}$$

by setting

$$f[A] := [f(a_{jk})]_{1 \leq j \leq m,\ 1 \leq k \leq n} \in \mathbb{C}^{m \times n}.$$

Below, we allow the dimensions m and n to vary, while keeping the uniform notation $f[-]$.

We also let $\mathbf{1}_{m\times n}$ denote the $m \times n$ matrix with each entry equal to one. Note that $\mathbf{1}_{n\times n} \in \mathcal{P}_n(\mathbb{R})$.

In this survey, we explore the following overarching question in several different settings.

Which functions preserve positive semidefiniteness when applied entrywise to a class of positive matrices?

This question was first asked by Pólya and Szegö in their well-known book [70]. The authors observed that Schur's product theorem, together with the fact that the positive matrices form a closed convex cone, has the following consequence: if $f(x)$ is any power series with non-negative Maclaurin coefficients that converges on a domain $I \subset \mathbb{R}$, then f preserves positivity (that is, preserves positive semidefiniteness) when applied entrywise to positive semidefinite matrices with entries in I. Pólya and Szegö then asked if there are any other functions that possess this property. As discussed above, Schoenberg's theorem 2.12 provides a definitive answer to their question (together with the improvements by Rudin or Christensen–Ressel to remove the continuity hypothesis). Thanks to Pólya and Szegö's observation, Schoenberg's result may be considered as a rather challenging converse to the Schur product theorem.

In a similar vein, Rudin [77] observed that if one moves to the complex setting, then the conjugation map also preserves positivity when applied entrywise to positive semidefinite complex matrices. Therefore the maps

$$z \mapsto z^j \overline{z}^k \qquad (j, k \geq 0)$$

preserve positivity when applied entrywise to complex matrices of all dimensions, again by the Schur product theorem. The same property is now satisfied by non-negative linear combinations of these functions. In [77], Rudin made this observation and conjectured, *à la* Pólya–Szegö, that these are all of the preservers. This was proved by Herz in 1963.

Theorem 3.2 (Herz [49]) *Let $D(0, 1)$ denote the open unit disc in $\mathbb{C}$, and suppose $f : D(0, 1) \to \mathbb{C}$. The entrywise map $f[-]$ preserves positivity on $\mathcal{P}_n\big(D(0, 1)\big)$ for all $n \geq 1$, if and only if*

$$f(z) = \sum_{j,k\geq 0} c_{jk} z^j \overline{z}^k \qquad \textit{for all } z \in D(0, 1),$$

where $c_{jk} \geq 0$ for all $j, k \geq 0$.

Akin to the above results by Schoenberg, Rudin, Christensen and Ressel, and Herz, we mention one more Schoenberg-type theorem, for matrices with positive entries. The following result again demonstrates the rigid principle that analyticity and absolute monotonicity follow from the preservation of positivity in all dimensions.

Theorem 3.3 (Vasudeva [93]) *Let $f : (0,\infty) \to \mathbb{R}$. Then $f[-]$ preserves positivity on $\mathcal{P}_n\big((0,\infty)\big)$ for all $n \geq 1$, if and only if $f(x) = \sum_{k=0}^{\infty} c_k x^k$ on $(0,\infty)$, where $c_k \geq 0$ for all $k \geq 0$.*

3.2 The Horn–Loewner Necessary Condition in Fixed Dimension

The previous section contains several variants of a "dimension-free" result: namely, the classification of entrywise maps that preserve positivity on test sets of matrices of all sizes. In the next section, we discuss a dimension-free result that parallels Rudin's work in [77], by approaching the problem via preservers of moment sequences for positive measures on the real line. In other words, we will work with Hankel instead of Toeplitz matrices.

In the later part of this survey, we focus on entrywise functions that preserve positivity when the test set consists of matrices of a fixed size. For both of these settings, the starting point is an important result found in the PhD thesis of Roger Horn, which he attributes to his advisor, Charles Loewner.

Theorem 3.4 ([50]) *Let $f : (0,\infty) \to \mathbb{R}$ be continuous. Fix a positive integer n and suppose $f[-]$ preserves positivity on $\mathcal{P}_n\big((0,\infty)\big)$. Then $f \in C^{n-3}(I)$,*

$$f^{(k)}(x) \geq 0 \qquad \text{whenever } x \in (0,\infty) \text{ and } 0 \leq k \leq n-3,$$

and $f^{(n-3)}$ is a convex non-decreasing function on $(0,\infty)$. Furthermore, if $f \in C^{n-1}\big((0,\infty)\big)$, then $f^{(k)}(x) \geq 0$ whenever $x \in (0,\infty)$ and $0 \leq k \leq n-1$.

This result and its variations are the focus of the present section.

Theorem 3.4 is remarkable for several reasons.

(1) Modulo variations, it remains to this day the only known criterion for a general entrywise function to preserve positivity in a fixed dimension. Later on, we will see more precise conclusions drawn when f is a polynomial or a power function, but for a general function there are essentially no other known results.
(2) While Theorem 3.4 is a fixed-dimension result, it can be used to prove some of the aforementioned dimension-free characterizations. For instance, if $f[-]$ preserves positivity on $\mathcal{P}_n\big((0,\infty)\big)$ for all $n \geq 1$, then, by Theorem 3.4, the function f is absolutely monotonic on $(0,\infty)$. A classical result of Bernstein on absolutely monotonic functions now implies that f is necessarily given by a power series with non-negative coefficients, which is precisely Vasudeva's Theorem 3.3.

    In the next section, we will outline an approach to prove a stronger version of Schoenberg's Theorem 2.12 (in the spirit of Theorem 2.16 by Rudin), starting from Theorem 3.3.

(3) Theorem 3.4 is also significant because there is a sense in which it is sharp. We elaborate on this when studying polynomial and power-function preservers; this is discussed in the second part of the survey: see [10] or [9, Chapters 4 and 6].

Remark 3.5 There are other, rather unexpected consequences of Theorem 3.4 as well. It was recently shown that the key determinant computation underlying Theorem 3.4 can be generalized to yield a new class of symmetric function identities for any formal power series. The only such identities previously known were for the case $f(x) = \frac{1-cx}{1-x}$. This is discussed in the second part of this survey [10] and the full version [9, Section 4.6].

We next explain the steps behind the proof of the Horn–Loewner Theorem 3.4. These also help in proving certain strengthenings of Theorem 3.4, which are mentioned below. In turn, these strengthenings additionally serve to clarify the nature of the Horn–Loewner necessary condition.

Proof of Theorem 3.4 The proof by Loewner is in two steps. First he assumes f to be smooth and shows the result by induction on n. The base case of $n = 1$ is immediate, and for the induction step one proceeds as follows. Fix $a > 0$, choose any vector $\mathbf{u} = (u_1, \dots, u_n)^T \in \mathbb{R}^n$ with distinct coordinates, and define

$$\Delta(t) := \det[f(a + tu_j u_k)]_{j,k=1}^n = \det f[a\mathbf{1}_{n\times n} + \mathbf{u}\mathbf{u}^T] \qquad (0 < t \ll 1).$$

Then Loewner shows that

$$\begin{aligned} \Delta(0) = \Delta'(0) = \cdots = \Delta^{\binom{n}{2}-1}(0) = 0, \\ \Delta^{\binom{n}{2}}(0) = cf(a)f'(a)\cdots f^{(n-1)}(a) \qquad \text{for some } c > 0. \end{aligned} \tag{3.1}$$

(See Remark 3.5 above.)

Returning to the proof of Theorem 3.4 for smooth functions: apply the above treatment not to f but to $g_\tau(x) := f(x) + \tau x^n$, where $\tau > 0$. By the Schur product theorem, g_τ satisfies the hypotheses, whence $\Delta(t)/t^{\binom{n}{2}} \geq 0$ for $t > 0$. Taking $t \to 0^+$, by L'Hôpital's rule we obtain

$$g_\tau(a)g_\tau'(a)\cdots g_\tau^{(n-1)}(a) \geq 0, \qquad \text{for all } \tau > 0.$$

Finally, the induction hypothesis implies that $f, f', \dots, f^{(n-2)}$ are non-negative at a, whence $g_\tau(a), \dots, g_\tau^{(n-2)}(a) > 0$. It follows that $g_\tau^{(n-1)}(a) \geq 0$ for all $\tau > 0$, and hence, $f^{(n-1)}(a) \geq 0$, as desired.

Remark 3.6 The above argument is amenable to proving more refined results. For example, it can be used to prove the positivity of the first n non-zero derivatives of a smooth preserver f; see Theorem 3.10.

The second step of Loewner's proof begins by using mollifiers. Suppose f is continuous; approximate it by a mollified family $f_\delta \to f$ as $\delta \to 0^+$. Thus f_δ is smooth and its first n derivatives are non-negative on $(0, \infty)$. By the mean-value theorem for divided differences, this implies that the divided differences of each f_δ, of orders up to $n-1$ are non-negative. Since f is continuous, the same holds for f.

Now one invokes a rather remarkable result by Boas and Widder [17], which can be viewed as a converse to the mean-value theorem for divided differences. It asserts that given an integer $k \geq 2$ and an open interval $I \subset \mathbb{R}$, if all kth order "equi-spaced" forward differences (whence divided differences) of a continuous function $f : I \to \mathbb{R}$ are non-negative on I, then f is $k-2$ times differentiable on I; moreover, $f^{(k-2)}$ is continuous and convex on I, with non-decreasing left- and right-hand derivatives. Applying this result for each $2 \leq k \leq n-1$ concludes the proof of Theorem 3.4. □

Note that this proof only uses matrices of the form $a\mathbf{1}_{n\times n} + t\mathbf{u}\mathbf{u}^T$, and the arguments are all local. Thus it is unsurprising that strengthened versions of the Horn–Loewner theorem can be found in the literature; see [7, 47], for example. We present here the stronger of these variants.

Theorem 3.7 (See [7, Section 3]) *Suppose* $0 < \rho \leq \infty$, $I = (0, \rho)$, *and* $f : I \to \mathbb{R}$. *Fix* $u_0 \in (0, 1)$ *and an integer* $n \geq 1$, *and define* $\mathbf{u} := (1, u_0, \dots, u_0^{n-1})^T$. *Suppose* $f[A] \in \mathcal{P}_2(\mathbb{R})$ *for all* $A \in \mathcal{P}_2(I)$, *and also that* $f[A] \in \mathcal{P}_n(\mathbb{R})$ *for all Hankel matrices* $A = a\mathbf{1}_{n\times n} + t\mathbf{u}\mathbf{u}^T$, *with* $a, t \geq 0$ *such that* $a + t \in I$. *Then the conclusions of Theorem 3.4 hold.*

Beyond the above strengthenings, the notable feature here is that the continuity hypothesis has been removed, akin to the Rudin and Christensen–Ressel results. We reproduce here an elegant argument to show continuity; this can be found in Vasudeva's paper [93], and uses only the test set $\mathcal{P}_2(I)$. By considering $f[A]$ for $A = \begin{bmatrix} a & b \\ b & a \end{bmatrix}$ with $0 < b < a < \rho$, it follows that f is non-negative and non-decreasing on I. One also shows that f is either identically zero or never zero on I. In the latter case, considering $f[A]$ for $A = \begin{bmatrix} a & \sqrt{ab} \\ \sqrt{ab} & b \end{bmatrix} \in \mathcal{P}_2(I)$ shows that f is *multiplicatively mid-convex*: the function

$$g(y) := \log f(e^y) \qquad (y < \log \rho)$$

is midpoint convex and locally bounded on the interval $\log I$. Now the following classical result [74, Theorem 71.C] shows that g is continuous on $\log I$, so f is continuous on I.

Proposition 3.8 *Let* U *be a convex open set in a real normed linear space. If* $g : U \to \mathbb{R}$ *is midpoint convex on* U *and bounded above in an open neighborhood of a single point in* U, *then* g *is continuous, so convex, on* U.

We now move to variants of the Horn–Loewner result. Notice that Theorems 3.4 and 3.7 are results for arbitrary positivity preservers $f(x)$. When more is known about f, such as smoothness or even real analyticity, stronger conclusions can be drawn from smaller test sets of matrices. A recent variant is the following lemma, shown by evaluating $f[-]$ at matrices $(tu_j u_k)_{j,k=1}^n$ and using the invertibility of "generic" generalized Vandermonde matrices.

Lemma 3.9 (Belton–Guillot–Khare–Putinar [6] and Khare–Tao [58]) *Let $n \geq 1$ and $0 < \rho \leq \infty$. Suppose $f(x) = \sum_{k \geq 0} c_k x^k$ is a convergent power series on $I = [0, \rho)$ that is positivity preserving entrywise on rank-one matrices in $\mathcal{P}_n(I)$. Further assume that $c_{m'} < 0$ for some m'.*

(1) *If $\rho < \infty$, then we have $c_m > 0$ for at least n values of $m < m'$. (In particular, the first n non-zero Maclaurin coefficients of f, if they exist, must be positive.)*
(2) *If instead $\rho = \infty$, then we have $c_m > 0$ for at least n values of $m < m'$ and at least n values of $m > m'$. (In particular, if f is a polynomial, then the first n non-zero coefficients and the last n non-zero coefficients of f, if they exist, are all positive.)*

Notice that this lemma (a) talks about the derivatives of f at 0 and not in $(0, \rho)$; and moreover, (b) considers not the first few derivatives, but the first few *non-zero* derivatives. Thus, it is morally different from the preceding two theorems, and one naturally seeks a common unification of these three results. This was recently achieved.

Theorem 3.10 (Khare [57]) *Let $0 \leq a < \infty, \epsilon \in (0, \infty)$, $I = [a, a + \epsilon)$, and let $f : I \to \mathbb{R}$ be smooth. Fix integers $n \geq 1$ and $0 \leq p \leq q \leq n$, with $p = 0$ if $a = 0$, and such that $f(x)$ has $q - p$ non-zero derivatives at $x = a$ of order at least p. Now let*

$$m_0 := 0, \quad \ldots \quad m_{p-1} := p - 1;$$

suppose further that

$$p \leq m_p < m_{p+1} < \cdots < m_{q-1}$$

are the lowest orders (above p) of the first $q - p$ non-zero derivatives of $f(x)$ at $x = a$.

Also fix distinct scalars $u_1, \ldots, u_n \in (0, 1)$, and let $\mathbf{u} := (u_1, \ldots, u_n)^T$. If $f[a\mathbf{1}_{n\times n} + t\mathbf{u}\mathbf{u}^T] \in \mathcal{P}_n(\mathbb{R})$ for all $t \in [0, \epsilon)$, then the derivative $f^{(k)}(a)$ is non-negative whenever $0 \leq k \leq m_{q-1}$.

Notice that varying p allows one to control the number of initial derivatives versus the number of subsequent non-zero derivatives of smallest order. In particular, if $p = q = n$, then the result implies the "stronger" Horn–Loewner Theorem 3.7 (and so Theorem 3.4) pointwise at every $a > 0$. At the other extreme is the special case of $p = 0$ (at any $a \geq 0$), which strengthens the conclusions of Theorems 3.4 and 3.7 for smooth functions.

Corollary 3.11 *Suppose a, ϵ, I, f, n, and $\mathbf{u}$ are as in Theorem 3.10. If $f[a\mathbf{1}_{n\times n} + t\mathbf{u}\mathbf{u}^T] \in \mathcal{P}_n(\mathbb{R})$ for all $t \in [0, \epsilon)$, then the first n non-zero derivatives of $f(x)$ at $x = a$ are positive.*

Remark 3.12 Theorem 3.10 further clarifies the nature of the Horn–Loewner result and its proof. The reduction from arbitrary functions, to continuous functions, to smooth functions, requires an open domain $(0, \rho)$, in order to use mollifiers, for example. However, the result for smooth functions actually holds pointwise, as shown by Theorem 3.10.

The proof of Theorem 3.10 combines novel arguments together with the previously mentioned techniques of Loewner. The refinement of the determinant computations (3.1) is of particular note; see the second part of this survey ([10] or [9, Section 4.6]).

3.3 Schoenberg Redux: Moment Sequences and Hankel Matrices

In this section, we outline another approach to proving Schoenberg's Theorem 2.12, which yields a stronger version parallel to the strengthening by Rudin of Theorem 2.16. The present section reveals connections between positivity preservers, totally non-negative Hankel matrices, moment sequences of positive measures on the real line, and also a connection to semi-algebraic geometry.

We begin with Rudin's Theorem 2.16 and the family (2.5). Notice that the positive definite sequences in (2.5) give rise to the Toeplitz matrices $A(n, \alpha, \beta, \theta)$ with (j, k) entry equal to $\alpha + \beta \cos\big((j-k)\theta\big)$. From the elementary identity

$$\cos(p-q) = \cos p \cos q + \sin p \sin q \qquad (p, q \in \mathbb{R}),$$

it follows that these Toeplitz matrices have rank at most three:

$$A(n, \alpha, \beta, \theta) = \alpha \mathbf{1}_{n\times n} + \beta \mathbf{u}\mathbf{u}^T + \beta \mathbf{v}\mathbf{v}^T, \tag{3.2}$$

where

$$\mathbf{u} := \big(\cos\theta, \cos(2\theta), \ldots, \cos(n\theta)\big)^T \text{ and } \mathbf{v} := \big(\sin\theta, \sin(2\theta), \ldots, \sin(n\theta)\big)^T.$$

In particular, Rudin's work (see Theorem 2.16 and the subsequent discussion) implies the following result.

Proposition 3.13 *Let $\theta \in \mathbb{R}$ such that θ/π is irrational. An entrywise map $f : \mathbb{R} \to \mathbb{R}$ preserves positivity on the set of Toeplitz matrices*

$$\{A(n, \alpha, \beta, \theta) : n \geq 1,\ \alpha, \beta > 0\}$$

if and only if $f(x) = \sum_{k=0}^{\infty} c_k x^k$ is a convergent power series on $\mathbb{R}$, with $c_k \geq 0$ for all $k \geq 0$.

Thus, one can significantly reduce the set of test matrices.

Proof Given $0 < \rho < \infty$, let the restriction $f_\rho := f|_{(-\rho,\rho)}$. Observe from the discussion following Theorem 2.16 that Rudin's work explicitly shows the result for f_1, whence for any f_ρ by a change of variables. Thus,

$$f_\rho(x) = \sum_{k=0}^{\infty} c_{k,\rho} x^k, \qquad c_{k,\rho} \geq 0 \text{ for all } k \geq 0 \text{ and } \rho > 0.$$

Given $0 < \rho < \rho' < \infty$, it follows by the identity theorem that $c_{k,\rho} = c_{k,\rho'}$ for all k. Hence $f(x) = \sum_{k \geq 0} c_{k,1} x^k$ (which was Rudin's $f_1(x)$), now on all of $\mathbb{R}$. □

In a parallel vein to Rudin's results and Proposition 3.13, the following strengthening of Schoenberg's result can be shown, using a different (and perhaps more elementary) approach than those of Schoenberg and Rudin.

Theorem 3.14 (Belton–Guillot–Khare–Putinar [7]) *Suppose* $0 < \rho \leq \infty$ *and* $I = (-\rho, \rho)$. *Then the following are equivalent for a function* $f : I \to \mathbb{R}$.

(1) *The entrywise map* $f[-]$ *preserves positivity on* $\mathcal{P}_n(I)$, *for all* $n \geq 1$.
(2) *The entrywise map* $f[-]$ *preserves positivity on the Hankel matrices in* $\mathcal{P}_n(I)$ *of rank at most* 3, *for all* $n \geq 1$.
(3) *The function* f *is real analytic on* I *and absolutely monotonic on* $(0, \rho)$. *In other words,* $f(x) = \sum_{k \geq 0} c_k x^k$ *on* I, *with* $c_k \geq 0 \ \forall k$.

Remark 3.15 Recall the alternate notion of positive definite functions discussed in Remark 2.13. In [68] and related works, Pinkus and other authors study this alternate notion of positive definite functions on H. Notice that such matrices form precisely the set of positive semidefinite symmetric matrices of rank at most $\dim H$. In particular, Theorem 3.14 and the far earlier 1959 paper [77] of Rudin both provide a characterization of these functions, on every Hilbert space of dimension 3 or more.

Parallel to the discussions of the proofs of Schoenberg's and Rudin's results (see the previous chapter), we now explain how to prove Theorem 3.14. Clearly, (3) $\implies$ (1) $\implies$ (2) in the theorem. We first outline how to weaken the condition (2) even further and still imply (3). The key idea is to consider *moment sequences* of certain non-negative measures on the real line. This parallels Rudin's considerations of Fourier–Stieltjes coefficients of non-negative measures on the circle.

Definition 3.16 A measure μ with support in $\mathbb{R}$ is said to be *admissible* if $\mu \geq 0$ on $\mathbb{R}$, and all moments of μ exist and are finite:

$$s_k(\mu) := \int_{\mathbb{R}} x^k \, \mathrm{d}\mu(x) < \infty \qquad (k \geq 0).$$

The sequence $\mathbf{s}(\mu) := \big(s_k(\mu)\big)_{k=0}^{\infty}$ is termed the *moment sequence* of μ. Corresponding to μ and this moment sequence is the *moment matrix* of μ:

$$H_\mu := \begin{bmatrix} s_0(\mu) & s_1(\mu) & s_2(\mu) & \cdots \\ s_1(\mu) & s_2(\mu) & s_3(\mu) & \cdots \\ s_2(\mu) & s_3(\mu) & s_4(\mu) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix};$$

note that $H_\mu = [s_{i+j}(\mu)]_{i,j\geq 0}$ is a semi-infinite Hankel matrix. Finally, a function $f : \mathbb{R} \to \mathbb{R}$ acts entrywise on moment sequences, to yield real sequences:

$$f[\mathbf{s}(\mu)] := \big(f\big(s_0(\mu)\big), \ldots, f\big(s_k(\mu)\big), \ldots\big).$$

We are interested in understanding which entrywise functions preserve the space of moment sequences of admissible measures. The connection to positive semidefinite matrices is made through Hamburger's theorem, which says that a real sequence $(s_0, s_1, \ldots)$ is the moment sequence of an admissible measure on $\mathbb{R}$ if and only if every (finite) principal minor of the moment matrix H_μ is positive semidefinite. For simplicity, this last will be reformulated below to saying that H_μ is positive semidefinite.

The weakening of Theorem 3.14(2) is now explained: it suffices to consider the reduced test set of those Hankel matrices, which arise as the moment matrices of admissible measures supported at three points. Henceforth, let δ_x denote the Dirac probability measure supported at $x \in \mathbb{R}$. It is not hard to verify that the m-point measure $\mu = \sum_{j=1}^m c_j \delta_{x_j}$ has Hankel matrix H_μ with rank no more than m:

$$\begin{aligned} s_k(\mu) &= \sum_{j=1}^{m} c_j x_j^k \qquad (k \geq 0) \\ \implies \quad H_\mu &= \sum_{j=1}^{m} c_j \mathbf{u}_j \mathbf{u}_j^T, \text{ where } \mathbf{u}_j := (1, x_j, x_j^2, \ldots)^T. \end{aligned} \tag{3.3}$$

Thus, a further strengthening of Schoenberg's result is as follows.

Theorem 3.17 (Belton–Guillot–Khare–Putinar [7]) *In the setting of Theorem 3.14, the three assertions contained therein are also equivalent to*

(4) *For each measure*

$$\mu = a\delta_1 + b\delta_{u_0} + c\delta_{-1}, \quad \text{with } u_0 \in (0,1),\ a, b, c \geq 0,\ a+b+c \in (0, \rho), \tag{3.4}$$

there exists an admissible measure σ_μ on $\mathbb{R}$ such that $f\big(s_k(\mu)\big) = s_k(\sigma_\mu)$ for all $k \geq 0$.

In fact, we will see in Sect. 3.4 below that this assertion (4) can be simplified to just assert that $f[H_\mu]$ is positive semidefinite, and so completely avoid the use of Hamburger's theorem.

We now discuss the proof of these results, working with $\rho = \infty$ for ease of exposition. The first observation is that the strengthening of the Horn–Loewner Theorem 3.7, together with the use of Bernstein's theorem (see remark (2) following Theorem 3.4), implies the following "stronger" form of Vasudeva's Theorem 3.3:

Theorem 3.18 (See [7]) *Suppose $I = (0, \infty)$ and $f : I \to \mathbb{R}$. Also fix $u_0 \in (0, 1)$. The following are equivalent:*

(1) *The entrywise map $f[-]$ preserves positivity on $\mathcal{P}_n(I)$ for all $n \geq 1$.*
(2) *The entrywise map $f[-]$ preserves positivity on all moment matrices H_μ for $\mu = a\delta_1 + b\delta_{u_0},\ a, b > 0$.*
(3) *The function f equals a convergent power series $\sum_{k=0}^{\infty} c_k x^k$ for all $x \in I$, with the Maclaurin coefficients $c_k \geq 0$ for all $k \geq 0$.*

Notice that the test matrices in assertion (2) are all Hankel, and of rank at most two. This severely weakens Vasudeva's original hypotheses.

Now suppose the assertion in Theorem 3.17(4) holds. By the preceding result, $f(x)$ is given on $(0, \infty)$ by an absolutely monotonic function $\sum_{k \geq 0} c_k x^k$. The next step is to show that f is continuous. For this, we will crucially use the following "integration trick." Suppose for each admissible measure μ as in (3.4), there is a non-negative measure σ_μ supported on $[-1, 1]$ such that $f\big(s_k(\mu)\big) = s_k(\sigma_\mu)$ for all $k \geq 0$. (Note here that it is not immediate that the support is contained in $[-1, 1]$.)

Now let $p(t) = \sum_{k \geq 0} b_k t^k$ be a polynomial that takes non-negative values on $[-1, 1]$. Then,

$$0 \leq \int_{-1}^{1} p(t)\,\mathrm{d}\sigma_\mu(t) = \sum_{k=0}^{\infty} \int_{-1}^{1} b_k t^k\,\mathrm{d}\sigma_\mu(t) = \sum_{k=0}^{\infty} b_k s_k(\sigma_\mu) = \sum_{k=0}^{\infty} b_k f\big(s_k(\mu)\big). \tag{3.5}$$

Remark 3.19 For example, suppose $p(t) = 1 - t^d$ for some $d \geq 1$. If $\mu = a\delta_1 + b\delta_{u_0} + c\delta_{-1}$, where $u_0 \in (0, 1)$ and $a, b, c > 0$, then the inequality (3.5) gives that

$$0 \leq f\big(s_0(\mu)\big) - f\big(s_d(\mu)\big) = f(a + b + c) - f(a + bu_0^d + c(-1)^d).$$

It is not clear *a priori* how to deduce this inequality using the fact that $f[-]$ preserves matrix positivity and the Hankel moment matrix of μ. The explanation, which we provide in Sect. 3.4 below, connects moment problems, matrix positivity, and real algebraic geometry.

We now outline how (3.5) can be used to prove the continuity of f. First note that $|s_k(\mu)| \leq s_0(\mu)$ for μ as above and all $k \geq 0$. This fact and the easy observation that f is bounded on compact subsets of $\mathbb{R}$ together imply that all moments of σ_μ are uniformly bounded. From this we deduce that σ_μ is necessarily supported on $[-1, 1]$.

The inequality (3.5) now gives the left-continuity of f at $-\beta$, for every $\beta \geq 0$. Fix $u_0 \in (0, 1)$, and let

$$\mu_b := (\beta + bu_0)\delta_{-1} + b\delta_{u_0} \qquad (b > 0).$$

Applying (3.5) to the polynomials $p_{\pm,1}(t) := (1 \pm t)(1 - t^2)$, we deduce that

$$f\big(\beta + b(1 + u_0)\big) - f\big(\beta + b(u_0 + u_0^2)\big) \geq |f(-\beta) - f\big(-\beta - bu_0(1 - u_0^2)\big)|.$$

Letting $b \to 0^+$, the left continuity of f at $-\beta$ follows. Similarly, to show that f is right continuous at $-\beta$, we apply the integral trick to $p_{\pm,1}(t)$ and to $\mu_b' := (\beta + bu_0^3)\delta_{-1} + b\delta_{u_0}$ instead of μ_b.

Having shown continuity, to prove the stronger Schoenberg theorem, we next assume that f is smooth on $\mathbb{R}$. For all $a \in \mathbb{R}$, define the function

$$H_a : \mathbb{R} \to \mathbb{R};\ x \mapsto f(a + e^x).$$

The function H_a satisfies the estimates

$$|H_a^{(n)}(x)| \leq H_{|a|}^{(n)}(x) \qquad (a, x \in \mathbb{R},\ n \in \mathbb{Z}_+). \tag{3.6}$$

This is shown by another use of the integration trick (3.5), this time for the polynomials $p_{\pm,n}(t) := (1 \pm t)(1 - t^2)^n$ for all $n \geq 0$. In turn, the estimates (3.6) lead to showing that H_a is real analytic on $\mathbb{R}$, for all $a \in \mathbb{R}$. Now composing H_{-a} for $a > |x|$ with the function $L_a(y) := \log(a + y)$ shows that $f(x)$ is real analytic on $\mathbb{R}$ and agrees with $\sum_{k\geq 0} a_k x^k$ on $(0, \infty)$. This concludes the proof for smooth functions.

Finally, to pass from smooth functions to continuous functions, we again use a mollified family $f_\delta \to f$ as $\delta \to 0^+$. Each f_δ is the restriction of an entire function, say $\widetilde{f_\delta}$, and the family $\{\widetilde{f_{1/n}} : n \geq 1\}$ forms a normal family on each open disc $D(0, r)$. It follows from results by Montel and Morera that $\widetilde{f_{1/n}}(z)$ converges uniformly to a function g_r on each closed disc $\overline{D(0, r)}$, and g_r is analytic. Since g_r restricts to f on $(-r, r)$, it follows that f is necessarily also real analytic on $\mathbb{R}$, and we are done.

3.4 *The Integration Trick and Positivity Certificates*

Observe that the inequality (3.5) can be written more generally as follows.

Given a polynomial $p(t) = \sum_{k\geq 0} b_k t^k$ *which takes non-negative values on* $[-1, 1]$, *as well as a positive semidefinite Hankel matrix* $H = (s_{i+j})_{i,j\geq 0}$, *we have that*

$$\sum_{k\geq 0} b_k s_k \geq 0. \tag{3.7}$$

As shown in (3.5), this assertion is clear via an application of Hamburger's theorem. We now demonstrate how the assertion can instead be derived from first principles, with interesting connections to positivity certificates.

First note that the inequality (3.7) holds if $p(t)$ is the square of a polynomial. For instance, if $p(t) = (1 - 3t)^2 = 1 - 6t + 9t^2$ on $[-1, 1]$, then

$$s_0 - 6s_1 + 9s_2 = (e_0 - 3e_1)^T H(e_0 - 3e_1), \tag{3.8}$$

where $e_0 = (1, 0, 0, \ldots)$ and $e_1 = (0, 1, 0, 0, \ldots)$. The non-negativity of (3.8) now follows immediately from the positivity of the matrix H. The same reasoning applies if $p(t)$ is a sum of squares of polynomials, or even the limit of a sequence of sums of squares. Thus, one approach to showing the inequality (3.7) for an arbitrary polynomial $p(t)$ which is non-negative on $[-1, 1]$ is to seek a *limiting sum-of-squares representation*, which is also known as a *positivity certificate*, for p.

If a d-variate real polynomial is a sum of squares of real polynomials, then it is clearly non-negative on $\mathbb{R}^d$, but the converse is not true for $d > 1$.[3] Even when $d = 1$, while a sum-of-squares representation is an equivalent characterization for one-variable polynomials that are non-negative on $\mathbb{R}$, here we are working on the compact semi-algebraic set $[-1, 1]$. We now give three proofs of the existence of such a positivity certificate in the setting used above.

Proof 1. A result of Berg, Christensen, and Ressel (see the end of [12]) shows more generally that, for every dimension $d \geq 1$, any non-negative polynomial on $[-1, 1]^d$ has a limiting sum-of-squares representation. □

Proof 2. The only polynomials used in proving the stronger form of Schoenberg's theorem, Theorems 3.14 and 3.17, appear following (3.6):

$$p_{\pm,n}(t) := (1 \pm t)(1 - t^2)^n \qquad (n \geq 0).$$

Each of these polynomials is composed of factors of the form $p_{\pm,0}(t) = 1 \pm t$, so it suffices to produce a limiting sum-of-squares representation for these two polynomials on $[-1, 1]$. Note that

$$\begin{aligned}
\frac{1}{2}(1 \pm t)^2 &= \frac{1}{2} \pm t + \frac{t^2}{2}, \\
\frac{1}{4}(1 - t^2)^2 &= \frac{1}{4} - \frac{t^2}{2} + \frac{t^4}{4}, \\
\frac{1}{8}(1 - t^4)^2 &= \frac{1}{8} - \frac{t^4}{4} + \frac{t^8}{8},
\end{aligned}$$

[3] This is connected to semi-algebraic geometry and to Hilbert's seventeenth problem: recall the famous result of Motzkin that there are non-negative polynomials on $\mathbb{R}^d$ that are not sums of squares, such as $x^4y^2 + x^2y^4 - 3x^2y^2 + 1$. Such phenomena have been studied in several settings, including polytopes (by Farkas, Handelman, and Pólya) and more general semi-algebraic sets (by Putinar, Schmüdgen, Stengel, Vasilescu, and others).

and so on. Adding the first n equations shows that $(1 \pm t) + 2^{-n}(t^{2^n} - 1)$ is a sum-of-squares polynomial for all n. Taking $n \to \infty$ finishes the proof. $\square$

Proof 3. In fact, for any $d \geq 1$ and any compact set $K \subset \mathbb{R}^d$, if f is a non-negative continuous function on K, then f has a positivity certificate. The Stone–Weierstrass theorem gives a sequence of polynomials which converges to $\sqrt{f}$, and the squares of these polynomials then provide the desired limiting representation for f. This is a simpler proof than Proof 1 from [12], but the convergence here is uniform, whereas the convergence in [12] is stronger. $\square$

Remark 3.20 In (3.5), we used $H = H_{\sigma_\mu}$, which was positive semidefinite by assumption. The previous discussion shows that Theorem 3.17(4) can be further weakened, by requiring only that $f[H_\mu]$ is positive semidefinite, as opposed to being equal to H_σ for some admissible measure σ. Hence we do not require Hamburger's theorem in order to prove the strengthening of Schoenberg's theorem that uses the test set of low-rank Hankel matrices.

3.5 Variants of Moment-Sequence Transforms

We now present a trio of results on functions which preserve moment sequences.

For $K \subset \mathbb{R}$, let $\mathcal{M}(K)$ denote the set of moment sequences corresponding to admissible measures with support in K. We say that F maps $\mathcal{M}(K)$ into $\mathcal{M}(L)$, where $K, L \subset \mathbb{R}$, if for every admissible measure μ with support in K there exists an admissible measure σ with support in L such that

$$F(s_k(\mu)) = s_k(\sigma) \quad \text{for all } k \in \mathbb{Z}_+,$$

where $s_k(\mu)$ is the kth-power moment of μ, as in Definition 3.16.

Theorem 3.21 *A function $F : \mathbb{R} \to \mathbb{R}$ maps $\mathcal{M}([-1, 1])$ into itself if and only if F is the restriction to $\mathbb{R}$ of an absolutely monotonic entire function.*

Theorem 3.22 *A function $F : \mathbb{R}_+ \to \mathbb{R}$ maps $\mathcal{M}([0, 1])$ into itself if and only if F is absolutely monotonic on $(0, \infty)$ and $0 \leq F(0) \leq \lim_{\epsilon \to 0^+} F(\epsilon)$.*

Theorem 3.23 *A function $F : \mathbb{R} \to \mathbb{R}$ maps $\mathcal{M}([-1, 0])$ into $\mathcal{M}((-\infty, 0])$ if and only if there exists an absolutely monotonic entire function $\widetilde{F} : \mathbb{C} \to \mathbb{C}$ such that*

$$F(x) = \begin{cases} \widetilde{F}(x) & \text{if } x \in (0, \infty), \\ 0 & \text{if } x = 0, \\ -\widetilde{F}(-x) & \text{if } x \in (-\infty, 0). \end{cases}$$

It is striking to observe the possibility of a discontinuity at the origin which may occur in the latter two of these three theorems.

We will content ourselves here with sketching the proof of the second result. For the others, see [7], noting that the first of the results follows from Theorems 3.14 and 3.17 for $\rho = \infty$.

Proof of Theorem 3.22 Note that the moment matrix corresponding to an element of $\mathcal{M}([0, 1])$ has a zero entry if and only if $\mu = a\delta_0$ for some $a \geq 0$. This and the Schur product theorem give one implication.

For the converse, suppose F preserves $\mathcal{M}([0, 1])$. Fix finitely many scalars c_j, $t_j > 0$ and an integer $n \geq 0$, and set

$$p(t) = (1 - t)^n \quad \text{and} \quad \mu = \sum_j e^{-t_j \alpha} c_j \delta_{e^{-t_j h}}, \tag{3.9}$$

where $\alpha > 0$ and $h > 0$. If $g(x) := \sum_j c_j e^{-t_j x}$ then the integration trick (3.5), but working on $[0, 1]$, shows that the forward finite differences of $F \circ g$ alternate in sign:

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} F\Big(\sum_j c_j e^{-t_j(\alpha + kh)}\Big) \geq 0,$$

so $(-1)^n \Delta_h^n (F \circ g)(\alpha) \geq 0$. As this holds for all $\alpha, h > 0$ and all $n \geq 0$, it follows that $F \circ g : (0, \infty) \to (0, \infty)$ is completely monotonic. The weak density of measures of the form μ, together with Bernstein's Theorem (2.1), gives that $F \circ g$ is completely monotonic on $(0, \infty)$ for every completely monotonic function $g : (0, \infty) \to (0, \infty)$. Finally, a theorem of Lorch and Newman [61, Theorem 5] now gives that $F : (0, \infty) \to (0, \infty)$ is absolutely monotonic. □

3.6 Multivariable Positivity Preservers and Moment Families

We now turn to the multivariable case, and begin with two results of FitzGerald, Micchelli, and Pinkus [33]. We first introduce some notation and a piece of terminology.

Fix $I \subset \mathbb{C}$ and an integer $m \geq 1$, and let

$$A^k = (a_{ij}^k)_{i,j=1}^N \in I^{N \times N} \quad \text{for } k = 1, \ldots, m.$$

For any function $f : I^m \to \mathbb{C}$, we have the $N \times N$ matrix

$$f(A^1, \ldots, A^m) := \big(f(a_{ij}^1, \ldots, a_{ij}^m)\big)_{i,j=1}^N \in \mathbb{C}^{N \times N}.$$

We say that $f : \mathbb{R}^m \to \mathbb{R}$ is *real positivity preserving* if

$$f(A^1, \ldots, A^m) \in \mathcal{P}_N(\mathbb{R}) \text{ for all } A^1, \ldots, A^m \in \mathcal{P}_N(\mathbb{R}) \text{ and all } N \geq 1,$$

where, as above $\mathcal{P}_N(\mathbb{R})$ is the collection of $N \times N$ positive semidefinite matrices with real entries. Similarly, we say that $f : \mathbb{C}^m \to \mathbb{C}$ is positivity preserving if

$$f(A^1, \dots, A^m) \in \mathcal{P}_N \qquad \text{for all } A^1, \dots, A^m \in \mathcal{P}_N \text{ and all } N \geq 1,$$

where $\mathcal{P}_N$ is the collection of $N \times N$ positive semidefinite matrices with complex entries. Finally, recall that a function $f : \mathbb{R}^m \to \mathbb{R}$ is said to be *real entire* if there exists an entire function $F : \mathbb{C}^m \to \mathbb{C}$ such that $F|_{\mathbb{R}^m} = f$. We will also use the multi-index notation

$$\mathbf{x}^\alpha := x_1^{\alpha_1} \cdots x_m^{\alpha_m} \quad \text{if } \mathbf{x} = (x_1, \dots, x_m) \text{ and } \alpha = (\alpha_1, \dots, \alpha_m).$$

The following theorems are natural extensions of Schoenberg's theorem and Herz's theorem, respectively.

Theorem 3.24 ([33, Theorem 2.1]) *Let $f : \mathbb{R}^m \to \mathbb{R}$, where $m \geq 1$. Then f is real positivity preserving if and only f is real entire of the form*

$$f(\mathbf{x}) = \sum_{\alpha \in \mathbb{Z}_+^m} c_\alpha \mathbf{x}^\alpha \qquad (\mathbf{x} \in \mathbb{R}^m),$$

where $c_\alpha \geq 0$ for all $\alpha \in \mathbb{Z}_+^m$.

Theorem 3.25 ([33, Theorem 3.1]) *Let $f : \mathbb{C}^m \to \mathbb{C}$, where $m \geq 1$. Then f is positivity preserving if and only f is of the form*

$$f(\mathbf{z}) = \sum_{\alpha, \beta \in \mathbb{Z}_+^m} c_{\alpha\beta} \mathbf{z}^\alpha \overline{\mathbf{z}}^\beta \qquad (\mathbf{z} \in \mathbb{C}^m),$$

where $c_{\alpha\beta} \geq 0$ for all $\alpha, \beta \in \mathbb{Z}_+^m$ and the power series converges absolutely for all $\mathbf{z} \in \mathbb{C}$.

We now consider the notion of moment family for measures on $\mathbb{R}^d$. As above, a measure on $\mathbb{R}^d$ is said to be *admissible* if it is non-negative and has moments of all orders. Given such a measure μ, we define the *moment family*

$$s_\alpha(\mu) := \int \mathbf{x}^\alpha \, \mathrm{d}\mu(\mathbf{x}) \qquad \text{for all } \alpha \in \mathbb{Z}_+^m.$$

In line with the above, we let $\mathcal{M}(K)$ denote the set of all moment families of admissible measures supported on $K \subset \mathbb{R}^d$.

Note that a measure μ is supported in $[-1, 1]^d$ if and only if its moment family is uniformly bounded:

$$\sup\{|s_\alpha(\mu)| : \alpha \in \mathbb{Z}_+^m\} < \infty.$$

Theorem 3.26 ([7, Theorem 8.1]) *A function $F : \mathbb{R} \to \mathbb{R}$ maps $\mathcal{M}\big([-1, 1]^d\big)$ to itself if and only if F is absolutely monotonic and entire.*

Proof Since $[-1, 1]$ can be identified with $[-1, 1] \times \{0\}^{d-1} \subset [-1, 1]^d$, the forward implication follows from the one-dimensional result, Theorem 3.21.

For the converse, we use the fact [73] that a collection of real numbers $(s_\alpha)_{\alpha \in \mathbb{Z}_+^d}$ is an element of $\mathcal{M}\big([-1, 1]^d\big)$ if and only if the weighted Hankel-type kernels on $\mathbb{Z}_+^d \times \mathbb{Z}_+^d$

$$(\alpha, \beta) \mapsto s_{\alpha+\beta} \quad \text{and} \quad (\alpha, \beta) \mapsto s_{\alpha+\beta} - s_{\alpha+\beta+2\mathbf{1}_j} \quad (1 \le j \le d)$$

are positive semidefinite, where

$$\mathbf{1}_j := (0, \dots, 0, 1, 0, \dots, 0) \in \mathbb{Z}_+^d$$

with 1 in the jth position. Now suppose F is absolutely monotonic and entire; given a family $(s_\alpha)_{\alpha \in \mathbb{Z}_+^d}$ subject to these positivity constraints, we have to verify that the family $(F(s_\alpha))_{\alpha \in \mathbb{Z}_+^d}$ satisfies them as well.

Theorem 3.14 gives that $(\alpha, \beta) \mapsto F(s_{\alpha+\beta})$ and $(\alpha, \beta) \mapsto F(s_{\alpha+\beta+2\mathbf{1}_j})$ are positive semidefinite, so we must show that

$$(\alpha, \beta) \mapsto F(s_{\alpha+\beta}) - F(s_{\alpha+\beta+2\mathbf{1}_j})$$

is positive semidefinite for $j = 1, \dots, d$. As F is absolutely monotonic and entire, it suffices to show that

$$(\alpha, \beta) \mapsto (s_{\alpha+\beta})^{\circ n} - (s_{\alpha+\beta+2\mathbf{1}_j})^{\circ n}$$

is positive semidefinite for any $n \ge 0$, but this follows from the Schur product theorem: if $A \ge B \ge 0$, then

$$A^{\circ n} \ge A^{\circ(n-1)} \circ B \ge A^{\circ(n-2)} \circ B^{\circ 2} \ge \cdots \ge B^{\circ n}.$$

□

We next consider characterizations of real-valued multivariable functions which map tuples of moment sequences to moment sequences.

Let $K_1, \dots, K_m \subset \mathbb{R}$. A function $F : \mathbb{R}^m \to \mathbb{R}$ acts on tuples of moment sequences of (admissible) measures $\mathcal{M}(K_1) \times \cdots \times \mathcal{M}(K_m)$ as follows:

$$F[\mathbf{s}(\mu_1), \dots, \mathbf{s}(\mu_m)]_k := F\big(s_k(\mu_1), \dots, s_k(\mu_m)\big) \quad \text{for all } k \ge 0. \tag{3.10}$$

Given $I \subset \mathbb{R}^m$, a function $F : I \to \mathbb{R}$ is *absolutely monotonic* if F is continuous on I, and for all interior points $\mathbf{x} \in I$ and $\alpha \in \mathbb{Z}_+^m$, the mixed partial derivative

$D^\alpha F(\mathbf{x})$ exists and is non-negative, where

$$D^\alpha F(\mathbf{x}) := \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \cdots \partial x_m^{\alpha_m}} F(x_1, \dots, x_m) \quad \text{and } |\alpha| := \alpha_1 + \cdots + \alpha_m.$$

With this definition, the multivariable analogue of Bernstein's theorem is as one would expect; see [20, Theorem 4.2.2].

To proceed further, it is necessary to introduce the notion of a *facewise absolutely monotonic function* on $\mathbb{R}_+^m$. Observe that the orthant $\mathbb{R}_+^m$ is a convex polyhedron, and is therefore the disjoint union of the relative interiors of its faces. These faces are in one-to-one correspondence with subsets of $[m] := \{1, \dots, m\}$:

$$J \mapsto \mathbb{R}_+^J := \{(x_1, \dots, x_m) \in \mathbb{R}_+^m : x_i = 0 \text{ if } i \notin J\}; \tag{3.11}$$

note that this face has relative interior $\mathbb{R}_{>0}^J := (0, \infty)^J \times \{0\}^{[m]\setminus J}$.

Definition 3.27 A function $F : \mathbb{R}_+^m \to \mathbb{R}$ is *facewise absolutely monotonic* if, for every $J \subset [m]$, there exists an absolutely monotonic function g_J on $\mathbb{R}_+^J$ which agrees with F on $\mathbb{R}_{>0}^J$.

Thus a facewise absolutely monotonic function is piecewise absolutely monotonic, with the pieces being the relative interiors of the faces of the orthant $\mathbb{R}_+^m$. See [7, Example 8.4] for further discussion. In the special case $m = 1$, this broader class of functions (than absolutely monotonic functions on $\mathbb{R}_+$) coincides precisely with the maps which are absolutely monotonic on $(0, \infty)$ and have a possible discontinuity at the origin, as in Theorem 3.22 above.

This definition allows us to characterize the preservers of m-tuples of elements of $\mathcal{M}\big([0, 1]\big)$; the preceding observation shows that Theorem 3.22 is precisely the $m = 1$ case.

Theorem 3.28 ([7, Theorem 8.5]) *Let $F : \mathbb{R}_+^m \to \mathbb{R}$, where the integer $m \geq 1$. The following are equivalent.*

(1) *F maps $\mathcal{M}([0, 1])^m$ into $\mathcal{M}([0, 1])$.*
(2) *F is facewise absolutely monotonic, and the functions $\{g_J : J \subset [m]\}$ are such that $0 \leq g_J \leq g_K$ on $\mathbb{R}_+^J$ whenever $J \subset K \subset [m]$.*
(3) *F is such that*

$$F\big(\sqrt{x_1 y_1}, \dots, \sqrt{x_m y_m}\big)^2 \leq F(x_1, \dots, x_m) F(y_1, \dots, y_m)$$

for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}_+^m$ and there exists some $\mathbf{z} \in (0, 1)^m$ such that the products $\mathbf{z}^\alpha := z_1^{\alpha_1} \cdots z_m^{\alpha_m}$ are distinct for all $\alpha \in \mathbb{Z}_+^m$ and F maps $\mathcal{M}\big(\{1, z_1\}\big) \times \cdots \times \mathcal{M}\big(\{1, z_m\}\big) \cup \mathcal{M}(\{0, 1\})^m$ to $\mathcal{M}(\mathbb{R})$.

The heart of Theorem 3.28 can be deduced from the following result on positivity preservation on tuples of low-rank Hankel matrices. In a sense, it is the multidimensional generalization of the "stronger Vasudeva Theorem" 3.18.

Fix $\rho \in (0, \infty]$, an integer $m \geq 1$ and a point $\mathbf{z} \in (0, 1)^m$ with distinct products, as in Theorem 3.28(3). For all $N \geq 1$, let

$$\mathcal{H}_N := \{a\mathbf{1}_{N\times N} + b\mathbf{u}_{l,N}\mathbf{u}_{l,N}^T : a \in (0, \rho),\ b \in [0, \rho - a),\ 1 \leq l \leq m\},$$

where $\mathbf{u}_{l,N} := (1, z_l, \ldots, z_l^{N-1})^T$.

Theorem 3.29 ([7, Theorem 8.6]) *If $F : (0, \rho)^m \to \mathbb{R}$ preserves positivity on $\mathcal{P}_2\big((0, \rho)\big)^m$ and $\mathcal{H}_N^m$ for all $N \geq 1$, then F is absolutely monotonic and is the restriction of an analytic function on the polydisc $D(0, \rho)^m$.*

The notion of facewise absolute monotonicity emerges from the study of positivity preservers of tuples of moment sequences. If one focuses instead on maps preserving positivity of tuples of all positive semidefinite matrices, or even all Hankel matrices, then this richer class of maps does not appear.

Proposition 3.30 *Suppose $\rho \in (0, \infty]$ and $F : [0, \rho)^m \to \mathbb{R}$. The following are equivalent.*

(1) *$F[-]$ preserves positivity on the space of m-tuples of Hankel matrices with entries in $[0, \rho)$.*
(2) *F is absolutely monotonic on $[0, \rho)^m$.*
(3) *$F[-]$ preserves positivity on the space of m-tuples of all matrices with entries in $[0, \rho)$.*

Proof Clearly (2) $\implies$ (3) $\implies$ (1), so suppose (1) holds. It follows from Theorem 3.29 that F is absolutely monotonic on the domain $(0, \rho)^m$ and agrees there with an analytic function $g : D(0, \rho)^m \to \mathbb{C}$. To see that $F \equiv g$ on $[0, \rho)^m$, we use induction on m, with the $m = 1$ case being left as an exercise (see [7, Proof of Proposition 7.3]).

Now suppose $m > 1$, let $\mathbf{c} = (c_1, \ldots, c_m) \in [0, \rho)^m \setminus (0, \rho)^m$ and define

$$H := \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix} \qquad \text{and} \qquad A_i := \begin{cases} \mathbf{1}_{3\times 3} & \text{if } c_i > 0, \\ H & \text{if } c_i = 0. \end{cases}$$

Choosing $\mathbf{u}_n = (u_{1,n}, \ldots, u_{m,n}) \in (0, \rho)^m$ such that $\mathbf{u}_n \to \mathbf{c}$, it follows that

$$\lim_{n\to\infty} F[u_{1,n}A_1, \ldots, u_{m,n}A_m] = \begin{bmatrix} g(\mathbf{c}) & F(\mathbf{c}) & g(\mathbf{c}) \\ F(\mathbf{c}) & g(\mathbf{c}) & g(\mathbf{c}) \\ g(\mathbf{c}) & g(\mathbf{c}) & g(\mathbf{c}) \end{bmatrix} \in \mathcal{P}_3,$$

where the $(1, 2)$ and $(2, 1)$ entries are as claimed by the induction hypothesis. The determinants of the first and last principal minors now give that

$$g(\mathbf{c}) \geq 0 \qquad \text{and} \qquad -g(\mathbf{c})\big(g(\mathbf{c}) - F(\mathbf{c})\big)^2 \geq 0,$$

whence $F(\mathbf{c}) = g(\mathbf{c})$. □

Having considered functions defined on the positive orthant, we now look at the situation for functions defined over the whole of $\mathbb{R}^m$.

Theorem 3.31 ([7, Theorem 8.9]) *Suppose* $F : \mathbb{R}^m \to \mathbb{R}$ *for some integer* $m \geq 1$. *The following are equivalent.*

(1) *F maps* $\mathcal{M}\big([-1, 1]\big)^m$ *into* $\mathcal{M}(\mathbb{R})$.
(2) *The function F is real positivity preserving.*
(3) *The function F is absolutely monotonic on* $\mathbb{R}_+^m$ *and agrees with an entire function on* $\mathbb{R}^m$.

As before, the proof reveals that verifying positivity preservation for tuples of low-rank Hankel matrices suffices. The following notation and corollary make this precise.

For all $u \in (0, \infty)$, let $\mathcal{M}_u := \mathcal{M}\big(\{-1, u, 1\}\big)$ and

$$\mathcal{M}_{[u]} := \bigcup \big\{\mathcal{M}\big(\{s_1, s_2\}\big) : s_1 \in \{-1, 0, 1\},\ s_2 \in \{-u, 0, u\}\big\}.$$

Corollary 3.32 ([7, Theorem 8.10]) *The hypotheses in Theorem 3.31 are also equivalent to the following.*

(4) There exist $u_0 \in (0, 1)$ *and* $\epsilon > 0$ *such that F maps*

$$\mathcal{M}_{[u_0]}^m \cup \bigcup \big\{\mathcal{M}_{v_1} \times \cdots \times \mathcal{M}_{v_m} : v_1, \dots, v_m \in (0, 1 + \epsilon)\big\}$$

into $\mathcal{M}(\mathbb{R})$.

4 Totally Non-negative Matrices and Positivity Preservers

In this chapter, we discuss variant notions of matrix positivity that are well studied in the literature, *total positivity* and *total non-negativity*, and characterize the maps which preserve these properties.

Definition 4.1 A real matrix A is said to be *totally non-negative* or *totally positive* if every minor of A is non-negative or positive, respectively. We will denote these matrices, as well as the property, by TN and TP.

In older texts, such matrices were called *totally positive* and *strictly totally positive*, respectively.

To introduce the theory of total positivity, we can do no better than quote from the preface of Karlin's magisterial book [54]: "Total positivity is a concept of considerable power that plays an important role in various domains of mathematics, statistics and mechanics." Karlin goes on to list "problems involving convexity, moment spaces, eigenvalues of integral operators, ... oscillation properties of solutions of linear differential equations ... the theory of approximations ... statistical decision procedures ... discerning uniformly most powerful tests for hypotheses ... ascertaining optimal policy for inventory and production processes ... analysis of diffusion-type stochastic processes, and ... coupled mechanical systems."

Perhaps the earliest result on total positivity is due to Fekete, in correspondence with Pólya [32] published in 1912. Schoenberg observed the variation-diminishing properties of TP matrices in 1930 [80], and published a series of papers on Pólya frequency functions, which are defined in terms of total positivity, in the 1950s [87–89]. Independently of Schoenberg, Krein's investigation of ordinary differential equations led him to the total positivity of Green's functions for certain differential operators, and in the mid-1930s his works with Gantmacher looked at spectral and other properties of totally positive matrices and kernels; see [36] and [54, Section 10.6].

For more on these four authors, one may consult the afterwork of Pinkus's book on total positivity [69], which also contains a wealth of results on totally positive and totally non-negative matrices. For a modern collection of applications of the theory of total positivity, see the book edited by Gasca and Micchelli [37].

More recently, total positivity has had a major impact on Lie theory. Lusztig extended the theory of total positivity to the setting of linear algebraic groups; see [62] for an exposition of this work. This led Fomin and Zelevinsky to investigate the combinatorics of Lusztig's theory [34] and resulted in the invention of cluster algebras [35]. These objects have generated an enormous amount of activity in a short period of time, with connections across a wide range of areas within representation theory, combinatorics, geometry, and mathematical physics. For the latter, we will mention only the totally non-negative Grassmannian [72], its connections with scattering amplitudes for quantum field theories [1], and the work by Kodama and Williams on regular soliton solutions of the Kadomtsev–Petviashvili equation [59].

Example 4.2 Perhaps the most well-known class of totally positive matrices consists of the *(generalized) Vandermonde matrices*: for real numbers $0 < x_1 < \cdots < x_m$ and $\alpha_1 < \cdots < \alpha_n$, the $m \times n$ matrix

$$A := [x_j^{\alpha_k}]_{1 \leq j \leq m,\ 1 \leq k \leq n}$$

is totally positive. Indeed, it suffices to show the positivity of any such matrix determinant $\det A$ when $m = n$. That $\det A$ is non-zero follows from Laguerre's extension of Descartes' rule of signs (see [51]) and by fixing the x_j and considering a linear homotopy from $(0, 1, \ldots, n-1)$ to $(\alpha_1, \ldots, \alpha_n)$, one obtains a continuous non-vanishing function from the usual Vandermonde determinant $\prod_{1 \leq j < k \leq n}(x_k - x_j)$ (which is positive) to $\det A$.

Example 4.3 Another prominent class of symmetric totally positive matrices consists of the Hankel moment matrices $H_\mu := [s_{j+k}(\mu)]_{j,k \geq 0}$ corresponding to admissible measures μ; see Definition 3.16.

4.1 Totally Non-negative and Totally Positive Kernels

An important generalization of TN and TP matrices is given by the following functional form.

Definition 4.4 Let X and Y be totally ordered sets, and let $K : X \times Y \to \mathbb{R}$ be a kernel.

(1) The kernel K is *totally positive of order* r, denoted TP_r, if, for any n-tuples of points $x_1 < \cdots < x_n$ in X and $y_1 < \cdots < y_n$ in Y, where $1 \leq n \leq r$, the matrix

$$[K(x_j, y_k)]_{j,k=1}^n$$

has positive determinant.

(2) The kernel K is *totally positive* if K is TP_r for all $r \geq 1$.

(3) Similarly, one defines TN_r kernels and totally non-negative kernels by replacing the word "positive" in the above by "non-negative."

If $X = \{1, \ldots, m\}$ and $Y = \{1, \ldots, n\}$, we recover the earlier notions of totally positive and totally non-negative matrices. When X and Y are taken to be real intervals, TN and TP kernels can be thought of as continuous analogues of TN and TP matrices. In fact, one has a continuous analogue of the Cauchy–Binet formula, which generalizes its traditional version.

Theorem 4.5 (Basic Composition Lemma, See, e.g., [54, 55]) *Suppose $X, Y, Z \subset \mathbb{R}$ and let μ be a non-negative Borel measure on Y. Suppose $K : X \times Y \to \mathbb{R}$ and $L : Y \times Z \to \mathbb{R}$ are pointwise Borel measurable with respect to Y, and let*

$$M : X \times Z \to \mathbb{R};\ (x, z) \mapsto \int_Y K(x, y)L(y, z)\, \mathrm{d}\mu(y).$$

If M is well defined on the whole of $X \times Z$, then

$$\det \begin{bmatrix} M(x_1, z_1) & \ldots & M(x_1, z_m) \\ \vdots & \ddots & \vdots \\ M(x_m, z_1) & \ldots & M(x_m, z_m) \end{bmatrix}$$

$$= \int \cdots \int_{y_1 < y_2 < \cdots < y_m \in Y} \det[K(x_i, y_j)]_{i,j=1}^m \det[L(y_j, z_k)]_{j,k=1}^m \prod_{j=1}^m \mathrm{d}\mu(y_j).$$

As an immediate consequence, we have the following corollary.

Corollary 4.6 *In the setting of Theorem 4.5, if the kernels K and L are both TN_r or TP_r for some $r \geq 1$, then M has the same property. In particular, if K and L are both TN or TP, then so is M.*

We conclude this part with an observation of Pólya that connects to a class of well-studied functions, and also implies the positive definiteness of the Gaussian kernel. Recall from the proof of Theorem 2.4 above that this latter property was crucially used by Schoenberg in characterizing metric space embeddings into Hilbert space; however, its proof above was only outlined (via the more sophisticated machinery of Fourier analysis and Bochner's theorem).

Lemma 4.7 (Pólya) *The Gaussian kernel* $K : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ *given by* $K(x, y) := \exp(-(x-y)^2)$ *is totally positive.*

Proof It suffices to show that every square matrix generated from the kernel has positive determinant. Given real numbers $x_1 < \cdots < x_n$ and $y_1 < \cdots < y_n$, we observe the following factorization:

$$[\exp(-(x_j - y_k)^2)]_{j,k=1}^n = \operatorname{diag}[\exp(-x_j^2)]_{j=1}^n [\exp(2x_j y_k)]_{j,k=1}^n \operatorname{diag}[\exp(-y_k^2)]_{k=1}^n.$$

The proof concludes by observing that all three matrices on the right-hand side have positive determinants, the second because it is a Vandermonde matrix $[p_j^{\alpha_k}]$ with $p_j = \exp(2x_j)$ and $\alpha_k = y_k$. □

Example 4.8 The Gaussian function $f(x) = \exp(-x^2)$ is thus an example of a *Pólya frequency function*, that is, one for which $f(x-y)$ is a TP kernel on $\mathbb{R} \times \mathbb{R}$. As noted above, these functions were intensively studied by Schoenberg, and continue to be much studied in mathematics and statistics; two of the classic references are [22, 27].

The case of the multivariate Gaussian kernel follows immediately from the one-dimensional version.

Corollary 4.9 *For all* $d \geq 1$, *the Gaussian kernel*

$$\mathbb{R}^d \times \mathbb{R}^d \to (0, \infty); \ (\mathbf{x}, \mathbf{y}) \mapsto K(\mathbf{x}, \mathbf{y}) := \exp(-\|\mathbf{x} - \mathbf{y}\|^2)$$

is positive semidefinite on $\mathbb{R}^d \times \mathbb{R}^d$. *In other words, the matrix* $[\exp(-\|\mathbf{x}_j - \mathbf{x}_k\|^2)]_{j,k=1}^n$ *is positive semidefinite for all* $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathbb{R}^d$.

Proof The $d = 1$ case is a direct consequence of Lemma 4.7, and the case of general d follows from this by using the Schur product theorem. □

4.2 Entrywise Preservers of Totally Non-negative Matrices

The TN property is very rigid when it comes to entrywise operations, as the following result makes clear.

Theorem 4.10 ([8, Theorem 2.1]) *Let $F : \mathbb{R}_+ \to \mathbb{R}$ be a function and let $d := \min(m, n)$, where m and n are positive integers. The following are equivalent.*

(1) *F preserves TN entrywise on $m \times n$ matrices.*
(2) *F preserves TN entrywise on $d \times d$ matrices.*
(3) *F is either a non-negative constant or*

  (a) $(d = 1)$ $F(x) \geq 0$*;*
  (b) $(d = 2)$ $F(x) = cx^\alpha$ *for some* $c > 0$ *and some* $\alpha \geq 0$*;*
  (c) $(d = 3)$ $F(x) = cx^\alpha$ *for some* $c > 0$ *and some* $\alpha \geq 1$*;*
  (d) $(d \geq 4)$ $F(x) = cx$ *for some* $c > 0$.

Proof That (1) $\iff$ (2) is immediate, as is the equivalence of (2) and (3) when $d = 1$. For larger values of d, we sketch the implication (2) $\implies$ (3).

For $d = 2$, let the totally non-negative matrices

$$A(x, y) := \begin{bmatrix} x & xy \\ 1 & y \end{bmatrix} \quad \text{and} \quad B(x, y) := \begin{bmatrix} xy & x \\ y & 1 \end{bmatrix} \quad (x, y \geq 0). \tag{4.1}$$

If the non-constant function F preserves TN entrywise for 2×2 matrices, then the non-negativity of the determinants of $F[A(x, y)]$ and $F[B(x, y)]$ gives that

$$F(xy)F(1) = F(x)F(y) \qquad \text{for all } x, y \geq 0. \tag{4.2}$$

It follows that F is strictly positive. Applying Vasudeva's argument, as set out before Proposition 3.8, now implies that F is continuous on $(0, \infty)$. Since the identity (4.2) shows that $x \mapsto F(x)/F(1)$ is multiplicative, there exists an exponent $\alpha \in \mathbb{R}_+$ such that $F(x) = F(1)x^\alpha$ for all $x > 0$. The final details are left as an exercise.

For $d = 3$, note that the 3×3 matrix $A \oplus 0$ is totally non-negative if and only if the 2×2 matrix A is. Hence the previous working gives that $F(x) = cx^\alpha$ for some $c > 0$ and $\alpha \geq 0$. Looking at $\det F[C]$ for the totally non-negative matrix

$$C := \begin{bmatrix} 1 & 1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 1 & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} & 1 \end{bmatrix} \tag{4.3}$$

shows that we must have $\alpha \geq 1$.

The argument to rule out the possibility that $\alpha \in [1, 2)$ when $d \geq 4$ is more involved, but makes use of an example of Fallat, Johnson, and Sokal [31, Example 5.8]. Full details are provided in [8]. □

If our totally non-negative matrices are also required to be symmetric, and so positive semidefinite, then the classes of preservers are enlarged somewhat, but still fairly restrictive.

Theorem 4.11 ([8, Theorem 2.3]) *Let $F : \mathbb{R}_+ \to \mathbb{R}$ and let d be a positive integer. The following are equivalent.*

(1) *F preserves TN entrywise on symmetric $d \times d$ matrices.*
(2) *F is either a non-negative constant or*

  (a) *$(d = 1)$ $F \geq 0$;*
  (b) *$(d = 2)$ F is non-negative, non-decreasing, and multiplicatively mid-convex, that is, $F(\sqrt{xy})^2 \leq F(x)F(y)$ for all $x, y \in [0, \infty)$, so continuous;*
  (c) *$(d = 3)$ $F(x) = cx^\alpha$ for some $c > 0$ and some $\alpha \geq 1$;*
  (d) *$(d = 4)$ $F(x) = cx^\alpha$ for some $c > 0$ and some $\alpha \in \{1\} \cup [2, \infty)$;*
  (e) *$(d \geq 5)$ $F(x) = cx$ for some $c > 0$.*

4.3 Entrywise Preservers of Totally Positive Matrices

In moving from total non-negativity to total positivity, we face two significant technical challenges. Firstly, the idea of realizing totally non-negative $d \times d$ matrices as submatrices of totally non-negative $(d + 1) \times (d + 1)$ matrices, by padding with zeros, does not transfer to the TP setting. Secondly, it is no longer possible to use Vasudeva's idea to establish multiplicative mid-point convexity, since the test matrices used for this are not always totally positive.

The first issue leads us into the domain of *totally positive completion problems* [30]. It is possible to do this generality, using parametrizations of TP matrices [34] or exterior bordering [29, Chapter 9], but the following result has the advantage of providing an explicit embedding into a well-known class of matrices.

Lemma 4.12 ([8, Lemma 3.2]) *Any totally positive 2×2 matrix may be realized as the leading principal submatrix of a positive multiple of a rectangular totally positive generalized Vandermonde matrix of any larger size.*

Remark 4.13 ([8, Remark 3.4]) Lemma 4.12 can be strengthened to the following completion result: given integers $m, n \geq 2$, an arbitrary 2×2 matrix A occurs as a minor in a totally positive $m \times n$ matrix at any given position (that is, in a specified pair of rows and pair of columns) if and only if A is totally positive.

The other tool which will be vital to our deliberations is the following result of Whitney.

Theorem 4.14 ([97, Theorem 1]) *The set of totally positive $m \times n$ matrices is dense in the set of totally non-negative $m \times n$ matrices.*

With these tools in hand, we are able to provide a complete classification of the entrywise TP preservers of each fixed size, akin to the results in the preceding section.

Theorem 4.15 ([8, Theorem 3.1]) *Let* $F : (0, \infty) \to \mathbb{R}$ *be a function and let* $d := \min(m, n)$*, where* m *and* n *are positive integers. The following are equivalent.*

(1) F *preserves total positivity entrywise on* $m \times n$ *matrices.*
(2) F *preserves total positivity entrywise on* $d \times d$ *matrices.*
(3) *The function* F *satisfies*

   (a) $(d = 1)$ $F(x) > 0$*;*
   (b) $(d = 2)$ $F(x) = cx^\alpha$ *for some* $c > 0$ *and some* $\alpha > 0$*;*
   (c) $(d = 3)$ $F(x) = cx^\alpha$ *for some* $c > 0$ *and some* $\alpha \geq 1$*;*
   (d) $(d \geq 4)$ $F(x) = cx$ *for some* $c > 0$.

Proof We sketch the proof that (2) $\implies$ (3) when $d = 2$ and $d \geq 3$. For the first case, working with the matrix

$$\begin{bmatrix} y & x \\ x & y \end{bmatrix} \qquad (y > x > 0)$$

shows that F takes positive values and is increasing, so is Borel measurable and continuous except on a countable set. We now fix a point of continuity a and use the totally positive matrices

$$A(x, y, \epsilon) := \begin{bmatrix} ax & axy \\ a - \epsilon & ay \end{bmatrix} \quad \text{and} \quad B(x, y, \epsilon) := \begin{bmatrix} axy & ax \\ ay & a + \epsilon \end{bmatrix}$$

to show that

$$0 \leq \lim_{\epsilon \to 0^+} \det F[A(x, y, \epsilon)] = F(ax)F(ay) - F(axy)F(a)$$

$$\text{and} \quad 0 \leq \lim_{\epsilon \to 0^+} \det F[B(x, y, \epsilon)] = F(a)F(axy) - F(ax)F(ay)$$

for all $x, y > 0$. Hence $G : x \mapsto F(ax)/F(a)$ is such that

$$G(xy) = G(x)G(y) \qquad \text{for all } x, y > 0,$$

so G is a measurable solution of the Cauchy functional equation. It follows that $G(x) = x^\alpha$ for some $\alpha \in \mathbb{R}$. As F, and so G, is increasing, we must have $\alpha > 0$.

Finally, if $d \geq 3$, then the embedding of Lemma 4.12 and the previous working give positive constants c and α such that $F(x) = cx^\alpha$. In particular, the function F admits a continuous extension $\tilde{F}$ to $\mathbb{R}_+$. The density of TP in TN, that is, Theorem 4.14, implies that $\tilde{F}$ preserves TN entrywise on $d \times d$ matrices. Theorem 4.10 now establishes the form of $\tilde{F}$, and so of F. □

We may consider a version of the previous theorem which restricts to the case of totally positive matrices which are symmetric. A moment's thought leads to the consideration of a symmetric version of the matrix completion problem.

Lemma 4.16 ([8, Lemma 3.7]) *Any symmetric totally positive* 2×2 *matrix occurs as the leading principal submatrix of a totally positive* $d \times d$ *Hankel matrix, where* $d \geq 2$ *can be taken arbitrary large.*

Proof It suffices to embed the matrix

$$\begin{bmatrix} 1 & a \\ a & b \end{bmatrix} \qquad (0 < a < \sqrt{b})$$

into such a Hankel matrix. It is an exercise to prove the existence of a continuous function $f : [0, 1] \to \mathbb{R}_+;\ x \mapsto cx^s$ such that

$$\int_0^1 f(x)\,\mathrm{d}x = a \qquad \text{and} \qquad \int_0^1 f(x)^2\,\mathrm{d}x = b,$$

and then setting

$$a_{jk} := \int_0^1 f(x)^{j+k}\,\mathrm{d}x \qquad (j, k \geq 0)$$

gives a Hankel matrix A as required. The verification of total positivity may be made with the help of Andréief's identity,

$$\det\left[\int \phi_i(x)\psi_j(x)\,\mathrm{d}x\right]_{i,j=1}^k = \frac{1}{k!}\int\cdots\int \det(\phi_i(x_j))_{i,j=1}^k \det(\psi_i(x_j))_{i,j=1}^k\,\mathrm{d}x_1\cdots\mathrm{d}x_k,$$

where $\phi_i(x) = f(x)^{\alpha_i - 1}$ and $\psi_j(x) = f(x)^{\beta_j - 1}$, with

$$1 \leq \alpha_1 < \cdots < \alpha_k \leq d \qquad \text{and} \qquad 1 \leq \beta_1 < \cdots < \beta_k \leq d,$$

together with the total positivity of generalized Vandermonde matrices. □

We remark here that the preceding result can be further strengthened to have the symmetric TP 2×2 matrix occur in any "symmetric" position inside a larger square symmetric TP Hankel matrix, in the spirit of Remark 4.13. See [8, Theorem 3.9] for details.

We now state the symmetric version of Theorem 4.15.

Theorem 4.17 ([8, Theorem 3.6]) *Let* $F : (0, \infty) \to \mathbb{R}$ *and let* d *be a positive integer. The following are equivalent.*

(1) *F preserves total positivity entrywise on symmetric $d \times d$ matrices.*
(2) *The function F satisfies*

(a) $(d = 1)$ $F(x) > 0$;
(b) $(d = 2)$ F *is positive, increasing, and multiplicatively mid-convex, that is,* $F(\sqrt{xy})^2 \leq F(x)F(y)$ *for all* $x, y \in (0, \infty)$, *so continuous;*
(c) $(d = 3)$ $F(x) = cx^\alpha$ *for some* $c > 0$ *and some* $\alpha \geq 1$;
(d) $(d = 4)$ $F(x) = cx^\alpha$ *for some* $c > 0$ *and some* $\alpha \in \{1\} \cup [2, \infty)$.
(e) $(d \geq 5)$ $F(x) = cx$ *for some* $c > 0$.

Although we have developed the key ingredients to prove this theorem, we content ourselves with referring the interested reader to [8].

References

1. N. Arkani-Hamed, J.L. Bourjaily, F. Cachazo, A.B. Goncharov, A. Postnikov, J. Trnka. Scattering amplitudes and the positive Grassmannian (2012). Preprint. Available at http://arxiv.org/abs/1212.5605
2. V.S. Barbosa, V.A. Menegatto, Strictly positive definite kernels on compact two-point homogeneous spaces. Math. Inequal. Appl. **19**(2), 743–756 (2016)
3. V.S. Barbosa, V.A. Menegatto, Strict positive definiteness on products of compact two-point homogeneous spaces. Integr. Transf. Spec. Funct. **28**(1), 56–73 (2017)
4. R.K. Beatson, W. zu Castell, Dimension hopping and families of strictly positive definite zonal basis functions on spheres. J. Approx. Theory **221**(C), 22–37 (2017)
5. R.K. Beatson, W. zu Castell, Y. Xu, A Pólya criterion for (strict) positive-definiteness on the sphere. IMA J. Numer. Anal. **34**(2), 550–568 (2014)
6. A. Belton, D. Guillot, A. Khare, M. Putinar, Matrix positivity preservers in fixed dimension. I. Adv. Math. **298**, 325–368 (2016)
7. A. Belton, D. Guillot, A. Khare, M. Putinar, Moment-sequence transforms (2016). Preprint. Available at http://arxiv.org/abs/1610.05740
8. A. Belton, D. Guillot, A. Khare, M. Putinar, Total-positivity preservers (2017). Preprint. Available at http://arxiv.org/abs/1711.10468
9. A. Belton, D. Guillot, A. Khare, M. Putinar, A panorama of positivity (2018). Preprint. Available at http://arxiv.org/abs/1812.05482
10. A. Belton, D. Guillot, A. Khare, M. Putinar, A panorama of positivity. II: fixed dimension, in *Complex Analysis and Spectral Theory: Thomas Ransford Festschrift*, ed. by J. Mashreghi. CRM Proceedings and Lecture Notes Series (American Mathematical Society, Providence, to appear)
11. C. Berg, E. Porcu, From Schoenberg coefficients to Schoenberg functions. Constr. Approx. **45**(2), 217–241 (2017)
12. C. Berg, J.P.R. Christensen, P. Ressel, Positive definite functions on abelian semigroups. Math. Ann. **223**(3), 253–274 (1976)
13. C. Berg, A.P. Peron, E. Porcu, Schoenberg's theorem for real and complex Hilbert spheres revisited. J. Approx. Theory **228**, 58–78 (2018)
14. S. Bernstein, Sur les fonctions absolument monotones. Acta Math. **52**(1), 1–66 (1929)
15. G. Birkhoff, Metric foundations of geometry. I. Trans. Am. Math. Soc. **55**, 465–492 (1944)
16. L.M. Blumenthal, *Theory and Applications of Distance Geometry*, 2nd edn. (Chelsea Publishing Co., Bronx, 1970)
17. R.P. Boas, Jr., D.V. Widder, Functions with positive differences. Duke Math. J. **7**(1), 496–503 (1940)
18. S. Bochner, Monotone Funktionen, Stieltjessche Integrale und harmonische Analyse. Math. Ann. **108**(1), 378–410 (1933)
19. S. Bochner, Hilbert distances and positive definite functions. Ann. Math. **42**(3), 647–656 (1941)

20. S. Bochner, *Harmonic Analysis and the Theory of Probability* (University of California Press, Berkeley, 1955)
21. R.N. Bonfim, J.C. Guella, V.A. Menegatto, Strictly positive definite functions on compact two-point homogeneous spaces: the product alternative. SIGMA Symmetry Integrability Geom. Methods Appl. **14**, 112 (2018)
22. F. Brenti, Unimodal, log-concave and Pólya frequency sequences in combinatorics. Mem. Am. Math. Soc. **81**(413), viii+106 (1989)
23. D. Chen, V.A. Menegatto, X. Sun, A necessary and sufficient condition for strictly positive definite functions on spheres. Proc. Am. Math. Soc. **131**(9), 2733–2740 (2003)
24. J.P.R. Christensen, P. Ressel, Functions operating on positive definite matrices and a theorem of Schoenberg. Trans. Am. Math. Soc. **243**, 89–95 (1978)
25. H. Cohn, Y. Zhao, Sphere packing bounds via spherical codes. Duke Math. J. **163**(10), 1965–2002 (2014)
26. J. Dattorro, Equality relating Euclidean distance cone to positive semidefinite cone. Linear Algebra Appl. **428**(11–12), 2597–2600 (2008)
27. B. Efron, Increasing properties of Pólya frequency functions. Ann. Math. Stat. **36**(1), 272–279 (1965)
28. J. Emonds, H. Führ, Strictly positive definite functions on compact abelian groups. Proc. Am. Math. Soc. **139**(3), 1105–1113 (2011)
29. S.M. Fallat, C.R. Johnson, *Totally Nonnegative Matrices*. Princeton Series in Applied Mathematics (Princeton University Press, Princeton, 2011)
30. S.M. Fallat, C.R. Johnson, R.L. Smith, The general totally positive matrix completion problem with few unspecified entries. Electron. J. Linear Algebra **7**, 1–20 (2000)
31. S.M. Fallat, C.R. Johnson, A.D. Sokal, Total positivity of sums, Hadamard products and Hadamard powers: results and counterexamples. Linear Algebra Appl. **520**, 242–259 (2017)
32. M. Fekete, Über ein problem von Laguerre. Rend. Circ. Math. Palermo **34**, 89–120 (1912)
33. C.H. FitzGerald, C.A. Micchelli, A. Pinkus, Functions that preserve families of positive semidefinite matrices. Linear Algebra Appl. **221**, 83–102 (1995)
34. S. Fomin, A. Zelevinsky, Total positivity: tests and parametrizations. Math. Intell. **22**(1), 23–33 (2000)
35. S. Fomin, A. Zelevinsky, Cluster algebras. I. Foundations. J. Am. Math. Soc. **15**(2), 497–529 (2002)
36. F.R. Gantmacher, M.G. Krein, *Oscillation Matrices and Kernels and Small Vibrations of Mechanical Systems*, ed. by A. Eremenko, revised edn. (AMS Chelsea Publishing, New York, 2002)
37. M. Gasca, C.A. Micchelli (eds.), *Total Positivity and its Applications*. Mathematics and Its Applications, vol. 359 (Springer, Utrecht, 1996)
38. I.M. Gelfand, Normierte Ringe. Rec. Math. [Mat. Sbornik] N. S. **9**(51), 3–24 (1941)
39. T. Gneiting, Strictly and non-strictly positive definite functions on spheres. Bernoulli **19**(4), 1327–1349 (2013)
40. J.C. Guella, V.A. Menegatto, Strictly positive definite kernels on a product of spheres. J. Math. Anal. Appl. **435**(1), 286–301 (2016)
41. J.C. Guella, V.A. Menegatto, Strictly positive definite kernels on the torus. Constr. Approx. **46**(2), 271–284 (2017)
42. J.C. Guella, V.A. Menegatto, A.P. Peron, An extension of a theorem of Schoenberg to products of spheres. Banach J. Math. Anal. **10**(4), 671–685 (2016)
43. J.C. Guella, V.A. Menegatto, A.P. Peron, Strictly positive definite kernels on a product of spheres II. SIGMA Symmetry Integrability Geom. Methods Appl. **12**, 15 (2016)
44. J.C. Guella, V.A. Menegatto, A.P. Peron, Strictly positive definite kernels on a product of circles. Positivity **21**(1), 329–342 (2017)
45. D. Guillot, B. Rajaratnam, Retaining positive definiteness in thresholded matrices. Linear Algebra Appl. **436**(11), 4143–4160 (2012)
46. D. Guillot, B. Rajaratnam, Functions preserving positive definiteness for sparse matrices. Trans. Am. Math. Soc. **367**(1), 627–649 (2015)

47. D. Guillot, A. Khare, B. Rajaratnam, Preserving positivity for rank-constrained matrices. Trans. Am. Math. Soc. **369**(9), 6105–6145 (2017)
48. H. Helson, J.-P. Kahane, Y. Katznelson, W. Rudin, The functions which operate on Fourier transforms. Acta Math. **102**(1–2), 135–157 (1959)
49. C.S. Herz, Fonctions opérant sur les fonctions définies-positives. Ann. Inst. Fourier (Grenoble) **13**(1), 161–180 (1963)
50. R.A. Horn, The theory of infinitely divisible matrices and kernels. Trans. Am. Math. Soc. **136**, 269–286 (1969)
51. G.J.O. Jameson, Counting zeros of generalised polynomials: descartes' rule of signs and Laguerre's extensions. Math. Gaz. **90**(518), 223–234 (2006)
52. J.-P. Kahane, Sur un théorème de Wiener–Lévy. C. R. Acad. Sci. Paris **246**, 1949–1951 (1958)
53. J.-P. Kahane, W. Rudin, Caractérisation des fonctions qui opèrent sur les coefficients de Fourier-Stieltjes. C. R. Acad. Sci. Paris **247**, 773–775 (1958)
54. S. Karlin, *Total Positivity*, vol. I (Stanford University Press, Palo Alto, 1968)
55. S. Karlin, Y. Rinott, A generalized Cauchy–Binet formula and applications to total positivity and majorization. J. Multivar. Anal. **27**(1), 284–299 (1988)
56. Y. Katznelson, Sur les fonctions opérant sur l'algèbre des séries de Fourier absolument convergentes. C. R. Acad. Sci. Paris **247**, 404–406 (1958)
57. A. Khare, Smooth entrywise positivity preservers, a Horn–Loewner master theorem, and Schur polynomials (2018). Preprint. Available at http://arxiv.org/abs/1809.01823
58. A. Khare, T. Tao, On the sign patterns of entrywise positivity preservers in fixed dimension (2017). Preprint. Available at http://arxiv.org/abs/1708.05197
59. Y. Kodama, L. Williams, KP solitons and total positivity for the Grassmannian. Invent. Math. **198**(3), 637–699 (2014)
60. L. Liberti, C. Lavor, N. Maculan, A. Mucherino, Euclidean distance geometry and applications. SIAM Rev. **56**(1), 3–69 (2014)
61. L. Lorch, D.J. Newman, On the composition of completely monotonic functions and completely monotonic sequences and related questions. J. Lond. Math. Soc. **28**(1), 31–45 (1983)
62. G. Lusztig, Introduction to total positivity, in *Positivity in Lie Theory: Open Problems*, ed. by J. Hilgert, J.D. Lawson, K.-H. Neeb, E.B. Vinberg. De Gruyter Expositions in Mathematics, vol. 26 (Walter de Gruyter & Co., Berlin, 1998), pp. 133–145
63. K. Menger, Untersuchungen über allgemeine Metrik. Math. Ann. **100**(1), 75–163 (1928)
64. K. Menger, New foundation of euclidean geometry. Am. J. Math. **53**(4), 721–745 (1931)
65. J. Møller, M. Nielsen, E. Porcu, E. Rubak, Determinantal point process models on the sphere. Bernoulli **24**(2), 1171–1201 (2018)
66. O.R. Musin, The kissing number in four dimensions. Ann. Math. **168**(1), 1–32 (2008)
67. O.R. Musin, Multivariate positive definite functions on spheres, in *Discrete Geometry and Algebraic Combinatorics*. Contemporary Mathematics, vol. 625 (American Mathematical Society, Providence, 2014), pp. 177–190
68. A. Pinkus, Strictly positive definite functions on a real inner product space. Adv. Comput. Math. **20**(4), 263–271 (2004)
69. A. Pinkus, *Totally Positive Matrices*. Cambridge Tracts in Mathematics, vol. 181 (Cambridge University Press, Cambridge, 2010)
70. G. Pólya, G. Szegő, *Aufgaben und Lehrsätze aus der Analysis. Band II: Funktionentheorie, Nullstellen, Polynome Determinanten, Zahlentheorie* (Springer, Berlin, 1971)
71. E. Porcu, A. Alegria, R. Furrer, Modeling temporally evolving and spatially globally dependent data. Int. Stat. Rev. **86**(2), 344–377 (2018)
72. A. Postnikov, D. Speyer, L. Williams, Matching polytopes, toric geometry, and the totally non-negative Grassmannian. J. Algebraic Combin. **30**(2), 173–191 (2009)
73. M. Putinar, Positive polynomials on compact semi-algebraic sets. Indiana Univ. Math. J. **42**(3), 969–984 (1993)

74. A.W. Roberts, D.E. Varberg, *Convex Functions*. Pure and Applied Mathematics, vol. 57 (Academic, New York, 1973)
75. A.J. Rothman, E. Levina, J. Zhu, Generalized thresholding of large covariance matrices. J. Am. Stat. Assoc. **104**(485), 177–186 (2009)
76. W. Rudin, Transformations des coefficients de Fourier. C. R. Acad. Sci. Paris **243**, 638–640 (1956)
77. W. Rudin, Positive definite sequences and absolutely monotonic functions. Duke Math. J. **26**(4), 617–622 (1959)
78. W. Rudin, Some theorems on Fourier coefficients. Proc. Am. Math. Soc. **10**(6), 855–859 (1959)
79. W. Rudin, A strong converse of the Wiener-Levy theorem. Can. J. Math. **14**(4), 694–701 (1962)
80. I.J. Schoenberg, Über variationsvermindernde lineare Transformationen. Math. Z. **32**(1), 321–328 (1930)
81. I.J. Schoenberg, Remarks to Maurice Fréchet's article "Sur la définition axiomatique d'une classe d'espace distanciés vectoriellement applicable sur l'espace de Hilbert". Ann. Math. **36**(3), 724–732 (1935)
82. I.J. Schoenberg, On certain metric spaces arising from Euclidean spaces by a change of metric and their imbedding in Hilbert space. Ann. Math. **38**(4), 787–793 (1937)
83. I.J. Schoenberg, Metric spaces and completely monotone functions. Ann. Math. **39**(4), 811–841 (1938)
84. I.J. Schoenberg, Metric spaces and positive definite functions. Trans. Am. Math. Soc. **44**(3), 522–536 (1938)
85. I.J. Schoenberg, On metric arcs of vanishing Menger curvature. Ann. Math. **41**(4), 715–726 (1940)
86. I.J. Schoenberg, Positive definite functions on spheres. Duke Math. J. **9**(1), 96–108 (1942)
87. I.J. Schoenberg, On Pólya frequency functions. II. Variation-diminishing integral operators of the convolution type. Acta Sci. Math. Szeged **12**, 97–106 (1950)
88. I.J. Schoenberg, On Pólya frequency functions. I. The totally positive functions and their Laplace transforms. J. Anal. Math. **1**(1), 331–374 (1951)
89. I.J. Schoenberg, A.M. Whitney, On Pólya frequency functions. III. The positivity of translation determinants with an application to the interpolation problem by spline curves. Trans. Am. Math. Soc. **74**(2), 246–259 (1953)
90. I. Schur, Bemerkungen zur Theorie der beschränkten Bilinearformen mit unendlich vielen Veränderlichen. J. Reine Angew. Math. **140**, 1–28 (1911)
91. I. Steinwart, On the influence of the kernel on the consistency of support vector machines. J. Mach. Learn. Res. **2**(1), 67–93 (2002)
92. V.N. Vapnik, The nature of statistical learning theory, in *Statistics for Engineering and Information Science*, 2nd edn. (Springer, New York, 2000)
93. H.L. Vasudeva, Positive definite matrices and absolutely monotonic functions. Indian J. Pure Appl. Math. **10**(7), 854–858 (1979)
94. J. von Neumann, I.J. Schoenberg, Fourier integrals and metric geometry. Trans. Am. Math. Soc. **50**(2), 226–251 (1941)
95. H.-C. Wang, Two-point homogeneous spaces. Ann. Math. **55**(1), 177–191 (1952)
96. H. Weyl, Harmonics on homogeneous manifolds. Ann. Math. **35**(3), 486–499 (1934)
97. A.M. Whitney, A reduction theorem for totally positive matrices. J. Anal. Math. **2**(1), 88–92 (1952)
98. Y. Xu, Positive definite functions on the unit sphere and integrals of Jacobi polynomials. Proc. Am. Math. Soc. **146**(5), 2039–2048 (2018)
99. Y. Xu, E.W. Cheney, Strictly positive definite functions on spheres. Proc. Am. Math. Soc. **116**(4), 977–981 (1992)
100. J. Ziegel, Convolution roots and differentiability of isotropic positive definite functions on spheres. Proc. Am. Math. Soc. **142**(6), 2063–2077 (2014)

Table of Contents from Part II of the Survey

Inner Functions in Reproducing Kernel Spaces

Raymond Cheng, Javad Mashreghi, and William T. Ross

In memory of S. Shimorin

Abstract In Beurling's approach to inner functions for the shift operator S on the Hardy space H^2, a function f is inner when $f \perp S^n f$ for all $n \geqslant 1$. Inspired by this approach, this paper develops a notion of an inner vector $\mathbf{x}$ for any operator T on a Hilbert space, via the analogous condition $\mathbf{x} \perp T^n\mathbf{x}$ for all $n \geqslant 1$. We study these inner vectors in a variety of settings. Using Birkhoff–James orthogonality, we extend this notion of inner vector for an operator on a Banach space. We then apply this development of inner function to recast a theorem of Shapiro and Shields to discuss the zero sets for functions in Hilbert spaces, as well as obtain a corresponding result for zero sets for a wide class of Banach spaces.

Keywords BJ-orthogonality · Inner functions

2010 Mathematics Subject Classification Primary: 30A31, 30A38, 30A78; Secondary: 30A98, 30C15

This work was supported by NSERC (Canada).

R. Cheng
Department of Mathematics and Statistics, Old Dominion University, Norfolk, VA, USA
e-mail: rcheng@odu.edu

J. Mashreghi
Département de Mathématiques et de Statistique, Université Laval, Québec, QC, Canada
e-mail: javad.mashreghi@mat.ulaval.ca

W. T. Ross (✉)
Department of Mathematics and Computer Science, University of Richmond, Richmond, VA, USA
e-mail: wross@richmond.edu

A. Aleman et al. (eds.), *Analysis of Operators on Function Spaces*,
Trends in Mathematics, https://doi.org/10.1007/978-3-030-14640-5_6

1 Introduction

Inspired by Beurling's analysis of the structure of the shift invariant subspaces of the classical Hardy space H^2 [4, 11], and by similar analysis in other settings [1, 3, 22–24, 27], we explored a notion of "inner function" in the sequence space ℓ^p_A and used it to characterize its zero sets [8, 10]. As this "Beurling approach" seems to be ubiquitous, we will survey a method from [27] to the setting of reproducing kernel Hilbert spaces of analytic functions, as we head towards an analogous result for Banach spaces of analytic functions.

Broadly speaking, we start with a Banach space $\mathscr{X}$ of analytic functions on a bounded planar domain Ω for which, among some mild technical conditions (see Sect. 4), the shift operator $(S_{\mathscr{X}} f)(z) = zf(z)$ is well defined and continuous. We will examine a notion of "orthogonality" $f \perp_{\mathscr{X}} g$ for $f, g \in \mathscr{X}$ due to Birkhoff and James [18] (see Sect. 7) and use this orthogonality to define an $S_{\mathscr{X}}$*-inner function* to be an $f \in \mathscr{X} \setminus \{0\}$ for which

$$f \perp_{\mathscr{X}} S^n_{\mathscr{X}} f, \quad n \geqslant 1.$$

When Ω is the open unit disk $\mathbb{D}$ and $\mathscr{X}$ is the classical Hardy space H^2, basic Fourier analysis will show that an S_{H^2}-inner function is a bounded analytic function on $\mathbb{D}$ for which the radial boundary function has constant modulus almost everywhere, in agreement with the classical and well-known notion of inner. Similarly defined inner functions were explored in other spaces [1, 3, 10, 26]. As a topic to be explored in future work, a more general notion of T-inner vector will be presented in this paper, in which T is a bounded linear transformation on a Banach space $\mathscr{X}$, and a vector $\mathbf{x} \in \mathscr{X}$ is said to be T-inner if $\mathbf{x} \perp_{\mathscr{X}} T^n\mathbf{x}$ for all $n \geqslant 1$.

This abstract notion of "inner" arises naturally in prediction theory for norm-stationary processes. We say that a nonzero sequence $\{X_k\}_{k\in\mathbb{Z}}$ in a Banach space $\mathscr{X}$ is *norm stationary* when

$$\Big\| \sum_{j=1}^{m} a_j X_{k_j} \Big\| = \Big\| \sum_{j=1}^{m} a_j X_{k_j+t} \Big\| \tag{1.1}$$

for all $t \in \mathbb{Z}$, coefficients $a_j \in \mathbb{C}$, and indices $k_j \in \mathbb{N}$. The identity in (1.1) induces an isometry T on

$$\mathscr{M} := \bigvee \{X_0, X_1, X_2, \ldots\},$$

the closed linear span of the sequence $\{X_k\}_{k\geqslant 0}$, for which

$$TX_k = X_{k+1}, \quad k \geqslant 0.$$

Writing $\widehat{X}_0$ for a metric projection (nearest point) of X_0 onto $T\mathscr{M}$, one can show that the vector $X_0 - \widehat{X}_0$ is T-inner on $\mathscr{M}$. This construction appears in studies involving norm-stationary processes with infinite variance [6, 7, 20], extending, in part, the extensive literature on stationary Gaussian processes. In particular, the results from [20] seek to find a Wold-like decomposition in this setting.

This paper is structured as follows. In Sect. 2 we discuss a general notion of a T-inner vector, where T is a bounded linear transformation on a Hilbert space, and give a variety of examples, and encourage the reader to investigate further. In Sect. 3 we develop some basic properties of T-inner vectors and show in Proposition 3.1 that all T-inner vectors take a particular form.

In Sect. 4 we apply this notion of T-inner to recast some work of Shapiro and Shields [27] (in which the concept of inner also has its roots in the work of Beurling), in terms of $S_{\mathscr{H}}$-inner functions, to characterize the zero sets of a Hilbert space of analytic functions on a bounded planar domain (see Theorem 4.12). This will lead us in several directions. First, we explore whether the $S_{\mathscr{H}}$-inner function associated with a polynomial has extra zeros. Indeed, with the Hardy space H^2, the inner factor of a function in H^2 has exactly all of the zeros of the original function, and no others. In Sect. 5 we develop conditions (see Theorem 5.1) for which the $S_{\mathscr{H}}$-inner function J associated with an $f \in \mathscr{H}$ (where $\mathscr{H}$ is a Hilbert space of analytic functions on a bounded planar domain) has only the zeros of f, and no others. In particular, our result applies to the shift operator on the well-known Dirichlet space (see Corollary 5.4) as well as shift operator on a space studied by Korenblum (see Corollary 5.5).

Second, we investigate the connection between inner functions and zero sets. In particular, we encounter the phenomenon of an $S_{\mathscr{H}}$-inner function J having "extra zeros," that is, zeros in addition to a prescribed set. The existence of such extra zeros was first demonstrated in [15], where $\mathscr{H}$ was a weighted Bergman space. In Sect. 6 we give a large class of spaces $\mathscr{H}$ for which the $S_{\mathscr{H}}$-inner function associated with a linear polynomial has extra zeros.

Third, so far, we have focused on Hilbert spaces. In our final two sections we develop, via Birkhoff–James orthogonality, notions of "inner" for operators on Banach spaces. Our concept of inner will coincide with the classical definition for the Hardy classes H^p, when $p \in (1, \infty)$. In addition, we discuss the zero sets for Banach spaces of analytic functions on a planar domain, and prove an extension of the Shapiro-Shields result.

2 Inner Vectors in Hilbert Spaces

Let us begin with a discussion of T-inner vectors for Hilbert space operators T, where one can take a very broad approach. We will see later in the Banach space setting that some restrictions become necessary in order for the definitions to make sense.

Let $\mathscr{H}$ be a complex Hilbert space with inner product $\langle\cdot,\cdot\rangle$, and let T be a bounded linear operator on $\mathscr{H}$. We say a vector $\mathbf{v} \in \mathscr{H} \setminus \{0\}$ is *T-inner* when

$$\mathbf{v} \perp T^n\mathbf{v}, \quad n \geqslant 1.$$

For a vector $\mathbf{w} \in \mathscr{H}$, let

$$[\mathbf{w}]_T := \bigvee\{\mathbf{w}, T\mathbf{w}, T^2\mathbf{w}, \ldots\} \tag{2.1}$$

denote the T-invariant subspaces generated by $\mathbf{w}$. When the context is clear we will use $[\mathbf{w}]$ in place of $[\mathbf{w}]_T$. Observe that $\mathbf{v}$ is T-inner precisely when $\mathbf{v} \perp [T\mathbf{v}]_T$. Here are a few examples of T-inner vectors.

Example 2.2 Suppose that T is the shift operator $(Tf)(z) = zf(z)$ on the classical Hardy space H^2 [11]. Via standard theory of radial boundary values, the inner product on H^2 can be written as the integral

$$\langle f, g\rangle = \int_0^{2\pi} f(e^{i\theta})\overline{g(e^{i\theta})}\frac{d\theta}{2\pi}, \quad f, g \in H^2. \tag{2.3}$$

Thus a function (vector) $f \in H^2 \setminus \{0\}$ is T-inner precisely when

$$0 = \langle f, T^n f\rangle = \int_0^{2\pi} |f(e^{i\theta})|^2 e^{-in\theta}\frac{d\theta}{2\pi}, \quad n \geqslant 1.$$

The equation above, along with its complex conjugate, shows f is T-inner precisely when all but the zeroth Fourier coefficients of $|f|^2$ vanish. This implies that the function $\theta \mapsto |f(e^{i\theta})|$ is constant almost everywhere. The condition "$|f|$ has constant radial limit values almost everywhere on the unit circle" is the classical definition of inner [11]—though one usually normalizes things so that inner means $|f(e^{i\theta})| = 1$ for almost every θ. We will refer to this notion of inner as *classical inner*.

Example 2.4 Suppose that $(Tf)(z) = z^2 f(z)$, the *square* of the unilateral shift on H^2. Then, with a similar analysis as in the previous example, $f \in H^2$ is T-inner when

$$\int_0^{2\pi} |f(e^{i\theta})|^2 e^{2ik\theta}\frac{d\theta}{2\pi} = 0, \quad k \in \mathbb{Z} \setminus \{0\},$$

though it is somewhat unclear what to glean from this condition. Certainly any classical inner function is a T-inner function. However, functions like $f(z) = a + bz$, which are not classical inner when a and b are both nonzero, is a T-inner function. Observe that this class of T-inner functions is closed under multiplication by classical inner functions.

With a little extra effort, and transferring the problem to a different venue, we can describe the T-inner functions more explicitly. Indeed, if

$$H^2 \oplus H^2 := \{f \oplus g : f, g \in H^2\}$$

with norm

$$\|f \oplus g\|^2_{H^2 \oplus H^2} := \int_0^{2\pi} |f(e^{i\theta})|^2 \frac{d\theta}{2\pi} + \int_0^{2\pi} |g(e^{i\theta})|^2 \frac{d\theta}{2\pi},$$

then the operator

$$U : H^2 \to H^2 \oplus H^2,$$

defined by

$$U\Big(\sum_{n=0}^{\infty} a_n z^n\Big) = \Big(\sum_{n=0}^{\infty} a_{2n} z^n, \sum_{n=0}^{\infty} a_{2n+1} z^n\Big) \tag{2.5}$$

is unitary. Furthermore, if $(Sf)(z) = zf(z)$ is the shift on H^2, we have

$$S \oplus S : H^2 \oplus H^2 \to H^2 \oplus H^2, \quad (S \oplus S)(f \oplus g) = (Sf) \oplus (Sg),$$

and one can show that $US^2 = (S \oplus S)U$. Thus $f \in H^2$ is S^2-inner, if and only if $Uf \in H^2 \oplus H^2$ is $S \oplus S$-inner. If $Uf = f_1 \oplus f_2$ as in (2.5), then f is S^2-inner when

$$\begin{aligned} 0 &= \langle (f_1 \oplus f_2, (S \oplus S)^n (f_1 \oplus f_2) \rangle_{H^2 \oplus H^2} \\ &= \langle f_1 \oplus f_2, (S^n f_1) \oplus (S^n f_2) \rangle_{H^2 \oplus H^2} \\ &= \int_0^{2\pi} |f_1(e^{i\theta})|^2 e^{-in\theta} \frac{d\theta}{2\pi} + \int_0^{2\pi} |f_2(e^{i\theta})|^2 e^{-in\theta} \frac{d\theta}{2\pi}. \end{aligned}$$

The above equation, along with its complex conjugate, shows that $|f_1|^2 + |f_2|^2$ is (almost everywhere) constant on the circle. We leave it to the reader to show that $U^{-1}(f_1 \oplus f_2)$ is equal to $f_1(z^2) + zf_2(z^2)$ and thus $f \in H^2$ is S^2-inner if and only if

$$f(z) = f_1(z^2) + zf_2(z^2),$$

where $f_1, f_2 \in H^2$ with $|f_1|^2 + |f_2|^2$ is constant almost everywhere on $\mathbb{T}$.

This example only scratches the surface of a much wider (and deeper) theory of shifts of higher multiplicity and the well-developed Beurling–Lax theorem [16].

Example 2.6 The previous example can be extended even further to $T = T_\phi$, $\phi \in H^\infty$ is an analytic Toeplitz operator on H^2 with symbol ϕ, i.e., $T_\phi f = \phi f$. Here $f \in H^2 \setminus \{0\}$ is T_ϕ-inner when

$$\int_0^{2\pi} |f(e^{i\theta})|^2 \overline{\phi(e^{i\theta})}^n \frac{d\theta}{2\pi} = 0, \quad n \geqslant 1.$$

Of course, when $\phi(0) = 0$, then any (classical) inner function is T_ϕ inner, and this class is also closed under multiplication by classical inner functions. In general, what are the T_ϕ-inner functions?

Let us work out a particular example. Suppose that ϕ is a Riemann map from $\mathbb{D}$ onto a simply connected domain G with smooth boundary Γ. Then, with $ds_\mathbb{T}$ denoting arc length measure on $\mathbb{T}$, ds_Γ denoting arc length measure on Γ, and $\psi = \phi^{-1}$, we see, via a change of variables, that a unit vector $f \in H^2$ is T_ϕ-inner when

$$\begin{aligned} 0 &= \int_\mathbb{T} |f(\zeta)|^2 \overline{\phi(\zeta)}^n ds_\mathbb{T}(\zeta) \\ &= \int_\Gamma |f(\psi(w)|^2 |\psi'(w)| \overline{w}^n ds_\Gamma(w), \quad n \geqslant 1. \end{aligned}$$

Using the (harmless) assumption that f is a unit vector, we see that

$$\int_\Gamma (|f(\psi(w)|^2 |\psi'(w)| - 1) \overline{w}^n ds_\Gamma(w) = 0, \quad n \geqslant 0.$$

Taking the complex conjugate of the above expression we see the measure

$$(|f \circ \psi|^2 |\psi'| - 1) ds_\Gamma$$

annihilates w^n and $\overline{w}^n$ for all $n \geqslant 0$. Standard harmonic analysis will show that this measure must be the zero measure and so

$$|f \circ \psi|^2 |\psi'| = 1$$

almost everywhere on Γ. Consequently, we see that

$$|f|^2 |\psi' \circ \phi| = 1$$

almost everywhere on $\mathbb{T}$. But since

$$\psi' \circ \phi = \frac{1}{\phi'}$$

we see that $f/\sqrt{\phi'}$ is a classical inner function. In summary, f is T_ϕ-inner if and only if $f/\sqrt{\phi'}$ is a classical inner function. We thank Dima Khavinson for pointing this out to us.

For a particularly simple example, consider the case where

$$\phi(z) = \frac{z-w}{1-\overline{w}z}, \quad w \in \mathbb{D}.$$

Here ϕ is a simple Blaschke factor (which is an automorphism of $\mathbb{D}$). Since

$$\phi'(z) = \frac{1-|w|^2}{(1-\overline{w}z)^2},$$

the T_ϕ inner functions in this case take the form

$$C\frac{j(z)}{1-\overline{w}z},$$

where $C \in \mathbb{C}$ and j is a classical inner function.

Example 2.7 If $(Tf)(x) = xf(x)$ on $L^2[0,1]$, it is an easy exercise to show that there are no (nonzero) T-inner vectors. Indeed, if

$$\langle f, x^n f\rangle = \int_0^1 x^n |f(x)|^2 dx = 0, \quad n \geqslant 1,$$

then all the polynomials annihilate the measure $x|f(x)|^2dx$ and an argument using the Weierstrass approximation and the Riesz representation theorems will show that $f = 0$ (almost everywhere).

Example 2.8 Let

$$(Tf)(x) = \int_0^x f(t)dt,$$

be the *Volterra operator* on $L^2[0,1]$. Let us establish that there are no nonzero T-inner vectors. By a well-known result [25], every invariant subspace of the Volterra operator takes the form $\chi_{[a,1]}L^2[0,1]$ for some $a \in [0,1]$. Thus

$$[Tf] = \chi_{[a,1]}L^2[0,1]$$

for some $a \in [0,1]$. By the Lebesgue differentiation theorem, $f = \frac{d}{dx}Tf$ almost everywhere and so $f \in \chi_{[a,1]}L^2[0,1]$. In other words, $f \in [Tf]$, and since f is T-inner, we have $f \perp f$. This forces $f = 0$, and so there are no T-inner functions.

Example 2.9 Let T denote the *compressed shift* $Tf = P_\Theta(zf)$ on the model space $(\Theta H^2)^\perp$, where Θ is a classical inner function as in Example 2.2. These compressed shifts have been well studied and serve as models for certain types of contractions on Hilbert spaces [14, Ch, 9]. Here an $f \in (\Theta H^2)^\perp$ is T-inner when

$$\begin{aligned} 0 &= \langle f, T^n f\rangle \\ &= \langle f, P_\Theta(z^n f)\rangle \\ &= \langle P_\Theta f, z^n f\rangle \\ &= \langle f, z^n f\rangle \\ &= \int_0^{2\pi} |f(e^{i\theta})|^2 e^{-in\theta}\frac{d\theta}{2\pi}, \quad n \geqslant 1. \end{aligned}$$

As in Example 2.2, this says that f must have constant modulus on the unit circle and thus be a classical inner function. However, f must also belong to the model space $(\Theta H^2)^\perp$. This extra condition places a restriction on Θ, namely $\Theta(0) = 0$, and on f, namely f must be an inner divisor of Θ/z [14, p. 177].

Example 2.10 Continuing with Example 2.9, one can consider the special case where $\Theta(z) = z^n$, $n \geqslant 1$. Here the model space takes the form

$$(z^n H^2)^\perp = \bigvee\{1, z, z^2, \ldots, z^{n-1}\}$$

and the matrix representation of the compressed shift $Tf = P_\Theta(zf)$ with respect to the orthonormal basis $\{1, z, z^2, \ldots, z^{n-1}\}$ for $(z^n H^2)^\perp$ becomes

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

(see [14]). The powers of the above matrix just move the 1s on the sub-diagonal to the succeeding sub-diagonals (until the matrix becomes the zero matrix) and from here one can see that the T-inner vectors are $\mathbf{v} = c\,\mathbf{e}_j$, for $j = 0, 1, \ldots, n-2$, where $\mathbf{e}_j$ is the standard basis vector. Notice how this corresponds to the T-inner vectors

$$f(z) = cz^k, \quad k = 0, 1, \ldots, n-1.$$

from the previous example (the inner divisors of z^{n-1}).

Example 2.11 In the previous example if $\Theta(z) = z^4$, then the model space becomes $(z^4 H^2)^{\perp} = \bigvee\{1, z, z^2, z^3\}$ and the matrix of the compressed shift is

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

If T is the *square* of the compressed shift, then T has matrix representation

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

If $\mathbf{v} = (z_1, z_2, z_3, z_4) \in \mathbb{C}^4$, one can quickly check that $\mathbf{v}$ is T-inner if and only if

$$z_3\overline{z_1} + z_4\overline{z_2} = 0.$$

In terms of a function in the model space, this says, for example, that $f(z) = a + bz^3$ is T-inner for any $a, b \in \mathbb{C}$.

Example 2.12 Let $(Tf)(z) = zf(z)$ be the unilateral shift on the Dirichlet space $\mathscr{D}$ of analytic functions $f(z) = \sum_{n \geqslant 0} a_n z^n$ on $\mathbb{D}$ for which

$$\sum_{n \geqslant 0} (1 + n)|a_n|^2 < \infty. \tag{2.13}$$

The above quantity defines the square of the norm on $\mathscr{D}$. In [26, 27] they discussed the T-inner functions. The reproducing kernel for $\mathscr{D}$ is

$$k_w(z) = \frac{1}{\overline{w}z} \log\left(\frac{1}{1 - \overline{w}z}\right), \quad w, z \in \mathbb{D},$$

and the function

$$f(z) = k_w(w) - k_w(z)$$

is T-inner.

Example 2.14 Let $(Tf)(z) = zf(z)$ be the unilateral shift on the Bergman space $\mathscr{B}$ of analytic functions $f(z) = \sum_{n \geqslant 0} a_n z^n$ on $\mathbb{D}$ for which

$$\sum_{n \geqslant 0} \frac{|a_n|^2}{n + 1} < \infty.$$

The above quantity defines the square of the norm on $\mathscr{B}$. The T-inner functions were discussed in [1]. As in the Dirichlet space example, if

$$k_w(z) = \frac{1}{(1 - \overline{w}z)^2}$$

denotes the reproducing kernel for $\mathscr{B}$, then $k_w(w) - k_w(z)$ is a T-inner function.

Example 2.15 Consider the space H_1^2 of analytic functions $f \in H^2$ whose first derivative f' also belongs to H^2. This space, along with other associated spaces, was studied by Korenblum in [19] in his work on ideals of algebras of analytic functions. The quantity

$$|f(0)|^2 + \sum_{n \geqslant 1} n^2 |a_n|^2$$

defines the square of the norm on this space. This is a reproducing kernel Hilbert space with kernel

$$k_w(z) = 1 + \sum_{n \geqslant 1} \frac{\overline{w}^n z^n}{n^2}.$$

The shift operator $(Tf)(z) = zf(z)$ turns out to be continuous on H_1^2 and, as with previous two examples, $k_w(w) - k_w(z)$ is a T-inner function.

Example 2.16 We point out that T-inner functions for $(Tf)(z) = zf(z)$ in other weighted Hardy spaces were studied in [3].

Observe that in the four previous examples of the shift on the Dirichlet space, the Bergman space, H_1^2, and other weighted spaces, the respective T-inner functions look quite different.

3 Elementary Properties

Here are some routine but nevertheless interesting facts about T-inner vectors. Recall the definition of $[T\mathbf{v}]$ from (2.1).

Proposition 3.1 *Suppose that T is a bounded linear transformation on a Hilbert space $\mathscr{H}$ and $\mathbf{v}$ is any vector in $\mathscr{H}$. Let $P_{[T\mathbf{v}]}$ be the orthogonal projection onto the subspace $[T\mathbf{v}]$. Then the vector $\mathbf{v} - P_{[T\mathbf{v}]}\mathbf{v}$ is T-inner (or zero), and every T-inner vector arises in this way.*

Proof Observe that for any two vectors $\mathbf{u}, \mathbf{v}$ in a Hilbert space $\mathscr{H}$ we have

$$\mathbf{u} \perp \mathbf{v} \iff \|\mathbf{u} + \alpha \mathbf{v}\| \geqslant \|\mathbf{u}\|, \quad \alpha \in \mathbb{C}. \tag{3.2}$$

To see this, use the Pythagorean theorem for one direction and the definition of the orthogonal projection of $\mathbf{u}$ onto $\mathbf{v}$ for the other.

By the definition of the orthogonal projection $P_{[T\mathbf{v}]}$, we know that

$$\mathbf{v} - P_{[T\mathbf{v}]}\mathbf{v} \perp [T\mathbf{v}]$$

and so for any $n \geqslant 1$ we can use (3.2) to see that

$$\|(\mathbf{v} - P_{[T\mathbf{v}]}\mathbf{v}) - \alpha T^n(\mathbf{v} - P_{[T\mathbf{v}]}\mathbf{v})\| \geqslant \|\mathbf{v} - P_{[T\mathbf{v}]}\mathbf{v}\|, \quad \alpha \in \mathbb{C}.$$

Another application of (3.2) yields

$$\mathbf{v} - P_{[T\mathbf{v}]}\mathbf{v} \perp T^n(\mathbf{v} - P_{[T\mathbf{v}]}\mathbf{v})$$

which says that $\mathbf{v} - P_{[T\mathbf{v}]}\mathbf{v}$ is T-inner.

Now suppose that $\mathbf{v}$ is T-inner. By the definition of T-inner, $\mathbf{v} \perp \mathbf{z}$ for all $\mathbf{z} \in [T\mathbf{v}]$ which implies

$$\|\mathbf{v}\| \leqslant \|\mathbf{v} - \mathbf{z}\|, \quad \mathbf{z} \in [T\mathbf{v}].$$

By the uniqueness of $P_{[T\mathbf{v}]}\mathbf{v}$ as a vector satisfying the above inequality, we see that $P_{[T\mathbf{v}]}\mathbf{v} = \mathbf{0}$ and so the T-inner vector $\mathbf{v}$ has the desired form $\mathbf{v} = \mathbf{v} - P_{[T\mathbf{v}]}\mathbf{v}$. □

Remark 3.3 This proposition suggests a possible avenue to describe the T-inner vectors. Indeed, if $\{\mathbf{u}_1, \mathbf{u}_2, \ldots\}$ is an orthonormal basis for $[T\mathbf{v}]$, then Proposition 3.1 says that every T-inner function can be described as

$$\mathbf{v} - \sum_{j \geqslant 1} \langle \mathbf{v}, \mathbf{u}_j \rangle \mathbf{u}_j. \tag{3.4}$$

Though this approach might seem initially appealing, this is not always a tractable problem. For example, when $T = T_\phi$, $\phi \in H^\infty$ is an analytic Toeplitz operator on H^2, as in Example 2.6, the above analysis requires a description of

$$[T_\phi f] = \bigvee \{\phi f, \phi^2 f, \phi^3 f, \ldots\}$$

which can be extremely complicated.

When $\phi(z) = z$, things become much easier in that Beurling's theorem [11] says that $[zf] = zI_f H^2$, where I_f is the (classical) inner factor of f. Moreover, due to the fact that each of the functions $z^{n+1} I_f$ has unimodular boundary values, along with Beurling's theorem, the set $\{z^{n+1} I_f : n \geqslant 0\}$ is an orthonormal basis for $zI_f H^2$. Furthermore, following the formula in (3.4), we have

$$\langle f, z^{n+1} I_f \rangle = \widehat{O_f}(n+1),$$

where $\widehat{O_f}(n+1)$ is $(n+1)$st Fourier coefficient of the outer factor O_f of f. Thus we obtain the curious fact that

$$f - \sum_{n=0}^{\infty} \widehat{O_f}(n+1) z^{n+1} I_f = \widehat{O}_f(0) I_f \tag{3.5}$$

is inner (in the classical sense) for any nonzero $f \in H^2$ and moreover, any inner function arises in this fashion. Note that when f is inner then $\widehat{O_f}(n+1) = 0$ for all $n \geqslant 0$ and so the expression in (3.5) simply reduces to f. When f is outer, then $I_f = 1$ and (3.5) becomes the constant function $\widehat{O_f}(0)$ which, according to our definitions, is inner.

Proposition 3.6 *A vector* $\mathbf{v} \in \mathscr{H}$ *is* T*-inner if and only if* $\mathbf{v}$ *is* T^**-inner.*

Proof For any $n \geqslant 1$ we have

$$\langle \mathbf{v}, T^n \mathbf{v} \rangle = \langle T^{*n} \mathbf{v}, \mathbf{v} \rangle.$$

This shows that $\mathbf{v}$ is T-inner if and only if $\mathbf{v}$ is T^*-inner. □

Though the proposition above seems to be a triviality, we mention it since in the Banach space setting the T-inner vectors and the T^*-inner vectors are from different spaces (see Proposition 7.7).

4 Application: Zero Sets for Reproducing Kernel Hilbert Spaces

In exploring the zero sets of functions in the Dirichlet space $\mathscr{D}$ (recall the definition from (2.13)), Shapiro and Shields [27] constructed solutions to certain extremal problems. As a consequence of their investigations, they developed necessary and sufficient conditions on a sequence of points in $\mathbb{D}$ to be the set of zeros of a non-trivial function from $\mathscr{D}$. (Towards a Banach space generalization of this, see Sect. 7.) We now recast the Shapiro-Shields construction in the language of S-inner functions on a more general class of Hilbert spaces of analytic functions and obtain a characterization of zero sets. We will also begin to examine when these S-inner functions have extra zeros.

Suppose Ω is a bounded domain in $\mathbb{C}$ with $0 \in \Omega$. Also suppose that $\mathscr{H}$ is a Hilbert space of (scalar-valued) analytic functions on Ω satisfying the following properties:

For every nonnegative integer j, and every $w \in \Omega$, there exists a constant $C = C(j, w)$ such that

$$|f^{(j)}(w)| \leqslant C \|f\|, \quad f \in \mathscr{H}; \tag{4.1}$$

$$f \in \mathscr{H} \implies zf(z) \in \mathscr{H}; \tag{4.2}$$

$$\bigvee \{z^j : j \geqslant 0\} = \mathscr{H}; \tag{4.3}$$

$$w \in \Omega, f \in \mathscr{H} \implies (Q_w f)(z) := \frac{f(z) - f(w)}{z - w} \in \mathscr{H} \tag{4.4}$$

The first property (4.1) says that for each $w \in \Omega$, the point evaluation at w of the jth order derivative of f is continuous and so, by the Riesz representation theorem for Hilbert spaces, there is a $k_{j,w} \in \mathscr{H}$ (called a *reproducing kernel [21]* for $\mathscr{H}$) for which

$$f^{(j)}(w) = \langle f, k_{j,w} \rangle, \quad f \in \mathscr{H}.$$

When $j = 0$ we write k_w in place of $k_{0,w}$.

The closed graph theorem, together with the second property (4.2), shows that the shift operator

$$S_{\mathscr{H}} : \mathscr{H} \to \mathscr{H}, \quad (S_{\mathscr{H}} f)(z) = zf(z),$$

is well defined and continuous on $\mathscr{H}$. We included the hypothesis that Ω was a bounded domain from the beginning. However, the continuity of $S_{\mathscr{H}}$ along with the existence of reproducing kernels k_w, $w \in \Omega$, automatically gives us that Ω is a bounded domain. Indeed, it is a straightforward computation to show that

$$S_{\mathscr{H}}^* k_w = \overline{w} k_w, \quad w \in \Omega.$$

It follows that $\{\overline{w} : w \in \Omega\}$ must belong to the spectrum of $S_{\mathscr{H}}^*$, which, by basic functional analysis, is a bounded set. Thus, at the end of the day, Ω is a bounded domain anyway.

Furthermore, the list of hypotheses (4.1)–(4.4) is actually redundant in that we can deduce the first condition from the other three. To see this, let $f \in \mathscr{H}$, and $w \in \Omega$. By (4.3), $\mathscr{H}$ contains the constant function 1, and so

$$\begin{aligned} |f(w)| &= \frac{\|f(w)\|}{\|1\|} \\ &\leqslant \frac{\|f - f(w)\| + \|f\|}{\|1\|} \\ &\leqslant \frac{\|(S_{\mathscr{H}} - wI)Q_w f\| + \|f\|}{\|1\|} \\ &\leqslant \frac{\|S_{\mathscr{H}} - wI\| \|Q_w f\| + \|f\|}{\|1\|} \\ &\leqslant \frac{\|S_{\mathscr{H}} - wI\| \|Q_w\| + 1}{\|1\|} \|f\|. \end{aligned}$$

From the Taylor series of f about w, we see that

$$(Q_w f)(z) = f'(w) + \frac{f''(w)}{2!}(z-w) + \cdots .$$

This shows that $(Q_w f)(w) = f'(w)$. By the boundedness of Q_w, and of point evaluation as shown above, it must be that point evaluation at a derivative is bounded. This result extends to derivatives of all orders, and (4.1) follows.

We point out that many of the known Hilbert spaces of analytic functions (Hardy, Bergman, Dirichlet, etc.) discussed previously satisfy conditions (4.1)–(4.4).

If $(w_j)_{j\geqslant 1}$ is a sequence of points in Ω (repetitions allowed), then we say, for fixed $g \in \mathscr{H}$, that

$$Z(g) = (w_j)_{j\geqslant 1},$$

when w_j has multiplicity $r_j \geqslant 1$,

$$g(w_j) = g'(w_j) = \cdots = g^{(r_j-1)}(w_j) = 0$$

and

$$g^{(r_j)}(w_j) \neq 0$$

and

$$g(w) \neq 0 \ \text{ when } \ w \notin (w_j)_{j\geqslant 1}.$$

We say that $(w_j)_{j\geqslant 1} \subseteq \Omega$ is a *zero set* for $\mathscr{H}$ if $Z(g) \supseteq (w_j)_{j\geqslant 1}$ for some $g \in \mathscr{H} \setminus \{0\}$. Here, g may have zeros in addition to the prescribed points $(w_j)_{j\geqslant 1}$. Obviously $(w_j)_{j\geqslant 1}$ cannot be a zero set for Ω if it has an accumulation point in Ω.

Lemma 4.5 *Suppose p is a polynomial whose zeros*

$$W = \{w_1, w_2, \ldots, w_n\},$$

repeated according to their multiplicity, belong to Ω. Then

$$[p] := \bigvee\{S^j_{\mathscr{H}} p : j \geqslant 0\} = \{g \in \mathscr{H} : Z(g) \supseteq W\}.$$

Proof By property (4.1) we see that since $Z(p) = W$ then

$$\bigvee\{S^j_{\mathscr{H}} p : j \geqslant 0\} \subseteq \{g \in \mathscr{H} : Z(g) \supseteq W\}. \tag{4.6}$$

For the other inclusion, let $g \in \mathscr{H}$ with $Z(g) \supseteq W$. Observe that n applications of property (4.4) shows that $g/p \in \mathscr{H}$. Now use condition (4.3), the density of the polynomials in $\mathscr{H}$, to produce a sequence of polynomials q_n so that $q_n \to g/p$ in the norm of $\mathscr{H}$. Using the continuity of $S_{\mathscr{H}}$ (really the continuity of $p(S_{\mathscr{H}})$) we see that $pq_n \to g$ in $\mathscr{H}$. This yields $\supseteq$ in (4.6) which completes the proof. □

Sticking to the same notation as before, taking into account the multiplicities of the $w \in W$, we use the notation

$$\bigvee\{k_w : w \in W\}$$

to include the linear span of k_w along with k_{s,w_j} for $0 \leqslant s \leqslant r_w - 1$.

Lemma 4.7 *Suppose p is a polynomial whose zeros*

$$W = \{w_1, w_2, \dots, w_n\},$$

repeated according to their multiplicity, belong to Ω. Then

$$[p] = \Big(\bigvee\{k_w : w \in W\}\Big)^{\perp}.$$

Proof Suppose that

$$g \in \Big(\bigvee\{k_w : w \in W\}\Big)^{\perp}.$$

The reproducing property of the kernels k_w will show that $Z(g) \supseteq W$ and so Lemma 4.5 yields $g \in [p]$. Conversely, if $g \in [p]$ then $Z(g) \supseteq W$ and so g has zeros with at least the correct multiplicities at the $w \in W$ and so $0 = \langle g, k_w\rangle$. Thus $g \perp k_w$ for all $w \in W$ which proves the reverse inclusion. □

We now recast a result of Shapiro and Shields [27] to develop a criterion, based on $S_{\mathscr{H}}$-inner functions, for an infinite sequence $(w_j)_{j\geqslant 1} \subseteq \Omega \setminus \{0\}$ to be a zero set for $\mathscr{H}$. To this end, let

$$W_n = \{w_1, w_2, \dots, w_n\}$$

and

$$f_n(z) = \prod_{j=1}^{n}\Big(1 - \frac{z}{w_j}\Big),$$

which belongs to $\mathscr{H}$ by (4.3). Define the function

$$J_n = f_n - P_{[zf_n]} f_n,$$

where $P_{[zf_n]}$ is the orthogonal projection of $\mathscr{H}$ onto

$$[zf_n] = \bigvee\{z^j f_n : j \geqslant 1\},$$

and note that Proposition 3.1 shows that J_n is $S_{\mathscr{H}}$-inner. For notational convenience we are using $[zf_n]$ in place of the more cumbersome $[S_{\mathscr{H}} f]$.

To compute J_n somewhat explicitly, let

$$v_1, v_2, \ldots, v_n$$

denote the Gram-Schmidt normalization of the kernel functions

$$k_{w_1}, \ldots, k_{w_n},$$

where, as discussed earlier in this section, we include $k_{s,w}$ for $0 \leqslant s \leqslant r_w - 1$ if the multiplicity of w is more than one. Note that

$$\bigvee\{k_{w_j} : 1 \leqslant j \leqslant n\} = \bigvee\{v_j : 1 \leqslant j \leqslant n\}$$

and by Lemma 4.7,

$$\left(\bigvee\{v_j : 1 \leqslant j \leqslant n\}\right)^{\perp} = \{f \in \mathscr{H} : Z(f) \supseteq W_n\}.$$

Now define

$$v_0 = \frac{k_0 - \sum_{j=1}^{n} \langle k_0, v_j \rangle v_j}{\| k_0 - \sum_{j=1}^{n} \langle k_0, v_j \rangle v_j \|}.$$

Observe that $v_0 \neq 0$, since

$$k_0 \notin \bigvee\{k_{w_j} : 1 \leqslant j \leqslant n\},$$

and that

$$v_0, v_1, \ldots, v_n$$

is an orthonormal basis for

$$\bigvee\{k_0, k_{w_1}, \ldots, k_{w_n}\}.$$

By Lemmas 4.5 and 4.7,

$$\begin{aligned}\left(\bigvee\{v_j : 0 \leqslant j \leqslant n\}\right)^{\perp} &= \{g \in \mathscr{H} : Z(g) \supseteq W_n \cup \{0\}\} \\ &= [zf_n].\end{aligned}$$

Basic linear algebra shows that

$$\begin{aligned} P_{[zf_n]} f_n &= f_n - \sum_{j=0}^{n} \langle f_n, v_j \rangle v_j \\ &= f_n - \langle f_n, v_0 \rangle v_0 \end{aligned}$$

and thus

$$\begin{aligned} J_n &= f_n - P_{[zf_n]} \qquad (4.8) \\ &= \langle f_n, v_0 \rangle v_0 \\ &= \left\langle f_n, \frac{k_0 - \sum_{j=1}^{n} \langle k_0, v_j \rangle v_j}{\|k_0 - \sum_{j=1}^{n} \langle k_0, v_j \rangle v_j\|} \right\rangle \frac{k_0 - \sum_{j=1}^{n} \langle k_0, v_j \rangle v_j}{\|k_0 - \sum_{j=1}^{n} \langle k_0, v_j \rangle v_j\|} \\ &= \frac{k_0 - \sum_{j=1}^{n} \langle k_0, v_j \rangle v_j}{\|k_0 - \sum_{j=1}^{n} \langle k_0, v_j \rangle v_j\|^2}. \qquad (4.9) \end{aligned}$$

In the above calculation note the use of the facts that $\langle f_n, v_j \rangle = 0$ for all $1 \leqslant j \leqslant n$ and $\langle f_n, k_0 \rangle = f_n(0) = 1$. This says that

$$\begin{aligned} \|J_n\|^2 &= \frac{1}{\|k_0 - \sum_{j=1}^{n} \langle k_0, v_j \rangle v_j\|^2} \\ &= \frac{1}{\|k_0\|^2 - \sum_{j=1}^{n} |\langle k_0, v_j \rangle|^2}. \qquad (4.10) \end{aligned}$$

By Bessel's inequality, applied to the denominator of the expression above, we have $\|J_n\| > 1/\|k_0\|$, and that $\|J_n\|$ is a non-decreasing sequence in n.

Let Φ_n be the co-projection of k_0 onto $\{g \in \mathscr{H} : Z(g) \supseteq W_n\}$. Again, linear algebra will show that

$$\Phi_n = \sum_{j=1}^{n} \langle k_0, v_j \rangle v_j$$

and Eqs. (4.9) and (4.10) yield the identity

$$\Phi_n = k_0 - \frac{J_n}{\|J_n\|^2}.$$

By Bessel's inequality we have

$$\|\Phi_n\|^2 = \sum_{j=1}^{n} |\langle k_0, v_j \rangle|^2 \leqslant \|k_0\|^2.$$

We now present a technical lemma.

Lemma 4.11 *With the notation above, $(w_j)_{j\geqslant 1}$ is a zero set for $\mathscr{H}$ if and only if*

$$\sup\{\|\Phi_n\| : n \geqslant 1\} < \|k_0\|^2.$$

Proof Let $W = (w_j)_{j\geqslant 1}$ and

$$\mathscr{H}(W) := \{g \in \mathscr{H} : Z(g) \supseteq W\}.$$

From our previous discussions we now see that

$$\bigvee\{v_j : j \geqslant 1\} = \bigvee\{k_{w_j} : j \geqslant 1\}$$

and

$$(\bigvee\{k_{w_j} : j \geqslant 1\})^{\perp} = \mathscr{H}(W).$$

Also observe that

$$\sup\{\|\Phi_n\| : n \geqslant 1\} = \sum_{j\geqslant 1} |\langle k_0, v_j\rangle|^2 = \|k_0\|^2$$

if and only if

$$k_0 \in \bigvee\{k_{w_j} : j \geqslant 1\}$$

if and only if

$$f \in \mathscr{H}(W) \implies f(0) = 0.$$

Thus if $\mathscr{H}(W) \neq \{0\}$ then for some $n \geqslant 0$, $f(z)/z^n$ belongs to $\mathscr{H}(W)$ (note the use of property (4.4)) and does not vanish at the origin. The result now follows. □

Finally we note that $1 = J_n(0) = \langle J_n, k_0\rangle$ and so

$$\begin{aligned}\|\Phi_n\|^2 &= \langle \Phi_n, \Phi_n\rangle \\ &= \left\langle k_0 - \frac{J_n}{\|J_n\|^2}, k_0 - \frac{J_n}{\|J_n\|^2}\right\rangle \\ &= \|k_0\|^2 - \frac{1}{\|J_n\|^2}.\end{aligned}$$

Putting this all together, we obtain the identity

$$(\|k_0\|^2 - \|\Phi_n\|^2)\|J_n\|^2 = 1,$$

which means that $(w_j)_{j\geqslant 1}$ is a zero set for $\mathscr{H}$ if and only if

$$\sup\{\|J_n\| : n \geqslant 1\} < \infty.$$

This leads to the following result of Shapiro and Shields [27], expressed in terms of $S_{\mathscr{H}}$-inner functions, and extended to a wide class of reproducing kernel Hilbert spaces of analytic functions.

Theorem 4.12 *Let* $(w_j)_{j\geqslant 1} \subseteq \Omega \setminus \{0\}$ *and*

$$f_n = \prod_{j=1}^{n}\Big(1 - \frac{z}{w_j}\Big), \qquad J_n = f_n - P_{[zf_n]} f_n.$$

Then

(1) Each J_n is an $S_{\mathscr{H}}$-inner function;
(2) the sequence $\|J_n\|$ is a non-decreasing sequence;
(3) $(w_j)_{j\geqslant 1}$ is a zero sequence for $\mathscr{H}$ if and only if

$$\sup\{\|J_n\| : n \geqslant 1\} < \infty.$$

Example 4.13 Suppose $\mathscr{H} = H^2$. A result of Takenaka [14, p. 120] shows that if w_j are the proposed zeros, then the Gram-Schmidt process applied to the first n Cauchy kernels $k_{w_1}, \dots, k_{w_n}$ yields

$$v_1 = \frac{\sqrt{1-|w_1|^2}}{1-\overline{w_1}z};$$

$$v_2 = \frac{\sqrt{1-|w_2|^2}}{1-\overline{w_2}z}\frac{w_1 - z}{1-\overline{w_1}z};$$

$$v_3 = \frac{\sqrt{1-|w_3|^2}}{1-\overline{w_3}z}\frac{w_1 - z}{1-\overline{w_1}z}\frac{w_2 - z}{1-\overline{w_2}z};$$

and so on. The condition to be a zero set is then

$$\sup\{\|J_n\| : n \geqslant 1\} < \infty$$

which, by the previous analysis, translates to

$$\inf\left\{\|k_0\|^2 - \sum_{j=1}^{n} |\langle k_0, v_j\rangle|^2 : n \geqslant 1\right\} > 0.$$

A calculation shows that

$$|\langle k_0, v_j\rangle|^2 = (1 - |w_j|^2) \prod_{i=1}^{j-1} |w_i|^2.$$

Furthermore, by telescoping series,

$$\|k_0\|^2 - \sum_{j=1}^{n} |\langle k_0, v_j\rangle|^2 = \prod_{j=1}^{n} |w_j|^2.$$

Thus we have

$$\inf\left\{\|k_0\|^2 - \sum_{j=1}^{n} |\langle k_0, v_j\rangle|^2 : n \geqslant 1\right\} = \inf\left\{\prod_{j=1}^{n} |w_j|^2 : n \geqslant 1\right\}$$

and the above infimum being positive is equivalent to the standard Blaschke condition

$$\sum_{j\geqslant 1} \big(1 - |w_j|\big) < \infty.$$

This confirms that the nontrivial zero sets of H^2 are exactly the Blaschke sequences.

Example 4.14 Let us compute the $S_{\mathscr{H}}$-inner function J corresponding to a one point zero set. Suppose that $\mathscr{H}$ is a reproducing kernel Hilbert space satisfying our assumptions and

$$f(z) = 1 - \frac{z}{w}, \qquad w \in \Omega \setminus \{0\}.$$

Following the procedure in the derivation of Theorem 4.12, we define

$$v_w(z) = \frac{k_w(z)}{\sqrt{k_w(w)}},$$

the normalized reproducing kernel at w. By the formula (4.9) for J (the inner function corresponding to f) we have

$$\begin{aligned} J &= \frac{k_0 - \langle k_0, v_w\rangle v_w}{\|k_0\|^2 - |\langle k_0, v_w\rangle|^2} \\ &= \frac{k_0 - \overline{v_w(0)} v_w}{k_0(0) - |v_w(0)|^2} \end{aligned}$$

$$= \frac{k_0 - \frac{\overline{k_w(0)}}{\sqrt{k_w(w)}} \frac{k_w}{\sqrt{k_w(w)}}}{k_0(0) - \frac{|k_w(0)|^2}{k_w(w)}}$$

$$= \frac{k_w(w)k_0 - \overline{k_w(0)}k_w}{k_w(w)k_0(0) - |k_w(0)|^2}. \tag{4.15}$$

Any nonzero constant multiple of an $S_{\mathscr{H}}$-inner function is also $S_{\mathscr{H}}$-inner, and so

$$k_w(w)k_0 - \overline{k_w(0)}k_w$$

is always an $S_{\mathscr{H}}$-inner function.

In the H^2 case we have

$$k_w(z) = \frac{1}{1 - \overline{w}z}$$

and so (4.15) yields

$$J = \frac{1}{w} \frac{z - w}{1 - \overline{w}z},$$

which, as expected by classical theory, is a constant multiple of a Blaschke factor.

In the Dirichlet space case, the reproducing kernel is

$$k_w(z) = \frac{1}{\overline{w}z} \log\left(\frac{1}{1 - \overline{w}z}\right)$$

and (4.15) yields

$$J(z) = \frac{\log(1 - |w|^2) - \frac{w}{z} \log(1 - \overline{w}z)}{\log(1 - |w|^2) - |w|^2}.$$

In the Bergman space $\mathscr{B}^2$, we have

$$k_w(z) = \frac{1}{(1 - \overline{w}z)^2}$$

and (4.15) yields

$$J = \frac{1 - \frac{(1-|w|^2)^2}{(1-\overline{w}z)^2}}{1 - (1 - |w|^2)^2}.$$

Notice the concept of "inner" yields different types of functions in each Hardy, Dirichlet, and Bergman setting. In the above analysis we see that the expression

$$k_w(w)k_0 - \overline{k_w(0)}k_w \tag{4.16}$$

is always an $S_{\mathscr{H}}$-inner function. This can also be verified directly from the calculation

$$\langle k_w(w)k_0 - \overline{k_w(0)}k_w, S^n_{\mathscr{H}}(k_w(w)k_0 - \overline{k_w(0)}k_w)\rangle = 0, \quad n \geqslant 1.$$

These next two results provide an interesting link between the zero set for $\mathscr{H}$ and the property that $\|S_{\mathscr{H}}\|$ or $\|Q_0\| = 1$.

Theorem 4.17 *Suppose that $\mathscr{H}$ is a RKHS of analytic functions on $\mathbb{D}$ satisfying conditions* (4.1)–(4.4)*. If $\|S_{\mathscr{H}}\| \leqslant 1$, then the union of a zero set with a Blaschke sequence is again a zero set for $\mathscr{H}$.*

Proof For notational convenience let $S = S_{\mathscr{H}}$. First, suppose that J is S-inner, and $w \in \mathbb{D} \setminus \{0\}$. Since J is S-inner, we have $J \perp S^k J$ for all $k \geqslant 1$. Let

$$F(z) = \sum_{k=0}^{d} F_k z^k$$

be any polynomial with $F_0 = 1$. By the linearity of $\perp$ (in the second slot) in a Hilbert space, and the Pythagorean Theorem,

$$\begin{aligned}
\|JF\|^2 &= \|J + F_1 SJ + F_2 S^2 J + \cdots + F_d S^d J\|^2 \\
&= \|J\|^2 + \|F_1 SJ + F_2 S^2 J + \cdots + F_d S^d J\|^2 \\
&\leqslant \|J\|^2 + \|S\|^2 \|F_1 J + F_2 SJ + \cdots + F_d S^{d-1} J\|^2 \\
&\leqslant \|J\|^2 + \|S\|^2 |F_1|^2 \|J\|^2 + \|S\|^2 \|F_2 SJ + \cdots + F_d S^{d-1} J\|^2 \\
&\leqslant \cdots \\
&\leqslant \|J\|^2 \big(1 + |F_1|^2 \|S\|^2 + |F_2|^2 \|S\|^{2\cdot 2} + \cdots + |F_d|^2 \|S\|^{2\cdot d}\big) \\
&= \|J\|^2 \big(1 + |F_1|^2 + |F_2|^2 + \cdots + |F_d|^2\big).
\end{aligned}$$

The final expression in parentheses is the square of the norm in ℓ^2_A of $F(z)$. The inequality remains true if F is the Blaschke factor that vanishes at w, normalized so that $F(0) = 1$, i.e.,

$$F(z) = \frac{1}{w} \frac{w - z}{1 - \overline{w} z}.$$

This function has norm in $\ell^2_A = H^2$ given by $1/|w|$.

Now let W be any zero set for $\mathscr{H}$, and let $\{w_1, w_2, w_3, \ldots\} \in \mathbb{D} \setminus \{0\}$. Let J_W be the S-inner function associated with W with $J(0) = 1$, i.e., $J = f - \widehat{f}$, where

$f \in \mathscr{H}$ and f has zeros W (according to multiplicity) and $f(0) = 1$. By repeated application of the above argument, we find that

$$\|J_{W\cup\{w_1,w_2,\ldots,w_m\}}\| \leqslant \frac{\|J_W\|}{|w_1 w_2 \cdots w_m|}, \quad m \geqslant 1.$$

This, in conjunction with Theorem 4.12, proves the assertion. □

Theorem 4.18 *Suppose that $\mathscr{H}$ is a RKHS of analytic functions on $\mathbb{D}$ satisfying conditions (4.1)–(4.4). If $\|Q_0\| = 1$, J is an S-inner function with zero set W, and $f \in [J] \setminus \{0\}$, then the zero set for f is the union of W and a Blaschke sequence.*

Proof For any $g \in \mathscr{H}$ observe that

$$g = Q_0 S g$$

and so, since $\|Q_0\| = 1$ by assumption,

$$\|g\| \leqslant \|Q_0\|\|Sg\| = \|Sg\|.$$

Apply this identity k times to get

$$\|S^k g\| \geqslant \|g\|, \quad k \geqslant 1. \tag{4.19}$$

Suppose that $f \in [J]$. By the inner property of J, along with repeated use of (4.19),

$$\begin{aligned}
\|JF\|^2 &= \|JF_0\|^2 + \|SJF_1 + S^2JF_2 + \cdots\|^2 \\
&= \|JF_0\|^2 + \|S(JF_1 + SJF_2 + \cdots)\|^2 \\
&\geqslant \|JF_0\|^2 + \|JF_1\|^2 + \|SJF_2 + \cdots)\|^2 \\
&\;\;\vdots \\
&= \|J\|^2\big(|F_0|^2 + |F_1|^2 + |F_2|^2 + \cdots\big)
\end{aligned}$$

for any polynomial F. The bound is true for any sequence of polynomials F_m such that JF_m tends to f in $\mathscr{H}$. This tells us that f is the product of J and a function in H^2. The claim follows. □

5 Zeros of S-Inner Functions

In the Hardy space H^2, we know that when $f \in H^2 \setminus \{0\}$, the classical inner part I_f of f takes the form $I_f = BS_\mu$, where B is the Blaschke product and S_μ is an inner function. The Blaschke factor contains all the zeros of f in $\mathbb{D}$ (and no others) while the inner factor S_μ has no zeros in $\mathbb{D}$. This means that the inner factor I_f has *precisely* the same zeros as f (counting multiplicity). How ubiquitous is this phenomenon? In other words, if $\mathscr{H}$ is a Hilbert space of analytic functions satisfying conditions (4.1)–(4.4) and $f \in \mathscr{H} \setminus \{0\}$, does the $S_{\mathscr{H}}$-inner function

$$J = f - P_{[zf]}f$$

have any "extra" zeros inside $\mathbb{D}$? Certainly J has *at least* the zeros of f. Does it have any others? A result of Hedenmalm and Zhu shows that in the weighted Bergman space of analytic functions f on $\mathbb{D}$ for which $f \in L^2((1-|z|^2)^\alpha dA)$, where dA is planar Lebesgue measure, it is possible, when $\alpha > 4$, for the inner function J corresponding to the linear function $f(z) = 1 - z/w$ to have an extra zero in $\mathbb{D}$. So, indeed, the "no extra zeros" property for $S_{\mathscr{H}}$-inner functions is not ubiquitous. In this section we obtain lower bounds for these extra zeros and show that they must lie somewhat close to the boundary. Moreover, we will see that in some situations such extra zeros do not exist at all.

From condition (4.4), we know that for each $w \in \Omega$, the operator

$$Q_w : \mathscr{H} \to \mathscr{H}, \quad Q_w f(z) = \frac{f(z) - f(w)}{z - w}$$

is well defined and continuous. Our criterion that the $S_{\mathscr{H}}$-inner function J has no extra zeros will be stated in terms of the norm of the operator Q_0. This operator

$$(Q_0 f)(z) = \frac{f(z) - f(0)}{z}$$

is often called the *backward shift operator* since if $\Omega = \mathbb{D}$, then Q_0 acts on the Taylor series of f (about the origin) by shifting all of the coefficients backwards and dropping the constant term, i.e.,

$$Q_0(a_0 + a_1 z + a_2 z^2 + \cdots) = a_1 + a_2 z + a_3 z^2 + \cdots .$$

Theorem 5.1 *Let $f \in \mathscr{H} \setminus \{0\}$, and let $J = f - P_{[zf]}f$ be the $S_{\mathscr{H}}$-inner function corresponding to f. If $w \in \Omega \setminus \{0\}$ is a zero of J that is not a zero of f, then*

$$|w| \geqslant \frac{[1 + \|S_{\mathscr{H}}\|^2 \|Q_w\|^2]^{1/2}}{\|Q_0\| \|S_{\mathscr{H}}\| \|Q_w\|}.$$

Towards the proof of this theorem, we start with the following.

Proposition 5.2 *Let $f \in \mathscr{H} \setminus \{0\}$ and let $J = f - P_{[zf]}f$. If w is a zero of J that is not a zero of f, then $Q_w J \in [f]$.*

Proof By hypothesis, there are polynomials ϕ_n such that $\phi_n f$ converges in norm to J. It follows that $Q_w(\phi_n f)$ converges in norm to $Q_w J$, i.e.,

$$\frac{\phi_n(z)f(z) - \phi_n(w)f(w)}{z - w} \longrightarrow Q_w J.$$

Since evaluation at w is bounded, we may further conclude that

$$\phi_n(w)f(w) \to J(w) = 0,$$

and hence $Q_w J$ is the limit in norm of

$$\begin{aligned}
&\frac{\phi_n(z)f(z) - \phi_n(w)f(w)}{z - w}\\
&= \frac{\phi_n(z)f(z) - \phi_n(w)f(z) + \phi_n(w)f(z) - \phi_n(w)f(w)}{z - w}\\
&= \frac{\phi_n(z) - \phi_n(w)}{z - w} f(z) + \frac{f(z) - f(w)}{z - w}\phi_n(w).
\end{aligned}$$

The last term above tends to zero which says that $Q_w J \in [f]$. □

Proof of Theorem 5.1 Observe that

$$\begin{aligned}
\left\| \frac{J(z)}{1 - \frac{z}{w}} \right\|^2 &= \left\| \frac{J(z)}{1 - \frac{z}{w}} \left(1 - \frac{z}{w} + \frac{z}{w}\right) \right\|^2\\
&= \left\| J(z) + \frac{z}{w} \frac{J(z)}{1 - \frac{z}{w}} \right\|^2.
\end{aligned}$$

Now apply Proposition 5.2 and the Pythagorean Theorem to get

$$\begin{aligned}
\left\| \frac{J(z)}{1 - \frac{z}{w}} \right\|^2 &= \|J(z)\|^2 + \left\| \frac{z}{w} \frac{J(z)}{1 - \frac{z}{w}} \right\|^2\\
\left\| Q_0\left(\frac{zJ(z)}{1 - \frac{z}{w}}\right) \right\|^2 &= \|J(z)\|^2 + \frac{1}{|w|^2} \left\| \frac{zJ(z)}{1 - \frac{z}{w}} \right\|^2\\
\left(\|Q_0\|^2 - \frac{1}{|w|^2}\right) \left\| \frac{zJ(z)}{1 - \frac{z}{w}} \right\|^2 &\geqslant \|J(z)\|^2\\
\left(\|Q_0\|^2 - \frac{1}{|w|^2}\right) \|S_{\mathscr{H}}\|^2 \|Q_w\|^2 |w|^2 \|J(z)\|^2 &\geqslant \|J(z)\|^2\\
\left(\|Q_0\|^2 - \frac{1}{|w|^2}\right) \|S_{\mathscr{H}}\|^2 \|Q_w\|^2 |w|^2 &\geqslant 1
\end{aligned}$$

$$|w|^2 \geqslant \frac{1 + \|S_{\mathscr{H}}\|^2 \|Q_w\|^2}{\|Q_0\|^2 \|S_{\mathscr{H}}\|^2 \|Q_w\|^2}.$$

□

As a corollary to this theorem we note that if Q_0 is contractive, and $\Omega = \mathbb{D}$, then J will have no extra zeros.

Corollary 5.3 *Let $\mathscr{H}$ be a RKHS of analytic functions on $\mathbb{D}$. If Q_0 is contractive, then the $S_{\mathscr{H}}$-inner function J corresponding to f will have no extra zeros.*

Proof If $\|Q_0\| \leqslant 1$, then

$$\frac{[1 + \|S_{\mathscr{H}}\|^2 \|Q_w\|^2]^{1/2}}{\|Q_0\| \|S_{\mathscr{H}}\| \|Q_w\|} \geqslant \frac{[1 + \|S_{\mathscr{H}}\|^2 \|Q_w\|^2]^{1/2}}{\|S_{\mathscr{H}}\| \|Q_w\|} \geqslant 1.$$

By Theorem 5.1, any extra zero $w \in \mathbb{D}$ must satisfy $|w| \geqslant 1$. □

It is easy to see that for the Hardy space H^2, the operator Q_0 (which is just the well-known backward shift operator) satisfies $\|Q_0\| = 1$ and so the S-inner function J of corresponding to f, which in this case is the classical inner factor of f, never has extra zeros. Slightly more work is that on the Dirichlet space $\mathscr{D}$ (See Example 2.12), the operator Q_0 also has norm equal to one [25]. This gives us the following.

Corollary 5.4 *For any $f \in \mathscr{D}$, the corresponding $S_{\mathscr{D}}$-inner function J has no extra zeros in $\mathbb{D}$.*

We point out here that this result, in a way, is known. As shown in [23], every shift invariant subspace M of the Dirichlet space has the property that $M \ominus zM = \mathbb{C}\phi$ and this function generates M, in that $\bigvee\{\phi, z\phi, z^2\phi, \ldots\} = M$. Applying this fact to a vector $f \in \mathscr{D}$ and $M = [f]$, we see that $[J] = [f]$ and so J cannot have any extra zeros.

For the Bergman space $\mathscr{B}$ from Example 2.14, Q_0 has norm $\sqrt{2}$ and so we are unable to apply Corollary 5.3. However, it is known, for different reasons [1], that J has no extra zeros. On the other hand, for the space H_1^2 from Example 2.15, one can quickly check (using power series) that Q_0 is contractive on H_1^2 and thus we have the following.

Corollary 5.5 *For any $f \in H_1^2$, the corresponding inner function J has no extra zeros in $\mathbb{D}$.*

6 Extra Zeros Abound

In the previous section it was shown that if an S-inner function J corresponding to a given function f has extra zeros, then those extra zeros must be bounded away from the origin. When $\Omega = \mathbb{D}$, this gave rise to a sufficient condition on the space

$\mathscr{H}$ for the S-inner functions to have no extra zeros. In the present section we shall see that extra zeros are nonetheless quite abundant. A large class of spaces will be constructed for which certain S-inner functions will have extra zeros.

We begin by presenting another description of the zero sets for a RKHS $\mathscr{H}$ satisfying our hypotheses. This description is due to Shapiro and Shields [27].

Let $W_n := \{w_1, w_2, \ldots, w_n\}$, $w_j \in \Omega \setminus \{0\}$, and define

$$f_n(z) = \Big(1 - \frac{z}{w_1}\Big)\Big(1 - \frac{z}{w_2}\Big) \cdots \Big(1 - \frac{z}{w_n}\Big),$$

and $J_n = f_n - P_{[zf_n]} f_n$. For notational simplicity, let k_j be the reproducing kernel for w_j and let k_0 be the reproducing kernel at the origin.

From (4.9) we know that J_n has the representation

$$J_n(z) = c_{n,0}k_0 + c_{n,1}k_1 + c_{n,2}k_2 + \cdots + c_{n,n}k_n, \tag{6.1}$$

where the coefficients $c_{n,j}$ are uniquely determined by the conditions

$$J_n(w_1) = J_n(w_2) = \cdots = J_n(w_n) = 0, \quad J_n(0) = 1.$$

Indeed, the coefficients are the unique solutions to the matrix equation

$$\begin{bmatrix} G_{0,0} & G_{0,1} & G_{0,2} & \ldots & G_{0,n} \\ G_{1,0} & G_{1,1} & G_{1,2} & \ldots & G_{1,n} \\ G_{2,0} & G_{2,1} & G_{2,2} & \ldots & G_{2,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ G_{n,0} & G_{n,1} & G_{n,2} & \ldots & G_{n,n} \end{bmatrix} \begin{bmatrix} c_{n,0} \\ c_{n,1} \\ c_{n,2} \\ \vdots \\ c_{n,n} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

where

$$G = G^{(n)} = G[k_0, k_1, k_2, \ldots, k_n]$$

is the Gramian matrix for the vectors $k_0, k_1, k_2, \ldots, k_n$, and

$$G_{s,t} := \langle k_t, k_s \rangle.$$

Since a finite set of reproducing kernels is linearly independent, the Gramian determinant is nonzero, and hence the matrix G is invertible, guaranteeing a unique solution for the coefficients.

Continuing from the above equation, we can write

$$J_n(z) = \left[\, k_0(z)\ k_1(z)\ k_2(z)\ \cdots\ k_n(z) \,\right] \begin{bmatrix} c_{n,0} \\ c_{n,1} \\ c_{n,2} \\ \cdots \\ c_{n,n} \end{bmatrix}$$

$$= \left[\, k_0(z)\ k_1(z)\ k_2(z)\ \cdots\ k_n(z) \,\right] G^{(n)^{-1}} \begin{bmatrix} 1 \\ 0 \\ 0 \\ \cdots \\ 0 \end{bmatrix}$$

$$= \left[A_{0,0}k_0(z) - A_{0,1}k_1(z) + \cdots + (-1)^n A_{0,n}k_n(z)\right] / \det G^{(n)},$$

where $A_{m,n}$ is the (m,n)th cofactor of $G^{(n)}$. But the last quantity in square brackets is itself the determinant of a certain matrix, yielding

$$J_n(z) \det G^{(n)} = \det \begin{bmatrix} k_0(z) & k_1(z) & k_2(z) & \ldots & k_n(z) \\ G_{1,0} & G_{1,1} & G_{1,2} & \ldots & G_{1,n} \\ G_{2,0} & G_{2,1} & G_{2,2} & \ldots & G_{2,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ G_{n,0} & G_{n,1} & G_{n,2} & \ldots & G_{n,n} \end{bmatrix} \tag{6.2}$$

Let

$$d_n := \inf \left\| k_0 - (c_1k_1 + c_2k_2 + \cdots c_nk_n) \right\| \tag{6.3}$$

where the infimum is over the coefficients $c_1, c_2, \ldots, c_n$. It is well known that

$$d_n^2 = \frac{\det G[k_0, k_1, k_2, \ldots, k_n]}{\det G[k_1, k_2, \ldots, k_n]}. \tag{6.4}$$

A proof of this appears in [13, Lemma 4.2.4].

Furthermore, Oppenheim's inequality (see, for example, [17]) tells us that for nonnegative definite square matrices $A = (a_{s,t})$ and $B = (b_{s,t})$, the Hadamard product of A and B, i.e., $(a_{s,t}b_{s,t})$, satisfies

$$\det[a_{s,t}b_{s,t}] \geqslant \left(\det[a_{s,t}]\right)\left(\prod_t b_{t,t}\right). \tag{6.5}$$

This enables us to derive, as was done in [27], the following sufficient condition for a zero set of $\mathscr{H}$ (see also [13] for an exposition of this).

Theorem 6.6 *Let* $\{w_1, w_2, w_3, \ldots\} \subseteq \mathbb{D} \setminus \{0\}$ *be a sequence of distinct points. If the matrix*

$$\left[1 - k_0(w_s)\overline{k_0(w_t)}/k_{w_t}(w_s)k_0(0)\right]_{1 \leqslant s,t \leqslant n} \tag{6.7}$$

is nonnegative definite for all $n \geqslant 1$, *and*

$$\inf_n \prod_{m=1}^{n} \left[1 - \frac{|k_0(w_m)|^2}{k_{w_m}(w_m)k_0(0)}\right] > 0, \tag{6.8}$$

then there exists a nonzero $f \in \mathscr{H}$ *such that* $f(w_n) = 0$ *for all* $n \geqslant 1$.

Proof By (4.10) and Theorem 4.12, it is enough to show that the quantity d_n from (6.3) satisfies $\inf d_n > 0$. Let us examine $\det G^{(n)}$, with a view towards applying (6.4). This determinant is unchanged if the multiple of any row is added to a different row. Suppose that $G_{s,0}/G_{0,0}$ times the 0th row (the rows and columns are indexed from 0 to n) is added to the sth row, for all $1 \leqslant s \leqslant n$. The result is that

$$\begin{aligned}
\det G^{(n)} &= G_{0,0} \det \left[G_{s,t} - G_{s,0}G_{0,t}/G_{0,0}\right]_{1 \leqslant s,t \leqslant n} \\
&= G_{0,0} \det \left[G_{s,t}(1 - G_{s,0}G_{0,t}/G_{s,t}G_{0,0})\right]_{1 \leqslant s,t \leqslant n} \\
&\geqslant G_{0,0}\Big(\det G[k_1, k_2, \ldots, k_n]\Big)\Big(\prod_{t=1}^{n}[1 - G_{t,0}G_{0,t}/G_{t,t}G_{0,0}]\Big),
\end{aligned}$$

where in the last step we applied (6.5). The claim now follows from invoking (6.4), and writing out $G_{t,t}$ in terms of the kernel functions. □

Example 6.9 When $\mathscr{H} = H^2$, the matrix in (6.7) takes the form

$$[\bar{w}_t w_s]_{1 \leqslant s,t \leqslant n}\,,$$

which is obviously positive definite.

Now the zero set criterion (6.8) is

$$\begin{aligned}
0 &< \inf_n \prod_{m=1}^{n} \left[1 - \frac{|k_0(w_m)|^2}{k_{w_m}(w_m)k_0(0)}\right] \\
&= \inf_n \prod_{m=1}^{n} [1 - \frac{1}{1/(1 - |w_m|^2)}] \\
&= \inf_n \prod_{m=1}^{n} |w_m|^2.
\end{aligned}$$

Of course, this is equivalent to the Blaschke condition.

In [27] the zero sets of functions in the Dirichlet space $\mathscr{D}$ (and other related spaces) were discussed. The Dirichlet space can be viewed as the weighted ℓ^2 space with weights $1, 2, 3, \ldots, n+1, \ldots$. We now construct a large class of such weighted spaces for which the corresponding matrices 6.7 are nonnegative definite, and hence lie within the scope of Theorem 6.6.

Example 6.10 Fix $\Omega = \mathbb{D}$, and let $w_1, w_2, w_3, \ldots$ be a sequence of distinct nonzero points in $\mathbb{D}$. Suppose that $\Lambda := (\lambda_n)_{n\geqslant 0}$ is a sequence of positive numbers with $\lambda_0 = 1$, and define $\ell^2(\Lambda)$ to be the Hilbert space of sequences $f = (f_n)_{n\geqslant 0}$ such that

$$\|f\| = \Big(\sum_{n=0}^{\infty} |f_n|^2\lambda_n\Big)^{1/2} < \infty.$$

Provided that the weights λ_n do not decay to zero too rapidly, each member of $\ell^2(\Lambda)$ can be identified with the analytic function

$$f(z) = \sum_{n=0}^{\infty} f_n z^n$$

on $\mathbb{D}$. (For example, if the weights decay exponentially, then $\ell^2(\Lambda)$ will contain some coefficient sequences that increase exponentially; such functions will not necessarily be analytic in all of $\mathbb{D}$.) The reproducing kernel function

$$k_w(z) := \sum_{n=0}^{\infty} \frac{(\bar{w}z)^n}{\lambda_n}$$

implements point evaluation at $w \in \mathbb{D}$. Again, if the weights λ_n do not decay too rapidly, the kernel function will be analytic in $\mathbb{D}$. Notice that point evaluation at the origin corresponds to the constant kernel 1.

Let us determine sufficient conditions on the sequence Λ of weights for the matrix in (6.7) to be nonnegative definite. We claim that for any $a > 0$, and positive integers m and n, the matrix

$$M := \big[a(\bar{w}_t w_s)^m\big]_{1\leqslant s,t\leqslant n}$$

is nonnegative definite. This is because

$$C^* M C = a\big|c_1 w_1^m + c_2 w_2^m + \cdots + c_n w_n^m\big|^2 \geqslant 0$$

for any column vector C with $C^* = [\bar{c}_1\ \bar{c}_2\ \ldots \bar{c}_n]$. For n fixed the sum of any such matrices is also nonnegative definite. In particular, if $(a_m)_{m\geqslant 1}$ is a sequence of nonnegative numbers with $a_1 > 0$ and $\sum_{m=1}^{\infty} a_m \leqslant 1$, the matrix

$$\left[\sum_{m=1}^{\infty} a_m (\bar{w}_t w_s)^m\right]_{1 \leqslant s,t \leqslant n}$$

is nonnegative definite.

It is clear that the function of z defined by

$$\Phi(z) := \frac{1}{1 - \sum_{m=1}^{\infty} a_n z^m}$$

is analytic in $\mathbb{D}$, and has a convergent power series

$$\Phi(z) = 1 + \sum_{n=1}^{\infty} b_n z^n \tag{6.11}$$

in $\mathbb{D}$. By expressing Φ as the geometric series

$$\Phi(z) = 1 + \Big(\sum_{m=1}^{\infty} a_n z^m\Big) + \Big(\sum_{m=1}^{\infty} a_n z^m\Big)^2 + \cdots,$$

and using the assumption that $a_1 > 0$, we find that each b_n is positive (see also the Kaluza lemma [13, p. 69]).

Thus, with the identification $b_n = 1/\lambda_n$ for all $n \geqslant 1$, the function $\Phi(\bar{w}z)$ is the reproducing kernel in the weighted space $\ell^2(\Lambda)$ for $w \in \mathbb{D}$. It then follows that

$$\begin{aligned}
1 - k_0(w_s)\overline{k_0(w_t)}/k_{w_t}(w_s)k_0(0) &= 1 - 1/k_{w_t}(w_s) \\
&= 1 - \Big(1 + \sum_{n=1}^{\infty} b_n (\bar{w}_t w_s)^n\Big)^{-1} \\
&= 1 - \Big(1 - \sum_{m=1}^{\infty} a_n (\bar{w}_t w_s)^m\Big) \\
&= \sum_{m=1}^{\infty} a_n (\bar{w}_t w_s)^m
\end{aligned}$$

That is to say, for the weighted space $\ell^2(\Lambda)$, the matrix in (6.7) is nonnegative definite.

According to Theorem 6.6, a sequence $w_1, w_2, w_3, \ldots$ of distinct nonzero points of $\mathbb{D}$ is the zero set of some nontrivial function $f \in \ell^2(\Lambda)$ if

$$\inf_{n \geqslant 1} \prod_{m=1}^{n} \left[1 - \frac{1}{1 - \sum_{j=1}^{\infty} a_j |w_m|^{2j}} \right] > 0.$$

This provides a sufficient condition for a zero set of $\ell^2(\Lambda)$.

Example 6.12 With the definitions of Example 6.10, it was shown in [27] that if the sequence $(b_n)_{n \geqslant 0}$ (the reciprocals of the weights of the space $\ell^2(\Lambda)$) satisfies

$$b_n^2 \leqslant b_{n+1} b_{n-1}$$

for all $n \geqslant 1$, then the matrices given by 6.7 are nonnegative definite, and thus Theorem 6.6 applies. This class of examples includes the Dirichlet space $\mathscr{D}$.

Here is another way to see how extra zeros may arise. Recall the formula from (6.1)

$$J_n(z) = c_{n,0} k_0 + c_{n,1} k_1 + c_{n,2} k_2 + \cdots + c_{n,n} k_n$$

for expressing the $S_{\mathscr{H}}$-inner function of a finite zero set in terms of the corresponding kernel functions.

Lemma 6.13 *The $S_{\mathscr{H}}$-inner function J_{n-1} has an extra zero at the point w_n if and only if the coefficient $c_{n,n}$ vanishes.*

Proof Suppose that $c_{n,n} = 0$. Then J_n has the following properties:

$$J_n(0) = 1,\ J_n(w_1) = \cdots = J_n(w_{n-1}) = 0,$$

and

$$\langle z^m f_{n-1}, J_n \rangle = 0, \quad m \geqslant 1,$$

This forces the identification $J_n = J_{n-1}$. Since $J_{n-1}(w_n) = 0$, it can be said that w_n is an extra zero of J_{n-1}.

Conversely suppose that w_n is an extra zero of J_{n-1}. First, for any $m \geqslant 1$,

$$\begin{aligned} \langle z^m f_n(z), J_{n-1} \rangle &= \langle z^m f_{n-1}(z), J_{n-1} \rangle - \langle z^{m+1} (1/w_n) f_{n-1}(z), J_{n-1} \rangle \\ &= 0. \end{aligned}$$

Furthermore,

$$J_{n-1}(w_1) = J_{n-1}(w_2) = \cdots = J_{n-1}(w_n) = 0, \quad J_{n-1}(0) = 1.$$

This implies that $J_{n-1} = J_n$. Since the representations (6.1) are unique, it must be that $c_{n,n} = 0$. □

Let us calculate $c_{n,n}$. Let H be the matrix $G^{(n)}$ with its nth column (the columns are indexed 0 through n) replaced by $[1\ 0\ 0\ \dots\ 0]^T$. By Cramer's Rule,

$$c_{n,n} = \frac{\det H}{\det G^{(n)}}.$$

Since the last column of H is such a special form, taking the determinant of H results in $(-1)^n$ times the determinant of the following submatrix of $G^{(N)}$:

$$R := \begin{bmatrix} G_{1,0} & G_{1,1} & G_{1,2} & \dots & G_{1,n-1} \\ G_{2,0} & G_{2,1} & G_{2,2} & \dots & G_{2,n-1} \\ \dots & \dots & \dots & \dots & \dots \\ G_{n,0} & G_{n,1} & G_{n,2} & \dots & G_{n,n-1} \end{bmatrix}.$$

Proposition 6.14 *The inner function J_{n-1} corresponding to the zero set $w_1, w_2, \dots, w_{n-1}$ has an extra zero at w_n precisely when* $\det R = 0$.

By comparing this situation to the representation (6.2), we can confirm that this is a way of expressing $J_{n-1}(w_n) = 0$.

When $n = 2$ this gives us a simple criterion for deciding whether the inner function corresponding to a linear polynomial has an extra zero. In this situation,

$$\det R = \det \begin{bmatrix} G_{1,0} & G_{1,1} \\ G_{2,0} & G_{2,1} \end{bmatrix} = \langle k_0, k_1 \rangle \langle k_2, k_1 \rangle - \langle k_0, k_2 \rangle \langle k_1, k_1 \rangle.$$

Thus by another route we have arrived at the inner function identified in (4.16).

Example 6.15 Consider the case $\mathscr{H} = H^2$. The S-inner functions are the classical inner functions, which have no extra zeros. Let us confirm this for linear polynomials, using Proposition 6.14. Let r and s be distinct nonzero points in $\mathbb{D}$. Then the inner part of the linear polynomial

$$f(z) = 1 - \frac{z}{r}$$

has the extra zero s precisely if

$$1 \cdot \frac{1}{1 - \bar{s}r} = 1 \cdot \frac{1}{1 - |r|^2}.$$

Of course, this never happens when $r \neq s$, reflecting that the Blaschke factor vanishing at r vanishes nowhere else. To rule out the possibility of a double root at r, we use the kernel function

$$k_{1,r} = \frac{1}{(1 - \bar{r}z)^2},$$

for evaluation of a derivative at r. The criterion then becomes

$$1 \cdot \frac{1}{(1-|r|^2)^2} = 1 \cdot \frac{1}{1-|r|^2},$$

which is also impossible.

Finally, we demonstrate that there are numerous spaces for which there exist S-inner functions with extra zeros.

Example 6.16 Let us return to the weighted spaces $\ell^2(\Lambda)$ of Example 6.10, and consider the special case that the weights arise in connection with the choice

$$\Phi(z) = \frac{1}{1 - a_1 z - a_2 z^2},$$

where $a_1 + a_2 \leqslant 1$, and $a_2 > 4a_1 > 0$. Then, by use of the geometric series formula we find that

$$\Phi(z) = 1 + \sum_{n=1}^{\infty} b_n z^n,$$

with

$$\begin{aligned} b_{2n-1} &= \binom{2n-1}{0} a_1^{2n-1} + \binom{2n-2}{1} a_1^{2n-3} a_2 + \binom{2n-3}{2} a_1^{2n-5} a_2^2 \\ &\quad + \cdots + \binom{n}{n-1} a_1 a_2^{n-1} \\ &\leqslant (a_1^2 + a_2)^{2n}/a_1 \\ &\leqslant 1/a_1 \\ b_{2n} &= \binom{2n}{0} a_1^{2n} + \binom{2n-1}{1} a_1^{2n-2} a_2 + \binom{2n-2}{2} a_1^{2n-4} a_2^2 \\ &\quad + \cdots + \binom{n}{n} a_2^n \\ &\leqslant (a_1^2 + a_2)^{2n} \\ &\leqslant 1. \end{aligned}$$

for all $n \geqslant 1$.

Each coefficient b_n is positive, and so we may define the weights $\lambda_0 = 1$, and $\lambda_n = 1/b_n$, $n \geqslant 1$. The weights are bounded away from zero, and therefore the functions belonging to $\ell^2(\Lambda)$ are analytic in $\mathbb{D}$. Furthermore, point evaluation at $w \in \mathbb{D}$ arises from the reproducing kernel function

$$k_w(z) = \Phi(\bar{w}z) = \frac{1}{1 - a_1\bar{w}z - a_2(\bar{w}z)^2},$$

which is obviously analytic in $\mathbb{D}$.

The inner function associated with the polynomial $1 - z/w$ has an extra zero ζ, distinct from w, provided that

$$\begin{aligned}
\langle k_0, k_w\rangle\langle k_\zeta, k_w\rangle &= \langle k_0, k_\zeta\rangle\langle k_w, k_w\rangle \\
1 \cdot \frac{1}{1 - a_1\bar{w}\zeta - a_2(\bar{w}\zeta)^2} &= 1 \cdot \frac{1}{1 - a_1\bar{w}w - a_2(\bar{w}w)^2} \\
a_1\bar{w}(\zeta - w) + a_2\bar{w}^2(\zeta - w)^2 &= 0 \\
a_1\bar{w} + a_2\bar{w}^2(\zeta + w) &= 0 \\
\zeta &= -\frac{a_1 + a_2|w|^2}{a_2\bar{w}}
\end{aligned}$$

But by assumption $a_2 > 4a_1$, so we can choose $w \in \mathbb{D}$ so that

$$|w| - |w|^2 > a_1/a_2,$$

which in turn implies that $\zeta \in \mathbb{D}$.

We have thus constructed a family of spaces $\mathscr{H}$ of analytic functions on $\mathbb{D}$ for which there exist S-inner functions having extra zeros. This shows that the phenomenon of extra zeros is in some way unexceptional.

7 Inner Vectors in Banach Spaces

Recall from Sect. 2 that a vector $\mathbf{v}$ in a Hilbert space is T-inner if

$$\langle \mathbf{v}, T^n\mathbf{v}\rangle = 0, \ n \geqslant 1. \tag{7.1}$$

We want to extend the definition of T-inner vectors to Banach spaces. However, first we need a notion of "orthogonality" so we can make sense of the very definition in a Banach space. Indeed, what do we mean by $\mathbf{v} \perp T^n\mathbf{v}$ when there is no inner product?

Before jumping into our definition of orthogonality, we need to review a few necessary facts. See [5] for the details. For a complex Banach space $\mathscr{X}$ with norm $\|\cdot\|$, we say that $\mathscr{X}$ is *smooth* if given any $\mathbf{x} \in \mathscr{X} \setminus \{\mathbf{0}\}$ there is a *unique* $\ell \in \mathscr{X}^*$

(the norm dual space of $\mathscr{X}$) such that $\|\ell\| = 1$ and $\ell(\mathbf{x}) = \|\mathbf{x}\|$. Though not relevant to our discussion here, there is an equivalent definition of smoothness of a Banach space involving the Gâteaux derivative of the norm. It is important to point out that the Hahn-Banach theorem yields the existence of a *norming functional* $\ell_{\mathbf{x}}$ for each $\mathbf{x} \in \mathscr{X}$. The *uniqueness* of the above norming functional for every $\mathbf{x} \in \mathscr{X} \setminus \{\mathbf{0}\}$ is what makes $\mathscr{X}$ smooth. Hilbert spaces are smooth, as are the Lebesgue spaces $L^p(X, \mu)$ when $p \in (1, \infty)$. The spaces $L^1(X, \mu)$ and $L^\infty(X, \mu)$ are not smooth.

A Banach space $\mathscr{X}$ is *uniformly convex* if given $\epsilon > 0$, there is a $\delta = \delta(\epsilon) > 0$ such that

$$\|\mathbf{x}\| \leqslant 1,\ \|\mathbf{y}\| \leqslant 1,\ \|\mathbf{x} - \mathbf{y}\| \geqslant \epsilon \implies \|\tfrac{1}{2}(\mathbf{x} + \mathbf{y})\| \leqslant 1 - \delta.$$

A Hilbert space is uniformly convex and Clarkson's inequalities imply that $L^p(X, \mu)$ is uniformly convex when $p \in (1, \infty)$ [5, p. 107]. A uniformly convex Banach space turns out to be reflexive. Important to this paper is the fact that uniformly convex spaces enjoy the *unique nearest point property* in that for a closed subspace (or more generally a closed convex set) Y of $\mathscr{X}$ and a vector $\mathbf{x} \in \mathscr{X}$, there is a unique vector $\widehat{\mathbf{x}} \in Y$ for which

$$\|\mathbf{x} - \widehat{\mathbf{x}}\| \leqslant \|\mathbf{x} - \mathbf{y}\|, \quad \mathbf{y} \in Y. \tag{7.2}$$

This unique nearest point $\widehat{\mathbf{x}}$ is called the *metric projection* of $\mathbf{x}$ onto Y. When $\mathscr{X}$ is a Hilbert space, $\widehat{\mathbf{x}}$ turns out to be the orthogonal projection of $\mathbf{x}$ onto Y and the mapping $\mathbf{x} \mapsto \widehat{\mathbf{x}}$ is linear. For a general Banach space, the mapping $\mathbf{x} \mapsto \widehat{\mathbf{x}}$ is not necessarily linear.

We now follow [2, 18] and define what it means for vectors to be "orthogonal" in a Banach space. For vectors $\mathbf{x}$ and $\mathbf{y}$ in a Banach space $\mathscr{X}$ we say that $\mathbf{x}$ is *orthogonal to* $\mathbf{y}$ *in the Birkhoff–James sense* if

$$\|\mathbf{x} + \beta\mathbf{y}\|_{\mathscr{X}} \geqslant \|\mathbf{x}\|_{\mathscr{X}} \tag{7.3}$$

for all $\beta \in \mathbb{C}$. In this situation we write $\mathbf{x} \perp_{\mathscr{X}} \mathbf{y}$. A little exercise will show that if $\mathscr{X}$ is a Hilbert space, then $\mathbf{x} \perp \mathbf{y} \iff \mathbf{x} \perp_{\mathscr{X}} \mathbf{y}$. In this generality the relation $\perp_{\mathscr{X}}$ is generally neither symmetric nor linear in either argument. However, in a *smooth* Banach space, the relation $\perp_{\mathscr{X}}$ is linear in its second slot, meaning that

$$\mathbf{x} \perp_{\mathscr{X}} \mathbf{y},\ \ \mathbf{x} \perp_{\mathscr{X}} \mathbf{z}, \implies \mathbf{x} \perp_{\mathscr{X}} (\alpha\mathbf{y} + \beta\mathbf{z}), \quad \alpha, \beta \in \mathbb{C}.$$

See [18] for a proof of this.

When $\mathscr{X}$ is a smooth Banach space and $\mathbf{x} \in \mathscr{X}$, we let $\ell_{\mathbf{x}} \in \mathscr{X}^*$ denote the unique norming functional for $\mathbf{x}$ (recall $\ell_{\mathbf{x}}(\mathbf{x}) = \|x\|$). By [2, Cor. 4.2], we can state Birkhoff–James orthogonality equivalently as

$$\mathbf{x} \perp_{\mathscr{X}} \mathbf{y} \iff \ell_{\mathbf{x}}(\mathbf{y}) = 0. \tag{7.4}$$

This condition can be expressed move tangibly for $L^p(X, \mu)$ spaces as (see [18]): For $f, g \in L^p(X, \mu)$,

$$f \perp_{L^p(X,\mu)} g \iff \int_X |f|^{p-2}\overline{f} g d\mu = 0. \tag{7.5}$$

In the above integral, we interpret any instance of $|0|^{p-2}0$ to be zero. We have used Birkhoff–James orthogonality in several recent papers to discuss problems involving the ℓ^p_A spaces of analytic functions whose power series coefficients belong to the sequence space ℓ^p. In [8] we use this orthogonality to give some new bounds on the zeros of an analytic function while in [10] we use this orthogonality, and the concept of an ℓ^p_A-inner function, to describe the zero sets of ℓ^p_A. Still further, we use orthogonality in [7] to give a factorization theorem for ℓ^p_A functions. Though perhaps not using explicitly, by name, the authors in [12] use the above orthogonality to discuss zero sets, via extremal functions, for the L^p Bergman spaces.

With these preliminary remarks, we are ready to define a notion of inner elements. We make the following assumption for the rest of the paper:

$$\mathscr{X} \text{ is a uniformly convex, smooth, complex Banach space.}$$

For a bounded linear transformation $T : \mathscr{X} \to \mathscr{X}$ and a nonzero vector $\mathbf{x} \in \mathscr{X}$, we say that $\mathbf{x}$ is T*-inner* when

$$\mathbf{x} \perp_{\mathscr{X}} T^n\mathbf{x}, \quad n \geqslant 1.$$

By the linearity of the relation $\perp_{\mathscr{X}}$ in the second slot (which follows from our assumptions on $\mathscr{X}$), we see that $\mathbf{x}$ is T-inner if and only if $\mathbf{x} \perp_{\mathscr{X}} [T\mathbf{x}]$, where, as a reminder,

$$[T\mathbf{x}] = \bigvee\{T\mathbf{x}, T^2\mathbf{x}, T^3\mathbf{x}, \ldots\}.$$

If we let $\widehat{\mathbf{x}}$ denote the metric projection (nearest point) of $\mathbf{x}$ onto the subspace $[T\mathbf{x}]$, equivalently, $\widehat{\mathbf{x}}$ is the unique vector satisfying

$$\|\mathbf{x} - \widehat{\mathbf{x}}\| \leqslant \|\mathbf{x} - \mathbf{y}\|, \quad \mathbf{y} \in [T\mathbf{x}],$$

the proof of Proposition 3.1 yields the following.

Proposition 7.6 *If T is a bounded linear transformation on $\mathscr{X}$ and $\mathbf{x} \in \mathscr{X}$, then $\mathbf{x} - \widehat{\mathbf{x}}$ is T-inner (or zero) and every T-inner vector arises in this manner.*

Recall that if T is a bounded linear transformation on $\mathscr{X}$, then the Banach space adjoint operator T^*, i.e., $(T^*\ell)(\mathbf{x}) = \ell(T\mathbf{x})$ for all $\mathbf{x} \in \mathscr{X}$ and $\ell \in \mathscr{X}^*$, is a bounded linear transformation on $\mathscr{X}^*$.

Proposition 7.7 *Suppose that T is a bounded linear transformation on $\mathscr{X}$. If $\mathbf{x} \in \mathscr{X}$ is T-inner, and $\ell_{\mathbf{x}}$ is the unique norming functional of $\mathbf{x}$, then $\ell_{\mathbf{x}}$ is T^*-inner in $\mathscr{X}^*$.*

Proof The assumption of uniform smoothness implies that each nonzero element of $\mathscr{X}$ has a unique norming functional. The hypotheses further imply that $\mathscr{X}$ is reflexive and that $\mathscr{X}^*$ is strictly convex and smooth [5]. Therefore we may speak of unique norming functionals for both $\mathscr{X}$ and $\mathscr{X}^*$.

Suppose that $\|\mathbf{x}\| = 1$, and

$$\mathbf{x} \perp T^k\mathbf{x}, \quad k \geqslant 1.$$

This implies that $\ell_{\mathbf{x}}(T^k\mathbf{x}) = 0$ and $T^{*k}\ell_{\mathbf{x}}(\mathbf{x}) = 0$ for all $k \geqslant 1$. Note that $\mathbf{x}/\|\mathbf{x}\|$ can be viewed as the norming functional for $\ell_{\mathbf{x}}$, since it has norm 1 and

$$\ell_{\mathbf{x}}(\mathbf{x}/\|\mathbf{x}\|) = \frac{1}{\|\mathbf{x}\|}\|\mathbf{x}\| = 1 = \|\ell_{\mathbf{x}}\|.$$

It follows that $\ell_{\mathbf{x}} \perp T^{*k}\ell_{\mathbf{x}}$ for all $k \geqslant 1$. This says that the vector $\ell_{\mathbf{x}} \in \mathscr{X}^*$ is T^*-inner. □

Example 7.8 For the Hardy spaces H^p, $1 < p < \infty$, which we can regard as closed subspaces of $L^p(d\theta/2\pi)$, we can use (7.5) to see that

$$f \perp_{H^p} g \iff \int_0^{2\pi} |f(e^{i\theta})|^{p-2}\overline{f(e^{i\theta})}g(e^{i\theta})\frac{d\theta}{2\pi} = 0.$$

If, as in Example 2.2, $(Tf)(z) = zf(z)$ is the unilateral shift on H^p, then f is T-inner precisely when

$$f \perp_{H^p} z^n f \iff \int_0^{2\pi} |f(e^{i\theta})|^p e^{in\theta}\frac{d\theta}{2\pi} = 0, \quad n \geqslant 1.$$

Again, this shows that f has constant modulus on the circle, i.e., inner in the classical sense.

Example 7.9 For the Bergman spaces $\mathscr{A}^p$ of analytic functions f on $\mathbb{D}$ for which

$$\int_{\mathbb{D}} |f(z)|^p dA < \infty$$

(which is a closed subspace of $L^p(\mathbb{D}, dA)$), we can use (7.5) to extend Example 2.14 and say that $f \in \mathscr{A}^p$ is T-inner ($Tf = zf$) when

$$\int_{\mathbb{D}} |f(z)|^p z^n dA = 0, \quad n \geqslant 1.$$

Example 7.10 For the space

$$\ell^p_A = \left\{ f(z) = \sum_{k \geqslant 0} a_k z^k : \sum_{k \geqslant 0} |a_k|^p < \infty \right\},$$

which turns out to be a well-studied space Banach space of analytic functions on $\mathbb{D}$ (see [9] for a survey), the Birkhoff–James orthogonality becomes

$$f \perp_{\ell^p_A} g \iff \sum_{k \geqslant 0} |a_k|^{p-2} \overline{a_k} b_k = 0.$$

The unilateral shift $(Tf)(z) = zf$ is an isometry on ℓ^p_A and the notion of T-inner was studied in [10]. The condition for $f \in \ell^p_A$ to be T-inner is

$$\sum_{k \geqslant 0} |a_k|^{p-2} \overline{a_k} a_{N+k} = 0, \quad N \geqslant 1,$$

but this condition can be difficult to work with. One can see functions such as $f(z) = z^n$ are inner. When $w \in \mathbb{D} \setminus \{0\}$ an analysis in [7] shows that

$$f(z) = \frac{1 - z/w}{1 - |w|^{p-2} \overline{w} z}$$

is inner. Notice how when $p = 2$ this function becomes a constant times the single Blaschke factor

$$\frac{z - w}{1 - \overline{w} z}.$$

8 Application: Zero Sets for Banach Spaces of Analytic Functions

In this section we develop the analog of Theorem 4.12 for Banach spaces of analytic functions. Let $\mathscr{X}$ be a uniformly convex, smooth, complex Banach space of analytic functions on a domain Ω that satisfies the following conditions.

$$\text{Point evaluation of derivatives of any order is bounded;} \tag{8.1}$$

$$f \in \mathscr{X} \implies zf(z) \in \mathscr{X}; \tag{8.2}$$

$$\bigvee \{z^j : j \geqslant 0\} = \mathscr{X}; \tag{8.3}$$

$$w \in \Omega,\, f \in \mathscr{X} \implies \frac{f(z) - f(w)}{z - w} \in \mathscr{X} \tag{8.4}$$

For some positive constants r and K,

$$f \perp_{\mathscr{X}} g \implies \|f + g\|^r \geqslant \|f\|^r + K\|g\|^r. \tag{8.5}$$

Just as in the Hilbert space case, condition (8.1) can be deduced from conditions (8.2)–(8.4). Furthermore, conditions (8.1), (8.2), and the closed graph theorem show that the operator

$$S_{\mathscr{X}} : \mathscr{X} \to \mathscr{X}, \quad (S_{\mathscr{X}} f)(z) = zf(z),$$

is a bounded linear operator on $\mathscr{X}$.

Note that $\mathscr{X}$ is reflexive and enjoys the unique nearest point property in the sense of (7.2). Furthermore, each nonzero vector $f \in \mathscr{X}$ has a unique norming functional ℓ_f from which it follows from our general discussion in the previous section that $f \perp_{\mathscr{X}} g$ if and only if $\ell_f(g) = 0$. Since evaluation at each point $w \in \Omega$ is continuous, it is given by a functional $k_w \in \mathscr{X}^*$, i.e.,

$$f(w) = k_w(f).$$

Unlike the Hilbert space case discussed earlier, where k_λ belonged to the Hilbert space (equating Hilbert space with its dual space in the natural way via the Riesz representation theorem), here k_w belongs to the dual space $\mathscr{X}^*$ which is not necessarily a space of analytic functions (and for which we don't use the notation $k_w(z)$ as we did for the Hilbert space case).

Condition (8.5) is a "Pythagorean inequality," and it was shown in [6] that all L^p spaces with $p \in (1, \infty)$ satisfy this condition for a range of parameter values r and K. Furthermore, the inequality holds in reverse for other values of r and K.

Important to the development of the analog of Theorem 4.12 for Banach spaces is the following projection lemma, which makes use of the Pythagorean inequality from (8.5).

Lemma 8.6 *Let $\mathscr{X}$ be a smooth Banach space satisfying (8.5). For each $n \in \mathbb{N}$, suppose that $\mathscr{X}_n$ is a subspace of $\mathscr{X}$, such that*

$$\mathscr{X}_1 \subseteq \mathscr{X}_2 \subseteq \mathscr{X}_3 \subseteq \cdots .$$

Define $\mathscr{X}_\infty = \overline{\bigcup_{n=1}^\infty \mathscr{X}_n}$. If P_n is the metric projection mapping from $\mathscr{X}$ to $\mathscr{X}_n$, for all $n \in \mathbb{N} \cup \{\infty\}$, then for any $\mathbf{x} \in \mathscr{X}$, $P_n\mathbf{x}$ converges to $P_\infty\mathbf{x}$ in norm.

Proof By hypothesis, $\mathscr{X}$ is uniformly convex (and hence has unique nearest points), and satisfies the Pythagorean inequality

$$\|\mathbf{x} + \mathbf{y}\|^r \geqslant \|\mathbf{x}\|^r + K\|\mathbf{y}\|^r$$

whenever $\mathbf{x} \perp_{\mathscr{X}} \mathbf{y}$. Let $\mathbf{x} \in \mathscr{X}$. By the definition of metric projection, whenever $m < n$, we have

$$\begin{aligned}\|\mathbf{x} - P_m\mathbf{x}\| &= \inf\{\|\mathbf{x} - \mathbf{z}\| : \mathbf{z} \in \mathscr{X}_m\} \\ &\geqslant \inf\{\|\mathbf{x} - \mathbf{z}\| : \mathbf{z} \in \mathscr{X}_n\} \\ &= \|\mathbf{x} - P_n\mathbf{x}\| \\ &\geqslant \|\mathbf{x} - P_\infty\mathbf{x}\|.\end{aligned}$$

Thus, as a sequence indexed by n, $\|\mathbf{x} - P_n\mathbf{x}\|$ is monotone nonincreasing, and bounded below. Accordingly, it converges.

Next, for $m < n$, the vector $P_n\mathbf{x} - P_m\mathbf{x}$ lies in $\mathscr{X}_n$ (the larger space), and hence the co-projection $\mathbf{x} - P_n\mathbf{x}$ is Birkhoff–James orthogonal to it. Consequently, the Pythagorean inequality says that

$$\|\mathbf{x} - P_m\mathbf{x}\|^r \geqslant \|\mathbf{x} - P_n\mathbf{x}\|^r + K\|P_n\mathbf{x} - P_m\mathbf{x}\|^r.$$

Since the (positive) difference $\|\mathbf{x}-P_m\mathbf{x}\|^r - \|\mathbf{x}-P_n\mathbf{x}\|^r$ can be made arbitrarily small by choosing m sufficiently large, it follows that $\{P_m\mathbf{x}\}_{m\geqslant 1}$ is a Cauchy sequence in norm, and converges to some vector $\mathbf{z}$. It is clear that $\mathbf{z} \in \mathscr{X}_\infty$, and hence

$$\|\mathbf{x} - \mathbf{z}\| \geqslant \|\mathbf{x} - P_\infty\mathbf{x}\|.$$

Next, let $\epsilon > 0$. There exists an N such that

$$\|\mathbf{x} - \mathbf{y}\| \leqslant \|\mathbf{x} - P_\infty\mathbf{x}\| + \epsilon$$

for some $\mathbf{y} \in \mathscr{X}_N$. But then

$$\|\mathbf{x} - \mathbf{z}\| \leqslant \|\mathbf{x} - P_n\mathbf{x}\| \leqslant \|\mathbf{x} - \mathbf{y}\| \leqslant \|\mathbf{x} - P_\infty\mathbf{x}\| + \epsilon.$$

Since this is true for arbitrary ϵ, we conclude that

$$\|\mathbf{x} - \mathbf{z}\| \leqslant \|\mathbf{x} - P_\infty\mathbf{x}\|.$$

Equality holds in these norms, so finally uniqueness of nearest points forces $\mathbf{z} = P_\infty\mathbf{x}$. □

With the above setup we are now ready to develop a version of Theorem 4.12 for Banach spaces satisfying the conditions (8.1)–(8.5). Fix an infinite sequence $W = (w_1, w_2, w_3, \ldots) \subseteq \Omega \setminus \{0\}$, and for each $n \geqslant 1$, define

$$f_n(z) = \left(1 - \frac{z}{w_1}\right) \cdots \left(1 - \frac{z}{w_n}\right)$$

and, by Proposition 7.6, the $S_{\mathscr{X}}$-inner function

$$J_n = f_n - \widehat{f_n},$$

where $\widehat{f}$ stands for the metric projection of f onto $[zf]$. Note that $\widehat{f}$ exists and is unique, by uniform convexity. (When $\mathscr{X}$ is a Hilbert space, the metric projection coincides with the orthogonal projection.)

Let $k_j \in \mathscr{X}^*$ denote the evaluation functional at w_j, $j \geqslant 1$ and k_0 denote the evaluation functional at the origin. The analogous argument used to prove Lemma 4.7 shows that

$$\{f \in \mathscr{X} : f(0) = f(w_j) = 0, 1 \leqslant j \leqslant n\} = [zf_n].$$

Next, suppose that $\lambda = \ell_{J_n} \in \mathscr{X}^*$ is the norming functional for J_n. From

$$J_n \perp_{\mathscr{X}} z^j f_n, \quad 1 \leqslant j \leqslant n$$

and (7.4) we see that

$$\lambda(z^j f_n) = 0, \quad 1 \leqslant j \leqslant n.$$

That is, $\lambda \in [zf_n]^\perp = \bigvee\{k_j : 0 \leqslant j \leqslant n\}$. We may therefore express λ as

$$\lambda = c_0 k_0 + c_1 k_1 + \cdots + c_n k_n$$

for some complex coefficients c_0, c_1,..., c_n. By definition of norming functional this says that

$$\begin{aligned}\|J_n\| &= \lambda(J_n) \\ &= c_0 k_0(J_n) + c_1 k_1(J_n) + \cdots + c_n k_n(J_n) \\ &= c_0 \cdot 1 + 0 + \cdots + 0 \\ &= c_0,\end{aligned}$$

since $k_0(J_n) = J_n(0) = f_n(0) = 1$.

Finally, the condition

$$k_j(J_n) = 0, \ 1 \leqslant j \leqslant n$$

can be interpreted as saying that

$$\lambda \perp_{\mathscr{X}^*} k_j, \ 1 \leqslant j \leqslant n.$$

That is, λ solves the infimum problem

$$\inf \|c_0 k_0 + c_1' k_1 + \cdots + c_n' k_n\|$$

where $c_0 = \|J_n\|$ is fixed, and $c_1', \ldots, c_n'$ are varied.

By renaming the constants, we have shown that

$$1 = \|\lambda\| = \|J_n\| \inf \|k_0 + b_1 k_1 + \cdots + b_n k_n\|,$$

or

$$\|J_n\| = \left[\inf \|k_0 + b_1 k_1 + \cdots + b_n k_n\| \right]^{-1}.$$

As n tends to infinity, the infimum is over a larger set, and thus decreases monotonically, while $\|J_n\|$ must therefore be nondecreasing monotonically.

Suppose W is the zero set of some nontrivial function $f \in \mathscr{X}$. By dividing by z a suitable number of times, we can assume that $f(0) \neq 0$. Then

$$\begin{aligned} \|k_0 + b_1 k_1 + \cdots + b_n k_n\| &\geqslant |\langle k_0 + b_1 k_1 + \cdots + b_n k_n, f\rangle| / \|f\| \\ &= |f(0)| / \|f\| \end{aligned}$$

is bounded from zero, and consequently $\|J_n\|$ is bounded above.

Conversely, if W fails to be the zero set of some nontrivial function of $\mathscr{X}$, then by the Lemma 8.6, there exists an element Λ of $\mathscr{X}^*$ such that the following infimum is attained:

$$\|\Lambda\| = \inf\{\|k_0 + b_1 k_1 + \cdots + b_n k_n\| : b_1, \ldots, b_n \in \mathbb{C}, n \geqslant 1\}.$$

Indeed, Lemma 8.6 tells us that Λ is the k_0 minus its metric projection onto

$$\bigvee \{k_1, k_2, k_3, \ldots\}.$$

Let $\Phi \in \mathscr{X}$ be the norming functional of Λ. Then the infimum condition assures that

$$\Lambda \perp_{\mathscr{X}^*} k_j, \ 1 \leqslant j;$$

that is,

$$k_j(\Phi) = 0$$

for all $j \geqslant 1$. This shows that W is a zero set for Φ. The only way this can happen is if Φ is identically zero, which implies that

$$\lim_{n\to\infty} \inf \|k_0 + b_1 k_1 + \cdots + b_n k_n\| = 0.$$

We memorialize these findings as follows, obtaining an extension of Theorem 4.12 to certain Banach spaces of analytic functions.

Theorem 8.7 *Let $\mathscr{X}$ be a uniformly convex, smooth, complex Banach space of analytic functions on a domain Ω satisfying conditions (8.1)–(8.5). Let $(w_j)_{j\geqslant 1} \subseteq \Omega \setminus \{0\}$ and*

$$f_n = \prod_{j=1}^{n}\Big(1 - \frac{z}{w_j}\Big), \quad J_n = f_n - P_{[zf_n]} f_n.$$

Then

(1) Each J_n is an $S_{\mathscr{X}}$-inner function;
(2) the sequence $\|J_n\|$ is a non-decreasing sequence;
(3) $(w_j)_{j\geqslant 1}$ is a zero sequence for $\mathscr{X}$ if and only if

$$\sup\{\|J_n\| : n \geqslant 1\} < \infty.$$

Spaces for which this applies (i.e., they satisfy the conditions of the abstract Banach space along with the Pythagorean inequality) include the L^p Bergman spaces and ℓ^p_A spaces. A proof specifically tailored for ℓ^p_A was developed in [10].

References

1. A. Aleman, S. Richter, C. Sundberg, Beurling's theorem for the Bergman space. Acta Math. **177**(2), 275–310 (1996)
2. J. Alonso, H. Martini, S. Wu, On Birkhoff orthogonality and isosceles orthogonality in normed linear spaces. Aequationes Math. **83**(1–2), 153–189 (2012)
3. C. Beneteau, M. Fleeman, D. Seco, A. Sola, Remarks on inner functions and optimal approximants. Can. Math. Bull. **61**(4), 704–716 (2018)
4. A. Beurling, On two problems concerning linear transformations in Hilbert space. Acta Math. **81**, 239–55 (1949)
5. N.L. Carothers, in *A Short Course on Banach Space Theory*. London Mathematical Society Student Texts, vol. 64 (Cambridge University Press, Cambridge, 2005)
6. R. Cheng, W.T. Ross, Weak parallelogram laws on Banach spaces and applications to prediction. Period. Math. Hung. **71**(1), 45–58 (2015)
7. R. Cheng, W.T. Ross, An inner-outer factorization in ℓ^p with applications to ARMA processes. J. Math. Anal. Appl. **437**, 396–418 (2016)
8. R. Cheng, J. Mashreghi, W.T. Ross, Birkhoff–James orthogonality and the zeros of an analytic function. Comput. Methods Funct. Theory **17**(3), 499–523 (2017)

9. R. Cheng, J. Mashreghi, W.T. Ross, Multipliers of sequence spaces. Concr. Oper. **4**, 76–108 (2017)
10. R. Cheng, J. Mashreghi, William T. Ross, Inner functions and zero sets for ℓ^p_A (2018). Preprint
11. P.L. Duren, in *Theory of H^p Spaces*. Pure and Applied Mathematics, vol. 38 (Academic, New York/London, 1970)
12. P. Duren, D. Khavinson, H.S. Shapiro, Extremal functions in invariant subspaces of Bergman spaces. Ill. J. Math. **40**(2), 202–210 (1996)
13. O. El-Fallah, K. Kellay, J. Mashreghi, T. Ransford, *A Primer on the Dirichlet Space* (Cambridge University Press, New York, 2014)
14. S.R. Garcia, J. Mashreghi, W.T. Ross, *Introduction to Model Spaces and Their Operators*. Cambridge Studies in Advanced Mathematics, vol. 148 (Cambridge University Press, Cambridge, 2016)
15. H.K. Hedenmalm, K.H. Zhu, On the failure of optimal factorization for certain weighted Bergman spaces. Complex Variables Theory Appl. **19**(3), 165–176 (1992)
16. K. Hoffman, *Banach Spaces of Analytic Functions* (Dover Publications, New York, 1988). Reprint of the 1962 original
17. R.A. Horn, C.R. Johnson, *Matrix Analysis* (Cambridge University Press, New York, 2012)
18. R.C. James, Orthogonality and linear functionals in normed linear spaces. Trans. Am. Math. Soc. **61**, 265–292 (1947)
19. B.I. Korenbljum, Invariant subspaces of the shift operator in a weighted Hilbert space. Mat. Sb. (N.S.) **89**(131), 110–137, 166 (1972)
20. A.G. Miamee, M. Pourahmadi, Wold decomposition, prediction and parameterization of stationary processes with infinite variance. Probab. Theory Relat. Fields **79**(1), 145–164 (1988)
21. V.I. Paulsen, M. Raghupathi, in *An Introduction to the Theory of Reproducing Kernel Hilbert Spaces*. Cambridge Studies in Advanced Mathematics, vol. 152 (Cambridge University Press, Cambridge, 2016)
22. S. Richter, Invariant subspaces of the Dirichlet shift. J. Reine Angew. Math. **386**, 205–220 (1988)
23. S. Richter, C. Sundberg, Multipliers and invariant subspaces in the Dirichlet space. J. Operator Theory **28**(1), 167–186 (1992)
24. S. Richter, C. Sundberg, Invariant subspaces of the Dirichlet shift and pseudocontinuations. Trans. Am. Math. Soc. **341**(2), 863–879 (1994)
25. D. Sarason, Invariant subspaces. Math. Surveys **13**, 1–47 (1974)
26. D. Seco, A characterization of Dirichlet inner functions. Compl. Anal. Oper. Theory (to appear)
27. H.S. Shapiro, A.L. Shields, On the zeros of functions with finite Dirichlet integral and some related function spaces. Math. Z. **80**, 217–229 (1962)

Spherically Quasinormal Pairs of Commuting Operators

Raúl E. Curto and Jasang Yoon

Abstract We first discuss the spherical Aluthge and spherical Duggal transforms for commuting pairs of operators on Hilbert space. Second, we study the fixed points of these transforms, which are the spherically quasinormal commuting pairs. In the case of commuting 2-variable weighted shifts, we prove that spherically quasinormal pairs are intimately related to spherically isometric pairs. We show that each spherically quasinormal 2-variable weighted shift is completely determined by a subnormal unilateral weighted shift (either the 0-th row or the 0-th column in the weight diagram). We then focus our attention on the case when this unilateral weighted shift is recursively generated (which corresponds to a finitely atomic Berger measure). We show that in this case the 2-variable weighted shift is also recursively generated, with a finitely atomic Berger measure that can be computed from its 0-th row or 0-th column. We do this by invoking the relevant Riesz functionals and the functional calculus for the columns of the associated moment matrix.

Keywords Aluthge transform · Duggal transform · Spherically quasinormal · Recursively generated 2-variable weighted shifts

2000 Mathematics Subject Classification Primary 47B20, 47B37, 47A13, 28A50; Secondary 44A60, 47-04, 47A20

The author "Raúl E. Curto" was partially supported by NSF Grant-1302666. The author "Jasang Yoon" was partially supported by a grant from the University of Texas System and the Consejo Nacional de Ciencia y Tecnología de México (CONACYT).

R. E. Curto (✉)
Department of Mathematics, The University of Iowa, Iowa City, IA, USA
e-mail: raul-curto@uiowa.edu; rcurto@math.uiowa.edu

J. Yoon
School of Mathematical and Statistical Sciences, The University of Texas Rio Grande Valley, Edinburg, TX, USA
e-mail: jasang.yoon@utrgv.edu

A. Aleman et al. (eds.), *Analysis of Operators on Function Spaces*,
Trends in Mathematics, https://doi.org/10.1007/978-3-030-14640-5_7

1 Introduction

Let $\mathcal{H}$ be a complex Hilbert space and let $\mathcal{B}(\mathcal{H})$ denote the algebra of bounded linear operators on $\mathcal{H}$. We say that $T \in \mathcal{B}(\mathcal{H})$ is *normal* if $T^*T = TT^*$; *quasinormal* if T commutes with T^*T; *subnormal* if $T = N|_{\mathcal{H}}$, where N is normal and $N(\mathcal{H}) \subseteq \mathcal{H}$; and *hyponormal* if $T^*T \geq TT^*$. It is well known that

$$\text{normal} \Longrightarrow \text{quasinormal} \Longrightarrow \text{subnormal} \Longrightarrow \text{hyponormal}.$$

For $T \in \mathcal{B}(\mathcal{H})$, consider now the *canonical polar decomposition* of T, $T \equiv VP$, where V is a partial isometry, $P := (T^*T)^{\frac{1}{2}}$, and $\ker T = \ker V = \ker P$. The *Aluthge transform* $\widehat{T}$ is $\widehat{T} := P^{\frac{1}{2}}VP^{\frac{1}{2}}$; on the other hand, the *Duggal transform* is $\widehat{T}^D := PV$. The Aluthge transform was first introduced in [1] and it has attracted considerable attention over the last two decades (see, for instance, [2, 9, 27, 32, 33, 36] and [41]).

It is well known that $T \in \mathcal{B}(\mathcal{H})$ is quasinormal if and only if T commutes with the positive factor P in the canonical polar decomposition $T \equiv VP$; equivalently, if V commutes with P. It follows easily that T is quasinormal if and only if $T = \widehat{T}$, that is, if and only if T is a fixed point for the Aluthge transform. One can similarly establish that the fixed points of the Duggal transform are also the quasinormal operators.

To study the bivariate situation, we need some notation. The class of commuting pairs of operators on Hilbert space will be denoted by $\mathfrak{C}_0$; the subclass of commuting pairs of subnormal operators will be denoted by $\mathfrak{H}_0$; and the subclass of jointly subnormal pairs by $\mathfrak{H}_\infty$.

For $S, T \in \mathcal{B}(\mathcal{H})$ let $[S, T] := ST - TS$. We say that a commuting pair $\mathbf{T} = (T_1, T_2)$ of operators on $\mathcal{H}$ is (jointly) *hyponormal* if the operator matrix

$$[\mathbf{T}^*, \mathbf{T}] := \begin{pmatrix} [T_1^*, T_1] & [T_2^*, T_1] \\ [T_1^*, T_2] & [T_2^*, T_2] \end{pmatrix}$$

is positive on the direct sum of two copies of $\mathcal{H}$ (cf. [3, 12]). A commuting pair $\mathbf{T}$ is said to be *normal* if $\mathbf{T}$ is commuting and each T_i is normal, and *subnormal* if $\mathbf{T}$ is the restriction of a normal pair to a common invariant subspace. For $k \geq 1$, a commuting pair $\mathbf{T} \equiv (T_1, T_2)$ is said to be *k-hyponormal* [25] if

$$\mathbf{T}(k) := \left(T_1, T_2, T_1^2, T_2T_1, T_2^2, \cdots, T_1^k, T_2T_1^{k-1}, \cdots, T_2^k\right)$$

is hyponormal.

For $k \geq 1$, we let $\mathfrak{H}_k$ denote the class of k-hyponormal pairs in $\mathfrak{H}_0$. It is now clear that

$$\mathfrak{H}_\infty \subseteq \mathfrak{H}_k \subseteq \mathfrak{H}_0 \subseteq \mathfrak{C}_0.$$

The multivariable Bram-Halmos Theorem states that $\mathbf{T}$ is subnormal if and only if $\mathbf{T}$ is k-hyponormal for all $k \geq 1$ [25, Theorem 2.3]; that is, $\mathfrak{H}_\infty = \bigcap_{k\geq 1} \mathfrak{H}_k$.

We next consider a suitable polar decomposition (and corresponding Aluthge and Duggal transforms) for $\mathbf{T} = (T_1, T_2) \in \mathfrak{C}_0$. Given a commuting pair $\mathbf{T} \equiv (T_1, T_2)$ of operators acting on $\mathcal{H}$, let

$$Q := (T_1^*T_1 + T_2^*T_2)^{\frac{1}{2}}. \tag{1.1}$$

Clearly, $\ker Q = \ker T_1 \bigcap \ker T_2$. For $x \in \ker Q$, let $V_i x := 0$, and for $y \in \operatorname{Ran} Q$, say $y = Qx$, let $V_i y := T_i x$ $(i = 1, 2)$. It is easy to see that V_1 and V_2 are well defined. We then have

$$\begin{pmatrix} T_1 \\ T_2 \end{pmatrix} = \begin{pmatrix} V_1 \\ V_2 \end{pmatrix} Q,$$

as operators from $\mathcal{H}$ to $\mathcal{H} \oplus \mathcal{H}$. Moreover, this is the unique canonical polar decomposition of $\begin{pmatrix} T_1 \\ T_2 \end{pmatrix}$. It follows that $\begin{pmatrix} V_1 \\ V_2 \end{pmatrix}$ is a partial isometry from $(\ker Q)^\perp$ onto $\overline{\operatorname{Ran} \begin{pmatrix} T_1 \\ T_2 \end{pmatrix}}$.

Following [5] and [30], we say that $\mathbf{T}$ is *(jointly) quasinormal* if T_i commutes with $T_j^*T_j$ for all $i, j = 1, 2$; and *spherically quasinormal* if T_i commutes with Q for $i = 1, 2$. By [5], for all $k \geq 1$, one has

$$\begin{aligned} \text{normal} &\Longrightarrow \text{(jointly) quasinormal} \Longrightarrow \text{spherically quasinormal} \\ &\Longrightarrow \text{subnormal} \Longrightarrow k\text{-hyponormal}. \end{aligned} \tag{1.2}$$

On the other hand, the results in [25] and [30] show that the reverse implications in (1.2) do not necessarily hold.

We are now ready to introduce two bivariate operator transforms.

Definition 1.1 (cf. [23, 24]) With $\mathbf{T}$, V_1, V_2 and Q as above, the spherical Aluthge transform of $\mathbf{T}$ is

$$\widehat{\mathbf{T}} \equiv (\widehat{T_1}, \widehat{T_2}),$$

where

$$\widehat{T_i} := Q^{\frac{1}{2}} V_i Q^{\frac{1}{2}} \ (i = 1, 2). \tag{1.3}$$

Lemma 1.2 ([7, 24]) $\widehat{\mathbf{T}}$ *is commutative.*

Definition 1.3 (cf. [35]) With $\mathbf{T}$, V_1, V_2 and Q as above, the spherical Duggal transform of $\mathbf{T}$ is

$$\widehat{\mathbf{T}}^D := (\widehat{T}_1^D, \widehat{T}_2^D),$$

where

$$\widehat{T}_i^D := QV_i \ (i = 1, 2). \tag{1.4}$$

A simple application of Lemma 1.2 together with the fact that $\ker V_1 \bigcap \ker V_2 = \ker Q$ readily implies the following result.

Lemma 1.4 ([35]) $\widehat{\mathbf{T}}^D$ *is commutative.*

Remark 1.5 Note that, in general, $\widehat{T_i} \equiv (\widehat{\mathbf{T}})_i$ (resp. $\widehat{T}_i^D \equiv (\widehat{\mathbf{T}}^D)_i$) is *not* the Aluthge (resp. Duggal) transform of T_i $(i = 1, 2)$. □

The spherical Aluthge transform was introduced in [23]; its general theory was developed in [24]. In this paper we focus on the spherical quasinormal pairs, which are the fixed points of the spherical Aluthge and Duggal transforms. After characterizing the spherically quasinormal 2-variable weighted shifts, we study the case when a row or column in the weight diagram corresponds to a recursively generated unilateral weighted shift, that is, a weighted shift with finitely atomic Berger measure.

The organization of this paper is as follows. In Sect. 2 we will characterize the fixed points of the Aluthge and Duggal bivariate operator transforms; these are the spherically quasinormal pairs, that is, those commuting pairs for which T_i commutes with $T_1^*T_1 + T_2^*T_2$ for all $i = 1, 2$. In Sect. 3 we characterize the spherically quasinormal 2-variable weighted shifts. In Sect. 4 we provide a concrete construction of spherically quasinormal 2-variable weighted shifts, in terms of the 0-th row or 0-th column in their weight diagram (see Fig. 1i). In Sect. 5 we focus our

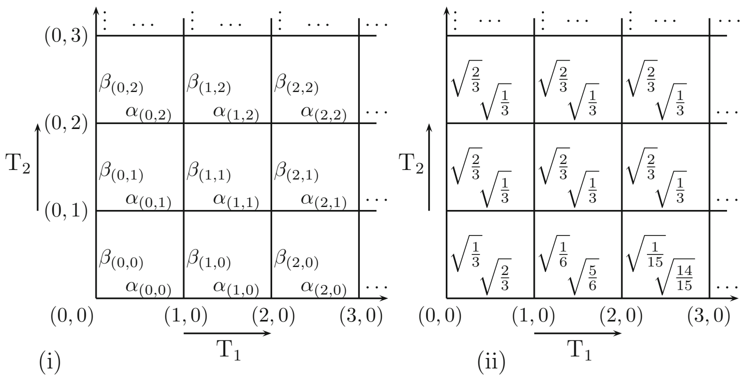


Fig. 1 Weight diagram of a generic 2-variable weighted shift and weight diagram of the 2-variable weighted shift in Example 5.10, respectively

attention on the case when the 0-th row or 0-th column corresponds to a recursively generated subnormal unilateral weighted shift. Finally, we list in Appendix some known results which are needed somewhere else in the paper.

We devote the rest of this section to establishing some additional basic terminology and notation. For $\omega \equiv \{\omega_n\}_{n=0}^{\infty}$ a bounded sequence of positive real numbers (called *weights*), let $W_\omega \equiv \text{shift}(\omega_0, \omega_1, \cdots) : \ell^2(\mathbb{Z}_+) \to \ell^2(\mathbb{Z}_+)$ be the associated unilateral weighted shift, defined by $W_\omega e_n := \omega_n e_{n+1}$ (all $n \geq 0$), where $\{e_n\}_{n=0}^{\infty}$ is the canonical orthonormal basis in $\ell^2(\mathbb{Z}_+)$. The moments of $\omega \equiv \{\omega_n\}_{n=0}^{\infty}$ are given as

$$\gamma_k \equiv \gamma_k(W_\omega) := \begin{cases} 1, & \text{if } k = 0 \\ \omega_0^2 \cdots \omega_{k-1}^2, & \text{if } k > 0. \end{cases} \tag{1.5}$$

The (unweighted) unilateral shift is $U_+ := \text{shift}(1, 1, 1, \cdots)$. For $0 < a < 1$ we let $S_a := \text{shift}(a, 1, 1, \cdots)$.

We now recall a well-known characterization of subnormality for unilateral weighted shifts, due to Berger (cf. [10, III.8.16]) and independently established by Gellar and Wallen [29]: W_ω is subnormal if and only if there exists a probability measure σ supported in $[0, \|W_\omega\|^2]$ (called the *Berger measure* of W_ω) such that $\gamma_k(\omega) = \gamma_k(W_\omega)\, \omega_0^2 \cdot \ldots \cdot \omega_{k-1}^2 = \int t^k d\sigma(t)$ $(k \geq 1)$.

Observe that U_+ and S_a are subnormal unilateral weighted shifts, with Berger measures δ_1 and $(1 - a^2)\delta_0 + a^2\delta_1$, respectively. (Here δ_p denotes the point-mass probability measure with support the singleton set $\{p\}$.)

Similarly, consider double-indexed positive bounded sequences $\alpha_{\mathbf{k}}, \beta_{\mathbf{k}} \in \ell^\infty(\mathbb{Z}_+^2)$, $\mathbf{k} \equiv (k_1, k_2) \in \mathbb{Z}_+^2$ and let $\ell^2(\mathbb{Z}_+^2)$ be the Hilbert space of square-summable complex sequences indexed by $\mathbb{Z}_+^2$. (Recall that $\ell^2(\mathbb{Z}_+^2)$ is canonically isometrically isomorphic to $\ell^2(\mathbb{Z}_+) \bigotimes \ell^2(\mathbb{Z}_+)$.) We define the 2-variable weighted shift $\mathbf{T} \equiv (T_1, T_2) = W_{(\alpha,\beta)}$ by

$$T_1 e_{\mathbf{k}} := \alpha_{\mathbf{k}} e_{\mathbf{k}+\varepsilon_1} \text{ and } T_2 e_{\mathbf{k}} := \beta_{\mathbf{k}} e_{\mathbf{k}+\varepsilon_2}, \tag{1.6}$$

where $\varepsilon_1 := (1, 0)$ and $\varepsilon_2 := (0, 1)$. Clearly,

$$T_1 T_2 = T_2 T_1 \Longleftrightarrow \beta_{\mathbf{k}+\varepsilon_1} \alpha_{\mathbf{k}} = \alpha_{\mathbf{k}+\varepsilon_2} \beta_{\mathbf{k}} \ \left(\text{all } \mathbf{k} \in \mathbb{Z}_+^2\right). \tag{1.7}$$

Moreover, for $\mathbf{k} \in \mathbb{Z}_+^2$ we have

$$T_1^* e_{(0,k_2)} = 0 \ \text{ and } \ T_1^* e_{\mathbf{k}} = \alpha_{\mathbf{k}-\varepsilon_1} e_{\mathbf{k}-\varepsilon_1} \ (k_1 \geq 1); \tag{1.8}$$

$$T_2^* e_{(k_1,0)} = 0 \ \text{ and } \ T_2 e_{\mathbf{k}} := \beta_{\mathbf{k}-\varepsilon_2} e_{\mathbf{k}-\varepsilon_2} \ (k_2 \geq 1). \tag{1.9}$$

In an entirely similar way one can define multivariable weighted shifts (see. [21, 22]).

We now recall the definition of moments for a 2-variable weighted shift $\mathbf{T} \equiv (T_1, T_2) = W_{(\alpha,\beta)} \in \mathfrak{C}_0$. Given $\mathbf{k} \equiv (k_1, k_2) \in \mathbb{Z}_+^2$, the moment of $\mathbf{T} \equiv (T_1, T_2) = W_{(\alpha,\beta)}$ of order $\mathbf{k}$ is

$$\gamma_{\mathbf{k}} \equiv \gamma_{\mathbf{k}}(W_{(\alpha,\beta)}) := \begin{cases} 1, & \text{if } k_1 = 0 \text{ and } k_2 = 0 \\ \alpha_{(0,0)}^2 \cdots \alpha_{(k_1-1,0)}^2, & \text{if } k_1 \geq 1 \text{ and } k_2 = 0 \\ \beta_{(0,0)}^2 \cdots \beta_{(0,k_2-1)}^2, & \text{if } k_1 = 0 \text{ and } k_2 \geq 1 \\ \alpha_{(0,0)}^2 \cdots \alpha_{(k_1-1,0)}^2 \beta_{(k_1,0)}^2 \cdots \beta_{(k_1,k_2-1)}^2, & \text{if } k_1 \geq 1 \text{ and } k_2 \geq 1. \end{cases} \tag{1.10}$$

We remark that, due to the commutativity condition (1.7), $\gamma_{\mathbf{k}}$ can be computed using any nondecreasing path from $(0, 0)$ to (k_1, k_2). Given a 2-variable weighted shift $\mathbf{T} \equiv (T_1, T_2) = W_{(\alpha,\beta)} \in \mathfrak{C}_0$, and given $k_1, k_2 \geq 0$, we let

$$W_{k_2} := \text{shift}(\alpha_{(0,k_2)}, \alpha_{(1,k_2)}, \cdots) \tag{1.11}$$

be the k_2-th horizontal slice of T_1; similarly, we let

$$V_{k_1} := \text{shift}(\beta_{(k_1,0)}, \beta_{(k_1,1)}, \cdots) \tag{1.12}$$

be the k_1-th vertical slice of T_2. (Clearly, W_0 and V_0 are the unilateral weighted shifts associated with the 0-th row and 0-column in the weight diagram for $\mathbf{T}$, resp.) By the commutativity condition (1.7), we note that

$$\gamma_{(k_1,k_2)}(W_{(\alpha,\beta)}) = \frac{\gamma_{k_1}\left(W_{k_2}\right)}{\beta_{(0,0)}^2 \cdots \beta_{(0,k_2-1)}^2}, \tag{1.13}$$

where $\gamma_{k_1}\left(W_{k_2}\right)$ is given by (1.5). A similar identity holds for V_{k_1}.

A straightforward generalization of the above-mentioned Berger-Gellar-Wallen result was proved in [31]. That is, a commuting 2-variable weighted shift $\mathbf{T} \equiv (T_1, T_2) = W_{(\alpha,\beta)}$ admits a commuting normal extension if and only if there is a probability measure μ (which we call the Berger measure of $\mathbf{T}$) defined on the 2-dimensional rectangle $R = [0, a_1] \times [0, a_2]$ (where $a_i := \|T_i\|^2$) such that $\gamma_{\mathbf{k}} = \int_R s^{k_1} t^{k_2} d\mu(s, t)$, for all $\mathbf{k} \in \mathbb{Z}_+^2$.

In the single variable case, if W_ω is subnormal with Berger measure σ_ω and $h \geq 1$, and if we let $\mathcal{L}_h := \bigvee\{e_n : n \geq h\}$ denote the invariant subspace obtained by removing the first h vectors in the canonical orthonormal basis of $\ell^2(\mathbb{Z}_+)$, then the Berger measure of $W_\omega|_{\mathcal{L}_h}$ is $\frac{s^h}{\gamma_h} d\sigma_\omega(s)$; alternatively, if $S : \ell^\infty(\mathbb{Z}_+) \to \ell^\infty(\mathbb{Z}_+)$ is defined by

$$S(\omega)(n) := \omega(n + 1) \ (\omega \in \ell^\infty(\mathbb{Z}_+), n \geq 0), \tag{1.14}$$

then

$$d\sigma_{S(\omega)}(s) = \frac{s}{\omega_0^2} d\sigma_\omega(s). \tag{1.15}$$

2 The Spherical Aluthge and Duggal Transforms

In [32], Jung et al. proved that an operator $T \in B(\mathcal{H})$ with dense range has a nontrivial invariant subspace if and only if $\widehat{T}$ does. On the other hand, one can show that T has a nontrivial invariant subspace if and only if $\widehat{T}^D$ does, where $\widehat{T}^D$ is the Duggal transform for T. In [34, 35], the authors studied the common invariant subspaces between the spherical Aluthge (resp. Duggal) transform and its original pair. By Lemmas 1.2 and 1.4, we know that $\widehat{\mathbf{T}}, \widehat{\mathbf{T}}^D \in \mathfrak{C}_0$ whenever $\mathbf{T} \in \mathfrak{C}_0$ (cf. [23, 35]). In [34, 35], the authors showed that for $\mathbf{T} \in \mathfrak{C}_0$ with dense ranges, $\mathbf{T}$ has a common nontrivial invariant subspace if and only if $\widehat{\mathbf{T}}$ does if and only if $\widehat{\mathbf{T}}^D$ does.

In [33], Jung et al. also proved that T and $\widehat{T}$ have the same spectrum. This result can be extended to pairs $\mathbf{T} \in \mathfrak{C}_0$ (cf. [7, 26]). That is, one can show that for a commuting pair $\mathbf{T} \equiv (T_1, T_2)$

$$\sigma_T(\widehat{\mathbf{T}}) = \sigma_T(\mathbf{T}), \tag{2.1}$$

where $\sigma_T(\mathbf{T})$ is the Taylor spectrum of $\mathbf{T}$. (For more information on the notion of Taylor spectrum and related results, the reader is referred to [11, 13, 39, 40]).

Related to the above-mentioned results, it is well known, and easy to prove, that if $T \in \mathcal{B}(\mathcal{H})$ is invertible, then $\widehat{T}$ is also invertible. In this case, $\widehat{T} = |T|^{\frac{1}{2}} T |T|^{-\frac{1}{2}}$. Similarly, for $\mathbf{T} \in \mathfrak{C}_0$, one can use a bit of homological algebra applied to the appropriate Koszul complexes to prove directly that $\widehat{\mathbf{T}}$ is Taylor invertible when $\mathbf{T}$ is Taylor invertible. If $\mathbf{T} \equiv (T_1, T_2)$ is Taylor invertible and we represent it as a column matrix, then one can see that Q is also invertible, and in this case,

$$\widehat{\mathbf{T}} = Q^{\frac{1}{2}} \mathbf{T} \left(Q^{-\frac{1}{2}} \oplus Q^{-\frac{1}{2}} \right).$$

We next consider the structure of commuting pairs which are fixed points of the spherical Aluthge and Duggal transform. It is known that T is quasinormal if and only if $T = \widehat{T}$ if and only if $T = \widehat{T}^D$. We will extend this result to the case of commuting pairs $\mathbf{T} \equiv (T_1, T_2)$. First, we need an auxiliary result.

Lemma 2.1 *For $i = 1, 2$, T_i commutes with Q if and only if V_i commutes with Q.*

Proof Recall that, for $i = 1, 2$, $T_i = V_i Q$. If T_i commutes with Q, then $V_i Q^2 = (V_i Q) Q = T_i Q = Q T_i = Q(V_i Q)$, and as a consequence $(V_I Q - Q V_i) Q = 0$; that is, V_i commutes with Q on Ran Q. On the other hand, $V_i Q - Q V_i$ vanishes on ker Q. It now easily follows that V_i commutes with Q. The converse is trivial. □

We next consider spherical quasinormality for commuting pairs. Suppose a commuting pair $\mathbf{T}$ is spherically quasinormal. Since for $i = 1, 2$, T_i commutes with $T_1^*T_1 + T_2^*T_2$, then for $i = 1, 2$ T_i commutes with Q (by the continuous functional calculus for Q). Observe now that

$$\widehat{(T_1, T_2)}\sqrt{Q} = \left(\sqrt{Q}V_1\sqrt{Q}, \sqrt{Q}V_2\sqrt{Q}\right)\sqrt{Q} = \left(\sqrt{Q}T_1, \sqrt{Q}T_2\right) = (T_1, T_2)\sqrt{Q},$$

so that

$$\widehat{(T_1, T_2)} = (T_1, T_2) \text{ on } \overline{\operatorname{Ran}\sqrt{Q}} \; (= \overline{\operatorname{Ran} Q}). \tag{2.2}$$

On the other hand, since $\ker Q = \ker T_1 \bigcap \ker T_2$, it follows easily that

$$\widehat{(T_1, T_2)} = (T_1, T_2) \text{ on } \ker Q. \tag{2.3}$$

Since $\mathcal{H} = \left(\overline{\operatorname{Ran} Q}\right) \oplus \ker Q$, we can combine (2.2) and (2.3) to prove that $\widehat{(T_1, T_2)} = (T_1, T_2)$.

We are now ready to state the main result of this section.

Theorem 2.2 *Let* $\mathbf{T} \equiv (T_1, T_2) \in \mathfrak{C}_0$. *The following statements are equivalent.*

(i) $\mathbf{T}$ *is spherically quasinormal.*
(ii) $\widehat{(T_1, T_2)} = (T_1, T_2)$.
(iii) $\widehat{(T_1, T_2)}^D = (T_1, T_2)$.

Proof (i) $\Rightarrow$ (ii): This follows from the discussion preceding the statement of Theorem 2.2.
(ii) $\Rightarrow$ (iii):

$$\begin{aligned}
\widehat{\mathbf{T}} = \mathbf{T} &\Longrightarrow \left(\sqrt{Q}V_1\sqrt{Q}, \sqrt{Q}V_2\sqrt{Q}\right) = (V_1 Q, V_2 Q) \\
&\Longrightarrow \left(\sqrt{Q}T_1, \sqrt{Q}T_2\right) = \left(T_1\sqrt{Q}, T_2\sqrt{Q}\right) \\
&\Longrightarrow T_i \text{ commutes with } \sqrt{Q} \; (i = 1, 2) \\
&\Longrightarrow T_i \text{ commutes with } Q \; (i = 1, 2) \\
&\Longrightarrow V_i \text{ commutes with } Q \; (i = 1, 2) \\
&\Longrightarrow \widehat{\mathbf{T}}^D = \mathbf{T}.
\end{aligned}$$

(iii) $\Rightarrow$ (i): Assume that $\widehat{\mathbf{T}}^D = \mathbf{T}$. It follows that V_i commutes with Q $(i = 1, 2)$. As a consequence, T_i commutes with Q, which implies that T_i commutes with Q^2 $(i = 1, 2)$, as desired. □

3 A Characterization of Spherically Quasinormal 2-Variable Weighted Shifts

In this section we present a characterization of spherical quasinormality for 2-variable weighted shifts. The following theorem was announced in [23]. Before we state it, we list a simple fact about quasinormality for 2-variable weighted shifts.

Remark 3.1 A 2-variable weighted shift $\mathbf{T} \equiv (T_1, T_2) = W_{(\alpha,\beta)}$ is (jointly) quasinormal if and only if $\alpha_{(k_1,k_2)} = \alpha_{(0,0)}$ and $\beta_{(k_1,k_2)} = \beta_{(0,0)}$ for all $k_1, k_2 \geq 0$. This can be seen via a simple application of (1.7) and (1.8). As a result, up to a scalar multiple in each component, a quasinormal 2-variable weighted shift is identical to the so-called Helton-Howe shift; that is, the shift that corresponds to the pair of multiplications by the coordinate functions in the Hardy space $H^2(\mathbb{T} \times \mathbb{T})$ of the 2-torus, with respect to arclength measure on each circle $\mathbb{T}$ (cf. [30]). This fact is entirely consistent with the one-variable result: a unilateral weighted shift W_ω is quasinormal if and only if $W_\omega = cU_+$ for some $c > 0$. □

Theorem 3.2 *Let* $\mathbf{T} \equiv (T_1, T_2) = W_{(\alpha,\beta)} \in \mathfrak{C}_0$ *be a* 2*-variable weighted shift. Then the following statements are equivalent.*

(i) $\mathbf{T} \equiv (T_1, T_2)$ *is spherically quasinormal.*

(ii) There exists a constant $c > 0$ *such that for all* $\mathbf{k} \equiv (k_1, k_2) \in \mathbb{Z}_+^2$,

$$\alpha_{(k_1,k_2)}^2 + \beta_{(k_1,k_2)}^2 = c.$$

(iii) $T_1^*T_1 + T_2^*T_2 = c\, I.$

Proof (i) $\Longrightarrow$ (ii): Assume that $\mathbf{T} \equiv (T_1, T_2)$ is spherically quasinormal. Then, $\widehat{(T_1, T_2)} = (T_1, T_2)$, where $\widehat{(T_1, T_2)}$ is the spherical Aluthge transform of $\mathbf{T}$. Thus, we have

$$\left(\sqrt{Q}V_1\sqrt{Q}, \sqrt{Q}V_2\sqrt{Q}\right) = (V_1Q, V_2Q) \Longrightarrow \left(\sqrt{Q}T_1, \sqrt{Q}T_2\right) = \left(T_1\sqrt{Q}, T_2\sqrt{Q}\right),$$

that is, for all $i = 1, 2$, T_i commutes with $\sqrt{Q}$. Hence, the continuous functional calculus imposes that T_1 and T_2 commute with Q. We now consider the following: for all $\mathbf{k} \equiv (k_1, k_2) \in \mathbb{Z}_+^2$, $\alpha_{(k_1,k_2)}^2 + \beta_{(k_1,k_2)}^2 = c$. If we fix an orthonormal basis vector $e_{\mathbf{k}}$, then by (1.6) and (1.1) we have

$$T_1 e_{\mathbf{k}} = \alpha_{\mathbf{k}} e_{\mathbf{k}+\varepsilon_1}, \quad T_2 e_{\mathbf{k}} := \beta_{\mathbf{k}} e_{\mathbf{k}+\varepsilon_2},$$

and

$$Q e_{\mathbf{k}} = \sqrt{\alpha_{(k_1,k_2)}^2 + \beta_{(k_1,k_2)}}.$$

We thus obtain

$$QT_1 e_{\mathbf{k}} = \alpha_{(k_1,k_2)} \sqrt{\alpha^2_{(k_1+1,k_2)} + \beta^2_{(k_1+1,k_2)}} \tag{3.1}$$

$$T_1 Q e_{\mathbf{k}} = \sqrt{\alpha^2_{(k_1,k_2)} + \beta^2_{(k_1,k_2)}}\, \alpha_{(k_1,k_2)}. \tag{3.2}$$

It follows that

$$\sqrt{\alpha^2_{(k_1+1,k_2)} + \beta^2_{(k_1+1,k_2)}} = \sqrt{\alpha^2_{(k_1,k_2)} + \beta^2_{(k_1,k_2)}}. \tag{3.3}$$

Similarly, we have

$$QT_2 e_{\mathbf{k}} = \beta_{(k_1,k_2)} \sqrt{\alpha^2_{(k_1,k_2+1)} + \beta^2_{(k_1,k_2+1)}} \tag{3.4}$$

$$T_2 Q e_{\mathbf{k}} = \sqrt{\alpha^2_{(k_1,k_2)} + \beta^2_{(k_1,k_2)}}\, \beta_{(k_1,k_2)}. \tag{3.5}$$

Hence,

$$\sqrt{\alpha^2_{(k_1,k_2+1)} + \beta^2_{(k_1,k_2+1)}} = \sqrt{\alpha^2_{(k_1,k_2)} + \beta^2_{(k_1,k_2)}}. \tag{3.6}$$

Therefore, by (3.3) and (3.6), for all $\mathbf{k} \equiv (k_1, k_2) \in \mathbb{Z}^2_+$ we obtain

$$\sqrt{\alpha^2_{(k_1,k_2)} + \beta^2_{(k_1,k_2)}} = \sqrt{\alpha^2_{(k_1+1,k_2)} + \beta^2_{(k_1+1,k_2)}} = \sqrt{\alpha^2_{(k_1,k_2+1)} + \beta^2_{(k_1,k_2+1)}};$$

that is, for all $\mathbf{k} \equiv (k_1, k_2) \in \mathbb{Z}^2_+$ we have

$$\alpha^2_{(k_1,k_2)} + \beta^2_{(k_1,k_2)} = c := \alpha^2_{(0,0)} + \beta^2_{(0,0)} > 0,$$

as desired.

(ii) $\Longrightarrow$ (iii): We assume that $\alpha^2_{(k_1,k_2)} + \beta^2_{(k_1,k_2)} = c > 0$ for all $\mathbf{k} \equiv (k_1, k_2) \in \mathbb{Z}^2_+$. Then, by (3.2) and (3.5), we clearly get that

$$T_1^* T_1 + T_2^* T_2 = c \cdot I.$$

(iii) $\Longrightarrow$ (i): We assume that $T_1^* T_1 + T_2^* T_2 = c \cdot I$. Then, for all $i = 1, 2$, we have

$$T_i \left(T_1^* T_1 + T_2^* T_2\right) = c \cdot T_i = \left(T_1^* T_1 + T_2^* T_2\right) T_i,$$

so that we get that T_1 and T_2 commute with Q. Thus, by the same argument in the proof of Theorem 2.2, we have that $\widehat{(T_1, T_2)} = (T_1, T_2)$. Therefore, by Theorem 2.2, $\mathbf{T}$ is spherically quasinormal. $\square$

Remark 3.3 If $\mathbf{T} \equiv (T_1, T_2) = W_{(\alpha,\beta)} \in \mathfrak{C}_0$ is a spherically quasinormal 2-variable weighted shift, then Q is injective, so that by the continuous functional calculus, we have that $(T_1, T_2) \in \mathfrak{C}_0$ is spherically quasinormal if and only if each T_i is commute with $Q^2 = T_1^* T_1 + T_2^* T_2$ for all $i = 1, 2$. Observe that in the case of arbitrary commuting pairs of operators, we always have $Q^2 = T_1^* T_1 + T_2^* T_2 = Q(V_1^* V_1 + V_2^* V_2) Q$; thus, when Q is injective, we obtain $V_1^* V_1 + V_2^* V_2 = I$. □

We now investigate the weight diagrams of $\widehat{\mathbf{T}}$ and $\widehat{\mathbf{T}}^D$ for a 2-variable weighted shift $\mathbf{T} \equiv (T_1, T_2) = W_{(\alpha,\beta)}$.

Proposition 3.4 *Let* $\mathbf{T} \equiv (T_1, T_2) = W_{(\alpha,\beta)}$ *be a 2-variable weighted shift. Then*

$$\widehat{T}_1 e_{\mathbf{k}} = \alpha_{\mathbf{k}} \frac{(\alpha^2_{\mathbf{k}+\epsilon_1} + \beta^2_{\mathbf{k}+\epsilon_1})^{1/4}}{(\alpha^2_{\mathbf{k}} + \beta^2_{\mathbf{k}})^{1/4}} e_{\mathbf{k}+\epsilon_1}; \ \widehat{T}_2 e_{\mathbf{k}} = \beta_{\mathbf{k}} \frac{(\alpha^2_{\mathbf{k}+\epsilon_2} + \beta^2_{\mathbf{k}+\epsilon_2})^{1/4}}{(\alpha^2_{\mathbf{k}} + \beta^2_{\mathbf{k}})^{1/4}} e_{\mathbf{k}+\epsilon_2} \tag{3.7}$$

and

$$\widehat{T}_1^D e_{\mathbf{k}} = \alpha_{\mathbf{k}} \frac{(\alpha^2_{\mathbf{k}+\epsilon_1} + \beta^2_{\mathbf{k}+\epsilon_1})^{1/2}}{(\alpha^2_{\mathbf{k}} + \beta^2_{\mathbf{k}})^{1/2}} e_{\mathbf{k}+\epsilon_1}; \ \widehat{T}_2^D e_{\mathbf{k}} = \beta_{\mathbf{k}} \frac{(\alpha^2_{\mathbf{k}+\epsilon_2} + \beta^2_{\mathbf{k}+\epsilon_2})^{1/2}}{(\alpha^2_{\mathbf{k}} + \beta^2_{\mathbf{k}})^{1/2}} e_{\mathbf{k}+\epsilon_2} \tag{3.8}$$

for all $\mathbf{k} \in \mathbb{Z}^2_+$.

Proof Straightforward from (1.1), (1.3), (1.4) and (1.6). □

Remark 3.5 By (3.7) and (3.8) in Proposition 3.4, if $\widehat{W_{(\alpha,\beta)}} = \widehat{W_{(\alpha,\beta)}}^D$, then for all $\mathbf{k} \equiv (k_1, k_2) \in \mathbb{Z}^2_+$, $\alpha^2_{(k_1,k_2)} + \beta^2_{(k_1,k_2)} = c > 0$, so that $W_{(\alpha,\beta)}$ is a spherically quasinormal. Thus, consistent with Theorem 3.2, we see that $W_{(\alpha,\beta)}$ is spherically quasinormal if and only if $\widehat{W_{(\alpha,\beta)}} = \widehat{W_{(\alpha,\beta)}}^D$. □

We now recall the class of spherically isometric commuting pairs of operators (cf. [4–6, 28, 30]).

Definition 3.6 A commuting pair $\mathbf{T} \equiv (T_1, T_2)$ is a spherical isometry if $T_1^* T_1 + T_2^* T_2 = I$.

The following result is a straightforward application of Definition 3.6.

Lemma 3.7 *A 2-variable weighted shift* $\mathbf{T} \equiv (T_1, T_2) = W_{(\alpha,\beta)}$ *is a spherical isometry if and only if*

$$\alpha^2_{\mathbf{k}} + \beta^2_{\mathbf{k}} = 1$$

for all $\mathbf{k} \in \mathbb{Z}^{\mathbf{2}}_+$.

By Theorem 3.2, we have:

Corollary 3.8 *A 2-variable weighted shift* $\mathbf{T} \equiv (T_1, T_2)$ *is spherically quasinormal if and only if there exists* $c > 0$ *such that* $\frac{1}{\sqrt{c}}\mathbf{T}$ *is a spherical isometry, that is,* $T_1^* T_1 + T_2^* T_2 = I$.

We pause to recall an important result about spherical isometries.

Lemma 3.9 ([28]) *Any spherical isometry is subnormal.*

Combining Corollary 3.8 and Lemma 3.9, we easily obtain the following result.

Theorem 3.10 *Any quasinormal* 2*-variable weighted shift is subnormal.*

Remark 3.11 (cf. [23, Remark 2.14])

(i) A. Athavale and S. Poddar have recently proved that a commuting spherically quasinormal pair is always subnormal [5, Proposition 2.1]; this provides a different proof of Theorem 3.10.

(ii) In a different direction, let $Q_{\mathbf{T}}(X) := T_1^* X T_1 + T_2^* X T_2$. By induction, it is easy to prove that if $\mathbf{T}$ is spherically quasinormal, then $Q_{\mathbf{T}}^n(I) = (Q_{\mathbf{T}}(I))^n$ $(n \geq 0)$; by [8, Remark 4.6], $\mathbf{T}$ is subnormal. □

4 Construction of Spherically Quasinormal 2-Variable Weighted Shifts

As observed in [24], within the class of 2-variable weighted shifts there is a simple description of spherical isometries, in terms of the weight sequences $\alpha \equiv \{\alpha_{(k_1,k_2)}\}$ and $\beta \equiv \{\beta_{(k_1,k_2)}\}$. Indeed, since spherical isometries are (jointly) subnormal, we know that the unilateral weighted shift associated with the 0-th row in the weight diagram must be subnormal. Thus, without loss of generality, we can always assume that the 0-th row corresponds to a subnormal unilateral weighted shift, and denote its weights by $\{\alpha_{(k,0)}\}_{k=0,1,2,\ldots}$. Also, in view of Corollary 3.8 we can assume that $c = 1$. Using the identity

$$\alpha_{\mathbf{k}}^2 + \beta_{\mathbf{k}}^2 = 1 \quad (\mathbf{k} \in \mathbb{Z}_+^2) \tag{4.1}$$

and the above-mentioned 0-th row, we can compute $\beta_{(k,0)} := \sqrt{1 - \alpha_{k,0}^2}$ for $k = 0, 1, 2, \cdots$. With these new values at our disposal, we can use the commutativity property (1.7) to generate the values of α in the first row (see Fig. 1i); that is,

$$\alpha_{(k,1)} := \alpha_{(k,0)}\beta_{(k+1,0)}/\beta_{(k,0)}.$$

We can now repeat the algorithm, and calculate the weights $\beta_{(k,1)}$ for $k = 0, 1, 2, \cdots$, again using the identity (4.1). This in turn leads to the α weights for the second row, and so on.

This simple construction of spherically isometric 2-variable weighted shifts will allow us to study properties like recursiveness (tied to the existence of finitely atomic Berger measures) and propagation of recursive relations. We pursue this in Sect. 5 below.

5 Recursively Generated Spherically Quasinormal 2-Variable Weighted Shifts

We begin by recalling some terminology and basic results from [15] and [16]. A subnormal unilateral weighted shift W_ω is said to be *recursively generated* if the sequence of moments $\gamma_n(W_\omega)$ admits a finite-step recursive relation; that is, if there exists an integer $k \geq 1$ and real coefficients $\varphi_0, \varphi_1, \cdots, \varphi_{k-1}$ such that

$$\gamma_{n+k} = \varphi_0\gamma_n + \varphi_1\gamma_{n+1} + \cdots + \varphi_{k-1}\gamma_{n+k-1} \quad (\text{all } n \geq 0). \tag{5.1}$$

In conjunction with (5.1) we consider the generating function

$$g_\omega(s) := s^k - (\varphi_0 + \varphi_1 s + \cdots + \varphi_{k-1}s^{k-1}). \tag{5.2}$$

The following result characterizes recursively generated subnormal unilateral weighted shifts.

Lemma 5.1 ([17]) *Let W_ω be a subnormal unilateral weighted shift. The following statements are equivalent.*

(i) W_ω is recursively generated.

(ii) The Berger measure μ of W_ω is finitely atomic, and $\operatorname{supp}\mu \subseteq \mathcal{Z}(g_\omega)$, *where $\mathcal{Z}(g_\omega)$ denotes the zero set of g_ω, that is, the set of roots of the equation $g_\omega = 0$.*

Our first result in this section establishes the propagation of a recursive relation from the 0-th row of a spherically quasinormal 2-variable weighted shift to the first row. Given a 2-variable weighted shift $\mathbf{T} \equiv (T_1, T_2) = W_{(\alpha,\beta)}$, recall from (1.11) the notation W_0 and W_1.

Theorem 5.2 *Let $\mathbf{T}$ be a spherically quasinormal 2-variable weighted shift, and assume that W_0 is recursively generated, with coefficients $\varphi_0, \varphi_1, \ldots, \varphi_{n-1}$; that is,*

$$\gamma_{n+k}(W_0) = \varphi_0\gamma_k(W_0) + \varphi_1\gamma_{k+1}(W_0) + \cdots + \varphi_{n-1}\gamma_{n+k-1}(W_0) \quad (\textit{all } k \geq 0). \tag{5.3}$$

Then W_1 is recursively generated, with the same recursion coefficients.

Proof Since $\mathbf{T} \equiv (T_1, T_2) \in \mathfrak{C}_0$ is spherically quasinormal, by Theorem 3.2, for all $k \geq 0$ observe that

$$\begin{aligned}
\beta^2_{(0,0)}\gamma_{k_1}(W_1) &= \beta^2_{(k_1,0)}\gamma_{k_1}(W_0) \\
&= \left(c - \alpha^2_{(k_1,0)}\right)\gamma_{k_1}(W_0) \\
&= c\gamma_{k_1}(W_0) + \gamma_{k_1+1}(W_0),
\end{aligned}$$

Thus, we have

$$\beta_{(0,0)}^2 \gamma_{n+k_1}(W_1) = c\gamma_{n+k_1}(W_0) - \gamma_{n+k_1+1}(W_0) \tag{5.4}$$

$$= c\left(\sum_{i=0}^{n-1} \varphi_i \gamma_{k_1+i}(W_0)\right) - \sum_{i=0}^{n-1} \varphi_i \gamma_{k_1+i+1}(W_0) \tag{5.5}$$

$$= \sum_{i=0}^{n-1} \varphi_i \gamma_{k_1+i}(W_0)\left(c - \alpha_{(k_1+i,0)}^2\right) \tag{5.6}$$

$$= \sum_{i=0}^{n-1} \varphi_i \gamma_{k_1+i}(W_0)\, \beta_{(k_1+i,0)}^2 \tag{5.7}$$

$$= \beta_{(0,0)}^2 \sum_{i=0}^{n-1} \varphi_i \gamma_{k_1+i}(W_1). \tag{5.8}$$

It follows from (5.8) that

$$\gamma_{n+k_1}(W_1) = \varphi_0 \gamma_{k_1}(W_1) + \varphi_1 \gamma_{k_1+1}(W_1) + \cdots + \varphi_{n-1}\gamma_{n+k_1-1}(W_1). \tag{5.9}$$

Thus, we see that W_1 is a recursively generated weighted shift with the same recursion coefficients; that is, (5.3) holds for W_1. □

A straightforward induction argument yields the following result.

Corollary 5.3 *Let* $\mathbf{T}$ *be a spherically quasinormal* 2*-variable weighted shift, and assume that* W_0 *is recursively generated, with coefficients* $\varphi_0, \varphi_1, \ldots, \varphi_{n-1}$*, and let* $k_2 > 1$*. Then* W_{k_2} *is recursively generated, with the same recursion coefficients.*

In view of Theorem 5.2, one is naturally led to the following question. If W_0 is recursively generated, is it also the case that V_0 is recursively generated? To study this question, we will take advantage of the theory of truncated moment problems in two real variables. (The reader is referred to [18–20] for terminology and basic results.) Here we will only make use of the moment matrix associated with $W_{(\alpha,\beta)}$; that is, the infinite matrix $M(\alpha, \beta)$ whose rows and columns are indexed by $\mathbf{k} \in \mathbb{Z}_+^2$ and whose $(\mathbf{i}, \mathbf{j})$-entry is given by $\gamma_{\mathbf{i}+\mathbf{j}}$. As typically done in the theory of truncated real moment problems, it is natural to label the rows and columns of $M(\alpha, \beta)$ using the homogenous monomials of ascending degree $1, S, T, S^2, ST, T^2, S^2, S^2T, ST^2, T^3, \cdots$. For instance, when we refer to the entry in the position $((1, 2), (0, 1))$, we mean the entry corresponding to row $(1, 2)$ and column $(0, 1)$, that is, the row labeled by the monomial ST^2 and the column labeled by the monomial T.

The proof of the following result is straightforward.

Lemma 5.4 *Let $W_{(\alpha,\beta)}$ be a 2-variable weighted shift, let $c > 0$ and fix $\mathbf{k} \in \mathbb{Z}_+^2$. The following statements are equivalent.*

(i) $\alpha_{\mathbf{k}}^2 + \beta_{\mathbf{k}}^2 = c$.
(ii) $\gamma_{\mathbf{k}+\varepsilon_1} + \gamma_{\mathbf{k}+\varepsilon_2} = c\gamma_{\mathbf{k}}$.

Corollary 5.5 *Let $W_{(\alpha,\beta)}$ be a spherically quasinormal 2-variable weighted shift, with constant $c > 0$. Then the columns of the moment matrix $M(\alpha, \beta)$ satisfy the linear relation $S + T = c\ 1$.*

Corollary 5.6 *Let $W_{(\alpha,\beta)}$ be a spherically quasinormal 2-variable weighted shift, with constant $c > 0$, and let σ and τ be the Berger measures of W_0 and V_0, respectively. Then* $\operatorname{supp} \tau = c - \operatorname{supp} \sigma := \{c - s : s \in \operatorname{supp} \sigma\}$.

Proof Since the columns of the moment matrix $M(\alpha, \beta)$ satisfy the linear relation $S + T = c\ 1$, the Riesz functionals Λ_α and Λ_β for σ and τ (resp.) satisfy the condition

$$\Lambda_\beta(p(t)) = \Lambda_\alpha(p(c - t)),$$

for every polynomial p in one real variable. This immediately leads to the desired result about the supports of the Berger measures. □

We are now ready to prove that for spherically quasinormal 2-variable weighted shifts the property of being recursively generated transfers from the 0-th row in the weight diagram to the 0-th column.

Theorem 5.7 *Let $W_{(\alpha,\beta)}$ be a spherically quasinormal 2-variable weighted shift, with constant $c > 0$, and assume that the unilateral weighted shift W_0 (which corresponds to the 0-th row in the weight diagram of $W_{(\alpha,\beta)}$) is recursively generated. Then the unilateral weighted shift V_0 ((which corresponds to the 0-th column) is also recursively generated.*

Proof The proof is based on a simple observation at the level of the Riesz functional associated with the moment matrix $M \equiv M(\alpha, \beta)$. Since $S+T = c\ 1$ in the column space of M, it follows that, at the level of polynomials in the indeterminates s and t, one can replace any occurrence of s by $c - t$. As a consequence, the same holds for the columns of M, by the functional calculus introduced and studied in [17, 18] and [19]. Thus, the linear relation

$$S^k = \varphi_0 1 + \varphi_1 S + \cdots + \varphi_{k-1} S^{k-1}$$

can be rewritten (in terms of T) as

$$(c\ 1 - T)^k = \varphi_0 1 + \varphi_1 (c\ 1 - T) + \cdots + \varphi_{k-1} (c\ 1 - T)^{k-1}. \tag{5.10}$$

Inspection of (5.10) already shows that T^k can be expressed in terms of columns labeled by monomials of degree up to $k-1$. In what follows we make the recursive relation explicit. Recall that, by the Binomial Theorem,

$$(c\,1-T)^p=\sum_{j=0}^{p}(-1)^j\binom{p}{j}c^{p-j}T^j.$$

As a result, (5.10) becomes

$$\sum_{j=0}^{k}(-1)^j\binom{k}{j}c^{k-j}T^j=\sum_{i=0}^{k-1}\varphi_i\sum_{j=0}^{i}(-1)^j\binom{i}{j}c^{i-j}T^j \tag{5.11}$$

$$=\sum_{j=0}^{k-1}\left[\sum_{i=j}^{k-1}(-1)^j\binom{i}{j}\varphi_i c^{i-j}\right]T^j \tag{5.12}$$

It follows that

$$(-1)^kT^k+\sum_{j=0}^{k-1}\left[(-1)^j\binom{k}{j}c^{k-j}\right]T^j=\sum_{j=0}^{k-1}(-1)^j\left[\sum_{i=j}^{k-1}\binom{i}{j}\varphi_i c^{i-j}\right]T^j, \tag{5.13}$$

and therefore

$$(-1)^kT^k=\sum_{j=0}^{k-1}(-1)^j\left[\sum_{i=j}^{k-1}\binom{i}{j}\varphi_i c^{i-j}-\binom{k}{j}c^{k-j}\right]T^j, \tag{5.14}$$

so that

$$T^k=\sum_{j=0}^{k-1}(-1)^{k-j}\psi_jT^j, \tag{5.15}$$

where

$$\psi_j:=\sum_{i=j}^{k-1}\binom{i}{j}\varphi_i c^{i-j}-\binom{k}{j}c^{k-j}.$$

We have thus found explicitly the recursive coefficients for the moments associated with V_0. This completes the proof. □

Corollary 5.8 *Let $W_{(\alpha,\beta)}$ be a spherically quasinormal 2-variable weighted shift, with constant $c>0$, and assume that the unilateral weighted shift W_0 (which* *corresponds to the* 0*-th row in the weight diagram of* $W_{(\alpha,\beta)}$*) is recursively generated. Let* σ *be the Berger measure of* W_0*, and let* μ *be the Berger measure of* $W_{(\alpha,\beta)}$*. Then*

(i) $\operatorname{supp}\mu \subseteq \operatorname{supp}\sigma \times (c - \operatorname{supp}\sigma)$*; and*
(ii) μ *is finitely atomic.*

Proof Recall that σ and τ are the marginal measures of μ (cf. Definition 6.1). By Lemma 6.2, we know that

$$\operatorname{supp}\mu \subseteq \operatorname{supp}\sigma \times \operatorname{supp}\tau.$$

By Corollary 5.6, we obtain (i). Since σ is finitely atomic, (ii) is now immediate. □

Remark 5.9 Let $W_{(\alpha,\beta)}$ be a spherically quasinormal 2-variable weighted shift, with constant $c > 0$, and assume that the unilateral weighted shift W_0 (which corresponds to the 0-th row in the weight diagram of $W_{(\alpha,\beta)}$) is recursively generated. By Theorem 5.2, W_1 is also recursively generated, and let $\sigma^{(1)}$ be its (finitely atomic) Berger measure. Although the recursive coefficients transfer from W_0 to W_1, it is not necessarily true that σ and $\sigma^{(1)}$ have the same support. By Lemma 6.4 in Appendix, we know that $\sigma^{(1)} \ll \sigma$, so $\operatorname{supp}\sigma^{(1)} \subseteq \operatorname{supp}\sigma$. Example 5.10 below shows that this inclusion may be proper. □

First, we need some terminology.

Given three positive numbers a, b, c such that $0 < a < b < c$, we recall Stampfli's result on the existence of a subnormal unilateral weighted shift, denoted by $W_{(\sqrt{a},\sqrt{b},\sqrt{c})^\wedge}$, whose first three weights are a, b and c [38]. Here we will briefly recall the approach to Stampfli's result presented in [15, 16] and [17]. As proved in those papers, the Berger measure σ of $W_{(\sqrt{a},\sqrt{b},\sqrt{c})^\wedge}$ is finitely atomic, and the coefficients of recursion are given by

$$\varphi_0 = -\frac{ab(c-b)}{b-a} \text{ and } \varphi_1 = \frac{b(c-a)}{b-a}; \tag{5.16}$$

cf. [14, Section 1, p. 81], [15, Example 3.12], [16, Section 3]. Moreover, the atoms t_0 and t_1 are the roots of the equation

$$s^2 - (\varphi_0 + \varphi_1 s) = 0, \tag{5.17}$$

and the densities ρ_0 and ρ_1 uniquely solve the system of equations

$$\begin{cases} \rho_0 + \rho_1 &= 1 \\ \rho_0 s_0 + \rho_1 s_1 &= \alpha_0^2, \end{cases} \tag{5.18}$$

where

$$s_0 := \frac{\varphi_1 - \sqrt{\varphi_1^2 + 4\varphi_0}}{2}, s_1 := \frac{\varphi_1 + \sqrt{\varphi_1^2 + 4\varphi_0}}{2}, \rho_0 := \frac{s_1 - a}{s_1 - s_0}, \text{ and } \rho_1 := \frac{a - s_0}{s_1 - s_0}. \tag{5.19}$$

We can now easily see that $\varphi_0 < 0$, $\varphi_1 > 0$, and $s_0 < s_1 < \varphi_1$. We thus obtain $\sigma = \rho_0 \delta_{s_0} + \rho_1 \delta_{s_1}$, which is the Berger measure of $W_{(\sqrt{a},\sqrt{b},\sqrt{c})^\wedge}$. The recursive relation, at the level of the weights, is

$$\alpha_{k+1}^2 = \varphi_1 + \frac{\varphi_0}{\alpha_k^2} \quad (k \geq 0). \tag{5.20}$$

In view of the preservation of the recursive relation from W_0 to W_1, one might be tempted to claim that all unilateral weighted shifts W_{k_2} corresponding to horizontal rows have Berger measures $\sigma^{(k_2)}$ with the same support. This is not true. What is actually true is that $\operatorname{supp} \sigma^{(k_2)} = \operatorname{supp} \sigma^{(1)}$ for all $k_2 > 1$. The support of $\sigma^{(1)}$, however, might be strictly smaller than the support of σ. We will exhibit this behavior in the following concrete example. Notice that in this example, the 2-variable weighted shift is actually a spherical isometry.

Example 5.10 Consider the 2-variable weighted shift $\mathbf{T} \equiv (T_1, T_2) = W_{(\alpha,\beta)} \in \mathfrak{C}_0$ whose weight diagram is given in Fig. 1ii. That is, W_0 is the Stampfli subnormal completion of the initial segment of weights $\{\sqrt{\frac{2}{3}}, \sqrt{\frac{5}{6}}, \sqrt{\frac{14}{15}}\}$. Using (5.16) one gets at once

$$\varphi_0 = -\frac{\alpha_{(0,0)}^2 \alpha_{(1,0)}^2 \left(\alpha_{(2,0)}^2 - \alpha_{(1,0)}^2\right)}{\alpha_{(1,0)}^2 - \alpha_{(0,0)}^2} = -\frac{1}{3} \text{ and } \varphi_1 = \frac{\alpha_{(1,0)}^2 \left(\alpha_{(2,0)}^2 - \alpha_{(0,0)}^2\right)}{\alpha_{(1,0)}^2 - \alpha_{(0,0)}^2} = \frac{4}{3}. \tag{5.21}$$

It follows that W_0 is subnormal with Berger measure

$$\sigma = \frac{1}{2}\delta_{\frac{1}{3}} + \frac{1}{2}\delta_1.$$

Since

$$\beta_{(k_1,0)} := \sqrt{1 - \alpha_{(k_1,0)}^2} \quad (k_1 \geq 0),$$

direct calculation yields

$$\beta_{(k_1,0)} = \sqrt{\frac{2}{3\left(3^{k_1} + 1\right)}} \quad k_1 \geq 0.$$

Theorem 5.2 says that $W_1 = \text{shift}(\alpha_{(0,1)}, \alpha_{(1,1)}, \cdots)$ is also a recursively generated weighted shift with the same recursion coefficients φ_0 and φ_1. Moreover, the generating function

$$g(t) := t^2 - (\varphi_1 t + \varphi_0)$$

has two distinct real roots

$$0 < s_0 \equiv \frac{1}{3} < s_1 \equiv 1.$$

Let

$$V := \begin{pmatrix} 1 & 1 \\ s_0 & s_1 \end{pmatrix}$$

and let

$$\begin{pmatrix} \rho_0(W_1) \\ \rho_1(W_1) \end{pmatrix} = V^{-1} \begin{pmatrix} \gamma_0(W_1) \\ \gamma_1(W_1) \end{pmatrix}.$$

We then have

$$\sigma^{(1)} = \rho_0(W_1)\,\delta_{s_0} + \rho_1(W_1)\,\delta_{s_1},$$

where $\sigma^{(1)}$ is the Berger measure of $W_1 \equiv \text{shift}(\sqrt{\frac{1}{3}}, \sqrt{\frac{1}{3}}, \cdots)$. A straightforward calculation yields $\rho_0(W_1) = 1$ and $\rho_1(W_1) = 0$. It follows that $\sigma^{(1)} = \delta_{\frac{1}{3}}$, as desired. Moreover, for $k_1 \geq 0$ we have $\beta_{(k_1,1)} = \sqrt{\frac{2}{3}}$. Now, $W_2 = \text{shift}(\sqrt{\frac{1}{3}}, \sqrt{\frac{1}{3}}, \cdots) = W_1$ and, more generally, $W_{k_2} = W_1$ for all $k_2 \geq 1$. We have thus shown that even within the class of spherically isometric 2-variable weighted shifts it is indeed possible to shrink the support of σ as we move from the 0-th row to the remaining rows in the weight diagram. □

We will now state and prove an improved version of Lemma 6.4. We have known this fact for many years, as it was implicit in the proof of [21, Theorem 3.1]. We have also referred to it in research presentations, but somehow we have never had the occasion to give a formal proof. Let us first recall that for $h \geq 1$ we let $\mathcal{L}_h := \bigvee\{e_n : n \geq h\}$ denote the invariant subspace obtained by removing the first h vectors in the canonical orthonormal basis of $\ell^2(\mathbb{Z}_+)$. Thus, if W_ω is subnormal, then the Berger measure of $W_\omega|_{\mathcal{L}_h}$ is $\frac{1}{\gamma_h} s^h d\sigma_\omega(s)$, where $W_\omega|_{\mathcal{L}_h}$ means the restriction of W_ω to the invariant subspace $\mathcal{L}_h$. We can extend this result to the case of 2-variable weighted shifts. We first recall that for an arbitrary 2-variable weighted shift $W_{(\alpha,\beta)}$, we let $\mathcal{M}_j$ (resp. $\mathcal{N}_i$) be the subspace of $\ell^2(\mathbb{Z}_+^2)$ spanned by the canonical orthonormal basis associated to indices $\mathbf{k} = (k_1, k_2)$ with $k_1 \geq 0$ and $k_2 \geq j$ (resp. $k_1 \geq i$ and

$k_2 \geq 0$). If $W_{(\alpha,\beta)}$ is subnormal with Berger measure μ, then the Berger measure of $W_\omega|_{\mathcal{M}_j}$ is $\frac{t^j}{\gamma_{0j}(W_{(\alpha,\beta)})} d\mu(s,t)$. We then have:

Theorem 5.11 *Consider the* 2*-variable weighted shift* $\mathbf{T} \equiv (T_1, T_2) = W_{(\alpha,\beta)}$ *given by Fig. 1i. Let* $\sigma^{(j)}$ *and* $\tau^{(i)}$ *be as in Lemma 6.4. If* $\mathbf{T} \equiv (T_1, T_2) = W_{(\alpha,\beta)}$ *is subnormal, then* $\sigma^{(j+1)} \simeq \sigma^{(j)}$ *and* $\tau^{(i+1)} \simeq \tau^{(i)}$ *for* $(i, j \geq 1)$*, where* $\simeq$ *indicates that the two relevant measures are mutually absolutely continuous.*

Proof Let $R := X \times Y \equiv [0, a_1] \times [0, a_2]$, where $a_k := \|T_k\|$ $(k = 1, 2)$. By Lemma 6.4, we only need to show the following implication:

$$\text{for } j \geq 1,\ \sigma^{(j)} \ll \sigma^{(j+1)};\ \text{ that is, } \sigma^{(j+1)}(E) = 0 \Longrightarrow \sigma^{(j)}(E) = 0 \ \ (\text{for all } E \subseteq X).$$

Since $W_{(\alpha,\beta)} \equiv (T_1, T_2)$ is subnormal, we let μ be the Berger measure of $W_{(\alpha,\beta)}$. Then, by Lemma 6.4, for $j \geq 1$ $d\mu_j(s,t) := \frac{1}{\gamma_{0j}(W_{(\alpha,\beta)})} t^j d\mu(s,t)$, and as a result

$$\begin{aligned} d\mu_{j+1}(s,t) &= \frac{1}{\gamma_{0j}\left(W_{(\alpha,\beta)}\right)} t^{j+1} d\mu(s,t) \\ &= \frac{\gamma_{0j}\left(W_{(\alpha,\beta)}\right)}{\gamma_{0j+1}\left(W_{(\alpha,\beta)}\right)} t\, d\mu_j(s,t) \\ &= \frac{\gamma_{0j-1}\left(W_{(\alpha,\beta)}\right)}{\gamma_{0j+1}\left(W_{(\alpha,\beta)}\right)} t^2 d\mu_{j-1}(s,t). \end{aligned}$$

Suppose now that for $j \geq 1$ and for $E \subseteq X$, $\sigma^{(j+1)}(E) = 0$. Then, by Lemma 6.6, we have that

$$\begin{aligned} \sigma^{(j+1)}(E) &= \mu^X_{j+1}(E) = \mu_{j+1}(E \times Y) = \int_{E\times Y} d\mu_{j+1}(s,t) = 0 \\ &\Longrightarrow \int_{E\times Y} t^2 d\mu_{j-1}(s,t) = 0. \end{aligned}$$

Since $t^2 \geq 0$, we know that $t^2 = 0$ a.e. $[\mu_{j-1}]$ on $E \times Y$; it follows that $t = 0$ a.e. $[\mu_{j-1}]$ on $E \times Y$. Then

$$\int_{E\times Y} t\, d\mu_{j-1}(s,t) = 0 \Longrightarrow \int_{E\times Y} d\mu_j(s,t) = 0 \Longrightarrow \sigma^{(j)}(E) = 0.$$

This completes the proof. □

Remark 5.12 (i) We now refer back to the 2-variable weighted shift $\mathbf{T} \equiv (T_1, T_2) = W_{(\alpha,\beta)}$ constructed in Example 5.10. Since $W_{(\alpha,\beta)}$ is quasinormal, it is also subnormal by Theorem 3.10; let μ be its Berger measure. It is easy to see that

$\mathbf{T}|_{\mathcal{M}} \in \mathfrak{H}_\infty$ and that the Berger measure of $\mathbf{T}|_{\mathcal{M}}$ is $\mu_{\mathcal{M}} = \delta_{\left(\frac{1}{3},\frac{2}{3}\right)}$. By Lemma 6.3 and Fig. 1ii, we can see that $(\mu_{\mathcal{M}})_{ext} = \delta_{\left(\frac{1}{3},\frac{2}{3}\right)}$ and hence it follows that

$$(\mu_{\mathcal{M}})^X_{ext} = \delta_{\frac{1}{3}}.$$

Since $\beta_{00}^2 = \frac{1}{3}$, Lemma 6.3 shows that

$$\mu = \frac{1}{2}\left(\delta_{\left(\frac{1}{3},\frac{2}{3}\right)} + \delta_{(1,0)}\right). \qquad \square$$

The following problem arises naturally.

Problem 5.13 Consider a spherically quasinormal 2-variable weighted shift $\mathbf{T} \equiv (T_1, T_2) = W_{(\alpha,\beta)}$ and let σ be the Berger measure of W_0. Since $W_{(\alpha,\beta)}$ is subnormal by Theorem 3.10, let μ be the Berger measure of $W_{(\alpha,\beta)}$.

(i) Describe μ in terms of σ.
(ii) Assume that W_0 is recursively generated. By Corollary 5.8, we know that μ is finitely atomic, and that $\operatorname{supp}\mu \subseteq \operatorname{supp}\sigma \times (c - \operatorname{supp}\sigma)$. What else can we say? Can we give a concrete formula for the atoms and densities of μ?

In Problem 5.13(ii) we know that W_0 carries all the information about $W_{(\alpha,\beta)}$; therefore, we know that the atoms and densities of μ must algorithmically be obtained from those of σ. Thus, the question refers to finding such algorithm. In Example 5.14 below, we show how one might go about finding a concrete formula for μ.

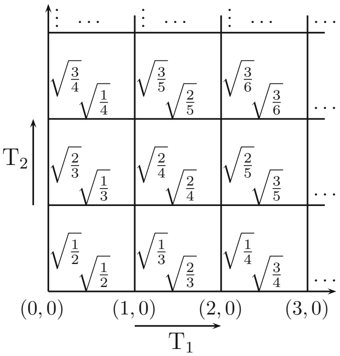
Example 5.14 In Problem 5.13, assume that σ is 2-atomic, and write $\sigma \equiv \lambda_0\delta_{s_0} + \lambda_1\delta_{s_1}$, with $0 \le s_0 < s_1 \le 1$ and $\lambda_0, \lambda_1 > 0$. From Corollary 5.8 we know that

$$\operatorname{supp}\mu \subseteq \{(s_0, c - s_0), (s_0, c - s_1), (s_1, c - s_0), (s_1, c - s_1)\}.$$

Moreover, $\operatorname{supp}\mu$ must have at least two atoms, because σ (and τ) are 2-atomic. Thus, we can postulate that $\mu = \rho_{00}\delta_{(s_0,c-s_0)} + \rho_{01}\delta_{(s_0,c-s_1)} + \rho_{10}\delta_{(s_1,c-s_0)} + \rho_{11}\delta_{(s_1,c-s_1)}$, with $\rho_{ij} \ge 0$ $(i, j = 1, 2)$. We now write the moment equations as follows:

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ s_0 & s_0 & s_1 & s_1 \\ c - s_0 & c - s_1 & c - s_0 & c - s_1 \\ s_0(c - s_0) & s_0(c - s_1) & s_1(c - s_0) & s_1(c - s_1) \end{pmatrix} \begin{pmatrix} \rho_{00} \\ \rho_{01} \\ \rho_{10} \\ \rho_{11} \end{pmatrix} = \begin{pmatrix} \gamma_{(0,0)} \\ \gamma_{(0,1)} \\ \gamma_{(1,0)} \\ \gamma_{(1,1)} \end{pmatrix}. \tag{5.22}$$

Denote the 4×4 matrix in (5.22) by V. A calculation using *Mathematica* [37] shows that $\det V = -(s_1 - s_0)^4 < 0$. It follows that we can always find real numbers $\rho_{00}, \rho_{01}, \rho_{10}, \rho_{11}$ satisfying the moment equations. However, that is not sufficient, since we need to guarantee that these four numbers are nonnegative. We



The 2-variable weighted shift $W_{(\alpha,\beta)}$ whose weight diagram is shown on the left has the following properties:

(i) $\alpha^2_{(k_1,k_2)} + \beta^2_{(k_1,k_2)} = 1$ for all $(k_1, k_2) \in \mathbb{Z}^2_+$.

(ii) $\widehat{W_{(\alpha,\beta)}} = \widehat{W_{(\alpha,\beta)}}^D = W_{(\alpha,\beta)}$.

(iii) $W_{(\alpha,\beta)}$ is a spherical isometry.

Fig. 2 The weight diagram on the left corresponds to the 2-variable weighted shift in Question 5.15

do know that $\gamma_{(0,0)} = 1$, $\gamma_{(0,1)} = \lambda_0 s_0 + \lambda_1 s_1$, $\gamma_{(1,0)} = \lambda_0(c - s_0) + \lambda_1(c - s_1)$ and $\gamma_{(1,1)} = \gamma_{(1,0)}\beta^2_{(1,0)}$. We also know that $\beta^2_{(1,0)} = c - \alpha^2_{(1,0)} = c - \frac{\gamma_{(2,0)}}{\gamma(1,0)}$. Using this information, a calculation with *Mathematica* reveals that $\rho_{01} = \rho_{10} = 0$, and that $\rho_{00} = \lambda_0$ and $\rho_{11} = 1 - \lambda_0$. It follows that

$$\mu = \lambda_0 \delta_{(s_0, c-s_0)} + (1 - \lambda_0)\delta_{(s_1, c-s_1)}.$$

In particular, μ is always 2-atomic. For instance, the Berger measure of the spherical isometry built in Example 5.10 is

$$\mu = \frac{1}{2}(\delta_{(\frac{1}{3},\frac{2}{3})} + \delta_{(1,0)}).$$

This formula for μ is entirely consistent with Remark 5.12. □

We conclude this section with an intriguing question.

Question 5.15 *Let* W_0 *be the Bergman shift* shift$(\sqrt{\frac{1}{2}}, \sqrt{\frac{2}{3}}, \sqrt{\frac{3}{4}}, \cdots)$, *and use Sect. 4 to build a spherically quasinormal* 2*-variable weighted shift* W *(cf. Fig. 2). For this shift the* j*-th row is identical to the* j*-column, for every* $j \geq 0$. *Note also that* W *is a close relative of the Drury-Arveson* 2*-variable weighted shift, in that the* j*-row of* W *is the Agler* A_{j+2} *shift. What is the Berger measure of* W*?*

6 Appendix

For the reader's convenience, in this section, we gather several well-known auxiliary results which are needed for the proofs of the main results in this article. To check subnormality of 2-variable weighted shifts, we introduce some definitions [22, Proposition 3.10].

Definition 6.1

(i) Let μ and ν be two positive Borel measures on a set X. We say that $\mu \le \nu$ on X, if $\mu(E) \le \nu(E)$ for each Borel subset $E \subseteq X$; equivalently, $\mu \le \nu$ if and only if $\int f d\mu \le \int f d\nu$ for all $f \in C(X)$ such that $f \ge 0$ on X.

(ii) Let μ be a probability Borel measure on $X \times Y$, and assume that $\frac{1}{t} \in L^1(\mu)$. The *extremal measure* μ_{ext} (which is also a probability Borel measure) on $X \times Y$ is given by $d\mu_{ext}(s,t) := (1-\delta_0(t))\frac{1}{t\left\|\frac{1}{t}\right\|_{L^1(\mu)}}d\mu(s,t)$.

(iii) Given a Borel measure μ on $X \times Y$, the *marginal measure* μ^X is given by $\mu^X := \mu \circ \pi_X^{-1}$, where $\pi_X : X \times Y \to X$ is the canonical projection onto X. Thus $\mu^X(E) = \mu(E \times Y)$, for every $E \subseteq X$.

Lemma 6.2 *Let μ be a probability Borel measure on $X \times Y$, and let μ^X and μ^Y be the two marginal measures. Then*

$$\operatorname{supp}\mu \subseteq \operatorname{supp}\sigma \times \operatorname{supp}\tau \subseteq X \times Y.$$

Lemma 6.3 ([22, Proposition 3.9] Subnormal Backward Extension) *Assume that $W_{(\alpha,\beta)} \in \mathfrak{H}_0$ (see Fig. 1i) and that $W_{(\alpha,\beta)}|_{\mathcal{M}}$ is subnormal with associated measure $\mu_{\mathcal{M}}$. Then $W_{(\alpha,\beta)}$ is subnormal if and only if the following conditions hold:*

(i) $\frac{1}{t} \in L^1(\mu_{\mathcal{M}})$;

(ii) $\beta_{00}^2 \le \left(\left\|\frac{1}{t}\right\|_{L^1(\mu_{\mathcal{M}})}\right)^{-1}$;

(iii) $\beta_{00}^2 \left\|\frac{1}{t}\right\|_{L^1(\mu_{\mathcal{M}})} (\mu_{\mathcal{M}})_{ext}^X \le \sigma$.

Moreover, if $\beta_{00}^2 \left\|\frac{1}{t}\right\|_{L^1(\mu_{\mathcal{M}})} = 1$, then $(\mu_{\mathcal{M}})_{ext}^X = \sigma$. In the case when $W_{(\alpha,\beta)}$ is subnormal, the Berger measure μ of $W_{(\alpha,\beta)}$ is given by

$$\mu = \beta_{00}^2 \left\|\frac{1}{t}\right\|_{L^1(\mu_{\mathcal{M}})} (\mu_{\mathcal{M}})_{ext} + \left(\sigma - \beta_{00}^2 \left\|\frac{1}{t}\right\|_{L^1(\mu_{\mathcal{M}})} (\mu_{\mathcal{M}})_{ext}^X\right) \times \delta_0. \tag{6.1}$$

Recall that given two positive regular Borel measures μ and ω, μ is said to be absolutely continuous with respect to ω (in symbols, $\mu \ll \omega$) if for every Borel set E, $\omega(E) = 0 \Rightarrow \mu(E) = 0$.

Lemma 6.4 ([21, Theorem 3.3]) *Consider the 2-variable weighted shift $\mathbf{T} \equiv (T_1, T_2) = W_{(\alpha,\beta)}$ given by Fig. 1i. If $\mathbf{T} \equiv (T_1, T_2) = W_{(\alpha,\beta)}$ is subnormal, then $\sigma^{(j+1)} \ll \sigma^{(j)}$ and $\tau^{(i+1)} \ll \tau^{(i)}$ $(i, j \ge 0)$, where $\sigma^{(j)}$ (resp. $\tau^{(i)}$) is the Berger measure of the j-th horizontal slice of T_1 (resp. the i-th vertical slice of T_2).*

Lemma 6.5 ([21]) *Let μ and ν be two regular Borel measures on R, and assume that $\mu \ll \nu$. Then $\mu^X \ll \nu^X$ and $\mu^Y \ll \nu^Y$.*

Lemma 6.6 ([21]) *Let μ be the Berger measure of a subnormal 2-variable weighted shift, and for $j \geq 0$ let $\sigma^{(j)}$ be as in Lemma 6.4. Then $\sigma^{(j)} = \mu_j^X$, where $d\mu_j(s,t) := \frac{1}{\gamma_{0j}(W_{(\alpha,\beta)})} t^j d\mu(s,t)$.*

References

1. A. Aluthge, On p-hyponormal Operators for $0 < p < 1$. Integr. Equ. Oper. Theory **13**, 307–315 (1990)
2. T. Ando, Aluthge transforms and the convex hull of the spectrum of a Hilbert space operator, in *Recent Advances in Operator Theory and Its Applications*. Operator Theory: Advances and Applications, vol. 160 (Springer, Berlin, 2005), pp. 21–39
3. A. Athavale, On joint hyponormality of operators. Proc. Am. Math. Soc. **103**, 417–423 (1988)
4. A. Athavale, On the intertwining of joint isometries. J. Operator Theory **23**, 339–350 (1990)
5. A. Athavale, S. Poddar, On the reflexivity of certain operator tuples. Acta Math. Sci. (Szeged) **81**, 285–291 (2015)
6. K.R.M. Attele, A.R. Lubin, Commutant lifting for jointly isometric operators-a geometrical approach. J. Funct. Anal. **140**, 300–311 (1996)
7. C. Benhida, R. Curto, Spectral properties of the spherical Aluthge transform (in final stages of preparation)
8. S. Chavan, V. Sholapurkar, Rigidity theorems for spherical hyperexpansions. Complex Anal. Oper. Theory **7**, 1545–1568 (2013)
9. M. Chō, I.B. Jung, W.Y. Lee, On Aluthge Transforms of p-hyponormal operators. Integr. Equ. Oper. Theory **53**, 321–329 (2005)
10. J. Conway, in *The Theory of Subnormal Operators*. Mathematical Surveys and Monographs, vol. 36 (American Mathematical Society, Providence, 1991)
11. R. Curto, On the connectedness of invertible n-tuples. Indiana Univ. Math. J. **29**, 393–406 (1980)
12. R. Curto, Joint hyponormality: a bridge between hyponormality and subnormality. Proc. Sympos. Pure Math. **51**, 69–91 (1990)
13. R. Curto, Spectral theory of elementary operators, in *Elementary Operators and Applications*, ed. by M. Mathieu (World Scientific, River Edge, 1992), pp. 3–52
14. R. Curto, An operator-theoretic approach to truncated moment problems, in *Linear Operators*, vol. 38 (Banach Center Publications, Warszawa, 1997), pp. 75–104
15. R. Curto, L. Fialkow, Recursively generated weighted shifts and the subnormal completion problem. Integr. Equ. Oper. Theory **17**, 202–246 (1993)
16. R. Curto, L. Fialkow, Recursively generated weighted shifts and the subnormal completion problem, II. Integr. Equ. Oper. Theory **18**, 369–426 (1994)
17. R. Curto, L. Fialkow, Solution of the truncated complex moment problem with flat data, in *Memoirs of the American Mathematical Society*, vol. 568 (American Mathematical Society, Providence, 1996)
18. R. Curto, L. Fialkow, Flat extensions of positive moment matrices: recursively generated relations, in *Memoirs of the American Mathematical Society*, vol. 648 (American Mathematical Society, Providence, 1998)
19. R. Curto, L. Fialkow, The truncated complex K-moment problem. Trans. Am. Math. Soc. **352**, 2825–2855 (2000)
20. R. Curto, L. Fialkow, Solution of the singular quartic moment problem, J. Operator Theory **48**, 315–354 (2002)

21. R. Curto, J. Yoon, Disintegration-of-measure techniques for multivariable weighted shifts. Proc. Lond. Math. Soc. **93**, 381–402 (2006)
22. R. Curto, J. Yoon, Jointly hyponormal pairs of subnormal operators need not be jointly subnormal. Trans. Am. Math. Soc. **358**, 5139–5159 (2006)
23. R. Curto, J. Yoon, Toral and spherical Aluthge transforms for 2-variable weighted shifts. C. R. Acad. Sci. Paris **354**, 1200–1204 (2016)
24. R. Curto, J. Yoon, Aluthge transforms of 2-variable weighted shifts. Integr. Equ. Oper. Theory **90**, 52 (2018)
25. R. Curto, S.H. Lee, J. Yoon, k-hyponormality of multivariable weighted shifts. J. Funct. Anal. **229**, 462–480 (2005)
26. R. Curto, S.H. Lee, J. Yoon, Quasinormality of powers of commuting pairs of bounded operators (preprint 2019; 18 pp., in preprint form)
27. K. Dykema, H. Schultz, Brown measure and iterates of the Aluthge transform for some operators arising from measurable actions. Trans. Am. Math. Soc. **361**, 6583–6593 (2009)
28. J. Eschmeier, M. Putinar, Some remarks on spherical isometries. Operator Theory Adv. Appl. **129**, 271–291 (2001)
29. R. Gellar, L.J. Wallen, Subnormal weighted shifts and the Halmos-Bram criterion. Proc. Jpn. Acad. **46**, 375–378 (1970)
30. J. Gleason, Quasinormality of Toeplitz tuples with analytic symbols. Houst. J. Math. **32**, 293–298 (2006)
31. N.P. Jewell, A.R. Lubin, Commuting weighted shifts and analytic function theory in several variables. J. Operator Theory **1**, 207–223 (1979)
32. I.B. Jung, E. Ko, C. Pearcy, Aluthge transform of operators. Integr. Equ. Oper. Theory **37**, 437–448 (2000)
33. I.B. Jung, E. Ko, C. Pearcy, Spectral pictures of Aluthge transforms of operators. Integr. Equ. Oper. Theory **40**, 52–60 (2001)
34. J. Kim, J. Yoon, Aluthge transforms and common invariant subspaces for a commuting n-tuple of operators. Integr. Equ. Oper. Theory **87**, 245–262 (2017)
35. J. Kim, J. Yoon, Taylor spectra and common invariant subspaces through the Duggal and generalized Aluthge transforms for commuting n-tuples of operators. J. Operator Theory **81**(1), 81–105 (2019)
36. S.H. Lee, W.Y. Lee, J. Yoon, Subnormality of Aluthge transform of weighted shifts. Integr. Equ. Oper. Theory **72**, 241–251 (2012)
37. Mathematica, Version 8.0 (Wolfram Research, Champaign 2011)
38. J. Stampfli, Which weighted shifts are subnormal? Pac. J. Math. **17**, 367–379 (1966)
39. J.L. Taylor, A joint spectrum for several commuting operators. J. Funct. Anal. **6**, 172–191 (1970)
40. J.L. Taylor, The analytic functional calculus for several commuting operators. Acta Math. **125**, 1–48 (1970)
41. T. Yamazaki, An expression of spectral radius via Aluthge transformation. Proc. Am. Math. Soc. **130**, 1131–1137 (2002)

Remarks on the Interplay Between Algebra and PDE

Dmitry Khavinson

Dedicated to the memory of S. Shimorin, an extraordinary mathematician and a kind and gentle man

Abstract We discuss Hesse's conjecture for homogeneous polynomials and Korenblum's conjecture on algebras of harmonic functions from the standpoint of nonlinear first-order PDE. Also, we extend a recent theorem of McKinley and Shekhtman for homogeneous polynomial partial differential operators to a wider class of linear PDE with entire coefficients.

1 Hesse's Conjecture

In 1859, Hesse [5] conjectured that if a homogenous polynomial u of $N > 1$ variables has a vanishing Hessian $\left(\frac{\partial^2 u}{\partial x_j \partial x_k}\right)_{j,k=1}^N$, then the partial derivatives $\frac{\partial u}{\partial x_i}$, $j = 1, \ldots, N$ are linearly dependent. In other words, Hess $u \equiv 0 \Leftrightarrow \nabla u := grad\, u : \mathbb{C}^N \to$ hyperplane. For example, let $N = 2$ and $u(x, y)$ is a homogeneous C^2-function of degree of homogeneity $k+1$, such that Hess $u = u_{xx}u_{yy} - u_{xy}^2 \equiv 0$. Let $u_x = f$, $u_y = g$, f, g are homogeneous of degree k. Then, $f_x g_y - f_y g_x = 0$ implies $\frac{f_x}{g_x} = \frac{f_y}{g_y} := \lambda$, while by homogeneity, $x f_x + y f_y = k\, f$ and $x g_x + y g_y = kg = \frac{1}{\lambda} x f_x + \frac{1}{\lambda} y\, f_y = \frac{k}{\lambda} f$. So, $f = \lambda g$ and $f_x = \lambda_x g + \lambda g_x$. Hence, $\lambda_x \equiv 0$ and, similarly, $\lambda_y \equiv 0$. Thus, $\lambda \equiv \text{const} = c$, $u_x = c u_y$ and ∇u maps $\mathbb{C}^2$ into a line.

Gordan and Nöther [2] showed that Hesse's conjecture holds for $N = 2, 3, 4$ but is false for $n \geq 5$ in view of the following example of a cubic in 5 variables: $u(x_1, \ldots, x_5) = x_1 x_4^4 + x_2 x_4 x_5 + x_3 x_5^2$. Indeed, denoting $D_j u = \frac{\partial u}{\partial x_j}$, we have $(D_1 u)(D_3 u) - (D_2 u)^2 \equiv 0$. Hence, the components of ∇u are algebraically

D. Khavinson (✉)
Department of Mathematics and Statistics, University of South Florida, Tampa, FL, USA
e-mail: dkhavins@usf.edu

A. Aleman et al. (eds.), *Analysis of Operators on Function Spaces*,
Trends in Mathematics, https://doi.org/10.1007/978-3-030-14640-5_8

dependent, so Hess $u \equiv 0$ and $\nabla u : \mathbb{C}^5 \to \{x_1x_3 - x_2^2 = 0\}$—cf. [10] for further discussion.

Note u also satisfies a linear PDE, $D_1D_3\,u - D_2^2\,u = 0$. In other words, if we denote by $P\,(x_1, x_2, x_3) = x_1x_3 - x_2^2$, a quadratic homogeneous polynomial, then Gordan–Nöther quintic satisfies two equations: a nonlinear one, $P(\nabla u) = 0$; and a linear one, $P(D)(u) = 0$. We shall return to this point later in the discussion—cf. [10].

2 The Higher Ground: General Nonlinear First-Order PDE

Looking for the higher ground, one might ask whether if u, a holomorphic function, satisfies a "purely" nonlinear equation $F(\nabla u) = 0$, with $F : \mathbb{C}^N \to \mathbb{C}$ being an entire or a meromorphic function with no linear factors, then the choices for u to be a global solution of such nonlinear equation are severely limited—e.g., perhaps forcing u to be linear. The Gordan–Nöther example, though crashing such hopes in general, is not overly satisfying since their u is a function of 5 variables while $F := x_1x_3 - x_2^2$ is a function of only 3 variables so F vanishes on the 2-dimensional linear subspace $\{(x_4, x_5)\}$. The following result is relevant to our discussion.

Theorem 1 (Khavinson [7]) *If an entire function u solves the (eiconal) equation* $u_x^2 + y_y^2 - 1 = 0$*, then u is linear.*

The proof was based on some elementary trick, thus missing the "correct," much more general, theorem.

Theorem 2 (Hemmati (Guerra) [4]) *If* $F : \mathbb{C}^2 \to \mathbb{C}$ *is a meromorphic, purely nonlinear (cf. above) function and u is a meromorphic in* $\mathbb{C}^2$ *solution of* $F\left(u_x, u_y\right) = 0$*, then u is a linear function.*

Thus, in particular, if for a meromorphic in $\mathbb{C}^2$ function u, the gradient map, grad $u : \mathbb{C}^2 \to V$, maps $\mathbb{C}^2$ into an algebraic, nonlinear and irreducible variety V, u is a linear function and the grad u is a constant map—cf. [4]. This is, of course, a far-reaching generalization of Hesse's conjecture for $N = 2$. We refer to the survey [6] for recent extensions and generalizations of the Hemmati (Guerra) theorem.

Remark 1

(i) Not only is Theorem 2 more general than its predecessor, Theorem 1, but its proof is much shorter and more to the point. Namely, it is easy to check [4] that the characteristics for $F\left(u_x, u_y\right)$ are all straight lines. Also, u_x, u_y stay constant on characteristics while nonlinearity implies that these characteristic lines have different slopes. This yields multivaluedness of u_x, u_y at the intersection points, thus implying that those functions have branching singularities and, hence, cannot be meromorphic.

(ii) Also, as another illustration of the failure of Hesse's conjecture in higher dimensions, Theorem 1 already fails in $\mathbb{C}^3$. The function $z - \varphi(x + iy) := u(x, y, z)$ satisfies the eiconal in $\mathbb{C}^3$ for any entire function φ of one variable. Moreover, in higher dimensions there are more and more opportunities for entire solutions of the eiconal $\sum_1^N \left(\frac{\partial u}{\partial z_j}\right)^2 = 1$. Take in $\mathbb{C}^5$, for example, $u = \varphi\,(z_1 + i\,z_2) + \psi\,(z_3 + i\,z_u) + z_5$ with entire φ, ψ, etc.

(iii) It is worth noticing that nonlinear equations in $\mathbb{R}^N$ are even more rigid. For example, as is well-known (cf. the references in [4, 7]), any C^1 solution u in $\mathbb{R}^N$ of $\sum_1^N u_{x_j}^2 = 1$ that is real-valued is linear. Indeed, the eiconal equation describes the velocity of light moving along the normals to the level surface with the constant speed (=1). If the level surfaces of u have nontrivial curvatures, the normals will intersect causing for the solution to become multivalued.

3 Korenblum's Conjecture

What happens when a solution of a linear PDE generates an algebra of solutions? Consider the following example.

Example 1 Let $P = \sum_1^N x_j^2$, $x_j \in \mathbb{R}$, so $P(D) = \Delta$. If $\Delta u = 0$, and $\Delta u^2 = 0$, then $\Delta u^2 = 2u\Delta u + 2\sum_1^N \left(\frac{\partial u}{\partial x_j}\right)^2 = 0$, thus implying that u also satisfies a nonlinear equation $(\text{grad } u)^2 = 0$, a similar equation to the eiconal. In the latter case, one can easily check that for all $k \in \mathbb{N}$, $\Delta\left(u^k\right) = 0$, thus u generates an algebra of harmonic functions. For example, $\Delta u^3 = u\Delta u^2 + u^2\Delta u + 2u(\text{grad } u)^2 = 0$, etc. In two variables, $\sum_1^2 \left(\frac{\partial u}{\partial x_j}\right)^2 = 0$ is equivalent to either $\frac{\partial u}{\partial x_1} + i\,\frac{\partial u}{\partial x_2} = 0$, or $\frac{\partial u}{\partial x_1} - i\,\frac{\partial u}{\partial x_2} = 0$, thus making u either a holomorphic or an anti-holomorphic function.

Korenblum [9] in the late 1970s conjectured that if $u \in C^2(\Omega)$, $\Omega \subset \mathbb{R}^3$ is a domain, and $\Delta u = \Delta u^2 = 0$ (and then $\Delta u^k = 0$, $k \in \mathbb{N}$), then, after an appropriate rotation of coordinates, u must be either an analytic or an anti-analytic complex-valued function in two dimensions. Korenblum announced several proofs of the conjecture, all of which contained gaps.

The reason was that, as stated, the conjecture is false and the intensely developing theory of harmonic morphisms (cf., e. g., [1]) provides many counterexamples.

However, if we consider a global version of the conjecture, it might as well be true.

The following unpublished result by the author verifies the conjecture in the category of polynomials.

Theorem 3 (DK 1992, Unpublished) *If u is a polynomial in $\mathbb{R}^3$ and $\Delta u = \Delta u^2 = 0$, then after an appropriate rotation of the coordinates, u must become an analytic or an anti-analytic complex polynomial in 2 dimensions.*

The proof rests on the Lemma (DK 1992, unpublished), characterizing carriers of singularities of harmonic functions in $\mathbb{C}^3$.

Lemma 1 ([8, Prop. 20.1]) *Let $\varphi = \varphi(z_1, z_2, z_3)$ be a homogeneous polynomial of degree m such that the variety $\Gamma := \left\{z \in \mathbb{C}^3 : \varphi(z) = 0\right\}$ is everywhere characteristic (cf., e.g. [8, pp. 16, 53, 151] with respect to $\Delta := \sum_1^3 \frac{\partial^2}{\partial z_i^2}$, i. e., $\sum_1^3 \left(\frac{\partial \varphi}{\partial z_i}\right)^2 = 0$ on Γ. Then, up to a constant factor, either $\varphi(z_1, z_2, z_3) = \left(\sum_1^3 \alpha_j z_j\right)^m$, where $\alpha_j \in \mathbb{C}$ are constants such that $\sum_1^3 \alpha_j^2 = 0$, i.e., Γ is a characteristic (w.r.t. Δ) plane, or $\varphi(z_1, z_2, z_3) = \left(\sum_1^3 z_j^2\right)^{\frac{m}{2}}$, and Γ is an isotropic cone.*

For our purposes, we need the following obvious corollary.

Corollary 1 *If $\varphi(z_1, z_2, z_3)$ is a homogeneous polynomial of degree m satisfying an "eiconal" equation $\sum_1^3 \left(\frac{\partial \varphi}{\partial z_i}\right)^2 = 0$ in $\mathbb{C}^3$, then, up to a constant factor, $\varphi = \left(\sum_1^3 \alpha_j z_j\right)^m$, $\sum_1^3 \alpha_1^2 = 0$.*

We shall sketch the proof of the lemma later. Now, let us finish the proof of the theorem.

Let $u = u_0 + \cdots + u_m$, where u_j are homogeneous harmonic polynomials of degree $j \leq m$. Then, clearly, the senior term u_m satisfies $\Delta u_m^2 = 0$, hence $\sum_1^3 \left(\frac{\partial u_m}{\partial z_i}\right)^2 \equiv 0$ and, by Corollary 1, $u_m = \left(\sum_1^3 \alpha_j z_i\right)^m$, with $\sum_1^3 \alpha_j^2 = 0$. Rotating the coordinate system in $\mathbb{C}^3$ we can assume without loss of generality that $u_m = c_m (z_1 + i\, z_2)^m$, where c is a constant. Now $u_{m-1} u_m$ is harmonic as well as the second senior term in the expansion of u^2 and since u_m^2 is harmonic. Therefore, $0 = \Delta(u_{m-1} u_m) = u_{m-1} \Delta u_m + u_m \Delta u_{m-1} + 2 \operatorname{grad} u_{m-1} \cdot c_m (1, i)(z_1 + i\, z_2)^{m-1} = 2C_m \left(\frac{\partial u_{m-1}}{\partial z_1} + i \frac{\partial u_{m-1}}{\partial z_2}\right)(z_1 + i\, z_2)^{m-1}$. Hence, $\frac{\partial u_{m-1}}{\partial z_1} + i \frac{\partial u_{m-1}}{\partial z_2} = 0$, yielding $u_{m-1} = c_{m-1} (z_1 + i\, z_2)^{m-1} + b\, z_3^{m-1}$. But $\Delta u_{m-1} = b(m-1)(m-2)\, z_3^{m-3} = 0$, yielding $b = 0$ and $u_{m-1} = c_{m-1} (z_1 + i\, z_2)^{m-1}$. Continuing this "backward" induction, we conclude that $u = P(z_1 + i\, z_2)$, where $P(u)$ is a polynomial of degree m of one variable. Thus, it remains to indicate the proof of the lemma.

Here are the main steps—cf. [8, Ch. 20, Sec. 2].

1. Solving $\varphi(z) = 0$ for one of the variables, say z_3, we obtain on $\Gamma = \{\varphi(z) = 0\}$, $z_3 = \psi(z_1, z_2)$, whose ψ, as is easily-verified, satisfies an eiconal equation $\left(\psi_j := \frac{\partial\varphi}{\partial z_j}, j = 1, 2\right)$, $\psi_1^2 + \psi_2^2 = -1$. φ is homogeneous of order m, so $\sum_1^3 z_j \varphi_j = m\varphi$, and the implicit differentiation yields $\psi_j = -\frac{\varphi_j}{\varphi_3}$, $j = 1, 2$, so $-z_1\varphi_3\psi_1 - z_2\varphi_3\psi_2 + z_3\varphi_3 = m\,\varphi = 0$ on Γ.
2. Substituting $z_3 = \phi(z_1, z_2)$, we conclude that $z_1\psi_1 + z_2\psi_2 = \psi$, i.e., ϕ is homogeneous of order 1 function in 2 variables. Switching to polar coordinates r, θ we can write $\phi = rf(\theta)$ and it is easy to check that f satisfies an ODE $\left(f'\right)^2 + f^2 = 1$. Differentiating the latter equation we obtain a second-order ODE that factors easily producing two solutions: (I) $f = \pm 1$, in which case Γ is an isotropic cone $\left\{\sum_1^3 z_j^2 = 0\right\}$, or (II) $f = \beta_1 \cos\theta + \beta_2 \sin\theta$, $\beta_1^2 + \beta_2^2 = 1$, in which case Γ is a plane.

Remarks

(i) In view of the results on global solutions of the eiconal equations in 2D described in Sect. 2, Korenblum's conjecture holds for entire functions u in $\mathbb{C}^3$ as well. Indeed, as before, $\Delta u^2 = 0 \Rightarrow \sum_1^3 \left(\frac{\partial u}{\partial z_i}\right)^2 \equiv 0$, so on a level surface $\{u = c\}$, writing $z_3 = \psi(z_1, z_2)$, we have $\left(\psi_{z_1}\right)^2 + \left(\psi_{z_2}\right)^2 = -1$, i.e., ψ is a "global" solution of an eiconal, and hence must be linear. Therefore, all level surfaces of u are planes, and after a rotation, we conclude that $u = f(z_1 \pm i\, z_2)$, where f is an entire function of one variable.

(ii) With appropriate modifications one can show that an extended Korenblum's conjecture holds for polynomials in N variables but the statement must be adjusted, and loses its esthetic appeal. For example, in $\mathbb{C}^4$, $u = f(z_1 + i\, z_2) + g(z_3 - i\, z_4)$, where f, g are analytic functions of one variable, satisfy $\Delta u = \Delta u^2 = \cdots = \Delta u^k = \cdots = 0$. In higher dimensions there are even more opportunities to group the variables according to the same principle by taking corresponding vectors $(0, \ldots, \alpha_1, 0, \ldots, 0, \alpha_k, 0, \ldots)$, $k \le N$, $\sum_1^k \alpha_1^2 = 0$ in the isotropic cone $\Gamma_0 = \left\{z : \sum_1^N z_j^2 = 0\right\}$ and applying functions of one variable to dot products of these vectors with $z = (z_1, \ldots, z_N)$. We leave it to the interested reader to draw out the corresponding statements.

4 The McKinley–Shekhtman Conjecture

Recall that the Gordan–Nöther homogeneous cubic u in Sect. 1 satisfies two homogeneous equations: $P(\text{grad } u) = 0$ (first-order nonlinear equation), where $P(z_1, \ldots, z_5) = z_1 z_3 - z_2^2$, and a linear equation $P(D)\, u = D_1 D_3 u - D_2^2 u = 0$. In a recent elegant paper [10], McKinley and Shekhtman suggested that this is part of a general phenomenon.

Conjecture 1 (McKinley–Shekhtman [10]) Let P, u be homogeneous polynomials. If $P(\text{grad } u) = 0$, then $P(D)\, u = 0$.

The conjecture is based on the general feeling, underscored in Sect. 2, that global solutions of the first-order nonlinear PDE are quite special and scarce.

Example 2 As was shown in Sect. 3, a homogeneous polynomial u in $\mathbb{C}^3$ satisfying an "eiconal" $\sum_1^3 \left(\frac{\partial u}{\partial z_i}\right)^2 \equiv 0$ has a very special form $c \left(\sum_1^3 \alpha_j z_i\right)^m$, $\sum_1^3 \alpha_j^2 = 0$, thus obviously satisfying $\Delta u = 0$. We refer the reader to [10], where several special cases of the above conjecture are verified.

Also in [10], the following weak converse to the M–S conjecture is proved.

Theorem 4 ([10]) *Let P be a homogeneous polynomial while u is a polynomial. If $P(D)\left[f^k\right] \equiv 0$ for all $k \in \mathbb{N}$, then $P(\text{grad } f) \equiv 0$.*

The proof in [10] is based on clever algebraic manipulations. Theorem 4 unexpectedly has a nice implication in the approximation theory based on the following result by Pinkus and Wajnryb [11].

Theorem 5 ([11]) *Let $f \in \mathbb{C}[z_1, \ldots, z_n)$ be a polynomial, then the following are equivalent:*

(i) $\mathscr{P}(f) := \text{Span}\left\{[f(\cdot + b)]^k : b \in \mathbb{C}^N, k \in \mathbb{N}\right\} \neq \mathbb{C}[z_1, \ldots, z_N]$.
(ii) *$\exists$ polynomial $P : P(D)\left[f^k\right] = 0$, for all $k \in \mathbb{N}$*
(iii) *$\overline{\mathscr{P}(f)} \neq C\left(\mathbb{C}^N\right)$ with respect to the usual topology of convergence on compact subsets of $\mathbb{C}^N$. Invoking this, a nice corollary to Theorem 4 is given in [10].*

Invoking this, a nice corollary to Theorem 4 is given in [10].

Corollary 2 *Let f be a homogeneous polynomial. If $P(f) \neq C[z_1, \ldots, z_N]$, then there exists a homogeneous polynomial $P : P(\text{grad } f) \equiv 0$, i.e., $\text{grad } f : \mathbb{C}^N \to \mathbb{C}^N$ maps $\mathbb{C}^N$ into an algebraic variety and, hence, $\text{Hess } f \equiv 0$.*

For the proof one just takes a senior homogeneous part of a polynomial guaranteed by Theorem 5 and applies Theorem 4. However, applying the standard classical result in PDE known as the Delassus–Le Roux theorem (cf. [8, pp. 22, 153], one can substantially expand Theorem 4 in [10] and the proof becomes much more straightforward and transparent.

Theorem 6 *Let* $P(D) := \sum_{|\alpha|\le m} a_\alpha(z)\, D^\alpha$, $z = (z_1, \dots, z_N)$, $\alpha = (\alpha_1, \dots, \alpha_N)$, $\alpha_j \in \mathbb{N} \cup \{0\}$, $D^\alpha = \left(\frac{\partial}{\partial z_1}\right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial z_N}\right)^{\alpha_N}$ *be a linear differential operator with entire coefficients* a_α. *Let* $u : \mathbb{C}^N \to \mathbb{C}$ *be an entire function and* $P(D)\left[u^k\right] = 0$, *for all* k *in some arithmetic progression (e.g.,* $k \in \mathbb{N}$, *or* $k = 2n + 1$, $n \in \mathbb{N}$, *etc.). Then,* $\sum_{|\alpha|=m} a_\alpha(z)(\text{grad } u)^\alpha = \sum_{|\alpha|=m} a_\alpha(z) \left(\frac{\partial u}{\partial z_1}\right)^{\alpha_1} \cdots \left(\frac{\partial u}{\partial z_N}\right)^{\alpha_N} \equiv 0$. *Thus,* grad u *maps* $\mathbb{C}^N$ *into an analytic hypersurface and, hence,* Hess $u \equiv 0$.

(Theorem 4 follows at once from Theorem 6 when $P = \sum_{|\alpha|=m} a_\alpha z^\alpha$, a homogeneous polynomial, i.e., $P(D)$ is a constant coefficients operator.)

The following result of Delassus–Le Roux is the key.

Lemma 2 (cf. [8, pp. 22, 153 and the References there]) *Let* $\Gamma := \{z : \varphi(z) = 0\}$ *be a non-singular analytic hypersurface in* $\mathbb{C}^N$ *and* v *be a holomorphic solution of* $P(D)\, v = 0$ *in* $\mathbb{C}^N \setminus \Gamma$ *and* v *is singular everywhere on* Γ. *Then,* Γ *is everywhere characteristic with respect to* $P(D)$, *i.e.,* $\sum_{|\alpha|=m} a_\alpha(z)(\text{grad } \varphi)^\alpha \equiv 0$ *on* Γ.

Remark 2 The Delassus–Le Roux theorem says simply that the singularities of solutions of linear analytic PDE "propagate" through $\mathbb{C}^N$ exclusively along characteristic surfaces.

From Lemma 2, Theorem 6 follows almost at once.

Proof First assume, for the sake of clarity, $P(D)\left[u^k\right] = 0$, for all $k \in N$. For any $c \in C$, in an open neighborhood where $|u| < |c|$ we have $f := \frac{1}{c-u} = 1/c \sum_0^\infty \frac{u^k}{c^k}$, and the series converges. Hence, by the hypothesis, $P(D)(f) = 0$ in that neighborhood, and by analytic continuation everywhere in $\mathbb{C}^N \setminus \{u = c\}$. By Lemma 2, $\Gamma_c := \{u = c\}$ must be everywhere characteristic with respect to $P(D)$, i.e., $\sum_{|\alpha|=m} a_\alpha(z)(\text{grad}\, u)^\alpha \equiv 0$ on Γ_c. But taking a continual family of Γ_c, c runs over an open set in $\mathbb{C}$, we arrive at the conclusion of the theorem.

The proof is easily modified to establish the theorem in full generality. Indeed, if the hypothesis holds for $k = n\ell + d$, $\ell, d \in \mathbb{N}$, fixed, $n = 1, 2, \dots$ we can always write

$$f_c := \frac{u^d}{c^\ell - u^\ell} = \sum_{n=0}^{\infty} \frac{u^d}{c^\ell} \left(\frac{u}{c}\right)^{n\ell} = \sum_{r=0}^{\infty} \frac{u^{n\ell+d}}{c^{(n+n\ell)}}$$

and then proceed exactly as before.

Remarks

(i) The classical ("calculus") proof of the Delassus–Le Roux theorem can be found in [3, Ch. 3]. A modern proof based on the elementary but far-reaching extension of the Cauchy–Kovalevskaya theorem due to Zerner (1971) is in [8, pp. 22, 153].

(ii) Instead of the family of functions $\left\{\frac{1}{u-c}, c \in \mathbb{C}\right\}$, one can, of course, take dilations of any function $f(u)$ with finitely many singularities on the circle of convergence of its Taylor series. We leave the straightforward details of formulating the corresponding result to the reader.

Acknowledgement The author's research was supported by the Simons Foundation Grant 513381.

References

1. P. Baird, J.C. Wood, Harmonic morphisms between Riemannian manifolds, in *London Mathematical Society Monographs. New Series*, vol. 29 (The Clarendon Press, Oxford University Press, Oxford, 2003). http://dx.doi.org/10.1093/acprof:oso/9780198503620.001.0001
2. P. Gordan, M. Nöther, Ueber die algebraischen Formen, deren Hesse'sche Determinante identisch verschwindet. Math. Ann. **10**(4), 547–568 (1876). http://dx.doi.org/10.1007/BF01442264
3. J. Hadamard, *Lectures on Cauchy's Problem in Linear Partial Differential Equations* (Dover Publications, New York, 1953)
4. J.E. Hemmati, Entire solutions of first-order nonlinear partial differential equations. Proc. Am. Math. Soc. **125**(5), 1483–1485 (1997). http://dx.doi.org/10.1090/S0002-9939-97-03881-1
5. O. Hesse, Zur Theorie der ganzen homogenen Functionen. J. Reine Angew. Math. **56**, 263–269 (1859). http://dx.doi.org/10.1515/crll.1859.56.263
6. P.C. Hu, B.Q. Li, On complex solutions of certain partial differential equations. Complex Variables Elliptic Equ. **62**(10), 1492–1505 (2017). http://dx.doi.org/10.1080/17476933.2017.1316978
7. D. Khavinson, A note on entire solutions of the eiconal [eikonal] equation. Am. Math. Mon. **102**(2), 159–161 (1995). http://dx.doi.org/10.2307/2975351
8. D. Khavinson, E. Lundberg, Linear holomorphic partial differential equations and classical potential theory, in *Mathematical Surveys and Monographs*, vol. 232 (American Mathematical Society, Providence, 2018)
9. B.I. Korenblum, Unknown. Oral Communication (2018)
10. T. McKinley, B. Shekhtman, On polynomials with vanishing Hessians and some density problems, in *Approximation Theory XV: San Antonio 2016*. Springer Proceedings in Mathematics and Statistics, vol. 201 (Springer, Cham, 2017), pp. 269–277
11. A. Pinkus, B. Vaĭnrib, On a problem of approximation by means of multidimensional polynomials. Uspekhi Mat. Nauk **50**(2(302)), 89–110 (1995). http://dx.doi.org/10.1070/RM1995v050n02ABEH002057

Which Quartic Polynomials Have a Hyperbolic Antiderivative?

Rajesh Pereira

In memory of Serguei Shimorin

Abstract Every linear, quadratic or cubic polynomial having all real zeros is the derivative of a polynomial having all real zeros. The statement is false for higher degree polynomials. In particular, not every fourth degree polynomial with real zeros is the derivative of a polynomial having all real zeros. We derive a necessary and sufficient condition for a quartic polynomial to be the derivative of a polynomial having all real zeros. This condition is a single quadratic form inequality involving the zeros of the quartic polynomial.

Keywords Geometry of polynomials · Hyperbolic polynomials · Quartics

Mathematics Subject Classification (2010) Primary 26C10; Secondary 26D05

1 Introduction

The relationship between the zeros of a polynomial and those of its derivative has been of significant interest to mathematicians for at least three centuries. Serguei Shimorin has worked in this area [3]. In this paper, we will study polynomials having all of their zeros on the real line; these are sometimes called hyperbolic polynomials. It is a simple consequence of Rolle's theorem that the derivative of a hyperbolic polynomial is a hyperbolic polynomial.

This work supported by NSERC and the hospitality of the Mittag-Leffler Institute.

R. Pereira (✉)
Department of Mathematics and Statistics, University of Guelph, Guelph, ON, Canada
e-mail: pereirar@uoguelph.ca

A. Aleman et al. (eds.), *Analysis of Operators on Function Spaces*,
Trends in Mathematics, https://doi.org/10.1007/978-3-030-14640-5_9

The converse to this is false. A hyperbolic polynomial of degree three or less always has a hyperbolic antiderivative. However for $n \geq 4$, there are nth degree hyperbolic polynomials which have no hyperbolic antiderivatives at all. The example $p(x) = (x-1)^2(x-4)^2$ was given in [1].

It would be desirable to have a systematic test for this. Suppose $Q(z)$ is an $(n+1)$th degree monic hyperbolic polynomial with zeros $z_1 \geq z_2 \geq z_3 \geq \ldots \geq z_n \geq z_{n+1}$. Then $Q(z)$ is nonnegative on $[z_{2j+1}, z_{2j}]$ and nonpositive on $[z_{2j}, z_{2j-1}]$ for all $j : 1 \leq j \leq \frac{n}{2}$. Now let $w_1 \geq w_2 \geq w_3 \geq \ldots \geq w_n$ be the zeros of $Q'(z)$. By Rolle's theorem $z_{j+1} \leq w_j \leq z_j$ and hence $Q(w_{2j}) \geq 0$ and $Q(w_{2j-1}) \leq 0$ for all j. Conversely if Q is a real monic $(n+1)$th degree polynomial and there exists $w_1 \geq w_2 \geq w_3 \geq \ldots \geq w_n$ with $Q'(w_j) = 0$, $Q(w_{2j}) \geq 0$ and $Q(w_{2j-1}) \leq 0$ for all j, then Q is hyperbolic by the Intermediate Value theorem.

This is a nice characterization, however it is in terms of $Q(z)$. If we start with the polynomial $p(z)$ and find an antiderivative, there is no guarantee that we will get the particular $Q(x)$ which has all of its zeros real, we will instead get $P(z) = Q(z) + c$ for some arbitrary real c. This will shift everything by c which gives us the criterion of Souroujon and Stoyanov.

Lemma 1.1 ([4]) *Let $\{w_k\}_{k=1}^n$ be real numbers with $w_1 \geq w_2 \geq \ldots \geq w_{n-1} \geq w_n$. Let $p(x) = \prod_{k=1}^n (x - w_k)$ and let $P(x)$ be any antiderivative of $p(x)$, then there exists $c \in \mathbb{R}$ such that $P(x) - c$ has all zeros real if and only if* $\max\{P(w_k) : k \text{ odd}\} \leq \min\{P(w_k) : k \text{ even}\}$ *in which case we can take any choice of c such that* $\max\{P(w_k) : k \text{ odd}\} \leq c \leq \min\{P(w_k) : k \text{ even}\}$.

We can restate this Lemma in a more convenient form.

Lemma 1.2 *Let $\{w_k\}_{k=1}^n$ be real numbers with $w_1 \geq w_2 \geq \ldots \geq w_{n-1} \geq w_n$. Let $p(x) = \prod_{k=1}^n (x - w_k)$ and let $P(x)$ be any antiderivative of $p(x)$, then there exists $c \in \mathbb{R}$ such that $P(x) - c$ has all zeros real if and only if $P(w_j) \geq P(w_k)$ whenever j is even and k is odd and $|j - k| \geq 3$.*

We note that if $j - k = 1$, then $p(x) > 0$ on the interval (w_k, w_j) and if $k - j = 1$, then $p(x) < 0$ on the interval (w_j, w_k); therefore in both cases, we automatically get $P(w_j) \geq P(w_k)$ which is why we can drop these as conditions in Lemma 1.2. (Interestingly, while this fact will not play a role in this paper, these inequalities are the only conditions on the ordered set $\{P(w_j)\}$ for arbitrary hyperbolic polynomials P. See [2] for the exact statement, proof and discussion of this fact.)

2 Quartic Polynomials

A simple induction shows that the number of inequalities in Lemma 1.2 is $\lfloor (\frac{n}{2} - 1)^2 \rfloor$ when $n \geq 2$. In particular, we see that for fourth degree polynomials the existence of a hyperbolic antiderivative essentially is equivalent to a single condition. We state this special case of Lemma 1.2.

Corollary 2.1 *Let* $\{w_k\}_{k=1}^{n}$ *be real numbers with* $w_1 \geq w_2 \geq w_3 \geq w_4$. *Let* $p(x) = \prod_{k=1}^{4}(x - w_k)$ *and let* $P(x)$ *be any antiderivative of* $p(x)$, *then there exists* $c \in \mathbb{R}$ *such that* $P(x) - c$ *has all zeros real if and only if* $P(w_4) \geq P(w_1)$.

We note that if a and b are real numbers with $a \neq 0$ then $p(ax+b)$ is a hyperbolic polynomial with a hyperbolic antiderivative if and only if $p(x)$ is. We may therefore apply the transformation $ax + b$ which maps w_1 to 1 and w_4 to -1 and consider quartic polynomials having $1, -1, s$ and t as zeros with $s, t \in [-1, 1]$. In this case, we get a very simple condition in terms of the zeros of p.

Theorem 2.2 *Let* $s, t \in [-1, 1]$ *and let* $p(x) = (x - 1)(x - s)(x - t)(x + 1)$. *Then* $p(x)$ *has a hyperbolic antiderivative if and only if* $st \geq -\frac{1}{5}$.

Proof Since $p(x) = x^4 - (s + t)x^3 + (st - 1)x^2 + (s + t)x - st$, we get $60P(x) = 12x^5 - 15(s + t)x^4 + 20(st - 1)x^3 + 30(s + t)x^2 - 60stx$ where $P(x)$ is an antiderivative of $p(x)$. Now $60(P(1) - P(-1)) = 2(12 + 20(st - 1) - 60st) = 8(3 + 5(st - 1) - 15st) = 8(-2 - 10st) = -16(1 + 5st)$, which means p has a hyperbolic antiderivative if and only if $st \geq -\frac{1}{5}$. □

We note that the mapping $\frac{w_1-w_4}{2}x - \frac{w_1+w_4}{2}$ maps the numbers $-1, s, t, 1$ (where $s = \frac{2w_2-w_1-w_4}{w_1-w_4}$ and $t = \frac{2w_3-w_1-w_4}{w_1-w_4}$) to w_1, w_2, w_3, w_4. The inequality $st \geq -\frac{1}{5}$ is equivalent to $5(2w_2 - w_1 - w_4)(2w_3 - w_1 - w_4) + (w_1 - w_4)^2 \geq 0$. After some algebra, we can restate this condition as follows:

Theorem 2.3 *Let* $\{w_i\}_{i=1}^{4}$ *be real numbers with* $w_1 \geq w_2 \geq w_3 \geq w_4$ *and let* $p(x) = (x - w_1)(x - w_2)(x - w_3)(x - w_4)$. *Then* $p(x)$ *has a hyperbolic antiderivative if and only if* $w^t Aw \geq 0$ *where* $w = (w_1, w_2, w_3, w_4)$ *and where*

$$A = \begin{bmatrix} 6 & -5 & -5 & 4 \\ -5 & 0 & 10 & -5 \\ -5 & 10 & 0 & -5 \\ 4 & -5 & -5 & 6 \end{bmatrix} .$$

We can also reformulate this result in terms of the gaps between the zeros. Let $g_j = w_j - w_{j+1}$ for $j = 1, 2, 3$. Then $5(2w_2 - w_1 - w_4)(2w_3 - w_1 - w_4) + (w_1 - w_4)^2 = 5(-g_1 + g_2 + g_3)(-g_1 - g_2 + g_3) + (g_1 + g_2 + g_3)^2 = 5(g_3 - g_1)^2 - 5g_2^2 + (g_1 + g_2 + g_3)^2$. This gives us the following result.

Theorem 2.4 *Let* $\{w_i\}_{i=1}^{4}$ *be real numbers with* $w_1 \geq w_2 \geq w_3 \geq w_4$ *and let* $p(x) = (x - w_1)(x - w_2)(x - w_3)(x - w_4)$. *Then* $p(x)$ *has a hyperbolic antiderivative if and only if* $v^t Bv \geq 0$ *where* $v = (w_1 - w_2, w_2 - w_3, w_3 - w_4)$ *is the vector of distances between adjacent zeros of* p *and where*

$$B = \begin{bmatrix} 6 & 1 & -4 \\ 1 & -4 & 1 \\ -4 & 1 & 6 \end{bmatrix} .$$

This suggests the problem of finding the characterization of the zero sets of higher degree polynomials which have hyperbolic antiderivatives. It is clear that the characterization will be a set of homogeneous polynomial inequalities of the form $S_n = \{w_1 \geq w_2 \geq \ldots \geq w_n : p_{n,k}(w_1, w_2, \ldots, w_n) \geq 0;\ 1 \leq k \leq \lfloor (\frac{n}{2} - 1)^2 \rfloor \}$. A characterization of the degrees of these polynomials in terms of n would be a good start.

References

1. G.G. Bilodeau, Generating polynomials all of whose roots are real. Math. Mag. **64**(4), 263–270 (1991)
2. C. Davis, Problem 4714. Am. Math. Mon. **63**(10), 729–729 (1956)
3. D. Khavinson, R. Pereira, M. Putinar, E.B. Saff, S. Shimorin, Borcea's variance conjectures on the critical points of polynomials, in *Notions of Positivity and the Geometry of Polynomials* (Springer, Berlin, 2011), pp. 283–309
4. D.M. Souroujon, T.S. Stoyanov, About the primitive polynomials of polynomials with real zeros. J. Anal. Appl. **14**(1), 21–31 (2016)

Positive Integral Kernels for Polar Derivatives

Mihai Putinar and Serguei Shimorin

Abstract The non-negativity on the unit disk of the real part of the polar derivative of a polynomial is proved via an integral representation with a positive kernel, or as a consequence of a weighted sum of hermitian squares decomposition.

Keywords Grace Theorem · Walsh Coincidence Theorem · Positive polynomial · Sums of squares · Polar derivative

MSC Codes: 30C15, 26C10, 42A16

1 Introduction

The polar derivative $dp(z) - (z - \alpha)p'(z)$ of a complex polynomial $p(z)$ of degree d is one of the wonder constructs in function theory; it has impacted deep results for a century and a half and continues to do so. Laguerre proved that, depending on the locus of the parameter α, the zeros of p and those of the polar derivative of p, or equivalently, the critical points of $p(z)/(z - \alpha)^d$, cannot be separated by a circular region. This is the key ingredient in Grace's apolarity theorem and the coincidence theorem of Walsh, both referring to the multi-affine symmetrizations of the polynomial p, see for ample comments [4, 6, 7].

When speaking about the geometry of zeros or critical points of complex analytic functions inequalities of various kinds naturally appear. In general their proofs rely on integral representations with a positive kernel, or purely algebraic completion of

M. Putinar (✉)
Department of Mathematics, University of California at Santa Barbara, Santa Barbara, CA, USA

Mathematics Department, Newcastle University, Newcastle upon Tyne, UK
e-mail: mputinar@math.ucsb.edu; mihai.putinar@ucl.ac.uk

S. Shimorin
Mathematics Department, The Royal Institute of Technology, Stockholm, Sweden

A. Aleman et al. (eds.), *Analysis of Operators on Function Spaces*,
Trends in Mathematics, https://doi.org/10.1007/978-3-030-14640-5_10

squares, see Chapter 4 of [6] for many revealing examples. Directly related to the polar derivative of a polynomial are inequalities discovered by van der Corput and Schaake [8] and later generalized by Hörmander [1].

Regarding the polar derivative operation from the point of view of harmonic analysis on the unit circle one unveils a representation of the polar derivative via an integral with positive kernel, and from there one derives an algebraic identity expressing the pull back of the polar derivative by a covering map of the circle in terms of the original polynomial. Both representations imply and explain from different perspectives van der Corput-Schaake and Hörmander inequalities. A new observation resulting from our computations states that, modulo a coordinate change, a symmetric multi-affine trigonometric polynomial which is non-negative on the torus is equal to a sum of hermitian squares. The general theory provides such a decomposition only for strictly positive polynomials [3].

This note is dedicated to the second author, a refined mathematician and talented photographer, who left us too early due to an accident in the Caucasus mountains.

2 Analytic Positivity

The complex variable $z \in \mathbb{C}$ defines the unit disk $\mathbb{D} = \{z; |z| < 1\}$ and the torus, or unit circle, $\mathbb{T} = \{z; |z| = 1\}$.

The *polar derivative* of a polynomial $p \in \mathbb{C}[z]$ is defined as follows:

$$D_{\alpha,n}[p](z) = (nI - (z-\alpha)\frac{d}{dz})[p](z) = np(z) - (z-\alpha)p'(z).$$

Note the dependence of the order n, usually equal to the degree of p, and of the parameter $\alpha \in \mathbb{C}$.

Poisson's kernel for the disk is

$$P_\beta(\zeta) = \frac{1-|\beta|^2}{|1-\overline{\beta}\zeta|^2}, \quad |\zeta| = 1, \quad |\beta| < 1.$$

Also we consider *Fejér's kernel* or order n:

$$\Phi_{n-1}(z) = \sum_{k=-n}^{n} (1 - \frac{|k|}{n}) z^n, \quad n \geq 1.$$

A direct computation proves that the latter is a hermitian square on the unit circle:

$$\Phi_{n-1}(\zeta) = \frac{1}{n}|1+\zeta+\zeta^2+\cdots+\zeta^{n-1}|^2, \quad |\zeta| = 1. \tag{2.1}$$

Consequently the product kernel

$$K_{\alpha,n}(\zeta,\eta) = \Phi_{n-1}(\overline{\eta}\zeta) P_{\overline{\alpha}\zeta}(\overline{\eta}^n \zeta^n)$$

has positive values for $|\zeta| = |\eta| = 1$ and $|\alpha| < 1$.

Proposition 2.1 *Let $p \in \mathbb{C}[z]$ be a polynomial of degree $n \geq 0$ and let $\alpha \in \mathbb{D}$. Then*

$$D_{\alpha,n}[p](\zeta) = n \int_{\mathbb{T}} K_{\alpha,n}(\zeta,\eta) p(\eta) \frac{|d\eta|}{2\pi}.$$

Proof Write

$$p(z) = \sum_{k=0}^{n} \hat{p}(k) z^k$$

and remark that

$$D_{\alpha,n}[p](\zeta) = n \sum_{k=0}^{n} (1 - \frac{k}{n}) \hat{p}(k) \zeta^k + \alpha\overline{\zeta} \sum_{k=0}^{n} k \hat{p}(k) \zeta^k. \tag{2.2}$$

On the other hand, the power series expansion of Poisson's kernel yields

$$K_{\alpha,n}(\zeta,\eta) = \sum_{\ell=0}^{\infty} \Phi_{n-1}(\overline{\eta}\zeta)(\alpha\overline{\zeta}\overline{\eta}^n\zeta^n)^\ell + \sum_{m=0}^{\infty} \Phi_{n-1}(\overline{\eta}\zeta)(\overline{\alpha}\zeta\overline{\eta}^n\zeta^n)^m.$$

When integrating $p(\eta)$ against $K_{\alpha,n}(\zeta,\eta)$ we notice that among all the terms above only those corresponding to $\ell = 0, 1$ are possibly non-zero. Hence

$$n \int_{\mathbb{T}} K_{\alpha,n}(\zeta,\eta) p(\eta) \frac{|d\eta|}{2\pi} =$$

$$n \int_{\mathbb{T}} \Phi_{n-1}(\overline{\eta}\zeta) p(\eta) \frac{|d\eta|}{2\pi} + n\alpha\overline{\zeta} \int_{T} \Phi_{n-1}(\overline{\eta}\zeta)\overline{\eta}^n \zeta^n p(\eta) \frac{|d\eta|}{2\pi}.$$

But this is exactly the decomposition (2.2). □

Note that the kernel $K_{\alpha,n}(\zeta,\eta)$ has the projection property on constants:

$$\int_{\mathbb{T}} K_{\alpha,n}(\zeta,\eta) \frac{|d\eta|}{2\pi} = 1.$$

Laguerre's Theorem asserts that, for all $\alpha \in \mathbb{D}$, $D_{\alpha,n}[p]$ does not vanish in $\mathbb{D}$ if the polynomial p has degree n and does not vanish on $\mathbb{D}$, see [4, 6] for ample comments on this important fact. The following is a variation on Laguerre's theme.

Corollary 2.2 *Let $\alpha \in \mathbb{D}$ and $n \geq 1$. For a polynomial $p \in \mathbb{C}[z]$ of degree n satisfying* Re $p \geq 0$ *on* $\mathbb{T}$, *one has* Re $D_{\alpha,n}[p] \geq 0$ *on* $\mathbb{T}$.

The *multi-affine symmetrization* of a polynomial $p \in \mathbb{C}[z]$ of degree n can be defined directly or by an iteration of the polar derivative operations:

$$P(z_1, \cdots, z_n) = \frac{1}{n!} D_{z_2,2} D_{z_3,3} \cdots D_{z_n,n}[p](z_1).$$

In any case the result being the substitution of the monomial z^k by the symmetric polynomial $\sigma_k(z_1, \cdots, z_n)$ of degree k. The relationship between the values of the affine symmetrization and those of the original polynomial is at the core of Grace-Walsh-Szegö Theorem, or rather phenomenon, in a variety of equivalent statements, see again [4, 6]. In this direction, algebraic inequalities responsible for geometric facts were recurrently unveiled.

The following classical result immediately follows from the preceding computations.

Theorem 2.3 *Let $f : \mathbb{C} \longrightarrow \mathbb{C}^N$ be a polynomial map of degree n and denote by $F : \mathbb{C}^n \longrightarrow \mathbb{C}^N$ its multi-affine symmetrization. If $M \subseteq \mathbb{C}^N$ is a convex set containing $f(\mathbb{D})$, then it also contains $F(\mathbb{D} \times \cdots \times \mathbb{D})$.*

Proof Apply Corollary 2.2 to an affine functional of the form $\mathrm{Re}\langle f(z), b\rangle + a$. □

For the original, quite different proofs, see [1, 8].

3 Algebraic Positivity

A foundational theorem of real algebra, going back to the early discoveries of Tarski, asserts that polynomial inequalities are consequences of completion of squares identities. For the relevance to real algebraic geometry and modern optimization theory of this is Ansatz, see, for instance, [5]. We mention from this context Riesz and Fejér Lemma which asserts that every non-negative trigonometric polynomial is equal to a hermitian square. More precisely, if

$$\sum_{k=-n}^{n} c_k e^{ik\theta} \geq 0, \quad \theta \in [-\pi, \pi],$$

then there exists a polynomial $p \in \mathbb{C}[z]$ with the property

$$\sum_{k=-n}^{n} c_k e^{ik\theta} = |p(e^{i\theta}|^2, \quad \theta \in [-\pi, \pi].$$

A similar phenomenon was unveiled by Quillen on the odd dimensional spheres: if a polynomial $P(z_1, \cdots, z_d, \overline{z}_1, \cdots, \overline{z}_d)$ is strictly positive on the sphere of equation $|z_1|^2 + \cdots + |z_d|^2 = 1$, then there exists a vector valued, complex polynomial map $\Pi : \mathbb{C}^d \longrightarrow \mathbb{C}^m$, such that

$$P(z_1, \cdots, z_d, \overline{z}_1, \cdots, \overline{z}_d) = \|\Pi(z_1, \cdots, z_d)\|^2, \quad |z_1|^2 + \cdots + |z_d|^2 = 1.$$

Then we say that P is equal to a *sum of hermitian squares* along the sphere. For a characterization of real ideals of the complex polynomial algebra on which similar decompositions hold, see [3].

We will show below that our harmonic analysis approach implies such a universal positivity certificate for the multi-affine symmetrization of a polynomial with non-negative real part on the unit circle. We emphasize that the sums of hermitian squares decomposition for strictly positive polynomials on the d-torus (the natural support of the symmetrization) follows from general theory, see [3]. The novelty in the statement below being the relaxed assumption allowing the non-negative polynomials to vanish on the respective tori.

Theorem 3.1 *Let $p \in \mathbb{C}[z]$ be a complex polynomial of degree $d \geq 1$ with* $\mathrm{Re}\, p(\zeta) \geq 0, \quad |\zeta| = 1$. *Denote by $P(z_1, \cdots, z_d)$ the multi-affine symmetrization of p. Then* $\mathrm{Re}\, P(\zeta_1^{d!}, \cdots, \zeta_d^{d!})$ *is equal to a sum of hermitian squares on the d-torus: $|\zeta_k| = 1, 1 \leq k \leq d$.*

Proof The multi-affine symmetrization P is obtained from p by an iteration of the polar derivative operation:

$$P(z_1, \cdots, z_d) = \frac{1}{d!} D_{z_2,2} D_{z_3,3} \cdots D_{z_d,d}[p](z_1).$$

We perform partial symmetrizations:

$$p_{z_d}(z) = \frac{1}{d} D_{z_d,d}[p](z), \;\; p_{z_{d-1},z_d}(z) = \frac{1}{d-1} D_{z_{d-1},d-1}[p_{z_d}](z), \ldots,$$

so that $P(z_1, \cdots, z_d) = p_{z_2,z_3,\cdots,z_d}(z_1)$.

Let $\ell_k = \frac{d!}{(k-1)!}, \;\; 1 \leq k \leq d$. We will prove by descending induction that

$$\mathrm{Re}\, p_{\zeta_k^{\ell_k},\cdots,\zeta_d^{\ell_d}}(\zeta^{\ell_k})$$

is a sum of hermitian squares on the torus $|\zeta| = |\zeta_d| = \cdots = |\zeta_k| = 1$.

Riesz-Fejér Lemma proves step zero, that is $\mathrm{Re}\, p(\zeta)$ is a hermitian square on the circle. The induction step is implied by the following statement.

Lemma 3.2 *Let $q \in \mathbb{C}[z, z_1, \cdots, z_k]$ with the property that there exists a positive integer m, so that* $\mathrm{Re}\, q(\zeta^m, \zeta_1, \cdots, \zeta_k)$ *is a sum of hermitian squares on the torus $|\zeta| = |\zeta_1| = \cdots = |\zeta_k| = 1$. Assume* $\deg_z(q) \leq d$. *Then the polynomial*

$Q(z, w, z_1, \cdots, z_k) = D(w, d)[p(\cdot, z_1, \cdots, z_k)](z)$ *has the property that*

$$\text{Re}\, Q(\zeta^{dm}, \sigma^{dm}, \zeta_1, \cdots, \zeta_k)$$

is a sum of hermitian squares on the torus $|\sigma| = |\zeta| = |\zeta_1| = \cdots = |\zeta_k| = 1$.

To prove Lemma we use Poisson's formula, written below in the sense of distributions:

$$\frac{1}{2\pi} \sum_{n \in \mathbb{Z}} e^{in(t-\alpha)} = \sum_{\ell \in \mathbb{Z}} \delta_{\alpha+2\pi\ell}(t),$$

or, for a fixed positive integer d:

$$\frac{1}{2\pi} \sum_{n \in \mathbb{Z}} e^{in(dt-\alpha)} = \frac{1}{d} \sum_{\ell \in \mathbb{Z}} \delta_{\frac{\alpha+2\pi\ell}{d}}(t) = \sum_{s \in \mathbb{Z}} \frac{1}{d} \sum_{k=0}^{d-1} \delta_{\frac{\alpha+2\pi k}{d}+2\pi s}(t).$$

Consider a single variable polynomial $f(z)$ of degree at most d, and with non-negative real part $h = \text{Re}\, f$ on the unit circle. According to Proposition 2.1 we have

$$\text{Re}\, D_{\sigma,d}[p](\zeta) = d \int_{\mathbb{T}} K_{\sigma,d}(\zeta, \eta) h(\eta) \frac{|d\eta|}{2\pi}$$

and we focus on the real integral in the second term. Write for $\zeta \in \mathbb{T}$

$$h(\zeta) = \sum_{k=-d}^{d} \hat{h}(k) \zeta^k,$$

and define

$$J_d[h](\sigma, \zeta) = \int_{\mathbb{T}} K_{\sigma,d}(\zeta, \eta) h(\eta) \frac{|d\eta|}{2\pi}.$$

Formula (2.2) yields, for $|\sigma| = |\zeta| = 1$:

$$J_d(h)(\sigma, \zeta) =$$

$$\sum_{k=-d}^{d} (1 - \frac{|k|}{d}) \hat{h}(k) \zeta^k + \sigma \sum_{k=0}^{d-1} \frac{k+1}{d} \hat{h}(k+1) \zeta^k + \overline{\sigma} \sum_{k=0}^{d-1} \frac{k+1}{d} \hat{h}(-k-1) \zeta^{-k}.$$

We prove, using Poisson's summation formula, that $J_d(h)(\sigma^d, \zeta^d)$ is a sum of hermitian squares on the bi-torus (whenever $h = \text{Re}\, f \geq 0$). To this aim, it will be convenient to switch to periodic real variables

$$\sigma = e^{ix},\ \zeta = e^{i\varphi}.$$

To simplify notation we set $\Phi_{d-1}(t) = \Phi_{d-1}(e^{it})$ and similarly $h(x) = h(e^{ix})$. By an identification of the Fourier coefficients appearing in the expression of $J_d(h)$, or a rewriting of the formula proved in Proposition 2.1 in real coordinates, we find:

$$J_d(e^{ix}, e^{i\varphi}) = \sum_{n\in\mathbb{Z}} \int_{-\pi}^{\pi} h(x-t)e^{in(\varphi-x)}e^{indt}\Phi_{d-1}(t)\frac{dt}{2\pi} =$$

$$\int_{-\pi}^{\pi} h(x-t)\Phi_{d-1}(t)\sum_{n\in\mathbb{Z}} e^{i(dt-x+\varphi)}\frac{dt}{2\pi} =$$

$$\int_{-\pi}^{\pi} h(x-t)\Phi_{d-1}(t)\sum_{j\in\mathbb{Z}}\frac{1}{d}\sum_{k=0}^{d-1}\delta_{\frac{x-\varphi+2\pi k}{d}+2\pi j}(t) =$$

$$\frac{1}{d}\sum_{k=0}^{d-1} h(\frac{d-1}{d}x + \frac{\varphi}{d} - \frac{2\pi k}{d})\Phi(\frac{x-\varphi}{d} + \frac{2\pi k}{d}).$$

In conclusion, returning to complex coordinates, we obtain:

$$J_d[h](\sigma^d, \zeta^d) = \frac{1}{d}\sum_{k=0}^{d-1} h(\sigma^{d-1}\zeta e^{-2\pi ki/d})\Phi_{d-1}(\overline{\zeta}\sigma e^{2\pi ki/d}). \tag{3.1}$$

Since the function h is non-negative and Φ is a sum of hermitian squares on the unit circle, we obtain the conclusion in the statement. □

As a matter of fact, the above theorem can be directly formulated in terms of a symmetric multi-affine function $F(z_1, \cdots, z_d)$ which is non-negative on the d-torus.

Returning to polar derivatives, the proof above implies the following algebraic identity.

Corollary 3.3 *Let $p \in \mathbb{C}[z]$ be a polynomial of degree less than or equal to $d \geq 1$. Then*

$$D_{\sigma^d,d}[p](\zeta^d) = \frac{1}{d}\sum_{k=0}^{d} p(\sigma^{d-1}\zeta\epsilon^{-k})|1 + \overline{\zeta}\sigma\epsilon^k + \cdots + (\overline{\zeta}\sigma\epsilon^k)^{d-1}|^2,$$

where $\sigma, \zeta \in \mathbb{T}$ and $\epsilon = e^{2\pi i/d}$.

Proof The two terms have equal real parts by (3.1). Then there difference is a purely imaginary constant. By taking $\sigma = \zeta$ we find that this constant is equal to zero. □

We can of course reverse the order and prove directly the above corollary by taking p to be a monomial. Denote as before by P the multi-affine symmetrization of the complex polynomial p. The iterative process

$$P(\zeta_1^{d!}, \zeta_2^{d!/1!}, \zeta_3^{d!/2!}, \cdots, \zeta_d^d) = \frac{1}{d!} D_{\zeta_2^{d!/1!},2} D_{\zeta_3^{d!/2!},3} \cdots D_{\zeta_d^d,d}[p](\zeta_1^{d!})$$

provides a weighted sum of hermitian squares decomposition for the left-hand side.

References

1. L. Hörmander, On a theorem of Grace. Math. Scand. **2**, 55–64 (1954)
2. D. Khavinson, R. Pereira, M. Putinar, E.B. Saff, S. Shimorin, Borcea's variance conjectures on the critical points of polynomials, in *Notions of Positivity and the Geometry of Polynomials*, ed. by P. Brändén et al. Trends in Mathematics (Birkhäuser, Basel, 2011), pp 283–309.
3. M. Putinar, C. Scheiderer, Quillen property of real algebraic varieties. Münster J. Math. **7**, 671–696 (2014)
4. Q.I. Rahman, G. Schmeisser, *Analytic Theory of Polynomials* (Oxford University Press, Oxford, 2002)
5. C. Scheiderer, Positivity and sums of squares: a guide to recent results, in *Emerging Applications of Algebraic Geometry*. The IMA Volumes in Mathematics and its Applications (Springer, New York, 2009), pp. 271–324
6. T. Sheil-Small, *Complex Polynomials* (Cambridge University Press, Cambridge, 2002)
7. G. Szegö, Bemerkungen zu einem Satz von J. H. Grace über die Wurzeln algebraischer Gleichungen. Math. Z. **13**, 28–55 (1922)
8. J.G. van der Corput, G. Schaake, Ungleichungen für Polynome und trigonometrische Polynome. Compos. Math. **2**, 321–361 (1935)
9. L. Walsh, On the location of the roots of certain types of polynomials. Trans. Am. Math. Soc. **24**, 163–180 (1922)

The Weak Type Estimates of Two Different Martingale Transforms Coincide

Alexander Reznikov and Alexander Volberg

Abstract We consider several weak type estimates for dyadic singular operators using the Bellman function approach. We write the precise formula for the unweighted weak type estimate Bellman function. We prove that the weak norms of two different martingale transforms coincide. The proof uses the precise form of the Bellman function of the weak type estimate of martingale transform.

Keywords Martingale transforms · Weak type · Bellman function · Monge–Ampère equation · A_1 weights

1991 Mathematics Subject Classification 30E20, 47B37, 47B40, 30D55

1 Introduction

Maria Reguera and Christoph Thiele disproved the Muckenhoupt–Wheeden conjecture [6, 7], which asked whether the Hilbert transform maps $L^1(Mw)$ into $L^{1,\infty}(w)$. It has been suggested in Pérez' paper [5] that there should exist such a counterexample, also [5] has several very interesting positive results, where Mw

AV was supported by the NSF grant DMS 1600065.

A. Reznikov
Center for Constructive Approximation, Department of Mathematics, Florida State University, Tallahassee, FL, USA

A. Volberg (✉)
Department of Mathematics, Michigan State University, East Lansing, MI, USA
e-mail: volberg@math.msu.edu

A. Aleman et al. (eds.), *Analysis of Operators on Function Spaces*,
Trends in Mathematics, https://doi.org/10.1007/978-3-030-14640-5_11

is replaced by a slightly bigger maximal function, in particular, by $M^2 w$ (which is equivalent to a certain Orlicz maximal function).

There exists a related A_1 conjecture of Muckenhoupt and Wheeden, sometimes called a weak Muckenhoupt–Wheeden conjecture. In this conjecture the weight w is assumed to be in A_1, thus, one has the pointwise inequality

$$Mw \le [w]_{A_1} w\,.$$

Under this assumption, Muckenhoupt and Wheeden asked whether the norm of the Hilbert transform (or martingale transform) from $L^1(w)$ into $L^{1,\infty}(w)$ is bounded by $C[w]_{A_1}$ (linear bound). The reader can get acquainted with the best so far positive result on A_1 conjecture in the paper [1].

In paper [4] we strengthened the Reguera and Reguera–Thiele results by disproving this A_1 conjecture.

We prove that the linear estimate in the weak Muckenhoupt–Wheeden conjecture is impossible, and, moreover, the growth of the weak norm of the Martingale transform and the weak norm of the Hilbert transform from $L^1(w)$ into $L^{1,\infty}(w)$ is at least $c\,[w]_{A_1} \log^{\frac{1}{3}}[w]_{A_1}$. Paper [1] gives an estimate from above for such a norm: it is $\le C\,[w]_{A_1} \log[w]_{A_1}$. Our method relies on Bellman function technique [2, 3].

2 Unweighted Weak Type of 0 Shift

The unweighted problem is much easier than the weighted problem. However, a glance at a simpler problem helps us to set up a more difficult one and to understand the difficulties. So we start with unweighted martingale transform, and briefly recall for the reader the setup and some of the results of preprint [8].

We are on $I_0 := [0, 1]$. As always D denotes the dyadic lattice. We consider the operator

$$\varphi \to \sum_{I \subseteq I_0, I \in \mathscr{D}} \epsilon_I (\varphi, h_I) h_I\,,$$

where $-1 \le \epsilon_I \le 1$. Notice that the sum does not contain the constant term.

Put

$$F := \langle |\varphi| \rangle_I\,,\ f := \langle \varphi \rangle_I\,,$$

and introduce the following function:

$$B(F, f, \lambda) := \sup \frac{1}{|I|} |\{x \in I : \sum_{J \subseteq I, J \in \mathscr{D}} \epsilon_J (\varphi, h_J) h_J(x) > \lambda\}|\,, \tag{2.1}$$

where the *sup* is taken over all $\epsilon_J = \pm 1$, $J \in D$, $J \subseteq I$, and over all $\varphi \in L^1(I)$ such that $F := \langle|\varphi|\rangle_I$, $f := \langle\varphi\rangle_I$, h_I are normalized in $L^2(\mathbb{R})$ Haar function of the cube (interval) I, and $|\cdot|$ denotes Lebesgue measure. Recall that

$$h_I(x) := \begin{cases} \frac{1}{\sqrt{|I|}}, & x \in I_+ \\ -\frac{1}{\sqrt{|I|}}, & x \in I_- \end{cases}$$

This function is defined in a convex domain $\Omega \subset \mathbb{R}^3$: $\Omega := \{(F, f, \lambda) \in \mathbb{R}^3 : |f| \leq F\}$.

Now we would like to introduce another Bellman function. It is assigned to a "slightly" different martingale transform.

Put

$$F := \langle|\varphi|\rangle_I\,,\ f := \langle\varphi\rangle_I\,,$$

and introduce the following function:

$$\widetilde{B}(F, f, \lambda) := \sup \frac{1}{|I|}|\{x \in I : \sum_{J\subseteq I, J\in\mathscr{D}} \epsilon_J(\varphi, h_J)h_J(x) > \lambda\}|\,, \tag{2.2}$$

where the *sup* is taken over all $-1 \leq \epsilon_J \leq 1$, $J \in D$, $J \subseteq I$, and over all $\varphi \in L^1(I)$ such that $F := \langle|\varphi|\rangle_I$, $f := \langle\varphi\rangle_I$.

Remark The functions B, $\widetilde{B}$ should not be indexed by I because they do not depend on I. We will use this soon.

2.1 The Main Inequality

Theorem 2.1 *Let $P, P_+, P_- \in \Omega$, $P = (F, f, \lambda)$, $P_+ = (F + \alpha, f + \beta, \lambda + \beta)$, $P_- = (F - \alpha, f - \beta, \lambda - \beta)$. Then*

$$B(P) - \frac{1}{2}(B(P_+) + B(P_-)) \geq 0\,. \tag{2.3}$$

At the same time, if $P, P_+, P_- \in \Omega$, $P = (F, f, \lambda)$, $P_+ = (F + \alpha, f + \beta, \lambda - \beta)$, $P_- = (F - \alpha, f - \beta, \lambda + \beta)$. Then

$$B(P) - \frac{1}{2}(B(P_+) + B(P_-)) \geq 0\,. \tag{2.4}$$

Proof Fix $P, P_+, P_- \in \Omega$, $P = (F, f, \lambda)$, $P_+ = (F+\alpha, f+\beta, \lambda+\beta)$, $P_- = (F - \alpha, f - \beta, \lambda - \beta)$. Let φ_+, φ_- be functions giving the supremum in $B(P_+)$, $B(P_-)$, respectively, up to a small number $\eta > 0$. Using the remark above we think that φ_+ is on I_+ and φ_- is on I_-. Consider

$$\varphi(x) := \begin{cases} \varphi_+(x)\,, \ x \in I_+ \\ \varphi_-(x)\,, \ x \in I_- \end{cases}$$

Notice that then

$$(\varphi, h_I) \cdot \frac{1}{\sqrt{|I|}} = \beta\,. \tag{2.5}$$

Then it is easy to see that

$$\langle|\varphi|\rangle_I = F = P_1, \ \ \langle\varphi\rangle_I = f = P_2\,. \tag{2.6}$$

Fix $P, P_+, P_- \in \Omega$, $P = (F, f, \lambda)$, $P_+ = (F + \alpha, f + \beta, \lambda + \beta)$, $P_- = (F - \alpha, f - \beta, \lambda - \beta)$. Let φ_+, φ_- be functions giving the supremum in $B(P_+)$, $B(P_-)$, respectively, up to a small number $\eta > 0$. Using the remark above we think that φ_+ is on I_+ and φ_- is on I_-. Consider

$$\varphi(x) := \begin{cases} \varphi_+(x)\,, \ x \in I_+ \\ \varphi_-(x)\,, \ x \in I_- \end{cases}$$

Notice that then

$$(\varphi, h_I) \cdot \frac{1}{\sqrt{|I|}} = \beta\,. \tag{2.7}$$

Then it is easy to see that

$$\langle|\varphi|\rangle_I = F = P_1, \ \ \langle\varphi\rangle_I = f = P_2\,. \tag{2.8}$$

If $\epsilon_I = -1$, then for $x \in I_+$, we get (we use (2.7) here)

$$\frac{1}{|I|}|\{x \in I_+ : \sum_{J\subseteq I, J\in\mathcal{D}} \epsilon_J(\varphi, h_J)h_J(x) > \lambda\}| = \frac{1}{|I|}|\{x \in I_+ : \sum_{J\subseteq I_+, J\in\mathcal{D}} \epsilon_J(\varphi, h_J)h_J(x) > \lambda + \beta\}|$$

$$= \frac{1}{2|I_+|}|\{x \in I_+ : \sum_{J\subseteq I_+, J\in\mathcal{D}} \epsilon_J(\varphi_+, h_J)h_J(x) > P_{+,3}\}| \ge \frac{1}{2}B(P_+) - \eta\,.$$

Similarly, for $x \in I_-$ using (2.7), we get if $\epsilon_I = -1$

$$\frac{1}{|I|}|\{x \in I_- : \sum_{J \subseteq I, J \in \mathcal{D}} \epsilon_J (\varphi, h_J) h_J(x) > \lambda\}| = \frac{1}{|I|}|\{x \in I_- : \sum_{J \subseteq I_+, J \in \mathcal{D}} \epsilon_J (\varphi, h_J) h_J(x) > \lambda - \beta\}|$$

$$= \frac{1}{2|I_-|}|\{x \in I_- : \sum_{J \subseteq I_-, J \in \mathcal{D}} \epsilon_J (\varphi_-, h_J) h_J(x) > P_{-,3}\}| \geq \frac{1}{2} B(P_-) - \eta\,.$$

Combining the two left-hand sides we obtain for $\epsilon_I = -1$

$$\frac{1}{|I|}|\{x \in I_+ : \sum_{J \subseteq I, J \in \mathcal{D}} \epsilon_J (\varphi, h_J) h_J(x) > \lambda\}| \geq \frac{1}{2}(B(P_+) + B(P_-)) - 2\eta\,.$$

Let us use now the simple information (2.8): if we take the supremum in the left-hand side over all functions φ, such that $\langle|\varphi|\rangle_I = F$, $\langle\varphi\rangle_I = f$, and supremum over all $\epsilon_J \in [-1, 1]$ (only $\epsilon_I = -1$ stays fixed), we get a quantity smaller or equal than the one, where we have the supremum over all functions φ, such that $\langle|\varphi|\rangle = F$, $\langle\varphi\rangle_I = f$, and an unrestricted supremum over all $\epsilon_J \in [-1, 1]$. The latter quantity is of course $B(F, f, \lambda)$. So we proved (2.3).

To prove (2.4) we repeat verbatim the same reasoning, only keeping now $\epsilon_I = 1$. We are done. □

Denote

$$T\varphi := \sum_{J \subseteq I, J \in \mathcal{D}} \epsilon_J (\varphi, h_J) h_J(x)\,.$$

It is a dyadic singular operator (actually, it is a family of operators enumerated by sequences of $\epsilon_I \in [-1, 1]$). To prove that it is of weak type is the same as to prove

$$B(F, f, \lambda) \leq \frac{C\,F}{\lambda},\ \lambda > 0\,. \tag{2.9}$$

Our B satisfies (2.3), (2.4). We consider these two conditions as special concavity conditions.

Let us make the change of variables, $(F, f, \lambda) \to (F, y_1, y_2)$:

$$y_1 := \frac{1}{2}(\lambda + f)\,,\ \ y_2 := \frac{1}{2}(\lambda - f)\,.$$

Denote

$$M(F, y_1, y_2) := B(F, y_1 - y_2, y_1 + y_2) = B(F, f, \lambda).$$

In terms of function M Theorem 2.1 reads as follows:

Theorem 2.2 *The function M is defined in the domain $G := \{(F, y_1, y_2) : |y_1 - y_2| \le F\}$, and for each fixed y_2, $M(F, y_1, \cdot)$ is concave and for each fixed y_1, $M(F, \cdot, y_2)$ is concave.*

Abusing the language we will call by the same letter B (correspondingly, M) *any* function satisfying (2.3), (2.4) (correspondingly satisfying Theorem 2.2).

It is not difficult to obtain one more condition, the so-called *obstacle condition*:

Lemma 2.3

$$\text{If } \lambda < f \text{ then } B(F, f, \lambda) = 1. \tag{2.10}$$

Proof

a) Let us first consider the case $f = F$, which can be viewed as the case of non-negative functions φ. Fix λ_0 and $\epsilon > 0$, let φ be a non-negative function on $I = [0, 1]$ such that it looks like $(\lambda_0 + \epsilon)\delta_0$, and $F = f := \int_0^1 \varphi\, dx = \lambda_0 + \epsilon > \lambda_0$. Namely, φ is zero on the set of measure $1 - \tau$, and an almost δ function times $\lambda_0 + \epsilon$ on a small interval of measure τ.

   Since it looks like a multiple of delta function, it can be written down as $(\lambda_0 + \epsilon)\mathbf{1_I} + H$, where H is a combination of Haar functions. Then consider a special martingale transform of φ, namely, $-H$. Then $-H = \lambda_0 + \epsilon > \lambda_0$ on a set of measure $1 - \tau$ with an arbitrary small τ (the smallness is independent of λ_0 and ϵ). Then the example of this φ shows that

$$B(\lambda_0 + \epsilon, \lambda_0 + \epsilon, \lambda_0) \ge 1 - \tau$$

   with an arbitrary $\tau > 0$.

b) We have to consider the case of $f < F$ as well. But if f $f > \lambda_0$, the construction is the same. Namely, consider $\Phi := \varphi + aS$, where S is a Haar function with very small support in a small dyadic interval ℓ (say, of measure smaller than τ) and normalized in L^1, let ℓ be contained in the set, where φ is small (φ is small essentially on almost the whole interval, because it looks like a positive multiple of the delta function), and ensure that $\int S\, dx = 0$, and $\int |S|\, dx = 1$. Then the example of φ shows that

$$B(\int_0^1 |\Phi|\, dx, \lambda_0 + \epsilon, \lambda_0) \ge 1 - 2\tau\,.$$

   By varying a from 0 to ∞ we can reach $\int |\Phi|\, dx = F$ for any $F \ge \lambda_0 + \epsilon$. Therefore, making first $\tau \to 0$ and then $\epsilon \to 0$, we prove (2.10).

□

Remark 2.4

$$\text{If } \lambda < f \text{ or } \lambda < 0, \text{ then } B(F, f, \lambda) = 1. \tag{2.11}$$

Of course if $\lambda < 0$ this is obvious because we can take a test function equal to constant. The rest is $0 \geq \lambda < f$, then parts a) and b) of the proof of Lemma 2.3 show that

Theorem 2.5 *Let us have $B \geq 0$ that satisfies (2.3), (2.4). (Equivalently, let the corresponding $M \geq 0$ be concave in (F, y_1) and in (F, y_2).) Let B satisfy (2.9), or, equivalently,*

$$M(F, y_1, y_2) \leq \frac{C\,F}{y_1 + y_2},\ y_1 + y_2 > 0\,. \tag{2.12}$$

Let $B(F, f, \lambda) = 1$ if $\lambda < 0$. Then we have the weak type estimate with constant at most C for all T uniformly in $\epsilon_I \in \{-1, 1\}$.

Proof Just by reversing the argument of Theorem 2.1. □

Remark Notice that the Bellman function B defined above satisfies by definition $B(F, f, \lambda) = B(F, -f, \lambda)$. Therefore, Lemma 2.3 claims in particular that $B(F, f, \lambda) = 1$ if $\lambda < 0$ (and we saw that it also satisfies (2.3), (2.4)).

3 The Main Result

We are interested in weak norm of two martingale transforms:

$$\phi \to \sum_{I \in \mathscr{D}(J)} \epsilon_I (\phi, h_I) h_I\,,$$

the first one T_1^ϵ is when ϵ_I are allowed to be only ± 1, and the second one T_2^ϵ is when ϵ_I runs freely in $[-1, 1]$.

Of course for every number $t_I \in [-1, 1]$ we can write it as a convex combination

$$t_I = \sum_{k=1}^{\infty} 2^{-k} t_{k,I}\,,$$

where $t_{k,I} \in \{-1, 1\}$. Thus, for any sequence $\epsilon = \{\epsilon_I\}_{I \in \mathscr{D}}$, $\epsilon_I \in [-1, 1]$, there will be sequences $\epsilon_k = \{\epsilon_{k,I}\}_{I \in \mathscr{D}}$, $\epsilon_{k,I} \in \{-1, 1\}$ such that

$$T_2^\epsilon = \sum_{k=1}^{\infty} 2^{-k} T_1^{\epsilon_k}\,.$$

If we were interested in the estimate of T_2^ϵ in a Banach space X (say, $X = L^p$, $p > 1$, then this convex combination representation would show that

$$\sup_{\epsilon:\, \epsilon_I \in [-1,1]} \|T_2^\epsilon\|_X = \sup_{\epsilon:\, \epsilon_I \in \{-1,1\}} \|T_1^\epsilon\|_X\,. \tag{3.1}$$

However, we are interested in the case $X = L^{1,\infty}$. Here one can use the Lemma of Stein–Weiss:

Lemma 3.1 *Let $\{g_j\}$ be a sequence of non-negative measurable functions, such that $\|g_j\|^{L^{1,\infty}} \le 1$ for all j. Let $\{c_j\}$ be a sequence of non-negative scalars such that $\sum c_j = 1$ and $\sum c_j \log \frac{1}{c_j} = K < \infty$. Then*

$$\|\sum_j c_j g_j\|_{L^{1,\infty}} \le 2(K+2)\,.$$

See [9] for the proof. From this lemma, we would conclude that

$$\sup_{\epsilon:\, \epsilon_I \in [-1,1]} \|T_2^\epsilon\|_{L^{1,\infty}} \le 2(2 + \sum_{k=1}^{\infty} k2^{-k}) \sup_{\epsilon:\, \epsilon_I \in \{-1,1\}} \|T_1^\epsilon\|_{L^{1,\infty}}\,.$$

However, Theorem 4.1 gives immediately a better result. This is the main result of this note.

Theorem 3.2

$$\sup_{\epsilon:\, \epsilon_I \in [-1,1]} \|T_2^\epsilon\|_{L^{1,\infty}} = \sup_{\epsilon:\, \epsilon_I \in \{-1,1\}} \|T_1^\epsilon\|_{L^{1,\infty}}\,. \tag{3.2}$$

4 The Second Type of Martingale Transform: Why It has the Same Weak Norm?

In this section we are concerned with the second type of the martingale transform, the one where the numbers ϵ_J, $J \in \mathcal{D}$ are allowed to run over the interval $[-1, 1]$. In other words, we are interested to find the function $\widetilde{B}$ from (2.2) rather than function B from (2.1). The latter is found in Theorem 5.1.

It turns out that these two functions are equal, but this requires a proof.

Theorem 4.1 $\widetilde{B}(F, f, \lambda) = B(F, f, \lambda)$.

4.1 The Proof of Theorem 4.1

It is immediate by definition that $\widetilde{B} \geq B$. Indeed, the supremum in the definition of $\widetilde{B}$ is taken over a larger set. It is easy to see that to prove the equality $\widetilde{B} = B$, one needs to prove the following concavity of the function B found in Theorem 5.1:

$$d^2 B(F, f, \lambda) \geq 0, \text{ if } |d\lambda| \leq |df| \text{ and } \lambda > F \geq |f|\,.$$

Recall that in coordinates

$$y_1 := \frac{1}{2}(\lambda + f), \;\; y_2 := \frac{1}{2}(\lambda - f)$$

$\lambda > |f|$ can be written down as $y_1 y_2 > 0$, and we write the function B as $M(y_1, y_2, F)$. The above mentioned property becomes

$$-d^2 M(y_1, y_2, F) \geq 0, \text{ if } dy_1, dy_2 \text{ are of the opposite sign and } y_1 + y_2 > F, y_1 y_2 > 0\,. \tag{4.1}$$

The formula for B is found in Theorem 5.1, thus, we know the formula for M:

$$4(1 - M(y_1, y_2, F)) = \frac{(y_1 + y_2 - F)^2}{y_1 y_2} =: \Phi(y_1, y_2, F)\,.$$

So we need to prove the appropriate version of "convexity" of Φ:

$$d^2 \Phi(y_1, y_2, F) \geq 0, \text{ if } dy_1, dy_2 \text{ of opposite sign and } y_1 + y_2 > F, y_1 y_2 > 0\,. \tag{4.2}$$

As

$$\Phi(y_1, y_2, F) = \frac{y_1}{y_2} + \frac{y_2}{y_1} - \frac{2F}{y_1} - \frac{2F}{y_2} + \frac{F^2}{y_1 y_2},$$

we can calculate its Hessian:

$$A \cdot d^2\Phi \cdot A = \frac{1}{y_1 y_2} \begin{bmatrix} 2(y_2 - F)^2, & -(y_1^2 + y_2^2 - F^2), & 2(y_2 - F) \\ -(y_1^2 + y_2^2 - F^2), & 2(y_1 - F)^2, & 2(y_1 - F) \\ 2(y_2 - F), & 2(y_1 - F), & 2 \end{bmatrix},$$

where $A = \text{diag}(y_1, y_2, 1)$ is a diagonal non-singular matrix.

As we know that $y_1 y_2 > 0$ we need to prove that the quadratic form of the matrix in the right-hand side is positive as soon as it is computed on vectors (dy_1, dy_2, dF), where $dy_1 dy_2 < 0$. Consider vectors $(-\eta, \xi, dF)$, where ξ, η are of the same sign, and let us write down the quadratic form on such a vector.

This is (up to constant 2)

$$\Big[(y_2-F)^2\eta^2+(y_1-F)^2\xi^2+(y_1^2+y_2^2-F^2)\xi\eta\Big]+\Big[(dF)^2+2\,(-(y_2-F)\eta+2(y_1-F)\xi)\,dF\Big]\,.$$

What we have is at least

$$(-(y_2-F)\eta+2(y_1-F)\xi)^2+\Big[2(y_2-F)(y_1-F)\xi\eta+(y_1^2+y_2^2-F^2)\xi\eta\Big]+$$
$$(dF)^2+2\,(-(y_2-F)\eta+2(y_1-F)\xi)\,dF\,.$$

Expanding [...] and using the fact that $\xi\eta\geq 0$:

$$[\dots]=(y_1^2+y_2^2+2y_1y_2-F^2+2F^2-2(y_1+y_2)F)\xi\eta=(y_1+y_2-F)^2\xi\eta\geq 0\,.$$

Hence, our quadratic form is at least

$$(-(y_2-F)\eta+2(y_1-F)\xi)^2+(dF)^2+2\,(-(y_2-F)\eta+2(y_1-F)\xi)\,dF\,.$$

But this is also a full square. We are done, Theorem 4.1 is proved.

5 Unweighted Case: The Exact Bellman Function

Consider first finding the unweighted Bellman function on the boundary $F=f$. This means that we are working with positive functions φ. If we denote as before

$$y_1=\frac{1}{2}(\lambda+f)\,,\;\; y_2=\frac{1}{2}(\lambda-f)$$

then we need a 0-homogeneous function in y_1, y_2, that is a function $b(\frac{y_1}{y_2})$, such that $b(t)$ is defined for $t\in(1,\infty]$ (which corresponds to $\lambda\in[f,\infty)$) such that

$$\left(b\left(\frac{y_1}{y_2}\right)\right)''_{y_1y_1}\leq 0\,,\;\;\left(b\left(\frac{y_1}{y_2}\right)\right)''_{y_2y_2}\leq 0 \tag{5.1}$$

and knowing that when $\lambda\to\infty$ we have $\frac{y_1}{y_2}\to 1$, we obtain that

$$b(1)=0\,. \tag{5.2}$$

On the other hand, we will use soon that

$$b(t)\to 1,\; t\to\infty\,. \tag{5.3}$$

But (5.3) is just Lemma 2.3 in the case $f = F$ and the continuity of the function $B(F, f, \lambda)$ on the border $\lambda = f$ of the domain $\lambda < f$ considered in Lemma 2.3 and the domain $\lambda > f$ considered here.

From (5.1) we derive

$$b'' \le 0,\ t^2 b''(t) + 2tb'(t) \le 0,\ \ t \in [1, \infty)\,.$$

Make the second inequality an equation. Then we have its concave solution

$$b(t) = c_1 + \frac{c_2}{t}, c_2 < 0\,,$$

the last assumption ($c_2 < 0$) follows from the concavity requirement on b. Now (5.2) and (5.3) give us our desired biconcave function:

$$b(t) = 1 - \frac{1}{t} \Rightarrow M(y_1 - y_2, y_1, y_2) = 1 - \frac{y_2}{y_1} = 1 - \frac{\lambda - f}{\lambda + f} = \frac{2f}{\lambda + f}\,.$$

Here is the Bellman function for unweighted weak type inequality for martingale transform, see [8].

Theorem 5.1

$$B(F, f, \lambda) = \begin{cases} 1, & \text{if } \lambda \le F\,, \\ 1 - \frac{(\lambda - F)^2}{\lambda^2 - f^2} & \text{if } \lambda > F\,. \end{cases} \tag{5.4}$$

As above we always denote

$$y_1 = \frac{1}{2}(\lambda + f),\ \ y_2 = \frac{1}{2}(\lambda - f)\,. \tag{5.5}$$

We are looking for a function M, which is defined in $\Omega := \{(F, y_1, y_2) : F \ge |y_1 - y_2|\}$, bi-concave in variables (F, y_1) and (F, y_2), satisfies the first boundary condition:

$$\text{If } F = f := y_1 - y_2 \ge 0\,,\ y_2 > 0 \Rightarrow B(F, f, \lambda) := \frac{2f}{\lambda + f} \Rightarrow M(y_1 - y_2, y_1, y_2) = 1 - \frac{y_2}{y_1}\,. \tag{5.6}$$

It also satisfies, see (2.10), the second boundary condition:

$$y_2 + y_1 \le F \Rightarrow M = 1\,. \tag{5.7}$$

In particular, if we recall that M is defined only in the domain $\Omega := \{(F, y_1, y_2) \in \mathbb{R}^3 : |y_1 - y_2| \le F\}$ we conclude that (see (5.8))

$$M(F, y_1, 0) = 1,\ F \ge |y_1|\,. \tag{5.8}$$

Now let us consider the section Ω_{y_1} of the domain of definition $\Omega = \{(F, y_1, y_2) : F \geq |y_1 - y_2|\}$ given by the hyperplane $y_1 =$ fixed. We want to find an M satisfying concavity in this hyperplane. We are going to look for M (and we will check later that it is concave) that solves Monge–Ampère equation in Ω_{y_1} with boundary conditions (5.6) and (5.7). That is

$$\det D^2_{F,y_2} M = 0,\ (F, y_1, y_2) \in \Omega_{y_1}\,. \tag{5.9}$$

In Ω_{y_1}, there is a point $P := (y_1, y_1, 0)$. Let us make a guess that the characteristics (which we know by Pogorelov's theorem form the foliation of Ω_{y_1} by straight lines) of our Monge–Ampère equation in Ω_{y_1} form the fan of lines with common point $P = (y_1, y_1, 0)$.

We denote the collection of straight line segments emanating from $P = (y_1, y_1, 0)$ and foliating the domain Ω_{y_1} by $\mathcal{L} = \{L_t\}_{t\geq 0}$. The parameter $t = 0$ corresponds to the line segment $F = -y_2 + y_1$, $0 \leq y_2 \leq y_1$, which is a boundary segment of Ω_{y_1}.

By Pogorelov's theorem we also know that there exist functions t_1, t_2, t constant on characteristics such that

$$M = t_1 F + t_2 y_2 + t\,, \tag{5.10}$$

such that $t_1 = t_1(t; y_1), t_2 = t_2(t; y_1)$ (we think that y_1 is a parameter) that

$$0 = (t_1)'_t F + (t_2)'_t y_2 + 1\,, \tag{5.11}$$

and, moreover, that one can choose

$$t_1 = \frac{\partial M(F, \cdot, y_2)}{\partial F}\,,\ t_2 = \frac{\partial M(F, \cdot, y_2)}{\partial y_2}\,. \tag{5.12}$$

Remark 5.2 Incidentally, the reader can easily see that (5.11) and (5.10) imply (5.12). In fact, differentiate (5.10) in F. Then

$$\frac{\partial M}{\partial F} = t_1 + ((t_1)'_t F + (t_2)'_t y_2 + 1)\frac{\partial t}{\partial F}\,,$$

and now we use (5.11). Relationship (5.11) gives the straight lines $\ell_t := (F, y_2) : (t_1)'_t F + (t_2)'_t y_2 + 1 = 0$ in domain Ω_{y_1}. These are exactly the lines L_t (up to reparametrization of t).

Remark 5.3 Actually (5.10), (5.12) follow from Monge–Ampère equation on M. In fact, the fact that M satisfies the Monge–Ampère equation in variables F, y_2 obviously means that the level curves of $\frac{\partial M(F,\cdot,y_2)}{\partial F}$ and $\frac{\partial M(F,\cdot,y_2)}{\partial y_2}$ are the same. The reader can see this by checking that the normals to the level curves of these two functions are just rows of Hessian matrix $D^2_{F,y_2}M$, and these rows are proportional by the Monge–Ampère equation (5.9).

If two functions share the family of level sets, then any of them is a function of the other one. So the functions t_1, t_2 (partial derivatives) are functions of each other in (5.12). It is very easy to check (chain rule) that if, for a smooth function $\Phi(x, y)$, its partial derivatives Φ_x, Φ_y share all the level curves, then $t(x, y) := \Phi(x, y) - \Phi_x(x, y)x - \Phi_y(x, y)y$ has also the same set of level curves, that is, Φ_x, Φ_y are some functions of t. This is exactly relationship (5.10).

Extend a segment L_t from P till $y_2 = y_1$. The latter is the vertical line in Ω_{y_1}, which is the intersection of Ω_{y_1} with $f = 0$. Our function B is even with respect to f, being (as we will see shortly) smooth it has one more boundary condition: $\frac{\partial B}{\partial f}(F, 0, \lambda) = 0$, that is

$$y_2 = y_1 \Rightarrow \frac{\partial M}{\partial y_2} = \frac{\partial M}{\partial y_1}\,. \tag{5.13}$$

Or if we denote the intersection of L_t with the hyperplane $y_2 = y_1$ by

$$(F(t), y_1, y_1) \tag{5.14}$$

we get

$$\frac{\partial M}{\partial y_1}(F(t), y_1, y_1) = t_2(t; y_1)\,. \tag{5.15}$$

We want to prove now that

$$F(t)t_1(t) + 2y_1t_2(t) = 0\,. \tag{5.16}$$

In fact, our M is 0 homogeneous. So everywhere $FM'_F + y_1M'_{y_1} + y_2M'_{y_2} = 0$. Apply this to point $(F(t), y_1, y_1)$, where we can use (5.15) to get $F(t)t_1 + y_1t_2 + y_1t_2 = 0$, which is (5.16).

So far we did not use our guess that the L_t fan from $P = (y_1, y_1, 0)$. Let us use this now. Plug that coordinates into $0 = (t_1)'_t F + (t_2)'_t y_2 + 1$, which is (5.11). Then we get the crucial (and trivial) ODE

$$t_1'(t) = -\frac{1}{y_1} \Rightarrow t_1(t) = -\frac{1}{y_1}t + C_1(y_1)\,. \tag{5.17}$$

By our convention the boundary line

$$\begin{cases} F = y_1 - u \\ y_2 = u \end{cases}$$

corresponds to $t = t_0 = 0$.

It is time to use the **boundary condition** (5.6). We use (5.10) and (5.6):

$$\Big(-\frac{1}{y_1}t_0 + C_1(y_1)\Big)(y_1 - u) + t_2 u + t_0 = 1 - \frac{u}{y_1}\,.$$

Using (5.16) we can plug t_2 expressed via $F(t)$. But by definition (5.14) $F(t_0) = 0$, and hence by (5.16) $t_2(t_0) = 0$. So we get (plug $t_0 = 0$)

$$C_1(y_1)(y_1 - u) = 1 - \frac{u}{y_1}\,.$$

So we get $C_1(y_1) = \frac{1}{y_1}$. Now from (5.17) we get

$$t_1(t) = \frac{1}{y_1}(1 - t)\,. \tag{5.18}$$

After that, (5.11) at the point $(F(t), y_1, y_1)$ and (5.16) become the system of two linear algebro-differential equations in $F(t)$ and $t_2(t)$:

$$\begin{cases} -\frac{1}{y_1}F(t) + y_1 t_2'(t) + 1 = 0 \\ 2y_1 t_2(t) + F(t)\frac{1}{y_1}(1 - t) = 0\,. \end{cases} \tag{5.19}$$

Hence we found

$$(1 - t)y_1 t_2'(t) + 2y_1 t_2(t) + (1 - t) = 0\,.$$

So $y_1 t_2 = C(1 - t)^2 - (1 - t)$. Plug this t_2 into the second equation of (5.19) and then plug in $t = 0$. Taking into account, that L_0 is the boundary segment $F = -y_2 + y_1,\ 0 \le y_2 \le y_1$, we conclude that

$$F(0) = 0\,.$$

This allows us to find $C = 1$, hence

$$t_2 = -\frac{1}{y_1}(1 - t)t\,. \tag{5.20}$$

Plugging in (5.18), (5.20) into (5.11) yields the following (this is the form of the straight line L_t):

$$-\frac{1}{y_1}F + \frac{1}{y_1}(2t - 1)y_2 + 1 = 0\,. \tag{5.21}$$

From this we find at last the function $t(F, y_2; y_1)$ from (5.10):

$$t = \frac{1}{2}\frac{F - (y_1 - y_2)}{y_2}\,. \tag{5.22}$$

Remark Notice that in (F, f, λ) coordinates, we will get $t = \frac{F-f}{\lambda-f}$.

Plug this into (5.10), in which we know already $t_1(t)$ and $t_2(t)$ (see (5.18), (5.20)). In other words, in the formula

$$M(F, y_1, y_2) = t_1(t)F + t_2(t)y_2 + t \tag{5.23}$$

we know now all three functions t, t_1, t_2 as functions of F and y_2 (and of parameter y_1).

We get

$$M = \begin{cases} 1\,, \; y_2 + y_1 \le F\,, \\ 1 - \frac{(y_2+y_1-F)^2}{4y_2y_1}\,, \; y_2 + y_1 > F\,. \end{cases} \tag{5.24}$$

This is the same as

$$B(F, f, \lambda) = 1 - \frac{(\lambda - F)^2}{\lambda^2 - f^2}\,. \tag{5.25}$$

6 Verification

It is not difficult to check now that $M(F, y_1, y_2) = B(F, y_1 - y_2, y_1 + y_2)$ is bi-concave and satisfies all boundary conditions. To verify that actually this is the exact Bellman function of the unweighted weak type estimate we need to prove that for every $(F, y_1, y_2) \in \Omega_{y_1}$ we can find the sequence of functions giving the right estimate of the measure of the level sets of their martingale transform. The reader can find these functions in [8].

References

1. A. Lerner, S. Ombrosi, C. Pérez, A1 bounds for Calderón-Zygmund operators related to a problem of Muckenhoupt and Wheeden. Math. Res. Lett. **16**(1), 149–156 (2009)
2. F. Nazarov, A. Volberg, The Bellman function, the two-weight Hilbert transform, and embedding of the model spaces K_θ, volume in honor of Thomas Wolff. J. Anal. Math. **87**, 385–414 (2002)
3. F. Nazarov, S. Treil, A. Volberg, Bellman function in stochastic control and harmonic analysis, in *Systems, Approximation, Singular Integral Operators, and Related Topics. Operator Theory: Advances and Applications* vol 129 (Birkhauser, Basel, 2001) pp. 393–423

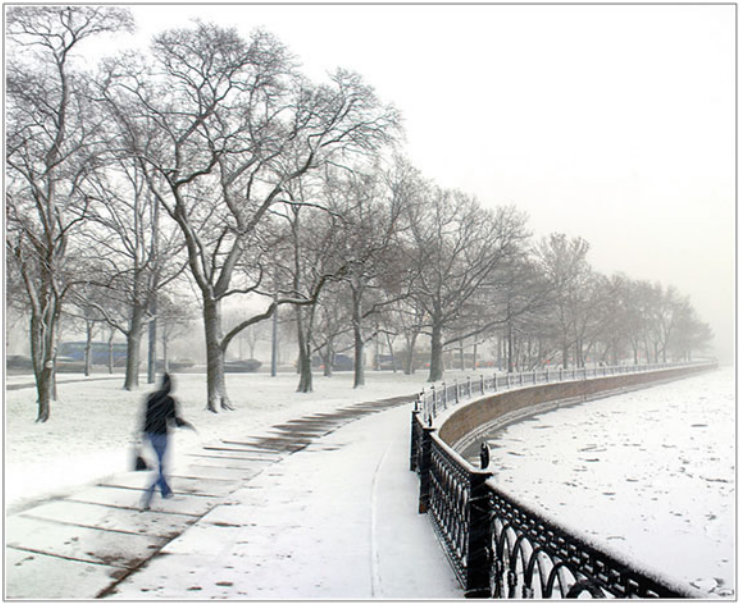
4. F. Nazarov, A. Reznikov, V. Vasyunin, A. Volberg, On weak weighted estimates of the martingale transform and a dyadic shift. Anal. PDE **11**(8), 2089–2109 (2018)
5. C. Pérez, Weighted norm inequalities for singular integral operators. J. Lond. Math. Soc. (2) **49**(2), 296–308 (1994)
6. M.C. Reguera, *On Muckenhoupt–Wheeden Conjecture*, pp. 1–14 (2010). Preprint
7. M.C. Reguera, C. Thiele, The Hilbert transform does not map $L^1(Mw)$ to $L^{1,\infty}(w)$ (2010). arXiv:1011.1767
8. A. Reznikov, V. Vasyunin, A. Volberg, Extremizers and the Bellman function for the weak type estimate of the martingale transform (2013). arxiv:1311.2133
9. E.M. Stein, N.J. Weiss, On the convergence of Poisson integrals. Trans. AMS **140**, 35–54 (1969)

Photographs by Serguei Shimorin

Serguei Shimorin—The Photographer

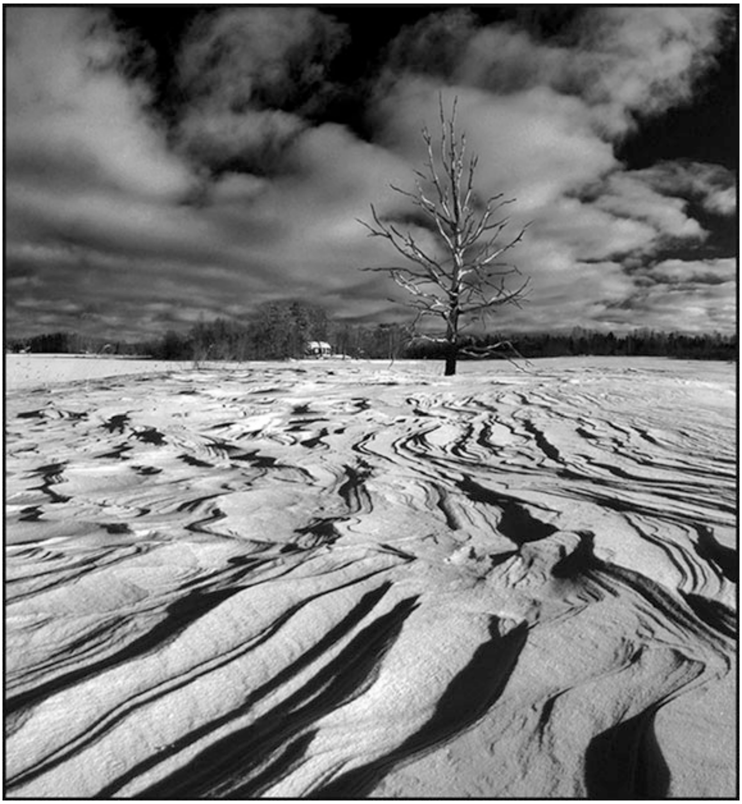
A. Aleman et al. (eds.), *Analysis of Operators on Function Spaces*,
Trends in Mathematics, https://doi.org/10.1007/978-3-030-14640-5

Winter evening Stockholm



Snowstorm

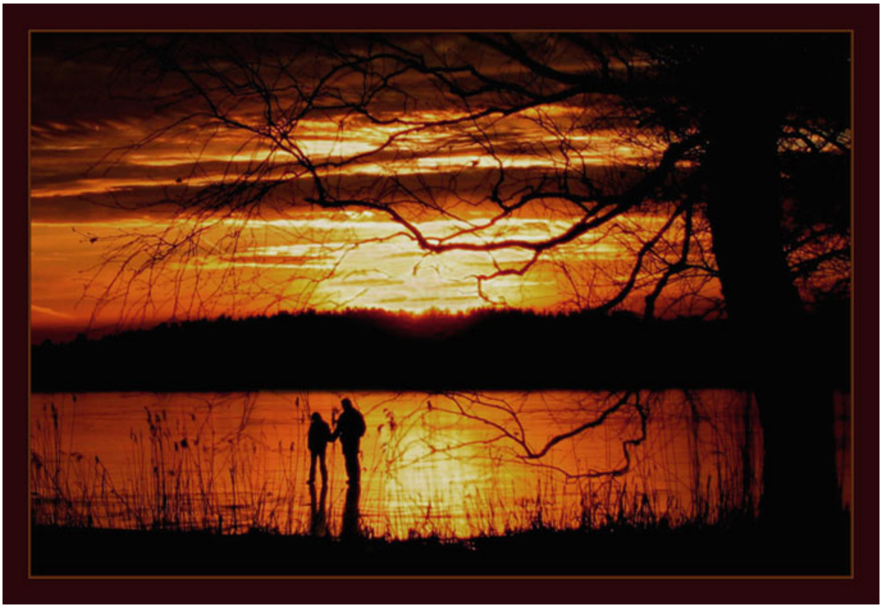
Valley of ten peaks



February loneliness

Golden autumn

Swing



Valentine's day