Operator Theory
Advances and Applications
282

M. Amélia Bastos
Luís Castro
Alexei Yu. Karlovich
Editors

Operator Theory, Functional Analysis and Applications

Birkhäuser

Operator Theory: Advances and Applications

Volume 282

Founded in 1979 by Israel Gohberg

More information about this series at http://www.springer.com/series/4850

M. Amélia Bastos • Luís Castro •
Alexei Yu. Karlovich
Editors

Operator Theory, Functional Analysis and Applications

Birkhäuser

Editors
M. Amélia Bastos
Instituto Superior Técnico
Universidade de Lisboa
Lisboa, Portugal

Luís Castro
Departamento de Matemática
Universidade de Aveiro
Aveiro, Portugal

Alexei Yu. Karlovich
Faculdade de Ciências e Tecnologia
Universidade Nova de Lisboa
Lisboa, Portugal

ISSN 0255-0156 ISSN 2296-4878 (electronic)
Operator Theory: Advances and Applications
ISBN 978-3-030-51944-5 ISBN 978-3-030-51945-2 (eBook)
https://doi.org/10.1007/978-3-030-51945-2

Mathematics Subject Classification: 47-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 volume is dedicated to the 30th International Workshop on Operator Theory and its Applications, IWOTA 2019, where a wide range of topics on the recent developments in Operator Theory and Functional Analysis was presented and discussed.

The book is composed of 30 articles covering the different scientific areas of IWOTA 2019. Namely:

- Group representations and determinantal hypersurfaces. Dilation theory of completely positive maps and semigroups as well as the operator algebraic approach to dilation theory. Tight and cover-to-join representations of semilattices and inverse semigroups. Compact sequences in quasifractal algebras. Representable and continuous functionals of Banach quasi *-algebras. Langlands reciprocity for C^*-algebras.
- Tau functions and linear systems. Bishop operators, spectral properties, and spectral invariant subspaces. The operator Jensen–Mercer inequality. Birkhoff–James orthogonality.
- Extended Heinz and Jensen type inequalities and rearrangements. d-Modified Riesz potentials on central Campanato spaces. K-inner functions and K-contractions for unitarily invariant reproducing kernel functions. Products of unbounded block functions. Riesz–Fischer maps and frames in rigged Hilbert spaces.
- Minimality properties of Sturm–Liouville problems. Solutions of inhomogeneous ill-posed problems in Banach space. Quantum graph with Rashba Hamiltonian. Unbounded operators on the Segal–Bargmann space. Solvability of the Caffarelli–Silvestre extension problem.
- Semiclassical elliptic pseudodifferential operators and periodic coherent states decomposition. Time-dependent approach to the Sommerfeld solution of a diffraction problem. A numerical approach for approximating variable-order fractional integral operator. Inner–outer factorization of rational matrix valued functions.

- Toeplitz plus Hankel operators and their close relatives. Fourier convolution operators with slowly oscillating symbols. Singular integral operators with Cauchy and Mellin kernels. Noncommutative C^*-algebras generated by Toeplitz operators on the unit sphere.
- Groups of orthogonal matrices, all orbits of which generate lattices. The inverse characteristic polynomial problem for trees.

IWOTA 2019 was held from 22nd to 26th of July 2019 at Instituto Superior Técnico, University of Lisbon, Portugal. It was focused on the latest developments in Functional Analysis, Operator Theory, and related fields and was organized by M. Amélia Bastos (IST, UL), António Bravo (IST, UL), Catarina Carvalho (IST, UL), Luís Castro (UA), Alexei Karlovich (FCT, UNL), and Helena Mascarenhas (IST, UL).

IWOTA 2019 had 471 participants from all over the world. The program consisted of 11 plenary lectures:

- J. Ball, Input/State/Output Linear Systems and Their Transfer Functions: From Single-Variable to Multivariable to Free Noncommutative Function Theory;
- A. Böttcher, Lattices from Equiangular Tight Frames;
- K. Davidson, Noncommutative Choquet Theory;
- R. Exel, Statistical Mechanics on Markov Spaces with Infinitely Many States;
- H. Feichtinger, Classical Fourier Analysis and the Banach Gelfand Triple;
- R. Kaashoek, Inverting Structured Operators and Solving Related Inverse Problems;
- I. Klep, Bianalytic Maps between Matrix Convex Sets;
- L.-E. Persson, My Life with Hardy and His Inequalities;
- P. Semrl, Automorphisms of Effect Algebras;
- B. Silbermann, Invertibility Issues for Toeplitz Plus Hankel Operators;
- C. Tretter, Spectra and Essential Spectra of Non-Self-Adjoint Operators;

16 semi-plenary lectures:

- P. Ara, Separated Graphs and Dynamics;
- S. Belinschi, Analytic Transforms of Noncommutative Distributions;
- G. Blower, Linear Systems in Random Matrix Theory;
- A. Caetano, Function Spaces Techniques in Problems of Scattering by Fractal Screens;
- A. B. Cruzeiro, On Some Stochastic Partial Differential Equations Obtained by a Variational Approach;
- R. Duduchava, Boundary Value Problems on Hypersurfaces and Γ-Convergence;
- P. Freitas, Spectral Determinants of Elliptic Operators: Dependence on Spatial Dimension and Order of the Operator;
- E. Gallardo, Invariant Subspaces for Bishop Operators and Beyond;
- Yu. Karlovich, Algebras of Singular Integral Operators with Piecewise Quasi-continuous Coefficients and Non-Smooth Shifts;
- S. Petermichl, Change of Measure;
- S. Roch, On Quasifractal Algebras;

- O. M. Shalit, Dilation Theory: Fresh Directions with New Applications;
- F.-O. Speck, Advances in General Wiener–Hopf Factorization;
- I. Spitkovsky, One Hundred Years of... Numerical Range;
- N. Vasilevski, Algebras Generated by Toeplitz Operators on the Hardy Space $H^2(S^{2n-1})$;
- N. Zorboska, Toeplitz Operators on the Bergman Space with BMO^p Symbols and the Berezin Transform;

and more than 400 contributed talks distributed among 22 special sessions:

- Analysis and Algebraic Geometry for Operator Variables (organized by I. Klep and V. Vinnikov);
- Analysis and Synthesis for Operator Algebras (organized by A. Dor-On, R. Exel, E. Katsoulis, and D. Pitts);
- Free Analysis and Free Probability (organized by S. Belinschi, M. Popa, and R. Speicher);
- Functional Calculus, Spectral Sets and Constants (organized by L. Kosinski, F. Schwenninger, and M. Wojtylak);
- Gabor Analysis and Noncommutative Geometry (organized by F. Luef and I. Nikolaev);
- Geometry of Linear Operators and Operator Algebras (organized by K. Paul and D. Sain);
- Hypercomplex Analysis and Operator Theory (organized by D. Alpay, F. Colombo, and U. Kähler);
- Integral Operators and Applications (organized by R. Hagger, K.-M. Perfekt, and J. Virtanen);
- Linear Operators and Function Spaces (organized by M. C. Câmara and M. Ptak);
- Matrix Theory and Control (organized by M. Dodig and S. M. Furtado);
- Multivariable Operator Theory (organized by J. Ball and V. Bolotnikov);
- Operators of Harmonic Analysis, Related Function Spaces, and Applications — dedicated to Lars-Erik Persson on his 75th birthday (organized by H. Rafeiro and N. Samko);
- Operators on Reproducing Kernel Hilbert Spaces (organized by N. Vasilevski and K. Zhu);
- Operator Theoretical Methods in Mathematical Physics (organized by L. Castro and F.-O. Speck);
- Orders Preserving Operators on Cones and Applications (organized by M. Akian, S. Gaubert, A. Peperko, and G. Vigeral);
- Preserver Problems in Operator Theory and Functional Analysis (organized by F. Botelho and G. Geher);
- Random Matrix Theory (organized by H. Hedenmalm and J. Virtanen);
- Representation Theory of Algebras and Groups (organized by C. André, S. Lopes, and A. P. Santana);
- Semigroups and Evolution Equations (organized by C. Budde and C. Seifert);
- Spectral Theory and Differential Operators (organized by A. Khrabustovskyi, O. Post, and C. Trunk);

- Toeplitz Operators, Convolution Type Operators, and Operator Algebras — dedicated to Yuri Karlovich on his 70th birthday (organized by M. A. Bastos and A. Karlovich);
- Truncated Moment Problems (organized by M. Infusino and S. Kuhlmann).

Finally, the editors of this volume express their gratitude to the IWOTA 2019 sponsors: the Center for Functional Analysis, Linear Structures and Applications supported by the projects CEAFEL-UID/MAT/04721/2013, UID/MAT/04721/2019; the Center for Research and Development in Mathematics and Applications supported by the project CIDMA-UID/MAT/04106/2019; the Centre for Mathematics and Applications supported by the project CMA-UID/MAT/00297/2019; and to the Rector of Lisbon University and the Portuguese Foundation for Science and Technology.

Lisboa, Portugal — M. Amélia Bastos
Aveiro, Portugal — Luís Castro
Lisboa, Portugal — Alexei Yu. Karlovich
May 2020

Contents

Extended Heinz and Jensen Type Inequalities and Rearrangements

Shoshana Abramovich

Dedicated to Lars-Erik Persson on the occasion of his 75th birthday.

Abstract In this paper we extend the well known Heinz inequality which says that $2\sqrt{a_1a_2} \leq H(t) \leq a_1+a_2, a_1, a_2 > 0, 0 \leq t \leq 1$, where $H(t) = a_1^t a_2^{1-t} + a_1^{1-t}a_2^t$. We discuss the bounds of $H(t)$ in the intervals $t \in [1,2]$ and $t \in [2,\infty)$ using the subquadracity and the superquadracity of $\varphi(x) = x^t$, $x \geq 0$ respectively. Further, we extend $H(t)$ to get results related to $\sum_{i=1}^n H_i(t) = \sum_{i=1}^n \left(a_i^t a_{i+1}^{1-t} + a_i^{1-t}a_{i+1}^t\right)$, $a_{n+1} = a_1$, $a_i > 0$, $i = 1,\ldots,n$, where $H_1(t) = H(t)$. These results, obtained by using rearrangement techniques, show that the minimum and the maximum of the sum $\sum_{i=1}^n H_i(t)$ for a given t, depend only on the specific arrangements called circular alternating order rearrangement and circular symmetrical order rearrangement of a given set $(\mathbf{a}) = (a_1, a_2, \ldots, a_n)$, $a_i > 0, i = 1,2,\ldots,n$. These lead to extended Heinz type inequalities of $\sum_{i=1}^n H_i(t)$ for different intervals of t. The results may also be considered as special cases of Jensen type inequalities for concave, convex, subquadratic and superquadratic functions, which are also discussed in this paper.

Keywords Rearrangements · Heinz inequality · Jensen inequality · Convexity · Superquadracity

Mathematics Subject Classification (2010) 26D15, 26A51, 47A50, 47A60

S. Abramovich (✉)
Department of Mathematics, University of Haifa, Mount Carmel, Haifa, Israel
e-mail: abramos@math.haifa.ac.il

M. A. Bastos et al. (eds.), *Operator Theory, Functional Analysis and Applications*,
Operator Theory: Advances and Applications 282,
https://doi.org/10.1007/978-3-030-51945-2_1

1 Introduction and Basic Results

In this paper we extend the well known Heinz inequality

$$2\sqrt{a_1a_2} \le H(t) \le a_1 + a_2, \quad a_1, a_2 > 0, \quad 0 \le t \le 1, \tag{1.1}$$

where

$$H(t) = a_1^t a_2^{1-t} + a_1^{1-t} a_2^t.$$

In the past few years attention has been put toward refining or reversing this inequality. Recently, in 2019, new refinements of (1.1) have been proved in [5].

We start by discussing the bounds of $H(t)$ in the intervals $t \in [1, 2]$ and $t \in [2, \infty)$ using the subquadracity and the superquadracity of $\varphi(x) = x^t$, $x \ge 0$ respectively.

Then, we continue extending $H(t)$ to get results related to

$$\sum_{i=1}^{n} H_i(t) = \sum_{i=1}^{n} \left(a_i^t a_{i+1}^{1-t} + a_i^{1-t} a_{i+1}^t \right), \quad a_{n+1} = a_1, \quad a_i > 0, \quad i = 1, \dots, n,$$

where $H_1(t) = H(t)$ and $t \in \mathbb{R}$.

These results, obtained by using rearrangement techniques, show that the minimum and the maximum of the sum $\sum_{i=1}^{n} H_i(t)$ for a specific t, depend only on the specific arrangements called circular alternating order rearrangement and circular symmetrical order rearrangement of a given set $(\mathbf{a}) = (a_1, a_2, \dots, a_n)$, $a_i > 0$, $i = 1, 2, \dots, n$. These lead to extended Heinz type inequalities of $\sum_{i=1}^{n} H_i(t)$ for different intervals of t. The results may also be considered as special cases of Jensen type inequalities for concave, convex, subquadratic and superquadratic functions, which are also discussed in this paper.

We start with a lemma which easily leads to (1.1):

Lemma 1.1 *Let $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$ be a non-negative increasing function and let $G_p(a, b)$ be the function*

$$G_p(a, b) = a \left(\frac{b}{a} \right)^p \varphi\left(\frac{b}{a} \right) + b \left(\frac{a}{b} \right)^p \varphi\left(\frac{a}{b} \right), \quad p \ge \frac{1}{2}, \quad a, b > 0.$$

Then,

$$G_p(a, b) \ge a^{\frac{1}{2}} b^{\frac{1}{2}} \left(\varphi\left(\frac{b}{a} \right) + \varphi\left(\frac{a}{b} \right) \right). \tag{1.2}$$

Proof Computing the derivative of $G_p(a,b)$ we see that

$$\frac{\partial G_p(a,b)}{\partial p} = \left(a\left(\frac{b}{a}\right)^p \varphi\left(\frac{b}{a}\right) - b\left(\frac{a}{b}\right)^p \varphi\left(\frac{a}{b}\right)\right)\ln\frac{b}{a} \geq 0, \tag{1.3}$$

when without loss of generality we assume that $b \geq a > 0$. Therefore (1.2) follows from (1.3). □

When $\varphi(x) \equiv 1$ we get from Lemma 1.1:

Corollary 1.2 *Let $a, b, t \in \mathbb{R}$ and let the symmetric function around $t = \frac{1}{2}$ be*

$$H(t) = a^{1-t}b^t + a^t b^{1-t}, \qquad a, b > 0.$$

Then $H(t)$ is symmetric around $t = \frac{1}{2}$, increasing when $t \geq \frac{1}{2}$, decreasing when $t < \frac{1}{2}$ and the inequality

$$H(t) = a^{1-t}b^t + a^t b^{1-t} \geq 2a^{\frac{1}{2}}b^{\frac{1}{2}}$$

holds for any $t \in \mathbb{R}$.

Also, because of the monotonicity and the symmetry of $H(t)$ we get that for $a, b > 0$,

$$2a^{\frac{1}{2}}b^{\frac{1}{2}} \leq a^{1-t}b^t + a^t b^{1-t} \leq a + b, \qquad 0 < t < 1.$$

When $n \leq t \leq n+1$ and $n = 1, 2, \ldots,$

$$a^{1-n}b^n + a^n b^{1-n} \leq a^{1-t}b^t + a^t b^{1-t} \leq a^{-n}b^{n+1} + a^{n+1}b^{-n}.$$

When $n \leq t \leq n+1$ and $n = -1, -2, \ldots,$

$$a^{n+1}b^{-n} + a^{-n}b^{n+1} \leq a^{1-t}b^t + a^t b^{1-t} \leq a^n b^{-n+1} + a^{-n+1}b^n.$$

In order to obtain more Heinz type inequalities we introduce the definitions of superquadratic and subquadratic functions (see for instance [2, 4] and [7]) which include the functions $f(x) = x^t$, $x \geq 0$, when $t \geq 2$ and $1 \leq t \leq 2$, respectively.

Definition 1.3 A function $\varphi : [0, B) \to \mathbb{R}, 0 < B \leq \infty$ is *superquadratic* provided that for all $x \in [0, B)$ there exists a constant $C_\varphi(x) \in \mathbb{R}$ such that the inequality

$$\varphi(y) \geq \varphi(x) + C_\varphi(x)(y - x) + \varphi(|y - x|)$$

holds for all $y \in [0, B)$, (see [2, Definition 2.1], there $[0, \infty)$ instead $[0, B)$).

The function φ is called *subquadratic* if $-\varphi$ is superquadratic.

Proposition 1.4 *Suppose that f is superquadratic. Let $0 \leq x_i < B$, $i = 1, \ldots, n$ and let $\overline{x} = \sum_{i=1}^{n} a_i x_i$, where $a_i \geq 0$, $i = 1, \ldots, n$ and $\sum_{i=1}^{n} a_i = 1$. Then*

$$\sum_{i=1}^{n} a_i f(x_i) - f(\overline{x}) \geq \sum_{i=1}^{n} a_i f(|x_i - \overline{x}|). \tag{1.4}$$

If f is non-negative, it is also convex and the inequality refines Jensen's inequality. In particular, the functions $f(x) = x^r$, $x \geq 0$, $r \geq 2$ are superquadratic and convex, and equality holds in inequality (1.4) when $r = 2$.

Similarly, suppose that f is subquadratic. Let $0 \leq x_i < B$, $i = 1, \ldots, n$ and let $\overline{x} = \sum_{i=1}^{n} a_i x_i$, where $a_i \geq 0$, $i = 1, \ldots, n$ and $\sum_{i=1}^{n} a_i = 1$. Then

$$\sum_{i=1}^{n} a_i f(x_i) - f(\overline{x}) \leq \sum_{i=1}^{n} a_i f(|x_i - \overline{x}|). \tag{1.5}$$

In Theorem 1.5 we employ for $t \geq 2$ the inequality satisfied by the superquadratic function $\phi(x) = x^t$.

Theorem 1.5 *Let $a, b > 0$ and $t \geq 2$. Then*

$$\begin{aligned} a^t b^{1-t} + a^{1-t} b^t &\geq (a+b) + \left(a^{1-t} + b^{1-t}\right) |b-a|^t \\ &\geq \max\left\{(a+b)\left(1 + \left(\frac{2|b-a|}{a+b}\right)^t\right), \frac{a^2}{b} + \frac{b^2}{a}\right\}. \end{aligned} \tag{1.6}$$

Proof From the superquadracity of $\phi(x) = x^t$, $t \geq 2$, $n = 2$, we get that

$$\begin{aligned} b\left(\frac{a}{b}\right)^t + a\left(\frac{b}{a}\right)^t &= a^t b^{1-t} + a^{1-t} b^t \\ &\geq (a+b)\left(1 + \frac{a}{a+b}\left|\frac{b}{a} - 1\right|^t + \frac{b}{a+b}\left|1 - \frac{a}{b}\right|^t\right) \\ &= (a+b)\left(1 + \frac{a^{1-t} + b^{1-t}}{a+b}|b-a|^t\right) \\ &\geq \max\left\{(a+b)\left(1 + \left(\frac{2|b-a|}{a+b}\right)^t\right), \frac{a^2}{b} + \frac{b^2}{a}\right\}. \end{aligned} \tag{1.7}$$

Indeed, the first inequality in (1.7) follows from the superquadracity of $\phi(x) = x^t$, $t \geq 2$, $x \geq 0$. The second inequality follows from the convexity of $\phi(x) = x^t$, $t \geq 1$, $x \geq 0$, and the monotonicity of $a^t b^{1-t} + a^{1-t} b^t$.

The proof is complete. □

It is easy to verify that there are cases by which the superquadratic property leads to a better result in (1.6) than the monotonicity property of $H(t)$ and vice versa.

In Theorem 1.6 we employ for $1 \leq t \leq 2$ the inequality satisfied by the subquadratic function $\phi(x) = x^t$.

Theorem 1.6 *Let $a, b > 0$ and $1 < t < 2$. Then*

$$\begin{aligned} a+b &\leq a^t b^{1-t} + a^{1-t} b^t \\ &\leq \min\left\{(a+b) + \left(a^{1-t} + b^{1-t}\right)|b-a|^t, \frac{a^2}{b} + \frac{b^2}{a}\right\} \end{aligned} \tag{1.8}$$

holds.

Proof The left hand-side of (1.8) follows from the monotonicity of $a^t b^{1-t} + a^{1-t} b^t$, $x \geq 0$, $t \geq \frac{1}{2}$. The right hand-side follows from the subquadracity of $\phi(x) = x^t$, $x \geq 0$, $1 \leq t \leq 2$, and the monotonicity of $a^t b^{1-t} + a^{1-t} b^t$, $1 \leq t \leq 2$.

The proof is complete. □

In the sequel we use the following lemma.

Lemma 1.7 *Let φ be a continuous function on $x \geq 0$ which is twice differentiable on $x > 0$, with $\varphi(0) = 0$ and*

$$\lim_{x \to 0^+} x\varphi'(x) = 0.$$

If φ is convex, then $x\varphi'(x) - \varphi(x) \geq 0$ on $x > 0$.
If φ is concave, then $x\varphi' x) - \varphi(x) \leq 0$ on $x > 0$.

Proof Assuming φ is convex, we have $\varphi''(x) \geq 0$ for $x > 0$, hence

$$(x\varphi'(x) - \varphi(x))' = x\varphi''(x) \geq 0$$

for $x > 0$, and $x\varphi'(x) - \varphi(x)$ is increasing. Also

$$\lim_{x \to 0^+} (x\varphi'(x) - \varphi(x)) = 0,$$

and therefore $x\varphi'(x) - \varphi(x) \geq 0$ when $x > 0$.

Similarly we prove the concavity case. □

2 Rearrangements and New Jensen and Heinz Type Inequalities

In [3] the authors deal with Jensen type inequalities and rearrangements. The following definitions appear first in [6]. Theorem 2.6 below appears in [3].

Definition 2.1 ([6]) An ordered set $(\mathbf{x}) = (x_1, \ldots, x_n)$ of n real numbers is arranged in *symmetrical decreasing order* if

$$x_1 \leq x_n \leq x_2 \leq \cdots \leq x_{[(n+2/2)]}$$

or if

$$x_n \leq x_1 \leq x_{n-1} \leq \cdots \leq x_{[(n+1/2)]}.$$

Definition 2.2 ([6]) A *circular rearrangement* of an ordered set $(\mathbf{x})$ is a cyclic rearrangement of $(\mathbf{x})$ or a cyclic rearrangement followed by inversion.

For example, the circular rearrangements of the ordered set $(1, 2, 3, 4)$ are the sets

$$(1, 2, 3, 4), (2, 3, 4, 1), (3, 4, 1, 2), (4, 1, 2, 3),$$
$$(4, 3, 2, 1), (1, 4, 3, 2), (2, 1, 4, 3), (3, 2, 1, 4).$$

Definition 2.3 ([6]) A set $(\mathbf{x})$ is arranged in *circular symmetrical order* if one of its circular rearrangements is symmetrically decreasing.

The alternating order of $(\mathbf{x})$ as defined in Definition 2.4 below was introduced and proved in [8] to be the minimum of $\sum_{i=1}^{n} x_i x_{i+1}$ under rearrangement. Here the minimum and maximum of

$$\sum_{i=1}^{n} \left(x_i \varphi \left(\frac{x_{i+1}}{x_i} \right) + x_{i+1} \varphi \left(\frac{x_i}{x_{i+1}} \right) \right)$$

and of

$$\sum_{i=1}^{n} \left(x_i^t x_{i+1}^{1-t} + x_i^{1-t} x_{i+1}^t \right)$$

under rearrangement of $(\mathbf{x})$ is obtained, which leads to Heinz and Jensen type inequalities.

Definition 2.4 An ordered set $(\mathbf{x}) = (x_1, \ldots, x_n)$ of n real numbers is arranged in *alternating order* if

$$x_1 \leq x_{n-1} \leq x_3 \leq x_{n-3} \leq \cdots \leq x_{\left[\frac{n+1}{2}\right]} \leq \cdots \leq x_4 \leq x_{n-2} \leq x_2 \leq x_n, \quad (2.1)$$

or if

$$x_n \leq x_2 \leq x_{n-2} \leq x_4 \leq \cdots \leq x_{\left[\frac{n+1}{2}\right]} \leq \cdots \leq x_{n-3} \leq x_3 \leq x_{n-1} \leq x_1. \quad (2.2)$$

Definition 2.5 A set $(\mathbf{x})$ is arranged in *circular alternating order* if one of its circular rearrangements is arranged in an alternating order.

Theorem 2.6 ([3]) *Let $F(u, v)$ be differentiable and symmetric real function defined on (α, β), $-\infty \leq \alpha < \beta \leq \infty$ and $\alpha \leq u, v, w \leq \beta$. Assume that*

$$\frac{\partial F(v, u)}{\partial v} \leq \frac{\partial F(v, w)}{\partial v}$$

for $u \leq \min\{w, v\}$. Then, for any set $(\mathbf{x}) = (x_1, x_2, \ldots, x_n)$, $\alpha \leq x_i \leq \beta$, $i = 1, \ldots, n$ given except its arrangements

$$\sum_{i=1}^{n} F(x_i, x_{i+1}), \qquad x_{n+1} = x_1$$

is maximal if $(\mathbf{x})$ is arranged in circular symmetrical order and minimal if $(\mathbf{x})$ is arranged in circular alternating order as defined above.

It is proved in [1] that the maximal arrangement of $\sum_{i=1}^{n} F(x_i, x_{i+1})$, $x_{n+1} = x_1$ is attained when $(\mathbf{x})$ is arranged in circular symmetrical order.

An outline of the minimal arrangement of $\sum_{i=1}^{n} F(x_i, x_{i+1})$, $x_{n+1} = x_1$ as stated in Theorem 2.6 and proved in [3] is as follows:

We denote a given set of n real numbers according to their increasing order $(\mathbf{a}) = (a_1, a_2, \ldots, a_{n-1}, a_n)$, where $a_1 \leq a_2 \leq \cdots \leq a_{n-1} \leq a_n$. We start with an arbitrary permutation of $(\mathbf{a})$ called $(\mathbf{b}) = (b_1, \ldots, b_n)$. As $F(u, v)$ is symmetric and we are interested in $\sum_{i=1}^{n} F(b_i, b_{i+1})$, $b_{n+1} = b_1$, which is clearly invariant under all circular rearrangements, we can assume that $b_1 = a_1$. Now we go through three permutations which bring us from $(\mathbf{b}) \to (\mathbf{c}) \to (\mathbf{d}) \to (\mathbf{e})$ in which

$$\sum_{i=1}^{n} F(b_i, b_{i+1}) \geq \sum_{i=1}^{n} F(c_i, c_{i+1}) \geq \sum_{i=1}^{n} F(d_i, d_{i+1}) \geq \sum_{i=1}^{n} F(e_i, e_{i+1}),$$

and we make sure that the two first and two last numbers in $(\mathbf{e})$ are $e_1 = a_1$, $e_2 = a_{n-1}$, $e_{n-1} = a_2$, $e_n = a_n$ which are already the two first and the two last in the rearrangements of the alternating order of type (2.1). We realize also that when we

check $(e_2, e_3, \ldots, e_{n-2}, e_{n-1})$ we already have that e_2 and e_{n-1} are the largest and the smallest numbers respectively in $(e_2, e_3, \ldots, e_{n-2}, e_{n-1})$.

Now we use the induction procedure: We assume the validity of Theorem 2.6 for the set of $n-2$ numbers and show that this implies its validity for the set of n numbers. More specifically, the $n-2$ numbers if rearranged in alternating order of (2.2) give, according to the induction hypothesis, the smallest value of $\sum_{i=2}^{n-2} F(e_i, e_{i+1})$ and in the same time we get that $(e_1, \ldots, e_n)$ is arranged in alternating order too, this time according to (2.1) and therefore the proof by induction for n numbers is obtained.

In Theorem 2.7 we introduce the symmetric function $F(u, v) = u\varphi\left(\frac{v}{u}\right) + v\varphi\left(\frac{u}{v}\right)$ and show when it satisfies Theorem 2.6. Similarly, in Theorem 2.8 we introduce the symmetric function $F(u, v) = u^s v^t + u^t v^s$ and show when it satisfies Theorem 2.6.

We denote in the sequel $(\widetilde{\mathbf{x}}) = (\widetilde{x}_1, \widetilde{x}_2, \ldots, \widetilde{x}_n)$ and $(\widehat{\mathbf{x}}) = (\widehat{x}_1, \widehat{x}_2, \ldots, \widehat{x}_n)$ to be the circular alternating order rearrangement and circular symmetrical order rearrangement of a given set $(\mathbf{x}) = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$.

Theorem 2.7 *Assume φ is a concave differentiable function on $\mathbb{R}_+$ and*

$$\lim_{x \to 0^+} (x\varphi'(x) - \varphi(x)) = 0.$$

Then the inequalities

$$\begin{aligned} &\sum_{i=1}^{n} \left(\widetilde{x}_i \varphi\left(\frac{\widetilde{x}_{i+1}}{\widetilde{x}_i} \right) + \widetilde{x}_{i+1} \varphi\left(\frac{\widetilde{x}_i}{\widetilde{x}_{i+1}} \right) \right) \\ &\quad \le \sum_{i=1}^{n} \left(x_i \varphi\left(\frac{x_{i+1}}{x_i} \right) + x_{i+1} \varphi\left(\frac{x_i}{x_{i+1}} \right) \right) \\ &\quad \le \sum_{i=1}^{n} \left(\widehat{x}_i \varphi\left(\frac{\widehat{x}_{i+1}}{\widehat{x}_i} \right) + \widehat{x}_{i+1} \varphi\left(\frac{\widehat{x}_i}{\widehat{x}_{i+1}} \right) \right) \end{aligned} \tag{2.3}$$

hold where $x_i > 0$, $i = 1, 2, \ldots, n$ and $\widetilde{x}_{n+1} = \widetilde{x}_1$, $x_{n+1} = x_1$, $\widehat{x}_{n+1} = \widehat{x}_1$.

Assume φ is a convex differentiable function on $\mathbb{R}_+$ and

$$\lim_{x \to 0^+} (x\varphi'(x) - \varphi(x)) = 0.$$

Then the reverse of inequalities (2.3) hold.

Proof Using Theorem 2.6 we define

$$F(u, v) = u\varphi\left(\frac{v}{u}\right) + v\varphi\left(\frac{u}{v}\right).$$

Then, for $u \leq v, w$

$$\frac{\partial F(u,v)}{\partial v} = \varphi'\left(\frac{v}{u}\right) + \varphi\left(\frac{u}{v}\right) - \frac{u}{v}\varphi'\left(\frac{u}{v}\right)$$
$$\leq \varphi'\left(\frac{v}{w}\right) + \varphi\left(\frac{w}{v}\right) - \frac{w}{v}\varphi'\left(\frac{w}{v}\right) = \frac{\partial F(w,v)}{\partial v}.$$

Indeed, φ is concave, that is, φ' is decreasing, and

$$\varphi'\left(\frac{v}{u}\right) \leq \varphi'\left(\frac{v}{w}\right).$$

Moreover,

$$\varphi\left(\frac{u}{v}\right) - \frac{u}{v}\varphi'\left(\frac{u}{v}\right) \leq \varphi\left(\frac{w}{v}\right) - \frac{w}{v}\varphi'\left(\frac{w}{v}\right)$$

according to Lemma 1.7. Hence, according to Theorem 2.6 inequalities (2.3) hold. Similarly, we get that the reversed (2.3) hold when φ is convex. □

Theorem 2.6 above is used in Theorem 2.8 to prove an extension of Heinz type inequalities.

Theorem 2.8 *Let* $F(x, y) = x^s y^t + x^t y^s$.

(a) *If* $x, y, s, t \in \mathbb{R}_+$, *then for* $(\mathbf{x}) \in \mathbb{R}_+^n$, *the inequalities*

$$\sum_{i=1}^{n}\left(\widetilde{x}_i^t \widetilde{x}_{i+1}^s + \widetilde{x}_{i+1}^t \widetilde{x}_i^s\right) \leq \sum_{i=1}^{n}\left(x_i^s x_{i+1}^t + x_{i+1}^t x_i^s\right)$$
$$\leq \sum_{i=1}^{n}\left(\widehat{x}_i^t \widehat{x}_{i+1}^s + \widehat{x}_{i+1}^t \widehat{x}_i^s\right), \tag{2.4}$$

hold, where $\widetilde{x}_{n+1} = \widetilde{x}_1$, $x_{n+1} = x_1$, $\widehat{x}_{n+1} = \widehat{x}_1$, *and* $(\widetilde{\mathbf{x}})$ *is the circular alternating order rearrangement of* $(\mathbf{x})$ *and* $(\widehat{\mathbf{x}})$ *is the circular symmetrical order rearrangement of* $(\mathbf{x})$. *In particular,* (2.4) *holds when* $t + s = 1$ *and* $0 \leq t \leq 1$.

(b) *If* $s \leq 0$, $t \geq 0$ *and* $x_i > 0$, $i = 1, 2, \ldots, n$, *the reverse of* (2.4) *holds. In particular, the reverse of* (2.4) *holds when* $t + s = 1$ *and* $t \geq 1$.

Proof We use Theorem 2.6, which guarantees that (2.4) holds for $F(x, y) = x^s y^t + x^t y^s$ if

$$\frac{\partial F(u,v)}{\partial v} \leq \frac{\partial F(w,v)}{\partial v} \quad \text{when } u \leq \min(v, w).$$

Since

$$\frac{\partial F(u,v)}{\partial v}=\frac{\partial\left(u^s v^t+u^t v^s\right)}{\partial v}=tu^s v^{t-1}+su^t v^{s-1},$$

we have to show that

$$tu^s v^{t-1}+su^t v^{s-1}\le tw^s v^{t-1}+sw^t v^{s-1}\quad\text{when } u\le\min(v,w).$$

This holds because

$$tv^{t-1}\left(u^s-w^s\right)\le 0\le sv^{s-1}\left(w^t-u^t\right)\quad\text{when } u\le\min(v,w). \tag{2.5}$$

This proves Case (a).

In the same way as Case (a) is proved for $s,t>0$, we see that in Case (b) we get the reverse of (2.5) and therefore we get the reverse of (2.4).

The proof is complete. □

Remark 2.9 Inequalities (2.4) when $0\le t\le 1$ and $t+s=1$, and the reverse of inequalities (2.4) when $t\ge 1$ and $t+s=1$ follow also from Theorem 2.7 when the function is $\varphi(x)=x^t$, $x>0$.

In Theorem 2.10 we combine the results obtained in Theorem 2.7 concerning rearrangement with inequalities satisfied by concave functions, convex functions, subquadratic functions and superquadratic functions to get new Jensen type inequalities.

Theorem 2.10 *Let $x_i>0$, $i=1,2,\ldots,n$ and let $(\widetilde{\mathbf{x}})$ and $(\widehat{\mathbf{x}})$ be the circular alternating order rearrangement and the circular symmetrical order rearrangement respectively of $(\mathbf{x})=(x_1,x_2,\ldots,x_n)$. Let $\widetilde{x}_{n+1}=\widetilde{x}_1$, $x_{n+1}=x_1$, $\widehat{x}_{n+1}=\widehat{x}_1$. Then*

(a) *For a concave function $\varphi:\mathbb{R}_+\to\mathbb{R}$ the following inequalities hold:*

$$\begin{aligned}
&\sum_{i=1}^{n}\left(\widetilde{x}_i\varphi\left(\frac{\widetilde{x}_{i+1}}{\widetilde{x}_i}\right)+\widetilde{x}_{i+1}\varphi\left(\frac{\widetilde{x}_i}{\widetilde{x}_{i+1}}\right)\right)\\
&\quad\le\sum_{i=1}^{n}\left(x_i\varphi\left(\frac{x_{i+1}}{x_i}\right)+x_{i+1}\varphi\left(\frac{x_i}{x_{i+1}}\right)\right)\\
&\quad\le\sum_{i=1}^{n}\left(\widehat{x}_i\varphi\left(\frac{\widehat{x}_{i+1}}{\widehat{x}_i}\right)+\widehat{x}_{i+1}\varphi\left(\frac{\widehat{x}_i}{\widehat{x}_{i+1}}\right)\right)\\
&\quad\le 2\varphi(1)\sum_{i=1}^{n}x_i.
\end{aligned}\tag{2.6}$$

(b) *For a convex and subquadratic function φ the inequalities*

$$\begin{aligned}
&2\varphi(1)\sum_{i=1}^{n}\widetilde{x}_i+\sum_{i=1}^{n}\widetilde{x}_i\varphi\left(\left|\frac{\widetilde{x}_i-\widetilde{x}_{i+1}}{\widetilde{x}_i}\right|\right)+\widetilde{x}_{i+1}\varphi\left(\left|\frac{\widetilde{x}_i-\widetilde{x}_{i+1}}{\widetilde{x}_{i+1}}\right|\right)\\
&\quad\geq\sum_{i=1}^{n}\widetilde{x}_i\varphi\left(\frac{\widetilde{x}_{i+1}}{\widetilde{x}_i}\right)+\widetilde{x}_{i+1}\varphi\left(\frac{\widetilde{x}_i}{\widetilde{x}_{i+1}}\right)\\
&\quad\geq\sum_{i=1}^{n}x_i\varphi\left(\frac{x_{i+1}}{x_i}\right)+x_{i+1}\varphi\left(\frac{x_i}{x_{i+1}}\right)\\
&\quad\geq\sum_{i=1}^{n}\widehat{x}_i\varphi\left(\frac{\widehat{x}_{i+1}}{\widehat{x}_i}\right)+\widehat{x}_{i+1}\varphi\left(\frac{\widehat{x}_i}{\widehat{x}_{i+1}}\right)\\
&\quad\geq 2\varphi\,(1)\sum_{i=1}^{n}x_i
\end{aligned}\tag{2.7}$$

hold.

(c) *For a superquadratic function φ the inequalities*

$$\begin{aligned}
&\sum_{i=1}^{n}\left(\widetilde{x}_i\varphi\left(\frac{\widetilde{x}_{i+1}}{\widetilde{x}_i}\right)+\widetilde{x}_{i+1}\varphi\left(\frac{\widetilde{x}_i}{\widetilde{x}_{i+1}}\right)\right)\\
&\quad\geq\sum_{i=1}^{n}\left(x_i\varphi\left(\frac{x_{i+1}}{x_i}\right)+x_{i+1}\varphi\left(\frac{x_i}{x_{i+1}}\right)\right)\\
&\quad\geq\sum_{i=1}^{n}\left(\widehat{x}_i\varphi\left(\frac{\widehat{x}_{i+1}}{\widehat{x}_i}\right)+\widehat{x}_{i+1}\varphi\left(\frac{\widehat{x}_i}{\widehat{x}_{i+1}}\right)\right)\\
&\quad\geq\sum_{i=1}^{n}\left(\widehat{x}_i\varphi\left(\left|\frac{\widehat{x}_i-\widehat{x}_{i+1}}{\widehat{x}_i}\right|\right)+\widehat{x}_{i+1}\varphi\left(\left|\frac{\widehat{x}_i-\widehat{x}_{i+1}}{\widehat{x}_{i+1}}\right|\right)\right)\\
&\qquad+2\varphi(1)\sum_{i=1}^{n}\widehat{x}_i
\end{aligned}\tag{2.8}$$

hold.

Proof We start with the proof of Case (a). From the concavity of φ we get that

$$x_i\varphi\left(\frac{x_{i+1}}{x_i}\right)+x_{i+1}\varphi\left(\frac{x_i}{x_{i+1}}\right)\leq\varphi(1)\,(x_i+x_{i+1}),\ i=1,\ldots,n,\ x_{n+1}=x_1.$$

Therefore,

$$\sum_{i=1}^{n}\left(x_i\varphi\left(\frac{x_{i+1}}{x_i}\right)+x_{i+1}\varphi\left(\frac{x_i}{x_{i+1}}\right)\right)\leq 2\varphi(1)\sum_{i=1}^{n}x_i.$$

Together with Theorem 2.7 we get inequalities (2.6).

Case (b): Inequality (1.5) says that for $n=2$ when φ is subquadratic, the inequality

$$\begin{aligned}&x_i\varphi\left(\frac{x_{i+1}}{x_i}\right)+x_{i+1}\varphi\left(\frac{x_i}{x_{i+1}}\right)\\&\quad\leq(x_i+x_{i+1})\,\varphi(1)+x_i\varphi\left(\left|1-\frac{x_{i+1}}{x_i}\right|\right)+x_{i+1}\varphi\left(\left|1-\frac{x_i}{x_{i+1}}\right|\right)\end{aligned}$$

holds. Therefore summing up, we get

$$\begin{aligned}&2\varphi(1)\sum_{i=1}^{n}x_i+\sum_{i=1}^{n}\left(x_i\varphi\left(\left|1-\frac{x_{i+1}}{x_i}\right|\right)+x_{i+1}\varphi\left(\left|1-\frac{x_i}{x_{i+1}}\right|\right)\right)\\&\quad\geq\sum_{i=1}^{n}\left(x_i\varphi\left(\frac{x_{i+1}}{x_i}\right)+x_{i+1}\varphi\left(\frac{x_i}{x_{i+1}}\right)\right),\end{aligned}$$

and together with Theorem 2.7 we get that inequalities (2.7) hold.

Case (c) follows similarly from (1.4) and Theorem 2.7. The proof is complete. □

Theorem 2.11 combines the results related to rearrangements proved in Theorem 2.7 with the concavity of x^t, $x\geq 0$, $0<t\leq 1$, convexity and subquadracity properties of x^t, $x\geq 0$, for $1\leq t\leq 2$ and its superquadracity properties for $t\geq 2$, to get what we call: "Extended Heinz type inequalities".

Alternatively the results of Theorem 2.11 can also be obtained from Theorem 2.8 for $0\leq t<1$, $s=1-t$ and for $t\geq 1$ and $s=1-t$, combined with the concavity of x^t, $x\geq 0$, $0<t\leq 1$, convexity and subquadracity properties of x^t, $x\geq 0$, for $1\leq t\leq 2$ and its superquadracity properties for $t\geq 2$, to get the "Extended Heinz type inequalities".

Theorem 2.11 *Let* $(\mathbf{x})\in\mathbb{R}^n_+$, $(\widetilde{\mathbf{x}})$ *and* $(\widehat{\mathbf{x}})$ *its circular alternating order rearrangement and its circular symmetrical order rearrangement respectively, then when*

$\widetilde{x}_{n+1} = \widetilde{x}_1$, $x_{n+1} = x_1$, $\widehat{x}_{n+1} = \widehat{x}_1$, *we get three cases of Extended Heinz type inequalities:*

(a) *For* $0 \le t < 1$, *the inequalities*

$$\begin{aligned} 2\sum_{i=1}^{n} \widetilde{x}_i^{\frac{1}{2}} \widetilde{x}_{i+1}^{\frac{1}{2}} &\le \sum_{i=1}^{n} \left(\widetilde{x}_i^t \widetilde{x}_{i+1}^{1-t} + \widetilde{x}_{i+1}^t \widetilde{x}_i^{1-t} \right) \\ &\le \sum_{i=1}^{n} \left(x_i^t x_{i+1}^{1-t} + x_{i+1}^t x_i^{1-t} \right) \\ &\le \sum_{i=1}^{n} \left(\widehat{x}_i^t \widehat{x}_{i+1}^{1-t} + \widehat{x}_{i+1}^t \widehat{x}_i^{1-t} \right) \\ &\le \sum_{i=1}^{n} 2x_i \end{aligned}$$

hold.

(b) *For* $1 \le t \le 2$ *and* $x_i > 0$, $i = 1, 2, \ldots, n$, *the inequalities*

$$\begin{aligned} 2\sum_{i=1}^{n} \widetilde{x}_i + \sum_{i=1}^{n} \left(\widetilde{x}_i^{1-t} + \widetilde{x}_{i+1}^{1-t} \right) (|\widetilde{x}_i - \widetilde{x}_{i+1}|)^t &\ge \sum_{i=1}^{n} \left(\widetilde{x}_i^t \widetilde{x}_{i+1}^{1-t} + \widetilde{x}_{i+1}^t \widetilde{x}_i^{1-t} \right) \\ &\ge \sum_{i=1}^{n} \left(x_i^{1-t} x_{i+1}^t + x_{i+1}^t x_i^{1-t} \right) \\ &\ge \sum_{i=1}^{n} \left(\widehat{x}_i^t \widehat{x}_{i+1}^{1-t} + \widehat{x}_{i+1}^t \widehat{x}_i^{1-t} \right) \\ &\ge 2\sum_{i=1}^{n} x_i \end{aligned}$$

hold.

(c) *For* $t \ge 2$ *and* $x_i > 0$, $i = 1, 2, \ldots, n$, *the inequalities*

$$\begin{aligned} \sum_{i=1}^{n} \left(\widetilde{x}_i^t \widetilde{x}_{i+1}^{1-t} + \widetilde{x}_{i+1}^t \widetilde{x}_i^{1-t} \right) &\ge \sum_{i=1}^{n} \left(x_i^{1-t} x_{i+1}^t + x_{i+1}^t x_i^{1-t} \right) \\ &\ge \sum_{i=1}^{n} \left(\widehat{x}_i^t \widehat{x}_{i+1}^{1-t} + \widehat{x}_{i+1}^t \widehat{x}_i^{1-t} \right) \\ &\ge 2\sum_{i=1}^{n} \widehat{x}_i + \sum_{i=1}^{n} \left(\widehat{x}_i^{1-t} + \widehat{x}_{i+1}^{1-t} \right) (|\widehat{x}_i - \widehat{x}_{i+1}|)^t \end{aligned}$$

hold.

Proof From (2.6) in Theorem 2.10 for $\varphi(x) = x^t$, $0 \leq t \leq 1$ together with the left hand-side inequality (1.1) we get that Case (a) holds.

From (2.7) in Theorem 2.10 for $\varphi(x) = x^t$, $1 \leq t \leq 2$ we get that Case (b) holds.

From (2.8) in Theorem 2.10 for $\varphi(x) = x^t$, $t \geq 2$ we get that Case (c) holds.

Alternatively we obtain the result of the theorem using Theorem 2.8 and the concavity, convexity, subquadracity and superquadracity of x^t, $x > 0$ in the relevant intervals of t. The proof is complete. □

References

1. S. Abramovich, The increase of sums and products dependent on $(y_1, \ldots, y_n)$ by rearrangement of this set. Israel J. Math. **5**, 177–181 (1967)
2. S. Abramovich, G. Jameson, G. Sinnamon, Refining Jensen's inequality. Bull. Math. Soc. Sci. Math. Roumanie (N.S.) **47(95)**(1–2), 3–14 (2004)
3. S. Abramovich, L.-E. Persson, Rearrangements and Jensen type inequalities related to convexity, superquadracity, strong convexity and 1-quasiconvexity. J. Math. Inequal. **14**, 641–659 (2020).
4. S. Abramovich, L.-E. Persson, N. Samko, On γ-quasiconvexity, superquadracity and two sided reversed Jensen type inequalities. Math. Inequal. Appl. **18**, 615–628 (2015)
5. F. Kittaneh, M.S. Moslehian, M. Sababheh, Quadratic interpolation of the Heinz mean. Math. Inequal. Appl. **21**, 739–757 (2018)
6. A.L. Lehman, Results on rearrangements. Israel J. Math. **1**, 22–28 (1963)
7. C.P. Niculescu, L.-E. Persson, *Convex Functions and Their Applications - A Contemporary Approach*, 2nd edn. (Springer, Cham, 2018)
8. H. Yu, Circular rearrangement inequality. J. Math. Inequal. **12**, 635–643 (2018)

On Some Applications of Representable and Continuous Functionals of Banach Quasi *-Algebras

Maria Stella Adamo

Abstract This survey aims to highlight some of the consequences that representable (and continuous) functionals have in the framework of Banach quasi *-algebras. In particular, we look at the link between the notions of *-semisimplicity and full representability in which representable functionals are involved. Then, we emphasize their essential role in studying *-derivations and representability properties for the tensor product of Hilbert quasi *-algebras, a special class of Banach quasi *-algebras.

Keywords Representable functionals · Banach and Hilbert quasi *-algebras · Weak derivations on Banach quasi *-algebras · Tensor product of Hilbert quasi *-algebras

Mathematics Subject Classification (2010) Primary 46L08; Secondary 46A32, 46L57, 46L89, 47L60

1 Introduction and Preliminaries

The investigation of (locally convex) quasi *-algebras was undertaken around the beginning of the ’80s, in the last century, to give a solution to specific problems concerning quantum statistical mechanics and quantum field theory, that required instead a representation of observables as *unbounded* operators, see, e.g., [9, 28]. They were introduced by G. Lassner in the series of papers [21] and [22] in 1988.

A particular interest has been shown for the theory of *-representations of quasi *-algebras in a specific family of unbounded densely defined and closable operators. In this framework, a central role is played by *representable functionals*, i.e., those functionals that admit a GNS-like construction. In the process of looking

M. S. Adamo (✉)
Dipartimento di Matematica, Università di Roma “Tor Vergata”, Roma, Italy
e-mail: adamo@axp.uniroma2.it; msadamo@unict.it

M. A. Bastos et al. (eds.), *Operator Theory, Functional Analysis and Applications*,
Operator Theory: Advances and Applications 282,
https://doi.org/10.1007/978-3-030-51945-2_2

at structural properties of locally convex quasi *-algebras, *-semisimplicity and full representability are the critical notions involved, making an extensive use of representable (and continuous) functionals [4, 7, 15, 30].

The goal of this survey is to point out some of the various connections that representable and continuous functionals for Banach quasi *-algebras have and examine their applications, e.g., the study of unbounded *-derivations and tensor products. The importance of investigating these themes stem from the study of physical phenomena. Moreover, very little is known about unbounded *-derivations and tensor products in the framework of unbounded operator algebras. The reader is referred to as [16–18, 31–33].

Banach quasi *-algebras constitute a particular subclass of locally convex *-algebras. In this context, the interesting question is to understand whether representable functionals are continuous. The reasons to get to know about the continuity of these functionals are several. Among them, the continuity is a crucial feature for representable functionals, since it would reflect on the sesquilinear forms and the *-representations associated with them through the GNS-like construction (see [30]). Furthermore, no example of representable functional that is *not continuous* is known in the literature.

A positive answer to this question has been given for the space $L^2(I, d\lambda)$, where $I = [0, 1]$ and λ is the Lebesgue measure, over continuous or essentially bounded functions, and more in general, for commutative *Hilbert* quasi *-algebras under certain conditions. In the case of L^p-spaces for $p \geq 1$, we observe a discontinuous behaviour in the quantity of representable and continuous functional when $p \geq 1$ gets bigger. For $p \geq 2$, the L^p-spaces turn out to be *fully representable* and *-*semisimple* Banach quasi *-algebras and in this case these notions *coincide* (see Sect. 2, [4, 7, 8, 10, 15]).

Hilbert quasi *-algebras constitute a class of *-semisimple and fully representable quasi *-algebras. In this case, representable and continuous functionals are in 1-1 correspondence with weakly positive and bounded elements. This result allows us to get the representability for the tensor product of two representable and continuous functionals in the framework of tensor product of Hilbert quasi *-algebras (refer to [2, 3]).

*-Semisimplicity allows us to define a proper notion of *-derivation in the case of Banach quasi *-algebras and prove a result characterizing infinitesimal generators of one-parameter group of *-automorphisms in the Banach quasi *-algebras case, extending results of Bratteli–Robinson for C*-algebras (see, e.g., [11]).

The survey is structured as follows. Firstly, we give some preliminaries about (Banach) quasi *-algebras, representable (and continuous) functionals and the GNS-construction. For these quasi *-algebras, in Sect. 2 we recall the notions of *-semisimplicity and full representability, summing up the main results about their link in the Banach case and the characterization we have for Hilbert quasi *-algebras. In Sect. 3, we concentrate on the case of *-semisimple Banach quasi *-algebras and deal with weak *-derivations and related results. In the last section, we look at the construction of the tensor product Hilbert quasi *-algebra and look at the

properties of tensor products of representable functionals and the sesquilinear forms involved in the definition of *-semisimplicity.

For the reader's convenience, we recall some preliminary notions for future use. Further details can be found in [7].

Definition 1.1 A *quasi *-algebra* $(\mathfrak{A}, \mathfrak{A}_0)$ (or over $\mathfrak{A}_0$) is a pair consisting of a vector space $\mathfrak{A}$ and a *-algebra $\mathfrak{A}_0$ contained in $\mathfrak{A}$ as a subspace and such that

(i) the left multiplication ax and the right multiplication xa of an element $a \in \mathfrak{A}$ and $x \in \mathfrak{A}_0$ are always defined and bilinear;
(ii) $(xa)y = x(ay)$ and $a(xy) = (ax)y$ for each $x, y \in \mathfrak{A}_0$ and $a \in \mathfrak{A}$;
(iii) an involution $*$, which extends the involution of $\mathfrak{A}_0$, is defined in $\mathfrak{A}$ with the property $(ax)^* = x^*a^*$ and $(xa)^* = a^*x^*$ for all $a \in \mathfrak{A}$ and $x \in \mathfrak{A}_0$.

We say that a quasi *-algebra $(\mathfrak{A}, \mathfrak{A}_0)$ has a *unit*, if there is an element in $\mathfrak{A}_0$, denoted by $\mathbb{1}$, such that $a\mathbb{1} = a = \mathbb{1}a$ for all $a \in \mathfrak{A}$. If a unit exists, then it is always unique.

Definition 1.2 Let $(\mathfrak{A}, \mathfrak{A}_0)$ be a quasi *-algebra. A linear functional $\omega : \mathfrak{A} \to \mathbb{C}$ satisfying

(L.1) $\omega(x^*x) \geq 0$ for all $x \in \mathfrak{A}_0$;
(L.2) $\omega(y^*a^*x) = \overline{\omega(x^*ay)}$ for all $x, y \in \mathfrak{A}_0, a \in \mathfrak{A}$;
(L.3) for all $a \in \mathfrak{A}$, there exists $\gamma_a > 0$ such that

$$|\omega(a^*x)| \leq \gamma_a \omega(x^*x)^{\frac{1}{2}}, \quad \forall x \in \mathfrak{A}_0,$$

is called *representable* on the quasi *-algebra $(\mathfrak{A}, \mathfrak{A}_0)$.

The family of all representable functionals on the quasi *-algebra $(\mathfrak{A}, \mathfrak{A}_0)$ will be denoted by $\mathcal{R}(\mathfrak{A}, \mathfrak{A}_0)$.

This definition is justified by the following Theorem 1.4, proving the existence of a GNS-like construction of a *-representation π_ω and a Hilbert space $\mathcal{H}_\omega$ for a representable functional ω on $\mathfrak{A}$.

Let $\mathcal{H}$ be a Hilbert space and let $\mathcal{D}$ be a dense linear subspace of $\mathcal{H}$. We denote by $\mathcal{L}^\dagger(\mathcal{D}, \mathcal{H})$ the following family of closable operators:

$$\mathcal{L}^\dagger(\mathcal{D}, \mathcal{H}) = \left\{ X : \mathcal{D} \to \mathcal{H} : \mathcal{D}(X) = \mathcal{D}, \mathcal{D}(X^*) \supset \mathcal{D} \right\}.$$

$\mathcal{L}^\dagger(\mathcal{D}, \mathcal{H})$ is a $\mathbb{C}$-vector space with the usual sum and scalar multiplication. If we define the involution $\dagger$ and partial multiplication $\Box$ as

$$X \mapsto X^\dagger \equiv X^* \restriction_{\mathcal{D}} \quad \text{and} \quad X \Box Y = X^\dagger * Y,$$

then $\mathcal{L}^\dagger(\mathcal{D}, \mathcal{H})$ is a partial *-algebra defined in [7].

Definition 1.3 A **-representation* of a quasi *-algebra $(\mathfrak{A}, \mathfrak{A}_0)$ is a *-homomorphism $\pi : \mathfrak{A} \to \mathcal{L}^\dagger(\mathcal{D}_\pi, \mathcal{H}_\pi)$, where $\mathcal{D}_\pi$ is a dense subspace of the Hilbert space $\mathcal{H}_\pi$, with the following properties:

(i) $\pi(a^*) = \pi(a)^\dagger$ for all $a \in \mathfrak{A}$;
(ii) if $a \in \mathfrak{A}$ and $x \in \mathfrak{A}_0$, then $\pi(a)$ is a left multiplier of $\pi(x)$ and

$$\pi(a)\Box\pi(x) = \pi(ax).$$

A *-representation π is

- *cyclic* if $\pi(\mathfrak{A}_0)\xi$ is dense in $\mathcal{H}_\pi$ for some $\xi \in \mathcal{D}_\pi$;
- *closed* if π coincides with its closure $\widetilde{\pi}$ defined in [30, Section 2].

If $(\mathfrak{A}, \mathfrak{A}_0)$ has a unit $\mathbb{1}$, then we suppose that $\pi(\mathbb{1}) = I_\mathcal{D}$, the identity operator of $\mathcal{D}$.

Theorem 1.4 ([30]) *Let $(\mathfrak{A}, \mathfrak{A}_0)$ be a quasi *-algebra with unit $\mathbb{1}$ and let ω be a linear functional on $(\mathfrak{A}, \mathfrak{A}_0)$ that satisfies the conditions* (L.1)–(L.3) *of Definition* 1.2. *Then, there exists a closed cyclic *-representation π_ω of $(\mathfrak{A}, \mathfrak{A}_0)$, with cyclic vector ξ_ω such that*

$$\omega(a) = \langle \pi_\omega(a)\xi_\omega | \xi_\omega \rangle\,, \quad \forall a \in \mathfrak{A}.$$

This representation is unique up to unitary equivalence.

1.1 Normed Quasi *-Algebras

Definition 1.5 A quasi *-algebra $(\mathfrak{A}, \mathfrak{A}_0)$ is called a *normed quasi *-algebra* if a norm $\|\cdot\|$ is defined on $\mathfrak{A}$ with the properties

(i) $\|a^*\| = \|a\|, \quad \forall a \in \mathfrak{A}$;
(ii) $\mathfrak{A}_0$ is dense in $\mathfrak{A}$;
(iii) for every $x \in \mathfrak{A}_0$, the map $R_x : a \in \mathfrak{A} \to ax \in \mathfrak{A}$ is continuous in $\mathfrak{A}$.

The continuity of the involution implies that

(iii') for every $x \in \mathfrak{A}_0$, the map $L_x : a \in \mathfrak{A} \to xa \in \mathfrak{A}$ is continuous in $\mathfrak{A}$.

Definition 1.6 If $(\mathfrak{A}, \|\cdot\|)$ is a Banach space, we say that $(\mathfrak{A}, \mathfrak{A}_0)$ is a *Banach quasi *-algebra*.

The norm topology of $\mathfrak{A}$ will be denoted by τ_n.

An important class of Banach quasi *-algebras is given by *Hilbert quasi *-algebras*.

Definition 1.7 Let $\mathfrak{A}_0$ be a *-algebra which is also a pre-Hilbert space with respect to the inner product $\langle\cdot|\cdot\rangle$ such that:

(1) the map $y \mapsto xy$ is continuous with respect to the norm defined by the inner product;
(2) $\langle xy|z\rangle = \langle y|x^*z\rangle$ for all $x, y, z \in \mathfrak{A}_0$;
(3) $\langle x|y\rangle = \langle y^*|x^*\rangle$ for all $x, y \in \mathfrak{A}_0$;
(4) $\mathfrak{A}_0^2$ is total in $\mathfrak{A}_0$.

Such a *-algebra $\mathfrak{A}_0$ is said to be a *Hilbert algebra*. If $\mathcal{H}$ denotes the Hilbert space completion of $\mathfrak{A}_0$ with respect to the inner product $\langle\cdot|\cdot\rangle$, then $(\mathcal{H}, \mathfrak{A}_0)$ is called a *Hilbert quasi *-algebra*.

2 Analogy Between *-Semisimplicity and Full Representability

Representable functionals constitute a valid tool for investigating structural properties of Banach quasi *-algebras. Indeed, the notions of *-semisimplicity and full representability are strongly related to these functionals. They turn out to be the same notion when dealing with Banach quasi *-algebras that verify the condition (P). For further reading, see [4, 7, 8, 15, 30].

If $\omega \in \mathcal{R}(\mathfrak{A}, \mathfrak{A}_0)$, then we can associate with it the sesquilinear form φ_ω defined on $\mathfrak{A}_0 \times \mathfrak{A}_0$ as

$$\varphi_\omega(x, y) := \omega(y^*x), \quad x, y \in \mathfrak{A}_0. \tag{2.1}$$

If $(\mathfrak{A}[\tau_n], \mathfrak{A}_0)$ is a normed quasi *-algebra, we denote by $\mathcal{R}_c(\mathfrak{A}, \mathfrak{A}_0)$ the subset of $\mathcal{R}(\mathfrak{A}, \mathfrak{A}_0)$ consisting of continuous functionals.

Let $\omega \in \mathcal{R}_c(\mathfrak{A}, \mathfrak{A}_0)$. Then ω is continuous on $\mathfrak{A}$, but φ_ω is not necessarily continuous on $\mathfrak{A}_0 \times \mathfrak{A}_0$. φ_ω is said to be *closable* if for every sequence of elements $\{x_n\}$ in $\mathfrak{A}_0$ such that

$$x_n \xrightarrow{\tau_n} 0 \quad \text{and} \quad \varphi_\omega(x_n - x_m, x_n - x_m) \to 0 \quad \text{as} \quad n \to \infty \tag{2.2}$$

then $\varphi_\omega(x_n, x_n) \to 0$ as $n \to \infty$.

By the condition (2.2), the closure of φ_ω, denoted by $\overline{\varphi}_\omega$, is a well-defined sesquilinear form on $\mathcal{D}(\overline{\varphi}_\omega) \times \mathcal{D}(\overline{\varphi}_\omega)$ as

$$\overline{\varphi}_\omega(a, a) := \lim_{n\to\infty} \varphi_\omega(x_n, x_n),$$

where $\mathcal{D}(\overline{\varphi}_\omega)$ is the following dense domain

$$\mathcal{D}(\overline{\varphi}_\omega) = \{a \in \mathfrak{A} : \exists\{x_n\} \subset \mathfrak{A}_0 \text{ s.t. } x_n \xrightarrow{\tau_n} a \text{ and } \varphi_\omega(x_n - x_m, x_n - x_m) \to 0\}.$$

For a locally convex quasi *-algebra $(\mathfrak{A}, \mathfrak{A}_0)$, $\overline{\varphi}_\omega$ always exists, [15]. Nevertheless, it is unclear whether $\mathcal{D}(\overline{\varphi}_\omega)$ is the whole space $\mathfrak{A}$. We show in Proposition 2.1 below that $\mathcal{D}(\overline{\varphi}_\omega)$ is $\mathfrak{A}$ in the case of a Banach quasi *-algebra.

Proposition 2.1 ([4]) *Let $(\mathfrak{A}, \mathfrak{A}_0)$ be a Banach quasi *-algebra with unit $\mathbb{1}$, $\omega \in \mathcal{R}_c(\mathfrak{A}, \mathfrak{A}_0)$ and φ_ω the associated sesquilinear form on $\mathfrak{A}_0 \times \mathfrak{A}_0$ defined as in (2.1). Then $\mathcal{D}(\overline{\varphi}_\omega) = \mathfrak{A}$; hence $\overline{\varphi}_\omega$ is everywhere defined and bounded.*

Set now

$$\mathfrak{A}_0^+ := \left\{ \sum_{k=1}^{n} x_k^* x_k,\ x_k \in \mathfrak{A}_0,\ n \in \mathbb{N} \right\}.$$

Then $\mathfrak{A}_0^+$ is a wedge in $\mathfrak{A}_0$ and we call the elements of $\mathfrak{A}_0^+$ *positive elements of* $\mathfrak{A}_0$. As in [15], we call positive elements of $\mathfrak{A}$ the elements of $\overline{\mathfrak{A}_0^+}^{\tau_n}$. We set $\mathfrak{A}^+ := \overline{\mathfrak{A}_0^+}^{\tau_n}$.

Definition 2.2 A linear functional on $\mathfrak{A}$ is *positive* if $\omega(a) \geq 0$ for every $a \in \mathfrak{A}^+$. A family of positive linear functionals $\mathcal{F}$ on $(\mathfrak{A}, \mathfrak{A}_0)$ is called *sufficient* if for every $a \in \mathfrak{A}^+$, $a \neq 0$, there exists $\omega \in \mathcal{F}$ such that $\omega(a) > 0$.

Definition 2.3 A normed quasi *-algebra $(\mathfrak{A}, \mathfrak{A}_0)$ is called *fully representable* if $\mathcal{R}_c(\mathfrak{A}, \mathfrak{A}_0)$ is sufficient and $\mathcal{D}(\overline{\varphi}_\omega) = \mathfrak{A}$ for every ω in $\mathcal{R}_c(\mathfrak{A}, \mathfrak{A}_0)$.

We denote by $\mathcal{Q}_{\mathfrak{A}_0}(\mathfrak{A})$ the family of all sesquilinear forms $\Omega : \mathfrak{A} \times \mathfrak{A} \to \mathbb{C}$ such that

(i) $\Omega(a, a) \geq 0$ for every $a \in \mathfrak{A}$;
(ii) $\Omega(ax, y) = \Omega(x, a^* y)$ for every $a \in \mathfrak{A}$, $x, y \in \mathfrak{A}_0$;

We denote by $\mathcal{S}_{\mathfrak{A}_0}(\mathfrak{A})$ the subset of $\mathcal{Q}_{\mathfrak{A}_0}(\mathfrak{A})$ consisting of all continuous sesquilinear forms having the property that also

(iii) $|\Omega(a, b)| \leq \|a\| \|b\|$, for all $a, b \in \mathfrak{A}$.

Definition 2.4 A normed quasi *-algebra $(\mathfrak{A}, \mathfrak{A}_0)$ is called **-semisimple* if, for every $0 \neq a \in \mathfrak{A}$, there exists $\Omega \in \mathcal{S}_{\mathfrak{A}_0}(\mathfrak{A})$ such that $\Omega(a, a) > 0$.

Proposition 2.1 is useful to show the following result, clarifying the link between *-semisimplicity and full representability. We need first to introduce the following condition of positivity

$$a \in \mathfrak{A} \text{ and } \omega(x^* a x) \geq 0 \ \ \forall \omega \in \mathcal{R}_c(\mathfrak{A}, \mathfrak{A}_0) \text{ and } x \in \mathfrak{A}_0 \ \ \Rightarrow \ \ a \in \mathfrak{A}^+. \qquad (P)$$

Theorem 2.5 ([4]) *Let $(\mathfrak{A}, \mathfrak{A}_0)$ be a Banach quasi *-algebra with unit $\mathbb{1}$. The following statements are equivalent.*

(i) $\mathcal{R}_c(\mathfrak{A}, \mathfrak{A}_0)$ *is sufficient.*
(ii) $(\mathfrak{A}, \mathfrak{A}_0)$ *is fully representable.*

If the condition (P) *holds,* (i) *and* (ii) *are equivalent to the following*

(iii) $(\mathfrak{A}, \mathfrak{A}_0)$ *is* **-semisimple.*

The condition (P) is not needed to show (iii) $\Rightarrow$ (ii) of Theorem 2.5.

Theorem 2.5 shows the deep connection between full representability and *-semisimplicity for a Banach quasi *-algebra. Under the condition of positivity (P), the families of sesquilinear forms involved in the definitions of full representability and *-semisimplicity can be identified.

For a Hilbert quasi *-algebra $(\mathcal{H}, \mathfrak{A}_0)$, representable and continuous functionals are in 1-1 correspondence with a certain family of elements in $\mathcal{H}$.

Definition 2.6 Let $(\mathcal{H}, \mathfrak{A}_0)$ be a Hilbert quasi *-algebra. An element $\xi \in \mathcal{H}$ is called

(i) *weakly positive* if the operator $L_\xi : \mathfrak{A}_0 \to \mathcal{H}$ defined as $L_\xi(x) := \xi x$ is positive.
(ii) *bounded* if the operator $L_\xi : \mathfrak{A}_0 \to \mathcal{H}$ is bounded.

The set of all weakly positive (resp. bounded) elements will be denoted as $\mathcal{H}_w^+$ (resp. $\mathcal{H}_b$), see [4, 29].

Theorem 2.7 ([4]) *Suppose that* $(\mathcal{H}, \mathfrak{A}_0)$ *is a Hilbert quasi* *-*algebra. Then* $\omega \in \mathcal{R}_c(\mathcal{H}, \mathfrak{A}_0)$ *if, and only if, there exists a unique weakly positive bounded element* $\eta \in \mathcal{H}$ *such that*

$$\omega(\xi) = \langle \xi | \eta \rangle, \quad \forall \xi \in \mathcal{H}.$$

2.1 Case of $L^p(I, d\lambda)$ for $p \geq 1$

Consider $\mathfrak{A}_0$ to be $L^\infty(I, d\lambda)$, where I is a compact interval of the real line and λ is the Lebesgue measure. Let τ_n be the topology generated by the p-norm

$$\|f\|_p := \left(\int_I |f|^p d\lambda \right)^{\frac{1}{p}}, \quad \forall f \in L^\infty(I, d\lambda),$$

for $p \geq 1$. Then, the completion of $L^\infty(I, d\lambda)$ with respect to the $\|\cdot\|_p$-norm is given by $L^p(I, d\lambda)$.

We conclude that, for $p \geq 1$, $(L^p(I, d\lambda), L^\infty(I, d\lambda))$ is a Banach quasi *-algebra. The same conclusion holds if we consider the *-algebra of continuous functions over $I = [0, 1]$, denoted by $\mathcal{C}(I)$.

For $p \geq 2$, $(L^p(I, d\lambda), L^\infty(I, d\lambda))$ and $(L^p(I, d\lambda), \mathcal{C}(I))$ are fully representable and *-semisimple Banach quasi *-algebras. For $1 \leq p < 2$, we have $\mathcal{R}_c(L^p(I, d\lambda), \mathfrak{A}_0) = \{0\}$ for both $\mathfrak{A}_0 = L^\infty(I, d\lambda)$ or $\mathfrak{A}_0 = \mathcal{C}(I)$. Note that the Banach quasi *-algebras $(L^p(I, d\lambda), L^\infty(I, d\lambda))$ and $(L^p(I, d\lambda), \mathcal{C}(I))$ verify the condition (P) for all $p \geq 1$.

The absence of representable and continuous functionals that we observe passing the threshold $p = 2$ and the lack in the literature of an example of representable functional that is not continuous led us to pose the following question.

Question Is every representable functional on a Banach quasi *-algebra continuous?

The answer to this question is positive in the case of the Hilbert quasi *-algebras $(L^2(I, d\lambda), L^\infty(I, d\lambda))$ and $(L^2(I, d\lambda), \mathcal{C}(I))$, see [4].

Theorem 2.8 ([4]) *Let ω be a representable functional on the Hilbert quasi *-algebra $(L^2(I, d\lambda), L^\infty(I, d\lambda))$. Then there exists a unique bounded finitely additive measure ν on I which vanish on subsets of I of zero λ-measure and a unique bounded linear operator $S : L^2(I, d\lambda) \to L^2(I, d\nu)$ such that*

$$\omega(f) = \int_I (Sf) d\nu, \quad \forall f \in L^2(I, d\lambda).$$

The operator S satisfies the following conditions:

$$S(f\phi) = (Sf)\phi = \phi(Sf) \quad \forall f \in L^2(I, d\lambda), \phi \in L^\infty(I, d\lambda);$$
$$S\phi = \phi, \quad \forall \phi \in L^\infty(I, d\lambda).$$

Thus, every representable functional ω on $\left(L^2(I, d\lambda), L^\infty(I, d\lambda)\right)$ is continuous.

Theorem 2.9 ([4]) *Let ω be a representable functional on the Hilbert quasi *-algebra $(L^2(I, d\lambda), \mathcal{C}(I))$. Then there exists a unique Borel measure μ on I and a unique bounded linear operator $T : L^2(I, d\lambda) \to L^2(I, d\mu)$ such that*

$$\omega(f) = \int_I (Tf) d\mu, \quad \forall f \in L^2(I, d\lambda).$$

The operator T satisfies the following conditions:

$$T(f\phi) = (Tf)\phi = \phi(Tf) \quad \forall f \in L^2(I, d\lambda), \phi \in \mathcal{C}(I);$$
$$T\phi = \phi, \quad \forall \phi \in \mathcal{C}(I).$$

Thus, every representable functional ω on $\left(L^2(I, d\lambda), \mathcal{C}(I)\right)$ is continuous. Moreover, μ is absolutely continuous with respect to λ.

The above theorems can be extended to the case of a commutative Hilbert quasi *-algebra, under certain hypotheses.

Theorem 2.10 ([4]) *Let $(\mathcal{H}, \mathfrak{A}_0)$ be a commutative Hilbert quasi *-algebra with unit $\mathbb{1}$. Assume that $\mathfrak{A}_0[\|\cdot\|_0]$ is a Banach *-algebra and that there exists an element x of $\mathfrak{A}_0$ such that the spectrum $\sigma(\overline{R}_x)$ of the bounded operator $\overline{R}_x$ of right*

multiplication by x consists only of its continuous part $\sigma_c(\overline{R}_x)$. If ω is representable on $(\mathfrak{A}, \mathfrak{A}_0)$, then ω is bounded.

The main tools to get these results are the Riesz-Markov Representation Theorem for continuous functions on a compact space and the intertwining theory on Hilbert spaces (see [26]). Unfortunately, these tools turn out to be unsuitable to the case of a (non-commutative) Hilbert quasi *-algebra or a Banach quasi *-algebra.

3 First Application: Derivations and Their Closability

-Derivations have been widely employed to describe the dynamics for a quantum phenomenon. For a quantum system of finite volume V, the Hamiltonian belongs to the local C-algebra and implements the inner *-derivation for the dynamics. Nevertheless, the thermodynamical limit in general fails to exist, see [7].

Under some assumptions, the limit turns out to be a *weak *-derivation* generating a one parameter group of *weak *-automorphisms*, defined for a *-semisimple Banach quasi *-algebra, as we will investigate in the following section. For detailed discussion, see [5, 6, 20].

The derivation $\delta : \mathfrak{A}_0[\|\cdot\|] \to \mathfrak{A}[\|\cdot\|]$ is densely defined. If δ is closable, then its closure $\overline{\delta}$ as a linear map is *not* a derivation in general.

- In this section, we only consider *-semisimple Banach quasi *-algebras, if not otherwise specified.

For these Banach quasi *-algebras, define a weaker multiplication in $\mathfrak{A}$ as in [29].

Definition 3.1 Let $(\mathfrak{A}, \mathfrak{A}_0)$ be a Banach quasi *-algebra. Let $a, b \in \mathfrak{A}$. We say that the *weak multiplication* $a\Box b$ is well-defined if there exists a (necessarily unique) $c \in \mathfrak{A}$ such that:

$$\varphi(bx, a^*y) = \varphi(cx, y), \quad \forall\, x, y \in \mathfrak{A}_0, \quad \forall\, \varphi \in \mathcal{S}_{\mathfrak{A}_0}(\mathfrak{A}).$$

In this case, we put $a\Box b := c$.

Let $a \in \mathfrak{A}$. The space of right (resp. left) weak multipliers of a, i.e., the space of all the elements $b \in \mathfrak{A}$ such that $a\Box b$ (resp. $b\Box a$) is well-defined, will be denoted as $R_w(a)$ (resp. $L_w(a)$). We indicate by $R_w(\mathfrak{A})$ (resp. $L_w(\mathfrak{A})$) the space of universal right (resp. left) multipliers of $\mathfrak{A}$, i.e., all the elements $b \in \mathfrak{A}$ such that $b \in R_w(a)$ (resp. $b \in L_w(a)$) for every $a \in \mathfrak{A}$. Clearly, $\mathfrak{A}_0 \subseteq R_w(\mathfrak{A}) \cap L_w(\mathfrak{A})$.

Let $(\mathfrak{A}, \mathfrak{A}_0)$ be a Banach quasi *-algebra. For every $a \in \mathfrak{A}$, two linear operators L_a and R_a are defined in the following way

$$L_a : \mathfrak{A}_0 \to \mathfrak{A} \quad L_a(x) = ax \quad \forall x \in \mathfrak{A}_0 \tag{3.1}$$

$$R_a : \mathfrak{A}_0 \to \mathfrak{A} \quad R_a(x) = xa \quad \forall x \in \mathfrak{A}_0. \tag{3.2}$$

An element $a \in \mathfrak{A}$ is called *bounded* if the operators L_a and R_a defined in (3.1) and (3.2) are $\|\cdot\|$-continuous and thus extendible to the whole space $\mathfrak{A}$. As for Hilbert quasi *-algebras in Definition 2.6, the set of all bounded elements will be denoted by $\mathfrak{A}_b$.

Lemma 3.2 ([5, Lemma 2.16]) *If $(\mathfrak{A}, \mathfrak{A}_0)$ is a *-semisimple Banach quasi *-algebra with unit $\mathbb{1}$, the set $\mathfrak{A}_b$ of bounded elements coincides with the set $R_w(\mathfrak{A}) \cap L_w(\mathfrak{A})$.*

Let $\theta : \mathfrak{A} \to \mathfrak{A}$ be a linear bijection. We say that θ is a *weak *-automorphism of* $(\mathfrak{A}, \mathfrak{A}_0)$ if

(i) $\theta(a^*) = \theta(a)^*$, for every $a \in \mathfrak{A}$;
(ii) $\theta(a)\Box\theta(b)$ is well defined if, and only if, $a\Box b$ is well defined and, in this case,

$$\theta(a\Box b) = \theta(a)\Box\theta(b).$$

Definition 3.3 Let β_t be a weak *-automorphism of $\mathfrak{A}$ for every $t \in \mathbb{R}$. If

(i) $\beta_0(a) = a$, $\forall a \in \mathfrak{A}$
(ii) $\beta_{t+s}(a) = \beta_t(\beta_s(a))$, $\forall a \in \mathfrak{A}$

then we say that β_t is a *one-parameter group of weak *-automorphisms of* $(\mathfrak{A}, \mathfrak{A}_0)$. If τ is a topology on $\mathfrak{A}$ and the map $t \mapsto \beta_t(a)$ is τ-continuous, for every $a \in \mathfrak{A}$, we say that β_t is a *τ-continuous* weak *-automorphism group.

In this case

$$\mathcal{D}(\delta_\tau) = \left\{ a \in \mathfrak{A} : \lim_{t\to 0} \frac{\beta_t(a) - a}{t} \text{ exists in } \mathfrak{A}[\tau] \right\}$$

and

$$\delta_\tau(a) = \tau - \lim_{t\to 0} \frac{\beta_t(a) - a}{t}, \quad a \in \mathcal{D}(\delta_\tau).$$

Then δ_τ is called the *infinitesimal generator* of the one-parameter group $\{\beta_t\}$ of weak *-automorphisms of $(\mathfrak{A}, \mathfrak{A}_0)$.

Definition 3.4 ([5]) Let $(\mathfrak{A}, \mathfrak{A}_0)$ be a *-semisimple Banach quasi *-algebra and δ a linear map of $\mathcal{D}(\delta)$ into $\mathfrak{A}$, where $\mathcal{D}(\delta)$ is a partial *-algebra with respect to the weak multiplication $\Box$. We say that δ is a *weak *-derivation* of $(\mathfrak{A}, \mathfrak{A}_0)$ if

(i) $\mathfrak{A}_0 \subset \mathcal{D}(\delta)$
(ii) $\delta(x^*) = \delta(x)^*$, $\forall x \in \mathfrak{A}_0$
(iii) if $a, b \in \mathcal{D}(\delta)$ and $a\Box b$ is well defined, then $a\Box b \in \mathcal{D}(\delta)$ and

$$\varphi(\delta(a\Box b)x, y) = \varphi(bx, \delta(a)^* y) + \varphi(\delta(b)x, a^* y),$$

for all $\varphi \in \mathcal{S}_{\mathfrak{A}_0}(\mathfrak{A})$, for every $x, y \in \mathfrak{A}_0$.

Analogously to the case of a C*-algebra, to a *uniformly bounded* norm continuous weak *-automorphisms group there corresponds a closed weak *-derivation that generates the group (see [11]).

Theorem 3.5 ([5]) *Let $\delta : \mathcal{D}(\delta) \to \mathfrak{A}[\|\cdot\|]$ be a weak *-derivation on a *-semisimple Banach quasi *-algebra $(\mathfrak{A}, \mathfrak{A}_0)$. Suppose that δ is the infinitesimal generator of a uniformly bounded, τ_n-continuous group of weak *-automorphisms of $(\mathfrak{A}, \mathfrak{A}_0)$. Then δ is closed; its resolvent set $\rho(\delta)$ contains $\mathbb{R} \setminus \{0\}$ and*

$$\|\delta(a) - \lambda a\| \geq |\lambda| \, \|a\|, \quad a \in \mathcal{D}(\delta), \lambda \in \mathbb{R}.$$

Theorem 3.6 ([5]) *Let $\delta : \mathcal{D}(\delta) \subset \mathfrak{A}_b \to \mathfrak{A}[\|\cdot\|]$ be a closed weak *-derivation on a *-semisimple Banach quasi *-algebra $(\mathfrak{A}, \mathfrak{A}_0)$. Suppose that δ verifies the same conditions on its spectrum of Theorem 3.5 and $\mathfrak{A}_0$ is a core for every multiplication operator $\hat{L}_a$ for $a \in \mathfrak{A}$, i.e. $\hat{L}_a = \overline{L_a}$. Then δ is the infinitesimal generator of a uniformly bounded, τ_n-continuous group of weak *-automorphisms of $(\mathfrak{A}, \mathfrak{A}_0)$.*

In Theorem 3.6, we assumed further conditions, for instance that the domain $\mathcal{D}(\delta)$ of the weak *-derivation is contained in $\mathfrak{A}_b$, which turns out to be satisfied in some interesting situations such as the weak derivative in L^p-spaces.

*Example 3.7 (Inner Weak *-Derivations for Unbounded Hamiltonian h)* Let $(\mathfrak{A}, \mathfrak{A}_0)$ be a *-semisimple Banach quasi *-algebra. Let $h \in \mathfrak{A}$ be a self-adjoint unbounded element, i.e., $h = h^*$ and $\sigma(h) \subset \mathbb{R}$. We define the following derivation

$$\delta_h : \mathfrak{A}_0 \to \mathfrak{A}, \quad \text{defined as} \quad \delta_h(x) = i[h, x], \quad x \in \mathfrak{A}_0.$$

Then, for every fixed $t \in \mathbb{R}$, $\beta_t(a) = e^{ith} \square a \square e^{-ith}$ is a well-defined weak *-automorphism of $(\mathfrak{A}, \mathfrak{A}_0)$ since e^{ith}, e^{-ith} are bounded elements in $\mathfrak{A}$ (see [5]) and thus $(e^{ith} \square a) \square e^{-ith} = e^{ith} \square (a \square e^{-ith})$ for every $a \in \mathfrak{A}$. Moreover, β_t is a uniformly bounded norm continuous group of weak *-automorphisms. The infinitesimal generator is given by

$$\overline{\delta}_h(a) := \lim_{t \to 0} \frac{\beta_t(a) - a}{t} = \lim_{t \to 0} \frac{e^{ith} \square a \square e^{-ith} - a}{t} = i[h \square a - a \square h],$$

when *a is bounded*.

In the following Proposition, the *-semisimplicity is automatically given by assuming the existence of a representable and continuous functional with associated faithful *-representation (see [1, 11]).

Proposition 3.8 ([1]) *Let $(\mathfrak{A}, \mathfrak{A}_0)$ be a Banach quasi *-algebra with unit $\mathbb{1}$ and let δ be a weak *-derivation of $(\mathfrak{A}, \mathfrak{A}_0)$ such that $\mathcal{D}(\delta) = \mathfrak{A}_0$. Suppose that there exists a representable and continuous functional ω with $\omega(\delta(x)) = 0$ for $x \in \mathfrak{A}_0$ and let $(\mathcal{H}_\omega, \pi_\omega, \lambda_\omega)$ be the GNS-construction associated with ω. Suppose that π_ω*

*is a faithful *-representation of* $(\mathfrak{A}, \mathfrak{A}_0)$. *Then there exists an element* $H = H^\dagger$ *of* $\mathcal{L}^\dagger(\lambda_\omega(\mathfrak{A}_0))$ *such that*

$$\pi_\omega(\delta(x)) = -i[H, \pi_\omega(x)], \quad \forall x \in \mathfrak{A}_0$$

and δ *is closable.*

4 Second Application: Tensor Product of Hilbert Quasi *-Algebras

In this section, we recall the construction of the tensor product Hilbert quasi *-algebra of two given Hilbert quasi *-algebras $(\mathcal{H}_1, \mathfrak{A}_0)$ and $(\mathcal{H}_2, \mathfrak{B}_0)$, giving some results about the relationship between the representability properties for the tensor product and those for the factors. For further reading on the algebraic and topological tensor product, see [12–14, 19, 23, 24], for the tensor product Hilbert quasi *-algebras refer to [2, 3].

For convenience, we assume that $(\mathcal{H}_1, \mathfrak{A}_0)$ and $(\mathcal{H}_2, \mathfrak{B}_0)$ are unital Hilbert quasi *-algebras.

The algebraic tensor product $\mathfrak{A}_0 \otimes \mathfrak{B}_0$ is the tensor product *-algebra of $\mathfrak{A}_0$ and $\mathfrak{B}_0$, it is endowed with the canonical multiplication and involution and it is considered as a subspace of the tensor product vector space $\mathcal{H}_1 \otimes \mathcal{H}_2$.

If $\mathcal{H}_1$ and $\mathcal{H}_2$ are endowed with the inner product $\langle\cdot|\cdot\rangle_1$ and $\langle\cdot|\cdot\rangle_2$ respectively, then $\mathfrak{A}_0 \otimes \mathfrak{B}_0$ satisfies the requirements of Definition 1.7 if we endow it with the following well-defined inner product

$$\langle z|z'\rangle_h := \sum_{i=1}^{n}\sum_{j=1}^{m}\left\langle x_i|x'_j\right\rangle_1\left\langle y_i|y'_j\right\rangle_2, \quad \forall z, z' \in \mathfrak{A}_0 \otimes \mathfrak{B}_0, \tag{4.1}$$

where $z = \sum_{i=1}^{n} x_i \otimes y_i$ and $z' = \sum_{j=1}^{m} x'_j \otimes y'_j$ (see [24, 25, 27]). Then, the completion of $\mathfrak{A}_0 \otimes \mathfrak{B}_0$ with respect to the norm $\|\cdot\|_h$ induced by the inner product in (4.1) is a Hilbert quasi *-algebra. Since $\mathfrak{A}_0$, $\mathfrak{B}_0$ are respectively dense in $\mathcal{H}_1$, $\mathcal{H}_2$ and $\|\cdot\|_h$ is a *cross-norm*, i.e. $\|x \otimes y\|_h = \|x\|_1\|y\|_2$ for all $x \otimes y \in \mathfrak{A}_0 \otimes \mathfrak{B}_0$, the tensor product *-algebra $\mathfrak{A}_0 \otimes \mathfrak{B}_0$ is $\|\cdot\|_h$-dense in $\mathcal{H}_1 \otimes_h \mathcal{H}_2$. We conclude that

$$\mathfrak{A}_0 \widehat{\otimes}^h \mathfrak{B}_0 \equiv \mathcal{H}_1 \widehat{\otimes}^h \mathcal{H}_2$$

and $(\mathcal{H}_1 \widehat{\otimes}^h \mathcal{H}_2, \mathfrak{A}_0 \otimes \mathfrak{B}_0)$ is a Hilbert quasi *-algebra, see [2, 3].

For the tensor product Hilbert quasi *-algebra $(\mathcal{H}_1 \widehat{\otimes}^h \mathcal{H}_2, \mathfrak{A}_0 \otimes \mathfrak{B}_0)$ we can apply all the known results for Banach quasi *-algebras about representability presented in Sect. 2, see also [4, 5]. In particular, we know that $(\mathcal{H}_1 \widehat{\otimes}^h \mathcal{H}_2, \mathfrak{A}_0 \otimes \mathfrak{B}_0)$ is always a *-

semisimple and fully representable Hilbert quasi *-algebra, applying Theorems 2.5 and 2.7.

Employing Theorem 2.7 and Lemma 4.1 in [2], we can give an alternative proof for Theorem 4.1. The proof will be given below after the proof of Theorem 4.2.

Theorem 4.1 ([2]) *Let $(\mathcal{H}_1\widehat{\otimes}^h\mathcal{H}_2, \mathfrak{A}_0 \otimes \mathfrak{B}_0)$ be the tensor product Hilbert quasi *-algebra of $(\mathcal{H}_1, \mathfrak{A}_0)$ and $(\mathcal{H}_2, \mathfrak{B}_0)$. Then, if ω_1, ω_2 are representable and continuous functionals on $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively, then $\omega_1 \otimes \omega_2$ extends to a representable and continuous functional Ω on the tensor product Hilbert quasi *-algebra $\mathcal{H}_1\widehat{\otimes}^h\mathcal{H}_2$.*

We now look at what happens to the sesquilinear forms in $\mathcal{S}_{\mathfrak{A}_0}(\mathcal{H}_1)$ and $\mathcal{S}_{\mathfrak{B}_0}(\mathcal{H}_2)$.

Theorem 4.2 *Let $(\mathcal{H}_1\widehat{\otimes}^h\mathcal{H}_2, \mathfrak{A}_0 \otimes \mathfrak{B}_0)$ be the tensor product Hilbert quasi *-algebra of $(\mathcal{H}_1, \mathfrak{A}_0)$ and $(\mathcal{H}_2, \mathfrak{B}_0)$. Let $\phi_1 \in \mathcal{S}_{\mathfrak{A}_0}(\mathcal{H}_1)$ and $\phi_2 \in \mathcal{S}_{\mathfrak{B}_0}(\mathcal{H}_2)$. Then, $\phi_1\widehat{\otimes}\phi_2 \in \mathcal{S}_{\mathfrak{A}_0\otimes\mathfrak{B}_0}(\mathcal{H}_1\widehat{\otimes}^h\mathcal{H}_2)$, i.e., it satisfies the conditions* (i), (ii) *and* (iii) *after the Definition 2.3.*

Proof Let $\phi_1 \in \mathcal{S}_{\mathfrak{A}_0}(\mathcal{H}_1)$ and $\phi_2 \in \mathcal{S}_{\mathfrak{B}_0}(\mathcal{H}_2)$. These sesquilinear forms are continuous, thus by the representation theorem for bounded sesquilinear forms over a Hilbert space, there exist unique bounded operators $T_1 : \mathcal{H}_1 \to \mathcal{H}_1$ and $T_2 : \mathcal{H}_2 \to \mathcal{H}_2$ such that

$$\phi_1(\xi, \xi') = \langle \xi | T_1 \xi' \rangle \quad \text{and} \quad \phi_2(\eta, \eta') = \langle \eta | T_2 \eta' \rangle,$$

for all $\xi, \xi' \in \mathcal{H}_1$, $\eta, \eta' \in \mathcal{H}_2$, and $\|T_i\| = \|\phi_i\| \leq 1$ for $i = 1, 2$. Moreover, T_1 and T_2 are positive operators on $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively.

On the pre-completion $\mathcal{H}_1 \otimes^h \mathcal{H}_2$, define the tensor product of ϕ_1 and ϕ_2 as

$$\phi_1 \otimes \phi_2(\zeta, \zeta') := \sum_{i,j=1}^{n} \phi_1(\xi_i, \xi_j')\phi_2(\eta_i, \eta_j')$$

for all $\zeta = \sum_{i=1}^n \xi_i \otimes \eta_i$, $\zeta' = \sum_{j=1}^n \xi_j' \otimes \eta_j'$ in $\mathcal{H}_1 \otimes^h \mathcal{H}_2$.

By [19], it is known that $T_1 \otimes T_2$ extends to a bounded operator $T_1\widehat{\otimes}T_2$ on the completion $\mathcal{H}_1\widehat{\otimes}^h\mathcal{H}_2$ such that $\|T_1\widehat{\otimes}T_2\| \leq \|T_1\|\|T_2\| \leq 1$.

We show that $\phi_1 \otimes \phi_2$ is represented by $T_1 \otimes T_2$ on $\mathcal{H}_1 \otimes^h \mathcal{H}_2$. Thus, $\phi_1 \otimes \phi_2$ is continuous and can be extended to $\mathcal{H}_1\widehat{\otimes}^h\mathcal{H}_2$. Its extension will be denoted by $\phi_1\widehat{\otimes}\phi_2$ and it corresponds to the bounded operator $T_1\widehat{\otimes}T_2$. Indeed, by the definition of $\phi_1 \otimes \phi_2$, we have

$$\begin{aligned}
\phi_1 \otimes \phi_2(\zeta, \zeta') &= \sum_{i,j=1}^{n} \phi_1(\xi_i, \xi_j')\phi_2(\eta_i, \eta_j') \\
&= \sum_{i,j=1}^{n} \left\langle \xi_i | T_1 \xi_j' \right\rangle \left\langle \eta_i, T_2 \eta_j' \right\rangle \\
&= \left\langle \zeta | (T_1 \otimes T_2)\zeta' \right\rangle.
\end{aligned}$$

To conclude the proof, we show that $\phi_1 \widehat{\otimes} \phi_2$ is in $\mathcal{S}_{\mathfrak{A}_0 \otimes \mathfrak{B}_0}(\mathcal{H}_1 \widehat{\otimes}^h \mathcal{H}_2)$. Indeed, $\phi_1 \widehat{\otimes} \phi_2$ is a positive sesquilinear form, since $T_1 \widehat{\otimes} T_2$ is a positive operator as a tensor product of positive operators on a Hilbert space. Thus, the condition (i) is verified.

The condition (ii) can be easily verified using the corresponding properties of ϕ_1 and ϕ_2. For (iii), we know that

$$\|\phi_1 \widehat{\otimes} \phi_2\| = \|T_1 \widehat{\otimes} T_2\| \leq \|T_1\| \|T_2\| \leq 1.$$

Hence, all the conditions for $\phi_1 \widehat{\otimes} \phi_2$ to belong to $\mathcal{S}_{\mathfrak{A}_0 \otimes \mathfrak{B}_0}(\mathcal{H}_1 \widehat{\otimes}^h \mathcal{H}_2)$ are verified. □

Using the same argument as in Theorem 2.5 (refer to [4] for a complete proof), the sesquilinear form $\overline{\varphi}_\omega$ associated with a representable and continuous functional ω in a Hilbert quasi *-algebra $(\mathcal{H}, \mathfrak{A}_0)$ with unit $\mathbb{1}$ is bounded and in $\mathcal{Q}_{\mathfrak{A}_0}(\mathcal{H})$. Hence, it is represented by a bounded operator S_ω such that

$$\omega(\xi) = \overline{\varphi}_\omega(\xi, \mathbb{1}) = \langle \xi | S_\omega \mathbb{1} \rangle, \quad \xi \in \mathcal{H}.$$

We want to show that $S_\omega \mathbb{1}$ is a weakly positive and bounded element. Indeed, for all $x \in \mathfrak{A}_0$, we have

$$\left\langle x | R_{S_\omega \mathbb{1}} x \right\rangle = \langle x | x(S_\omega \mathbb{1}) \rangle = \left\langle x^* x | S_\omega \mathbb{1} \right\rangle = \omega(x^* x) \geq 0,$$

thus $S_\omega \mathbb{1}$ is weakly positive. Moreover, $S_\omega \mathbb{1}$ is bounded by the condition (L.3) of representability in Definition 1.2. Indeed, (L.3) means that for every $\xi \in \mathcal{H}$, there exists $\gamma_\xi > 0$ such that

$$\left|\omega(\xi^* x)\right| \leq \gamma_\xi \omega(x^* x)^{\frac{1}{2}},$$

for all $x \in \mathfrak{A}_0$. By Proposition 2.1, $\overline{\varphi}_\omega$ is everywhere defined and bounded, hence

$$\left|\omega(\xi^* x)\right| = \left|\overline{\varphi}_\omega(x, \xi)\right| \leq c_\xi \|x\|,$$

for some positive constant c_ξ. Hence, by the Riesz representation theorem, there exists $\chi_\xi \in \mathcal{H}$ such that $\omega(\xi^* x) = \left\langle x | \chi_\xi \right\rangle$ for all $x \in \mathfrak{A}_0$. Therefore, the weak product (in the sense of [4, Definition 4.4]) $\xi \square S_\omega \mathbb{1}$ is well-defined for all $\xi \in \mathcal{H}$. Indeed, for $x, y \in \mathfrak{A}_0$, we have

$$\left\langle \xi^* x | S_\omega \mathbb{1} y \right\rangle = \left\langle \xi^* x y^* | S_\omega \mathbb{1} \right\rangle = \left\langle x y^* | \chi_\xi \right\rangle = \left\langle x | \chi_\xi y \right\rangle.$$

A similar argument shows that $S_\omega \mathbb{1} \square \xi$ is well-defined for all $\xi \in \mathcal{H}$. Recall that an element $\xi \in \mathcal{H}$ is bounded if, and only if, $R_w(\xi) = L_w(\xi) = \mathcal{H}$, where $R_w(\xi)$ (resp. $L_w(\xi)$) is the space of universal right (resp. left) weak multipliers of ξ (see [4, Proposition 4.10]). Then, $S_\omega \mathbb{1}$ is a bounded element.

By Theorem 2.7, we have that $S_\omega \mathbb{1}$ is the weakly positive and bounded element in $\mathcal{H}$ corresponding to the representable and continuous functional ω.

For what we just argued, we can give an alternative proof of Theorem 4.1. The proof tells us explicitly what the elements in $\mathcal{H}_1 \widehat{\otimes}^h \mathcal{H}_2$ are corresponding to representable and continuous functionals of the form $\omega_1 \widehat{\otimes} \omega_2$ for $\omega_1 \in \mathcal{R}_c(\mathcal{H}_1, \mathfrak{A}_0)$, $\omega_2 \in \mathcal{R}_c(\mathcal{H}_2, \mathfrak{B}_0)$.

Proof of Theorem 4.1 Let ω_1, ω_2 be representable and continuous functionals on $\mathcal{H}_1$, $\mathcal{H}_2$ respectively. Then, for what we discussed above, $\omega_1(\xi) = \langle \xi | S_{\omega_1} \mathbb{1}_{\mathcal{H}_1} \rangle$ for all $\xi \in \mathcal{H}_1$ and $\omega_2(\eta) = \langle \eta | S_{\omega_2} \mathbb{1}_{\mathcal{H}_2} \rangle$ for $\eta \in \mathcal{H}_2$.

Looking at their tensor product on $\mathcal{H}_1 \otimes^h \mathcal{H}_2$, we have

$$\omega_1 \otimes \omega(\zeta) = \sum_{i=1}^{n} \langle \xi_i \otimes \eta_i | S_{\omega_1} \mathbb{1}_{\mathcal{H}_1} \otimes S_{\omega_2} \mathbb{1}_{\mathcal{H}_2} \rangle,$$

for every $\zeta = \sum_{i=1}^n \xi_i \otimes \eta_i$ in $\mathcal{H}_1 \otimes^h \mathcal{H}_2$.

Since $\overline{\varphi}_{\omega_1}$ and $\overline{\varphi}_{\omega_2}$ are in $\mathcal{Q}_{\mathfrak{A}_0}(\mathcal{H}_1)$ and $\mathcal{Q}_{\mathfrak{B}_0}(\mathcal{H}_2)$ respectively, with the same argument of Theorem 4.2, $\overline{\varphi}_{\omega_1} \widehat{\otimes} \overline{\varphi}_{\omega_2}$ is continuous and belongs to $\mathcal{Q}_{\mathfrak{A}_0 \otimes \mathfrak{B}_0}(\mathcal{H}_1 \widehat{\otimes}^h \mathcal{H}_2)$. Furthermore, it is represented by $S_{\omega_1} \widehat{\otimes} S_{\omega_2}$. Hence

$$S_{\omega_1} \otimes S_{\omega_2}(\mathbb{1}_{\mathcal{H}_1} \otimes \mathbb{1}_{\mathcal{H}_2}) = S_{\omega_1} \mathbb{1}_{\mathcal{H}_1} \otimes S_{\omega_2} \mathbb{1}_{\mathcal{H}_2}$$

is weakly bounded and positive. Hence, by Theorem 2.7, $\omega_1 \otimes \omega_2$ is representable and continuous. Furthermore, the continuous extension of $\omega_1 \otimes \omega_2$ is representable. Indeed, let us denote this extension as Ω, then

$$\begin{aligned} \Omega(\psi) &= \lim_{n \to +\infty} \omega_1 \otimes \omega_2 (z_n) \\ &= \lim_{n \to +\infty} \langle z_n | S_{\omega_1} \otimes S_{\omega_2}(\mathbb{1}_{\mathcal{H}_1} \otimes \mathbb{1}_{\mathcal{H}_2}) \rangle \\ &= \langle \psi | S_{\omega_1} \otimes S_{\omega_2}(\mathbb{1}_{\mathcal{H}_1} \otimes \mathbb{1}_{\mathcal{H}_2}) \rangle, \end{aligned}$$

where $\{z_n\}$ is a sequence of elements in $\mathfrak{A}_0 \otimes \mathfrak{B}_0$ $\| \cdot \|_h$-converging to $\psi \in \mathcal{H}_1 \widehat{\otimes} \mathcal{H}_2$. □

It would be of interest to look at the same questions in the more general framework of Banach quasi *-algebras. This is work in progress with Maria Fragouloupoulou. In this case, we need further assumptions on the considered cross-norm to get some of the properties that we showed in this work, see [3].

Acknowledgments The author is grateful to the organizers of the International Workshop on Operator Theory and its Applications 2019, especially to the Organizers of the section entitled "Linear Operators and Function Spaces", for this interesting and delightful conference and the Instituto Superior Técnico of Lisbon for its hospitality. The author was financially supported by the ERC Advanced Grant no. 669240 QUEST "Quantum Algebraic Structures and Models".

The author wishes to thank the anonymous referees for their useful suggestions that improved the presentation of this manuscript.

References

1. M.S. Adamo, The interplay between representable functionals and derivations on Banach quasi *-algebras, in *Proceedings of the International Conference on Topological Algebras and Their Applications – ICTAA 2018*, ed. by M. Abel. Mathematics Studies (Tartu), 7 (Estonian Mathematical Society, Tartu, 2018), pp. 48–59
2. M.S. Adamo, About tensor product of Hilbert quasi *-algebras and their representability, (accepted for publication in the Proceedings of OT27, Theta 2020)
3. M.S. Adamo, M. Fragoulopoulou, Tensor products of normed and Banach quasi *-algebras. J. Math. Anal. Appl. **490**, 2 (2020)
4. M.S. Adamo, C. Trapani, Representable and continuous functionals on a Banach quasi *-algebra. Mediterr. J. Math. **14**, 157 (2017)
5. M.S. Adamo, C. Trapani, Unbounded derivations and *-automorphisms groups of Banach quasi *-algebras. Ann. Mat. Pura Appl. (4) **198**, 1711–1729 (2019)
6. J.-P. Antoine, A. Inoue, C. Trapani, O*-dynamical systems and *-derivations of unbounded operator algebras. Math. Nachr. **204**, 5–28 (1999)
7. J.-P. Antoine, A. Inoue, C. Trapani, *Partial *-Algebras and Their Operator Realizations* (Kluwer Academic, Dordrecht, 2003)
8. F. Bagarello, C. Trapani, CQ*-algebras: structure properties. Publ. RIMS Kyoto Univ. **32**, 85–116 (1996)
9. F. Bagarello, C. Trapani, The Heisenberg dynamics of spin systems: a quasi *-algebra approach. J. Math. Phys. **37**, 4219–4234 (1996)
10. F. Bagarello, C. Trapani, L^p-spaces as quasi *-algebras. J. Math. Anal. Appl. **197**, 810–824 (1996)
11. O. Bratteli, D.W. Robinson, Unbounded derivations of C*-algebras. Commun. Math. Phys. **42**, 253–268 (1975)
12. L. Chambadal, J.L. Ovaert, *Algèbre Linéaire et Algèbre Tensorielle* (Dunod Université, Paris, 1968)
13. A. Defant, K. Floret, *Tensor Norms and Operator Ideals* (North-Holland, Amsterdam, 1993)
14. M. Fragoulopoulou, *Topological Algebras with Involution* (North-Holland, Amsterdam, 2005)
15. M. Fragoulopoulou, C. Trapani, S. Triolo, Locally convex quasi *-algebras with sufficiently many *-representations. J. Math. Anal. Appl. **388**, 1180–1193 (2012)
16. M. Fragoulopoulou, A. Inoue, M. Weigt, Tensor products of generalized B^*-algebras. J. Math. Anal. Appl. **420**, 1787–1802 (2014)
17. M. Fragoulopoulou, A. Inoue, M. Weigt, Tensor products of unbounded operator algebras. Rocky Mt. **44**, 895–912 (2014)
18. W.-D. Heinrichs, Topological tensor products of unbounded operator algebras on Frèchet domains. Publ. RIMS Kyoto Univ. **33**, 241–255 (1997)
19. A.Ya. Helemskii, *Lectures and Exercises on Functional Analysis* (American Mathematical Society, Providence, 2006)
20. E. Hille, R. Phillips, *Functional Analysis and Semi-groups* (American Mathematical Society, Providence, 1996)
21. G. Lassner, Topological algebras and their applications in quantum statistics. Wiss. Z. KMU-Leipzig, Math. Naturwiss. R. **30**, 572–595 (1981)
22. G. Lassner, Algebras of unbounded operators and quantum dynamics. Physica A **124**, 471–480 (1984)
23. K.B. Laursen, Tensor products of Banach algebras with involution. Trans. Am. Math. Soc. **136**, 467–487 (1969)

24. A. Mallios, *Topological Algebras. Selected Topics* (North-Holland, Amsterdam, 1986)
25. G.J. Murphy, *C*-Algebras and Operator Theory* (Academic, Boston, 2014)
26. A.M. Sinclair, *Automatic Continuity of Linear Operators*. London Mathematical Society, Lecture Notes Series 21 (Cambridge University Press, Cambridge, 1976)
27. M. Takesaki, *Theory of Operator Algebras I* (Springer, New York, 1979)
28. C. Trapani, Quasi *-algebras of operators and their applications. Rev. Math. Phys. **7**, 1303–1332 (1995)
29. C. Trapani, Bounded elements and spectrum in Banach quasi *-algebras. Stud. Math. **172**, 249–273 (2006)
30. C. Trapani, *-Representations, seminorms and structure properties of normed quasi *-algebras. Stud. Math. **186**, 47–75 (2008)
31. M. Weigt, Derivations of τ-measurable operators. Oper. Theory Adv. Appl. **195**, 273–286 (2009)
32. M. Weigt, I. Zarakas, Unbounded derivations of GB*-algebras. Oper. Theory Adv. Appl. **247**, 69–82 (2015)
33. M. Weigt, I. Zarakas, Derivations of Fréchet nuclear GB*-algebras. Bull. Aust. Math. Soc. **92**, 290–301 (2015)

Minimality Properties of Sturm-Liouville Problems with Increasing Affine Boundary Conditions

Yagub N. Aliyev

Abstract We consider Sturm-Liouville problems with a boundary condition linearly dependent on the eigenparameter. We concentrate the study on the cases where non-real or non-simple (multiple) eigenvalues are possible. We prove that the system of root (i.e. eigen and associated) functions of the corresponding operator, with an arbitrary function removed, form a minimal system in $L_2(0, 1)$, except some cases where this system is neither complete nor minimal. The method used is based on the determination of the explicit form of the biorthogonal system. These minimality results can be extended to basis properties in $L_2(0, 1)$.

Keywords Sturm-Liouville · Eigenparameter-dependent boundary conditions · Minimal system · Root functions

Mathematics Subject Classification (2010) Primary 34B24; Secondary 34L10

1 Introduction

Consider the following spectral problem

$$-y'' + q(x)y = \lambda y, \quad 0 < x < 1, \tag{1.1}$$

$$y(0)\cos\beta = y'(0)\sin\beta, \quad 0 \le \beta < \pi, \tag{1.2}$$

$$y(1) = (c\lambda + d)y'(1), \tag{1.3}$$

This work was completed with the support of ADA University Faculty Research and Development Fund.

Y. N. Aliyev (✉)
School of IT and Engineering, ADA University, Baku, Azerbaijan
e-mail: yaliyev@ada.edu.az

M. A. Bastos et al. (eds.), *Operator Theory, Functional Analysis and Applications*,
Operator Theory: Advances and Applications 282,
https://doi.org/10.1007/978-3-030-51945-2_3

where c, d are real constants and $c > 0$, λ is the spectral parameter, $q(x)$ is a real valued and continuous function over the interval $[0, 1]$.

The present article is about the minimality properties in $L_2(0, 1)$ of the system of root functions of the boundary value problem (1.1)–(1.3).

It was proved in [2] (see also [3]) that the eigenvalues of the boundary value problem (1.1)–(1.3) form an infinite sequence accumulating only at $+\infty$ and only the following cases are possible:

(a) all the eigenvalues are real and simple;
(b) all the eigenvalues are real and all, except one double, are simple;
(c) all the eigenvalues are real and all, except one triple, are simple;
(d) all the eigenvalues are simple and all, except a conjugate pair of non-real, are real.

Let $\{v_n\}_{n=0}^{\infty}$ be a sequence of elements from $L_2(0, 1)$ and V_k the closure (in the norm of $L_2(0, 1)$) of the linear span of $\{v_n\}_{n=0,\ n\neq k}^{\infty}$. The system $\{v_n\}_{n=0}^{\infty}$ is called minimal in $L_2(0, 1)$ if $v_k \notin V_k$ for all $k = 0, 1, 2, \ldots$.

The eigenvalues λ_n $(n \geq 0)$ will be considered to be listed according to non-decreasing real part and repeated according to algebraic multiplicity. Asymptotics of eigenvalues and oscillation of eigenfunctions of the boundary value problem (1.1)–(1.3), with linear function in the boundary condition, replaced by general rational function, were studied in paper [3]. The case $c < 0$ in our problem does not involve any non-real or non-simple eigenvalues and can be found as a special case in papers [5, 6].

The main objective of the present paper is to show that the set of root functions with an arbitrary function removed form a minimal system in $L_2(0, 1)$, except some cases where this system is neither complete nor minimal. In more general form this result was obtained in [8, 9]. The main advantage of the current method is that it does not use any heavy machinery or extensions into more general operators defined over $L_2 \oplus C$. Also the special associated functions y^*_{k+1}, $y^{\#}_{k+1}$, $y^{\#}_{k+2}$, which are called in the paper as auxiliary associated functions and used to describe the necessary and sufficient conditions for minimality in a nice way, are very helpful when a concrete boundary value problem is studied. One such worked out example is given at the end of the current paper. The mere number of possibilities that arise in the simple case of linear dependence shows how different the minimality properties are from corresponding properties in the cases of boundary conditions free of any eigenvalue parameters. These minimality results can be used to prove basis properties of the root functions in $L_2(0, 1)$.

2 Inner Products and Norms of Eigenfunctions

The proofs in this and subsequent sections are similar to the corresponding results in [1] so we will skip some of them. Let $y(x, \lambda)$ be a non-zero solution of (1.1), satisfying initial conditions $y(0) = \sin\beta$ and $y'(0) = \sin\beta$. Then we can write the

characteristic equation as

$$\varpi(\lambda) = y(1,\lambda) - (c\lambda + d)y'(1,\lambda). \tag{2.1}$$

By (1.3), λ_n is an eigenvalue of (1.1)–(1.3) if $\varpi(\lambda_n) = 0$. It is a simple eigenvalue if $\varpi(\lambda_n) = 0 \neq \varpi'(\lambda_n)$. Similarly, λ_k is a double eigenvalue if

$$\varpi(\lambda_k) = \varpi'(\lambda_k) = 0 \neq \varpi''(\lambda_k),$$

and a triple eigenvalue if

$$\varpi(\lambda_k) = \varpi'(\lambda_k) = \varpi''(\lambda_k) = 0 \neq \varpi'''(\lambda_k).$$

We also note that $y(x,\lambda) \to y(x,\lambda_n)$, uniformly, for $x \in [0,1]$, as $\lambda \to \lambda_n$, and the function $y_n(x) = y(x,\lambda_n)$ is an eigenfunction of (1.1)–(1.3) corresponding to eigenvalue λ_n.

By (1.1)–(1.3) we have,

$$-y_n'' + q(x)y_n = \lambda_n y_n, \quad y_n'(0)\sin\beta = y_n(0)\cos\beta, \quad y_n(1) = (c\lambda_n + d)y_n'(1).$$

We denote by $(\cdot,\cdot)$ the inner product in $L_2(0,1)$.

Lemma 2.1 *Let y_n, y_m be eigenfunctions corresponding to eigenvalues λ_n, λ_m $(\lambda_n \neq \overline{\lambda_m})$. Then*

$$(y_n, y_m) = cy_n'(1)\overline{y_m'(1)}. \tag{2.2}$$

Proof To begin we note that

$$\frac{d}{dx}(y(x,\lambda)\overline{y'(x,\mu)} - y'(x,\lambda)\overline{y(x,\mu)}) = (\lambda - \overline{\mu})y(x,\lambda)\overline{y(x,\mu)}.$$

By integrating this identity from 0 to 1, we obtain

$$(\lambda - \overline{\mu})(y(\cdot,\lambda), y(\cdot,\mu)) = (y(x,\lambda)\overline{y'(x,\mu)} - y'(x,\lambda)\overline{y(x,\mu)})\Big|_0^1. \tag{2.3}$$

From (1.2), we obtain

$$y(0,\lambda)\overline{y'(0,\mu)} - y'(0,\lambda)\overline{y(0,\mu)} = 0. \tag{2.4}$$

By (2.1),

$$\begin{aligned} &y(1,\lambda)\overline{y'(1,\mu)} - y'(1,\lambda)\overline{y(1,\mu)} \\ &\quad = c(\lambda - \overline{\mu})y'(1,\lambda)\overline{y'(1,\mu)} - y'(1,\lambda)\overline{\varpi(\mu)} + \overline{y'(1,\mu)}\varpi(\lambda). \end{aligned} \tag{2.5}$$

From (2.3)–(2.5), it follows that for $\lambda \neq \overline{\mu}$,

$$(y(\cdot,\lambda), y(\cdot,\mu)) = cy'(1,\lambda)\overline{y'(1,\mu)} - y'(1,\lambda)\frac{\overline{\varpi(\mu)}}{\lambda-\overline{\mu}} + \overline{y'(1,\mu)}\frac{\varpi(\lambda)}{\lambda-\overline{\mu}}. \tag{2.6}$$

Note the fact that $\varpi(\lambda_n) = \varpi(\lambda_m) = 0$. The required equality (2.2) is now obvious if we substitute the parameters λ, μ by λ_n, λ_m, respectively. □

Remark 2.2 Since λ_n, λ_m are the eigenvalues of (1.1)–(1.3), it is also possible to prove (2.2) by letting $\lambda \to \lambda_n$ ($\overline{\mu} \neq \lambda_n$) and then letting $\mu \to \lambda_m$ in (2.6). It is also possible to substitute the parameters λ, μ, respectively by λ_n, λ_m at the beginning of the proof, in (2.3). Then the proof would be much simpler. But we chose the present proof because we will need (2.6) in the subsequent arguments.

In the remaining part of this section we will collect some simple facts about the inner products and norms of the eigenfunctions.

Lemma 2.3 *If λ_n is a real eigenvalue then*

$$\|y_n\|_2^2 = (y_n, y_n) = cy_n'(1)^2 + y_n'(1)\varpi'(\lambda_n). \tag{2.7}$$

Corollary 2.4 *If λ_k is a multiple eigenvalue then*

$$\|y_k\|_2^2 = (y_k, y_k) = cy_k'(1)^2. \tag{2.8}$$

An immediate corollary of (2.2) is the following.

Corollary 2.5 *If λ_r is a non-real eigenvalue then*

$$\|y_r\|_2^2 = c|y_r'(1)|^2. \tag{2.9}$$

For the eigenfunction y_n define

$$B_n = \|y_n\|_2^2 - c|y_n'(1)|^2. \tag{2.10}$$

The following corollary of (2.7)–(2.9) will be needed (cf. [2, Theorem 4.3]).

Corollary 2.6 *$B_n \neq 0$ if and only if the corresponding eigenvalue λ_n is real and simple.*

If λ_k is a multiple (double or triple) eigenvalue ($\lambda_k = \lambda_{k+1}$) then $B_k = y_k'(1)\omega'(\lambda_k) = 0$ and B_{k+1} is not defined, so we set $B_{k+1} = y_k'(1)\omega''(\lambda_k)/2$. If λ_k is a triple eigenvalue ($\lambda_k = \lambda_{k+1} = \lambda_{k+2}$) then $B_{k+1} = 0$ and B_{k+2} is not defined, so we set $B_{k+2} = y_k'(1)\omega'''(\lambda_k)/6$.

We conclude this section with the following.

Lemma 2.7 *If λ_r and $\lambda_s = \overline{\lambda_r}$ are a conjugate pair of non-real eigenvalues then*

$$(y_r, y_s) = c y_r'(1)^2 + y_r'(1)\varpi'(\lambda_r). \tag{2.11}$$

Since in the following text $\varpi'(\lambda_r)$ will appear in denominators of some of the fractions, it is useful to note here that all non-real eigenvalues of (1.1)–(1.3) are simple and therefore $\varpi'(\lambda_r) \neq 0$ in (2.11).

3 Inner Products and Norms of Associated Functions

In the previous section we collected some simple facts about the inner products and norms of the eigenfunctions. But in the case of multiple eigenvalues there are also some associated functions. In this section we will find formulae for inner products and norms involving associated functions. These cases appear only for the real eigenvalues so, throughout these sections we assume that all the eigenvalues and eigenfunctions are real. In particular, we will not write complex conjugate sign that appeared in the previous formulae.

If λ_k is a multiple eigenvalue ($\lambda_k = \lambda_{k+1}$) then for a first order associated function y_{k+1} corresponding to the eigenfunction y_k, following relations hold true [7, p. 28]:

$$-y_{k+1}'' + q(x)y_{k+1} = \lambda_k y_{k+1} + y_k,$$

$$y_{k+1}(0)\cos\beta = y_{k+1}'(0)\sin\beta,$$

$$y_{k+1}(1) = (c\lambda_k + d)y_{k+1}'(1) + c y_k'(1).$$

If λ_k is a triple eigenvalue ($\lambda_k = \lambda_{k+1} = \lambda_{k+2}$) then together with the first order associated function y_{k+1} there exists a second order associated function y_{k+2} for which

$$-y_{k+2}'' + q(x)y_{k+2} = \lambda_k y_{k+2} + y_{k+1},$$

$$y_{k+2}(0)\cos\beta = y_{k+2}'(0)\sin\beta,$$

$$y_{k+2}(1) = (c\lambda_k + d)y_{k+2}'(1) + c y_{k+1}'(1).$$

The following well known properties of the associated functions play an important role in our investigation. The functions $y_{k+1} + C y_k$ and $y_{k+2} + D y_k$, where C and D are arbitrary constants, are also associated functions of the first and second order, respectively. Next, we observe that if we replace the associated function y_{k+1} by $y_{k+1} + C y_k$, then the associated function y_{k+2} changes to $y_{k+2} + C y_{k+1}$.

By differentiating (1.1), (1.2) and (2.1) with respect to λ we obtain

$$-y''_\lambda(x,\lambda)+q(x)y_\lambda(x,\lambda)=\lambda y_\lambda(x,\lambda)+y(x,\lambda),$$

$$y_\lambda(0,\lambda)\cos\beta=y'_\lambda(0,\lambda)\sin\beta,$$

$$\varpi'(\lambda)=y_\lambda(1,\lambda)-(c\lambda+d)y'_\lambda(1,\lambda)-cy'(1,\lambda),$$

where the subscript denotes differentiation with respect to λ.

Let λ_k be a multiple (double or triple) eigenvalue of (1.1)–(1.3). Since $\varpi(\lambda_k)=\varpi'(\lambda_k)=0$, it follows that $y(x,\lambda)\to y_k$, $y_\lambda(x,\lambda)\to\tilde{y}_{k+1}$, uniformly with respect to $x\in[0,1]$, as $\lambda\to\lambda_k$, where $\tilde{y}_{k+1}$ is one of the associated functions of the first order, and it is obvious that $\tilde{y}_{k+1}=y_{k+1}+\tilde{C}y_k$, for a certain constant $\tilde{C}$. Note that $\tilde{C}=(\tilde{y}'_{k+1}(1)-y'_{k+1}(1))/y'_k(1)$.

Similarly, if we differentiate (1.1), (1.2) and (2.1) with respect to λ again we obtain

$$-y''_{\lambda\lambda}(x,\lambda)+q(x)y_{\lambda\lambda}(x,\lambda)=\lambda y_{\lambda\lambda}(x,\lambda)+2y_\lambda(x,\lambda),$$

$$y_{\lambda\lambda}(0,\lambda)\cos\beta=y'_{\lambda\lambda}(0,\lambda)\sin\beta,$$

$$\varpi''(\lambda)=y'_{\lambda\lambda}(1,\lambda)-(c\lambda+d)y'_{\lambda\lambda}(1,\lambda)-2cy'_\lambda(1,\lambda).$$

We note again that if λ_k is a triple eigenvalue of (1.1)–(1.3) then $\varpi''(\lambda_k)=0$, hence $y_{\lambda\lambda}\to 2\tilde{y}_{k+2}$, uniformly with respect to $x\in[0,1]$, as $\lambda\to\lambda_k$, where $\tilde{y}_{k+2}$ is one of the associated functions of the second order corresponding to the first associated function $\tilde{y}_{k+1}$, and it is obvious that $\tilde{y}_{k+2}=y_{k+2}+\tilde{C}y_{k+1}+\tilde{D}y_k$, for a certain constant $\tilde{D}$. Note that $\tilde{D}=(\tilde{y}'_{k+2}(1)-y'_{k+2}(1)-\tilde{C}y'_{k+1}(1))/y'_k(1)$.

Lemma 3.1 *If λ_k is a multiple eigenvalue and $\lambda_n\neq\lambda_k$ then*

$$(y_{k+1},y_n)=cy'_{k+1}(1)y'_n(1). \tag{3.1}$$

Proof Differentiating (2.6) with respect to λ we obtain

$$\begin{aligned}(y_\lambda(\cdot,\lambda),y(\cdot,\mu))=&cy'_\lambda(1,\lambda)y'(1,\mu)-y'_\lambda(1,\lambda)\frac{\varpi(\mu)}{\lambda-\mu}+y'(1,\lambda)\frac{\varpi(\mu)}{(\lambda-\mu)^2}\\&+y'(1,\mu)\frac{\varpi'(\lambda)}{\lambda-\mu}-y'(1,\mu)\frac{\varpi(\lambda)}{(\lambda-\mu)^2}.\end{aligned} \tag{3.2}$$

Letting $\mu\to\lambda_n$ ($\lambda\neq\lambda_n$) and then $\lambda\to\lambda_k$ in (3.2) we obtain that (3.1) is true (cf. [1]). We used the fact that the differentiation and the subsequent passage to the limit within the integrals are meaningful because all the involved functions are continuous with respect to both x and λ [4, Ch. 3, §4, Theorems 1, 2]. □

The same result can be achieved if we started with the identity

$$\frac{d}{dx}\left(y_{k+1}y_n' - y_{k+1}'y_n\right) = (\lambda_k - \lambda_n)y_{k+1}y_n + y_k y_n,$$

which can be easily derived using the definition of y_{k+1}. By integrating this equality from 0 to 1 we will obtain

$$(\lambda_k - \lambda_n)(y_{k+1}, y_n) + (y_k, y_n) = \left(y_{k+1}y_n' - y_{k+1}'y_n\right)\Big|_0^1.$$

By using the boundary conditions for y_{k+1} and y_n, and the fact that $(y_k, y_n) = cy_k'(1)y_n'(1)$ we obtain (3.1) again.

Lemma 3.2 *If λ_k is a multiple eigenvalue then*

$$(y_{k+1}, y_k) = cy_{k+1}'(1)y_k'(1) + y_k'(1)\frac{\varpi''(\lambda_k)}{2}. \tag{3.3}$$

Let us apply the procedure mentioned in the comments following the proof of (3.1) to the functions y_{k+1} and y_k. By integrating the identity

$$\frac{d}{dx}\left(y_{k+1}y_k' - y_{k+1}'y_k\right) = y_k^2,$$

from 0 to 1, we will obtain

$$\|y_k\|_2^2 = \left(y_{k+1}y_k' - y_{k+1}'y_k\right)\Big|_0^1.$$

By using the boundary conditions for y_{k+1} and y_k in the last equality we obtain (2.8) again.

Lemma 3.3 *If λ_k is a multiple eigenvalue then*

$$\|y_{k+1}\|_2^2 = cy_{k+1}'(1)^2 + \widehat{y}_{k+1}'(1)\frac{\varpi''(\lambda_k)}{2} + y_k'(1)\frac{\varpi'''(\lambda_k)}{6}. \tag{3.4}$$

where $\widehat{y}_{k+1} = y_{k+1} - \tilde{C}y_k$.

Proof Differentiating (3.2) with respect to μ we obtain

$$\begin{aligned}(y_\lambda(\cdot,\lambda), y_\mu(\cdot,\mu)) &= cy_\lambda'(1,\lambda)y_\mu'(1,\mu) - y_\lambda'(1,\lambda)\frac{\varpi'(\mu)}{\lambda-\mu} - y_\lambda'(1,\lambda)\frac{\varpi(\mu)}{(\lambda-\mu)^2}\\ &+ y'(1,\lambda)\frac{\varpi'(\mu)}{(\lambda-\mu)^2} + y'(1,\lambda)\frac{2\varpi(\mu)}{(\lambda-\mu)^3} + y_\mu'(1,\mu)\frac{\varpi'(\lambda)}{\lambda-\mu}\\ &+ y'(1,\mu)\frac{\varpi'(\lambda)}{(\lambda-\mu)^2} - y_\mu'(1,\mu)\frac{\varpi(\lambda)}{(\lambda-\mu)^2} - y'(1,\mu)\frac{2\varpi(\lambda)}{(\lambda-\mu)^3}.\end{aligned}$$

Letting $\mu \to \lambda_k$ ($\lambda \neq \lambda_k$) (cf. [1]) and then $\lambda \to \lambda_k$ we obtain (3.4). □

Lemma 3.4 *If λ_k is a triple eigenvalue and $\lambda_n \neq \lambda_k$ then*

$$(y_{k+2}, y_n) = c y'_{k+2}(1) y'_n(1). \tag{3.5}$$

Proof Differentiating (3.2) with respect to λ, letting $\lambda \to \lambda_k$ ($\mu \neq \lambda_k$) we obtain

$$\begin{aligned}(\tilde{y}_{k+2}, y(\cdot, \mu)) =& c\tilde{y}'_{k+2}(1) y'(1, \mu) - \tilde{y}'_{k+2}(1) \frac{\varpi(\mu)}{\lambda_k - \mu} \\ &+ \tilde{y}'_{k+1}(1) \frac{\varpi(\mu)}{(\lambda_k - \mu)^2} - y'_k(1) \frac{\varpi(\mu)}{(\lambda_k - \mu)^3},\end{aligned}$$

from which (3.5) easily follows (cf. [1]). □

Again, the same result can be achieved if we started with the identity

$$\frac{d}{dx}\left(y_{k+2} y'_n - y'_{k+2} y_n\right) = (\lambda_k - \lambda_n) y_{k+2} y_n + y_{k+1} y_n,$$

which can be again easily derived using the definition of y_{k+2}. By integrating this equality from 0 to 1 we will obtain

$$(\lambda_k - \lambda_n)(y_{k+2}, y_n) + (y_{k+1}, y_n) = \left(y_{k+2} y'_n - y'_{k+2} y_n\right)\Big|_0^1.$$

By using the boundary conditions for y_{k+2} and y_n, and (3.1) we obtain a new proof for (3.5).

Lemma 3.5 *If λ_k is a triple eigenvalue then*

$$(y_{k+2}, y_k) = c y'_{k+2}(1) y'_k(1) + y'_k(1) \frac{\varpi'''(\lambda_k)}{6}. \tag{3.6}$$

Again, by applying the above mentioned procedure to the functions y_{k+2} and y_k we obtain

$$\frac{d}{dx}\left(y_{k+2} y'_k - y'_{k+2} y_k\right) = y_{k+1} y_k.$$

By integrating this equality from 0 to 1 we will obtain

$$(y_{k+1}, y_k) = \left(y_{k+2} y'_k - y'_{k+2} y_k\right)\Big|_0^1.$$

By using the boundary conditions for y_{k+2} and y_k, we obtain

$$(y_{k+1}, y_k) = c y'_{k+1}(1) y'_k(1),$$

which is in perfect agreement with (3.3), because in triple eigenvalue case $\varpi''(\lambda_k) = 0$.

Lemma 3.6 *If λ_k is a triple eigenvalue then*

$$(y_{k+2}, y_{k+1}) = cy'_{k+2}(1)y'_{k+1}(1) + \widehat{y}'_{k+1}(1)\frac{\varpi'''(\lambda_k)}{6} + y'_k(1)\frac{\varpi^{IV}(\lambda_k)}{24}.$$

By applying the above mentioned procedure one more time but now to the functions y_{k+2} and y_{k+1} we obtain

$$\frac{d}{dx}\left(y_{k+2}y'_{k+1} - y'_{k+2}y_{k+1}\right) = y_{k+1}^2 - y_{k+2}y_k.$$

Integration of this equality from 0 to 1 will give us

$$\|y_{k+1}\|_2^2 - (y_{k+2}, y_k) = \left(y_{k+2}y'_{k+1} - y'_{k+2}y_{k+1}\right)\Big|_0^1.$$

By using the boundary conditions for y_{k+2} and y_{k+1}, we obtain

$$\|y_{k+1}\|_2^2 - (y_{k+2}, y_k) = cy'_{k+1}(1)^2 - cy'_{k+2}(1)y'_k(1),$$

which is again in perfect agreement with (3.4) and (3.6), because $\varpi''(\lambda_k) = 0$.

Lemma 3.7 *If λ_k is a triple eigenvalue then*

$$\begin{aligned}\|y_{k+2}\|_2^2 =& cy'_{k+2}(1)^2 + \widehat{y}'_{k+2}(1)\frac{\varpi'''(\lambda_k)}{6} \\ &+ \widehat{y}'_{k+1}(1)\frac{\varpi^{IV}(\lambda_k)}{24} + y'_k(1)\frac{\varpi^{V}(\lambda_k)}{120},\end{aligned}$$

where $\widehat{y}_{k+2} = y_{k+2} - \tilde{C}\widehat{y}_{k+1} - \tilde{D}y_k$.

4 Existence of Auxiliary Associated Functions

By comparing the equalities in Sects. 2 and 3 we can see that there are some fundamental differences between the formulae for the inner products and norms of the eigenfunctions and the corresponding formulae for the associated functions. In this section we will prove that it is possible to find special associated functions (we call them *auxiliary associated functions*) whose properties in inner products make them more close to the eigenfunctions than the other associated functions. In the last section, these functions will play a crucial role in our description of minimality properties.

Lemma 4.1 *If λ_k is a double eigenvalue then there exists an associated function $y^*_{k+1} = y_{k+1} + C_1 y_k$, where*

$$C_1 = -\frac{y'_k(1)\varpi'''(\lambda_k) + 3\widehat{y}'_{k+1}(1)\varpi''(\lambda_k)}{3y'_k(1)\varpi''(\lambda_k)},$$

such that

$$(y^*_{k+1}, y_{k+1}) = c(y^*_{k+1})'(1)y'_{k+1}(1). \tag{4.1}$$

Here, it should be pointed out that $(y^*_{k+1})'(1) = 0$ if and only if

$$\varpi'''(\lambda_k) = 3\tilde{C}\varpi''(\lambda_k).$$

Before proceeding, we also note that for $\lambda_n \neq \lambda_k$,

$$(y^*_{k+1}, y_n) = c(y^*_{k+1})'(1)y'_n(1), \tag{4.2}$$

$$(y^*_{k+1}, y_k) = c(y^*_{k+1})'(1)y'_k(1) + y'_k(1)\frac{\varpi''(\lambda_k)}{2}.$$

We shall now concentrate on the triple eigenvalue case. Although we will not need the function y^*_{k+1} in the triple eigenvalue case, it is still worthwhile to note that such a function does not exist in this case. Instead, we will need other associated functions of y_k which will be denoted by $y^\#_{k+1}$ and $y^\#_{k+2}$.

Lemma 4.2 *If λ_k is a triple eigenvalue then there exists an associated function $y^\#_{k+1} = y_{k+1} + C_2 y_k$, where*

$$C_2 = -\frac{y'_k(1)\varpi^{IV}(\lambda_k) + 4\widehat{y}'_{k+1}(1)\varpi'''(\lambda_k)}{4y'_k(1)\varpi'''(\lambda_k)},$$

for which

$$(y^\#_{k+1}, y_{k+2}) = c(y^\#_{k+1})'(1)y'_{k+2}(1).$$

It is worthwhile to note that $(y^\#_{k+1})'(1) = 0$ if and only if $\varpi^{IV}(\lambda_k) = 4\tilde{C}\varpi'''(\lambda_k)$. We now indicate some relations between $y^\#_{k+1}$ and the other root functions:

$$(y^\#_{k+1}, y_n) = c(y^\#_{k+1})'(1)y'_n(1),\ (n \neq k+1,\ k+2),$$

$$(y^\#_{k+1}, y_{k+1}) = c(y^\#_{k+1})'(1)y'_{k+1}(1) + y'_k(1)\frac{\varpi'''(\lambda_k)}{6}.$$

Note that the function y^*_{k+2}, defined by $y^*_{k+2} = y_{k+2} + C_2 y_{k+1}$, where C_2 is the same constant, also enjoys similar properties:

$$(y^*_{k+2}, y_{k+1}) = c(y^*_{k+2})'(1)y'_{k+1}(1), \tag{4.3}$$

$$(y^*_{k+2}, y_n) = c(y^*_{k+2})'(1)y'_n(1), \ (n \neq k, \ k+1, \ k+2), \tag{4.4}$$

$$(y^*_{k+2}, y_k) = c(y^*_{k+2})'(1)y'_k(1) + y'_k(1)\frac{\varpi'''(\lambda_k)}{6}. \tag{4.5}$$

Lemma 4.3 *If λ_k is a triple eigenvalue then there exists an associated function $y^\#_{k+2} = y^*_{k+2} + D_1 y_k$, where D_1 is a constant, for which*

$$(y^\#_{k+2}, y_{k+1}) = c(y^\#_{k+2})'(1)y'_{k+1}(1), \tag{4.6}$$

$$(y^\#_{k+2}, y_{k+2}) = c(y^\#_{k+2})'(1)y'_{k+2}(1). \tag{4.7}$$

Proof Note first that

$$(y^*_{k+2}, y_{k+2}) = c(y^*_{k+2})'(1)y'_{k+2}(1) + Q_k,$$

where

$$\begin{aligned} Q_k =& \widehat{y}'_{k+2}(1)\frac{\varpi'''(\lambda_k)}{6} + \widehat{y}'_{k+1}(1)\frac{\varpi^{IV}(\lambda_k)}{24} + y'_k(1)\frac{\varpi^{V}(\lambda_k)}{120} \\ &+ C_2\left(\widehat{y}'_{k+1}(1)\frac{\varpi'''(\lambda_k)}{6} + y'_k(1)\frac{\varpi^{IV}(\lambda_k)}{24}\right). \end{aligned}$$

It is not difficult to check that for the function $y^\#_{k+2} = y^*_{k+2} + D_1 y_k$, where

$$D_1 = -\frac{6Q_k}{y'_k(1)\varpi'''(\lambda_k)},$$

both equalities (4.6) and (4.7) hold true. □

Note also that for the function $y^\#_{k+2}$, equalities like (4.4), (4.5) are also true:

$$(y^\#_{k+2}, y_n) = c(y^\#_{k+2})'(1)y'_n(1), \ (n \neq k, \ k+1, \ k+2);$$

$$(y^\#_{k+2}, y_k) = c(y^\#_{k+2})'(1)y'_k(1) + y'_k(1)\frac{\varpi'''(\lambda_k)}{6}.$$

We remark that $(y^\#_{k+2})'(1) = 0$ if and only if

$$5\varpi^{IV}(\lambda_k)\left(\varpi^{IV}(\lambda_k) - 4\tilde{C}\varpi'''(\lambda_k)\right) = 4\varpi'''(\lambda_k)\left(\varpi^{V}(\lambda_k) - 20\tilde{D}\varpi'''(\lambda_k)\right).$$

5 Minimality of the System of Root Functions

In this section we will consider all possible cases of the choice of the root function which will be deleted from the system to obtain a minimal system. In each case we will construct explicitly a biorthogonal system.

5.1 Case (a)

Theorem 5.1 *If all the eigenvalues of* (1.1)–(1.3) *are real and simple then the system*

$$\{y_n\} \ (n = 0,\ 1, \ldots; n \neq l), \tag{5.1}$$

where l is a non-negative integer, is minimal in $L_2(0, 1)$.

Proof It suffices to show the existence of a system

$$\{u_n\} \ (n = 0,\ 1, \ldots; n \neq l), \tag{5.2}$$

biorthogonal to the system (5.1). Noting the relation $B_n \neq 0$ we define elements of the system (5.2) by

$$u_n(x) = \frac{1}{B_n y_l'(1)} \begin{vmatrix} y_n(x) & y_n'(1) \\ y_l(x) & y_l'(1) \end{vmatrix}. \tag{5.3}$$

It remains to see, noting (2.2), (2.7) and (2.10), that

$$(u_n, y_m) = \delta_{nm},$$

where δ_{nm} $(n,\ m = 0,\ 1, \ldots;\ n,\ m \neq l)$ denotes as usually, Kronecker's symbol: $\delta_{nm} = 0$ if $n \neq m$ and $\delta_{nn} = 1$. □

5.2 Case (b)

Theorem 5.2 *If λ_k is a double eigenvalue then the system*

$$\{y_n\} \ (n = 0,\ 1, \ldots; n \neq k + 1),$$

is minimal in $L_2(0, 1)$.

Proof In this case the biorthogonal system is defined by

$$u_n(x) = \frac{1}{B_n y_k'(1)} \begin{vmatrix} y_n(x) & y_n'(1) \\ y_k(x) & y_k'(1) \end{vmatrix} \quad (n \neq k,\ k+1), \tag{5.4}$$

$$u_k(x) = \frac{1}{B_{k+1} y_k'(1)} \begin{vmatrix} y_{k+1}(x) & y_{k+1}'(1) \\ y_k(x) & y_k'(1) \end{vmatrix}. \qquad \square$$

Theorem 5.3 *If λ_k is a double eigenvalue, and if $(y_{k+1}^*)'(1) \neq 0$ then the system*

$$\{y_n\} \quad (n = 0,\ 1, \ldots; n \neq k), \tag{5.5}$$

is minimal in $L_2(0, 1)$.

Proof The elements of the biorthogonal system are defined as follows

$$u_n(x) = \frac{1}{B_n (y_{k+1}^*)'(1)} \begin{vmatrix} y_n(x) & y_n'(1) \\ y_{k+1}^*(x) & (y_{k+1}^*)'(1) \end{vmatrix} \quad (n \neq k,\ k+1),$$

$$u_{k+1}(x) = \frac{1}{B_{k+1} (y_{k+1}^*)'(1)} \begin{vmatrix} y_k(x) & y_k'(1) \\ y_{k+1}^*(x) & (y_{k+1}^*)'(1) \end{vmatrix}. \qquad \square$$

Before proceeding we comment on the condition $(y_{k+1}^*)'(1) \neq 0$ above. Let $(y_{k+1}^*)'(1) = 0$, then by (4.1), (4.2) the function y_{k+1}^* is orthogonal to all the elements of the system (5.5). Therefore the system (5.5) is not complete in $L_2(0, 1)$.

Theorem 5.4 *If λ_k is a double eigenvalue then the system*

$$\{y_n\} \quad (n = 0,\ 1, \ldots; n \neq l),$$

where $l \neq k,\ k+1$ is a non-negative integer, is minimal in $L_2(0, 1)$.

Proof The biorthogonal system is given by the formula (5.3) for $n \neq k,\ k+1$, and

$$u_{k+1}(x) = \frac{1}{B_{k+1} y_l'(1)} \begin{vmatrix} y_k(x) & y_k'(1) \\ y_l(x) & y_l'(1) \end{vmatrix},$$

$$u_k(x) = \frac{1}{B_{k+1} y_l'(1)} \begin{vmatrix} y_{k+1}^*(x) & (y_{k+1}^*)'(1) \\ y_l(x) & y_l'(1) \end{vmatrix}. \qquad \square$$

5.3 Case (c)

Theorem 5.5 *If λ_k is a triple eigenvalue then the system*

$$\{y_n\} \quad (n = 0,\ 1, \ldots; n \neq k+2),$$

is minimal in $L_2(0, 1)$.

Proof The biorthogonal system is given by the formula (5.4) for $n \neq k,\ k+1,\ k+2$, and

$$u_{k+1}(x) = \frac{1}{B_{k+2}y_k'(1)} \begin{vmatrix} y_{k+1}(x) & y_{k+1}'(1) \\ y_k(x) & y_k'(1) \end{vmatrix},$$

$$u_k(x) = \frac{1}{B_{k+2}y_k'(1)} \begin{vmatrix} y_{k+2}^{\#}(x) & (y_{k+2}^{\#})'(1) \\ y_k(x) & y_k'(1) \end{vmatrix}. \qquad \square$$

Theorem 5.6 *If λ_k is a triple eigenvalue, and if $(y_{k+1}^{\#})'(1) \neq 0$ then the system*

$$\{y_n\}\ (n = 0,\ 1, \ldots; n \neq k+1), \tag{5.6}$$

is minimal in $L_2(0, 1)$.

Proof In this case the elements of the biorthogonal system are

$$u_n(x) = \frac{1}{B_n(y_{k+1}^{\#})'(1)} \begin{vmatrix} y_n(x) & y_n'(1) \\ y_{k+1}^{\#}(x) & (y_{k+1}^{\#})'(1) \end{vmatrix}\ (n \neq k,\ k+1,\ k+2),$$

$$u_{k+2}(x) = \frac{1}{B_{k+2}(y_{k+1}^{\#})'(1)} \begin{vmatrix} y_k(x) & y_k'(1) \\ y_{k+1}^{\#}(x) & (y_{k+1}^{\#})'(1) \end{vmatrix},$$

$$u_k(x) = \frac{1}{B_{k+2}(y_{k+1}^{\#})'(1)} \begin{vmatrix} y_{k+2}^{\#}(x) & (y_{k+2}^{\#})'(1) \\ y_{k+1}^{\#}(x) & (y_{k+1}^{\#})'(1) \end{vmatrix}. \qquad \square$$

In analogy with Theorem 5.3, we may show that if $(y_{k+1}^{\#})'(1) = 0$ then the function $y_{k+1}^{\#}(x)$ is orthogonal to all the elements of the system (5.6); hence the system (5.6) is not complete.

Theorem 5.7 *If λ_k is a triple eigenvalue, and if $(y_{k+2}^{\#})'(1) \neq 0$ then the system*

$$\{y_n\}\ (n = 0,\ 1, \ldots; n \neq k), \tag{5.7}$$

is minimal in $L_2(0, 1)$.

Proof We define the elements of the biorthogonal system by

$$u_n(x) = \frac{1}{B_n(y_{k+2}^{\#})'(1)} \begin{vmatrix} y_n(x) & y_n'(1) \\ y_{k+2}^{\#}(x) & (y_{k+2}^{\#})'(1) \end{vmatrix}\ (n \neq k,\ k+1,\ k+2),$$

$$u_{k+2}(x) = \frac{1}{B_{k+2}(y_{k+2}^{\#})'(1)} \begin{vmatrix} y_k(x) & y_k'(1) \\ y_{k+2}^{\#}(x) & (y_{k+2}^{\#})'(1) \end{vmatrix},$$

$$u_{k+1}(x) = \frac{1}{B_{k+2}(y_{k+2}^{\#})'(1)} \begin{vmatrix} y_{k+1}(x) & y_{k+1}'(1) \\ y_{k+2}^{\#}(x) & (y_{k+2}^{\#})'(1) \end{vmatrix}. \qquad \square$$

If $(y_{k+2}^{\#})'(1) = 0$ then the system (5.7) is not complete.

Theorem 5.8 *If λ_k is a triple eigenvalue then the system*

$$\{y_n\}\ (n = 0,\ 1, \ldots; n \neq l),$$

where $l \neq k,\ k+1,\ k+2$ is a non-negative integer, is minimal in $L_2(0, 1)$.

Proof The elements of the biorthogonal system can be defined by (5.3) for $n \neq k,\ k+1,\ k+2,\ l$ and

$$u_{k+2}(x) = \frac{1}{B_{k+2}y_l'(1)} \begin{vmatrix} y_k(x) & y_k'(1) \\ y_l(x) & y_l'(1) \end{vmatrix},$$

$$u_{k+1}(x) = \frac{1}{B_{k+2}y_l'(1)} \begin{vmatrix} y_{k+1}^{\#}(x) & (y_{k+1}^{\#})'(1) \\ y_l(x) & y_l'(1) \end{vmatrix},$$

$$u_k(x) = \frac{1}{B_{k+2}y_l'(1)} \begin{vmatrix} y_{k+2}^{\#}(x) & (y_{k+2}^{\#})'(1) \\ y_l(x) & y_l'(1) \end{vmatrix}.$$

□

5.4 Case (d)

Theorem 5.9 *If λ_r and $\lambda_s = \overline{\lambda_r}$ are a conjugate pair of non-real eigenvalues then each of the systems*

$$\{y_n\}\ (n = 0,\ 1, \ldots; n \neq r), \tag{5.8}$$

$$\{y_n\}\ (n = 0,\ 1, \ldots; n \neq l), \tag{5.9}$$

where $l \neq r, s$ is a non-negative integer, is minimal in $L_2(0, 1)$.

Proof The biorthogonal system of (5.8) is as follows

$$u_n(x) = \frac{1}{B_n y_s'(1)} \begin{vmatrix} y_n(x) & y_n'(1) \\ y_s(x) & y_s'(1) \end{vmatrix}\ (n \neq r,\ s),$$

$$u_s(x) = \frac{1}{y_r'(1)y_s'(1)\varpi'(\lambda_r)} \begin{vmatrix} y_r(x) & y_r'(1) \\ y_s(x) & y_s'(1) \end{vmatrix}.$$

The biorthogonal system of (5.9) is defined by (5.3) for $n \neq r, s, l$ and

$$u_r(x) = \frac{1}{y_s'(1)y_l'(1)\varpi'(\lambda_s)} \begin{vmatrix} y_s(x) & y_s'(1) \\ y_l(x) & y_l'(1) \end{vmatrix},$$

$$u_s(x) = \frac{1}{y_r'(1)y_l'(1)\varpi'(\lambda_r)} \begin{vmatrix} y_r(x) & y_r'(1) \\ y_l(x) & y_l'(1) \end{vmatrix}.$$

□

Remark 5.10 Using the method of the paper [5] these minimality results can also be extended to basis properties. Then the sufficient conditions

$$(y_{k+1}^{*})'(1) \neq 0, \quad (y_{k+1}^{\#})'(1) \neq 0, \quad (y_{k+2}^{\#})'(1) \neq 0$$

in Theorems 5.3, 5.6 and 5.7, respectively, will be necessary conditions, too.

6 Example

As an illustration of the above theory, we present a particular result for the following problem

$$-y'' = \lambda y, \quad 0 < x < 1, \quad y(0) = 0, \quad y(1) = \left(\frac{\lambda}{3} + 1\right) y'(1).$$

For this problem $\lambda_0 = \lambda_1 = 0$ is a double eigenvalue. The other eigenvalues $\lambda_2 < \lambda_3 < \dots$ are the solutions of the equation $\tan\sqrt{\lambda} = \sqrt{\lambda}\left(\frac{\lambda}{3} + 1\right)$. Eigenfunctions are $y_0 = x$, $y_n = \sin\sqrt{\lambda_n}x$ $(n \geq 2)$ and an associated function corresponding to y_0 is $y_1 = -\frac{1}{6}x^3 + Cx$, where C is an arbitrary constant. We look for an auxiliary associated function of the form $y_1^{*} = -\frac{1}{6}x^3 + C'x$. In other words, $C_1 = C' - C$. By (4.1),

$$\int_0^1 \left(-\frac{1}{6}x^3 + Cx\right)\left(-\frac{1}{6}x^3 + C'x\right) dx = \frac{1}{3}\left(-\frac{1}{2} + C\right)\left(-\frac{1}{2} + C'\right).$$

From this equality we obtain that $C' = -C + \frac{25}{42}$, so

$$(y_1^{*})'(x) = -\frac{1}{6}x^3 + \left(-C + \frac{25}{42}\right)x.$$

Consequently, $(y_1^{*})'(1) = \frac{2}{21} - C$. Therefore the above condition $(y_1^{*})'(1) \neq 0$ in Theorem 5.3 is equivalent to $C \neq \frac{2}{21}$. So, the system

$$\{y_n\}\ (n = 1,\ 2, \dots) = \left\{\sin\sqrt{\lambda_n}x\right\}\ (n = 2,\ 3, \dots) \cup \left\{-\frac{1}{6}x^3 + Cx\right\},$$

from which only the eigenfunction y_0 is excluded but the associated function y_1 is included, is minimal in $L_2(0, 1)$ if $C \neq \frac{2}{21}$.

We shall now apply a different method to this problem. Note that for this problem $y(x, \lambda) = \frac{\sin\sqrt{\lambda}x}{\sqrt{\lambda}}$. We need $\sqrt{\lambda}$ in the denominator to make sure that

$\lim_{\lambda\to 0} y(x,\lambda) = y_0 = x$. Then

$$y_\lambda(x,\lambda) = \frac{x\cos\sqrt{\lambda}x}{2\lambda} - \frac{\sin\sqrt{\lambda}x}{2\lambda\sqrt{\lambda}}$$

and therefore $\tilde{y}_1 = \lim_{\lambda\to 0} y_\lambda(x,\lambda) = -\frac{x^3}{6}$. Let $y_1 = -\frac{1}{6}x^3 + C$. Then $\tilde{C} = -C$. Note also that $\varpi(\lambda) = \sin\sqrt{\lambda} - \left(\frac{\lambda}{3}+1\right)\sqrt{\lambda}\cos\sqrt{\lambda}$, then

$$\varpi(0) = \lim_{\lambda\to 0}\varpi(\lambda) = 0, \qquad \varpi'(0) = \lim_{\lambda\to 0}\varpi'(\lambda) = 0,$$
$$\varpi''(0) = \lim_{\lambda\to 0}\varpi''(\lambda) = 4/15, \ \varpi'''(0) = \lim_{\lambda\to 0}\varpi'''(\lambda) = -8/105.$$

As was pointed out in the comments following the proof of Lemma 4.1, the condition $(y_1^*)'(1) \neq 0$ is equivalent to $\varpi'''(\lambda_k) \neq 3\tilde{C}\varpi''(\lambda_k)$. Since $C = -\tilde{C}$, we obtain, once again, $C \neq \frac{2}{21}$.

If $C = \frac{2}{21}$ then we obtain $y_1^*(x) = -\frac{x^3}{6} + \frac{x}{2}$ which is orthogonal to all the elements of the system $\{y_n\}$ $(n = 1,\ 2, \ldots)$.

Acknowledgments Many thanks to ADA University and especially School of IT and Engineering for their constant support.

References

1. Y.N. Aliyev, Minimality of the system of root functions of Sturm-Liouville problems with decreasing affine boundary conditions. Colloq. Math. **109**, 147–162 (2007)
2. P.A. Binding, P.J. Browne, Application of two parameter eigencurves to Sturm-Liouville problems with eigenparameter-dependent boundary conditions. Proc. R. Soc. Edinb. **125A**, 1205–1218 (1995)
3. P.A. Binding, P.J. Browne, B.A. Watson, Equivalence of inverse Sturm-Liouville problems with boundary conditions rationally dependent on the eigenparameter. J. Math. Anal. Appl. **291**, 246–261 (2004)
4. A.P. Kartashev, B.L. Rojdestvenskiy, *Ordinary Differential Equations and Foundations of Calculus of Variations* (Nauka, Moscow, 1980, in Russian)
5. N.B. Kerimov, Y.N. Aliyev, The basis property in L_p of the boundary value problem rationally dependent on the eigenparameter. Stud. Math. **174**, 201–212 (2006)
6. N.B. Kerimov, V.S. Mirzoev, On the basis properties of one spectral problem with a spectral parameter in boundary conditions. Sib. Math. J. **44**, 813–816 (2003)
7. M.A. Naimark, *Linear Differential Operators*, 2nd edn. (Nauka, Moscow, 1969, in Russian). English trans. of 1st edn., Parts I, II (Ungar, New York, 1967, 1968)
8. A.A. Shkalikov, Boundary value problems for ordinary differential equations with a parameter in the boundary conditions. Tr. Semin. Im. I. G. Petrovskogo **9**, 190–229 (1983, in Russian)
9. A.A. Shkalikov, Basis properties of root functions of differential operators with spectral parameter in the boundary conditions. Differ. Equ. **55**, 631–643 (2019)

Scattering, Spectrum and Resonance States Completeness for a Quantum Graph with Rashba Hamiltonian

Irina V. Blinova, Igor Y. Popov, and Maria O. Smolkina

Abstract Quantum graphs consisting of a ring with two semi-infinite edges attached to the same point of the ring is considered. We deal with the Rashba spin-orbit Hamiltonian on the graph. A theorem concerning to completeness of the resonance states on the ring is proved. Due to use of a functional model, the problem reduces to factorization of the characteristic matrix-function. The result is compared with the corresponding completeness theorem for the Schrödinger, Dirac and Landau quantum graphs.

Keywords Spectrum · Resonance · Completeness · Spin-orbit interaction

Mathematics Subject Classification (2010) Primary 81U20; Secondary 46N50

1 Introduction

The problem of resonances was studied during a long time starting with the famous Rayleigh work on the Helmholtz resonator. The completeness problem for resonance states is younger. Mathematicians dealt with it during a half of century. The problem is related to the stability of the completeness under some perturbation. Namely, one starts with a system with purely discrete spectrum and complete system of eigenstates, e.g., a closed resonator. Then one perturbs the system in such a way that eigenvalues turn to resonances and eigenstates—to resonance states, e.g., the resonator with a boundary window [2, 7, 8, 10]. Correspondingly, a natural question appears: is the system of quasi-eigenstates complete? One of the approaches to this problem is related to the Sz.-Nagy functional model [14, 21, 25]. Starting with work [1], it is known that the scattering matrix is the same as the characteristic function from the functional model. In particular, root vectors in the functional model

I. V. Blinova · I. Y. Popov (✉) · M. O. Smolkina
ITMO University, St. Petersburg, Russia
e-mail: irin-a@yandex.ru

M. A. Bastos et al. (eds.), *Operator Theory, Functional Analysis and Applications*,
Operator Theory: Advances and Applications 282,
https://doi.org/10.1007/978-3-030-51945-2_4

correspond to resonance states in scattering theory. The completeness problem for the system of root vectors is related to the factorization problem for the characteristic function as a function from the Hardy space. We will use this relation to study the completeness. Particularly, for the finite-dimensional case, this approach gives one an effective completeness criterion [21].

The simplest model for an open resonator is based on a quantum graph, well-developed model in quantum theory [3, 9, 18]. The corresponding completeness problem was considered in [4, 11, 23, 24].

Spin-orbit interaction attracts great attention last time due to possible perspectives of applications in nanoelectronics, particularly, in spintronics and quantum computing [5, 6, 13, 15, 16, 19, 20]. Full spin-orbit Hamiltonian is very complicated even for quantum graph. Usually, theoreticians consider its particular cases, Rashba or Dresselhaus Hamiltonians. In this paper we deal with the Rashba Hamiltonian. We investigate this operator on a quantum graph with loop imbedded in a plane in $\mathbb{R}^3$. The graph is posed in a constant magnetic field orthogonal to the graph plane. We choose such type of the graph (a loop touching a line) because the Schrödinger, Dirac and Landau operators on such graph lead to incompleteness of the resonance states [12]. In some sense, this graph presents a degenerate case and any small perturbation, e.g., point-like potential at the vertex, restores the completeness. It is interesting that the Rashba Hamiltonian leads to the completeness.

1.1 Scattering, Functional Model and Completeness Criterion

It is convenient to consider resonances in the framework of the Lax-Phillips scattering theory [17]. It is based on non-stationary problem. Let H be the Hamiltonian for the quantum graph, i.e. the Rashba Hamiltonian on the ring and the Schrödinger operator on the half-axes. The wave function is a two-component vector. Different components correspond to different spin directions. Consider the Cauchy problem for the time-dependent problem on the graph Γ:

$$\begin{cases} i\hbar u'_t = Hu, \\ u(x,0) = u^0(x), \ x \in \Gamma. \end{cases}$$

Below, we will use the system of units in which $\hbar = 1$. The standard Lax-Phillips approach is applied to the wave (acoustic) equation. There is a close relation between the Schrödinger and wave cases. We will describe it briefly following [17, Section 6.4]. Namely, it is necessary to consider the operator A^2 (here A is the generator of the evolution group $U(t)$ for the wave equation, see below):

$$A^2 = \begin{pmatrix} H & 0 \\ 0 & H \end{pmatrix}.$$

One can see that A^2 acts as the Hamiltonian H on each component of the data of the acoustic problem. This allows one to use the "acoustic" construction for the Schrödinger case and, as a result, comes to the relation between the Schrödinger (S^{Schr}) and the acoustic (S) scattering matrices:

$$S^{Schr}(z) = S(\sqrt{z}).$$

Let us describe briefly the acoustic case. Consider the Cauchy problem for the wave equation

$$\begin{cases} u''_{tt} = u''_{xx}, \\ u(x,0) = u_0(x), u'_t(x,0) = u_1(x), x \in \Gamma. \end{cases} \tag{1.1}$$

Here u, u_0, u_1 are two-component vectors. Let $\mathcal{E}$ be the Hilbert space of four-component functions (u_0, u_1) on the graph with finite energy

$$\|(u_0,u_1)\|^2_{\mathcal{E}} = 2^{-1}\int_\Gamma (|u'_0|^2 + |u_1|^2)dx.$$

The pair (u_0, u_1) is called the Cauchy data. The unitary (in $\mathcal{E}$) group $U(t)$, $U(t)(u_0,u_1) = (u(x,t), u'_t(x,t))$, solves the problem (1.1). The unitary group $U(t)|_{t\in\mathbb{R}}$ has two orthogonal (in $\mathcal{E}$) subspaces, D_- and D_+, called, correspondingly, incoming and outgoing subspaces.

Definition 1.1 The outgoing (incoming) subspace $D_+(D_-)$ is a subspace of $\mathcal{E}$ having the following properties:

(a) $U(t)D_+ \subset D_+$ for $t > 0$; $U(t)D_- \subset D_-$ for $t < 0$,
(b) $\cap_{t>0}U(t)D_+ = \{0\}$; $\overline{\cap_{t<0}U(t)D_-} = \{0\}$
(c) $\cup_{t<0}U(t)D_+ = \mathcal{E}$, $\overline{\cup_{t>0}U(t)D_-} = \mathcal{E}$.

The existence of the incoming and outgoing subspaces is related to the spectral properties of the operator H. The property (c) from the definition is fulfilled if the spectrum of the operator is continuous [17]. The operator H is self-adjoint. Using the spectral expansion, one can obtain the continuous subspace $\mathcal{E}_c$ by removing the discrete subspace (i.e. eigenspaces) from $\mathcal{E}$. Below we will deal with $H|_{\mathcal{E}_c}$ (we will not introduce a new notation for this operator with the continuous spectrum). The choice of incoming and outgoing subspaces is not unique. For the graph Γ, one can choose the subspace D_+ containing four-component functions vanishing at the ring Γ_0 and satisfying the radiation condition on all leads (infinite edges). Due to the radiating condition, one has only outgoing exponential at the edges. It gives one properties (a), (b) for the subspace. As for (c), it takes place for $\mathcal{E}_c$ due to the self-adjointness of the operator H:

$$\overline{\cup_{t<0}U(t)D_+} = \mathcal{E}_c.$$

The definition of the subspace D_- is analogous. Property (c) for D_- takes the form

$$\overline{\cup_{t>0} U(t) D_-} = \mathcal{E}_c.$$

Let $P_\pm$ be the orthogonal projection of $\mathcal{E}$ onto the orthogonal complement of $D_\pm$. Consider the semigroup $\{Z(t)\}|_{t\geq 0}$ of operators on $\mathcal{E}$ defined by

$$Z(t) = P_+ U(t) P_-, \quad t \geq 0.$$

The square roots of the eigenvalues of the generator B of $Z(t)$ are resonances. The following subspace $K = \mathcal{E} \ominus (D_- \oplus D_+)$ is very important for the construction. The operators $\{Z(t)\}|_{t\geq 0}$ map the subspace K into itself. Lax and Phillips proved the following theorem [17].

Theorem 1.2 *There is a pair of isometric maps $T_\pm : \mathcal{E} \to L_2(\mathbb{R}, N)$ (the outgoing and incoming spectral representations), N is an auxiliary space, having the following properties:*

$$T_\pm U(t) = e^{ikt} T_\pm, \quad T_\pm D_\pm = H^2_\pm(N), \quad T_- D_+ = S H^2_+(N),$$

where $H^2_\pm(N)$ is the Hardy space in the upper (lower) half-plane, the matrix-function S is an inner function in $\mathbb{C}_+$, and

$$K_- = T_- K = H^2_+ \ominus S H^2_+, \quad T_- Z(t)|_K = P_{K_-} e^{ikt} T_-|_{K_-}.$$

Remark 1.3 S is known in functional analysis as characteristic function. At the same time, it is the well-known physical object—the S-matrix.

We deal with the completeness of the system of resonance states (i.e. root vectors of B). There is an interesting relation between the completeness problem and the factorization of the scattering matrix. Namely, as an inner operator-function, S can be represented in the form $S = \Pi\Theta$, where Π is a Blaschke-Potapov product and Θ is a singular inner function [14, 21, 25]. The next theorem shows this relation.

Theorem 1.4 (Completeness Criterion from [21]) *The following statements are equivalent:*

1. *The system of root vectors of the operator B is complete;*
2. *The system of root vectors of the operator B^* is complete;*
3. *S is a Blaschke-Potapov product.*

There is a simple criterion for the absence of the singular inner factor in the case $\dim N < \infty$ (in the general operator case there is no such simple criterion).

Theorem 1.5 ([21]) *Let* $\dim N < \infty$. *The following statements are equivalent:*

1. *S is a Blaschke-Potapov product;*
2.

$$\lim_{r\to 1}\int\limits_{C_r} \ln|\det S(k)|\,\frac{2i}{(k+i)^2}dk = 0, \tag{1.2}$$

where C_r *is the image of* $|\zeta| = r$ *under the inverse Cayley transform.*

Remark 1.6 The theorem in [21] is formulated for the unit disk. We use the Cayley transform which maps the upper half-plane to the unit disk:

$$W(z) = \frac{z-\mathrm{i}}{z+\mathrm{i}},$$

whereas the inverse Cayley transform maps the unit disk to the upper half-plane:

$$w(\zeta) = \mathrm{i}\frac{1+\zeta}{1-\zeta}.$$

One notable property of the Cayley transform is that it injectively maps $\mathbb{R}$ into the unit circle. Another important property we are going to use is that the Cayley transform preserves circles.

The integration curve can be parameterized as

$$C_r = \{\mathsf{R}(r)\mathrm{e}^{it} + i\mathsf{C}(r) \mid t \in [0, 2\pi)\}$$

where

$$\mathsf{C}(r) = \frac{1+r^2}{1-r^2}, \quad \mathsf{R}(r) = \frac{2r}{1-r^2}.$$

It should be noted that $\mathsf{R} \to \infty$ corresponds to $r \to 1$.

For brevity, we define

$$s(k) = |\det S(k)|\,,$$

and after throwing away constants which are irrelevant for convergence, we obtain the final form of the criterion (1.2), which is convenient for us and will be used afterwards:

$$\lim_{r\to 1}\int\limits_0^{2\pi} \frac{\mathsf{R}(r)\ln(s(\mathsf{R}(r)\mathrm{e}^{it} + i\mathsf{C}(r)))}{(\mathsf{R}(r)\mathrm{e}^{it} + i\mathsf{C}(r) + i)^2}\mathrm{e}^{it}dt = 0. \tag{1.3}$$

Our main theorem is as follows.

Theorem 1.7 *The system of resonance states is complete on* $L^2(\Gamma_0) \oplus L^2(\Gamma_0)$.

Remark 1.8 We recall that two components correspond to two directions of spin in respect to magnetic field. It is interesting to compare the result with the corresponding theorems for other operators. If one deals with the same geometric graph with the Shrödinger operator, the Dirac operator or the Landau operator (the Schrödinger operator with a magnetic field) and the Kirchhoff coupling condition at the vertex then there is an incompleteness. But small perturbation of the coupling condition (delta-potential at the vertex) restores the completeness. In our case the spin-orbit interaction plays a role of such perturbation.

2 Scattering for Rashba Hamiltonian

2.1 Rashba Operator

We consider a quantum graph shown in Fig. 1. The quantum ring of radius a is located in the plane xy in the presence of Rashba spin-orbit interaction and a perpendicular magnetic field B. We assume that the magnetic field B is relatively weak. In such case the interaction between the electron spin and the field (Zeeman term) can be treated as a perturbation and the corresponding dimensionless Hamiltonian reads

$$\hat{H} = \left[\left(-i\frac{\partial}{\partial\varphi} + \frac{\omega_{SO}}{2\Omega}\sigma_r - \frac{\Phi}{\Phi_0}\right)^2 - \frac{\omega_{SO}^2}{4\Omega^2}\right] + \frac{\omega_L}{\Omega}\sigma_z,$$

where φ is the azimuthal angle of a point on the ring, Φ denotes the magnetic flux encircled by the ring, $\Phi_0 = \frac{\hbar}{e}$, e is the electron charge, $\omega_{SO} = \frac{\alpha}{\hbar a}$ is the frequency associated with the spin-orbit interaction (with interaction constant α),

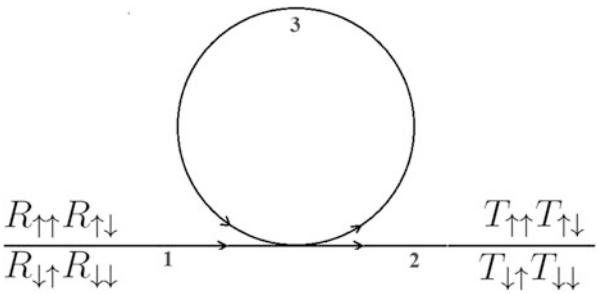


Fig. 1 Task 1 (upper row), Task 2 (lower row); directions of edges are shown

$\hbar\Omega = \dfrac{\hbar^2}{2m^* a^2}$, m^* is the effective mass of the electron. The radial spin operator is given by

$$\sigma_r = \sigma_x \cos\varphi + \sigma_y \sin\varphi = \begin{pmatrix} 0 & \mathrm{e}^{-i\varphi} \\ \mathrm{e}^{i\varphi} & 0 \end{pmatrix}.$$

The constant in Zeeman term is as follows $\omega_L = \frac{g^* eB}{4m}$ with g^* and m being the effective gyromagnetic ratio and the free electron mass, respectively.

2.2 Solution on the Line

To construct the scattering matrix for the graph Γ we solve two scattering problems differing in solutions on the lines.

Task 1:

$$\psi^{\mathrm{e}_1\uparrow}(x) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \mathrm{e}^{ikx} + \begin{pmatrix} R_{\uparrow\uparrow} \\ R_{\uparrow\downarrow} \end{pmatrix} \mathrm{e}^{-ikx}, \qquad \psi^{\mathrm{e}_2\uparrow}(x) = \begin{pmatrix} T_{\uparrow\uparrow} \\ T_{\uparrow\downarrow} \end{pmatrix} \mathrm{e}^{ikx}.$$

Task 2:

$$\psi^{\mathrm{e}_1\downarrow}(x) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \mathrm{e}^{ikx} + \begin{pmatrix} R_{\downarrow\uparrow} \\ R_{\downarrow\downarrow} \end{pmatrix} \mathrm{e}^{-ikx}, \qquad \psi^{\mathrm{e}_2\downarrow}(x) = \begin{pmatrix} T_{\downarrow\uparrow} \\ T_{\downarrow\downarrow} \end{pmatrix} \mathrm{e}^{ikx}.$$

2.3 Solution on the Ring

For the both tasks one has the following form of the solution on the ring:

$$\psi(\varphi) = \sum_{\mu=1,2} \sum_{j=1,2} a_j^\mu \mathrm{e}^{K_j^\mu \varphi} \chi^{(\mu)}(\varphi),$$

where

$$\chi^{(1)}(\varphi) = \frac{1}{\sqrt{2\pi}} \begin{pmatrix} \mathrm{e}^{-i\frac{\varphi}{2}} \cos\frac{\theta}{2} \\ \mathrm{e}^{i\frac{\varphi}{2}} \sin\frac{\theta}{2} \end{pmatrix}, \qquad \chi^{(2)}(\varphi) = \frac{1}{\sqrt{2\pi}} \begin{pmatrix} \mathrm{e}^{-i\frac{\varphi}{2}} \sin\frac{\theta}{2} \\ -\mathrm{e}^{i\frac{\varphi}{2}} \cos\frac{\theta}{2} \end{pmatrix}.$$

Here $\theta = \arctan(-\alpha)$,

$$K_j^\mu = \frac{1}{2} + \frac{\Phi_{AC}^\mu}{2\pi} + \frac{\Phi}{2\pi} + (-1)^{\mu+j+1} \left[\frac{\alpha^2}{4} + \frac{E}{\hbar\Omega} \right]^{1/2}, \quad k = \left(\frac{2m^* E}{\hbar^2} \right)^{1/2}.$$

The value of $j = 1$ and $j = 2$ corresponds to the clockwise and counterclockwise motions of an electron through the ring, respectively. Finally,

$$\Phi_{AC}^{\mu} = -\pi[1 + (-1)^{\mu}\sqrt{1+\alpha^2}]$$

is the so-called Aharonov-Casher (AC) phase.

2.4 Vertex Coupling Conditions

We assume the standard coupling condition at vertex v for the spin-orbit Hamiltonian: the continuity of the wave functions and vanishing net spin current densities [22]. For our graph (see Fig. 1) it has the following form:

$$\sum_{e_j}(-1)^{[e_j]}D_j\psi_j(v) = 0, \quad \psi^{e_1}(v) = \psi^{e_2}(v) = \cdots = \psi^{e_j}(v),$$

where summation is over all edges e_j adjacent to the vertex v, $[e_j] = 0$ for the outgoing edge and $[e_j] = 1$ for the incoming edge. The third edge (see Fig. 1) gives one two terms (outgoing and incoming) with

$$D_j = \left(\frac{\partial}{\partial\varphi} - i\left(\frac{\alpha}{2}\boldsymbol{\sigma}_r - \frac{\Phi}{\Phi_0}\right)\right), \quad \boldsymbol{\sigma}_r = \begin{pmatrix} 0 & \mathrm{e}^{-i\varphi} \\ \mathrm{e}^{i\varphi} & 0 \end{pmatrix}$$

(in our case for the third edge, one has $\phi = 0$ for the outgoing term and $\phi = 2\pi$ for the incoming term). For straight semi-infinite edges (the first and the second edges in Fig. 1) $D_j = \frac{d}{dx}$.

Remark 2.1 From mathematical point of view, these conditions ensure the self-adjointness of the operator on the graph. In physical literature the conditions are known as Griffith conditions (see, e.g., [13, 20]).

These conditions take the following form for our graph.

For Task 1:

$$\psi_1^{\mathrm{e}_1\uparrow}(0) = \psi_1^{\mathrm{e}_2\uparrow}(0) = \psi_1^{\mathrm{e}_3}(0) = \psi_1^{\mathrm{e}_3}(2\pi a),$$

$$\psi_2^{\mathrm{e}_1\uparrow}(0) = \psi_2^{\mathrm{e}_2\uparrow}(0) = \psi_2^{\mathrm{e}_3}(0) = \psi_2^{\mathrm{e}_3}(2\pi a),$$

$$\frac{\partial\psi(0)}{\partial\varphi} - \boldsymbol{\beta}\psi(0) - \frac{\partial\psi(2\pi a)}{\partial\varphi} + \boldsymbol{\beta}\psi(2\pi a) - ik\begin{pmatrix}1\\0\end{pmatrix} + ik\begin{pmatrix}T_{\uparrow\uparrow}\\T_{\uparrow\downarrow}\end{pmatrix} + ik\begin{pmatrix}R_{\uparrow\uparrow}\\R_{\uparrow\downarrow}\end{pmatrix} = 0,$$

where

$$\boldsymbol{\beta} = i\left(\frac{\alpha}{2}\boldsymbol{\sigma}_r - \frac{\Phi}{\Phi_0}\right),$$

ψ_1 and ψ_2 denote, correspondingly, the first and the second component of the vector ψ, $2\pi a$ is the length of the ring.

For Task 2:

$$\psi_1^{\mathrm{e}_1\downarrow}(0) = \psi_1^{\mathrm{e}_2\downarrow}(0) = \psi_1^{\mathrm{e}_3}(0) = \psi_1^{\mathrm{e}_3}(2\pi a),$$

$$\psi_2^{\mathrm{e}_1\downarrow}(0) = \psi_2^{\mathrm{e}_2\downarrow}(0) = \psi_2^{\mathrm{e}_3}(0) = \psi_2^{\mathrm{e}_3}(2\pi a),$$

$$\frac{\partial\psi(0)}{\partial\varphi} - \boldsymbol{\beta}\psi(0) - \frac{\partial\psi(2\pi a)}{\partial\varphi} + \boldsymbol{\beta}\psi(2\pi a) + ik\begin{pmatrix}T_{\downarrow\uparrow}\\ T_{\downarrow\downarrow}\end{pmatrix} - ik\begin{pmatrix}0\\ 1\end{pmatrix} + ik\begin{pmatrix}R_{\downarrow\uparrow}\\ R_{\downarrow\downarrow}\end{pmatrix} = 0.$$

2.5 S-Matrix

Solutions of Task 1 and Task 2 give us the entries of the S-matrix (see [13]):

$$\mathbf{S}(k) = \begin{pmatrix} R_{\uparrow\uparrow} & R_{\downarrow\uparrow} & T_{\uparrow\uparrow} & T_{\downarrow\uparrow} \\ R_{\uparrow\downarrow} & R_{\downarrow\downarrow} & T_{\uparrow\downarrow} & T_{\downarrow\downarrow} \\ T_{\uparrow\uparrow} & T_{\downarrow\uparrow} & R_{\uparrow\uparrow} & R_{\downarrow\uparrow} \\ T_{\uparrow\downarrow} & T_{\downarrow\downarrow} & R_{\uparrow\downarrow} & R_{\downarrow\downarrow} \end{pmatrix}$$

where

$$\begin{aligned} R_{\uparrow\uparrow} &= \varrho^{(1)}\cos^2\frac{\theta}{2} + \varrho^{(2)}\sin^2\frac{\theta}{2} - 1, \\ R_{\uparrow\downarrow} &= (\varrho^{(1)} - \varrho^{(2)})\sin\frac{\theta}{2}\cos\frac{\theta}{2}, \\ R_{\downarrow\uparrow} &= R_{\uparrow\downarrow}, \\ R_{\downarrow\downarrow} &= \varrho^{(1)}\sin^2\frac{\theta}{2} + \varrho^{(2)}\cos^2\frac{\theta}{2} - 1 \end{aligned} \tag{2.1}$$

and

$$\begin{aligned} T_{\uparrow\uparrow} &= \tau^{(1)}\cos^2\frac{\theta}{2} + \tau^{(2)}\sin^2\frac{\theta}{2}, \\ T_{\uparrow\downarrow} &= (\tau^{(1)} - \tau^{(2)})\sin\frac{\theta}{2}\cos\frac{\theta}{2}, \\ T_{\downarrow\uparrow} &= T_{\uparrow\downarrow}, \\ T_{\downarrow\downarrow} &= \tau^{(1)}\sin^2\frac{\theta}{2} + \tau^{(2)}\cos^2\frac{\theta}{2}, \end{aligned} \tag{2.2}$$

where

$$\varrho^{(\mu)} = \frac{4k^2a^2}{\hat{y}^{(\mu)}} iq^{(\mu)} \sin(2q^{(\mu)}\pi), \quad \tau^{(\mu)} = \frac{4ikaq^{(\mu)}}{\hat{y}^{(\mu)}} \sin(2\pi q^{(\mu)}),$$

$$\hat{y}^{(\mu)} = 4ikaq^{(\mu)} \sin(2q^{(\mu)}\pi) - 4(q^{(\mu)})^2 \cos\left(\left((-1)^{\mu+1}\omega - 2\frac{\Phi}{\Phi_0}\right)\pi\right) - 4(q^{(\mu)})^2 \cos(2q^{(\mu)}\pi),$$

$$q^{(\mu)} = \sqrt{q^2 + (-1)^{\mu}\frac{\omega_L}{\Omega}\frac{1}{\omega}}, \quad q = \sqrt{\left(\frac{\omega_{SO}}{2\Omega}\right)^2 + \frac{E}{\hbar\Omega}}, \quad E = \frac{\hbar^2k^2}{2m^*},$$

$$\omega = \sqrt{1 + \frac{\omega_{SO}^2}{\Omega^2}} \quad \omega_L = \frac{g^*eB}{4m}, \quad \omega_{SO} = \frac{\alpha}{\hbar a}.$$

2.6 Completeness

The completeness criterion (see above) is related to the absence or presence of an exponential in k factor in $\det S(k)$. Such a factor gives one a linear growth of $\ln\det S(k)$ for $k \to \infty$ (i.e. $r \to 1$) which leads to the destroying of integral convergence to zero in (1.3), i.e. to incompleteness of resonance states. One more possible reason for destroying of such convergence is the possibility for singularities of $\ln\det S(k)$ to be placed at the integration curve. These singularities are roots and singularities of $\det S(k)$. Taking into account the obtained expressions for the entries of the matrix $S(k)$, one concludes that $\det S(k)$ is a ratio of two analytic functions. Correspondingly, the roots have no accumulation points on the complex plane. It means that there are not more than a finite number of singularities of $\ln\det S(k)$ at the integration curve. Moreover, these singularities are logarithmic, i.e. integrable. Hence, these singularities do not destroy the convergence (1.3). To reveal whether there is an exponential factor in $\det S(k)$, one can look after the behaviour of $\det S(k)$ on the imaginary axis. If there is an exponential factor, one has an exponential growth at one of the imaginary half-axes. One can check that all entries of the S-matrix (see the explicit expressions (2.1) and (2.2)) have no exponential growth. It means that one has the convergence (1.3) in the completeness criterion and, correspondingly, the system of resonance states is complete. This finalizes the proof of Theorem 1.7.

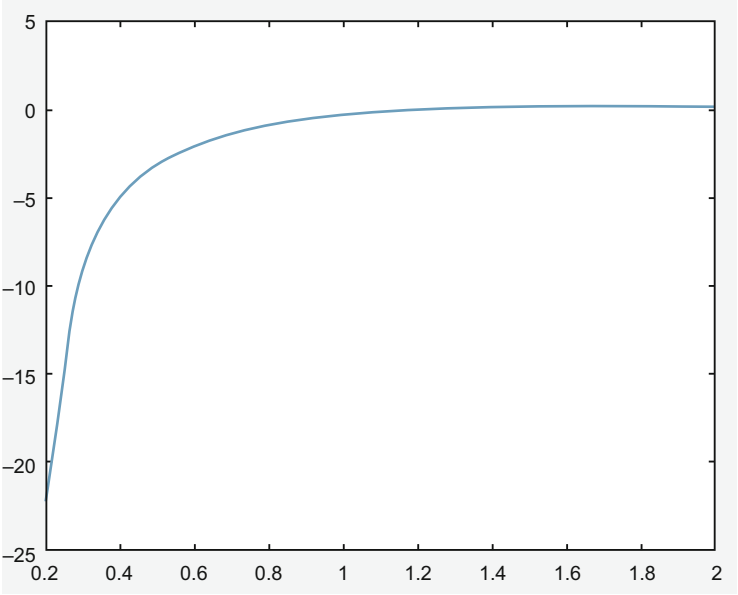


Fig. 2 Behaviour of the integrand of (1.3)

2.7 Numerical Results

Having explicit expression for the S-matrix, one can simply demonstrate the behaviour of the integrand of (1.3) at infinity. It is shown in Fig. 2 for the main part of the integrand:

$$F(k) = \frac{\ln(|\det(S(k))|)}{|k|}.$$

Acknowledgments This work was partially financially supported by the Government of the Russian Federation (grant 08-08), grant 16-11-10330 of Russian Science Foundation and grant 19-31-90164 of Russian Foundation for Basic Researches.

References

1. V.M. Adamyan, D.Z. Arov, On a class of scattering operators and characteristic operator-functions of contractions. Dokl. Akad. Nauk SSSR **160**, 9–12 (1965) (in Russian)
2. A. Aslanyan, L. Parnovski, D. Vassiliev, Complex resonances in acoustic waveguides. Q. J. Mech. Appl. Math. **53**, 429–447 (2000)
3. G. Berkolaiko, P. Kuchment, *Introduction to Quantum Graphs* (AMS, Providence, 2012)

4. I.V. Blinova, I.Y. Popov, Quantum graph with the Dirac operator and resonance states completeness. Oper. Theor. Adv. Appl. **268**, 111–124 (2018)
5. A. Chatterjee, I.Y. Popov, M.O. Smolkina, Persistent current in a chain of two Holstein–Hubbard rings in the presence of Rashba spin-orbit interaction. Nanosyst. Phys. Chem. Math. **10**, 50–62 (2019)
6. E. Dehghana, D.S. Khoshnoud, A.S. Naeimi, Logical spin-filtering in a triangular network of quantum nanorings with a Rashba spin-orbit interaction. Physica B **529**, 21–26 (2018)
7. P. Duclos, P. Exner, B. Meller, Open quantum dots: resonances from perturbed symmetry and bound states in strong magnetic fields. Rep. Math. Phys. **47**, 253–267 (2001)
8. J. Edward, On the resonances of the Laplacian on waveguides. J. Math. Anal. Appl. **272**, 89–116 (2002)
9. P. Exner, P. Keating, P. Kuchment, T. Sunada, A. Teplyaev (eds.), *Analysis on Graphs and Its Applications* (AMS, Providence, 2008)
10. P. Exner, V. Lotoreichik, M. Tater, On resonances and bound states of Smilansky Hamiltonian. Nanosyst. Phys. Chem. Math. **7**, 789–802 (2016)
11. D.A. Gerasimov, I.Y. Popov, Completeness of resonance states for quantum graph with two semi-infinite edges. Complex Var. Elliptic Equ. **63**, 996–1010 (2018)
12. D. Gerasimov, I. Popov, I. Blinova, A. Popov, Incompleteness of resonance states for quantum ring with two semi-infinite edges. Anal. Math. Phys. **9**, 1287–1302 (2019)
13. O. Kálmán, P. Földi, M.G. Benedict, F.M. Peeters, Magnetoconductance of rectangular arrays of quantum rings. Phys. Rev. B **78**, 125306 (2008)
14. S.V. Khrushchev, N.K. Nikol'skii, B.S. Pavlov, Unconditional bases of exponentials and of reproducing kernels, in *Complex Analysis and Spectral Theory (Leningrad, 1979/1980)*, vol. 864. Lecture Notes in Mathematics (Springer, Berlin, 1981), pp. 214–335
15. V.K. Kozin, I.V. Iorsh, O.V. Kibis, I.A. Shelykh, Quantum ring with the Rashba spin-orbit interaction in the regime of strong light-matter coupling. Phys. Rev. B **97**, 155434 (2018)
16. V.V. Kudryashov, A.V. Baran, Rashba spin-orbit interaction in a circular quantum ring in the presence of a magnetic field. Nonlinear Phenom. Complex Syst. Minsk **14**(1), 89–95 (2011)
17. P.D. Lax, R.S. Phillips, *Scattering Theory* (Academic Press, New York, 1967)
18. J. Lipovsky, Quantum graphs and their resonance properties. Acta Physica Slovaca **66**, 265–363 (2016)
19. A.S. Naeimi, L. Eslami, M. Esmaeilzadeh, A wide range of energy spin-filtering in a Rashba quantum ring using S-matrix method. J. Appl. Phys. **113**, 044316 (2013)
20. A.S. Naeimi, L. Eslami, M. Esmaeilzadeh, M.R. Abolhassani, Spin transport properties in a double quantum ring with Rashba spin-orbit interaction. J. Appl. Phys. **113**, 014303 (2013)
21. N. Nikol'skii, *Treatise on the Shift Operator: Spectral Function Theory* (Springer, Berlin, 1986)
22. K. Pankrashkin, Localization effects in a periodic quantum graph with magnetic field and spin-orbit interaction. J. Math. Phys. **47**, 112105 (2006)
23. I.Y. Popov, A.I. Popov, Quantum dot with attached wires: Resonant states completeness. Rep. Math. Phys. **80**, 1–10 (2017)
24. I.Y. Popov, A.I. Popov, Line with attached segment as a model of Helmholtz resonator: resonant states completeness. J. King Saud Univ. Sci. **29**, 133–136 (2017)
25. B. Sz-Nagy, C. Foias, H. Bercovici, L. Kerchy, *Harmonic Analysis of Operators on Hilbert Space*, 2nd edn. (Springer, Berlin, 2010)

Tau Functions Associated with Linear Systems

Gordon Blower and Samantha L. Newsham

Abstract Let $(-A, B, C)$ be a linear system in continuous time $t > 0$ with input and output space $\mathbb{C}$ and state space H. The function $\phi_{(x)}(t) = Ce^{-(t+2x)A}B$ determines a Hankel integral operator $\Gamma_{\phi_{(x)}}$ on $L^2((0,\infty);\mathbb{C})$; if $\Gamma_{\phi_{(x)}}$ is trace class, then the Fredholm determinant $\tau(x) = \det(I + \Gamma_{\phi_{(x)}})$ defines the tau function of $(-A, B, C)$. Such tau functions arise in Tracy and Widom's theory of matrix models, where they describe the fundamental probability distributions of random matrix theory. Dyson considered such tau functions in the inverse spectral problem for Schrödinger's equation $-f'' + uf = \lambda f$, and derived the formula for the potential $u(x) = -2\frac{d^2}{dx^2}\log\tau(x)$ in the self-adjoint scattering case (Commun Math Phys 47:171–183, 1976). This paper introduces a operator function R_x that satisfies Lyapunov's equation $\frac{dR_x}{dx} = -AR_x - R_xA$ and $\tau(x) = \det(I + R_x)$, without assumptions of self-adjointness. When $-A$ is sectorial, and B, C are Hilbert–Schmidt, there exists a non-commutative differential ring $\mathcal{A}$ of operators in H and a differential ring homomorphism $\lfloor\ \rfloor : \mathcal{A} \to \mathbb{C}[u, u', \dots]$ such that $u = -4\lfloor A\rfloor$, which extends the multiplication rules for Hankel operators considered by Pöppe and McKean (Cent Eur J Math 9:205–243, 2011).

Keywords Integrable systems · Fredholm determinant · Inverse scattering

Mathematics Subject Classification (2010) Primary 47B3; Secondary 5, 34B27

1 Introduction

This paper is concerned with Fredholm determinants which arise in the theory of linear systems and their application to inverse spectral problem for Schrödinger's equation. For $\phi \in L^2((0,\infty);\mathbb{R})$, the Hankel integral operator corresponding to ϕ

G. Blower (✉) · S. L. Newsham
Department of Mathematics and Statistics, Lancaster University, Lancaster, UK
e-mail: g.blower@lancaster.ac.uk

M. A. Bastos et al. (eds.), *Operator Theory, Functional Analysis and Applications*,
Operator Theory: Advances and Applications 282,
https://doi.org/10.1007/978-3-030-51945-2_5

is Γ_ϕ where

$$\Gamma_\phi f(x) = \int_0^\infty \phi(x+y) f(y)\, dy \qquad (f \in L^2((0,\infty);\mathbb{C}).$$

Using the Laguerre system of orthogonal functions as in [30], one can express Γ_ϕ as a matrix $[\gamma_{j+k}]_{j,k=1}^\infty$ on ℓ^2, which has the characteristic shape of a Hankel matrix, and one can establish criteria for the operator to be bounded on $L^2((0,\infty);\mathbb{C})$. Megretskii et al. [25] determined the possible spectrum and spectral multiplicity function that can arise from a bounded and self-adjoint Hankel operator. Thus they characterized the class of bounded self-adjoint Hankel operators up to unitary equivalence. Their method involved introducing suitable linear systems on a state space H, and this motivated the approach of our paper.

Following earlier works by Faddeev and others in the Russian literature, Dyson [8] considered the inverse spectral problem for Schrödinger's equation $-f''+uf = \lambda f$, for $u \in C^2(\mathbb{R};\mathbb{R})$ that decays rapidly as $x \to \pm\infty$. From the asymptotic solutions, he introduced a scattering function ϕ, considered the translations $\phi_{(x)}(y) = \phi(y+2x)$, and established connections with eigenvalue distributions in random matrix theory which are described in [37]. He showed that the potential can be recovered from the scattering data by means of the formula

$$u(x) = -2\frac{d^2}{dx^2} \log\det(I + \Gamma_{\phi_{(x)}}), \tag{1.1}$$

These results were developed further by Ercolani and McKean [10] and others [13, 38, 39] to describe the inverse spectral problem for self-adjoint Schrödinger operators on $\mathbb{R}$. Grudsky and Rybkin [17] describes the inverse scattering theory of the KdV equation in terms of Hankel and Toeplitz operators. The latter paper uses Sarason's algebra $H^\infty + C$ on the unit disc to describe compact Hankel operators. In the current paper, we use Hankel operators within the setting of linear systems in continuous time.

Remarkably, some of the methods of inverse scattering theory do not really need self-adjointness. However, a significant obstacle in this approach is that Hankel operators do not have a natural product structure, so it is unclear as to how one can fully exploit the multiplicative properties of determinants. This paper seeks to address this issue, by realizing Hankel operators from linear systems, and then introducing algebras of operators on state space that reflect the properties of Hankel operators and their Fredholm determinants. As in [25], the Lyapunov differential equation is fundamental to the development of the theory.

Definition 1.1 (Lyapunov Equation)

(i) Let H be a complex Hilbert space, known as the state space, and $\mathcal{L}(H)$ the algebra of bounded linear operators on H with the usual operator norm. Let $(e^{-tA})_{t\geq 0}$ be a strongly continuous (C_0) semigroup of bounded linear operators on H such that $\|e^{-tA}\|_{\mathcal{L}(H)} \leq M$ for all $t \geq 0$ and some $M < \infty$. Let $\mathcal{D}(A)$

be the domain of the generator $-A$ so that $\mathcal{D}(A)$ is itself a Hilbert space for the graph norm

$$\|\xi\|^2_{\mathcal{D}(A)} = \|\xi\|^2_H + \|A\xi\|^2_H,$$

and let $A^\dagger$ be the adjoint of A. Let $R : (0, \infty) \to \mathcal{L}(H)$ be a differentiable function. The Lyapunov equation is

$$-\frac{dR_z}{dz} = AR_z + R_z A \qquad (z > 0), \tag{1.2}$$

where the right-hand side is to be interpreted as a bounded bilinear form on $\mathcal{D}(A) \times \mathcal{D}(A^\dagger)$. (This is a modified form of the version in [30, p. 502].)

(ii) *(Operator ideals).* Let $\mathcal{L}^2(H)$ be the space of Hilbert–Schmidt operators on H, and $\mathcal{L}^1(H)$ be the space of trace class operators on H, so $\mathcal{L}^1(H) = \{T : T = VW; V, W \in \mathcal{L}^2(H)\}$ and let det be the Fredholm determinant defined on $\{I + T : T \in \mathcal{L}^1(H)\}$; see [24].

Definition 1.2

(i) *(Linear system).* Let H_0 be a complex separable Hilbert space which serves as the input and output spaces; let $B : H_0 \to H$ and $C : H \to H_0$ be bounded linear operators. The continuous-time linear system $(-A, B, C)$ is

$$\frac{dX}{dt} = -AX + BU, \quad Y = CX.$$

(ii) *(Scattering function).* The scattering function is $\phi(x) = Ce^{-xA}B$, which is a bounded and weakly continuous function $\phi : (0, \infty) \to \mathcal{L}(H_0)$. The terminology is justified by [10, p. 493]. In control theory, the transfer function is the Laplace transform of ϕ; see [30, p. 467].

(iii) *(Hankel operator).* Suppose that $\phi \in L^2((0, \infty); \mathcal{L}(H_0))$. Then the corresponding Hankel operator is Γ_ϕ on $L^2((0, \infty); H_0)$, where

$$\Gamma_\phi f(x) = \int_0^\infty \phi(x + y) f(y)\, dy;$$

see [28, 30] for boundedness criteria.

Definition 1.3 (Admissible Linear System) Let $(-A, B, C)$ be a linear system as above; suppose furthermore that the observability operator $\Theta_0 : L^2((0, \infty); H_0) \to H$ is bounded, where

$$\Theta_0 f = \int_0^\infty e^{-sA^\dagger} C^\dagger f(s)\, ds; \tag{1.3}$$

suppose that the controllability operator $\Xi_0 : L^2((0,\infty); H_0) \to H$ is also bounded, where

$$\Xi_0 f = \int_0^\infty e^{-sA} B f(s)\, ds.$$

(i) Then $(-A, B, C)$ is an admissible linear system. See [30, p. 469].
(ii) Suppose furthermore that Θ_0 and Ξ_0 belong to the ideal $\mathcal{L}^2$ of Hilbert–Schmidt operators. Then we say that $(-A, B, C)$ is (2, 2)-admissible.

The scattering map associates to any (2, 2) admissible linear system $(-A, B, C)$ the corresponding scattering function $\phi(x) = Ce^{-xA}B$. The inverse scattering problem involves recovering data about u from ϕ, as in (1.1). In Sect. 2 of this paper, we analyze the existence and uniqueness problem for the Lyapunov equation, and show that for any (2, 2) admissible linear system, the operator, as in [1, 5],

$$R_x = \int_x^\infty e^{-tA} BC e^{-tA}\, dt$$

is trace class and gives the unique solution to (1.2) with the initial condition

$$\left(\frac{dR_x}{dx}\right)_{x=0} = -AR_0 - R_0 A = -BC. \tag{1.4}$$

Also, $R_x \in \mathcal{L}^1(H)$ and the Fredholm determinant satisfies

$$\det(I + \lambda R_x) = \det(I + \lambda \Gamma_{\phi_{(x)}}) \qquad (x > 0, \lambda \in \mathbb{C}). \tag{1.5}$$

Definition 1.4 (Tau Function) Given a (2, 2) admissible linear system $(-A, B, C)$, we define

$$\tau(x) = \det(I + R_x).$$

Using this general definition of τ, we can unify several results from the scattering theory of ordinary differential equations. Under circumstances discussed in [17] and [34], this becomes the well-known Hitota tau function of soliton theory. Such tau functions are also strongly analogous to the tau functions introduced by Miwa et al. [26] to describe the isomonodromy of rational differential equations and generalize classical results on theta functions. The connection between Fredholm determinants and rational differential equations is further described in [11] and [37]; see also [20].

The Gelfand–Levitan–Marchenko equation [12] provides the linkage between ϕ and u via R_x. Consider

$$T(x, y) + \Phi(x + y) + \mu \int_x^\infty T(x, z)\Phi(z + y)\, dz = 0 \qquad (0 < x < y) \tag{1.6}$$

where $T(x, y)$ and $\Phi(x + y)$ are $m \times m$ matrices with scalar entries. In the context of $(-A, B, C)$ we assume that $\Phi(x) = Ce^{-xA}B$ is known and aim to find $T(x, y)$. In section two, we use R_x to construct solutions to the associated Gelfand–Levitan equation (1.6), and introduce a potential

$$u(x) = -2\frac{d^2}{dx^2}\log\det(I + R_x).$$

In section three, we obtain a differential equation linking $\Phi(x)$ to $u(x)$. In examples of interest in scattering theory, one can calculate $\det(I + \lambda R_x)$ more easily than the Hankel determinant of $\Gamma_{\Phi(x)}$ directly [10], since R_x has additional properties that originate from Lyapunov's equation. In section four, we introduce a differential algebra of operators on the state space, and a homomorphism to the differential algebra $\mathbb{C}[u, u', \dots]$ that is generated by the potential. In section five, we describe the connection between this algebra and the stationary KdV hierarchy. There is a fundamental connection between theta functions and equations of KdV and KP type; see [27].

2 τ Functions in Terms of Lyapunov's Equation and the Gelfand–Levitan Equation

The following section proves existence and uniqueness of solutions of the Lyapunov equation (1.2), in a style suggested by Peller [30, p. 503]. Peller discusses scattering functions that produce bounded self-adjoint Hankel operators Γ_ϕ, and their realization in terms of continuous time linear systems. He observes that in some cases one needs a bounded semigroup with unbounded generator $(-A)$. We prove the uniqueness results for bounded and strongly continuous semigroups, then specialize to holomorphic semigroups. The main application is to the Gelfand–Levitan equation (1.6), and associated determinants.

Proposition 2.1 *Let $(e^{-tA})_{t\geq 0}$ be a strongly continuous and weakly asymptotically stable semigroup on a complex Hilbert space H, so $e^{-tA}f \to 0$ weakly as $t \to \infty$ for all $f \in H$. Then*

(i) *$S_t : R \mapsto e^{-tA}Re^{-tA}$ for $t \geq 0$ defines a strongly continuous semigroup on $\mathcal{L}^1(H)$, which has generator $(-L)$, with dense domain of definition $\mathcal{D}(L)$ such that*

$$L(R) = AR + RA \qquad (R \in \mathcal{D}(L)).$$

(ii) *The linear operator* $L : \mathcal{D}(L) \to \mathcal{L}^1(H)$ *is injective, and for each* $R_0 \in \mathcal{D}(L)$ *with* $L(R_0) = X$, *there exists a weakly convergent integral*

$$R_0 = \int_0^\infty e^{-tA} X e^{-tA}\, dt. \tag{2.1}$$

(iii) *Suppose moreover that* $\|e^{-t_0 A}\|_{\mathcal{L}(H)} < 1$ *for some* $t_0 > 0$. *Then* $L : \mathcal{D}(L) \to \mathcal{L}^1(H)$ *is surjective, the integral* (2.1) *converges absolutely in* $\mathcal{L}^1(H)$ *and* R_0 *gives the unique solution to* $AR_0 + R_0 A = X$.

Proof

(i) First observe that by the uniform boundedness theorem, there exists M such that $\|e^{-tA}\|_{\mathcal{L}(H)} \le M$ for all $t \ge 0$, so $(e^{-tA})_{t\ge 0}$ is uniformly bounded. Also, the adjoint semigroup $(e^{-tA^\dagger})_{t\ge 0}$ is also strongly continuous and uniformly bounded, so A and $A^\dagger$ have dense domains $\mathcal{D}(A)$ and $\mathcal{D}(A^\dagger)$ in H. Now $\mathcal{L}^1(H) = H \hat{\otimes} H$, the projective tensor product, so for all $X \in \mathcal{L}^1(H)$, there exists a nuclear decomposition

$$X = \sum_{j=1}^\infty B_j C_j$$

where $B_j, C_j \in H$ satisfy $\|X\|_{\mathcal{L}^1(H)} = \sum_{j=1}^\infty \|B_j\|_H \|C_j\|_H$. Then

$$S_t(X) - X = \sum_{j=1}^\infty (e^{-tA} B_j C_j e^{-tA} - B_j C_j e^{-tA}) + \sum_{j=1}^\infty (B_j C_j e^{-tA} - B_j C_j)$$

where (e^{-tA}) is bounded, $\|e^{-tA} B_j - B_j\|_H \to 0$ and $\|e^{-tA^\dagger} C_j - C_j\|_H \to 0$ as $t \to 0+$; so $\|S_t(X) - X\|_{\mathcal{L}^1(H)} \to 0$ as $t \to 0+$; so $(S_t)_{t\ge 0}$ is strongly continuous on $\mathcal{L}^1(H)$. By the Hille–Yoshida theorem [15, p. 16], there exists a dense linear subspace $\mathcal{D}(L)$ of $\mathcal{L}^1(H)$ such that $S_t(R)$ is differentiable at $t = 0+$ for all $R \in \mathcal{D}(L)$, and $(d/dt)_{t=0+} S_t(R) = -AR - RA$, so the generator is $(-L)$, where $L(R) = AR + RA$.

(ii) Certainly $\mathcal{D}(L)$ contains $\mathcal{D}(A^\dagger) \hat{\otimes} \mathcal{D}(A)$ in $\mathcal{L}^1(H) = H \hat{\otimes} H$. Choosing $f \in \mathcal{D}(A)$ and $g \in \mathcal{D}(A^\dagger)$, we find that

$$\frac{d}{dt} \langle e^{-tA} R_0 e^{-tA} f, g \rangle = -\langle e^{-tA} (AR_0 + R_0 A) e^{-tA} f, g \rangle = -\langle e^{-tA} X e^{-tA} f, g \rangle$$

a continuous function of $t > 0$; so integrating we obtain

$$\langle R_0 f, g \rangle - \langle e^{-sA} R_0 e^{-sA} f, g \rangle = \int_0^s \langle e^{-tA} X e^{-tA} f, g \rangle\, dt.$$

We extend this identity to all $f, g \in H$ by joint continuity; then we let $s \to \infty$ and observe that $R_0 : H \to H$ is trace class and hence is completely continuous, hence R_0 maps the weakly null family $(e^{-sA} f)_{s\to\infty}$ to the norm convergent family $(R_0 e^{-sA} f)_{s\to\infty}$, so $\langle e^{-sA} R_0 e^{-sA} f, g\rangle \to 0$ as $s \to \infty$; hence we have a weakly convergent improper integral

$$\big\langle R_0 f, g\big\rangle = \lim_{s\to\infty} \int_0^s \big\langle e^{-tA} X e^{-tA} f, g\big\rangle\, dt \qquad (f, g \in H).$$

(iii) The function $t \mapsto e^{-tA} X e^{-tA}$ takes values in the separable space $\mathcal{L}^1(H)$ and is weakly continuous, hence strongly measurable, by Pettis's theorem. By considering the spectral radius, Engel and Nagel [9] show that there exist $\delta > 0$ and $M_\delta > 0$ such that $\|e^{-tA}\|_{\mathcal{L}(H)} \le M_\delta e^{-\delta t}$ for all $t \ge 0$; hence (2.1) converges as a Bochner–Lebesgue integral with

$$\|R_x\|_{\mathcal{L}^1(H)} \le \int_x^\infty M_\delta^2 \|X\|_{\mathcal{L}^1(H)} e^{-2\delta t}\, dt \le \frac{M_\delta^2}{2\delta} \|X\|_{\mathcal{L}^1(H)} e^{-2\delta x}.$$

Furthermore, A is a closed linear operator and satisfies

$$\begin{aligned} A\int_x^s e^{-tA} X e^{-tA}\, dt + \int_x^s e^{-tA} X e^{-tA}\, dt A &= \int_x^s -\frac{d}{dt}\big(e^{-tA} X e^{-tA}\big)\, dt \\ &= e^{-xA} X e^{-xA} - e^{-sA} X e^{-sA} \\ &\to e^{-xA} X e^{-xA} \end{aligned}$$

as $s \to \infty$ where

$$\int_x^s e^{-tA} X e^{-tA} dt \to R_x;$$

so $AR_x + R_x A = e^{-xA} X e^{-xA}$ for all $x \ge 0$. We deduce that $x \mapsto R_x$ is a differentiable function from $(0, \infty)$ to $\mathcal{L}^1(H)$ and that the modified Lyapunov equation (1.2) holds. □

The hypotheses (i) and (ii) are symmetrical under the adjoint $(A, R_0) \mapsto (A^\dagger, R_0^\dagger)$; however, the hypothesis (iii) is rather stringent, and in many applications one only needs existence of the integral (2.1).

Definition 2.2 ((2, 2) Admissible Linear Systems)

(i) Let H be a complex Hilbert space and let $\Sigma = (-A, B, C)$ be a linear system with state space H. Suppose that the integral

$$W_c = \int_0^\infty e^{-tA} B B^\dagger e^{-tA^\dagger}\, dt$$

converges weakly and defines a bounded linear operator on H; then W_c is the controllability Gramian. Suppose further that the integral

$$W_o = \int_0^\infty e^{-tA^\dagger} C^\dagger C e^{-tA}\, dt$$

converges weakly and defines a bounded linear operator on H; then W_o is the observability Gramian.

(ii) Then as in [5, p. 318] we define R_x to be the bounded linear operator on H determined by the weakly convergent integral

$$R_x = \int_x^\infty e^{-tA} BC e^{-tA}\, dt. \tag{2.2}$$

(iii) Then Σ satisfying (i) is said to be balanced if $W_c = W_o$ and $\ker(W_c) = 0$; see [30, p. 499].

(iv) Also, Σ satisfying (i) is said to be $(2, 2)$ admissible if W_c and W_o are trace class, or equivalently Θ_0 and Ξ_0 are Hilbert-Schmidt; see [5].

(v) We introduce the scattering function $\phi(t) = Ce^{-tA}B$ and the shifted scattering function $\phi_{(x)}(t) = \phi(t + 2x)$ for $x, t > 0$.

(vi) *(Sectorial operator).* For $0 < \theta \leq \pi$, we introduce the sector

$$S_\theta = \{z \in \mathbb{C} \setminus \{0\} : |\arg z| < \theta\}.$$

A closed and densely defined linear operator $-A$ is sectorial [9, 15] if there exists $\pi/2 < \theta < \pi$ such that S_θ is contained in the resolvent set of $-A$ and $|\lambda| \|(\lambda I + A)^{-1}\|_{\mathcal{L}(H)} \leq M$ for all $\lambda \in S_\theta$. Let $\mathcal{D}(A)$ be the domain of A and $\mathcal{D}(A^\infty) = \cap_{n=0}^\infty \mathcal{D}(A^n)$. See [15, p. 37].

(vii) For $\pi/2 < \delta < \pi$, we introduce $X_\delta = \{\zeta \in S_\delta : -\zeta \in S_\delta\}$ which is an open set, symmetrical about $i\mathbb{R}$ and bounded by lines passing through 0.

Theorem 2.3 *Let $(-A, B, C)$ be a linear system such that $\|e^{-t_0 A}\|_{\mathcal{L}(H)} < 1$ for some $t_0 > 0$, and that B and C are Hilbert–Schmidt operators such that $\|B\|_{\mathcal{L}^2(H_0;H)} \|C\|_{\mathcal{L}^2(H;H_0)} \leq 1$. Suppose further that $-A$ is sectorial on S_θ for some $\pi/2 < \theta < \pi$.*

(i) *Then $(-A, B, C)$ is $(2, 2)$-admissible, so the trace class operators $(R_x)_{x>0}$ give the solution to Lyapunov's equation (1.2) for $x > 0$ that satisfies the initial condition (1.4), and the solution to (1.2) with (1.4) is unique.*

(ii) *The function $\tau(x) = \det(I + R_x)$ is differentiable for $x \in (0, \infty)$.*

(iii) *Then R_z extends to a holomorphic function that satisfies (1.2) on $S_{\theta-\pi/2}$, and $R_z \to 0$ as $z \to \infty$ in $S_{\theta-\varepsilon-\pi/2}$ for all $0 < \varepsilon < \theta - \pi/2$.*

Proof

(i) Since $BC \in \mathcal{L}^1(H)$, the integrand of (2.2) takes values in $\mathcal{L}^1(H)$, and we can apply Proposition 2.1(iii) to $X = BC$.

(ii) The Fredholm determinant $R \mapsto \det(I+R)$ is a continuous function on $\mathcal{L}^1(H)$. Also the integral

$$R_x = \int_x^\infty e^{-tA} BC e^{-tA}\, dt$$

belongs to $\mathcal{D}(L)$ and gives a differentiable function of $x > 0$ with values in $\mathcal{L}^1(H)$.

(iii) By classical results of Hille [15, p. 34], $(e^{-zA})_{z\in S_{\theta-\pi/2}}$ defines an analytic semigroup on $S_{\theta-\pi/2}$, bounded on S_ν for all $0 < \nu < \theta-\pi/2$, so we can define $R_z = e^{-zA} R_0 e^{-zA}$ and obtain an analytic solution to Lyapunov's equation. For all $0 < \varepsilon < \theta-\pi/2$, there exists M'_ε such that $\|e^{-zA}\|_{\mathcal{L}(H)} \le M'_\varepsilon$ for all $z \in S_\delta$ where $\delta = \theta - \varepsilon - \pi/2$. Now for $z \in S_{\delta/2}$, we write $z = x/2 + (x/2 + iy)$ with $x/2 + iy \in S_\delta$ and use the bound

$$\|e^{-zA}\|_{\mathcal{L}(H)} \le \|e^{-xA/2}\|_{\mathcal{L}(H)} \|e^{-(x/2+iy)A}\|_{\mathcal{L}(H)}$$

to obtain $\|e^{-zA}\|_{\mathcal{L}(H)} \le M'^2_\varepsilon \|e^{-t_0 A}\|^{x/(4t_0)}_{\mathcal{L}(H)}$, so $\|e^{-zA}\|_{\mathcal{L}(H)} \to 0$ exponentially fast as $z \to \infty$ in the sector $S_{\delta/2}$. Hence R_z is holomorphic and bounded on $S_{(\theta-\varepsilon-\pi/2)}$ and by (2.2), $R_z \to 0$ as $z \to \infty$ in $S_{(\theta-\varepsilon-\pi/2)/2}$. □

Example

(i) Let $\Delta = -d^2/dx^2$ be the usual Laplace operator which is essentially self-adjoint and non-negative on $C_c^\infty(\mathbb{R};\mathbb{C})$ in $L^2(\mathbb{R};\mathbb{C})$. We introduce $A = \sqrt{I+\Delta}$ which is given by the Fourier multiplier $\mathcal{F}Af(\xi) = \sqrt{1+\xi^2}\mathcal{F}f(\xi)$. Then (e^{-zA}) and (e^{-zA^2}) give bounded holomorphic semigroups on H, as in Theorem 2.3, on the right half-plane $\{z \in \mathbb{C} : \Re z \ge 0\}$, which is the closure of $S_{\pi/2}$. On the imaginary axis, we have unitary groups (e^{itA}) and (e^{-itA^2}). By classical results from wave equations, we can write $e^{itA} + e^{-itA} = 2\cos(tA)$ where $u(x,t) = \cos(tA) f(x)$ for $f \in C_c^\infty(\mathbb{R};\mathbb{C})$ is given by

$$u(x,t) = \frac{1}{2}\big(f(x+t) + f(x-t)\big) + \frac{t}{2}\int_{x-t}^{x+t} f(y) \frac{J_0'(\sqrt{t^2-(x-s)^2})}{\sqrt{t^2-(x-s)^2}} ds,$$

where J_0 is Bessel's function of the first kind of order zero, and u satisfies

$$\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial t^2} = u(x,t), \quad u(x,0) = f(x), \quad \frac{\partial u}{\partial t}(x,0) = 0.$$

See [15, p. 121]. Note that $(\exp(t(iA)^{2j-1}))$ gives a unitary group on H for $j = 0, 1, 2, \ldots$. This can be used to deform the linear system in the sense of Proposition 2.5(iii). Unitary deformation groups for tau functions are considered in [26]

(ii) In section 4 of [5], we introduced linear systems to describe Schrödinger's equation when the potential is smooth and localized. In [17], the authors obtain detailed results about the corresponding Hankel operator.

Definition 2.4 (Block Hankel Operators)

(i) Say that $\Gamma \in \mathcal{L}(H)$ is block Hankel if there exists $1 \leq m < \infty$ such that Γ is unitarily equivalent to the block matrix $[A_{j+k-2}]_{j,k=1}^{\infty}$ on $\ell^2(\mathbb{C}^m)$ where $A_j \in \mathbb{C}^{m\times m}$ for $j = 0, 1, \ldots$.

(ii) Let $(-A, B, C)$ be a $(2, 2)$ admissible linear system with input and output space H_0, where the dimension of H_0 over $\mathbb{C}$ is $m < \infty$. Then m is the number of outputs of the system, and systems with finite $m > 1$ are known as MIMO for multiple input, multiple output, and give rise to block Hankel operators with $\Phi(x) = Ce^{-xA}B$.

(iii) The Gelfand–Levitan integral equation for $(-A, B, C)$ as in (ii) is (1.6), where $T(x, y)$ and $\Phi(x + y)$ are $m \times m$ matrices with scalar entries, and $\mu \in \mathbb{C}$. We proceed to obtain a solution.

Proposition 2.5

(i) *In the notation of Theorem 2.3, there exists $x_0 > 0$ such that*

$$T_\mu(x, y) = -Ce^{-xA}(I + \mu R_x)^{-1}e^{-yA}B$$

satisfies the integral equation (1.6) for $x_0 < x < y$ and $|\mu| < 1$.

(ii) *The determinant satisfies* $\det(I + \mu R_x) = \det(I + \mu\Gamma_{\Phi_{(x)}})$ *and*

$$\mu \operatorname{trace} T_\mu(x, x) = \frac{d}{dx} \log\det(I + \mu R_x).$$

(iii) *Suppose that $t \mapsto U(t)$ is a continuous function $[0, 1] \to \mathcal{L}(H)$ such that $U(t)A = AU(t)$ and $\|U(t)\|_{\mathcal{L}(H)} \leq 1$. Then there is a family of $(2, 2)$ admissible linear systems*

$$\Sigma(t) = (-A, U(t)B, CU(t)) \qquad (t \in [0, 1]);$$

the corresponding tau function $\tau(x, t)$ is continuous for $(x, t) \in (0, \infty) \times [0, 1]$.

Proof

(i) We choose x_0 so large that $e^{\delta x_0} \geq M_\delta/2\delta$, then by Theorem 2.3(iii), we have $|\mu|\|R_x\|_{\mathcal{L}(H)} < 1$ for $x > x_0$, so $I + \mu R_x$ is invertible. Substituting $T_\mu(x, y)$ into the integral equation (1.6), we obtain

$$Ce^{-(x+y)A}B - Ce^{-xA}(I + \mu R_x)^{-1}e^{-yA}B$$
$$- \mu Ce^{-xA}(I + \mu R_x)^{-1}\int_x^\infty e^{-zA}BCe^{-zA}\,dz e^{-yA}B$$

$$
\begin{aligned}
&= Ce^{-(x+y)A}B - Ce^{-xA}(I+\mu R_x)^{-1}e^{-yA}B \\
&\quad - \mu Ce^{-xA}(I+\mu R_x)^{-1}R_x e^{-yA}B \\
&= 0.
\end{aligned}
$$

(ii) As in (1.3), the operator $\Theta_x : L^2(0,\infty) \to H$ is Hilbert–Schmidt; likewise $\Xi_x : L^2(0,\infty) \to H$ is Hilbert–Schmidt; so $(-A, B, C)$ is $(2,2)$-admissible. Hence $\Gamma_{\Phi_{(x)}} = \Theta_x^\dagger \Xi_x$ and $R_x = \Xi_x \Theta_x^\dagger$ are trace class, $(I+\mu R_x)$ is a holomorphic function of x on some sector S_δ as in Theorem 2.3 and

$$\det(I+\mu R_x) = \det(I+\mu \Xi_x \Theta_x^\dagger) = \det(I+\mu \Theta_x^\dagger \Xi_x) = \det(I+\mu \Gamma_{\Phi_{(x)}}).$$

By the Riesz functional calculus, $(I+\mu R_x)^{-1}$ is meromorphic for x in some S_δ. Correcting a typographic error in [5, p. 324], we rearrange terms and calculate the derivative

$$
\begin{aligned}
\mu T_\mu(x,x) &= -\mu \operatorname{trace}\left(Ce^{-xA}(I+\mu R_x)^{-1}e^{-xA}B\right) \\
&= -\mu \operatorname{trace}\left((I+\mu R_x)^{-1}e^{-xA}BCe^{-xA}\right) \\
&= \mu \operatorname{trace}\left((I+\mu R_x)^{-1}\frac{dR_x}{dx}\right) \\
&= \frac{d}{dx}\operatorname{trace}\log(I+\mu R_x).
\end{aligned}
$$

This identity is proved for $|\mu| < 1$ and extends by analytic continuation to the maximal domain of $T_\mu(x,x)$.

(iii) Since A commutes with $U(t)$, the domain $\mathcal{D}(A)$ is invariant under $U(t)$, and the multiplications $B \mapsto U(t)B$ and $C \mapsto CU(t)$ preserve the hypotheses of Theorem 2.3, so $(-A, U(t)B, CU(t))$ is $(2,2)$ admissible. By commutativity, we have $\tau(x,t) = \det(I + U(t)R_x U(t))$, which depends continuously on (x,t). □

3 The Baker–Akhiezer Function of an Admissible Linear System

In this section, we consider the Darboux addition rule for potentials and analyze the transformation $(-A, B, C) \mapsto (-A, B, -C)$ and the effect on the ratios and derivatives of τ functions. This generalizes [10, section 3.4], and allows us to introduce a version of the Baker–Akhiezer function for a family of linear systems with properties that are similar to the classical case, as presented in [3] and [22].

Definition 3.1 (Baker–Akhiezer Function)

(i) Let $(-A, B, C)$ be as in Theorem 2.3, and let

$$\Sigma_\zeta = (-A, (\zeta I + A)(\zeta I - A)^{-1}B, C) \qquad (\zeta \in \mathbb{C} \cup \{\infty\} \setminus \mathrm{Spec}(A))$$

so that Σ_ζ defines a $(2, 2)$ admissible linear systems for ζ in an open subset of $\mathbb{C} \cup \{\infty\}$ which includes $\{\zeta \in \mathbb{C} : -\zeta \in S_\theta\}$ for some $\pi/2 < \theta < \pi$. We identify Σ_∞ with $(-A, B, C)$, and Σ_0 with $(-A, B, -C)$.

(ii) Let τ_ζ be the tau function of Σ_ζ, and let the Baker–Akhiezer function for the family of linear systems be

$$\psi_\zeta(x) = \frac{\tau_\zeta(x)}{\tau_\infty(x)} \exp\big(\zeta x\big). \tag{3.1}$$

(iii) Let $\tau^*_\zeta(x) = \overline{\tau_{\bar\zeta}(\bar x)}$ as in Schwarz's reflection principle, and let

$$\Sigma^*_\zeta = (-A^\dagger, C^\dagger, B^\dagger(\zeta I + A^\dagger)(\zeta I - A^\dagger)^{-1}) \qquad (\zeta \in \mathbb{C} \cup \{\infty\} \setminus \mathrm{Spec}(A^\dagger))$$

so $\Sigma_\zeta \mapsto \Sigma^*_\zeta$ is an involution, and Σ^*_ζ has tau function τ^*.

The following result introduces a family of solutions of Schrödinger equation corresponding to the Σ_ζ with an addition rule in the style of Darboux.

Proposition 3.2 *Let $(-A, B, C)$ be as in Theorem 2.3.*

(i) *Then for $-\zeta \in S_\theta$, the linear system Σ_ζ is also $(2, 2)$ admissible, and the Baker–Akhiezer function satisfies*

$$-\frac{d^2}{dx^2}\psi_\zeta(x) + u_\infty(x)\psi_\zeta(x) = -\zeta^2\psi_\zeta(x). \tag{3.2}$$

(ii) *There exist $h_j \in C^\infty((0, \infty); \mathbb{C})$ such that there is an asymptotic expansion*

$$\psi_\zeta(x) \asymp e^{\zeta x}\Big(1 + \frac{h_1(x)}{\zeta} + \frac{h_2(x)}{\zeta^2} + \dots\Big)$$

as $\zeta \to \pm i\infty$, and the expansion is uniform for x in compact subsets of $(0, \infty)$.

Proof

(i) For all $\zeta \in \mathbb{C} \setminus \mathrm{Spec}(A)$, there exists $x_0(\zeta)$ such that

$$\|(\zeta I + A)(\zeta I - A)^{-1}R_x\|_{\mathcal{L}^1(H)} < 1$$

for all $x > x_0(\zeta)$, so that $\tau_\zeta(x)$ is continuously differentiable and non-zero as a function of $x \in (x_0(\zeta), \infty)$. In particular, suppose that $\Re\zeta < 0$, then $-\zeta \in S_\theta$ so $\zeta I - A$ is invertible. Using the R function for Σ_ζ, we write

$$\begin{aligned}\frac{\tau_\zeta(x)}{\tau_\infty(x)} &= \frac{\det\big(I+(\zeta I+A)(\zeta I-A)^{-1}R_x\big)}{\det\big(I+R_x\big)}\\ &= \frac{\det\big(I+(\zeta I-A)^{-1}((\zeta I-A)R_x+AR_x+R_xA)\big)}{\det\big(I+R_x\big)}\\ &= \frac{\det\big(I+R_x+(\zeta I-A)^{-1}(AR_x+R_xA)\big)}{\det\big(I+R_x\big)}\end{aligned} \tag{3.3}$$

so that when $AR_x + R_xA$ has rank one, the perturbing term $(\zeta I - A)^{-1}(AR_x + R_xA)$ has rank one; continuing we find

$$\begin{aligned}\frac{\tau_\zeta(x)}{\tau_\infty(x)} &= \det\big(I+(\zeta I-A)^{-1}e^{-xA}BCe^{-xA}(I+R_x)^{-1}\big)\\ &= \det\big(I+Ce^{-xA}(I+R_x)^{-1}(\zeta I-A)^{-1}e^{-xA}B\big)\\ &= 1+Ce^{-xA}(I+R_x)^{-1}(\zeta I-A)^{-1}e^{-xA}B,\end{aligned}$$

since $B : \mathbb{C} \to H$ and $C : H \to \mathbb{C}$ have rank one. Hence

$$\begin{aligned}\psi_\zeta(x) &= \frac{\tau_\zeta(x)}{\tau_\infty(x)}\exp\big(\zeta x\big)\\ &= \exp\big(\zeta x\big)+Ce^{-xA}(I+R_x)^{-1}(\zeta I-A)^{-1}e^{-xA}B\exp\big(\zeta\big)\\ &= \exp\big(\zeta x\big)-\int_x^\infty Ce^{-xA}(I+R_x)^{-1}e^{-yA}B\exp\big(\zeta y\big)\,dy\\ &= \exp\big(\zeta x\big)+\int_x^\infty T(x,y)\exp\big(\zeta y\big)\,dy.\end{aligned}$$

Here T satisfies the Gelfand–Levitan equation, and by integrating by parts, we see that

$$\frac{\partial^2 T}{\partial x^2}-\frac{\partial^2 T}{\partial y^2}=u(x)T(x,y)$$

where $u(x) = -2\frac{d^2}{dx^2}\log\tau(x)$. Then by integrating by parts, we see that ψ_ζ satisfies Schrödinger's equation.

The solutions of the differential equation depend analytically on ζ at those points where the potential depends analytically on ζ; note that $\zeta \mapsto \tau_\zeta(x)$ is holomorphic and non zero for $\|R_x\| < 1$ and $-\zeta \in S_\theta$. Then we continue the solutions analytically to all $-\zeta$ in the sector S_θ, on which $\psi_\zeta(x)$ is holomorphic as a function of ζ for $x > 0$.

(ii) Observe that $X_\theta = S_\theta \cap (-S_\theta)$ contains $i\mathbb{R} \setminus \{0\}$. For $\zeta \in S_\theta \cap (-S_\theta)$, by (i) there exist solutions $\psi_\zeta(x)$ and $\psi_{-\zeta}(x)$ to (3.2). In particular, ψ_{ik} and $\psi_{-ik}(x)$

are solutions for $k > 0$. We integrate by parts repeatedly

$$\begin{aligned} e^{-xA}(\zeta I - A)^{-1} =& e^{-xA}\int_0^\infty e^{\zeta s}e^{-sA}\,ds \\ =& \frac{e^{-xA}}{\zeta} + \frac{Ae^{-xA}}{\zeta^2} + \cdots + \frac{A^{k-1}e^{-xA}}{\zeta^k} \\ &+ \int_0^\infty \frac{A^k e^{-xA}}{\zeta^k} e^{\zeta s}e^{-sA}\,ds, \end{aligned}$$

where the integral converges by the hypothesis of Theorem 2.3. Also, (e^{-zA}) is an analytic semigroup in the sector $S_{\theta-\pi/2}$, so $\mathcal{D}(A^j)$ is a dense linear subspace of H for all $j = 1, 2, \ldots$ and $A^j e^{-xA} \in \mathcal{L}(H)$ and by Cauchy's estimates there exists $C > 0$ such that $\|A^j e^{-xA}\|_{\mathcal{L}(H)} \le Cj!/x^j$ for all $x > 0$. So we can generate an asymptotic expansion of (3.3) with terms

$$h_j(x) = Ce^{-xA}(I + R_x)^{-1}A^{j-1}e^{-xA}B$$

which are bounded on compact subsets of $(0, \infty)$. □

Definition 3.3 (Darboux Transforms) Let $(-A, B, C)$ be a $(2, 2)$ admissible linear system with tau function $\tau_\infty(x; \mu) = \det(I + \mu R_x)$. Define the Darboux transform of $(-A, B, C)$ to be $(-A, B, -C)$ with tau function transform $\tau_0(x; \mu) = \det(I - \mu R_x)$. Let

$$\begin{aligned} v &= \frac{1}{\mu}\frac{d}{dx}\log\frac{\tau_\infty}{\tau_0}, \qquad w = \frac{1}{\mu}\frac{d}{dx}\log\left(\tau_0\tau_\infty\right), \\ u_\infty &= -\frac{2}{\mu^2}\frac{d^2}{dx^2}\log\tau_\infty,\ u_0 = -\frac{2}{\mu^2}\frac{d^2}{dx^2}\log\tau_0. \end{aligned}$$

In the following result, we show how products and quotients of τ functions can be linked by the Gelfand–Levitan equation for 2×2 matrices, and satisfy the identities usually associated with Darboux transforms in the theory of integrable systems. See [23].

Theorem 3.4 *Let $(-A, B, C)$ be a $(2, 2)$-admissible linear system with input and output spaces $\mathbb{C}$, and let $\phi(x) = Ce^{-xA}B$.*

(i) *Then there exists $\delta > 0$ such that for all $\mu \in \mathbb{C}$ such that $|\mu| < \delta$, the Gelfand-Levitan equation (1.6) with*

$$T(x, y) = \begin{bmatrix} W(x, y) & V(x, y) \\ V(x, y) & W(x, y) \end{bmatrix}, \quad \Phi(x + y) = \begin{bmatrix} 0 & \phi(x + y) \\ \phi(x + y) & 0 \end{bmatrix}$$

has a solution such that

$$W(x,x)=\frac{1}{2\mu}\frac{d}{dx}\log\big(\tau_\infty(x;\mu)\tau_0(x;\mu)\big),\qquad V(x,x)=\frac{1}{2\mu}\frac{d}{dx}\log\frac{\tau_\infty(x;\mu)}{\tau_0(x;\mu)}$$

and

$$\frac{1}{2\mu}\frac{d}{dx}W(x,x)=-V(x,x)^2;$$

(ii) *also Toda's equation holds in the form*

$$\tau_0''\tau_\infty-2\tau_0'\tau_\infty'+\tau_0\tau_\infty''=0. \tag{3.4}$$

Proof

(i) Let

$$T_\infty(x,y)=-Ce^{-xA}(I+\mu R_x)^{-1}e^{-yA}B,$$
$$T_0(x,y)=Ce^{-xA}(I-\mu R_x)^{-1}e^{-yA}B$$

and

$$\Phi(x)=\begin{bmatrix}0&\phi(x)\\ \phi(x)&0\end{bmatrix}.$$

Now let

$$T(x,y)=\frac{1}{2}\begin{bmatrix}T_\infty+T_0 & T_\infty-T_0\\ T_\infty-T_0 & T_\infty+T_0\end{bmatrix}$$

so that

$$T(x,y)=-\begin{bmatrix}C&0\\0&C\end{bmatrix}\begin{bmatrix}e^{-xA}&0\\0&e^{-xA}\end{bmatrix}\begin{bmatrix}I&\mu R_x\\ \mu R_x&I\end{bmatrix}^{-1}\begin{bmatrix}e^{-yA}&0\\0&e^{-yA}\end{bmatrix}\begin{bmatrix}0&B\\B&0\end{bmatrix}$$

hence T satisfies the Gelfand–Levitan equation (1.6).

(ii) As in Proposition 2.5,

$$T_\infty(x,x)=\frac{1}{\mu}\frac{d}{dx}\log\tau_\infty(x),\qquad T_0(x,x)=\frac{1}{\mu}\frac{d}{dx}\log\tau_0(x);$$

hence (3.4) is equivalent to the condition

$$\frac{d}{dx}T_0(x,x)+\mu\big(T_0(x,x)-T_\infty(x,x)\big)^2+\frac{d}{dx}T_\infty(x,x)=0, \tag{3.5}$$

which we now verify. The left-hand side of (3.5) equals

$$\begin{aligned}
&Ce^{-xA}\Big(- A(I-\mu R_x)^{-1} - (I-\mu R_x)^{-1}\mu(AR_x + R_xA)(I-\mu R_x)^{-1} \\
&- (I-\mu R_x)^{-1}A\Big)e^{-xA}B \\
&+ Ce^{-xA}\Big((I-\mu R_x)^{-1} + (I+\mu R_x)^{-1}\Big)e^{-xA}\mu BCe^{-xA} \\
&\quad \times \Big((I-\mu R_x)^{-1} + (I+\mu R_x)^{-1}\Big)e^{-xA} \\
&+ Ce^{-xA}\Big(A(I+\mu R_x)^{-1} - (I+\mu R_x)^{-1}\mu(AR_x + R_xA)(I+\mu R_x)^{-1} \\
&+ (I+\mu R_x)^{-1}A\Big)e^{-xA}B. \qquad (3.6)
\end{aligned}$$

All of the terms begin with Ce^{-xA} and end with $e^{-xA}B$, and we can replace $e^{-xA}\mu BCe^{-xA}$ by $\mu(AR_x + R_xA)$ to obtain

$$\begin{aligned}
(3.6) = Ce^{-xA}\Big(&- 2(I-\mu R_x)^{-1}A(I-\mu R_x)^{-1} \\
&+ 4(I-\mu^2R_x^2)^{-1}\mu(AR_x + R_xA)(I-\mu^2R_x^2)^{-1} \\
&+ 2(I+\mu R_x)^{-1}A(I+\mu R_x)^{-1}\Big)e^{-xA}B \\
= 0.&
\end{aligned}$$

This proves (3.5), and one can easily check that (3.4) is equivalent to

$$u_0(x) = \frac{1}{\mu}\frac{dv}{dx} + v(x)^2, \quad v(x)^2 = -\frac{1}{\mu}\frac{dw}{dx}.$$

The entries of T satisfy the pair of coupled integral equations

$$\begin{aligned}
0 &= W(x,y) + \mu\int_x^\infty V(x,s)\phi(s+y)\,ds \\
0 &= V(x,y) + \phi(x+y) + \mu\int_x^\infty W(x,s)\phi(s+y)\,ds;
\end{aligned}$$

so W satisfies

$$\begin{aligned}
0 = -W(x,z) &+ \mu\int_x^\infty \phi(x+y)\phi(y+z)\,dy \\
&+ \mu^2\int_x^\infty W(x,s)\int_x^\infty \phi(s+y)\phi(y+z)\,dyds,
\end{aligned}$$

which explains how $\mu^2\Gamma_\phi^2$ enters into several determinant formulas [37]. □

Definition 3.5 (Darboux Addition)

(i) For $-\zeta \in S_\theta \cup \{0\}$ we define the Darboux addition rule on $(2, 2)$ admissible linear systems by

$$M_\zeta : (-A, B, C) \mapsto (-A, (\zeta I + A)(\zeta I - A)^{-1}B, C)$$

and correspondingly on potentials by

$$u_\infty \mapsto u_\zeta = u_\infty - 2(\log \psi_\zeta)''.$$

(ii) Let $\mathrm{Wr}(\varphi, \psi)$ be the Wronskian of $\psi, \varphi \in C^1((0, \infty); \mathbb{C})$.

Corollary 3.6 *The set* $\{M_\zeta, (\zeta \in X_\theta), M_0, M_\infty = I\}$ *generates a group such that* $M_0^2 = I$, $M_\zeta M_{-\zeta} = I$ *and* $M_\zeta M_\eta$ *corresponds to adding*

$$-2\frac{d^2}{dx^2} \log \mathrm{Wr}(\psi_\zeta, \psi_\eta)$$

to the potential.

Proof The definition is consistent with [10, p. 484]. In particular, $\psi_0(x) = \tau_0(x)/\tau_\infty(x)$, and $u_0(x) = u_\infty(x) - 2\frac{d^2}{dx^2} \log \psi_0(x)$, which is consistent with (3.5).

For $\zeta_1 \neq \zeta_2$, let $\Psi(x) = \mathrm{Wr}(\psi_{\zeta_1}, \psi_{\zeta_2})/\psi_{\zeta_2}$, and observe that

$$\Psi'' = \left(\zeta_2^2 + u_\infty - 2(\log \psi_{\zeta_1})''\right)\Psi.$$

This gives the basic composition rule for $M_{\zeta_2} M_{\zeta_1}$. The other statements follow from Proposition 3.2 and Theorem 3.4. See [24]. □

4 The State Ring Associated with an Admissible Linear System

Gelfand and Dikii [11] considered the algebra $\mathcal{A}_u = \mathbb{C}[u, u', u'', \dots]$ of complex polynomials in a smooth potential u and its derivatives. They showed that if u satisfies the stationary higher order KdV equations (5.1), then $\mathcal{A}_u$ is a Noetherian ring [2] and the associated Schrödinger equation is integrable by quadratures; see [7]. In this section, we introduce an analogue $\mathcal{A}_\Sigma$ for an admissible linear system.

We develop a calculus for R_x which is the counterpart of Pöppe's functional calculus for Hankel operators from [24, 31, 32]. As we see in other papers, our theory of state rings has wider scope for generalization.

Definition 4.1 (Differential Rings)

(1) Let $\mathcal{R}$ be a ring with ideal $\mathcal{J}$, and let $\partial : \mathcal{R} \to \mathcal{R}$ be a derivation. Then $\mathcal{R}_{\mathcal{J}} = \{r \in \mathcal{R} : \partial(r) \in \mathcal{J}\}$ gives a subring of $\mathcal{R}$, the ring of constants relative to $\mathcal{J}$. When $\mathcal{R}$ is an algebra over $\mathbb{C}$ and $\mathcal{J} = (0)$, we call $\mathcal{R}_0$ the constants; see [33].

(2) *(State ring of a linear system).* Let $(-A, B, C)$ be a linear system such that $A \in \mathcal{L}(H)$. Suppose that:

(i) $\mathcal{S}$ is a differential subring of $C^{\infty}((0, \infty); \mathcal{L}(H))$;
(ii) I, A and BC are constant elements of $\mathcal{S}$;
(iii) e^{-xA}, R_x and $F_x = (I + R_x)^{-1}$ belong to $\mathcal{S}$.

Then $\mathcal{S}$ is a state ring for $(-A, B, C)$.

Lemma 4.2 *Suppose that $(-A, B, C)$ is a linear system with $A \in \mathcal{L}(H)$ and that R_x gives a solution of Lyapunov's equation (1.2) such that $I + R_x$ is invertible for $x > 0$ with inverse F_x. Then the free associative algebra $\mathcal{S}$ generated by $I, R_0, A, F_0, e^{-xA}, R_x$ and F_x is a state ring for $(-A, B, C)$ on $(0, \infty)$. For all $t > 0$, there exists a ring homomorphism $S_t : \mathcal{S} \to \mathcal{S}$ given by $S_t : G(x) \mapsto G(x + t)$ such that S_t commutes with d/dx.*

Proof We can regard $\mathcal{S}$ as a subring of $C_b((0, \infty), \mathcal{L}(H)))$, so the multiplication is well defined. Then we note that $BC = AR_0 + R_0A$ belongs to $\mathcal{S}$, as required. We also note that $(d/dx)e^{-xA} = -Ae^{-xA}$ and that Lyapunov's equation (1.2) gives

$$\frac{d}{dx}(I + R_x)^{-1} = (I + R_x)^{-1}(AR_x + R_xA)(I + R_x)^{-1},$$

which implies

$$\frac{dF_x}{dx} = AF_x + F_xA - 2F_xAF_x,$$

with the initial condition

$$AF_0 + F_0A - 2F_0AF_0 = F_0BCF_0.$$

Hence $\mathcal{S}$ is a differential ring.

We can map $I \mapsto I$, $e^{-xA} \mapsto e^{-(x+t)A}$, $R_0 \mapsto e^{-tA}R_0e^{-tA}$, $R_x \mapsto e^{-tA}R_xe^{-tA}$ and $F_x \mapsto (I + e^{-tA}R_xe^{-tA})^{-1}$, and thus produce a ring homomorphism $G(x) \mapsto G(x + t)$ which satisfies

$$(d/dx)S_tG(x) = G'(x + t) = S_t(d/dx)G(x). \qquad \square$$

Definition 4.3 (Products and Brackets)

(i) Given a state ring $\mathcal{S}$ for $(-A, B, C)$, let $\mathcal{B}$ be any differential ring of functions from $(0, \infty) \to \mathcal{L}(H_0)$. Let

$$\mathcal{A}_\Sigma = \text{span}_{\mathbb{C}}\{A^{n_1}, A^{n_1} F_x A^{n_2} \dots F_x A^{n_r} : n_j \in \mathbb{N}\}.$$

(ii) On $\mathcal{S}$ we introduce the associative product $*$ by

$$P * Q = P(AF + FA - 2FAF)Q \qquad (P, Q \in \mathcal{S}),$$

which is distributive over the standard addition, and the derivation $\partial : \mathcal{S} \to \mathcal{S}$ by

$$\partial P = A(I - 2F)P + \frac{dP}{dx} + P(I - 2F)A \qquad (P \in \mathcal{S}). \tag{4.1}$$

(iii) Let $\lfloor \cdot \rfloor : \mathcal{S} \to \mathcal{B}$ be the linear map

$$\lfloor Y \rfloor = Ce^{-xA} F_x Y F_x e^{-xA} B \qquad (Y \in \mathcal{S}),$$

so that $x \mapsto \lfloor Y \rfloor$ is a differentiable function $(x_0, \infty) \to \mathcal{L}(H_0)$.

For $x_0 \geq 0$ and $0 < \phi < \pi$, let $S_\delta^{x_0}$ be the translated sector

$$S_\delta^{x_0} = \{z = x_0 + w : w \in \mathbb{C} \setminus \{0\}; |\arg w| < \delta\}$$

and let $H^\infty(S_\delta^{x_0})$ the bounded holomorphic complex functions on $S_\delta^{x_0}$. Then let $H_\infty^\infty = \cup_{x_0 > 0} H^\infty(S_\delta^{x_0})$ be the algebra of complex functions which are bounded on some translated sector $S_\delta^{x_0}$, with the usual pointwise multiplication.

Theorem 4.4 *Let $(-A, B, C)$ be a $(2, 2)$-admissible linear system with $H_0 = \mathbb{C}$ as in Theorem 2.3, so (e^{-zA}) for $z \in S_\phi^0$ is a bounded holomorphic semigroup on H. Let $\Theta_0 = \{P \in \mathcal{A}_\Sigma : \lfloor P \rfloor = 0\}$.*

(i) *Then $(\mathcal{A}_\Sigma, *, \partial)$ is a differential ring with bracket $\lfloor \cdot \rfloor$;*
(ii) *there is a homomorphism of differential rings $\lfloor \cdot \rfloor : (\mathcal{A}_\Sigma, *, \partial) \to (H_\infty^\infty, \cdot, d/dz)$;*
(iii) *Θ_0 is a differential ideal in $(\mathcal{A}_\Sigma, *, \partial)$ such that $\mathcal{A}_\Sigma/\Theta_0$ is a commutative differential ring, and an integral domain.*

Proof

(i) We can multiply elements in $\mathcal{S}$ by concatenating words and taking linear combinations. Since all words in $\mathcal{A}_\Sigma$ begin and end with A, we obtain words of the required form, hence $\mathcal{A}_\Sigma$ is a subring of $\mathcal{S}$. To differentiate a word in $\mathcal{A}_\Sigma$ we add words in which we successively replace each F_x by $AF_x + F_x A - 2F_x A F_x$, giving a linear combination of words of the required

form. The basic observation is that $dF/dx = AF + FA - 2FAF$, so one can check that

$$\partial(P * Q) = (\partial P) * Q + P * (\partial Q); \tag{4.2}$$

hence $(\mathcal{S}, *, \partial)$ is a differential ring with differential subring $(\mathcal{A}_\Sigma, *, \partial)$.

(ii) Now we verify that there is a homomorphism of differential rings $(\mathcal{A}_\Sigma, *, \partial) \to (\mathcal{B}, \cdot, d/dx)$ given by $P \mapsto \lfloor P \rfloor$. From the definition of R_x, we have $AR_x + R_x A = e^{-xA}BCe^{-xA}$, and hence

$$F_x e^{-xA} BC e^{-xA} F_x = AF_x + F_x A - 2F_x A F_x,$$

which implies

$$\begin{aligned} \lfloor P \rfloor \lfloor Q \rfloor &= Ce^{-xA} F_x P F_x e^{-xA} BC e^{-xA} F_x Q F_x e^{-xA} B \\ &= Ce^{-xA} F_x P (AF_x + F_x A - 2F_x A F_x) Q F_x e^{-xA} B \\ &= \lfloor P(AF_x + F_x A - 2F_x A F_x) Q \rfloor \\ &= \lfloor P * Q \rfloor. \end{aligned}$$

Moreover, the first and last terms in $\lfloor P \rfloor$ have derivatives

$$\frac{d}{dx} Ce^{-xA} F_x = Ce^{-xA} F_x A(I - 2F_x), \quad \frac{d}{dx} F_x e^{-xA} B = (I - 2F_x) A F_x e^{-xA} B,$$

so the bracket operation satisfies

$$\frac{d}{dx} \lfloor P \rfloor = \left\lfloor A(I - 2F_x)P + \frac{dP}{dx} + P(I - 2F_x)A \right\rfloor = \lfloor \partial P \rfloor. \tag{4.3}$$

In this case A is possibly unbounded as an operator, so we use the holomorphic semigroup to ensure that products (4.1) and brackets (4.2) are well defined. We observe that $\mathcal{A}_\Sigma$ has a grading $\mathcal{A}_\Sigma = \oplus_{n=1}^{\infty} A_n$, where A_n is the span of the elements that have total degree n when viewed as products of A and F. For $X_n \in A_n$ and $Y_m \in A_m$, we have $X_n * Y_m \in A_{n+m+2} \oplus A_{n+m+3}$ and $\partial X_n \in A_{n+1} \oplus A_{n+2}$.

Also we have $A^k e^{-zA} \in \mathcal{L}(H)$ for all $z \in S_\phi^0$ and $\|A^k e^{-zA}\|_{\mathcal{L}(H)} \to 0$ as $z \to \infty$ in S_ϕ^0; hence $R_z A^k \to 0$ and $A^k R_z \to 0$ in $\mathcal{L}(H)$ as $z \to \infty$ in S_ϕ^0. Hence there exists an increasing positive sequence $(x_k)_{k=0}^{\infty}$ such that $A^k F_z - A^k \in \mathcal{L}(H)$ for all $z \in S_\phi^{x_k}$ and $A^k F_z - A^k \to 0$ in $\mathcal{L}(H)$ as $z \to \infty$ in $S_\phi^{x_k}$. Let $X_n \in A_n$ and consider a typical summand $AF_z A^k F_z \ldots A$ in X_n; we replace each factor like $A^k F_z$ by the sum of $A^k(F_z - I)$ and A^k where $k \le n$; then we observe that there in an initial factor Ce^{-zA} and a final factor $e^{-zA}B$ in $\lfloor X_n \rfloor$; hence $\lfloor X_n \rfloor$ determines an element of $H^\infty(S_\phi^{x_n})$.

We can identify H_∞^∞ with the algebraic direct limit

$$H_\infty^\infty = \lim_{n\to\infty} H^\infty(S_\phi^{x_0+n}).$$

By the principle of isolated zeros, the multiplication on H_∞^∞ is consistently defined, and H_∞^∞ is an integral domain. Now each $f \in H_\infty^\infty$ gives $f \in H^\infty(S_\phi^{x_0})$ so $f' \in H^\infty(S_\phi^{x_0+1})$ by Cauchy's estimates, so $f' \in H_\infty^\infty$. From (i) we deduce that

$$\lfloor \cdot \rfloor : \oplus_{n=1}^\infty \mathcal{A}_n \to \cup_{n=1}^\infty H^\infty(S_\phi^{x_n})$$

is well-defined and the bracket is multiplicative with respect to $*$, and behaves naturally with respect to differentiation.

(iii) We check that $\lfloor \cdot \rfloor$ is commutative on $(\mathcal{A}_\Sigma, *, \partial)$, by computing

$$\begin{aligned}\lfloor P * Q \rfloor &= \operatorname{trace}\bigl(Ce^{-xA}FPFe^{-xA}BCe^{-xA}FQFe^{-xA}B\bigr)\\ &= \operatorname{trace}\bigl(Ce^{-xA}FQFe^{-xA}BCe^{-xA}FPFe^{-xA}B\bigr)\\ &= \lfloor Q * P \rfloor.\end{aligned}$$

Hence Θ_0 contains all the commutators $P * Q - Q * P$, and Θ_0 is the kernel of the homomorphism $\lfloor \cdot \rfloor$, hence is an ideal for $*$. Also, we observe that for all $Q \in \Theta_0$, we have $\partial Q \in \Theta_0$ since $\lfloor \partial Q \rfloor = (d/dx)\lfloor Q \rfloor = 0$. Hence Θ_0 is a differential ideal which contains the commutator subspace of $(\mathcal{A}_\Sigma, *)$, so $\mathcal{A}_\Sigma/\Theta_0$ is a commutative algebra. Also, ∂ determines a unique derivation $\bar\partial$ on $\mathcal{A}_\Sigma/\Theta_0$ by $\bar\partial Q = \partial Q + \Theta_0$ for all $Q \in \mathcal{A}_\Sigma$; hence $\mathcal{A}_\Sigma/\Theta_0$ is a differential algebra. We can identify $\mathcal{A}_\Sigma/\Theta_0$ with a subalgebra of H_∞^∞, which is an integral domain. □

Remark 4.5 Pöppe [31] introduced a linear functional $\lceil . \rceil$ on Fredholm kernels $K(x, y)$ on $L^2(0, \infty)$ by $\lceil K \rceil = K(0, 0)$. In particular, let K, G, H, L be integral operators on $L^2(0, \infty)$ that have smooth kernels of compact support, let $\Gamma = \Gamma_{\phi(x)}$ have kernel $\phi(s+t+2x)$, let $\Gamma' = \frac{d}{dx}\Gamma$ and $G = \Gamma_{\psi(x)}$ be another Hankel operator; then the trace satisfies

$$\lceil \Gamma \rceil = -\frac{d}{dx}\operatorname{trace}\Gamma, \tag{4.4}$$

$$\lceil \Gamma K G \rceil = -\frac{1}{2}\frac{d}{dx}\operatorname{trace}\Gamma K G, \tag{4.5}$$

$$\lceil (I+\Gamma)^{-1}\Gamma \rceil = -\operatorname{trace}\bigl((I+\Gamma)^{-1}\Gamma'\bigr), \tag{4.6}$$

$$\lceil K\Gamma \rceil \lceil GL \rceil = -\frac{1}{2}\lceil K(\Gamma' G + \Gamma G')L \rceil, \tag{4.7}$$

where (4.7) is known as the product formula. The easiest way to prove (4.4)–(4.7) is to observe that $\Gamma' G + \Gamma G'$ is the integral operator with kernel $-2\phi_{(x)}(s)\psi_{(x)}(t)$, which has rank one. These ideas were subsequently revived by McKean [24], and are implicit in some results of [37]. Our formulas (4.2) and (4.3) incorporate a similar idea, and are the basis of the proof of Theorem 4.4. The results we obtain appear to be more general than those of Pöppe, and extend to periodic linear systems [6].

For the remainder of this section, we let A be a $n \times n$ complex matrix with eigenvalues λ_j $(j = 1, \ldots, m)$ with geometric multiplicity n_j such that $\lambda_j + \lambda_k \neq 0$ for all $j, k \in \{1, \ldots, m\}$; let $\mathbb{K} = \mathbb{C}(e^{-\lambda_1 t}, \ldots, e^{-\lambda_m t}, t)$. Also, let $B \in \mathbb{C}^{n\times 1}$ and $C \in \mathbb{C}^{1\times n}$. The formula (4.9) resembles the expressions used to obtain soliton solutions of KdV, as in [19, (14.12.11)] and [16]. In [17, (6.25)], there is a discussion of how the scattering data evolve under the time evolution associated with the KdV flow.

Proposition 4.6

(i) *There exists a solution R_t to Lyapunov's equation (1.2) with $R_0 = BC$, such that the entries of R_t belong to $\mathbb{K}$, and $\tau(t) \in \mathbb{K}$;*

(ii) *$\phi \in \mathbb{K}$ satisfies a linear differential equation with constant coefficients.*

(iii) *Suppose further that all the eigenvalues of A are simple. Then there exists an invertible matrix S such that*

$$S^{-1}B = (b_j)_{j=1}^n \in \mathbb{C}^{n\times 1}, \quad CS = (c_j)_{j=1}^n \in \mathbb{C}^{1\times n}$$

and the tau function is given by

$$\begin{aligned}\tau(t) =& 1 + \sum_{j=1}^{n} \frac{b_j c_j e^{-2\lambda_j t}}{2\lambda_j} \\ &+ \sum_{(j,k),(m,p):j\neq m;k\neq p} (-1)^{j+k+m+p} \frac{b_j b_m c_k c_p e^{-(\lambda_j+\lambda_k+\lambda_m+\lambda_p)t}}{(\lambda_j+\lambda_m)(\lambda_k+\lambda_p)} + \ldots \\ &+ \prod_{j=1}^{n} \frac{b_j c_j}{2\lambda_j} \prod_{1\leq j<k\leq n} \frac{(\lambda_j-\lambda_k)^2}{(\lambda_j+\lambda_k)^2} e^{-2\sum_{j=1}^n \lambda_j t}.\end{aligned}$$

Proof

(i) By the hypothesis, we can introduce a chain of circles $\mathcal{C}$ that go once round each λ_j in the positive sense and have all the points $-\lambda_k$ in their exterior. Then by [4], the matrix

$$R_0 = \frac{-1}{2\pi i} \int_{\mathcal{C}} (A + \lambda I)^{-1} BC (A - \lambda I)^{-1} d\lambda$$

gives a solution to Sylvester's equation in the form $-AR_0 - R_0A = -BC$. To see this, one considers $(A + \lambda I)R_0 + R_0(A - \lambda I)$ and then uses the calculus of residues. By the Riesz functional calculus, we also have

$$e^{-tA} = \frac{1}{2\pi i}\int_{\mathcal{C}} \big(\lambda I - A\big)^{-1} e^{-t\lambda} d\lambda;$$

hence by Cauchy's residue theorem, there exist complex polynomials p_j and q_j, and integers $m_j \geq 0$ such that

$$e^{-tA} = \sum_{j=1}^{m} q_j(t) e^{-t\lambda_j} p_j(A), \tag{4.8}$$

where $q_j(t)$ is constant if the corresponding eigenvalue is simple. We let $R_t = e^{-tA} R_0 e^{-tA}$, which gives a solution to Lyapunov's equation with initial condition $-BC$. From (4.8), we see that all the entries of R_t belong to $\mathbb{K}$. By the Laplace expansion of the determinant, we see that all entries of $\tau(t) = \det(I + R_t)$ also belong to $\mathbb{K}$.

(ii) We have $\phi(t) = Ce^{-tA}B \in \mathbb{K}$ by (4.8). Also, we introduce the characteristic polynomial of $(-A)$ by $\det(\lambda I + A) = \sum_{j=0}^{n} a_j \lambda^j$. Then by the Cayley–Hamilton theorem, $\sum_{j=0}^{n} a_j \phi^{(j)}(t) = 0$.

(iii) There exists an invertible matrix S such that SAS^{-1} is the $n \times n$ diagonal matrix $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, and we observe that

$$R_t = \Big[\frac{b_j c_k e^{-(\lambda_j + \lambda_k)t}}{\lambda_j + \lambda_k}\Big]_{j,k=1}^{n}$$

satisfies

$$\frac{d}{dt} R_t = -[b_j c_k e^{-(\lambda_j+\lambda_k)t}]_{j,k=1}^{n}, \quad -DR_t - R_t D = -[b_j c_k e^{-(\lambda_j+\lambda_k)t}]_{j,k=1}^{n};$$

so R_t gives a solution of the Lyapunov equation with generator $-D$ and initial condition given by the rank-one matrix $-S^{-1}BCS = -[b_j c_k]_{j,k=1}^{n}$. Hence the tau function is given by $\tau(t) = \det(I + R_t)$ for this matrix, and there is an expansion

$$\det\Big[\delta_{jk} + \frac{b_j c_k e^{-(\lambda_j+\lambda_k)x}}{\lambda_j + \lambda_k}\Big]_{j,k=1}^{n} = \sum_{\sigma \subseteq \{1,\ldots,n\}} \det\Big[\frac{b_j c_k e^{-\lambda_j x - \lambda_k x}}{\lambda_j + \lambda_k}\Big]_{j,k\in\sigma} \tag{4.9}$$

in which each subset σ of $\{1,\dots,n\}$ of order $\sharp\sigma$, contributes a minor indexed by $j,k\in\sigma$. From the Cauchy determinant formula, we obtain the identity

$$\det\Big[\frac{b_jc_ke^{-\lambda_jx-\lambda_kx}}{\lambda_j+\lambda_k}\Big]_{j,k\in\sigma}=\prod_{j\in\sigma}\frac{b_jc_je^{-2\lambda_jx}}{2\lambda_j}\prod_{j,k\in\sigma:j\neq k}\frac{\lambda_j-\lambda_k}{\lambda_j+\lambda_k}.\qquad\square$$

5 Diagonal Green's Function and Stationary KdV Hierarchy

In this section, we obtain properties of $\mathcal{A}_\Sigma$ in terms of the brackets of odd powers of A. Thus we obtain some sufficient conditions for some differential equations to be integrable. Throughout this section, we suppose that the hypotheses of Theorem 4.4 are in force, so that any finite set of elements of $\mathcal{A}_\Sigma$ are holomorphic functions on a some sector Ω containing (x_0,∞) for some $x_0\geq 0$. We do not generally require u to be real valued, although in Theorem 5.4(iv) we impose this further condition so that we can compare our results with the classical spectral theory for the Schrödinger equation on the real line.

Definition 5.1 (Stationary KdV Hierarchy)

(i) Let $f_0=1$ and $f_1=(1/2)u$. Then the KdV recursion formula is

$$4\frac{d}{dx}f_{m+1}(x)=4f_1(x)\frac{d}{dx}f_m(x)+4\frac{d}{dx}\big(f_1(x)f_m(x)\big)-\frac{d^3}{dx^3}f_m(x).\qquad(5.1)$$

(ii) If u satisfies $f_m=0$ for all m greater than or equal to some m_0, then u satisfies the stationary KdV hierarchy and is said to be an algebro-geometric (finite gap) potential; see [10, 11, 13, 29, 35].

(iii) Suppose that $u(x)\to 0$ as $x\to\infty$, and likewise for all the partial derivatives $\partial^\ell u/\partial x^\ell$; suppose further that $f_j(x)\to 0$ as $x\to 0$ as $x\to\infty$ for all $j=1,2,\dots$. Then we say that the f_j are homogeneous solutions of the KdV hierarchy, and we consider cases where the system of differential equations (5.1) has no arbitrary constants of integration.

Proposition 5.2 *Let $\mathcal{A}_\Sigma$ be as in Theorem 4.4]. Then*

$$f_m=(-1)^m2\lfloor A^{2m-1}\rfloor,\quad m=1,2,\dots$$

satisfies the stationary KdV hierarchy (Novikov's equations), since

$$\begin{aligned}4\frac{d}{dx}\lfloor A^{2m+3}\rfloor=&\frac{d^3}{dx^3}\lfloor A^{2m+1}\rfloor+8\Big(\frac{d}{dx}\lfloor A\rfloor\Big)\lfloor A^{2m+1}\rfloor\\&+16\lfloor A\rfloor\Big(\frac{d}{dx}\lfloor A^{2m+1}\rfloor\Big).\end{aligned}\qquad(5.2)$$

Proof

(i) We have the basic identities

$$\lfloor A(I-2F)A(I-2F)X\rfloor = \lfloor A^2X\rfloor - 2\lfloor A\rfloor\lfloor X\rfloor; \tag{5.3}$$

$$-2A(AF+FA-2FAF) = A(I-2F)A(I-2F) - A^2 \tag{5.4}$$

and their mirror images. Hence

$$\frac{d}{dx}\lfloor A^{2m+1}\rfloor = \lfloor A(I-2F)A^{2m+1} + A^{2m+1}(I-2F)A\rfloor,$$

so

$$\begin{aligned}\frac{d^2}{dx^2}\lfloor A^{2m+1}\rfloor &= \lfloor A(I-2F)A(I-2F)A^{2m+1} + 2A(I-2F)A^{2m+1}(I-2F)A\\ &\quad + A^{2m+1}(I-2F)A(I-2F)A\\ &\quad - 2A(AF+AF-2FAF)A^{2m+1}\\ &\quad - 2A^{2m+1}(AF+FA-2FAF)A\rfloor\\ &= \lfloor A(I-2F)A(I-2F)A^{2m+1} + 2A(I-2F)A^{2m+1}(I-2F)A\\ &\quad + A^{2m+1}(I-2F)A(I-2F)A\\ &\quad + A(I-2F)A(I-2F)A^{2m+1} - A^{2m+3}\\ &\quad + A^{2m+1}(I-2F)A(I-2F)A - A^{2m+3}\rfloor\end{aligned}$$

and by the basic identities (5.3) and (5.4)

$$\begin{aligned}\frac{d^2}{dx^2}\lfloor A^{2m+1}\rfloor &= 2\lfloor A(I-2F)A^{2m+1}(I-2F)A\rfloor - 2\lfloor A^{2m+3}\rfloor\\ &\quad + 2\lfloor A(I-2F)A(I-2F)A^{2m+1}\rfloor\\ &\quad + 2\lfloor A^{2m+1}(I-2F)A(I-2F)A\rfloor\\ &= 2\lfloor A(I-2F)A^{2m+1}(I-2F)A\rfloor + 2\lfloor A^{2m+3}\rfloor\\ &\quad - 4\lfloor A^{2m+1}\rfloor\lfloor A\rfloor - 4\lfloor A\rfloor\lfloor A^{2m+1}\rfloor.\end{aligned}$$

Now we differentiate the first summand of the final term

$$\begin{aligned}&\frac{d}{dx}2\lfloor A(I-2F)A^{2m+1}(I-2F)A\rfloor\\ &\qquad = 2\lfloor A(I-2F)A(I-2F)A^{2m+1}(I-2F)A\rfloor\end{aligned}$$

$$
\begin{aligned}
&+ 2\lfloor A(I-2F)A^{2m+1}(I-2F)A(I-2F)A\rfloor \\
&- 4\lfloor A(AF+FA-2FAF)A^{2m+1}(I-2F)A\rfloor \\
&- 4\lfloor A(I-2F)A^{2m+1}(AF+FA-2FAF)A\rfloor
\end{aligned}
$$

$$
\begin{aligned}
= {} & 2\lfloor A(I-2F)A(I-2F)A^{2m+1}(I-2F)A\rfloor \\
&+ 2\lfloor A(I-2F)A^{2m+1}(I-2F)A(I-2F)A\rfloor \\
&+ 2\lfloor A(I-2F)A(I-2F)A^{2m+1}(I-2F)A\rfloor \\
&- 2\lfloor A^{2m+3}(I-2F)A\rfloor \\
&+ 2\lfloor A(I-2F)A^{2m+1}(I-2F)A(I-2F)A\rfloor \\
&- 2\lfloor A(I-2F)A^{2m+3}\rfloor
\end{aligned}
$$

thus we obtain

$$
\begin{aligned}
\frac{d^2}{dx^2}\lfloor A^{2m+1}\rfloor = {} & 4\lfloor A(I-2F)A(I-2F)A^{2m+1}(I-2F)A\rfloor \\
&+ 4\lfloor A(I-2F)A^{2m+1}(I-2F)A(I-2F)A\rfloor \\
&- 2\lfloor A(I-2F)A^{2m+3}+A^{2m+3}(I-2F)A\rfloor \\
= {} & -8\lfloor A\rfloor\lfloor A^{2m+1}(I-2F)A\rfloor + 4\lfloor A^{2m+3}(I-2F)A\rfloor \\
&- 8\lfloor A\rfloor\lfloor A(I-2F)A^{2m+1}\rfloor \\
&+ 4\lfloor A(I-2F)A^{2m+3}\rfloor - 2\frac{d}{dx}\lfloor A^{2m+3}\rfloor \\
= {} & -8\lfloor A\rfloor\lfloor A(I-2F)A^{2m+1}+A^{2m+1}(I-2F)A\rfloor \\
&+ 4\lfloor A(I-2F)A^{2m+3}+A^{2m+3}(I-2F)A\rfloor - 2\frac{d}{dx}\lfloor A^{2m+3}\rfloor \\
= {} & -8\lfloor A\rfloor\frac{d}{dx}\lfloor A^{2m+1}\rfloor + 2\frac{d}{dx}\lfloor A^{2m+3}\rfloor;
\end{aligned}
$$

hence

$$
\frac{d^3}{dx^3}\lfloor A^{2m+1}\rfloor = -8\lfloor A\rfloor\frac{d}{dx}\lfloor A^{2m+1}\rfloor + 4\frac{d}{dx}\lfloor A^{2m+3}\rfloor - 8\frac{d}{dx}\Big(\lfloor A\rfloor\lfloor A^{2m+1}\rfloor\Big);
$$

which gives the stated result (5.2). □

Definition 5.3 (Diagonal Green's Function) Let $(-A, B, C)$ be as in Theorem 2.3. Then the diagonal Green's function is $g_0(x;\zeta)/\sqrt{\zeta}$ where

$$g_0(x;\zeta) = (1/2) + \lfloor A(\zeta I - A^2)^{-1} \rfloor. \tag{5.5}$$

The notation $g_0(x;\zeta)$ is chosen to indicate a generating function and also the diagonal of a Green's function; now in Theorem 5.4(iv) we explain the latter connection. Let $\mathbb{C}_+ = \{\lambda \in \mathbb{C} : \Im\lambda > 0\}$.

Theorem 5.4 *Let $(-A, B, C)$ be as in Theorem 2.3.*

(i) *Then $g_0(x;\zeta)$ is bounded and continuously differentiable in x and has a unique asymptotic expansion depending on the bracketed odd powers of A,*

$$g_0(x;\zeta) \asymp \frac{1}{2} + \frac{\lfloor A \rfloor}{\zeta} + \frac{\lfloor A^3 \rfloor}{\zeta^2} + \frac{\lfloor A^5 \rfloor}{\zeta^3} + \dots \qquad (\zeta \to -\infty); \tag{5.6}$$

(ii) *$g_0(x;\zeta)$ satisfies Drach's equation*

$$\frac{d^3 g_0}{dx^3} = 4(u+\zeta)\frac{dg_0}{dx} + 2\frac{du}{dx} g_0 \qquad (x > x_0; -\zeta > \omega); \tag{5.7}$$

(iii) *there exists $x_1 > 0$ such that*

$$\psi_\pm(x,\zeta) = \sqrt{g_0(x,-\zeta)}\exp\Big(\mp\sqrt{-\zeta}\int_{x_1}^{x}\frac{dy}{2g_0(y;-\zeta)}\Big) \tag{5.8}$$

satisfies Schrödinger's equation

$$-\psi_\pm''(x;\zeta) + u(x)\psi_\pm(x,\zeta) = \zeta\psi_\pm(x;\zeta) \qquad (x > x_1, \zeta > \omega).$$

(iv) *Suppose that u is a continuous real function that is bounded below, and that $\psi_\pm$ from (iii) satisfy $\psi_+(x;\zeta) \in L^2((0,\infty);\mathbb{C})$ and $\psi_-(x;\zeta) \in L^2((-\infty,0);\mathbb{C})$ for all $\zeta \in \mathbb{C}_+$. Then $L = -\frac{d^2}{dx^2} + u(x)$ defines an essentially self-adjoint operator in $L^2(\mathbb{R};\mathbb{C})$, and the Green's function $G(x,y;\zeta)$ which represents $(\zeta I - L)^{-1}$ has a diagonal that satisfies*

$$G(x,x;\zeta) = \frac{g_0(x;-\zeta)}{\sqrt{-\zeta}}.$$

Proof

(i) Let $\pi - \theta < \arg\lambda < \theta$, so λ and $-\lambda$ both lie in S_θ, hence $\zeta = \lambda^2$ satisfies $2\pi - 2\theta < \arg\zeta < 2\theta$, so ζ lies close to $(-\infty, 0)$. Then $\zeta I - A^2$ is invertible

and $|\zeta|\|(\zeta I - A^2)^{-1}\|_{\mathcal{L}(H)} \le M$. The function

$$g_0(x;\zeta) = \frac{1}{2} + Ce^{-xA}(I+R_x)^{-1}A(\zeta I - A^2)^{-1}(I+R_x)^{-1}e^{-xA}B \quad (x>0)$$

is well defined by Theorem 2.3(iii).

To obtain the asymptotic expansion, we note that $e^{-xA}(I+R_x)^{-1}$ and $(I+R_x)e^{-xA}$ involve the factor e^{-xA}, where (e^{-zA}) is a holomorphic semigroup on $S_{\theta-\pi/2}$. Hence $A^{2j+1}e^{-xA} \in \mathcal{L}(H)$ and by Cauchy's estimates there exist $\delta, x_0, M_0 > 0$ such that $\|A^{2j+1}e^{-xA}\|_{\mathcal{L}(H)} \le M_0(2j+1)!$ for all $x \ge x_0 > 0$, and $\|e^{-sA}\|_{\mathcal{L}(H)} \le M_0 e^{-s\delta}$. As in Proposition 3.2, we have an asymptotic expansion of

$$\begin{aligned} &e^{-zA}\big((\lambda I - A)^{-1} - (\lambda I + A)^{-1}\big) \\ &\quad = -e^{-zA}\int_0^\infty e^{\lambda s}e^{-sA}\,ds - e^{-zA}\int_0^\infty e^{-\lambda s}e^{-sA}\,ds \\ &\quad = e^{-zA}\Big(\frac{A}{\lambda^2} + \frac{A^3}{\lambda^4} + \cdots + \frac{A^{2j-1}}{\lambda^{2j}}\Big) \\ &\qquad + \frac{e^{-zA}}{\lambda^{2j+1}}\int_0^\infty A^{2j+1}e^{-sA}(e^{s\lambda} - e^{-\lambda s})\,ds, \end{aligned}$$

in which all the summands are in $\mathcal{L}(H)$ due to the factor e^{-zA} for $z \in S_{\theta-\pi/2}$. Hence

$$Ce^{-xA}(I+R_x)^{-1}\int_0^\infty A^{2j+1}e^{-sA}(e^{s\lambda} - e^{-s\lambda})\,ds(I+R_x)^{-1}e^{-xA}B \to 0 \ (x>0)$$

as $\lambda \to i\infty$, or equivalently $\zeta \to -\infty$, so

$$\begin{aligned} g_0(x,\zeta) = Ce^{-xA}(I+R_x)^{-1}\Big(\frac{A}{\zeta} + \frac{A^3}{\zeta^2} + \cdots + \frac{A^{2j-1}}{\zeta^j}\Big)(I+R_x)^{-1}e^{-xA}B \\ + \frac{1}{2} + O\Big(\frac{1}{\zeta^{j+1}}\Big). \end{aligned}$$

This gives the asymptotic series; generally, the series is not convergent since the implied constants in the term $O(\zeta^{-(j+1)})$ involve $(2j+1)!$.

(ii) From Proposition 5.2 we have

$$\begin{aligned} 4\frac{d}{dx}\sum_{m=0}^\infty \frac{\lfloor A^{2m+3}\rfloor}{\zeta^{m+1}} = \frac{d^3}{dx^3}\sum_{m=0}^\infty \frac{\lfloor A^{2m+1}\rfloor}{\zeta^{m+1}} + 8\Big(\frac{d}{dx}\lfloor A\rfloor\Big)\sum_{m=0}^\infty \frac{\lfloor A^{2m+1}\rfloor}{\zeta^{m+1}} \\ + 16\lfloor A\rfloor\frac{d}{dx}\sum_{m=0}^\infty \frac{\lfloor A^{2m+1}\rfloor}{\zeta^{m+1}}; \end{aligned}$$

the required result follows on rearranging.

Conversely, suppose that g_0 as defined in (5.5) has an asymptotic expansion with coefficients in $C^\infty((0,\infty);\mathbb{C})$ as $\zeta \to -\infty$ and that $g_0(x;\zeta)$ satisfies (5.7). Then the coefficients of ζ^{-j} satisfy a recurrence relation which is equivalent to the systems of differential equations (5.1).

The asymptotic expansion is unique in the following sense. Suppose momentarily that $t \mapsto \lfloor Ae^{-tA^2}\rfloor$ is bounded and repeatedly differentiable on $(0,\infty)$, with $M, \omega > 0$ such that $|\lfloor Ae^{-tA^2}\rfloor| \le Me^{\omega t}$ for $t > 0$, and that there is a Maclaurin expansion

$$\lfloor Ae^{-tA^2}\rfloor = \lfloor A\rfloor - \lfloor A^3\rfloor t + \frac{\lfloor A^5\rfloor t^2}{2!} - \cdots + O(t^k)$$

on some neighbourhood of $0+$. Then by Watson's Lemma [36, p. 188], the integral

$$\int_0^\infty \lfloor Ae^{-tA^2}\rfloor e^{t\zeta}\, dt$$

has an asymptotic expansion as $\zeta \to -\infty$, where the coefficients give the formula (5.6).

(iii) Since $(e^{-tA})_{t>0}$ is a contraction semigroup on H, we have $\mathcal{D}(A^2) \subseteq \mathcal{D}(A)$ and $\|Af\|_H^2 \le 2\|A^2 f\|_H \|f\|_H$ for all $f \in \mathcal{D}(A^2)$ by the Hardy-Littlewood-Landau inequality [15, p. 65], so $\|\zeta f + A^2 f\|_H \ge \sqrt{\zeta}\|Af\|_H$ for $\zeta > 0$. We deduce that $A^2 - 2A + \zeta I$ is invertible for $\zeta > 9$ and generally for all $\zeta \in \mathbb{C}$ such that $\Re\zeta$ is sufficiently large. By Proposition 5.2 and the multiplicative property of the bracket, we have

$$\frac{1}{2g_0(x;-\zeta)} = 1 + \left\lfloor 2A(\zeta I + A^2 - 2A)^{-1}\right\rfloor,$$

and we observe that $g_0(x;-\zeta) \to 1/2$ as $x \to \infty$, so there exists $x_1 > 0$ such that $g_0(x,-\zeta) > 0$ for all $x > x_1$ and the differential equation integrates to

$$g_0\frac{d^2 g_0}{dx^2} - \frac{1}{2}\left(\frac{dg_0}{dx}\right)^2 = 2(u-\zeta)g_0^2 + \frac{\zeta}{2}. \tag{5.9}$$

So one can define $\psi(x;\zeta)$ as in (5.8), and then one verifies the differential equation for $\psi(x;\zeta)$ by using (5.9).

(iv) By a theorem of Weyl [18, 10.1.4], L is of limit point type at $\pm\infty$, and there exist nontrivial solutions $\psi_\pm(x;\zeta)$ to $-\psi_\pm''(x;\zeta) + u(x)\psi_\pm(x;\zeta) = \zeta\psi_\pm(x;\zeta)$ such that $\psi_+(x;\zeta) \in L^2(0,\infty)$ and $\psi_-(x;\zeta) \in L^2(-\infty,0)$, and these are unique up to constant multiples. Also the inverse operator $(-\zeta I + L)^{-1}$ may be represented as an integral operator in $L^2(\mathbb{R};\mathbb{C})$ with

kernel $G(x, y; \zeta)$, which has diagonal

$$G(x, x; \zeta) = \frac{\psi_+(x; \zeta)\psi_-(x; \zeta)}{\mathrm{Wr}(\psi_+(\,; \zeta), \psi_-(\,; \zeta))} \qquad (\Im\zeta > 0).$$

Given $\psi_\mp$ as in (iii), we can compute $\psi_+(x; \zeta)\psi_-(x; \zeta) = g_0(x; -\zeta)$ and their Wronskian is $\mathrm{Wr}(\psi_+, \psi_-) = \sqrt{-\zeta}$, hence the result. □

Remark 5.5

(i) The importance of the diagonal Green's function is emphasized in [14]. Gesztesy and Holden [13] obtain an asymptotic expansion of the diagonal $G(x, x; \zeta)$ which is consistent with Theorem 5.4(i). Under conditions discussed in Theorem 5.4, we have similar asymptotics as $-\zeta \to \infty$.

(ii) Drach observed that one can start with the differential equation (5.7), and produce the solutions (5.8); see [7]. He showed that Schrödinger's equation is integrable by quadratures, if and only if (5.7) can be integrated by quadratures for typical values of ζ, and Brezhnev translated his results into the modern theory of finite gap integration [7]. Having established integrability of Schrödinger's equation by quadratures, one can introduce the hyperelliptic spectral curve $\mathcal{E}$ with $g < \infty$ and proceed to express the solution in terms of the Baker–Akhiezer function. Hence one can integrate the equation and express the solution in terms of the Riemann's theta function on the Jacobian of $\mathcal{E}$, as in [3, 10, 13].

(iii) Kotani [21] has introduced the Baker–Akhiezer function and the τ function via the Weyl m-function for a suitable class of potentials that included multi-solitons and algebro-geometric potentials. There is a determinant formula for τ corresponding to (1.5) and (3.1), and the theory develops themes from [35].

(iv) The deformation theory for rational differential equations is discussed in [20].

Acknowledgments GB thanks Henry McKean for helpful conversations. SLN thanks EPSRC for financially supporting this research. The authors thank the referee for drawing attention to recent literature.

References

1. T. Aktosun, F. Demontis, C. van der Mee, Exact solutions to the focusing nonlinear Schrödinger equation. Inverse Prob. **23**, 2171–2195 (2007)
2. M.F. Atiyah, I.G. Macdonald, *Introduction to Commutative Algebra* (Addison Wesley, Reading, 1969)
3. H.F. Baker, *Abelian Functions: Abel's Theorem and the Allied Theory of Theta Functions* (Cambridge University Press, Cambridge, 1995)
4. R. Bhatia, P. Rosenthal, How and why to solve the operator equation $AX - XB = Y$. Bull. Lond. Math. Soc. **29**, 1–21 (1997)
5. G. Blower, Linear systems and determinantal random point fields. J. Math. Anal. Appl. **335**, 311–334 (2009)

6. G. Blower, On tau functions for orthogonal polynomials and matrix models. J. Phys. A **44**, 285202 (2011)
7. Y.V. Brezhnev, What does integrability of finite-gap or soliton potentials mean? Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. **366**, 923–945 (2008)
8. F.J. Dyson, Fredholm determinants and inverse scattering problems. Commun. Math. Phys. **47**, 171–183 (1976)
9. K.-J. Engel, R. Nagel, *One Parameter Semigroups for Linear Evolution Equations* (Springer, New York, 2000)
10. N. Ercolani, H.P. McKean, Geometry of KdV. IV: Abel sums, Jacobi variety and theta function in the scattering case. Invent. Math. **99**, 483–544 (1990)
11. I.M. Gelfand, L.A. Dikii, Integrable nonlinear equations and the Liouville theorem. Funct. Anal. Appl. **13**, 6–15 (1979)
12. I.M. Gelfand, B.M. Levitan, On the determination of a differential equation from its spectral function. Izvestiya Akad. Nauk SSSR Ser. Mat. **15**, 309–360 (1951)
13. F. Gesztesy, H. Holden, *Soliton Equations and Their Algebro-Geometric Solutions Volume I:* $(1+1)$*-Dimensional Continuous Models* (Cambridge University Press, Cambridge, 2003)
14. F. Gesztesy, B. Simon, The Xi function. Acta Math. **176**, 49–71 (1996)
15. J.A. Goldstein, *Semigroups of Linear Operators and Applications* (Oxford University Press, Oxford, 1985)
16. S. Grudsky, A. Rybkin, On classical solutions of the KdV equation. Proc. Lond. Math. Soc. **121**, 354–371 (2020)
17. S. Grudsky, A. Rybkin, Soliton theory and Hankel operators. SIAM J. Math. Anal. **47**, 2283–2323 (2015)
18. E. Hille, *Lectures on Ordinary Differential Equations* (Addison-Wesley, Reading, 1968)
19. V.G. Kac, *Infinite Dimensional Lie Algebras* (Cambridge University Press, Cambridge, 1985)
20. V. Katsnelson, D. Volok, Rational solutions of the Schlesinger system and isoprincipal deformations of rational matrix functions. I. Oper. Theory Adv. Appl. **149**, 291–348 (2004)
21. S. Kotani, Construction of KdV flow I. τ-function via Weyl function. Zh. Mat. Fiz. Anal. Geom. **14**, 297–335 (2018)
22. I.M. Krichever, The integration of nonlinear equations by the methods of algebraic geometry. Funct. Anal. Appl. **11**, 12–26 (1977)
23. V.B. Matveev, Darboux transformation and explicit solutions of the Kadomtcev–Petviaschvily equation, depending upon functional parameters. Lett. Math. Phys. **3**, 213–216 (1979)
24. H.P. McKean, Fredholm determinants. Cent. Eur. J. Math. **9**, 205–243 (2011)
25. A.V. Megretskii, V.V. Peller, S.R. Treil, The inverse spectral problem for self-adjoint Hankel operators. Acta Math. **174**, 241–309 (1995)
26. T. Miwa, M. Jimbo, E. Date, *Solitons: Differential Equations, Symmetries, and Infinite Dimensional Algebras* (Cambridge University Press, Cambridge, 2000)
27. M. Mulase, Cohomological structure in soliton equations and Jacobian varieties. J. Differ. Geom. **19**, 403–430 (1984)
28. N.K. Nikolski, *Operators, Functions and Systems: An Easy Reading*, vol. 1 (American Mathematical Society, Providence, 2002)
29. S. Novikov, S.V. Manakov, L.P. Pitaevskii, V.F. Zakharov, *Theory of Solitons, the Inverse Scattering Method* (Consultants Bureau, New York and London, 1984)
30. V.V. Peller, *Hankel Operators and Their Applications* (Springer, New York, 2003)
31. C. Pöppe, The Fredholm determinant method for the KdV equations. Phys. D **13**, 137–160 (1984)
32. C. Pöppe, D.H. Sattinger, Fredholm determinants and the τ function for the Kadomtsev–Petviashvili hierarchy. Publ. Res. Inst. Math. Sci. **24**, 505–538 (1988)
33. M. van der Put, M.F. Singer, *Galois Theory of Linear Differential Equations* (Springer, Berlin, 2003)
34. A. Rybkin, The Hirota τ-function and well-posedness of the KdV equation with an arbitrary step-like initial profile decaying on the right half line. Nonlinearity **24**, 2953–2990 (2011)

35. G. Segal, G. Wilson, Loop groups and equations of KdV type. Inst. Hautes Études Sci. Publ. Math. **61**, 5–65 (1985)
36. I.N. Sneddon, *The Use of Integral Transforms* (McGraw-Hill, New York, 1972)
37. C.A. Tracy, H. Widom, Fredholm determinants, differential equations and matrix models. Commun. Math. Phys. **163**, 33–72 (1994)
38. M. Trubowitz, The inverse problem for periodic potentials. Commun. Pure Appl. Math. **30**, 321–337 (1977)
39. T. Zhang, S. Venakides, Periodic limit of inverse scattering. Commun. Pure Appl. Math. **46**, 819–865 (1993)

Groups of Orthogonal Matrices All Orbits of Which Generate Lattices

Albrecht Böttcher

Abstract There are infinitely many finite groups of orthogonal matrices all orbits of which, including those of irrational vectors, span lattices, that is, discrete additive subgroups of the underlying Euclidean space. We show that, both up to isomorphism and up to orthogonal similarity, exactly eight of these groups are irreducible: the two trivial groups in one dimension, the cyclic groups of orders 3, 4, 6 in two dimensions, and the quaternion, binary dihedral, binary tetrahedral groups in four dimensions.

Keywords Matrix groups · Orthogonal matrices · Lattices in Euclidean space · Tight frames

Mathematics Subject Classification (2010) Primary 20H15; Secondary 11H06, 15B10, 42C15, 52C07

1 Introduction and Main Results

Let $k \geq 1$ and let $\mathcal{G}$ be a finite subgroup of $\mathrm{O}(k)$, that is, let $\mathcal{G}$ be a finite group of orthogonal $k \times k$ matrices. Suppose the order of $\mathcal{G}$ is $n \geq 1$ and write $\mathcal{G} = \{G_1, \ldots, G_n\}$. We think of $\mathbf{R}^k$ as a column space. For $f \in \mathbf{R}^k$, consider the $k \times n$ matrix

$$F = \begin{pmatrix} G_1 f & G_2 f & \ldots & G_n f \end{pmatrix} \tag{1}$$

A. Böttcher (✉)
Fakultät für Mathematik, TU Chemnitz, Chemnitz, Germany
e-mail: aboettch@math.tu-chemnitz.de

M. A. Bastos et al. (eds.), *Operator Theory, Functional Analysis and Applications*,
Operator Theory: Advances and Applications 282,
https://doi.org/10.1007/978-3-030-51945-2_6

whose columns are the orbit of f under $\mathcal{G}$. We denote by $\Lambda(\mathcal{G}, f) \subset \mathbf{R}^k$ the set of all linear combinations of the columns of F with integer coefficients,

$$\begin{aligned}\Lambda(\mathcal{G}, f) &= \mathrm{span}_{\mathbf{Z}}\{G_1 f, \ldots, G_n f\} = \mathrm{span}_{\mathbf{Z}}\{Gf : G \in \mathcal{G}\} \\ &= \{Fx : x \in \mathbf{Z}^n\} = F\mathbf{Z}^n.\end{aligned}$$

Inspired by Fukshansky et al. [2], we are interested in whether $\Lambda(\mathcal{G}, f)$ is a lattice, that is, a discrete additive subgroup of $\mathbf{R}^k$. Since $\Lambda(\mathcal{G}, f)$ is always a subgroup of $\mathbf{R}^k$, the problem is its discreteness. Clearly, whether $\Lambda(\mathcal{G}, f)$ is a lattice or not depends on both $\mathcal{G}$ and f. In contrast to [2], we here embark on the question whether there are groups $\mathcal{G}$ for which $\Lambda(\mathcal{G}, f)$ is a lattice *independently* of f. We call such groups *lattice generating*. Thus, $\mathcal{G}$ is lattice generating if $\Lambda(\mathcal{G}, f)$ is a lattice for every $f \in \mathbf{R}^k$.

Obviously, the two subgroups $\{1\}$ and $\{1, -1\}$ of $\mathrm{O}(1)$ are lattice generating. The trivial group $\{I\} \subset \mathrm{O}(k)$ is also lattice generating. The interesting cases are $k \geq 2$ and $n \geq 2$.

For each G_j, we have

$$\begin{aligned}G_j\Lambda(\mathcal{G}, f) &= G_j\mathrm{span}_{\mathbf{Z}}\{Gf : G \in \mathcal{G}\} = \mathrm{span}_{\mathbf{Z}}\{G_jGf : G \in \mathcal{G}\} \\ &= \mathrm{span}_{\mathbf{Z}}\{Gf : G \in \mathcal{G}\} = \Lambda(\mathcal{G}, f).\end{aligned}$$

It follows that if $\Lambda(\mathcal{G}, f)$ is a nontrivial lattice for at least one f, then each $G_j \in \mathcal{G}$ must leave this lattice invariant. This indicates that we have to search for lattice generating groups within the crystallographic point groups.

The 10 crystallographic point groups in $\mathbf{R}^2$ are the rotation groups C_ℓ for $\ell = 1, 2, 3, 4, 6$ and the dihedral groups D_ℓ for $\ell = 1, 2, 3, 4, 6$. It is not difficult to check that the 7 groups $C_1, C_2, C_3, C_4, C_6, D_1, D_2$ are lattice generating, whereas the groups D_3, D_4, D_6 are not lattice generating. It can be shown that exactly half of the 32 crystallographic point groups in $\mathbf{R}^3$ are lattice generating. The table lists the 32 groups (in Schoenflies notation). The 16 groups in boldface are the lattice generating groups, the other 16 groups are not lattice generating.

$$\begin{aligned}&\mathbf{C_1}, \mathbf{C_2}, \mathbf{C_3}, \mathbf{C_4}, \mathbf{C_6}, \\ &\mathbf{C_{1h}}, \mathbf{C_{2h}}, \mathbf{C_{3h}}, \mathbf{C_{4h}}, \mathbf{C_{6h}}, \\ &\mathbf{C_{2v}}, C_{3v}, C_{4v}, C_{6v}, \\ &\mathbf{D_2}, D_3, D_4, D_6, \mathbf{D_{2h}}, D_{3h}, D_{4h}, D_{6h}, D_{2d}, D_{3d}, \\ &\mathbf{S_2}, \mathbf{S_4}, \mathbf{S_6}, \\ &T, T_d, T_h, O, O_h.\end{aligned}$$

Thus, taking direct sums of these $2 + 7 + 16 = 25$ lattice generating groups we obtain lots of lattice generating groups in every dimension. However, the truly

interesting groups are the irreducible ones. A finite group $\mathcal{G} \subset \mathrm{O}(k)$ is said to be *irreducible* if the members of the group do not share a common invariant subspace except for $\{0\}$ and all of $\mathbf{R}^k$. Equivalently, $\mathcal{G}$ is irreducible if and only if $\mathrm{span}_{\mathbf{R}}\{Gf : G \in \mathcal{G}\} = \mathbf{R}^k$ for every $f \neq 0$, which in turn is equivalent to the requirement that the rank of the matrix (1) is k for every $f \neq 0$. An irreducible group must in particular have $n \geq k$ elements.

The direct sums of the 25 lattice generating groups we found do not yield irreducible groups. But what about these groups themselves? It turns out that C_3, C_4, C_6 are the only irreducible lattice generating subgroups of O(2) and that none of the 16 lattice generating subgroups of O(3) is irreducible (because each of the latter leaves a rotation axis invariant). Tensor products behave better with regard to irreducibility. Do we obtain more lattice generating groups in this way? Herewith our first result.

Theorem 1.1 *The tensor product of two irreducible and lattice generating groups acting on at least two-dimensional spaces is never both irreducible and lattice generating.*

Note that the theorem is not true but becomes a triviality if one of the groups acts on $\mathbf{R}^1$: we have $\{1\} \otimes \mathcal{G} = \mathcal{G}$ and $\{1, -1\} \otimes \mathcal{G} = -\mathcal{G} \cup \mathcal{G}$. Clearly, $-\mathcal{G} \cup \mathcal{G}$ is irreducible if and only if so is $\mathcal{G}$, and $-\mathcal{G} \cup \mathcal{G}$ is lattice generating if and only if $\mathcal{G}$ has this property. Thus, to search for irreducible lattice generating groups we have to proceed differently.

All unitary irreducible representations of a finite abelian group are in U(1) and hence all irreducible representations of a finite abelian group in $\mathrm{O}(k)$ must have degree $k \leq 2$. Consequently, if $k \geq 3$, the $\mathrm{O}(k)$ does not contain irreducible finite abelian groups and thus all the more does not contain irreducible and lattice generating abelian groups. So we are left with non-abelian groups.

Theorem 1.2 *If $k \geq 2$ and a subgroup of $\mathrm{O}(k)$ is irreducible and lattice generating, then it is actually contained in $\mathrm{SO}(k)$.*

In particular, irreducible Coxeter (reflection) groups are never lattice generating.

Theorem 1.3 *If $k \geq 3$ is odd, then $\mathrm{SO}(k)$ does not contain irreducible lattice generating groups.*

Eventually we have the following.

Theorem 1.4 *If $k \geq 6$ is even, then there are no irreducible and lattice generating groups in $\mathrm{SO}(k)$. If a subgroup of $\mathrm{SO}(4)$ is irreducible and lattice generating, then it is either isomorphic to the quaternion group Q_8 of order 8 or to the binary dihedral group Q_{12} of order 12 or to the binary tetrahedral group $2T$ of order 24. The groups $Q_8, Q_{12}, 2T$ have faithful irreducible and lattice generating representations in $\mathrm{SO}(4)$.*

The quaternion group Q_8 may be given by $\langle a, b, c : a^2 = b^2 = c^2 = abc \rangle$. The binary dihedral group Q_{12} also goes under the notations Dic_{12}, Dic_6, or Dic_3. It has

the group presentation

$$\langle a, b, c : a^3 = b^2 = c^2 = abc \rangle.$$

The binary tetrahedral group $2T$ is isomorphic to $\mathrm{SL}(2, 3)$, that is, to the 2×2 matrices over the field $\mathbf{F}_3$ with determinant 1. It is also isomorphic to the group of units in the ring of Hurwitz integers. A group presentation of $2T$ is

$$\langle a, b, c : a^2 = b^3 = c^3 = abc \rangle.$$

We will construct faithful irreducible and lattice generating representations of these three groups in SO(4) when proving Theorem 1.4.

Two subgroups $\mathcal{G}$ and $\mathcal{H}$ of $\mathrm{SO}(k)$ are said to be *orthogonally similar* (or simply to be equivalent) if the there is an orthogonal matrix U such that $G \mapsto UGU^{-1}$ is a bijection of $\mathcal{G}$ onto $\mathcal{H}$. From the character tables of Q_8, Q_{12}, $2T$ which can be found in [4, 6] it follows that the faithful irreducible representations of Q_8 and Q_{12} in SO(4) are all orthogonally similar but that $2T$ has exactly two classes of orthogonally similar faithful representations in SO(4), which are called the quaternionic and the complex representations. We here prove the following.

Theorem 1.5 *The quaternionic representations of* $2T$ *in* SO(4) *are lattice generating whereas the complex representations of* $2T$ *in* SO(4) *are not lattice generating.*

Thus, not only up to isomorphism but also up to orthogonal similarity, there exist exactly eight irreducible and lattice generating groups.

The results on groups we will use in the following are all well known and can be found in the basic literature. All we use from representation theory is some general facts that are in the classic [7] or in the recent text [3], for example, and two character tables of real representations which are explicitly given in [4, 6].

2 Proofs

We denote by $(\cdot, \cdot)$ and $\|\cdot\|$ the usual scalar product and norm in the Euclidean space $\mathbf{R}^k$. The transpose of a matrix A is denoted by A'.

Assertion (a) of the following theorem is Corollary 10.5 of [8]. The other assertions of the theorem were established in [1] for $k \in \{2, 3\}$ and subsequently in [2] for general $k \geq 2$. For the reader's convenience, we include the proof from [2].

Theorem 2.1 *Suppose* $n \geq k \geq 2$. *Let* $\mathcal{G} \subset \mathrm{O}(k)$ *be a finite irreducible group of order* n, *let* $f \in \mathbf{R}^k \setminus \{0\}$, *and let* F *be the* $k \times n$ *matrix* (1). *Then the following hold.*

(a) *The matrix* FF' *is a nonzero scalar multiple of the identity matrix,*

$$FF' = \gamma I.$$

(b) *The set* $\Lambda(\mathcal{G}, f)$ *is a lattice if and only if the Gram matrix*

$$F'F = ((G_j f, G_k f))_{j,k=1}^{n}$$

is a scalar multiple of a rational matrix.

(c) *In case* $\Lambda(\mathcal{G}, f)$ *is a lattice, we actually have* $F'F \in \gamma \mathbf{Q}^{n\times n}$.

(d) *If* $\|f\| = 1$, *then* $\Lambda(\mathcal{G}, f)$ *is a lattice if and only if* $F'F$ *is a rational matrix.*

Proof

(a) Corollary 10.5 of [8] states that if $\mathcal{G}$ is a finite irreducible subgroup of $\mathrm{O}(k)$, then the matrix F is a tight frame for every $f \in \mathbf{R}^k \setminus \{0\}$, which is equivalent to saying that $FF' = \gamma I$ with some nonzero real γ.

(b) Suppose $\mu F'F \in \mathbf{Q}^{n\times n}$ for some nonzero $\mu \in \mathbf{R}$. Then $\mu d F'F \in \mathbf{Z}^{n\times n}$ for some positive integer d. It follows that

$$\mu^2 d^2 \|Fx\|^2 = \mu d(\mu d F'Fx, x) \in \mu d\mathbf{Z}$$

for all $x \in \mathbf{Z}^n$. Hence, if $x \in \mathbf{Z}^n$, then either $Fx = 0$ of $\|Fx\|^2 \geq 1/(\mu d)$. This proves that $\Lambda(\mathcal{G}, f) = F\mathbf{Z}^n$ is discrete and thus a lattice.

Conversely, suppose $\Lambda(\mathcal{G}, f)$ is a lattice. We know from (a) that $FF' = \gamma I$. This implies that $(1/\sqrt{\gamma})F$ is built by the first k rows of an orthogonal matrix, and hence the rank of F is k. Thus, the columns of F span all of $\mathbf{R}^k$. Let $B \in \mathbf{R}^{k\times k}$ be a basis matrix for $\Lambda(\mathcal{G}, f)$, that is, B is invertible and $\Lambda(\mathcal{G}, f) = \{Bx : x \in \mathbf{Z}^k\}$. It follows that there is a matrix $Z \in \mathbf{Z}^{k\times n}$ such that $F = BZ$. We obtain that $BZZ'B' = FF' = \gamma I$ and hence $ZZ' = \gamma B^{-1}(B')^{-1} = \gamma(B'B)^{-1}$. Consequently, $F'F = Z'B'BZ = Z'(\gamma(ZZ')^{-1})Z \in \gamma\mathbf{Q}^{n\times n}$.

(c) We just showed that if $\Lambda(\mathcal{G}, f)$ is a lattice, then $F'F \in \gamma\mathbf{Q}^{n\times n}$.

(d) If f has norm 1, then $1 = (f, f)$ is an entry of $F'F \in \gamma\mathbf{Q}^{n\times n}$, which implies that γ must be rational. Thus, $F'F \in \mathbf{Q}^{n\times n}$. □

Every orthogonal matrix G is of the form $G = USU'$ with an orthogonal matrix U and a real block-diagonal matrix S. The blocks of S are either 2×2 matrices of the form

$$\begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} \quad \text{with} \quad \alpha^2 + \beta^2 = 1,\ \beta > 0 \tag{2}$$

or the 1×1 matrices (± 1). Note that since

$$\begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

the restriction $\beta > 0$ can always be achieved by changing U. We say that G *has equal blocks* if all blocks of S are equal to each other. The trivial cases are $G = UIU' = I$ and $G = U(-I)U' = -I$.

Lemma 2.2 *If $\mathcal{G}$ is an irreducible and lattice generating subgroup of $\mathrm{O}(k)$, then each matrix in $\mathcal{G}$ has equal blocks.*

Proof Let $\mathcal{G} = \{G_1, \ldots, G_n\}$. From Theorem 2.1(c) we infer that

$$(G_j f, G_k f) \in \gamma_f \mathbf{Q}$$

with some $\gamma_f \neq 0$ for every f and all j, k. Taking $G_k = I$, we arrive at the conclusion that $(Gf, f) \in \gamma_f \mathbf{Q}$ for every $G \in \mathcal{G}$ and every f. It follows that $(Gf, f)/(f, f) \in \mathbf{Q}$ for all $f \neq 0$. Let now $G = USU'$. Since

$$\frac{(Gf, f)}{(f, f)} = \frac{(USU'f, f)}{(UU'f, f)} = \frac{(SU'f, U'f)}{(U'f, U'f)},$$

and every $h \neq 0$ is of the form $U'f$ with $f \neq 0$, we conclude that $(Sh, h)/(h, h)$ must be rational for every $h \neq 0$.

Thus, if $G = USU'$, then (Sh, h) must be rational for every $h \in \mathbf{R}^k$ of norm 1. Suppose S has two 2×2 blocks

$$B = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}, \quad C = \begin{pmatrix} \gamma & -\delta \\ \delta & \gamma \end{pmatrix}.$$

We may without loss of generality assume that these are the first two blocks of S, that is, $S = \mathrm{diag}(B, C, \ldots)$. We take $h = (f_1, f_2, g_1, g_2, 0, \ldots, 0)' \in \mathbf{R}^k$ of norm 1 and put $f = (f_1, f_2)'$, $g = (g_1, g_2)'$. Thus, $\|h\|^2 = \|f\|^2 + \|g\|^2 = 1$. Then

$$\begin{aligned}(Sh, h) &= (Bf, f) + (Cg, g) = \alpha(f_1^2 + f_2^2) + \gamma(g_1^2 + g_2^2) \\ &= \alpha\|f\|^2 + \gamma\|g\|^2 = \gamma + (\alpha - \gamma)\|f\|^2,\end{aligned}$$

and if $\alpha \neq \gamma$, we may clearly choose f so that $\gamma + (\alpha - \gamma)\|f\|^2$ is irrational. Thus, we must have $\alpha = \gamma$. Since $\beta^2 = \delta^2 = 1 - \alpha^2$ and $\beta, \delta > 0$, it follows that $\beta = \delta$. If S has only one 2×2 block, $S = \mathrm{diag}(B, \pm 1, \ldots)$, we get analogously with $h = (f_1, f_2, g, 0, \ldots)'$ and $f_1^2 + f_2^2 + g^2 = 1$ that

$$(Sh, h) = (Bf, f) \pm g^2 = \alpha(f_1^2 + f_2^2) \pm g^2 = \alpha - (\alpha \mp 1)g^2.$$

Since $-1 < \alpha < 1$, there are g such that this is irrational. Thus, the presence of one 1×1 block forces all blocks to be 1×1. Finally, if S has two different 1×1 blocks, $S = \mathrm{diag}(1, -1, \ldots)$, then $h = (f, g, 0, \ldots)'$ with $f^2 + g^2 = 1$ gives $(Sh, h) = f^2 - g^2 = 2f^2 - 1$, which may also be irrational. □

Proof of Theorems 1.2 and 1.3 Let $\mathcal{G} \subset \mathrm{O}(k)$ be irreducible and lattice generating. If $\det G = -1$ for some $G = USU' \in \mathcal{G}$, then S must contain a 1×1 block (-1), so Lemma 2.2 implies that $G = -I$, and thus $\mathcal{G} = \{I, -I\}$. But this group is not

irreducible. If k is an odd number, then S must have a 1×1 block (± 1) and we arrive again at the reducible groups $\mathcal{G} = \{I\}$ or $\mathcal{G} = \{I, -I\}$. □

If $\mathcal{G} \subset \mathrm{O}(k)$ is a finite group, then each element $G \in \mathcal{G}$ is of finite order (as an element of the group). If $G = USU'$ has equal blocks, then G is of finite order if and only if $G \in \{I, -I\}$ or if the block is (2) with $\alpha \in \{-1/2, 0, 1/2\}$ (Niven's theorem; see [5, pp. 37–41]). In these cases the order of G is 1, 2, 3, 4, 6, respectively. Let us denote the blocks with $\alpha = -1/2, 0, 1/2$ by Δ_3, Δ_4, Δ_6, respectively.

Corollary 2.3 *Let $\mathcal{G} \subset \mathrm{O}(k)$ be irreducible and lattice generating. If $-I \notin \mathcal{G}$, then*

$$\mathcal{G} = \{I, B_1, \ldots, B_q\} \tag{3}$$

where each B_j is of order 3. If $-I \in \mathcal{G}$, then

$$\mathcal{G} = \{I, -I, A_1, \ldots, A_s, -A_1, \ldots, -A_s, B_1, \ldots, B_q, -B_1, \ldots, -B_q\} \tag{4}$$

($s, q \geq 0, s + q > 0$) where the $\pm A_j$ are of order 4 and satisfy $(\pm A_j)^2 = -I$ and $-A_j = A_j^3$, each B_j is of order 6 and satisfies $B_j^3 = -I$, and each $-B_j$ is of order 3 and satisfies $-B_j = B_j^4$.

Proof We know from Lemma 2.2 that each $G \in \mathcal{G} \setminus \{I\}$ has equal blocks. This block is (-1), Δ_6, Δ_4, or Δ_3. If $-I \notin \mathcal{G}$, then the block cannot be Δ_6 or Δ_4, since then G^3 or G^2 would be $-I$. Thus, in this case $\mathcal{G}$ is of the form (3) in which the matrices B_j have the block Δ_3. Now suppose $-I \in \mathcal{G}$ and let $G \in \mathcal{G} \setminus \{I, -I\}$. If the block of G is Δ_4, then G is of order 4 with $G^2 = -I$ and $G^3 = -G$, which gives us the elements $\pm A_j$ in (4). If the block is Δ_6, then G is of order 6 and we have $G^3 = -I$ and $G^4 = -G$. These elements G are the $\pm B_j$ in (4). Matrices with the block Δ_3 are of the order 3. The map $G \mapsto -G$ is a bijection of every group containing I and $-I$ and it changes the orders 4 and 3 to the orders 4 and 6, respectively. Thus, the number r of elements of order 3 is equal to the number q of elements of the order 6, and since the elements $-B_1, \ldots, -B_q$ have the order 3, it follows that G must be among these elements. □

Proof of Theorem 1.1 By virtue of Theorem 1.3 we may suppose that the groups are subgroups of $\mathrm{O}(2m)$ and $\mathrm{O}(2k)$ with $m, k \geq 1$. Lemma 2.2 shows that the trace of an element G of an irreducible and lattice generating subgroup of $\mathrm{O}(2\ell)$ is ℓ if the order of G is 6, equals 0 if G has the order 4, and is $-\ell$ if G is of order 3. Let $\mathcal{G} \subset \mathrm{O}(2m)$ and $\mathcal{H} \subset \mathrm{O}(2k)$ be irreducible and lattice generating groups. Assume $\mathcal{G} \otimes \mathcal{H}$ is an irreducible and lattice generating subgroup of $\mathrm{O}(4mk)$. We have $\mathrm{tr}\,(G \otimes H) = (\mathrm{tr}\,G)(\mathrm{tr}\,H)$ for $G \in \mathcal{G}$ and $H \in \mathcal{H}$. Thus if both $\mathcal{G}$ and $\mathcal{H}$ contain elements G, H of orders 3 or 6, then $\mathrm{tr}\,(G \otimes H) = \pm mk$ whereas the possible traces of elements of $\mathcal{G} \otimes \mathcal{H}$ are $\pm 4mk$ (for $I, -I$), $\pm 2mk$, and 0. It follows that at least one of the groups, say $\mathcal{G}$, has no elements of the orders 3 and 6. By Corollary 2.3, $\mathcal{G}$ contains a G of order 4 such that $G^2 = -I$. If in the other group there is an element H of order 3 or 6, then $G \otimes H$ has the order 12, which is impossible. Consequently,

again by Corollary 2.3, the other group has an H of order 4 satisfying $H^2 = -I$. But then $G \otimes H$ satisfies $(G \otimes H)^2 = (-I) \otimes (-I) = I \otimes I$, which shows that $\mathcal{G} \otimes \mathcal{H}$ has two elements of order 2. However, Corollary 2.3 tells us that this cannot happen. □

Lemma 2.4 *An irreducible lattice generating group cannot contain a subgroup isomorphic to* $\mathbf{Z}_3^2$.

Proof Let $\mathcal{G}$ be irreducible and lattice generating and $\mathcal{H}$ be a subgroup isomorphic to $\mathbf{Z}_3^2$. Let X be a nontrivial invariant subspace of $\mathcal{H}$. Since all matrices in $\mathcal{H}$ are orthogonal, $X^\perp$ is also a nontrivial invariant subspace of $\mathcal{H}$. Choose orthonormal bases in X and $X^\perp$, take their union, and represent the matrices in $\mathcal{H}$ in this new bases. The new matrices are composed by two diagonal blocks (of the same sizes for all members of $\mathcal{H}$), are again orthogonal and result from the original matrices by a transformation $H \mapsto WHW'$ with an orthogonal matrix W. Such a transformation leads to an isomorphic group, and it does not violate irreducibility or the property of being lattice generating. By repeating this construction until all blocks are irreducible, we eventually have an orthogonal matrix W and numbers $m_j \geq 1$ such that the matrices in $W\mathcal{H}W'$ are all of the form $\mathrm{diag}(H_1, \ldots, H_\ell)$ with $H_j \in \mathrm{O}(m_j)$. The map $\mathcal{H} \to \mathrm{O}(m_j)$ given by $H \mapsto H_j$ is a group homomorphism. Hence the image of this map is a subgroup $\mathcal{M}_j$ of $\mathrm{O}(m_j)$. This subgroup is irreducible, lattice generating, and isomorphic to a subgroup of $\mathbf{Z}_3^2$.

Since all $\mathcal{M}_j$ are abelian, we conclude that $m_j \leq 2$ for all j. Now the fact that $\mathcal{G}$ is both irreducible and lattice generating comes into play. Lemma 2.2 implies that all matrices in $\mathcal{G}$, and in particular those in $\mathcal{H}$, must have equal blocks. Consequently, if $m_j = 1$ for one j, then $m_j = 1$ for all j and it follows that $W\mathcal{H}W' \subset \{I, -I\}$, which is impossible because $W\mathcal{H}W'$ is isomorphic to $\mathbf{Z}_3^2$. Hence $m_j = 2$ for all j, and for each $H \in \mathcal{H}$ the blocks of H_j are Δ_3. Let A and B be two generators of $\mathcal{H}$. We so have

$$W\mathsf{A}W' = \mathrm{diag}(U_j \Delta_3 U_j'), \quad W\mathsf{B}W' = \mathrm{diag}(V_j \Delta_3 V_j').$$

Put $U = \mathrm{diag}(U_j)$. Then

$$A := U'W\mathsf{A}W'U = \mathrm{diag}(\Delta_3) =: \mathrm{diag}(A_j),$$
$$B := U'W\mathsf{B}W'U = \mathrm{diag}(Z_j \Delta_3 Z_j') =: \mathrm{diag}(B_j)$$

with $Z_j \in \mathrm{O}(2)$. But if $Z \in \mathrm{O}(2)$, then $Z\Delta_3 Z' = \Delta_3$ if $\det Z = 1$ and $Z\Delta_3 Z' = \Delta_3^2$ if $\det Z = -1$. It follows that each block B_j is Δ_3 or Δ_3^2. If $B_j = \Delta_3$, then the jth block of $U'W'\mathsf{AB}W'U$ is $A_j B_j = \Delta_3 \cdot \Delta_3 = \Delta_3^2$, while if $B_j = \Delta_3^2$, the jth block of $U'W'\mathsf{AB}W'U$ is $A_j B_j = \Delta_3 \cdot \Delta_3^2 = I$. Thus, as the number of blocks is at least two, AB does not have equal blocks if B contains two different blocks. Therefore either $B = \mathrm{diag}(\Delta_3)$ or $B = \mathrm{diag}(\Delta_3^2)$. We arrive at the conclusion that $\mathcal{H}$ is isomorphic

to the group generated by Δ_3 alone, that is, to $\mathbf{Z}_3$. Since $\mathcal{H}$ is isomorphic to $\mathbf{Z}_3^2$, this is a contradiction. □

The following lemma proves part of Theorem 1.4.

Lemma 2.5 *If $k \geq 4$ is even and $\mathcal{G} \subset \mathrm{SO}(k)$ is an irreducible and lattice generating group, the $\mathcal{G}$ is either isomorphic to the quaternion group Q_8 of order 8 or to the binary dihedral group Q_{12} of order 12 or to the binary tetrahedral group $2T$ of order 24.*

Proof If $\mathcal{G}$ is a finite group and a prime number p divides the order of $\mathcal{G}$, then $\mathcal{G}$ contains an element of order p (Cauchy's theorem). By Corollary 2.3, the orders of the elements of our group are 3, 4, 6, and hence the order of $\mathcal{G}$ must be $n = 2^r 3^s$. The Sylow theorems imply that if $\mathcal{G}$ is a finite group and p^ℓ is a prime power dividing the order of $\mathcal{G}$, then $\mathcal{G}$ has at least one subgroup of order p^ℓ. Thus, if $r \geq 4$, then $\mathcal{G}$ contains a subgroup of order 16. The groups Q_{16} and $\mathbf{Z}_{16}$ have an element of order 8 and the other 8 groups of order 16 all have at least two elements of order 2. By Corollary 2.3, this is impossible. If $s \geq 2$, the $\mathcal{G}$ contains a subgroup of order 9. Lemma 2.4 shows that this subgroup cannot be isomorphic to $\mathbf{Z}_3^2$, and it is also impossible that it is isomorphic to $\mathbf{Z}_9$, which contains elements of order 9. The remaining possible orders are $\{2, 3, 4, 6, 8, 12, 24\}$. In what follows we permanently employ Corollary 2.3 without mentioning this each time.

***n* = 2, 3, 4** These groups are abelian and hence not irreducible.
***n* = 6** C_6 is abelian and the symmetric group S_3 has 3 elements of order 2.
***n* = 8** C_8, $C_4 \times C_2$, C_2^3 are abelian and D_4 has more than one element of order 2. The only group remaining is Q_8.
***n* = 12** We could rule out the abelian groups immediately, but with the case $n = 24$ in mind, we argue as follows: C_{12} contains an element of order 12 and $C_3 \times C_2^2$ has 3 elements of order 2. The three non-abelian groups are A_4, the dihedral group D_6, and the binary dihedral group Q_{12}. The first two of them have more than one element of order 2, so that only Q_{12} is left.
***n* = 24** We may exclude the three abelian groups. There are 12 non-abelian groups. Except for $2T$, each of the remaining 11 groups contains a subgroup of order 12. From our arguments to settle the case $n = 12$ we know that, this time with the exception of Q_{12}, each of these subgroups has an element of order 12 or more than one element of order 2. The only of the 11 groups having Q_{12} as a subgroup is $Q_{12} \times C_2$. But this group has at least three elements of order 2. Thus, $2T$ is the only possible group. □

Lemma 2.6 *If $k \geq 6$ is even and a group $\mathcal{G} \subset \mathrm{O}(k)$ is isomorphic to Q_8, Q_{12}, or $2T$, then $\mathcal{G}$ is reducible.*

Proof The degrees of the irreducible representations over $\mathbf{C}$ of the three groups Q_8, Q_{12}, $2T$ are

$$Q_8 : 1, 1, 1, 1, 2, \qquad Q_{12} : 1, 1, 1, 1, 2, 2, \qquad 2T : 1, 1, 1, 2, 2, 2, 3, \tag{5}$$

and degrees of the irreducible representations over $\mathbf{R}$ of these groups are

$$Q_8 : 1, 1, 1, 1, 4, \qquad Q_{12} : 1, 1, 2, 2, 4 \qquad 2T : 1, 2, 3, 4, 4. \tag{6}$$

The assertion of the lemma is therefore immediate from (6). The lists (5) and the list for Q_8 in (6) are well known. The lists for Q_{12} and $2T$ in (6) are explicitly on the website [4] and in the lecture notes [6]. An alternative proof of the lemma based solely on (5) and the list for Q_8 in (6) is as follows.

If, for a finite group, the maximal degree of an irreducible representation over $\mathbf{C}$ is ℓ, then the maximal degree k of an irreducible representation over $\mathbf{R}$, that is, in $\mathrm{O}(k)$, satisfies $k \le 2\ell$. It follows that Q_8 and Q_{12} have no irreducible representations in $\mathrm{O}(k)$ for even $k \ge 6$ and that $2T$ has no irreducible representations in $\mathrm{O}(k)$ for even $k \ge 8$. So consider $2T \subset \mathrm{O}(6)$.

Let $\mathcal{H}$ be a subgroup of $\mathrm{O}(6)$ which is isomorphic to Q_8. We claim that this group has a 5-dimensional invariant subspace. The degrees of the irreducible representations of Q_8 in $\mathrm{O}(k)$ are $1, 1, 1, 1, 4$. Thus, $\mathcal{H}$ is reducible in dimensions different from 1 and 4. Taking this into account, we may proceed as in the proof of Lemma 2.4 to get an orthogonal matrix W such that each element of $W\mathcal{H}W'$ is of the form $\mathrm{diag}(H_1, \dots, H_\ell)$ with $H_j \in \mathrm{O}(m_j)$ and $m_j \in \{1, 4\}$ for all j. Accordingly,

$$\mathbf{R}^6 = X_1 \oplus \cdots \oplus X_\ell, \tag{7}$$

where each X_j is an m_j-dimensional invariant subspace of $W\mathcal{H}W'$. Since $m_1 + \cdots + m_\ell = 6$, we necessarily have $m_j = 1$ for some j. Replacing the X_j on the right of (7) by $\{0\}$, we get a 5-dimensional invariant subspace V of $W\mathcal{H}W'$, and hence $W'V$ is a 5-dimensional subspace of $\mathcal{H}$.

Suppose now $\mathcal{G} \subset \mathrm{O}(6)$ is isomorphic to $2T$. It is well known that $2T \cong Q_8 \rtimes \mathbf{Z}_3$, that is, every $G \in \mathcal{G}$ may (uniquely) be written as $G = HC_3^j$ with $j \in \{0, 1, 2\}$, H in a subgroup $\mathcal{H}$ isomorphic to Q_8, and an element $C_3 \in \mathcal{G}$ of order 3. From what was just proved we conclude that the subgroup $\mathcal{H}$ has a 5-dimensional invariant subspace $U := W'V \subset \mathbf{R}^6$. The image space $C_3U \subset \mathbf{R}^6$ also has the dimension 5. Consequently, $\dim(U \cap C_3U) \ge 4$. As C_3^2U is of dimension 5 as well, it follows that $U \cap C_3U \cap C_3^2U$ is of dimension at least 3. Take $u = C_3v = C_3^2w \ne 0$ from $U \cap C_3U \cap C_3^2U$. For every $H \in \mathcal{H}$ we then have $Hu \in U$, $HC_3u = Hw \in U$, and $HC_3^2u = Hv \in U$. Consequently, the orbit of u does not span $\mathbf{R}^6$, which proves that $\mathcal{G} \cong 2T$ cannot be irreducible in $\mathrm{O}(6)$. □

We finally prove the second half of Theorem 1.4.

Lemma 2.7 *The groups Q_8, Q_{12}, $2T$ have faithful irreducible and lattice generating representations in* SO(4).

Proof The quaternion group Q_8 is faithfully represented in SO(4) by

$$\begin{pmatrix} \mathbf{1} & 0 \\ 0 & \mathbf{1} \end{pmatrix}, \quad \begin{pmatrix} 0 & -\mathbf{1} \\ \mathbf{1} & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & \boldsymbol{i} \\ \boldsymbol{i} & 0 \end{pmatrix}, \quad \begin{pmatrix} -\boldsymbol{i} & 0 \\ 0 & \boldsymbol{i} \end{pmatrix},$$
$$\begin{pmatrix} -\mathbf{1} & 0 \\ 0 & -\mathbf{1} \end{pmatrix}, \quad \begin{pmatrix} 0 & \mathbf{1} \\ -\mathbf{1} & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & -\boldsymbol{i} \\ -\boldsymbol{i} & 0 \end{pmatrix}, \quad \begin{pmatrix} \boldsymbol{i} & 0 \\ 0 & -\boldsymbol{i} \end{pmatrix}$$

with

$$\mathbf{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \boldsymbol{i} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

This group contains $I, -I$ and six elements of order 4. This is the case $s = 3$ and $q = 0$ in (4). For $f = (a, b, c, d)' \in \mathbf{R}^4 \setminus \{0\}$ the matrix

$$F = \begin{pmatrix} G_1 f & G_2 f & \dots & G_8 f \end{pmatrix}$$

is

$$F = \begin{pmatrix} a & -c & -d & b & -a & c & d & -b \\ b & -d & c & -a & -b & d & -c & a \\ c & a & -b & -d & -c & -a & b & d \\ d & b & a & c & -d & -b & -a & -c \end{pmatrix}.$$

Multiplying F from the right by $\xi = (x, y, z, w, p, q, r, s)' \in \mathbf{Z}^8$ and putting

$$X = x - p, \quad Y = y - q, \quad Z = z - r, \quad W = w - s,$$

we get after some elementary computations that

$$F\xi = \begin{pmatrix} a & -c & -d & b \\ b & -d & c & -a \\ c & a & -b & -d \\ d & b & a & c \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ W \end{pmatrix} =: Av.$$

The matrix A is $\sqrt{a^2 + b^2 + c^2 + d^2}$ times an orthogonal matrix and therefore invertible. Thus,

$$\|F\xi\|^2 = \|Av\|^2 \geq m\|v\|^2$$

with some $m > 0$, and since $m\|v\|^2$ does not assume values in $(0, m)$, it follows that $F\mathbf{Z}^8$ is a lattice. Finally, the span of $\{Gf : G \in \mathcal{G}\}$ is the span of the columns of F.

The first four columns form the matrix A, and since this matrix is invertible, F has the rank 4, which shows that the columns of F span all of $\mathbf{R}^4$. Thus, $\mathcal{G}$ is irreducible.

The group $Q_{12} \subset \mathrm{SO}(4)$ is faithfully represented by the 12 matrices

$$\begin{aligned}
I &= \begin{pmatrix} \mathbf{1} & 0 \\ 0 & \mathbf{1} \end{pmatrix}, & -I &= \begin{pmatrix} -\mathbf{1} & 0 \\ 0 & -\mathbf{1} \end{pmatrix}, \\
A_1 &= \begin{pmatrix} 0 & -\mathbf{1} \\ \mathbf{1} & 0 \end{pmatrix}, & -A_1 &= \begin{pmatrix} 0 & \mathbf{1} \\ -\mathbf{1} & 0 \end{pmatrix}, \\
A_2 &= \begin{pmatrix} 0 & -\boldsymbol{\omega} \\ -\boldsymbol{\omega}^2 & 0 \end{pmatrix}, & -A_2 &= \begin{pmatrix} 0 & \boldsymbol{\omega} \\ \boldsymbol{\omega}^2 & 0 \end{pmatrix}, \\
A_3 &= \begin{pmatrix} 0 & -\boldsymbol{\omega}^2 \\ -\boldsymbol{\omega} & 0 \end{pmatrix}, & -A_3 &= \begin{pmatrix} 0 & \boldsymbol{\omega}^2 \\ \boldsymbol{\omega} & 0 \end{pmatrix}, \\
B &= \begin{pmatrix} \boldsymbol{\omega} & 0 \\ 0 & -\boldsymbol{\omega}^2 \end{pmatrix}, & B^2 &= \begin{pmatrix} \boldsymbol{\omega}^2 & 0 \\ 0 & -\boldsymbol{\omega} \end{pmatrix}, \\
B^4 &= \begin{pmatrix} -\boldsymbol{\omega} & 0 \\ 0 & \boldsymbol{\omega}^2 \end{pmatrix}, & B^5 &= \begin{pmatrix} -\boldsymbol{\omega}^2 & 0 \\ 0 & \boldsymbol{\omega} \end{pmatrix},
\end{aligned}$$

where

$$\mathbf{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \boldsymbol{\omega} = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}, \quad \alpha = \frac{1}{2}, \quad \beta = -\frac{\sqrt{3}}{2}.$$

This group contains $I, -I$, six elements of order 4, two elements of order 6 (B and B^5), and two elements of order 3 (B^2 and B^4), that is, we have the situation $s = 3$ and $q = 2$ in (4).

That Q_{12} is irreducible and lattice generating can be proved as for Q_8. Given $f = (a, b, c, d)' \in \mathbf{R}^4 \setminus \{0\}$, let F be the matrix

$$F = \begin{pmatrix} G_1 f & G_2 f & \dots & G_{12} f \end{pmatrix},$$

where the ordering $G_1, \dots, G_{12}$ is

$$I, A_1, A_2, A_3, -A_1, -A_2, -A_3, B, B^2, B^3 = -I, B^4, B^5.$$

The first four columns of F are a scalar multiple of an orthogonal matrix, which implies that F has rank 4 and hence that Q_{12} is irreducible. If

$$\xi = (g, x, y, z, p, q, r, s, t, u, v, w)' \in \mathbf{Z}^{12},$$

then, again after elementary calculations,

$$F\xi = \begin{pmatrix} a & c & d & b \\ b & d & -c & -a \\ c & -a & b & -d \\ d & -b & -a & c \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ W \end{pmatrix} =: CV$$

with

$$X = g - u + \alpha(s + w - t - v), \quad Y = p - x + \alpha(z + q - y - r),$$
$$Z = \beta(y + z - q - r), \quad W = \beta(v + w - s - t).$$

The matrix C is a scalar multiple of an orthogonal matrix. It follows that

$$\|F\xi\|^2 = \|CV\|^2 \geq m\|V\|^2$$

with some $m > 0$. But

$$\begin{aligned} \|V\|^2 =& X^2 + Y^2 + Z^2 + W^2 \\ =& \frac{3}{4}(q + r - y - z)^2 + \frac{3}{4}(s + t - v - w)^2 \\ &+ \left(g - u + \frac{1}{2}(s + w - t - v)\right)^2 + \left(p - x + \frac{1}{2}(z + q - y - r)\right)^2 \end{aligned}$$

is either 0 or at least $1/4$, which shows that $F\mathbf{Z}^{12}$ is discrete and thus a lattice.

The group $2T$ is faithfully represented in SO(4) by the eight matrices

$$I = \begin{pmatrix} \mathbf{1} & 0 \\ 0 & \mathbf{1} \end{pmatrix}, \quad A_1 = \begin{pmatrix} 0 & -\mathbf{1} \\ \mathbf{1} & 0 \end{pmatrix}, \; -I, \quad -A_1,$$
$$A_2 = \begin{pmatrix} 0 & \boldsymbol{i} \\ \boldsymbol{i} & 0 \end{pmatrix}, \; A_3 = \begin{pmatrix} -\boldsymbol{i} & 0 \\ 0 & \boldsymbol{i} \end{pmatrix}, \; -A_2, \; -A_3,$$

with

$$\mathbf{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \boldsymbol{i} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

and by the sixteen matrices

$$\frac{1}{2}(\pm I \pm A_1 \pm A_2 \pm A_3).$$

We let $B_1, \dots, B_8$ denote the latter matrices with I having the $+$ sign and ordered according the lexicographic order $+++, ++-, +-+, +--, \dots$ of the remaining signs. The other eight matrices then are $-B_1, \dots, -B_8$. With these notations we are in the situation of Corollary 2.3 with $s = 3$ and $q = 8$. The rest of the proof is as for Q_8 and Q_{12}. Given $f = (a, b, c, d)' \in \mathbf{R}^4 \setminus \{0\}$, let F be the matrix

$$F = \begin{pmatrix} G_1 f & G_2 f & \dots & G_{24} f \end{pmatrix}$$

where $G_1, \dots, G_{24}$ are the matrices

$$I, A_1, A_2, A_3, B_1, \dots, B_8, -I, -A_1, -A_2, -A_3, -B_1, \dots, -B_8$$

in this order. The first four columns of F are the same as the first four columns we had for Q_8. Thus, they form scalar multiple of an orthogonal matrix, which implies that F has rank 4 and hence that $\mathcal{G}$ is irreducible. Let $\xi \in \mathbf{Z}^{24}$ be the vector

$$\begin{aligned} \xi = &(g_1, x_1, y_1, z_1, p_1, q_1, r_1, s_1, t_1, u_1, v_1, w_1, \\ &g_2, x_2, y_2, z_2, p_2, q_2, r_2, s_2, t_2, u_2, v_2, w_2)' \end{aligned}$$

and put

$$g = g_1 - g_2, \ x = x_1 - x_2, \ \dots, \ w = w_1 - w_2.$$

After elementary calculations we arrive at the equality

$$2F\xi = \begin{pmatrix} a & c & d & b \\ b & d & -c & -a \\ c & -a & b & -d \\ d & -b & -a & c \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ W \end{pmatrix} =: CV$$

with

$$\begin{aligned} X &= 2g + p + q + r + s + t + u + v + w, \\ Y &= -2x - p - q - r - s + t + u + v + w, \\ Z &= -2y - p - q + r + s - t - u + v + w, \\ W &= 2z + p - q + r - s + t - u + v - w. \end{aligned}$$

The matrix C is a scalar multiple of an orthogonal matrix. It follows that

$$4\|F\xi\|^2 = \|CV\|^2 \geq m\|V\|^2$$

with some $m > 0$ and since $\|V\|^2$ is either 0 or at least 1, we see that $F\mathbf{Z}^{24}$ is discrete and thus a lattice. □

As said, Lemmas 2.5, 2.6, 2.7 prove Theorem 1.4.

Proof of Theorem 1.5 The representation of $2T$ we used in the proof of Theorem 1.4 is the quaternionic representation. The complex representation over $\mathbf{R}$ can be constructed as follows. Put

$$E = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad I = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad J = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \quad K = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}$$

and let $B_1, \ldots, B_8$ be the matrices $(1/2)(E \pm I \pm J \pm K)$ listed according the lexicographic order of $+$ and $-$. Thus,

$$B_1 = \frac{1}{2}(E+I+J+K), \ B_2 = \frac{1}{2}(E+I+J-K), \ B_2 = \frac{1}{2}(E+I-J+K), \ \ldots.$$

Then the 24 matrices $\pm E, \pm I, \pm J, \pm K, \pm B_1, \ldots, \pm B_8$ are a faithful irreducible representation of $2T$ in U(2). We denote this matrix group simply by $2T$. The seven conjugacy classes are $\{E\}, \{-E\}, \{\pm I, \pm J, \pm K\}$,

$$\mathcal{C}_1 := \{B_1, B_4, B_6, B_7\}, \quad \mathcal{C}_2 := \{B_2, B_3, B_5, B_8\},$$

and $-\mathcal{C}_1, -\mathcal{C}_2$. Let $\omega = -1/2 + i\sqrt{3}/2$ and let $\varrho : 2T \to \mathbf{C}$ be the homomorphism sending $\pm E, \pm I, \pm J, \pm K$ to 1, the elements of $-\mathcal{C}_1 \cup \mathcal{C}_1$ to ω, and the elements of $-\mathcal{C}_2 \cup \mathcal{C}_2$ to ω^2. This is a representation of $2T$ in U(1). It follows that the map $\tau : 2T \to \mathrm{U}(2), G \mapsto \varrho(G)G$ is a faithful representation of $2T$. The representing matrices are complex of the form $A + iB$. Replacing each such matrix by

$$\begin{pmatrix} A & -B \\ B & A \end{pmatrix},$$

we get a faithful representation $\sigma : 2T \to \mathrm{O}(4)$. This is what is called a complex representation of $2T$ over $\mathbf{R}$. The matrices $\sigma(E), \sigma(I), \sigma(J), \sigma(K)$ are

$$\sigma(E) = \begin{pmatrix} E & 0 \\ 0 & E \end{pmatrix}, \quad \sigma(I) = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix},$$
$$\sigma(J) = \begin{pmatrix} 0 & -W \\ W & 0 \end{pmatrix}, \ \sigma(K) = \begin{pmatrix} 0 & -S \\ S & 0 \end{pmatrix}$$

with

$$W = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad S = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

The matrix $\sigma(B_1)$ equals

$$\sigma(B_1) = \begin{pmatrix} \delta & -\delta & -\gamma & \gamma \\ -\gamma & -\gamma & -\delta & -\delta \\ \gamma & -\gamma & \delta & -\delta \\ \delta & \delta & -\gamma & -\gamma \end{pmatrix}, \quad \gamma = \frac{\sqrt{3}+1}{4}, \quad \delta = \frac{\sqrt{3}-1}{4}.$$

Now consider the matrix

$$F = \begin{pmatrix} \sigma(E)f & \sigma(I)f & \sigma(J)f & \sigma(K)f & \sigma(B_1)f & \dots \end{pmatrix}$$

with $f = (a, b, c, d)' \in \mathbf{R}^4$. (Incidentally, again the first columns of F form an orthogonal matrix, which shows that $\sigma(2T)$ is irreducible.) Taking f equal to $(1, 0, 0, 0)'$ we get

$$F = \begin{pmatrix} 1 & 0 & 0 & 0 & \delta & \dots \\ 0 & 1 & 0 & 0 & -\gamma & \dots \\ 0 & 0 & 0 & -1 & \gamma & \dots \\ 0 & 0 & 1 & 0 & \delta & \dots \end{pmatrix}.$$

Multiplying F from the right by $\xi = (x, y, z, w, p, 0, \dots, 0)' \in \mathbf{Z}^{24}$ we obtain

$$F\xi = (x + \delta p, y - \gamma p, -w + \gamma p, z + \delta p)'.$$

By Kronecker's theorem, $\mu\mathbf{Z}$ modulo 1 is dense in $(0, 1)$ whenever μ is irrational. Consequently, given $m \in \mathbf{N}$, there are

$$\xi_m = (x_m, y_m, z_m, w_m, p_m, 0, \dots, 0)' \in \mathbf{Z}^{24}$$

such that $\|F\xi_m\| \in (0, 1/m)$, which shows that $F\mathbf{Z}^{24}$ is not discrete and proves that $\sigma(2T)$ is not lattice generating. □

3 Groups of Unitary Matrices

The topic is trivial for unitary matrices, that is, for finite subgroups of $\mathrm{U}(k)$. A group $\mathcal{G} \subset \mathrm{U}(k)$ of order n is said to be lattice generating if, with F given by (1), $\Lambda(\mathcal{G}, f) = F\mathbf{Z}[i]^n$ is a discrete additive subgroup of $\mathbf{C}^k$ for every $f \in \mathbf{C}^k$.

Theorem 3.1 *The only irreducible and lattice generating subgroups of* $\mathrm{U}(k)$ *are* $\mathcal{G} = \{1, i, -1, -i\} \subset \mathrm{U}(1)$ *and its subgroups* $\{1\}$ *and* $\{1, -1\}$.

Proof Let $\mathcal{G} \subset \mathrm{U}(k)$ be an irreducible and lattice generating group. In the complex case, Theorem 2.1 reads as follows. *Suppose* $n \geq k \geq 1$. *Let* $\mathcal{G} \subset \mathrm{U}(k)$ *be a finite*

irreducible group of order n, let $f \in \mathbf{C}^k \setminus \{0\}$, and let F be the $k \times n$ matrix (1). Then the following hold.

(a) *We have $FF^* = \gamma I$ with some real $\gamma > 0$.*
(b) *The set $\Lambda(\mathcal{G}, f)$ is a lattice if and only if the Gram matrix $F^*F = ((G_j f, G_k f))_{j,k=1}^n$ belongs to $\mu\mathbf{Q}[i]^{n\times n}$ for some nonzero $\mu \in \mathbf{C}$.*
(c) *In case $\Lambda(\mathcal{G}, f)$ is a lattice, we actually have $F^*F \in \gamma\mathbf{Q}[i]^{n\times n}$.*
(d) *If $\|f\| = 1$, then $\Lambda(\mathcal{G}, f)$ is a lattice if and only if $F^*F \in \mathbf{Q}[i]^{n\times n}$.*

The reasoning of the proof of Lemma 2.2 shows that if $G = USU^*$ with a diagonal matrix S, then S must have equal entries, that is, $S = \omega I$ with some $\omega \in \mathbf{T}$. Consequently, $\mathcal{G}$ is composed of scalar multiples of the identity matrix. It follows that $\mathcal{G} = \{I, \omega I, \ldots, \omega^{n-1} I\}$ with $\omega = e^{2\pi i/n}$, which is reducible for $k \geq 2$. So let $k = 1$. For $|f| = 1$, the Gram matrix is $F^*F = (\overline{\omega}^k \omega^j)_{j,k=0}^{n-1}$. Since ω is an entry of this matrix, it must be in $\mathbf{Q}[i]$. The only such roots of unity are $\omega = 1, -1, i$. Conversely, it is clear, that the (irreducible) subgroups of $\{1, i, -1, -i\}$ are indeed lattice generating. □

Acknowledgments I sincerely thank Lenny Fukshansky, Christian Lehn, Josiah Park, and Dmytro Shklyarov for stimulating and helpful discussions.

References

1. A. Böttcher, L. Fukshansky, Addendum to "Lattices from equiangular tight frames". Linear Algebra Appl. **531**, 592–601 (2017)
2. L. Fukshansky, D. Needell, J. Park, Y. Xin, Lattices from tight frames and vertex transitive graphs. Electron. J. Comb. **26**(3), #P3.49, 30 (2019)
3. C. Gruson, V. Serganova, *A Journey Through Representation Theory* (Springer, Basel, 2018)
4. J. Montaldi, *Real Representations.* http://www.maths.manchester.ac.uk/~jm/wiki/Representations/Representations
5. I.M. Niven, *Irrational Numbers* (Wiley, New York, 1956)
6. M. Reif, in *Groups and Representations*. Lecture Notes for a Course Held (2018). http://homepages.warwick.ac.uk/~masda/MA3E1/GrRepns_2018.pdf
7. J.-P. Serre, *Linear Representations of Finite Groups* (Springer, New York, 1977)
8. S.F.D. Waldron, *An Introduction to Finite Tight Frames* (Birkhäuser, New York, 2018)

Invertibility Issues for Toeplitz Plus Hankel Operators and Their Close Relatives

Victor D. Didenko and Bernd Silbermann

Abstract The paper describes various approaches to the invertibility of Toeplitz plus Hankel operators in Hardy and l^p-spaces, integral and difference Wiener-Hopf plus Hankel operators and generalized Toeplitz plus Hankel operators. Special attention is paid to a newly developed method, which allows to establish necessary, sufficient and also necessary and sufficient conditions of invertibility, one-sided and generalized invertibility for wide classes of operators and derive efficient formulas for the corresponding inverses. The work also contains a number of problems whose solution would be of interest in both theoretical and applied contexts.

Keywords Toeplitz plus Hankel operators · Wiener-Hopf plus Hankel operators · Invertibility · Inverses

Mathematics Subject Classification (2010) Primary 47B35, 47B38; Secondary 47B33, 45E10

This work was supported by the Special Project on High-Performance Computing of the National Key R&D Program of China (Grant No. 2016YFB0200604), the National Natural Science Foundation of China (Grant No. 11731006) and the Science Challenge Project of China (Grant No. TZ2018001).

V. D. Didenko (✉)
Department of Mathematics, SUSTech International Center for Mathematics, Southern University of Science and Technology, Shenzhen, China

B. Silbermann
Technische Universität Chemnitz, Fakultät für Mathematik, Chemnitz, Germany
e-mail: silbermn@mathematik.tu-chemnitz.de

M. A. Bastos et al. (eds.), *Operator Theory, Functional Analysis and Applications*,
Operator Theory: Advances and Applications 282,
https://doi.org/10.1007/978-3-030-51945-2_7

1 Introduction

Toeplitz $T(a)$ and Hankel $H(b)$ operators appear in various fields of mathematics, physics, statistical mechanics and they have been thoroughly studied [7, 42, 47]. Toeplitz plus Hankel operators $T(a)+H(b)$ and Wiener-Hopf plus Hankel operators $W(a)+H(b)$ play an important role in random matrix theory [1, 6, 30] and scattering theory [34–36, 43, 45, 46, 56]. Although Fredholm properties and index formulas for such operators acting on different Banach and Hilbert spaces are often known—cf. [7, 13, 38–41, 48–52], their invertibility is little studied. So far progress has been made only in rare special cases. In this work, we want to present an approach, which allows to treat invertibility problem for a wide classes of Toeplitz plus Hankel operators on classic Hardy spaces and also for their close relatives: Toeplitz plus Hankel operators on l^p-spaces, generalized Toeplitz plus Hankel operators and integral and difference Wiener-Hopf plus Hankel operators. The operators acting on classical Hardy spaces are discussed in more details, whereas for other classes of operators we only provide a brief overview of the corresponding results.

2 Toeplitz and Hankel Operators on Hardy Spaces

Let $\mathbb{T} := \{t \in \mathbb{C} : |t| = 1\}$ be the counterclockwise oriented unit circle in the complex plane $\mathbb{C}$ and let $p \in [1, \infty]$. Consider the Hardy spaces

$$H^p = H^p(\mathbb{T}) := \{f \in L^p(\mathbb{T}) : \widehat{f_n} = 0 \quad \text{for all} \quad n < 0\},$$
$$\overline{H^p} = \overline{H^p}(\mathbb{T}) := \{f \in L^p(\mathbb{T}) : \widehat{f_n} = 0 \quad \text{for all} \quad n > 0\},$$

where $\widehat{f_n}$, $n \in \mathbb{N}$ are the Fourier coefficients of function f. Moreover, let I denote the identity operator, $P : L^p(\mathbb{T}) \to H^p(\mathbb{T})$ the projection defined by

$$P : \sum_{n=-\infty}^{\infty} \widehat{f_n} e^{in\theta} \mapsto \sum_{n=0}^{\infty} \widehat{f_n} e^{in\theta}$$

and $Q = I - P$. If $p \in (1, \infty)$, the Riesz projection P is bounded and $\operatorname{im} P = H^p$. Note that here we consider operators acting on the spaces L^p, H^p or l^p. In this connection, let us agree that whenever p and q appears in the text, they are related as $1/p + 1/q = 1$.

On the space H^p, $1 < p < \infty$, any function $a \in L^\infty$ generates two operators—viz. the Toeplitz operator $T(a) : \; : f \mapsto Paf$ and the Hankel operator $H(a) : f \mapsto PaQJf$, where $J : L^p \mapsto L^p$ is the flip operator,

$$(Jf)(t) := t^{-1} f\left(t^{-1}\right), \quad t \in \mathbb{T}.$$

We note that

$$J^2 = I, \quad JPJ = Q, \quad JQJ = P, \quad JaJ = \widetilde{a}, \quad \widetilde{a}(t) := a(1/t),$$

and the operators $T(a)$ and $H(b)$ are related to each other as follows

$$\begin{aligned} T(ab) &= T(a)T(b) + H(a)H(\widetilde{b}), \\ H(ab) &= T(a)H(b) + H(a)T(\widetilde{b}). \end{aligned} \tag{2.1}$$

We now consider the invertibility of Toeplitz plus Hankel operators $T(a) + H(b)$ generated by L^∞-functions a and b and acting on a Hardy space $H^p(\mathbb{T})$. Observe that the matrix representation of such operators in the standard basis $\{t^n\}_{n=0}^\infty$ of $H^p(\mathbb{T})$ is

$$T(a) + H(b) \sim (\widehat{a}_{k-j} + \widehat{b}_{k+j+1})_{k,j=0}^\infty.$$

There are a variety of approaches to the study of their invertibility and we briefly discuss some of them.

2.1 Classical Approach: I. Gohberg, N. Krupnik and G. Litvinchuk

Let $\mathcal{L}(X)$ and $\mathcal{F}(X)$ be, respectively, the sets of linear bounded and Fredholm operators on the Banach space X. Besides, if $\mathcal{A}$ is a unital Banach algebra, then $G\mathcal{A}$ stands for the group of all invertible elements in $\mathcal{A}$.

Assume that $a \in GL^\infty$, $b \in L^\infty$ and set

$$V(a,b) := \begin{pmatrix} a - b\widetilde{b}\widetilde{a}^{-1} & d \\ -c & \widetilde{a}^{-1} \end{pmatrix},$$

where $c := \widetilde{b}\widetilde{a}^{-1}$, $d := b\widetilde{a}^{-1}$. Writing $T(a) \pm H(b) + Q$ for $(T(a) \pm H(b))P + Q$, we consider the operator $\operatorname{diag}(T(a) + H(b) + Q, T(a) - H(b) + Q)$ on $L^p(\mathbb{T}) \times L^p(\mathbb{T})$ and represent it in the form

$$\begin{aligned} &\begin{pmatrix} T(a) + H(b) + Q & 0 \\ 0 & T(a) - H(b) + Q \end{pmatrix} \\ &\qquad = A(T(V(a,b)) + \operatorname{diag}(Q, Q))B \end{aligned} \tag{2.2}$$

with invertible operators A and B. More precisely,

$$B = \begin{pmatrix} I & 0 \\ \widetilde{b}I & \widetilde{a}I \end{pmatrix} \begin{pmatrix} I & I \\ J & -J \end{pmatrix}.$$

and the operator A is also known but its concrete form is not important now.

An immediate consequence of Eq. (2.2) is that both operators $T(a) \pm H(b)$ are Fredholm if and only if the block Toeplitz operator $T(V(a,b))$ is Fredholm. This representation is of restricted use because there are piecewise continuous functions a, b such that only one of the operators $T(a) \pm H(b)$ is Fredholm. In addition, both operators $T(a) \pm H(b)$ can be Fredholm but may have different indices. Therefore, more efficient methods for studying Toeplitz plus Hankel operators are needed, especially for discontinuous generating functions a and b.

Let us recall that there is a well-developed Fredholm theory for the operators $T(a) + H(b)$ with generating functions a, b from the set of piecewise continuous functions $PC = PC(\mathbb{T})$—cf. [7] for operators acting on the space H^2 and [50] for the ones on H^p, $p \neq 2$. However, the defect numbers of these operators, conditions for their invertibility, and inverse operators can be rarely determined directly from Eq. (2.2).

2.2 Basor-Ehrhardt Approach

This approach is aimed at the study of defect numbers of $T(a) + H(b) \in \mathcal{F}(H^p)$ by means of a factorization theory. It is well-known that for $b = 0$, the problem can be solved by Wiener-Hopf factorization. Since this notion is important for what follows, we recall the definition here. Note that from now on, all operators are considered in the spaces L^p or H^p for $p \in (1, \infty)$. Recall that p and q are related as $1/p + 1/q = 1$.

Definition 2.1 We say that a function $a \in L^\infty$ admits a Wiener-Hopf factorization in L^p if it can be represented in the form

$$a = a_- \chi_m a_+, \tag{2.3}$$

where $a_+ \in H^q$, $a_+^{-1} \in H^p$, $a_- \in \overline{H^p}$, $a_-^{-1} \in \overline{H^q}$, $\chi_m(t) := t^m$, $m \in \mathbb{Z}$, the term $a_+^{-1} P(a_+\varphi)$ belongs to L^p for any φ from the set of all trigonometrical polynomials $\mathcal{P} = \mathcal{P}(\mathbb{T})$ and there is a constant c_p such that

$$||a_+^{-1} P(a_+\varphi)||_p \leq c_p ||\varphi||_p \quad \text{for all} \quad \varphi \in \mathcal{P}.$$

Theorem 2.2 (Simonenko [53]) *If $a \in L^\infty$, then $T(a) \in \mathcal{F}(H^p)$ if and only if $a \in GL^\infty$ and admits the Wiener-Hopf factorization (2.3) in L^p. In this case*

$$\operatorname{ind} T(a) = -m.$$

This result extends to the case of matrix-valued generating functions as follows.

Theorem 2.3 *If $a \in L^\infty_{N\times N}$, then $T(a) \in \mathcal{F}(H^p_N)$ if and only if $a \in GL^\infty_{N\times N}$ and admits a factorization $a = a_- d a_+$, where*

$$a_+ \in H^q_{N\times N},\ a_+^{-1} \in H^p_{N\times N},\ a_- \in \overline{H}^p_{N\times N},\ a_-^{-1} \in \overline{H}^q_{N\times N},$$
$$d = \operatorname{diag}(\chi_{k_1}, \chi_{k_2}, \cdots, \chi_{k_N}), \quad \kappa_1, \kappa_2, \cdots, \kappa_N \in \mathbb{Z},$$

the term $a_+^{-1} P(a_+\varphi)$ belongs to L^p_N for any $\varphi \in \mathcal{P}_N$ and there is a constant c_p such that

$$||a_+^{-1} P(a_+\varphi)||_p \le c_p ||\varphi||_p \quad \textit{for all} \quad \varphi \in \mathcal{P}_N.$$

Moreover, if $T(a) \in \mathcal{F}(H^p_N)$, then

$$\dim\ker T(a) = -\sum_{\kappa_j<0} \kappa_j, \quad \dim\operatorname{coker} T(a) = -\sum_{\kappa_j>0} \kappa_j.$$

The numbers κ_j, called partial indices, are uniquely defined. Moreover, in some sense, the Wiener-Hopf factorization is unique if it exists. For example, for $N = 1$ the uniqueness of factorization can be ensured by the condition $a_-(\infty) = 1$. We also note that if $T(a) \in \mathcal{F}(H^p)$ and $\operatorname{ind} T(a) \ge 0$ ($\operatorname{ind} T(a) \le 0$) then $T(a)$ is right-invertible (left-invertible) and if $\kappa := \operatorname{ind} T(a) > 0$, the functions

$$a_+^{-1}\chi_j, \quad j = 0, \ldots, \kappa - 1$$

form a basis in the space $\ker T(a)$ and one of the right-inverses has the form $T^{-1}(a\chi_\kappa)T(\chi_\kappa)$.

A comprehensive information about Wiener-Hopf factorization is provided in [11, 44] and in books, which deal with singular integral and convolution operators [7, 32, 32, 33]. In particular, Wiener-Hopf factorization furnishes conditions for invertibility of related operators. However, generally there are no efficient methods for constructing such factorizations and computing partial indices even for continuous matrix-functions. Therefore, in order to study the invertibility of Toeplitz plus Hankel operators, we have to restrict ourselves to suitable classes of generating functions.

In the beginning of this century, Ehrhardt [28, 29] developed a factorization theory to study invertibility for large classes of convolution operators with flips. Toeplitz plus Hankel operators are included in this general framework. In particular,

it was shown that an operator $T(a) + H(b)$, $a \in GL^\infty$, $b \in L^\infty$ is Fredholm if and only if the matrix-function

$$V^\tau(a,b) = \begin{pmatrix} b\widetilde{a}^{-1} & a - b\widetilde{b}\widetilde{a}^{-1} \\ \widetilde{a}^{-1} & -\widetilde{a}^{-1}\widetilde{b} \end{pmatrix}$$

admits a certain type of antisymmetric factorization. Moreover, the defect numbers of the operator $T(a)+H(b)$ can be expressed via partial indices of this factorization. However, it is not known how the partial indices of general matrix-functions can be determined.

It was already mentioned that there are functions $a, b \in L^\infty$ such that the operator $T(a) + H(b)$ is Fredholm but $T(a) - H(b)$ is not. In this case we can use the representation (2.2) and conclude that the matrix $V^\tau(a,b)$ does not admit a Wiener-Hopf factorization. Thus, if $V(a,b)$ admits a Wiener-Hopf factorization, then $V^\tau(a,b)$ has the antisymmetric factorization mentioned but not vice versa.

This discussion shows that one has to select a class of generating functions a and b such that the defect numbers of operators $T(a) + H(b) \in \mathcal{F}(H^p)$ can be determined. An important class of suitable generating functions $a \in GL^\infty(\mathbb{T})$, $b \in L^\infty(\mathbb{T})$ is given by the condition

$$a\widetilde{a} = b\widetilde{b}. \tag{2.4}$$

This class of pairs of functions first appears in [4] and [13]. Equation (2.4) was latter called the matching condition and the corresponding duo (a,b) a matching pair—cf. [16]. Let us note that Toeplitz plus Hankel operators of the form

$$T(a) \pm H(a),\ T(a) - H(at^{-1}),\ T(a) + H(at) \tag{2.5}$$

appear in random ensembles [1, 6, 30] and in numerical methods for singular integral equations on intervals [37]. The generating functions of the operators (2.5) clearly satisfy the matching condition (2.4).

It is notable that the Fredholmness of the operator $T(a) + H(b)$ implies that $a \in GL^\infty(\mathbb{T})$. Therefore, the term b in the matching pair (a,b) is also invertible in $L^\infty(\mathbb{T})$ and one can introduce another pair (c,d), called the subordinated pair for (a,b), with the functions c and d defined by

$$c := a/b = \widetilde{b}/\widetilde{a},\ d := a/\widetilde{b} = b/\widetilde{a}.$$

An important property of these functions is that $c\widetilde{c} = d\widetilde{d} = 1$. In what follows, any function $g \in L^\infty(\mathbb{T})$ satisfying the equation $g\widetilde{g} = 1$ is referred to as a matching function.

Let us point out that the sets of matching functions and matching pairs are quite large. In particular, we have:

1. Let $\mathbb{T}^+ := \{t \in \mathbb{T} : \Im t > 0\}$ be the upper half-circle. If an element $g_0 \in GL^\infty$, then

$$g(t) := \begin{cases} g_0(t) & \text{if } t \in \mathbb{T}^+ \\ g_0^{-1}(1/t) & \text{if } t \in \mathbb{T} \setminus \mathbb{T}^+ \end{cases},$$

   is a matching function.
2. If g_1, g_2 are matching functions, then the product $g = g_1 g_2$ is also a matching function.
3. If g is a matching function, then for any $a \in GL^\infty$ the duo (a, ag) is a matching pair.
4. Any matching pair (a, b), $a \in GL^\infty$ can be represented in the form (a, ag), where $g = \widetilde{a}\widetilde{b}^{-1}$ is a matching function.

In this section we discuss the Basor-Ehrhardt approach to the study of defect numbers of the operators $T(a) + H(b) \in \mathcal{F}(H^p(\mathbb{T}))$ if a and b are piecewise continuous functions satisfying the condition (2.4). Then we present an explicit criterion for the Fredholmness of such operators. Recall that the circle $\mathbb{T}$ is counterclockwise oriented and $f \in PC$ if and only if for any $t \in \mathbb{T}$, the one-sided limits

$$f^\pm(t) := \lim_{\varepsilon \to \pm 0} f(te^{i\varepsilon})$$

exist. Without loss of generality we assume that $a, b \in GPC$.

Theorem 2.4 (Basor and Ehrhardt [4]) *Let $a, b \in GPC$ form a matching pair and let (c, d) be the subordinated pair. The operator $T(a) + H(b)$ is Fredholm on the space H^p if and only if the following conditions hold:*

$$\frac{1}{2\pi} \arg c^-(1) \notin \left\{\frac{1}{2} + \frac{1}{2p} + \mathbb{Z}\right\}, \quad \frac{1}{2\pi} \arg \widetilde{d}^-(1) \notin \left\{\frac{1}{2} + \frac{1}{2q} + \mathbb{Z}\right\}, \tag{2.6}$$

$$\frac{1}{2\pi} \arg c^-(-1) \notin \left\{\frac{1}{2p} + \mathbb{Z}\right\}, \quad \frac{1}{2\pi} \arg \widetilde{d}^-(-1) \notin \left\{\frac{1}{2q} + \mathbb{Z}\right\}, \tag{2.7}$$

$$\frac{1}{2\pi} \arg \left(\frac{c^-(\tau)}{c^+(\tau)}\right) \notin \left\{\frac{1}{p} + \mathbb{Z}\right\}, \quad \frac{1}{2\pi} \arg \left(\frac{(\widetilde{d})^-(\tau)}{(\widetilde{d})^+(\tau)}\right) \notin \left\{\frac{1}{q} + \mathbb{Z}\right\}, \ \forall \tau \in \mathbb{T}^+. \tag{2.8}$$

The definition of d in [4] differs from the one used here. In fact, d in [4] corresponds to $\widetilde{d}$ here. We keep this notation for sake of easy comparability of the results.

Theorem 2.4 already shows the exceptional role of the endpoints $+1$ and -1 of the upper semicircle $\mathbb{T}^+$. To establish the index formula mentioned in [4], a more geometric interpretation of the Fredholm conditions (2.6)–(2.8) is needed. Here there are a few details from [4].

For $z_1, z_2 \in \mathbb{C}$ and $\theta \in (0,1)$, we consider the open arc $\mathcal{A}(z_1, z_2, \theta)$ connecting the points z_1 and z_2 and defined by

$$\mathcal{A}(z_1, z_2, \theta) := \left\{ z \in \mathbb{C} \setminus \{z_1, z_2\} : \frac{1}{2\pi} \arg\left(\frac{z - z_1}{z - z_2}\right) \in \{\theta + \mathbb{Z}\} \right\}.$$

For $\theta = 1/2$ this arc becomes a line segment and if $z_1 = z_2$ it is an empty set. Assuming that $a, b \in GPC$, $a\widetilde{a} = b\widetilde{b}$ and using the auxiliary functions c and $\widetilde{d}$, one can show that the conditions (2.6)–(2.8) mean that any of the arcs

$$\mathcal{A}\left(1, c^+(1); \frac{1}{2} + \frac{1}{2p}\right), \ \mathcal{A}\left(c^-(\tau), c^+(\tau); \frac{1}{p}\right), \ \mathcal{A}\left(c^-(-1), 1; \frac{1}{2p}\right),$$
$$\mathcal{A}\left(1, (\widetilde{d})^+(1); \frac{1}{2} + \frac{1}{2q}\right), \ \mathcal{A}\left((\widetilde{d})^-(\tau), (\widetilde{d})^+(\tau); \frac{1}{q}\right), \ \mathcal{A}\left((\widetilde{d})^-(-1), 1; \frac{1}{2q}\right),$$

where $\tau \in \mathbb{T}^+$, does not cross the origin. It is clear that one has to take into account only the jump discontinuity points since $c, \widetilde{d} \in GPC$ and consequently $c^{\pm}(\tau) \neq 0$ and $(\widetilde{d})^{\pm}(\tau) \neq 0$.

The functions c and $\widetilde{d}$ satisfy the condition $c\widetilde{c} = d\widetilde{d} = 1$, so that they are effectively defined by their values on $\mathbb{T}^+$ only. If we let τ run along $\mathbb{T}^+$ from $\tau = 1$ to $\tau = -1$, the image of the function c is the curve with possible jump discontinuities, starting at the point $c^+(1)$ and terminating at $c^-(-1)$. We now add the arcs $\mathcal{A}(c^-(\tau), c^+(\tau); 1/p)$ to any discontinuity point τ of c located on $\mathbb{T}^+$. Besides, if necessary we also add the arcs $\mathcal{A}(1, c^+(1); 1/2 + 1/(2p))$ and $\mathcal{A}(c^-(-1), 1; 1/(2p))$ connecting the endpoints $c^+(1)$ and $c^-(-1)$ with the point $\tau = 1$, respectively. That way, we obtain a closed oriented curve. If the operator $T(a) + H(b)$ is Fredholm, the curve does not cross the origin and we consider its winding number $\operatorname{wind}(c^{\#,p}) \in \mathbb{Z}$. Similar constructions lead to the curve $\widetilde{d}^{\#,q}$ with a winding number $\operatorname{wind}(\widetilde{d}^{\#,q}) \in \mathbb{Z}$. Now we can conclude that $T(a) + H(b) \in \mathcal{F}(H^p)$ if and only if the origin does not belong to the curve $c^{\#,p}$ or $\widetilde{d}^{\#,q}$.

Theorem 2.5 (Basor and Ehrhardt [4]) *Assume that $a, b \in GPC$ is a matching pair with the subordinated pair (c, d) such that the conditions (2.6)–(2.8) hold. Then $T(a) + H(b) \in \mathcal{F}(H^p)$ with the Fredholm index*

$$\operatorname{ind}(T(a) + H(b)) = \operatorname{wind}(\widetilde{d}^{\#,q}) - \operatorname{wind}(c^{\#,p}).$$

For what follows we need a definition.

Definition 2.6 (Basor and Ehrhardt [4]) A matching pair (a, b), $a, b \in L^\infty$ with the subordinated pair (c, d) satisfies the basic factorization condition in H^p if c and d admit the factorization of the form

$$c(t) = c_+(t) t^{2n} c_+^{-1}(t^{-1}), \quad n \in \mathbb{Z}, \tag{2.9}$$
$$\widetilde{d}(t) = (\widetilde{d})_+(t) t^{2m} (\widetilde{d})_+^{-1}(t^{-1}), \quad m \in \mathbb{Z} \tag{2.10}$$

and

$$(1+t)c_+(t) \in H^q, \quad (1-t)c_+^{-1}(t) \in H^p,$$
$$(1+t)(\widetilde{d})_+(t) \in H^p,\ (1-t)(\widetilde{d})_+^{-1}(t) \in H^q.$$

Note that the indices m, n are uniquely determined and the functions c_+ and d_+ are also unique up to a multiplicative constant. The representations (2.9), and (2.10) are called antisymmetric factorization of the functions c and $\widetilde{d}$, respectively.

Theorem 2.7 (Basor and Ehrhardt [4]) *If (a, b), $a, b \in PC$ is a matching pair and $T(a) + H(b) \in \mathcal{F}(H^p)$, then (a, b) satisfies the basic factorization condition.*

The next result allows to determine the defect numbers of the operators $T(a) + H(b)$ in certain situations.

Theorem 2.8 (Basor and Ehrhardt [4]) *Assume that the matching pair (a, b), $a, b \in L^\infty$ satisfies the basic factorization condition in H^p with $m, n \in \mathbb{Z}$. If $T(a) + H(b) \in \mathcal{F}(H^p)$, then*

$$\dim\ker(T(a)+H(b)) = \begin{cases} 0 & \text{if} \quad n > 0, m \le 0, \\ -n & \text{if} \quad n \le 0, m \le 0, \\ m-n & \text{if} \quad n \le 0, m > 0, \\ \dim\ker A_{n,m} & \text{if} \quad n > 0, m > 0, \end{cases}$$

$$\dim\ker(T(a)+H(b))^* = \begin{cases} 0 & \text{if} \quad m > 0, n \le 0, \\ -m & \text{if} \quad m \le 0, n \le 0, \\ n-m & \text{if} \quad m \le 0, n > 0, \\ \dim\ker(A_{n,m})^{\mathsf{T}} & \text{if} \quad m > 0, n > 0, \end{cases}$$

Therein, in case $n > 0$, $m > 0$,

$$A_{n,m} := \left[\rho_{i-j} + \rho_{i+j}\right]_{i=0,j=0}^{n-1,m-1}$$

and

$$\rho(t) := t^{-m-n}(1+t)(1+t^{-1})\widetilde{c}_+\widetilde{d}_+b^{-1} \in L^1(\mathbb{T}).$$

In particular, the Fredholm index of $T(a) + H(b)$ is $m - n$.

Let us now briefly discuss the notion of the adjoint operator used in [4]. If we identify the dual to H^p with H^q via the mapping $g \in H^q \mapsto < g, \cdot > \in (H^p)'$,

$$< g, f > := \frac{1}{2\pi}\int_0^{2\pi} g(e^{-i\theta})f(e^{i\theta})\,d\theta,$$

then the adjoint operator to $T(a) + H(b)$ has the form $T(\widetilde{a}) + H(b)$, so that

$$\dim\ker(T(a) + H(b))^* = \dim\ker(T(\widetilde{a}) + H(b)).$$

A natural question to ask is whether the Fredholmness of $T(a) + H(b)$, $a, b \in L^\infty$ implies the existence of the antisymmetric factorizations of the functions c and $\widetilde{d}$. This problem has been also discussed in [4] and an example there shows that this is not true in general.

The theory above has been used to examine the operators (2.5) and also the operators $I + H(b)$ with matching functions b. We are not going to discuss this approach here. However, in what follows, we present a simple method to handle these operators.

2.3 Classical Approach Revisited

Now we turn attention to another method based on the classical approach and recently developed by the authors of this paper. Let us start with a special factorization of the operator $T(V(a, b))$ for generating functions a and b constituting a matching pair. If $a \in GL^\infty(\mathbb{T})$, $b \in L^\infty(\mathbb{T})$ satisfy the matching condition (2.4), the matrix function $V(a, b)$ in (2.2) has the form

$$V(a,b) = \begin{pmatrix} 0 & d \\ -c & \widetilde{a}^{-1} \end{pmatrix}, \quad c = \frac{a}{b}, \quad d = \frac{a}{\widetilde{b}}. \tag{2.11}$$

It follows that the corresponding block Toeplitz operator $T(V(a, b))$ with the generating matrix-function (2.11) can be represented in the form

$$\begin{aligned} T(V(a,b)) &= \begin{pmatrix} 0 & T(d) \\ -T(c) & T(\widetilde{a}^{-1}) \end{pmatrix} \\ &= \begin{pmatrix} -T(d) & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} 0 & -I \\ I & T(\widetilde{a}^{-1}) \end{pmatrix} \begin{pmatrix} -T(c) & 0 \\ 0 & I \end{pmatrix}, \end{aligned} \tag{2.12}$$

with the invertible operator

$$\begin{pmatrix} 0 & -I \\ I & T(\widetilde{a}^{-1}) \end{pmatrix} : H^p \times H^p \to H^p \times H^p.$$

This representations turns out to be extremely useful in the study of Toeplitz plus Hankel operators. To show this we start with the Coburn-Simonenko theorem. For Toeplitz operators this theorem indicates that for any non-zero $a \in L^\infty(\mathbb{T})$ one has

$$\min\{\dim\ker T(a), \dim\operatorname{coker} T(a)\} = 0.$$

It follows that Fredholm Toeplitz operators with index zero are invertible. However, in general for block-Toeplitz and for Toeplitz plus Hankel operators, Coburn-Simonenko theorem is not true. This causes serious difficulties when studying the invertibility of the operators involved. Nevertheless, the following theorem holds.

Theorem 2.9 *Let $a \in GL^{\infty}(\mathbb{T})$ and A refer to one of the operators $T(a) + H(at)$, $T(a) - H(at^{-1})$, $T(a) \pm H(a)$. Then*

$$\min\{\dim\ker A, \dim\operatorname{coker} A\} = 0.$$

Proof For $a \in PC(\mathbb{T})$ this result goes back to Basor and Ehrhardt [2, 3, 28] with involved proofs. However, there is an extremely simple proof—cf. [16], based on the representation (2.2) and valid for generating functions $a \in GL^{\infty}$. We would like to sketch this proof here. Thus one of the consequences of Eq. (2.2) is that there is an isomorphism between the kernels of the operators $T(V(a, b))$ and $\operatorname{diag}(T(a) + H(b), T(a) - H(b))$. Let us start with the operators $T(a) \pm H(a)$. The corresponding subordinated pairs (c, d) have the form $(\pm 1, a\widetilde{a}^{-1})$, the third operator in Eq. (2.12) is $\operatorname{diag}(\mp I, I)$, so that it does not influence the kernel and the image of $T(V(a, \pm a))$. Considering the two remaining operators in (2.12), we note that the Coburn-Simonenko theorem is valid for the block Toeplitz operator $T(V(a, \pm a))$ and hence for $T(a) \pm H(a)$.

Consider now the operators $T(a) + H(at)$. The duo (a, at) is a matching pair with the subordinated pair (t^{-1}, d), $d = a\widetilde{a}^{-1}t$. The operator $T(t^{-1})$ is Fredholm and $\operatorname{ind} T(t^{-1}) = 1$, so that $\operatorname{im} T(t^{-1}) = H^p$. Besides, a direct check shows that the function $e := e(t) = 1$, $t \in \mathbb{T}$ belongs to the kernels of both operator $T(t^{-1})$ and $T(a) - H(at)$. Assuming that $\dim\ker T(d) > 0$ and using Coburn-Simonenko theorem for Toeplitz operators, we note that $\operatorname{im} T(d)$ is dense in H^p. On the other hand, the factorization (2.12) and Eq. (2.2) yield that both spaces $\operatorname{im} T(V(a, at))$ and $\operatorname{im}\operatorname{diag}(T(a) + H(at), T(a) - H(at))$ are dense in $H^p \times H^p$. Hence,

$$\operatorname{coker}(T(a) + H(at)) = \operatorname{coker}(T(a) - H(at) = \{0\}.$$

Passing to the case $\dim\ker T(d) = 0$, we first note that

$$1 = \dim\ker T(t^{-1})) = \dim\ker\operatorname{diag}((T(a) + H(at), T(a) - H(at)),$$

and since the kernel of the operator $T(a) - H(at)$ contains the function $e(t) = 1$, it follows that

$$\ker(T(a) + H(at)) = \{0\}.$$

Thus if $\dim\ker T(a\widetilde{a}^{-1}t) > 0$, then $\operatorname{coker}(T(a) + H(at)) = \{0\}$, otherwise $\ker(T(a) + H(at)) = \{0\}$ and the Coburn-Simonenko theorem is proved for the operators $T(a) + H(at)$. The operators $T(a) - H(at^{-1})$ can be considered analogously [16]. □

Theorem 2.9 can be extended in a few directions—cf. Proposition 3.9 and Corollary 3.10 below. It is also valid for Toeplitz plus Hankel operators acting on l^p-spaces [14] and for Wiener-Hopf plus Hankel integral operators acting on $L^p(\mathbb{R}^+)$-spaces [21].

3 Kernel Representations

As was already mentioned, Eq. (2.2) is of limited use in studying the Fredholmness of the operators $T(a) + H(b)$. However, if the generating functions a and b satisfy the matching condition, the representation (2.2) allows to determine defect numbers and obtain efficient representations for the kernels and cokernels of Toeplitz plus Hankel operators. In order to present the method, we recall relevant results for operators acting on Hardy spaces—cf. [16]. Let us start with the connections between the kernels of Toeplitz plus/minus operators and the kernels of the corresponding block Toeplitz operators. The following lemma is a direct consequence of Eq. (2.2) and is valid even if a and b do not constitute a matching pair.

Lemma 3.1 *If $a \in GL^\infty$, $b \in L^\infty$ and the operators $T(a) \pm H(b)$ are considered on the space H^p, $1 < p < \infty$, then the following relations hold:*

- *If $(\varphi, \psi)^T \in \ker T(V(a,b))$, then*

$$(\Phi, \Psi)^T = \frac{1}{2}(\varphi - JQc\varphi + JQ\widetilde{a}^{-1}\psi, \varphi + JQc\varphi - JQ\widetilde{a}^{-1}\psi)^T \in \ker \operatorname{diag}(T(a) + H(b), T(a) - H(b)). \tag{3.1}$$

- *If $(\Phi, \Psi)^T \in \ker \operatorname{diag}(T(a) + H(b), T(a) - H(b))$, then*

$$(\Phi + \Psi, P(\widetilde{b}(\Phi + \Psi) + \widetilde{a}JP(\Phi - \Psi))^T \in \ker T(V(a,b)). \tag{3.2}$$

Moreover, the operators

$$E_1 : \ker T(V(a,b)) \to \ker \operatorname{diag}(T(a) + H(b), T(a) - H(b)),$$
$$E_2 : \ker \operatorname{diag}(T(a) + H(b), T(a) - H(b)) \to \ker T(V(a,b)),$$

defined, respectively, by Eqs. (3.1) *and* (3.2) *are mutually inverses to each other.*

3.1 Subordinated Operators and Kernels of $T(a) + H(b)$

Thus, if the kernel of the operator $T(V(a,b))$ is known, the kernels of both operators $T(a) + H(b)$ and $T(a) - H(b)$ can be also determined. However, the kernels of the

operators $T(V(a,b))$ are known only for a few special classes of matrices $V(a,b)$, and in the case of general generating functions $a, b \in L^\infty$ the kernel $T(V(a,b))$ is not available. The problem becomes more manageable if a and b form a matching pair. In this case, $V(a,b)$ is a triangular matrix—cf. (2.11) and the subordinated functions c and d satisfy the equations

$$c\widetilde{c} = 1, \quad d\widetilde{d} = 1.$$

Moreover, it follows from Eq. (2.12) that for any function $\varphi \in \ker T(c)$, the vector $(\varphi, 0)^\intercal$ belongs to the kernel of the operator $T(V(a,b))$ and the first assertion in Lemma 3.1 shows that

$$\begin{aligned} \varphi - JQcP\varphi &\in \ker(T(a) + H(b)), \\ \varphi + JQcP\varphi &\in \ker(T(a) - H(b)). \end{aligned} \tag{3.3}$$

Another remarkable fact is that both functions $\varphi + JQcP\varphi$ and $\varphi - JQcP\varphi$ also belong to the kernel of the operator $T(c)$. In order to show this, we need an auxiliary result.

Lemma 3.2 *Let $g \in L^\infty$ satisfy the relation $g\widetilde{g} = 1$ and $f \in \ker T(g)$. Then*

$$JQgPf \in \ker T(g) \quad \text{and} \quad (JQgP)^2 f = f. \tag{3.4}$$

Proof We only check the first relation (3.4)). Thus

$$T(g)(JQgPf) = PgPJQgPf = JQ\widetilde{g}QgPf = JQ\widetilde{g}gPf - JQ\widetilde{g}PgPf = 0,$$

so that $JQgPf \in \ker T(g)$. □

Considering the operators $\mathbf{P}_g^\pm : \ker T(g) \to \ker T(g)$,

$$\mathbf{P}_g^\pm := \frac{1}{2}(I \pm JQgP)\Big|_{\ker T(g)},$$

we observe that according to Lemma 3.2, one has $(\mathbf{P}_g^\pm)^2 = \mathbf{P}_g^\pm$. Therefore, $\mathbf{P}_g^\pm$ are complementary projections. This property and Eqs. (3.3) lead to the following conclusion.

Corollary 3.3 *If (c,d) is the subordinated pair for the matching pair $(a,b) \in L^\infty \times L^\infty$, then*

$$\begin{aligned} \ker T(c) &= \operatorname{im} \mathbf{P}_c^- \dotplus \operatorname{im} \mathbf{P}_c^+, \\ \operatorname{im} \mathbf{P}_c^- &\subset \ker(T(a) + H(b)), \\ \operatorname{im} \mathbf{P}_c^+ &\subset \ker(T(a) - H(b)). \end{aligned}$$

Corollary 3.3 shows the impact of the operator $T(c)$ on the kernels of $T(a) + H(b)$ and $T(a) - H(b)$. The impact of the operator $T(d)$ on $\ker(T(a) \pm H(b))$ is much more involved. Thus if $T(c)$ is right-invertible and $s \in \ker T(d)$, then

$$(T_r^{-1}(c)T(\widetilde{a}^{-1})s, s)^{\intercal} \in \ker T(V(a,b)),$$

where $T_r^{-1}(c)$ is one of the right-inverses for $T(c)$.

We now can obtain the following representation of the kernel of the operator $T(V(a,b))$.

Lemma 3.4 *Let (a,b) be a matching pair such that the operator $T(c)$ is invertible from the right. Then*

$$\ker T(V(a,b)) = \Omega(c) \dotplus \widehat{\Omega}(d)$$

where

$$\Omega(c) := \left\{(\varphi, 0)^T : \varphi \in \ker T(c)\right\},$$
$$\widehat{\Omega}(d) := \left\{(T_r^{-1}(c)T(\widetilde{a}^{-1})s, s)^T : s \in \ker T(d)\right\}.$$

Taking into account this representation and using the first assertion in Lemma 3.1, we obtain that if $s \in \ker T(d)$, then

$$\begin{aligned} 2\varphi_{\pm}(s) = T_r^{-1}(c)T(\widetilde{a}^{-1})s \mp JQcPT_r^{-1}(c)T(\widetilde{a}^{-1})s \pm JQ\widetilde{a}^{-1}s \\ \in \ker(T(a) \pm H(b)) \end{aligned} \tag{3.5}$$

The operators $\varphi_{\pm}$ can be referred as transition operators, since they transform the kernel of $T(d)$ into the kernels of the operators $T(a) \pm H(b)$. An important property of these operators $\varphi_{\pm}$ is expressed by the following lemma.

Lemma 3.5 *Let (c,d) be the subordinated pair for a matching pair $(a,b) \in L^{\infty} \times L^{\infty}$. If the operator $T(c)$ is right-invertible, then for every $s \in \ker T(d)$ the following relations*

$$(P\widetilde{b}P + P\widetilde{a}JP)\varphi_+(s) = \mathbf{P}_d^+(s),$$
$$(P\widetilde{b}P - P\widetilde{a}JP)\varphi_-(s) = \mathbf{P}_d^-(s),$$

hold. Thus the transition operators φ_+ and φ_- are injections on the spaces $\operatorname{im}\mathbf{P}_d^+$ *and* $\operatorname{im}\mathbf{P}_d^-$*, respectively.*

Using Lemmas 3.1–3.5, one can obtain a complete description for the kernel of the operator $T(a) + H(b)$ if (a,b) is a matching pair and $T(c)$ is right-invertible.

Proposition 3.6 (VD and BS [16]) *Let (c, d) be the subordinated pair for the matching pair $(a, b) \in L^\infty \times L^\infty$. If the operator $T(c)$ is right-invertible, then the kernels of the operators $T(a) \pm H(b)$ can be represented in the form*

$$\ker(T(a) + H(b)) = \varphi_+(\operatorname{im} \mathbf{P}_d^+) \dotplus \operatorname{im} \mathbf{P}_c^-,$$
$$\ker(T(a) - H(b)) = \varphi_-(\operatorname{im} \mathbf{P}_d^-) \dotplus \operatorname{im} \mathbf{P}_c^+.$$

Remark 3.7 It was shown in [16] that the operators φ_+ and φ_- send the elements of the spaces $\operatorname{im} \mathbf{P}_d^-$ and $\operatorname{im} \mathbf{P}_d^+$ into $\operatorname{im} \mathbf{P}_c^-$ and $\operatorname{im} \mathbf{P}_c^+$, respectively. Therefore,

$$\varphi_+ : \operatorname{im} \mathbf{P}_d^- \to \operatorname{im} \mathbf{P}_c^-, \quad \varphi_- : \operatorname{im} \mathbf{P}_d^+ \to \operatorname{im} \mathbf{P}_c^+$$

are well-defined linear operators. If $\operatorname{im} \mathbf{P}_c^- = \{0\}$ ($\operatorname{im} \mathbf{P}_c^+ = \{0\}$), then $\varphi_+(s_-) = 0$ for all $s_- \in \operatorname{im} \mathbf{P}_d^-$ ($\varphi_-(s_+) = 0$ for all $s_+ \in \operatorname{im} \mathbf{P}_d^+$), which yields

$$\varphi_-(s_-) = T_r^{-1}(c) T(\widetilde{a}^{-1}) s_-, \quad s_- \in \operatorname{im} \mathbf{P}_d^-,$$
$$\left(\varphi_+(s_+) = T_r^{-1}(c) T(\widetilde{a}^{-1}) s_+, \quad s_+ \in \operatorname{im} \mathbf{P}_d^+\right).$$

Indeed, assume for instance that $s_- \in \operatorname{im} \mathbf{P}_d^-$. Then $\varphi_+(s_-) = 0$ leads to

$$0 = 2\varphi_+(s_-) = T_r^{-1}(c) T(\widetilde{a}^{-1}) s_- - J Q c P T_r^{-1}(c) T(\widetilde{a}^{-1}) s_- + J Q \widetilde{a}^{-1} s_-.$$

Hence,

$$T_r^{-1}(c) T(\widetilde{a}^{-1}) s_- = J Q c P T_r^{-1}(c) T(\widetilde{a}^{-1}) s_- - J Q \widetilde{a}^{-1} s_-.$$

Thus

$$\begin{aligned} 2\varphi_-(s_-) &= T_r^{-1}(c) T(\widetilde{a}^{-1}) s_- + J Q c P T_r^{-1}(c) T(\widetilde{a}^{-1}) s_- - J Q \widetilde{a}^{-1} s_- \\ &= 2T_r^{-1}(c) T(\widetilde{a}^{-1}) s_- \end{aligned}$$

and the claim follows.

These representations of the transition operators are simpler than (3.5) and it would be interesting to find out which conditions ensure that the restrictions $\varphi_+ \big|_{\operatorname{im} \mathbf{P}_d^-}$ and $\varphi_- \big|_{\operatorname{im} \mathbf{P}_d^+}$ become zero functions.

Thus in order to obtain an efficient description of the spaces $\ker(T(a) + H(b))$ and $\ker(T(a) - H(b))$, one has to characterize the projections $\mathbf{P}_c^\pm$ and $\mathbf{P}_d^\pm$ first. Such a characterization can be derived from the Wiener-Hopf factorization (2.3) of the subordinated functions c and d. The Wiener-Hopf factorization of these functions can be described in more details, which yields a very comprehensive representation of the kernels of $T(c)$, $T(d)$ and the related projections $\mathbf{P}_c^\pm$, $\mathbf{P}_d^\pm$.

We first consider related constructions for a matching function g such that the operator $T(g)$ is Fredholm on H^p with the index n. One can show that under the condition $g_-(\infty) = 1$, the factorization (2.3) takes the form

$$g = \sigma(g) g_+ t^{-n} \widetilde{g}_+^{-1}, \tag{3.6}$$

where $\sigma(g) = g_+(0) = \pm 1$ and $g_- = \sigma(g)\widetilde{g}_+^{-1}$. The term $\sigma(g)$ is called factorization signature. The finding of $\sigma(g)$ is a non-trivial problem but if $T(g)$ is invertible and g is continuous at $t = 1$ or $t = -1$, then $n = 0$ and $\sigma(g) = g(1)$ or $\sigma(g) = g(-1)$, respectively. For piecewise continuous functions g, the term $\sigma(g)$ can be also determined.

Notice that $T(a) - H(b)$ can be also written as Toeplitz plus Hankel operator $T(a) + H(-b)$. Thereby, the duo $(a, -b)$ is again a matching pair with the subordinated pair $(-c, -d)$ and for the factorization signatures we have $\sigma(-c) = -\sigma(c)$, $\sigma(-d) = -\sigma(d)$. This observation shows that we can restrict ourselves to the study of Toeplitz plus Hankel operators. Nevertheless, in some cases it is preferable to consider the operator $T(a) - H(b)$ too. But then the leading role still belongs to the operator $T(a)+H(b)$ since the notions of matching pair, subordinated pair and factorization signature are associated with this operator.

Let g stand for the subordinated function c or d, so that $g\widetilde{g} = 1$. If $T(g)$ is Fredholm, then the factorization (3.6) exists with a function g_+ satisfying the conditions for factorization (2.3) and we can describe the spaces $\operatorname{im} \mathbf{P}_g^{\pm}$. This description depends on the evenness of the index of $T(g)$.

Theorem 3.8 (VD and BS [16]) *Assume that g is a matching function, the operator $T(g)$ is Fredholm,* $\operatorname{ind} T(g) = n \geq 0$ *and g_+ is the plus factor in the Wiener-Hopf factorization* (3.6) *of g in H^p. Then*

- *For $n = 2r$, $r \in \mathbb{N}$, the systems of functions*

$$\mathcal{B}_{\pm}(g) := \{g_+^{-1}(t^{r-k-1} \pm \sigma(g) t^{r+k}) : k = 0, 1, \cdots, r-1\},$$

  form bases in the spaces $\operatorname{im} \mathbf{P}_g^{\pm}$ *and* $\dim \ker \mathbf{P}_g^{\pm} = r$.
- *For $n = 2r + 1$, $r \in \mathbb{Z}_+$, the systems of functions*

$$\mathcal{B}_{\pm}(g) := \{g_+^{-1}(t^{r+k} \pm \sigma(g) t^{r-k}) : k = 0, 1, \cdots, r\} \setminus \{0\},$$

  form bases in the spaces $\operatorname{im} \mathbf{P}_g^{\pm}$ *and* $\dim \ker \mathbf{P}_g^{\pm} = r + (1 \pm \sigma(g))/2$.

Thus if $T(c), T(d)$ are Fredholm and $T(c)$ is right-invertible, Proposition 3.6 provides complete description of the spaces $\ker(T(a) \pm H(b))$. On the other hand, if $T(c)$ is Fredholm but not right-invertible, the representation

$$T(a) + H(b) = (T(at^{-n}) + H(bt^n))T(t^n)$$

can be used to study $\ker(T(a)+H(b))$. This is because for any matching pair (a,b) the duo (at^{-n}, bt^n) is also a matching pair with the subordinated pair (ct^{-2n}, d). A suitable choice of n leads to the right-invertibility of the operator $T(ct^{-2n})$ and we consequently obtain

$$\ker(T(a)+H(b)) = \ker(T(at^{-n})+H(bt^n)) \cap \operatorname{im} T(t^n). \tag{3.7}$$

The representation (3.7) has been used in [16], to derive the description of the kernels of the operators $T(a) \pm H(b)$. It can be also employed to study one-sided or generalized invertibility of Toeplitz plus Hankel operators and to construct the corresponding one-sided and generalized inverses [17, 19, 22]. In the forthcoming sections, invertibility problems will be discussed in more details. In this connection, we note that the (familiar) adjoint operator $(T(a)+H(b))^*$ can be identified with the operator $T(\overline{a})+H(\overline{\widetilde{b}})$ acting on the space H^q, $1/p+1/q=1$. It is easily seen that for any matching pair (a,b), the duo $(\overline{a}, \overline{\widetilde{b}})$ is also a matching pair with the subordinated pair $(\overline{d}, \overline{c})$ and $\boldsymbol{\sigma}(\overline{c}) = \boldsymbol{\sigma}(c)$, $\boldsymbol{\sigma}(\overline{d}) = \boldsymbol{\sigma}(d)$. Therefore, cokernel description can be determined directly from the previous results for the kernels of Toeplitz plus Hankel operators.

In some cases the approach above allows omitting the condition of Fredholmness of the operator $T(d)$. We note a few results, which can be viewed as an extension of Coburn-Simonenko Theorem 2.9.

Proposition 3.9 (VD and BS [16]) *Let $(a,b) \in L^\infty \times L^\infty$ be a matching pair with the subordinated pair (c,d), and let $T(c)$ be a Fredholm operator. Then:*

(a) If $\operatorname{ind} T(c) = 1$ *and* $\boldsymbol{\sigma}(c) = 1$, *then*

$$\min\{\dim\ker(T(a)+H(b)), \dim\operatorname{coker}(T(a)+H(b))\} = 0.$$

(b) If $\operatorname{ind} T(c) = -1$ *and* $\boldsymbol{\sigma}(c) = 1$, *then*

$$\min\{\dim\ker(T(a)-H(b)), \dim\operatorname{coker}(T(a)-H(b))\} = 0.$$

(c) If $\operatorname{ind} T(c) = 0$, *then*

$$\min\{\dim\ker(T(a)\pm H(b)), \dim\operatorname{coker}(T(a)\pm H(b))\} = 0.$$

An immediate consequence of Proposition 3.9 concerns the Toeplitz plus Hankel operators of the form $I \pm H(b)$.

Corollary 3.10 *Let $b \in L^\infty$ be a matching function such that $T(\widetilde{b})$ is a Fredholm operator. Then:*

(a) If $\operatorname{ind} T(\widetilde{b}) = 1$ *and* $\boldsymbol{\sigma}(\widetilde{b}) = 1$, *then*

$$\min\{\dim\ker(I+H(b)), \dim\operatorname{coker}(I+H(b))\} = 0.$$

(b) If $\operatorname{ind} T(\widetilde{b}) = -1$ *and* $\boldsymbol{\sigma}(\widetilde{b}) = 1$, *then*

$$\min\{\dim\ker(I - H(b)), \dim\operatorname{coker}(I - H(b))\} = 0.$$

(c) If $\operatorname{ind} T(\widetilde{b}) = 0$, *then*

$$\min\{\dim\ker(I \pm H(b)), \dim\operatorname{coker}(I \pm H(b))\} = 0.$$

3.2 *Kernels of* $T(a) + H(b)$ *for Piecewise Continuous Generating Functions*

If more information about generating functions is available, then the kernel of the Fredholm operator $T(a) + H(b) \in \mathcal{L}(H^p)$ can be studied under weaker conditions. Thus for piecewise continuous functions a and b, the assumption that the subordinated operators $T(c), T(d) \in \mathcal{L}(H^p)$ are Fredholm can be dropped. In order to show this we need a few facts from [16].

Let A be an operator defined on all spaces L^p for $1 < p < \infty$. Consider the set

$$A_F := \{p \in (1, \infty) \text{ such that the operator } A : H^p \to H^p \text{ is Fredholm}\}.$$

Proposition 3.11 (Šneĭberg [54]) *The set* A_F *is open. Moreover, for each connected component* $\gamma \in A_F$, *the index of the operator* $A : L^p \to L^p$, $p \in \gamma$ *is constant.*

This result also holds for operators A acting on the spaces H^p, $1 < p < \infty$, since any operator $A\colon H^p \to H^p$ can be identified with the operator $AP + Q$ acting already on L^p. For Toeplitz operators the structure of the set A_F can be characterized as follows.

Proposition 3.12 (Spitkovskiĭ [55]) *Let* G *be an invertible matrix–function with entries from* PC *and let* $A := T(G)$. *Then there is an at most countable subset* $S_A \subset (1, \infty)$ *with the only possible accumulation points* $t = 1$ *and* $t = \infty$ *such that* $A_F = (1, \infty) \setminus S_A$.

Clearly, if G is piecewise continuous with only a finite number of discontinuities, then S_A is at most finite. This result can be used to describe the corresponding set A_F for Toeplitz plus Hankel operators with PC-generating functions.

Corollary 3.13 *Let* $a, b \in PC$ *and*

$$A := \operatorname{diag}(T(a) + H(b), T(a) - H(b))\colon H^p \times H^p \to H^p \times H^p.$$

Then there is at most countable subset $S_A \subset (1, \infty)$ *with only possible accumulation points* $t = 1$ *and* $t = \infty$ *such that* $A_F = (1, \infty) \setminus S_A$.

Proof It follows directly from Proposition 3.12 since $\operatorname{diag}(T(a)+H(b), T(a)-H(b))$ is Fredholm if and only if so is the operator $T(V(a,b))$. □

Thus if $a, b \in PC$ and the operator $T(a)+H(b)$ is Fredholm on H^p, there is an interval (p', p'') containing p such that $T(a)+H(b)$ is Fredholm on all spaces H^r, $r \in (p', p'')$ and the index of this operator does not depend on r. Moreover, there is an interval $(p, p_0) \subset (p', p'')$, $p < p_0$ such that $T(a)-H(b)$ is Fredholm on H^r, $r \in (p, p_0)$ and its index does not depend on r. Recalling that for $\nu < s$ the space H^s is continuously embedded into H^ν, we note that the kernel of $T(a)+H(b) : H^r \to H^r$ does not depend on $r \in (p', p'')$. The same is also true for $\ker(T(a)-H(b))$, $r \in (p, p_0)$. We want to note that both claims are based on the following well-known result.

Lemma 3.14 (Gohberg and Feldman [31]) *Let a Banach space $\mathfrak{B}_1$ be continuously and densely embedded into a Banach space $\mathfrak{B}_2$. Further, let A_1 and A_2 be bounded Fredholm operators which respectively act on $\mathfrak{B}_1$ and $\mathfrak{B}_2$ and have equal indices. If A_2 is an extension of A_1, then*

$$\ker A_1 = \ker A_2.$$

Hence, the kernel of the operator $T(a)+H(b)$ acting on the space H^p coincides with the kernel of this operator acting on the space H^r for $r \in (p, p_0)$ and the latter can be studied by the approach above. Therefore, if $T(a)+H(b) \in \mathcal{L}(H^p)$ is Fredholm and $a, b \in PC$ form a matching pair, the kernel of the operator $T(a)+H(b)$ can be described.

4 Generalized Invertibility, One-Sided Invertibility, and Invertibility

Let $a, b \in L^\infty$ be a matching pair with the subordinated pair (c, d). The pair is called a Fredholm matching pair (with respect to H^p) if the operators $T(c), T(d) \in \mathcal{L}(H^p)$ are Fredholm. We write $T(a)+H(b) \in \mathcal{F}^p_{TH}(\kappa_1, \kappa_2)$ if (a, b) is a Fredholm matching pair with the subordinated pair (c, d) such that $\operatorname{ind} T(c) = \kappa_1$, $\operatorname{ind} T(d) = \kappa_2$. It was first noted in [15] that if $\kappa_1 \geq 0, \kappa_2 \geq 0$ or if $\kappa_1 \leq 0, \kappa_2 \leq 0$, then (2.2) and (2.12) yield one-sided invertibility of the operator $T(a)+H(b)$. However, if $\kappa_1\kappa_2 < 0$, the invertibility issues become more involved. We start this section by considering the generalized invertibility of the operators $T(a)+H(b)$, $a, b \in L^\infty$. Set

$$B := \mathcal{P}V(a,b)\mathcal{P} + \mathcal{Q},$$

where $\mathcal{P} := \operatorname{diag}(P, P)$, $\mathcal{Q} := \operatorname{diag}(Q, Q)$.

Theorem 4.1 (VD and BS [17]) *Assume that (a, b) is a matching pair with the subordinated pair (c, d) and B is generalized invertible operator, which has a generalized inverse B^{-1} of the form*

$$B^{-1} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{D} & 0 \end{pmatrix} + \mathcal{Q}, \tag{4.1}$$

where $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{D}$ are operators acting in the space H^p. Then the operator $T(a) + H(b) : H^p \to H^p$ is also generalized invertible and one of its generalized inverses has the form

$$\begin{aligned}(T(a) + H(b))_g^{-1} = &-H(\widetilde{c})(\mathbf{A}(I - H(d)) - \mathbf{B}H(\widetilde{a}^{-1})) \\ &+ H(a^{-1})\mathbf{D}(I - H(d)) + T(a^{-1}).\end{aligned}$$

This result can now be used to construct generalized inverses for the operator $T(a) + H(b)$ in the following cases—cf. [17]:

(a) If $\kappa_1 \geq 0$ and $\kappa_2 \geq 0$, then B has generalized inverse of the form (4.1) with $\mathbf{A} = T_r^{-1}(c)T(\widetilde{a}^{-1})T_r^{-1}(d)$, $\mathbf{B} = -T_r^{-1}(c)$ and $\mathbf{D} = T_r^{-1}(d)$.
(b) If $\kappa_1 \leq 0$ and $\kappa_2 \leq 0$, then B has generalized inverse of the form (4.1) with $\mathbf{A} = T_l^{-1}(c)T(\widetilde{a}^{-1})T_l^{-1}(d)$, $\mathbf{B} = -T_l^{-1}(c)$ and $\mathbf{D} = T_l^{-1}(d)$.
(c) If $\kappa_1 \geq 0$ and $\kappa_2 \leq 0$, then B has generalized inverse of the form (4.1) with $\mathbf{A} = T_r^{-1}(c)T(\widetilde{a}^{-1})T_l^{-1}(d)$, $\mathbf{B} = -T_r^{-1}(c)$ and $\mathbf{D} = T_l^{-1}(d)$.

It is clear that in the cases (a) and (b), generalized inverses are, respectively, right and left inverses. We also note that under conditions of assertion (a), a right inverse of $T(a) + H(b)$ can be written in a simpler form—cf. [19]

$$B := (I - H(\widetilde{c}))T_r^{-1}(c)T(\widetilde{a}^{-1})T_r^{-1}(d) + H(a^{-1})T_r^{-1}(d). \tag{4.2}$$

The proof of this result is straightforward—i.e. one can use the relations (2.1) to verify that $(T(a) + H(b))B = I$. On the other hand, under conditions of (b), a simpler representation of the left-inverse of $T(a) + H(b)$ can be derived from (4.2) by passing to the adjoint operator.

In addition to the cases considered, there is one more situation—viz.

(d) $\kappa_1 < 0, \kappa_2 > 0$.

This case is much more involved and we will deal with it later on. At the moment, we focus on invertibility of operators from $\mathcal{F}_{TH}^p(\kappa_1, \kappa_2)$, $1 < p < \infty$.

Theorem 4.2 (VD and BS [19]) *Assume that $T(a) + H(b) \in \mathcal{F}_{TH}^p(\kappa_1, \kappa_2)$ is invertible. Then:*

(i) If $\kappa_1 \geq \kappa_2$ or $\kappa_1\kappa_2 \geq 0$, then

$$|\kappa_1| \leq 1, \quad |\kappa_2| \leq 1. \tag{4.3}$$

(ii) If $\kappa_1 < 0$ and $\kappa_2 > 0$, then

(a) If κ_1 and κ_2 are even numbers, then $\kappa_2 = -\kappa_1$.
(b) If κ_1 is an odd number and κ_2 is an even one, then $\kappa_2 = -\kappa_1 + \boldsymbol{\sigma}(c)$.
(c) If κ_1 is an even number and κ_2 is an odd one, then $\kappa_2 = -\kappa_1 - \boldsymbol{\sigma}(d)$.
(d) If κ_1 and κ_2 are odd numbers, then $\kappa_2 = -\kappa_1 + \boldsymbol{\sigma}(c) - \boldsymbol{\sigma}(d)$.

Our next goal is to obtain sufficient invertibility conditions for the invertibility of the operators from $\mathcal{F}^p_{TH}(\kappa_1, \kappa_2)$ and to provide effective representations for their inverses. We assume first that κ_1 and κ_2 satisfy conditions (4.3).

Theorem 4.3 (VD and BS [19]) *Assume that $T(a) + H(b) \in \mathcal{F}^p_{TH}(\kappa_1, \kappa_2)$ and the indices of the subordinated operators $T(c)$, $T(d)$ and the factorization signatures of c and d satisfy one of the following conditions:*

(i) $\kappa_1 = 0, \kappa_2 = 0$;
(ii) $\kappa_1 = 1, \kappa_2 = 0$ *and* $\boldsymbol{\sigma}(c) = 1$;
(iii) $\kappa_1 = 0, \kappa_2 = 1$ *and* $\boldsymbol{\sigma}(d) = -1$;
(iv) $\kappa_1 = 1, \kappa_2 = 1$ *and* $\boldsymbol{\sigma}(c) = 1, \boldsymbol{\sigma}(d) = -1$;
(v) $\kappa_1 = 0, \kappa_2 = -1$ *and* $\boldsymbol{\sigma}(d) = 1$;
(vi) $\kappa_1 = -1, \kappa_2 = 0$ *and* $\boldsymbol{\sigma}(c) = -1$;
(vii) $\kappa_1 = -1, \kappa_2 = -1$ *and* $\boldsymbol{\sigma}(c) = -1, \boldsymbol{\sigma}(d) = 1$;
(viii) $\kappa_1 = 1, \kappa_2 = -1$ *and* $\boldsymbol{\sigma}(c) = 1, \boldsymbol{\sigma}(d) = 1$;
(ix) $\kappa_1 = -1, \kappa_2 = 1$ *and* $\boldsymbol{\sigma}(c) = -1, \boldsymbol{\sigma}(d) = -1$.

Then the operator $T(a) + H(b)$ is invertible. Moreover, we have:

1. *Under conditions (i)–(iv), the inverse operator has the form (4.2), where right-inverses of $T(c)$ or/and $T(d)$ shall be replaced by the corresponding inverses.*
2. *Under conditions (v)–(vii), the inverse operator has the form*

$$\begin{aligned}(T(a) + H(b))^{-1} = &- H(\widetilde{c})(T_l^{-1}(c)T(\widetilde{a}^{-1})T_l^{-1}(d)(I - H(d))\\ &+ T_l^{-1}(c)H(\widetilde{a}^{-1})) + H(a^{-1})T_l^{-1}(d)(I - H(d)) + T(a^{-1}).\end{aligned}$$

3. *Under condition (viii), the inverse operator has the form*

$$\begin{aligned}(T(a) + H(b))^{-1} = &- H(\widetilde{c})(T_r^{-1}(c)T(\widetilde{a}^{-1})T_l^{-1}(d)(I - H(d))\\ &+ T_r^{-1}(c)H(\widetilde{a}^{-1})) + H(a^{-1})T_l^{-1}(d)(I - H(d)) + T(a^{-1}).\end{aligned}$$

4. *Under condition (ix), the inverse operator has the form*

$$\begin{aligned}&(T(a) + H(b))^{-1} = T(t^{-1})(I - c_+^{-1}tQt^{-1})\\ &\times [(I - H(t^2\widetilde{c}))T_r^{-1}(t^{-2}c)T(\widetilde{a}^{-1}t^{-1})T_r^{-1}(d) + H(a^{-1}t)T_r^{-1}(d)],\end{aligned}$$

where c_+ is the plus factor in factorization (3.6) for the function c.

Theorem 4.3 is, in fact, the collection of various results from [19]. On the other hand, conditions (i)–(ix) are not necessary for the invertibility of $T(a) + H(b)$ in the case (4.3). Thus if $\kappa_1 = -1$, $\kappa_2 = 1$, then the operator $T(a) + H(b)$ can be invertible even if $\boldsymbol{\sigma}(c)$ and $\boldsymbol{\sigma}(d)$ do not satisfy condition (ix). This case, however, requires special consideration and has been not yet studied.

Consider now the situation (ix) in more detail. This is a subcase of assertion (ii) in Theorem 4.2 and a closer inspection shows substantial difference from the other cases in Theorem 4.3. What makes it very special is the presence of the factorization factor of c in the representation of the inverse operator. It is also worth noting that the construction of $(T(a) + H(b))^{-1}$ is more involved and requires the following result.

Lemma 4.4 (VD and BS [19, 22]) *Let C, D be operators acting on a Banach space X. If $A = CD$ is an invertible operator, then C and D are, respectively, right- and left-invertible operators. Moreover, the operator $D : X \to \operatorname{im} D$ is invertible and if $D_0^{-1} : \operatorname{im} D \to X$ is the corresponding inverse, then the operator A^{-1} can be represented in the form*

$$A^{-1} = D_0^{-1} P_0 C_r^{-1},$$

where P_0 is the projection from X onto $\operatorname{im} D$ parallel to $\ker C$ and C_r^{-1} is any right-inverse of C.

Theorem 4.2(ii) provides necessary conditions for the invertibility of $T(a) + H(b) \in \mathcal{F}^p_{TH}(\kappa_1, \kappa_2)$ for negative κ_1 and positive κ_2. In the next section, we take a closer look at the condition (iia). The related ideas can be seen as a model to study invertibility in cases (iib)–(iid) of Theorem 4.2.

Now we would like to discuss a few examples.

Example 4.5 Let us consider the operator $T(a) + H(b)$ in the case where $a = b$. In this situation $c(t) = 1$ and $d(t) = a(t)\widetilde{a}^{-1}(t)$. Hence, $H(\widetilde{c}) = 0$, $T(c) = I$ and if $\operatorname{ind} T(d) = 0$, then the operator $T(a) + H(a)$ is also invertible and

$$(T(a) + H(a))^{-1} = (T(\widetilde{a}^{-1}) + H(a^{-1}))T^{-1}(a\widetilde{a}^{-1}).$$

Example 4.6 Similar approach show that the operator $T(a) + H(\widetilde{a})$ is invertible if $T(c)$, $c(t) = a(t)\widetilde{a}^{-1}(t)$ is invertible and

$$(T(a) + H(\widetilde{a}))^{-1} = (I - H(\widetilde{a}a^{-1}))T^{-1}(a\widetilde{a}^{-1})T(\widetilde{a}^{-1}) + H(a^{-1}).$$

Example 4.7 Consider the operator $I + H(\phi_0 b)$, where $b(t)\widetilde{b}(t) = 1$ and $\phi_0(t) = t$, $\phi_0(t) = -t^{-1}$ or $\varphi_0(t) = \pm 1$ for all $t \in \mathbb{T}$. Assume that the operator $T(\widetilde{b})$ is Fredholm. Corollary 3.10 indicates that if one of the conditions

(a) $\operatorname{ind} T(\widetilde{b}) = 0$ and $\varphi_0(t) = \pm 1$,
(b) $\operatorname{ind} T(\widetilde{b}) = 0$, $\boldsymbol{\sigma}(\widetilde{b}) = 1$ and $\varphi_0(t) = t$,
(c) $\operatorname{ind} T(\widetilde{b}) = 0$, $\boldsymbol{\sigma}(\widetilde{b}) = 1$ and $\varphi_0(t) = -t^{-1}$,

holds, then

$$\min\{\dim\ker(I + H(\varphi_0 b)), \dim\operatorname{coker}(I + H(\varphi_0 b))\} = 0.$$

Therefore, if $I+H(\varphi_0 b)$ is Fredholm with index zero, then this operator is invertible. However, for $b \in PC$, the Fredholmness of the operators $T(\widetilde{b})$ and $I + H(\varphi_0 b)$ can be studied by Theorems 2.4 and 2.5. It is also possible to construct the inverse operator using the corresponding results on the factorization of PC-functions. However, instead of going into details, we would like to observe that if $T(b)$, $b \in L^\infty$ is invertible, then $I + T(\varphi_0 b)$ is also invertible under the conditions of Theorem 4.3(i), (viii), and (ix), respectively. Moreover, the inverse operator can be explicitly constructed.

Using a distinct method, Basor and Ehrhardt [5] also proved the invertibility of this operator on H^2 under the condition that $T(\widetilde{b}) : H^2 \to H^2$ is invertible. For the H^2-space, the invertibility of $T(b)$ automatically follows from that of $T(\widetilde{b})$. However, if $p \neq 2$, this is not true and the corresponding examples can be found already among operators with piecewise continuous generating functions. It is interesting enough that in each case $\varphi_0(t) = \pm 1$, $\varphi_0(t) = t$ or $\varphi_0(t) = -t^{-1}$, the inverse operator can be represented in the form

$$(I + H(\varphi_0 b))^{-1} = (T(\widetilde{b}) + H(\varphi_0))^{-1}(I + H(\varphi_0\widetilde{b}))(T(b) + H(\varphi_0))^{-1}.$$

5 Invertibility of Operators from $\mathcal{F}^p_{TH}(-2n, 2n)$, $n \in \mathbb{N}$

Theorem 4.2(ii) provides necessary conditions for invertibility of the operator $T(a)+H(b) \in \mathcal{F}^p_{TH}(\kappa_1, \kappa_2)$ if $\kappa_1 < 0$ and $\kappa_2 > 0$. There are four different situations to consider. Here we focus only on the case $T(a) + H(b) \in \mathcal{F}^p_{TH}(-2n, 2n)$, $n \in \mathbb{N}$, but the reader can handle the remaining cases using the ideas below. Let us start with the simplest case $T(a) + H(b) \in \mathcal{F}^p_{TH}(-2, 2)$ and let (c, d) be the subordinated pair for the matching pair $(a, b) \in L^\infty \times L^\infty$. In passing note that if $T(a) + H(b) \in \mathcal{F}^p_{TH}(-2, 2)$, then the adjoint operator $(T(a) + H(b))^* = T(\overline{a}) + H(\widetilde{\overline{b}})$ belongs to the set $\mathcal{F}^q_{TH}(-2, 2)$.

According to (3.6), the Wiener-Hopf factorization of the function d is

$$d(t) = \sigma(d) d_+(t)\, t^{-2}\, \widetilde{d}_+^{-1}(t).$$

It is easily seen that the operator $T(a) + H(b)$ can be represented in the form

$$T(a) + H(b) = (T(a_1) + H(b_1))T(t), \tag{5.1}$$

where $a_1 = at^{-1}$ and $b_1 = tb$. The duo $(a_1, b_1) = (at^{-1}, bt)$ is a matching pair with the subordinated pair $(c_1, d_1) = (ct^{-2}, d)$. Hence, $T(a_1)+H(b_1) \in \mathcal{F}^p_{TH}(0, 2)$

and we note that $T(c_1)$ is invertible. The invertibility of $T(c_1)$ implies that both projections $\mathbf{P}^+_{c_1}$ are $\mathbf{P}^-_{c_1}$ are the zero operators. According to Remark 3.7, the functions $\varphi_\pm$ admit the representations

$$\varphi_\pm(s_\pm) = T_r^{-1}(c)T(\widetilde{a}^{-1})s_\pm, \quad s_\pm \in \operatorname{im}\mathbf{P}^\pm_d.$$

so that

$$\ker(T(a_1) \pm H(b_1)) = T_r^{-1}(c_1)T(\widetilde{a}_1^{-1})(\operatorname{im}\mathbf{P}^\pm_d).$$

Further, we also note that

$$\operatorname{im}\mathbf{P}^\pm_d = \{\nu d_+^{-1}(1 \pm \boldsymbol{\sigma}(d)t : \nu \in \mathbb{C}\}$$

is a one-dimensional subspace of $\ker T(d)$.

By $\omega^{a,b,\pm}$ we denote the functions

$$\omega^{a,b,\pm}(t) = T^{-1}(ct^{-2})T(\widetilde{a}^{-1}t^{-1})(d_+^{-1}(t) \pm \boldsymbol{\sigma}(d)d_+^{-1}(t)t), \tag{5.2}$$

which respectively belong to the kernels of the operators $T(a_1) \pm H(b_1)$. It is clear that $\omega^{a,b,\pm}$ also depend on p.

Representation (5.1) shows that $T(a) + H(b)$ has trivial kernel if and only if

$$\ker(T(a_1) + H(b_1)) \cap \operatorname{im} T(t) = \{0\}.$$

It is possible only if

$$\widehat{\omega}_0^{a,b,+} \neq 0,$$

where $\widehat{\omega}_0^{a,b,+}$ is the zero Fourier coefficient of the function $\omega^{a,b,+}(t)$. Similar result is valid for the operator $T(a) - H(b)$. Note that if $T(a) + H(b)$ belongs to the set $\mathcal{F}^p_{TH}(-2, 2)$, then so is the operator $T(a) - H(b)$, since $T(a) - H(b) = T(a) + H(-b)$.

Theorem 5.1 (VD and BS [22]) *Let $T(a) + H(b) \in \mathcal{F}^p_{TH}(-2, 2)$.*

(a) The operator $T(a) + H(b)$ $(T(a) - H(b))$ is left-invertible if and only if $\widehat{\omega}_0^{a,b,+} \neq 0$ $(\widehat{\omega}_0^{a,b,-} \neq 0)$.

(b) The operator $T(a) + H(b)$ $(T(a) - H(b))$ is right-invertible if and only if $\widehat{\omega}_0^{\overline{a},\widetilde{\overline{b}},+} \neq 0$ $(\widehat{\omega}_0^{\overline{a},\widetilde{\overline{b}},-} \neq 0)$.

(c) The operator $T(a)+H(b)$ $(T(a)-H(b))$ is invertible if and only if $\widehat{\omega}_0^{a,b,+} \neq 0$ and $\widehat{\omega}_0^{\overline{a},\widetilde{\overline{b}},+} \neq 0$ $(\widehat{\omega}_0^{a,b,-} \neq 0$ and $\widehat{\omega}_0^{\overline{a},\widetilde{\overline{b}},-} \neq 0)$.

(d) If $\widehat{\omega}_0^{a,b,+} \neq 0$ and $\widehat{\omega}_0^{a,b,-} \neq 0$, then both operators $T(a) + H(b)$ and $T(a) - H(b)$ are invertible.

Let us sketch the proof of Assertion (d). It follows from the representations (2.2) and (2.12) that

$$\operatorname{ind}(T(a)+H(b))+\operatorname{ind}(T(a)-H(b))=0. \tag{5.3}$$

By Assertion (a), both operators $T(a)+H(b)$ and $T(a)-H(b)$ are left-invertible. Therefore, $\operatorname{ind}(T(a)+H(b))\le 0$, $\operatorname{ind}(T(a)-H(b))\le 0$ and taking into account (5.3), we obtain that

$$\operatorname{ind}(T(a)+H(b))=\operatorname{ind}(T(a)-H(b))=0,$$

which yields the invertibility of both operators under consideration.

If an operator $T(a)+H(b)\in\mathcal{F}^p_{TH}(-2,2)$ is invertible, we can construct its inverse by using Lemma 4.4. In particular, we have.

Theorem 5.2 (VD and BS [22]) *If $T(a)+H(b)\in\mathcal{F}^p_{TH}(-2,2)$ is invertible, then the inverse operator can be represented in the form*

$$\begin{aligned}&(T(a)+H(b))^{-1}\\ &=T(t^{-1})\left(I-\frac{1}{\widehat{\omega}_0^{a,b,+}}T^{-1}(ct^{-2})T(\widetilde{a}^{-1}t^{-1})d_+^{-1}(t)(1+\boldsymbol{\sigma}(d)t)tQt^{-1}\right)\\ &\quad\times\left((I-H(t^2\widetilde{c}))T^{-1}(t^{-2}c)T(\widetilde{a}^{-1}t^{-1})T_r^{-1}(d)+H(a^{-1}t)T_r^{-1}(d)\right).\end{aligned} \tag{5.4}$$

Example 5.3 We now consider the operator $A=T(a)+H(t^{-2}a)$, defined by the function

$$a(t):=(1-\gamma t^{-1})(1-\gamma t)^{-1},$$

where γ is a fixed number in the interval $(0,1)$.

The duo (a,at^{-2}) is a matching pair with the subordinated pair $(c,d)=(t^2,a^{-1}t^2)$ and $A\in\mathcal{F}^p_{TH}(-2,2)$. The corresponding Wiener-Hopf factorizations of a and d are the same in all H^p. More precisely, we have

$$\begin{aligned}a(t)&=a_+(t)\widetilde{a}_+^{-1}(t), & a_+(t)&=(1-\gamma t)^{-1},\\ d(t)&=1\cdot d_+(t)t^{-2}\widetilde{d}_+^{-1}(t), & d_+(t)&=(1-\gamma t)^{-2}.\end{aligned}$$

Hence, $\boldsymbol{\sigma}(d)=1$ and computing the zero Fourier coefficients of the corresponding functions (5.2), we obtain

$$\widehat{w}_0^{a,at^{-2},+}=\widehat{w}_0^{a,at^{-2},-}=\gamma^2-\gamma+1.$$

It is easily seen that for any $\gamma \in (0, 1)$ these coefficients are not equal to zero, so that by Theorem 5.1(c), the operator $T(a) + H(at^{-2})$ is invertible. The corresponding inverse operator, which is constructed according to the representation (5.4), has the form

$$\begin{aligned}
&(T(a) + H(t^{-2}a))^{-1} \\
&= T(t^{-1}) \left(I - \frac{1}{\gamma^2 - \gamma + 1} P \left((1 - \gamma t^{-1})(1 - \gamma t)(1 + t) \right) Qt^{-1} \right) \\
&\quad \times \left(T \left(\frac{(1 - \gamma t^{-1})t^{-1}}{1 - \gamma t} \right) + H \left(\frac{(1 - \gamma t)t}{1 - \gamma t^{-1}} \right) \right) \\
&\quad \times T((1 - \gamma t)^2) T((1 - \gamma t^{-1})^{-2}) T(t^2).
\end{aligned}$$

Consider now the invertibility of the operators $T(a) + H(b)$ from the set $\mathcal{F}^p_{TH}(-2n, 2n)$ for n greater than 1. Thus we assume that the subordinated operator $T(c) = T(ab^{-1})$ and $T(d) = T(a\widetilde{b}^{-1})$ are Fredholm and

$$\operatorname{ind} T(c) = -2n, \quad \operatorname{ind} T(d) = 2n, \ n > 1.$$

Considering the functions

$$\omega^{a,b}_k(t) := T^{-1}(ct^{-2n}) T(\widetilde{a}^{-1} t^{-n}) d_+^{-1}(t) (t^{n-k-1} + \boldsymbol{\sigma}(d) t^{n+k})$$

for $k = 0, 1, \cdots, n-1$, we introduce the matrix

$$W_n(a, b) = (\omega^{a,b}_{jk})^{n-1}_{k,j=0},$$

with the entries

$$\omega^{a,b}_{jk} = \frac{1}{2\pi} \int_0^{2\pi} \omega^{a,b}_k(e^{i\theta}) e^{-ij\theta} \, d\theta, \quad j, k = 0, 1, \cdots, n-1,$$

and the terms d_+ and $\boldsymbol{\sigma}(d)$ defined by the Wiener-Hopf factorization

$$d(t) = \boldsymbol{\sigma}(d) d_+(t) \, t^{-2n} \, \widetilde{d}_+^{-1}(t),$$

with respect to L^p. Notice that the functions $\omega^{a,b}_k$ form a basis in $\varphi_+(\operatorname{im} \mathbf{P}^+_d)$, where $\varphi_+ : \operatorname{im} \mathbf{P}^+_d \to \ker(T(a_1) + H(b_1))$ is defined by

$$\varphi_+ = T^{-1}(ct^{-2n}) T(\widetilde{a}^{-1} t^{-n})).$$

The invertibility of the operators from $\mathcal{F}^p_{TH}(-2n, 2n)$ is described by the following theorem.

Theorem 5.4 (VD and BS [22]) *If $T(a)+H(b)\in\mathcal{F}^p_{TH}(-2n,2n)$, then:*

(a) $T(a)+H(b)$ is left-invertible if and only if $W_n(a,b)$ is a non-degenerate matrix.
(b) $T(a)+H(b)$ is right-invertible if and only if $W_n(\overline{a},\widetilde{\overline{b}})$ is a non-degenerate matrix.
(c) $T(a)+H(b)$ is invertible if and only if $W_n(a,b)$ and $W_n(\overline{a},\widetilde{\overline{b}})$ are non-degenerate matrices.
(d) If $W_n(a,b)$ and $W_n(a,-b)$ are non-degenerate matrices, then both operators $T(a)+H(b)$ and $T(a)-H(b)$ are invertible.

Example 5.5 Consider operator $T(a)+H(at^{-2n})$, $n\in\mathbb{N}$ and

$$a(t)=\frac{1-\gamma t^{-1}}{1-\gamma t},\ \gamma\in(0,1). \tag{5.5}$$

It was shown in [22] that for any $n\in\mathbb{N}$, the above operator is left-invertible in any space H^p, $1<p<\infty$. Moreover, since $H(at^{-2n})$ is compact and $\operatorname{ind}T(a)=0$, the operator at hand is even invertible. The inverse of $T(a)+H(at^{-2n})$ can be constructed in explicit form.

Remark 5.6 If $m,n\in\mathbb{N}$ and $m\neq n$, the set $\mathcal{F}^p_{TH}(-2m,2n)$ does not contain invertible operators, but it still includes one-sided invertible operators.

6 Toeplitz Plus Hankel Operators on l^p-Spaces

A substantial portion of the results presented in Sects. 2–5 can be extended to Toeplitz plus Hankel operators acting on l^p-spaces, $1<p<\infty$. Such an extension is highly non-trivial because many tools perfectly working in classical Hardy spaces H^p, are not available for operators on l^p-spaces. In particular, a big problem is the absence of a Wiener-Hopf factorization, which plays outstanding role in the study of Toeplitz plus Hankel operators on classical H^p-spaces.

6.1 Spaces and Operators

Let $l^p(\mathbb{Z})$ denote the complex Banach space of all sequences $\xi=(\xi_n)_{n\in\mathbb{N}}$ of complex numbers with the norm

$$||\xi||_p=\left(\sum_{n\in\mathbb{Z}}|\xi_n|^p\right)^{1/p},\quad 1\le p<\infty.$$

As usual, $\mathbb{Z}$ denotes the set of all integers. If we replace $\mathbb{Z}$ by the set of all non-negative integers $\mathbb{Z}^+$, we get another Banach space $l^p(\mathbb{Z}^+)$. It can be viewed as a subspace of $l^p(\mathbb{Z})$ and we will often write l^p for $l^p(\mathbb{Z}^+)$. By P, we now denote the canonical projection from $l^p(\mathbb{Z})$ onto $l^p(\mathbb{Z}^+)$ and let $Q := I - P$. Further, let J refer to the operator on $l^p(\mathbb{Z})$ defined by

$$(J\xi)_n = \xi_{-n-1}, \quad n \in \mathbb{Z}.$$

The operators J, P and Q are connected with each other by the relations

$$J^2 = I, \quad JPJ = Q, \quad JQJ = P.$$

For each $a \in L^p = L^p(\mathbb{T})$, let $(\widehat{a}_k)_{k\in\mathbb{Z}}$ be the sequence of its Fourier coefficients. The Laurent operator $L(a)$ associated with $a \in L^\infty(\mathbb{T})$ acts on the space $l^0(\mathbb{Z})$ of all finitely supported sequences on $\mathbb{Z}$ by

$$(L(a)x)_k := \sum_{m\in\mathbb{Z}} \widehat{a}_{k-m}x_m, \tag{6.1}$$

where the sum in the right-hand side of (6.1) contains only a finite number of non-zero terms for every $k \in \mathbb{Z}$. We say that a is a multiplier on $l^p(\mathbb{Z})$ if $L(a)x \in l^p(\mathbb{Z})$ for any $x \in l_0(\mathbb{Z})$ and if

$$||L(a)|| := \sup\{||L(a)x||_p : x \in l_0(\mathbb{Z}),\ ||x||_p = 1\}$$

is finite. In this case, $L(a)$ extends to a bounded linear operator on $l^p(\mathbb{Z})$, which is again denoted by $L(a)$. The set M^p of all multipliers on $l^p(\mathbb{Z})$ is a Banach algebra under the norm $||a||_{M_p} := ||L(a)||$—cf. [7]. It is well-known that $M^2 = L^\infty(\mathbb{T})$. Moreover, every function $a \in L^\infty(\mathbb{T})$ with bounded total variation $\mathrm{Var}\,(a)$ is in M^p for every $p \in (1, \infty)$ and the Stechkin inequality

$$||a||_{M^p} \le c_p(||a||_\infty + \mathrm{Var}\,(a))$$

holds with a constant c_p independent of a. In particular, every trigonometric polynomial and every piecewise constant function on $\mathbb{T}$ are multipliers on any space $l^p(\mathbb{Z})$, $p \in (1, \infty)$. By C_p and PC_p we, respectively, denote the closures of algebras of all trigonometric polynomials $\mathcal{E}$ and all piecewise constant functions PC in M^p. Note that C_2 is just the algebra $C(\mathbb{T})$ of all continuous functions on $\mathbb{T}$, and PC_2 is the algebra $PC(\mathbb{T})$ of all piecewise continuous functions on $\mathbb{T}$. We also note that the Wiener algebra W of the functions with absolutely converging Fourier series is also a subalgebra of M^p and

$$W \subset C_p \subset PC_p \subset PC \quad \text{and} \quad M^p \subset L^\infty.$$

For this and other properties of multiplier cf. [7]. We also recall the equation $JL(a)J = L(\widetilde{a})$ often used in what follows.

Let $a \in M^p$. The operators $T(a) : l^p \to l^p$, $x \mapsto PL(a)x$ and $H(a) : l^p \to l^p$, $x \mapsto PL(a)Jx = PL(a)QJ$ are, respectively, called Toeplitz and Hankel operators with generating function a. It is well-known that $||T(a)|| = ||a||_{M^p}$ and $||H(a)|| \leq ||a||_{M^p}$ for every multiplier $a \in M^p$. In this section we also use the notation $T_p(a)$ or $H_p(a)$ in order to underline that the corresponding Toeplitz or Hankel operator is considered on the space l^p for a fixed $p \in (1, \infty)$. The action of the operators $T_p(a)$ and $H_p(a)$ on the elements from l^p can be written as follows

$$T(a) : (\xi_j)_{j\in\mathbb{Z}_+} \to \left(\sum_{k\in\mathbb{Z}_+} \widehat{a}_{j-k}\xi_k \right)_{j\in\mathbb{Z}_+},$$

$$H(a) : (\xi_j)_{j\in\mathbb{Z}_+} \to \left(\sum_{k\in\mathbb{Z}_+} \widehat{a}_{j+k+1}\xi_k \right)_{j\in\mathbb{Z}_+}.$$

Let GM^p denote the group of invertible elements in M^p.

Lemma 6.1 (cf. Ref. [7]) *Let $p \in (1, \infty)$.*

1. *If $T_p(a)$ is Fredholm, then $a \in GM^p$.*
2. *If $a \in M^p$, then one of the kernels of the operators $T_p(a)$ or $T_q^*(a)$, $1/p + 1/q = 1$ is trivial.*
3. *If $a \in GM^p$, the operator $T_p(a)$ is Fredholm, and $\operatorname{ind} T_p(a) = 0$, then $T(a)$ is invertible on l^p.*

6.2 Kernels of a Class of Toeplitz Plus Hankel Operators

The goal of this subsection is to present a method on how to study certain problems for Toeplitz plus Hankel operator $T_p(a) + H_p(b)$ defined on l^p via known results obtained in Sects. 2–5. Since $M^p \subset L^\infty(\mathbb{T})$, for any given elements $a, b \in M^p$ the operator $T_p(a) + H_p(b) \in \mathcal{L}(l^p)$ generates the operator $\mathbf{T}_s(a) + \mathbf{H}_s(b) \in \mathcal{L}(H^s)$, $1 < s < \infty$ in an obvious manner. We denote by $\mathcal{L}^{p,q}(l^p)$ the collection of all Toeplitz plus Hankel operators acting on the space l^p such that the following conditions hold:

(a) $a, b \in M^p$.
(b) $T_p(a) + H_p(b) \in \mathcal{L}(l^p)$ is Fredholm.
(c) $\mathbf{T}_q(a) + \mathbf{H}_q(b) \in \mathcal{L}(H^q)$, $1/p + 1/q = 1$ is Fredholm.
(d) Both operators acting on the spaces mentioned have the same Fredholm index.

The following observation is crucial for what follows. The famous Hausdorff-Young Theorem connects the spaces l^p and H^q, $1/p + 1/q = 1$ via Fourier transform $\mathcal{F}(a) = (\widehat{a}_n)_{n\in\mathbb{Z}}$, $a \in H^q$.

Theorem 6.2 (Hausdorff and Young [27])

(a) If $g \in H^p$ and $1 \le p \le 2$, then $\mathcal{F}g \in l^q$ and

$$||\mathcal{F}g||_q \le ||g||_p.$$

(b) If $\varphi = (\varphi_n)_{n\in\mathbb{Z}} \in l^p(\mathbb{Z})$ and $1 \le p \le 2$, then the series $\sum \varphi_n e^{int}$ converges in L^q to a function $\breve{\varphi}$ and

$$||\breve{\varphi}||_q \le ||\varphi||_p.$$

Theorem 6.2 gives rise to the following construction. Let $\widehat{H}^p$, $p \in (1,\infty)$ be the set of all sequences $(\chi_n)_{n\in\mathbb{Z}_+}$ such that there exists a function $h \in H^p$ with the property $\mathcal{F}h = (\chi_n)_{n\in\mathbb{Z}_+}$. The set $\widehat{H}^p$ equipped with the norm $||(g_k)_{k\in\mathbb{Z}_+}|| := ||g||_{H^p}$, becomes a Banach space isometrically isomorphic to H^p. Part (a) of the Hausdorff-Young Theorem assures that $\widehat{H}^p$ is densely continuously embedded in the space l^q and part (b) claims that l^p, $1 \le p \le 2$ is continuously and densely embedded into $\widehat{H}^q$.

Theorem 6.3 (VD and BS [21]) *Let $p \in (1,\infty)$. If $T_p(a) + H_p(b) \in \mathcal{L}^{p,q}$, then the Fourier transform $\mathcal{F}$ is an isomorphism between the spaces* $\ker(\mathbf{T}_q(a) + \mathbf{H}_q(b))$ *and* $\ker(T_p(a) + H_p(b))$.

Let us give a sketch of the proof. The matrix representation $[\mathbf{T}_q(a) + \mathbf{H}_q(b)]$ of the operator $\mathbf{T}_q(a) + \mathbf{H}_q(b)$ in the standard basis $(t^n)_{n\in\mathbb{Z}^+}$ induces a linear bounded operator on $\widehat{H}^q$, so that

$$[\mathbf{T}_q(a) + \mathbf{H}_q(b)] = (\widehat{a}_{i-j} + \widehat{b}_{i+j+1})_{i,j=0}^{\infty}$$

The operators $\mathbf{T}_q(a) + \mathbf{H}_q(b) \in \mathcal{L}(H^q)$ and $[\mathbf{T}_q(a) + \mathbf{H}_q(b)] \in \mathcal{L}(\widehat{H}^q)$ have the same Fredholm properties as $T_p(a) + H_p(b) \in \mathcal{L}(l^p)$ for $1 < q \le 2$. Moreover, the operator $T_p(a) + H_p(b) \in \mathcal{L}(l^p)$ is the extension of $[\mathbf{T}_q(a) + \mathbf{H}_q(b)]$ with the same index. However, in this case Lemma 3.14 then indicates that

$$\ker[\mathbf{T}_q(a) + \mathbf{H}_q(b)] = \ker(T_p(a) + H_p(b)), \quad T_p(a) + H_p(b) \in \mathcal{L}(l^p),$$

and for $p \ge 2$ the assertion follows. The case $1 < p \le 2$ can be treated analogously.

Thus the main problem in using Theorem 6.2 is whether it is known that $T_p(a) + H_p(b) \in \mathcal{L}^{p,q}$. This is a non-trivial fact but the following result holds.

Proposition 6.4 (VD and BS [21]) *Let $a, b \in PC_p$, $1 < p < \infty$. If the operator $T_p(a) + H_p(b) \in \mathcal{L}(l^p)$ is Fredholm, then $T_p(a) + H_p(b) \in \mathcal{L}^{p,q}$.*

The proof of this proposition can be carried out using ideas from [50] and [51]. However, a simpler proof can be obtained if the generating function a and b satisfy the matching condition $a\widetilde{a} = b\widetilde{b}$ and the operators $T_p(c), T_p(d)$ are Fredholm say in l^p. Then $T_p(c), T_p(d) \in \mathcal{L}^{p,q}$—cf. [7, Chapters 4 and 6], and using classical approach, which also works in l^p situation, one obtains the result.

Thus, it is now clear how to extend the results of Sects. 2–5 to Toeplitz plus Hankel operators acting on spaces l^p. Let us formulate just one such result without going into much details.

Theorem 6.5 *Let $(a, b) \in PC_p \times PC_p$ be a Fredholm matching pair with the subordinated pair (c, d), and let $\widehat{c}_{+,j}^{-1}$, $j \in \mathbb{Z}_+$ be the Fourier coefficients of the function c_+^{-1}, where c_+ is the plus factor in the Wiener-Hopf factorization (3.6) of the function c in H^q. If $\kappa_1 := \operatorname{ind} T_p(c) > 0$, $\kappa_2 := \operatorname{ind} T_p(d) \leq 0$, then the kernel of the operator $T_p(a) + H_p(b)$ admits the following representation:*

(a) If $\kappa_1 = 1$ and $\sigma(c) = 1$, then

$$\ker(T_p(a) + H_p(b)) = \{0\}.$$

(b) If $\kappa_1 = 1$ and $\sigma(c) = -1$, then

$$\ker(T_p(a) + H_p(b)) = \operatorname{lin\,span}\{(\widehat{c}_{+,j}^{-1})_{j\in\mathbb{Z}_+}\}.$$

(c) If $\kappa_1 > 1$ is odd, then

$$\ker(T_p(a) + H_p(b)) = \operatorname{lin\,span}\{(\widehat{c}_{+,j-(\kappa_1-1)/2-l}^{-1} - \boldsymbol{\sigma}(c)\widehat{c}_{+,j-(\kappa_1-1)/2+l}^{-1})_{j\in\mathbb{Z}_+} : l = 0, \cdots, (\kappa_1 - 1)/2\}.$$

(d) If κ_1 is even, then

$$\ker(T_p(a) + H_p(b)) = \operatorname{lin\,span}\{(\widehat{c}_{+,j-\kappa_1/2+l+1}^{-1} - \boldsymbol{\sigma}(c)\widehat{c}_{+,j-\kappa_1/2-l}^{-1})_{j\in\mathbb{Z}_+} : l = 0, 1, \cdots, \kappa_1/2 - 1\}.$$

Remark 6.6 Sometime the study of invertibility of Toeplitz plus Hankel operators $T_p(a) + H_p(b)$ can be carried out even if it is not known, whether this operator belongs to $\mathcal{L}^{p,q}$. Thus l^p-versions of Theorem 2.9, Proposition 3.9 and Corollary 3.10 can be directly proved.

7 Wiener-Hopf Plus Hankel Operators

This section is devoted to Wiener-Hopf plus Hankel operators. Let χ_E refer to the characteristic function of the subset $E \subset \mathbb{R}$ and let $\mathcal{F}$ and $\mathcal{F}^{-1}$ be the direct and

inverse Fourier transforms—i.e.

$$\mathcal{F}\varphi(\xi) := \int_{-\infty}^{\infty} e^{i\xi x}\varphi(x)\,dx, \quad \mathcal{F}^{-1}\psi(x) := \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\xi x}\psi(\xi)\,d\xi, \quad x \in \mathbb{R}.$$

In what follows, we identify the spaces $L^p(\mathbb{R}^+)$ and $L^p(\mathbb{R}^-)$, $1 \le p \le \infty$ with the subspaces $P(L^p(\mathbb{R}))$ and $Q(L^p(\mathbb{R}))$ of the space $L^p(\mathbb{R})$, where P and Q are the projections in $L^p(\mathbb{R})$ defined by $Pf := \chi_{\mathbb{R}^+} f$ and $Q := I - P$, respectively.

Consider the set G of functions g having the form

$$g(t) = (Fk)(t) + \sum_{j\in\mathbb{Z}} g_j e^{i\delta_j t}, \quad t \in \mathbb{R}, \tag{7.1}$$

where $k \in L(\mathbb{R})$, $\delta_j \in \mathbb{R}$, $g_j \in \mathbb{C}$ and the series in the right-hand side of (7.1) is absolutely convergent. Any function $g \in G$ generates an operator $W^0(a)\colon L^p(\mathbb{R}) \to L^p(\mathbb{R})$ and operators $W(g), H(g)\colon L^p(\mathbb{R}^+) \to L^p(\mathbb{R}^+)$ defined by

$$W^0(g) := \mathcal{F}^{-1}g\mathcal{F}\varphi, \quad W(g) := PW^0(g), \quad H(g) := PW^0(g)QJ.$$

Here and throughout this section, $J\colon L^p(\mathbb{R}) \to L^p(\mathbb{R})$ is the reflection operator defined by $J\varphi := \widetilde{\varphi}$ and $\widetilde{\varphi}(t) := \varphi(-t)$ for any $\varphi \in L^p(\mathbb{R})$, $p \in [1, \infty]$. Operators $W(g)$ and $H(g)$ are, respectively, called Wiener-Hopf and Hankel operators on the semi-axis $\mathbb{R}^+$. It is well-known [31] that for $g \in G$, all three operators above are bounded on any space L^p, $p \in [1, \infty)$.

In particular, the operator $W(g)$ has the form

$$W(g)\varphi(t) = \sum_{j=-\infty}^{\infty} g_j B_{\delta_j}\varphi(t) + \int_0^{\infty} k(t-s)\varphi(s)\,ds, \quad t \in \mathbb{R}^+,$$

where

$$B_{\delta_j}\varphi(t) = \varphi(t-\delta_j) \quad \text{if} \quad \delta_j \le 0,$$

$$B_{\delta_j}\varphi(t) = \begin{cases} 0, & 0 \le t \le \delta_j \\ \varphi(t-\delta_j), & t > \delta_j \end{cases} \quad \text{if} \quad \delta_j > 0.$$

Moreover, for $g = \mathcal{F}k$ the operator $H(a)$ acts as

$$H(g)\varphi(t) = \int_0^{\infty} k(t+s)\varphi(s)\,ds$$

and for $g = e^{\delta t}$, one has $H(a)\varphi(t) = 0$ if $\delta \leq 0$ and

$$H(g)\varphi(t) = \begin{cases} \varphi(\delta - t), & 0 \leq t \leq \delta, \\ 0, & t > \delta, \end{cases}$$

if $\delta > 0$.

For various classes of generating functions, the Fredholm properties of operators $W(a)$ are well studied [7, 8, 12, 24–26, 31]. In particular, Fredholm and semi-Fredholm Wiener-Hopf operators are one-sided invertible, and for $a \in G$ there is efficient description of the kernels and cokernels of $W(a)$ and formulas for the corresponding one-sided inverses.

The study of Wiener-Hopf plus Hankel operators

$$\mathbb{W}(a, b) = W(a) + H(b), \quad a, b \in L^\infty(\mathbb{R}). \tag{7.2}$$

is much more involved. The invertibility and Fredholmness of such operators in the space L^2 is probably first considered by Lebre et al. [43]. In particular, the invertibility of $\mathbb{W}(a, b)$ has been connected with the invertibility of a block Wiener-Hopf operator $W(G)$. Assuming that a admits canonical Wiener-Hopf factorization in L^2 and the Wiener-Hopf factorization of the matrix G is known, Lebre et al. provided a formula for $\mathbb{W}^{-1}(a, b)$. Nevertheless, the difficulties with Wiener-Hopf factorization of matrix G influence the efficiency of the method. For piecewise continuous generating functions, the conditions of Fredholmness are obtained in [51]. A different method, called equivalence after extension, has been applied to Wiener-Hopf plus Hankel operators $\mathbb{W}(a, a) = W(a) + H(a)$ with generating function a belonging to various functional classes [9, 10]. The Fredholmness, one-sided invertibility and invertibility of such operators are equivalent to the corresponding properties of the Wiener-Hopf operator $W(a\widetilde{a}^{-1})$. Therefore, known results about the invertibility and Fredholmness of $W(a\widetilde{a}^{-1})$ can be retranslated to the operator $\mathbb{W}(a, a)$. On the other hand, the equivalence after extension has not been used to establish any representations of the corresponding inverses. Another approach has been exploited in [38, 39] to study the essential spectrum and the index of the operators $I + H(b)$.

We now consider Wiener-Hopf plus Hankel operators (7.2) with generating functions $a, b \in G$ satisfying the matching condition (2.4), where $\widetilde{a}(t)$ and $\widetilde{b}(t)$ denote the functions $a(-t)$ and $b(-t)$, respectively. Thus we now assume that

$$a(t)a(-t) = b(t)b(-t) \tag{7.3}$$

and define the subordinated functions c and d by

$$c(t) := a(t)b^{-1}(t), \quad d(t) := a(t)\widetilde{b}^{-1}(t) = a(t)b^{-1}(-t), \quad t \in \mathbb{R}.$$

The assumption (7.3) allows us to employ the method developed in Sects. 2–5 and establish necessary and also sufficient conditions for invertibility and one-sided invertibility of the operators under consideration and obtain efficient representations for the corresponding inverses. Note that here we only provide the main results. For more details the reader can consult [14, 20, 23].

Let $G^+ \subset G$ and $G^- \subset G$ denote the sets of functions, which admit holomorphic extensions to the upper and lower half-planes, respectively. If $g \in G$ is a matching function—i.e. $g\widetilde{g} = 1$, then according to [14, 20], it can be represented in the form

$$g(t) = \left(\boldsymbol{\sigma}(g)\,\widetilde{g}_+^{-1}(t)\right) e^{i\nu t}\Big(\frac{t-i}{t+i}\Big)^n g_+(t),$$

where $\nu = \nu(g) \in \mathbb{R}$, $n = n(g) \in \mathbb{N}$, $\boldsymbol{\sigma}(g) = (-1)^n g(0)$, $\widetilde{g}_+^{\pm 1}(t) \in G^-$ and $\widetilde{g}_+(\infty) = 1$. This representation is unique and the numbers $\nu(g)$ and $n(g)$ in the representation (7.1) are defined as follows:

$$\nu(g) := \lim_{l\to\infty} \frac{1}{2l}[\arg g(t)]_{-l}^{l}, \quad n(g) := \frac{1}{2\pi}[\arg(1 + g^{-1}(t)(\mathcal{F}k)(t)]_{t=-\infty}^{\infty}.$$

Moreover, following the considerations of Sects. 2–3, for any right-invertible operator $W(g)$ generated by a matching function g, we can introduce complementary projections $\mathbf{P}_g^{\pm}$ on $\ker W(g)$. More precisely, if $\nu < 0$ or if $\nu = 0$ and $n < 0$, then $W(g)$ is right-invertible and the projections $\mathbf{P}_g^{\pm}$ are defined by

$$\mathbf{P}_g^{\pm} := (1/2)(I \pm JQPW^0(g)P)\colon \ \ker W(g) \to \ker W(g).$$

In addition, we also need the translation operator φ^+ defined by the subordinated functions c and d. Assume that $W(c)$ is right-invertible operator. Let $W_r^{-1}(c)$ be any of its right-inverses and consider the following function:

$$\varphi^+ = \varphi^+(a,b) := \frac{1}{2}(W_r^{-1}(c)W(\tilde{a}^{-1}) - JQW^0(c)PW_r^{-1}(c)W(\tilde{a}^{-1})) + \frac{1}{2}JQW^0(\tilde{a}^{-1}),$$

We are ready to discuss the invertibility of Wiener-Hopf plus Hankel operators starting with necessary conditions in the case where at least one of the indices $\nu_1 := \nu(c)$ or $\nu_2 := \nu(d)$ is not equal to zero. The situation $\nu_1 = \nu_2 = 0$ will be considered later on. Let n_1 and n_2 denote the indices $n(c)$ and $n(d)$, respectively.

Theorem 7.1 (VD and BS [23]) *If $a, b \in G$, the operator $W(a) + H(b)$ is one-sided invertible in $L^p(\mathbb{R}^+)$ and at least one of the indices ν_1 or ν_2 is not equal to zero, then:*

(i) Either $\nu_1\nu_2 \geq 0$ or $\nu_1 > 0$ and $\nu_2 < 0$.
(ii) If $\nu_1 = 0$ and $\nu_2 > 0$, then $n_1 > -1$ or $n_1 = -1$ and $\boldsymbol{\sigma}(c) = -1$.
(iii) If $\nu_1 < 0$ and $\nu_2 = 0$, then $n_2 < 1$ or $n_2 = 1$ and $\boldsymbol{\sigma}(d) = -1$.

Consider now the case $\nu_1 > 0$ and $\nu_2 < 0$ in more detail. It can be specified as follows.

Theorem 7.2 (VD and BS [23]) *Let $\nu_1 > 0$, $\nu_2 < 0$, $n_1 = n_2 = 0$ and $\mathfrak{N}_\nu^p$, $\nu > 0$ denote the set of functions $f \in L^p(\mathbb{R}^+)$ such that $f(t) = 0$ for $t \in (0, \nu)$.*

(i) If the operator $W(a) + H(b)\colon L^p(\mathbb{R}^+) \to L^p(\mathbb{R}^+)$, $1 < p < \infty$ is invertible from the left, then

$$\varphi^+(\mathbf{P}_d^+) \cap \mathfrak{N}_{\nu_1/2}^p = \{0\},$$

where $\varphi^+ = \varphi^+(ae^{-i\nu_1 t/2}, be^{i\nu_1 t/2})$.

(ii) If the operator $W(a) + H(b)\colon L^p(\mathbb{R}^+) \to L^p(\mathbb{R}^+)$, $1 < p < \infty$ is invertible from the right, then

$$\varphi^+(\mathbf{P}_{\widetilde{c}}^+) \cap \mathfrak{N}_{-\nu_2/2}^p = \{0\}, \tag{7.4}$$

where $\varphi^+ = \varphi^+(\overline{a}e^{i\nu_2 t/2}, \overline{\widetilde{b}}e^{-i\nu_2 t/2})$.

Passing to the case $\nu_1 = \nu_2 = 0$, we note that now the indices n_1 and n_2 take over.

Theorem 7.3 (VD and BS [23]) *Let $a, b \in G$, $\nu_1 = \nu_2 = 0$ and the operator $W(a) + H(b)$ is invertible from the left. Then:*

(i) In the case $n_2 \geq n_1$, the index n_1 satisfies the inequality

$$n_1 \geq -1$$

and if $n_1 = -1$, then $\boldsymbol{\sigma}(c) = -1$ and $n_2 > n_1$.

(ii) In the case $n_1 > n_2$, the index n_1 satisfies the inequality

$$n_1 \geq 1,$$

and the index n_2 is either non-negative or $n_2 < 0$ and $n_1 \geq -n_2$.

The necessary conditions of the right-invertibility have similar form and we refer the reader to [23] for details. The proof of the above results is based on the analysis of the kernel and cokernel of the operator $W(a)+T(b)$. In particular, if $W(a)+T(b)$ is left-invertible and one of the corresponding conditions is not satisfied, then this operator should have a non-zero kernel, which cannot be true.

The sufficient conditions of one-sided invertibility can be also formulated in terms of indices ν_1, ν_2, n_1 and n_2. For example, the following theorem holds.

Theorem 7.4 (VD and BS [23]) *Let $a, b \in G$ and indices ν_1, ν_2, n_1 and n_2 satisfy any of the following conditions:*

(i) $\nu_1 < 0$ *and* $\nu_2 < 0$.
(ii) $\nu_1 > 0$, $\nu_2 < 0$, $n_1 = n_2 = 0$, *operator* $W(a) + H(b)$ *is normally solvable and satisfies the condition* (7.4).
(iii) $\nu_1 < 0$, $\nu_2 = 0$ *and* $n_2 < 1$ *or* $n_2 = 1$ *and* $\sigma(d) = -1$.
(iv) $\nu_1 = 0$, $n_1 \leq 0$ *and* $\nu_2 < 0$.
(v) $\nu_1 = 0$ *and* $\nu_2 = 0$

(a) $n_1 \leq 0$, $n_2 < 1$;
(b) $n_1 \leq 0$, $n_2 = 1$ *and* $\sigma(d) = -1$;

Then the operator $W(a) + H(b)$ is right-invertible.

The sufficient conditions for left-invertibility of $W(a) + H(b)$ can be obtained by passing to the adjoint operator. Here we are not going to discuss this problem in whole generality. However, we use assumptions, which allow to derive simple formulas for left- or right-inverses. These conditions are not necessary for one-sided invertibility and the corresponding inverses can be also constructed even if the conditions above are not satisfied.

Theorem 7.5 (VD and BS [23]) *Let $(a, b) \in G \times G$ be a matching pair. Then:*

1. *If $W(c)$ and $W(d)$ are left-invertible, the operator $W(a) + H(b)$ is also left-invertible and one of its left-inverses has the form*

$$\begin{aligned}(W(a) + H(b))_l^{-1} =& W_l^{-1}(c) W(\tilde{a}^{-1}) W_l^{-1}(d)(I - H(\tilde{d})) \\ &+ W_l^{-1}(c) H(\tilde{a}^{-1}).\end{aligned} \tag{7.5}$$

2. *If $W(c)$ and $W(d)$ are right-invertible, the operator $W(a) + H(b)$ is also right-invertible and one of its right-inverses has the form*

$$\begin{aligned}(W(a) + H(b))_r^{-1} =& (I - H(\tilde{c})) W_r^{-1}(c) W(\tilde{a}^{-1}) W_r^{-1}(d) \\ &+ H(a^{-1}) W_r^{-1}(d).\end{aligned} \tag{7.6}$$

For $g \in G$, the corresponding one-sided inverse of the operator $W(g)$ can be written by using Wiener-Hopf factorization of g [31]. We also note that if $W(c)$ and $W(d)$ are invertible, formulas (7.5) and (7.6) can be used to write the inverse operator for $W(a) + H(b)$. They also play an important role when establishing inverse operators in a variety of situations not covered by Theorem 7.5. The corresponding proofs run similar to considerations of Sects. 4 and 5, but there are essential technical differences because the corresponding kernel spaces can be infinite dimensional.

It is worth noting that the generalized invertibility of the operators $W(a) + H(b)$ can also be studied—cf. [23].

8 Generalized Toeplitz Plus Hankel Operators

Here we briefly discuss generalized Toeplitz plus Hankel operators. These operators are similar to classical Toeplitz plus Hankel operators considered in Sect. 2, but the flip operator J is replaced by another operator J_α generated by a linear fractional shift α. It turns out that the classical approach of Sects. 2–5 can also be used but the application of the method is not straightforward and requires solving various specific problems. Therefore, in this section we mainly focus on the description of the kernels and cokernels of the corresponding operators. These results lay down foundation for the invertibility study. We also feel that the Basor-Ehrhardt method can be realized in this situation, but we are not going to pursue this problem here.

The following construction is based on the considerations of [18]. Let $\mathbb{S}$ denote the Riemann sphere. We consider a mapping $\alpha\colon \mathbb{S} \to \mathbb{S}$ defined by

$$\alpha(z) := \frac{z-\beta}{\overline{\beta}z-1}, \tag{8.1}$$

where β is a complex number such that $|\beta| > 1$.

Let us recall basic properties of α.

1. The mapping $\alpha\colon \mathbb{S}^2 \to \mathbb{S}^2$ is one-to-one, $\alpha(\mathbb{T}) = \mathbb{T}$, and if $\mathbb{D} := \{z \in \mathbb{C} : |z| < 1\}$ is the interior of the unite circle $\mathbb{T}$ and $\overline{\mathbb{D}} := \mathbb{D} \cup \mathbb{T}$, then

$$\alpha(\mathbb{D}) = \mathbb{S}^2 \setminus \overline{\mathbb{D}}, \quad \alpha(\mathbb{S}^2 \setminus \overline{\mathbb{D}}) = \mathbb{D}.$$

   We note that α is an automorphism of the Riemann sphere and the mappings $H^p \to \overline{H^p}$, $h \mapsto h \circ \alpha$ and $\overline{H^p} \to H^p$, $h \mapsto h \circ \alpha$ are well-defined isomorphisms.
2. The mapping $\alpha\colon \mathbb{T} \to \mathbb{T}$ changes the orientation of $\mathbb{T}$, satisfies the Carleman condition $\alpha(\alpha(t)) = t$ for all $t \in \mathbb{T}$, and possesses two fixed points—viz.

$$t_+ = (1+\lambda)/\overline{\beta} \quad \text{and} \quad t_- = (1-\lambda)/\overline{\beta}, \tag{8.2}$$

   where $\lambda := i\sqrt{|\beta|^2 - 1}$.
3. The mapping α admits the factorization

$$\alpha(t) = \alpha_+(t)\, t^{-1}\, \alpha_-(t)$$

   with the factorization factors

$$\alpha_+(t) = \frac{t-\beta}{\lambda}, \quad \alpha_-(t) = \frac{\lambda t}{\overline{\beta}t - 1}.$$

4. On the space L^p, $1 < p < \infty$, the mapping α generates a bounded linear operator J_α, called weighted shift operator and defined by

$$J_\alpha \varphi(t) := t^{-1}\alpha_-(t)\varphi(\alpha(t)), \quad t \in \mathbb{T}.$$

Further, for $a \in L^\infty$ let a_α denote the composition of the functions a and α—i.e.

$$a_\alpha(t) := a(\alpha(t)), \quad t \in \mathbb{T}.$$

The operators J_α, P, Q and aI are connected with each other by the equations

$$J_\alpha^2 = I, \quad J_\alpha aI = a_\alpha J_\alpha, \quad J_\alpha P = QJ_\alpha, \quad J_\alpha Q = PJ_\alpha,$$

and for any $n \in \mathbb{Z}$, one has $(a^n)_\alpha = (a_\alpha)^n := a_\alpha^n$. In addition to the Toeplitz operator $T(a)$, any element $a \in L^\infty$ defines another operator $H_\alpha := P\alpha QJ_\alpha$, called generalized Hankel operator. Generalized Hankel operators are similar to classical Hankel operators $H(a)$. For example, analogously to (2.1) operators H_α are connected with Toeplitz operators by the relations

$$T(ab) = T(a)T(b) + H_\alpha(a)H_\alpha(b_\alpha),$$
$$H_\alpha(ab) = T(a)H_\alpha(b) + H_\alpha(a)T(b_\alpha).$$

On the space L^p, we now consider the operators of the form $T(a) + H_\alpha(b)$ and call them generalized Toeplitz plus Hankel operators generated by the functions a, b and the shift α.

The classical approach of Sect. 2 can be also employed to describe the kernels and cokernels of $T(a) + H_\alpha(b)$. To this aim we develop a suitable framework, which is not a straightforward extension of the methods of Sect. 2. Let us now assume that a belongs to the group of invertible elements GL^∞ and the duo a, b satisfy the condition

$$aa_\alpha = bb_\alpha. \tag{8.3}$$

Relation (8.3) is again called the matching condition and the duo (a, b) are α-matching functions. To each matching pair, one can assign another α-matching pair $(c, d) := (ab^{-1}, ab_\alpha^{-1})$ called the subordinated pair for (a, b). It is easily seen that $cc_\alpha = dd_\alpha = 1$. In what follows, any element $g \in L^\infty$ satisfying the relation $gg_\alpha = 1$ is referred to as α-matching function. The functions c and d can also be expressed in the form $c = b_\alpha a_\alpha^{-1}$, $d = b_\alpha^{-1}a$. Besides, if (c, d) is the subordinated pair for an α-matching pair (a, b), then $(\overline{d}, \overline{c})$ is the subordinated pair for the matching pair $(\overline{a}, \overline{b}_\alpha)$, which defines the adjoint operator

$$(T(a) + H_\alpha(b))^* = T(\overline{a}) + H_\alpha(\overline{b}_\alpha).$$

Rewrite the operator $J_\alpha : L^p \mapsto L^p$ in the form

$$J_\alpha \varphi(t) := \chi^{-1}(t)\varphi(\alpha(t)),$$

where $\chi(t) = t/(\alpha_-(t))$, and note that if α is the shift (8.1), then:

1. $\chi \in H^\infty$ is an α-matching function and wind $\chi = 1$, where wind χ denotes the winding number of the function χ with respect to the origin.
2. The function $\chi_\alpha \in \overline{H^\infty}$ and $\chi_\alpha(\infty) = 0$.
3. If $a, b \in L^\infty$ and n is a positive integer, then

$$T(a) + H_\alpha(b) = (T(a\chi^{-n}) + H_\alpha(b\chi^n))T(\chi^n).$$

These properties allow us to establish the following version of the Coburn-Simonenko Theorem for generalized Toeplitz plus Hankel operators.

Theorem 8.1 *Let $a \in GL^\infty$ and let A denote any of the operators $T(a) - H_\alpha(a\chi^{-1})$, $T(a) + H_\alpha(a\chi)$, $T(a) + H_\alpha(a)$, $T(a) - H_\alpha(a)$. Then*

$$\min\{\dim\ker A, \dim\operatorname{coker} A\} = 0.$$

Further development runs similar to the one presented in Sect. 2.3 and all results are valid if the operator $H(b)$ is replaced by $H_\alpha(b)$, $\widetilde{a}$ by a_α and b by b_α. For example, if (a, b) is an α-matching pair with the subordinated pair (c, d), then the block Toeplitz operator $T(V(a, b))$ with the matrix function

$$V(a, b) = \begin{pmatrix} 0 & d \\ -c & a_\alpha^{-1} \end{pmatrix},$$

can be represented in the form

$$\begin{aligned} T(V(a, b)) &= \begin{pmatrix} 0 & T(d) \\ -T(c) & T(a_\alpha^{-1}) \end{pmatrix} \\ &= \begin{pmatrix} -T(d) & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} 0 & -I \\ I & T(a_\alpha^{-1}) \end{pmatrix} \begin{pmatrix} -T(c) & 0 \\ 0 & I \end{pmatrix}. \end{aligned}$$

Moreover, assuming that $T(c)$ is invertible from the right and $T_r^{-1}(c)$ is one of the right inverses, we can again represent the kernel of $T(V(a, b))$ as the direct sum

$$\ker T(V(a, b)) = \Omega(c) \dotplus \widehat{\Omega}(d),$$

where

$$\Omega(c) := \left\{(\varphi, 0)^T : \varphi \in \ker T(c)\right\},$$

$$\widehat{\Omega}(d) := \left\{(T_r^{-1}(c)T(a_\alpha^{-1})s, s)^T : s \in \ker T(d)\right\}.$$

Further, we observe that the operators

$$\mathbf{P}_\alpha^{\pm}(c) := \frac{1}{2}(I \pm J_\alpha Qc P) : \ker T(c) \to \ker T(c),$$

$$\mathbf{P}_\alpha^{\pm}(d) := \frac{1}{2}(I \pm J_\alpha Qd P) : \ker T(d) \to \ker T(d),$$

are projections on the corresponding spaces and consider the translation operators

$$\varphi_\alpha^{\pm} : \mathbf{P}_\alpha^{\pm}(d) \to \ker(T(a) \pm H_\alpha(b)),$$

which are defined similar to (3.5), but with J and $\widetilde{a}^{-1}$ replaced by J_α and a_α^{-1}, respectively.

In this situation, Proposition 3.6 reads as follows.

Proposition 8.2 *Assume that $(a, b) \in L^\infty \times L^\infty$ is a Fredholm matching pair. If the operator $T(c)$ is right-invertible, then*

$$\ker(T(a) + H_\alpha(b)) = \operatorname{im} \mathbf{P}_\alpha^{-}(c) \dotplus \varphi_\alpha^{+}(\operatorname{im} \mathbf{P}_\alpha^{+}(d)),$$

$$\ker(T(a) - H_\alpha(b)) = \operatorname{im} \mathbf{P}_\alpha^{+}(c) \dotplus \varphi_\alpha^{-}(\operatorname{im} \mathbf{P}_\alpha^{-}(d)).$$

Thus there is a one-to-one correspondence between $\ker T(V(a, b))$ and $\ker \operatorname{diag}(T(a) + H_\alpha(b), T(a) - H_\alpha(b))$ and in order to establish the latter, we need an efficient description of the spaces $\operatorname{im} \mathbf{P}_\alpha^{+}(c)$ and $\operatorname{im} \mathbf{P}_\alpha^{+}(d)$.

Let g stand for one of the functions c or d. Then g is a matching function and $T(g)$ is a Fredholm Toeplitz operator. We may assume that $\operatorname{ind} T(g) > 0$, since only in this case at least one of the spaces $\operatorname{im} \mathbf{P}_\alpha^{+}(g)$ or $\operatorname{im} \mathbf{P}_\alpha^{-}(g)$ is non-trivial. The following factorization result is crucial for the description of $\operatorname{im} \mathbf{P}_\alpha^{\pm}(g)$.

Theorem 8.3 *If $g \in L^\infty$ satisfies the matching condition $gg_\alpha = 1$ and $\operatorname{wind} g = n$, $n \in \mathbb{Z}$, then g can be represented in the form*

$$g = \xi g_+ \chi^{-n}(g_+^{-1})_\alpha,$$

where g_+ and n occur in the Wiener-Hopf factorization

$$g = g_- t^{-n} g_+, \quad g_-(\infty) = 1,$$

of the function g, whereas $\xi \in \{-1, 1\}$ and is defined by

$$\xi = \left(\frac{\lambda}{\overline{\beta}}\right)^n g_+^{-1}\left(\frac{1}{\overline{\beta}}\right). \tag{8.4}$$

Definition 8.4 The number ξ in (8.4) is called the α-factorization signature, or simply, α-signature of g and is denoted by $\boldsymbol{\sigma}_\alpha(g)$.

The α-signature is used to describe the kernels of the operators $T(a) + H_\alpha(b)$ and for certain classes of generating functions it can be defined with relative ease. For example, assuming that the operator $T(g)$ is Fredholm, $n := \operatorname{ind} T(g)$ and g is Hölder continuous at the fixed point t_+ or t_- of (8.2), one can show that $\boldsymbol{\sigma}_\alpha(g) = g(t_+)$ or $\boldsymbol{\sigma}_\alpha(g) = g(t_-)(-1)^n$. For piecewise continuous functions, the situation is more complicated but still can be handled—cf. [18].

Theorem 8.5 *Let $g \in L^\infty$ be an α-matching function such that the operator $T(g) : H^p \to H^p$ is Fredholm and $n := \operatorname{ind} T(g) > 0$. If $g = g_- t^{-n} g_+$, $g_-(\infty) = 1$ is the corresponding Wiener–Hopf factorization of g in H^p, then the following systems of functions $\mathcal{B}_\alpha^\pm(g)$ form bases in the spaces $\operatorname{im} \mathbf{P}_\alpha^\pm(g)$:*

1. *If $n = 2m$, $m \in \mathbb{N}$, then*

$$\mathcal{B}_\alpha^\pm(g) := \{g_+^{-1}(\chi^{m-k-1} \pm \boldsymbol{\sigma}_\alpha(g)\chi^{m+k}) : k = 0, 1, \cdots, m-1\},$$

   and $\dim \operatorname{im} \mathbf{P}_\alpha^\pm(g) = m$.
2. *If $n = 2m + 1$, $m \in \mathbb{Z}_+$, then*

$$\mathcal{B}_\alpha^\pm(g) := \{g_+^{-1}(\chi^{m+k} \pm \boldsymbol{\sigma}_\alpha(g)\chi^{m-k}) : k = 0, 1, \cdots, m\} \setminus \{0\},$$
$$\dim \operatorname{im} \mathbf{P}_\alpha^\pm(g) = m + (1 \pm \boldsymbol{\sigma}_\alpha(g))/2\,,$$

   and the zero element belongs to one of the sets $\mathcal{B}_\alpha^+(g)$ or $\mathcal{B}_\alpha^-(g)$—viz. for $k = 0$ one of the terms $\chi^m(1 \pm \boldsymbol{\sigma}_\alpha(g))$ is equal to zero.

The above considerations provide a powerful tool for the study of invertibility of generalized Toeplitz plus Hankel operators and it should be clear that the results obtained in Sects. 4 and 5 can be extended to this class of operators. However, additional studies may be needed in certain situations. Nevertheless, let us mention one of such results here.

Proposition 8.6 *Assume that $(a, b) \in L^\infty \times L^\infty$ is a Fredholm matching pair and the operators $T(c)$ and $T(d)$ are right-invertible. Then $T(a) + H_\alpha(b)$ and $T(a) - H_\alpha(b)$ are also right-invertible and corresponding right inverses are given by*

$$(T(a) \pm H_\alpha(b))_r^{-1} = (I \mp H_\alpha(c_\alpha))T_r^{-1}(c)T(a_\alpha^{-1})T_r^{-1}(d) \pm H_\alpha(a^{-1})T_r^{-1}(d).$$

In conclusion of this section, we note the works [40, 41] where more general operators with the shift (8.1) are considered. However, the conditions imposed on coefficient functions are more restrictive and the results obtained are less complete.

9 Final Remarks

We considered various approaches to the study of invertibility of Toeplitz plus Hankel operators and their close relatives. Before concluding this survey, we would like to mention two more problems of special interest. The first one is the construction of Wiener-Hopf factorizations for multipliers acting on the spaces l^p and also for those on $L^p(\mathbb{R})$-spaces. Some ideas on how to proceed with this problem in the l^p-context have been noted in [21].

Another problem of interest is the study of the kernels and cokernels of Wiener-Hopf plus Hankel operators acting on $L^p(\mathbb{R})$-spaces, $p \neq 2$ and generated by multipliers from sets more involved than the set G considered in Sect. 7, such as piecewise continuous multipliers, for example.

In should be clear that the list of open problems in the theory of Toeplitz plus Hankel operators is not limited to those mentioned in this work and it is up to the interested reader to single out a one for further consideration.

Acknowledgments The authors express their sincere gratitude to anonymous referees for insightful comments and suggestions that helped to improve the paper.

References

1. J. Baik, E.M. Rains, Algebraic aspects of increasing subsequences. Duke Math. J. **109**, 1–65 (2001)
2. E.L. Basor, T. Ehrhardt, On a class of Toeplitz + Hankel operators. New York J. Math. **5**, 1–16 (1999)
3. E.L. Basor, T. Ehrhardt, Factorization theory for a class of Toeplitz + Hankel operators. J. Oper. Theory **51**, 411–433 (2004)
4. E.L. Basor, T. Ehrhardt, Fredholm and invertibility theory for a special class of Toeplitz + Hankel operators. J. Spectral Theory **3**, 171–214 (2013)
5. E.L. Basor, T. Ehrhardt, Asymptotic formulas for determinants of a special class of Toeplitz + Hankel matrices, in *Large Truncated Toeplitz Matrices, Toeplitz Operators, and Related Topics*, vol. 259. The Albrecht Böttcher Anniversary Volume. Operator Theory: Advances and Applications (Birkhäuser, Basel, 2017), pp. 125–154
6. E. Basor, Y. Chen, T. Ehrhardt, Painlevé V and time-dependent Jacobi polynomials. J. Phys. A **43**, 015204, 25 (2010)
7. A. Böttcher, B. Silbermann, *Analysis of Toeplitz Operators*, 2nd edn. Springer Monographs in Mathematics (Springer, Berlin, 2006)
8. A. Böttcher, Y.I. Karlovich, I.M. Spitkovsky, *Convolution Operators and Factorization of Almost Periodic Matrix Functions* (Birkhäuser, Basel, 2002)

9. L.P. Castro, A.P. Nolasco, A semi-Fredholm theory for Wiener-Hopf-Hankel operators with piecewise almost periodic Fourier symbols. J. Oper. Theory **62**, 3–31 (2009)
10. L.P. Castro, A.S. Silva, Wiener-Hopf and Wiener-Hopf-Hankel operators with piecewise-almost periodic symbols on weighted Lebesgue spaces. Mem. Diff. Equ. Math. Phys. **53**, 39–62 (2011)
11. K. Clancey, I. Gohberg, *Factorization of Matrix Functions and Singular Integral Operators* (Birkhäuser, Basel, 1981)
12. L.A. Coburn, R.G. Douglas, Translation operators on the half-line. Proc. Nat. Acad. Sci. USA **62**, 1010–1013 (1969)
13. V.D. Didenko, B. Silbermann, Index calculation for Toeplitz plus Hankel operators with piecewise quasi-continuous generating functions. Bull. London Math. Soc. **45**, 633–650 (2013)
14. V.D. Didenko, B. Silbermann, The Coburn-Simonenko Theorem for some classes of Wiener–Hopf plus Hankel operators. Publ. de l'Institut Mathématique **96**(110), 85–102 (2014)
15. V.D. Didenko, B. Silbermann, Some results on the invertibility of Toeplitz plus Hankel operators. Ann. Acad. Sci. Fenn. Math. **39**, 443–461 (2014)
16. V.D. Didenko, B. Silbermann, Structure of kernels and cokernels of Toeplitz plus Hankel operators. Integr. Equ. Oper. Theory **80**, 1–31 (2014)
17. V.D. Didenko, B. Silbermann, Generalized inverses and solution of equations with Toeplitz plus Hankel operators. Bol. Soc. Mat. Mex. **22**, 645–667 (2016)
18. V.D. Didenko, B. Silbermann, Generalized Toeplitz plus Hankel operators: kernel structure and defect numbers. Compl. Anal. Oper. Theory **10**, 1351–1381 (2016)
19. V.D. Didenko, B. Silbermann, Invertibility and inverses of Toeplitz plus Hankel operators. J. Oper. Theory **72**, 293–307 (2017)
20. V.D. Didenko, B. Silbermann, Kernels of Wiener-Hopf plus Hankel operators with matching generating functions, in *Recent Trends in Operator Theory and Partial Differential Equations*, vol. 258. The Roland Duduchava Anniversary Volume. Operator Theory: Advances and Applications (Birkhäuser, Basel, 2017), pp. 111–127
21. V.D. Didenko, B. Silbermann, Kernels of a class of Toeplitz plus Hankel operators with piecewise continuous generating functions, in *Contemporary Computational Mathematics – A Celebration of the 80th Birthday of Ian Sloan*, ed. by J. Dick, F.Y. Kuo, H. Woźniakowski (eds). (Springer, Cham, 2018), pp. 317–337
22. V.D. Didenko, B. Silbermann, The invertibility of Toeplitz plus Hankel operators with subordinated operators of even index. Linear Algebra Appl. **578**, 425–445 (2019)
23. V.D. Didenko, B. Silbermann, Invertibility issues for a class of Wiener-Hopf plus Hankel operators. J. Spectral Theory **11** (2021)
24. R.V. Duduchava, Wiener-Hopf integral operators with discontinuous symbols. Dokl. Akad. Nauk SSSR **211**, 277–280 (1973) (in Russian)
25. R.V. Duduchava, Integral operators of convolution type with discontinuous coefficients. Math. Nachr. **79**, 75–98 (1977)
26. R.V. Duduchava, *Integral Equations with Fixed Singularities* (BSB B. G. Teubner Verlagsgesellschaft, Leipzig, 1979)
27. R. Edwards, *Fourier Series. A Modern Introduction*, vol. 1. Graduate Texts in Mathematics, vol. 85 (Springer, Berlin, 1982)
28. T. Ehrhardt, *Factorization theory for Toeplitz+Hankel operators and singular integral operators with flip*. Habilitation Thesis, Technische Universität Chemnitz (2004)
29. T. Ehrhardt, Invertibility theory for Toeplitz plus Hankel operators and singular integral operators with flip. J. Funct. Anal. **208**, 64–106 (2004)
30. P.J. Forrester, N.E. Frankel, Applications and generalizations of Fisher-Hartwig asymptotics. J. Math. Phys. **45**, 2003–2028 (2004)
31. I.C. Gohberg, I.A. Feldman, *Convolution Equations and Projection Methods for Their Solution* (American Mathematical Society, Providence, 1974)
32. I. Gohberg, N. Krupnik, *One-Dimensional Linear Singular Integral Equations. I*, vol. 53. Operator Theory: Advances and Applications (Birkhäuser Verlag, Basel, 1992)

33. I. Gohberg, N. Krupnik, *One-Dimensional Linear Singular Integral Equations. II*, vol. 54. Operator Theory: Advances and Applications (Birkhäuser Verlag, Basel, 1992)
34. S. Grudsky, A. Rybkin, On positive type initial profiles for the KdV equation. Proc. Am. Math. Soc. **142**, 2079–2086 (2014)
35. S. Grudsky, A. Rybkin, Soliton theory and Hankel operators. SIAM J. Math. Anal. **47**, 2283–2323 (2015)
36. S.M. Grudsky, A.V. Rybkin, On the trace-class property of Hankel operators arising in the theory of the Korteweg-de Vries equation. Math. Notes **104**, 377–394 (2018)
37. P. Junghanns, R. Kaiser, A note on Kalandiya's method for a crack problem. Appl. Numer. Math. **149**, 52–64 (2020)
38. N.K. Karapetiants, S.G. Samko, On Fredholm properties of a class of Hankel operators. Math. Nachr. **217**, 75–103 (2000)
39. N.K. Karapetiants, S.G. Samko, *Equations with Involutive Operators* (Birkhäuser Boston Inc., Boston, 2001)
40. V.G. Kravchenko, A.B. Lebre, J.S. Rodríguez, Factorization of singular integral operators with a Carleman shift via factorization of matrix functions: the anticommutative case. Math. Nachr. **280**, 1157–1175 (2007)
41. V.G. Kravchenko, A.B. Lebre, J.S. Rodríguez, Factorization of singular integral operators with a Carleman backward shift: the case of bounded measurable coefficients. J. Anal. Math. **107**, 1–37 (2009)
42. N.Y. Krupnik, *Banach Algebras with Symbol and Singular Integral Operators*, vol. 26. Operator Theory: Advances and Applications (Birkhäuser Verlag, Basel, 1987)
43. A.B. Lebre, E. Meister, F.S. Teixeira, Some results on the invertibility of Wiener-Hopf-Hankel operators. Z. Anal. Anwend. **11**, 57–76 (1992)
44. G.S. Litvinchuk, I.M. Spitkovskii, *Factorization of Measurable Matrix Functions*, vol. 25. Operator Theory: Advances and Applications (Birkhäuser Verlag, Basel, 1987)
45. E. Meister, F. Penzel, F.-O. Speck, F.S. Teixeira, Two-media scattering problems in a half-space, in *Partial Differential Equations with Real Analysis*. Dedicated to Robert Pertsch Gilbert on the occasion of his 60th birthday (Longman Scientific & Technical, Harlow; John Wiley & Sons, Inc., New York, 1992), pp. 122–146
46. E. Meister, F.-O. Speck, F.S. Teixeira, Wiener-Hopf-Hankel operators for some wedge diffraction problems with mixed boundary conditions. J. Integral Equ. Appl. **4**, 229–255 (1992)
47. V.V. Peller, *Hankel Operators and Their Applications*. Springer Monographs in Mathematics (Springer, New York, 2003)
48. S.C. Power, C*-algebras generated by Hankel operators and Toeplitz operators. J. Funct. Anal. **31**, 52–68 (1979)
49. S. Roch, B. Silbermann, *Algebras of convolution operators and their image in the Calkin algebra*, vol. 90. Report MATH (Akademie der Wissenschaften der DDR Karl-Weierstrass-Institut für Mathematik, Berlin, 1990)
50. S. Roch, B. Silbermann, *A handy formula for the Fredholm index of Toeplitz plus Hankel operators*. Indag. Math. **23**, 663–689 (2012)
51. S. Roch, P.A. Santos, B. Silbermann, *Non-Commutative Gelfand Theories. A Tool-Kit for Operator Theorists and Numerical Analysts*. Universitext (Springer, London, 2011)
52. B. Silbermann, The C^**-algebra generated by Toeplitz and Hankel operators with piecewise quasicontinuous symbols*. Integr. Equ. Oper. Theory **10**, 730–738 (1987)
53. I.B. Simonenko, Some general questions in the theory of Riemann boundary problem. Math. USSR Izvestiya **2**, 1091–1099 (1968)
54. I.J. Šneĭberg, Spectral properties of linear operators in interpolation families of Banach spaces. Mat. Issled. **9**, 2(32), 214–229 (1974) (in Russian)
55. I.M. Spitkovskiĭ, The problem of the factorization of measurable matrix-valued functions. Dokl. Akad. Nauk SSSR **227**, 576–579 (1976) (in Russian)
56. F.S. Teixeira, Diffraction by a rectangular wedge: Wiener-Hopf-Hankel formulation. Integr. Equ. Oper. Theory **14**, 436–454 (1991)

K-Inner Functions and K-Contractions

Jörg Eschmeier and Sebastian Toth

Abstract For a large class of unitarily invariant reproducing kernel functions K on the unit ball $\mathbb{B}_d$ in $\mathbb{C}^d$, we characterize the K-inner functions on $\mathbb{B}_d$ as functions admitting a suitable transfer function realization. We associate with each K-contraction $T \in L(H)^d$ a canonical operator-valued K-inner function and extend a uniqueness theorem of Arveson for minimal K-dilations to our setting. We thus generalize results of Olofsson for m-hypercontractions on the unit disc and of the first named author for m-hypercontractions on the unit ball.

Keywords K-inner functions · K-contractions · Wandering subspaces

Mathematics Subject Classification (2010) Primary 47A13; Secondary 47A20,47A45, 47A48

1 Introduction

Let $\mathbb{B}_d \subset \mathbb{C}^d$ be the open Euclidean unit ball and let

$$k : \mathbb{D} \to \mathbb{C}, \quad k(z) = \sum_{n=0}^{\infty} a_n z^n$$

be an analytic function without zeros on the unit disc $\mathbb{D}$ in $\mathbb{C}$ such that $a_0 = 1$, $a_n > 0$ for all $n \in \mathbb{N}$ and such that

$$0 < \inf_{n\in\mathbb{N}} \frac{a_n}{a_{n+1}} \leq \sup_{n\in\mathbb{N}} \frac{a_n}{a_{n+1}} < \infty.$$

J. Eschmeier (✉) · S. Toth
Fachrichtung Mathematik, Universität des Saarlandes, Saarbrücken, Germany
e-mail: eschmei@math.uni-sb.de; toth@math.uni-sb.de

M. A. Bastos et al. (eds.), *Operator Theory, Functional Analysis and Applications*,
Operator Theory: Advances and Applications 282,
https://doi.org/10.1007/978-3-030-51945-2_8

Since k has no zeros, the reciprocal function $1/k \in \mathcal{O}(\mathbb{D})$ admits a Taylor expansion

$$(1/k)(z) = \sum_{n=0}^{\infty} c_n z^n \qquad (z \in \mathbb{D}).$$

The reproducing kernel

$$K : \mathbb{B}_d \times \mathbb{B}_d \to \mathbb{C},\ K(z, w) = k(\langle z, w\rangle)$$

defines an analytic functional Hilbert space H_K such that the row operator $M_z : H_K^d \to H_K$ is bounded and has closed range [2, Theorem A.1]. Typical examples of functional Hilbert spaces of this type on the unit ball $\mathbb{B}_d$ are the Drury–Arveson space, the Dirichlet space, the Hardy space and the weighted Bergman spaces.

Let $T = (T_1, \ldots, T_d) \in L(H)^d$ be a commuting tuple of bounded linear operators on a complex Hilbert space H and let $\sigma_T : L(H) \to L(H)$ be the map defined by $\sigma_T(X) = \sum_{i=1}^d T_i X T_i^*$. The tuple T is called a K-contraction if the limit

$$\frac{1}{K}(T) = \mathrm{SOT}\text{-}\sum_{n=0}^{\infty} c_n \sigma_T^n(1_H) = \mathrm{SOT}\text{-}\lim_{N\to\infty} \sum_{|\alpha|\le N} c_{|\alpha|}\gamma_\alpha T^\alpha T^{*\alpha}$$

exists and defines a positive operator. Here and in the following we use the abbreviation $\gamma_\alpha = |\alpha|!/\alpha!$ for $\alpha \in \mathbb{N}^d$.

If $K(z, w) = 1/(1 - \langle z, w\rangle)$ is the Drury–Arveson kernel, then under a natural pureness condition the K-contractions coincide with the commuting row contractions of class $C_{\cdot 0}$. If m is a positive integer and $K_m(z, w) = 1/(1-\langle z, w\rangle)^m$, then the pure K_m-contractions are precisely the row-m-hypercontractions of class $C_{\cdot 0}$ [17, Theorem 3.49] and [12, Lemma 2].

An operator-valued analytic function $W : \mathbb{B}_d \to L(\mathscr{E}_*, \mathscr{E})$ with Hilbert spaces $\mathscr{E}$ and $\mathscr{E}_*$ is called K-inner if the map $\mathscr{E}_* \to H_K(\mathscr{E})$, $x \mapsto Wx$, is a well-defined isometry and

$$(W\mathscr{E}_*) \perp M_z^\alpha(W\mathscr{E}_*) \qquad \text{for all } \alpha \in \mathbb{N}^d \setminus \{0\}.$$

Here $H_K(\mathscr{E})$ is the $\mathscr{E}$-valued functional Hilbert space on $\mathbb{B}_d$ with reproducing kernel $K_{\mathscr{E}} : \mathbb{B}_d \times \mathbb{B}_d \to L(\mathscr{E}), (z, w) \mapsto K(z, w)1_{\mathscr{E}}$. A well known more explicit description of the functional Hilbert space $H_K(\mathscr{E})$ is given by

$$H_K(\mathscr{E}) = \left\{ f = \sum_{\alpha\in\mathbb{N}^d} f_\alpha z^\alpha \in \mathscr{O}(\mathbb{B}_d, \mathscr{E});\quad \|f\|^2 = \sum_{\alpha\in\mathbb{N}^d} \frac{\|f_\alpha\|^2}{a_{|\alpha|}\gamma_\alpha} < \infty \right\}.$$

It was shown by Olofsson [13] that, for $d = 1$ and the Bergman-type kernel

$$K_m : \mathbb{D} \times \mathbb{D} \to \mathbb{C}, \ K_m(z, w) = \frac{1}{(1 - z\overline{w})^m} \quad (m \in \mathbb{N} \setminus \{0\}),$$

the K_m-inner functions $W : \mathbb{D} \to L(\mathscr{E}_*, \mathscr{E})$ are precisely the functions of the form

$$W(z) = D + C \sum_{k=1}^{m} (1 - zT^*)^{-k} B,$$

where $T \in L(H)$ is a pure m-hypercontraction on some Hilbert space H and $B \in L(\mathscr{E}_*, H)$, $C \in L(H, \mathscr{E})$ and $D \in L(\mathscr{E}_*, \mathscr{E})$ are bounded operators satisfying the operator equations

$$C^*C = (1/K_m)(T), \quad D^*C + B^*\Delta_T T^* = 0, \quad D^*D + B^*\Delta_T B = 1_{\mathscr{E}_*}.$$

Here $(1/K_m)(T)$ is the m-th order defect operator of T and

$$\Delta_T = \sum_{k=0}^{m-1} (-1)^k \binom{m}{k+1} T^k T^{*k}.$$

The single variable results of Olofsson were extended in a system theoretic framework by Ball and Bolotnikov in [6] (see also [5]) to a large class of kernels $K(z, w) = \sum_{n \geq 0} a_n (z\overline{w})^n$ on the unit disc $\mathbb{D}$ in $\mathbb{C}$. In [6] Beurling–Lax type representations for M_z-invariant subspaces $M \subset H_k(\mathscr{E})$ in terms of suitable K-inner function families $(\theta_k)_{k \geq 0}$ together with transfer function realizations for the functions θ_k are obtained and operator models using K-inner characteristic function families are developed for quite general kernels K on the unit disc.

In [10] the result of Olofsson was extended to the unit ball by showing that a corresponding characterization holds for functions $W : \mathbb{B}_d \to L(\mathscr{E}_*, \mathscr{E})$ that are K_m-inner with respect to the generalized Bergman kernels

$$K_m : \mathbb{B}_d \times \mathbb{B}_d \to \mathbb{C}, \ K_m(z, w) = 1/(1 - \langle z, w \rangle)^m (m \in \mathbb{N} \setminus \{0\}).$$

In the present note we show that results of Olofsson from [13, 14] hold true for a large class of kernels

$$K : \mathbb{B}_d \times \mathbb{B}_d \to \mathbb{C}, \ K(z, w) = \sum_{n=0}^{\infty} a_n \langle z, w \rangle^n$$

including all complete Nevanlinna–Pick kernels and powers $K_\nu(z, w) = 1/(1 - \langle z, w \rangle)^\nu$ of the Drury–Arveson kernel with positive real exponents. To prove that each K-inner function admits a transfer function realization as described above we

extend a uniqueness result for minimal K-dilations due to Arveson to our class of kernels.

2 Wandering Subspaces

Let $T = (T_1, \ldots, T_d) \in L(H)^d$ be a K-contraction, that is, a commuting tuple of bounded linear operators on a complex Hilbert space H such that the limit

$$\frac{1}{K}(T) = \text{SOT}-\sum_{n=0}^{\infty} c_n \sigma_T^n(1_H) = \text{SOT}-\lim_{N\to\infty} \sum_{|\alpha|\leq N} c_{|\alpha|}\gamma_\alpha T^\alpha T^{*\alpha}$$

exists and defines a positive operator. A K-contraction $T \in L(H)^d$ is said to be pure if

$$\text{SOT}-\lim_{N\to\infty} 1_H - \sum_{n=0}^{N} a_n \sigma_T^n(\frac{1}{K}(T)) = 0.$$

Let us define the defect operator and the defect space of a K-contraction T by

$$C = \frac{1}{K}(T)^{\frac{1}{2}} \text{ and } \mathscr{D} = \overline{\text{Im}\, C}.$$

We call an isometric linear map $j\colon H \to H_K(\mathscr{E})$ which intertwines the tuples $T^* \in L(H)^d$ and $M_z^* \in L(H_K(\mathscr{E}))^d$ componentwise a K-dilation of T. By definition a K-dilation $j\colon H \to H_K(\mathscr{E})$ is minimal if the only reducing subspace of $M_z \in L(H_K(\mathscr{E}))^d$ that contains the image of j is $H_K(\mathscr{E})$.

Exactly as for row-m-hypercontractions of class $C_{\cdot 0}$, one can construct a canonical K-dilation for each K-contraction.

Theorem 2.1 *Let $T \in L(H)^d$ be a pure K-contraction. Then*

$$j\colon H \to H_K(\mathscr{D}), \quad j(h) = \sum_{\alpha\in\mathbb{N}^d} a_{|\alpha|}\gamma_\alpha C T^{*\alpha} h z^\alpha$$

is a well-defined isometry such that $jT_i^* = M_{z_i}^* j$ for $i = 1, \ldots, d$.

For a proof, see [17, Theorem 2.15].

For $h \in H$ and $f = \sum_{\alpha\in\mathbb{N}^d} f_\alpha z^\alpha \in H_K(\mathscr{D})$,

$$\langle h, j^* f\rangle = \sum_{\alpha\in\mathbb{N}^d} \langle CT^{*\alpha}h, f_\alpha\rangle = \sum_{\alpha\in\mathbb{N}^d} \langle h, T^\alpha C f_\alpha\rangle.$$

An application of the uniform boundedness principle shows that the adjoint $j^*\colon H_K(\mathscr{D}) \to H$ of the isometry j acts as

$$j^*\left(\sum_{\alpha\in\mathbb{N}^d} f_\alpha z^\alpha\right) = \sum_{\alpha\in\mathbb{N}^d} T^\alpha C f_\alpha.$$

Since j intertwines T^* and M_z^* componentwise, the space

$$M = H_K(\mathscr{D}) \ominus \operatorname{Im} j \subset H_K(\mathscr{D})$$

is invariant for $M_z \in L(H_K(\mathscr{D}))^d$.

In the following we show that the wandering subspace of M_z restricted to M can be described in terms of a suitable K-inner function. Recall that a closed subspace $\mathcal{W} \subset H$ is called a wandering subspace for a commuting tuple $S \in L(H)^d$ if

$$\mathcal{W} \perp S^\alpha \mathcal{W} \qquad (\alpha \in \mathbb{N}^d \setminus \{0\}).$$

The space $\mathcal{W}$ is called a generating wandering subspace for S if in addition $H = \bigvee(S^\alpha \mathcal{W};\, \alpha \in \mathbb{N}^d)$. For each closed S-invariant subspace $L \subset H$, the space

$$W_S(L) = L \ominus \sum_{i=1}^d S_i L$$

is a wandering subspace for S, usually called the wandering subspace associated with S on L. If $\mathcal{W}$ is a generating wandering subspace for S, then an elementary argument shows that necessarily $\mathcal{W} = W_S(H)$.

In the following we write

$$W(M) = M \ominus \left(\sum_{i=1}^d M_{z_i} M\right)$$

for the wandering subspace associated with the restriction of M_z to the invariant subspace $M = H_K(\mathscr{D}) \ominus \operatorname{Im} j$. Our main tool will be the matrix operator

$$M_z^* M_z = (M_{z_i}^* M_{z_j})_{1\le i,j\le d} \in L(H_K(\mathscr{D})^d).$$

Since the row operator $M_z\colon H_K(\mathscr{D})^d \to H_K(\mathscr{D})$ has closed range [2, Theorem A.1], the operator

$$M_z^* M_z\colon \operatorname{Im} M_z^* \to \operatorname{Im} M_z^*$$

is invertible. We denote its inverse by $(M_z^* M_z)^{-1}$. In the following we make essential use of the operators

$$\delta\colon H_K(\mathscr{D}) \to H_K(\mathscr{D}), \quad \delta\left(\sum_{n=0}^{\infty}\sum_{|\alpha|=n} f_\alpha z^\alpha\right) = f_0 + \sum_{n=1}^{\infty}\frac{a_n}{a_{n-1}}\sum_{|\alpha|=n} f_\alpha z^\alpha$$

and

$$\Delta\colon H_K(\mathscr{D}) \to H_K(\mathscr{D}), \quad \Delta\left(\sum_{n=0}^{\infty}\sum_{|\alpha|=n} f_\alpha z^\alpha\right) = \sum_{n=0}^{\infty}\frac{a_{n+1}}{a_n}\sum_{|\alpha|=n} f_\alpha z^\alpha.$$

By definition δ and Δ are diagonal operators with respect to the orthogonal decomposition $H_K(\mathscr{D}) = \bigoplus_{n=0}^{\infty} H_n(\mathscr{D})$ of $H_K(\mathscr{D})$ into the spaces $H_n(\mathscr{D})$ of all $\mathscr{D}$-valued homogenous polynomials of degree n. Our hypotheses on the sequence (a_n/a_{n+1}) imply that δ and Δ are invertible positive operators on $H_K(\mathscr{D})$. An elementary calculation shows that

$$\delta M_{z_i} = M_{z_i}\Delta$$

for $i = 1, \ldots, d$.

Lemma 2.2 *For $f \in H_K(\mathscr{D})$, we have*

$$(M_z^* M_z)^{-1}(M_z^* f) = M_z^*\delta f = (\oplus\Delta)\, M_z^* f.$$

In particular the row operator

$$\delta M_z\colon H_K(\mathscr{D})^d \to H_K(\mathscr{D})$$

defines the trivial extension of the operator

$$M_z\left(M_z^* M_z\right)^{-1}\colon \operatorname{Im} M_z^* \to H_K(\mathscr{D}).$$

Proof Since the column operator M_z^* annihilates the constant functions, to prove the first identity, we may suppose that $f(0) = 0$. With respect to the orthogonal decomposition $H_K(\mathscr{D}) = \bigoplus_{n=0}^{\infty} H_n(\mathscr{D})$ the operator $M_z M_z^*$ acts as (Lemma 4.3 in [11])

$$M_z M_z^*\left(\sum_{n=0}^{\infty} f_n\right) = \sum_{n=1}^{\infty}\left(\frac{a_{n-1}}{a_n}\right) f_n.$$

Hence $M_z M_z^* \delta f = f$ and

$$\left(M_z^* M_z\right)^{-1} M_z^* f = \left(M_z^* M_z\right)^{-1} \left(M_z^* M_z\right) M_z^* \delta f = M_z^* \delta f = (\oplus \Delta) M_z^* f.$$

Since any two diagonal operators commute, it follows in particular that $M_z \left(M_z^* M_z\right)^{-1} M_z^* = \delta \left(M_z M_z^*\right)$. Thus also the second assertion follows. □

The preceding proof shows in particular that the orthogonal projection of $H_K(\mathscr{D})$ onto $\operatorname{Im} M_z$ acts as

$$P_{\operatorname{Im} M_z} = M_z (M_z^* M_z)^{-1} M_z^* = \delta (M_z M_z^*) = P_{H_K(\mathscr{D}) \ominus \mathscr{D}},$$

where $\mathscr{D} \subset H_K(\mathscr{D})$ is regarded as the closed subspace consisting of all constant functions. As in the single-variable case we call the operator defined by $M_z' = \delta M_z \in L(H_K(\mathscr{D})^d, H_K(\mathscr{D}))$ the Cauchy dual of the multiplication tuple M_z.

We use the operator $\Delta_T \in L(H)$ defined by

$$\Delta_T = j^* \Delta j$$

to give a first description of the wandering subspace $W(M)$ of M_z restricted to the invariant subspace $M = (\operatorname{Im} j)^\perp$.

Theorem 2.3 *A function $f \in H_K(\mathscr{D})$ is an element of the wandering subspace $W(M)$ of $M = (\operatorname{Im} j)^\perp \in \operatorname{Lat}(M_z, H_K(\mathscr{D}))$ if and only if*

$$f = f_0 + M_z'(j x_i)_{i=1}^d$$

for some vectors $f_0 \in \mathscr{D}$, $x_1, \ldots, x_d \in H$ with $(j x_i)_{i=1}^d \in M_z^* H_K(\mathscr{D})$ and

$$C f_0 + T(\Delta_T x_i)_{i=1}^d = 0.$$

In this case $(j x_i)_{i=1}^d = M_z^* f$.

Proof Note that a function $f \in H_K(\mathscr{D})$ belongs to the wandering subspace $W(M) = M \ominus \sum_{i=1}^d z_i M$ of M_z on $M = \operatorname{Ker} j^* \in \operatorname{Lat}(M_z, H_K(\mathscr{D}))$ if and only if $j^* f = 0$ and $(1_{H_K(\mathscr{D})} - j j^*) M_{z_i}^* f = 0$ for $i = 1, \ldots, d$. Using the remark following Lemma 2.2, we obtain, for $(x_i)_{i=1}^d \in H^d$ and $f \in H_K(\mathscr{D})$ with $(j x_i)_{i=1}^d = M_z^* f$,

$$\begin{aligned} j^* f &= j^*(f(0) + \delta M_z M_z^* f) \\ &= C f(0) + j^* M_z (\Delta j x_i)_{i=1}^d \\ &= C f(0) + T(j^* \Delta j x_i)_{i=1}^d \\ &= C f(0) + T(\Delta_T x_i)_{i=1}^d. \end{aligned}$$

Thus if $f \in W(M)$, then $(x_i)_{i=1}^d = (j^* M_{z_i}^* f)_{i=1}^d$ defines a tuple in H^d with $(jx_i)_{i=1}^d = M_z^* f$ such that $Cf(0) + T(\Delta_T x_i)_{i=1}^d = j^* f = 0$ and

$$f = f(0) + (f - f(0)) = f(0) + M_z(M_z^* M_z)^{-1} M_z^* f = f(0) + M_z'(jx_i)_{i=1}^d.$$

Conversely, if $f = f_0 + M_z'(jx_i)_{i=1}^d$ with $f_0 \in \mathscr{D}$, $x_1, \ldots, x_d$ as in Theorem 2.3, then using Lemma 2.2 we find that

$$M_z^* f = M_z^* M_z (M_z^* M_z)^{-1} (jx_i)_{i=1}^d = (jx_i)_{i=1}^d.$$

Since j is an isometry, it follows that $jj^* M_{z_i}^* f = jx_i = M_{z_i}^* f$ for $i = 1, \ldots, d$. Since $j^* f = Cf(0) + T(\Delta_T x_i)_{i=1}^d = 0$, we have shown that $f \in W(M)$. □

Lemma 2.4 *Let $T \in L(H)^d$ be a pure K-contraction and let*

$$f = f_0 + M_z'(jx_i)_{i=1}^d$$

be a representation of a function $f \in W(M)$ as in Theorem 2.3. Then we have

$$\|f\|^2 = \|f_0\|^2 + \sum_{i=1}^d \langle \Delta_T x_i, x_i \rangle.$$

Proof Since by Lemma 2.2

$$\operatorname{Im} M_z' = M_z(M_z^* M_z)^{-1} M_z^* H_K(\mathscr{D}) = \operatorname{Im} M_z = H_K(\mathscr{D}) \ominus \mathscr{D},$$

it follows that

$$\begin{aligned} \|f\|^2 - \|f_0\|^2 &= \|M_z'(jx_i)_{i=1}^d\|^2 \\ &= \langle (M_z^* M_z)^{-1} M_z^* f, (jx_i)_{i=1}^d \rangle \\ &= \langle (\oplus j^*) M_z^* \delta f, (x_i)_{i=1}^d \rangle \\ &= \langle (j^* \Delta j x_i)_{i=1}^d, (x_i)_{i=1}^d \rangle. \end{aligned}$$

Since by definition $\Delta_T = j^* \Delta j$, the assertion follows. □

Let $T \in L(H)^d$ be a pure K-contraction. Then $\Delta_T = j^* \Delta j$ is a positive operator with

$$\langle \Delta_T x, x \rangle = \|\Delta^{\frac{1}{2}} jx\|^2 \geq \|\Delta^{-\frac{1}{2}}\|^{-2} \|jx\|^2 = \|\Delta^{-1}\|^{-1} \|x\|^2$$

for all $x \in H$. Hence $\Delta_T \in L(H)$ is invertible and

$$(x, y) = \langle \Delta_T x, y \rangle$$

defines a scalar product on H such that the induced norm $\|\cdot\|_T$ is equivalent to the original norm with

$$\|\Delta^{\frac{1}{2}}\|\|x\| \geq \|x\|_T \geq \|\Delta^{-\frac{1}{2}}\|^{-1}\|x\|$$

for $x \in H$. We write $\tilde{H}$ for H equipped with the norm $\|\cdot\|_T$. Then

$$I_T : H \to \tilde{H}, \quad x \mapsto x$$

is an invertible bounded operator such that

$$\langle I_T^* x, y \rangle = \langle \Delta_T x, y \rangle \quad (x \in \tilde{H}, y \in H).$$

Hence $I_T^* x = \Delta_T x$ for $x \in \tilde{H}$. Let $\tilde{T} = (\tilde{T}_1, \ldots, \tilde{T}_d) : \tilde{H}^d \to H$ be the row operator with components $\tilde{T}_i = T_i \circ I_T^* \in L(\tilde{H}, H)$. Then

$$\begin{aligned} \tilde{T}\tilde{T}^* &= \sum_{i=1}^{d} T_i (I_T^* I_T) T_i^* = \sigma_T(\Delta_T) = \sigma_T(j^* \Delta j) = j^* M_z(\oplus \Delta) M_z^* j \\ &= j^*(\delta M_z M_z^*) j = j^* P_{H_K(\mathscr{D}) \ominus \mathscr{D}} j \end{aligned}$$

and hence $\tilde{T}$ is a contraction. As in [13] we use its defect operators

$$\begin{aligned} D_{\tilde{T}} &= (1_{\tilde{H}^d} - \tilde{T}^* \tilde{T})^{1/2} \in L(\tilde{H}^d), \\ D_{\tilde{T}^*} &= (1_H - \tilde{T}\tilde{T}^*)^{1/2} = (j^* P_{\mathscr{D}} j)^{1/2} = C \in L(H). \end{aligned}$$

Here the identity $(j^* P_{\mathscr{D}} j)^{1/2} = C$ follows from the definition of j and the representation of j^* explained in the section following Theorem 2.1. We write $\mathscr{D}_{\tilde{T}} = \overline{D_{\tilde{T}} \tilde{H}^d} \subset \tilde{H}^d$ and $\mathscr{D}_{\tilde{T}^*} = \overline{D_{\tilde{T}^*} H} = \mathscr{D}$ for the defect spaces of $\tilde{T}$. As in the classical single-variable theory of contractions it follows that $\tilde{T} D_{\tilde{T}} = D_{\tilde{T}^*} \tilde{T}$ and that

$$U = \left(\begin{array}{c|c} \tilde{T} & D_{\tilde{T}^*} \\ \hline D_{\tilde{T}} & -\tilde{T}^* \end{array} \right) : \tilde{H}^d \oplus \mathscr{D}_{\tilde{T}^*} \to H \oplus \mathscr{D}_{\tilde{T}}$$

is a well-defined unitary operator. In the following we define an analytically parametrized family $W_T(z) \in L(\tilde{\mathscr{D}}, \mathscr{D})$ $(z \in \mathbb{B})$ of operators on the closed subspace

$$\tilde{\mathscr{D}} = \{ y \in \mathscr{D}_{\tilde{T}};\ (\oplus j I_T^{-1}) D_{\tilde{T}} y \in M_z^* H_K(\mathscr{D}) \} \subset \mathscr{D}_{\tilde{T}}$$

such that

$$W(M) = \{W_T x;\ x \in \tilde{\mathscr{D}}\},$$

where $W_T x : \mathbb{B}_d \to \mathscr{D}$ acts as $(W_T x)(z) = W_T(z)x$. We equip $\tilde{\mathscr{D}}$ with the norm $\|y\| = \|y\|_{\tilde{H}^d}$ that it inherits as a closed subspace $\tilde{\mathscr{D}} \subset \tilde{H}^d$.

Lemma 2.5 *Let $T \in L(H)^d$ be a pure K-contraction. Then a function $f \in H_K(\mathscr{D})$ belongs to the wandering subspace $W(M)$ of*

$$M = (\operatorname{Im} j)^{\perp} \in \operatorname{Lat}(M_z, H_K(\mathscr{D}))$$

if and only if there is a vector $y \in \tilde{\mathscr{D}}$ with

$$f = -\tilde{T} y + M_z'(\oplus j I_T^{-1}) D_{\tilde{T}} y.$$

In this case $\|f\|^2 = \|y\|^2_{\tilde{H}^d}$.

Proof By Theorem 2.3 a function $f \in H_K(\mathscr{D})$ belongs to $W(M)$ if and only if it is of the form

$$f = f_0 + M_z'(j x_i)_{i=1}^d$$

with $f_0 \in \mathscr{D}$ and $x_1, \ldots, x_d \in H$ such that $(j x_i)_{i=1}^d \in M_z^* H_K(\mathscr{D})$ and

$$\tilde{T}(I_T x_i)_{i=1}^d + D_{\tilde{T}^*} f_0 = 0.$$

Then $y = D_{\tilde{T}}(I_T x_i)_{i=1}^d - \tilde{T}^* f_0 \in \mathscr{D}_{\tilde{T}}$ is a vector with

$$U \begin{pmatrix} (I_T x_i) \\ f_0 \end{pmatrix} = \begin{pmatrix} 0 \\ y \end{pmatrix},$$

or equivalently, with

$$\begin{pmatrix} (I_T x_i) \\ f_0 \end{pmatrix} = U^* \begin{pmatrix} 0 \\ y \end{pmatrix} = \begin{pmatrix} D_{\tilde{T}} y \\ -\tilde{T} y \end{pmatrix}.$$

But then $y \in \tilde{\mathscr{D}}$ and $f = -\tilde{T} y + M_z'(\oplus j I_T^{-1}) D_{\tilde{T}} y$. Conversely, if f is of this form, then using the definitions of $\tilde{T}$, $\tilde{\mathscr{D}}$ and the intertwining relation $\tilde{T} D_{\tilde{T}} = D_{\tilde{T}^*} \tilde{T}$ one can easily show that the vectors defined by

$$f_0 = -\tilde{T} y \in \mathscr{D} \text{ and } (x_i)_{i=1}^d = (\oplus I_T^{-1}) D_{\tilde{T}} y \in H^d$$

yield a representation $f = f_0 + M'_z(jx_i)_{i=1}^d$ as in Theorem 2.3. By Lemma 2.4 and the definition of the scalar product on $\tilde{H}$ we find that

$$\begin{aligned}\|f\|^2 &= \|f_0\|^2 + \sum_{i=1}^d \langle \Delta_T x_i, x_i\rangle = \|\tilde{T}y\|^2 + \sum_{i=1}^d \|I_T x_i\|_{\tilde{H}}^2 \\ &= \|\tilde{T}y\|^2 + \|D_{\tilde{T}}y\|_{\tilde{H}^d}^2 = \|y\|_{\tilde{H}^d}^2. \qquad \square\end{aligned}$$

Recall that the reproducing kernel $K : \mathbb{B}_d \times \mathbb{B}_d \to \mathbb{C}$ is defined by $K(z, w) = k(\langle z, w\rangle)$, where

$$k\colon \mathbb{D} \to \mathbb{C}, \quad k(z) = \sum_{n=0}^{\infty} a_n z^n$$

is an analytic function with $a_0 = 1$, $a_n > 0$ for all n such that

$$0 < \inf_n \frac{a_n}{a_{n+1}} \leq \sup_n \frac{a_n}{a_{n+1}} < \infty.$$

Let us suppose in addition that the limit

$$r = \lim_{n\to\infty} \frac{a_n}{a_{n+1}}$$

exists. Then $r \in [1, \infty)$ is the radius of convergence of the power series defining k and by Theorem 4.5 in [11] the Taylor spectrum of $M_z \in L(H_K(\mathscr{D}))^d$ is given by

$$\sigma(M_z) = \{z \in \mathbb{C}^d;\ \|z\| \leq \sqrt{r}\}.$$

If $T \in L(H)^d$ is a pure K-contraction, then T^* is unitarily equivalent to a restriction of M_z^* and hence

$$\sigma(T^*) \subset \{z \in \mathbb{C}^d;\ \|z\| \leq \sqrt{r}\}.$$

The function $F\colon D_r(0) \to \mathbb{C}$, $F(z) = \sum_{n=0}^{\infty} a_{n+1} z^n$, is analytic on the open disc $D_r(0)$ with radius r and center 0 and satisfies

$$F(z) = \frac{k(z) - 1}{z} \quad (z \in D_r(0) \setminus \{0\}).$$

For $z \in \mathbb{B}_d$, let us denote by $Z\colon H^d \to H$, $(h_i)_{i=1}^d \mapsto \sum_{i=1}^d z_i h_i$, the row operator induced by z. As a particular case of a much more general analytic spectral

mapping theorem for the Taylor spectrum [9, Theorem 2.5.10] we find that

$$\sigma(ZT^*) = \{\sum_{i=1}^{d} z_i w_i;\ w \in \sigma(T^*)\} \subset D_r(0)$$

for $z \in \mathbb{B}_d$. Thus we can define an operator-valued function $F_T : \mathbb{B}_d \to L(H)$,

$$F_T(z) = F(ZT^*) = \sum_{n=0}^{\infty} a_{n+1} \left(\sum_{|\alpha|=n} \gamma_\alpha T^{*\alpha} z^\alpha \right).$$

Lemma 2.6 *For $(x_i)_{i=1}^d \in H^d$ and $z \in \mathbb{B}_d$,*

$$CF(ZT^*)Z(x_i)_{i=1}^d = (\delta M_z(jx_i)_{i=1}^d)(z).$$

Proof For $(x_i)_{i=1}^d \in H^d$,

$$\begin{aligned}
\delta M_z(jx_i)_{i=1}^d &= \sum_{i=1}^{d} \delta M_{z_i} \sum_{n=0}^{\infty} a_n \left(\sum_{|\alpha|=n} \gamma_\alpha CT^{*\alpha} x_i z^\alpha \right) \\
&= \sum_{i=1}^{d} \sum_{n=0}^{\infty} a_n \delta \left(\sum_{|\alpha|=n} \gamma_\alpha CT^{*\alpha} x_i z^{\alpha+e_i} \right) \\
&= \sum_{i=1}^{d} \sum_{n=0}^{\infty} a_{n+1} \sum_{|\alpha|=n} \gamma_\alpha CT^{*\alpha} x_i z^{\alpha+e_i},
\end{aligned}$$

where the series converge in $H_K(\mathscr{D})$. Since the point evaluations are continuous on $H_K(\mathscr{D})$, we obtain

$$\begin{aligned}
\left(\delta M_z(jx_i)_{i=1}^d\right)(z) &= \sum_{n=0}^{\infty} a_{n+1} \sum_{|\alpha|=n} \gamma_\alpha CT^{*\alpha} \left(\sum_{i=1}^{d} z_i x_i \right) z^\alpha \\
&= CF(ZT^*)Z(x_i)_{i=1}^d
\end{aligned}$$

for all $z \in \mathbb{B}_d$. □

By Lemma 2.6 the map $W_T : \mathbb{B}_d \to L(\tilde{\mathscr{D}}, \mathscr{D})$,

$$\begin{aligned}
W_T(z)(x) &= -T(\oplus \Delta_T I_T^{-1})x + CF(ZT^*)Z(\oplus I_T^{-1})D_{\tilde{T}}x \\
&= -\tilde{T}x + CF(ZT^*)Z(\oplus I_T^{-1})D_{\tilde{T}}x
\end{aligned}$$

defines an analytic operator-valued function.

Theorem 2.7 *Let $T \in L(H)^d$ be a pure K-contraction. Then*

$$W(M) = \{W_T x;\ x \in \tilde{\mathscr{D}}\}$$

and $\|W_T x\| = \|x\|$ for $x \in \tilde{\mathscr{D}}$.

Proof For $x \in \tilde{\mathscr{D}}$, Lemma 2.6 implies that

$$W_T = -\tilde{T}x + \delta M_z(\oplus j I_T^{-1}) D_{\tilde{T}} x = -\tilde{T}x + M_z'(\oplus j I_T^{-1}) D_{\tilde{T}} x.$$

Thus the assertion follows from Lemma 2.5. □

Since $W(M)$ is a wandering subspace for M_z, the map $W_T : \mathbb{B}_d \to L(\tilde{\mathscr{D}}, \mathscr{D})$ is an operator-valued analytic function such that $\tilde{\mathscr{D}} \to H_K(\mathscr{D})$, $x \mapsto W_T x$, is an isometry and

$$W_T(\tilde{\mathscr{D}}) \perp M_z^\alpha \left(W_T(\tilde{\mathscr{D}})\right) \text{ for all } \alpha \in \mathbb{N}^d \setminus \{0\}.$$

Thus $W_T : \mathbb{B}_d \to L(\tilde{\mathscr{D}}, \mathscr{D})$ is a K-inner function with $W_T(\tilde{\mathscr{D}}) = W(M)$. In the case that $M_z \in L(H_K)^d$ is a row contraction one can show that each K-inner function $W : \mathbb{B}_d \to L(\tilde{\mathscr{E}}, \mathscr{E})$ defines a contractive multiplier

$$M_W : H_d^2(\mathscr{E}) \to H_K, \quad f \mapsto Wf$$

from the $\mathscr{E}$-valued Drury–Arveson space $H_d^2(\mathscr{E})$ to $H_K(\tilde{\mathscr{E}})$ [3, Theorem 6.2].

3 K-Inner Functions

In the previous section we saw that the K-inner function $W_T : \mathbb{B}_d \to L(\tilde{\mathscr{D}}, \mathscr{D})$ associated with a pure K-contraction $T \in L(H)^d$ has the form

$$W_T(z) = D + CF(ZT^*)ZB,$$

where $C = \left(\frac{1}{K}(T)\right)^{\frac{1}{2}} \in L(H, \mathscr{D})$, $D = -\tilde{T} \in L(\tilde{\mathscr{D}}, \mathscr{D})$ and $B = (\oplus I_T^{-1}) D_{\tilde{T}} \in L(\tilde{\mathscr{D}}, H^d)$. An elementary calculation using the definitions and the intertwining relation $\tilde{T} D_{\tilde{T}} = D_{\tilde{T}^*} \tilde{T}$ shows that the operators T, B, C, D satisfy the conditions

$$\text{(K1)} \quad C^*C = \frac{1}{K}(T),$$

$$\text{(K2)} \quad D^*C + B^*(\oplus \Delta_T) T^* = 0,$$

$$\text{(K3)} \quad D^*D + B^*(\oplus \Delta_T)B = 1_{\tilde{\mathscr{D}}},$$

$$\text{(K4)} \quad \mathrm{Im}((\oplus j)B) \subset M_z^* H_K(\mathscr{D}).$$

If $\mathscr{E}$ is a Hilbert space and $C \in L(H, \mathscr{E})$ is any operator with $C^*C = \frac{1}{K}(T)$, then exactly as in the proof of Proposition 2.6 from [17] it follows that

$$j_C \colon H \to H_K(\mathscr{E}), \quad j_C(x) = \sum_{\alpha \in \mathbb{N}^d} a_{|\alpha|} \gamma_\alpha (CT^{*\alpha} x) z^\alpha$$

is a well defined isometry that intertwines the tuples $T^* \in L(H)^d$ and $M_z^* \in L(H_K(\mathscr{E}))^d$ componentwise. As in the section following Theorem 2.1 one can show that

$$j_C^* f = \sum_{\alpha \in \mathbb{N}^d} T^\alpha C^* f_\alpha$$

for $f = \sum_{\alpha \in \mathbb{N}^d} f_\alpha z^\alpha \in H_K(\mathscr{E})$. Hence we find that

$$\begin{aligned} j_C^* \Delta j_C x &= j_C^* \Delta \sum_{\alpha \in \mathbb{N}^d} a_{|\alpha|} \gamma_\alpha (CT^{*\alpha} x) z^\alpha \\ &= j_C^* \sum_{\alpha \in \mathbb{N}^d} a_{|\alpha|+1} \gamma_\alpha (CT^{*\alpha} x) z^\alpha \\ &= \sum_{\alpha \in \mathbb{N}^d} a_{|\alpha|+1} \gamma_\alpha (T^\alpha C^* C T^{*\alpha} x) \\ &= \sum_{\alpha \in \mathbb{N}^d} a_{|\alpha|+1} \gamma_\alpha (T^\alpha \frac{1}{K}(T) T^{*\alpha} x) \end{aligned}$$

for all $x \in H$. By performing the same chain of calculations with j_C replaced by the canonical K-dilation j of T from Theorem 2.1 we obtain that

$$j_C^* \Delta j_C = j^* \Delta j = \Delta_T.$$

Our next aim is to show that any matrix operator

$$\left(\begin{array}{c|c} T^* & B \\ \hline C & D \end{array}\right) : H \oplus \mathscr{E}_* \to H^d \oplus \mathscr{E},$$

where T is a pure K-contraction and T, B, C, D satisfy the conditions (K1)–(K3) with $(\tilde{\mathscr{D}}, \mathscr{D})$ replaced by $(\mathscr{E}_*, \mathscr{E})$ and

$$\text{(K4)} \quad \mathrm{Im}((\oplus j_C)B) \subset M_z^* H_K(\mathscr{E})$$

gives rise to a K-inner function $W\colon \mathbb{B}_d \to L(\mathscr{E}_*, \mathscr{E})$ defined as

$$W(z) = D + CF(ZT^*)ZB$$

and that, conversely, under a natural condition on the kernel K each K-inner function is of this form.

Theorem 3.1 *Let $W\colon \mathbb{B}_d \to L(\mathscr{E}_*, \mathscr{E})$ be an operator-valued function between Hilbert spaces $\mathscr{E}_*$ and $\mathscr{E}$ such that*

$$W(z) = D + CF(ZT^*)ZB \quad (z \in \mathbb{B}_d),$$

where $T \in L(H)^d$ is a pure K-contraction and the matrix operator

$$\left(\begin{array}{c|c} T^* & B \\ \hline C & D \end{array}\right) : H \oplus \mathscr{E}_* \to H^d \oplus \mathscr{E}$$

satisfies the condition (K1)–(K4). Then W is a K-inner function.

Proof The space $M = H_K(\mathscr{E}) \ominus \operatorname{Im} j_C \subset H_K(\mathscr{E})$ is a closed M_z-invariant subspace. Let $x \in \mathscr{E}_*$ be a fixed vector. By condition (K4) there is a function $f \in H_K(\mathscr{E})$ with $(\oplus j_C)Bx = M_z^* f$. Exactly as in the proof of Lemma 2.6 it follows that

$$CF(ZT^*)ZBx = \delta M_z(\oplus j_C)Bx(z) = \delta M_z M_z^* f(z)$$

for all $z \in \mathbb{B}_d$. Since $\delta(M_z M_z^*) = P_{\operatorname{Im} M_z}$ is an orthogonal projection and since $\delta M_z = M_z(\oplus\Delta)$, we find that

$$\begin{aligned} \|Wx\|^2_{H_K(\mathscr{E})} - \|Dx\|^2 &= \langle \delta M_z M_z^* f, f\rangle_{H_K(\mathscr{E})} \\ &= \langle \oplus(j_C^*\Delta j_C)Bx, Bx\rangle_{H^d} \\ &= \langle (\oplus\Delta_T)Bx, Bx\rangle_{H^d} \\ &= \langle (1_{\mathscr{E}_*} - D^*D)x, x\rangle \\ &= \|x\|^2 - \|Dx\|^2. \end{aligned}$$

Hence the map $\mathscr{E}_* \to H_K(\mathscr{E}), \quad x \mapsto Wx$, is a well-defined isometry. Using the second part of Lemma 2.2 we obtain

$$M_z^*(Wx) = M_z^*\delta M_z M_z^* f = M_z^* f = (\oplus j_C)Bx$$

and hence that $P_M M_{z_i}^*(Wx) = (1_{H_K(\mathscr{E})} - j_C j_C^*) M_{z_i}^*(Wx) = 0$ for $i = 1, \cdots, d$. To see that $W\mathscr{E}_* \subset M$ note that with x and f as above

$$\begin{aligned} j_C^*(Wx) &= C^* Dx + j_C^*(\delta M_z M_z^* f) \\ &= C^* Dx + j_C^*(M_z(\oplus\Delta) M_z^* f) \\ &= C^* Dx + T(\oplus j_C^* \Delta j_C) Bx \\ &= C^* Dx + T(\oplus\Delta_T) Bx \\ &= 0. \end{aligned}$$

Thus we have shown that $W\mathscr{E}_* \subset M \ominus \sum_{i=1}^d z_i M$ which implies that

$$W\mathscr{E}_* \perp z^\alpha(W\mathscr{E}_*)$$

for all $\alpha \in \mathbb{N}^d \setminus \{0\}$. □

To prove that conversely each K-inner function $W : \mathbb{B}_d \to L(\mathscr{E}_*, \mathscr{E})$ has the form described in Theorem 3.1 we make the additional assumption that the multiplication tuple $M_z \in L(H_K)^d$ is a K-contraction. This hypothesis is satisfied, for instance, if H_K is a complete Nevanlinna–Pick space such as the Drury–Arveson space or the Dirichlet space or if K is a power

$$K_\nu : \mathbb{B}_d \times \mathbb{B}_d \to \mathbb{C}, \quad K_\nu(z, w) = \frac{1}{(1 - \langle z, w\rangle)^\nu} \quad (\nu \in (0, \infty))$$

of the Drury–Arveson kernel (see the discussion following Theorem 4.2). In the proof we shall use a uniqueness result for minimal K-dilations whose proof we postpone to Sect. 4.

Theorem 3.2 *Let $M_z \in L(H_K)^d$ be a K-contraction. If $W : \mathbb{B}_d \to L(\mathscr{E}_*, \mathscr{E})$ is a K-inner function, then there exist a pure K-contraction $T \in L(H)^d$ and a matrix operator*

$$\left(\begin{array}{c|c} T^* & B \\ \hline C & D \end{array}\right) \in L(H \oplus \mathscr{E}_*, H^d \oplus \mathscr{E})$$

satisfying the conditions (K1)–(K4) *such that*

$$W(z) = D + CF(ZT^*)ZB \quad (z \in \mathbb{B}_d).$$

Proof Since W is K-inner, the space

$$\mathscr{W} = W\mathscr{E}_* \subset H_K(\mathscr{E})$$

is a generating wandering subspace for $M_z \in L(H_K(\mathscr{E}))^d$ restricted to

$$\mathscr{S} = \bigvee_{\alpha \in \mathbb{N}^d} M_z^\alpha \mathscr{W} \subset H_K(\mathscr{E}).$$

The compression $T = P_H M_z|_H$ of $M_z \in L(H_K(\mathscr{E}))^d$ to the M_z^*-invariant subspace $H = H_K(\mathscr{E}) \ominus \mathscr{S}$ is easily seen to be a pure K-contraction [17, Proposition 2.12 and Lemma 2.21]. Let $\mathscr{R} \subset H_K(\mathscr{E})$ be the smallest reducing subspace for $M_z \in L(H_K(\mathscr{E}))^d$ that contains H. By Lemma 4.4

$$\mathscr{R} = \bigvee_{\alpha \in \mathbb{N}^d} z^\alpha (\mathscr{R} \cap \mathscr{E}) = H_K(\mathscr{R} \cap \mathscr{E}).$$

Thus the inclusion map $i : H \to H_K(\mathscr{R} \cap \mathscr{E})$ is a minimal K-dilation for T. Let $j : H \to H_K(\mathscr{D})$ be the K-dilation of the pure K-contraction $T \in L(H)^d$ defined in Theorem 2.1. Since also j is a minimal K-dilation for T (Corollary 4.5), by Corollary 4.3 there is a unitary operator $U : \mathscr{D} \to \mathscr{R} \cap \mathscr{E}$ such that

$$i = (1_{H_K} \otimes U) j.$$

Define $\hat{\mathscr{E}} = \mathscr{E} \ominus (\mathscr{R} \cap \mathscr{E})$. By construction

$$H_K(\hat{\mathscr{E}}) = H_K(\mathscr{E}) \ominus H_K(\mathscr{R} \cap \mathscr{E}) = H_K(\mathscr{E}) \ominus \mathscr{R} \subset \mathscr{S}$$

is the largest reducing subspace for $M_z \in L(H_K(\mathscr{E}))^d$ contained in $\mathscr{S}$. In particular, the space $\mathscr{S}$ admits the orthogonal decomposition

$$\mathscr{S} = H_K(\hat{\mathscr{E}}) \oplus (\mathscr{S} \cap H_K(\hat{\mathscr{E}})^\perp) = H_K(\hat{\mathscr{E}}) \oplus (H_K(\mathscr{R} \cap \mathscr{E}) \ominus \mathscr{S}^\perp).$$

We complete the proof by comparing the given K-inner function $W : \mathbb{B}_d \to L(\mathscr{E}_*, \mathscr{E})$ with the K-inner function $W_T : \mathbb{B}_d \to L(\tilde{\mathscr{D}}, \mathscr{D})$ associated with the pure K-contraction $T \in L(H)^d$. For this purpose, let us define the M_z-invariant subspace

$$M = H_K(\mathscr{D}) \ominus \operatorname{Im} j$$

and its wandering subspace

$$W(M) = M \ominus \left(\sum_{i=1}^d z_i M \right)$$

as in Sect. 2. Using the identity $i = (1_{H_K} \otimes U) j$ one obtains that

$$1_{H_K} \otimes U : M \to H_K(\mathscr{R} \cap \mathscr{E}) \ominus \mathscr{S}^\perp = H_K(\mathscr{R} \cap \mathscr{E}) \cap \mathscr{S}$$

defines a unitary operator that intertwines the restrictions of M_z to both sides componentwise. Consequently we obtain the orthogonal decomposition

$$\begin{aligned}\mathscr{W} &= W_{M_z}(\mathscr{S}) = W_{M_z}(H_K(\hat{\mathscr{E}})) \oplus W_{M_z}(H_K(\mathscr{R} \cap \mathscr{E}) \cap \mathscr{S}) \\ &= \hat{\mathscr{E}} \oplus (1_{H_K} \otimes U)W(M).\end{aligned}$$

Let $W_T : \mathbb{B}_d \to L(\tilde{\mathscr{D}}, \mathscr{D})$ be the K-inner function associated with the pure K-contraction $T \in L(H)^d$. Then there is a matrix operator

$$\left(\begin{array}{c|c} T^* & B \\ \hline C & D \end{array}\right) \in L(H \oplus \tilde{\mathscr{D}}, H^d \oplus \mathscr{D})$$

such that

$$W_T(z) = D + CF(ZT^*)ZB \quad (z \in \mathbb{B}_d)$$

and $W(M) = \{W_T x;\ x \in \tilde{\mathscr{D}}\}$ (see the beginning of Sect. 3 and Theorem 2.7). Let us denote by

$$P_1 : \mathscr{W} \to \hat{\mathscr{E}} \text{ and } P_2 : \mathscr{W} \to (1_{H_K} \otimes U)W(M)$$

the orthogonal projections. The K-inner functions $W : \mathbb{B}_d \to L(\mathscr{E}_*, \mathscr{E})$ and $W_T : \mathbb{B}_d \to L(\tilde{\mathscr{D}}, \mathscr{D})$ induce unitary operators

$$\mathscr{E}_* \to \mathscr{W}, \quad x \mapsto Wx$$

and

$$\tilde{\mathscr{D}} \to W(M) \quad x \mapsto W_T x.$$

We define surjective bounded linear operators by

$$U_1 : \mathscr{E}_* \to \hat{\mathscr{E}}, \quad U_1 x = P_1 W x$$

and

$$U_2 : \mathscr{E}_* \to \tilde{\mathscr{D}}, \quad U_2 x = \tilde{x} \text{ if } (1_{H_K} \otimes U)W_T \tilde{x} = P_2 W x.$$

By construction the column operator

$$(U_1, U_2) : \mathscr{E}_* \to \hat{\mathscr{E}} \oplus \tilde{\mathscr{D}}$$

defines an isometry such that

$$W(z)x = U_1 x + U W_T(z) U_2 x = (U_1 + U D U_2)x + (UC)F(ZT^*)Z(BU_2)x$$

holds for $z \in \mathbb{B}_d$ and $x \in \mathscr{E}_*$. To complete the proof we show that the operators

$$T \in L(H^d, H), \quad \tilde{B} = BU_2 \in L(\mathscr{E}_*, H^d), \quad \tilde{C} = UC \in L(H, \mathscr{E})$$
$$\text{and } \tilde{D} = U_1 + U D U_2 \in L(\mathscr{E}_*, \mathscr{E})$$

satisfy the conditions (K1)–(K4). To see this note that

$$\tilde{C}^* \tilde{C} = C^* U^* U C = C^* C = \frac{1}{K}(T)$$

and

$$\begin{aligned} \tilde{D}^* \tilde{C} &= U_2^* D^* U^* U C = U_2^* D^* C \\ &= -U_2^* B^* (\oplus \Delta_T) T^* = -\tilde{B}^* (\oplus \Delta_T) T^*. \end{aligned}$$

To verify condition (K3) note that $\tilde{D}$ acts as the column operator

$$\tilde{D} = (U_1, U D U_2) \colon \mathscr{E}_* \to \mathscr{E} = \hat{\mathscr{E}} \oplus (R \cap \mathscr{E}).$$

Thus we obtain that

$$\begin{aligned} \tilde{D}^* \tilde{D} &= U_1^* U_1 + U_2^* D^* U^* U D U_2 \\ &= U_1^* U_1 + U_2^* U_2 - U_2^* B^* (\oplus \Delta_T) B U_2 \\ &= 1_{\mathscr{E}_*} - \tilde{B}^* (\oplus \Delta_T) \tilde{B}. \end{aligned}$$

Since $j_{\tilde{C}} = U j_C$, it follows that

$$\left(\oplus j_{\tilde{C}}\right) \tilde{B} x = (\oplus U)(\oplus j_C) B(U_2 x) \in M_z^* H_K(\mathscr{E})$$

holds for all $x \in \mathscr{E}_*$. Thus the K-inner function $W \colon \mathbb{B}_d \to L(\mathscr{E}_*, \mathscr{E})$ admits a matrix representation of the claimed form. □

4 Minimal K-Dilations

Let $\mathcal{A}$ be a unital subalgebra of a unital C^*-Algebra $\mathcal{B}$. A completely positive unital map $\varphi \colon \mathcal{B} \to L(H)$ is called an $\mathcal{A}$-morphism if $\varphi(ax) = \varphi(a)\varphi(x)$ for $a \in \mathcal{A}$ and

$x \in \mathcal{B}$. Under the condition that $\mathcal{B}$ is the norm-closed linear span

$$\mathcal{B} = \overline{\text{span}}^{\|\cdot\|}\{\mathcal{A}\mathcal{A}^*\}$$

Arveson proved in [1, Lemmma 8.6] that every unitary operator that intertwines two $\mathcal{A}$-morphisms $\varphi_i : \mathcal{B} \to L(H_i)$ $(i = 1, 2)$ pointwise on $\mathcal{A}$ extends to a unitary operator that intertwines the minimal Stinespring representations of φ_1 and φ_2.

Straightforward modifications of the arguments given in [1] show that Arveson's result remains true if $\mathcal{B}$ is a von Neumann algebra which is the w^*-closed linear span

$$\mathcal{B} = \overline{\text{span}}^{w^*}\{\mathcal{A}\mathcal{A}^*\}$$

and if the $\mathcal{A}$-morphisms $\varphi_i : \mathcal{B} \to L(H_i)$ $(i = 1, 2)$ are supposed to be w^*-continuous.

Theorem 4.1 *Let $\mathcal{B}$ be a von Neumann algebra and let $\mathcal{A} \subset \mathcal{B}$ be a unital subalgebra such that*

$$\mathcal{B} = \overline{\text{span}}^{w^*}\{\mathcal{A}\mathcal{A}^*\}.$$

For $i = 1, 2$, let $\varphi_i : \mathcal{B} \to L(H_i)$ be a w^*-continuous $\mathcal{A}$-morphism and let (π_i, V_i, H_{π_i}) be the minimal Stinespring representations for φ_i. For every unitary operator $U : H_1 \to H_2$ with

$$U\varphi_1(a) = \varphi_2(a)U \quad (a \in \mathcal{A}),$$

there is a unique unitary operator $W : H_{\pi_1} \to H_{\pi_2}$ with $WV_1 = V_2U$ and $W\pi_1(x) = \pi_2(x)W$ for all $x \in \mathcal{B}$.

Since this version of Arveson's result follows in exactly the same way as the original one [1, Lemma 8.6], we leave the details to the reader. As an application of Theorem 4.1 we show that, under suitable conditions on the kernel $K : \mathbb{B}_d \times \mathbb{B}_d \to \mathbb{C}$, minimal K-dilations are uniquely determined. Recall that a commuting tuple $T \in L(H)^d$ on a Hilbert space H is called essentially normal if $T_iT_i^* - T_i^*T_i$ is compact for $i = 1, \ldots, d$. If $T \in L(H)^d$ is essentially normal, then by the Fuglede–Putnam theorem also all cross commutators $T_iT_j^* - T_j^*T_i$ $(i, j = 1, \ldots, d)$ are compact. For our multiplication tuple $M_z \in L(H_K)^d$, essential normality is equivalent to the condition that [11, Corollary 4.4]

$$\lim_{n\to\infty}\left(\frac{a_n}{a_{n+1}} - \frac{a_{n-1}}{a_n}\right) = 0.$$

Theorem 4.2 *Suppose that $M_z \in L(H_K)^d$ is an essentially normal K-contraction. Then the von Neumann algebra generated by $M_{z_1}, \ldots, M_{z_d}$ is given by*

$$W^*(M_z) = \overline{\text{span}}^{w^*}\{M_z^\alpha M_z^{*\beta};\ \alpha, \beta \in \mathbb{N}^d\}.$$

Proof Define $\mathscr{L} = \overline{\text{span}}^{w^*}\{M_z^\alpha M_z^{*\beta};\ \alpha, \beta \in \mathbb{N}^d\}$. Obviously $\mathscr{L} \subset W^*(M_z)$. Since M_z is supposed to be a K-contraction,

$$P_{\mathbb{C}} = \tau_{\text{SOT}} - \sum_{n=0}^{\infty} c_n \sigma_{M_z}^n(1_{H_K}) \in \mathscr{L}.$$

For $\alpha, \beta \in \mathbb{N}^d$ and $w \in \mathbb{B}_d$, we obtain

$$M_z^\alpha P_{\mathbb{C}} M_z^{*\beta}(K(\cdot, w)) = \overline{w}^\beta z^\alpha = z^\alpha \otimes z^\beta(K(\cdot, w)).$$

Since the multiplication on $L(H_K)$ is separately w^*-continuous, it follows that $\mathscr{L}$ contains all compact operators

$$K(H_K) = \overline{\text{span}}^{\|\cdot\|}\{z^\alpha \otimes z^\beta;\ \alpha, \beta \in \mathbb{N}^d\} \subset \mathscr{L}.$$

But then the hypothesis that M_z is essentially normal implies that $\mathscr{L} \subset L(H_K)$ is a subalgebra. Since the involution on $L(H_K)$ is w^*-continuous, the algebra $\mathscr{L} \subset L(H_K)$ is a von Neumann algebra and hence $\mathscr{L} = W^*(M_z)$. □

The tuple $M_z \in L(H_K)^d$ is known to be a K-contraction if there is a natural number $p \in \mathbb{N}$ such that $c_n \geq 0$ for all $n \geq p$ or $c_n \leq 0$ for all $n \geq p$ [7, Lemma 2.2] or [17, Proposition 2.10]. The latter condition holds, for instance, if H_K is a complete Nevanlinna–Pick space or if K is a kernel of the form

$$K_\nu : \mathbb{B}_d \times \mathbb{B}_d \to \mathbb{C}, \quad K_\nu(z, w) = \frac{1}{(1 - \langle z, w\rangle)^\nu}$$

with a positive real number $\nu > 0$ (see [8, Lemma 2.1] and [17, Section 1.5.2] for these results and further examples).

Let $T \in L(H)^d$ be a commuting tuple and let $j : H \to H_K(\mathscr{E})$ be a K-dilation of T. We denote by $\mathcal{B} = W^*(M_z) \subset L(H_K)$ the von Neumann algebra generated by M_z and set $\mathcal{A} = \{p(M_z);\ p \in \mathbb{C}[z]\}$. The unital C^*-homomorphism

$$\pi : \mathcal{B} \to L(H_K(\mathscr{E})), \quad X \mapsto X \otimes 1_{\mathscr{E}}$$

together with the isometry $j : H \to H_K(\mathscr{E})$ is a Stinespring representation for the completely positive map

$$\varphi : \mathcal{B} \to L(H_K(\mathscr{E})), \quad \varphi(X) = j^*(X \otimes 1_{\mathscr{E}})j.$$

The map φ is an $\mathcal{A}$-morphism, since

$$\begin{aligned}\varphi(p(M_z)X) &= j^*(p(M_z)\otimes 1_{\mathscr{E}})X\otimes 1_{\mathscr{E}})j = j^*p(M_z\otimes 1_{\mathscr{E}})(jj^*)(X\otimes 1_{\mathscr{E}})j\\ &= \varphi(p(M_z))\varphi(X)\end{aligned}$$

for all $p\in\mathbb{C}[z]$ and $X\in\mathcal{B}$. Standard duality theory for Banach space operators shows that π is w^*-continuous. Indeed, as an application of Krein–Smulian's theorem (Theorem IV. 6.4 in [16]) one only has to check that $\tau_{w^*}-\lim_\alpha(X_\alpha\otimes 1_{\mathscr{E}}) = X\otimes 1_{\mathscr{E}}$ for each norm-bounded net (X_α) in $\mathcal{B}$ with $\tau_{w^*}-\lim_\alpha X_\alpha = X$. To complete the argument it suffices to recall that on norm-bounded sets the w^*-topology and the weak operator topology coincide. Thus we have shown that φ is a w^*-continuous $\mathcal{A}$-morphism with Stinespring representation π. By definition the K-dilation $j\colon H\to H_K(\mathscr{E})$ is minimal if and only if

$$\bigvee_{X\in W^*(M_z)}\pi(X)(jH) = H_K(\mathscr{E}),$$

hence if and only if π as a Stinespring representation of φ is minimal.

Corollary 4.3 *Suppose that $M_z\in L(H_K)^d$ is an essentially normal K-contraction. If $j_i\colon H\to H_K(\mathscr{E}_i)$ $(i=1,2)$ are two minimal K-dilations of a commuting tuple $T\in L(H)^d$, then there is a unitary operator $U\in L(\mathscr{E}_1,\mathscr{E}_2)$ with $j_2=(1_{H_K}\otimes U)j_1$*

Proof As before we denote by $\mathcal{B}=W^*(M_z)\subset L(H_K)$ the von Neumann algebra generated by $M_{z_1},\dots,M_{z_d}\in L(H_K)$ and define $\mathcal{A}=\{p(M_z);\, p\in\mathbb{C}[z]\}$. The remarks preceding the corollary show that the maps

$$\varphi_i\colon\mathcal{B}\to L(H),\quad \varphi_i(X)=j_i^*(X\otimes 1_{\mathscr{E}_i})j_i\quad (i=1,2)$$

are w^*-continuous $\mathcal{A}$-morphisms with minimal Stinespring representations

$$\pi_i\colon\mathcal{B}\to L(H_K(\mathscr{E}_i)),\quad \pi_i(X)=X\otimes 1_{\mathscr{E}_i}\quad (i=1,2).$$

Since

$$\varphi_i(p(M_z)) = j^*p(M_z\otimes 1_{\mathscr{E}})j = p(T)$$

for all $p\in\mathbb{C}[z]$ and $i=1,2$, Theorem 4.1 implies that there is a unitary operator $W\colon H_K(\mathscr{E}_1)\to H_K(\mathscr{E}_2)$ with $Wj_1=j_2$ and $W(X\otimes 1_{\mathscr{E}_1})=(X\otimes 1_{\mathscr{E}_2})W$ for all $X\in\mathcal{B}$. In particular, the unitary operator W satisfies the intertwining relations

$$W(M_{z_i}\otimes 1_{\mathscr{E}_1})=(M_{z_i}\otimes 1_{\mathscr{E}_2})W\quad (i=1,\dots,d)$$

A standard characterization of multipliers on reproducing kernel Hilbert spaces [4, Theorem 2.1] shows that there exist operator-valued functions $A\colon \mathbb{B}_d \to L(\mathscr{E}_1, \mathscr{E}_2)$ and $B\colon \mathbb{B}_d \to L(\mathscr{E}_2, \mathscr{E}_1)$ such that $Wf = Af$ and $W^*g = Bg$ for $f \in H_K(\mathscr{E}_1)$ and $g \in H_K(\mathscr{E}_2)$ (see also [17, Proposition 4.5]). It follows that $A(z)B(z) = 1_{\mathscr{E}_2}$ and $B(z)A(z) = 1_{\mathscr{E}_1}$ for $z \in \mathbb{B}_d$. Since

$$K(z, w)x = (WW^*K(\cdot, w)x)(z) = A(z)K(z, w)A(w)^*x$$

for $z, w \in \mathbb{B}_d$ and $x \in \mathscr{E}_2$, we find that $A(z)A(w)^* = 1_{\mathscr{E}_2}$ for $z, w \in \mathbb{B}_d$. But then the constant value $A(z) \equiv U \in L(\mathscr{E}_1, \mathscr{E}_2)$ is a unitary operator with $W = 1_{H_K} \otimes U$. □

We conclude this section by showing that the canonical K-dilation of a K-contraction $T \in L(H)^d$ defined in Theorem 2.1 is minimal. To prepare this result we first identify the M_z-reducing subspaces of $H_K(\mathscr{E})$.

Lemma 4.4 *Let $M \subset H_K(\mathscr{E})$ be a closed linear subspace. If M is reducing for $M_z \in L(H_K(\mathscr{E}))^d$, then $P_{\mathscr{E}}M \subset M$ and*

$$M = \bigvee_{\alpha\in\mathbb{N}^d} z^\alpha(M \cap \mathscr{E}) = H_K(M \cap \mathscr{E}).$$

Proof The hypothesis implies that M is reducing for the von Neumann algebra $W^*(M_z) \subset L(H_K(\mathscr{E}))$ generated by $M_{z_1}, \ldots M_{z_d} \in L(H_K(\mathscr{E}))$. Standard results on von Neumann algebras (Corollary 17.6 and Proposition 24.1 in [18]) show that

$$P_{\mathscr{E}} = P_{\bigcap \operatorname{Ker} M_{z_i}^*} \in W^*(M_z).$$

Hence $P_{\mathscr{E}}M \subset M$. Let $f = \sum_{\alpha\in\mathbb{N}^d} f_\alpha z^\alpha \in H_K(\mathscr{E})$ be arbitrary. An elementary calculation yields that

$$P_{\mathscr{E}}(M_z^{*\beta} f) \in (\mathbb{C} \setminus \{0\}) f_\beta \quad (\beta \in \mathbb{N}^d).$$

Hence, if $f \in M$, then $f_\beta \in M \cap \mathscr{E}$ for all $\beta \in \mathbb{N}^d$ and the observation that

$$f = \sum_{\alpha\in\mathbb{N}^d} f_\alpha z^\alpha \in \bigvee_{\alpha\in\mathbb{N}^d} z^\alpha(M \cap \mathscr{E}) = H_K(M \cap \mathscr{E})$$

completes the proof. □

Corollary 4.5 *Let $T \in L(H)^d$ be a pure K-contraction. Then the K-dilation*

$$j\colon H \to H_K(\mathscr{D}), \quad j(x) = \sum_{\alpha\in\mathbb{N}^d} a_{|\alpha|}\gamma_\alpha(CT^{*\alpha}x)z^\alpha$$

defined in Theorem 2.1 is minimal.

Proof Let $\operatorname{Im} j \subset M$ be a reducing subspace for $M_z \in L(H_K(\mathscr{D}))^d$. We know from Lemma 4.4 that

$$M = \bigvee_{\alpha \in \mathbb{N}^d} z^\alpha (M \cap \mathscr{D})$$

and that

$$CH = P_{\mathscr{D}}(\operatorname{Im} j) \subset P_{\mathscr{D}}(M) \subset M \cap \mathscr{D}.$$

It follows that $\mathscr{D} = \overline{CH} = M \cap \mathscr{D}$ and that $M = \bigvee_{\alpha \in \mathbb{N}^d} z^\alpha \mathscr{D} = H_K(\mathscr{D})$. □

It should be interesting to compare the uniqueness result proved in this section with a related result proved by Olofsson [15, Theorem 7.6] for single contractions $T \in L(H)$ satisfying a slightly different K-contractivity condition. In [15] it is shown that even in the non-pure case each dilation factors through a canonical defined dilation of T.

References

1. W. Arveson, Subalgebras of C^*-algebras. III: multivariable operator theory. Acta Math. **181**, 159–228 (1998)
2. W. Arveson, Quotients of standard Hilbert modules. Trans. Am. Math. Soc. **359**, 6027–6055 (2007)
3. M. Bhattacharjee, J. Eschmeier, D.K. Keshari, J. Sarkar, Dilations, wandering subspaces and inner functions. Linear Algebra Appl. **523**, 263–280 (2017)
4. C. Barbian, A characterization of multiplication operators on reproducing kernel Hilbert spaces. J. Oper. Theory **65**, 235–240 (2011)
5. J.A. Ball, V. Bolotnikov, Weighted Bergman spaces: shift-invariant subspaces and input/state/output linear systems. Integr. Equ. Oper. Theor. **76**, 301–356 (2013)
6. J.A. Ball, V. Bolotnikov, Weighted Hardy spaces: shift-invariant subspaces and coinvariant subspaces, linear systems and operator model theory. Acta Sci. Math. (Szeged) **79**, 623–686 (2013)
7. Y. Chen, Quasi-wandering subspaces in a class of reproducing analytic Hilbert spaces. Proc. Am. Math. Soc. **140**, 4235–4242 (2012)
8. R. Clouâtre, M. Hartz, Multiplier algebras of complete Nevanlinna–Pick spaces: dilations, boundary representations and hyperrigidity. J. Funct. Anal. **274**, 1690–1738 (2018)
9. J. Eschmeier, M. Putinar, *Spectral Decompositions and Analytic Sheaves*. London Mathematical Society Monographs, New Series, vol. 10 (Clarendon Press, Oxford, 1996)
10. J. Eschmeier, Bergman inner functions and m-hypercontractions. J. Funct. Anal. **275**, 73–102 (2018)
11. K. Guo, J. Hu, X. Xu, Toeplitz algebras, subnormal tuples and rigidity on reproducing $\mathbb{C}[z_1, \ldots, z_d]$-modules. J. Funct. Anal. **210**, 214–247 (2004)
12. V. Müller, F.-H. Vasilescu, Standard models for some commuting multioperators. Proc. Am. Math. Soc. **117**, 979–989 (1993)
13. A. Olofsson, A characteristic operator function for the class of n-hypercontractions. J. Funct. Anal. **236**, 517–545 (2006)

14. A. Olofsson, Operator-valued Bergman inner functions as transfer functions. Algebra Anal. **19**, 146–173 (2007); St. Petersberg Math. J. **19**, 603–623 (2008)
15. A. Olofsson, Parts of adjoint weighted shifts. J. Oper. Theory **74**, 249–280 (2015)
16. H.H. Schaefer, *Topological Vector Spaces* (Macmillan, New York, 1966)
17. D. Schillo, K-contractions and perturbations of Toeplitz operators. Ph.D. thesis, Saarland University, 2018
18. K. Zhu, *An Introduction to Operator Algebras*. Studies in Advanced Mathematics (CRC Press, Boca Raton, 1993)

Tight and Cover-to-Join Representations of Semilattices and Inverse Semigroups

Ruy Exel

Abstract We discuss the relationship between tight and cover-to-join representations of semilattices and inverse semigroups, showing that a slight extension of the former, together with an appropriate selection of codomains, makes the two notions equivalent. As a consequence, when constructing universal objects based on them, one is allowed to substitute cover-to-join for tight and vice-versa.

Keywords Semilattice · Inverse semigroup · Tight representation · Cover-to-join · Boolean algebra · Non-degenerate representation · Universal C*-algebra

Mathematics Subject Classification (2010) Primary 20M18, 20M30; Secondary 46L05

1 Introduction

Exactly 12 years ago, to be precise on March 7, 2007, I posted a paper on the arXiv [3] describing the notion of *tight representations* of semilattices and inverse semigroups, which turned out to have many applications and in particular proved to be useful to give a unified perspective to a significant number of C*-algebras containing a preferred generating set of partial isometries [1, 2, 4, 6–8, 12, 13].

The notion of tight representations (described below for the convenience of the reader) is slightly involving as it depends on the analysis of certain pairs of finite sets X and Y, but it becomes much simplified when X is a singleton and Y is empty (see [4, Proposition 11.8]). In this simplified form it has been rediscovered and used in many subsequent works (e.g. [2, 9–11]) under the name of *cover-to-join* representations.

R. Exel (✉)
Departamento de Matemática, Universidade Federal de Santa Catarina, Florianópolis, SC, Brazil

M. A. Bastos et al. (eds.), *Operator Theory, Functional Analysis and Applications*,
Operator Theory: Advances and Applications 282,
https://doi.org/10.1007/978-3-030-51945-2_9

The notion of cover-to-join representations, requiring a smaller set of conditions, is consequently weaker and, as it turns out, strictly weaker, than the original notion of tightness. Nevertheless, besides being easier to formulate, the notion of cover-to-join representations has the advantage of being applicable to representations taking values in *generalized* Boolean algebras, that is, Boolean algebras without a unit. Explicitly mentioning the operation of complementation, tight representations only make sense for unital Boolean algebras.

The goal of this note is to describe an attempt to reconcile the notions of tight and cover-to-join representations: slightly extending the former, and adjusting for the appropriate codomains, we show that, after all, the two notions coincide.

One of the main practical consequences of this fact is that the difference between the two notions becomes irrelevant for the purpose of constructing universal objects based on them, such as the completion of an inverse semigroup recently introduced in [11]. We are moreover able to fix a slight imprecision in the proof of [2, Theorem 2.2], at least as far as its consequence that the universal C*-algebras for tight vs. cover-to-join representations are isomorphic.

2 Generalized Boolean Algebras

We begin by recalling the well known notion of generalized Boolean algebras.

Definition 2.1 ([14, Definition 5]) A *generalized Boolean algebra* is a set B equipped with binary operations $\wedge$ and $\vee$, and containing an element 0, such that for every a, b and c in B, one has that

(i) (commutativity) $a \vee b = b \vee a$, and $a \wedge b = b \wedge a$,
(ii) (associativity) $(a \wedge b) \wedge c = a \wedge (b \wedge c)$,
(iii) (distributivity) $a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$,
(iv) $a \vee 0 = a$,
(v) (relative complement) if $a = a \wedge b$, there is an element x in B, such that $x \vee a = b$, and $x \wedge a = 0$,
(vi) $a \vee a = a = a \wedge a$.

It follows that (ii) and (iii) also hold with $\vee$ and $\wedge$ interchanged, meaning that $\vee$ is associative [14, Theorems 55 & 14], and that $\vee$ distributes over $\wedge$ [14, Theorems 55 & 11].

When $a = a \wedge b$, as in (v), one writes $a \leq b$. It is then easy to see that $\leq$ is a partial order on B.

The element x referred to in (v) is called the *relative complement* of a in b, and it is usually denoted $b \setminus a$.

Definition 2.2 (cf. [14, Theorem 56]) A generalized Boolean algebra B is called a *Boolean algebra* if there exists an element 1 in B, such that $a \wedge 1 = a$, for every a in B.

For Boolean algebras, the complement of an element a relative to 1 is often denoted $\neg a$.

Recall that an *ideal* of a generalized Boolean algebra B is any nonempty subset C of B which is closed under $\vee$, and such that

$$a \leq b \in C \Rightarrow a \in C.$$

Such an ideal is evidently also closed under $\wedge$ and under relative complements, so it is a generalized Boolean algebra in itself.

Given any nonempty subset S of B, notice that the subset C defined by

$$C = \left\{a \in B : a \leq \bigvee_{z \in Z} z, \text{ for some nonempty finite subset } Z \subseteq S\right\},$$

is an ideal of B and it is clearly the smallest ideal containing S, so we shall call it the *ideal generated by* S, and we shall denote it by $\langle S \rangle$.

3 Tight and Cover-to-Join Representations of Semilattices

From now on let us fix a meet semilattice E (always assumed to have a zero element).

Definition 3.1 A *representation* of E in a generalized Boolean algebra B is any map $\pi : E \to B$, such that

(i) $\pi(0) = 0$, and
(ii) $\pi(x \wedge y) = \pi(x) \wedge \pi(y)$, for every x and y in E.

In order to spell out the definition of the notion of *tight representations*, introduced in [4], let F be any subset of E. We then say that a given subset $Z \subseteq F$ is a *cover* for F, if for every nonzero x in F, there exists some z in Z, such that $z \wedge x \neq 0$.

Furthermore, if X and Y are finite subsets of E, we let

$$E^{X,Y} = \{z \in E : z \leq x, \ \forall x \in X, \text{ and } z \perp y, \ \forall y \in Y\}.$$

Definition 3.2 (cf. [4, Definition 11.6]) A representation π of E in a Boolean algebra B is said to be *tight* if, for any finite subsets X and Y of E, and for any finite cover Z for $E^{X,Y}$, one has that

$$\bigvee_{z \in Z} \pi(z) = \bigwedge_{x \in X} \pi(x) \wedge \bigwedge_{y \in Y} \neg \pi(y). \tag{3.1}$$

Observe that if Y is empty and X is a singleton, say $X = \{x\}$, then

$$E^{X,Y} = E^{\{x\},\emptyset} = \{z \in E : z \leq x\},$$

and if Z is a cover for this set, then (3.1) reads

$$\bigvee_{z\in Z} \pi(z) = \pi(x). \tag{3.2}$$

To check that a given representation is tight, it is not enough to verify (3.2), as it is readily seen by considering the example in which $E = \{0, 1\}$ and B is any Boolean algebra containing an element $x \neq 1$. Indeed, the map $\pi : E \to B$ given by $\pi(0) = 0$, and $\pi(1) = x$, satisfies all instances of (3.2) even thought it is not tight. The reader might wonder if the fact that π fails to preserve the unit is playing a part in this counter-example, but it is also easy to find examples of cover-to-join representations of non-unital semilattices which are not tight.

Representations π satisfying (3.2) whenever Z is a cover for $E^{\{x\},\emptyset}$ have been considered in [4, Proposition 11.8], and they have been called *cover-to-join* representations in [2].

It is a trivial matter to prove that a cover-to-join representation satisfies (3.1) whenever X is nonempty (see the proof of [4, Lemma 11.7]), so the question of whether a cover-to-join representation is indeed tight rests on verifying (3.1) when X is empty. In this case, and assuming that Z is a cover for $E^{\emptyset,Y}$, it is easy to see that $Z \cup Y$ is a cover for the whole of E. Should we be dealing with a semilattice not admitting any finite cover, this situation will therefore never occur, that is, one will never be required to check (3.1) for an empty set X, hence every cover-to-join representation is automatically tight.

This has in fact already been observed in [4, Proposition 11.8], which says that every cover-to-join representation is tight in case E does not admit any finite cover, as we have just discussed, but also if E contains a finite set X such that

$$\bigvee_{x\in X} \pi(x) = 1. \tag{3.3}$$

The latter condition is useful for dealing with characters, i.e. with representations of E in the Boolean algebra $\{0, 1\}$, because the requirement that a character be nonzero immediately implies (3.3), so again cover-to-join suffices to prove tightness.

On the other hand, an advantage of the notion of cover-to-join representations is that it makes sense for representations in generalized Boolean algebras, while the reference to the unary operation $\neg$ in (3.1) precludes it from being applied when the target algebra lacks a unit, that is, for a representation into a generalized Boolean algebra.

Again referring to the occurrence of $\neg$ in (3.1), observe that if X is nonempty, then the right hand side of (3.1) lies in the ideal of B generated by the range of

π. This is because, even though $\neg\,\pi(y)$ is not necessarily in $\langle \pi(E)\rangle$, this term will appear besides $\pi(x)$, for some x in X, and hence

$$\pi(x) \wedge \neg\pi(y) = \pi(x) \setminus \big(\pi(x) \wedge \pi(y)\big) \in \langle \pi(E)\rangle.$$

This means that:

Proposition 3.3 *If E is a semilattice not admitting any finite cover then, whenever X and Y are finite subsets of E, and Z is a finite cover of $E^{X,Y}$, the right hand side of (3.1) lies in $\langle \pi(E)\rangle$.*

As a consequence we see that Definition 3.2 may be safely applied to a representation of E in a generalized Boolean algebra, as long as E does not admit a finite cover: despite the occurrence of $\neg$ in (3.1), once its right hand side is expanded, it may always be expressed in terms of relative complements, hence avoiding the use of the missing unary operation $\neg$.

We may therefore consider the following slight generalization of the notion of tight representations:

Definition 3.4 A representation π of E in a generalized Boolean algebra B is said to be *tight* if, either B is a Boolean algebra and π is tight in the sense of Definition 3.2, or the following two conditions are verified:

(i) E admits no finite cover, and
(ii) (3.1) holds for any finite subsets X and Y of E, and for any finite cover Z for $E^{X,Y}$.

As already stressed, despite the occurrence of $\neg$ in (3.1), condition (ii) in Definition 3.4 will always make sense in a generalized Boolean algebra.

So here is a result that perhaps may be used to reconcile the notions of tightness and cover-to-join representations:

Theorem 3.5 *Let π be a representation of the semilattice E in the generalized Boolean algebra B. Then*

(i) *if π is tight then it is also cover-to-join,*
(ii) *if π is cover-to-join then there exists an ideal B' of B, containing the range of π, such that, once π is seen as a representation of E in B', one has that π is tight.*

Proof Point (i) being immediate, let us prove (ii). Under the assumption that E does not admit any finite cover, we have that π is tight as a representation into $B' = B$, by Exel [4, Proposition 11.8], or rather by its obvious adaptation to generalized Boolean algebras.

It therefore remains to prove (ii) in case E does admit a finite cover, say Z. Setting

$$e = \bigvee_{z\in Z} \pi(z), \tag{3.4}$$

we claim that

$$\pi(x) \leq e, \quad \forall x \in E. \tag{3.5}$$

To see this, pick x in E and notice that, since Z is a cover for E, we have in particular that the set

$$\{z \wedge x : z \in Z\}$$

is a cover for x, so the cover-to-join property of π implies that

$$\pi(x) = \bigvee_{z\in Z} \pi(z \wedge x) \leq \bigvee_{z\in Z} \pi(z) = e,$$

proving (3.5). We therefore let

$$B' = \{a \in B : a \leq e\},$$

which is evidently an ideal of B containing the range of π by (3.5).

By (3.4) we then have that π satisfies [4, Lemma 11.7.(i)], as long as we see π as a representation of E in B', whose unit is clearly e. The result then follows from [4, Proposition 11.8]. □

4 Non-Degenerate Representations of Semilattices

The following is perhaps the most obvious adaptation of the notion of non-degenerate representations extensively used in the theory of operator algebras [15, Definition 9.3].

Definition 4.1 We shall say that a representation π of a semilattice E in a generalized Boolean algebra B is *non-degenerate* if, for every a in B, there is a finite subset Z of E such that $a \leq \bigvee_{z\in Z} \pi(z)$. In other words, π is non-degenerate if and only if B coincides with the ideal generated by the range of π.

Observe that, if both E and B have a unit, and if π is a unital map, then π is evidently non-degenerate. More generally, if π satisfies (3.3), then the same is also clearly true.

The following result says that, by adjusting the codomain of a representation, we can always make it non-degenerate.

Proposition 4.2 *Let π be a representation of E in the generalized Boolean algebra B. Letting C be the ideal of B generated by the range of π, one has that π is a non-degenerate representation of E in C.*

Proof Obvious. □

For non-degenerate representations we have the following streamlined version of Theorem 3.5:

Corollary 4.3 *Let π be a non-degenerate representation of the semilattice E in the generalized Boolean algebra B. Then π is tight if and only if it is cover-to-join.*

Proof The "only if" direction being trivial, we concentrate on the "if" part, so let us assume that π is cover-to-join. By Theorem 3.5 there exists an ideal B' of B, containing the range of π, and such that π is tight as a representation in B'. Such an ideal will therefore contain the ideal generated by $\pi(E)$, which coincides with B by hypothesis. Therefore $B' = B$, and hence π is tight as a representation into its default codomain B. □

5 Representations of Inverse Semigroups

By its very nature, the concept of a tight representation pertains to the realm of semilattices and Boolean algebras. However, given the relevance of the study of semilattices in the theory of inverse semigroups, tight representations have had a strong impact on the latter.

Recall that a *Boolean inverse semigroup* (see [5] but please observe that this notion is not equivalent to the homonym studied in [9] and [16]) is an inverse semigroup whose idempotent semilattice $E(S)$ is a Boolean algebra. In accordance with what we have been discussing up to now, it is sensible to give the following:

Definition 5.1

(i) A *generalized Boolean inverse semigroup* is an inverse semigroup whose idempotent semilattice is a generalized Boolean algebra.

(ii) (cf. [4, Definition 13.1] and [5, Proposition 6.2]) If S is any inverse semigroup[1] and T is a generalized Boolean inverse semigroup, we say that a homomorphism $\pi : S \to T$ (always assumed to preserve zero) is *tight* if the restriction of π to $E(S)$ is a tight representation into $E(T)$, in the sense of Definition 3.4.

(iii) If π is as above, we say that π is *cover-to-join* if the restriction of π to $E(S)$ is cover-to-join.

Addressing the already mentioned slight imprecision in the proof of [2, Theorem 2.2], we then have the following version of Theorem 3.5 and Proposition 4.2:

[1] All inverse semigroups in this note are required to have a zero.

Corollary 5.2 *Let π be a representation of the inverse semigroup S in the generalized Boolean inverse semigroup T. Then*

(i) *if π is tight then it is also cover-to-join;*
(ii) *if π is cover-to-join then there exists a generalized Boolean inverse sub-semigroup T' of T, containing the range of π, such that, once π is seen as a representation of S in T', one has that π is tight;*
(iii) *if π is cover-to-join, and if the restriction of π to $E(S)$ is non-degenerate, then π is tight.*

Proof The proof is essentially contained in the proofs of Theorem 3.5 and Proposition 4.2, except maybe for the proof of (ii) under the assumption that $E(S)$ admits a finite cover, say Z. In this case, let e be as in (3.4) and put

$$T' = \{t \in T : t^*t \leq e,\ tt^* \leq e\},$$

observing that T' is clearly an inverse sub-semigroup of T, and that its idempotent semilattice is a Boolean algebra. Given any s in S, observe that s^*s lies in $E(S)$ and

$$\pi(s)^*\pi(s) = \pi(s^*s) \leq e,$$

where the last inequality above follows as in (3.5). By a similar reasoning one shows that also $\pi(s)\pi(s)^* \leq e$, so we see that $\pi(s)$ lies in T', and we may then think of π as a representation of S in T'. As in Theorem 3.5, one may now easily prove that π becomes a tight representation into T'. □

6 Conclusion

As a consequence of the above results, when defining universal objects (such as semigroups, algebras or C*-algebras) for a class of representations of inverse semigroups, one may safely substitute cover-to-join for tight and vice-versa. Given the widespread use of tight representations, there are many instances where the above principle applies. Below we spell out one such result to concretely illustrate our point, but similar results may be obtained as trivial reformulations of the following:

Theorem 6.1 *Let S be an inverse semigroup and let $C^*_{\text{tight}}(S)$ be the universal C*-algebra* [4, Theorem 13.3] *for tight Hilbert space representations of S* [4, Definition 13.1]. *Also let $C^*_{\substack{\text{cover-}\\\text{to-join}}}(S)$ be the universal C*-algebra for cover-to-join Hilbert space representations of S. Then*

$$C^*_{\text{tight}}(S) \simeq C^*_{\substack{\text{cover-}\\\text{to-join}}}(S).$$

Proof It suffices to prove that $C^*_{\text{tight}}(S)$ also has the universal property for cover-to-join representations. So let

$$\pi : S \to B(H)$$

be a cover-to-join representation of S on some Hilbert space H. Should the idempotent semilattice of S admit no finite covers, one has that π is tight so there is nothing to do. On the other hand, assuming that Z is a finite cover for $E(S)$, let e be as in (3.4).

Writing H_e for the range of e and letting $K = H_e^{\perp}$, we then obviously have that $H = H_e \oplus K$. It then follows from (3.5) that each $\pi(s)$ decomposes as a direct sum of operators

$$\pi(s) = \pi'(s) \oplus 0,$$

thus defining a representation π' of S on H_e which is clearly also cover-to-join. It is also clear that π' is non-degenerate on $E(S)$, so we have by Corollary 5.2(iii) that π' is tight. Therefore the universal property provides a *-representation φ' of $C^*_{\text{tight}}(S)$ on $B(H_e)$ coinciding with π' on the canonical image of S within $C^*_{\text{tight}}(S)$. It then follows that $\varphi := \varphi' \oplus 0$ coincides with π on S, concluding the proof. □

We therefore believe this clarifies [2, Corollaries 2.3 & 2.5].

References

1. G. Boava, G.G. de Castro, F.de L. Mortari, Inverse semigroups associated with labelled spaces and their tight spectra. Semigroup Forum **94**, 582–609 (2017)
2. A.P. Donsig, D. Milan, Joins and covers in inverse semigroups and tight C*-algebras. Bull. Aust. Math. Soc. **90**, 121–133 (2014)
3. R. Exel, *Inverse Semigroups and Combinatorial C*-algebras* (2007). arXiv:math/ 0703182v1 [math.OA]
4. R. Exel, Inverse semigroups and combinatorial C*-algebras. Bull. Braz. Math. Soc. (N.S.) **39**, 191–313 (2008)
5. R. Exel, Tight representations of semilattices and inverse semigroups. Semigroup Forum **79**, 159–182 (2009)
6. R. Exel, E. Pardo, Self-similar graphs, a unified treatment of Katsura and Nekrashevych C*-algebras. Adv. Math. **306**, 1046–1129 (2017)
7. R. Exel, C. Starling, Self-similar graph C^*-algebras and partial crossed products. J. Oper. Theory **75**, 299–317 (2016)
8. R. Exel, D. Goncalves, C. Starling, The tiling C*-algebra viewed as a tight inverse semigroup algebra. Semigroup Forum **84**, 229–240 (2012)
9. M.V. Lawson, Non-commutative Stone duality: inverse semigroups, topological groupoids and C*-algebras. Internat. J. Algebra Comput. **22**(6), 1250058, 47 (2012)
10. M.V. Lawson, D.G. Jones, Graph inverse semigroups: their characterization and completion. J. Algebra **409**, 444–473 (2014)

11. M.V. Lawson, A. Vdovina, The universal Boolean inverse semigroup presented by the abstract Cuntz-Krieger relations. J. Noncommut. Geom. (2019, to appear). arXiv:1902.02583v3 [math.OA], February, 28, 2019.
12. C. Starling, Boundary quotients of C^*-algebras of right LCM semigroups. J. Funct. Anal. **268**, 3326–3356 (2015)
13. C. Starling, Inverse semigroups associated to subshifts. J. Algebra **463**, 211–233 (2016)
14. M.H. Stone, Postulates for Boolean algebras and generalized Boolean algebras. Am. J. Math. **57**, 703–732 (1935)
15. M. Takesaki, *Theory of Operator Algebras. I* (Springer, Heidelberg, 1979)
16. F. Wehrung, in *Refinement Monoids, Equidecomposability Types, and Boolean Inverse Semigroups*. Lecture Notes in Mathematics, vol. 2188 (Springer, Berlin, 2017)

Calkin Images of Fourier Convolution Operators with Slowly Oscillating Symbols

C. A. Fernandes, A. Yu. Karlovich, and Yu. I. Karlovich

Abstract Let Φ be a C^*-subalgebra of $L^\infty(\mathbb{R})$ and $SO^\diamond_{X(\mathbb{R})}$ be the Banach algebra of slowly oscillating Fourier multipliers on a Banach function space $X(\mathbb{R})$. We show that the intersection of the Calkin image of the algebra generated by the operators of multiplication aI by functions $a \in \Phi$ and the Calkin image of the algebra generated by the Fourier convolution operators $W^0(b)$ with symbols in $SO^\diamond_{X(\mathbb{R})}$ coincides with the Calkin image of the algebra generated by the operators of multiplication by constants.

Keywords Fourier convolution operator · Fourier multiplier · Multiplication operator · Slowly oscillating function · Calkin algebra · Calkin image

Mathematics Subject Classification (2010) Primary 47G10, Secondary 42A45, 46E30

This work was partially supported by the Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) through the project UID/MAT/00297/2019 (Centro de Matemática e Aplicações). The third author was also supported by the SEP-CONACYT Project A1-S-8793 (México).

C. A. Fernandes (✉) · A. Yu. Karlovich
Centro de Matemática e Aplicações, Departamento de Matemática, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Quinta da Torre, Portugal
e-mail: caf@fct.unl.pt; oyk@fct.unl.pt

Yu. I. Karlovich
Centro de Investigación en Ciencias, Instituto de Investigación en Ciencias Básicas y Aplicadas, Universidad Autónoma del Estado de Morelos, Cuernavaca, México
e-mail: karlovich@uaem.mx

M. A. Bastos et al. (eds.), *Operator Theory, Functional Analysis and Applications*,
Operator Theory: Advances and Applications 282,
https://doi.org/10.1007/978-3-030-51945-2_10

1 Introduction

Let $\mathcal{F} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ denote the Fourier transform

$$(\mathcal{F}f)(x) := \widehat{f}(x) := \int_{\mathbb{R}} f(t)e^{itx}\,dt, \quad x \in \mathbb{R},$$

and let $\mathcal{F}^{-1} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ be the inverse of $\mathcal{F}$,

$$(\mathcal{F}^{-1}g)(t) = \frac{1}{2\pi}\int_{\mathbb{R}} g(x)e^{-itx}\,dx, \quad t \in \mathbb{R}.$$

It is well known that the Fourier convolution operator

$$W^0(a) := \mathcal{F}^{-1}a\mathcal{F} \tag{1.1}$$

is bounded on the space $L^2(\mathbb{R})$ for every $a \in L^\infty(\mathbb{R})$.

Let $X(\mathbb{R})$ be a Banach function space and $X'(\mathbb{R})$ be its associate space. Their technical definitions are postponed to Sect. 2.1. The class of Banach function spaces is very large. It includes Lebesgue, Orlicz, Lorentz spaces, variable Lebesgue spaces and their weighted analogues (see, e.g., [4, 6]). Let $\mathcal{B}(X(\mathbb{R}))$ denote the Banach algebra of all bounded linear operators acting on $X(\mathbb{R})$, let $\mathcal{K}(X(\mathbb{R}))$ be the closed two-sided ideal of all compact operators in $\mathcal{B}(X(\mathbb{R}))$, and let $\mathcal{B}^\pi(X(\mathbb{R})) = \mathcal{B}(X(\mathbb{R}))/\mathcal{K}(X(\mathbb{R}))$ be the Calkin algebra of the cosets $A^\pi := A+\mathcal{K}(X(\mathbb{R}))$, where $A \in \mathcal{B}(X(\mathbb{R}))$.

If $X(\mathbb{R})$ is separable, then $L^2(\mathbb{R}) \cap X(\mathbb{R})$ is dense in $X(\mathbb{R})$ (see Lemma 2.1 below). A function $a \in L^\infty(\mathbb{R})$ is called a Fourier multiplier on $X(\mathbb{R})$ if the convolution operator $W^0(a)$ defined by (1.1) maps $L^2(\mathbb{R}) \cap X(\mathbb{R})$ into $X(\mathbb{R})$ and extends to a bounded linear operator on $X(\mathbb{R})$. The function a is called the symbol of the Fourier convolution operator $W^0(a)$. The set $\mathcal{M}_{X(\mathbb{R})}$ of all Fourier multipliers on $X(\mathbb{R})$ is a unital normed algebra under pointwise operations and the norm

$$\|a\|_{\mathcal{M}_{X(\mathbb{R})}} := \left\|W^0(a)\right\|_{\mathcal{B}(X(\mathbb{R}))}.$$

For a unital C^*-subalgebra Φ of the algebra $L^\infty(\mathbb{R})$, we consider the quotient algebra $\mathcal{MO}^\pi(\Phi)$ consisting of the cosets

$$[aI]^\pi := aI + \mathcal{K}(X(\mathbb{R}))$$

of multiplication operators by functions in Φ:

$$\mathcal{MO}^\pi(\Phi) := \{[aI]^\pi : a \in \Phi\} = \{aI + \mathcal{K}(X(\mathbb{R})) : a \in \Phi\}.$$

For a unital Banach subalgebra Ψ of the algebra $\mathcal{M}_{X(\mathbb{R})}$, we also consider the quotient algebra $\mathcal{CO}^\pi(\Psi)$ consisting of the cosets

$$[W^0(b)]^\pi := W^0(b) + \mathcal{K}(X(\mathbb{R}))$$

of convolution operators with symbols in the algebra Ψ:

$$\mathcal{CO}^\pi(\Psi) := \{[W^0(b)]^\pi : b \in \Psi\} = \{W^0(b) + \mathcal{K}(X(\mathbb{R})) : b \in \Psi\}.$$

It is easy to see that $\mathcal{MO}^\pi(\Phi)$ and $\mathcal{CO}^\pi(\Psi)$ are commutative unital Banach subalgebras of the Calkin algebra $\mathcal{B}^\pi(X(\mathbb{R}))$. It is natural to refer to the algebras $\mathcal{MO}^\pi(\Phi)$ and $\mathcal{CO}^\pi(\Psi)$ as the Calkin images of the algebras

$$\mathcal{MO}(\Phi) = \{aI : a \in \Phi\} \subset \mathcal{B}(X(\mathbb{R})),\ \mathcal{CO}(\Psi) = \{W^0(b) : b \in \Psi\} \subset \mathcal{B}(X(\mathbb{R})),$$

respectively. The algebras $\mathcal{MO}(\Phi)$ and $\mathcal{CO}(\Psi)$ are building blocks of the algebra of convolution type operators

$$\mathcal{A}(\Phi, \Psi; X(\mathbb{R})) = \mathrm{alg}_{\mathcal{B}(X(\mathbb{R}))}\left\{aI, W^0(b) : a \in \Phi,\ b \in \Psi\right\},$$

the smallest closed subalgebra of $\mathcal{B}(X(\mathbb{R}))$ that contains the algebras $\mathcal{MO}(\Phi)$ and $\mathcal{CO}(\Psi)$.

Let $SO^\diamond$ be the C^*-algebra of slowly oscillating functions and $SO^\diamond_{X(\mathbb{R})}$ be the Banach algebra of all slowly oscillating Fourier multipliers on the space $X(\mathbb{R})$, which are defined below in Sects. 2.5–2.7. The third author proved in [22, Lemma 4.3] in the case of Lebesgue spaces $L^p(\mathbb{R}, w)$, $1 < p < \infty$, with Muckenhoupt weights $w \in A_p(\mathbb{R})$ that

$$\mathcal{MO}^\pi(SO^\diamond) \cap \mathcal{CO}^\pi(SO^\diamond_{L^p(\mathbb{R},w)}) = \mathcal{MO}^\pi(\mathbb{C}), \tag{1.2}$$

where

$$\mathcal{MO}^\pi(\mathbb{C}) := \{[cI]^\pi : c \in \mathbb{C}\}. \tag{1.3}$$

This result allowed him to describe the maximal ideal space of the commutative Banach algebra

$$\mathcal{A}^\pi(SO^\diamond, SO^\diamond_{L^p(\mathbb{R},w)}; L^p(\mathbb{R}, w)) = \mathcal{A}(SO^\diamond, SO^\diamond_{L^p(\mathbb{R},w)}; L^p(\mathbb{R}, w))/\mathcal{K}(L^p(\mathbb{R}, w))$$

(see [22, Theorem 3.1]). In turn, this description plays a crucial role in the study of the Fredholmness of operators in more general algebras of convolution type operators with piecewise slowly oscillating data on weighted Lebesgue space $L^p(\mathbb{R}, w)$ (see [22, 24, 25]).

Recall that the (non-centered) Hardy-Littlewood maximal function $\mathcal{M}f$ of a function $f \in L^1_{\mathrm{loc}}(\mathbb{R})$ is defined by

$$(\mathcal{M}f)(x) := \sup_{I \ni x} \frac{1}{|I|} \int_I |f(y)|\, dy,$$

where the supremum is taken over all intervals $I \subset \mathbb{R}$ of finite length containing x. The Hardy-Littlewood maximal operator $\mathcal{M}$ defined by the rule $f \mapsto \mathcal{M}f$ is a sublinear operator.

The aim of this paper is to extend (1.2) to the case of separable Banach function spaces such that the Hardy-Littlewood maximal operator $\mathcal{M}$ is bounded on $X(\mathbb{R})$ and on its associate space $X'(\mathbb{R})$ and to the case of arbitrary algebras of functions $\Phi \subset L^\infty(\mathbb{R})$ in place of $SO^\diamond$.

The following statement extends [22, Lemma 4.3].

Theorem 1.1 (Main Result) *Let $X(\mathbb{R})$ be a separable Banach function space such that the Hardy-Littlewood maximal operator $\mathcal{M}$ is bounded on the space $X(\mathbb{R})$ and on its associate space $X'(\mathbb{R})$. If Φ is a unital C^*-subalgebra of $L^\infty(\mathbb{R})$, then*

$$\mathcal{MO}^\pi(\Phi) \cap \mathcal{CO}^\pi(SO^\diamond_{X(\mathbb{R})}) = \mathcal{MO}^\pi(\mathbb{C}), \tag{1.4}$$

where $\mathcal{MO}^\pi(\mathbb{C})$ is defined by (1.3).

This result is one more step towards the study of Fredholm properties of convolution type operators with discontinuous data on Banach function spaces more general than weighted Lebesgue spaces initiated in the authors works [8–10].

One can expect, by analogy with the case of weighted Lebesgue spaces, that, for instance, $\mathcal{K}(X(\mathbb{R})) \subset \mathcal{A}(SO^\diamond, SO^\diamond_{X(\mathbb{R})}; X(\mathbb{R}))$ and that the quotient algebra

$$\mathcal{A}^\pi(SO^\diamond, SO^\diamond_{X(\mathbb{R})}; X(\mathbb{R})) = \mathcal{A}(SO^\diamond, SO^\diamond_{X(\mathbb{R})}; X(\mathbb{R}))/\mathcal{K}(X(\mathbb{R}))$$

is commutative. It seems, however, that the proofs of both hypotheses will require tools, which are not available in the setting of general Banach function spaces. We plan to return to these questions in a forthcoming work, restricting ourselves to particular Banach function spaces, like rearrangement-invariant spaces with Muckenhoupt weights or variable Lebesgue spaces, where interpolation theorems are available.

The paper is organized as follows. In Sect. 2, we collect necessary facts on Banach function spaces and Fourier multipliers on them. Further, we recall the definition of the C^*-algebra $SO^\diamond$ of slowly oscillating functions and introduce the Banach algebra of slowly oscillating Fourier multipliers $SO^\diamond_{X(\mathbb{R})}$ on a Banach function space $X(\mathbb{R})$. In Sect. 3, we discuss the structure of the maximal ideal spaces $M(SO^\diamond)$ and $M(SO^\diamond_{X(\mathbb{R})})$ of the C^*-algebra $SO^\diamond$ of slowly oscillating functions and the Banach algebra $SO^\diamond_{X(\mathbb{R})}$ of slowly oscillating Fourier multipliers on a Banach function space $X(\mathbb{R})$. In particular, we show that the fibers $M_t(SO^\diamond)$

of $M(SO^{\diamond})$ over the points $t \in \dot{\mathbb{R}} := \mathbb{R} \cup \{\infty\}$ can be identified with the fibers $M_t(SO_t)$, where SO_t is the C^*-algebra of all bounded continuous functions on $\dot{\mathbb{R}} \setminus \{t\}$ that slowly oscillate at the point t. An analogous result is also obtained for the fibers of the maximal ideal spaces of algebras of slowly oscillating Fourier multipliers on a Banach function space $X(\mathbb{R})$. In Sect. 4, we show that the maximal ideal spaces of the algebras $\mathcal{MO}^{\pi}(\Phi)$ and $\mathcal{CO}^{\pi}(\Psi)$ are homeomorphic to the maximal ideal spaces of the algebras Φ and Ψ, respectively, where Φ is a unital C^*-subalgebra of $L^{\infty}(\mathbb{R})$ and Ψ is a unital Banach subalgebra of $\mathcal{M}_{X(\mathbb{R})}$. In Sect. 5, we recall the definition of a limit operator (see [26] for a general theory of limit operators), as well as, a known fact about limit operators of compact operators acting on Banach function spaces. Further, we calculate the limit operators of the Fourier convolution operator $W^0(b)$ with a slowly oscillating symbol $b \in SO^{\diamond}_{X(\mathbb{R})}$. Finally, gathering the above mentioned results on limit operators, we prove Theorem 1.1.

2 Preliminaries

2.1 Banach Function Spaces

The set of all Lebesgue measurable complex-valued functions on $\mathbb{R}$ is denoted by $\mathfrak{M}(\mathbb{R})$. Let $\mathfrak{M}^+(\mathbb{R})$ be the subset of functions in $\mathfrak{M}(\mathbb{R})$ whose values lie in $[0, \infty]$. The Lebesgue measure of a measurable set $E \subset \mathbb{R}$ is denoted by $|E|$ and its characteristic function is denoted by χ_E. Following [4, Chap. 1, Definition 1.1], a mapping $\rho : \mathfrak{M}^+(\mathbb{R}) \to [0, \infty]$ is called a Banach function norm if, for all functions f, g, f_n $(n \in \mathbb{N})$ in $\mathfrak{M}^+(\mathbb{R})$, for all constants $a \geq 0$, and for all measurable subsets E of $\mathbb{R}$, the following properties hold:

(A1) $\rho(f) = 0 \Leftrightarrow f = 0$ a.e., $\quad \rho(af) = a\rho(f), \quad \rho(f + g) \leq \rho(f) + \rho(g)$,

(A2) $0 \leq g \leq f$ a.e. $\Rightarrow \rho(g) \leq \rho(f)$ (the lattice property),

(A3) $0 \leq f_n \uparrow f$ a.e. $\Rightarrow \rho(f_n) \uparrow \rho(f)$ (the Fatou property),

(A4) $|E| < \infty \Rightarrow \rho(\chi_E) < \infty$,

(A5) $|E| < \infty \Rightarrow \int_E f(x)\,dx \leq C_E \rho(f)$

with $C_E \in (0, \infty)$ which may depend on E and ρ but is independent of f. When functions differing only on a set of measure zero are identified, the set $X(\mathbb{R})$ of all functions $f \in \mathfrak{M}(\mathbb{R})$ for which $\rho(|f|) < \infty$ is called a Banach function space. For each $f \in X(\mathbb{R})$, the norm of f is defined by

$$\|f\|_{X(\mathbb{R})} := \rho(|f|).$$

Under the natural linear space operations and under this norm, the set $X(\mathbb{R})$ becomes a Banach space (see [4, Chap. 1, Theorems 1.4 and 1.6]). If ρ is a Banach function norm, its associate norm ρ' is defined on $\mathfrak{M}^+(\mathbb{R})$ by

$$\rho'(g) := \sup\left\{\int_{\mathbb{R}} f(x)g(x)\,dx \;:\; f \in \mathfrak{M}^+(\mathbb{R}),\ \rho(f) \le 1\right\}, \quad g \in \mathfrak{M}^+(\mathbb{R}).$$

It is a Banach function norm itself [4, Chap. 1, Theorem 2.2]. The Banach function space $X'(\mathbb{R})$ determined by the Banach function norm ρ' is called the associate space (Köthe dual) of $X(\mathbb{R})$. The associate space $X'(\mathbb{R})$ is naturally identified with a subspace of the (Banach) dual space $[X(\mathbb{R})]^*$.

2.2 Density of Nice Functions in Separable Banach Function Spaces

As usual, let $C_0^\infty(\mathbb{R})$ denote the set of all infinitely differentiable functions with compact support.

Lemma 2.1 ([8, Lemma 2.1] and [23, Lemma 2.12(a)]) *If $X(\mathbb{R})$ is a separable Banach function space, then the sets $C_0^\infty(\mathbb{R})$ and $L^2(\mathbb{R}) \cap X(\mathbb{R})$ are dense in the space $X(\mathbb{R})$.*

Let $\mathcal{S}(\mathbb{R})$ be the Schwartz space of rapidly decreasing smooth functions and let $\mathcal{S}_0(\mathbb{R})$ denote the set of functions $f \in \mathcal{S}(\mathbb{R})$ such that their Fourier transforms $\mathcal{F}f$ have compact support.

Theorem 2.2 ([10, Theorem 4]) *Let $X(\mathbb{R})$ be a separable Banach function space such that the Hardy-Littlewood maximal operator $\mathcal{M}$ is bounded on $X(\mathbb{R})$. Then the set $\mathcal{S}_0(\mathbb{R})$ is dense in the space $X(\mathbb{R})$.*

2.3 Banach Algebra $\mathcal{M}_{X(\mathbb{R})}$ of Fourier Multipliers

The following result plays an important role in this paper.

Theorem 2.3 ([21, Corollary 4.2] and [8, Theorem 2.4]) *Let $X(\mathbb{R})$ be a separable Banach function space such that the Hardy-Littlewood maximal operator $\mathcal{M}$ is bounded on $X(\mathbb{R})$ and on its associate space $X'(\mathbb{R})$. If $a \in \mathcal{M}_{X(\mathbb{R})}$, then*

$$\|a\|_{L^\infty(\mathbb{R})} \le \|a\|_{\mathcal{M}_{X(\mathbb{R})}}. \tag{2.1}$$

The constant 1 *on the right-hand side of* (2.1) *is best possible.*

Inequality (2.1) was established earlier in [18, Theorem 1] with some constant on the right-hand side that depends on the space $X(\mathbb{R})$.

Since (2.1) is available, an easy adaptation of the proof of [13, Proposition 2.5.13] leads to the following (we refer to the proof of [18, Corollary 1] for details).

Corollary 2.4 *Let $X(\mathbb{R})$ be a separable Banach function space such that the Hardy-Littlewood maximal operator $\mathcal{M}$ is bounded on $X(\mathbb{R})$ and on its associate space $X'(\mathbb{R})$. Then the set of Fourier multipliers $\mathcal{M}_{X(\mathbb{R})}$ is a Banach algebra under pointwise operations and the norm $\|\cdot\|_{\mathcal{M}_{X(\mathbb{R})}}$.*

2.4 Stechkin-Type Inequality

Let $V(\mathbb{R})$ be the Banach algebra of all functions $a : \mathbb{R} \to \mathbb{C}$ with finite total variation

$$V(a) := \sup \sum_{i=1}^{n} |a(t_i) - a(t_{i-1})|,$$

where the supremum is taken over all finite partitions

$$-\infty < t_0 < t_1 < \cdots < t_n < +\infty$$

of the real line $\mathbb{R}$ and the norm in $V(\mathbb{R})$ is given by

$$\|a\|_V = \|a\|_{L^\infty(\mathbb{R})} + V(a).$$

Theorem 2.5 *Let $X(\mathbb{R})$ be a separable Banach function space such that the Hardy-Littlewood maximal operator $\mathcal{M}$ is bounded on $X(\mathbb{R})$ and on its associate space $X'(\mathbb{R})$. If $a \in V(\mathbb{R})$, then the convolution operator $W^0(a)$ is bounded on the space $X(\mathbb{R})$ and*

$$\|W^0(a)\|_{\mathcal{B}(X(\mathbb{R}))} \le c_X \|a\|_V \tag{2.2}$$

where c_X is a positive constant depending only on $X(\mathbb{R})$.

This result follows from [17, Theorem 4.3].

For Lebesgue spaces $L^p(\mathbb{R})$, $1 < p < \infty$, inequality (2.2) is usually called Stechkin's inequality, and the constant c_{L^p} is calculated explicitly:

$$c_{L^p} = \|S\|_{\mathcal{B}(L^p(\mathbb{R}))} = \begin{cases} \tan\left(\frac{\pi}{2p}\right) \text{ if } 1 < p \le 2, \\ \cot\left(\frac{\pi}{2p}\right) \text{ if } 2 \le p < \infty, \end{cases} \tag{2.3}$$

where S is the Cauchy singular integral operator given by

$$(Sf)(x) := \frac{1}{\pi i} \lim_{\varepsilon \to 0} \int_{\mathbb{R} \setminus (x-\varepsilon, x+\varepsilon)} \frac{f(t)}{t-x}\, dt.$$

We refer to [7, Theorem 2.11] for the proof of (2.2) in the case of Lebesgue spaces $L^p(\mathbb{R})$ with $c_{L^p} = \|S\|_{\mathcal{B}(L^p(\mathbb{R}))}$ and to [12, Chap. 13, Theorem 1.3] for the calculation of the norm of S given in the second equality in (2.3). For Lebesgue spaces with Muckenhoupt weights $L^p(\mathbb{R}, w)$, the proof of Theorem 2.5 with $c_{L^p(w)} = \|S\|_{\mathcal{B}(L^p(\mathbb{R},w))}$ is contained in [5, Theorem 17.1]. Further, for variable Lebesgue spaces $L^{p(\cdot)}(\mathbb{R})$, Theorem 2.5 with $c_{L^{p(\cdot)}} = \|S\|_{\mathcal{B}(L^{p(\cdot)}(\mathbb{R}))}$ was obtained in [20, Theorem 2].

2.5 Slowly Oscillating Functions

Let $\dot{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$. For a set $E \subset \dot{\mathbb{R}}$ and a function $f : \dot{\mathbb{R}} \to \mathbb{C}$ in $L^\infty(\mathbb{R})$, let the oscillation of f over E be defined by

$$\operatorname{osc}(f, E) := \operatorname*{ess\,sup}_{s,t \in E} |f(s) - f(t)|.$$

Following [3, Section 4], [24, Section 2.1], and [25, Section 2.1], we say that a function $f \in L^\infty(\mathbb{R})$ is slowly oscillating at a point $\lambda \in \dot{\mathbb{R}}$ if for every $r \in (0, 1)$ or, equivalently, for some $r \in (0, 1)$, one has

$$\begin{aligned} \lim_{x \to 0+} \operatorname{osc}\big(f, \lambda + ([-x, -rx] \cup [rx, x])\big) &= 0 \text{ if } \lambda \in \mathbb{R}, \\ \lim_{x \to +\infty} \operatorname{osc}\big(f, [-x, -rx] \cup [rx, x]\big) &= 0 \qquad \text{if } \lambda = \infty. \end{aligned}$$

For every $\lambda \in \dot{\mathbb{R}}$, let SO_λ denote the C^*-subalgebra of $L^\infty(\mathbb{R})$ defined by

$$SO_\lambda := \left\{ f \in C_b(\dot{\mathbb{R}} \setminus \{\lambda\}) \ : \ f \text{ slowly oscillates at } \lambda \right\},$$

where $C_b(\dot{\mathbb{R}} \setminus \{\lambda\}) := C(\dot{\mathbb{R}} \setminus \{\lambda\}) \cap L^\infty(\mathbb{R})$.

Let $SO^\diamond$ be the smallest C^*-subalgebra of $L^\infty(\mathbb{R})$ that contains all the C^*-algebras SO_λ with $\lambda \in \dot{\mathbb{R}}$. The functions in $SO^\diamond$ are called slowly oscillating functions.

2.6 Banach Algebra SO^3_λ of Three Times Continuously Differentiable Slowly Oscillating Functions

For a point $\lambda \in \dot{\mathbb{R}}$, let $C^3(\mathbb{R} \setminus \{\lambda\})$ be the set of all three times continuously differentiable functions $a : \mathbb{R} \setminus \{\lambda\} \to \mathbb{C}$. Following [24, Section 2.4] and [25, Section 2.3], consider the commutative Banach algebras

$$SO^3_\lambda := \left\{ a \in SO_\lambda \cap C^3(\mathbb{R} \setminus \{\lambda\}) \ : \ \lim_{x \to \lambda} (D^k_\lambda a)(x) = 0, \ k = 1, 2, 3 \right\}$$

equipped with the norm

$$\|a\|_{SO^3_\lambda} := \sum_{k=0}^{3} \frac{1}{k!} \left\| D^k_\lambda a \right\|_{L^\infty(\mathbb{R})},$$

where $(D_\lambda a)(x) = (x - \lambda)a'(x)$ for $\lambda \in \mathbb{R}$ and $(D_\lambda a)(x) = xa'(x)$ for $\lambda = \infty$.

Lemma 2.6 *For every $\lambda \in \dot{\mathbb{R}}$, the set SO^3_λ is dense in the C^*-algebra SO_λ.*

Proof In view of [2, Lemma 2.3], the set

$$SO^\infty_\infty := \left\{ f \in SO_\infty \cap C^\infty_b(\mathbb{R}) \ : \ \lim_{x \to \infty} (D^k_\infty f)(x) = 0, \ k \in \mathbb{N} \right\} \tag{2.4}$$

is dense in the Banach algebra SO_∞. Here $C^\infty_b(\mathbb{R})$ denotes the set of all infinitely differentiable functions $f : \mathbb{R} \to \mathbb{C}$, which are bounded with all their derivatives. Note that SO^∞_∞ can be equivalently defined by replacing $C^\infty_b(\mathbb{R})$ in (2.4) by $C^\infty(\mathbb{R})$, because $f \in SO_\infty$ is bounded and its derivatives $f^{(k)}$ are bounded for all $k \in \mathbb{N}$ in view of $\lim_{x\to\infty}(D^k_\infty f)(x) = 0$. Since $SO^\infty_\infty \subset SO^3_\infty$, this completes the proof in the case $\lambda = \infty$.

If $\lambda \in \mathbb{R}$, then by Karlovich and Loreto Hernández [25, Corollary 2.2], the mapping $Ta = a \circ \beta_\lambda$, where $\beta_\lambda : \dot{\mathbb{R}} \to \dot{\mathbb{R}}$ is defined by

$$\beta_\lambda(x) = \frac{\lambda x - 1}{x + \lambda}, \tag{2.5}$$

is an isometric isomorphism of the algebra SO_λ onto the algebra SO_∞. Hence each function $a \in SO_\lambda$ can be approximated in the norm of SO_λ by functions $c_n = b_n \circ \beta^{-1}_\lambda$, where $b_n \in SO^\infty_\infty$ for $n \in \mathbb{N}$ and

$$\beta^{-1}_\lambda(y) = \frac{\lambda y + 1}{\lambda - y} = x, \quad x, y \in \dot{\mathbb{R}}. \tag{2.6}$$

It remains to show that $c_n \in SO^3_\lambda$. Taking into account (2.5)–(2.6), we obtain for $y = \beta_\lambda(x) \in \mathbb{R} \setminus \{\lambda\}$ and $x = \beta_\lambda^{-1}(y) \in \mathbb{R}$:

$$(D_\lambda c_n)(y) = b_n'\left(\beta_\lambda^{-1}(y)\right)\frac{\lambda^2+1}{y-\lambda} = -b_n'(x)(x+\lambda), \tag{2.7}$$

$$\begin{aligned}(D_\lambda^2 c_n)(y) =& b_n''\left(\beta_\lambda^{-1}(y)\right)\frac{(\lambda^2+1)^2}{y-\lambda} - b_n'\left(\beta_\lambda^{-1}(y)\right)\frac{\lambda^2+1}{y-\lambda}\\ =& -b_n''(x)(x+\lambda)(\lambda^2+1) + b_n'(x)(x+\lambda),\end{aligned} \tag{2.8}$$

$$\begin{aligned}(D_\lambda^3 c_n)(y) =& b_n'''\left(\beta_\lambda^{-1}(y)\right)\frac{(\lambda^2+1)^3}{y-\lambda} - 2b_n''\left(\beta_\lambda^{-1}(y)\right)\frac{(\lambda^2+1)^2}{y-\lambda}\\ &+ b_n'\left(\beta_\lambda^{-1}(y)\right)\frac{\lambda^2+1}{y-\lambda}\\ =& -b_n'''(x)(x+\lambda)(\lambda^2+1)^2 + 2b_n''(x)(x+\lambda)(\lambda^2+1)\\ &- b_n'(x)(x+\lambda).\end{aligned} \tag{2.9}$$

Since

$$\lim_{x\to\infty}(D_\infty^k b_n)(x) = 0 \quad \text{for} \quad k \in \{1,2,3\},$$

we see that

$$\lim_{x\to\infty} x^k b_n^{(k)}(x) = 0 \quad \text{for} \quad k \in \{1,2,3\}. \tag{2.10}$$

It follows from (2.7)–(2.10) that

$$\lim_{y\to\lambda}(D_\lambda^k c_n)(y) = 0 \quad \text{for} \quad k \in \{1,2,3\}.$$

Hence $c_n \in SO^3_\lambda$ for all $n \in \mathbb{N}$, which completes the proof. □

2.7 *Slowly Oscillating Fourier Multipliers*

The following result leads us to the definition of slowly oscillating Fourier multipliers.

Theorem 2.7 ([19, Theorem 2.5]) *Let $X(\mathbb{R})$ be a separable Banach function space such that the Hardy-Littlewood maximal operator $\mathcal{M}$ is bounded on $X(\mathbb{R})$ and on*

its associate space $X'(\mathbb{R})$. *If* $\lambda \in \dot{\mathbb{R}}$ *and* $a \in SO^3_\lambda$, *then the convolution operator* $W^0(a)$ *is bounded on the space* $X(\mathbb{R})$ *and*

$$\|W^0(a)\|_{\mathcal{B}(X(\mathbb{R}))} \le c_X \|a\|_{SO^3_\lambda}, \tag{2.11}$$

where c_X *is a positive constant depending only on* $X(\mathbb{R})$.

Let $SO_{\lambda,X(\mathbb{R})}$ denote the closure of SO^3_λ in the norm of $\mathcal{M}_{X(\mathbb{R})}$. Further, let $SO^\diamond_{X(\mathbb{R})}$ be the smallest Banach subalgebra of $\mathcal{M}_{X(\mathbb{R})}$ that contains all the Banach algebras $SO_{\lambda,X(\mathbb{R})}$ for $\lambda \in \dot{\mathbb{R}}$. The functions in $SO^\diamond_{X(\mathbb{R})}$ will be called slowly oscillating Fourier multipliers.

Lemma 2.8 *Let* $X(\mathbb{R})$ *be a separable Banach function space such that the Hardy-Littlewood maximal operator* $\mathcal{M}$ *is bounded on* $X(\mathbb{R})$ *and on its associate space* $X'(\mathbb{R})$. *Then*

$$SO^\diamond_{X(\mathbb{R})} \subset SO^\diamond_{L^2(\mathbb{R})} = SO^\diamond.$$

Proof The continuous embedding $SO^\diamond_{X(\mathbb{R})} \subset SO^\diamond_{L^2(\mathbb{R})}$ (with the embedding constant one) follows immediately from Theorem 2.3 and the definitions of the Banach algebras $SO^\diamond_{X(\mathbb{R})}$ and $SO^\diamond_{L^2(\mathbb{R})}$. It is clear that $SO^\diamond_{L^2(\mathbb{R})} \subset SO^\diamond$. The embedding $SO^\diamond \subset SO^\diamond_{L^2(\mathbb{R})}$ follows from Lemma 2.6. □

3 Maximal Ideal Spaces of the Algebras $SO^\diamond$ and $SO^\diamond_{X(\mathbb{R})}$

3.1 Extensions of Multiplicative Linear Functionals on C*-algebras

For a C^*-algebra (or, more generally, a Banach algebra) $\mathfrak{A}$ with unit e and an element $a \in \mathfrak{A}$, let $\mathrm{sp}_{\mathfrak{A}}(a)$ denote the spectrum of a in $\mathfrak{A}$. Recall that an element a of a C^*-algebra $\mathfrak{A}$ is said to be positive if it is self-adjoint and $\mathrm{sp}_{\mathfrak{A}}(a) \subset [0, \infty)$. A linear functional ϕ on $\mathfrak{A}$ is said to be a state if $\phi(a) \ge 0$ for all positive elements $a \in \mathfrak{A}$ and $\phi(e) = 1$. The set of all states of $\mathfrak{A}$ is denoted by $\mathfrak{S}(\mathfrak{A})$. The extreme points of $\mathfrak{S}(\mathfrak{A})$ are called pure states of $\mathfrak{A}$ (see, e.g., [15, Section 4.3]).

Following [1, p. 304], for a state ϕ, let

$$\mathcal{G}_\phi(\mathfrak{A}) := \{a \in \mathfrak{A} : |\phi(a)| = \|a\|_{\mathfrak{A}} = 1\}$$

and let $\mathcal{G}^+_\phi(\mathfrak{A})$ denote the set of all positive elements of $\mathcal{G}_\phi(\mathfrak{A})$. Let $\mathfrak{A}$ and $\mathfrak{B}$ be C^*-algebras such that $e \in \mathfrak{B} \subset \mathfrak{A}$. Let ϕ be a state of $\mathfrak{B}$. Following [1, p. 310], we say that $\mathfrak{A}$ is $\mathfrak{B}$-compressible modulo ϕ if for each $x \in \mathfrak{A}$ and each $\varepsilon > 0$ there is $b \in \mathcal{G}^+_\phi(\mathfrak{B})$ and $y \in \mathfrak{B}$ such that $\|bxb - y\|_{\mathfrak{A}} < \varepsilon$.

Since a nonzero linear functional on a commutative C^*-algebra is a pure state if and only if it is multiplicative (see, e.g., [15, Proposition 4.4.1]), we immediately get the following lemma from [1, Theorem 3.2].

Lemma 3.1 *Let $\mathfrak{B}$ be a C^*-subalgebra of a commutative C^*-algebra $\mathfrak{A}$. A nonzero multiplicative linear functional ϕ on $\mathfrak{B}$ admits a unique extension to a multiplicative linear functional ϕ' on $\mathfrak{A}$ if and only if $\mathfrak{A}$ is $\mathfrak{B}$-compressible modulo ϕ.*

3.2 Family of Positive Elements

For $t \in \dot{\mathbb{R}}$ and $\omega > 0$, let $\psi_{t,\omega}$ be a real-valued function in $C(\dot{\mathbb{R}})$ such that $0 \leq \psi_{t,\omega}(x) \leq 1$ for all $x \in \mathbb{R}$. Assume that for $t \in \mathbb{R}$,

$$\psi_{t,\omega}(s) = 1 \quad \text{if} \quad s \in (t-\omega, t+\omega), \quad \psi_{t,\omega}(s) = 0 \quad \text{if} \quad s \in \mathbb{R} \setminus (t-2\omega, t+2\omega),$$

and for $t = \infty$,

$$\psi_{\infty,\omega}(s) = 1 \quad \text{if} \quad s \in \mathbb{R} \setminus (-2\omega, 2\omega), \quad \psi_{\infty,\omega}(s) = 0 \quad \text{if} \quad s \in (-\omega, \omega).$$

Let $M(\mathfrak{A})$ denote the maximal ideal space of a commutative Banach algebra $\mathfrak{A}$.

Lemma 3.2 *For $t \in \dot{\mathbb{R}}$ and $\omega > 0$, the function $\psi_{t,\omega}$ is a positive element of the C^*-algebras $C(\dot{\mathbb{R}})$, SO_t, and $SO^\diamond$.*

Proof Since $M(C(\dot{\mathbb{R}})) = \dot{\mathbb{R}}$, it follows from the Gelfand theorem (see, e.g., [28, Theorem 2.1.3]) that $\mathrm{sp}_{C(\dot{\mathbb{R}})}(\psi_{t,\omega}) = [0, 1]$ for all $t \in \dot{\mathbb{R}}$ and all $\omega > 0$. Since $C(\dot{\mathbb{R}}) \subset SO_t \subset SO^\diamond$, we conclude that the functions $\psi_{t,\omega}$ for $t \in \dot{\mathbb{R}}$ and $\omega > 0$ are positive elements of the C^*-algebras $C(\dot{\mathbb{R}})$, SO_t, and $SO^\diamond$ because their spectra in each of these algebras coincide with $[0, 1]$ in view of [15, Proposition 4.1.5]. □

3.3 Maximal Ideal Space of the C*-algebra $SO^\diamond$

If $\mathfrak{B}$ is a Banach subalgebra of $\mathfrak{A}$ and $\lambda \in M(\mathfrak{B})$, then the set

$$M_\lambda(\mathfrak{A}) := \{\xi \in M(\mathfrak{A}) : \xi|_{\mathfrak{B}} = \lambda\}$$

is called the fiber of $M(\mathfrak{A})$ over $\lambda \in M(\mathfrak{B})$. Hence for every Banach algebra $\Phi \subset L^\infty(\mathbb{R})$ with $M(C(\dot{\mathbb{R}}) \cap \Phi) = \dot{\mathbb{R}}$ and every $t \in \dot{\mathbb{R}}$, the fiber $M_t(\Phi)$ is the set of all multiplicative linear functionals (characters) on Φ that annihilate the set $\{f \in C(\dot{\mathbb{R}}) \cap \Phi : f(t) = 0\}$. As usual, for all $a \in \Phi$ and all $\xi \in M(\Phi)$, we put $a(\xi) := \xi(a)$. We will frequently identify the points $t \in \dot{\mathbb{R}}$ with the evaluation functionals δ_t

defined by

$$\delta_t(f) = f(t) \quad \text{for} \quad f \in C(\dot{\mathbb{R}}), \quad t \in \dot{\mathbb{R}}.$$

Lemma 3.3 *For every point $t \in \dot{\mathbb{R}}$, the fibers $M_t(SO_t)$ and $M_t(SO^\diamond)$ can be identified as sets:*

$$M_t(SO_t) = M_t(SO^\diamond). \tag{3.1}$$

Proof Since $C(\dot{\mathbb{R}}) \subset SO_t \subset SO^\diamond$, by the restriction of a multiplicative linear functional defined on a bigger algebra to a smaller algebra, we have

$$M(SO^\diamond) \subset M(SO_t) \subset M(C(\dot{\mathbb{R}})), \quad t \in \dot{\mathbb{R}}. \tag{3.2}$$

Since

$$M(\Phi) = \bigcup_{t \in \dot{\mathbb{R}}} M_t(\Phi) \quad \text{for} \quad \Phi \in \{SO^\diamond, SO_\lambda : \lambda \in \dot{\mathbb{R}}\},$$

where

$$M_t(\Phi) = \{\zeta \in M(\Phi) \,:\, \zeta|_{C(\dot{\mathbb{R}})} = \delta_t\}, \quad t \in \dot{\mathbb{R}}, \tag{3.3}$$

it follows from (3.2) and (3.3) that

$$M_t(SO^\diamond) \subset M_t(SO_t), \quad t \in \mathbb{R}. \tag{3.4}$$

Now fix $t \in \dot{\mathbb{R}}$ and a multiplicative linear functional $\eta \in M_t(SO_t)$. Let us show that the C^*-algebra $SO^\diamond$ is SO_t-compressible modulo η. Take $\varepsilon > 0$. By the definition of $SO^\diamond$, for a function $x \in SO^\diamond$, there are a finite set $F \in \dot{\mathbb{R}}$ and a finite set $\{x_\lambda \in SO_\lambda : \lambda \in F\}$ such that

$$\left\| x - \sum_{\lambda \in F} x_\lambda \right\|_{L^\infty(\mathbb{R})} < \varepsilon.$$

If $t \neq \infty$, take ω such that

$$0 < \omega < \frac{1}{2} \min_{\lambda \in F \setminus \{t\}} |\lambda - t|$$

and $b := \psi_{t,\omega}$. Then

$$y := b \left(\sum_{\lambda \in F} x_\lambda \right) b \tag{3.5}$$

is equal to zero outside the interval $(t - 2\omega, t + 2\omega)$. Therefore, $y \in SO_t$.

If $t = \infty$, take ω such that

$$\omega > \max_{\lambda \in F \setminus \{\infty\}} |\lambda|$$

and $b := \psi_{\infty,\omega}$. Then the function y defined by (3.5) is equal to zero on $(-\omega, \omega)$ and $y \in SO_\infty$.

For $t \in \dot{\mathbb{R}}$, we have

$$\|bxb - y\|_{L^\infty(\mathbb{R})} = \left\| b\left(x - \sum_{\lambda \in F} x_\lambda\right) b \right\|_{L^\infty(\mathbb{R})} \leq \left\| x - \sum_{\lambda \in F} x_\lambda \right\|_{L^\infty(\mathbb{R})} < \varepsilon.$$

Since b is a positive element of SO_t in view of Lemma 3.2, we have $b \in \mathcal{G}_\eta^+(SO_t)$, which completes the proof of the fact that $SO^\diamond$ is SO_t-compressible modulo the multiplicative linear functional $\eta \in M_t(SO_t)$.

In view of Lemma 3.1, there exists a unique extension η' of the multiplicative linear functional η to the whole algebra $SO^\diamond$. By the definition of the fiber $M_t(SO^\diamond)$, we have $\eta' \in M_t(SO^\diamond)$. Thus, we can identify $M_t(SO_t)$ with a subset of $M_t(SO^\diamond)$:

$$M_t(SO_t) \subset M_t(SO^\diamond). \tag{3.6}$$

Combining (3.4) and (3.6), we arrive at (3.1). □

Corollary 3.4 *The maximal ideal space of the commutative C^*-algebra $SO^\diamond$ can be identified with the set*

$$\bigcup_{t \in \dot{\mathbb{R}}} M_t(SO_t).$$

3.4 Extensions of Multiplicative Linear Functionals on Banach Algebras

The following theorem in a slightly different form is contained in [29, Theorem 2.1.1] and [30, Theorem 3.10]. For the convenience of readers, we give its proof here.

Theorem 3.5 *Let $\mathfrak{A}, \mathfrak{B}, \mathfrak{C}$ be commutative unital Banach algebras with common unit and homomorphic imbeddings $\mathfrak{A} \subset \mathfrak{B} \subset \mathfrak{C}$, where $\mathfrak{A}$ is dense in $\mathfrak{B}$. If for each functional $\varphi \in M(\mathfrak{A})$ there exists a unique extension $\varphi' \in M(\mathfrak{C})$, then for every functional $\psi \in M(\mathfrak{B})$ there exists a unique extension $\psi' \in M(\mathfrak{C})$.*

Proof Let $\psi \in M(\mathfrak{B})$. Then $\psi_1 := \psi|_{\mathfrak{A}} \in M(\mathfrak{A})$. By the hypotheses, there exists a unique extension $\psi_3 := (\psi_1)' \in M(\mathfrak{C})$. Then $\psi_1(a) = \psi(a) = \psi_3(a)$ for all $a \in \mathfrak{A}$.

Let $\psi_2 := \psi_3|_{\mathfrak{B}} \in M(\mathfrak{B})$. Since $\mathfrak{A} \subset \mathfrak{B}$, it follows that

$$\psi(a) = \psi_2(a) \quad \text{for all} \quad a \in \mathfrak{A}. \tag{3.7}$$

On the other hand, functionals $\psi, \psi_2 \in M(\mathfrak{B})$ are continuous on $\mathfrak{B}$ (see, e.g., [16, Lemma 2.1.5]). Since $\mathfrak{A}$ is dense in $\mathfrak{B}$, for every $b \in \mathfrak{B}$ there exists a sequence $\{a_n\}_{n\in\mathbb{N}} \subset \mathfrak{A}$ such that $\|a_n - b\|_{\mathfrak{B}} \to 0$ as $n \to \infty$. It follows from this observation and (3.7) that for every $b \in \mathfrak{B}$,

$$\psi(b) = \lim_{n\to\infty} \psi(a_n) = \lim_{n\to\infty} \psi_2(a_n) = \psi_2(b) = \psi_3(b).$$

Thus $\psi_3 \in M(\mathfrak{C})$ is an extension of ψ. This extension is unique by construction. □

3.5 Maximal Ideal Space of the Banach Algebras $SO_{t,X(\mathbb{R})}$

We start with the following refinement of [25, Lemma 3.4].

Lemma 3.6 *Let $t \in \dot{\mathbb{R}}$. Then for each functional $\varphi \in M(SO_t^3)$ there exists a unique extension $\varphi' \in M(SO_t)$.*

The density of SO_t^3 in the Banach algebra SO_t essentially used in the proof of [25, Lemma 3.4] is justified in Lemma 2.6. Note that the uniqueness of an extension was not explicitly mentioned in [25, Lemma 3.4]. However, since $M(SO_t^3)$ and $M(SO_t)$ are Hausdorff spaces (see, e.g., [16, Theorem 2.2.3]), the uniqueness of an extension constructed in the proof of [25, Lemma 3.4] is a consequence of a standard fact from general topology (see, e.g., [27, Theorem IV.2(b)]).

The following lemma is analogous to [25, Lemma 3.5].

Lemma 3.7 *Let $X(\mathbb{R})$ be a separable Banach function space such that the Hardy-Littlewood maximal operator $\mathcal{M}$ is bounded on the space $X(\mathbb{R})$ and on its associate space $X'(\mathbb{R})$. If $t \in \dot{\mathbb{R}}$, then the maximal ideal spaces of the C^*-algebra SO_t and the Banach algebra $SO_{t,X(\mathbb{R})}$ can be identified as sets:*

$$M(SO_t) = M(SO_{t,X(\mathbb{R})}). \tag{3.8}$$

Proof It follows from Theorem 2.3 that $SO_t^3 \subset SO_{t,X(\mathbb{R})} \subset SO_t$, where the imbeddings are homomorphic. By the definition of the algebra $SO_{t,X(\mathbb{R})}$, the algebra SO_t^3 is dense in $SO_{t,X(\mathbb{R})}$ with respect to the norm of $\mathcal{M}_{X(\mathbb{R})}$. Taking into account these observations and Lemma 3.6, we see that the commutative Banach algebras

$$\mathfrak{A} = SO_t^3, \quad \mathfrak{B} = SO_{t,X(\mathbb{R})}, \quad \mathfrak{C} = SO_t$$

satisfy all the conditions of Theorem 3.5. By this theorem, every multiplicative linear functional on $SO_{t,X(\mathbb{R})}$ admits a unique extension to a multiplicative linear

functional on SO_t. Hence we can identify $M(SO_{t,X(\mathbb{R})})$ with a subset of $M(SO_t)$:

$$M(SO_{t,X(\mathbb{R})}) \subset M(SO_t). \tag{3.9}$$

On the other hand, since $SO_{t,X(\mathbb{R})} \subset SO_t$, by the restriction of a multiplicative linear functional defined on a bigger algebra to a smaller algebra, we have

$$M(SO_t) \subset M(SO_{t,X(\mathbb{R})}). \tag{3.10}$$

Combining inclusions (3.9) and (3.10), we immediately arrive at (3.8). □

The next lemma is analogous to Lemma 3.3.

Lemma 3.8 *Let $X(\mathbb{R})$ be a separable Banach function space such that the Hardy-Littlewood maximal operator $\mathcal{M}$ is bounded on the space $X(\mathbb{R})$ and on its associate space $X'(\mathbb{R})$. Then, for every point $t \in \dot{\mathbb{R}}$, the fibers $M_t(SO_{t,X(\mathbb{R})})$ and $M_t(SO^{\diamond}_{X(\mathbb{R})})$ can be identified as sets:*

$$M_t(SO_{t,X(\mathbb{R})}) = M_t(SO^{\diamond}_{X(\mathbb{R})}). \tag{3.11}$$

Proof Since $SO_{t,X(\mathbb{R})} \subset SO^{\diamond}_{X(\mathbb{R})}$ for every $t \in \dot{\mathbb{R}}$, we conclude by the restriction of a multiplicative linear functional defined on the bigger algebra to the smaller algebra that $M(SO^{\diamond}_{X(\mathbb{R})}) \subset M(SO_{t,X(\mathbb{R})})$. Hence

$$M_t(SO^{\diamond}_{X(\mathbb{R})}) \subset M_t(SO_{t,X(\mathbb{R})}). \tag{3.12}$$

On the other hand, in view of Lemma 3.7, any multiplicative linear functional $\xi \in M_t(SO_{t,X(\mathbb{R})})$ admits a unique extension $\xi' \in M(SO_t)$. Moreover, ξ' belongs to $M_t(SO_t)$ as well. By Lemma 3.3, the functional $\xi' \in M_t(SO_t)$ admits a unique extension $\xi'' \in M_t(SO^{\diamond})$. It is clear that the restriction of ξ'' to $SO^{\diamond}_{X(\mathbb{R})}$ belongs to $M_t(SO^{\diamond}_{X(\mathbb{R})})$. Thus $M_t(SO_{t,X(\mathbb{R})})$ can be identified with a subset of $M_t(SO^{\diamond}_{X(\mathbb{R})})$:

$$M_t(SO_{t,X(\mathbb{R})}) \subset M_t(SO^{\diamond}_{X(\mathbb{R})}). \tag{3.13}$$

Combining (3.12) and (3.13), we arrive at (3.11). □

3.6 Maximal Ideal Space of the Banach Algebra $SO^{\diamond}_{X(\mathbb{R})}$

Now we are in a position to prove that the maximal ideal spaces of the commutative Banach algebra $SO^{\diamond}_{X(\mathbb{R})}$ and the C^*-algebra $SO^{\diamond}$ can be identified as sets.

Theorem 3.9 *Let $X(\mathbb{R})$ be a separable Banach function space such that the Hardy-Littlewood maximal operator $\mathcal{M}$ is bounded on the space $X(\mathbb{R})$ and on its associate*

space $X'(\mathbb{R})$. Then the maximal ideal space of the Banach algebra $SO^{\diamond}_{X(\mathbb{R})}$ can be identified with the maximal ideal space of the C^*-algebra $SO^{\diamond}$:

$$M(SO^{\diamond}_{X(\mathbb{R})}) = M(SO^{\diamond}).$$

Proof It follows from Lemmas 3.8, 3.7, and 3.3 that for every $t \in \dot{\mathbb{R}}$,

$$M_t(SO^{\diamond}_{X(\mathbb{R})}) = M_t(SO_{t,X(\mathbb{R})}) = M_t(SO_t) = M_t(SO^{\diamond}).$$

Hence

$$M(SO^{\diamond}_{X(\mathbb{R})}) = \bigcup_{t\in\dot{\mathbb{R}}} M_t(SO^{\diamond}_{X(\mathbb{R})}) = \bigcup_{t\in\dot{\mathbb{R}}} M_t(SO^{\diamond}) = M(SO^{\diamond}),$$

which completes the proof. □

4 Maximal Ideal Spaces of the Calkin Images of the Banach Algebras $\mathcal{MO}(\Phi)$ and $\mathcal{CO}(\Psi)$

4.1 Maximal Ideal Space of the Algebra $\mathcal{MO}^{\pi}(\Phi)$

We start with the following known result [14, Theorem 2.4] (see also [9, Theorem 3.1]).

Theorem 4.1 *Let $X(\mathbb{R})$ be a separable Banach function space and $a \in L^{\infty}(\mathbb{R})$. Then the multiplication operator aI is compact on the space $X(\mathbb{R})$ if and only if $a = 0$ almost everywhere on $\mathbb{R}$.*

The next theorem says that one can identify the maximal ideal spaces of the algebras $\mathcal{MO}^{\pi}(\Phi)$ and Φ for an arbitrary unital C^*-subalgebra of $L^{\infty}(\mathbb{R})$.

Theorem 4.2 *Let $X(\mathbb{R})$ be a separable Banach function space. If Φ is a unital C^*-subalagebra of $L^{\infty}(\mathbb{R})$, then the maximal ideal spaces of the commutative Banach algebra $\mathcal{MO}^{\pi}(\Phi)$ and the commutative C^*-algebra Φ are homeomorphic:*

$$M(\mathcal{MO}^{\pi}(\Phi)) = M(\Phi).$$

Proof Consider the mapping $F : \Phi \to \mathcal{MO}^{\pi}(\Phi)$ defined by $F(a) = [aI]^{\pi}$ for every $a \in \Phi$. It is clear that this mapping is surjective. If $[aI]^{\pi} = [bI]^{\pi}$ for some $a, b \in \Phi$, then $(a - b)I \in \mathcal{K}(X(\mathbb{R}))$. It follows from Theorem 4.1 that $a = b$ a.e. on $\mathbb{R}$. This implies that the mapping F is injective. Thus, $F : \Phi \to \mathcal{MO}^{\pi}(\Phi)$ is an algebraic isomorphism of commutative Banach algebras. It follows from [16, Lemma 2.2.12] that the maximal ideal spaces $M(\mathcal{MO}^{\pi}(\Phi))$ and $M(\Phi)$ are homeomorphic. □

4.2 Maximal Ideal Space of the Algebra $\mathcal{CO}^{\pi}(\Psi)$

The following analogue of Theorem 4.1 for Fourier convolution operators was obtained recently by the authors [8, Theorem 1.1].

Theorem 4.3 *Let $X(\mathbb{R})$ be a separable Banach function space such that the Hardy-Littlewood maximal operator $\mathcal{M}$ is bounded on $X(\mathbb{R})$ and on its associate space $X'(\mathbb{R})$. Suppose that $b \in \mathcal{M}_{X(\mathbb{R})}$. Then the Fourier convolution operator $W^0(a)$ is compact on the space $X(\mathbb{R})$ if and only if $b = 0$ almost everywhere on $\mathbb{R}$.*

The next theorem is an analogue of Theorem 4.2 for Fourier multipliers.

Theorem 4.4 *Let $X(\mathbb{R})$ be a separable Banach function space such that the Hardy-Littlewood maximal operator $\mathcal{M}$ is bounded on $X(\mathbb{R})$ and on its associate space $X'(\mathbb{R})$. If Ψ is a unital Banach subalagebra of $\mathcal{M}_{X(\mathbb{R})}$, then the maximal ideal spaces of the commutative Banach algebras $\mathcal{CO}^{\pi}(\Psi)$ and Ψ are homeomorphic:*

$$M(\mathcal{CO}^{\pi}(\Psi)) = M(\Psi).$$

Proof The proof is analogous to the proof of Theorem 4.2. Consider the mapping $F : \Psi \to \mathcal{CO}^{\pi}(\Psi)$ defined by $F(a) = [W^0(a)]^{\pi}$ for every $a \in \Psi$. It is obvious that this mapping is surjective. If $[W^0(a)]^{\pi} = [W^0(b)]^{\pi}$ for some $a, b \in \Psi$, then $W^0(a - b) = W^0(a) - W^0(b) \in \mathcal{K}(X(\mathbb{R}))$. By Theorem 4.3, we conclude that $a = b$ a.e. on $\mathbb{R}$. Therefore, the mapping F is injective. Thus, $F : \Psi \to \mathcal{CO}^{\pi}(\Psi)$ is an algebraic isomorphism of commutative Banach algebras. In this case it follows from [16, Lemma 2.2.12] that the maximal ideal spaces $M(\mathcal{CO}^{\pi}(\Psi))$ and $M(\Psi)$ are homeomorphic. □

5 Applications of the Method of Limit Operators

5.1 Known Result about Limit Operators on Banach Function Spaces

Let $X(\mathbb{R})$ be a Banach function space. For a sequence of operators $\{A_n\}_{n\in\mathbb{N}} \subset \mathcal{B}(X(\mathbb{R}))$, let

$$\operatorname*{s-lim}_{n\to\infty} A_n$$

denote the strong limit of this sequence, if it exists. For $\lambda, x \in \mathbb{R}$, consider the function

$$e_{\lambda}(x) := e^{i\lambda x}.$$

Let $T \in \mathcal{B}(X(\mathbb{R}))$ and let $h = \{h_n\}_{n\in\mathbb{N}}$ be a sequence of numbers $h_n > 0$ such that $h_n \to +\infty$ as $n \to \infty$. The strong limit

$$T_h := \text{s-}\lim_{n\to\infty} e_{h_n} T e_{h_n}^{-1} I$$

is called the limit operator of T related to the sequence $h = \{h_n\}_{n\in\mathbb{N}}$, if it exists.

In our previous paper [9] we calculated the limit operators for all compact operators.

Lemma 5.1 ([9, Lemma 3.2]) *Let $X(\mathbb{R})$ be a separable Banach function space and K be a compact operator on $X(\mathbb{R})$. Then for every sequence $\{h_n\}_{n\in\mathbb{N}}$ of positive numbers satisfying $h_n \to +\infty$ as $n \to \infty$, one has*

$$\text{s-}\lim_{n\to\infty} e_{h_n} K e_{h_n}^{-1} I = 0$$

on the space $X(\mathbb{R})$.

5.2 Limit Operators for Fourier Convolution Operators with Symbols in the Algebra $SO^{\diamond}_{X(\mathbb{R})}$

Now we will calculate the limit operators for the Fourier convolution operator with a slowly oscillating symbol.

Theorem 5.2 *Let $X(\mathbb{R})$ be a separable Banach function space such that the Hardy-Littlewood maximal operator $\mathcal{M}$ is bounded on the space $X(\mathbb{R})$ and on its associate space $X'(\mathbb{R})$. If $b \in SO^{\diamond}_{X(\mathbb{R})}$, then for every $\xi \in M_\infty(SO^{\diamond})$ there exists a sequence $\{h_n\}_{n\in\mathbb{N}}$ of positive numbers such that $h_n \to +\infty$ as $n \to \infty$ and*

$$\text{s-}\lim_{n\to\infty} e_{h_n} W^0(b) e_{h_n}^{-1} I = b(\xi) I \tag{5.1}$$

on the space $X(\mathbb{R})$.

Proof This statement is proved by analogy with [25, Lemma 5.1]. In view of Lemma 2.8, $SO^{\diamond}_{X(\mathbb{R})} \subset SO^{\diamond}$. Therefore every $\xi \in M_\infty(SO^{\diamond})$ is a multiplicative linear functional on $SO^{\diamond}_{X(\mathbb{R})}$, that is, $b(\xi)$ is well defined. By the definition of $SO^{\diamond}_{X(\mathbb{R})}$, if $b \in SO^{\diamond}_{X(\mathbb{R})}$, then there is a sequence

$$b_m = \sum_{\lambda\in F_m} b_{m,\lambda}, \quad m \in \mathbb{N},$$

where $F_m \subset \dot{\mathbb{R}}$ are finite sets and $b_{m,\lambda} \in SO^3_\lambda$ for $\lambda \in F_m$ and all $m \in \mathbb{N}$, such that

$$\lim_{m\to\infty} \|b_m - b\|_{\mathcal{M}_{X(\mathbb{R})}} = 0. \tag{5.2}$$

By Lemma 3.3, $M_\infty(SO^\diamond) = M_\infty(SO_\infty)$. Fix $\xi \in M_\infty(SO^\diamond) = M_\infty(SO_\infty)$. Assume first that the set

$$B_\infty := \{b_{m,\infty} \in SO^3_\infty \ : \ m \in \mathbb{N}\}$$

is not empty. Since the set B_∞ is at most countable, it follows from [2, Corollary 3.3] or [25, Proposition 3.1] that there exists a sequence $\{h_n\}_{n\in\mathbb{N}}$ such that $h_n \to +\infty$ as $n \to \infty$ and

$$\xi(b_{m,\infty}) = \lim_{n\to\infty} b_{m,\infty}(h_n) \quad \text{for all} \quad b_{m,\infty} \in B_\infty. \tag{5.3}$$

As the functions $b_{m,\lambda}$ are continuous at ∞ if $\lambda \neq \infty$, we see that

$$\xi(b_{m,\lambda}) = b_{m,\lambda}(\infty) = \lim_{n\to\infty} b_{m,\lambda}(h_n) \quad \text{for all} \quad \lambda \in \bigcup_{m\in\mathbb{N}} F_m \setminus \{\infty\}. \tag{5.4}$$

Combining (5.3) and (5.4), for every $m \in \mathbb{N}$, we get

$$\begin{aligned} \xi(b_m) &= \sum_{\lambda\in F_m} \xi(b_{m,\lambda}) = \sum_{\lambda\in F_m} \lim_{n\to\infty} b_{m,\lambda}(h_n) \\ &= \lim_{n\to\infty} \sum_{\lambda\in F_m} b_{m,\lambda}(h_n) = \lim_{n\to\infty} b_m(h_n). \end{aligned} \tag{5.5}$$

If the set B_∞ is empty, we can take an arbitrary sequence $\{h_n\}_{n\in\mathbb{N}}$ such that $h_n \to +\infty$ as $n \to \infty$.

Let $f \in \mathcal{S}_0(\mathbb{R})$. Then, by a smooth version of Urysohn's lemma (see, e.g., [11, Proposition 6.5]), there is a function $\psi \in C^\infty_0(\mathbb{R})$ such that $0 \le \psi \le 1$, $\operatorname{supp}\mathcal{F}f \subset \operatorname{supp}\psi$ and $\psi|_{\operatorname{supp}\mathcal{F}f} = 1$. Therefore, for all $n \in \mathbb{N}$,

$$\begin{aligned} e_{h_n} W^0(b) e^{-1}_{h_n} f - b(\xi) f &= W^0[b(\cdot + h_n)]f - \xi(b) f \\ &= \mathcal{F}^{-1}[b(\cdot + h_n) - \xi(b)]\psi\mathcal{F}f \end{aligned}$$

and

$$\left\| \left(e_{h_n} W^0(b) e^{-1}_{h_n} - b(\xi)\right) f \right\|_{X(\mathbb{R})} \le \left\| \left[b(\cdot + h_n) - \xi(b)\right]\psi \right\|_{\mathcal{M}_{X(\mathbb{R})}} \|f\|_{X(\mathbb{R})}. \tag{5.6}$$

Since $\mathcal{M}_{X(\mathbb{R})}$ is translation-invariant and $\xi \in M_\infty(SO^\diamond)$ is a multiplicative linear functional on $SO^\diamond_{X(\mathbb{R})}$, we infer for all $m, n \in \mathbb{N}$ that

$$\begin{aligned} \left\| \left[b(\cdot + h_n) - \xi(b)\right]\psi \right\|_{\mathcal{M}_{X(\mathbb{R})}} \le{}& \left\| \left[b(\cdot + h_n) - b_m(\cdot + h_n)\right]\psi \right\|_{\mathcal{M}_{X(\mathbb{R})}} \\ &+ \left\| \left[b_m(\cdot + h_n) - \xi(b_m)\right]\psi \right\|_{\mathcal{M}_{X(\mathbb{R})}} \end{aligned}$$

$$\begin{aligned} &+ \big\|\xi(b_m) - \xi(b)\big]\psi\|_{\mathcal{M}_{X(\mathbb{R})}} \\ \leq &2\|b - b_m\|_{\mathcal{M}_{X(\mathbb{R})}}\|\psi\|_{\mathcal{M}_{X(\mathbb{R})}} \\ &+ \big\|\big[b_m(\cdot + h_n) - \xi(b_m)\big]\psi\|_{\mathcal{M}_{X(\mathbb{R})}}. \end{aligned} \tag{5.7}$$

Fix $\varepsilon > 0$. By Theorem 2.5, $\|\psi\|_{\mathcal{M}_{X(\mathbb{R})}} < \infty$. It follows from (5.2) that there exists a sufficiently large number $m \in \mathbb{N}$ (which we fix until the end of the proof) such that

$$2\|b - b_m\|_{\mathcal{M}_{X(\mathbb{R})}}\|\psi\|_{\mathcal{M}_{X(\mathbb{R})}} < \varepsilon/2. \tag{5.8}$$

Let

$$\Lambda := \begin{cases} \max\limits_{\lambda \in F_m \setminus \{\infty\}} |\lambda| & \text{if} \quad F_m \setminus \{\infty\} \neq \emptyset, \\ 0 & \text{if} \quad F_m \setminus \{\infty\} = \emptyset, \end{cases}$$

let $K := \operatorname{supp}\psi$ and

$$k := \max\{-\inf K, \sup K\} \in [0, \infty).$$

For $x \in K$ and $n \in \mathbb{N}$, let $I_n(x)$ be the segment with the endpoints h_n and $x + h_n$. Then $I_n(x) \subset [h_n - k, h_n + k]$. Since $h_n \to +\infty$ as $n \to \infty$, there exists $N_1 \in \mathbb{N}$ such that for all $n > N_1$, one has

$$I_n(x) \subset [h_n - k, h_n + k] \subset (\Lambda, \infty).$$

For all $n > N_1$, we have

$$\begin{aligned} \big\|\big[b_m(\cdot + h_n) - \xi(b_m)\big]\psi\big\|_{\mathcal{M}_{X(\mathbb{R})}} \leq &\big\|\big[b_m(\cdot + h_n) - b_m(h_n)\big]\psi\big\|_{\mathcal{M}_{X(\mathbb{R})}} \\ &+ \big|b_m(h_n) - \xi(b_m)\big|\|\psi\|_{\mathcal{M}_{X(\mathbb{R})}}, \end{aligned} \tag{5.9}$$

where the functions $\big[b_m(\cdot + h_n) - b_m(h_n)\big]\psi$ for $n > N_1$ belong to SO^3_∞ because they are three times continuously differentiable functions of compact support.

By (5.5), there exists $N_2 \in \mathbb{N}$ such that $N_2 \geq N_1$ and for all $n > N_2$,

$$\big|b_m(h_n) - \xi(b_m)\big|\|\psi\|_{\mathcal{M}_{X(\mathbb{R})}} < \varepsilon/4. \tag{5.10}$$

On the other hand, since $\big[b_m(\cdot + h_n) - b_m(h_n)\big]\psi \in SO^3_\infty$ for all $n > N_2$, it follows from Theorem 2.7 that there exists a constant $c_X > 0$ depending only on the space

$X(\mathbb{R})$ such that for all $n > N_2$,

$$\begin{aligned}
&\left\|\left[b_m(\cdot + h_n) - b_m(h_n)\right]\psi\right\|_{\mathcal{M}_{X(\mathbb{R})}} \\
&\quad \le c_X \left\|\left[b_m(\cdot + h_n) - b_m(h_n)\right]\psi\right\|_{SO^3_\infty} \\
&\quad = c_X \sum_{j=0}^{3} \frac{1}{j!} \left\| D^j_\infty \left(\left[b_m(\cdot + h_n) - b_m(h_n)\right]\psi\right)\right\|_{L^\infty(\mathbb{R})}. \qquad (5.11)
\end{aligned}$$

For all $j \in \{0, 1, 2, 3\}$, we have

$$\begin{aligned}
&D^j_\infty \left(\left[b_m(\cdot + h_n) - b_m(h_n)\right]\psi\right) \\
&\quad = \sum_{\nu=0}^{j} \binom{j}{\nu} \left(D^\nu_\infty\left[b_m(\cdot + h_n) - b_m(h_n)\right]\right) \left(D^{j-\nu}_\infty \psi\right). \qquad (5.12)
\end{aligned}$$

It follows from the mean value theorem that

$$\begin{aligned}
\left\|\left[b_m(\cdot + h_n) - b_m(h_n)\right]\chi_K\right\|_{L^\infty(\mathbb{R})} &= \sup_{x\in K} \left| \int_{h_n}^{x+h_n} b'_m(t)\, dt \right| \\
&= \sup_{x\in K} \left| \int_{h_n}^{x+h_n} t b'_m(t)\, \frac{dt}{t} \right| \\
&\le \sup_{x\in K} \int_{I_n(x)} \left|(D_\infty b_m)(t)\right| \frac{dt}{t} \\
&\le \sup_{t\in[h_n-k,h_n+k]} \left|(D_\infty b_m)(t)\right| \int_{h_n-k}^{h_n+k} \frac{dt}{t} \\
&\le \ln \frac{h_n + k}{h_n - k} \left\| D_\infty b_m \right\|_{L^\infty(\mathbb{R})}. \qquad (5.13)
\end{aligned}$$

It is easy to see that for $x \in K$,

$$\left(D_\infty\left[b_m(\cdot + h_n) - b_m(h_n)\right]\right)(x) = \frac{x}{x + h_n}(D_\infty b_m)(x + h_n), \qquad (5.14)$$

$$\begin{aligned}
&\left(D^2_\infty\left[b_m(\cdot + h_n) - b_m(h_n)\right]\right)(x) \\
&\quad = \frac{x^2}{(x + h_n)^2}(D^2_\infty b_m)(x + h_n) + \frac{x h_n}{(x + h_n)^2}(D_\infty b_m)(x + h_n), \qquad (5.15)
\end{aligned}$$

and

$$\begin{aligned}\big(D_\infty^3\big[b_m(\cdot+h_n)-b_m(h_n)\big]\big)(x)&\\ =\frac{x^3}{(x+h_n)^3}(D_\infty^3 b_m)(x+h_n)+\frac{3x^2h_n}{(x+h_n)^3}(D_\infty^2 b_m)(x+h_n)&\\ +\frac{xh_n^2-x^2h_n}{(x+h_n)^3}(D_\infty b_m)(x+h_n).&\end{aligned}\tag{5.16}$$

It follows from (5.14)–(5.16) that for all $n > N_2$,

$$\big\|\big(D_\infty\big[b_m(\cdot+h_n)-b_m(h_n)\big]\big)\chi_K\big\|_{L^\infty(\mathbb{R})}\le\frac{k}{h_n-k}\|D_\infty b_m\|_{L^\infty(\mathbb{R})},\tag{5.17}$$

$$\begin{aligned}&\Big\|\big(D_\infty^2\big[b_m(\cdot+h_n)-b_m(h_n)\big]\big)\chi_K\Big\|_{L^\infty(\mathbb{R})}\\ &\quad\le\frac{k^2}{(h_n-k)^2}\left\|D_\infty^2 b_m\right\|_{L^\infty(\mathbb{R})}+\frac{kh_n}{(h_n-k)^2}\|D_\infty b_m\|_{L^\infty(\mathbb{R})},\end{aligned}\tag{5.18}$$

and

$$\begin{aligned}&\Big\|\big(D_\infty^3\big[b_m(\cdot+h_n)-b_m(h_n)\big]\big)\chi_K\Big\|_{L^\infty(\mathbb{R})}\\ &\quad\le\frac{k^3}{(h_n-k)^3}\left\|D_\infty^3 b_m\right\|_{L^\infty(\mathbb{R})}+\frac{3k^2h_n}{(h_n-k)^3}\left\|D_\infty^2 b_m\right\|_{L^\infty(\mathbb{R})}\\ &\qquad+\frac{kh_n^2+k^2h_n}{(h_n-k)^3}\|D_\infty b_m\|_{L^\infty(\mathbb{R})}.\end{aligned}\tag{5.19}$$

Since

$$\max_{j\in\{0,1,2,3\}}\left\|D_\infty^j\psi\right\|_{L^\infty(\mathbb{R})}<\infty,$$

it follows from (5.12)–(5.13) and (5.17)–(5.19) that for all $j \in \{0, 1, 2, 3\}$,

$$\lim_{n\to\infty}\left\|D_\infty^j\big(\big[b_m(\cdot+h_n)-b_m(h_n)\big]\psi\big)\right\|_{L^\infty(\mathbb{R})}=0.\tag{5.20}$$

We deduce from (5.11) and (5.20) that there exists $N_3 \in \mathbb{N}$ such that $N_3 \ge N_2$ and for all $n > N_3$,

$$\big\|\big[b_m(\cdot+h_n)-b_m(h_n)\big]\psi\big\|_{\mathcal{M}_{X(\mathbb{R})}}<\varepsilon/4.\tag{5.21}$$

Combining (5.7)–(5.10) and (5.21), we see that for every $f \in \mathcal{S}_0(\mathbb{R})$ and every $\varepsilon > 0$ there exists $N_3 \in \mathbb{N}$ such that for all $n > N_3$,

$$\left\| \left(e_{h_n} W^0(b) e_{h_n}^{-1} - b(\xi) \right) f \right\|_{X(\mathbb{R})} < \varepsilon \| f \|_{X(\mathbb{R})},$$

whence for all $f \in \mathcal{S}_0(\mathbb{R})$,

$$\lim_{n\to\infty} \left\| \left(e_{h_n} W^0(b) e_{h_n}^{-1} I - b(\xi) I \right) f \right\|_{X(\mathbb{R})} = 0.$$

Since $\mathcal{S}_0(\mathbb{R})$ is dense in $X(\mathbb{R})$ (see Theorem 2.2), this equality immediately implies (5.1) in view of [28, Lemma 1.4.1(ii)], which completes the proof. □

5.3 *Proof of Theorem 1.1*

Since the function $e_0 \equiv 1$ belongs to Φ and $\Psi_{SO^\diamond_{X(\mathbb{R})}}$, we see that the set of all constant functions is contained in Φ and in $SO^\diamond_{X(\mathbb{R})}$. Therefore

$$\mathcal{MO}^\pi(\mathbb{C}) \subset \mathcal{MO}^\pi(\Phi) \cap \mathcal{CO}^\pi(SO^\diamond_{X(\mathbb{R})}). \tag{5.22}$$

Let $A^\pi \in \mathcal{MO}^\pi(\Phi) \cap \mathcal{CO}^\pi(SO^\diamond_{X(\mathbb{R})})$. Then $A^\pi = [aI]^\pi = [W^0(b)]^\pi$, where $a \in \Phi$ and $b \in SO^\diamond_{X(\mathbb{R})}$. Therefore, there is an operator $K \in \mathcal{K}(X(\mathbb{R}))$ such that

$$aI = W^0(b) + K. \tag{5.23}$$

By Theorem 5.2, for every $\xi \in M_\infty(SO^\diamond)$ there exists a sequence $\{h_n\}_{n\in\mathbb{N}}$ of positive numbers such that $h_n \to +\infty$ as $n \to \infty$ and

$$\operatorname*{s-lim}_{n\to\infty} e_{h_n} W^0(b) e_{h_n}^{-1} I = b(\xi) I. \tag{5.24}$$

Equalities (5.23)–(5.24) and Lemma 5.1 imply that

$$aI = \operatorname*{s-lim}_{n\to\infty} e_{h_n} (aI) e_{h_n}^{-1} I = \operatorname*{s-lim}_{n\to\infty} e_{h_n} (W^0(b) + K) e_{h_n}^{-1} I = b(\xi) I.$$

Hence $[aI]^\pi = [b(\xi)I]^\pi \in \mathcal{MO}^\pi(\mathbb{C})$ and

$$\mathcal{MO}^\pi(\Phi) \cap \mathcal{CO}^\pi(SO^\diamond_{X(\mathbb{R})}) \subset \mathcal{MO}^\pi(\mathbb{C}). \tag{5.25}$$

Combining (5.22) and (5.25), we arrive at (1.4). □

Acknowledgments We would like to thank the anonymous referee for pointing out a gap in the original version of the paper. To fill in this gap, we strengthened the hypotheses in the main result.

References

1. J. Anderson, Extensions, restrictions, and representations of states on C^*-algebras. Trans. Am. Math. Soc. **249**, 303–329 (1979)
2. M.A. Bastos, A. Bravo, Y.I. Karlovich, Convolution type operators with symbols generated by slowly oscillating and piecewise continuous matrix functions. Oper. Theory Adv. Appl. **147**, 151–174 (2004)
3. M.A. Bastos, C.A. Fernandes, Y.I. Karlovich, C^*-algebras of integral operators with piecewise slowly oscillating coefficients and shifts acting freely. Integr. Equ. Oper. Theory **55**, 19–67 (2006)
4. C. Bennett, R. Sharpley, *Interpolation of Operators* (Academic Press, Boston, 1988)
5. A. Böttcher, Y.I. Karlovich, I.M. Spitkovsky, *Convolution Operators and Factorization of Almost Periodic Matrix Functions* (Birkhäuser, Basel, 2002)
6. D. Cruz-Uribe, A. Fiorenza, *Variable Lebesgue Spaces* (Birkhäuser/Springer, New York, 2013)
7. R.V. Duduchava, *Integral Equations with Fixed Singularities* (Teubner, Leipzig, 1979)
8. C.A. Fernandes, A.Y. Karlovich, Y.I. Karlovich, Noncompactness of Fourier convolution operators on Banach function spaces. Ann. Funct. Anal. AFA **10**, 553–561 (2019)
9. C.A. Fernandes, A.Y. Karlovich, Y.I. Karlovich, Algebra of convolution type operators with continuous data on Banach function spaces. Banach Center Publ. **119**, 157–171 (2019)
10. C.A. Fernandes, A.Y. Karlovich, Y.I. Karlovich, Fourier convolution operators with symbols equivalent to zero at infinity on Banach function spaces, in *Proceedings of ISAAC* (2019, to appear). arXiv:1909.13538 [math.FA]
11. G.B. Folland, *A Guide to Advanced Real Analysis* (The Mathematical Association of America, Washington, 2009)
12. I. Gohberg, N. Krupnik, *One-Dimensional Linear Singular Integral Equations*, vol. II (Birkhäuser, Basel, 1992)
13. L. Grafakos, *Classical Fourier Analysis*, 3rd ed. (Springer, New York, 2014)
14. H. Hudzik, R. Kumar, R. Kumar, Matrix multiplication operators on Banach function spaces. Proc. Indian Acad. Sci. Math. Sci. **116**, 71–81 (2006)
15. R.V. Kadison, J.R. Ringrose, *Fundamentals of the Theory of Operator Algebras*, in *Elementary Theory*, vol. I, 2nd ed. (American Mathematical Society, Providence, 1997)
16. E. Kaniuth, *A Course in Commutative Banach Algebras* (Springer, New York, 2009)
17. A.Y. Karlovich, Maximally modulated singular integral operators and their applications to pseudodifferential operators on Banach function spaces. Contemp. Math. **645**, 165–178 (2015)
18. A.Y. Karlovich, Banach algebra of the Fourier multipliers on weighted Banach function spaces. Concr. Oper. **2**, 27–36 (2015)
19. A.Y. Karlovich, Commutators of convolution type operators on some Banach function spaces. Ann. Funct. Anal. AFA **6**, 191–205 (2015)
20. A.Y. Karlovich, The Stechkin inequality for Fourier multipliers on variable Lebesgue spaces. Math. Inequal. Appl. **18**, 1473–1481 (2015)
21. A. Karlovich, E. Shargorodsky, When does the norm of a Fourier multiplier dominate its L^∞ norm? Proc. London Math. Soc. **118**, 901–941 (2019)
22. Y.I. Karlovich, Algebras of convolution-type operators with piecewise slowly oscillating data on weighted Lebesgue spaces. Mediterr. J. Math. **14**, paper no. 182, 20 (2017)
23. A.Y. Karlovich, I.M. Spitkovsky, The Cauchy singular integral operator on weighted variable Lebesgue spaces. Oper. Theory Adv. Appl. **236**, 275–291 (2014)

24. Y.I. Karlovich, I. Loreto Hernández, Algebras of convolution type operators with piecewise slowly oscillating data. I: Local and structural study. Integr. Equ. Oper. Theory **74**, 377–415 (2012)
25. Y.I. Karlovich, I. Loreto Hernández, On convolution type operators with piecewise slowly oscillating data. Oper. Theory Adv. Appl. **228**, 185–207 (2013)
26. V. Rabinovich, S. Roch, B. Silbermann, *Limit Operators and Their Applications in Operator Theory* (Birkhäuser, Basel, 2004)
27. M. Reed, B. Simon, *Methods of Modern Mathematical Physics. I: Functional Analysis* (Academic Press, New York, 1980)
28. S. Roch, P.A. Santos, B. Silbermann, *Non-Commutative Gelfand Theories. A Tool-kit for Operator Theorists and Numerical Analysts* (Springer, Berlin, 2011)
29. I.B. Simonenko, *Local Method in the Theory of Shift Invariant Operators and Their Envelopes* (Rostov University Press, Rostov on Don, 2007, in Russian)
30. I.B. Simonenko, C.N. Min, *Local Method in the Theory of One-Dimensional Singular Integral Equations with Piecewise Continuous Coefficients. Noetherity* (Rostov University Press, Rostov on Don, 1986, in Russian)

Inner Outer Factorization of Wide Rational Matrix Valued Functions on the Half Plane

A. E. Frazho and A. C. M. Ran

Abstract The main purpose of this note is to use operator methods to solve a rational inner-outer factorization problem for wide functions. It is believed that this will provide valuable insight into the inner-outer factorization problem. Our approach involves Wiener–Hopf operators, Hankel operators and invariant subspaces for the backward shift. It should be emphasized that the formulas for the inner and outer factor are derived in a computational manner.

Keywords Inner-outer factorization · Matrix valued function · Wiener–Hopf operators · State space representation

Mathematics Subject Classification (2010) Primary 47B35, 47A68; Secondary 30J99

1 Introduction

In this note we will use operator techniques to develop a method to compute the inner-outer factorization for certain matrix valued rational functions defined on the closure of the right half plane. We shall focus on the "wide" case, i.e., the case where the matrix function has more columns than rows (or an equal number, thereby also including the square case). The "tall" case, where the matrix function has more

A. E. Frazho
Department of Aeronautics and Astronautics, Purdue University, West Lafayette, IN, USA
e-mail: frazho@ecn.purdue.edu

A. C. M. Ran (✉)
Department of Mathematics, Faculty of Science, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

Research Focus: Pure and Applied Analytics, North-West University, Potchefstroom, South Africa
e-mail: a.c.m.ran@vu.nl

M. A. Bastos et al. (eds.), *Operator Theory, Functional Analysis and Applications*,
Operator Theory: Advances and Applications 282,
https://doi.org/10.1007/978-3-030-51945-2_11

rows than columns (or an equal number, again including the square case) is well understood, and presented in, e.g., [3, 6, 18] and elsewhere.

It should be emphasized that we obtain explicit formulas for the inner and outer factors in terms of the matrices appearing in a state space realization of the original rational matrix valued function. Finally, we shall always assume that the rational matrix function is proper, i.e., it has a finite value at infinity.

To set the stage, let $\mathcal{E}, \mathcal{U}$ and $\mathcal{Y}$ be finite dimensional, complex vector spaces and $\dim \mathcal{Y} \leq \dim \mathcal{U}$. The Hilbert space of all Lebesgue measurable square integrable functions over $[0, \infty)$ with values in $\mathcal{E}$ is denoted by $L^2_+(\mathcal{E})$. Throughout $H^\infty(\mathcal{U}, \mathcal{E})$ is the set of all analytic functions G in the open right hand plane $\{s : \Re(s) > 0\}$ such that

$$\|G\|_\infty = \sup\{\|G(s)\| : \Re(s) > 0\} < \infty.$$

Recall that a function G_i is *inner* if G_i is a function in $H^\infty(\mathcal{E}, \mathcal{Y})$ and $G_i(i\omega)$ is almost everywhere an isometry. (In particular, $\dim \mathcal{E} \leq \dim \mathcal{Y}$.) Equivalently (see, e.g., [6, 18]), G_i in $H^\infty(\mathcal{E}, \mathcal{Y})$ is an inner function if and only if the Wiener–Hopf operator T_{G_i} mapping $L^2_+(\mathcal{E})$ into $L^2_+(\mathcal{Y})$ is an isometry. (The Wiener–Hopf operator is defined in (2.1) below.) A function G_o is *outer* if G_o is a function in $H^\infty(\mathcal{U}, \mathcal{E})$ and the range of the Wiener–Hopf operator T_{G_o} is dense in $L^2_+(\mathcal{E})$.

Let G be a function in $H^\infty(\mathcal{U}, \mathcal{Y})$. Then G admits a unique inner-outer factorization of the form $G(s) = G_i(s)G_o(s)$ where $G_i(s)$ is an inner function in $H^\infty(\mathcal{E}, \mathcal{Y})$ and $G_o(s)$ is an outer function in $H^\infty(\mathcal{U}, \mathcal{E})$ for some intermediate space $\mathcal{E}$ (see [14]). Because $G_i(e^{i\omega})$ is almost everywhere an isometry, $\dim \mathcal{E} \leq \dim \mathcal{Y}$. Since G_o is outer, $G_o(e^{i\omega})$ is almost everywhere onto $\mathcal{E}$, and thus, $\dim \mathcal{E} \leq \dim \mathcal{U}$. By unique we mean that if $G(s) = F_i(s)F_o(s)$ is another inner-outer factorization of G where F_i is an inner function in $H^\infty(\mathcal{L}, \mathcal{Y})$ and F_o is an outer function in $H^\infty(\mathcal{U}, \mathcal{L})$, then there exists a constant unitary operator V mapping $\mathcal{E}$ onto $\mathcal{L}$ such that $G_i = F_i V$ and $V G_o = F_o$; see [2, 6, 7, 18–20] for further details.

Throughout we assume that $\mathcal{U}$, $\mathcal{E}$ and $\mathcal{Y}$ are all finite dimensional. We say that G_i in $H^\infty(\mathcal{E}, \mathcal{Y})$ is a *square inner function* if $\mathcal{E}$ and $\mathcal{Y}$ have the same dimension. So if $G_i G_o$ is an inner-outer factorization of G where G_i is square, then without loss of generality we can assume that $\mathcal{E} = \mathcal{Y}$.

We say that the inner-outer factorization $G = G_i G_o$ is *full rank* if G_i is a square inner function in $H^\infty(\mathcal{Y}, \mathcal{Y})$ and the range of T_{G_o} equals $L^2_+(\mathcal{Y})$. An inner-outer factorization $G = G_i G_o$ is full rank if and only if G_i is a square inner function and the range of T_G is closed. Finally, if G in $H^\infty(\mathcal{U}, \mathcal{Y})$ admits a full rank inner-outer factorization, then $\dim \mathcal{Y} \leq \dim \mathcal{U}$.

Here we are interested in computing the inner-outer factorization for full rank rational functions G in $H^\infty(\mathcal{U}, \mathcal{Y})$. So throughout we assume that $\dim \mathcal{Y} \leq \dim \mathcal{U}$. Computing inner-outer factorizations when G does not admit a full rank factorization is numerically sensitive. This is already apparent in the scalar case. For example, take an outer function $G(s) = \frac{p(s)}{d(s)}$ with zeros on the imaginary axis. A slight perturbation of the coefficients of $p(s)$ could yield an inner part or even

an invertible outer function. To avoid this we simply assume that $G(s)$ is full rank. Moreover, from an operator perspective if $G(s)$ is not full rank, then one has to invert an unbounded Wiener–Hopf operator which is difficult. Furthermore, if G does not admit a full rank inner-outer factorization, then a small H^∞ perturbation of G does admit such a factorization. First we will present necessary and sufficient conditions to determine when G admits a full rank inner-outer factorization. Then we will give a state space algorithm to compute G_i and then G_o.

Inner-outer factorization for the "wide" case has received attention before. In several papers procedures to find the inner and outer factors have been derived. For instance, in [4] the factors were derived by considering $G^*G + \varepsilon^2 I$, and deducing the outer factor from this by solving a reduced Riccati equation. In [21] the reduced Riccati equation is replaced by a different approach, avoiding in fact an approximation procedure. In that paper, first the inner-outer factorization for $G(\overline{s})^*$ is computed, and then balanced coordinates are used to further factor the inner part of $G(\overline{s})^*$. Then they compute another inner-outer factorization of the remaining part. In [16], see also [13] a different approach is chosen, [16] uses state space methods along with some interesting state space decompositions to derive algorithms for the inner and outer factorization in a general rational setting. Their methods are algebraic in nature and quite different from our operator approach. Yet another approach is taken in [17]. In all of these papers state space formulas are given and their approach is based on finite dimensional techniques with properties of Riccati equations. A common method is to compute the outer factor first and then solve for the inner factor using the outer factor. Our approach is different: we derive a formula for the inner factor from the fact that the subspace $\ker T_G^*$ is shift invariant, and thus, by the Beurling–Lax–Halmos theorem (see[14]) there is an inner function G_i such that $\ker T_{G_i}^* = \ker T_G^*$. It is this inner function which we construct first, providing a state space formula for G_i, and then deduce the outer factor from $G_o = G_i^* G$ on the imaginary axis.

Finally, it is emphasized that the previous methods to compute the inner-outer factorization rely mainly on state space methods. Roughly speaking, the previous methods involve solving Riccati equations, Hamilton methods, decomposing the state space with balanced coordinates or other decompositions, or eigenvectors and eigenvalues for certain state space operators; see for instance [4, 16, 17, 21]. Our approach is quite different. We solve the inner-outer factorization problem by using operator techniques for the general formulas and then use state space methods to write down an explicit formula in the rational case. The operator methods give us valuable insight into the underlying framework. Once the foundation is constructed, then one simply uses state space techniques to formulate a realization of the inner and outer part.

2 Preliminaries

In this paper we shall make heavy use of the results and methods from [10]. Let Φ be a proper rational matrix function with values in $\mathfrak{L}(\mathcal{U}, \mathcal{Y})$ and no poles on the imaginary axis $i\mathbb{R}$. Let φ be the Lebesgue integrable (continuous) matrix function on the imaginary axis determined by Φ via the Fourier transform, that is,

$$\Phi(i\omega) = \Phi(\infty) + \int_{-\infty}^{\infty} e^{-i\omega t}\varphi(t)\,dt.$$

The *Wiener–Hopf operator* T_Φ associated with Φ and the *Hankel operator* H_Φ, both mapping $L^2_+(\mathcal{U})$ into $L^2_+(\mathcal{Y})$, are defined as follows for $t \geq 0$ and f in $L^2_+(\mathcal{U})$:

$$(T_\Phi f)(t) = \Phi(\infty)f(t) + \int_0^t \varphi(t-\tau)f(\tau)d\tau \tag{2.1}$$

$$(H_\Phi f)(t) = \int_0^\infty \varphi(t+\tau)f(\tau)d\tau. \tag{2.2}$$

In the sequel we shall freely use the basic theory of Wiener–Hopf and Hankel operators which can be found in Chapters XII and XIII of [11]. Note that in [11] the Fourier transform is taken with respect to the real line instead of the imaginary axis as is done here.

Now let G be the stable rational function in $H^\infty(\mathcal{U}, \mathcal{Y})$ given by the following state space realization:

$$G(s) = D + C(sI - A)^{-1}B. \tag{2.3}$$

Here A is stable operator acting on a finite dimensional space $\mathcal{X}$ and B maps $\mathcal{U}$ into $\mathcal{X}$, while C maps $\mathcal{X}$ into $\mathcal{Y}$ and D maps $\mathcal{U}$ into $\mathcal{Y}$. (By stable we mean that all the eigenvalues of A are contained in the open left hand half-plane of $\mathbb{C}$.) The realization for G in (2.3) is denoted by $\{A, B, C, D\}$. It is noted that $G(s)$ is also given by the following Laplace transform:

$$G(s) = D + \int_0^\infty e^{-st}Ce^{At}Bdt \qquad (\mathrm{Re}\, s \geq 0).$$

Hence the corresponding Wiener–Hopf operator T_G and the Hankel operator H_G can be expressed in terms of the matrices appearing in the realization of G as follows

$$(T_G f)(t) = Df(t) + \int_0^t Ce^{A(t-\tau)}Bf(\tau)d\tau \qquad (\text{for } t \geq 0) \tag{2.4}$$

$$(H_G f)(t) = \int_0^\infty Ce^{A(t+\tau)}Bf(\tau)d\tau \qquad (\text{for } t \geq 0). \tag{2.5}$$

Throughout this note we assume that G is full row rank and DD^* is invertible. A realization is called *controllable* if $\vee_{j=0}^{\infty}\mathrm{Im}\, A^jB$ equals the state space $\mathcal{X}$, and a realization is called *observable* if $\cap_{j=0}^{\infty}\ker CA^j = \{0\}$. It is well known from systems theory that the dimension of the state space $\mathcal{X}$ is as small as possible (over all possible realizations) if and only if the realization is both observable and controllable. Such a realization is called minimal, and it will be assumed throughout that the realization (2.3) for G is minimal.

With G we also associate the rational matrix function R given by $R(s) = G(s)G(-\overline{s})^*$. Notice that R is a proper rational matrix function with values in $\mathfrak{L}(\mathcal{Y},\mathcal{Y})$ (the set of bounded linear operators mapping $\mathcal{Y}$ to itself) and has no poles on the imaginary axis. By T_R we denote the corresponding Wiener–Hopf operator acting on $L^2_+(\mathcal{Y})$. It is well-known (see, e.g., formula (24) in Section XII.2 of [11]) that

$$T_R = T_G T_G^* + H_G H_G^*. \tag{2.6}$$

Let P be the controllability Grammian for the pair $\{A, B\}$. In other words,

$$P = \int_0^\infty e^{At}BB^*e^{A^*t}\,dt.$$

Moreover, P is the unique solution to the Lyapunov equation

$$AP + PA^* + BB^* = 0. \tag{2.7}$$

Because the pair $\{A, B\}$ is controllable P is strictly positive. Let W_c be the controllability operator mapping $L^2_+(\mathcal{U})$ into the state space $\mathcal{X}$ defined by

$$W_c u = \int_0^\infty e^{At}Bu(t)\,dt \qquad (u \in L^2_+(\mathcal{U})).$$

It is noted that the controllability Grammian is given by $P = W_cW_c^*$.

Let W_o be the observability operator mapping $\mathcal{X}$ into $L^2_+(\mathcal{Y})$ corresponding to the pair $\{C, A\}$ defined by

$$W_o x = Ce^{At}x \qquad (x \in \mathcal{X}).$$

Because the pair $\{C, A\}$ is observable, W_o is one to one.

Let us introduce the operator Γ mapping $\mathcal{Y}$ into $\mathcal{X}$ defined by

$$\Gamma = BD^* + PC^*.$$

Now consider the algebraic Riccati equation:

$$A^*Q + QA + (C - \Gamma^*Q)^*(DD^*)^{-1}(C - \Gamma^*Q) = 0. \tag{2.8}$$

Let A_o on $\mathcal{X}$ and C_o mapping $\mathcal{X}$ into $\mathcal{Y}$ be the operators defined by

$$A_o = A - \Gamma C_o \quad \text{and} \quad C_o = (DD^*)^{-1}(C - \Gamma^* Q). \tag{2.9}$$

We say that Q is a *stabilizing solution* to the algebraic Riccati equation (2.8) if Q is a positive solution to (2.8) and the operator A_o is stable. Finally, the stabilizing solution is unique.

Proposition 2.1 in [10] states that T_R is a strictly positive operator on $L^2_+(\mathcal{Y})$ if and only if there exists a stabilizing solution to the algebraic Riccati equation (2.8). Moreover, in this case, the stabilizing solution Q is given by

$$Q = W_o^* T_R^{-1} W_o;$$

see Proposition 2.2, equation (2.11) in [10]. Because $\{C, A\}$ is observable, W_o is one to one, and thus, Q is strictly positive. Formula (2.14) in [10] states that

$$T_R^{-1} W_o x = C_o e^{A_o t} x \qquad (x \in \mathcal{X}).$$

In other words, $T_R^{-1} W_o$ equals the observability operator $C_o e^{A_o t}$ mapping $\mathcal{X}$ into $L^2_+(\mathcal{Y})$ determined by the pair $\{C_o, A_o\}$. Since $T_R^{-1} W_o$ is one to one, the pair $\{C_o, A_o\}$ is observable.

The Hankel operator H_G can be written as $H_G = W_o W_c$; see (2.5). Using $P = W_c W_c^*$, we have $H_G H_G^* = W_o P W_o^*$. This, with Eq. (2.6) implies that

$$T_G T_G^* = T_R - W_o P W_o^*.$$

Hence $\ker T_G^* = \ker(T_R - W_o P W_o^*)$. By consulting Lemma 4.1 in [9], restated as Lemma 4.4 at the end of this paper for the readers convenience, we have

$$\dim \ker T_G^* = \dim \ker(T_R - W_o P W_o^*) = \dim \ker(I - QP).$$

Later we will see that the McMillan degree of the inner part G_i of G equals $\dim \ker(I - QP)$.

The next lemma characterizes the existence of a full rank inner-outer factorization. The proof of this lemma is essentially the same as that of Lemma 3.1 in [9], and is therefore omitted.

Lemma 2.1 *Let G be a rational function in $H^\infty(\mathcal{U}, \mathcal{Y})$ where $\mathcal{U}$ and $\mathcal{Y}$ are finite dimensional spaces satisfying $\dim \mathcal{Y} \leq \dim \mathcal{U}$. Then G admits a full rank inner-outer factorization if and only if*

$$G(i\omega)G(i\omega)^* \geq \epsilon I \qquad (\textit{for some } \epsilon > 0),$$

or equivalently, the Toeplitz operator T_R is strictly positive.

3 Inner Functions

First we consider realizations of rational inner functions. Recall that if G_i is a rational function with values in $\mathfrak{L}(\mathcal{Y}, \mathcal{Y})$, then G_i is an inner function if and only if G_i is a function $H^\infty(\mathcal{Y}, \mathcal{Y})$ and $G_i(i\omega)$ is a unitary operator for all ω.

Consider a rational function G_i in $H^\infty(\mathcal{Y}, \mathcal{Y})$. Let $\{A_i \text{ on } \mathcal{X}_i, B_i, C_i, D_i\}$ be a minimal realization for G_i. In particular,

$$G_i(s) = D_i + C_i(sI - A_i)^{-1}B_i.$$

Because G_i is a rational function in $H^\infty(\mathcal{Y}, \mathcal{Y})$, and the realization is minimal, the state space operator A_i is stable. Let Q_i be the observability Grammian for the pair $\{C_i, A_i\}$. In other words, Q_i is the unique solution to the Lyapunov equation

$$A_i^* Q_i + Q_i A_i + C_i^* C_i = 0. \tag{3.1}$$

Then G_i is an inner function in $H^\infty(\mathcal{Y}, \mathcal{Y})$ if and only if

$$D_i \text{ is a unitary operator and } Q_i B_i = -C_i^* D_i.$$

(The state space formula for an inner function is classical; for example, see [1, 3, 21] and also Theorem 19.15 in [5].) In this case, Q_i^{-1} is the controllability Grammian for the pair $\{A_i, B_i\}$, that is, Q_i^{-1} is the unique solution to the Lyapunov equation

$$A_i Q_i^{-1} + Q_i^{-1} A_i^* + B_i B_i^* = 0. \tag{3.2}$$

Let W_i mapping $\mathcal{X}_i$ into $L^2_+(\mathcal{Y})$ be the *observability operator determined by* $\{C_i, A_i \text{ on } \mathcal{X}_i\}$ defined by $W_i x = C_i e^{A_i t} x$ for x in $\mathcal{X}_i$.

Next, we introduce some notation and present some well-known results. If Θ is an inner function in $H^\infty(\mathcal{E}, \mathcal{Y})$, then $\mathfrak{H}(\Theta)$ is the subspace of $L^2_+(\mathcal{Y})$ defined by

$$\mathfrak{H}(\Theta) = L^2_+(\mathcal{Y}) \ominus T_\Theta L^2_+(\mathcal{E}) = \ker T^*_\Theta.$$

Let S^*_τ be the backward shift on $L^2_+(\mathcal{Y})$ for $\tau \geq 0$, that is, $(S^*_\tau f)(t) = f(t + \tau)$ for f in $L^2_+(\mathcal{Y})$. It is noted that $\mathfrak{H}(\Theta)$ is an invariant subspace for the backward shift S^*_τ for all $\tau \geq 0$. According to the Beurling–Lax–Halmos Theorem if $\mathfrak{H}$ is any invariant subspace for the backward shift S^*_τ for all $\tau \geq 0$, then there exists a unique inner function Θ in $H^\infty(\mathcal{E}, \mathcal{Y})$ such that $\mathfrak{H} = \mathfrak{H}(\Theta)$. By unique we mean that if $\mathfrak{H} = \mathfrak{H}(\Psi)$ where Ψ is an inner function in $H^\infty(\mathcal{L}, \mathcal{Y})$, then there exists a constant unitary operator V from $\mathcal{E}$ onto $\mathcal{L}$ such that $\Theta = \Psi V$; see [2], Section 3.9 for the half plane case we use here, see also [6, 12, 15, 18–20] for further details in the unit disc case.

Now we present the following classical result; see Theorem 7.1 in [8], Sections 4.2 and 4.3 in [6] and Section XXVIII.7 in [12] for the disc case, and [2], Lemma 3.45 and Lemma 3.46 for the half plane case.

Lemma 3.1 *Let Θ be an inner function in $H^\infty(\mathcal{Y}, \mathcal{Y})$ where $\mathcal{Y}$ is finite dimensional. Then the Hankel operator H_Θ is a contraction. Moreover,*

$$\mathfrak{H}(\Theta) = \overline{\operatorname{Im} H_\Theta} = \ker T_\Theta^*.$$

Furthermore, the following holds.

(i) *The subspace $\mathfrak{H}(\Theta)$ is finite dimensional if and only if Θ is rational.*
(ii) *The dimension of $\mathfrak{H}(\Theta)$ equals the McMillan degree of Θ.*
(iii) *Let $\{A_i \text{ on } \mathcal{X}_i, B_i, C_i, D_i\}$ be a minimal realization for a rational inner function Θ. Then $\mathfrak{H}(\Theta)$ equals the range of the observability operator W_i mapping $\mathcal{X}_i$ into $L^2_+(\mathcal{Y})$.*

Let $\{C_i, A_i \text{ on } \mathcal{X}_i\}$ be a finite dimensional stable observable pair. Then we can compute operators B_i mapping $\mathcal{Y}$ into $\mathcal{X}_i$ and D_i on $\mathcal{Y}$ such that $\{A_i, B_i, C_i, D_i\}$ is a minimal realization for an inner function Θ. In this case, $\mathfrak{H}(\Theta)$ equals the range of W_i, the observability operator determined by $\{C_i, A_i\}$. To accomplish this, first compute the solution Q_i to the observability Lyapunov equation in (3.1). Note that Q_i is automatically strictly positive because A_i is stable and $\{C_i, A_i\}$ is observable. Then compute B_i from the Lyapunov equation (3.2). Now use $Q_i B_i B_i^* Q_i = C_i^* C_i$ to find a unitary D_i with $Q_i B_i = -C_i^* D_i$. We shall call $\{B_i, D_i\}$ the *complementary operators for* $\{C_i, A_i\}$. According to Lemma 3.1, the subspace $\mathfrak{H}(\Theta) = \operatorname{Im} W_i = \ker T_\Theta^*$. Finally, the complementary operators $\{B_i, D_i\}$ are not unique. However, the corresponding subspace $\mathfrak{H}(\Theta) = \operatorname{Im} W_i$ they determine is unique.

4 Main Result

Let us combine several results of [10] into one theorem, which is the right half plane analogue of Theorem 3.2 in [9] in the open unit disc case.

Theorem 4.1 *Let $\{A \text{ on } \mathcal{X}, B, C, D\}$ be a minimal realization for a rational function G in $H^\infty(\mathcal{U}, \mathcal{Y})$ where $\dim \mathcal{Y} \leq \dim \mathcal{U}$. Let R be the function in $L^\infty(\mathcal{Y}, \mathcal{Y})$ defined by $R(i\omega) = G(i\omega)G(i\omega)^*$. Let P the controllability Grammian for the pair $\{A, B\}$; see (2.7). Then the following statements are equivalent.*

(i) *The function G admits a full rank inner-outer factorization;*
(ii) *The Wiener–Hopf operator T_R is strictly positive.*
(iii) *There exists a (unique) stabilizing solution Q to the algebraic Riccati equation (2.8).*

In this case, the solution Q is given by $Q = W_o^* T_R^{-1} W_o$ and Q is strictly positive. Moreover, the following holds.

(iv) *The eigenvalues of QP are real numbers contained in the interval* $[0, 1]$.
(v) *If G_i is the inner factor of G, then the dimension of $\mathfrak{H}(G_i)$ is given by*

$$\dim \mathfrak{H}(G_i) = \dim \ker T^*_{G_i} = \dim \ker T^*_G = \dim \ker(I - QP),$$

and the McMillan degree of G_i is given by

$$\dim \mathfrak{H}(G_i) = \dim \ker(I - QP).$$

In particular, the McMillan degree of G_i is less than or equal to the McMillan degree of G.
(vi) *The operator $T_R^{-1} W_o$ mapping $\mathcal{X}$ into $L^2_+(\mathcal{Y})$ is equal to the observability operator generated by $\{C_o, A_o\}$, that is,*

$$T_R^{-1} W_o x = C_o e^{A_o t} x \qquad (x \in \mathcal{X}). \tag{4.1}$$

Finally, because $\{C, A\}$ is observable, $T_R^{-1} W_o$ is one to one and $\{C_o, A_o\}$ is a stable observable pair.

The construction of the state space realization for the inner G_i and outer factor G_o from the realization of G is analogous to the approach presented in [9] for the disc case. The crucial observation is that if $G = G_i G_o$ is an inner-outer factorization, then $\mathfrak{H}(G_i) = \ker T^*_{G_i} = \ker T^*_G$ because T_{G_o} is onto. This observation can be used to find a realization for G_i, and then G_o is computed from $G_o = G_i^* G$ on the imaginary axis. Note that this is quite the other way around as in the tall case, where we first find the outer factor G_o from the spectral factorization $G^* G = G_o^* G_o$, and then compute G_i from $G_i = G G_o^{-1}$.

To describe a state space realization for the inner G_i and outer G_o factors of G, we first introduce an isometry U as in [9]. To be precise, let

$$k = \dim \ker(I - QP).$$

Let U be any isometry mapping $\mathbb{C}^k$ onto $\ker(I - QP)$. Note, such a U can be constructed explicitly from the singular value decomposition of $I - QP$. The analogue of Theorem 3.4 in [9] for the half plane now reads as follows.

Theorem 4.2 *Let $\{A, B, C, D\}$ be a minimal realization for a rational function G. Moreover, assume that the algebraic Riccati equation (2.8) has a stabilizing solution Q, and let A_o and C_o be as defined in (2.9). Let A_i on $\mathbb{C}^k$ and C_i mapping $\mathbb{C}^k$ into $\mathcal{Y}$ be the operators defined by*

$$A_i = U^* Q A_o P U \quad \text{and} \quad C_i = C_o P U.$$

Let B_i and D_i be complementary operators to $\{C_i, A_i\}$. Then

$$G_i(s) = D_i + C_i(sI - A_i)^{-1} B_i \tag{4.2}$$

is the inner factor for G. The outer factor G_o for G is given by

$$G_o(s) = D_i^* D + (D_i^* C + B_i^* U^*)(sI - A)^{-1} B.$$

Finally, G is an outer function if and only if $k = 0$. In this case, $G_i(s) = I$.

It is noted that the state space realization for the outer function G_o given in the Theorem 4.2 is not necessarily minimal.

The proof is analogous to the proof in the disc case. According to Lemma 4.1 in [9], restated for convenience as Lemma 4.4 at the end of this section, we have

$$\mathfrak{H}(G_i) = \ker T_{G_i^*} = \ker T_G^* = \ker(T_R - W_o P W_o^*).$$

This readily implies that

$$\dim \mathfrak{H}(G_i) = \dim \ker(I - QP) = k.$$

The following lemma is a crucial part of the proof. We shall give two proofs, one algebraic in nature, and a second operator theoretic proof using the fact that $\mathfrak{H}(G_i)$ is an invariant subspace for the backward shift.

Lemma 4.3 *Assume that the hypotheses of Theorem 4.2 hold. Then there exists a unique stable operator A_i on $\mathbb{C}^k$ such that $A_o PU = PUA_i$. Finally, $A_i = U^* Q A_o PU$.*

Algebraic Proof First we prove the existence of an operator A_i such that $QA_o PU = QPUA_i$. We claim that $\operatorname{Im} QA_o PU \subset \operatorname{Im} QPU = \operatorname{Im} U$. Since $\operatorname{Im} U$ equals $\ker(I - QP)$, we have $U = QPU$. Let us show that

$$(I - QP)QA_o PU = 0. \tag{4.3}$$

Formula (3.8) of [10] states that

$$A_o^*(Q - QPQ) + (Q - QPQ)A_o + C_1^* C_1 = 0$$

where $C_1 = D^* C_o + B^* Q$. (The formula for C_1 is not needed in our proof.) Using this we obtain

$$(I - QP)QA_o PU = -A_o^*(Q - QPQ)PU - C_1^* C_1 PU.$$

By employing $U = QPU$, we have $(I - QP)QA_o PU = -C_1^* C_1 PU$. Multiplying both sides by $U^* P$ and using

$U^* = U^* PQ$, we obtain

$$0 = U^* P(I - QP)QA_o PU = -U^* PC_1^* C_1 PU.$$

In other words, $C_1 PU = 0$. Hence (4.3) holds, and thus, the range of QA_0PU is contained in the kernel of $I - QP$. So there exists an operator A_i on $\mathbb{C}^k$ such that $QA_0PU = UA_i$. Multiplying by U^* on the left yields, $U^*QA_0PU = A_i$. In particular, A_i is uniquely determined by our choice of the isometry U.

Equation (4.3) also implies that

$$QA_oPU = QPQA_oPU = QPUA_i.$$

Because Q is invertible, we have

$$A_oPU = PUA_i.$$

Notice that $\mathcal{X}_1 = \operatorname{Im} PU$ is an invariant subspace for A_o. Since P is invertible, A_i is similar to the stable operator $A_o|\mathcal{X}_1$ on $\mathcal{X}_1$. Because A_o is stable, A_i is also stable. □

An Operator Theoretic Proof of Lemma 4.3 By Lemma 4.4 below with (4.1), we see that $T_R^{-1}W_oPU$ is the operator from $\mathbb{C}^k$ onto $\mathfrak{H}(G_i)$ given by

$$T_R^{-1}W_oPUx = C_oe^{A_0t}PUx \qquad (x \in \mathbb{C}^k)$$
$$\mathfrak{H}(G_i) = \operatorname{Im}\left(T_R^{-1}W_oPU\right). \tag{4.4}$$

Since $\mathfrak{H}(G_i)$ is an invariant subspace of the backward shift on $L^2_+(\mathcal{Y})$ (see [2, Chapter 3]), we have for all $t > 0$ and $t_1 > 0$ and all $x \in \mathbb{C}^k$

$$C_oe^{A_o(t+t_1)}PUx \in \mathfrak{H}(G_i).$$

This readily implies that $C_oe^{A_o(t+t_1)}PUx = C_oe^{A_ot}PUF(t_1)x$ for some linear operator $F(t_1)$ on $\mathbb{C}^k$. One easily checks that F satisfies the semigroup property $F(t_1 + t_2) = F(t_1)F(t_2)$. Hence there exists an operator A_i on $\mathbb{C}^k$ such that $F(t) = e^{A_it}$. In other words,

$$C_oe^{A_ot}e^{A_ot_1}PU = C_oe^{A_o(t+t_1)}PU = C_oe^{A_ot}PUe^{A_it_1}.$$

Since the observability operator $C_oe^{A_ot}$ is one to one, we have $e^{A_ot}PU = PUe^{A_it}$ (for all t). It is now clear that A_i must be stable. To see this simply note that P is invertible, U is an isometry, and A_o is stable. Therefore e^{A_it} converges to zero as t tends to infinity, and thus, A_i is stable. By taking the derivative and then letting t approach 0, we see that $A_oPU = PUA_i$ as desired. Multiplying both sides by U^*Q with $U = QPU$, we obtain $A_i = U^*QA_oPU$. □

Proof of Theorem 4.2 Note that using the definition of $C_i = C_oPU$ and A_i we have

$$C_oe^{A_ot}PU = C_ie^{A_it}. \tag{4.5}$$

Since U is an isometry, P is invertible and $\{C_o, A_o\}$ is an observable pair it follows that $\{C_i, A_i\}$, is an observable pair. Because $T_R^{-1} W_o P U$ is onto $\mathfrak{H}(G_i)$, the subspace $\{C_i e^{A_i t} x \mid x \in \mathbb{C}^k\}$ equals $\mathfrak{H}(G_i)$; see Eqs. (4) and (4.5).

Now let Q_i be the solution to

$$A_i^* Q_i + Q_i A_i + C_i^* C_i = 0,$$

and let B_i and D_i be the complementary operators for the pair $\{C_i, A_i\}$. Finally, $\{A_i, B_i, C_i, D_i\}$ is the state space realization for our inner function $G_i(s)$; see (4.2). By construction $\mathfrak{H}(G_i) = \ker T_{G_i}^*$.

By a direct calculation or consulting formula (2.12) in [10], we also have

$$A_o^* Q + Q A + C_o^* C = 0. \tag{4.6}$$

The formula for the outer factor $G_o(s)$ can now be derived in the same way as the last section of the paper [9]. Indeed, for $s \in i\mathbb{R}$, we have

$$G_i(s)^* = D_i^* - B_i^*(sI + A_i^*)^{-1} C_i^*.$$

This yields

$$\begin{aligned} G_i(s)^* G(s) = &\Big(D_i^* - B_i^*(sI + A_i^*)^{-1} C_i^*\Big)\Big(D + C(sI - A)^{-1} B\Big) \\ = &D_i^* D - B_i^*(sI + A_i^*)^{-1} C_i^* D + D_i^* C(sI - A)^{-1} B \\ &- B_i^*(sI + A_i^*)^{-1} C_i^* C(sI - A)^{-1} B. \end{aligned}$$

Using $C_i = C_o P U$ with (4.6), we obtain

$$C_i^* C = U^* P C_o^* C = -U^* P(A_o^* Q + Q A) = -U^* P A_o^* Q - U^* P Q A.$$

By Lemma 4.3 we have $U^* P A_o^* = A_i^* U^* P$. This and $U^* P Q = U^*$ yields

$$C_i^* C = -A_i^* U^* P Q - U^* P Q A = -A_i^* U^* - U^* A. \tag{4.7}$$

Rewrite this as

$$C_i^* C = -(sI + A_i^*) U^* + U^*(sI - A).$$

By taking the appropriate inverses, we obtain

$$(sI + A_i^*)^{-1} C_i^* C (sI - A)^{-1} = -U^*(sI - A)^{-1} + (sI + A_i^*)^{-1} U^*.$$

This readily implies that

$$\begin{aligned} G_i(s)^*G(s) &= \Big(D_i^* - B_i^*(sI + A_i^*)^{-1}C_i^*\Big)\Big(D + C(sI - A)^{-1}B\Big) \\ &= D_i^*D - B_i^*(sI + A_i^*)^{-1}C_i^*D + D_i^*C(sI - A)^{-1}B \\ &\quad + B_i^*U^*(sI - A)^{-1}B - B_i^*(sI + A_i^*)^{-1}U^*B \\ &= D_i^*D + (D_i^*C + B_i^*U^*)(sI - A)^{-1}B \\ &\quad - B_i^*(sI + A_i^*)^{-1}(U^*B + C_i^*D). \end{aligned}$$

Now observe that

$$U^*B + C_i^*D = U^*PQB + U^*PC_o^*D = U^*P(QB + C_o^*D) = U^*PC_1^*,$$

where $C_1 = B^*Q + D^*C_o$; see also formula (3.3) in [10]. Now use equation (3.8) in [10], which states that

$$A_o^*(Q - QPQ) + (Q - QPQ)A_o + C_1^*C_1 = 0.$$

Multiplying by U^*P on the left and by PU on the right, with $QPU = U$, yields

$$U^*PC_1^*C_1PU = 0 \quad \text{and thus} \quad U^*PC_1^* = 0.$$

In other words, $U^*B + C_i^*D = 0$. It follows that for $s \in i\mathbb{R}$, we have

$$G_o(s) = G_i(s)^*G(s) = D_i^*D + (D_i^*C + B_i^*U^*)(sI - A)^{-1}B.$$

Clearly, this holds for all s except the eigenvalues for A. □

For the tall case we refer to Theorem 17.26 in [3].

Note that C_i can be computed a bit more explicitly:

$$C_i = C_oPU = (DD^*)^{-1}(C - \Gamma^*Q)PU,$$

where $\Gamma^* = CP + DB^*$. This with $U = QPU$, yields

$$(C - \Gamma^*Q)PU = (C - CPQ - DB^*Q)PU = -DB^*U.$$

In other words, $C_i = -(DD^*)^{-1}DB^*U$.

Recall that $C_i^*C = -A_i^*U^* - U^*A$; see (4.7). Multiplying by U on the right, we obtain $C_i^*CU = -A_i^* - U^*AU$. In other words, $A_i = -U^*C^*C_i - U^*A^*U$. Using $C_i = -(DD^*)^{-1}DB^*U$, we have

$$\begin{aligned} A_i &= -U^*A^*U + U^*C^*(DD^*)^{-1}DB^*U \\ &= -U^*\big(A - BD^*(DD^*)^{-1}C\big)^*U. \end{aligned}$$

This leads to an alterative formula for the inner factor

$$G_i(s) = D_i - (DD^*)^{-1}DB^*U\Big(sI + U^*\big(A - BD^*(DD^*)^{-1}C\big)^*U\Big)^{-1}B_i.$$

In turn, this may be rewritten as follows:

$$G_i(s) = D_i - (DD^*)^{-1}DB^*\Big(sI + UU^*\big(A - BD^*(DD^*)^{-1}C\big)^*\Big)^{-1}UB_i.$$

The latter formula compares well with the formula for the inner factor in the tall case as presented in [3],Theorem 17.26.

Finally, Lemma 4.1 in [9] is restated in the following Lemma.

Lemma 4.4 *Let T be a strictly positive operator on a Hilbert space $\mathcal{H}$ and P a strictly positive operator on a Hilbert space $\mathcal{X}$. Let W be an operator mapping $\mathcal{X}$ into $\mathcal{H}$ and set $Q = W^*T^{-1}W$. Then the following two assertions hold.*

(i) *Let $\mathfrak{X}$ and $\mathfrak{H}$ be the subspaces defined by*

$$\mathfrak{X} = \ker(I - QP) \quad \textit{and} \quad \mathfrak{H} = \ker\big(T - WPW^*\big).$$

Then the operators

$$\Lambda_1 = W^*|\mathfrak{H} : \mathfrak{H} \to \mathfrak{X} \quad \textit{and} \quad \Lambda_2 = T^{-1}WP|\mathfrak{X} : \mathfrak{X} \to \mathfrak{H}$$

are both well defined and invertible. Moreover, $\Lambda_1^{-1} = \Lambda_2$. In particular, $\mathfrak{X}$ and $\mathfrak{H}$ have the same dimension.

(ii) *The operator $T - WPW^*$ is positive if and only if $P^{-1} - Q$ is positive, or equivalently, $P^{\frac{1}{2}}QP^{\frac{1}{2}}$ is a contraction. In this case, the spectrum of QP is contained in* $[0, 1]$. *In particular, if $\mathcal{X}$ is finite dimensional, then the eigenvalues for QP are contained in* $[0, 1]$.

References

1. D. Alpay, I. Gohberg, Unitary rational matrix functions, in *Topics in Interpolation Theory of Rational Matrix-valued Functions*. Operator Theory: Advances and Applications, vol. 33 (Birkhäuser Verlag, Basel, 1988), pp. 175–222
2. D.Z. Arov, H. Dym, *J-Contractive Matrix Valued Functions and Related Topics* (Cambridge University Press, Cambridge, 2008)
3. H. Bart, I. Gohberg, M.A. Kaashoek, A.C.M. Ran, *A State Space Approach to Canonical Factorization: Convolution Equations and Mathematical Systems*. Operator Theory: Advances and Applications, vol. 200 (Birkhäuser Verlag, Basel, 2010)
4. T. Chen, B.A. Francis, Spectral and inner-outer-factorizations of rational matrices. SIAM J. Matrix Anal. Appl. **10**, 1–17 (1989)
5. H. Dym, *Linear Algebra in Action*. Graduate Studies in Mathematics, vol. 78 (American Mathematical Society, Providence, 2007)

6. A.E. Frazho, W. Bosri, *An Operator Perspective on Signals and Systems*. Operator Theory: Advances and Applications, vol. 204 (Birkhäuser Verlag, Basel, 2010)
7. C. Foias, A. Frazho, *The Commutant Lifting Approach to Interpolation Problems*. Operator Theory: Advances and Applications, vol. 44 (Birkhäuser Verlag, Basel, 1990)
8. C. Foias, A. Frazho, I. Gohberg, M.A. Kaashoek, *Metric Constrained Interpolation, Commutant Lifting and Systems*. Operator Theory: Advances and Applications, vol. 100 (Birkhäuser Verlag, Basel, 1998)
9. A.E. Frazho, A.C.M. Ran, A note on inner-outer factorization for wide matrix-valued functions, in: *Operator Theory, Analysis and the State Space Approach*. Operator Theory: Advances and Applications, vol. 271 (Birkhäuser Verlag, Basel, 2018), pp. 201–214
10. A.E. Frazho, M.A. Kaashoek, A.C.M. Ran, Rational matrix solutions of a Bezout type equation on the half plane, in: *Advances in Structured Operator Theory and Related Areas*. Operator Theory: Advances and Applications, vol. 237 (Birkhäuser Verlag, Basel, 2013), pp. 145–160
11. I. Gohberg, S. Goldberg, M.A. Kaashoek, *Classes of Linear Operators, Volume I*. Operator Theory: Advances and Applications, vol. 49 (Birkhäuser Verlag, Basel, 1990)
12. I. Gohberg, S. Goldberg, M.A. Kaashoek, *Classes of Linear Operators, Volume II*. Operator Theory: Advances and Applications, vol. 63 (Birkhäuser Verlag, Basel, 1993)
13. V. Ionescu, C. Oară, M. Weiss, *Generalized Riccati Theory and Robust Control, A Popov Function Approach* (Wiley, Chichester, 1999)
14. P.D. Lax, Translation invariant spaces. Acta Math. **101**, 163–178 (1959)
15. N.K. Nikol'skii, *Treatise on the Shift Operator*. Grundlehren, vol. 273 (Springer, Berlin, 1986)
16. C. Oară, A. Varga, Computation of general inner–outer and spectral factorizations. IEEE Trans. Autom. Control **45**, 2307–2325 (2000)
17. T. Reis, M. Voigt, Inner-outer factorization for differential-algebraic systems. Math. Control Signals Syst. **30**, Art. 15, 19pp. (2018)
18. M. Rosenblum, J. Rovnyak, *Hardy Classes and Operator Theory* (Oxford University Press, Oxford, 1985)
19. B. Sz.-Nagy, C. Foias, *Harmonic Analysis of Operators on Hilbert Space* (North-Holland, Amsterdam, 1970)
20. B. Sz.-Nagy, C. Foias, H. Bercovici, L. Kérchy, *Harmonic Analysis of Operators on Hilbert Space* (Springer, New York, 2010)
21. F.-B. Yeh, L.-F. Wei, Inner-outer factorizations of right-invertible real-rational matrices. Syst. Control Lett. **14**, 31–36 (1990)

Convergence Rates for Solutions of Inhomogeneous Ill-posed Problems in Banach Space with Sufficiently Smooth Data

Matthew A. Fury

Abstract We consider the inhomogeneous, ill-posed Cauchy problem

$$u'(t) = Au + h(t),\ 0 < t < T, \quad u(0) = \varphi$$

where $-A$ is the infinitesimal generator of a holomorphic semigroup of angle θ in Banach space. As in conventional regularization methods, certain auxiliary well-posed problems and their associated C_0 semigroups are applied in order to approximate a known solution u. A key property however, that the semigroups adhere to requisite growth orders, may fail depending on the value of the angle θ. Our results show that an approximation of u may be still be established in such situations as long as the data of the original problem is sufficiently smooth, i.e. in a small enough domain. Our results include well-known examples applied in the approach of quasi-reversibility as well as other types of approximations. The outcomes of the paper may be applied to partial differential equations in L^p spaces, $1 < p < \infty$ defined by strongly elliptic differential operators.

Keywords Ill-posed problem · Regularizing family of operators · Holomorphic semigroup · Continuous dependence of solutions

Mathematics Subject Classification (2010) Primary 47A52; Secondary 47D06

M. A. Fury (✉)
Penn State Abington, Department of Mathematics, Abington, PA, USA
e-mail: maf44@psu.edu

M. A. Bastos et al. (eds.), *Operator Theory, Functional Analysis and Applications*,
Operator Theory: Advances and Applications 282,
https://doi.org/10.1007/978-3-030-51945-2_12

1 Introduction

In this paper, we consider the generally ill-posed, linear Cauchy problem

$$\begin{aligned} u'(t) &= Au + h(t), \quad 0 < t < T, \\ u(0) &= \varphi \end{aligned} \tag{1.1}$$

where A is a closed, linear operator in a Banach space X, $\varphi \in X$, and h is a function from $[0, T]$ into X. Ill-posed problems, whose solutions either do not exist for φ in a dense subspace of X, are not unique, or do not depend continuously upon φ, appear abundantly in several fields, most notably as backward parabolic problems (cf. [23, 25, 28, 34]). For instance, let $A = -\Delta$ defined by $\Delta u = \sum_{i=1}^{n} \frac{\partial^2 u}{\partial x_i^2}$ for all $u \in X = L^p(\mathbb{R}^n)$, $1 < p < \infty$, whose generalized derivatives up to order 2 are also in X. Then (1.1) becomes the classic ill-posed problem that is the backward heat equation. Related to ill-posed problems, inverse problems are also of significant interest with applications such as parameter identification in medical imaging and source identification in groundwater pollution (cf. [19, 22, 30, 33, 34, 36, 39]).

Due to the instability of ill-posed problems, it is common practice to estimate a supposed solution $u(t)$ of (1.1) via some approximation method. For instance, several regularization techniques have been applied to (1.1) [2, 17, 18, 24, 25, 28, 35] which utilize the generator of a C_0 semigroup of linear operators, that is a family of bounded linear operators $\{E(t)\}_{t\geq 0}$ on X satisfying

$$E(0) = I, \quad E(s+t) = E(s)E(t) \quad \text{for all} \quad s, t \geq 0, \quad \text{and}$$

$$E(t)x \to x \quad \text{as} \quad t \to 0^+ \quad \text{for all} \quad x \in X;$$

formally, we regard $E(t)$ as $E(t) = e^{tB}$ where B is its generator (cf. [29, Section 1.2]). While the literature contains many results in Hilbert space, recently including concrete numerical experiments which validate the theory [21, 37, 38], fewer results are available in pure Banach space. Generalizing the example $A = -\Delta$, authors have considered the case in Banach space where $-A$ generates a holomorphic semigroup e^{-tA} which extends analytically into a sector of the complex plane. The most well-known and cited approximations here are $f_\beta(A) = A - \beta A^2$ [20] and $f_\beta(A) = A(I + \beta A)^{-1}$ [35], though each of these regularizations depends on the sector angle θ of the semigroup e^{-tA} (cf. [2, 5, 6, 10, 17, 18, 23, 35]). For $f_\beta(A) = A - \beta A^2$, the restriction $\theta \in (\frac{\pi}{4}, \frac{\pi}{2}]$ is required in order for $f_\beta(A)$ to generate a semigroup. In this case, under appropriate conditions on h, the approximate Cauchy problem

$$\begin{aligned} v'(t) &= f_\beta(A)v + h(t), \quad 0 < t < T, \\ v(0) &= \varphi \end{aligned} \tag{1.2}$$

is well-posed with unique solution $v_\beta(t) = e^{tf_\beta(A)}\varphi + \int_0^t e^{(t-s)f_\beta(A)}h(s)ds$. The sum $-A + f_\beta(A) = -\beta A^2$ generates a C_0 semigroup as well, and in fact its growth order does not depend on β. Consequently, regularization is established as $v_\beta(t)$ then converges to $u(t)$ for each $t \in [0, T]$ as $\beta \to 0$ (cf. [24, Section 3.1.1]). As for $f_\beta(A) = A(I + \beta A)^{-1}$, regardless of the angle θ, this approximation yields well-posedness of (1.2) being a bounded operator. The auxiliary operator $-A + f_\beta(A)$ generates a C_0 semigroup as well, but as shown in the literature, its growth order is independent of β only when $\theta \in (\frac{\pi}{4}, \frac{\pi}{2}]$ (cf. [2, 10, 18, 35]).

Recently, authors have investigated possibilities where the restriction on the angle θ of the semigroup may be relaxed. Huang and Zheng [17] show that the approximation $f_\beta(A) = A - \beta A^2$ may be modified by use of the fractional power of A [3]. Here, even if $0 < \theta < \frac{\pi}{4}$, then both $f_\beta(A) = A - \beta A^\sigma$ and $-A + f_\beta(A) = -\beta A^\sigma$ still generate C_0 semigroups for a suitable $1 < \sigma < 2$. Also, a logarithmic approximation

$$f_\beta(A) = -\frac{1}{pT}\ln(\beta + e^{-pTA}), \quad 0 < \beta < 1, \quad p \geq 1 \tag{1.3}$$

first introduced by Boussetila and Rebbani [4] in Hilbert space, and later modified by Tuan and Trong [38], may be applied in Banach space (cf. [16], [6], [12]). The approximation (1.3) has lately received significant attention since it induces an error that is less severe than those of $A - \beta A^2$ and $A(I + \beta A)^{-1}$, both of which satisfy

$$\|(-A + f_\beta(A))x\| \leq \beta\|A^2 x\| \quad \text{for} \quad x \in Dom(A^2), \tag{1.4}$$

and $\|e^{tf_\beta(A)}\| \leq e^{tC/\beta}$. Recently, the author [11] investigated two additional approximations $f_\beta(A) = Ae^{-\beta A}$ and $f_\beta(A) = (\ln 2)^{-1}A\ \mathrm{Log}(1 + e^{-\beta A})$ which satisfy these same properties, but in such a way that the calculations do not rely on the value of the angle θ.

In this paper, we generalize the results of [11] in order to unify all four of the approximations

$$A - \beta A^\sigma, \quad A(I + \beta A)^{-1}, \quad Ae^{-\beta A}, \quad (\ln 2)^{-1}A\ \mathrm{Log}(1 + e^{-\beta A}),$$

and prove convergence estimates between $u(t)$ and $v_\beta(t)$ without needing to restrict the angle θ (Proposition 2.6 and Theorem 2.7 below). For this, our method relies on an assumed smoothness of the data φ and h in (1.1). For example, if $\varphi \in Dom(e^{QA})$ for large enough $Q > T$ then $e^{tA}\varphi$ behaves like $e^{-(Q-t)A}(e^{QA}\varphi)$ where $e^{-(Q-t)A}$ is bounded for each t. We note that such a requirement is not out of the ordinary since, for example, a solution $u(t)$ of (1.1) with $h \equiv 0$ exists if and only if $u(t) \in Dom(e^{tA})$ for each t (cf. [17, Introduction]). Also, many of the results associated with the logarithmic approximation (1.3) demonstrate a stricter property than (1.4), that is $\|(-A + f_\beta(A))x\| \leq \beta\|e^{\tau A}x\|$ for all $x \in Dom(e^{\tau A})$ where τ may be larger than T (cf. [37, Definition 1], [13, Lemma 1], [12, Proposition 6]). It is notable

that our intention, to regularize problem (1.1) when favorable C_0 semigroups are unavailable, likens to deLaubenfels's motivation of C-semigroups, for example $\{Ce^{tA}\}_{t\geq 0}$ where C is a bounded, injective operator such that $Ce^{tA} = e^{tA}C$ is a bounded linear operator for every $t \geq 0$ [7, 8].

Below, $B(X)$ denotes the space of all bounded linear operators on X. For a linear operator A in X, $\rho(A)$ denotes the resolvent set of A and $\sigma(A)$, the spectrum of A, is the complement of $\rho(A)$ in $\mathbb{C}$. Also, a strong solution of (1.1) is a function u which is differentiable almost everywhere on $[0, T]$ such that $u' \in L^1((0, T) : X)$, $u(0) = \varphi$, and $u'(t) = Au(t) + h(t)$ almost everywhere on $[0, T]$ (cf. [29, Definition 4.2.8]).

2 A Unifying Condition for Well-Posed Approximation

Assume $-A$ is the infinitesimal generator of a bounded holomorphic semigroup e^{-tA} of angle $\theta \in (0, \frac{\pi}{2}]$ on a Banach space X with $0 \in \rho(A)$. By definition, then e^{-tA} extends to an analytic function e^{-wA} defined in the open sector $S_\theta = \{w \neq 0 \in \mathbb{C} : |\arg w| < \theta\}$ which is uniformly bounded in every sector S_{θ_1} with $\theta_1 \in (0, \theta)$. Also, we have the following equivalence in terms of resolvent operators.

Theorem 2.1 ([31, Theorem X.52]) *For a closed operator A on a Banach space X, $-A$ is the infinitesimal generator of a bounded holomorphic semigroup of angle θ if and only if for each $\theta_1 \in (0, \theta)$, there exists a constant $M_1 > 0$ such that*

$$w \notin \overline{S}_{\frac{\pi}{2}-\theta_1} \implies w \in \rho(A),$$

and

$$\|(w - A)^{-1}\| \leq \frac{M_1}{dist(w, \bar{S}_{\frac{\pi}{2}-\theta_1})} \quad \textit{for all} \quad w \notin \overline{S}_{\frac{\pi}{2}-\theta_1}. \tag{2.1}$$

We note that by a simple geometric argument, the condition (2.1) is equivalent to

$$\|(w - A)^{-1}\| \leq \frac{M_1'}{|w|} \quad \text{for all} \quad w \notin \bar{S}_{\frac{\pi}{2}-\theta_1} \tag{2.2}$$

for a possibly different constant M_1' depending on θ_1.

Under our assumptions, (1.1) is generally ill-posed with instability of solutions as described in Sect. 1. To ensure that (1.1) has a well-posed approximation, we call upon the following result.

Theorem 2.2 ([29, Corollary 4.2.10]) *For each $0 < \beta < 1$, let $f_\beta(A)$ be the infinitesimal generator of a C_0 semigroup of bounded linear operators on X. If h is differentiable almost everywhere on $[0, T]$ and $h' \in L^1((0, T) : X)$, then (1.2) has*

a unique strong solution given by

$$v_\beta(t) = e^{tf_\beta(A)}\varphi + \int_0^t e^{(t-s)f_\beta(A)}h(s)ds$$

for every $\varphi \in Dom(f_\beta(A))$.

Unless the representation $f_\beta(A)$ is elementary, the operator $f_\beta(A)$ will be defined by the Dunford integral for functions $f_\beta(w)$ bounded and holomorphic in a sector of the complex plane, particularly for us in most cases

$$f_\beta(A) = \frac{1}{2\pi i}\int_{\Gamma_\alpha} f_\beta(w)(w - A)^{-1}dw$$

where $\Gamma_\alpha = \partial S_\alpha$, $\frac{\pi}{2} - \theta < \alpha < \frac{\pi}{2}$, oriented so that $\mathrm{Im}(w)$ decreases as w travels along Γ_α (cf. [9, Chapter VII] and also [15, Chapter 2]).

The relationship between A and $f_\beta(A)$ will need to be scrutinized in order to estimate known solutions of problem (1.1). Indeed, we aim to incorporate a condition similar to (1.4) where $f_\beta(A)x \to Ax$ as $\beta \to 0$ for x in an appropriate domain. Ultimately, this condition establishes continuous dependence on modeling since our final goal is to show a similar convergence for the corresponding solutions $v_\beta(t)$ and $u(t)$. To this end, we endorse the approximation condition, Condition (A) of Ames and Hughes [2, Definition 1].

Definition 2.3 Let $-A$ be the infinitesimal generator of a bounded holomorphic semigroup of angle $\theta \in (0, \frac{\pi}{2}]$ on a Banach space X with $0 \in \rho(A)$, and suppose that for each $0 < \beta < 1$, $f_\beta(A)$ is the infinitesimal generator of a C_0 semigroup. Then $f_\beta(A)$ satisfies Condition (A) if there exists a positive constant p independent of β such that $Dom(f_\beta(A)) \supseteq Dom(A^{1+p})$ and

$$\|(-A + f_\beta(A))x\| \leq \beta\|A^{1+p}x\| \tag{2.3}$$

for all $x \in Dom(A^{1+p})$.

Note that in the definition of Condition (A), p need not be an integer, in which case A^{1+p} is defined by the fractional power of A [3] (see also [29, Section 2.6]). Nevertheless, we still have $Dom(A^{1+p}) \subseteq Dom(A)$ so that (2.3) is valid.

As we will see for our main results below (Proposition 2.6 and Theorem 2.7), not only do we require that $f_\beta(A)$ satisfy the stipulations of Condition (A), but also the operator $-A + f_\beta(A)$, with domain $Dom(A) \cap Dom(f_\beta(A))$, must have stable enough properties in a certain sense. Therefore, we assume a stronger condition.

Definition 2.4 Let $-A$ be the infinitesimal generator of a bounded holomorphic semigroup of angle $\theta \in (0, \frac{\pi}{2}]$ on a Banach space X with $0 \in \rho(A)$, and suppose

that for each $0 < \beta < 1$, $f_\beta(A)$ is the infinitesimal generator of a C_0 semigroup. Then $f_\beta(A)$ satisfies Condition (A^+) if the following are satisfied:

(i) $Dom(f_\beta(A)) \supseteq Dom(A^2)$ and there exists a constant R independent of β such that

$$\|(-A + f_\beta(A))x\| \leq \beta R\|A^2 x\| \quad \text{for all} \quad x \in Dom(A^2),$$

(ii) There exist $\frac{\pi}{2} - \theta < \nu < \frac{\pi}{2}$ and $C \geq 0$ independent of β such that

$$Re(-w + f_\beta(w)) \leq C|w| \quad \text{for all } w \text{ in the open sector}$$
$$S_\nu = \{w \neq 0 \in \mathbb{C} \ : \ |\arg w| < \nu\}.$$

Remark 2.5 Conventionally, a more desirable condition is one similar to Condition (A^+) (ii) but the constant C being negative. In that case, the sum $-A + f_\beta(A)$ generates a C_0 semigroup whose growth order $Me^{\omega t}$ does not depend on β, a property which is required for standard regularization arguments [17, 18, 23]. In fact, this is possible even if $C < 0$ depends on β, for instance with the example $-A + f_\beta(A) = -\beta A^\sigma$ (cf. [17, Theorem 3.1] and also [10, Proposition 3.4]). In any case, in our general framework, we may not have this desired property, for example (as noted in Sect. 1) if $f_\beta(A) = A(I + \beta A)^{-1}$ and $0 < \theta < \frac{\pi}{4}$.

Despite Remark 2.5, we are able to prove regularization-type calculations via Theorem 2.7 below if we assume strong enough smoothness properties on φ and h. Proposition 2.6 demonstrates an initial estimate that may be established by choosing data in a small enough domain. As each operator e^{-tA}, $t \geq 0$ is a bounded, *injective* operator (cf. [7, Lemma 3.1]), we find an appropriate Q with φ and $Ran(h)$ contained in $Dom(e^{QA}) = Dom((e^{-QA})^{-1})$.

Proposition 2.6 *Let $-A$ be the infinitesimal generator of a bounded holomorphic semigroup of angle $\theta \in (0, \frac{\pi}{2}]$ on a Banach space X with $0 \in \rho(A)$ and assume $f_\beta(A)$ satisfies Condition (A^+). For every $0 < \epsilon < 1$ and $\frac{\pi}{2} - \theta < \alpha < \nu$, there exists a constant $Q = Q(\epsilon, \alpha) > T$ such that if $x \in Dom(e^{QA})$, then for each integer $n \geq 0$,*

$$\|A^n(e^{tA}x - e^{tf_\beta(A)}x)\| \leq \beta K(T - t + \epsilon)^{-(n+1)}\|e^{QA}x\| \quad \textit{for} \quad 0 \leq t \leq T$$

where K is a constant depending on n and α but independent of β, ϵ, and t.

Proof Let $0 < \epsilon < 1$ and $\frac{\pi}{2} - \theta < \alpha < \nu$ where ν and C are the constants from Condition (A^+) (ii). Further, let Q be any constant larger than T satisfying the inequality $(Q - T)\cos\alpha \geq C(T + \epsilon)$. If $x \in Dom(e^{QA})$, then by Condition (A^+) (i) and standard semigroup properties (cf. [29, Theorem 1.2.4]),

$$\begin{aligned}\|A^n(e^{tA}x - e^{tf_\beta(A)}x)\| &= \|(I - e^{tf_\beta(A)}e^{-tA})A^n e^{tA}x\| \\ &= \|(I - e^{tf_\beta(A)}e^{-tA})A^n(e^{tA}e^{-QA}(e^{QA}x))\|\end{aligned}$$

$$= \left\| - \int_0^t \left(\frac{\partial}{\partial \tau} e^{\tau f_\beta(A)} e^{-\tau A} \right) A^n e^{tA} e^{-QA} (e^{QA} x)\, d\tau \right\|$$

$$= \left\| - \int_0^t (-A + f_\beta(A)) e^{\tau f_\beta(A)} e^{-\tau A} A^n e^{tA} e^{-QA} (e^{QA} x)\, d\tau \right\|$$

$$\leq \beta R \int_0^t \| A^{2+n} e^{\tau f_\beta(A)} e^{-\tau A} e^{tA} e^{-QA} (e^{QA} x) \|\, d\tau. \tag{2.4}$$

By the estimate $(Q - T)\cos\alpha \geq C(T + \epsilon)$, then for $w = re^{\pm i\alpha} \in \Gamma_\alpha = \partial S_\alpha$,

$$\begin{aligned}
&\| A^{2+n} e^{\tau f_\beta(A)} e^{-\tau A} e^{tA} e^{-QA} (e^{QA} x) \| \\
&\quad = \left\| \frac{1}{2\pi i} \int_{\Gamma_\alpha} w^{2+n} e^{\tau(-w + f_\beta(w))} e^{-(Q-t)w} (w - A)^{-1} (e^{QA} x)\, dw \right\| \\
&\quad \leq \frac{1}{2\pi} \int_{\Gamma_\alpha} |w|^{2+n} e^{\tau Re(-w + f_\beta(w))} e^{-(Q-t)Re(w)} \| (w - A)^{-1} \|\, \| e^{QA} x \|\, |dw| \\
&\quad \leq \frac{1}{\pi} \int_0^\infty r^{1+n} e^{\tau C r} e^{-(Q-t) r \cos\alpha} M_\alpha \| e^{QA} x \|\, dr \\
&\quad \leq \frac{M_\alpha}{\pi} \int_0^\infty r^{1+n} e^{\tau C r} e^{-(Q-T) r \cos\alpha} \| e^{QA} x \|\, dr \\
&\quad \leq \frac{M_\alpha}{\pi} \int_0^\infty r^{1+n} e^{C(\tau - (T+\epsilon)) r} \| e^{QA} x \|\, dr \\
&\quad = \frac{M_\alpha}{\pi} (n+1)! C^{-(2+n)} (T - \tau + \epsilon)^{-(2+n)} \| e^{QA} x \|
\end{aligned} \tag{2.5}$$

where M_α is a constant due to (2.2). Returning to (2.4),

$$\begin{aligned}
&\| A^n (e^{tA} x - e^{t f_\beta(A)} x) \| \\
&\quad \leq \beta R \frac{M_\alpha}{\pi} (n+1)! C^{-(2+n)} \int_0^t (T - \tau + \epsilon)^{-(2+n)} \| e^{QA} x \|\, d\tau \\
&\quad = \beta R \frac{M_\alpha}{\pi} (n+1)! C^{-(2+n)} \frac{(T+\epsilon)^{n+1} - (T - t + \epsilon)^{n+1}}{(T - t + \epsilon)^{n+1} (T + \epsilon)^{n+1}} \| e^{QA} x \| \\
&\quad \leq \beta R \frac{M_\alpha}{\pi} (n+1)! C^{-(2+n)} (T - t + \epsilon)^{-(n+1)} \| e^{QA} x \|.
\end{aligned}$$

□

Now, we prove our main approximation theorem concerning strong solutions of (1.1) and (1.2).

Theorem 2.7 *Let $-A$ be the infinitesimal generator of a bounded holomorphic semigroup of angle $\theta \in (0, \frac{\pi}{2}]$ on a Banach space X with $0 \in \rho(A)$ and assume $f_\beta(A)$ satisfies Condition (A^+). Fix $0 < \epsilon < 1$ and $\frac{\pi}{2} - \theta < \alpha < \nu$, and let*

$Q = Q(\epsilon, \alpha) > T$ be as in Proposition 2.6. Assume that $u(t)$ and $v_\beta(t)$ are strong solutions of (1.1) and (1.2) respectively where h satisfies the conditions of Theorem 2.2, and as well that initial data φ and $Ran(h)$ are contained in $Dom(e^{QA})$ with $e^{QA}h \in L^1((0,T) : X)$. Then there exist constants $\tilde{C}$ and M each independent of β and ϵ such that for $0 \le t < T$,

$$\|u(t) - v_\beta(t)\| \le \tilde{C}\beta^{1-\omega(t)}M^{\omega(t)}\epsilon^{-2}N \tag{2.6}$$

where

$$N = \|e^{QA}\varphi\| + \int_0^T \|e^{QA}h(t)\|dt, \tag{2.7}$$

and $\omega(\zeta)$ is a harmonic function which is bounded and continuous on the bent strip

$$\Lambda = \{t + re^{\pm i\gamma} \mid 0 \le t \le T,\ r \ge 0\}, \quad \gamma \in (0, \theta),$$

satisfying $0 = \omega(0) \le \omega(t) < \omega(T) = 1$ for all $0 \le t < T$.

Proof Fix $\gamma \in (0, \theta)$. By the definition of the semigroup e^{-tA}, we know that e^{-wA} is uniformly bounded in S_γ (cf. [31, Theorem X.52]). Following [2, 5, 26], we extend the solutions $u(t)$ and $v_\beta(t)$ into the bent strip Λ and apply Carleman's inequality (cf. [14, p. 346]). For $\zeta = t + re^{\pm i\gamma} \in \Lambda$, define

$$w(\zeta) = e^{-(re^{\pm i\gamma})A}(u(t) - v_\beta(t)).$$

Also following [1], define

$$F(\zeta) = w(\zeta) + \frac{1}{\pi}\int_0^\infty\int_0^T \bar{\partial}w(z)\left(\frac{1}{z-\zeta} + \frac{1}{\bar{z}+1+\zeta}\right)d\xi d\eta \tag{2.8}$$

where $z = \xi + \eta e^{\pm i\gamma}$ and $\bar{\partial}$ denotes the Cauchy-Riemann operator (cf. [32]). Denote

$$M' = \left(\max_{r\ge 0}\|e^{-(re^{\pm i\gamma})A}\|\right).$$

By Proposition 2.6 and (2.7),

$$\begin{aligned}\|w(\zeta)\| &\le M'\|u(t) - v_\beta(t)\| \\ &\le M'\left(\|e^{tA}\varphi - e^{tf_\beta(A)}\varphi\| + \int_0^t \|(e^{(t-s)A} - e^{(t-s)f_\beta(A)})h(s)\|ds\right) \\ &\le M'\Big(\beta K(T - t + \epsilon)^{-1}\|e^{QA}\varphi\|\end{aligned}$$

$$+ \int_0^t \beta K(T-(t-s)+\epsilon)^{-1}\|e^{QA}h(s)\|ds\Big)$$
$$\le M'\beta K\epsilon^{-1}\left(\|e^{QA}\varphi\| + \int_0^T \|e^{QA}h(y)\|dy\right)$$
$$= M'\beta K\epsilon^{-1}N. \tag{2.9}$$

Next, by the definition of $w(\zeta)$, a straightforward calculation shows that

$$\bar\partial w(\zeta) = \frac{1}{2i\,\sin(\pm\gamma)}\left(e^{\pm i\gamma}\frac{\partial}{\partial t}w(\zeta) - \frac{\partial}{\partial \eta}w(\zeta)\right)$$
$$= \frac{e^{\pm i\gamma}}{2i\,\sin(\pm\gamma)}e^{-(re^{\pm i\gamma})A}\left((Au(t) - f_\beta(A)v_\beta(t)) + (Au(t) - Av_\beta(t))\right).$$

Again using Proposition 2.6 and (2.7),

$$\|Au(t) - Av_\beta(t)\|$$
$$= \left\| A\left(e^{tA}\varphi + \int_0^t e^{(t-s)A}h(s)ds\right)\right.$$
$$\left. + A\left(e^{tf_\beta(A)}\varphi + \int_0^t e^{(t-s)f_\beta(A)}h(s)ds\right)\right\|$$
$$\le \|A(e^{tA} - e^{tf_\beta(A)})\varphi\| + \int_0^t \|A(e^{(t-s)A} - e^{(t-s)f_\beta(A)})h(s)\|ds$$
$$\le \beta K(T-t+\epsilon)^{-2}\|e^{QA}\varphi\| + \int_0^t \beta K(T-(t-s)+\epsilon)^{-2}\|e^{QA}h(s)\|ds$$
$$\le \beta K\epsilon^{-2}N.$$

Also, by Condition (A$^+$) (i), (2.5), and (2.7),

$$\|Av_\beta(t) - f_\beta(A)v_\beta(t)\| = \|(-A + f_\beta(A))v_\beta(t)\|$$
$$\le \beta R\|A^2 v_\beta(t)\|$$
$$= \beta R\left\|A^2\left(e^{tf_\beta(A)}\varphi + \int_0^t e^{(t-s)f_\beta(A)}h(s)ds\right)\right\|$$

$$\le \beta R\left(\|A^2 e^{tf_\beta(A)}\varphi\| + \int_0^t \|A^2 e^{(t-s)f_\beta(A)}h(s)\|ds\right)$$
$$= \beta R\Big(\|A^2 e^{tf_\beta(A)}e^{-tA}e^{-(Q-t)A}(e^{QA}\varphi)\|$$

$$+ \int_0^t \|e^{(t-s)f_\beta(A)} e^{-(t-s)A} e^{-(Q-(t-s))A}(e^{QA}h(s))\|ds\Bigg)$$
$$\leq \beta RL\epsilon^{-2}N$$

where L is a constant independent of β, ϵ, and t. Altogether, we have shown

$$\|\bar{\partial}w(\zeta)\| \leq \beta C'\epsilon^{-2}N, \tag{2.10}$$

where C' is a constant independent of β, ϵ, and ζ.

If x^* is a member of the dual space of X with $\|x^*\| \leq 1$, then it follows that $x^*(F(\zeta))$ as defined by (2.8) is a bounded, continuous function on Λ which is also analytic on the interior of Λ (cf. [1, 5]). Furthermore, there exists a constant L' independent of ζ such that

$$|x^*F(\zeta)| \leq \|x^*\| \left(\|w(\zeta)\| + L' \max_{\zeta\in\Lambda} \|\bar{\partial}(w(\zeta))\| \right).$$

By (2.9) and (2.10), together with the fact that $0 < \epsilon < 1$, then

$$|x^*F(\zeta)| \leq \beta M\epsilon^{-2}N\|x^*\| \tag{2.11}$$

where M is a constant independent of β, ϵ, and ζ. Hence Carleman's Inequality implies

$$|x^*F(t)| \leq M(0)^{1-\omega(t)}M(T)^{\omega(t)}, \tag{2.12}$$

for $0 \leq t \leq T$, where

$$M(t) = \max_{r\geq 0} |x^*F(t + re^{\pm i\gamma})|$$

and ω is a harmonic function which is bounded and continuous on Λ, satisfying $0 = \omega(0) \leq \omega(t) < \omega(T) = 1$ for all $0 \leq t < T$ (cf. [26], [14, p. 346]). Note by (2.10),

$$\begin{aligned} M(0) &\leq \|x^*\| \left(\|w(re^{\pm i\gamma})\| + L' \max_{r\geq 0} \|\bar{\partial}(re^{\pm i\gamma})\| \right) \\ &= \|x^*\| \left(\|e^{-(re^{\pm i\gamma})A}(\varphi - \varphi)\| + L' \max_{r\geq 0} \|\bar{\partial}(re^{\pm i\gamma})\| \right) \\ &= L' \max_{r\geq 0} \|\bar{\partial}(re^{\pm i\gamma})\| \|x^*\| \\ &\leq L'\beta C'\epsilon^{-2}N\|x^*\|. \end{aligned}$$

Also, from (2.11) and the fact that $0 < \beta < 1$,

$$M(T) \leq M\epsilon^{-2}N\|x^*\|.$$

Returning to (2.12), then

$$|x^*F(t)| \leq (L'\beta C')^{1-\omega(t)}M^{\omega(t)}\epsilon^{-2}N\|x^*\|$$

and taking the supremum over $x^* \in X^*$ with $\|x^*\| \leq 1$, we obtain

$$\|F(t)\| \leq \tilde{C}\beta^{1-\omega(t)}M^{\omega(t)}\epsilon^{-2}N$$

for $0 \leq t \leq T$ where $\tilde{C}$ and M are constants each independent of β and ϵ. Then for $0 \leq t \leq T$,

$$\begin{aligned}
\|u(t) - v_\beta(t)\| &= \|w(t)\| \\
&\leq \|F(t)\| + L' \max_{0\leq t\leq T} \|\bar{\partial}w(t)\| \\
&\leq \tilde{C}\beta^{1-\omega(t)}M^{\omega(t)}\epsilon^{-2}N + L'\beta C'\epsilon^{-2}N \\
&= (\tilde{C} + L'\beta^{\omega(t)}C'M^{-\omega(t)})\beta^{1-\omega(t)}M^{\omega(t)}\epsilon^{-2}N \\
&\leq \tilde{C}\beta^{1-\omega(t)}M^{\omega(t)}\epsilon^{-2}N
\end{aligned}$$

for a possibly different constant $\tilde{C}$ independent of β and ϵ. □

3 Examples Illustrating Condition (A^+)

Here we outline several examples which satisfy Condition (A^+), so that Theorem 2.7 may be applied. As mentioned in Sect. 1, each of these approximations may be applied to backward problems defined in $L^p(\mathbb{R}^n)$, $1 < p < \infty$. For specifics as well as further applications, see Remark 4.2 below.

Example (1) $f_\beta(A) = A(I + \beta A)^{-1}$. Property (2.2) implies $(I + \beta A)^{-1} \in B(X)$ and

$$\|(I + \beta A)^{-1}\| = \beta^{-1}\|(-\beta^{-1} - A)^{-1}\| \leq \beta^{-1}\frac{R}{\beta^{-1}} = R \tag{3.1}$$

for some constant R. Then

$$f_\beta(A) = \beta^{-1}((I + \beta A) - I)(I + \beta A)^{-1} = \beta^{-1}(I - (I + \beta A)^{-1})$$

is also in $B(X)$ being a linear combination of two bounded operators, and by (3.1), $\|f_\beta(A)\| \leq (1+R)\beta^{-1}$. Therefore, $f_\beta(A)$ generates a uniformly continuous semigroup satisfying

$$\|e^{tf_\beta(A)}\| \leq e^{t(1+R)\beta^{-1}}$$

for each $t \geq 0$.

Note Condition (A^+) (i) is satisfied as

$$\begin{aligned} -A + A(I+\beta A)^{-1} &= (-A(I+\beta A) + A)(I+\beta A)^{-1} \\ &= -\beta A^2(I+\beta A)^{-1} = -\beta(I+\beta A)^{-1}A^2 \end{aligned}$$

on $Dom(A^2)$ and by (3.1),

$$\| -\beta(I+\beta A)^{-1}A^2 x\| \leq \beta R\|A^2 x\|$$

for all $x \in Dom(A^2)$.

Next, let $\frac{\pi}{2} - \theta < \nu < \frac{\pi}{2}$. Consider for $w = re^{\pm i\alpha} \in S_\nu$,

$$\begin{aligned} Re(-w + f_\beta(w)) &= Re(-w + \frac{w}{1+\beta w}) \\ &= -Re(\frac{\beta w^2}{1+\beta w}) \\ &= -Re\left(\frac{\beta r^2\cos(2\alpha) \pm i\beta r^2\sin(2\alpha)}{(1+\beta r\cos\alpha) \pm i\beta r\sin\alpha}\right) \\ &= -\frac{\beta r^2\cos(2\alpha) + \beta^2 r^3(\cos(2\alpha)\cos\alpha + \sin(2\alpha)\sin(\alpha))}{(1+\beta r\cos\alpha)^2 + (\beta r\sin\alpha)^2} \\ &= -\frac{\beta r^2\cos(2\alpha) + \beta^2 r^3\cos\alpha}{1+2\beta r\cos\alpha + \beta^2 r^2} := g(r). \end{aligned} \tag{3.2}$$

We claim that $r - g(r) \geq 0$ for all $r \geq 0$. Indeed,

$$\begin{aligned} r - g(r) &= \frac{r + 2\beta r^2\cos\alpha + \beta^2 r^3 + \beta r^2\cos(2\alpha) + \beta^2 r^3\cos\alpha}{1+2\beta r\cos\alpha + \beta^2 r^2} \\ &= \frac{r + (2\cos\alpha + \cos(2\alpha))\beta r^2 + (1+\cos\alpha)\beta^2 r^3}{1+2\beta r\cos\alpha + \beta^2 r^2}. \end{aligned}$$

Define

$$k(t) = 2\cos t + \cos(2t), \quad 0 \leq t \leq \frac{\pi}{2}.$$

Then $k'(t) = -2(\sin t + \sin(2t)) \leq 0$ since both $\sin t$ and $\sin(2t)$ are nonnegative on $[0, \frac{\pi}{2}]$. Consequently, $k(t)$ is a decreasing function, and thus is never smaller than $k(\frac{\pi}{2}) = -1$. We have

$$r + (2\cos\alpha + \cos(2\alpha))\beta r^2 + (1 + \cos\alpha)\beta^2 r^3$$
$$\geq r - \beta r^2 + \beta^2 r^3 = r(1 - \beta r + \beta^2 r^2) = r\left(\left(\beta r - \frac{1}{2}\right)^2 + \frac{3}{4}\right) \geq 0$$

and so $r - g(r) \geq 0$ as well. We have shown

$$|w| - Re(-w + f_\beta(w)) \geq 0$$

and so Condition (A^+) (ii) is satisfied with $C = 1$.

Remark 3.1 As pointed out in [18, Remark 3.3], if $\theta > \frac{\pi}{4}$, so that $\frac{\pi}{2} - \theta < \frac{\pi}{4}$, then ν can be chosen less than $\frac{\pi}{4}$ and so $g(r)$ in (3.2) is strictly negative. In this case, $-A + f_\beta(A)$ generates a C_0 semigroup whose growth order is independent of β [18, Theorem 2.1] and therefore, sharper estimates than (2.6) may be used to establish regularization as in the literature cited in Sect. 1. Nevertheless, in our framework, Condition (A^+) (ii) is still satisfied, this time with $C = 0$.

Example (2) $f_\beta(A) = A - \beta A^\sigma$ where σ and ν satisfy $\frac{\pi}{2} - \theta < \nu < \sigma\nu < \frac{\pi}{2}$. Note, σ is necessarily larger than 1 but its upper bound is determined by the value of θ. For instance, if $\frac{\pi}{2} - \theta < \frac{\pi}{4}$, then ν can be chosen less than $\frac{\pi}{4}$ and $\sigma = 2$ suffices. In this case, Condition (A^+) (i) is easily satisfied as

$$\|(-A + f_\beta(A))x\| = \| - \beta A^2 x\| = \beta\|A^2 x\|$$

for all $x \in Dom(A^2)$. Condition (A^+) (ii) is also satisfied with $C = 0$ since $2\nu < \frac{\pi}{2}$ implies

$$Re(-w + f_\beta(w)) = Re(-\beta w^2) = -\beta r^2 \cos(2\alpha) < -\beta r^2 \cos(2\nu) < 0$$

for all $w = re^{\pm i\alpha} \in S_\nu$.

In the case that $\frac{\pi}{2} - \theta \geq \frac{\pi}{4}$, we must revise our calculations. We can choose $1 < \sigma < \frac{\pi}{2\nu}$ so that $\sigma\nu < \frac{\pi}{2}$ but in this case σ is not an integer since $\nu > \frac{\pi}{4}$. Therefore, $f_\beta(A) = A - \beta A^\sigma$ is defined by the fractional power of A, that is

$$A^{-\sigma} = \frac{1}{2\pi i}\int_\Gamma w^{-\sigma}(w - A)^{-1} dw, \quad A^\sigma = (A^{-\sigma})^{-1}$$

where Γ is a contour similar to $\Gamma_\alpha = \partial S_\alpha$ but avoids both the origin and the negative real axis [29, Section 2.6]. In fact, since $0 \in \rho(A)$, there exists $\delta > 0$ such that the closed disk of radius δ, centered at the origin is contained in $\rho(A)$. Hence, we may

revise Γ_α to be the contour

$$\Gamma'_\alpha = \Gamma_1 \cup \Gamma_2 \cup \Gamma_3,$$
$$\Gamma_1 = \{re^{i\alpha} : r \geq \delta\},$$
$$\Gamma_2 = \{\delta e^{i\phi} : -\alpha \leq \phi \leq \alpha\},$$
$$\Gamma_3 = \{re^{-i\alpha} : r \geq \delta\},$$

oriented so that $\mathrm{Im}(w)$ decreases as w travels along Γ'_α (cf. [17]). But Γ_2 is bounded and does not affect the convergence of the contour integral, so the calculations in Proposition 2.6 remain fundamentally unchanged. Henceforth, using the properties (cf. [29, Theorem 2.6.8, Lemma 2.6.3])

$$Dom(A^2) \subseteq Dom(A^\sigma) \subseteq Dom(A) \quad \text{for} \quad 1 < \sigma < 2,$$

$$A^{\sigma_1}(A^{\sigma_2})x = A^{\sigma_1+\sigma_2}x = A^{\sigma_2}(A^{\sigma_1})x \quad \text{for} \quad x \in Dom(A^{\max\{\sigma_1,\sigma_2,\sigma_1+\sigma_2\}}),$$

$$\|A^{-\sigma}\| \leq \kappa \quad \text{for} \quad 0 < \sigma < 1,$$

Condition (A$^+$) (i) is satisfied by

$$\|(-A + f_\beta(A)x)\| = \beta\|A^\sigma x\| = \beta\|A^{-(2-\sigma)}A^2 x\| \leq \beta\kappa\|A^2 x\|$$

for all $x \in Dom(A^2)$. Also similar to the case $\sigma = 2$, Condition (A$^+$) (ii) is satisfied with $C = 0$ since $\sigma\nu < \frac{\pi}{2}$ implies

$$Re(-w + f_\beta(w)) = Re(-\beta w^\sigma) = -\beta r^\sigma \cos(\sigma\alpha) < -\beta r^\sigma \cos(\sigma\nu) < 0.$$

Finally, we point out that in either case whether $\sigma = 2$ or $1 < \sigma < 2$, while $f_\beta(A)$ and $-A + f_\beta(A)$ are unbounded operators, it may be shown that both generate C_0 semigroups. Again, if σ is not an integer, careful revisions to Γ_α must be taken near the origin. In either case, one may show the semigroup generated by $f_\beta(A)$ satisfies

$$\|e^{tf_\beta(A)}\| \leq Pe^{tP'\beta^{(1-\sigma)^{-1}}}$$

for some constants P, P' independent of β (cf. [17, Theorem 3.2]). Furthermore, the semigroup generated by $-A + f_\beta(A)$ has growth order independent of β (see the earlier Remark 2.5), and so as noted in Remark 3.1, one may prove regularization the standard way.

Example (3) $f_\beta(A) = Ae^{-\beta A}$. This approximation makes direct use of the holomorphic semigroup e^{-tA} generated by $-A$. By classic semigroup theory, $f_\beta(A) \in B(X)$ and satisfies $\|f_\beta(A)\| \leq \frac{M}{\beta}$ where M is a constant independent of

β (cf. [29, Theorem 2.5.2 (d)]). Therefore, $f_\beta(A)$ generates a uniformly continuous semigroup satisfying

$$\|e^{tf_\beta(A)}\| \leq e^{tM\beta^{-1}}$$

for each $t \geq 0$. Next, since e^{-tA} is uniformly bounded,

$$\|(-A + Ae^{-\beta A})x\| = \|(I - e^{-\beta A})Ax\| \leq \int_0^\beta \|e^{-tA}A^2x\|dt \leq \beta M'\|A^2x\|$$

for all $x \in Dom(A^2)$. Therefore, Condition (A$^+$) (i) is satisfied. Also, for $w = re^{\pm i\alpha} \in S_\nu$,

$$\begin{aligned} Re(-w + f_\beta(w)) &= Re(-w + we^{-\beta w}) \\ &= -r(\cos\alpha - e^{-\beta r\cos\alpha}(\cos(\alpha - \beta r\sin\alpha))) \\ &\leq re^{-\beta r\cos\alpha}\cos(\alpha - \beta r\sin\alpha) \leq r \end{aligned}$$

and so Condition (A$^+$) (ii) is satisfied with $C = 1$ and any $\frac{\pi}{2} - \theta < \nu < \frac{\pi}{2}$.

Example (4) $f_\beta(A) = (\ln 2)^{-1}A\ \mathrm{Log}(1 + e^{-\beta A})$. Here we define $f_\beta(A)$ by the Dunford integral

$$f_\beta(A) = \frac{1}{2\ln 2\,\pi i}\int_{\Gamma_\alpha} w\ Log(1 + e^{-\beta w})(w - A)^{-1}dw$$

where the principal branch of the logarithm is taken [9, 27]. The calculations for Condition (A$^+$) here are involved and so are omitted, but we can refer to [11] for a complete analysis. There, it is found that $f_\beta(A)$ is a bounded operator with $\|f_\beta(A)\| \leq \frac{C_\alpha}{\beta}$ where C_α is a constant independent of β but dependent on α. Further, it may be shown that $f_\beta(A)$ satisfies Condition (A$^+$) (i) with a constant R independent of β and also Condition (A$^+$) (ii) with $C = \frac{\pi}{2\ln 2}$ and any $\frac{\pi}{2} - \theta < \nu < \frac{\pi}{2}$ [11, Lemma 2.4 and Proposition 3.1].

4 Comments on Regularization and Applications

Traditionally, regularization arguments are made where $u(t)$ is any solution of (1.1) with limited assumptions on the data (cf. [17, Definition 4.1]). As is the theme throughout this paper, our calculations rely on sufficient smoothness of the data φ and h as in (2.7). We outline a regularization-type argument under the assumptions of Theorem 2.7.

First, note that in each of the four examples listed in Sect. 3, with the exception of the case $1 < \sigma < 2$ in Example (2), $f_\beta(A)$ generates a C_0 semigroup satisfying $\|e^{tf_\beta(A)}\| \leq Pe^{tP'\beta^{-1}}$ for some constants P, P' independent of β. Therefore, if $v_\beta^\delta(t)$ is the solution of the approximate well-posed problem with perturbed data

$$\begin{aligned} v'(t) &= f_\beta(A)v + h(t), \quad 0 < t < T, \\ v(0) &= \varphi_\delta \end{aligned}$$

satisfying $\|\varphi - \varphi_\delta\| \leq \delta$, then

$$\begin{aligned} \|v_\beta(t) - v_\beta^\delta(t)\| =& \left\| \left(e^{tf_\beta(A)}\varphi + \int_0^t e^{(t-s)f_\beta(A)}h(s)ds \right) \right. \\ & \left. - \left(e^{tf_\beta(A)}\varphi_\delta + \int_0^t e^{(t-s)f_\beta(A)}h(s)ds \right) \right\| \\ =& \|e^{tf_\beta(A)}(\varphi - \varphi_\delta)\| \\ \leq & Pe^{TP'\beta^{-1}}\delta. \end{aligned}$$

Now, let $u(t)$ be a strong solution of (1.1) that satisfies the hypotheses of Theorem 2.7. Choosing $\beta = -2TP'(\ln\delta)^{-1}$ we have $\beta \to 0$ as $\delta \to 0$, and

$$\begin{aligned} \|u(t) - v_\beta^\delta(t)\| &\leq \|u(t) - v_\beta(t)\| + \|v_\beta(t) - v_\beta^\delta(t)\| \\ &\leq \tilde{C}\beta^{1-\omega(t)}M^{\omega(t)}\epsilon^{-2}N + Pe^{TP'\beta^{-1}}\delta \\ &= \tilde{C}\beta^{1-\omega(t)}M^{\omega(t)}\epsilon^{-2}N + P\sqrt{\delta} \\ &\to 0 \quad \text{as} \quad \delta \to 0 \end{aligned}$$

for all $0 \leq t < T$. Also, the case for $t = T$ may be addressed separately. Indeed, similar to the calculation (2.9), we obtain

$$\begin{aligned} \|u(T) - v_\beta^\delta(T)\| &\leq \|u(T) - v_\beta(T)\| + \|v_\beta(T) - v_\beta^\delta(T)\| \\ &\leq \beta K\epsilon^{-1}N + Pe^{TP'\beta^{-1}}\delta \\ &= \beta K\epsilon^{-1}N + P\sqrt{\delta} \\ &\to 0 \quad \text{as} \quad \delta \to 0. \end{aligned}$$

In the case of Sect. 3 Example (2) where $1 < \sigma < 2$, the only difference is that $e^{tf_\beta(A)}$ satisfies

$$\|e^{tf_\beta(A)}\| \leq Pe^{tP'\beta^{-(\sigma-1)^{-1}}}.$$

Nevertheless, a similar calculation may be made if instead we choose

$$\beta = (-2TP'(\ln \delta)^{-1})^{\sigma-1}.$$

Remark 4.1 While convergence depends continuously on β (and δ) in our calculations, the estimate (2.6) of Theorem 2.7 may depend severely upon ϵ which is related to the choice for Q in the estimate $(Q-T)\cos\alpha \geq C(T+\epsilon)$ as in the proof of Proposition 2.6. For instance, suppose $C = 1$ and $\cos\alpha = \frac{1}{2}$ (which is possible in Examples 1 and 3). If $T \geq 1$ and $\epsilon = 0.5$, then we may choose $Q = 4T$ so that the conditions φ and $Ran(h)$ in $Dom(e^{4TA})$ with $e^{4TA}h \in L^1((0,T) : X)$ are required. If $\epsilon = 0.1$, then a smaller value for Q may be chosen which alleviates the restriction on φ and h. However, in this case the factor ϵ^{-2} in (2.6) obviously becomes much larger. In a different case, if $C = 0$ (which is possible in Example 2), then regardless of the value of α, we may choose Q arbitrarily close to T.

Remark 4.2 While our assumptions of the paper include that $-A$ generate a *bounded* holomorphic semigroup of angle θ with $0 \in \rho(A)$ for convenience, this is not the most general case. However, it is known that our assumptions imply that for $\theta' \in (0,\theta)$, there exists $\lambda \in \mathbb{R}$ such that $-A+\lambda$ is the infinitesimal generator of a bounded holomorphic semigroup of angle θ' with $0 \in \rho(A-\lambda)$ (cf. [31, Theorem X.53], [7]). Therefore, as in [10, 11, 17], we may still apply the results to the ill-posed backward heat equation in $L^p(\mathbb{R}^n)$, $1 < p < \infty$ ($A = -\Delta$) and more generally to PDE's defined by strongly elliptic differential operators of even order (cf. [29], [17, Section 5]).

Acknowledgments The author would like to thank the editors for the proceedings of the IWOTA 2019 and the referee for their helpful suggestions.

References

1. S. Agmon, L. Nirenberg, Properties of solutions of ordinary differential equations in Banach space. Comm. Pure Appl. Math. **16**, 121–151 (1963)
2. K.A. Ames, R.J. Hughes, Structural stability for ill-posed problems in Banach space. Semigroup Forum **70**, 127–145 (2005)
3. A.V. Balakrishnan, Fractional powers of closed operators and the semigroups generated by them. Pacific. J. Math. **10**, 419–437 (1960)
4. N. Boussetila, F. Rebbani, A modified quasi-reversibility method for a class of ill-posed Cauchy problems. Georgian Math. J. **14**, 627–642 (2007)
5. B. Campbell Hetrick, R.J. Hughes, Continuous dependence on modeling for nonlinear ill-posed problems. J. Math. Anal. Appl. **349**, 420–435 (2009)
6. D. Chen, B. Hofmann, J. Zou, Regularization and convergence for ill-posed backward evolution equations in Banach spaces. J. Differ. Eq. **265**, 3533–3566 (2018)
7. R. deLaubenfels, Entire solutions of the abstract Cauchy problem. Semigroup Forum **42**, 83–105 (1991)
8. R. deLaubenfels, C-semigroups and the Cauchy problem. J. Funct. Anal. **111**, 44–61 (1993)
9. N. Dunford, J. Schwartz, *Linear Operators, Part I* (Wiley, New York 1957)

10. M.A. Fury, Nonautonomous ill-posed evolution problems with strongly elliptic differential operators. Electron. J. Differ. Eq. **2013**(92), 1–25 (2013)
11. M.A. Fury, A class of well-posed approximations for ill-posed problems in Banach spaces. Commun. Appl. Anal. **23**(1), 97–14 (2019)
12. M.A. Fury, Logarithmic well-posed approximation of the backward heat equation in Banach space. J. Math. Anal. Appl. **475**, 1367–1384 (2019)
13. M. Fury, B. Campbell Hetrick, W. Huddell, Continuous dependence on modeling in Banach space using a logarithmic approximation, in *Mathematical and Computation Approaches in Advancing Modern Science and Engineering* (Springer, Cham, 2016)
14. A. Gorny, Contribution à l'étude des fonctions dérivables d'une variable réelle. Acta Math. **71**, 317–358 (1993)
15. M. Haase, *The Functional Calculus for Sectorial Operators* (Birkäuser Verlag, Basel, 2006)
16. Y. Huang, Modified quasi-reversibility method for final value problems in Banach spaces. J. Math. Anal. Appl. **340**, 757–769 (2008)
17. Y. Huang, Q. Zheng, Regularization for ill-posed Cauchy problems associated with generators of analytic semigroups. J. Differ. Eq. **203**, 38–54 (2004)
18. Y. Huang, Q. Zheng, Regularization for a class of ill-posed Cauchy problems. Proc. Amer. Math. Soc. **133–10**, 3005–3012 (2005)
19. K. Ito, B. Jin, *Inverse Problems: Tikhonov Theory and Algorithms* (World Scientific, Singapore, 2014)
20. R. Lattes, J.L. Lions, *The Method of Quasi-reversibility, Applications to Partial Differential Equations* (Elsevier, New York, 1969)
21. N.T. Long, A.P.N. Dinh, Approximation of a parabolic non-linear evolution equation backwards in time. Inverse Probl. **10**, 905–914 (1994)
22. A. Lorenzi, I.I. Vrabie, An identification problem for a linear evolution equation in a Banach space and applications. Discrete Contin. Dyn. Syst Ser. S **4**, 671–691 (2011)
23. I.V. Mel'nikova, General theory of the ill-posed Cauchy problem. J. Inverse Ill-posed Probl. **3**, 149–171 (1995)
24. I.V. Mel'nikova, A.I. Filinkov, *Abstract Cauchy Problems: Three Approaches*. Chapman & Hall/CRC Monographs and Surveys in Pure and Applied Mathematics, vol. 120 (Chapman & Hall, Boca Raton, 2001)
25. K. Miller, Stabilized quasi-reversibility and other nearly-best-possible methods for non-well-posed problems, in *Symposium on Non-Well-Posed Problems and Logarithmic Convexity*. Lecture Notes in Mathematics, vol. 316 (Springer, Berlin, 1973), pp. 161–176
26. K. Miller, Logarithmic convexity results for holomorphic semigroups. Pacific J. Math. **58**, 549–551 (1975)
27. V. Nollau, Über den logarithmus abgeschlossener operatoren in Banachschen Räumen. Acta Sci. Math. **30**, 161–174 (1969)
28. L.E. Payne, *Improperly Posed Problems in Partial Differential Equations*. CBMS Regional Conference Series in Applied Mathematics, vol. 22 (Society for Industrial and Applied Mathematics, Philadelphia, 1975)
29. A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations* (Springer, New York, 1983)
30. A.I. Prilepko, D.G. Orlovsky, I.A. Vasin, *Methods for Solving Inverse Problems in Mathematical Physics* (Dekker, New York, 2000)
31. M. Reed, B. Simon, *Methods of Modern Mathematical Physics, Vol. II: Fourier Analysis, Self-Adjointness* (Academic, New York, 1975)
32. W. Rudin, *Real and Complex Analysis*, 3rd edn. (McGraw-Hill, New York, 1987)
33. O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, F. Lenzen, *Variational Methods in Imaging*. Applied Mathematical Sciences, vol. 167 (Springer, New York, 2009)
34. T. Schuster, B. Kaltenbacher, B. Hofmann, K.S. Kazimierski, *Regularization Methods in Banach Spaces* (Walter de Gruyter, Berlin, 2012)
35. R.E. Showalter, The final value problem for evolution equations. J. Math. Anal. Appl. **47**, 563–572 (1974)

36. T.H. Skaggs, Z.J. Kabala, Recovering the release history of a groundwater contaminant. Water Resour. Res. **30**, 71–79 (1994)
37. D.D. Trong, N.H. Tuan, Stabilized quasi-reversibility method for a class of nonlinear ill-posed problems. Electron. J. Differ. Eq. **2008**(84), 1–12 (2008)
38. N.H. Tuan, D.D. Trong, On a backward parabolic problem with local Lipschitz source. J. Math. Anal. Appl. **414**, 678–692 (2014)
39. N.H. Tuan, D.D. Trong, T.H. Thong, N.D. Minh, Identification of the pollution source of a parabolic equation with the time-dependent heat conduction. J. Inequal. Appl. **2014**, 1–15 (2014)

A Closer Look at Bishop Operators

Eva A. Gallardo-Gutiérrez and Miguel Monsalve-López

Abstract The purpose of this work is to provide a survey, essentially self-contained, of those results mainly concerned with the study of *Bishop operators*, their (local) spectral properties and spectral invariant subspaces, whenever they do exist. Finally, we will discuss *Bishop-type operators*, addressing some open questions in this context.

Keywords Bishop operators · Invariant subspace problem · Spectral subspaces

Mathematics Subject Classification (2010) 47A15, 47B37, 47B38

1 Introduction

Given an irrational number $\alpha \in (0, 1)$, the *Bishop operator* T_α is defined on $L^p[0, 1)$, $1 \leq p \leq \infty$, by

$$T_\alpha f(t) = t f(\{t + \alpha\}), \qquad t \in [0, 1),$$

where $\{\cdot\}$ denotes the fractional part. Clearly every Bishop operator T_α is the product of two simple and well-understood operators, namely the multiplication operator M_t by the independent variable in $L^p[0, 1)$ and the composition operator C_{τ_α} induced by the symbol $\tau_\alpha(t) = \{t + \alpha\}$. Nevertheless, the structure of T_α is largely unknown for every irrational $\alpha \in (0, 1)$. In particular, it is unknown

E. A. Gallardo-Gutiérrez (✉)
Dpto. de Análisis Matemático y Matemática Aplicada, Facultad de Ciencias Matemáticas, Universidad Complutense de Madrid, Madrid, Spain
e-mail: eva.gallardo@mat.ucm.es

M. Monsalve-López
Instituto de Ciencias Matemáticas ICMAT (CSIC-UAM-UC3M-UCM), Madrid, Spain
e-mail: migmonsa@ucm.es

M. A. Bastos et al. (eds.), *Operator Theory, Functional Analysis and Applications*, Operator Theory: Advances and Applications 282,
https://doi.org/10.1007/978-3-030-51945-2_13

whether T_α has non-trivial closed invariant subspaces in $L^p[0,1)$, $1 \leq p < \infty$, for every irrational $\alpha \in (0,1)$. Indeed, according to Davie [15], these examples were suggested by Errett Bishop in the fifties as possible candidates for operators without non-trivial closed invariant subspaces, and therefore, for counterexamples to the *Invariant Subspace Problem*.

Bishop operators are particular instances of the so-called *weighted translation operators* defined in a more general measurable setting as follows: if $(X, \mathcal{F}, \mu)$ is a non-atomic measure space, $\phi \in L^\infty(X)$ and τ a measurable measure-preserving mapping from X to itself which is invertible with measurable inverse, then the weighted translation operator $W_{\phi,\tau}$ is defined by the equation

$$W_{\phi,\tau} f = \phi\,(f \circ \tau).$$

The class consisting of weighted translations operators is a large one, containing as particular examples bilateral weighted shifts and, therefore, roughly speaking, model operators.

Weighted translation operators were firstly studied by Parrott [26] in his Ph.D. thesis, analyzing the spectrum, numerical range and reducing subspaces of such operators. Indeed, Parrott computed the spectrum of Bishop operators showing, in particular, that it is the disc

$$\sigma(T_\alpha) = \left\{z \in \mathbb{C} : |z| \leq e^{-1}\right\},$$

independently of the irrational $\alpha \in (0,1)$ and moreover, T_α lacks point spectrum for every α. In 1973, Bastian [6] gave unitary invariants for some weighted translation operators and studied properties like subnormality and hyponormality among them. Later on, Petersen [27] showed some results on the commutant of weighted translation operators in an attempt to get a deeper insight in the general context. Nevertheless, many open questions are left and complete answers seem to be far away.

As far as Bishop operators concerns, one of the most striking result was proved by Davie [15] in 1974, who by means of a functional calculus approach, was able to show the existence of non-trivial closed invariant subspaces in $L^p[0,1)$ for T_α whenever α is a *non-Liouville irrational number* in $(0,1)$. Recall that an irrational α is a *Liouville number* if for every $n \in \mathbb{N}$ there exists an irreducible rational number p_n/q_n such that

$$\left|\alpha - \frac{p_n}{q_n}\right| < \frac{1}{q_n^n}.$$

Observe that by Jarník-Besicovitch Theorem (see [11, Section 5.5], for instance), Liouville irrationals form a set of vanishing Hausdorff dimension; so for almost every $\alpha \in (0,1)$, Davie's Theorem states that T_α has non-trivial closed invariant subspaces in $L^p[0,1)$. More indeed, T_α has non-trivial closed *hyperinvariant*

subspaces, that is, closed subspaces invariant under every linear bounded operator in the commutant of T_α.

In the nineties, extensions strengthening Davie's Theorem were due to Blecher and Davie [9] and MacDonald in [22] for Bishop-type operators, that is, weighted translation operators in $L^p[0,1)$ where $\tau = \tau_\alpha$:

$$W_{\phi,\tau_\alpha} f(t) = \phi(t)\, f(\{t+\alpha\}), \qquad t \in [0,1).$$

Once again, the brick wall consisted of Liouville irrationals, and despite of the efforts, the interesting extensions regarding the non-invertible weighted translation operators included many weights ϕ but neither Liouville number.

In 2008, Flattot [18], by refining the functional calculus approach, was able to provide a large class of irrationals $\alpha \in (0,1)$ including some Liouville numbers for which T_α has non-trivial closed hyperinvariant subspaces in $L^p[0,1)$ -for instance, the classical example of a Liouville number $\sum_{n=1}^{\infty} 10^{-n!}$.

Recently, in [12] the authors have extended the class of irrationals $\alpha \in (0,1)$ such that T_α has non-trivial closed hyperinvariant subspaces in $L^p[0,1)$ by considering arithmetical techniques to strengthen the analysis of certain functions associated to the functional calculus model. Indeed, for these Liouville numbers, Gallardo-Gutiérrez and Monsalve-López [20] have recently shown the existence of non-trivial *spectral subspaces*, revealing, indeed, the spectral nature of the hyperinvariant subspaces. Moreover, those Liouville numbers α left, that is, those for which it is open whether T_α has a non-trivial closed invariant subspace, are so extreme that the functional calculus approach fails to produce invariant subspaces (see [12, Theorem 4.1]).

The aim of this work is to survey the recent results at this regard and show how techniques in Operator Theory, Analytic Number Theory or Spectral Theory are linked together to produce, when it succeeds, invariant subspaces for such a *simple* family of operators as Bishop operators are. To that end, the rest of the manuscript is organized as follows. In Sect. 2 we deal with a first approach by studying the behaviour of the norm of the iterates T_α^n (clearly, they tend to 0 by means of Parrott's characterization of $\sigma(T_\alpha)$). In Sect. 3, we detail the main techniques and approaches previously mentioned to provide invariant subspaces for Bishop operators recalling the updated results in this context. In Sect. 4, our approach deals with local spectral properties fulfilled by Bishop operators T_α, independently of the irrational $\alpha \in (0,1)$. Note that though all Bishop operators T_α share the same spectrum, not all of them are known to possess invariant subspaces; and hence, a deeper insight in $\sigma(T_\alpha)$ could lead to study invariant subspaces by considering local spectral properties like the Dunford property (Property (C)) or the Bishop property (β). Indeed, in Sect. 5 the local spectral analysis is pushed further to provide *spectral subspaces* for all those Bishop operators that, up to now, are known to have non-trivial closed hyperinvariant subspaces. Finally, in Sect. 6, we deal with Bishop-type operators and their properties, ending up with some open questions in this context.

2 A First Approach: Understanding the Behaviour of $\|T_\alpha^n\|$

In this section, and as a preliminary stage, we determine explicitly the norm of the iterates of T_α acting on $L^p[0,1)$ for $1 \leq p \leq \infty$. In particular, it provides an insight of the behaviour of T_α depending on the irrational α.

Let n be a positive integer and denote by T_α^n the n-th iterate of T_α. A simple computation shows that

$$T_\alpha^n f(t) = t\,\{t+\alpha\}\cdots\{t+(n-1)\alpha\}\; f\left(\{t+n\alpha\}\right),$$

for any $L^p[0,1)$. Hence,

$$\|T_\alpha^n\| = S_{n-1}(\alpha)$$

where

$$S_n(\alpha) := \operatorname*{ess\,sup}_{t\in[0,1)}\; t\{t+\alpha\}\cdots\{t+n\alpha\}.$$

In Figs. 1 and 2, we represent the plots of $S_n(\alpha)$ for small values of n.

It turns out that the minimum of $S_m(\alpha)$ is reached at $\alpha = \frac{1}{m+1}$ and its value is

$$S_n\left(\frac{1}{n+1}\right) = \lim_{t\to 1^-}\; t\,\left\{t+\frac{1}{n+1}\right\}\cdots\left\{t+\frac{n}{n+1}\right\} = \frac{n!}{(n+1)^n}.$$

So, upon applying Stirling Formula, one deduces

$$\min_{\alpha\in[0,1)} S_n(\alpha) = \frac{n!}{(n+1)^n} \sim e^{-(n+1)}\sqrt{2\pi n}.$$

Accordingly, the following straightforward proposition follows:

Proposition 2.1 *For every irrational $\alpha \in [0,1)$ and n a positive integer*

$$\frac{(n-1)!}{n^{n-1}} \leq \|T_\alpha^n\| = S_{n-1}(\alpha) \leq \exp\left(\operatorname*{ess\,sup}_{t\in[0,1)} \sum_{j=0}^{n-1} \log\{t+j\alpha\}\right)$$

Observe that, by means of Birkoff Ergodic Theorem, it is possible to compute the

$$\lim_n \exp\left(\operatorname*{ess\,sup}_{t\in[0,1)} \frac{1}{n}\sum_{j=0}^{n-1} \log\{t+j\alpha\}\right)$$

and hence deduce that the spectral radius of T_α

$$r(T_\alpha) = \lim_n \|T_\alpha^n\|^{1/n} = e^{-1}$$

as Parrott showed in [26].

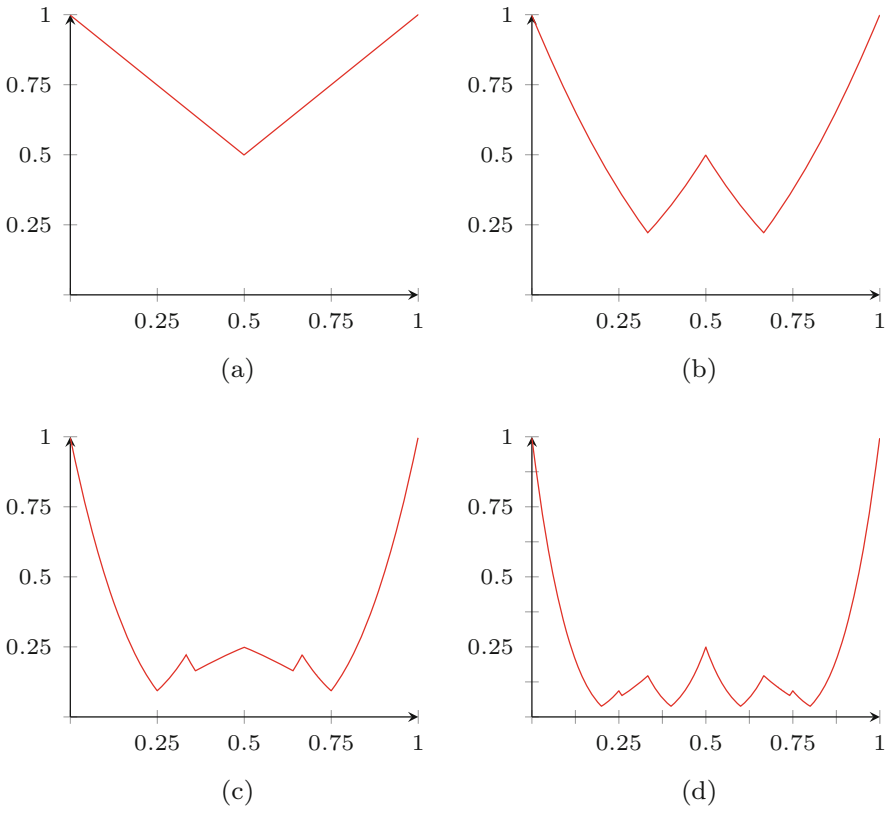


Fig. 1 $S_m(\alpha)$ for $m = 1, \ldots, 4$. (**a**) $S_1(\alpha)$ for $\alpha \in [0, 1)$. (**b**) $S_2(\alpha)$ for $\alpha \in [0, 1)$. (**c**) $S_3(\alpha)$ for $\alpha \in [0, 1)$. (**d**) $S_4(\alpha)$ for $\alpha \in [0, 1)$

3 Invariant Subspaces of Bishop Operators and a Theorem of Atzmon

The first goal of this section will be giving a detailed outline of the main techniques and approaches used to find invariant subspaces for Bishop operators. This will serve as a context in order to, thereupon, present in more detail our results, in which we considerably enlarge the set of known values α such that the Bishop operator T_α has invariant subspaces on each $L^p[0, 1)$. On the other hand, at the end of the section, we will show that in some sense, when our approach fails to produce invariant subspaces for T_α it is actually because the standard techniques cannot be applied anymore.

All the original results appearing in this section are included in the article [12].

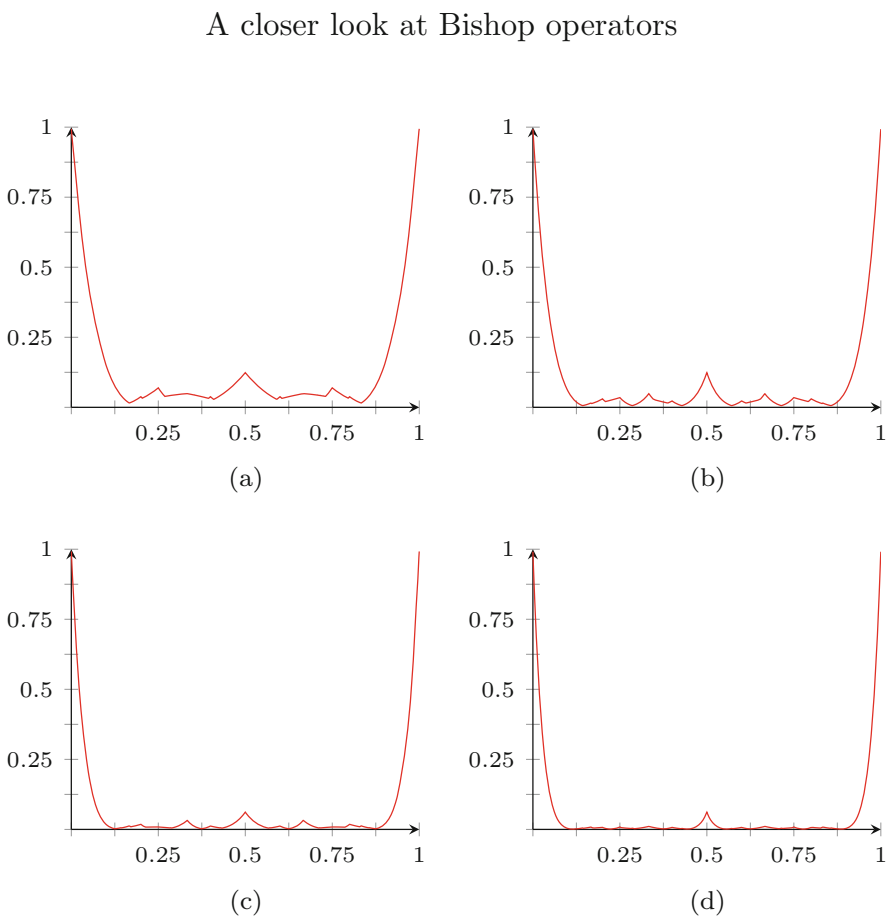


Fig. 2 $S_m(\alpha)$ for $m = 5, \ldots, 8$. (**a**) $S_5(\alpha)$ for $\alpha \in [0, 1)$. (**b**) $S_6(\alpha)$ for $\alpha \in [0, 1)$. (**c**) $S_7(\alpha)$ for $\alpha \in [0, 1)$. (**d**) $S_8(\alpha)$ for $\alpha \in [0, 1)$

3.1 Beurling Algebras and a Theorem of Atzmon

As it was aforementioned above, many authors have addressed the problem of seeking non-trivial closed invariant subspaces for all Bishop operators [12, 15, 18]. Ignoring the differences among all such distinct approaches within the literature, all of them rely essentially on the same idea: a functional calculus based on regular Beurling algebras which was properly formalized by Atzmon [4] in the mid eighties.

Given a sequence $\rho := (\rho_n)_{n\in\mathbb{Z}}$ in $[1, +\infty)$ with $\rho_0 = 1$ and such that

$$\rho_{m+n} \leq \rho_m \, \rho_n \quad \text{for every } m, n \in \mathbb{Z}, \tag{3.1}$$

$$\lim_{|n|\to\infty} \rho_n^{1/n} = 1; \tag{3.2}$$

we may consider its corresponding *Beurling algebra* $\mathcal{A}_\rho$ consisting of all continuous functions $f : \mathbb{T} \to \mathbb{C}$ with norm defined by

$$\|f\|_\rho := \sum_{n\in\mathbb{Z}} |\widehat{f}(n)|\, \rho_n < \infty,$$

where $(\widehat{f}(n))_{n\in\mathbb{Z}}$ denotes the sequence of Fourier coefficients of f. Endowed with the norm $\|\cdot\|_\rho$, the Beurling algebra $\mathcal{A}_\rho$ is a semi-simple commutative Banach algebra.

One of the most remarkable results regarding Beurling algebras $\mathcal{A}_\rho$ is the subsequent sufficient criterion, which seems to date back to Beurling [8], to determine the regularity of $\mathcal{A}_\rho$. Recall that a function algebra $\mathcal{A}$ on a compact space K is said to be *regular* if for all $p \in K$ and all compact subset $M \subseteq K$ with $p \notin M$, there exists $f \in \mathcal{A}$ such that $f(p) = 1$ and $f = 0$ on M.

Theorem 3.1 (Beurling [8]) *Let $\rho := (\rho_n)_{n\in\mathbb{Z}}$ be a real sequence satisfying both (3.1) and (3.2). Then, the Banach algebra $\mathcal{A}_\rho$ is regular whenever*

$$\sum_{n\in\mathbb{Z}} \frac{\log \rho_n}{1+n^2} < \infty. \tag{3.3}$$

Condition (3.3) is usually known as *Beurling condition* and it is closely related to the Denjoy-Carleman theorem on quasi-analytic classes [28]. Keeping the previous result in mind, it is natural to consider the following definition:

Definition 3.2 A sequence of real numbers $(\rho_n)_{n\in\mathbb{Z}}$ such that $\rho_0 = 1$ and $\rho_n \geq 1$ for all $n \in \mathbb{Z}$, is called a *Beurling sequence* if

$$\rho_{m+n} \leq \rho_m\, \rho_n \quad (\forall m, n \in \mathbb{Z}) \text{ and } \sum_{n\in\mathbb{Z}} \frac{\log \rho_n}{1+n^2} < +\infty.$$

One advantage of regularity in a function algebra is that it enables to construct two non-zero functions whose product is identically zero. This idea, combined with a functional calculus argument, provides a powerful method for obtaining invariant subspaces. Such a strategy, firstly studied by Wermer [29] for invertible operators, was lately refined by Atzmon [4] to the non-invertible case:

Theorem 3.3 (Atzmon [4]) *Let $T \in \mathcal{B}(X)$ be an operator on a complex Banach space X and suppose there exist two sequences $(x_n)_{n\in\mathbb{Z}}$ in X and $(y_n)_{n\in\mathbb{Z}}$ in X^* such that $x_0 \neq 0$, $y_0 \neq 0$ and*

$$T\, x_n = x_{n+1} \text{ and } T^*\, y_n = y_{n+1} \qquad (\forall n \in \mathbb{Z}).$$

Suppose further that both sequences $(\|x_n\|)_{n\in\mathbb{Z}}$ *and* $(\|y_n\|)_{n\in\mathbb{Z}}$ *are dominated by a Beurling sequence, and there is at least a* $\lambda \in \mathbb{T}$ *at which the following functions* G_x *and* G_y *do not both possess analytic continuation into a neighbourhood of* λ*:*

$$G_x(z) := \begin{cases} \sum_{n=1}^{\infty} x_{-n}\, z^{n-1} & \text{if } |z| < 1; \\ -\sum_{n=-\infty}^{0} x_{-n}\, z^{n-1} & \text{if } |z| > 1. \end{cases}$$

$$G_y(z) := \begin{cases} \sum_{n=1}^{\infty} y_{-n}\, z^{n-1} & \text{if } |z| < 1; \\ -\sum_{n=-\infty}^{0} y_{-n}\, z^{n-1} & \text{if } |z| > 1. \end{cases}$$

Then, either T *is a multiple of the identity or it has a non-trivial closed hyperinvariant subspace.*

As it was announced at the beginning of this subsection, Atzmon's Theorem turns out to be a useful machinery to produce invariant subspaces for many Bishop operators. Nevertheless, two things are still necessary: a right choice of the sequences $(x_n)_{n\in\mathbb{Z}}$ and $(y_n)_{n\in\mathbb{Z}}$ and a careful analysis on the growth of their norms. All this will be pursued in the following subsection.

3.2 Invariant Subspaces of Bishop Operators

Applying a novel version of these methods, Davie [15] was the first in obtaining positive results on the existence of invariant subspaces for some classes of Bishop operators. In particular, he showed that whenever $\alpha \in (0, 1)$ is a non-Liouville number, the corresponding Bishop operator T_α has a non-trivial hyperinvariant subspace on each $L^p[0, 1)$.

Further refinements due to Flattot [18] enlarged the class of such irrational numbers, embracing some Liouville numbers. More specifically, using the language of continued fractions, if $(a_j/q_j)_{j\geq 0}$ denote the convergents of α, the limit of Flattot's result ensures the existence of invariant subspaces for T_α on each $L^p[0, 1)$ whenever

$$\log q_{j+1} = O\big(q_j^{1/2-\varepsilon}\big) \quad \text{for every } \varepsilon > 0. \tag{3.4}$$

Turning to Bishop operators, up to date, our approach yields the most general result regarding the existence of invariant subspaces [12]:

Theorem 3.4 (Chamizo et al. [12]) *Let* $\alpha \in (0, 1)$ *be any irrational and* $(a_j/q_j)_{j\geq 0}$ *the convergents in its continuous fraction. If the following condition holds*

$$\log q_{j+1} = O\Big(\frac{q_j}{\log^3 q_j}\Big); \tag{3.5}$$

then, the Bishop operator T_α *has a non-trivial closed hyperinvariant subspace in* $L^p[0, 1)$ *for every* $1 \leq p \leq \infty$.

Again, the idea behind our address is exactly the same: a careful application of Atzmon's Theorem. For this, it shall be convenient to work with the operators

$$\widetilde{T}_\alpha := e\, T_\alpha.$$

This scalar multiple of T_α, far from being arbitrary, is determined by the fact $\sigma(\widetilde{T}_\alpha) = \overline{\mathbb{D}}$. Now, without going into much details, an insight in the proof may be the following: given any irrational $\alpha \in (0,1)$, in the sequel, consider the real functions

$$L_n(t) := \sum_{j=0}^{n-1} \big(1 + \log(\{t + j\alpha\})\big), \qquad (n \in \mathbb{N}).$$

It is plain that the functions $L_n(t)$ play a fundamental role in the understanding of the iterates of both T_α and T_α^*, since they codify their behaviours via the equations

$$\begin{aligned} \widetilde{T}_\alpha^n f(t) &= e^{L_n(t)} f(\{t + n\alpha\}), \\ \widetilde{T}_\alpha^{-n} f(\{t + n\alpha\}) &= e^{-L_n(t)} f(t), \end{aligned} \qquad (n \in \mathbb{N}), \tag{3.6}$$

and, in a similar way, for T_α^*. Note that Eq. (3.6) makes sense for every L^p-function despite of the fact that $\widetilde{T}_\alpha$ has not a bounded inverse.

In light of (3.6), it is not hard to figure out that, in order to control the growth of both iterates $\widetilde{T}_\alpha^n f$ and $(\widetilde{T}_\alpha^*)^n f$, it could be a good idea to construct ad hoc a function which kills each of the singularities arising from each summand of $L_n(t)$. To accomplish that, we may consider, for instance, the characteristic function of the set

$$\mathcal{B}_\alpha := \left\{ \frac{1}{20} < t < \frac{19}{20} \,:\, \langle t - n\alpha \rangle > \frac{1}{20n^2} \text{ for every } n \in \mathbb{Z} \setminus \{0\} \right\},$$

where $\langle t \rangle := \min\big(\{t\}, 1 - \{t\}\big)$ denotes the distance from t to the closest integer. First, note that, since the set $\mathcal{B}_\alpha$ has positive Lebesgue measure, it follows that $1_{\mathcal{B}_\alpha}$ does not vanish identically as an element of $L^p[0,1)$. On the other hand, it turns out that $|L_n(t)|$ can be bound properly in terms of the sequence $(q_j)_{j\geq 0}$ for $t \in \mathcal{B}_\alpha$, when $|n|$ is comparable to q_{j+1}.

We prefer not to deal with technicalities which are out of the scope of this survey. Anyway, as explained in [12], after some arithmetical estimations, we have:

Proposition 3.5 *Let $\alpha \in (0,1)$ be any irrational and $(a_j/q_j)_{j\geq 0}$ the convergents in its continuous fraction. Then, for every $n \in \mathbb{Z}$ with $|n| \leq q_{j+1}^{3/2}$, we have*

$$\begin{aligned} &\log\big(1 + \big\|\widetilde{T}_\alpha^n 1_{\mathcal{B}_\alpha}\big\|_\infty + \big\|(\widetilde{T}_\alpha^*)^n 1_{\mathcal{B}_\alpha}\big\|_\infty\big) \\ &\ll q_j + \frac{|n|}{q_j} \log q_j + \frac{|n| + q_{j+1}}{q_{j+1}} \log(|n| + 1). \end{aligned} \tag{3.7}$$

Now, as a consequence of the asymptotic bound (3.7), if we impose a suitable control on the growth of q_{j+1} in terms of q_j, we manage to find a Beurling sequence which dominates both $(\|T_\alpha^n 1_{\mathcal{B}_\alpha}\|)_{n\in\mathbb{Z}}$ and $(\|(\widetilde{T}_\alpha^*)^n 1_{\mathcal{B}_\alpha}\|)_{n\in\mathbb{Z}}$. More specifically, it may be seen that if condition (3.5) is satisfied, then the application of Atzmon's Theorem to Bishop operators can be achieved from picking the Beurling sequence

$$\rho_n = \exp\left(\frac{C\,|n|}{\log(2+|n|)(\log\log(4+|n|))^2}\right) \qquad (n\in\mathbb{Z}).$$

Observe that Theorem 3.4 relaxes significantly the condition imposed on α by Flattot (3.4), allowing the exponent 1 instead of 1/2 and quantifying the role of ε. Surprisingly, as we shall see hereafter, Theorem 3.4 is essentially the best possible result attainable from Atzmon's Theorem approach and any improvement seems to require different functional analytical techniques.

To conclude this subsection, we remark that it is possible to measure the difference between those cases covered by Davie [15], Flattot [18] and Theorem 3.4. For this, it suffices to consider the logarithmic Hausdorff dimension via the family of functions $(|\log x|^{-s})_{s\geq 0}$ (instead of the usual $(x^s)_{s\geq 0}$). With such a dimension, as a consequence of [11, Thm. 6.8], one has that the set of exceptions in Davie, Flattot and our case have dimension ∞, 4 and 2 respectively.

3.3 The Limits of Atzmon Theorem

Apart from being a refinement of Flattot/Davie approaches, the importance of Theorem 3.4 is evinced when it is understood in conjunction with the following statement:

Theorem 3.6 (Chamizo et al. [12]) *Let $\alpha\in(0,1)$ be an irrational not in*

$$\mathcal{M} := \left\{\alpha\in(0,1) \,:\, \log q_{j+1} = O\Big(\frac{q_j}{\log q_j}\Big)\right\}$$

and consider T_α on $L^p[0,1)$ for some $1\leq p<\infty$. Then, for every non-zero $f\in L^p[0,1)$, we have

$$\sum_{n\in\mathbb{Z}} \frac{\log\left(1+\left\|\widetilde{T}_\alpha^n f\right\|_p\right)}{1+n^2} = \infty.$$

Here, as an abuse of notation, for $f\in L^p[0,1)$ and $n>0$, $\|\widetilde{T}_\alpha^{-n} f\|$ denotes the norm of the n-th backward iterate $\widetilde{T}_\alpha^{-n} f$ whenever it belongs to $L^p[0,1)$ or ∞, otherwise.

Observe that this latter result shows that we cannot hope to improve much on Theorem 3.4, since at most, we shall be able to gain a power in $\log q_j$. Hence, it

establishes a threshold limit in the growth of the denominators of the convergents of α for the application of Atzmon's Theorem in order to find invariant subspaces for T_α.

Heuristically, in order to prove Theorem 3.6, we must show that, independently of $f \in L^p[0,1)$, one of the quantities $\|\widetilde{T}_\alpha^n f\|_p$ or $\|\widetilde{T}_\alpha^{-n} f\|_p$ is large enough for many values of $n \in \mathbb{N}$, inducing the series

$$\sum_{m\in\mathbb{Z}} \frac{\log\left(1+\left\|\widetilde{T}_\alpha^m f\right\|_p\right)}{1+m^2}$$

to be divergent. To do so, we consider the identity

$$\|\widetilde{T}_\alpha^n f\|_p^p + \|\widetilde{T}_\alpha^{-n} f\|_p^p = \int_0^1 \left(e^{pL_n(t-n\alpha)} + e^{-pL_n(t)}\right)|f(t)|^p\,dt, \tag{3.8}$$

and observe that, $\alpha \notin \mathcal{M}$ means exactly that α is pretty close to some rationals a/q. Roughly speaking, this implies that whenever n is comparable to q and $q \mid n$, we have a very accurate approximation between

$$L_n(t-n\alpha) \approx L_n\left(t-n\,\frac{a}{q}\right) = L_n(t).$$

But, in such a case, it results that the integral in (3.8) must be large unless $|L_n(t)|$ is small, which turns out to happen rarely.

4 Local Spectral Properties of Bishop Operators

The aim of the current section is identifying which local spectral properties are shared by all Bishop operators T_α, independently of the irrational $\alpha \in (0,1)$.

This apparent digression is justified by the facts presented in the preceding sections: while all Bishop operators T_α share the same spectrum, only a few of them are known to possess invariant subspaces; hence, might a deeper insight in $\sigma(T_\alpha)$ help us to deal with those cases not covered by Atzmon's Theorem? Actually, this naive idea of studying invariant subspaces via spectral subsets rarely succeeds; nevertheless, as we shall argue at the end of the survey, this will not be completely senseless in the case of Bishop operators.

All the original results presented in Sect. 4.2 may be found in [19]; on the other hand, those appearing in Sect. 4.3 are stated in [12].

4.1 Some Preliminaries on Local Spectral Theory

We begin this part by recalling some preliminaries regarding local spectral theory. In what follows, X will stand for an arbitrary complex Banach space, $\mathcal{L}(X)$ will denote the class of linear operators on X and $\mathcal{B}(X)$ the Banach algebra of linear bounded operators on X.

Given any operator $T \in \mathcal{B}(X)$, let $\sigma_T(x)$ denote the *local spectrum* of T at $x \in X$; i.e., the complement of the set of all $\lambda \in \mathbb{C}$ for which there exists an open neighbourhood $U_\lambda \ni \lambda$ and an analytic function $f : U_\lambda \to X$ such that

$$(T - zI) f(z) = x \qquad \text{for every } z \in U_\lambda. \tag{4.1}$$

The complement of $\sigma_T(x)$, which is denoted by $\rho_T(x)$, stands for the *local resolvent* of T at $x \in X$; and a function verifying (4.1) is called a *local resolvent function* nearby $\lambda \in \mathbb{C}$ for $x \in X$. In general, the uniqueness of the local resolvent function cannot be assumed. Thus, T satisfies the *single-valued extension property* (abbreviated as SVEP) if, fixed any $x \in X$ and $\lambda \in \rho_T(x)$, there exists a unique local resolvent function on a sufficiently small neighbourhood of λ. We highlight that any operator with $\sigma_p(T) = \emptyset$ enjoys the SVEP.

The notion of local spectrum allows us to gain a further knowledge on what constitutes each part of $\sigma(T)$, this may be accomplished via local spectral manifolds: given any subset $\Omega \subseteq \mathbb{C}$, the *local spectral manifold* $X_T(\Omega)$ is defined as

$$X_T(\Omega) := \big\{x \in X \,:\, \sigma_T(x) \subseteq \Omega\big\}.$$

In general, $X_T(\Omega)$ is always a T-hyperinvariant linear manifold; nevertheless, its closeness cannot be assured, even when $\Omega \subseteq \mathbb{C}$ is closed. Those operators $T \in \mathcal{B}(X)$ for which $X_T(F)$ is norm-closed in X for every closed subset $F \subseteq \mathbb{C}$ are said to satisfy the *Dunford property* or *Property* (C). We remark that Dunford property (C) implies SVEP.

An important class of Banach space-operators devised to resemble the spectral behaviour of normal operators is the class of decomposable operators. An operator $T \in \mathcal{B}(X)$ is said to be *decomposable* if given any open cover $\{U, V\}$ of $\mathbb{C}$, there exist two closed invariant subspaces $Y, Z \subseteq X$ such that

$$\sigma(T|Y) \subseteq U \text{ and } \sigma(T|Z) \subseteq V,$$

with $X = Y + Z$. Related to decomposability, two other weaker local spectral properties arise: the Bishop's property (β) and the decomposition property (δ). An operator $T \in \mathcal{B}(X)$ has the *property* (β) if, for every open set $U \subseteq \mathbb{C}$ and any sequence of analytic functions $f_n : U \to X$ such that

$$(T - zI) f_n(z) \to 0 \quad \text{as } n \to \infty$$

locally uniformly on U, then $f_n \to 0$ locally uniformly on U as well. We point out that property (β) entails both property (C) and SVEP. On the other hand, an operator $T \in \mathcal{B}(X)$ satisfies *property* (δ) if, for every open cover $\{U, V\}$ of $\mathbb{C}$ we have

$$X = \mathcal{X}_T(\overline{U}) + \mathcal{X}_T(\overline{V});$$

where, for each closed $F \subseteq \mathbb{C}$, the linear manifold $\mathcal{X}_T(F)$ is formed by all $x \in X_T(F)$ whose local resolvent function is globally defined on $\mathbb{C} \setminus F$. It may be seen that an operator is decomposable if and only if verifies both properties (β) and (δ).

To conclude this part, just mention that one of the most remarkable results in local spectral theory, due to Albrecht and Eschmeier [2], claims that properties (β) and (δ) are duals of each other. Furthermore, it happens that both properties (β) and (δ) can be comprehended in terms of decomposability: property (β) characterizes the restrictions of decomposable operators to invariant subspaces, while property (δ) characterizes the quotients of decomposable operators by invariant subspaces.

Finally, we disclose that properties (β) and (δ) can be combined with the Scott Brown techniques to produce invariant subspaces [16], extending substantially the classical result by S. W. Brown [10] on hyponormal operators.

4.2 Power-Regularity and the Local Spectral Radius of Bishop Operators

The programme to followed up along the subsequent pages consists on checking which of the aforemention local spectral properties are satisfied by all Bishop operators. In order to accomplish such a task, we shall make use of several estimations and previous results. The simplest case turns out to be the SVEP, since it follows immediately from the fact $\sigma_p(T_\alpha) = \emptyset$:

Proposition 4.1 *Let $\alpha \in (0, 1)$ be any irrational and T_α acting on $L^p[0, 1)$ for $1 \leq p \leq \infty$. Then, T_α satisfies the SVEP.*

Not surprisingly, regarding the remainder of local spectral properties, we shall need to work significantly harder and provide ourselves with new estimations concerning the behaviour of Bishop operators.

Given any operator $T \in \mathcal{B}(X)$, the Gelfand's formula for the spectral radius asserts that the sequence $(\|T^n\|^{1/n})_{n\geq 0}$ is always convergent and

$$r(T) := \max\left\{|\lambda| : \lambda \in \sigma(T)\right\} = \lim_{n\to\infty} \|T^n\|^{1/n}.$$

Nevertheless, if we aspire to wield a local version of Gelfand's formula, we will need to perform some modifications, since it turns out that $(\|T^n x\|^{1/n})_{n\geq 0}$ may be non-convergent. At this regard, Atzmon [5] introduced the notion of power-regularity: we recall that $T \in \mathcal{B}(X)$ is called *power-regular* if the sequence $(\|T^n x\|^{1/n})_{n\geq 1}$

is convergent for every $x \in X$. Furthermore, Atzmon proved a general criterion showing that a wide class of Banach space-operators, including decomposable operators, are power-regular. This forces to define the *local spectral radius* of T at $x \in X$ as

$$r_T(x) := \limsup_{n\to\infty} \|T^n x\|^{1/n}.$$

Not surprisingly, likewise in Gelfand's formula, the local spectral radius and the local spectrum may be related via the inequality

$$r_T(x) \geq \max\left\{|\lambda| : \lambda \in \sigma_T(x)\right\},$$

which can be replaced by an equality whenever T has the SVEP.

A general result due to Müller [24], asserts that given any $T \in \mathcal{B}(X)$, the equality $r_T(x) = r(T)$ must hold on a dense subset of X, which is indeed of second category [14]. Our next proposition exhibits that Bishop operators are extreme examples in this sense; more precisely, it shows that T_α is always power-regular with $r_{T_\alpha}(f) = r(T_\alpha)$ for every non-zero f:

Theorem 4.2 (Gallardo-Gutiérrez and Monsalve-López [19]) *Let $\alpha \in (0, 1)$ be any irrational and consider T_α acting on $L^p[0, 1)$ for $1 \leq p < \infty$. Then, for every non-zero $f \in L^p[0, 1)$, we have*

$$\lim_{n\to\infty} \|T_\alpha^n f\|_p^{1/n} = r(T_\alpha) = e^{-1}.$$

Moreover, the same holds for T_α^*.

We point out that the latter proposition can be proved for a considerably larger family of weighted translation operators [19, Cor. 2.3]. On the other hand, it is worth mentioning that its proof is mainly based on an application of ergodic theorems. Anyway, it entails the following striking corollary:

Corollary 4.3 *Let $\alpha \in (0, 1)$ be any irrational and $T_\alpha \in \mathcal{B}\big(L^p[0, 1)\big)$ for $1 \leq p < \infty$. Suppose M is a non-zero closed invariant subspace for T_α, then*

$$r(T_\alpha|M) = e^{-1}.$$

Given any Bishop operator T_α, as a somewhat direct consequence of the preceding results and the asymptotic bound provided by Proposition 3.5, we will be able to discard the fulfilment of the rest of the local spectral properties. In this subsection, we shall exclusively focus on the strongest ones: decomposability, property (β) and property (δ):

Theorem 4.4 (Gallardo-Gutiérrez and Monsalve-López [19]) *Let $\alpha \in (0, 1)$ be any irrational and consider T_α acting on $L^p[0, 1)$ for some $1 \leq p \leq \infty$. Then, T_α is not decomposable.*

Once we know Corollary 4.3, the Proof of Theorem 4.4 is pretty simple: suppose that T_α is decomposable in $L^p[0,1)$; then, fixed arbitrary $0 < r < s < e^{-1}$, we may consider the open cover

$$U = D(0,s) \quad \text{and} \quad V = \mathbb{C} \setminus \overline{D(0,r)}.$$

By means of Corollary 4.3, the unique closed invariant subspace $Y \subseteq L^p[0,1)$ such that $\sigma(T_\alpha|Y) \subseteq U$ would be $Y = \{0\}$; but this leads us to a contradiction, since $\sigma(T_\alpha) \not\subseteq V$.

Furthermore, it is possible to obtain a slightly stronger result by combining a similar argument to the previous one with the duality principle [2] between properties (β) and (δ):

Theorem 4.5 (Gallardo-Gutiérrez and Monsalve-López [19]) *Let $\alpha \in (0,1)$ be any irrational and consider T_α on $L^p[0,1)$ for $1 < p < \infty$. Then, T_α does not satisfy either property (β) or (δ).*

As a final remark, note that one could reasonably argue that, in a very vague sense, Proposition 4.2 and Corollary 4.3 seem to be suggesting that the most significant part of $\sigma(T_\alpha)$ is precisely its boundary. This intuition shall be explored more deeply and confirmed in more concrete terms throughout the rest of the article.

4.3 Dunford Property and a Dense Local Spectral Manifold

Throughout this subsection, we shall prove that Bishop operators can neither satisfy property (C); nevertheless, along the way, something much more stronger will be shown. More precisely, we shall see that, unlike for subsets $\Omega \subseteq \operatorname{int}(\sigma(T_\alpha))$, for which Proposition 4.2 ensured that $X_{T_\alpha}(\Omega) = \{0\}$; the local spectral manifold $X_{T_\alpha}(\partial\sigma(T_\alpha))$ turns out to be dense for every irrational α. To do so, we are going to split the proof into two parts: firstly, using (3.7), we will check that $\sigma_{T_\alpha}(1_{\mathcal{B}_\alpha})$ must be confined into $\partial\sigma(T_\alpha)$; on the other hand, an argument with L^∞ functions will let us to elucidate the density of $X_{T_\alpha}(\partial\sigma(T_\alpha))$.

Proposition 4.6 *Let $\alpha \in (0,1)$ be any irrational number and consider T_α acting on $L^p[0,1)$ for some $1 \leq p < \infty$. Then, both local spectra*

$$\sigma_{T_\alpha}(1_{\mathcal{B}_\alpha}) \subseteq \partial\sigma(T_\alpha) \text{ and } \sigma_{T_\alpha^*}(1_{\mathcal{B}_\alpha}) \subseteq \partial\sigma(T_\alpha).$$

In order to prove Proposition 4.6, consider the $L^p[0,1)$-valued analytic function given by

$$g_\alpha(z) := \sum_{n=1}^{\infty} \left(\widetilde{T}_\alpha^{-n} 1_{\mathcal{B}_\alpha}\right) \cdot z^{n-1}.$$

Now, note that g_α must be analytic on the open disk $|z| < 1$, since, by means of the asymptotic estimate (3.7), we have

$$\limsup_{n\to\infty} \left\| \widetilde{T}_\alpha^{-n} 1_{\mathcal{B}_\alpha} \right\|_p^{1/n} \leq 1.$$

Finally, it is routine to check that $(\widetilde{T}_\alpha - zI)g_\alpha(z) = 1_{\mathcal{B}_\alpha}$ for every $|z| < 1$.

An important remark is in order: given an operator with $\sigma_p(T) = \sigma_p(T^*) = \emptyset$, it may be seen that if Atzmon's Theorem may be applied to the sequences

$$x_n := T^n x \text{ and } y_n := (T^*)^n y \quad (n \in \mathbb{Z})$$

for a pair $x \in X$ and $y \in X^*$, the corresponding local spectra $\sigma_T(x)$ and $\sigma_{T^*}(y)$ must be inside $\partial\mathbb{D}$. Nevertheless, the converse is not true. This evinces the main gain of the asymptotic (3.7) with respect to the ones in the previous works [15] and [18], since Proposition 3.5 can be applied independently of the irrational α.

Now, once we are aware of $1_{\mathcal{B}_\alpha} \in X_{T_\alpha}(\partial\sigma(T_\alpha))$, we are in position to take another step ahead and prove the density of $X_{T_\alpha}(\partial\sigma(T_\alpha))$:

Theorem 4.7 (Chamizo et al. [12]) *Let $\alpha \in (0, 1)$ be any irrational number. Then, the local spectral manifold*

$$X_{T_\alpha}\big(\partial\sigma(T_\alpha)\big) = \big\{ f \in L^p[0, 1) \,:\, \sigma_{T_\alpha}(f) \subseteq \partial\sigma(T_\alpha) \big\}$$

is norm-dense in $L^p[0, 1)$ for each $1 \leq p < \infty$. In particular, T_α does not satify the Dunford property (C) on $L^p[0, 1)$ for $1 \leq p < \infty$.

In the argument of Theorem 4.7, we repeatedly exploit the fact that $1_{\mathcal{B}_\alpha}$ is a characteristic function. Firstly, note that standard bounds yield the inclusion

$$\left\{ g(t)1_{\mathcal{B}_\alpha}(t) \,:\, \operatorname*{ess\,sup}_{t\in[0,1)} |g(t)| < \infty \right\} \subseteq X_{T_\alpha}(\partial\sigma(T_\alpha)).$$

Therefore, by the density of L^∞ into L^p for each $1 \leq p < \infty$ and taking into account that the support of $1_{\mathcal{B}_\alpha}$ is precisely the set $\mathcal{B}_\alpha$, we have

$$\big\{ f \in L^p[0, 1) \,:\, \operatorname{supp}(f) \subseteq \mathcal{B}_\alpha \big\} \subseteq \overline{X_{T_\alpha}(\partial\sigma(T_\alpha))}.$$

Now, since $T_\alpha 1_{\mathcal{B}_\alpha} \in X_{T_\alpha}(\partial\sigma(T_\alpha))$ as well, we deduce a similar inclusion:

$$\left\{ tg(t)1_{\mathcal{B}_\alpha}(\{t + \alpha\}) \,:\, \operatorname*{ess\,sup}_{t\in[0,1)} |g(t)| < \infty \right\} \subseteq X_{T_\alpha}(\partial\sigma(T_\alpha)),$$

but, as the multiplication operator M_t is of dense-range, this again entails

$$\big\{ f \in L^p[0, 1) \,:\, \operatorname{supp}(f) \subseteq \tau_\alpha^{-1}(\mathcal{B}_\alpha) \big\} \subseteq \overline{X_{T_\alpha}(\partial\sigma(T_\alpha))}.$$

Repeating a similar discussion with $T_\alpha^n 1_{\mathcal{B}_\alpha}$ for each $n \in \mathbb{Z}$, we have

$$\operatorname*{span}_{m=-N,\ldots,N} \left\{ f \in L^p[0,1) \,:\, \operatorname{supp}(f) \subseteq \tau_\alpha^m(\mathcal{B}_\alpha) \right\} \subseteq \overline{X_{T_\alpha}(\partial\sigma(T_\alpha))},$$

but, since $\mathcal{B}_\alpha$ has strictly positive measure and τ_α is ergodic, the proof is done.

As it was promised before beginning this subsection, we may clearly appreciate that the meaning behind each part of $\sigma(T_\alpha)$ differs significantly. In this sense, it is plain that the boundary of $\sigma(T_\alpha)$ stores much more information about T_α than the interior. This idea will be raised into a new level in the following section, showing that those already known closed invariant subspaces for T_α can be characterized as the norm-closure of some local spectral manifolds related to $\partial\sigma(T_\alpha)$.

5 Spectral Decompositions of Bishop Operators

It is plain from the preceding section, that we cannot expect to encounter invariant subspaces for T_α when we consider certain local spectral manifolds. For example, if we pick an arbitrary set Ω inside in $\mathbb{C} \setminus \partial\sigma(T_\alpha)$, we fall too short since, in such a case, the associated local spectral manifold is trivial:

$$X_{T_\alpha}(\Omega) = \{0\};$$

on the other hand, if $\partial\sigma(T_\alpha) \subseteq \Omega$, we go too far since, in such case, the associated local spectral manifold is dense:

$$\overline{X_{T_\alpha}(\Omega)} = L^p[0,1),$$

for the corresponding $1 \leq p < \infty$. These two facts restrict significantly our quest of invariant subspaces for T_α via local spectral manifolds and indicate us that, our unique hope could be choosing subsets Ω which intersect the boundary of $\sigma(T_\alpha)$ but not covering it entirely. Surprisingly, at least for many values of α, this new attempt turns out to be successful, bringing about local spectral manifolds verifying

$$\{0\} \neq \overline{X_{T_\alpha}(\Omega)} \neq L^p[0,1).$$

Nevertheless, not everything shall be good news, since our address will depend on a local spectral variant of Atzmon's Theorem and the restriction imposed by Theorem 3.6 still remains. Anyway, considering these ideas, it seems fully justified to conjecture about the possibility of addressing the invariant subspace problem for general Bishop operators T_α using techniques borrowed from local spectral theory.

This section is mainly based on the article [20].

5.1 The Spectral Meaning Behind Atzmon's Theorem

As before, X will be any infinite-dimensional Banach space, $\mathcal{L}(X)$ will stand for the class of linear operators on X while $\mathcal{B}(X)$ will denote the Banach algebra of linear bounded operators on X. Such distinction between linear and linear bounded operators will be of particular interest along the current section.

As the title announces, throughout the first half of this section, our aim is unravelling the spectral meaning hidden behind the statement of Atzmon's Theorem (and related results in [3] and [7]) in order to, further on, apply it in the study of Bishop operators.

The preliminary step before understanding properly Atzmon's Theorem from a spectral perspective should be the following decomposability version of Wermer's theorem [29] given by Colojoară and Foiaş [13]:

Theorem 5.1 (Colojoară and Foiaş [13]) *Let $T \in \mathcal{B}(X)$ be an invertible operator on a complex Banach space X with $\sigma(T) \subseteq \mathbb{T}$ satisfying*

$$\sum_{n\in\mathbb{Z}} \frac{\log \|T^n\|}{1+n^2} < \infty. \tag{5.1}$$

Then, T is decomposable. In particular, if $\sigma(T)$ is not reduced to a singleton, the operator T has a non-trivial closed hyperinvariant subspace in X.

A scheme of the proof may be the following: firstly, note that the real sequence $\rho := (\|T^n\|)_{n\in\mathbb{Z}}$ is submultiplicative (3.1), satisfies the limit-type condition (3.2) and, as a byproduct of (5.1), defines a Beurling sequence. Therefore, its corresponding Beurling algebra $\mathcal{A}_\rho$ is regular. Thus, we may consider the continuous algebraic homomorphism determined by

$$\begin{aligned} \phi : \mathcal{A}_\rho &\to \mathcal{B}(X) \\ e^{\imath nt} &\mapsto T^n \end{aligned} \tag{5.2}$$

for every $n \in \mathbb{Z}$. Now, given any open cover $\{U, V\}$ of $\sigma(T)$, due to the regularity of $\mathcal{A}_\rho$, we can find a function $h \in \mathcal{A}_\rho$ such that

$$h \equiv 1 \text{ on } (\mathbb{C} \setminus \overline{V}) \cap \sigma(T) \quad \text{and} \quad h \equiv 0 \text{ on } (\mathbb{C} \setminus \overline{U}) \cap \sigma(T);$$

but, the manner in which h has been chosen causes that the ranges of the operators $\phi(h)$ and $I - \phi(h)$ split the spectrum of T in the following way:

$$\sigma\big(T \,|\, \overline{\phi(h)(X)}\big) \subseteq U \quad \text{and} \quad \sigma\big(T \,|\, \overline{(I-\phi(h))(X)}\big) \subseteq V,$$

and consequently, as we intended to show, T turns out to be decomposable.

At a glance, it sounds reasonable to believe that Atzmon's Theorem looks like a localized version of Colojoară and Foiaş decomposability theorem. Hence,

one may rightly conjecture that Atzmon's Theorem should also involve a sort of decomposition on the spectrum $\sigma(T)$; nevertheless, in this case, such spectral decomposition shall be much more subtle and less manageable. More specifically, when the invertibility of T is missing, we are forced to consider a decomposition of $\sigma(T)$ only with respect to a proper linear submanifold of X. For the sake of completeness, we explain it in more accurate terms: in the sequel, we are intended to modify the Proof of Theorem 5.1 with the aim of fitting it into the hypothesis required by Atzmon's Theorem.

Let $\rho := (\rho_n)_{n\in\mathbb{Z}}$ be an arbitrary Beurling sequence. Since T needs not to be invertible anymore, the algebraic homomorphism (5.2) must be replaced by

$$\begin{aligned} \phi : \mathcal{A}_\rho &\to \mathcal{L}(X) \\ e^{int} &\mapsto T^n \end{aligned} \qquad (n \in \mathbb{Z})$$

embracing some non-bounded linear operators. Nevertheless, in order to retrieve some appropriate spectral properties, we must restrict ourselves to work on a linear submanifold of X in which ϕ can be properly controlled; this is the goal behind the definition of the *continuity core*:

$$\mathcal{D}_\phi := \big\{x \in X \ : \ \text{the map } \mathcal{A}_\rho \to X,\ f \mapsto \phi(f)\,x \text{ is bounded}\big\},$$

which, trivially, must be constrained between

$$\{0\} \subseteq \mathcal{D}_\phi \subseteq \bigcap_{n\in\mathbb{Z}} \mathrm{Dom}(T^n).$$

By definition, the vectors of $\mathcal{D}_\phi$ satisfy the asymptotic bound

$$\|T^n\,x\| \ll \rho_n, \quad (n \in \mathbb{Z});$$

and, on addition, the Beurling condition

$$\sum_{n\in\mathbb{Z}} \frac{\log \rho_n}{1+n^2} < \infty$$

allows us to employ again the regularity of the Beurling algebra $\mathcal{A}_\rho$. Thus, for any open cover $\{U, V\}$ of $\sigma(T)$, the regularity of $\mathcal{A}_\rho$ enables us to pick a function $h \in \mathcal{A}_\rho$ such that

$$h \equiv 1 \text{ on } (\mathbb{C} \setminus \overline{V}) \cap \sigma(T) \quad \text{and} \quad h \equiv 0 \text{ on } (\mathbb{C} \setminus \overline{U}) \cap \sigma(T).$$

Now, given an arbitrary $x \in \mathcal{D}_\phi$, an adjustment of the previous estimations show that the local spectra

$$\sigma_T\big(\phi(h)\,x\big) \subseteq U \text{ and } \sigma_T\big((I-\phi(h))\,x\big) \subseteq V.$$

This clearly provides a local spectral decomposition of T with respect to the elements in $\mathcal{D}_\phi$. We resume the above discussion in the following proposition:

Proposition 5.2 *Let $\mathcal{A}_\rho$ be a regular Beurling algebra, X a complex Banach space and $\phi : \mathcal{A}_\rho \to \mathcal{L}(X)$ a algebra action with continuity core $\mathcal{D}_\phi$. Consider the operator $T := \phi(e^{it}) \in \mathcal{B}(X)$. Then, for every closed $F \subseteq \mathbb{C}$ with $\mathrm{int}(F) \neq \emptyset$,*

$$X_T(F) \supseteq \big\{\phi(h)(\mathcal{D}_\phi) \,:\, h \in \mathcal{A}_\rho \text{ with } \mathrm{supp}(h) \subseteq F\big\}.$$

Moreover, the inclusion

$$\mathcal{D}_\phi \subseteq X_T(U_1) + \ldots + X_T(U_n) \tag{5.3}$$

holds for every open cover $\{U_1, \ldots, U_n\}$ of $\mathbb{C}$.

Some remarks are in order: firstly, note that Eq. (5.3) reveals us that Atzmon's Theorem entails a sort of weak decomposability for T with respect to the submanifold $\mathcal{D}_\phi$. On the other hand, following the philosophy of [25], it is not hard to extend Proposition 5.2 to a much wider class of functional calculi. Such extension, performed using the language of Gelfand theory [20], allows us to consider arbitrary commutative Banach algebras as the model of the operator T.

But, certainly, the reader may wisely argue that Atzmon's Theorem granted the existence of non-trivial closed hyperinvariant subspaces for T while Proposition 5.2 is still far from guaranteeing them. Actually, and exclusively when the continuity core $\mathcal{D}_\phi \neq \{0\}$, Proposition 5.2 ensures the existence of non-zero local spectral manifolds. So, maybe the next question must be: how long is Atzmon's Theorem from Proposition 5.2? The keypoint is that, when the preceding techniques can be applied to both T and its adjoint T^* for two non-zero $x \in X$ and $y \in X^*$, we manage to construct non-trivial closed invariant subspaces due to the duality relation between local spectral manifolds [21, Prop. 2.5.1] given by

$$X_T(F) \subseteq {}^\perp X^*_{T^*}(G),$$

valid whenever T has the SVEP and for every pair of disjoint closed sets $F, G \subseteq \mathbb{C}$; where, as usual, ${}^\perp N$ denotes the *preannihilator* of a subset in $N \subseteq X^*$, i.e.

$${}^\perp N := \big\{x \in X \,:\, \varphi(x) = 0 \text{ for every } \varphi \in N\big\}.$$

As a consequence, we deduce the following local spectral version of Atzmon's Theorem:

Theorem 5.3 (Gallardo-Gutiérrez and Monsalve-López [20]) *Let $T \in \mathcal{B}(X)$ be an operator on a Banach space X such that $\sigma_p(T) = \sigma_p(T^*) = \emptyset$. Assume that there exist non-zero $x \in X$ and $y \in X^*$ for which*

$$\big\|T^n x\big\| \ll \rho_n \text{ and } \big\|(T^*)^n y\big\| \ll \rho_n \qquad (n \in \mathbb{Z})$$

for some Beurling sequence $\rho := (\rho_n)_{n\in\mathbb{Z}}$. Then, if $\sigma_T(x)\cup\sigma_{T^*}(y)$ is not a singleton, for every open subset $U \subseteq \mathbb{C}$ such that $U \cap \sigma_T(x) \neq \emptyset$ and $\sigma_{T^*}(y) \setminus \overline{U} \neq \emptyset$, we have

$$\{0\} \neq \overline{X_T(U)} \neq X.$$

In particular, T has a non-trivial closed hyperinvariant subspace.

In the same way as for Proposition 5.2, one has that Theorem 5.3 stands as the particular case regarding regular Beurling algebras of a much more general result [20, Thm. 2.5], aimed to extend Atzmon's Theorem to a wider class of models based on arbitrary commutative Banach algebras.

Unsurprisingly, our motivation in Atzmon's Theorem and its spectral implications are focused mainly on Bishop operators. This is the goal of the following subsection: analyse carefully what is happening to $\sigma(T_\alpha)$ and relate it, if possible, to the invariant subspaces of T_α.

5.2 Local Spectral Decomposition of Bishop Operators

According to the exposition held in Sect. 4, the lack of any profitable local spectral property seems to be an evident feature regarding Bishop operators. Nevertheless, as it shall be discussed along the following lines with the aid of the results exposed in the preceding subsection, this turns out to be just a kind of illusion (at least in some cases) caused by a misguided choice of the local spectral manifolds. Of course, the cornerstone in the development of such richer spectral theory for Bishop operators shall be the aforementioned local spectral version of Atzmon's Theorem.

Again, as we aim to apply Atzmon's Theorem to Bishop operators, we recall that the condition

$$\log q_{j+1} = O\Big(\frac{q_j}{\log^3 q_j}\Big), \tag{5.4}$$

which shall play a prominent role in the sequel. By means of Theorem 3.4, we know that given an irrational $\alpha \in (0, 1)$ satisfying (5.4), there exists a Beurling sequence $\rho = (\rho_n)_{n\in\mathbb{Z}}$ such that

$$\big\|\widetilde{T}_\alpha^n 1_{\mathcal{B}_\alpha}\big\| \ll \rho_n \text{ and } \big\|(\widetilde{T}_\alpha^*)^n 1_{\mathcal{B}_\alpha}\big\| \ll \rho_n.$$

In addition, by similarity, it may be checked that for every $\ell, m \in \mathbb{Z}$,

$$\sigma_{\widetilde{T}_\alpha}(e^{2\pi\imath\ell t}1_{\mathcal{B}_\alpha}) = e^{2\pi\imath\ell\alpha}\sigma_{\widetilde{T}_\alpha}(1_{\mathcal{B}_\alpha}),$$
$$\sigma_{\widetilde{T}_\alpha^*}(e^{2\pi\imath m t}1_{\mathcal{B}_\alpha}) = e^{-2\pi\imath m\alpha}\sigma_{\widetilde{T}_\alpha^*}(1_{\mathcal{B}_\alpha}).$$

Thus, due to the irrationality of α, it is plain that provided any open subset $U \subseteq \mathbb{C}$ such that $U \cap \mathbb{T} \neq \mathbb{T}$ and $\mathbb{T} \setminus \overline{U} \neq \emptyset$, we may find $\ell, m \in \mathbb{Z}$ satisfying both conditions

$$U \cap \sigma_{\widetilde{T}_\alpha}(e^{2\pi\imath\ell t} 1_{\mathcal{B}_\alpha}) \neq \emptyset \text{ and } \sigma_{\widetilde{T}^*_\alpha}(e^{2\pi\imath m t} 1_{\mathcal{B}_\alpha}) \setminus \overline{U} \neq \emptyset.$$

Therefore, as a corollary of Theorem 5.3, the above discussion leads us to the following local spectral version of Theorem 3.4:

Theorem 5.4 (Gallardo-Gutiérrez and Monsalve-López [20]) *Let $\alpha \in (0, 1)$ be an irrational satisfying (5.4). Suppose that T_α acts on $L^p[0, 1)$ for a fixed $1 \leq p < \infty$. Then, given any open subset $U \subseteq \mathbb{C}$ such that $U \cap \partial\sigma(T_\alpha) \neq \emptyset$ and $\partial\sigma(T_\alpha) \setminus \overline{U} \neq \emptyset$, we have*

$$\{0\} \neq \overline{X_{T_\alpha}(U)} \neq L^p[0, 1).$$

In particular, T_α has a non-trivial closed hyperinvariant subspace.

Although, in mere quantitative terms both Theorems 3.4 and 5.4 cover exactly the same cases; clearly, the advantage of the latter result with respect to Theorem 3.4 is that it allows us to identify the spectral nature of the invariant subspaces involved.

In addition, an application of Proposition 5.2 in conjunction with some of the arguments appearing in the Proof of Theorem 4.7, shows the following sort of weak decomposability fulfilled by all Bishop operators:

Theorem 5.5 (Gallardo-Gutiérrez and Monsalve-López [20]) *Let $\alpha \in (0, 1)$ be an irrational number satisfying (5.4). Then, for every open cover $\{U_1, \ldots, U_n\}$ of $\partial\sigma(T_\alpha)$, the algebraic sum*

$$X_{T_\alpha}(U_1) + \ldots + X_{T_\alpha}(U_n)$$

is norm-dense in $L^p[0, 1)$ for every $1 \leq p < \infty$.

Due to the absence of the Dunford property (C) for each Bishop operator T_α, the feature described within Theorem 5.5 cannot be properly considered a spectral decomposition. Anyway, it clearly suggests that Bishop operators verify much more interesting spectral properties than it seemed a priori and invites to study certain weaker local spectral decompositions for T_α than the ones mentioned in Sect. 4.

Finally, we remind that in Theorem 3.6 we stated that, regarding Bishop operators T_α, there exists a threshold limit on α from which Atzmon's Theorem cannot be applied anymore. In the language of Proposition 5.2, such result may be translated saying that if α does not belong to

$$\mathcal{M} := \left\{\alpha \in (0, 1) \,:\, \log q_{j+1} = O\Big(\frac{q_j}{\log q_j}\Big)\right\}$$

then, for every algebra action $\phi : \mathcal{A}_\rho \to \mathcal{L}(L^p[0,1))$ with $\phi(1) = I$ and $\phi(e^{int}) = \widetilde{T}_\alpha^n$ from a regular Beurling algebra $\mathcal{A}_\rho$, the corresponding continuity core must be $\mathcal{D}_\phi = \{0\}$.

On the other hand, it seems reasonable to expect that, for those cases T_α uncovered by Atzmon's Theorem (despite the absence of a profitable functional calculus), their local spectral properties remain identical and rich enough to construct closed invariant subspaces. In this sense, we suggest that one feasible route to solve the invariant subspace problem for all Bishop operators could be understanding in depth their local spectral manifolds $X_{T_\alpha}(U)$ arising from open sets U intersecting the boundary of $\sigma(T_\alpha)$. To support our question in the positive, we recall that there exist decomposable operators $T \in \mathcal{B}(X)$ whose spectral behaviour cannot be described in terms of a functional calculus from a suitable algebra [1]. We pose it as an open question:

Question 1 Let $\alpha \in (0,1)$ be any irrational number and $\varepsilon > 0$ sufficiently small. Is the local spectral manifold

$$X_{T_\alpha}\big(D(e^{-1}, \varepsilon)\big)$$

non-trivial and non-dense in $L^p[0,1)$ for every $1 \leq p < \infty$?

6 Bishop-Type Operators

In order to complete the survey, our purpose throughout this last section is discussing analogous results to the ones described above in a more general framework. In particular, we shall show that many Bishop-type operators also admit a rich spectral theory.

We begin by recalling that MacDonald in [22] was able to determine the spectrum of many weighted translations operators. In order to state his result, let us fix $(\Omega, \mathcal{B}, \mu)$ a non-atomic probability space which arises from the Borel sets of a compact metrizable space. Recall that any invertible measure-preserving transformation $\tau : \Omega \to \Omega$ is called *ergodic* for $(\Omega, \mathcal{B}, \mu)$ if given any $B \in \mathcal{B}$ such that $\tau^{-1}(B) = B$ then either $\mu(B) = 0$ or $\mu(B) = 1$ and, it is called *uniquely ergodic* whenever it is continuous and μ is the unique Ω-probability measure for which τ is ergodic. As an important instance, observe that whenever $\Omega = [0,1)$ is endowed with the usual Lebesgue measure, having in mind the natural identification with $\mathbb{T} = \{e^{2\pi i t} : t \in [0,1)\}$, the transformation $\tau_\alpha(t) := \{t + \alpha\}$ with $\alpha \notin \mathbb{Q}$ is uniquely ergodic.

We are now in position to state MacDonald's result. Recall that, given an arbitrary $\phi \in L^\infty[0,1)$ and τ any measure-preserving transformation with measurable inverse, the *Bishop-type operator* $T_{\phi,\tau}$ is defined as

$$T_{\phi,\tau} f := \phi \cdot (f \circ \tau)$$

on $L^p(\Omega, \mu)$ for each $1 \leq p \leq \infty$.

Theorem 6.1 (MacDonald [22]) *Let $(\Omega, \mathcal{B}, \mu, \tau)$ be a uniquely ergodic system and $\phi \in L^\infty(\Omega, \mu)$ continuous μ-almost everywhere. Consider $T_{\phi,\tau} \in \mathcal{L}\big(L^p(\Omega, \mu)\big)$ for $1 \leq p < \infty$ fixed. Then:*

- *If $0 \in ess\, ran(\phi)$, the spectrum of $T_{\phi,\tau}$ is*

$$\sigma(T_{\phi,\tau}) = \left\{\lambda \in \mathbb{C} : |\lambda| \leq \exp\left(\int_\Omega \log|\phi|\, d\mu\right)\right\}.$$

- *If $0 \notin ess\, ran(\phi)$, the spectrum of $T_{\phi,\tau}$ is*

$$\sigma(T_{\phi,\tau}) = \left\{\lambda \in \mathbb{C} : |\lambda| = \exp\left(\int_\Omega \log|\phi|\, d\mu\right)\right\}.$$

Clearly the assumptions on ϕ and τ determine the spectrum of $T_{\phi,\tau}$, but the general case is open and the following difficult question arises naturally:

Question 2 Determine the spectrum of weighted translation operators.

On the other hand, following Atzmon's theorem approach, MacDonald was able to establish the existence of non-trivial closed hyperinvariant subspaces for a large class of Bishop-type operators T_{ϕ,τ_α}, for the sake of brevity denoted by $T_{\phi,\alpha}$ (see [22, Theorem 2.5, Theorem 2.6]). Consequently, in the spirit of Sect. 5, for those cases covered by MacDonald, it is possible to obtain spectral decompositions similar to those appearing in Theorems 5.4 and 5.5. We address that in the following subsection.

6.1 Spectral Decomposition of Bishop-Type Operators

In order to state the results accurately, we must dwell on some minor technicalities for a moment: given a positive real number M, we recall from [22] the class of step functions

$$\mathcal{S}_M = \left\{S = \sum_{j=1}^{\ell} r_j \chi_{I_j} : \ r_j \in \mathbb{R},\ I_j \text{ intervals and } \sum_{j=1}^{\ell} |r_j| \leq M\right\}.$$

In addition, let $\mathcal{L}$ denote the class of all real functions $f \in L^\infty[0, 1)$ for which there exists $\gamma > 0$ and $K_f > 0$ (K_f depending exclusively on f) such that

$$\inf\{\|f - S\|_\infty : \ S \in \mathcal{S}_M\} < K_f \frac{1}{M}$$

for each positive real M. Thus, one immediately has:

Theorem 6.2 (Gallardo-Gutiérrez and Monsalve-López [20]) *Given any $\phi \in L^\infty[0,1)$ with $\log|\phi| \in \mathcal{L}$ and $\alpha \in (0,1)$ a non-Liouville irrational number, let $T_{\phi,\alpha}$ be the induced Bishop-type operator on $L^p[0,1)$ for some $1 < p < \infty$. Then, given any open set $U \subset \mathbb{C}$ such that $U \cap \partial\sigma(T_{\phi,\alpha}) \neq \emptyset$ and $\partial\sigma(T_{\phi,\alpha}) \setminus \overline{U} \neq \emptyset$, we have*

$$\{0\} \neq \overline{X_{T_{\phi,\alpha}}(U)} \neq L^p[0,1).$$

Theorem 6.3 (Gallardo-Gutiérrez, Monsalve-López [20]) *Given any $\phi \in L^\infty[0,1)$ with $\log|\phi| \in \mathcal{L}$ and $\alpha \in (0,1)$ a non-Liouville irrational number, let $T_{\phi,\alpha}$ be the induced Bishop-type operator on $L^p[0,1)$ for some $1 < p < \infty$. Then, for every open cover $\{U_1, \ldots, U_n\}$ of $\partial\sigma(T_{\phi,\alpha})$, the sum*

$$X_{T_{\phi,\alpha}}(U_1) + \ldots + X_{T_{\phi,\alpha}}(U_n) \tag{6.1}$$

is norm-dense in $L^p[0,1)$ for $1 < p < \infty$.

Note that, by means of Colojoară and Foiaş decomposability theorem, somewhat stronger conclusions hold for Theorem 6.3 whenever the operator $T_{\phi,\alpha}$ is invertible. In such a case, it happens that the operator $T_{\phi,\alpha}$ turns out to be decomposable and the algebraic sum (6.1) comprises the whole space $L^p[0,1)$. Moreover, further refinements established as well by MacDonald [23], showed decomposability for some Bishop-type operators $T_{\phi,\alpha}$ whenever the condition

$$\sum_{j=0}^{\infty} \frac{1}{q_j} \log\left(\frac{q_{j+1}}{q_j}\right) < \infty$$

is satisfied by the convergents $(a_j/q_j)_{j\geq 0}$ of the irrational α. Note that, as a matter of fact, the latter condition embraces some Liouville numbers.

In view of the above discussion, as it was previously posed for Bishop operators, it might happen that addressing of the invariant subspace problem for general Bishop-type operators based on identifying some of their local spectral manifolds was a fruitful approach. Similarly, we pose it as an open question:

Question 3 Let $\alpha \in (0,1)$ be any irrational number and $U \subseteq \mathbb{C}$ a *"sufficiently small"* open set intersecting $\partial\sigma(T_{\phi,\alpha})$. Is the local spectral manifold $X_{T_{\phi,\alpha}}(U)$ non-trivial and non-dense in $L^p[0,1)$ for every $1 \leq p < \infty$?

Acknowledgments Authors "Eva A. Gallardo-Gutiérrez and Miguel Monsalve-López" are partially supported by Plan Nacional I+D grant nos. MTM2016-77710-P (Spain) and PID2019-105979GB-I 00 (Spain) and by "Severo Ochoa Programme for Centres of Excellence in R&D" (SEV-2015-0554). In addition, M. Monsalve-López also acknowledges support of the grant *Ayudas de la Universidad Complutense de Madrid para contratos predoctorales de personal investigador en formación,* ref. no. CT27/16.

References

1. E. Albrecht, On two questions of I. Colojoară and C. Foiaş. Manuscripta Math. **25**, 1–15 (1978)
2. E. Albrecht, J. Eschmeier, Functional models and local spectral theory. Proc. London Math. Soc. **75**, 323–348 (1997)
3. A. Atzmon, Operators which are annihilated by analytic functions and invariant subspaces. Acta Math. **144**, 27–63 (1980)
4. A. Atzmon, On the existence of hyperinvariant subspaces. J. Operator Theor. **11**, 4–40 (1984)
5. A. Atzmon, Power-regular operators. Trans. Amer. Math. Soc. **347**, 3101–3109 (1995)
6. J.J. Bastian, Decomposition of weighted translation operators. Ph.D. Dissertation. Indiana University, 1973
7. B. Beauzamy, Sous-espaces invariants de type fonctionnel dans les espaces de Banach. Acta Math. **144**, 65–82 (1980)
8. A. Beurling, Sur les intégrales de Fourier absolument convergentes et leur application à une transformation fonctionelle, in *Ninth Scandinavian Mathematical Congress* (1938), pp. 345–366
9. D.P. Blecher, A.M. Davie, Invariant subspaces for an operator on $L^2(\Pi)$ composed of a multiplication and a translation. J. Operator Theory. **23**, 115–123 (1990)
10. S.W. Brown, Hyponormal operators with thick spectra have invariant subspaces. Ann. Math. **125**, 93–103 (1987)
11. Y. Bugeaud, *Approximation by Algebraic Numbers* (Cambridge University Press, Cambridge, 2004)
12. F. Chamizo, E.A. Gallardo-Gutiérrez, M. Monsalve-López, A. Ubis, Invariant subspaces for Bishop operators and beyond. Adv. Math. **375**, 107365 (2020)
13. I. Colojoară, C. Foiaş, *Theory of Generalized Spectral Operators* (CRC Press, Boca Raton, 1968)
14. J. Daneš, On local spectral radius. Časopis Pešt. Mat. **112**, 177–187 (1987)
15. A.M. Davie, Invariant subspaces for Bishop's operators. Bull. London Math. Soc. **6**, 343–348 (1974)
16. J. Eschmeier, B. Prunaru, Invariant subspaces for operators with property (β) and thick spectrum. J. Funct. Anal. **94**, 196–222 (1990)
17. J. Eschmeier, B. Prunaru, Invariant subspaces and localizable spectrum. Integr. Equ. Oper. Theory **42**, 461–471 (2002)
18. A. Flattot, Hyperinvariant subspaces for Bishop-type operators. Acta. Sci. Math. **74**, 689–718 (2008)
19. E. Gallardo-Gutiérrez, M. Monsalve-López, Power-regular Bishop operators and spectral decompositions. J. Operator Theory (in press). https://doi.org/10.7900/jot.2019sep21.2256
20. E. Gallardo-Gutiérrez, M. Monsalve-López, Spectral decompositions arising from Atzmon's hyperinvariant subspace theorem. (under review)
21. K. Laursen, M. Neumann, *An Introduction to Local Spectral Theory* (Clarendon Press, Oxford, 2000)
22. G.W. MacDonald, Invariant subspaces for Bishop-type operators. J. Funct. Anal. **91**, 287–311 (1990)
23. G.W. MacDonald, Decomposable weighted rotations on the unit circle. J. Operator Theory **35**, 205–221 (1996)
24. V. Müller, Local spectral radius formula for operators in Banach spaces. Czechoslovak Math. J. **38**, 726–729 (1988)
25. M. Neumann, Banach algebras, decomposable convolution operators, and a spectral mapping property, in *Function Spaces*. Marcel Dekker Series in Pure and Applied Mathematics, vol. 136 (Dekker, New York, 1992), pp. 307–323
26. S.K. Parrott, Weighted translation operators. ProQuest LLC, Ann Arbor, MI. Thesis (Ph.D)-University of Michigan, 1965

27. K. Petersen, The spectrum and commutant of a certain weighted translation operator. Math. Scand. **37**, 297–306 (1975)
28. W. Rudin, *Real and Complex Analysis* (Tata McGraw-Hill Education, New York, 2006)
29. J. Wermer, The existence of invariant subspaces. Duke Math. J. **19**, 615–622 (1952)

Products of Unbounded Bloch Functions

Daniel Girela

Abstract We give new constructions of pair of functions (f, g), analytic in the unit disc, with $g \in H^\infty$ and f an unbounded Bloch function, such that the product $g \cdot f$ is not a Bloch function.

Keywords Bloch function · Normal function · Blaschke product · Inner function · Minimal Besov space · Analytic mean Lipschitz spaces

Mathematics Subject Classification (2010) Primary 30D45; Secondary 30H30

1 Introduction and Statements of the Results

Let $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$ denote the open unit disc in the complex plane $\mathbb{C}$. The space of all analytic functions in $\mathbb{D}$ will be denoted by $\mathcal{H}ol(\mathbb{D})$.

For $0 < p \leq \infty$, the classical Hardy space H^p is defined as the set of all $f \in \mathcal{H}ol(\mathbb{D})$ for which

$$\|f\|_{H^p} \stackrel{\text{def}}{=} \sup_{0<r<1} M_p(r, f) < \infty,$$

where, for $0 < r < 1$ and $f \in \mathcal{H}ol(\mathbb{D})$,

$$M_p(r, f) = \left(\frac{1}{2\pi}\int_0^{2\pi} |f(re^{i\theta})|^p d\theta\right)^{1/p}, \ (0 < p < \infty);$$
$$M_\infty(r, f) = \sup_{\theta\in\mathbb{R}} |f(re^{i\theta})|.$$

D. Girela (✉)
Análisis Matemático, Facultad de Ciencias, Universidad de Málaga, Málaga, Spain
e-mail: girela@uma.es

M. A. Bastos et al. (eds.), *Operator Theory, Functional Analysis and Applications*,
Operator Theory: Advances and Applications 282,
https://doi.org/10.1007/978-3-030-51945-2_14

We mention [7] as a general reference for the theory of Hardy spaces.

A function $f \in \mathcal{H}ol(\mathbb{D})$ is said to be a Bloch function if

$$\|f\|_{\mathcal{B}} \stackrel{\text{def}}{=} |f(0)| + \sup_{z\in\mathbb{D}} (1-|z|^2)|f'(z)| < \infty.$$

The space of all Bloch functions is denoted by $\mathcal{B}$, it is a Banach space with the just defined norm $\|\cdot\|_{\mathcal{B}}$. It is well known that

$$H^\infty \subsetneq \mathcal{B}.$$

A typical example of an unbounded Bloch function is the function f defined by

$$f(z) = \log \frac{1}{1-z}, \quad z \in \mathbb{D}.$$

We mention [1] as a general reference for the theory of Bloch functions.

A function f which is meromorphic in $\mathbb{D}$ is said to be a normal function in the sense of Lehto and Virtanen [15] if

$$\sup_{z\in\mathbb{D}} (1-|z|^2)\frac{|f'(z)|}{1+|f(z)|^2} < \infty.$$

For simplicity, we shall let $\mathcal{N}$ denote the set of all holomorphic normal functions in $\mathbb{D}$. It is clear that any Bloch function is a normal function, that is, we have $\mathcal{B} \subset \mathcal{N}$. We refer to [1, 15] and [16] for the theory of normal functions. In particular, we remark here that if $f \in \mathcal{N}$, $\xi \in \partial\mathbb{D}$ and f has the asymptotic value L at ξ, (that is, there exists a curve γ in $\mathbb{D}$ ending at ξ such that $f(z) \to L$, as $z \to \xi$ along γ) then f has the non-tangential limit L at ξ.

Let us recall that if a sequence of points $\{a_n\}$ in the unit disc satisfies the *Blaschke condition*:

$$\sum_{n=1}^{\infty} (1-|a_n|) < \infty,$$

the corresponding Blaschke product B is defined as

$$B(z) = \prod_{n=1}^{\infty} \frac{|a_n|}{a_n} \frac{a_n - z}{1-\overline{a_n}z}.$$

Such a product is analytic in $\mathbb{D}$. In fact, it is an inner function, that is, an H^∞-function with radial limit of absolute value 1 at almost every point of $\partial\mathbb{D}$ (cf. [7, Chapter 2]).

If $\{a_n\}$ is a Blaschke sequence and there exists $\delta > 0$ such that

$$\prod_{m \neq n} \left| \frac{a_n - a_m}{1 - \overline{a}_n a_m} \right| \geq \delta, \quad \text{for all } n,$$

we say that the sequence $\{a_n\}$ is *uniformly separated* and that B is an *interpolating Blaschke product*. Equivalently,

$$B \text{ is an interpolating Blaschke product} \Leftrightarrow \inf_{n \geq 1} (1 - |a_n|^2)|B'(a_n)| > 0.$$

We refer to [7, Chapter 9] and [9, Chapter VII] for the basic properties of interpolating Blaschke products. In particular, we recall that an exponential sequence is uniformly separated and that the converse holds if all the a_k's are positive.

Lappan [14, Theorem 3] proved that if B is an interpolating Blaschke product and f is a normal analytic function in $\mathbb{D}$, the product $B \cdot f$ need not be normal. Lappan used this to show that $\mathcal{N}$ is not a vector space.

Lappan's result is a consequence of the following easy fact: if B is an interpolating Blaschke product whose sequence of zeros is $\{a_n\}$ and G is an analytic function in $\mathbb{D}$ with $G(a_n) \to \infty$, then $f = B \cdot G$ is not a normal function (and hence it is not a Bloch function either). This result has been used by several authors (see [3, 5, 10, 11, 17, 18]) to construct distinct classes of non-normal functions.

The author and Suárez proved in [12] a result of this kind dealing with Blaschke products with zeros in a Stolz angle but not necessarily interpolating, improving a result of [13]. Namely, Theorem 1 of [12] is the following.

Theorem A *Let B be an infinite Blaschke product whose sequence of zeros $\{a_n\}$ is contained in a Stolz angle with vertex at 1 and let G be analytic in $\mathbb{D}$ with $G(z) \to \infty$, as $z \to 1$. Then the function $f = B \cdot G$ is not a normal function.*

It is natural to ask whether it is possible to prove results similar to those described, substituting "Blaschke products" by some other classes of H^∞-functions. Our first result in this paper deals with the atomic singular inner function.

Theorem 1.1 *Let S be the atomic singular inner function defined by*

$$S(z) = \exp\left(-\frac{1+z}{1-z}\right), \quad z \in \mathbb{D}, \tag{1.1}$$

and let f be a Bloch function with

$$\lim_{z \to 1} |f(z)| = \infty.$$

Then the function F defined by $F(z) = S(z) f(z)$ is not a normal function (hence, it is not a Bloch function).

In particular, the function F defined by $F(z) = S(z) \cdot \log \frac{1}{1-z}$ $(z \in \mathbb{D})$ is not normal.

A Bloch function f satisfies

$$M_\infty(r, f) = \mathrm{O}\left(\log \frac{1}{1-r}\right)$$

and, consequently,

$$|f(r)| = \mathrm{o}\left(\exp \frac{1+r}{1-r}\right), \quad \text{as} \quad r \to 1^-.$$

Thus, Theorem 1.1 follows from the following result.

Theorem 1.2 *Let S be the singular inner function defined by (1.1) and let f be an analytic function in $\mathbb{D}$ satisfying:*

(i) $\lim_{z\to 1} |f(z)| = \infty$.
(ii) $|f(r)| = \mathrm{o}\left(\exp \frac{1+r}{1-r}\right)$, *as* $r \to 1^-$.

Then the function F defined by $F(z) = S(z) f(z)$ is not a normal function (hence, it is not a Bloch function).

For $g \in \mathcal{H}ol(\mathbb{D})$, the multiplication operator M_g is defined by

$$M_g(f)(z) \stackrel{\text{def}}{=} g(z) f(z), \quad f \in \mathcal{H}ol(\mathbb{D}), \ z \in \mathbb{D}.$$

Let us recall that if X and Y are two spaces of analytic function in $\mathbb{D}$ and $g \in \mathcal{H}ol(\mathbb{D})$ then g is said to be a multiplier from X to Y if $M_g(X) \subset Y$. The space of all multipliers from X to Y will be denoted by $M(X, Y)$ and $M(X)$ will stand for $M(X, X)$. Brown and Shields [4] characterized the space of multipliers of the Bloch space $M(\mathcal{B})$ as follows.

Theorem B *A function $g \in \mathcal{H}ol(\mathbb{D})$ is a multiplier of the Bloch space if and only if $g \in H^\infty \cap \mathcal{B}_{\log}$, where $\mathcal{B}_{\log}$ is the Banach space of those functions $f \in \mathcal{H}ol(\mathbb{D})$ which satisfy*

$$\|f\|_{B_{\log}} \stackrel{def}{=} |f(0)| + \sup_{z\in\mathbb{D}} (1 - |z|^2) \left(\log \frac{2}{1-|z|^2}\right) |f'(z)| < \infty.$$

Thus, if $g \in H^\infty \setminus \mathcal{B}_{\log}$ there exists a function $f \in \mathcal{B} \setminus H^\infty$ such that $g \cdot f \notin \mathcal{B}$. It is easy to see that the analytic Lipschitz spaces Λ_α $(0 < \alpha \leq 1)$ and the mean Lipschitz spaces Λ_α^p $(1 < p < \infty, 1/p < \alpha \leq 1)$ are contained in $M(\mathcal{B})$, we refer to [7, Chapter 5] for the definitions of these spaces. Let us simply recall here that

$$\Lambda_1^1 = \{f \in \mathcal{H}ol(\mathbb{D}) : f' \in H^1\}.$$

On the other hand, Theorem 1 of [8] shows the existence of a Jordan domain Ω with rectifiable boundary and $0 \in \Omega$, and such that the conformal mapping g from $\mathbb{D}$ onto Ω with $g(0) = 0$ and $g'(0) > 0$ does not belong to $\mathcal{B}_{\log}$. For this function g we have that $g \in \Lambda^1_1$ but g is not a multiplier of $\mathcal{B}$. Thus we have:

$$\Lambda^1_1 \not\subset M(\mathcal{B}).$$

In view of this and the results involving Blaschke products that we have mentioned above, it is natural to ask the following question:

Question 1.3 Is it true that for any given $f \in \mathcal{B} \setminus H^\infty$ there exists a function $g \in \Lambda^1_1$ such that $g \cdot f \notin \mathcal{B}$?

We shall show that the answer to this question is affirmative. Actually we shall prove a stronger result.

We let B^1 denote the minimal Besov space which consists of those functions $f \in \mathcal{H}ol(\mathbb{D})$ such that

$$\int_{\mathbb{D}} |f''(z)|\, dA(z) < \infty.$$

Here dA denotes the area measure on $\mathbb{D}$. Alternatively, the space B^1 can be characterized as follows (see [2]):

For $f \in \mathcal{H}ol(\mathbb{D})$, we have that $f \in B^1$ if and only there exist a sequence of points $\{a_k\}_{k=1}^\infty \subset \mathbb{D}$ and a sequence $\{\lambda_k\}_{k=0}^\infty \in \ell^1$ such that

$$f(z) = \lambda_0 + \sum_{k=1}^{\infty} \lambda_k \varphi_{a_k}(z), \quad z \in \mathbb{D}.$$

Here, for $a \in \mathbb{D}$, $\varphi_a : \mathbb{D} \to \mathbb{D}$ denotes the Möbius transformation defined by

$$\varphi_a(z) = \frac{a - z}{1 - \overline{a}z}, \quad z \in \mathbb{D}.$$

It is well known that $B^1 \subset \Lambda^1_1$ (see [2, 6]) and then our next result implies that the answer to question 1.3 is affirmative.

Theorem 1.4 *If $f \in \mathcal{B} \setminus H^\infty$ then there exists $g \in B^1$ such that $g \cdot f \notin \mathcal{B}$.*

The proofs of Theorems 1.2 and 1.4 will be presented in Sect. 2. We close this section noticing that throughout the paper we shall be using the convention that $C = C(p, \alpha, q, \beta, \dots)$ will denote a positive constant which depends only upon the displayed parameters $p, \alpha, q, \beta, \dots$ (which often will be omitted) but not necessarily the same at different occurrences. Moreover, for two real-valued functions E_1, E_2 we write $E_1 \lesssim E_2$, or $E_1 \gtrsim E_2$, if there exists a positive constant C independent of the arguments such that $E_1 \leq C E_2$, respectively $E_1 \geq C E_2$. If

we have $E_1 \lesssim E_2$ and $E_1 \gtrsim E_2$ simultaneously then we say that E_1 and E_2 are equivalent and we write $E_1 \asymp E_2$.

2 The Proofs

2.1 *Proof of Theorem 1.2*

For $0 < a < 1$, set $\Gamma_a = \{z \in \mathbb{D} : |z - a| = 1 - a\}$. If $z \in \Gamma_a$ then

$$\operatorname{Re} \frac{1+z}{1-z} = \frac{a}{1-a}$$

and, hence,

$$|S(z)| = \exp\left(\frac{-a}{1-a}\right), \quad z \in \Gamma_a.$$

This, together with (i), implies that

$$F(z) \to \infty, \quad \text{as } z \to 1 \text{ along } \Gamma_a.$$

Hence F has the asymptotic value ∞ at 1. On the other hand, (ii) implies that F has the radial limit 0 at 1. Then it follows that F is not normal. □

2.2 *Proof of Theorem 1.4*

Take $f \in \mathcal{B} \setminus H^\infty$. Set

$$\varphi(r) = M_\infty(r, f), \quad 0 < r < 1.$$

Clearly, $\varphi(r) \to \infty$, as $r \to 1$ and it is well known that

$$\varphi(r) = \mathrm{O}\left(\log \frac{1}{1-r}\right).$$

This implies that

$$(1-r)^2 \varphi(r) \to 0, \quad \text{as } r \to 1. \tag{2.1}$$

Choose a sequence of numbers $\{r_n\} \subset (0, 1)$ satisfying the following properties:

(i) $\{r_n\}$ is increasing.

(ii) $(1 - r_n)^2 \varphi(r_n) = o\left(\left(\frac{1 - r_{n-1}}{n}\right)^2\right)$, as $n \to \infty$.

(iii) $\varphi(r_n) \geq 2\varphi(r_{n-1})$, for all n.

(iv) $\frac{1 - r_{n+1}}{1 - r_n} \to 0$, as $n \to \infty$.

The existence of such a sequence is clear, bearing in mind (2.1) and the fact that $\varphi(r) \to \infty$, as $r \to 1$.

Set

$$\lambda_k = \varphi(r_k)^{-1/2}, \quad k = 1, 2, \ldots.$$

For each k, take $a_k \in \mathbb{D}$ with $|a_k| = r_k$ and $|f(a_k)| = \varphi(r_k)$. Using (iii), it follows that

$$\sum_{k=1}^{\infty} \lambda_k < \infty. \tag{2.2}$$

Define

$$g(z) = \sum_{k=1}^{\infty} \lambda_k \varphi_{a_k}(z), \quad z \in \mathbb{D}. \tag{2.3}$$

Using (2.2) we see that the sum in (2.3) defines an analytic function in $\mathbb{D}$ which belongs to B^1. Set

$$F(z) = g(z) f(z), \quad z \in \mathbb{D}.$$

Since $g \in H^\infty$ and $f \in \mathcal{B}$ we see that

$$|g(a_n) f'(a_n)| = O\left(\frac{1}{1 - |a_n|}\right). \tag{2.4}$$

On the other hand,

$$|g'(a_n) f(a_n)| \gtrsim I - II - III, \tag{2.5}$$

where

$$I = |f(a_n)|\lambda_n|\varphi'_{a_n}(a_n)|,$$

$$II \lesssim |f(a_n)| \sum_{k=1}^{n-1} \lambda_k \frac{1-|a_k|^2}{|1-\overline{a_k}\, a_n|^2},$$

$$III \lesssim |f(a_n)| \sum_{k=n+1}^{\infty} \lambda_k \frac{1-|a_k|^2}{|1-\overline{a_k}\, a_n|^2}.$$

Clearly,

$$I = |f(a_n)|\lambda_n|\varphi'_{a_n}(a_n)| \asymp \frac{\varphi(r_n)^{1/2}}{1-r_n}. \tag{2.6}$$

Using the definitions, the facts that φ and the sequence $\{r_n\}$ are increasing, and (ii), we obtain

$$\begin{aligned} II &\lesssim |f(a_n)| \sum_{k=1}^{n-1} \lambda_k \frac{1-|a_k|^2}{|1-\overline{a_k}\, a_n|^2} \\ &\lesssim \varphi(r_n) \sum_{k=1}^{n-1} \varphi(r_k)^{-1/2} \frac{1-|a_k|}{[(1-|a_k|)+(1-|a_n|)]^2} \\ &\lesssim \varphi(r_n) \sum_{k=1}^{n-1} \frac{1}{\varphi(r_k)^{1/2}(1-r_k)} \\ &\lesssim \frac{n\varphi(r_n)}{1-r_{n-1}} \\ &= \frac{\varphi(r_n)^{1/2}}{1-r_n}\, \varphi(r_n)^{1/2}\, \frac{n(1-r_n)}{1-r_{n-1}} \\ &= o\left(\frac{\varphi(r_n)^{1/2}}{1-r_n}\right). \end{aligned} \tag{2.7}$$

Likewise, using the definitions, the facts that φ and the sequence $\{r_n\}$ are increasing, (iii), and (iv), we obtain

$$\begin{aligned} III &\lesssim \varphi(r_n) \sum_{k=n+1}^{\infty} \frac{\varphi(r_k)^{-1/2}(1-r_k)}{[(1-r_k)+(1-r_n)]^2} \\ &\lesssim \varphi(r_n) \sum_{k=n+1}^{\infty} \varphi(r_k)^{-1/2} \frac{1-r_k}{(1-r_n)^2} \end{aligned}$$

$$\lesssim \varphi(r_n)\frac{1-r_{n+1}}{(1-r_n)^2}\sum_{k=n+1}^{\infty}\varphi(r_k)^{-1/2}$$
$$\lesssim \frac{\varphi(r_n)^{1/2}}{1-r_n}\cdot\frac{1-r_{n+1}}{1-r_n}$$
$$= o\left(\frac{\varphi(r_n)^{1/2}}{1-r_n}\right). \tag{2.8}$$

Using (2.5)–(2.8), and the fact that $\lim \varphi(r_n) = \infty$, we deduce that

$$(1-|a_n|)|g'(a_n)\,f(a_n)| \to \infty, \quad \text{as } n \to \infty.$$

This and (2.4) imply that F is not a Bloch function. □

Acknowledgments I wish to thank the referees for their careful reading of the article and for their suggestions to improve it.

This research is supported in part by a grant from "El Ministerio de Economía y Competitividad", Spain (PGC2018-096166-B-I00) and by grants from la Junta de Andalucía (FQM-210 and UMA18-FEDERJA-002).

References

1. J.M. Anderson, J. Clunie, Ch. Pommerenke, On Bloch functions and normal functions. J. Reine Angew. Math. **270**, 12–37 (1974)
2. J. Arazy, S.D. Fisher, J. Peetre, Möbius invariant function spaces. J. Reine Angew. Math. **363**, 110–145 (1985)
3. O. Blasco, D. Girela, M.A. Márquez, Mean growth of the derivative of analytic functions, bounded mean oscillation, and normal functions. Indiana Univ. Math. J. **47**, 893–912 (1998)
4. L. Brown, A.L. Shields, Multipliers and cyclic vectors in the Bloch space. Michigan Math. J. **38**, 141–146 (1991)
5. D.M. Campbell, Nonnormal sums and products of unbounded normal functions. II. Proc. Amer. Math. Soc. **74**, 202–203 (1979)
6. J.J. Donaire, D. Girela, D. Vukotić, On univalent functions in some Möbius invariant spaces. J. Reine Angew. Math. **553**, 43–72 (2002)
7. P.L. Duren, *Theory of H^p Spaces* (Academic, New York, 1970; Reprint: Dover, Mineola-New York, 2000)
8. P. Galanopoulos, D. Girela, R. Hernández, Univalent functions, VMOA and related spaces. J. Geom. Anal. **21**, 665–682 (2011)
9. J.B. Garnett, *Bounded Analytic Functions* (Academic, New York, 1981)
10. D. Girela, On a theorem of Privalov and normal funcions. Proc. Amer. Math. Soc. **125**, 433–442 (1997)
11. D. Girela, Mean Lipschitz spaces and bounded mean oscillation. Illinois J. Math. **41**, 214–230 (1997)
12. D. Girela, D. Suárez, On Blaschke products, Bloch functions and normal functions. Rev. Mat. Complut. **24**, 49–57 (2011)
13. D. Girela, C. González, J.A. Peláez, Multiplication and division by inner functions in the space of Bloch functions. Proc. Amer. Math. Soc. **134**, 1309–1314 (2006)

14. P. Lappan, Non-normal sums and products of unbounded normal function. Michigan Math. J. **8**, 187–192 (1961)
15. O. Lehto, K.I. Virtanen, Boundary behaviour and normal meromorphic functions. Acta Math. **97**, 47–65 (1957)
16. Ch. Pommerenke, *Univalent Functions* (Vandenhoeck und Ruprecht, Göttingen, 1975)
17. S. Yamashita, A nonnormal function whose derivative has finite area integral of order $0 < p < 2$. Ann. Acad. Sci. Fenn. Ser. A I Math. **4**(2), 293–298 (1979)
18. S. Yamashita, A nonnormal function whose derivative is of Hardy class H^p, $0 < p < 1$. Canad. Math. Bull. **23**, 499–500 (1980)

Birkhoff–James Orthogonality and Applications: A Survey

Priyanka Grover and Sushil Singla

Abstract In the last few decades, the concept of Birkhoff–James orthogonality has been used in several applications. In this survey article, the results known on the necessary and sufficient conditions for Birkhoff–James orthogonality in certain Banach spaces are mentioned. Their applications in studying the geometry of normed spaces are given. The connections between this concept of orthogonality, and the Gateaux derivative and the subdifferential set of the norm function are provided. Several interesting distance formulas can be obtained using the characterizations of Birkhoff–James orthogonality, which are also mentioned. In the end, some new results are obtained.

Keywords Orthogonality · Tangent hyperplane · Smooth point · Faces of unit ball · Gateaux differentiability · Subdifferential set · State on a C^*-algebra · Cyclic representation · Norm-parallelism · Conditional expectation

Mathematics Subject Classification (2010) Primary 15A60, 41A50, 46B20, 46L05, 46L08; Secondary 46G05, 47B47

1 Introduction

Let $(V, \|\cdot\|)$ be a normed space over the field $\mathbb{R}$ or $\mathbb{C}$. For normed spaces V_1, V_2, let $\mathscr{B}(V_1, V_2)$ denotes the space of bounded linear operators from V_1 to V_2 endowed with the *operator norm*, and let $\mathscr{B}(V)$ denotes $\mathscr{B}(V, V)$. Let $K(V_1, V_2)$ denotes the space of compact operators from V_1 to V_2. Let $\mathcal{H}$ be a Hilbert space over $\mathbb{R}$ or $\mathbb{C}$. If the underlying field is $\mathbb{C}$, the inner product on $\mathcal{H}$ is taken to be linear in the first coordinate and conjugate linear in the second coordinate. The notations $M_n(\mathbb{R})$ and $M_n(\mathbb{C})$ stand for $n \times n$ real and complex matrices, respectively.

P. Grover (✉) · S. Singla
Department of Mathematics, Shiv Nadar University, Dadri, Uttar Pradesh, India
e-mail: priyanka.grover@snu.edu.in; ss774@snu.edu.in

M. A. Bastos et al. (eds.), *Operator Theory, Functional Analysis and Applications*,
Operator Theory: Advances and Applications 282,
https://doi.org/10.1007/978-3-030-51945-2_15

Normed spaces provide a natural setting for studying geometry in the context of vector spaces. While inner product spaces capture the concept of the measure of an angle, orthogonality of two vectors can be described without knowing the notion of measure of angle. For example, a vector v is orthogonal to another vector u in $\mathbb{R}^n$ if and only if there exists a rigid motion T fixing the origin such that the union of rays $\overrightarrow{0u}, \overrightarrow{0T(v)}, \overrightarrow{0T(u)}$ minus the open ray $\overrightarrow{0v}$ is the one dimensional subspace generated by u.[1] This description of orthogonality by just using the notion of distance in $\mathbb{R}^n$ motivates to try and define orthogonality in normed spaces. In this approach, one can use the intuition about orthogonality in $\mathbb{R}^n$ to guess the results in general normed spaces and then prove them algebraically. This has been done in [3, 16, 44, 78].

One of the definitions for orthogonality in a normed space suggested by Roberts [78], known as *Roberts orthogonality*, is defined as follows: elements u and v are said to be (Roberts) orthogonal if $\|v + tu\| = \|v - tu\|$ for all scalars t. In [44, Example 2.1], it was shown that this definition has a disadvantage that there exist normed spaces in which two elements are Roberts orthogonal implies that one of the element has to be zero. In [44], two more inequivalent definitions of orthogonality in normed spaces were introduced. One of them is *isosceles orthogonality* which says that v is isosceles orthogonal to u if $\|v + u\| = \|v - u\|$. The other one is called *Pythagorean orthogonality*, that is, v is Pythagorean orthogonal to u if $\|v\|^2 + \|u\|^2 = \|v - u\|^2$. Note that if V is an inner product space, all the above mentioned definitions are equivalent to the usual orthogonality in an inner product space. Isosceles and Pythagorean orthogonalities have geometric intuitions for the corresponding definitions. In $\mathbb{R}^n$, two vectors are isosceles perpendicular if and only if their sum and difference can be sides of an isosceles triangle, and two vectors are Pythagorean perpendicular if there is a right triangle having the two vectors as legs. In [44], it was also proved that if u and v are two elements of a normed space, then there exist scalars a and b such that v is isosceles orthogonal to $av + u$ (see [44, Theorem 4.4]) and v is Pythagorean orthogonal to $bv + u$ (see [44, Theorem 5.1]). So these definitions don't have the above mentioned weakness of Roberts orthogonality.

In an inner product space, the following properties are satisfied by orthogonality. Let $u, u_1, u_2, v, v_1, v_2 \in V$.

1. *Symmetry:* If $v \perp u$, then $u \perp v$.
2. *Homogeneity:* If $v \perp u$, then $av \perp bu$ for all scalars a and b.
3. *Right additivity:* If $v \perp u_1$ and $v \perp u_2$, then $v \perp (u_1 + u_2)$.
4. *Left additivity:* If $v_1 \perp u$ and $v_2 \perp u$, then $(v_1 + v_2) \perp u$.
5. There exists a scalar a such that $v \perp av + u$. (In $\mathbb{R}^n$, this corresponds to saying that any plane containing a vector v contains a vector perpendicular to v.)

It is a natural question to study the above properties for any given definition of orthogonality. All the above definitions clearly satisfy symmetry. James [44, Theorem 4.7, Theorem 4.8, Theorem 5.2, Theorem 5.3] proved that if isosceles

[1] We learnt this characterization of orthogonality in $\mathbb{R}^n$ from Amber Habib.

or Phythagorean orthogonality satisfy homogeneity or (left or right) additivity, then V has to be an inner product space. These orthogonalities have been extensively studied in [3, 44, 78].

In [16], Birkhoff defined a concept of orthogonality, of which several properties were studied by James in [45]. An element v is said to be *Birkhoff–James* orthogonal to u if $\|v\| \leq \|v + ku\|$ for all scalars k. The analogy in $\mathbb{R}^n$ is that if two lines L_1 and L_2 intersect at p, then $L_1 \perp L_2$ if and only if the distance from a point of L_2 to a given point q of L_1 is never less than the distance from p to q. This definition clearly satisfies the homogeneity property. In [45, Corollary 2.2], it was shown that this definition also satisfies (4). But it lacks symmetry, for example, in $(\mathbb{R}^2, \|\cdot\|_{\max})$, where $\|(t_1, t_2)\|_{\max} = \max\{|t_1|, |t_2|\}$, take $v = (1, 1)$ and $u = (1, 0)$ ($v \perp u$ but $u \not\perp v$). It is not right additive, for example, in $(\mathbb{R}^2, \|\cdot\|_{\max})$, take $v = (1, 1)$, $u_1 = (1, 0)$ and $u_2 = (0, 1)$. It is also not left additive, for example, in $(\mathbb{R}^2, \|\cdot\|_{\max})$, take $v_1 = (1, 1)$, $v_2 = (0, -1)$ and $u = (1, 0)$.

Let W be a subspace of V. Then an element $v \in V$ is said to be Birkhoff–James orthogonal to W if v is Birkhoff–James orthogonal to w for all $w \in W$. A closely related concept is that of a *best approximation to a point in a subspace*. An element $w_0 \in W$ is said to be a best approximation to v in W if $\|v - w_0\| \leq \|v - w\|$ for all $w \in W$. Note that w_0 is a best approximation to v in W if and only if $v - w_0$ is Birkhoff–James orthogonal to W. These are also equivalent to saying that $\text{dist}(v, W) := \inf\{\|v - w\| : w \in W\} = \|v - w_0\|$. So v is Birkhoff–James orthogonal to W if and only if $\text{dist}(v, W)$ is attained at 0. Therefore the study of these concepts go hand in hand (see [89]). This is one of the reasons that this definition of orthogonality, even though not symmetric, is still being extensively studied in literature. Henceforward, orthogonality will stand for Birkhoff–James orthogonality.

Recently, a lot of work has been done in the form of applications of this concept of orthogonality and the main goal of this survey article is to bring all the related work under one roof. In Sect. 2, we mention the connections between orthogonality and geometry of normed spaces. We also deal with the question as to when the orthogonality is symmetric or (left or right) additive. This leads us to the study of various related notions like characterizations of smooth points and extreme points, subdifferential set, φ-Gateaux derivatives etc. In Sect. 3, characterizations of orthogonality in various Banach spaces are discussed along with some applications. In Sect. 4, these characterizations are used to obtain distance formulas in some Banach spaces. Some of the stated results are new and will appear in more detail in [35]. Theorems 4.5, 4.6 and 4.7 are the new results given with proofs only here.

2 Orthogonality and Geometry of Normed Spaces

A *hyperplane* is a closed subspace of codimension one. A connection between the concept of orthogonality and hyperplanes is given in the next theorem. An element v is orthogonal to a subspace W if and only if there exists a linear functional f on

V such that $\|f\| = 1$, $f(w) = 0$ for all $w \in W$ and $f(v) = \|v\|$ (see [89, Theorem 1.1, Ch. I]). This is equivalent to the following.

Theorem 2.1 ([45, Theorem 2.1]) *Let W be a subspace of V. Let $v \in V$. Then v is orthogonal to W if and only if there is a hyperplane H with v orthogonal to H and $W \subseteq H$.*

By the Hahn–Banach theorem and Theorem 2.1, it is easy to see that any element of a normed space is orthogonal to some hyperplane (see [45, Theorem 2.2]). The relation between orthogonality and hyperplanes is much deeper. We first recall some definitions. For $v \in V$, we say $S \subseteq V$ *supports* the closed ball $D[v, r] := \{x \in V : \|x - v\| \leq r\}$ if $\text{dist}(S, D[v, r]) = 0$ and $S \cap \text{Int } D[v, r] = \emptyset$. This is also equivalent to saying that $\text{dist}(v, S) = r$ (see [89, Lemma 1.3, Ch. I]). Let v_0 be an element of the boundary of $D[v, r]$. A hyperplane H is called a *support hyperplane to $D[v, r]$ at v_0* if H passes through v_0 and supports $D[v, r]$, and it is called a *tangent hyperplane to $D[v, r]$ at v_0* if H is the only support hyperplane to $D[v, r]$ at v_0. A real hyperplane is a hyperplane in V, when V is considered as a real normed space.

Theorem 2.2 ([89, Theorem 1.2, Ch. I]) *Let W be a subspace of V. Let $v \in V$. Then v is orthogonal to W if and only if there exists a support hyperplane to $D[v, r]$ at 0 passing through W if and only if there exists a real hyperplane which supports the closed ball $D[v, \|v\|]$ at 0 and passes through W.*

A direct consequence follows. If W is a non-trivial subspace of V, then 0 is the unique best approximation of v in W if and only if there exists a tangent hyperplane to $D[v, r]$ at 0 passing through W (see [89, Corollary 1.5, Ch. I]).

The above results are related to the questions as to when the orthogonality is (left or right) additive or symmetric. It was shown in [45, Theorem 5.1] that orthogonality is right additive in V if and only if for any unit vector $v \in V$, there is a tangent hyperplane to $D[v, \|v\|]$ at 0. There are other interesting characterizations for (left or right) additivity of orthogonality. To state them, some more definitions are required. A normed space V is called a *strictly convex* space if given any $v_1, v_2 \in V$, whenever $\|v_1\| + \|v_2\| = \|v_1 + v_2\|$ and $v_2 \neq 0$, then there exists a scalar k such that $v_1 = kv_2$. This is also equivalent to saying that if $\|v_1\| = \|v_2\| = 1$ and $v_1 \neq v_2$, then $\|v_1 + v_2\| < 2$. The norm $\|\cdot\|$ is said to be *Gateaux differentiable* at v if

$$\lim_{h \to 0} \frac{\|v + hu\| - \|v\|}{h}$$

exists for all $u \in V$.

Now we have the following characterizations for the orthogonality to be right additive in V.

Theorem 2.3 *The following statements are equivalent.*

1. *Orthogonality is right additive.*
2. *Norm is Gateaux differentiable at each nonzero point.*

3. *For $v \in V$, there exists a unique functional f of norm one on V such that $f(v) = \|v\|$.*
4. *For $v \in V$, there is a tangent hyperplane to $D[v, \|v\|]$ at 0.*

If V is a reflexive space, then the above are also equivalent to the following statements.

(5) *Any bounded linear functional on a given subspace of V has a unique norm preserving Hahn–Banach extension on V.*
(6) *The dual space V^* is strictly convex.*

Proof Equivalence of (1) and (2) is proved in [45, Theorem 4.2] and equivalence of (1), (3) and (4) is proved in [45, Theorem 5.1]. For a reflexive space, equivalence of (1) and (5) is given in [45, Theorem 5.7]. Equivalence of (5) and (6) is a routine exercise in functional analysis. □

Characterization of inner product spaces of dimension three or more can be given in terms of (left or right) additivity or symmetry of orthogonality. Birkhoff [16] gave a necessary and sufficient condition for a normed space of dimension at least three to be an inner product space, and examples to justify the restriction on the dimension. James [45, Theorem 6.1] showed that a normed space of dimension at least three is an inner product space if and only if orthogonality is right additive and symmetric if and only if the normed space is strictly convex and orthogonality is symmetric. Later, James improved his result and proved a much stronger theorem.

Theorem 2.4 ([46, Theorem 1, Theorem 2]) *Let V be a normed space of dimension at least three. Then V is an inner product space if and only if orthogonality is symmetric or left additive.*

A characterization of orthogonality to be symmetric or left additive in a normed space of dimension two can be found in [2]. Several other necessary and sufficient conditions for a normed space to be an inner product space are given in [2, 44]. This problem has also been extensively studied in [3, 89].

An element v is called a *smooth point* of $D[0, \|v\|]$ if there exists a hyperplane tangent to $D[v, \|v\|]$ at 0. We say v is a smooth point if it is a smooth point of $D[0, \|v\|]$. Equivalently, v is a smooth point if there exists a unique affine hyperplane passing through v which supports $D[0, \|v\|]$ at v (such an affine hyperplane is called the affine hyperplane tangent to $D[0, \|v\|]$ at v). A normed space is called *smooth* if all its vectors are smooth points. By Theorem 2.3, we get that orthogonality in a normed space is right additive if and only if the normed space is smooth. We also have that v is a smooth point if and only if the norm function is Gateaux differentiable at v:

Theorem 2.5 *Let $v \in V$. The norm function is Gateaux differentiable at v if and only if there is a unique $f \in V^*$ such that $\|f\| = 1$ and $f(v) = \|v\|$. In this case, the Gateaux derivative of the norm at v is given by Re $f(u)$ for all $u \in V$. In addition, for $u \in V$, we have that v is orthogonal to u if and only if $f(u) = 0$.*

Smooth points and this connection with Gateaux differentiability was studied in [1, 22, 52, 53] and many interesting results can be obtained as their applications. Let the space of continuous functions on a compact Hausdorff space X be denoted by $C(X)$ and let the space of bounded continuous functions on a normal space Ω be denoted by $C_b(\Omega)$. Kečkić [53, Corollary 2.2, Corollary 3.2] gave characterizations of smooth points in $C(X)$ and $C_b(\Omega)$. A characterization of smooth points in $\mathscr{B}(\mathcal{H})$ was given in [52, Corollary 3.3]. For $\mathcal{H}$ separable, Abatzoglou [1, Corollary 3.1] showed that the operators in $\mathscr{B}(\mathcal{H})$ of unit norm which are also smooth points are dense in the unit sphere of $\mathscr{B}(\mathcal{H})$. In $K(\mathcal{H})$, this result was first proved by Holub [42, Corollary 3.4]. Heinrich [40, Corollary 2.3] generalized this result for $K(V_1, V_2)$, where V_1 is a separable reflexive Banach space and V_2 is any normed space. He proved that the operators which attain their norm at a unique unit vector (upto scalar multiplication) are dense in $K(V_1, V_2)$.

In this paragraph, $\mathcal{H}$ is a separable Hilbert space. Schatten [86] proved that $D[0, 1]$ in $K(\mathcal{H})$ has no extreme points. In [42], the geometry of $K(\mathcal{H})$ and its dual $\mathscr{B}_1(\mathcal{H})$, the *trace class*, was studied by characterizing the smooth points and extreme points of their closed unit balls. It was shown in [42, Corollary 3.1] that the trace class operators of rank one and unit norm are exactly the extreme points of $D[0, 1]$ in $\mathscr{B}_1(\mathcal{H})$. The space $\mathscr{B}_1(\mathcal{H})$ is predual of $\mathscr{B}(\mathcal{H})$ and hence is isometrically isomorphic to a subspace of $\mathscr{B}(\mathcal{H})^*$. An interesting result in [1, Corollary 3.3] is that all the trace class operators of rank one and unit norm are also extreme points of $D[0, 1]$ in $\mathscr{B}(\mathcal{H})^*$. In [40], this study was continued to understand the geometry of $K(V_1, V_2)$, $\mathscr{B}(V_1, V_2)$ and the weak tensor product of V_1 and V_2, where V_1 and V_2 are Banach spaces. Characterizations of Gateaux differentiability and Fréchet differentiability of the norm at an operator T in these spaces were obtained. For Schatten classes of $\mathcal{H}$, this problem was addressed in [1, Theorem 2.2, Theorem 2.3]. In [1, Theorem 3.1], another characterization of Fréchet differentiability of the norm at T in $\mathscr{B}(\mathcal{H})$ was given, an alternative proof of which can be found in [72, Theorem 4.6]. In [40, Corollary 2.2], a necessary and sufficient condition for $0 \neq T \in K(V_1, V_2)$ to be a smooth point is obtained, where V_1 is a reflexive Banach space and V_2 is any Banach space. It is shown that such a T is a smooth point if and only if T attains its norm on the unique unit vector x_0 (up to a scalar factor) and Tx_0 is a smooth point. (This was proved for $K(\mathcal{H})$ in [42, Theorem 3.3].) Recently, as an application of orthogonality, it was shown in [72, Theorem 4.1, Theorem 4.2] that this characterization also holds when V_2 is any normed space (not necessarily complete).

If $T \in \mathscr{B}(V_1, V_2)$ attains its norm on the unique unit vector x_0 (up to a scalar factor) and Tx_0 is a smooth point of V_2, then T is said to satisfy *Holub's condition* (see [39]). Then Theorem 4.1 and Theorem 4.2 in [72] say that for a reflexive Banach space V_1 and any normed space V_2, smooth points of $K(V_1, V_2)$ are exactly those operators which satisfy Holub's condition. This characterization may not hold if T is not compact (see [39, Example (a)]) or when V_1 is not a reflexive space (see [39, Example (b), Example (c)]). In the case when V_1 is not a reflexive space, usually some extra condition is needed along with Holub's condition to characterize smooth points. For example, Corollary 1 in [37] states that for $1 < p, r < \infty$,

a necessary and sufficient condition for $T \in \mathscr{B}(l^p, l^r)$ to be a smooth point is that T satisfies Holub's condition and $\text{dist}(T, K(l^p, l^r)) < \|T\|$. As an application of orthogonality, it is proved in [60, Theorem 4.5] that for any normed spaces V_1, V_2, if $T \in \mathscr{B}(V_1, V_2)$ attains its norm and is a smooth point, then T satisfies Holub's condition and $\text{dist}(T, K(V_1, V_2)) < \|T\|$. The converse is true when V_1 is a reflexive Banach space and V_2 is any Banach space and $K(V_1, V_2)$ is an M-ideal in $\mathscr{B}(V_1, V_2)$ (see [60, Theorem 4.6]). It is an open question whether or not these extra assumptions on V_1 and V_2 are required. Some sufficient conditions, along with Holub's condition, for an operator to be smooth are also known when the underlying field is $\mathbb{R}$. If V_1 is a real Banach space and V_2 is a real normed space, one such condition for smooth points in $\mathscr{B}(V_1, V_2)$ is given in [72, Theorem 4.3]. When V_1 and V_2 are any real normed spaces, such conditions are given in [82, Theorem 3.2] and [84, Theorem 3.4]. The extra condition which along with Holub's condition gives the characterization for smoothness of any non zero norm attaining operator $T \in \mathscr{B}(V_1, V_2)$ (for any real normed spaces V_1, V_2) is obtained in [84, Theorem 3.3]. For further study of smooth points, we refer the readers to [36, 57, 73–75, 102].

Extreme points of $D[0, 1]$ are important because of Krein–Milman theorem. Along with the extreme points, the faces of $D[0, 1]$ in any normed space have also been of interest. (Note that the extreme points are exactly faces with a single element.) Let $M_n(\mathbb{R})$ or $M_n(\mathbb{C})$ be equipped with any *unitarily invariant norm*, $|||\cdot|||$ (that is, for any matrix A and U, U' unitary, $|||UAU'||| = |||A|||$). Then there is a unique *symmetric gauge function* Φ on $\mathbb{R}^n$ such that $|||A||| = \Phi((s_1(A), \ldots, s_n(A))$, where $s_i(A)$ are singular values of A arranged as $s_1(A) \geq \cdots \geq s_n(A)$. Ziętak [108, Theorem 5.1] showed that a necessary and sufficient condition for a matrix A to be an extreme point of the closed unit ball in $(M_n(\mathbb{R}), |||\cdot|||)$ is that $(s_1(A), \ldots, s_n(A))$ is an extreme point of the closed unit ball in $(\mathbb{R}^n, \Phi)$. This result was extended to $M_n(\mathbb{C})$ in [90, Theorem 1] (these results also follow from the results in [11]). Li and Schneider [58, Proposition 4.1] characterized the extreme points of $D[0, 1]$ in $M_n(\mathbb{R})$ and $M_n(\mathbb{C})$, equipped with the dual of an induced norm. In $\mathscr{B}(\mathcal{H})$, the extreme points of $D[0, 1]$ are exactly the isometries and the coisometries (see [38, p. 263]). It was proved in [95, Theorem 2.5] that $A \in \mathscr{B}(\mathcal{H})$ is an isometry or a coisometry if and only if $\|A\| = 1$ and A is *right symmetric* (for definition, see [30]). So the extreme points of $D[0, 1]$ in $\mathscr{B}(\mathcal{H})$ are precisely those operators which are of unit norm and are also right symmetric. There is also a concept of a *left symmetric operator*, the study of which can be found in [30, 71, 81, 96].

Theorem 2, Theorem 3 and Theorem 4 in [90] give characterizations of proper closed faces in $M_n(\mathbb{C})$, equipped with Schatten p-norms. Theorem 4.1 in [109] and the discussion above it give a characterization of faces of $D[0, 1]$ in $(M_n(\mathbb{C}), |||\cdot|||)$ as follows: F is a face of $(M_n(\mathbb{C}), |||\cdot|||)$ if and only if there exists $A \in M_n(\mathbb{C})$ such that F is a face of $\partial|||A|||^*$, the *subdifferential set* of $|||\cdot|||^*$ at A, where $|||\cdot|||^*$ is the dual norm of $|||\cdot|||$. In a normed space V, the subdifferential set of a continuous

convex function $g : V \to \mathbb{R}$ at $v \in V$ is denoted by $\partial g(v)$, and is defined as the set of bounded linear functionals $f \in V^*$ satisfying the below condition:

$$g(u) - g(v) \geq \text{Re } f(u - v) \quad \text{for all } u \in V.$$

It is a non-empty weak* compact convex subset of V^*. The below two propositions are easy to check. We refer the readers to [33, 41] for more details.

Proposition 2.6 *Let $v \in V$. Then*

$$\partial \|v\| = \{f \in V^* : Re\ f(v) = \|v\|, \|f\| \leq 1\}.$$

In particular, for $A \in M_n(\mathbb{C})$,

$$\partial |||A||| = \{G \in M_n(\mathbb{C}) : \text{Re tr}(G^*A) = |||A|||, |||G|||^* \leq 1\}.$$

Proposition 2.7 *Let $u, v \in V$. Then we have*

$$\lim_{t \to 0^+} \frac{\|v + tu\| - \|v\|}{t} = \max\{Re\ f(u) : f \in V^*, \|f\| = 1, f(v) = \|v\|\}.$$

Using this, Watson [97, Theorem 4] gave a characterization of $\partial ||| \cdot |||$ in the space $(M_n(\mathbb{R}), ||| \cdot |||)$. Ziętak [109, Theorem 3.1, Theorem 3.2] improved this result and showed the following.

Theorem 2.8 ([109, Theorem 3.1, Theorem 3.2]) *For $A \in M_n(\mathbb{C})$,*

$$\begin{aligned}\partial |||A||| = &\{U diag(d_1, \ldots, d_n) U'^* : A = U \Sigma U'^* \text{ is a singular value} \\ &\text{decomposition of } A, \sum s_i(A) d_i = |||A||| = \Phi((s_1, \ldots, s_n)), \\ &\Phi^*((d_1, \ldots, d_n)) = 1\}.\end{aligned}$$

In [98, Theorem 1], the above result was proved using a different approach. For the operator norm $\| \cdot \|$ on $M_n(\mathbb{C})$, we have the following.

Corollary 2.9 ([98, Example 3]) *For $A \in M_n(\mathbb{C})$,*

$$\partial \|A\| = conv\ \{uv^* : \|u\| = \|v\| = 1, Av = \|A\|u\},$$

where conv S denotes the convex hull of a set S.

Along the similar lines of [97] (that is, by using Proposition 2.7), the subdifferential set of the *Ky Fan k-norms*, $\| \cdot \|_{(k)}$, on $M_n(\mathbb{C})$ was obtained in [34]. In [15, 32–34], the subdifferential set was used to obtain characterizations of orthogonality in $M_n(\mathbb{C})$, equipped with various norms.

Actually the right hand derivative has a deeper connection with orthogonality as explored by Kečkić [51], where the author introduced the notion of *φ-Gateaux derivatives*: for $u, v \in V$ and $\varphi \in [0, 2\pi)$, the φ-Gateaux derivative of norm at v in the direction u is defined as

$$D_{\varphi,v}(u) = \lim_{t \to 0^+} \frac{\|v + te^{\iota\varphi}u\| - \|v\|}{t}.$$

These always exist for any two vectors u and v (see [51, Proposition 1.2]). A characterization of orthogonality follows.

Theorem 2.10 ([51, Theorem 1.4]) *Let $u, v \in V$. Then v is orthogonal to u if and only if*

$$\inf_{0 \le \varphi \le 2\pi} D_{\varphi,v}(u) \ge 0.$$

In [52, Theorem 2.4], the expression for the φ-Gateaux derivative of the norm on $\mathscr{B}(\mathcal{H})$ was obtained. Using the above proposition, a characterization of orthogonality in $\mathscr{B}(\mathcal{H})$ was given in [52, Corollary 3.1], which was first proved in [14] using a completely different approach. This characterization of orthogonality and many of its generalizations are the main content of the next section.

3 Characterizations and Applications of Orthogonality

Bhatia and Šemrl [14] gave characterizations of orthogonality in $\mathscr{B}(\mathcal{H})$ in terms of orthogonality of vectors in the underlying Hilbert space $\mathcal{H}$. These are given in the next two theorems. An independent proof of Theorem 3.2 was also given by Paul [70].

Theorem 3.1 ([14, Theorem 1.1]) *Let $A, B \in M_n(\mathbb{C})$. Then A is orthogonal to B if and only if there exists a unit vector $x \in \mathbb{C}^n$ such that $\|Ax\| = \|A\|$ and $\langle Ax | Bx \rangle = 0$.*

Let $\|\cdot\|$ be any norm on $\mathbb{C}^n$ or $\mathbb{R}^n$ and let $\|\cdot\|'$ be the corresponding induced norm on $M_n(\mathbb{C})$ or $M_n(\mathbb{R})$, respectively. It was conjectured in [14, Remark 3.3] that a matrix A is orthogonal to another matrix B in $(M_n(\mathbb{C}), \|\cdot\|')$ if and only if there exists a unit vector $x \in \mathbb{C}^n$ such that $\|Ax\| = \|A\|'$ and Ax is orthogonal to Bx in $(\mathbb{C}^n, \|\cdot\|)$. Li and Schneider [58, Example 4.3] gave an example to show that the conjecture is false in $M_n(\mathbb{C})$ as well as in $M_n(\mathbb{R})$. In $(M_n(\mathbb{C}), \|\cdot\|')$ (or $(M_n(\mathbb{R}), \|\cdot\|')$), a matrix A is said to satisfy *BŠ property* if for any matrix B, whenever A is orthogonal to B, there exists a unit vector x such that $\|Ax\| = \|A\|'$ and Ax is orthogonal to Bx in $(\mathbb{C}^n, \|\cdot\|)$ (or $(\mathbb{R}^n, \|\cdot\|)$) (see [80, Definition 1.1]). It was proved in [13] that $(\mathbb{R}^n, \|\cdot\|)$ is an inner product space if and only if every $A \in M_n(\mathbb{R})$ satisfies BŠ property. In [80, Theorem 2.2], it was shown that if $(\mathbb{R}^n, \|\cdot\|)$ is a smooth space

and $A \in M_n(\mathbb{R})$ is such that $\{x \in \mathbb{R}^n : \|x\| = 1, \|Ax\| = \|A\|'\}$ is a countable set with more than two points, then A does not satisfy BŠ property. Example 4.3 in [58] for $M_n(\mathbb{R})$ is a special case of this. It was shown in [79, Corollary 2.1.1] that if $A \in M_n(\mathbb{R})$ attains its norm at exactly two points, then A satisfies BŠ property. A generalization of this theorem can be found in [101, Theorem 3.1]. In [79, Theorem 2.1], another sufficient condition for A to satisfy BŠ property was given. If $(\mathbb{R}^n, \|\cdot\|)$ is a strictly convex space, then the collection of the matrices which satisfy BŠ property are dense in $M_n(\mathbb{R})$ (see [80, Theorem 2.6]).

Theorem 3.2 ([14, Remark 3.1], [70, Lemma 2]) *Let $\mathcal{H}$ be a complex Hilbert space. Let $A, B \in \mathscr{B}(\mathcal{H})$. Then A is orthogonal to B if and only if there exists a sequence of unit vectors $h_n \in \mathcal{H}$ such that $\|Ah_n\| \to \|A\|$ and $\langle Ah_n | Bh_n \rangle \to 0$, as $n \to \infty$.*

When $\mathcal{H}$ is an infinite dimensional space, one can't expect to get a single vector h in Theorem 3.2 such that $\|Ah\| = \|A\|$ and $\langle Ah|Bh\rangle = 0$. In fact it was proved in [72, Theorem 3.1] that for $A \in \mathscr{B}(\mathcal{H})$, the following are equivalent.

(a) For $B \in \mathscr{B}(\mathcal{H})$, A is orthogonal to B if and only if there exists a unit vector $h \in \mathcal{H}$ such that $\|Ah\| = \|A\|$ and $\langle Ah|Bh\rangle = 0$.
(b) There exists a finite dimensional subspace $\mathcal{H}_0$ of $\mathcal{H}$ such that

$$\{h \in \mathcal{H} : \|h\| = 1, \|A\| = \|Ah\|\} = \{h \in \mathcal{H}_0 : \|h\| = 1\} \text{ and } \left\| A|_{\mathcal{H}_0^\perp} \right\| < \|A\|.$$

It was noted in [14] that Theorem 3.1 is equivalent to saying that for $A, B \in M_n(\mathbb{C})$,

$$\operatorname{dist}(A, \mathbb{C}B) = \max\left\{ \left| \langle Ax|y\rangle \right| : \|x\| = \|y\| = 1 \text{ and } y \perp Bx \right\}. \tag{3.1}$$

It is natural to expect that in the infinite dimensional case, we should have for $A, B \in B(\mathcal{H})$,

$$\operatorname{dist}(A, \mathbb{C}B) = \sup\left\{ \left| \langle Ax|y\rangle \right| : \|x\| = \|y\| = 1 \text{ and } y \perp Bx \right\}. \tag{3.2}$$

This was indeed shown to be true in [14] by using the approach given in [5, p. 207]. We would like to point out that the book [5] deals with only separable spaces. However the arguments can be modified by replacing the sequence of finite rank operators converging pointwise to the identity operator by a net with this property. Since the same proof as in [5, p. 207] was used in the proof of Theorem 2.4 of [96], a similar modification is required there too.

Later, several authors have used different methods to prove Theorem 3.2. One of these techniques was given in [7, Remark 2.2] using a different distance formula [7, Proposition 2.1]. Another approach in [52, Corollary 3.1] uses Theorem 2.10 and the expression for the φ-Gateaux derivative of the norm on $\mathscr{B}(\mathcal{H})$ (which is given

in [52, Theorem 2.4]). Using Theorem 2.10, Wójcik [99] extended Theorem 3.1 for compact operators between two reflexive Banach spaces over $\mathbb{C}$:

Theorem 3.3 ([99, Theorem 3.1]) *Let V_1 and V_2 be reflexive Banach spaces over $\mathbb{C}$. Suppose $A, B \in K(V_1, V_2)$ and $A \neq 0$. Then A is orthogonal to B if and only if*

$$\min\{\max\{D_{\varphi, Ay}(By) : \|y\| = 1, \|Ay\| = \|A\|\} : \varphi \in [0, 2\pi)\} \geq 0.$$

Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be Hilbert spaces. In $K(\mathcal{H}_1, \mathcal{H}_2)$, the above theorem reduces to saying that for $A, B \in K(\mathcal{H}_1, \mathcal{H}_2)$, A is orthogonal to B if and only if there is a unit vector $h \in \mathcal{H}_1$ such that $\|Ah\| = \|A\|$ and $\langle Ah | Bh \rangle = 0$. But this is not always the case with reflexive Banach spaces.

An alternate proof of Theorem 3.1 was given in [15] by first giving a characterization of $\|A\| \leq \|A + tB\|$ for all $t \in \mathbb{R}$ using Corollary 2.9, and then extend the result to complex scalars to obtain Theorem 3.1. In [96], $\|A\| \leq \|A + tB\|$ for all $t \in \mathbb{R}$ is termed as *A is r-orthogonal to B*, and the same characterization as in [15] is given for r-orthogonality using a different approach. Using the same idea as in [15, Theorem 2.7], a proof of Theorem 3.1 was given in [96, Corollary 2.2].

The technique of using the subdifferential set as done in [15] has advantages that it gives a way to generalize Theorem 3.1 to characterize orthogonality to a subspace of $M_n(\mathbb{C})$.

Theorem 3.4 ([32, Theorem 1]) *Let $A \in M_n(\mathbb{C})$. Let $m(A)$ denotes the multiplicity of maximum singular value $\|A\|$ of A. Let $\mathcal{B}$ be any (real or complex) subspace of $M_n(\mathbb{C})$. Then A is orthogonal to $\mathcal{B}$ if and only if there exists a density matrix P of complex rank at most $m(A)$ such that $A^*AP = \|A\|^2 P$ and $tr(APB^*) = 0$ for all $B \in \mathcal{B}$.*

Theorem 3.4 can be expressed in terms of *states* on $M_n(\mathbb{C})$. Let $\mathcal{A}$ be a unital C^*-algebra over $\mathbb{F}(= \mathbb{R}$ or $\mathbb{C})$ with the identity element $1_{\mathcal{A}}$. For $\mathbb{F} = \mathbb{C}$, a state on $\mathcal{A}$ is a linear functional ϕ on $\mathcal{A}$ which takes $1_{\mathcal{A}}$ to 1 and positive elements of $\mathcal{A}$ to non-negative real numbers. For $\mathbb{F} = \mathbb{R}$, an additional requirement for ϕ to be a state is that $\phi(a^*) = \phi(a)$ for all $a \in \mathcal{A}$. Let $S_{\mathcal{A}}$ denotes the set of states on $\mathcal{A}$. Recently, the authors noticed in [35] that if P is a density matrix such that $\text{tr}(A^*AP) = \|A\|^2$, then P is a matrix of complex rank atmost $m(A)$ such that $A^*AP = \|A\|^2 P$ (the proof of this fact is along the lines of proof of Theorem 1.1 in [14]). Due to this fact, the above theorem can be restated in terms of states on $M_n(\mathbb{C})$ as follows: A is orthogonal to $\mathcal{B}$ if and only if there exists $\phi \in S_{M_n(\mathbb{C})}$ such that $\phi(A^*A) = \|A\|^2$ and $\phi(AB^*) = 0$ for all $B \in \mathcal{B}$. In a general complex C^*-algebra $\mathcal{A}$, it was shown in [7, Theorem 2.7] that an element $a \in \mathcal{A}$ is orthogonal to another element $b \in \mathcal{A}$ if and only if there exists $\phi \in S_{\mathcal{A}}$ such that $\phi(a^*a) = \|a\|^2$ and $\phi(a^*b) = 0$. A different proof of this result was given in [15, Proposition 4.1]. Theorem 6.1 in [76] shows that if $\mathcal{B}$ is a unital C^*-subalgebra of a complex C^*-algebra $\mathcal{A}$ and if a Hermitian element a of $\mathcal{A}$ is orthogonal to $\mathcal{B}$, then there exists $\phi \in S_{\mathcal{A}}$ such that $\phi(a^2) = \|a\|^2$ and $\phi(ab + b^*a) = 0$ for all $b \in \mathcal{B}$. Recently, the authors have extended these results to any (real or complex) C^*-algebra $\mathcal{A}$ for any element $a \in \mathcal{A}$ and any subspace $\mathcal{B}$ of $\mathcal{A}$ (see [35]). These are given in the next theorem.

If $\mathcal{A}$ is a complex (or real) unital C^*-algebra, then the triple $(\mathcal{H}, \pi, \xi)$ denotes a cyclic representation of $\mathcal{A}$, where $\mathcal{H}$ is a complex (or real) Hilbert space and $\pi : \mathcal{A} \to \mathscr{B}(\mathcal{H})$ is a *-algebra map such that $\pi(1_{\mathcal{A}}) = 1_{\mathscr{B}(\mathcal{H})}$ and $\{\pi(a)\xi : a \in \mathcal{A}\}$ is dense in $\mathscr{B}(\mathcal{H})$. For $\phi \in S_{\mathcal{A}}$, there exists a cyclic representation $(\mathcal{H}, \pi, \xi)$ such that $\phi(a) = \langle \pi(a)\xi | \xi \rangle$ for all $a \in \mathcal{A}$ (see [21, p. 250], [31, Proposition 15.2]).

Theorem 3.5 ([35, Corollary 1.3]) *Let $a \in \mathcal{A}$. Let $\mathcal{B}$ be a subspace of $\mathcal{A}$. Then the following are equivalent.*

1. *a is orthogonal to $\mathcal{B}$.*
2. *There exists $\phi \in S_{\mathcal{A}}$ such that $\phi(a^*a) = \|a\|^2$ and $\phi(a^*b) = 0$ for all $b \in \mathcal{B}$.*
3. *There exists a cyclic representation $(\mathcal{H}, \pi, \xi)$ such that $\|\pi(a)\xi\| = \|a\|$ and $\langle \pi(a)\xi | \pi(b)\xi \rangle = 0$ for all $b \in \mathcal{B}$.*

When $\mathcal{A} = C(X)$, Theorem 3.5 and Riesz Representation Theorem yield the following theorem by Singer [89, Theorem 1.3, Ch. I].

Corollary 3.6 ([89, Theorem 1.3, Ch. I]) *Let $f \in C(X)$. Let $\mathcal{B}$ be a subspace of $C(X)$. Then f is orthogonal to $\mathcal{B}$ if and only if there exists a probability measure μ on X such that*

$$\operatorname{dist}(a, \mathcal{B})^2 = \int_X |f|^2 d\mu \quad \text{and} \quad \int_X \overline{f} h d\mu = 0$$

for all $h \in \mathcal{B}$.

The condition

$$\|f\|_\infty^2 = \operatorname{dist}(a, \mathcal{B})^2 = \int_X |f|^2 d\mu$$

is equivalent to saying that the support of μ is contained in the set $\{x \in X : |f(x)| = \|f\|_\infty\}$. When $\mathcal{B}$ is one dimensional, this was proved in [53, Corollary 2.1] using Theorem 2.10.

Characterizations of orthogonality have been studied in several normed spaces. Using Theorem 2.10, a characterization of orthogonality in $C_b(\Omega)$ was obtained in [53, Corollary 3.1]. In the Banach spaces $L^1(X, \nu)$ and c_0, Theorem 2.10 was used to obtain such characterizations in [51, Example 1.6, Example 1.7]. For a separable Hilbert space $\mathcal{H}$, expressions for φ-Gateaux derivative of the norms on $\mathscr{B}_1(\mathcal{H})$ and $K(\mathcal{H})$ were given in [51, Theorem 2.1, Theorem 2.6] and were used to give characterizations of orthogonality in these spaces in [51, Corollary 2.5, Corollary 2.8]. Using tools of subdifferential calculus, characterizations of orthogonality in $(M_n(\mathbb{C}), \|\cdot\|_{(k)})$ are given in [34, Theorem 1.1, Theorem 1.2]. A necessary condition for orthogonality of a matrix A to a subspace in $(M_n(\mathbb{C}), \|\cdot\|_{(k)})$ is given in [34, Theorem 1.3]. Under the condition that $s_k(A) > 0$, the same condition is shown to be sufficient also. Using [89, Theorem 1.1, Ch. II], a characterization of orthogonality in $M_n(\mathbb{C})$, with any norm, is given in [58, Proposition 2.1] in terms of the dual norm. Using this, orthogonality in $M_n(\mathbb{C})$, with induced norms, is

obtained in [58, Proposition 4.2]. In $M_n(\mathbb{C})$, with *Schatten p-norms* ($1 \leq p \leq \infty$), characterizations of orthogonality are given in [58, Theorem 3.2, Theorem 3.3]. For $1 < p < \infty$, this was also given in [14, Theorem 2.1].

Orthogonality has been characterized in more general normed spaces, namely, *Hilbert C^*-modules*. It was shown in [7, Theorem 2.7] that in a Hilbert C^*-module $\mathscr{E}$ over a complex unital C^*-algebra $\mathcal{A}$, an element $e_1 \in \mathscr{E}$ is orthogonal to another element $e_2 \in \mathscr{E}$ if and only if there exists $\phi \in S_{\mathcal{A}}$ such that $\phi(\langle e_1|e_1\rangle) = \|e_1\|^2$ and $\phi(\langle e_1|e_2\rangle) = 0$. Another proof of this result was given in [15, Theorem 4.4]. This can be generalized to obtain a characterization of orthogonality to subspaces of Hilbert C^*-modules as follows.

Theorem 3.7 ([35, Theorem 3.5]) *Let $\mathscr{E}$ be a Hilbert C^*-module over a unital complex C^*-algebra $\mathcal{A}$. Let $e \in \mathscr{E}$. Let $\mathscr{F}$ be a subspace of $\mathscr{E}$. Then e is orthogonal to $\mathscr{F}$ if and only if there exists $\phi \in S_{\mathcal{A}}$ such that $\phi(\langle e|e\rangle) = \|e\|^2$ and $\phi(\langle e|f\rangle) = 0$ for all $f \in \mathscr{F}$.*

A proof of Theorem 3.7 can be found in [35]. Alternatively, this can also be proved along the same lines of the proof of [7, Theorem 2.4] by finding a generalization of the distance formula [7, Proposition 2.3] to a subspace. This extension of the distance formula is mentioned in the next section.

We end this section with various directions of research happening around the concept of orthogonality, where it comes into play naturally. In Hilbert C^*-modules, the role of scalars is played by the elements of the underlying C^*-algebra. Using this fact, a strong version of orthogonality was introduced in Hilbert C^*-modules in [8]. For a left Hilbert C^*-module, an element $e_1 \in \mathscr{E}$ is said to be *strong orthogonal* to another element $e_2 \in \mathscr{E}$ if $\|e_1\| \leq \|e_1 + ae_2\|$ for all $a \in \mathcal{A}$. Clearly, if e_1 is strong orthogonal to e_2, then we have e_1 is orthogonal to e_2. However, strong orthogonality is weaker than inner product orthogonality in $\mathscr{E}$ (see [8, Example 2.4]). Necessary and sufficient conditions are studied in [7, Theorem 3.1] and [9, Theorem 3.5, Corollary 4.9], when any two of these three orthogonalities coincide in a full Hilbert C^*-module. In [10, Theorem 2.6], it was shown that in a full Hilbert C^*-module, strong orthogonality is symmetric if and only if Birkhoff–James orthogonality is symmetric if and only if strong orthogonality coincides with inner product orthogonality. Theorem 2.5 of [8] gives characterization of strong orthogonality in terms of Birkhoff–James orthogonality.

Characterizations of orthogonality are also useful in finding conditions for equality in triangle inequality in a normed space:

Proposition 3.8 ([7, Proposition 4.1]) *Let V be a normed space. Let $u, v \in V$. Then the following are equivalent.*

1. $\|u + v\| = \|u\| + \|v\|$.
2. *v is orthogonal to $\|u\|v - \|v\|u$.*
3. *u is orthogonal to $\|u\|v - \|v\|u$.*

This can be extended to the case of arbitrarily finite families of vectors (see [7, Remark 4.2]). As mentioned in [7, Remark 4.2], a characterization of triangle

equality in a pre-Hilbert C^*-module given in [6, Theorem 2.1] can be proved using Theorem 3.7 and Proposition 3.8. For the study of various other equivalent conditions for equality in triangle inequality or Pythagoras equality, the interested reader is referred to [6, 7, 69].

In a normed linear space V, an element u is said to be *norm-parallel* to another element v (denoted by $u \parallel v$) if $\|u + \lambda v\| = \|u\| + \|v\|$ for some $\lambda \in \mathbb{F}$ with $|\lambda| = 1$ [87]. In the case of inner product spaces, the norm-parallel relation is exactly the usual vectorial parallel relation, that is, $u \parallel v$ if and only if u and v are linearly dependent. Seddik [87] introduced this concept while studying elementary operators on a standard operator algebra. Interested readers for the work on elementary operators and orthogonality are referred to [4, 25, 26, 87, 88, 103], and also to [7, Theorem 4.7] and [51, Section 3].

As a direct consequence of Proposition 3.8, norm-parallelism can be characterized using the concept of orthogonality (this characterization is also given in [68, Theorem 2.4]). So the results on orthogonality can be used to find results on norm-parallelism, for example, see [68, Proposition 2.19] for $M_n(\mathbb{C})$ equipped with the Schatten p-norms and [34, Remark 2] for $(M_n(\mathbb{C}), \|\cdot\|_{(k)})$. The characterizations of norm-parallelism in $C(X)$ are given in [105], and in $\mathscr{B}_1(\mathcal{H})$ and $K(\mathcal{H})$ are given in [104]. Other results in $\mathscr{B}(V_1, V_2)$ (with restrictions on V_1 and V_2, and the operators) are given in [100, 107]. Some of these results can be obtained by using [68, Theorem 2.4] and the corresponding results on orthogonality. Some variants of the definition of norm-parallelism have been introduced in [61, 104, 106]. Concepts of approximate Birkhoff–James orthogonality and ε-Birkhoff orthogonality have been studied in [19, 20, 23, 24, 43, 60, 83]. The idea to define these concepts of approximate Birkhoff–James orthogonality and ε-Birkhoff orthogonality in a normed space is to generalize the concept of approximate orthogonality in inner product spaces, which is defined as $v \perp^{\varepsilon} u \iff \big|\langle v|u\rangle\big| \leq \varepsilon \|v\| \|u\|$. The latter has been studied in [18] and [101, Section 5.2].

4 Distance Formulas and Conditional Expectations

An important connection of orthogonality with distance formulas was noted in (3.1) and (3.2). From (3.1), we also get that for any $A \in M_n(\mathbb{C})$,

$$\operatorname{dist}(A, \mathbb{C}1_{M_n(\mathbb{C})}) = \max\left\{\big|\langle Ax|y\rangle\big| : \|x\| = \|y\| = 1 \text{ and } y \perp x\right\}.$$

Using this, one obtains

$$\begin{aligned}\operatorname{dist}(A, \mathbb{C}1_{M_n(\mathbb{C})}) &= 2\max\{\|U'AU'^* - UAU^*\| : U, U' \text{ unitary}\}\\ &= 2\max\{\|AU - UA\| : U \text{ unitary}\}\\ &= 2\max\{\|AT - TA\| : T \in M_n(\mathbb{C}), \|T\| = 1\}. \qquad (4.1)\end{aligned}$$

This was proved in [14, Theorem 1.2] and the discussion after that. The operator $\delta_A(T) = AT - TA$ on $M_n(\mathbb{C})$ is called an *inner derivation*. So this gives that $\|\delta_A\| = 2\,\mathrm{dist}(A, \mathbb{C}1_{M_n(\mathbb{C})})$. This was also extended to the infinite dimensional case in [14, Remark 3.2]. These results were first proved by Stampfli [91] using a completely different approach. Since all the derivations on $\mathscr{B}(\mathcal{H})$ are inner derivations (see [49, Theorem 9]), the norm of any derivation on $\mathscr{B}(\mathcal{H})$ is $2\,\mathrm{dist}(A, \mathbb{C}1_{\mathscr{B}(\mathcal{H})})$, for some $A \in \mathscr{B}(\mathcal{H})$. For $A, B \in \mathscr{B}(\mathcal{H})$, let $\delta_{A,B}(T) = AT - TB$ for all $T \in \mathscr{B}(\mathcal{H})$. In [91, Theorem 8], an expression for the norm of the elementary operator $\delta_{A,B}$ is given. In [91, Theorem 5], a distance formula was obtained in an irreducible unital C^*-algebra as follows.

Theorem 4.1 ([91, Theorem 5]) *Let $\mathcal{H}$ be a complex Hilbert space. Let $\mathcal{B}$ be an irreducible unital C^*-subalgebra of $\mathscr{B}(\mathcal{H})$. Let $A \in \mathcal{B}$. Then*

$$2\,\mathrm{dist}(A, \mathbb{C}1_{\mathscr{B}(\mathcal{H})}) = \sup\{\|AT - TA\| : T \in \mathcal{B} \text{ and } \|T\| = 1\} = \|\delta_A|_{\mathcal{B}}\|.$$

By the Russo–Dye theorem [17, II.3.2.15], under the assumptions of the above theorem, we obtain

$$2\,\mathrm{dist}(A, \mathbb{C}1_{\mathscr{B}(\mathcal{H})}) = \sup\{\|AU - UA\| : U \in \mathcal{B} \text{ and } U \text{ is unitary}\}. \tag{4.2}$$

Expressions for the norm of a derivation on von Neumann algebras can be found in [29]. The most important fact used here is that all the derivations on von Neumann algebras are inner. This was a conjecture by Kadison for a long time and was proved in [85]. More on derivations on a C^*-algebra can be found in [47, 48, 67, 110]. A lot of work has been done to answer the question when the range of a derivation is orthogonal to its kernel. It was proved in [4, Theorem 1.7] that if N is a normal operator in $\mathscr{B}(\mathcal{H})$, then the kernel of δ_N is orthogonal to the range of δ_N. In [54, Theorem 1], it was shown that the Hilbert–Schmidt operators in the kernel of δ_N are orthogonal to the Hilbert–Schmidt operators in the range of δ_N, in the usual Hilbert space sense. In [59, Theorem 3.2(a)], the Schatten p-class operators in the kernel of δ_N were shown to be orthogonal to the Schatten p-class operators in the range of δ_N, in the Schatten p-norm. A similar result for the orthogonality in unitarily invariant norms defined on the norm ideals of $K(\mathcal{H})$ is given in [55, Theorem 1]. For related study on derivations, elementary operators and orthogonality in these normed spaces, see [27, 50, 56, 62–66, 92–94].

Similar to (4.2), an expression for the distance of an element of a general C^*-algebra from a C^*-subalgebra can be obtained from the below theorem of Rieffel [76].

Theorem 4.2 ([76, Theorem 3.2]) *Let $\mathcal{A}$ be a C^*-algebra. Let $\mathcal{B}$ be a C^*-subalgebra of $\mathcal{A}$ which contains a bounded approximate identity for $\mathcal{A}$. Let $a \in \mathcal{A}$. Then there exists a cyclic representation $(\mathcal{H}, \pi, \xi)$ of $\mathcal{A}$ and a Hermitian as well as a unitary operator U on $\mathcal{H}$ such that $\pi(b)U = U\pi(b)$ for all $b \in \mathcal{B}$ and* $\mathrm{dist}(a, \mathcal{B}) = \frac{1}{2}\|\pi(a)U - U\pi(a)\|$.

By Theorem 4.2, we obtain

$$\begin{aligned} 2\,\mathrm{dist}(a,\mathcal{B}) = \max\{\, &\|\pi(a)U - U\pi(a)\| : U \in \mathscr{B}(\mathcal{H}), U = U^*, U^2 = 1_{\mathscr{B}(\mathcal{H})}, \\ &(\mathcal{H},\pi,\xi) \text{ is a cyclic representation of } \mathcal{A}, \text{ and} \\ &\pi(b)U = U\pi(b) \text{ for all } b \in \mathcal{B}\}. \end{aligned}$$

Looking at the last expression and (4.2), it is tempting to conjecture that if $\mathcal{A}$ is a unital irreducible C^*-algebra, $a \in \mathcal{A}$ and $\mathcal{B}$ is a unital C^*-subalgebra of $\mathcal{A}$, then

$$\begin{aligned} 2\,\mathrm{dist}(a,\mathcal{B}) = \sup\{\, &\|au - ua\| : u \in \mathcal{A},\ u \text{ is a unitary element,} \\ &\text{and } bu = ub \text{ for all } b \in \mathcal{B}\}. \end{aligned} \tag{4.3}$$

We note that it is not possible to prove (4.3) by proceeding along the lines of the proof of Theorem 4.1 given in [14], which uses (4.1). In particular, the following does not hold true in $M_n(\mathbb{C})$:

$$\mathrm{dist}(A,\mathcal{B}) = \max\left\{\left|\langle Ax|y\rangle\right| : \|x\| = \|y\| = 1 \text{ and } y \perp Bx \text{ for all } B \in \mathcal{B}\right\}.$$

For example, take $A = 1_{M_n(\mathbb{C})}$ and $\mathcal{B} = \{X \in M_n(\mathbb{C}) : \mathrm{tr}(X) = 0\}$. Then $1_{M_n(\mathbb{C})}$ is orthogonal to $\mathcal{B}$. Now if the above is true, then we would get unit vectors x, y such that $\left|\langle x|y\rangle\right| = 1$ and $\langle Bx|y\rangle = 0$ for all $B \in \mathcal{B}$. Let $P = xy^*$. Then rank $P = 1$ and $\mathrm{tr}(BP) = 0$ for all $B \in \mathcal{B}$. But $\mathrm{tr}(BP) = 0$ for all $B \in \mathcal{B}$ gives $P = \lambda 1_{M_n(\mathbb{C})}$ for some $\lambda \in \mathbb{C}$ (see [32, Remark 3]), which contradicts the fact that rank $P = 1$. So this approach to prove (4.3) does not work. However it would be interesting to know if (4.3) is true or not. This is an open question.

The above example contradicts Theorem 5.3 of [101], which says that for Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$, if $A \in K(\mathcal{H},\mathcal{K})$ and $\mathcal{B}$ is a finite dimensional subspace of $K(\mathcal{H},\mathcal{K})$, then

$$\mathrm{dist}(A,\mathcal{B}) = \sup\left\{\left|\langle Ax|y\rangle\right| : \|x\| = \|y\| = 1 \text{ and } y \perp Bx \text{ for all } B \in \mathcal{B}\right\}.$$

The proof of this theorem has a gap, after invoking Theorem 5.2, in [101].

As an application of Theorem 3.5, we obtain the following distance formula.

Theorem 4.3 *Let $a \in \mathcal{A}$. Let $\mathcal{B}$ be a subspace of $\mathcal{A}$. Suppose there is a best approximation to a in $\mathcal{B}$. Then*

$$\begin{aligned} \mathrm{dist}(a,\mathcal{B}) = \max\Big\{\Big|\langle \pi(a)\xi|\eta\rangle\Big| : &(\mathcal{H},\pi,\xi) \textit{ is a cyclic representation of } \mathcal{A}, \\ &\eta \in \mathcal{H},\ \|\eta\| = 1 \textit{ and } \langle \pi(b)\xi|\eta\rangle = 0 \textit{ for all } b \in \mathcal{B}\Big\}. \end{aligned} \tag{4.4}$$

Proof Clearly $RHS \leq LHS$. To prove equality, we need to find a cyclic representation $(\mathcal{H},\pi,\xi)$ of $\mathcal{A}$ and a unit vector $\eta \in \mathcal{H}$ such that $\mathrm{dist}(a,\mathcal{B}) = \left|\langle \pi(a)\xi|\eta\rangle\right|$ and $\langle \pi(b)\xi|\eta\rangle = 0$ for all $b \in \mathcal{B}$. Let b_0 be a best approximation to

a in $\mathcal{B}$. By Theorem 3.5, there exists $\phi \in S_{\mathcal{A}}$ such that

$$\phi((a-b_0)^*(a-b_0)) = \|a-b_0\|^2$$

and $\phi((a-b_0)^*b) = 0$ for all $b \in \mathcal{B}$. Now there exists a cyclic representation $(\mathcal{H}, \pi, \xi)$ such that $\phi(c) = \langle \pi(c)\xi|\xi\rangle$ for all $c \in \mathcal{A}$. So $\|\pi(a-b_0)\xi\| = \|a-b_0\|$ and $\langle \pi(a-b_0)\xi|\pi(b)\xi\rangle = 0$ for all $b \in \mathcal{B}$. Taking $\eta = \frac{1}{\|a-b_0\|}\pi(a-b_0)\xi$, we get the required result. □

The authors have recently observed in [35] that the above theorem also holds true without the existence of a best approximation to a in $\mathcal{B}$. Notice that the right hand side of (4.4) uses only algebraic structure of $\mathcal{A}$ (as cyclic representations are defined by the algebraic structure of $\mathcal{A}$). More such distance formulas using only the algebraic structure of $\mathcal{A}$ are also known. When $\mathcal{B} = \mathbb{C}1_{\mathcal{A}}$, Williams [103, Theorem 2] proved that for $a \in \mathcal{A}$,

$$\operatorname{dist}(a, \mathbb{C}1_{\mathcal{A}})^2 = \max\{\phi(a^*a) - |\phi(a)|^2 : \phi \in S_{\mathcal{A}}\}. \tag{4.5}$$

When $\mathcal{A} = M_n(\mathbb{C})$, another proof of (4.5) was given by Audenaert [12, Theorem 9]. Rieffel [77, Theorem 3.10] obtained (4.5), using a different method. In [77], it was also desired to have a generalization of (4.5) with $\mathbb{C}1_{\mathcal{A}}$ replaced by a unital C^*-subalgebra. For $\mathcal{A} = M_n(\mathbb{C})$, a formula in this direction was obtained in [32, Theorem 2]. An immediate application of Theorem 3.5 gives the following generalization of (4.5), when $\mathbb{C}1_{\mathcal{A}}$ is replaced by a subspace $\mathcal{B}$ of $\mathcal{A}$ and there is a best approximation to a in $\mathcal{B}$.

Theorem 4.4 ([35, Corollary 1.2]) *Let $a \in \mathcal{A}$. Let $\mathcal{B}$ be a subspace of $\mathcal{A}$. Let b_0 be a best approximation to a in $\mathcal{B}$. Then*

$$\operatorname{dist}(a, \mathcal{B})^2 = \max\{\phi(a^*a) - \phi(b_0^*b_0) : \phi \in S_{\mathcal{A}}, \phi(a^*b) = \phi(b_0^*b) \text{ for all } b \in \mathcal{B}\}.$$

For details, see [35]. Geometric interpretations of Theorems 3.5 and 4.4 have also been explained in [35].

Henceforward, C^*-algebras are assumed to be complex C^*-algebras. Another distance formula, which is a generalization of [7, Proposition 2.3], is given below. Some notations are in order. Given $\phi \in S_{\mathcal{A}}$, let $\mathcal{L} = \{c \in \mathcal{A} : \phi(c^*c) = 0\}$, and let $\langle a_1 + \mathcal{L}|a_2 + \mathcal{L}\rangle_{\mathcal{A}/\mathcal{L}} = \phi(a_1^*a_2)$, for all $a_1, a_2 \in \mathcal{A}$. Then $\mathcal{A}/\mathcal{L}$ is an inner product space. For $a \in \mathcal{A}$, let b_0 be a best approximation to a in $\mathcal{B}$. Let

$$M_{a,\mathcal{B}}(\phi) = \sup\{\phi((a-b_0)^*(a-b_0)) - \sum_{\alpha} |\phi((a-b_0)^*b_\alpha)|^2\},$$

where the supremum is taken over all orthonormal bases $\{b_\alpha + \mathcal{L}\}$ of $\mathcal{B}/\mathcal{L}$ in $\mathcal{A}/\mathcal{L}$.

Theorem 4.5 *Let $a \in \mathcal{A}$. Let $\mathcal{B}$ be a subspace of $\mathcal{A}$. Let b_0 be a best approximation to a in $\mathcal{B}$. Then*

$$\operatorname{dist}(a, \mathcal{B})^2 = \max\{M_{a,\mathcal{B}}(\phi) : \phi \in S_{\mathcal{A}}\}.$$

Proof Clearly $RHS \leq LHS$. For an orthonormal basis $\{b_\alpha + \mathcal{L}\}$ of $\mathcal{B}/\mathcal{L}$, we have

$$\phi((a - b_0)^*(a - b_0)) - \sum_\alpha |\phi((a - b_0)^* b_\alpha)|^2 \leq LHS.$$

And equality occurs because by Theorem 3.5, there exists $\phi \in S_{\mathcal{A}}$ such that $\phi((a - b_0)^*(a - b_0)) = \operatorname{dist}(a, \mathcal{B})^2$ and $\phi((a - b_0)^* b) = 0$ for all $b \in \mathcal{B}$. □

Now along the lines of the proof of [7, Theorem 2.4] and using Theorem 4.5, we get the next result. For a Hilbert C^*-module $\mathscr{E}$ over $\mathcal{A}$ and $\phi \in S_{\mathcal{A}}$, let $\mathcal{L} = \{e \in \mathscr{E} : \phi(\langle e|e\rangle) = 0\}$. On $\mathscr{E}/\mathcal{L}$, define an inner product as $\langle e_1 + \mathcal{L}|e_2 + \mathcal{L}\rangle_{\mathscr{E}/\mathcal{L}} = \phi(\langle e_1|e_2\rangle)$ for all $e_1, e_2 \in \mathscr{E}$. For $e \in \mathscr{E}$, let f_0 be a best approximation to e in $\mathscr{F}$. Let

$$M_{e,\mathscr{F}}(\phi) = \sup\{\phi(\langle e - f_0|e - f_0\rangle) - \sum_\alpha |\phi(\langle e - f_0|f_\alpha\rangle)|^2\},$$

where the supremum is taken over all orthonormal bases $\{f_\alpha + \mathcal{L}\}$ of $\mathscr{F}/\mathcal{L}$ in $\mathscr{E}/\mathcal{L}$.

Theorem 4.6 *Let $\mathscr{E}$ be a Hilbert C^*-module over $\mathcal{A}$. Let $e \in \mathscr{E}$. Let $\mathscr{F}$ be a subspace of $\mathscr{E}$. Let f_0 be a best approximation to e in $\mathscr{F}$. Then*

$$\operatorname{dist}(e, \mathscr{F}) = \max\{M_{e,\mathscr{F}}(\phi) : \phi \in S_{\mathcal{A}}\}.$$

Rieffel [77, p. 46] had questioned to have expressions of distance formulas in terms of *conditional expectations*. We end the discussion on distance formulas with our progress in this direction. For a C^*-algebra $\mathcal{A}$ and a C^*-subalgebra $\mathcal{B}$ of $\mathcal{A}$, a conditional expectation from $\mathcal{A}$ to $\mathcal{B}$ is a completely positive map $E : \mathcal{A} \to \mathcal{B}$ of unit norm such that $E(b) = b$, $E(ba) = bE(a)$ and $E(ab) = E(a)b$, for all $a \in \mathcal{A}$ and $b \in \mathcal{B}$ [17, p. 141]. In fact any projection $E : \mathcal{A} \to \mathcal{B}$ of norm one is a conditional expectation and vice-a-versa (see [17, Theorem II.6.10.2]). An interesting fact is that a map $E : \mathcal{A} \to \mathcal{B}$ is a conditional expectation if and only if E is idempotent, positive and satisfies $E(b_1 a b_2) = b_1 E(a) b_2$, for all $a \in \mathcal{A}$ and $b_1, b_2 \in \mathcal{B}$ (see [17, Theorem II.6.10.3]). Thus conditional expectations from $\mathcal{A}$ to $\mathcal{B}$ are also determined completely by the algebraic structure. A Banach space V_1 is said to be *injective* if for any inclusion of Banach spaces $V_3 \subseteq V_2$, every bounded linear mapping $f_0 : V_3 \to V_1$ has a linear extension $f : V_2 \to V_1$ with $\|f\| = \|f_0\|$. A Banach space is injective if and only if it is isometrically isomorphic to $C(X)$, where X is a compact Hausdorff space in which closure of any open set is an open set (see [28, p. 70]). For $v \in V$ and W a subspace of V, let $\langle v, W\rangle$ denote the subspace generated by v and W.

Theorem 4.7 *Let $a \in \mathcal{A}$. Let $\mathcal{B}$ be a subspace of $\mathcal{A}$ such that $\mathcal{B}$ is an injective Banach space and $1_{\mathcal{A}} \in \mathcal{B}$. Suppose there is a best approximation to a in $\mathcal{B}$. Then there exists $\phi \in S_{\mathcal{A}}$ and a projection $E : \mathcal{A} \to \mathcal{B}$ of norm atmost two such that $\phi \circ E = \phi$ and* $\text{dist}(a, \mathcal{B})^2 = \phi(a^*a) - \phi(E(a)^*E(a))$.

Proof Let b_0 be a best approximation to a in $\mathcal{B}$. We define $\tilde{E} : \langle a, \mathcal{B}\rangle \to \mathcal{B}$ as $\tilde{E}(b) = b$ for all $b \in \mathcal{B}$ and $\tilde{E}(a) = b_0$ and extend it linearly on $\langle a, \mathcal{B}\rangle$. Using Theorem 3.5, there exists $\phi \in S_{\mathcal{A}}$ such that $\text{dist}(a, \mathcal{B})^2 = \phi(a^*a) - \phi(b_0^*b_0)$ and $\phi(a^*b) = \phi(b_0^*b)$ for all $b \in \mathcal{B}$. Since $1_{\mathcal{A}} \in \mathcal{B}$, we get $\phi(a) = \phi(b_0) = \phi(E(a))$. And clearly $\phi(b) = \phi(\tilde{E}(b))$ for all $b \in \mathcal{B}$. Thus $\phi \circ \tilde{E} = \phi$. Since b_0 is a best approximation to a in $\mathcal{B}$, $\|a - b_0\| \leq \|a\|$. So $\|b_0\| \leq 2\|a\|$. Now let $b \in \mathcal{B}$ and $\alpha \in \mathbb{C}$. Then $\tilde{E}(\alpha a + b) = \alpha b_0 + b$ and $\alpha b_0 + b$ is a best approximation to $\alpha a + b$ in $\mathcal{B}$. Thus $\|\alpha b_0 + b\| \leq 2\|\alpha a + b\|$. Hence $\|\tilde{E}\| \leq 2$. Since $\mathcal{B}$ is injective, there exists a linear extension $E : \mathcal{A} \to \mathcal{B}$ with norm same as that of $\tilde{E}$. This E is the required projection. □

For any given conditional expectation E from $\mathcal{A}$ to $\mathcal{B}$, we can define a $\mathcal{B}$-valued inner product on $\mathcal{A}$ given by $\langle a_1|a_2\rangle_E = E(a_1^*a_2)$ (see [77]). In [35], we obtain a lower bound for $\text{dist}(a, \mathcal{B})$ as follows.

Theorem 4.8 *Let $a \in \mathcal{A}$. Let $\mathcal{B}$ be a C^*-subalgebra of $\mathcal{A}$ such that $1_{\mathcal{A}} \in \mathcal{B}$. Then*

$$\text{dist}(a, \mathcal{B})^2 \geq \sup\{\phi(\langle a - E(a)|a - E(a)\rangle_E) : \phi \in \mathcal{S}_{\mathcal{A}}, E \text{ is a conditional expectation from } \mathcal{A} \text{ onto } \mathcal{B}\}. \tag{4.6}$$

(Here we follow the convention that $\sup(\emptyset) = -\infty$*.)*

Remark 4.9 It is worth mentioning here that if in Theorem 4.7, we take $\mathcal{B}$ to be a C^*-algebra and we are able to find a projection of norm one, then we will get equality in (4.6), that is,

$$\text{dist}(a, \mathcal{B})^2 = \sup\{\phi(\langle a - E(a)|a - E(a)\rangle_E) : \phi \in \mathcal{S}_{\mathcal{A}}, E \text{ is a conditional expectation from } \mathcal{A} \text{ onto } \mathcal{B}\}.$$

This happens in the special case when $\mathcal{B} = \mathbb{C}1_{\mathcal{A}}$, because for any $c \in \mathcal{A}$, the norm of the best approximation of c to $\mathbb{C}1_{\mathcal{A}}$ is less than or equal to $\|c\|$, and thus the norm of the projection E in Theorem 4.7 is one.

Acknowledgments The authors would like to thank Amber Habib and Ved Prakash Gupta for many useful discussions. The authors would also like to acknowledge very helpful comments by the referees.

The research of P. Grover is supported by INSPIRE Faculty Award IFA14-MA-52 of DST, India, and by Early Career Research Award ECR/2018/001784 of SERB, India.

References

1. T.J. Abatzoglou, Norm derivatives on spaces of operators. Math. Ann. **239**, 129–135 (1979)
2. J. Alonso, Some properties of Birkhoff and isosceles orthogonality in normed linear spaces, in *Inner Product Spaces and Applications*. Pitman Research Notes in Mathematical Series, vol. 376 (Longman, Harlow, 1997), pp. 1–11
3. D. Amir, *Characterizations of Inner Product Spaces*. Operator Theory: Advances and Applications, vol. 20 (Birkhäuser Verlag, Basel, 1986)
4. J. Anderson, On normal derivations. Proc. Am. Math. Soc. **38**, 135–140 (1973)
5. C. Apostol, L.A. Fialkow, D.A. Herrero, D. Voiculescu, *Approximation of Hilbert Space Operators II*. Research Notes in Mathematics, vol. 102 (Pitman (Advanced Publishing Program), Boston, 1984)
6. L. Arambašić, R. Rajić, On some norm equalities in pre-Hilbert C^*-modules. Linear Algebra Appl. **414**, 19–28 (2006)
7. L. Arambašić, R. Rajić, The Birkhoff-James orthogonality in Hilbert C^*-modules. Linear Algebra Appl. **437**, 1913–1929 (2012)
8. L. Arambašić, R. Rajić, A strong version of the Birkhoff-James orthogonality in Hilbert C^*-modules. Ann. Funct. Anal. **5**, 109–120 (2014)
9. L. Arambašić, R. Rajić, On three concepts of orthogonality in Hilbert C^*-modules. Linear Multilinear Algebra **63**, 1485–1500 (2015)
10. L. Arambašić, R. Rajić, On symmetry of the (strong) Birkhoff-James orthogonality in Hilbert C^*-modules. Ann. Funct. Anal. **7**, 17–23 (2016)
11. J. Arazy, On the geometry of the unit ball of unitary matrix spaces. Integr. Equ. Oper. Theory **4**, 151–171 (1981)
12. K.M.R. Audenaert, Variance bounds, with an application to norm bounds for commutators. Linear Algebra Appl. **432**, 1126–1143 (2010)
13. C. Benítez, M. Fernández, M.L. Soriano, Orthogonality of matrices. Linear Algebra Appl. **422**, 155–163 (2007)
14. R. Bhatia, P. Šemrl, Orthogonality of matrices and some distance problems. Linear Algebra Appl. **287**, 77–85 (1999)
15. T. Bhattacharyya, P. Grover, Characterization of Birkhoff-James orthogonality. J. Math. Anal. Appl. **407**, 350–358 (2013)
16. G. Birkhoff, Orthogonality in linear metric spaces. Duke Math. J. **1**, 169–172 (1935)
17. B. Blackadar, *Operator Algebras-Theory of C^*-Algebras and von Neumann Algebras* (Springer, Berlin, 2006)
18. J. Chmieliński, Linear mappings approximately preserving orthogonality. J. Math. Anal. Appl. **304**, 158–169 (2005)
19. J. Chmieliński, On an ε-Birkhoff orthogonality. J. Inequal. Pure Appl. Math. **6**, 1–7 (2005)
20. J. Chmieliński, T. Stypula, P. Wójcik, Approximate orthogonality in normed spaces and its applications. Linear Algebra Appl. **531**, 305–317 (2017)
21. J.B. Conway, *A Course in Functional Analysis* (Springer, New York, 1990)
22. J. Diestel, *Geometry of Banach Spaces*. Lecture Notes in Mathematics, vol. 485 (Springer, Berlin, 1975)
23. S.S. Dragomir, On Approximation of Continuous Linear Functionals in Normed Linear Spaces. An. Univ. Timişoara Ser. Ştiint. Mat. **29**, 51–58 (1991)
24. S.S. Dragomir, Continuous linear functionals and norm derivatives in real normed spaces. Univ. Beograd. Publ. Elektrotehn. Fak. Ser. Mat. **3**, 5–12 (1992)
25. H.-K. Du, Another generalization of Anderson's theorem. Proc. Am. Math. Soc. **123**, 2709–2714 (1995)
26. B.P. Duggal, A remark on normal derivations. Proc. Am. Math. Soc. **126**, 2047–2052 (1998)
27. B.P. Duggal, Range-kernel orthogonality of the elementary operator $X \to \sum_{i=1}^{n} A_i X B_i - X$. Linear Algebra Appl. **337**, 79–86 (2001)

28. E.G. Effros, Z.-J. Ruan, *Operator Spaces* (Oxford University Press, New York, 2000)
29. P. Gajendragadkar, Norm of a derivation on a von Neumann algebra. Trans. Am. Math. Soc. **170**, 165–170 (1972)
30. P. Ghosh, D. Sain, K. Paul, On symmetry of Birkhoff-James orthogonality of linear operators. Adv. Oper. Theory **2**, 428–434 (2017)
31. K.R. Goodearl, *Notes on Real and Complex C^*-Algebras* (Shiva, Cambridge, 1982)
32. P. Grover, Orthogonality to matrix subspaces, and a distance formula. Linear Algebra Appl. **445**, 280–288 (2014)
33. P. Grover, Some problems in differential and subdifferential calculus of matrices. Ph.D. Thesis, Indian Statistical Institute (2014)
34. P. Grover, Orthogonality of matrices in the Ky Fan k-norms. Linear Multilinear Algebra **65**, 496–509 (2017)
35. P. Grover, S. Singla, Best approximations, distance formulas and orthogonality in C^*-algebras. J. Ramanujan Math. Soc. (to appear)
36. R. Grząślewicz, R. Younis, Smooth points of some operator spaces. Arch. Math. **57**, 402–405 (1991)
37. R. Grząślewicz, R. Younis, Smooth points and M-ideals. J. Math. Anal. Appl. **175**, 91–95 (1993)
38. P. Halmos, *A Hilbert Space Problem Book* (D. Van Nostrand, Princeton, 1967)
39. J. Hennefeld, *Smooth, compact operators*. Proc. Am. Math. Soc. **77**, 87–90 (1979)
40. S. Heinrich, The differentiability of the norm in spaces of operators. Funct. Anal. Appl. **9**, 360–362 (1975)
41. J.B. Hiriart-Urruty, C. Lemarèchal, *Fundamentals of Convex Analysis* (Springer, Berlin, 2000)
42. J.R. Holub, On the metric geometry of ideals of operators on Hilbert space. Math. Ann. **201**, 157–163 (1973)
43. T. Jahn, Orthogonality in generalized Minkowski spaces. J. Convex Anal. **26**, 49–76 (2019)
44. R.C. James, Orthogonality in normed linear spaces. Duke Math. J. **12**, 291–302 (1945)
45. R.C. James, Orthogonality and linear functionals in normed linear spaces. Trans. Am. Math. Soc. **61**, 265–292 (1947)
46. R.C. James, Inner products in normed linear spaces. Bull. Am. Math. Soc. **53**, 559–566 (1947)
47. B. Johnson, Characterization and norms of derivations on von Neumann algebras, in *Algèbres d'Opérateurs*. Lecture Notes in Mathematics, vol. 725 (Springer, Berlin, 1979), pp. 228–236
48. R.V. Kadison, Derivations of operator algebras. Ann. Math. **83**, 280–293 (1966)
49. I. Kaplansky, Modules over operator algebras. Am. J. Math. **75**, 839–858 (1953)
50. D.J. Kečkić, Orthogonality of the range and the kernel of some elementary operators. Proc. Am. Math. Soc. **128**, 3369–3377 (2000)
51. D.J. Kečkić, Orthogonality in $\mathfrak{S}_1$ and $\mathfrak{S}_\infty$ spaces and normal derivations. J. Oper. Theory **51**, 89–104 (2004)
52. D.J. Kečkić, Gateaux derivative of $B(H)$ norm. Proc. Am. Math. Soc. **133**, 2061–2067 (2005)
53. D.J. Kečkić, Orthogonality and smooth points in $C(K)$ and $C_b(\Omega)$. Eur. Math. J. **3**, 44–52 (2012)
54. F. Kittaneh, On normal derivations of Hilbert-Schmidt type. Glasg. Math. J. **29**, 245–248 (1987)
55. F. Kittaneh, Normal derivations in norm ideals. Proc. Am. Math. Soc. **123**, 1779–1785 (1995)
56. F. Kittaneh, Operators that are orthogonal to the range of a derivation. J. Math. Anal. Appl. **203**, 868–873 (1996)
57. F. Kittaneh, R. Younis, Smooth points of certain operator spaces. Integr. Equ. Oper. Theory **13**, 849–855 (1990)
58. C.K. Li, H. Schneider, Orthogonality of matrices. Linear Algebra Appl. **347**, 115–122 (2002)
59. P.J. Maher, Commutator approximants. Proc. Am. Math. Soc. **115**, 995–1000 (1992)
60. A. Mal, K. Paul, T.S.S.R.K. Rao, D. Sain, Approximate Birkhoff-James orthogonality and smoothness in the space of bounded linear operators. Monatsh. Math. **190**, 549–558 (2019)
61. A. Mal, D. Sain, K. Paul, On some geometric properties of operator spaces. Banach J. Math. Anal. **13**, 174–191 (2019)

62. A. Mazouz, On the range and the kernel of the operator $X \mapsto AXB - X$. Proc. Am. Math. Soc. **127**, 2105–2107 (1999)
63. S. Mecheri, On minimizing $\|S - (AX - XB)\|_p^p$. Serdica Math. J. **26**, 119–126 (2000)
64. S. Mecheri, Some versions of Anderson's and Maher's inequalities I. Int. J. Math. Math. Sci. **52**, 3281–3297 (2003)
65. S. Mecheri, Some versions of Anderson's and Maher's inequalities II. Int. J. Math. Math. Sci. **53**, 3355–3372 (2003)
66. S. Mecheri, M. Bounkhel, Some variants of Anderson's inequality in C_1-classes. JIPAM. J. Inequal. Pure Appl. Math. **4**, Article 24 (2003)
67. P. Miles, Derivations on B^* algebras. Pac. J. Math. **14**, 1359–1366 (1964)
68. M.S. Moslehian, A. Zamani, Norm-parallelism in the geometry of Hilbert C^*-modules. Indag. Math. **27**, 266–281 (2016)
69. R. Nakamoto, S. Takahasi, Norm equality condition in triangular inequality. Sci. Math. Jpn. **55**, 463–466 (2002)
70. K. Paul, Translatable radii of an operator in the direction of another operator. Sci. Math. **2**, 119–122 (1999)
71. K. Paul, A. Mal, P. Wójcik, Symmetry of Birkhoff-James orthogonality of operators defined between infinite dimensional Banach spaces. Linear Algebra Appl. **563**, 142–153 (2019)
72. K. Paul, D. Sain, P. Ghosh, Birkhoff-James orthogonality and smoothness of bounded linear operators. Linear Algebra Appl. **506**, 551–563 (2016)
73. T.S.S.R.K. Rao, Very smooth points in spaces of operators. Proc. Indian Acad. Sci. Math. Sci. **113**, 53–64 (2003)
74. T.S.S.R.K. Rao, Smooth points in spaces of operators. Linear Algebra Appl. **517**, 129–133 (2017)
75. T.S.S.R.K. Rao, Adjoints of operators as smooth points in spaces of compact operators. Linear Multilinear Algebra **66**, 668–670 (2018)
76. M.A. Rieffel, Leibniz seminorms and best approximation from C^*-subalgebras. Sci. China Math. **54**, 2259–2274 (2011)
77. M.A. Rieffel, Standard deviation is a strongly Leibniz seminorm. New York J. Math. **20**, 35–56 (2014)
78. B.D. Roberts, On geometry of abstract vector spaces. Tohoku Math. J. **39**, 42–59 (1934)
79. D. Sain, K. Paul, Operator norm attainment and inner product spaces. Linear Algebra Appl. **439**, 2448–2452 (2013)
80. D. Sain, K. Paul, S. Hait, Operator norm attainment and Birkhoff-James orthogonality. Linear Algebra Appl. **476**, 85–97 (2015)
81. D. Sain, P. Ghosh, K. Paul, On symmetry of Birkhoff-James orthogonality of linear operators on finite-dimensional real Banach spaces. Oper. Matrices **11**, 1087–1095 (2017)
82. D. Sain, K. Paul, A. Mal, A complete characterization of Birkhoff-James orthogonality in infinite dimensional normed space. J. Oper. Theory **80**, 399–413 (2018)
83. D. Sain, K. Paul, A. Mal, On approximate Birkhoff-James orthogonality and normal cones in a normed space. J. Convex Anal. **26**, 341–351 (2019)
84. D. Sain, K. Paul, A. Mal, A. Ray, A complete characterization of smoothness in the space of bounded linear operators. Linear Multilinear Algebra (2019). https://doi.org/10.1080/03081087.2019.1586824
85. S. Sakai, Derivations of W^*-algebras. Ann. Math. **83**, 273–279 (1966)
86. R. Schatten, The space of completely continuous operators on a Hilbert space. Math. Ann. **134**, 47–49 (1957)
87. A. Seddik, Rank one operators and norm of elementary operators. Linear Algebra Appl. **424**, 177–183 (2007)
88. A. Seddik, On the injective norm of $\sum_{i=1}^{n} A_i \otimes B_i$ and characterization of normaloid operators. Oper. Matrices **2**, 67–77 (2008)
89. I. Singer, *Best Approximation in Normed Linear Spaces by Elements of Linear Subspaces* (Springer, New York, 1970)

90. W. So, Facial structures of Schatten p-norms. Linear Multilinear Algebra **27**, 207–212 (1990)
91. J. Stampfli, The norm of a derivation. Pac. J. Math. **33**, 737–747 (1970)
92. A. Turnšek, Elementary operators and orthogonality. Linear Algebra Appl. **317**, 207–216 (2000)
93. A. Turnšek, Orthogonality in $\mathscr{C}_p$ classes. Monatsh. Math. **132**, 349–354 (2001)
94. A. Turnšek, Generalized Anderson's inequality. J. Math. Anal. Appl. **263**, 121–134 (2001)
95. A. Turnšek, On operators preserving James' orthogonality. Linear Algebra Appl. **407**, 189–195 (2005)
96. A. Turnšek, A remark on orthogonality and symmetry of operators in $\mathscr{B}(\mathscr{H})$. Linear Algebra Appl. **535**, 141–150 (2017)
97. G.A. Watson, Characterization of the subdifferential of some matrix norms. Linear Algebra Appl. **170**, 33–45 (1992)
98. G.A. Watson, On matrix approximation problems with Ky Fan k norms. Numer. Algorithms **5**, 263–272 (1993)
99. P. Wójcik, Gateaux derivative of norm in $\mathcal{K}(X; Y)$. Ann. Funct. Anal. **7**, 678–685 (2016)
100. P. Wójcik, Norm-parallelism in classical M-ideals. Indag. Math. **28**, 287–293 (2017)
101. P. Wójcik, Orthogonality of compact operators. Expo. Math. **35**, 86–94 (2017)
102. W. Werner, Smooth points in some spaces of bounded operators. Integr. Equ. Oper. Theory **15**, 496–502 (1992)
103. J.P. Williams, Finite operators. Proc. Am. Math. Soc. **26**, 129–136 (1970)
104. A. Zamani, The operator-valued parallelism. Linear Algebra Appl. **505**, 282–295 (2016)
105. A. Zamani, Characterizations of norm-parallelism in spaces of continuous functions. Bull. Iranian Math. Soc. **45**, 557–567 (2019)
106. A. Zamani, M.S. Moslehian, Exact and approximate operator parallelism. Can. Math. Bull. **58**, 207–224 (2015)
107. A. Zamani, M.S. Moslehian, M.-T. Chien, H. Nakazato, Norm-parallelism and the Davis-Wielandt radius of Hilbert space operators. Linear Multilinear Algebra **67**, 2147–2158 (2019)
108. K. Ziętak, On the characterization of the extremal points of the unit sphere of matrices. Linear Algebra Appl. **106**, 57–75 (1988)
109. K. Ziętak, Subdifferentials, faces, and dual matrices. Linear Algebra Appl. **185**, 125–141 (1993)
110. L. Zsidó, The norm of a derivation in a W^*-algebra. Proc. Am. Math. Soc. **38**, 147–150 (1973)

The Generalized ∂-Complex on the Segal–Bargmann Space

Friedrich Haslinger

Abstract We study certain densely defined unbounded operators on the Segal-Bargmann space, related to the annihilation and creation operators of quantum mechanics. We consider the corresponding D-complex and study properties of the corresponding complex Laplacian $\tilde{\Box}_D = DD^* + D^*D$, where D is a differential operator of polynomial type.

Keywords ∂-complex · Segal–Bargmann space

Mathematics Subject Classification (2010) Primary 30H20, 32A36, 32W50; Secondary 47B38.

1 Introduction

We consider the classical Segal–Bargmann space

$$A^2(\mathbb{C}^n, e^{-|z|^2}) = \left\{ u : \mathbb{C}^n \longrightarrow \mathbb{C} \text{ entire} : \int_{\mathbb{C}^n} |u(z)|^2 e^{-|z|^2}\, d\lambda(z) < \infty \right\}$$

with inner product

$$(u, v) = \int_{\mathbb{C}^n} u(z)\, \overline{v(z)}\, e^{-|z|^2}\, d\lambda(z)$$

and replace a single derivative with respect to z_j by a differential operator of the form $p_j(\frac{\partial}{\partial z_1}, \ldots, \frac{\partial}{\partial z_n})$, where p_j is a complex polynomial on $\mathbb{C}^n$ (see [5, 6]). We

F. Haslinger (✉)
Fakultät für Mathematik, Universität Wien, Wien, Austria
e-mail: friedrich.haslinger@univie.ac.at

M. A. Bastos et al. (eds.), *Operator Theory, Functional Analysis and Applications*,
Operator Theory: Advances and Applications 282,
https://doi.org/10.1007/978-3-030-51945-2_16

write $p_j(u)$ for $p_j(\frac{\partial}{\partial z_1}, \dots, \frac{\partial}{\partial z_n})u$, where $u \in A^2(\mathbb{C}^n, e^{-|z|^2})$, and consider the densely defined operators

$$Du = \sum_{j=1}^{n} p_j(u)\, dz_j, \tag{1.1}$$

where $u \in A^2(\mathbb{C}^n, e^{-|z|^2})$ and $p_j(\frac{\partial}{\partial z_1}, \dots, \frac{\partial}{\partial z_n})$ are polynomial differential operators with constant coefficients [3].

More generally we define

$$Du = {\sum_{|J|=p}}' \sum_{k=1}^{n} p_k(u_J)\, dz_k \wedge dz_J,$$

where $u = {\sum_{|J|=p}}' u_J\, dz_J$ is a $(p,0)$-form with coefficients in $A^2(\mathbb{C}^n, e^{-|z|^2})$, here $J = (j_1, \dots, j_p)$ is a multiindex and $dz_J = dz_{j_1} \wedge \cdots \wedge dz_{j_p}$ and the summation is taken only over increasing multiindices.

It is clear that $D^2 = 0$ and that we have

$$(Du, v) = (u, D^* v), \tag{1.2}$$

where

$$u \in \mathrm{dom}(D) = \left\{ u \in A^2_{(p,0)}(\mathbb{C}^n, e^{-|z|^2}) : Du \in A^2_{(p+1,0)}(\mathbb{C}^n, e^{-|z|^2}) \right\}$$

and

$$D^* v = {\sum_{|K|=p-1}}' \sum_{j=1}^{n} p_j^* v_{jK}\, dz_K$$

for $v = {\sum_{|J|=p}}' v_J\, dz_J$ and where $p_j^*(z_1, \dots, z_n)$ is the polynomial p_j with complex conjugate coefficients, taken as multiplication operator.

Now the corresponding D-complex has the form

$$A^2_{(p-1,0)}(\mathbb{C}^n, e^{-|z|^2}) \underset{D^*}{\overset{D}{\rightleftarrows}} A^2_{(p,0)}(\mathbb{C}^n, e^{-|z|^2}) \underset{D^*}{\overset{D}{\rightleftarrows}} A^2_{(p+1,0)}(\mathbb{C}^n, e^{-|z|^2}).$$

Similarly to the classical $\overline{\partial}$-complex (see [3]) we consider the generalized box operator $\tilde{\Box}_{D,p} := D^*D + DD^*$ as a densely defined self-adjoint operator on $A^2_{(p,0)}(\mathbb{C}^n, e^{-|z|^2})$ with

$$\mathrm{dom}(\tilde{\Box}_{D,p}) = \{ f \in \mathrm{dom}(D) \cap \mathrm{dom}(D^*) : Df \in \mathrm{dom}(D^*),\ D^* f \in \mathrm{dom}(D) \},$$

see [4] for more details.

The $(p,0)$-forms with polynomial components are dense in the space $A^2_{(p,0)}(\mathbb{C}^n, e^{-|z|^2})$. In addition we have the following.

Lemma 1.1 *The $(p,0)$-forms with polynomial components are also dense in* $\mathrm{dom}(D) \cap \mathrm{dom}(D^*)$ *endowed with the graph norm*

$$u \mapsto (\|u\|^2 + \|Du\|^2 + \|D^*u\|^2)^{1/2}.$$

Proof Let $u = \sum_{|J|=p}{}' u_J\, dz_J \in \mathrm{dom}(D) \cap \mathrm{dom}(D^*)$ and consider the partial sums of the Fourier series expansions of

$$u_J = \sum_\alpha u_{J,\alpha}\varphi_\alpha,$$

where

$$\varphi_\alpha(z) = \frac{z^\alpha}{\sqrt{\pi^n \alpha!}} \text{ and } \sum_\alpha |u_{J,\alpha}|^2 < \infty$$

and $\alpha! = \alpha_1! \ldots \alpha_n!$. We have that $Du \in A^2_{(p+1,0)}(\mathbb{C}^n, e^{-|z|^2})$ and $D^*u \in A^2_{(p-1,0)}(\mathbb{C}^n, e^{-|z|^2})$. Hence the partial sums of the Fourier series of the components of Du converge to the components of Du in $A^2_{(p+1,0)}(\mathbb{C}^n, e^{-|z|^2})$ and the partial sums of the Fourier series of the components of D^*u converge to the components of D^*u in $A^2_{(p-1,0)}(\mathbb{C}^n, e^{-|z|^2})$. □

2 The Basic Estimate

We want to find conditions under which $\tilde{\Box}_{D,1}$ has a bounded inverse. For this purpose we have to consider the graph norm $(\|u\|^2 + \|Du\|^2 + \|D^*u\|^2)^{1/2}$ on $\mathrm{dom}(D) \cap \mathrm{dom}(D^*)$. We refer to [4, Theorem 5.1] here in a slightly different improved form.

Theorem 2.1 *Suppose that there exists a constant $C > 0$ such that*

$$\|u\|^2 \le C \sum_{j,k=1}^n ([p_k, p_j^*]u_j, u_k), \tag{2.1}$$

for any $(1,0)$-form $u = \sum_{j=1}^n u_j dz_j$ with polynomial components. Then

$$\|u\|^2 \le C(\|Du\|^2 + \|D^*u\|^2), \tag{2.2}$$

for any $u \in \mathrm{dom}(D) \cap \mathrm{dom}(D^*)$.

Proof First we have

$$Du = \sum_{j<k} (p_j(u_k) - p_k(u_j))\, dz_j \wedge dz_k \quad \text{and} \quad D^*u = \sum_{j=1}^{n} p_j^* u_j,$$

hence

$$\begin{aligned}
\|Du\|^2 + \|D^*u\|^2 &= \int_{\mathbb{C}^n} \sum_{j<k} |p_k(u_j) - p_j(u_k)|^2\, e^{-|z|^2}\, d\lambda \\
&\quad + \int_{\mathbb{C}^n} \sum_{j,k=1}^{n} p_j^* u_j\, \overline{p_k^* u_k}\, e^{-|z|^2}\, d\lambda \\
&= \sum_{j,k=1}^{n} \int_{\mathbb{C}^n} |p_k(u_j)|^2\, e^{-|z|^2}\, d\lambda \\
&\quad + \sum_{j,k=1}^{n} \int_{\mathbb{C}^n} (p_j^* u_j\, \overline{p_k^* u_k} - p_k(u_j)\overline{p_j(u_k)})\, e^{-|z|^2}\, d\lambda \\
&= \sum_{j,k=1}^{n} \int_{\mathbb{C}^n} |p_k(u_j)|^2\, e^{-|z|^2}\, d\lambda \\
&\quad + \sum_{j,k=1}^{n} \int_{\mathbb{C}^n} [p_k, p_j^*] u_j \overline{u_k}\, e^{-|z|^2}\, d\lambda,
\end{aligned}$$

where we used (1.2). Note that the expression

$$\sum_{j,k=1}^{n} \int_{\mathbb{C}^n} |p_k(u_j)|^2\, e^{-|z|^2}\, d\lambda$$

is finite, since the components u_j are polynomials, and it follows that the expression $\sum_{j,k=1}^{n} ([p_k, p_j^*]u_j, u_k)$ is a real number.

Now the assumption (2.1) implies that (2.2) holds for $(1,0)$-forms with polynomial components and, by Lemma 1.1, we obtain (2.2) for any $u \in \mathrm{dom}(D) \cap \mathrm{dom}(D^*)$. □

Remark 2.2 In Theorem 2.1 we implicitly suppose that the expression

$$\sum_{j,k=1}^{n} ([p_k, p_j^*]u_j, u_k)$$

is nonnegative.

First we consider the one-dimensional case. Let p_m denote the polynomial differential operator

$$p_m = a_0 + a_1 \frac{\partial}{\partial z} + \cdots + a_m \frac{\partial^m}{\partial z^m},$$

with constant coefficients $a_0, a_1, \ldots, a_m \in \mathbb{C}$, and let p_m^* denote the polynomial

$$p_m^*(z) = \overline{a}_0 + \overline{a}_1 z + \cdots + \overline{a}_m z^m,$$

with the complex conjugate coefficients $\overline{a}_0, \overline{a}_1, \ldots, \overline{a}_m \in \mathbb{C}$.

We consider the densely defined operator

$$Du = p_m(u)\, dz,$$

where $u \in A^2(\mathbb{C}, e^{-|z|^2})$ and $p_m(u)\, dz$ is considered as a $(1,0)$-form.

It is clear that $D^2 = 0$, as all $(2,0)$-forms are zero if $n = 1$, and that we have

$$(Du, v) = (u, D^* v),$$

where $u \in \mathrm{dom}(D) = \{u \in A^2(\mathbb{C}, e^{-|z|^2}) : Du \in A^2_{(1,0)}(\mathbb{C}, e^{-|z|^2})\}$ and

$$D^* v\, dz = p_m^* v.$$

In the sequel we consider the generalized box operator

$$\tilde{\Box}_{D,1} := DD^*$$

as a densely defined self-adjoint operator on $A^2_{(1,0)}(\mathbb{C}, e^{-|z|^2})$ with

$$\mathrm{dom}(\tilde{\Box}_{D,1}) = \{f \in \mathrm{dom}(D^*) : D^* f \in \mathrm{dom}(D)\}.$$

Lemma 2.3 *Let u be an arbitrary polynomial. Then*

$$\begin{aligned} &\int_{\mathbb{C}} [p_m, p_m^*] u(z)\, \overline{u(z)}\, e^{-|z|^2}\, d\lambda(z) \\ &\quad = \sum_{\ell=1}^{m} \ell\, ! \int_{\mathbb{C}} \left| \sum_{k=\ell}^{m} \binom{k}{\ell} a_k u^{(k-\ell)}(z) \right|^2 e^{-|z|^2}\, d\lambda(z). \end{aligned} \tag{2.3}$$

Proof On the right hand side of (2.3) we have the integrand

$$\sum_{\ell=1}^{m} \ell\, ! \sum_{k,j=\ell}^{m} \binom{k}{\ell} \binom{j}{\ell} a_k \overline{a}_j u^{(k-\ell)} \overline{u}^{(j-\ell)}. \tag{2.4}$$

For the left hand side of (2.3) we first compute

$$[p_m, p_m^*]u = \sum_{k=0}^{m} a_k \frac{\partial^k}{\partial z^k}\Big(\sum_{j=0}^{m} \overline{a}_j z^j u\Big) - \Big(\sum_{j=0}^{m} \overline{a}_j z^j\Big)\Big(\sum_{j=0}^{m} a_j u^{(j)}\Big).$$

We use the Leibniz rule to get

$$\frac{\partial^k}{\partial z^k}\Big(\sum_{j=0}^{m} \overline{a}_j z^j u\Big) = \sum_{j=0}^{m} \overline{a}_j \frac{\partial^k}{\partial z^k}(z^j u) = \sum_{j=0}^{m} \overline{a}_j \sum_{\ell=0}^{j} \ell\,! \binom{k}{\ell}\binom{j}{\ell} u^{(k-\ell)} z^{j-\ell},$$

notice that $\binom{k}{\ell} = 0$, in the case $k < \ell$. Hence we obtain

$$[p_m, p_m^*]u = \sum_{j,k=1}^{m} a_k \overline{a}_j \sum_{\ell=1}^{j} \ell\,! \binom{k}{\ell}\binom{j}{\ell} u^{(k-\ell)} z^{j-\ell}.$$

After integration we obtain

$$\begin{aligned}\int_{\mathbb{C}} [p_m, p_m^*]u(z)\,\overline{u(z)}\,e^{-|z|^2}\,d\lambda(z) \\ = \sum_{j,k=1}^{m} a_k \overline{a}_j \sum_{\ell=1}^{j} \ell\,! \binom{k}{\ell}\binom{j}{\ell} \int_{\mathbb{C}} u^{(k-\ell)} \overline{u}^{(j-\ell)}\, e^{-|z|^2}\,d\lambda(z). \qquad (2.5)\end{aligned}$$

Now it is easy to show that integration of (2.4) coincides with (2.5) and we are done. □

Lemma 2.4 *There exists a constant $C > 0$ such that, for each $u \in \mathrm{dom}(D^*)$,*

$$\|u\| \le C\|D^* u\|.$$

Proof First let u be a polynomial and note that the last term in (2.3) equals

$$m! \int_{\mathbb{C}} |a_m u(z)|^2\, e^{-|z|^2}\,d\lambda(z) = m!\,|a_m|^2 \|u\|^2$$

and all the other terms are non-negative, see Lemma 2.3. Now we get that

$$\|D^* u\|^2 = (p_m^* u, p_m^* u) = ([p_m, p_m^*]u, u) + (p_m(u), p_m(u)) \ge \frac{1}{C}\,\|u\|^2,$$

where

$$C = \frac{1}{m!\,|a_m|^2},$$

if we suppose that $a_m \neq 0$, and we are done. Finally apply Lemma 1.1 to obtain the desired result. □

Theorem 2.5 *Let D be as in (1.1). Then $\tilde{\Box}_{D,1} = DD^*$ has a bounded inverse*

$$\tilde{N}_{D,1} : A^2_{(1,0)}(\mathbb{C}, e^{-|z|^2}) \longrightarrow \mathrm{dom}(\tilde{\Box}_{D,1}).$$

If $\alpha \in A^2_{(1,0)}(\mathbb{C}, e^{-|z|^2})$, then $u_0 = D^*\tilde{N}_{D,1}\alpha$ is the canonical solution of $Du = \alpha$, this means that $Du_0 = \alpha$ and $u_0 \in (\ker D)^\perp = \mathrm{im}D^*$, and

$$\|D^*\tilde{N}_{D,1}\alpha\| \leq C\|\alpha\|$$

for some constant $C > 0$ independent of α.

Proof Using Lemma 2.4 we obtain that

$$\tilde{\Box}_{D,1} : \mathrm{dom}(\tilde{\Box}_{D,1}) \longrightarrow A^2_{(1,0)}(\mathbb{C}, e^{-|z|^2})$$

is bijective and has the bounded inverse $\tilde{N}_{D,1}$, see [4, Theorem 5.1]. The rest follows from [4, Theorem 5.2]. □

3 Commutators

Let A_j and B_j, $j = 1, \ldots, n$ be operators satisfying

$$[A_j, A_k] = [B_j, B_k] = [A_j, B_k] = 0,\ j \neq k$$

and

$$[A_j, B_j] = I,\ j = 1, \ldots, n.$$

Let P and Q be polynomials of n variables and write $A = (A_1, \ldots, A_n)$ and $B = (B_1, \ldots, B_n)$. Then

$$Q(A)P(B) = \sum_{|\alpha| \geq 0} \frac{1}{\alpha!} P^{(\alpha)}(B) Q^{(\alpha)}(A), \tag{3.1}$$

where $\alpha = (\alpha_1, \ldots, \alpha_n)$ are multiindices and $|\alpha| = \alpha_1 + \cdots + \alpha_n$ and $\alpha! = \alpha_1! \ldots \alpha_n!$, see [7, 8].

The assumptions are satisfied, if one takes $A_j = \frac{\partial}{\partial z_j}$ and $B_j = z_j$ the multiplication operator. The inspiration for this comes from quantum mechanics, where the annihilation operator A_j can be represented by the differentiation with

respect to z_j on $A^2(\mathbb{C}^n, e^{-|z|^2})$ and its adjoint, the creation operator B_j, by the multiplication by z_j, both operators being unbounded densely defined (see [1, 2]). One can show that $A^2(\mathbb{C}^n, e^{-|z|^2})$ with this action of the B_j and A_j is an irreducible representation M of the Heisenberg group; by the Stone-von Neumann theorem it is the only one up to unitary equivalence. Physically M can be thought of as the Hilbert space of a harmonic oscillator with n degrees of freedom and Hamiltonian operator

$$H = \sum_{j=1}^{n} \frac{1}{2}(A_j B_j + B_j A_j).$$

Remark 3.1 If we apply (3.1) for the one-dimensional case of Lemma 2.3, we get

$$\int_{\mathbb{C}} [p_m, p_m^*]u(z)\,\overline{u(z)}\,e^{-|z|^2}\,d\lambda(z) = \sum_{\ell=1}^{m} \int_{\mathbb{C}} \left|p_m^{(\ell)}u(z)\right|^2 e^{-|z|^2}\,d\lambda(z),$$

which coincides with (2.3).

In the following we consider $\mathbb{C}^2$ and choose p_1 and p_2 to be polynomials of degree 2 in 2 variables.

Theorem 3.2 *Let p_1, p_2 be polynomials of degree 2. Suppose that*

$$p_2^{(e_1)*} p_1^{(e_1)} = \pm\, p_1^{(e_2)*} p_2^{(e_2)},\ \ p_1^{(e_1)*} p_2^{(e_1)} = \pm\, p_2^{(e_2)*} p_1^{(e_2)}, \tag{3.2}$$

where (e_1) and (e_2) denote the derivatives with respect to z_1 and z_2 respectively. In addition suppose that for all derivatives (α) of order 2 we have

$$p_j^{(\alpha)*} p_k^{(\alpha)} = \delta_{j,k}\, c_{j,\alpha}\ \ j, k = 1, 2, \tag{3.3}$$

where

$$C_1 = \sum_{|\alpha|=2} \frac{1}{\alpha!}\, c_{1,\alpha} > 0 \text{ and } C_2 = \sum_{|\alpha|=2} \frac{1}{\alpha!}\, c_{2,\alpha} > 0.$$

Then

$$\|u\|^2 \le \frac{1}{\min(C_1, C_2)} \sum_{j,k=1}^{2} ([p_k, p_j^*]u_j, u_k), \tag{3.4}$$

for any $(1,0)$-form $u = \sum_{j=1}^{2} u_j dz_j$ with polynomial components.

Proof Using (3.1) we obtain

$$([p_k, p_j^*]u_j, u_k) = \sum_{|\alpha|\geq 1} \frac{1}{\alpha!} (p_j^{(\alpha)*} p_k^{(\alpha)} u_j, u_k).$$

Now we use (3.2) and get for the first order derivatives

$$\begin{aligned}
&\sum_{j,k=1}^{2} [(p_j^{(e_1)*} p_k^{(e_1)} u_j, u_k) + (p_j^{(e_2)*} p_k^{(e_2)} u_j, u_k)] \\
&\quad = (p_1^{(e_1)*} p_1^{(e_1)} u_1, u_1) + (p_2^{(e_1)*} p_2^{(e_1)} u_2, u_2) \\
&\qquad \pm (p_2^{(e_2)*} p_1^{(e_2)} u_1, u_2) \pm (p_1^{(e_2)*} p_2^{(e_2)} u_2, u_1) \\
&\qquad + (p_1^{(e_2)*} p_1^{(e_2)} u_1, u_1) + (p_2^{(e_2)*} p_2^{(e_2)} u_2, u_2) \\
&\qquad \pm (p_2^{(e_1)*} p_1^{(e_1)} u_1, u_2) \pm (p_1^{(e_1)*} p_2^{(e_1)} u_2, u_1) \\
&\quad = (p_1^{(e_1)} u_1 \pm p_2^{(e_1)} u_2, p_1^{(e_1)} u_1 \pm p_2^{(e_1)} u_2) \\
&\qquad + (p_1^{(e_2)} u_1 \pm p_2^{(e_2)} u_2, p_1^{(e_2)} u_1 \pm p_2^{(e_2)} u_2) \\
&\quad = \| p_1^{(e_1)} u_1 \pm p_2^{(e_1)} u_2 \|^2 + \| p_1^{(e_2)} u_1 \pm p_2^{(e_2)} u_2 \|^2.
\end{aligned}$$

For the second order derivatives we obtain

$$\sum_{j,k=1}^{2} \sum_{|\alpha|=2} \frac{1}{\alpha!} (p_j^{(\alpha)*} p_k^{(\alpha)} u_j, u_k) = C_1 \|u_1\|^2 + C_2 \|u_2\|^2.$$

Hence we get

$$\begin{aligned}
\sum_{j,k=1}^{2} ([p_k, p_j^*]u_j, u_k) =& C_1 \|u_1\|^2 + C_2 \|u_2\|^2 \\
&+ \| p_1^{(e_1)} u_1 \pm p_2^{(e_1)} u_2 \|^2 + \| p_1^{(e_2)} u_1 \pm p_2^{(e_2)} u_2 \|^2,
\end{aligned}$$

which gives (3.4). □

In a similar way one shows the following.

Theorem 3.3 *Let p_1, p_2 be polynomials of degree 2. Suppose that*

$$p_2^{(e_1)*} p_1^{(e_1)} = \pm\, p_1^{(e_1)*} p_2^{(e_1)}, \;\; p_1^{(e_2)*} p_2^{(e_2)} = \pm\, p_2^{(e_2)*} p_1^{(e_2)}, \tag{3.5}$$

where (e_1) *and* (e_2) *denote the derivatives with respect to* z_1 *and* z_2 *respectively. In addition suppose that for all derivatives* (α) *of order* 2 *we have*

$$p_j^{(\alpha)*} p_k^{(\alpha)} = \delta_{j,k}\, c_{j,\alpha} \;\; j,k = 1,2, \tag{3.6}$$

where

$$C_1 = \sum_{|\alpha|=2} \frac{1}{\alpha!}\, c_{1,\alpha} > 0 \text{ and } C_2 = \sum_{|\alpha|=2} \frac{1}{\alpha!}\, c_{2,\alpha} > 0.$$

Then

$$\|u\|^2 \le \frac{1}{\min(C_1, C_2)} \sum_{j,k=1}^{2} ([p_k, p_j^*]u_j, u_k),$$

for any $(1,0)$*-form* $u = \sum_{j=1}^{2} u_j dz_j$ *with polynomial components.*

Finally we exhibit some examples, where conditions (3.2) and (3.3), or (3.5) and (3.6) are checked. Examples (a) and (c) are taken from [4], where $\sum_{j,k=1}^{n} ([p_k, p_j^*]u_j, u_k)$ was directly computed.

Example

(a) We take $p_1 = \frac{\partial^2}{\partial z_1 \partial z_2}$ and $p_2 = \frac{\partial^2}{\partial z_1^2} + \frac{\partial^2}{\partial z_2^2}$. Then $p_1^*(z) = z_1 z_2$ and $p_2^*(z) = z_1^2 + z_2^2$ and we see that (3.2) and (3.3) are satisfied:

$$p_2^{(e_1)*} p_1^{(e_1)} = p_1^{(e_2)*} p_2^{(e_2)}, \quad p_1^{(e_1)*} p_2^{(e_1)} = p_2^{(e_2)*} p_1^{(e_2)},$$

and we obtain

$$\sum_{j,k=1}^{2} ([p_k, p_j^*]u_j, u_k)$$
$$= \int_{\mathbb{C}^2} \left(|u_1|^2 + 4|u_2|^2 + \left| \frac{\partial u_1}{\partial z_1} + 2\frac{\partial u_2}{\partial z_2} \right|^2 + \left| \frac{\partial u_1}{\partial z_2} + 2\frac{\partial u_2}{\partial z_1} \right|^2 \right) e^{-|z|^2}\, d\lambda,$$

for $u = \sum_{j=1}^{2} u_j dz_j$ with polynomial components.

(b) Taking $p_1 = i\frac{\partial^2}{\partial z_1 \partial z_2}$ and $p_2 = \frac{\partial^2}{\partial z_1^2} + \frac{\partial^2}{\partial z_2^2}$ we have that $p_1^*(z) = -i z_1 z_2$ and $p_2^*(z) = z_1^2 + z_2^2$ and that (3.2) and (3.3) are satisfied:

$$p_2^{(e_1)*} p_1^{(e_1)} = -p_1^{(e_2)*} p_2^{(e_2)}, \quad p_1^{(e_1)*} p_2^{(e_1)} = -p_2^{(e_2)*} p_1^{(e_2)},$$

and we obtain

$$\sum_{j,k=1}^{2} ([p_k, p_j^*]u_j, u_k)$$
$$= \int_{\mathbb{C}^2} \left(|u_1|^2 + 4|u_2|^2 + \left| \frac{\partial u_1}{\partial z_1} + 2i \frac{\partial u_2}{\partial z_2} \right|^2 + \left| \frac{\partial u_1}{\partial z_2} + 2i \frac{\partial u_2}{\partial z_1} \right|^2 \right) e^{-|z|^2} \, d\lambda,$$

for $u = \sum_{j=1}^{2} u_j dz_j$ with polynomial components.

(c) Let $p_k = \frac{\partial^2}{\partial z_k^2}, k = 1, 2$. Then $p_j^*(z) = z_j^2, j = 1, 2$ and we see that (3.5) and (3.6) are satisfied and we have

$$\sum_{j,k=1}^{2} ([p_k, p_j^*]u_j, u_k) = \sum_{j,k=1}^{2} (2\delta_{j,k} u_j, u_k) + \sum_{j,k=1}^{2} (4\delta_{jk} z_j \frac{\partial u_j}{\partial z_k}, u_k)$$
$$= 2\|u\|^2 + 4 \sum_{j=1}^{2} \left\| \frac{\partial u_j}{\partial z_j} \right\|^2.$$

(d) For $p_1 = \frac{\partial^2}{\partial z_1^2} + \frac{\partial}{\partial z_2}$ and $p_2 = \frac{\partial}{\partial z_1} + \frac{\partial^2}{\partial z_2^2}$ we have $p_1^*(z) = z_1^2 + z_2$ and $p_2^*(z) = z_1 + z_2^2$ and we see that (3.2) and (3.5) are not satisfied. In particular,

$$\sum_{j,k=1}^{2} ([p_k, p_j^*]u_j, u_k) = 3(\|u_1\|^2 + \|u_2\|^2) + 4 \left(\left\| \frac{\partial u_1}{\partial z_1} \right\|^2 + \left\| \frac{\partial u_2}{\partial z_2} \right\|^2 \right)$$
$$+ 2(u_1, z_2 u_2) + 2(z_1 u_1, u_2) + 2(u_2, z_1 u_1) + 2(z_2 u_2, u_1)$$

for $u = \sum_{j=1}^{2} u_j dz_j$ with polynomial components.

Acknowledgments The author thanks the referees for several useful suggestions.
This project was partially supported by the Austrian Science Fund (FWF) project P28154.

References

1. L.D. Fadeev, O.A. Yakubovskii, *Lectures on Quantum Mechanics for Mathematics Students*. Student Mathematical Library, vol. 47 (American Mathematical Society, Providence, 2009)
2. G.B. Folland, *Harmonic Analysis in Phase Space*. Annals of Mathematics Studies, vol. 122 (Princeton University Press, Princeton, 1989)
3. F. Haslinger, The $\overline{\partial}$-Neumann Problem and Schrödinger Operators. de Gruyter Expositions in Mathematics, vol. 59 (Walter de Gruyter, Berlin, 2014)

4. F. Haslinger, The ∂-complex on the Segal-Bargmann space. Ann. Polon. Mat. **123**, 295–317 (2019)
5. D.J. Newman, H.S. Shapiro, Certain Hilbert spaces of entire functions. Bull. Am. Math. Soc. **72**, 971–977 (1966)
6. D.J. Newman, H.S. Shapiro, Fischer spaces of entire functions, in *Entire Functions and Related Parts of Analysis*. Proceedings of Symposium in Pure Mathematics La Jolla, 1966 (American Mathematical Society, Providence, 1968), pp. 360–369
7. D.G. Quillen, On the representation of Hermitian forms as sums of squares. Invent. Math. **5**, 237–242 (1968)
8. F. Treves, *Linear Partial Differential Equations with Constant Coefficients* (Gordon and Breach, New York, 1966)

The Inverse Characteristic Polynomial Problem for Trees

Charles R. Johnson and Emma Gruner

Abstract It is known that any real polynomial is attained as the characteristic polynomial of a real combinatorially symmetric matrix, whose graph is either a path or a star. We conjecture that the same is true for any tree (this is so for complex characteristic polynomials and complex matrices). Here, we constructively prove a very large portion of this conjecture by a method that mates the graph of the polynomial with a notion of balance of the tree relative to a choice of root for the tree. Included is the first constructive proof for the case of the path, as well as the case of any tree on fewer than ten vertices. It also includes the known case of polynomials with distinct real roots for any tree (in a new way). This work is motivated by, and lies in contrast to, the considerable study of possible multiplicity lists for the eigenvalues of real symmetric matrices, whose graph is a tree.

Keywords Inverse eigenvalue problem · Characteristic polynomial · Partial fractions · Graph theory · Matrix theory

Mathematics Subject Classification (2010) Primary 15A29, 05C50; Secondary 05C05

1 Introduction

During the last 20+ years, a theory of the possible multiplicities of the eigenvalues, of Hermitian matrices with a given graph, has developed. This is the most developed, and most interesting, in the case of trees, but much is now known for general graphs.

C. R. Johnson (✉)
The College of William and Mary, Williamsburg, VA, USA
e-mail: crjohn@wm.edu

E. Gruner
Gettysburg College, Gettysburg, PA, USA
e-mail: grunem01@gettysburg.edu

M. A. Bastos et al. (eds.), *Operator Theory, Functional Analysis and Applications*,
Operator Theory: Advances and Applications 282,
https://doi.org/10.1007/978-3-030-51945-2_17

Recently the theory has expanded to geometric multiplicities for the eigenvalues of a general (combinatorially symmetric) matrix; there are strong similarities and some notable differences. Though ongoing, much of this work is summarized in the recent book [6].

The question then naturally arises: "What about algebraic multiplicities of the eigenvalues of a matrix with a given graph?" Here, the field makes a big difference. In the case of the complex numbers, there is a complete answer because of the "additive inverse eigenvalue problem." Recall that the diagonal entries of a matrix are independent of its undirected graph. Over the complex numbers, for a given n-by-n matrix A and given desired complex numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$, there is a diagonal matrix D such that $A + D$ has eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ [3].

This is not true over the reals. There it is natural to ask which real polynomials occur for real matrices with a given undirected graph. There is good reason to guess that any real polynomial may occur when the graph is connected. But, this is not easy to prove. Initially, it is natural to consider trees. Recently, it has been shown [2] that for the path (tridiagonal matrices), any real polynomial can occur. But, the proof is existential. It is shown that a real matrix may be constructed for a given monic, degree n real polynomial when there is a degree $n - 1$ polynomial with a certain relationship to the former. Existence may be shown, but construction is difficult. For the star (all vertices pendent from a single vertex), it has also recently been shown [5] that any real polynomial occurs over $\mathbb{R}$. The proof is constructive and much more is shown, just for the star. We benefit here from the technology therein, although the technology does not easily transfer to other trees.

Here, we consider general trees and real matrices; for a tree T, $\mathscr{R}(T)$ denotes the set of real matrices whose graph is T (not necessarily symmetric, but combinatorially symmetric, and no restriction besides reality on the diagonal entries). A method is developed that, for each tree, realizes many characteristic polynomials in $\mathscr{R}(T)$, and for many trees, realizes all real characteristic polynomials. The method is constructive and gives all real polynomials for the path (the first time that has occurred), as well as many other trees. The smallest tree for which not all real polynomials are realized has ten vertices, and this is not because of the number of vertices, but because of a notion of balance of the tree.

2 Preliminaries

We assume that the reader is familiar with the basic terminology from linear algebra and graph theory: see [4] and [6]. However, we shall briefly elaborate on the concept of the *graph of a matrix* discussed in the introduction.

Given an n-by-n matrix $A = (a_{ij})$ over a field $\mathbb{F}$, we let $\mathscr{G}(A)$ denote the *graph of* A. This graph consists of vertices $\{1, 2, \ldots, n\}$ with an edge $\{i, j\}$, $i \neq j$, if and only if $a_{ij} \neq 0$. For $\mathscr{G}(A)$ to be undirected, A must be *combinatorially symmetric*; that is, $a_{ij} \neq 0$ if and only if $a_{ji} \neq 0$. This paper will only be concerned with real matrices; given an undirected graph G, we let $\mathscr{R}(G)$ denote all real matrices whose

graph is G. If v is a vertex of $\mathscr{G}(A)$, then we let $A(v)$ denote the principal submatrix of A resulting from the deletion of the row and column corresponding to v.

The matrices we construct will all be of a specific form. Suppose we have a tree T on n vertices, and let v be any vertex with degree s. If $s \geq 1$, let $u_1, u_2, \ldots, u_s$ be the neighboring vertices of v. Additionally, let $T_1, T_2, \ldots, T_s$ be the branches of T at v, in which each T_i contains the vertex u_i, and let $\ell_1, \ell_2, \ldots, \ell_s$ be the respective number of vertices in $T_1, T_2, \ldots, T_s$. From this point forward, we will refer to the number of vertices in a tree or a branch of a tree as its *weight*.

We say that $A \in \mathscr{R}(T)$ is *centered at* v if it is 1-by-1, or if it has the following form:

$$A = \begin{bmatrix} b & v_1 & \cdots & v_s \\ e_{1,1} & A_1 & 0 & 0 \\ \vdots & 0 & \ddots & \vdots \\ e_{1,s} & 0 & \cdots & A_s \end{bmatrix}.$$

In this form, $b \in \mathbb{R}$, $e_{1,i} = e_1 \in \mathbb{R}^{\ell_i}$, and $v_i = (a_i e_{1,i})^T$, with $a_i \in \mathbb{R} \setminus \{0\}$. (Note that e_1 represents the basic unit vector with a 1 in the first entry and 0's elsewhere.) Each A_i is a matrix in $\mathscr{R}(T_i)$ that is centered at u_i.

Define A to be *pseudosymmetric* if it is 1-by-1, or if it has the above form and each A_i is diagonally similar to a symmetric matrix. We will usually be constructing matrices that have this property.

The *neighbors formula* [6] provides a convenient recursive representation of the characteristic polynomial for a matrix in this form. Let p_A denote the characteristic polynomial of the matrix A. If $A = [b]$, then $p_A = (x - b)$. Otherwise, we have:

$$\begin{aligned} p_A &= (x-b)\prod_{i=1}^{s} p_{A_i} - \sum_{i=1}^{s} a_i\, p_{A_i(u_i)} \prod_{j=1, j\neq i}^{s} p_{A_j} \\ &= (x-b)\, p_{A(v)} - \sum_{i=1}^{s} a_i\, p_{A_i(u_i)} \prod_{j=1, j\neq i}^{s} p_{A_j}. \end{aligned} \tag{1}$$

We define the empty product, and the characteristic polynomial of the empty matrix, to be 1.

Given any vertex v of T and any matrix $A \in \mathscr{R}(T)$, there is a diagonal and/or permutation similarity that takes A to a matrix centered at v. Since the characteristic polynomial is similarity invariant, if we are trying to realize a particular polynomial for a particular tree, we can take our matrix to be centered at whatever vertex is most convenient for us.

Finally, we will provide an overview of the algebraic technique known as the partial fraction decomposition, or PFD.

The following theorem was adapted from a more general statement in [1]. The proof is straightforward and is thus omitted.

Theorem 2.1 ([1]) *Let* $g(x) = \prod_{i=1}^{k}(x - \gamma_i)$, *in which* $\gamma_1, \gamma_2, \ldots, \gamma_k$ *are real numbers such that* $\gamma_1 > \gamma_2 > \cdots > \gamma_k$. *Let* $Q_i(x) = \prod_{j=1, j\neq i}^{k}(x - \gamma_j)$. *Finally, let* $f(x)$ *be a real polynomial of degree less than* k. *Then we have*

$$\frac{f(x)}{g(x)} = \frac{a_1}{x - \gamma_1} + \frac{a_2}{x - \gamma_2} + \cdots + \frac{a_k}{x - \gamma_k}$$

in which each a_i *is a real number determined uniquely by the expression* $a_i = f(\gamma_i)/Q_i(\gamma_i)$.

Also of interest to us is the following lemma. The case in which the polynomials f and g are both monic is well-known, but the proof is easily generalized.

Lemma 2.2 *Let* $g(x)$, $Q_i(x)$ *and* $f(x)$ *be defined as in Theorem 2.1. The polynomial* $f(x)$ *has exactly* $k - 1$ *distinct real roots that strictly interlace the roots of* $g(x)$ *if and only if the coefficients of the partial fraction decomposition* $\frac{f(x)}{g(x)}$, *denoted* $a_1, a_2, \ldots, a_k$, *are either all positive or all negative.*

Proof Consider the coefficients a_i and a_{i+1} for $i \in \{1, 2, \ldots, k-1\}$. By the previous theorem, we have $a_i = f(\gamma_i)/Q_i(\gamma_i)$ and $a_{i+1} = f(\gamma_{i+1})/Q_{i+1}(\gamma_{i+1})$. Observe that $Q_i(\gamma_i)$ and $Q_{i+1}(\gamma_{i+1})$ will always be nonzero and of opposite signs; therefore, a_i and a_{i+1} will be nonzero and of the same sign if and only if $f(\gamma_i)$ and $f(\gamma_{i+1})$ are nonzero and of opposite signs. This will occur if and only if neither γ_i nor γ_{i+1} are roots of f, but f has an odd number of roots between γ_i and γ_{i+1}.

Therefore, if f has $k - 1$ distinct real roots that strictly interlace the numbers $\gamma_1, \gamma_2, \ldots, \gamma_k$, then the a_i's must either be all positive or all negative. Conversely, if the a_i's are all of the same sign, then f must have at least one real root lying strictly between each consecutive pair of γ_i's. This means that f has at least $k - 1$ distinct real roots; however, since $\deg f(x) < k$, f can have at most $k - 1$ roots, and the result follows. □

Finally, we provide the following result, which combines the techniques of partial fractions and the polynomial division algorithm.

Lemma 2.3 *Let* p *be a real monic polynomial of degree* n, *and let* $g(x) = \prod_{i=1}^{n-1}(x - \gamma_i)$, *in which* $\gamma_1, \gamma_2, \ldots, \gamma_{n-1}$ *are real numbers such that* $\gamma_1 > \gamma_2 > \cdots > \gamma_{n-1}$ *and* $p(\gamma_i) \neq 0$ *for all* i. *Let* $Q_i(x) = \prod_{j=1, j\neq i}^{n-1}(x - \gamma_j)$. *Then we have*

$$\frac{p(x)}{g(x)} = (x - b) - \frac{a_1}{x - \gamma_1} - \frac{a_2}{x - \gamma_2} - \cdots - \frac{a_{n-1}}{x - \gamma_{n-1}},$$

with $b \in \mathbb{R}$ *and* $a_i = -p(\gamma_i)/Q_i(\gamma_i)$. *Furthermore,* a_i *and* a_{i+1} *are of the same sign for any* $i \in \{1, 2, \ldots, n-2\}$ *if and only if* p *has an odd number of roots lying strictly between* γ_i *and* γ_{i+1}.

Proof By the division algorithm for polynomials, we can write

$$p(x) = q(x)g(x) - r(x),$$

in which $q(x)$ and $r(x)$ are real polynomials and $\deg r(x) < \deg g(x) = n - 1$. (While the division algorithm normally allows for the possibility of a zero remainder, this cannot happen in our case because p and g are relatively prime.)

Since p and g are both monic and they differ by only one degree, we know that $q(x) = x - b$ for some unique real number b. So we may write

$$\frac{p(x)}{g(x)} = \frac{(x-b)g(x) - r(x)}{g(x)} = (x-b) - \frac{r(x)}{g(x)}.$$

Since $\deg r(x) < \deg g(x)$, we can apply Theorem 2.1 to write

$$\frac{r(x)}{g(x)} = \frac{a_1}{x-\gamma_1} + \frac{a_2}{x-\gamma_2} + \cdots + \frac{a_{n-1}}{x-\gamma_{n-1}},$$

with $a_i = r(\gamma_i)/Q_i(\gamma_i)$. Since each γ_i is a root of g, we have $r(\gamma_i) = -p(\gamma_i)$ for all $i \in \{1, 2, \ldots, n-1\}$, so we can just as well write $a_i = -p(\gamma_i)/Q_i(\gamma_i)$. The last statement of the lemma follows from a proof similar to that of Lemma 2.2. □

3 Main Results

The ease of constructing a matrix for a particular polynomial and tree hinges largely on the possible signs of the coefficients in a partial fraction decomposition. Therefore, we would like to find a way to classify a polynomial according to this characteristic.

Suppose we have a monic real polynomial p of degree n. As x ranges from $-\infty$ to $+\infty$, $p(x)$ undergoes $n - d$ strict sign changes, in which d is an even integer at least 0 and at most n (if n is even) or $n - 1$ (if n is odd). We shall call this integer d the *root deficiency* of p.

The root deficiency of p is directly related to its complete factorization over the real polynomials. Suppose we have

$$p = w_1^{s_1} w_2^{s_2} \cdots w_\alpha^{s_\alpha} q_1^{t_1} q_2^{t_2} \cdots q_\beta^{t_\beta},$$

in which $w_1, \ldots, w_\alpha$ are distinct linear polynomials, $q_1, \ldots, q_\beta$ are distinct irreducible quadratic polynomials, and each s_i and t_i is a positive integer. (The fact that such a representation of p exists, and is unique, is a consequence of the fundamental theorem of algebra.) Define $\rho(s_i)$ as follows:

$$\rho(s_i) = \begin{cases} s_i, & \text{if } s_i \text{ is even} \\ s_i - 1, & \text{if } s_i \text{ is odd.} \end{cases}$$

We then have $d = \sum_{i=1}^{\alpha} \rho(s_i) + \sum_{j=1}^{\beta} 2t_j$.

Observe that $d = 0$ if and only if p has n distinct linear factors. On the other hand, d has maximal value if and only if 1) p is composed entirely of irreducible quadratics, or 2) at most one s_i is odd.

Let us also define a *signed root* to be a root at which p undergoes a strict sign change. These are precisely the roots whose corresponding linear terms have odd exponents, and there are $n - d$ of them.

Example 3.1 The polynomial $p_1(x) = x^3 + x^2 - 4x - 4 = (x + 1)(x - 2)(x + 2)$ has root deficiency 0. The signed roots of p_1 are -1, 2, and -2.

Example 3.2 The polynomial $p_2(x) = x^2(x-5)^3(x+1)(x^2+1)$ has root deficiency 6. The signed roots of p_2 are 5 and -1.

Let p be a monic real polynomial of degree n, and let $\gamma_1 > \gamma_2 > \cdots > \gamma_{n-1}$ be real numbers with $p(\gamma_i) \neq 0$ for each $i \in \{1, 2, \ldots, n-1\}$. Suppose we apply the combined division algorithm/PFD procedure to p and $g = \prod_{i=1}^{n-1}(x - \gamma_i)$ as in Lemma 2.3. The potential positive/negative breakdown of the a_i's we obtain is connected to the root deficiency of p.

Proposition 3.3 *Let p be a monic real polynomial of degree n and root deficiency d. Let $k_1 \geq k_2$ be nonnegative integers with $k_1 + k_2 = n - 1$. We can choose real numbers $\gamma_1 > \gamma_2 > \cdots > \gamma_{n-1}$ such that the representation of $\frac{p(x)}{\prod_{i=1}^{n-1}(x-\gamma_i)}$ shown in Lemma 2.3 yields k_1 negative a_i's and k_2 positive a_i's if and only if $k_2 \geq (d-2)/2$.*

Proof For the "only if" direction, assume that we have a selection of $(n - 1)$ γ_i's that yields k_1 negative a_i's and k_2 positive a_i's for a particular polynomial p, with $k_1 \geq k_2$. We shall prove that d must be at most $2k_2+2$, implying that $k_2 \geq (d-2)/2$.

Let $z = k_1 - k_2$. We claim that we must have at least $z - 1$ distinct consecutive pairs of negative a_i's. We consider two consecutive pairs distinct if they differ in at least one element: for example, the set $\{a_1, a_2, a_3\}$ consists of two consecutive pairs.

If we ignore all the positive a_i's, then we have k_1 negative a_i's, which consists of $k_1 - 1$ distinct consecutive pairs. Each positive a_i that we introduce can break up at most one of these consecutive pairs (none if it is terminal or adjacent to an already introduced a_i). When we have introduced all k_2 of these a_i's, we are left with at least $(k_1 - 1) - k_2 = z - 1$ consecutive pairs of negative a_i's. Since each consecutive pair of same-sign a_i's necessitates at least one signed root of p between the corresponding γ_i's (see Lemma 2.3), p must have at least $z - 1$ signed roots.

Recall that p has $n-d$ distinct signed roots, so $z-1 \leq n-d$. Substituting k_1-k_2 for z and $k_1 + k_2 + 1$ for n, we determine that $d \leq 2k_2 + 2$, or $k_2 \geq (d - 2)/2$, as claimed.

For the "if" direction, suppose we have k_1 and k_2 in mind such that $k_1+k_2 = n-1$ and $k_2 \geq (d - 2)/2$. When we choose γ_i's in the following steps, we assume that no γ_i is a root of p, signed or "unsigned."

Let $R_1, R_2, \ldots, R_{n-d}$ be the distinct signed roots of p, with $R_1 > R_2 > \cdots > R_{n-d}$. If $n - d > 0$, then take $R = R_1$; otherwise, take R to be any real number.

If $k_2 = k_1 = (n - 1)/2$, then choose all γ_i's to be larger than R. Then by Lemma2.3, the a_i's will alternate in sign, and with $n - 1$ being even we

will have equal numbers of positive and negative a_i's. Otherwise, choose only $\gamma_1, \gamma_2, \ldots, \gamma_{2k_2+1}$ to be larger than R. This choice will ensure that the corresponding a_i's alternate in sign. Note that $a_1 = \frac{-p(\gamma_1)}{Q_1(\gamma_1)}$, as defined in Lemma 2.3; since both $p(\gamma_1)$ and $Q_1(\gamma_1)$ are positive, a_1 will be negative, as will $a_3, a_5, \ldots, a_{2k_2+1}$. Likewise, $a_2, a_4, \ldots, a_{2k_2}$ will be positive. In total, we will have k_2+1 negative a_i's and k_2 positive a_i's.

Since we now have all the positive a_i's we want, we'd like the remaining a_i's, of which there are $(k_1 + k_2) - (2k_2 + 1) = k_1 - k_2 - 1 = z - 1$, to all be negative. But since $(d-2)/2 \leq k_2$, we have that $z - 1 \leq n - d$. Therefore, we can choose the remaining γ_i's such that

$$R_1 > \gamma_{2k_2+2} > R_2 > \gamma_{2k_2+3} > R_3 > \cdots > R_{z-1} > \gamma_{n-1}.$$

Since exactly one signed root occurs between γ_i and γ_{i+1} for $i \in \{2k_2+1, 2k_2+2, \ldots, n-2\}$, by Lemma 2.3 a sign change will not occur between the corresponding a_i and a_{i+1}, so the remaining a_i's will all be negative, as desired. □

Arguably the most relevant previous result to our work is the following theorem, first proved by Leal-Duarte in [7]. We present a slightly different proof, but the basic argument is the same.

Theorem 3.4 ([7]) *Let T be a tree on n vertices, and let v be any vertex of T. Suppose p and g are monic polynomials of degrees n and $n-1$ respectively, that both polynomials have all distinct, real roots, and that the roots of g strictly interlace the roots of p. Then there exists a matrix $A \in \mathscr{R}(T)$, centered at v, such that A has characteristic polynomial p and $A(v)$ has characteristic polynomial g. Furthermore, this matrix is diagonally similar to a symmetric matrix.*

Proof We proceed by induction on n. Suppose first that $n = 1$; then $g = 1$, $p = (x - b)$ for some real number b, and $A = [b]$ is the matrix we seek. Note that A is diagonally similar to a symmetric matrix: namely, itself, via the similarity matrix $D = [1]$. Now suppose $n > 1$, and the claim holds for t with $n > t \geq 1$. Then v must have degree $s \geq 1$; let $u_1, u_2, \ldots, u_s$ be the neighboring vertices of v, in which each u_i is contained in the branch T_i with weight ℓ_i.

Suppose $g = \prod_{i=1}^{n-1}(x - \mu_i)$ for $\mu_1 > \mu_2 > \cdots > \mu_{n-1}$, and write

$$\frac{p(x)}{g(x)} = (x - b) - \left[\frac{a_1}{x - \mu_1} + \frac{a_2}{x - \mu_2} + \cdots + \frac{a_{n-1}}{x - \mu_{n-1}}\right],$$

as in Lemma 2.3.

Since p and g have strictly interlacing roots, we can use the last statement of Lemma 2.3 to conclude that all a_i's are of the same sign. In this case, they are all positive, since $a_1 = -p(\mu_1)/Q_1(\mu_1)$ is positive.

Suppose we partition the $n-1$ terms of the PFD into s partial sums, with the ith partial sum containing ℓ_i terms of the original expression—the exact grouping

does not matter here. Then we combine the terms within each partition into a single rational expression $h_i(x)/P_i(x)$, with $\deg P_i(x) = \ell_i$ and $\deg h_i(x) < \ell_i$. We now have:

$$\frac{p(x)}{g(x)} = (x-b) - \left[\frac{h_1(x)}{P_1(x)} + \frac{h_2(x)}{P_2(x)} + \cdots + \frac{h_s(x)}{P_s(x)}\right].$$

By the uniqueness of the partial fraction decomposition, we know that if we performed a PFD of any individual expression $h_i(x)/P_i(x)$, we would simply regenerate ℓ_i terms of our original PFD. We already established that these terms all have positive numerators; therefore by Lemma 2.2 each $h_i(x)$ has $\ell_i - 1$ real roots that strictly interlace those of $P_i(x)$.

Furthermore, we can write each $h_i(x)$ as $\alpha_i \hat{h}_i(x)$, in which α_i is the leading coefficient of $h_i(x)$ and $\hat{h}_i(x)$ is the monic polynomial obtained by dividing $h_i(x)$ by α_i. Observe that α_i is just the sum of the numerators of the partial fraction terms that were consolidated to yield $h_i(x)/P_i(x)$; since these numerators are all positive, so is α_i.

Also observe that $\hat{h}_i$ has the same roots as h_i, which, as just discussed, strictly interlace those of P_i. Since $\deg P_i(x) = \ell_i < n$, by the inductive hypothesis there exists a matrix $A_i \in \mathscr{R}(T_i)$, centered at u_i, such that A_i has characteristic polynomial P_i and $A_i(u_i)$ has characteristic polynomial $\hat{h}_i$. Construct such an A_i for each pair of $\hat{h}_i$ and P_i.

Now, let

$$A = \begin{bmatrix} b & (\alpha_1 e_{1,1})^T & \cdots & (\alpha_s e_{1,s})^T \\ e_{1,1} & A_1 & 0 & 0 \\ \vdots & 0 & \ddots & \vdots \\ e_{1,s} & 0 & \cdots & A_s \end{bmatrix},$$

which is a matrix in $\mathscr{R}(T)$ and centered at v.

By the neighbors formula (1), we have:

$$\begin{aligned} p_A &= (x-b)\prod_{i=1}^{s} p_{A_i} - \sum_{i=1}^{s} a_i\, p_{A_i(u_i)} \prod_{j=1,\, j\neq i}^{s} p_{A_j} \\ &= (x-b)P_1(x)\cdots P_s(x) - \sum_{i=1}^{s} \alpha_i \hat{h}_i(x) \prod_{j=1,\, j\neq i}^{s} P_j(x) \\ &= (x-b)g(x) - \sum_{i=1}^{s} \alpha_i \hat{h}_i(x) \prod_{j=1,\, j\neq i}^{s} P_j(x) = p(x). \end{aligned}$$

Furthermore, the matrix $A(v)$ is equal to

$$A(v) = \begin{bmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & 0 & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & A_s \end{bmatrix},$$

which has characteristic polynomial

$$p_{A(v)} = \prod_{i=1}^{s} p_{A_i} = P_1(x) \cdots P_s(x) = g(x),$$

as claimed.

By our inductive hypothesis, we have that for each $i \in \{1, 2, \ldots, s\}$, $A_i = D_i^{-1} B_i D_i$, in which D_i is a diagonal matrix of appropriate size and B_i is symmetric. Since each α_i is positive, we can construct the diagonal matrix

$$D = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & \sqrt{\alpha_1} D_1 & 0 & \cdots & 0 \\ 0 & 0 & \sqrt{\alpha_2} D_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \sqrt{\alpha_s} D_s \end{bmatrix}.$$

Note that as long as $D_1, D_2, \ldots, D_s$ are constructed in this same inductive manner, they will each have the $(1, 1)$ entry equal to 1. Therefore, one can verify that

$$DAD^{-1} = \begin{bmatrix} b & (\sqrt{\alpha_1} e_{1,1})^T & \cdots & (\sqrt{\alpha_s} e_{1,s})^T \\ \sqrt{\alpha_1} e_{1,1} & B_1 & 0 & 0 \\ \vdots & & 0 & \ddots & \vdots \\ \sqrt{\alpha_s} e_{1,s} & 0 & \cdots & B_s \end{bmatrix},$$

which is symmetric. Therefore, our constructed matrix A is diagonally similar to a symmetric matrix, as claimed. □

We note that the original theorem and proof in [7] produce a matrix A that is itself symmetric, instead of just symmetric by similarity. However, we amended it to keep consistency with the canonical form we use.

We are now ready to prove our main result.

Theorem 3.5 *Let p be a monic real polynomial of degree n and root deficiency d, and let T be a tree on n vertices. Let $k = n - 1$ and $\ell = (d - 2)/2$. Then p has a pseudosymmetric realization $A \in \mathscr{R}(T)$ if there exists a vertex v of T with branches $T_1, \ldots, T_s$ with respective weights $\ell_1, \ldots, \ell_s$ such that some partial sum of $\{\ell_1, \ldots, \ell_s\} \in [\ell, k - \ell]$.*

Proof Let v be a vertex of T that meets the above specifications. Without loss of generality, assume that $\ell_1 + \ell_2 + \cdots + \ell_t = C$, with $t \leq s$ and $C \in [\ell, k-\ell]$. (Note that since $d \leq n$, we know $\ell < k/2$ and therefore the range $[\ell, k-\ell]$ contains at least two integers.) Since $k - C$ would also be in $[\ell, k-\ell]$, we can also assume without loss of generality that $C \leq k - C$. By Proposition 3.3, we can choose $\gamma_1, \gamma_2, \ldots, \gamma_k$ such that, if we let $g = \prod_{i=1}^{k}(x - \gamma_i)$ and represent p/g as in Lemma 2.3, we obtain C positive coefficients and $k - C$ negative coefficients among the a_i's.

Relabeling the γ_i's and a_i's if necessary, write

$$\frac{p(x)}{g(x)} = (x-b) - \left[\frac{a_1}{x-\gamma_1} + \frac{a_2}{x-\gamma_2} + \cdots + \frac{a_C}{x-\gamma_C} + \frac{a_{C+1}}{x-\gamma_{C+1}} + \cdots + \frac{a_k}{x-\gamma_k}\right],$$

such that $a_1, \ldots, a_C$ are positive and $a_{C+1}, \ldots, a_k$ are negative.

Now the proof resembles that of the last theorem. We can partition the first C terms of the PFD into t partial sums, and then the last $k - C$ terms into $s - t$ partial sums, with the ith partial sum containing ℓ_i terms of the original expression. Then we combine the terms within each partition into a single rational expression $h_i(x)/P_i(x)$, with $\deg P_i(x) = \ell_i$ and $\deg h_i(x) < \ell_i$. We now have:

$$\frac{p(x)}{g(x)} = (x-b) - \left[\frac{h_1(x)}{P_1(x)} + \cdots + \frac{h_s(x)}{P_s(x)}\right].$$

If we perform a PFD of any individual expression $h_i(x)/P_i(x)$, we would, again, simply regenerate ℓ_i terms of our original PFD. Because of how we combined the terms originally, the terms we regenerate will either have all positive numerators (if $i \leq t$) or all negative numerators (if $i > t$). Either way, by Lemma 2.2 each $h_i(x)$ has $\ell_i - 1$ real roots that strictly interlace those of $P_i(x)$.

We can again write each $h_i(x)$ as $\alpha_i \hat{h}_i(x)$, in which α_i is the leading coefficient of $h_i(x)$ and $\hat{h}_i(x)$ is the monic polynomial obtained by dividing $h_i(x)$ by α_i. Again, $\hat{h}_i$ has the same roots as h_i, which strictly interlace those of P_i; by Theorem 3.4 there exists a matrix $A_i \in \mathscr{R}(T_i)$, centered at u_i, such that A_i has characteristic polynomial P_i, $A_i(u_i)$ has characteristic polynomial $\hat{h}_i$, and A_i is diagonally similar to a symmetric matrix. Construct such an A_i for each pair of $\hat{h}_i$ and P_i.

Now, let

$$A = \begin{bmatrix} b & (\alpha_1 e_{1,1})^T & \cdots & (\alpha_s e_{1,s})^T \\ e_{1,1} & A_1 & 0 & 0 \\ \vdots & 0 & \ddots & \vdots \\ e_{1,s} & 0 & \cdots & A_s \end{bmatrix}.$$

Note that by construction, we have $A \in \mathscr{R}(T)$ and it is centered at v.

By the neighbors formula, we have:

$$\begin{aligned} p_A &= (x-b)\prod_{i=1}^{s} p_{A_i} - \sum_{i=1}^{s} a_i\, p_{A_i(u_i)} \prod_{j=1,\, j\neq i}^{s} p_{A_j} \\ &= (x-b)P_1(x)\cdots P_s(x) - \sum_{i=1}^{s} \alpha_i \hat{h}_i(x) \prod_{j=1,\, j\neq i}^{s} P_j(x) \\ &= (x-b)g(x) - \sum_{i=1}^{s} \alpha_i \hat{h}_i(x) \prod_{j=1,\, j\neq i}^{s} P_j(x) = p(x). \end{aligned}$$

Since each A_i is diagonally similar to a symmetric matrix, A is pseudosymmetric, as claimed. □

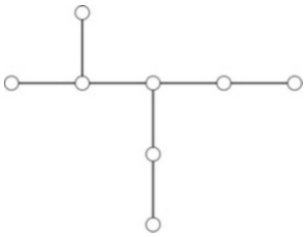
4 Constructive Examples

We present two examples that demonstrate the use of this algorithm. The first example uses a polynomial with real, distinct roots, which can be realized using the method presented in Theorem 3.4. The second example uses a polynomial with repeated and complex roots, which can be realized using the method presented in Theorem 3.5.

Example 4.1 Let p_1 be the polynomial

$$p_1(x) = (x+3)(x+2)(x+1)(x)(x-1)(x-2)(x-3)(x-4),$$

and let T be the following 8-vertex tree:



We wish to find a matrix $A \in \mathscr{R}(T)$ with characteristic polynomial p_1.

(Note: in the following calculations, all real numbers are rounded to the nearest thousandth.)

In order to use the algorithm presented in Theorem 3.4, we first need to choose a "central" vertex for our matrix A and a characteristic polynomial for $A(v)$.

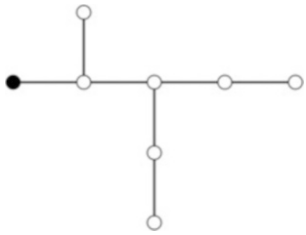


Fig. 1 Our tree T with highlighted vertex v

Since there are no restrictions on the central vertex in the real, distinct root case, let v be the far left vertex indicated in Fig. 1.

We need our characteristic polynomial for $A(v)$ to be of degree 7 and have real, distinct roots that strictly interlace those of p_1. Any such polynomial will suffice, so let us choose

$$g(x) = (x + 2.5)(x + 1.5)(x + 0.5)(x - 0.5)(x - 1.5)(x - 2.5)(x - 3.5).$$

Now, by Lemma 2.3 we can write

$$\begin{aligned}\frac{p_1}{g} =&(x - 0.5) - \left(\frac{0.733}{x + 2.5} + \frac{1.015}{x + 1.5} + \frac{1.154}{x + 0.5}\right) \\ &- \left(\frac{1.196}{x - 0.5} + \frac{1.154}{x - 1.5} + \frac{1.015}{x - 2.5} + \frac{0.733}{x - 3.5}\right).\end{aligned}$$

Observe that all the numerators in the PFD are positive, as the algorithm claims.

In this case T has only one branch at v - call it T_1 - so we may combine all partial fraction terms into a single rational expression:

$$\begin{aligned}\frac{p_1}{g} &= (x - 0.5) \\ &- \frac{7x^6 - 21x^5 - 48.125x^4 + 131.25x^3 + 60.183x^2 - 129.938x - 6.152}{g}.\end{aligned}$$

By pulling out the leading coefficient of the PFD numerator, and factoring the resulting monic polynomial (call it $\hat{h}_1$), we have:

$$\begin{aligned}\frac{p_1}{g} &= (x - 0.5) \\ &- \frac{7(x + 2.261)(x + 1.144)(x + 0.046)(x - 1.046)(x - 2.144)(x - 3.261)}{(x + 2.5)(x + 1.5)(x + 0.5)(x - 0.5)(x - 1.5)(x - 2.5)(x - 3.5)}.\end{aligned}$$

Observe that $\hat{h}_1$ is of degree 6 and has real, distinct roots that strictly interlace the roots of g. Letting u denote the sole neighboring vertex of v, we can construct a

matrix A_1 in $\mathscr{R}(T_1)$, centered at u, such that A_1 has characteristic polynomial g and $A_1(u)$ has characteristic polynomial $\hat{h}_1$. Our desired matrix A will be of the form:

$$A = \begin{bmatrix} 0.5 & (7e_1)^T \\ e_1 & A_1 \end{bmatrix}.$$

One can verify with the neighbors formula (1) that A has characteristic polynomial p_1.

Now, we can follow the same process to construct A_1. Dividing g by $\hat{h}_1$ as in Lemma 2.3 yields:

$$\begin{aligned}\frac{g}{\hat{h}_1} =& (x - 0.5) - \left(\frac{0.458}{x + 2.261} + \frac{0.588}{x + 1.144} + \frac{0.641}{x + 0.046}\right) \\ & - \left(\frac{0.641}{x - 1.046} + \frac{0.588}{x - 2.144} + \frac{0.458}{x - 3.261}\right).\end{aligned}$$

Now, T_1 has two branches at u, with weights 5 and 1; call these branches S_1 and S_2. We would like to combine the new PFD into two rational expressions such that the denominators have degrees 5 and 1. There are several ways we could do this, but let's consolidate the first five terms of the PFD and leave the last term alone.

$$\begin{aligned}\frac{g}{\hat{h}_1} =& (x - 0.5) \\ & - \frac{2.917(x + 1.941)(x + 0.687)(x - 0.535)(x - 1.767)}{(x + 2.261)(x + 1.144)(x + 0.046)(x - 1.046)(x - 2.144)} \\ & - \frac{0.458}{x - 3.261}.\end{aligned}$$

Let w_1 and w_2 denote the neighbors of v contained in S_1 and S_2 respectively. Let $\frac{2.917\hat{h}_2}{P_2}$ denote the first rational expression above. Again, since P_2 and $\hat{h}_2$ have strictly interlacing roots, we can construct a matrix B in $\mathscr{R}(S_1)$ centered at w_1 such that B has characteristic polynomial P_2 and $B(w_1)$ has characteristic polynomial $\hat{h}_2$. As for the last rational expression, $\frac{0.458}{x-3.261}$, the matrix $[3.621] \in \mathscr{R}(S_2)$ has characteristic polynomial $x - 3.261$, so we have reached the base case of our inductive algorithm.

Thus, our matrix A_1 will be of the form:

$$A_1 = \begin{bmatrix} 0.5 & (2.917e_1)^T & 0.458 \\ e_1 & B & 0 \\ 1 & 0 & 3.261 \end{bmatrix}.$$

By continuing this process, we can calculate

$$B = \begin{bmatrix} 0.066 & 1.027 & 0 & 1.204 & 0 \\ 1 & -0.082 & 3.438 & 0 & 0 \\ 0 & 1 & -0.093 & 0 & 0 \\ 1 & 0 & 0 & -0.076 & 0.373 \\ 0 & 0 & 0 & 1 & -0.077 \end{bmatrix},$$

which, by performing the appropriate substitutions, yields

$$A = \begin{bmatrix} 0.5 & 7 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0.5 & 2.917 & 0 & 0 & 0 & 0 & 0.458 \\ 0 & 1 & 0.066 & 1.027 & 0 & 1.204 & 0 & 0 \\ 0 & 0 & 1 & -0.082 & 3.438 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -0.093 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & -0.076 & 0.373 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -0.077 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 3.261 \end{bmatrix}.$$

While we will not perform the calculation here, one can use a similar inductive process to confirm that A is diagonally similar to a symmetric matrix, since all off-diagonal entries are positive.

Example 4.2 Let T be the same 8-vertex tree used in the previous example, and let p_2 be the following polynomial:

$$p_2(x) = x^2(x-5)^3(x+1)(x^2+1).$$

We wish to find a matrix $A \in \mathscr{R}(T)$ with characteristic polynomial p_2.

(In the following calculations, real numbers greater than 10 in absolute value are rounded to the nearest tenth. Numbers with magnitude between 1 and 10 are rounded to the nearest 100th, and numbers with magnitude less than 1 are rounded to the nearest 1000th.)

The polynomial p_2 has root deficiency 6, so let $\ell = (6-2)/2 = 2$. We wish to find a vertex v of T such that some partial sum of the branch weights of T at v lies in the range $[2, 5]$. Fortunately, the highlighted vertex v indicated in Fig. 2 suffices, since it has a branch of weight 2.

We will center our matrix A at v. Now, we need to construct an appropriate degree 7 polynomial g to be the characteristic polynomial of $A(v)$.

If $\gamma_1 > \gamma_2 > \cdots > \gamma_7$ are the roots of g, and 5 and -1 the signed roots of p_2, we would like to choose $\gamma_1, \gamma_2, \ldots, \gamma_5$ to be greater than 5, γ_6 to be between 5 and -1, and γ_7 to be less than -1. This will ensure that the PFD of $\frac{p_2}{g}$ will produce five

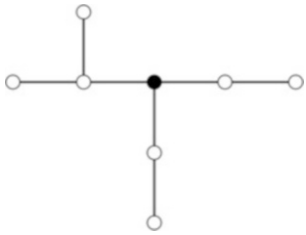


Fig. 2 Our tree T with highlighted vertex v

negative and two positive numerators, as shown in the proof of Proposition 3.3. So let

$$g(x) = (x-8)(x-7.5)(x-7)(x-6.5)(x-6)(x-1)(x+2).$$

Now we can write

$$\begin{aligned}\frac{p_2}{g} =&(x+20) - \left(\frac{-9627.4}{x-8} + \frac{1847.0}{x-7.5} + \frac{-11614.8}{x-7} + \frac{2638.4}{x-6.5}\right)\\ &- \left(\frac{-155.4}{x-6} + \frac{-0.011}{x-1} + \frac{-0.039}{x+2}\right).\end{aligned}$$

We see that T has three branches at v, which have weights 3, 2, and 2. Call these branches T_1, T_2, and T_3 respectively, and suppose they each contain the respective neighbors u_1, u_2, and u_3 of v.

We would like to consolidate the PFD terms into three rational expressions such that the denominators have degrees 3, 2, and 2. However, in order to apply our algorithm, we must only consolidate terms whose numerators have the same sign. So let's combine the first, third, and fifth terms into one expression, the second and fourth terms into another, and the last two terms into our final expression.

$$\begin{aligned}\frac{p_2}{g} =&(x+20) - \left(\frac{-21397.6(x-7.55)(x-6.01)}{(x-8)(x-7)(x-6)} + \frac{21108.4(x-6.63)}{(x-7.5)(x-6.5)}\right)\\ &- \left(\frac{-0.051(x+1.33)}{(x-1)(x+2)}\right)\\ =&(x+20) - \left(\frac{-21397.6\hat{h}_1}{P_1} + \frac{21108.4\hat{h}_2}{P_2} + \frac{-0.051\hat{h}_3}{P_3}\right).\end{aligned}$$

Since each $\hat{h}_i$ and P_i have strictly interlacing roots, we can construct a matrix $A_i \in \mathscr{R}(T_i)$, centered at u_i, such that A_i has characteristic polynomial P_i, $A_i(u_i)$

has characteristic polynomial $\hat{h}_i$, and A_i is diagonally similar to a symmetric matrix. Our desired matrix A will have the form

$$A_1 = \begin{bmatrix} -20 & (-21397.6e_1)^T & (21108.4e_1)^T & (-0.051e_1)^T \\ e_1 & A_1 & 0 & 0 \\ e_1 & 0 & A_2 & 0 \\ e_1 & 0 & 0 & A_3 \end{bmatrix}.$$

One can use the neighbors formula to verify that A does indeed have characteristic polynomial p_2.

Finally, we can construct A_1, A_2, and A_3 using the algorithm in Theorem 3.4 to yield the matrix

$$A = \begin{bmatrix} -20 & -21397.6 & 0 & 0 & 21108.4 & 0 & -0.051 & 0 \\ 1 & 7.44 & 0.249 & 0.012 & 0 & 0 & 0 & 0 \\ 0 & 1 & 7.55 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 6.01 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 7.38 & 0.109 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 6.63 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0.327 & 1.57 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1.33 \end{bmatrix}.$$

5 Corollaries, Applications, and Special Cases

To help narrow the scope of Theorem 3.5, and to pave the way for useful corollaries, we present a few entirely graph-theoretic results. These theorems will aid us in choosing the most appropriate vertex of the graph on which to center our constructed matrix.

Unless otherwise specified, let T be an arbitrary tree on n vertices. We let $k = n - 1$, ℓ be a nonnegative integer with $\ell < k/2$, and $m = k - 2\ell + 1$, the number of integers in the range $[\ell, k - \ell]$. Observe that since $\ell < k/2$, we must have $m \geq 2$.

With respect to ℓ, we say that a vertex v of T is *ideal* if the largest branch of T at v has weight $k - \ell$ or smaller.

If T is our tree and p is our desired polynomial with root deficiency $d \geq 2$, then the only vertices upon which we could center our constructed matrix A are the ideal vertices of T with respect to $\ell = (d - 2)/2$. If a vertex v is not ideal - that is, if T has a branch at v with weight greater than $k - \ell$ - then any partial sum of the branch weights of T at v will either be smaller than ℓ or larger than $k - \ell$.

Proposition 5.1 *T has at least one ideal vertex with respect to ℓ.*

Proof Let v be any vertex of T, and let M be the weight of the largest branch of T at v. We proceed by induction on M.

If $M \leq k - \ell$, then v is an ideal vertex, and we are done. So suppose $M > k - \ell$, and that the claim holds for all maximum branch weights c with $M > c \geq k - \ell$. Let $T_1, T_2, \ldots, T_s$ be the branches of T at v, in which each T_i has weight ℓ_i and contains u_i, a neighboring vertex of v. Without loss of generality, assume that $\ell_1 = M$. We must have $\ell_1 + \ell_2 + \cdots + \ell_s = k$, so if $\ell_1 > k - \ell$, then $\ell_2 + \cdots + \ell_s < \ell$. Now let $S_1, S_2, \ldots, S_t$ be the branches of T at u_1, and assume S_1 is the branch that contains v. By the previous statement, S_1 can have weight at most ℓ. Furthermore, the weights of $S_2, \ldots, S_t$ sum to $M - 1$, so no individual S_i can have weight greater than $M - 1$. This includes S_1; since $\ell < k/2$, we have $\ell < k - \ell < M$. The choice of our starting vertex was arbitrary, so by our inductive hypothesis, T has an ideal vertex. □

Note that the proof of the above theorem not only confirms the existence of an ideal vertex for any tree T, but also provides guidance for how to locate it. We also have the following useful lemma regarding the positions of the ideal vertices:

Lemma 5.2 *The ideal vertices of T with respect to ℓ, along with the edges connecting them, form a subtree of T.*

Proof The case when T only has one ideal vertex is trivial, so we assume that v and w are two distinct ideal vertices of T. Since T is a tree, there exists a unique path on T connecting v and w; it suffices to prove that every vertex along this path is ideal.

If v and w are the only vertices on this path (i.e. if v and w are neighbors), then the case is again trivial. So assume that the path contains a vertex u distinct from v and w. Let T_v and T_w be the (necessarily distinct) branches of T at u containing the vertices v and w respectively, and let $T_1, \ldots, T_s$ be the remaining branches of T at u. The vertex u is contained in some branch S with respect to v, and since v is ideal, S has weight $k - \ell$ or smaller. Therefore, the branches of T at u that do not contain v can have weights that sum to at most $k - \ell - 1$, implying that T_w and $T_1, \ldots, T_s$ all have weight less than $k - \ell$. However, since w is also ideal, we can apply a symmetric argument and conclude that T_v has weight less than $k - \ell$ as well. Since all branches of T at u have weight less than $k - \ell$, u must be ideal (Fig. 3). □

Now suppose that, for a given ℓ, we have an ideal vertex v for a tree T. If any branch of T at v has its weight in the interval $[\ell, k - \ell]$, then by Theorem 3.5 we can construct a pseudosymmetric matrix A for any polynomial of root deficiency $d = 2\ell + 2$. We need only be concerned if all branches of T at v have weight less than ℓ. In that case, we call v *deficient* with respect to ℓ. Not all trees have a deficient vertex. In fact, if a tree has more than one ideal vertex, then we can say with certainty that none of them are deficient.

Proposition 5.3 *If T has a deficient vertex v with respect to ℓ, then v is the only ideal vertex of T with respect to ℓ.*

Proof Let v be a deficient vertex of T, and let $u_1, u_2, \ldots, u_s$ be the neighboring vertices of v, in which u_i is contained in the branch T_i with weight ℓ_i. By Lemma 5.2, the ideal vertices of T form a subtree, so if T contained an additional ideal vertex, then at least one u_i would have to be ideal. However, by assumption

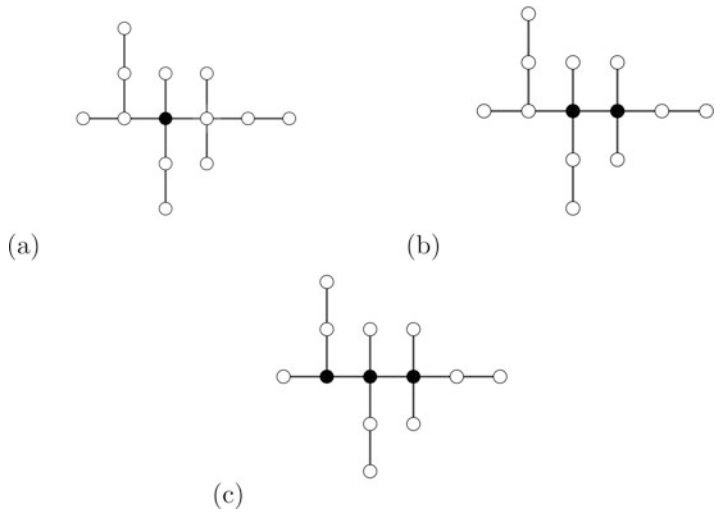


Fig. 3 Images (**a**), (**b**), and (**c**) denote the ideal vertices of this 13-vertex tree with respect to 5, 4, and 3 respectively

$\ell_i < \ell$ for each i, which also means that $\sum_{j=1,\, j\neq i}^{s} \ell_j > k - \ell$ for each i. Thus for any u_i, the branch containing v would have weight greater than $k - \ell$, so u_i cannot be ideal. □

Therefore, if our desired polynomial p has root deficiency d, and our tree T has a deficient vertex v with respect to $\ell = (d-2)/2$, then any pseudosymmetric matrix A would have to be centered at v. Fortunately, even in this case, we can still usually come up with an appropriate partial sum. We will present some sufficient conditions for that later.

We will conclude this graph theory discussion with a theorem that places an upper bound on the number of ideal vertices a tree can have with respect to a given ℓ.

This may not be the most useful result from a theoretical perspective—we only need to find one appropriate ideal vertex to know that a matrix can be constructed—but from a computational standpoint certain vertices may be better for centering than others. An upper bound—along with the fact that all ideal vertices must be connected—allows the algorithm user to know when he has found all possible ideal vertices, allowing him to make the best judgment on which one to use.

First, we introduce one more term and lemma. We call an ideal vertex v of T *ideal pendent* (with respect to ℓ) if it is a pendent vertex on the subtree of ideal vertices with respect to ℓ. It is well known that a tree on two or more vertices must have two or more of those vertices pendent. So as long as T has more than one ideal vertex, we are guaranteed at least two pendent ideal vertices.

Lemma 5.4 *Suppose v is an ideal pendent vertex of T, and let $T_1, \ldots, T_s$ be the branches of T at v with respective lengths $\ell_1, \ldots, \ell_s$. Without loss of generality, assume that T_1 is the only branch of T at v that contains an ideal vertex. Then $\ell_2 + \ell_3 + \cdots + \ell_s \geq \ell$.*

Proof If $\ell_2 + \ell_3 + \cdots + \ell_s < \ell$, then $\ell_1 > k - \ell$ and v would not be an ideal vertex. □

Now, we present and prove our upper-bound result:

Proposition 5.5 *T has at most $m = k - 2\ell + 1$ ideal vertices with respect to ℓ.*

Proof We proceed by induction on ℓ. If $\ell = 0$, then $m = n$, and the claim holds.

Now assume $\ell > 0$, and that the claim holds for s with $\ell > s \geq 0$. Suppose T has α ideal vertices with respect to $\ell - 1$ and β ideal vertices with respect to ℓ. Note that $\beta \leq \alpha$; any vertex that is ideal with respect to ℓ will also be ideal with respect to $\ell - 1$. By our inductive hypothesis, $\alpha \leq k - 2(\ell - 1) + 1 = m + 2$. If $\alpha \leq m$, then the claim holds automatically. So we only need to consider the cases when $\alpha = m + 1$ or $\alpha = m + 2$. Note that in either case, $\alpha > 2$ since $m \geq 2$.

In the case where $\alpha = m + 1$, we would have $n - (m + 1) = 2\ell - 1$ non-ideal vertices with respect to $\ell - 1$. However, recall that at least two of our α ideal vertices are ideal pendent, and by Lemma 5.4 they each have branches consisting entirely of non-ideal vertices whose weights sum to $\ell - 1$ or greater. (These branches are distinct since our graph is minimally connected.) If both of these sums were greater than $\ell - 1$, then the total number of non-ideal vertices would be 2ℓ or greater, a contradiction. Therefore, at least one of our ideal pendent vertices has its "non-ideal" branch weights summing to exactly $\ell - 1$, implying that the branch that *does* contain ideal vertices has weight $k - \ell + 1$. Therefore, when we change our parameter from $\ell - 1$ to ℓ, this particular vertex will no longer be ideal. So $\beta \leq \alpha - 1 = m$, as desired.

The case where $\alpha = m + 2$ is similar. In that case, both of our pendent ideal vertices with respect to $\ell - 1$ lose their ideal status upon changing the parameter to ℓ. We still find that $\beta \leq m$, as desired (Fig. 4). □

We can now combine our main theorem with the previous results on ideal vertices to draw some useful corollaries, targeting specific types of trees or polynomials that can be realized with our algorithm.

For the remainder of this section, we will let T be an arbitrary tree on n vertices, and p be an arbitrary polynomial of degree n and root deficiency $d \geq 2$. (The case when $d = 0$ merits no further study due to Theorem 3.4). We let $\ell = (d - 2)/2$, $k = n - 1$, and $m = k - 2\ell + 1$, the number of integers in the range $[\ell, k - \ell]$.

Fig. 4 With respect to $\ell = 4$, this 10-vertex tree has exactly $m = 2$ ideal vertices

Furthermore, observe that since d can be at most n, we always have $(d-2)/2 = \ell < (n-1)/2 = k/2$.

The first result we already mentioned, but we state it again formally.

Theorem 5.6 *If T has an ideal vertex v with respect to ℓ that is not deficient, then p has a pseudosymmetric realization $A \in \mathscr{R}(T)$.*

Proof By the definition of ideal vertex, T can have no branches at v with weight larger than $k-\ell$. Since v is not deficient, T must have at least one branch at v with weight ℓ or greater. The result follows from Theorem 3.5. □

As a natural follow-up to this corollary, we have:

Corollary 5.7 *If T has more than one ideal vertex with respect to ℓ, then p has a pseudosymmetric realization $A \in \mathscr{R}(T)$.*

Proof By Proposition 5.3, if T has more than one ideal vertex, then none of them can be deficient. See Theorem 5.6. □

Now the only case of interest is when T has a single, deficient ideal vertex v with respect to the appropriate ℓ. The next result covers a large number of cases:

Proposition 5.8 *Suppose T has a deficient ideal vertex v with respect to ℓ. The polynomial p has a pseudosymmetric realization $A \in \mathscr{R}(T)$ if at most two branches of T at v have weight greater than m.*

Proof Let $T_1, \ldots, T_s$ be the branches of T at v with respective weights $\ell_1, \ldots, \ell_s$. By definition of deficient vertex, each $\ell_i < \ell$. Without loss of generality, assume that $\ell_2, \ell_3, \ldots, \ell_{s-1}$ are all less than or equal to m. Let k_i denote the partial sum $\ell_1 + \ell_2 + \cdots + \ell_i$; clearly $k_1 < k_2 < \cdots < k_s$. Note that $k_1 = \ell_1 < \ell$, but $k_{s-1} = k - \ell_s > k - \ell$.

If there existed no $k_i \in [\ell, k-\ell]$, then for some $i \in [1, s-2]$ we would have $k_{i+1} - k_i > (k-\ell) - \ell + 1 = m$. But $k_{i+1} - k_i = \ell_{i+1}$, and by assumption $\ell_2, \ldots, \ell_{s-1}$ are all less than or equal to m; a contradiction. Therefore, there must exist some partial sum $k_i \in [\ell, k-\ell]$, which by Theorem 3.5 guarantees that p has a pseudosymmetric realization in $\mathscr{R}(T)$. □

This last result paves the way for a series of increasingly specific corollaries. The first one concerns *linear trees*. A linear tree T is a tree whose high-degree vertices all occur along a single induced path. (We define a *high-degree vertex* to be a vertex with degree at least three.) Furthermore, if P is the longest induced path that includes all of these vertices, then the *depth* of T is the maximum distance (in number of vertices) a vertex can be from this path P.

Proposition 5.9 *Suppose T is a linear tree with depth at most m. Then the polynomial p has a pseudosymmetric realization $A \in \mathscr{R}(T)$.*

Proof Let v be an ideal vertex of T. If v is not deficient, then the result follows from Theorem 5.6. If v is deficient, then it must be high-degree. Since $\ell < k/2$, if T

Fig. 5 A linear tree on 12 vertices with depth 2, with one deficient ideal vertex v with respect to $\ell = 5$. Note that the branch weights 4 and 1 sum to $5 \in [5, 6] = [\ell, k - \ell]$

had two or fewer branches at v, each with weight less than ℓ, then the sum of their weights would be less than k, a contradiction.

Since T is a linear tree, at most two branches of T at v can contain additional high-degree vertices, which means the remaining branches are all paths of length m or smaller. The result follows from Proposition 5.8. □

The previous result yields the following specific case:

Corollary 5.10 *Suppose T is a linear tree with depth at most 2. Then p has a pseudosymmetric realization $A \in \mathscr{R}(T)$.*

Proof Since $\ell < k/2$, we know that ℓ is strictly less than $k-\ell$, so the range $[\ell, k-\ell]$ must contain at least two integers. Therefore, $m = k - 2\ell + 1 \geq 2$. See the previous result (Fig. 5). □

This corollary includes the case of paths, which can be considered linear trees with depth 0. While the existence of arbitrary polynomial realizations for paths had already been proven in the real number case, this is the first proof that uses an explicit, constructive argument.

The next few results also follow from Proposition 5.8.

Proposition 5.11 *If $m \geq \ell - 1$, or, equivalently, if $n \geq (3/2)(d - 2) - 1$, then p has a pseudosymmetric realization $A \in \mathscr{R}(T)$.*

Proof Let v be an ideal vertex of T. If v is not deficient, then the result follows from Theorem 5.6. If v is deficient, then by definition all branches of T at v have weight less than ℓ, which in this case also means they have weight less than or equal to m. The result now follows from Proposition 5.8 (Fig. 6). □

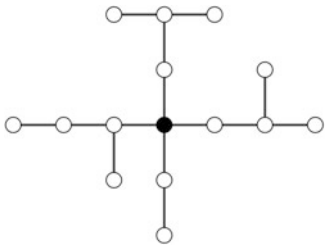


Fig. 6 A tree on 15 vertices, with one deficient ideal vertex v with respect to $\ell = 5$. Note that $m = 14 - 10 + 1 = 5 \geq \ell - 1 = 4$, and we have branch weights 4 and 2 that sum to $6 \in [5, 9] = [\ell, k - \ell]$

As direct consequence of Proposition 5.11, we have:

Corollary 5.12 *Suppose $n < 10$. Then p has a pseudosymmetric realization $A \in \mathscr{R}(T)$.*

Proof By the previous result, p is realizable for any tree if its degree n and root deficiency d satisfy the inequality $n \geq (3/2)(d-2) - 1$. If $n < 10$, then d can be either 0, 2, 4, 6, or 8. In each of these cases, substituting d into the right-hand side of the inequality yields a value less than or equal to d. Since n must be at least d, the inequality is satisfied. □

Note that the inequality $n \geq (3/2)(d-2) - 1$ is equivalent to $d \leq (2/3)(n+1) + 2$. This form is perhaps more useful, as it allows one to see the minimum number of signed roots needed for a degree n polynomial in order to guarantee a pseudosymmetric realization for every tree.

The last corollary we will state does not directly follow from any of the previous ones, but is an interesting observation nonetheless.

Corollary 5.13 *Suppose n is even, and T is composed of two distinct subtrees T_1 and T_2, each with weight $n/2$, that are joined by a single edge. Then p has a pseudosymmetric realization $A \in \mathscr{R}(T)$.*

Proof Let v be one of the two vertices on the edge that connects T_1 and T_2. Then T will have one branch at v with weight $n/2$ (equal to either T_1 or T_2), and additional branches whose weights sum to $n/2 - 1$. The root deficiency d of p can be at most n; therefore, $\ell = (d-2)/2$ will be at most $(n-2)/2 = n/2 - 1$, and $k - \ell$ will be at least $n/2$. Therefore, the branch partial sum $n/2 - 1$ is guaranteed to fall in the range $[\ell, k-\ell]$, and the result follows from Theorem 3.5 (Fig. 7). □

We emphasize that none of these corollaries are "if and only if" statements; a polynomial and tree pair may fall under none of the categories described by the corollaries yet still have a pseudosymmetric realization. Consider the polynomial $p(x) = x^{14}$ and the generalized star S shown in Fig. 8.

Since p has root deficiency 14, we have $\ell = 6$, $[\ell, k-\ell] = [6, 7]$, and $m = 2$. Here, S has a single deficient ideal vertex v, and T has four branches at v with weights 5, 4, 3, and 1. We can realize p with a pseudosymmetric matrix $A \in \mathscr{R}(S)$ because we have $5 + 1 = 6 \in [\ell, k-\ell]$. However, we have $m < \ell - 1$ (violating the

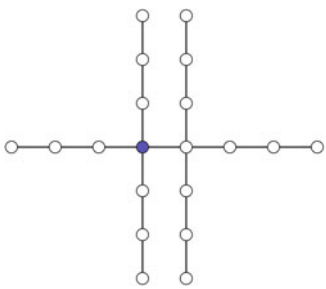


Fig. 7 A 20-vertex tree composed of two 10-vertex subtrees joined by a single edge. With respect to $\ell = 9$, we have an ideal vertex with a branch of length 10, and $10 \in [9, 10] = [\ell, k-\ell]$

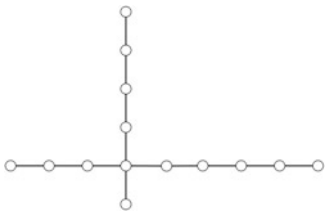


Fig. 8 A generalized star with arm lengths 5, 4, 3, and 1

conditions of Proposition 5.11), S has depth greater than m when viewed as a linear tree (violating the conditions of Proposition 5.9), and we have three branches with weight greater than $m = 2$ (violating the conditions of Proposition 5.8).

Acknowledgement This work was supported by the 2019 National Science Foundation grant DMS #1757603.

References

1. W. Adkins, M. Davidson, Synthetic partial fraction decompositions. Math. Mag. **81**(1), 16–26 (2008)
2. R.S. Cuestas-Santos, C.R. Johnson, Spectra of tridiagonal matrices over a field (2018). Preprint. arXiv:1807.08877
3. S. Friedland, Inverse eigenvalue problems. Linear Algebra Appl. **17**, 15–51 (1977)
4. R.A. Horn, C.R. Johnson, *Matrix Analysis*, 2nd edn. (Cambridge University Press, New York, 2013)
5. C.R. Johnson, A. Leal-Duarte, Complete spectral theory for matrices over a field whose graph is a star. Manuscript
6. C.R. Johnson, C.M. Saiago, *Eigenvalues, Multiplicities, and Graphs* (Cambridge University Press, New York, 2018)
7. A. Leal-Duarte, Construction of acyclic matrices from spectral data. Linear Algebra Appl. **113**, 173–182 (1989)

A Note on the Fredholm Theory of Singular Integral Operators with Cauchy and Mellin Kernels, II

Peter Junghanns and Robert Kaiser

Abstract Necessary and sufficient conditions for the Fredholmness of a class of singular integral operators in weighted $\mathbf{L}^p$-spaces on the interval $(0, 1)$ of the real line are formulated under weaker conditions than in Junghanns and Kaiser (Oper Theory Adv Appl 271:291–325, 2018). Moreover, results on the one-sided invertibility of the operators under consideration are proved.

Keywords Singular integral operators · Cauchy kernel · Mellin kernel · Fredholm theory · One-sided invertibility

Mathematics Subject Classification (2010) Primary 45E05; Secondary 45E10

1 Introduction

In this paper we consider linear integral operators, which are made up by multiplication operators, the Cauchy singular integral operator and Mellin type operators. More precisely, the operators under consideration are given by

$$\begin{aligned}(\mathcal{A}u)(x) :=& a(x)u(x) + \frac{b(x)}{\pi \mathbf{i}} \int_0^1 \frac{u(y)\,dy}{y-x} + c_+(x) \int_0^1 k_+\left(\frac{x}{y}\right) \frac{u(y)\,dy}{y} \\ &+ c_-(x) \int_0^1 k_-\left(\frac{1-x}{1-y}\right) \frac{u(y)\,dy}{1-y}, \quad 0 < x < 1,\end{aligned} \tag{1.1}$$

P. Junghanns (✉)
Chemnitz University of Technology, Faculty of Mathematics, Chemnitz, Germany
e-mail: peter.junghanns@mathematik.tu-chemnitz.de

R. Kaiser
TU Bergakademie Freiberg, Faculty of Mathematics and Computer Science, Freiberg, Germany
e-mail: xenim2001@web.de

M. A. Bastos et al. (eds.), *Operator Theory, Functional Analysis and Applications*,
Operator Theory: Advances and Applications 282,
https://doi.org/10.1007/978-3-030-51945-2_18

and are considered in weighted $\mathbf{L}^p$-spaces $\mathbf{L}^p_{\rho,\sigma} := \mathbf{L}^p_{v^{\rho,\sigma}}(0,1)$, where $a, b, c_\pm : [0,1] \to \mathbb{C}$ are piecewise continuous functions,

$$v^{\rho,\sigma}(x) = x^\rho (1+x)^\sigma , \quad x \in (0,1) , \quad \rho, \sigma > -1 ,$$

is a classical Jacobi weight, and the Mellin transforms of the functions $k_\pm : (0,\infty) \to \mathbb{C}$ are supposed to be continuous.

The paper continues the investigations in [5] and extends the results given there by using techniques from [2] and properties of Fourier convolution operators in weighted $\mathbf{L}^p$-spaces proved in [1] and [8]. To be more precise, in [5] the main theorem formulates necessary and sufficient conditions for the Fredholmness of the operator $\mathcal{A}$ as well as an index formula based on the winding number of a certain closed curve in $\mathbb{C}$ associated with $\mathcal{A}$. In the present paper we will state and prove this theorem again but under weaker conditions regarding the functions $k_\pm$. In order to do this, we basically use [1, Theorem 5.7] (cf. also [8, Theorem 1.3 and Theorem 3.2]).

Additionally to the results proved in [5], we investigate the regularity of solutions of equation $\mathcal{A}u = f$ and prove results concerning the one-sided invertibility of the operator $\mathcal{A}$ in $\mathbf{L}^p_{\rho,\sigma}$.

The paper is structured as follows. In Sect. 2, we define weighted $\mathbf{L}^p$-spaces on the real line and on the half line, collect useful properties of the Fourier transformation, and introduce specific classes of convolution operators. In the following Sect. 3 we are going to state results regarding the boundedness of operator $\mathcal{A}$ in the weighted $\mathbf{L}^p$-spaces $\mathbf{L}^p_{\rho,\sigma}$. To achieve this, we introduce the Mellin transformation at first and recall useful results from [5]. Section 4 starts with a result, which connects operators of form (1.1) and convolution operators introduced in Sect. 2, followed by the above mentioned Fredholm theorem, first in case $a, b, c_\pm$ are constants and later for piecewise continuous functions $a, b, c_\pm$. In Sect. 5, we investigate the smoothness of solutions of equation $\mathcal{A}u = f$ as well as their asymptotic behaviour near the boundary point 1. Here beside the case $c_-(x) \equiv 0$, we distinguish between the case of $b(x)$ being the zero function, that means, that the Cauchy singular integral operator does not appear, and the case of both functions $c_\pm(x)$ being identically zero, i.e., the Mellin-type operators $\mathcal{B}_{k_\pm}$ do not occur. Results and techniques from [2, Section 6] will be used and extended. The final Sect. 6 deals with the one-sided invertibility of $\mathcal{A}$ in $\mathbf{L}^p_{\rho,\sigma}$ in case c_+ or c_- vanish.

2 Preliminaries

In this section we first collect some properties of Fourier convolution operators. With that, we are going to introduce a specific class of convolution operators.

Subsequently, results concerning the one-sided invertibility of such operators are formulated.

For $1 < p < \infty$ and $-1 < \sigma < p - 1$, let us introduce the weighted $\mathbf{L}^p$-space $\mathbf{L}^p_\sigma(\mathbb{R})$ defined by the norm

$$\|u\|_{\sigma,p,\mathbb{R}} = \left(\int_{-\infty}^{\infty} |u(t)|^p \frac{|t|^\sigma}{(1+|t|)^\sigma}\, dt \right)^{\frac{1}{p}}. \tag{2.1}$$

We are going to use the abbreviations $\mathbf{L}^p(\mathbb{R}) := \mathbf{L}^p_0(\mathbb{R})$ and $\|u\|_{p,\mathbb{R}} := \|u\|_{0,p,\mathbb{R}}$.

For $u \in \mathbf{L}^1(\mathbb{R})$, the well-known Fourier transform is defined by

$$(\mathcal{F}u)(\eta) := \int_{-\infty}^{\infty} e^{-\mathbf{i}\eta t} u(t)\, dt, \quad \eta \in \mathbb{R}.$$

For $0 < R < \infty$ and $u \in \mathbf{L}^1(-R, R)$, set

$$(\mathcal{F}_R u)(\eta) = \int_{-R}^{R} e^{-\mathbf{i}\eta t} u(t)\, dt, \quad \eta \in \mathbb{R}, \tag{2.2}$$

and

$$\left(\mathcal{F}_R^- u\right)(t) = \frac{1}{2\pi} \int_{-R}^{R} e^{\mathbf{i}\eta t} u(\eta)\, d\eta, \quad t \in \mathbb{R}.$$

Let $1 < p \le 2$, $\frac{1}{p} + \frac{1}{q} = 1$ and $u \in \mathbf{L}^p(\mathbb{R})$. Then (cf. [9, p. 96, Theor. 74]) $\mathcal{F}_R u$ converges for $R \to \infty$ in $\mathbf{L}^q(\mathbb{R})$ to a function $\widetilde{u}$ and $\mathcal{F}_R^- \widetilde{u}$ converges in $\mathbf{L}^p(\mathbb{R})$ to u. Moreover, we have

$$\|\widetilde{u}\|_{q,\mathbb{R}} \le (2\pi)^{\frac{1}{q}} \|u\|_{p,\mathbb{R}}.$$

Obviously, if $u \in \mathbf{L}^1(\mathbb{R}) \cap \mathbf{L}^p(\mathbb{R})$, then, due to (2.2), $(\mathcal{F}u)(\eta) = \widetilde{u}(\eta)$ for almost every $\eta \in \mathbb{R}$. That is why, we will call the operator

$$\mathcal{F} : \mathbf{L}^p(\mathbb{R}) \to \mathbf{L}^q(\mathbb{R}), \quad u \mapsto \widetilde{u}$$

the Fourier transformation and the function $\mathcal{F}u$ the Fourier transform of $u \in \mathbf{L}^p(\mathbb{R})$. Of course, $\mathcal{F}^- : \mathbf{L}^p(\mathbb{R}) \to \mathbf{L}^q(\mathbb{R})$ is also well defined, where for $u \in \mathbf{L}^p(\mathbb{R})$, we set $\mathcal{F}^- u = \lim_{R\to\infty} \mathcal{F}_R^- u$ and take the limit in the $\mathbf{L}^q(\mathbb{R})$-sense. Note that, $\mathcal{F}v$ and $\mathcal{F}^- v$ are also defined for $v = \mathcal{F}^- u$ and $v = \mathcal{F}u$, respectively, where $\mathcal{F}v := \lim_{R\to\infty} \mathcal{F}_R v$ as well as $\mathcal{F}^- v := \lim_{R\to\infty} \mathcal{F}_R^- v$ and the limits are taken in the $\mathbf{L}^p(\mathbb{R})$-sense. Moreover, with these definitions we have

$$\mathcal{F}^- \mathcal{F} u = \mathcal{F}\mathcal{F}^- u = u \quad \forall\, u \in \mathbf{L}^p(\mathbb{R}),\ 1 < p \le 2.$$

In case $p = 2$, it's well-known that $\mathcal{F} : \mathbf{L}^2(\mathbb{R}) \to \mathbf{L}^2(\mathbb{R})$ is an isomorphism, where

$$\frac{1}{2\pi}\mathcal{F}^* = \mathcal{F}^{-1} = \mathcal{F}^-. \tag{2.3}$$

Let $f, g : \mathbb{R} \to \mathbb{C}$ be measurable functions. We assume that there is a set $N \subset \mathbb{R}$ of measure zero such that $f(t - \cdot)g(\cdot)$ is integrable for all $t \in \mathbb{R}\backslash N$. In this case, we define the convolution $f * g : \mathbb{R} \to \mathbb{C}$ of these functions by

$$(f * g)(t) := \begin{cases} \displaystyle\int_{-\infty}^{\infty} f(t-s)g(s)\,ds & : t \in \mathbb{R}\backslash N \\ 0 & : t \in N. \end{cases}$$

If $p \in (1, \infty)$, $\frac{1}{p} + \frac{1}{q} = 1$, $f \in \mathbf{L}^p(\mathbb{R})$, and $g \in \mathbf{L}^q(\mathbb{R})$, then

$$\left|(f * g)(t)\right| \le \|f\|_{p,\mathbb{R}}\, \|g\|_{q,\mathbb{R}}\,, \quad t \in \mathbb{R}.$$

This relation is generalized by the well known Young's inequality for convolutions

$$\|f * g\|_{r,\mathbb{R}} \le \|f\|_{p,\mathbb{R}}\, \|g\|_{q,\mathbb{R}}\,, \tag{2.4}$$

which is true for $f \in \mathbf{L}^p(\mathbb{R})$, $g \in \mathbf{L}^q(\mathbb{R})$ with $1 \le p, q \le \infty$ and $\frac{1}{p} + \frac{1}{q} \ge 1$ as well as $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$ $\left(\frac{1}{\infty} := 0\right)$. In particular, under these assumptions, $f * g \in \mathbf{L}^r(\mathbb{R})$. If

$$1 < p, q \le 2 \quad \text{and} \quad \frac{1}{p} + \frac{1}{q} > 1$$

hold true, then we have the convolution theorem

$$\mathcal{F}(f * g) = \mathcal{F}f \cdot \mathcal{F}g \quad \text{and} \quad \mathcal{F}^-(f * g) = 2\pi\, \mathcal{F}^- f \cdot \mathcal{F}^- g$$

as well as

$$f * g = \mathcal{F}^-(\mathcal{F}f \cdot \mathcal{F}g) \quad \text{and} \quad f * g = 2\pi\, \mathcal{F}(\mathcal{F}^- f \cdot \mathcal{F}^- g) \tag{2.5}$$

(cf. [9, p. 106, Theor. 78]).

By $\mathcal{S}_{\mathbb{R}}$ we denote the Cauchy singular integral operator

$$(\mathcal{S}_{\mathbb{R}} f)(s) = \frac{1}{\pi \mathbf{i}} \int_{-\infty}^{\infty} \frac{f(t)}{t - s}\, dt, \quad s \in \mathbb{R},$$

where the integral is considered as a Cauchy principal value integral. It is well known (see, for example, [3, Sections 1.2, 1.5]) that $\mathcal{S}_{\mathbb{R}} : \mathbf{L}^p_\sigma(\mathbb{R}) \to \mathbf{L}^p_\sigma(\mathbb{R})$ is a linear and bounded operator, since $1 < p < \infty$ and $-1 < \sigma < p - 1$.

Let $1 < p \le 2$, $\gamma \in \mathbb{R}$, and $b_\gamma(t) = e^{\mathbf{i}\gamma t}$. Then (cf. [2, Lemma 1.35])

$$b_\gamma^{-1} \mathcal{S}_{\mathbb{R}} b_\gamma u = -\mathcal{F}\mathrm{sgn}(\cdot - \gamma)\mathcal{F}^- u \quad \forall\, u \in \mathbf{L}^p(\mathbb{R}), \tag{2.6}$$

where $\mathrm{sgn} : \mathbb{R} \to \{-1, 0, 1\}$ denotes the sign function.

For $p \in (1, \infty)$ and $\rho \in (-1, p - 1)$, let us introduce the weighted $\mathbf{L}^p$-space $\widetilde{\mathbf{L}}^p_\rho(\mathbb{R}^+)$ defined by the norm

$$\|u\|_{\rho,p,\mathbb{R}^+,\sim} = \left(\int_0^\infty |u(t)|^p \, t^\rho \, dt \right)^{\frac{1}{p}}.$$

By $\mathcal{S}_{\mathbb{R}^+}$ we denote the Cauchy singular integral operator

$$(\mathcal{S}_{\mathbb{R}^+} f)(s) = \frac{1}{\pi \mathbf{i}} \int_0^\infty \frac{f(t)}{t - s} \, dt, \quad s \in \mathbb{R}^+,$$

where again the integral is considered as a Cauchy principal value integral. The operator $\mathcal{S}_{\mathbb{R}^+}$ is a linear and bounded operator in $\widetilde{\mathbf{L}}^p_\rho(\mathbb{R}^+)$ (see, for example, [3, Sections 1.2, 1.5]). For $\xi = \frac{1+\rho}{p}$, we introduce the mapping

$$\widetilde{\mathcal{Z}}_\xi : \widetilde{\mathbf{L}}^p_\rho(\mathbb{R}^+) \to \mathbf{L}^p(\mathbb{R}), \quad f \mapsto e^{-\xi \cdot} f(e^{-\cdot}),$$

which is an isometric isomorphism with $\left(\widetilde{\mathcal{Z}}_\xi^{-1} f\right)(y) = y^{-\xi} f(-\ln y)$. Now, a consequence of relation (2.6) is the formula

$$\widetilde{\mathcal{Z}}_\xi \mathcal{S}_{\mathbb{R}^+} \widetilde{\mathcal{Z}}_\xi^{-1} u = \mathcal{F} a \mathcal{F}^- u \quad \forall u \in \mathbf{L}^p(\mathbb{R}), \tag{2.7}$$

where $a(t) = -\mathbf{i} \cot(\pi\xi - \mathbf{i}\pi t)$, $t \in \mathbb{R}$, and $1 < p \le 2$.

For every $a \in \mathbf{L}^\infty(\mathbb{R})$ the operator

$$\mathcal{W}^0_a : \mathbf{L}^2(\mathbb{R}) \to \mathbf{L}^2(\mathbb{R}), \quad u \mapsto \mathcal{F} a \mathcal{F}^{-1} u$$

is well defined, linear, and bounded (cf. (2.3)). In particular, for $b_\gamma(t) = e^{\mathbf{i}\gamma t}$, $\gamma \in \mathbb{R}$, and $a_\gamma(t) = -\mathrm{sgn}(t - \gamma)$, due to (2.6), we have

$$\mathcal{W}^0_{a_\gamma} = b_\gamma^{-1} \mathcal{S}_{\mathbb{R}} b_\gamma \mathcal{I}. \tag{2.8}$$

Let $p \in (1,\infty)$, $-1 < \sigma < p-1$, and $a \in \mathbf{L}^\infty(\mathbb{R})$. If there is a finite constant $M = M(a,p,\sigma)$ such that

$$\left\| \mathcal{W}_a^0 u \right\|_{\sigma,p,\mathbb{R}} \le M \, \|u\|_{\sigma,p,\mathbb{R}} \quad \forall\, u \in \mathbf{L}^2(\mathbb{R}) \cap \mathbf{L}^p_\sigma(\mathbb{R}),$$

then the uniquely defined linear and continuous extension of $\mathcal{W}_a^0$ onto $\mathbf{L}^p_\sigma(\mathbb{R})$ is again denoted by $\mathcal{W}_a^0 : \mathbf{L}^p_\sigma(\mathbb{R}) \to \mathbf{L}^p_\sigma(\mathbb{R})$ and $a \in \mathbf{L}^\infty(\mathbb{R})$ is called a (p,σ)-multiplier. Let us denote the set of all (p,σ)-multipliers by $\mathbf{M}_{p,\sigma}$. In view of the above example given by (2.8) and the fact that $\mathcal{S}_\mathbb{R} \in \mathcal{L}(\mathbf{L}^p_\sigma(\mathbb{R}))$, we have $\mathcal{W}^0_{a_\gamma} \in \mathcal{L}(\mathbf{L}^p_\sigma(\mathbb{R}))$ for every $\gamma \in \mathbb{R}$. Since, for the characteristic function $\chi_{(a,b]}$ of the interval $(a,b] \subset \mathbb{R}$, we have the representation $\chi_{(a,b]}(t) = \operatorname{sgn}(t-a) - \operatorname{sgn}(t-b)$, the set of all finite linear combinations of characteristic functions of (also unbounded) intervals belongs to $\mathbf{M}_{p,\sigma}$.

As usual, by $\mathbf{V}_1(\mathbb{R})$ we denote the set of all functions $a : \mathbb{R} \to \mathbb{C}$ with bounded total variation and by $\operatorname{var}(a)$ the total variation of $a \in \mathbf{V}_1(\mathbb{R})$. It turns out (see [8, Theorem 1.3]) $\mathbf{V}_1(\mathbb{R}) \subset \mathbf{M}_{p,\sigma}$ and that, for all $a \in \mathbf{V}_1(\mathbb{R})$, the inequality

$$\|a\|_{\mathbf{M}_{p,\sigma}} \le \mathcal{C}_{p,\sigma} \, \|\mathcal{S}_\mathbb{R}\|_{\mathcal{L}\left(\mathbf{L}^p_\sigma(\mathbb{R})\right)} \left[\|a\|_\infty + \operatorname{var}(a) \right] \tag{2.9}$$

is true, where the constant $\mathcal{C}_{p,\sigma}$ does only depend on p and σ and not on $a \in \mathbf{V}_1$.

In what follows, $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ and $\dot{\mathbb{R}}$ denote the two-point compactification and the one-point compactification of $\mathbb{R}$, respectively. By $\mathbf{PC}(\mathbb{R})$ we refer to the algebra of all piecewise continuous functions $a : \mathbb{R} \to \mathbb{C}$ (i.e., $f \in \mathbf{PC}(\mathbb{R})$ if and only if all limits $a(t \pm 0), t \in \dot{\mathbb{R}}$, where $a(\infty \pm 0) := a(\mp\infty) = \lim_{t \to \mp\infty} a(t)$, exist and are finite). Moreover, by $\mathbf{PC}_{p,\sigma}$ we denote the completion of the set $\mathbf{PC}(\mathbb{R}) \cap \mathbf{V}_1(\mathbb{R})$ in the Banach algebra $\left(\mathbf{M}_{p,\sigma}, \|\cdot\|_{\mathbf{M}_{p,\sigma}}\right)$, where $\|a\|_{\mathbf{M}_{p,\sigma}} = \left\|\mathcal{W}_a^0\right\|_{\mathcal{L}(\mathbf{L}^p_\sigma(\mathbb{R}))}$ (cf. [1, p. 254]). Since $\|a\|_\infty = \left\|\mathcal{W}_a^0\right\|_{2,\mathbb{R}}$ (cf.[2, Proposition 2.3]), we have $\mathbf{PC}_{2,0} = (\mathbf{PC}(\mathbb{R}), \|\cdot\|_\infty)$. Note that $\mathbf{PC}_{p,\sigma} \subset \mathbf{PC}(\mathbb{R})$ (cf. [1]).

We introduce the space $\mathbf{L}^p_\sigma(\mathbb{R}^+)$ as the compression of $\mathbf{L}^p_\sigma(\mathbb{R})$ to the positive half line $\mathbb{R}^+$. Thus, $\mathbf{L}^p_\sigma(\mathbb{R}^+)$ may be identified with the image of the projection $\mathcal{P} : \mathbf{L}^p_\sigma(\mathbb{R}) \to \mathbf{L}^p_\sigma(\mathbb{R})$, which is defined by

$$(\mathcal{P}u)(t) = \begin{cases} u(t) : & t \in \mathbb{R}^+, \\ 0 & : t \in \mathbb{R} \setminus \mathbb{R}^+. \end{cases}$$

Due to (2.1), the norm in $\mathbf{L}^p_\sigma(\mathbb{R}^+)$ is given by

$$\|u\|_{\sigma,p,\mathbb{R}^+} = \left(\int_0^\infty |u(t)|^p \frac{t^\sigma}{(1+t)^\sigma} \, dt \right)^{\frac{1}{p}},$$

where we will use the abbreviation $\|u\|_{p,\mathbb{R}^+} := \|u\|_{0,p,\mathbb{R}^+}$. It is easily seen that

$$\|u\|_{p,\sigma,\sim} := \left(\int_0^\infty |u(t)|^p (1-e^{-t})^\sigma \, dt\right)^{\frac{1}{p}}$$

defines an equivalent norm on $\mathbf{L}^p_\sigma(\mathbb{R}^+)$, i.e., there are positive constants $\mathcal{C}_1$ and $\mathcal{C}_2$ such that

$$\mathcal{C}_1 \|u\|_{p,\sigma,\sim} \le \|u\|_{p,\sigma} \le \mathcal{C}_2 \|u\|_{p,\sigma,\sim} \quad \forall\, u \in \mathbf{L}^p_\sigma(\mathbb{R}^+)\,.$$

For $a \in \mathbf{M}_{p,\sigma}$, the so-called Wiener-Hopf integral operator $\mathcal{W}_a \in \mathcal{L}(\mathbf{L}^p_\sigma(\mathbb{R}^+))$ is defined as the restriction of $\mathcal{P}\mathcal{W}^0_a\mathcal{P}$ to the space $\mathbf{L}^p_\sigma(\mathbb{R}^+)$. For $1 < p < \infty$, note that, w.r.t. the dual product

$$\langle u, v\rangle_{\mathbb{R}} := \int_{\mathbb{R}} u(t)\overline{v(t)}\, dt$$

the dual space of $\mathbf{L}^p_\sigma(\mathbb{R})$ is equal to $\mathbf{L}^q_\mu(\mathbb{R})$ with $\mu = (1-q)\sigma$ and $q = \frac{p}{p-1}$. Analogously, $\left(\mathbf{L}^p_\sigma(\mathbb{R}^+)\right)^* = \mathbf{L}^q_\mu(\mathbb{R}^+)$.

Lemma 2.1 *Let $1 < p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$. Then $\mathbf{M}_{p,\sigma} = \mathbf{M}_{q,\mu}$. Moreover, if $a \in \mathbf{M}_{p,\sigma}$, then $\left(\mathcal{W}^0_a\right)^* = \mathcal{W}^0_{\overline{a}}$ as well as $(\mathcal{W}_a)^* = \mathcal{W}_{\overline{a}}$.*

Proof Define $\mathcal{R} : \mathbf{L}^p_\sigma(\mathbb{R}) \to \mathbf{L}^p_\sigma(\mathbb{R})$ by $(\mathcal{R}u)(t) = u(-t)$. Then, $\mathcal{R}\mathcal{F}u = \mathcal{F}\mathcal{R}u = 2\pi\mathcal{F}^- u$ for $u \in \mathbf{L}^p(\mathbb{R})$ and $1 < p \le 2$. Let $a \in \mathbf{M}_{p,\sigma}$ and $p \in (1,\infty)$. Taking into account (2.3), for $u \in \mathbf{L}^2(\mathbb{R}) \cap \mathbf{L}^p_\sigma(\mathbb{R})$ and $v \in \mathbf{L}^2(\mathbb{R}) \cap \mathbf{L}^q_\mu(\mathbb{R})$, we have

$$\left\langle \mathcal{W}^0_a u, v\right\rangle_{\mathbb{R}} = \left\langle u, \mathcal{F}\overline{a}\mathcal{F}^{-1} v\right\rangle_{\mathbb{R}} = \left\langle u, \mathcal{W}^0_{\overline{a}} v\right\rangle_{\mathbb{R}}\,.$$

Since $\mathcal{W}^0_a$ is continuous on $\left(\mathbf{L}^2(\mathbb{R}) \cap \mathbf{L}^p_\sigma(\mathbb{R}), \|.\|_{p,\sigma,\mathbb{R}}\right)$, the operator $\mathcal{W}^0_{\overline{a}}$ is continuous on $\left(\mathbf{L}^2(\mathbb{R}) \cap \mathbf{L}^q_\mu(\mathbb{R}), \|.\|_{q,\mu,\mathbb{R}}\right)$, which means $\overline{a} \in \mathbf{M}_{q,\mu}$. Using

$$\mathcal{R}\overline{\mathcal{F}v} = \mathcal{F}\overline{v} \quad \text{and} \quad \overline{\mathcal{F}^{-1}v} = (2\pi)^{-1}\mathcal{F}\overline{v} = \mathcal{F}^{-1}\mathcal{R}\overline{v} \quad \text{on} \quad \mathbf{L}^2(\mathbb{R}),$$

we get

$$\mathcal{R}\overline{\mathcal{F}\overline{a}\mathcal{F}^{-1}v} = \mathcal{F}a\overline{\mathcal{F}^{-1}v} = \mathcal{F}a\mathcal{F}^{-1}\mathcal{R}\overline{v} \quad \forall\, v \in \mathbf{L}^2(\mathbb{R}) \cap \mathbf{L}^q_\mu(\mathbb{R}),$$

which implies $a \in \mathbf{M}_{q,\mu}$. Hence, due to symmetry reasons, $\mathbf{M}_{p,\sigma} = \mathbf{M}_{q,\mu}$. By the way $\left(\mathcal{W}^0_a\right)^* = \mathcal{W}^0_{\overline{a}}$ and, consequently, $(\mathcal{W}_a)^* = \mathcal{P}\left(\mathcal{W}^0_a\right)^*\mathcal{P} = \mathcal{W}_{\overline{a}}$, since $\mathcal{P}^* = \mathcal{P}$. □

Let us finish this section with two lemmata on the triviality of the nullspace of a Wiener-Hopf integral operator or of its adjoint.

Lemma 2.2 ([2], Prop. 2.8) *Let us assume that* $1 < p < \infty$, $a \in \mathbf{M}_{p,0}$, *and* $a \neq 0$ *almost everywhere on* $\mathbb{R}$. *Then the nullspace of* $\mathcal{W}_a : \mathbf{L}^p(\mathbb{R}^+) \to \mathbf{L}^p(\mathbb{R}^+)$ *or the nullspace of the adjoint operator* $\mathcal{W}_{\widetilde{a}} : \mathbf{L}^q(\mathbb{R}^+) \to \mathbf{L}^q(\mathbb{R}^+)$ *are trivial, where* $\frac{1}{p} + \frac{1}{q} = 1$.

The following lemma generalizes the previous one to the case of weighted spaces using a stronger assumption on $a(t)$. Here, by $\Pi\mathbf{C}_{p,\sigma}$ we refer to the closure of the set of piecewise constant functions on $\dot{\mathbb{R}}$ (with finitely many jumps) in the Banach algebra $\left(\mathbf{M}_{p,\sigma}, \|.\|_{\mathbf{M}_{p,\sigma}}\right)$.

Lemma 2.3 ([8], Prop. 1.8) *If* $a \in \Pi\mathbf{C}_{p,\sigma}$ *and* $\inf_{t\in\mathbb{R}} |a(t)| > 0$, *then the homogeneous equation* $\mathcal{W}_a u = \Theta$ *in the space* $\mathbf{L}^p_\sigma(\mathbb{R}^+)$ *or the adjoint equation* $\mathcal{W}_{\widetilde{a}} v = \Theta$ *in the space* $\mathbf{L}^q_{(1-q)\sigma}(\mathbb{R}^+)$, $\frac{1}{p} + \frac{1}{q} = 1$, *have only the trivial solution.*

3 Boundedness

In this section we introduce the Mellin transformation and the Mellin type operators as well as the Cauchy singular integral operator on the finite interval $(0, 1)$. Useful properties of the Mellin transformation as well as a relation between the convolution operators from the previous section and the Mellin type operators are given. Furthermore, we formulate conditions for the Mellin operators to be bounded in weighted $\mathbf{L}^p$-spaces.

For $z \in \mathbb{C}$ and a measurable function $f : (0, \infty) \to \mathbb{C}$, for which $t^{z-1} f(t)$ is integrable on each compact subinterval of $(0, \infty)$, the Mellin transform $\widehat{f}(z)$ is defined as

$$\widehat{f}(z) = \lim_{R\to\infty} \int_{R^{-1}}^{R} t^{z-1} f(t)\, dt,$$

if this limit exists. For $f \in \widetilde{\mathbf{L}}^1_{\xi-1}(\mathbb{R}^+)$ and $\eta \in \mathbb{R}$, we have

$$\begin{aligned} \widehat{f}(\xi + \mathbf{i}\eta) &= \int_0^\infty t^{\xi+\mathbf{i}\eta-1} f(t)\, dt = \int_{-\infty}^\infty e^{-(\xi+\mathbf{i}\eta)s} f(e^{-s})\, ds \\ &= \left(\mathcal{F}\widetilde{\mathcal{Z}}_\xi f\right)(\eta) = 2\pi\left(\mathcal{F}^-\widetilde{\mathcal{Z}}_\xi f\right)(-\eta). \end{aligned} \tag{3.1}$$

Let $\Gamma_\xi := \{z \in \mathbb{C} : \Re z = \xi\}$ and

$$\mathbf{C}_0(\Gamma_\xi) = \left\{ u \in \mathbf{C}(\Gamma_\xi) : \lim_{|\eta|\to\infty} u(\xi + \mathbf{i}\eta) = 0 \right\}.$$

Since, for $u \in \mathbf{L}^1(\mathbb{R})$, $\mathcal{F}u \in \mathbf{C}(\mathbb{R})$ and $\lim_{|\eta|\to\infty}(\mathcal{F}u)(\eta) = 0$ (due to the Riemann-Lebesgue theorem), by (3.1) we see that, for $\xi \in \mathbb{R}$, the Mellin transformation

$$\mathcal{M}_\xi : \widetilde{\mathbf{L}}^1_{\xi-1}(\mathbb{R}^+) \to \mathbf{C}_0(\Gamma_\xi), \quad f \mapsto \widehat{f}$$

is well defined. Moreover, if $\widehat{f}(\xi + \mathbf{i}\cdot) \in \mathbf{L}^1(\mathbb{R})$ and $f \in \mathbf{C}(0,\infty)$, then (cf. [4, Lemma 2.8])

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} t^{-\xi-\mathbf{i}\eta}\,\widehat{f}(\xi+\mathbf{i}\eta)\,d\eta\,, \quad 0 < x < \infty\,. \tag{3.2}$$

By $\mathbf{L}^p(\Gamma_\xi)$ we denote the $\mathbf{L}^p$-space given by the norm

$$\|g\|_{\Gamma_\xi,p} = \left(\int_{-\infty}^{\infty} |g(\xi+\mathbf{i}\eta)|^p\,d\eta\right)^{\frac{1}{p}}.$$

Taking into account relation (3.1), we are able to define the Mellin transform for $f \in \widetilde{\mathbf{L}}^p_{p\xi-1}(\mathbb{R}^+)$ and $1 < p \le 2$ by

$$\widehat{f}(\xi+\mathbf{i}\cdot) := \mathcal{F}\widetilde{\mathcal{Z}}_\xi f = \lim_{R\to\infty} \mathcal{F}_R\widetilde{\mathcal{Z}}_\xi f = \lim_{R\to\infty}\int_{e^{-R}}^{e^R} t^{\xi+\mathbf{i}\cdot-1} f(t)\,dt, \tag{3.3}$$

where the limit is taken in the $\mathbf{L}^q(\Gamma_\xi)$-sense and where $\frac{1}{p}+\frac{1}{q}=1$.

Corollary 3.1 *Let $1 < p \le 2$ and $\frac{1}{p}+\frac{1}{q}$. For $\xi \in \mathbb{R}$, the Mellin transformation*

$$\mathcal{M}_\xi : \widetilde{\mathbf{L}}^p_{p\xi-1}(\mathbb{R}^+) \to \mathbf{L}^q(\Gamma_\xi), \quad f \mapsto \widehat{f},$$

defined by (3.3), is a linear and bounded operator.

For $-\infty < \alpha < \beta < \infty$, by $\Gamma_{\alpha,\beta}$ we refer to the strip

$$\Gamma_{\alpha,\beta} = \{z \in \mathbb{C} : \alpha < \Re z < \beta\}$$

of the complex plane. The following lemma modifies [4, Cor. 2.9].

Lemma 3.2 *Let $1 < p \le 2$, $\frac{1}{p}+\frac{1}{q}=1$, and $\alpha, \beta \in \mathbb{R}$ with $\alpha < \beta$. Moreover, let $k \in \widetilde{\mathbf{L}}^p_{p\xi-1}(\mathbb{R}^+)\cap\mathbf{C}(\mathbb{R}^+)$ for every $\xi \in (\alpha,\beta)$. By Corollary 3.1 we have $\widehat{k} \in \mathbf{L}^q(\Gamma_\xi)$ for all $\xi \in (\alpha,\beta)$. If, additionally, $\widehat{k}$ is holomorphic in the strip $\Gamma_{\alpha,\beta}$ satisfying*

$$M_{\alpha_0,\beta_0} := \sup\left\{(1+|z|)^{1+\delta}\left|\widehat{k}'(z)\right| : z \in \Gamma_{\alpha_0,\beta_0}\right\} < \infty \tag{3.4}$$

for all intervals $[\alpha_0, \beta_0] \subset (\alpha, \beta)$ *and some* $\delta = \delta(\alpha_0, \beta_0) > 0$, *then there hold*

$$k(t) = \frac{t^{\mu}}{2\pi(1-t^{2\mu})} \int_{-\infty}^{\infty} t^{-\xi-\mathbf{i}\eta} \left[\widehat{k}(\xi-\mu+\mathbf{i}\eta) - \widehat{k}(\xi+\mu+\mathbf{i}\eta)\right] d\eta \tag{3.5}$$

for all $\xi \in (\alpha, \beta)$, $x \in \mathbb{R}^+$, *and* $0 < \mu < \min\{\xi-\alpha, \beta-\xi\}$. *Moreover,* $\widehat{k}(\xi-\mu+\mathbf{i}\cdot) - \widehat{k}(\xi+\mu+\mathbf{i}\cdot) \in \mathbf{L}^1(\mathbb{R})$ *and*

$$\int_{-\infty}^{\infty} \left[\widehat{k}(\xi-\mu+\mathbf{i}\eta) - \widehat{k}(\xi+\mu+\mathbf{i}\eta)\right] d\eta = 0\,. \tag{3.6}$$

Proof For $\alpha < \xi \pm \mu < \beta$, $\eta \in \mathbb{R}$, and $0 < x < \infty$, we have

$$\widehat{k}(\xi+\mu+\mathbf{i}\eta) - \widehat{k}(\xi-\mu+\mathbf{i}\eta) = \int_{-\mu}^{\mu} \widehat{k}'(\xi+s+\mathbf{i}\eta)\, ds\,,$$

such that, due to (3.4),

$$\begin{aligned}
&\int_{-\infty}^{\infty} \left|\widehat{k}(\xi+\mu+\mathbf{i}\eta) - \widehat{k}(\xi-\mu+\mathbf{i}\eta)\right| d\eta \\
&\quad\le \int_{-\infty}^{\infty}\int_{-\mu}^{\mu} \left|\widehat{k}'(\xi+s+\mathbf{i}\eta)\right| ds\, d\eta \\
&\quad\le M_{\xi-\mu,\xi+\mu} \int_{-\infty}^{\infty}\int_{-\mu}^{\mu} \frac{ds\, d\eta}{(1+|\xi+s+\mathbf{i}\eta|)^{1+\delta}} < \infty
\end{aligned}$$

Hence, the continuous function $g : (0,\infty)$, $x \mapsto k(x)x^{-\mu} - k(x)x^{\mu}$, belongs to $\widetilde{\mathbf{L}}^p_{p\xi-1}$ (by assumption) and its Mellin transform

$$\widehat{g}(\xi+\mathbf{i}\cdot) = \widehat{k}(\xi-\mu+\mathbf{i}\cdot) - \widehat{k}(\xi+\mu+\mathbf{i}\cdot)$$

to $\mathbf{L}^1(\mathbb{R})$. Relation (3.2) yields (3.5), and relation (3.6) is a consequence of $g(1) = 0$. □

Corollary 3.3 *Under the assumptions of Lemma 3.2, we have*

$$\lim_{t\to+0} t^{\alpha+\varepsilon} k(t) = 0 \quad \textit{and} \quad \lim_{t\to\infty} t^{\beta-\varepsilon} k(t) = 0$$

for every $\varepsilon > 0$.

Proof For given $\varepsilon > 0$, choose $\xi \in (\alpha, \beta)$ and $\mu > 0$ such that $0 < \mu < \xi - \alpha < \beta - \xi$ and $\xi - \alpha - \mu < \varepsilon$. Then, for $0 < t < \frac{1}{2}$, by taking into account (3.5) and (3.6),

$$t^{\alpha+\varepsilon}|k(t)| \le \text{const}\, t^{\alpha+\varepsilon+\mu-\xi} \left\|\widehat{k}(\xi-\mu+\mathbf{i}\cdot) - \widehat{k}(\xi+\mu+\mathbf{i}\cdot)\right\|_{\mathbf{L}^1(\mathbb{R})} \to 0$$

if $t \to +0$. Analogously, we can choose $\xi \in (\alpha, \beta)$ and $\mu > 0$ such that $0 < \mu < \beta - \xi < \xi - \alpha$ and $\beta - \xi - \mu < \varepsilon$. Again by (3.5) and (3.6) we get, for $t > 2$,

$$t^{\beta-\varepsilon}|k(t)| \le \text{const}\, t^{\beta-\varepsilon-\mu-\xi} \left\|\widehat{k}(\xi - \mu + \mathbf{i}\cdot) - \widehat{k}(\xi + \mu + \mathbf{i}\cdot)\right\|_{\mathbf{L}^1(\mathbb{R})} \to 0$$

if $t \to \infty$. □

The following Lemma modifies Lemma 3.10 in [5].

Lemma 3.4 *Let $p \in [1, \infty)$, $-\infty < \alpha < \beta < \infty$, and*

$$k \in \widetilde{\mathbf{L}}^p_\alpha(\mathbb{R}^+) \cap \widetilde{\mathbf{L}}^p_\beta(\mathbb{R}^+),$$

Then we have $k \in \widetilde{\mathbf{L}}^1_{\frac{1+\rho}{p}-1}(\mathbb{R}^+)$ for all $\rho \in (\alpha, \beta)$.

Proof Let $\alpha < \rho < \beta$, $1 < p < \infty$, and $\frac{1}{p} + \frac{1}{q} = 1$. Then, due to Hölder's inequality,

$$\begin{aligned}\int_0^\infty t^{\frac{1+\rho}{p}-1}|f(t)|\,dt &= \int_0^1 t^{\frac{1+\rho-\alpha}{p}-1}|f(t)|\,t^{\frac{\alpha}{p}}\,dt + \int_1^\infty t^{\frac{1+\rho-\beta}{p}}|f(t)|\,t^{\frac{\beta}{p}}\,dt \\ &\le \left(\int_0^1 t^{\left(\frac{1+\rho-\alpha}{p}-1\right)q}\,dt\right)^{\frac{1}{q}} \|f\|_{\alpha,p,\mathbb{R}^+,\sim} \\ &\quad + \left(\int_1^\infty t^{\left(\frac{1+\rho-\beta}{p}-1\right)q}\,dt\right)^{\frac{1}{q}} \|f\|_{\beta,p,\mathbb{R}^+,\sim},\end{aligned}$$

where

$$\left(\frac{1+\rho-\alpha}{p} - 1\right)q > -1 \quad \text{and} \quad \left(\frac{1+\rho-\beta}{p} - 1\right)q < -1.$$

In case $p = 1$, we simply have

$$\begin{aligned}\int_0^\infty t^\rho|f(t)|\,dt &\le \int_0^1 t^\alpha|f(t)|\,dt + \int_1^\infty t^\beta|f(t)|\,dt \\ &\le \|f\|_{\alpha,1,\mathbb{R}^+,\sim} + \|f\|_{\beta,1,\mathbb{R}^+,\sim},\end{aligned}$$

and the corollary is proved. □

Corollary 3.5 *Let $k \in \widetilde{\mathbf{L}}^p_{p\alpha-1}(\mathbb{R}^+) \cap \widetilde{\mathbf{L}}^p_{p\beta-1}(\mathbb{R}^+)$ for some $p \in [1, \infty)$ and some real numbers α, β with $\alpha < \beta$. Then the Mellin transform $\widehat{k}$ is holomorphic in the strip $\Gamma_{\alpha,\beta}$.*

Proof By applying Lemma 3.4 we infer $k \in \widetilde{\mathbf{L}}^1_{\xi-1}(\mathbb{R}^+)$ for all $\xi \in (\alpha, \beta)$, and [5, Lemma 2.14] yields the assertion. □

Let $p \in [1, \infty)$, $\rho, \sigma \in \mathbb{R}$. We denote by $\mathbf{L}^p_{\rho,\sigma} := \mathbf{L}^p_{\rho,\sigma}(0,1)$ the weighted $\mathbf{L}^p$-space equipped with the norm

$$\|f\|_{\rho,\sigma,p} := \left(\int_0^1 |f(x)|^p \, v^{\rho,\sigma}(x)\, dx \right)^{1/p}, \quad v^{\rho,\sigma}(x) = x^\rho (1-x)^\sigma.$$

By $\mathcal{S}$ we denote the Cauchy singular integral operator given by

$$(\mathcal{S}f)(x) = \frac{1}{\pi \mathbf{i}} \int_0^1 \frac{f(y)}{y-x}\, dy, \quad x \in (0,1),$$

where the integral is considered as a Cauchy principal value one. It is well known that $\mathcal{S} : \mathbf{L}^p_{\rho,\sigma} \to \mathbf{L}^p_{\rho,\sigma}$ is a linear and bounded operator if and only if $1 < p < \infty$ and $-1 < \rho, \sigma < p-1$ (see, for example, [3, Sections 1.2, 1.5]). For a measurable function $k : \mathbb{R}^+ \to \mathbb{C}$ we define the Mellin operator $\mathcal{B}_k$ by

$$(\mathcal{B}_k f)(x) = \int_0^1 k\left(\frac{x}{y}\right) \frac{f(y)}{y}\, dy, \quad x \in (0,1).$$

Recall that, for $p \in [1, \infty)$ and $\rho \in \mathbb{R}$, the integral operator $\mathcal{B}_k : \mathbf{L}^p_{\rho,0} \to \mathbf{L}^p_{\rho,0}$ is bounded, if $k \in \widetilde{\mathbf{L}}^1_{\frac{1+\rho}{p}-1}(\mathbb{R}^+)$ (cf. [5, Lemma 3.7]).

Let $p \in [1, \infty)$, $\rho, \sigma \in \mathbb{R}$ and $\xi = \frac{1+\rho}{p}$. We introduce the mapping

$$\mathcal{Z}_\xi : \mathbf{L}^p_{\rho,\sigma} \to \mathbf{L}^p_\sigma(\mathbb{R}^+), \quad f \mapsto e^{-\xi \cdot} f(e^{-\cdot}),$$

which is a continuous isomorphism with $\left(\mathcal{Z}_\xi^{-1} f\right)(x) = x^{-\xi} f(-\ln x)$.

Lemma 3.6 *For $p \in (1, \infty)$, $\rho, \sigma \in (-1, p-1)$, and $a(t) = -\mathbf{i}\cot(\pi\xi - \mathbf{i}\pi t)$, the relation $\mathcal{Z}_\xi \mathcal{S} \mathcal{Z}_\xi^{-1} = \mathcal{W}_a$ holds true in $\mathbf{L}^p_\sigma(\mathbb{R}^+)$, where $\xi = \frac{1+\rho}{p}$.*

Proof From (2.7) and the boundedness of $\mathcal{S}_{\mathbb{R}^+} : \mathbf{L}^p_\sigma(\mathbb{R}^+) \to \mathbf{L}^p_\sigma(\mathbb{R}^+)$, we infer

$$\left\| \mathcal{F} a \mathcal{F}^{-1} u \right\|_{\sigma,p,\mathbb{R}} \le \|\mathcal{S}_{\mathbb{R}^+}\|_{\mathcal{L}(\mathbf{L}^p_\sigma(\mathbb{R}^+))} \, \|u\|_{\sigma,p,\mathbb{R}}$$

for all $u \in \mathbf{L}^2(\mathbb{R}) \cap \mathbf{L}^p_\sigma(\mathbb{R})$. Hence, $a \in \mathbf{M}_{p,\sigma}$. On the other hand, if $u \in \mathbf{L}^p_\sigma(\mathbb{R}^+)$ and $\widetilde{u} = \widetilde{\mathcal{P}} u$, then

$$\left(\mathcal{Z}_\xi \mathcal{S} \mathcal{Z}_\xi^{-1} u\right)(t) = \frac{e^{-\xi t}}{\pi \mathbf{i}} \int_0^\infty \frac{y^{-\xi} \widetilde{u}(-\ln y)\, dy}{y - e^{-t}}, \quad t \in \mathbb{R}^+,$$

i.e., $\mathcal{Z}_\xi \mathcal{S} \mathcal{Z}_\xi^{-1} = \mathcal{P}\widetilde{\mathcal{Z}}_\xi \mathcal{S}_{\mathbb{R}^+} \widetilde{\mathcal{Z}}_\xi^{-1}\widetilde{\mathcal{P}} = \mathcal{P}\mathcal{W}_a^0\widetilde{\mathcal{P}}$, where we again used relation (2.7). This proves the lemma. □

Lemma 3.7 *Let $1 < p \le 2$, $\rho \in \mathbb{R}$, and $\xi = \frac{1+\rho}{p}$, and set $a(t) = \widehat{k}(\xi - \mathbf{i}t)$. We assume that $k \in \widetilde{\mathbf{L}}^q_{q\xi-1}(\mathbb{R}^+)$ is satisfied for some $q \in \left[1, \frac{p}{p-1}\right) \cap [1, 2]$. Then we have*

$$\int_0^1 k\left(\frac{x}{y}\right) \frac{g(y)}{y}\, dy = x^{-\xi}\left(\mathcal{P}\mathcal{F}a\mathcal{F}^-\widetilde{\mathcal{P}}\mathcal{Z}_\xi g\right)(-\ln x) \tag{3.7}$$

for all $g \in \mathbf{L}^p_{\rho,0}$ and for almost all $x \in (0, 1)$.

Proof Let $\kappa(t) := \left(\widetilde{\mathcal{Z}}_\xi k\right)(t) = e^{-\xi t} k\left(e^{-t}\right)$, $t \in \mathbb{R}$. Then $\kappa \in \mathbf{L}^q(\mathbb{R})$ and, due to (2.4),

$$\mathcal{K}f := \kappa * f = \int_{-\infty}^{\infty} \kappa(\cdot - s) f(s)\, ds \in \mathbf{L}^r(\mathbb{R}), \quad r = \frac{1}{\frac{1}{p} + \frac{1}{q} - 1},$$

for all $f \in \mathbf{L}^p(\mathbb{R})$. Moreover, in virtue of (2.5),

$$\mathcal{K}f = 2\pi \mathcal{F}\left(\mathcal{F}^-\kappa \cdot \mathcal{F}^- f\right).$$

Hence $\mathcal{K}f = \mathcal{F}a\mathcal{F}^- f$ for all $f \in \mathbf{L}^p(\mathbb{R})$, where (cf. (3.1))

$$a(t) = 2\pi\left(\mathcal{F}^-\kappa\right)(t) = 2\pi\left(\mathcal{F}^-\widetilde{\mathcal{Z}}_\xi k\right)(t) = \widehat{k}(\xi - \mathbf{i}t).$$

On the other hand we have, for $f \in \mathbf{L}^p(\mathbb{R}^+)$ and for almost all $x \in (0, 1)$,

$$\begin{aligned} x^{-\xi}\left(\mathcal{P}\mathcal{K}\widetilde{\mathcal{P}} f\right)(-\ln x) &= \int_0^\infty k\left(e^s x\right) e^{\xi s} f(s)\, ds \\ &= \int_0^1 k\left(\frac{x}{y}\right) y^{-\xi} f(-\ln y)\, \frac{dy}{y} \\ &= \int_0^1 k\left(\frac{x}{y}\right) \left(\mathcal{Z}_\xi^{-1} f\right)(y)\, \frac{dy}{y}, \end{aligned}$$

and the lemma is proved, since every $g \in \mathbf{L}^p_{\rho,0}$ can be represented in the form $g = \mathcal{Z}_\xi^{-1} f$ with $f \in \mathbf{L}^p(\mathbb{R})$. □

Lemma 3.8 ([5], Prop. 3.13) *Let $p \in (1, \infty)$, $\rho \in (-1, p-1)$, and $k \in \mathbf{C}(\mathbb{R}^+)$. Moreover, we assume that there are real numbers α, β with $\alpha < \beta$ such that $\frac{1+\rho}{p} \in (\alpha, \beta)$ and such that*

$$\lim_{t\to+0} t^\alpha k(t) = 0 \quad \textit{and} \quad \lim_{t\to\infty} t^\beta k(t) = 0.$$

Then, for all $\sigma \in (-1, p-1)$, the integral operator $\mathcal{B}_k : \mathbf{L}^p_{\rho,\sigma} \to \mathbf{L}^p_{\rho,\sigma}$ is bounded.

Corollary 3.9 *Let $p \in (1,\infty)$, $\rho \in (-1, p-1)$, and $k \in \mathbf{C}(\mathbb{R}^+)$. Furthermore, we assume that there are real numbers α, β with $\alpha < \beta$ such that $\xi := \frac{1+\rho}{p} \in (\alpha,\beta)$ and $k \in \widetilde{\mathbf{L}}^q_{q\eta-1}(\mathbb{R}^+)$ for some $q \in (1,2]$ and all $\eta \in (\alpha,\beta)$. Moreover, we suppose that the Mellin transform $\widehat{k}$ is holomorphic in the strip $\Gamma_{\alpha,\beta}$ and fulfils (3.4). Then the integral operator $\mathcal{B}_k : \mathbf{L}^p_{\rho,\sigma} \to \mathbf{L}^p_{\rho,\sigma}$ is bounded for all $\sigma \in (-1, p-1)$.*

Proof Choose $\varepsilon > 0$ such that $\alpha + \varepsilon < \xi < \beta - \varepsilon$. Corollary 3.3 yields

$$\lim_{x\to+0} x^{\alpha+\varepsilon} k(x) = \lim_{x\to\infty} x^{\beta-\varepsilon} k(x) = 0\,,$$

and by Lemma 3.8 we get the assertion. □

Lemma 3.10 *Let $p \in (1,\infty)$, $\rho \in (-1, p-1)$, and $\xi = \frac{1+\rho}{p}$. We assume that $k \in \widetilde{\mathbf{L}}^q_{q\xi-1}(\mathbb{R}^+)$ for some $q \in \left[1, \frac{p}{p-1}\right) \cap [1,2]$ and $\widehat{k}(\xi + \mathbf{i}\cdot) \in \mathbf{V}_1(\mathbb{R})$. Then the operator $\mathcal{B}_k$ is bounded on the space $\mathbf{L}^p_{\rho,0} \cap \mathbf{L}^p_{\rho,\sigma}$ equipped with the norm $\|\cdot\|_{\rho,\sigma,p}$. Thus, the operator possesses a unique extension to the space $\mathbf{L}^p_{\rho,\sigma}$.*

Proof Let $a(t) = \widehat{k}(\xi - \mathbf{i}t)$. Relation (2.9) delivers $a \in \mathbf{M}_{p,\sigma}$, which means that the operator $\mathcal{W}_a : \mathbf{L}^p_\sigma(\mathbb{R}^+) \to \mathbf{L}^p_\sigma(\mathbb{R}^+)$ is well-defined and bounded. Thus $\mathcal{Z}_\xi^{-1}\mathcal{W}_a\mathcal{Z}_\xi : \mathbf{L}^p_{\rho,\sigma} \to \mathbf{L}^p_{\rho,\sigma}$ is bounded, too. Now from (3.7) follows

$$\mathcal{B}_k f = \mathcal{Z}_\xi^{-1}\mathcal{W}_a\mathcal{Z}_\xi f, \quad f \in \mathbf{L}^p_{\rho,0} \cap \mathbf{L}^p_{\rho,\sigma},$$

which completes the proof. □

4 Fredholm Properties

Here we derive necessary and sufficient conditions for the Fredholmness of the operators of the form (1.1) and also an index formula based on the winding number of a certain closed curve in $\mathbb{C}$ associated with such operators.

Lemma 4.1 *Let $p \in (1,\infty)$, $\frac{1}{p} + \frac{1}{q} = 1$ and $\rho, \sigma \in (-1, p-1)$ as well as $k \in \mathbf{C}(\mathbb{R}^+)$. For $\mathcal{A} \in \mathcal{L}\big(\mathbf{L}^p_{\rho,\sigma}\big)$ we define the adjoint operator by*

$$\int_0^1 (\mathcal{A}u)(x)\overline{v(x)}\,dx = \int_0^1 u(x)\overline{(\mathcal{A}^*v)(x)}\,dx, \quad \forall\, u \in \mathbf{L}^p_{\rho,\sigma},\ \forall\, v \in \mathbf{L}^q_{\rho',\sigma'}$$

with $\rho' = (1-q)\rho$ and $\sigma' = (1-q)\sigma$. If $\mathcal{B}_k \in \mathcal{L}\big(\mathbf{L}^p_{\rho,\sigma}\big)$ then

$$(\mathcal{S} + \mathcal{B}_k)^* = \mathcal{S} + \mathcal{B}_{k_1}, \quad k_1(t) = \overline{k(t^{-1})}t^{-1}.$$

Proof The proof is straightforward if one takes into account the commutation formula for Cauchy principal value and usual integrals (cf.[6, Chapter II, Prop. 4.4]). □

We recall that the assumption $k \in \widetilde{\mathbf{L}}^p_{p\alpha-1}(\mathbb{R}^+) \cap \widetilde{\mathbf{L}}^p_{p\beta-1}(\mathbb{R}^+)$ implies that the Mellin transform $\widehat{k} : \Gamma_{\alpha,\beta} \to \mathbb{C}$ is holomorphic (cf. Corollary 3.5), and we formulate the following condition for a function $k \in \mathbf{C}(\mathbb{R}^+)$ and $1 < p < \infty$ as well as $-1 < \rho < p-1$:

(A) There exist real numbers α and β with $\alpha < \beta$ such that $\xi = \frac{1+\rho}{p} \in (\alpha, \beta)$ and $k \in \widetilde{\mathbf{L}}^r_{r\alpha-1}(\mathbb{R}^+) \cap \widetilde{\mathbf{L}}^r_{r\beta-1}(\mathbb{R}^+)$ for some $r \in (1, 2]$ as well as (3.4) is satisfied.

In contrast to [5] we replaced condition

$$\sup\left\{(1+|z|)^{1+\ell}\left|\widehat{k}^{(\ell)}(z)\right| : z \in \Gamma_{\alpha,\beta}\right\} < \infty, \quad \ell = 0, 1, 2, \ldots$$

(cf. [5, equation (3.11)]) by (3.4), which is a weaker condition regarding the Mellin transform $\widehat{k}$.

Lemma 4.2 *Let $a, b \in \mathbb{C}$, $p \in (1, \infty)$, $\rho \in (-1, p-1)$, $\frac{1}{p}+\frac{1}{q} = 1$, and $k \in \mathbf{C}(\mathbb{R}^+)$ satisfy condition* (A). *Then, for all $\sigma \in (-1, p-1)$, we have the representations*

$$\mathcal{Z}_\xi(a\mathcal{I} + b\mathcal{S} + \mathcal{B}_k)\mathcal{Z}_\xi^{-1} = \mathcal{W}_{\mathbf{a}} \tag{4.1}$$

and

$$\mathcal{Z}_{1-\xi}(a\mathcal{I} + b\mathcal{S} + \mathcal{B}_k)^*\mathcal{Z}_{1-\xi}^{-1} = \mathcal{W}_{\overline{\mathbf{a}}} = (\mathcal{W}_{\mathbf{a}})^*, \tag{4.2}$$

in the spaces $\mathbf{L}^p_\sigma(\mathbb{R}^+)$ and $\mathbf{L}^q_{(1-q)\sigma}(\mathbb{R}^+)$, respectively, where

$$\mathbf{a}(t) = a - b\mathbf{i}\cot\pi(\xi - \mathbf{i}t) + \widehat{k}(\xi - \mathbf{i}t), \ t \in \mathbb{R}.$$

Proof Since the map $\mathcal{Z}_\xi : \mathbf{L}^p_{\rho,\sigma} \to \mathbf{L}^p_\sigma(\mathbb{R}^+)$ is a continuous isomorphism and $\mathcal{B}_k \in \mathcal{L}\big(\mathbf{L}^p_{\rho,\sigma}\big)$ for all $\sigma \in (-1, p-1)$ (cf. Corollary 3.9), Lemma 3.7 delivers immediately the relation

$$\widetilde{\mathcal{P}}\mathcal{Z}_\xi\mathcal{B}_k\mathcal{Z}_\xi^{-1}f = \mathcal{P}\mathcal{F}\widehat{k}(\xi - \mathbf{i}\cdot)\mathcal{F}^{-1}\widetilde{\mathcal{P}}f, \quad f \in \mathbf{L}^2(\mathbb{R}^+) \cap \mathbf{L}^p_\sigma(\mathbb{R}^+).$$

Due to (3.4), we have that $\widehat{k}(\xi - \mathbf{i}\cdot) \in \mathbf{M}_{p,\sigma}$ (cf. (2.9)). This yields $\mathcal{Z}_\xi\mathcal{B}_k\mathcal{Z}_\xi^{-1} = \mathcal{W}_{\mathbf{b}}$ with $\mathbf{b}(t) = \widehat{k}(\xi - \mathbf{i}t)$. Together with Lemma 3.6 we get (4.1). Relation (4.2) is now a consequence of $\mathcal{Z}_\xi^* = \mathcal{Z}_{1-\xi}^{-1}$ and Lemma 2.1. □

Lemma 4.3 ([2], Corollary 1.19) *Let $p \in (1, \infty)$, $\rho, \sigma \in (-1, p-1)$, and $a, b \in \mathbb{C}$. Moreover, let $\xi = \frac{1+\rho}{p}$, $\eta = \frac{1+\sigma}{p}$. Then the operator $\mathcal{A} = a\mathcal{I} + b\mathcal{S} : \mathbf{L}^p_{\rho,\sigma} \to$*

$\mathbf{L}^p_{\rho,\sigma}$ *is Fredholm if and only if* $a \pm b \neq 0$ *and there exists an integer* κ *satisfying*

$$-\xi < \Re\nu < 1-\xi \quad \text{and} \quad \eta - 1 < \Re\nu + \kappa < \eta\,, \tag{4.3}$$

where

$$\frac{a+b}{a-b} = e^{2\pi\mathbf{i}\nu}\,. \tag{4.4}$$

(*Note that* κ *can only take the values* $0,\ -1$, *and* 1). *In this case, we have* $\operatorname{ind}\mathcal{A} = \kappa$ *and the operator* $\mathcal{A}$ *is invertible, invertible from the left, or invertible from the right, if the index of* $\mathcal{A}$ *is zero,* -1, *or* 1, *respectively. The corresponding* (*one-sided*) *inverse is given by*

$$\left(\mathcal{A}^{(-1)}f\right)(x) = \frac{1}{a^2-b^2}\left[af(x) - \frac{b}{\pi\mathbf{i}}\int_0^1 \left(\frac{x}{y}\right)^{\nu}\left(\frac{1-y}{1-x}\right)^{\nu+\kappa}\frac{f(y)}{y-x}\,dy\right],$$

i.e., $\mathcal{A}^{(-1)} = \left(a^2-b^2\right)^{-1}\left[a\mathcal{I} - bv^{\nu,-\nu-\kappa}\mathcal{S}v^{-\nu,\nu+\kappa}\mathcal{I}\right]$.

Moreover, for $\frac{a+b}{a-b} = e^{2\pi\mathbf{i}\mu}$, $\mu \neq 0$, *and* $0 \leq \Re\mu < 1$,

$$a\,v^{\mu-1,-\mu}(x) + \frac{b}{\pi\mathbf{i}}\int_0^1 \frac{v^{\mu-1,-\mu}(y)\,dy}{y-x} = 0\,, \quad 0 < x < 1\,, \tag{4.5}$$

i.e., in case $\kappa = 1$, *we have* $\mu = 1+\nu$ *and the nullspace of the operator* $\mathcal{A} : \mathbf{L}^p_{\rho,\sigma} \to \mathbf{L}^p_{\rho,\sigma}$ *is spanned by* $v^{\nu,-\nu-1}$.

Corollary 4.4 *In* (4.4) *we can choose* ν *in such a way that* $0 \leq \Re\nu < 1$. *Then,*

$$\frac{a-b}{a+b} = e^{2\pi\mathbf{i}(1-\nu)}$$

and, due to (4.5),

$$a\,v^{-\nu,\nu-1}(x) - \frac{b}{\pi\mathbf{i}}\int_0^1 \frac{v^{-\nu,\nu-1}(y)\,dy}{y-x} = 0\,, \quad 0 < x < 1\,. \tag{4.6}$$

Moreover,

$$a\,v^{-\nu,\nu}(x) - \frac{b}{\pi\mathbf{i}}\int_0^1 \frac{v^{-\nu,\nu}(y)\,dy}{y-x} = \gamma_0\,, \quad 0 < x < 1\,, \tag{4.7}$$

and

$$a\,v^{1-\nu,\nu}(x) - \frac{b}{\pi\mathbf{i}}\int_0^1 \frac{v^{1-\nu,\nu}(y)\,dy}{y-x} = \delta_0 + \delta_1 x\,, \quad 0 < x < 1\,, \tag{4.8}$$

with certain constants $\gamma_0, \delta_0, \delta_1 \in \mathbb{C}$*, where* (4.7) *remains true also in case of* $-1 < \Re\nu < 0$.

Proof It remains to prove relations (4.7) and (4.8). By using (4.6), we get

$$\begin{aligned} & a\, v^{-\nu,\nu}(x) - \frac{b}{\pi \mathbf{i}} \int_0^1 \frac{v^{-\nu,\nu}(y)\, dy}{y-x} \\ &= (1-x)\left[a\, v^{-\nu,\nu-1}(x) - \frac{b}{\pi \mathbf{i}} \int_0^1 \frac{v^{-\nu,\nu-1}(y)\, dy}{y-x} \right] + \frac{b}{\pi \mathbf{i}} \int_0^1 v^{-\nu,\nu-1}(y)\, dy \\ &= \frac{b}{\pi \mathbf{i}} \int_0^1 v^{-\nu,\nu-1}(y)\, dy\,, \quad 0 < x < 1\,, \end{aligned}$$

and (4.7) is proved. Analogously,

$$\begin{aligned} & a\, v^{1-\nu,\nu}(x) - \frac{b}{\pi \mathbf{i}} \int_0^1 \frac{v^{1-\nu,\nu}(y)\, dy}{y-x} \\ &= x(1-x)\left[a\, v^{-\nu,\nu-1}(x) - \frac{b}{\pi \mathbf{i}} \int_0^1 \frac{v^{-\nu,\nu-1}(y)\, dy}{y-x} \right] \\ &\quad + \frac{b}{\pi \mathbf{i}} \int_0^1 (y+x-1) v^{-\nu,\nu-1}(y)\, dy \\ &= \frac{b}{\pi \mathbf{i}} \int_0^1 (y+x-1) v^{-\nu,\nu-1}(y)\, dy\,, \quad 0 < x < 1\,, \end{aligned}$$

which proves (4.8). To prove (4.7) in case $-1 < \Re\nu < 0$, we use $\frac{a-b}{a+b} = e^{2\pi \mathbf{i}(-\nu)}$ and get from (4.5) the relation

$$a\, v^{-\nu-1,\nu}(x) - \frac{b}{\pi \mathbf{i}} \int_0^1 \frac{v^{-\nu-1,\nu}(y)\, dy}{y-x} = 0\,, \quad 0 < x < 1\,,$$

and analogously as above we conclude (4.7) also in this case. □

Definition 4.5 For $p \in (1, \infty)$ and $\sigma \in (-1, p-1)$, we define the function $\sigma_\infty : \dot{\mathbb{R}} \to \mathbb{R}$ by

$$\sigma_\infty(t) := \begin{cases} \frac{1}{p} & : \ t \in \mathbb{R} \\ \frac{1+\sigma}{p} & : t = \infty. \end{cases}$$

Moreover, for $a \in \mathbf{PC}_{p,\sigma}$, we define the function $a_{p,\sigma} : \dot{\mathbb{R}} \times \overline{\mathbb{R}} \to \mathbb{C}$ by

$$\begin{aligned} a_{p,\sigma}(t,x) =& \frac{1}{2}\big[a(t+0)+a(t-0)\big] \\ & - \frac{\mathbf{i}}{2}\big[a(t+0)-a(t-0)\big]\cot\pi\big[\sigma_\infty(t)-\mathbf{i}x\big]. \end{aligned} \tag{4.9}$$

Then, the image of $a_{p,\sigma}$ defines a closed curve in the complex plane, which we denote by $\Gamma_{a_{p,\sigma}}$. Under the assumption

$$\inf\big\{\,|a_{p,\sigma}(t,x)| \,:\, t \in \mathbb{R}, x \in \mathbb{R}\,\big\} > 0 \tag{4.10}$$

the winding number $\operatorname{wind}\big(\Gamma_{a_{p,\sigma}}\big)$ is well defined, where the orientation of $\Gamma_{a_{p,\sigma}}$ is given by its inherent parametrization.

Lemma 4.6 ([1], Theorem 5.7, cf. also [8], Theorem 1.2) *Let $p \in (1,\infty)$, $\sigma \in (-1,\, p-1)$ and $a \in \mathbf{PC}_{p,\sigma}$. Then $\mathcal{W}_a$ is a Fredholm operator on $\mathbf{L}^p_\sigma(\mathbb{R}^+)$ if and only if (4.10) is satisfied. In this case, the Fredholm index of $\mathcal{W}_a$ is equal to $-\operatorname{wind}\big(\Gamma_{a_{p,\sigma}}\big)$.*

In the following, let $\xi = \frac{1+\rho}{p}$ for $p \in (1,\infty)$ and $\rho \in (-1,\, p-1)$.

Corollary 4.7 *Let $a, b \in \mathbb{C}$, $p \in (1,\infty)$, $\rho, \sigma \in (-1,\, p-1)$, and $k \in \mathbf{C}(\mathbb{R}^+)$ satisfy condition* (A). *Then the integral operator $\mathcal{A} := a\mathcal{I} + b\mathcal{S} + \mathcal{B}_k$ is Fredholm in the space $\mathbf{L}^p_{\rho,\sigma}$ if and only if the curve*

$$\begin{aligned} \Gamma_{\mathcal{A}} :=& \big\{a - b\mathbf{i}\cot\pi(\xi-\mathbf{i}t) + \widehat{k}(\xi-\mathbf{i}t) \,:\, t \in \mathbb{R}\big\} \\ & \cup \left\{a + b\mathbf{i}\cot\pi\left(\frac{1+\sigma}{p} - \mathbf{i}t\right) : t \in \overline{\mathbb{R}}\right\} \end{aligned} \tag{4.11}$$

does not run through the zero point. In this case, the Fredholm index of $\mathcal{A} : \mathbf{L}^p_{\rho,\sigma} \to \mathbf{L}^p_{\rho,\sigma}$ is equal to negative winding number of the curve $\Gamma_{\mathcal{A}}$, where the orientation is given by its inherent parametrization.

Proof From Lemma 4.2 follows

$$\mathcal{A} = \mathcal{Z}_\xi^{-1}\mathcal{W}_{\mathbf{a}}\mathcal{Z}_\xi \quad \text{in} \quad \mathbf{L}^p_{\rho,\sigma},$$

where $\mathbf{a}(t) = a - b\mathbf{i}\cot\pi(\xi-\mathbf{i}t) + \widehat{k}(\xi-\mathbf{i}t)$, $t \in \mathbb{R}$. Hence, $\mathbf{a}(\infty \pm 0) = a \mp b$, such that, due to (4.9),

$$\mathbf{a}_{p,\sigma}(t,x) = \begin{cases} \mathbf{a}(t) & : t \in \mathbb{R},\ x \in \overline{\mathbb{R}}, \\ a + \mathbf{i}b\cot\left(\frac{1+\sigma}{p} - \mathbf{i}x\right) & : t = \infty,\ x \in \overline{\mathbb{R}}. \end{cases}$$

Consequently, the curves $\Gamma_{\mathcal{A}}$ and $\Gamma_{\mathbf{a}_{p,\sigma}}$ coincide. Since furthermore

$$\mathbf{a} \in \mathbf{V}_1(\mathbb{R}) \cap \mathbf{PC}(\mathbb{R}) \subset \mathbf{PC}_{p,\sigma},$$

Lemma 4.6 delivers the assertion. □

Using Lemma 3.10, we can prove in the same way the following corollary.

Corollary 4.8 *Let $p \in (1,\infty)$, $\rho, \sigma \in (-1, p-1)$ and $a, b \in \mathbb{C}$. We assume that the function k belongs to $\widetilde{\mathbf{L}}^q_{q\xi-1}(\mathbb{R}^+)$ for some $q \in \left[1, \frac{p}{p-1}\right) \cap [1,2]$ and fulfils $\widehat{k}(\xi + \mathbf{i}\cdot) \in V_1(\mathbb{R})$. Moreover, for $t \in \mathbb{R}$, we set*

$$\mathbf{a}(t) := a - b\mathbf{i}\pi \cot(\xi - \mathbf{i}t) + \widehat{k}(\xi - \mathbf{i}t).$$

Then

$$\mathcal{Z}_\xi^{-1}\mathcal{W}_{\mathbf{a}}\mathcal{Z}_\xi u = (a\mathcal{I} + b\mathcal{S} + \mathcal{B}_k)u \quad \forall u \in \mathbf{L}^p_{\rho,0} \cap \mathbf{L}^p_{\rho,\sigma}$$

and $\mathcal{Z}_\xi^{-1}\mathcal{W}_{\mathbf{a}}\mathcal{Z}_\xi$ is Fredholm on the space $\mathbf{L}^p_{\rho,\sigma}$ if and only if the curve (4.11) does not run through the zero point. In this case, the Fredholm index of $\mathcal{Z}_\xi^{-1}\mathcal{W}_{\mathbf{a}}\mathcal{Z}_\xi : \mathbf{L}^p_{\rho,\sigma} \to \mathbf{L}^p_{\rho,\sigma}$ is equal to the negative winding number of this curve, where the orientation is given by its inherent parametrization.

For $\underline{a} = \left[a_j\right]_{j=1}^n \in \mathbf{PC}(\mathbb{R})^n$ and $\underline{b} = \left[b_j\right]_{j=1}^n \in \mathbf{PC}^n_{p,\sigma}$, now we consider operators of the form

$$\Psi_{\underline{a},\underline{b}} := \sum_{j=1}^n a_j \mathcal{W}^0_{b_j}. \tag{4.12}$$

Moreover, let

$$a_\pm(t) := \sum_{j=1}^n a_j(t) b_j(\pm\infty) \quad \text{and} \quad b_\pm(t) := \sum_{j=1}^n a_j(\pm\infty) b_j(t).$$

Definition 4.9 Let $n \in \mathbb{N}$ and $a_j \in \mathbf{PC}(\mathbb{R})$, $b_j \in \mathbf{PC}_{p,\sigma}$ for $j = 1, \cdots, n$. Under the assumptions

$$\inf\{|a_+(t)| : t \in \mathbb{R}\} > 0 \quad \text{and} \quad \inf\{|b_-(t)| : t \in \mathbb{R}\} > 0, \tag{4.13}$$

we define the functions $A_{p,\sigma} : \overline{\mathbb{R}} \times \overline{\mathbb{R}} \to \mathbb{C}$ and $B_{p,\sigma} : \overline{\mathbb{R}} \times \overline{\mathbb{R}} \to \mathbb{C}$ by

$$A_{p,\sigma}(t,x) := \frac{1}{2}\left[\frac{a_-(t+0)}{a_+(t+0)} + \frac{a_-(t-0)}{a_+(t-0)}\right] - \frac{\mathbf{i}}{2}\left[\frac{a_-(t+0)}{a_+(t+0)} - \frac{a_-(t-0)}{a_+(t-0)}\right]\cot\pi\left[\sigma_0(t) - \mathbf{i}x\right],\ t \in \mathbb{R},$$

$$A_{p,\sigma}(\pm\infty, x) := \frac{a_-(\pm\infty)}{a_+(\pm\infty)},$$

and

$$B_{p,\sigma}(t,x) := \frac{1}{2}\left[\frac{b_+(t+0)}{b_-(t+0)} + \frac{b_+(t-0)}{b_-(t-0)}\right] - \frac{\mathbf{i}}{2}\left[\frac{b_+(t+0)}{b_-(t+0)} - \frac{b_+(t-0)}{b_-(t-0)}\right]\cot\pi\left(\frac{1}{p} - \mathbf{i}x\right),\ t \in \mathbb{R},$$

$$B_{p,\sigma}(\pm\infty, x) := \frac{b_+(\pm\infty)}{b_-(\pm\infty)},$$

respectively, where

$$\sigma_0(t) := \begin{cases} \frac{1}{p} & : t \in \mathbb{R}\backslash\{0\} \\ \frac{1+\sigma}{p} & : \quad t = 0. \end{cases}$$

If the conditions

$$\inf\left\{\, |A_{p,\sigma}(t,x)| \,:\, t, \in \mathbb{R}, x \in \mathbb{R} \,\right\} > 0 \tag{4.14}$$

and

$$\inf\left\{\, |B_{p,\sigma}(t,x)| \,:\, t \in \mathbb{R}, x \in \mathbb{R} \,\right\} > 0 \tag{4.15}$$

are satisfied, we define $\kappa_{p,\sigma}$ as the increment of the function

$$\frac{1}{2\pi}\left[\arg A_{p,\sigma}(t,x) + \arg B_{p,\sigma}(t,x)\right],$$

where t runs over $\overline{\mathbb{R}}$, and at the points of discontinuity of $\frac{a_-}{a_+}$ or $\frac{b_+}{b_-}$, the variable x runs over $\overline{\mathbb{R}}$.

Lemma 4.10 ([1], Theorem 5.7, cf. also [8], Theorem 3.2) *Let $n \in \mathbb{N}$, $p \in (1,\infty)$, $\sigma \in (-1, p-1)$ and $a_j \in \mathbf{PC}(\mathbb{R})$, $b_j \in \mathbf{PC}_{p,\sigma}$ for $j = 1, \cdots, n$. Then $\Psi_{\underline{a},\underline{b}}$ is a Fredholm operator on $\mathbf{L}^p_\sigma(\mathbb{R})$ if and only if (4.13), (4.14), and (4.15) hold true. In this case, the Fredholm index of $\Psi_{\underline{a},\underline{b}} : \mathbf{L}^p_\sigma(\mathbb{R}) \to \mathbf{L}^p_\sigma(\mathbb{R})$ is equal to $-\kappa_{p,\sigma}$.*

A function $a : [0,1] \to \mathbb{C}$ is called piecewise continuous, if it is continuous at 0 and 1, if the one-sided limits $a(x \pm 0)$ exist for all $x \in (0,1)$, and if at least one of them coincides with $a(x)$. The set of all piecewise continuous functions $a : [0,1] \to \mathbb{C}$ having only a finite number of jumps is denoted by $\mathbf{PC}[0,1]$.

Corollary 4.11 *Let $p \in (1,\infty)$, $\rho, \sigma \in (-1, p-1)$ and $a, b \in \mathbf{PC}[0,1]$, and assume that $k \in \mathbf{C}(\mathbb{R}^+)$ satisfies condition* (A). *Then the integral operator $\mathcal{A} := a\mathcal{I} + b\mathcal{S} + \mathcal{B}_k$ is Fredholm on the space $\mathbf{L}^p_{\rho,\sigma}$ if and only if $a(x \pm 0) + b(x \pm 0) \neq 0$ for all $x \in (0,1)$, if $a(x) + b(x) \neq 0$ for $x \in \{0,1\}$, and if the closed curve*

$$\Gamma^d_{\mathcal{A}} := \widetilde{\Gamma}^d_0 \cup \Gamma^d_1 \cup \Gamma'^d_1 \cup \ldots \cup \Gamma^d_N \cup \Gamma'^d_N \cup \Gamma^d_{N+1} \cup \widetilde{\Gamma}^d_1$$

does not contain the point 0. Here N stands for the number of discontinuity points x_j, $j = 1, \ldots, N$, of the function

$$d(x) = \frac{a(x) - b(x)}{a(x) + b(x)}$$

chosen in such way that $x_0 := 0 < x_1 < \cdots < x_N < x_{N+1} := 1$. Using these x_j, the curves Γ^d_j, $j = 1, \ldots, N+1$ and Γ'^d_j, $j = 1, \ldots, N$ are given by

$$\Gamma^d_j := \big\{ d(x) : x_{j-1} < x < x_j \big\}$$

and

$$\Gamma'^d_j := \left\{ \frac{1}{2}\big[d(x_j + 0) + d(x_j - 0)\big] + \frac{\mathbf{i}}{2}\big[d(x_j + 0) - d(x_j - 0)\big] \cot \pi \left(\frac{1}{p} + \mathbf{i}t \right) : t \in \overline{\mathbb{R}} \right\},$$

respectively. The curves $\widetilde{\Gamma}^d_j$, $j \in \{0,1\}$, connecting the point 1 with one of the endpoints of Γ^d_1 and Γ^d_{N+1}, respectively, are given by the formulas

$$\widetilde{\Gamma}^d_0 := \left\{ \frac{a(0) - b(0)\mathbf{i}\cot \pi(\xi + \mathbf{i}t) + \widehat{k}(\xi + \mathbf{i}t)}{a(0) + b(0)} : t \in \overline{\mathbb{R}} \right\}$$

and

$$\widetilde{\Gamma}_1^d := \left\{ \frac{1}{2}\big[1+d(1)\big]+\frac{\mathbf{i}}{2}\big[1-d(1)\big]\cot\pi\left(\frac{1+\sigma}{p}+\mathbf{i}t\right) : t\in\overline{\mathbb{R}} \right\}.$$

In this case, the Fredholm index of $\mathcal{A}$ is equal to the winding number of the curve $\Gamma_{\mathcal{A}}$, where the orientation of $\Gamma_{\mathcal{A}}$ is due to the above given parametrization.

Proof Due to Lemma 4.2, we have the representation

$$\mathcal{A} = \mathcal{Z}_\xi^{-1}\big(\widetilde{a}\mathcal{I}+\widetilde{b}\mathcal{W}_{\mathbf{a}}+\mathcal{W}_{\mathbf{b}}\big)\mathcal{Z}_\xi \quad \text{in} \quad \mathbf{L}^p_{\rho,\sigma}$$

with $\mathbf{a}(t) = -\mathbf{i}\cot\pi(\xi-\mathbf{i}t)$, $\mathbf{b}(t) = \widehat{k}(\xi-\mathbf{i}t)$ and $\widetilde{a}(t) = a(e^{-t})$, $\widetilde{b}(t) = b(e^{-t})$. It is well known, that the Fredholm properties of the operator $\widetilde{a}\mathcal{I}+\widetilde{b}\mathcal{W}_{\mathbf{a}}+\mathcal{W}_{\mathbf{b}}$ in $\mathbf{L}^p_\sigma(\mathbb{R}^+)$ are equivalent to the Fredholm properties of

$$\widetilde{\mathcal{A}} = a_1\mathcal{I}+b_1\mathcal{W}^0_{\mathbf{a}}+\chi_+\mathcal{W}^0_{\mathbf{b}}+\chi_-\mathcal{I} \quad \text{in} \quad \mathbf{L}^p_\sigma(\mathbb{R}),$$

where

$$a_1(t) = \begin{cases} \widetilde{a}(t) : & t\in\mathbb{R}^+, \\ 0 & : \text{otherwise}, \end{cases} \qquad b_1(t) = \begin{cases} \widetilde{b}(t) : & t\in\mathbb{R}^+, \\ 0 & : \text{otherwise}, \end{cases}$$

$\chi_-(t) = 1-\chi_+(t)$, and

$$\chi_+(t) = \begin{cases} 1 : & t\in\mathbb{R}^+, \\ 0 : & \text{otherwise}. \end{cases}$$

Since $\widetilde{\mathcal{A}}$ is of the form (4.12), we can make use of Lemma 4.10. We have

$$a_+(t) = a_1(t)+b_1(t)+\chi_-(t), \quad a_-(t) = a_1(t)-b_1(t)+\chi_-(t)$$

and

$$b_+(t) = a(0)-b(0)\mathbf{i}\cot\pi(\xi-\mathbf{i}t)+\widehat{k}(\xi-\mathbf{i}t), \quad b_-(t) = 1.$$

Due to our assumptions, condition (4.13) is fulfilled, and hence we have, for $0 < t < \infty$,

$$A_{p,\sigma}(t,x) = \frac{1}{2}\left[\frac{\widetilde{a}(t+0)-\widetilde{b}(t+0)}{\widetilde{a}(t+0)+\widetilde{b}(t+0)}+\frac{\widetilde{a}(t-0)-\widetilde{b}(t-0)}{\widetilde{a}(t-0)+\widetilde{b}(t-0)}\right]$$
$$-\frac{\mathbf{i}}{2}\left[\frac{\widetilde{a}(t+0)-\widetilde{b}(t+0)}{\widetilde{a}(t+0)+\widetilde{b}(t+0)}-\frac{\widetilde{a}(t-0)-\widetilde{b}(t-0)}{\widetilde{a}(t-0)+\widetilde{b}(t-0)}\right]\cot\pi\left(\frac{1}{p}-\mathbf{i}x\right).$$

Moreover,

$$A_{p,\sigma}(0,x)=\frac{1}{2}\left[\frac{a(1)-b(1)}{a(1)+b(1)}+1\right]-\frac{\mathbf{i}}{2}\left[\frac{a(1)-b(1)}{a(1)+b(1)}-1\right]\cot\pi\left(\frac{1+\sigma}{p}-\mathbf{i}x\right),$$

$$A_{p,\sigma}(+\infty,x)=\frac{a(0)-b(0)}{a(0)+b(0)},$$

and, for $-\infty<t<0$,

$$A_{p,\sigma}(t,x)=A_{p,\sigma}(-\infty,x)=1\,.$$

Finally,

$$B_{p,\sigma}(\pm\infty,x)=a(0)\pm b(0)$$

and, for $t\in\mathbb{R}$,

$$B_{p,\sigma}(t,x)=a(0)-b(0)\mathbf{i}\cot\pi(\xi-\mathbf{i}t)+\widehat{k}(\xi-\mathbf{i}t)$$

Applying Lemma 4.10, we get, that $\widetilde{\mathcal{A}}$ is Fredholm on $\mathbf{L}^p_\sigma(\mathbb{R})$ if and only if

$$\inf\big\{\,|\widetilde{a}(t)+\widetilde{b}(t)|\,:\,t\in\mathbb{R}^+\,\big\}>0,$$

$$\inf\big\{\,|a(0)-b(0)\mathbf{i}\cot\pi(\xi-\mathbf{i}t)+\widehat{k}(\xi-\mathbf{i}t)|\,:\,t\in\mathbb{R}\,\big\}>0,$$

and

$$\inf\big\{\,|A_{p,\sigma}(t,x)|\,:\,t,\in\mathbb{R},x\in\mathbb{R}\,\big\}>0.$$

But this is obviously equivalent to our assumptions. In this case, the Fredholm index of $\widetilde{\mathcal{A}}$ is equal to $-\kappa_{p,\sigma}$, where $\kappa_{p,\sigma}$ is defined as the increment of the function

$$\frac{1}{2\pi}\left\{\arg A_{p,\sigma}(t,x)+\arg\frac{B_{p,\sigma}(t,x)}{a(0)+b(0)}\right\},$$

where t runs over $\overline{\mathbb{R}}$, and at the points of discontinuity of $(\widetilde{a}-\widetilde{b})/(\widetilde{a}+\widetilde{b})$, the variable x runs over $\overline{\mathbb{R}}$. But $\kappa_{p,\sigma}$ is equal to the negative winding number of the curve $\Gamma_{\mathcal{A}}$. This completes the proof. □

We define the operator $\mathcal{R}:\mathbf{L}^p_{\rho,\sigma}\to\mathbf{L}^p_{\sigma,\rho}$ by $(\mathcal{R}f)(x)=f(1-x)$. Moreover, we set

$$\mathcal{A}:=a\mathcal{I}+b\mathcal{S}+c_+\mathcal{B}_{k_+}+c_-\mathcal{R}\mathcal{B}_{k_-}\mathcal{R},\tag{4.16}$$

where $a,b,c_\pm\in\mathbf{L}^\infty(0,1)$ and $k_\pm\in\mathbf{C}(\mathbb{R}^+)$.

Theorem 4.12 *Let $p \in (1,\infty)$, $\sigma_\pm \in (-1, p-1)$ and $a, b, c_\pm \in \mathbf{PC}[0,1]$. If the functions $k_\pm \in \mathbf{C}(\mathbb{R}^+)$ satisfy condition* (A) *for $\xi = \xi_\pm := \frac{1+\sigma_\pm}{p}$, then the operator $\mathcal{A}$ defined in* (4.16) *is Fredholm on $\mathbf{L}^p_{\sigma_+,\sigma_-}$ if and only if $a(x \pm 0) - b(x \pm 0) \neq 0$ for all $x \in (0,1)$, if $a(x) - b(x) \neq 0$ for $x \in \{0,1\}$, and if the closed curve*

$$\Gamma^c_{\mathcal{A}} := \widetilde{\Gamma}^c_0 \cup \Gamma^c_1 \cup \Gamma'^c_1 \cup \ldots \cup \Gamma^c_N \cup \Gamma'^c_N \cup \Gamma^c_{N+1} \cup \widetilde{\Gamma}^c_1$$

does not contain the point 0. Here N stands for the number of discontinuity points x_j, $j = 1, \ldots, N$, of the function

$$c(x) = \frac{a(x) + b(x)}{a(x) - b(x)}$$

chosen in such way that $x_0 := 0 < x_1 < \cdots < x_N < x_{N+1} := 1$. Using these x_j, the curves Γ^c_j, $j = 1, \ldots, N+1$ and Γ'^c_j, $j = 1, \ldots, N$ are given by

$$\Gamma^c_j := \big\{c(x) : x_{j-1} < x < x_j\big\}$$

and

$$\Gamma'^c_j := \Big\{ \frac{1}{2}\big[c(x_j + 0) + c(x_j - 0)\big] - \frac{\mathbf{i}}{2}\big[c(x_j + 0) - c(x_j - 0)\big] \cot \pi \left(\frac{1}{p} - \mathbf{i}t\right) : t \in \overline{\mathbb{R}} \Big\}.$$

The curves $\widetilde{\Gamma}_j$, $j \in \{0,1\}$ connecting the point 1 with one of the endpoints of Γ^c_1 and Γ^c_{N+1}, respectively, are given by the formulas

$$\widetilde{\Gamma}^c_0 := \left\{ \frac{a(0) - b(0)\mathbf{i}\cot\pi(\xi_+ - \mathbf{i}t) + c{+}(0)\widehat{k}_+(\xi_+ - \mathbf{i}t)}{a(0) - b(0)} : t \in \overline{\mathbb{R}} \right\}$$

and

$$\widetilde{\Gamma}^c_1 := \left\{ \frac{a(1) + b(1)\mathbf{i}\cot\pi(\xi_- - \mathbf{i}t) + c_-(1)\widehat{k}_-(\xi_- - \mathbf{i}t)}{a(1) - b(1)} : t \in \overline{\mathbb{R}} \right\}.$$

In this case, the Fredholm index of $\mathcal{A}$ is equal to the negative winding number of the curve $\Gamma^c_{\mathcal{A}}$, where the orientation of $\Gamma^c_{\mathcal{A}}$ is due to the above given parametrization.

Proof At first, we consider the case $c_+ = 1$ and $c_- = 0$. Setting $\rho = \sigma_+$, $\sigma = \sigma_-$, $\xi = \xi_+ = \frac{1+\rho}{p}$, and $\xi_- = \frac{1+\sigma}{p}$. Having regard to Corollary 4.11, we have to show $0 \notin \Gamma^c_{\mathcal{A}}$ is equivalent to $0 \notin \Gamma^d_{\mathcal{A}}$ and that in this case $\operatorname{wind}\big(\Gamma^c_{\mathcal{A}}\big) = -\operatorname{wind}\big(\Gamma^d_{\mathcal{A}}\big)$.

But, this is seen by the relations

$$\Gamma_j^d = \left\{ \frac{1}{c(x)} : x_{j-1} < x < x_j \right\} = \left\{ z^{-1} : z \in \Gamma_j^c \right\},$$

$$\Gamma_j'^d = d(x_j + 0)d(x_j - 0) \left\{ \frac{1}{2}\left[c(x_j + 0) + c(x_j - 0)\right] \right.$$

$$\left. - \frac{\mathbf{i}}{2}\left[c(x_j + 0) - c(x_j - 0)\right] \cot \pi \left(\frac{1}{p} + \mathbf{i}t \right) : t \in \overline{\mathbb{R}} \right\},$$

$$\widetilde{\Gamma}_0^d = \frac{a(0) - b(0)}{a(0) + b(0)} \left\{ \frac{a(0 - b(0)\mathbf{i} \cot \pi(\xi + \mathbf{i}t) + \widehat{k}(\xi + \mathbf{i}t)}{a(0) - b(0)} : t \in \overline{\mathbb{R}} \right\},$$

$$\widetilde{\Gamma}_1^d = d(1) \left\{ \frac{1}{2}[1 + c(1)] - \frac{\mathbf{i}}{2}[1 - c(1)] \cot \pi \left(\frac{1 + \sigma}{p} + \mathbf{i}t \right) : t \in \overline{\mathbb{R}} \right\},$$

if one compares the increments of the argument along the respective pieces of the curves $\Gamma_{\mathcal{A}}^c$ and $\Gamma_{\mathcal{A}}^d$. □

5 Regularity Properties of Solutions of Eq. (1.1)

This section deals with some specific properties of the Mellin operators and the smoothness of the solutions of the equation

$$au + b\mathcal{S}u + c\mathcal{B}_k u = f \quad \text{in} \quad \mathbf{L}_{p,\sigma}^p.$$

In addition, their asymptotic behaviour near the end point 1 is investigated. Let $n \in \mathbb{N}_0$. For $k \in \mathbf{C}^n(\mathbb{R}^+)$ and $f : (0, 1) \to \mathbb{C}$ we define the operators $\partial_n \mathcal{B}_k$ by

$$(\partial_n \mathcal{B}_k f)(x) := \int_0^1 k^{(n)}\left(\frac{x}{y}\right) \frac{f(y)}{y^{n+1}}\, dy.$$

Lemma 5.1 *Let $p \in (1, \infty)$, $\sigma < p-1$, $\rho \in \mathbb{R}$, and $n \in \mathbb{N}_0$ as well as $k \in \mathbf{C}^n(\mathbb{R}^+)$. We assume that there is a $\beta > \frac{1+\rho}{p}$ such that the functions*

$$\widetilde{k}_\ell := [a, 1] \times [0, 1] \to \mathbb{C}, \quad (x, y) \mapsto y^{-(\beta+\ell)} k^{(\ell)}\left(\frac{x}{y}\right) \tag{5.1}$$

are continuous for all $a \in (0,1)$ and all $\ell \in \{0,1,\dots,n\}$. Then, for every $f \in \mathbf{L}^p_{\rho,\sigma}$, the function $\mathcal{B}_k f$ is n-times continuously differentiable on $(0,1]$, where

$$\frac{d^\ell}{dx^\ell}(\mathcal{B}_k f)(x) = (\partial_\ell \mathcal{B}_k f)(x), \quad x \in (0,1] \tag{5.2}$$

holds true.

Proof Let $\varepsilon > 0$. Then, there exists a $\delta = \delta(\varepsilon) > 0$, such that

$$\begin{aligned}
&\left|(\partial_\ell \mathcal{B}_k f)(x) - (\partial_\ell \mathcal{B}_k f)(x')\right| \\
&\quad \le \int_0^1 \left|\widetilde{k}_\ell(x,y) - \widetilde{k}_\ell(x',y)\right| y^{\beta-1} |f(y)|\, dy \\
&\quad \le \varepsilon \int_0^1 \upsilon^{\beta-1-\frac{\rho}{p},-\frac{\sigma}{p}}(y)\, \upsilon^{\frac{\rho}{p},\frac{\sigma}{p}}(y)\, |f(y)|\, dy \\
&\quad \le \varepsilon \left(\int_0^1 \upsilon^{\left(\beta-1-\frac{\rho}{p}\right)\frac{p}{p-1},-\frac{\sigma}{p-1}}(y)\, dy\right)^{\frac{p-1}{p}} \|f\|_{\rho,\sigma,p} \\
&\quad \le \text{const} \cdot \varepsilon
\end{aligned}$$

for all $x, x' \in [a,1]$, which satisfy $|x - x'| < \delta$. Here we took into account that $\left(\beta - 1 - \frac{\rho}{p}\right)\frac{p}{p-1} > -1$ is equivalent to $\beta > \frac{1+\rho}{p}$ and that $\sigma < p-1$. Hence $\partial_\ell \mathcal{B}_k f$ is continuous on $(0,1]$. Analogously, one can show that, for $x \in [a,1]$, the relations

$$\left|\partial_\ell \mathcal{B}_k f(x)\right| \le \text{const}\, \|f\|_{\rho,\sigma,p}\,, \tag{5.3}$$

are true, where the constant does only depend on $a \in (0,1)$. The differentiability of $\mathcal{B}_k f$ follows now from

$$\begin{aligned}
&\int_c^x (\partial_1 \mathcal{B}_k f)(\xi)\, d\xi = \int_c^x \int_0^1 k'\left(\frac{\xi}{y}\right)\frac{f(y)}{y^2}\, dy\, d\xi \\
&= \lim_{\varepsilon\to 0} \int_c^x \int_\varepsilon^1 k'\left(\frac{\xi}{y}\right)\frac{f(y)}{y^2}\, dy\, d\xi + \lim_{\varepsilon\to 0} \int_c^x \int_0^\varepsilon k'\left(\frac{\xi}{y}\right)\frac{f(y)}{y^2}\, dy\, d\xi \\
&= \lim_{\varepsilon\to 0} \int_\varepsilon^1 \int_c^x k'\left(\frac{\xi}{y}\right)\frac{f(y)}{y^2}\, d\xi\, dy + \lim_{\varepsilon\to 0} \int_c^x \int_0^\varepsilon k'\left(\frac{\xi}{y}\right)\frac{f(y)}{y^2}\, dy\, d\xi \\
&= \lim_{\varepsilon\to 0} \int_\varepsilon^1 \left[k\left(\frac{x}{y}\right)\frac{f(y)}{y} - k\left(\frac{c}{y}\right)\frac{f(y)}{y}\right] dy + \lim_{\varepsilon\to 0} \int_c^x \int_0^\varepsilon k'\left(\frac{\xi}{y}\right)\frac{f(y)}{y^2}\, dy\, d\xi,
\end{aligned}$$

where $x, c \in (0,1]$. With regard to (5.3), we can apply Lebesgue's dominated convergence theorem and get

$$\int_c^x (\partial_1 \mathcal{B}_k f)(\xi)\, d\xi = \int_0^1 \left[k\left(\frac{x}{y}\right) \frac{f(y)}{y} - k\left(\frac{c}{y}\right) \frac{f(y)}{y} \right] dy = (\mathcal{B}_k f)(x) + \text{const.}$$

Hence, $\frac{d}{dx}(\mathcal{B}_k f)(x) = (\partial_1 \mathcal{B}_k f)(x)$, $x \in (0,1]$. The general case follows now by induction. □

Remark 5.2 Note, that the function $\widetilde{k}_\ell$ defined in (5.1) is continuous for all $a \in (0,1)$ if the limit $\lim\limits_{t\to\infty} t^{\beta+\ell} k^{(\ell)}(t)$ exists and is finite.

Let $\psi, \zeta \geq 0$. By $\mathbf{BC}_{\psi,\zeta} = \mathbf{BC}_{\psi,\zeta}(0,1)$ we denote the set of all continuous functions $f : (0,1) \to \mathbb{C}$, for which the function $v^{\psi,\zeta} f$ is bounded on $(0,1)$. If we introduce the norm

$$\|f\|_{\psi,\zeta,\infty} = \sup\left\{ v^{\psi,\zeta}(x) |f(x)| : 0 < x < 1 \right\},$$

then $\mathbf{BC}_{\psi,\zeta}$ becomes a Banach space. Moreover by $\mathbf{C}_{\psi,\zeta} = \mathbf{C}_{\psi,\zeta}(0,1)$ we denote the set of all continuous functions $f : (0,1) \to \mathbb{C}$, for which the limits

$$\lim_{x\to 0} x^\psi f(x) \quad \text{and} \quad \lim_{x\to 1} (1-x)^\zeta f(x)$$

exist and if these limits are equal to zero if $\psi > 0$ or $\zeta > 0$, respectively. The space $\mathbf{C}_{\psi,\zeta}$ is a closed subspace of $\mathbf{BC}_{\psi,\zeta}$ and, consequently, also a Banach space.

Lemma 5.3 *Let $p \in (1,\infty)$, $\frac{1}{p} + \frac{1}{q} = 1$, $\sigma < p - 1$, and $\rho \in \mathbb{R}$, as well as $k \in \mathbf{C}(\mathbb{R}^+)$. Furthermore, we assume that there is a $\beta > \frac{1+\rho}{p}$ such that the function*

$$\widetilde{k} := [a,1] \times [0,1] \to \mathbb{C}, \quad (x,y) \mapsto y^{-\beta} k\left(\frac{x}{y}\right)$$

is continuous for all $a \in (0,1)$. If

$$\int_0^\infty |k(t)|^q t^{\left(\frac{\rho}{p}+1\right)q-2}\, dt < \infty$$

and if there are numbers $\chi \in \mathbb{R}$, $t_0 > 0$, and $c_0 > 0$ such that

$$|k(t)| \leq c_0 t^{-\chi}, \quad 0 < t < t_0, \tag{5.4}$$

then

$$\mathcal{B}_k \in \begin{cases} \mathcal{L}\left(\mathbf{L}^p_{\rho,\sigma}, \mathbf{BC}_{\psi,0}\right) : \max\left\{0, \frac{1+\rho}{p}, \chi\right\} \leq \psi, \\ \mathcal{L}\left(\mathbf{L}^p_{\rho,\sigma}, \mathbf{C}_{\psi,0}\right) : \max\left\{0, \frac{1+\rho}{p}, \chi\right\} < \psi. \end{cases}$$

Proof First we note that, due to Lemma 5.1, for $f \in \mathbf{L}^p_{\rho,\sigma}$, the function $(\mathcal{B}_k f)(x)$ is continuous on $(0, 1]$. Moreover, for $x \in (0, 1)$, using Hölder's inequality we can estimate

$$\begin{aligned}\left|v^{\psi,0}(x)\,(\mathcal{B}_k f)(x)\right| &= \left|x^{\psi}\int_0^1 k\left(\frac{x}{y}\right)\frac{f(y)\,dy}{y}\right| \\ &\le x^{\psi}\left(\int_0^1 \left|k\left(\frac{x}{y}\right)\right|^q \frac{v^{-\frac{\rho q}{p},-\frac{\sigma q}{p}}(y)\,dy}{y^q}\right)^{\frac{1}{q}} \|f\|_{\rho,\sigma,p} \\ &=: N(x)\,\|f\|_{\rho,\sigma,p}\ ,\end{aligned}$$

where, with the help of the substitution $t = \frac{x}{y}$,

$$\begin{aligned}N(x)^q &= x^{\psi q+1}\int_x^{\infty} |k(t)|^q \left(\frac{x}{t}\right)^{-\frac{\rho q}{p}-q}\left(1-\frac{x}{t}\right)^{-\frac{\sigma q}{p}}\frac{dt}{t^2} \\ &= x^{\psi q+1}\left(\int_x^{2x}+\int_{2x}^{\infty}\right)|k(t)|^q \left(\frac{x}{t}\right)^{-\frac{\rho q}{p}-q}\left(1-\frac{x}{t}\right)^{-\frac{\sigma q}{p}}\frac{dt}{t^2} \\ &=: N_1(x)+N_2(x)\,.\end{aligned}$$

Setting $M_0 = \max\left\{|k(t)| : \frac{t_0}{2} \le t \le 2\right\}$ and taking (5.4) into account, we get

$$N_1(x) \le x^{\psi q+1}\begin{cases} c_0^q \displaystyle\int_x^{2x} t^{-\chi q-2}\left(\frac{x}{t}\right)^{-\frac{\rho q}{p}-q}\left(1-\frac{x}{t}\right)^{-\frac{\sigma q}{p}} dt & : 2x < t_0\,, \\ M_0^q \displaystyle\int_x^{2x} t^{-2}\left(\frac{x}{t}\right)^{-\frac{\rho q}{p}-q}\left(1-\frac{x}{t}\right)^{-\frac{\sigma q}{p}} dt & : t_0 \le 2x\,,\end{cases}$$

$$\overset{t=xs}{=}\begin{cases} c_0^q x^{(\psi-\chi)q}\displaystyle\int_1^2 s^{\frac{\rho q}{p}+q-2+\frac{\sigma q}{p}-\chi q}(s-1)^{-\frac{\sigma q}{p}} ds & : 2x < t_0\,, \\ M_0^q x^{\psi q}\displaystyle\int_1^2 s^{\frac{\rho q}{p}+q-2+\frac{\sigma q}{p}}(s-1)^{-\frac{\sigma q}{p}} ds & : t_0 \le 2x\,,\end{cases}$$

and

$$N_2(x) \le \max\left\{1, 2^{\frac{\sigma q}{p}}\right\} x^{\psi q+1-\frac{\rho q}{p}-q}\int_0^{\infty} |k(t)|^q t^{\left(\frac{\rho}{p}+1\right)q-2}\,dt\,.$$

Since $\frac{\sigma q}{p} = \frac{\sigma}{p-1} < 1$ and since $\psi q + 1 - \frac{\rho q}{p} - q \ge (>)\ 0$ is equivalent to $\psi \ge (>)\ \frac{1+\rho}{p}$, we get the assertion. □

Writing

$$(\partial_\ell \mathcal{B}_k f)(x) = x^{-\ell} \int_0^1 k_\ell\left(\frac{x}{y}\right) \frac{f(y)\,dy}{y}$$

with $k_\ell(t) = k^{(\ell)}(t)t^\ell$, from Lemma 5.3 we get

Corollary 5.4 *Let $p \in (1,\infty)$, $\frac{1}{p} + \frac{1}{q} = 1$, $\sigma < p-1$, and $\rho \in \mathbb{R}$, $\ell \in \mathbb{N}_0$, as well as $k \in \mathbf{C}^\ell(\mathbb{R}^+)$. Furthermore, we assume that there is a $\beta > \frac{1+\rho}{p}$ such that the function $\widetilde{k}_\ell$ defined in (5.1) is continuous for all $a \in (0,1)$. If*

$$\int_0^\infty |k^{(\ell)}(t)|^q t^{\left(\frac{\rho}{p}+1\right)q+\ell q-2}\,dt < \infty$$

and if there are numbers $\chi \in \mathbb{R}$, $t_0 > 0$, and $c_0 > 0$ such that

$$|k^{(\ell)}(t)| \le c_0 t^{-\chi-\ell}, \quad 0 < t < t_0,$$

then

$$\partial_\ell \mathcal{B}_k \in \begin{cases} \mathcal{L}\left(\mathbf{L}^p_{\rho,\sigma}, \mathbf{BC}_{\psi,0}\right) : \ell + \max\left\{0, \frac{1+\rho}{p}, \chi\right\} \le \psi, \\ \mathcal{L}\left(\mathbf{L}^p_{\rho,\sigma}, \mathbf{C}_{\psi,0}\right) \ : \ell + \max\left\{0, \frac{1+\rho}{p}, \chi\right\} < \psi. \end{cases}$$

Corollary 5.5 *Let $p \in (1,\infty)$, $\sigma < p-1$, $\rho \in \mathbb{R}$, and $\ell \in \mathbb{N}_0$, $k \in \mathbf{C}^\ell(\mathbb{R}^+)$. If there are numbers $\alpha, \beta \in \mathbb{R}$ such that $\alpha < \frac{1+\rho}{p} < \beta$, that $t^{\alpha+\ell}k^{(\ell)}(t)$ is bounded for $t \to +0$, and that the finite limit $\lim\limits_{t\to\infty} t^{\beta+\ell}k^{(\ell)}(t)$ exists, then*

$$\partial_\ell \mathcal{B}_k \in \begin{cases} \mathcal{L}\left(\mathbf{L}^p_{\rho,\sigma}, \mathbf{BC}_{\psi,0}\right) : \ell + \max\left\{0, \frac{1+\rho}{p}\right\} \le \psi, \\ \mathcal{L}\left(\mathbf{L}^p_{\rho,\sigma}, \mathbf{C}_{\psi,0}\right) \ : \ell + \max\left\{0, \frac{1+\rho}{p}\right\} < \psi. \end{cases}$$

Proof Taking into account Remark 5.2, we can apply Corollary 5.4 with $\chi = \alpha$, since $\left(\frac{1}{p} + \frac{1}{q} = 1\right)$

$$\int_0^\infty |k^{(\ell)}(t)|^q t^{\left(\frac{\rho}{p}+1\right)q+\ell q-2}\,dt$$

$$\le c_1 \int_0^1 t^{\left(\frac{\rho}{p}+1\right)q-\alpha q-2}\,dt + c_2 \int_1^\infty t^{\left(\frac{\rho}{p}+1\right)q-\beta q-2}\,dt$$

and

$$\left(\frac{\rho}{p}+1\right)q - \alpha q - 2 = \left(\frac{1+\rho}{p} - \alpha - \frac{1}{q}\right)q > -1$$

as well as

$$\left(\frac{\rho}{p}+1\right)q-\beta q-2=\left(\frac{1+\rho}{p}-\beta-\frac{1}{q}\right)q<-1\,.$$ □

Corollary 5.6 *Let $a, c, f \in \mathbf{C}^n(0,1]$ and let the assumptions of Lemma 5.1 be fulfilled. If the operator equation $au + c\mathcal{B}_k u = f$ has a solution $u \in \mathbf{L}^p_{\rho,\sigma}$ and if $a(x) \neq 0$ for all $x \in (0,1]$, then $u \in \mathbf{C}^n(0,1]$.*

Proof With the help of Lemma 5.1, we get $u = a^{-1}(f - c\mathcal{B}_k)u \in \mathbf{C}^n(0,1]$. □

Corollary 5.7 *Let $a, c, f \in \mathbf{C}^n[0,1]$ and $a(x) \neq 0$ for all $x \in [0,1]$. If the conditions of Corollary 5.5 are satisfied and if the equation $au + c\mathcal{B}_k u = f$ has a solution $u \in \mathbf{L}^p_{\rho,\sigma}$, then*

$$u^{(\ell)} \in \mathbf{C}_{\frac{1+\rho}{p}+\ell+\varepsilon,0} \quad \text{for} \quad \ell \in \{0,\dots,n\} \quad \text{and all} \quad \varepsilon > 0.$$

Proof Due to the assumptions, Corollary 5.5 in combination with (5.2) deliver

$$u^{(\ell)} = \left[a^{-1}(f - c\mathcal{B}_k)u\right]^{(\ell)} \in \mathbf{C}_{\frac{1+\rho}{p}+\ell+\varepsilon,0}$$

for $\ell \in \{0,\dots,n\}$ and all $\varepsilon > 0$. □

Let $0 < \gamma < 1$, $m \in \mathbb{N}_0$, and $-\infty < a < b < \infty$. We denote by $\mathbf{C}^{m,\gamma}[a,b]$ the set of all m times differentiable functions $f : [a,b] \to \mathbb{C}$, for which

$$\sup\left\{\frac{|f^{(m)}(x) - f^{(m)}(y)|}{|x-y|^\gamma} : a \le x, y \le b,\ x \neq y\right\}$$

is finite. For $c \in [a,b]$, $m \in \mathbb{N}_0$, and $0 < \gamma < 1$, we set

$$\mathbf{C}^{m,\gamma}_{[a,b]}(c) = \left\{f : \exists \varepsilon > 0 \text{ with } f \in \mathbf{C}^{m,\gamma}\big([c-\varepsilon, c+\varepsilon] \cap [a,b]\big)\right\}$$

Moreover , let

$$\mathbf{C}^{m,0}[a,b] = \bigcup_{0<\gamma<1} \mathbf{C}^{m,\gamma}[a,b] \quad \text{and} \quad \mathbf{C}^{m,0}_{[a,b]}(c) = \bigcup_{0<\gamma<1} \mathbf{C}^{m,\gamma}_{[a,b]}(c)\,.$$

For $\gamma \in [0,\infty)$ and $-\infty < a < b < \infty$, the function class $\mathbf{H}^\gamma(a,b)$ is defined as the set of all functions $f : (a,b) \to \mathbb{C}$ belonging to $\mathbf{C}^{[\gamma],\gamma-[\gamma]}[c,d]$ for all intervals $[c,d] \subset (a,b)$, where $[\gamma]$ is the integer, which fulfills $\gamma - 1 < [\gamma] \le \gamma$. In the same manner we define the classes $\mathbf{H}^\gamma(a,b]$, $\mathbf{H}^\gamma[a,b)$, and $\mathbf{H}^\gamma[a,b]$. Of course, $\mathbf{H}^\gamma[a,b] = \mathbf{C}^{[\gamma],\gamma-[\gamma]}[a,b]$.

Lemma 5.8 *Let $p \in (1,\infty)$, $\sigma < p-1$, $n \in \mathbb{N}_0$ and $\rho, \beta \in \mathbb{R}$ with $\frac{1+\rho}{p} < \beta$. For $k \in \mathbf{C}(\mathbb{R}^+)$, assume that $t^\beta k(t)$ is bounded for $t \to \infty$ (i.e.,*

$m_a := \sup\left\{t^\beta k(t) : a \le t < \infty\right\} < \infty$ *for all* $a > 0$) *and that*

$$M_a := \sup\left\{\frac{\left|s^\beta k(s) - t^\beta k(t)\right|}{|s-t|^\gamma} : s, t \in [a, \infty),\ s \ne t\right\} < \infty$$

for some $\gamma \in (0,1)$ *satisfying* $\frac{1+\rho}{p} + \gamma < \beta$ *and for all* $a > 0$. *Then* $\mathcal{B}_k f \in \mathbf{H}^\gamma(0,1]$ *for all* $f \in \mathbf{L}^p_{\rho,\sigma}$.

Proof Let $f \in \mathbf{L}^p_{\rho,\sigma}$ and $0 < a \le x_1 < x_2 \le 1$. Then

$$\begin{aligned}
&\frac{|(\mathcal{B}_k f)(x_1) - (\mathcal{B}_k f)(x_2)|}{|x_1 - x_2|^\gamma} \\
&\quad\le \int_0^1 \frac{\left|k\left(\frac{x_1}{y}\right) - k\left(\frac{x_2}{y}\right)\right| \left(\frac{x_1}{y}\right)^\beta \left(\frac{x_1}{y}\right)^{-\beta}}{|\frac{x_1}{y} - \frac{x_2}{y}|^\gamma} \frac{|f(y)|\,dy}{y^{1+\gamma}} \\
&\quad\le \int_0^1 \frac{\left|\left(\frac{x_1}{y}\right)^\beta k\left(\frac{x_1}{y}\right) - \left(\frac{x_2}{y}\right)^\beta k\left(\frac{x_2}{y}\right)\right| \left(\frac{x_1}{y}\right)^{-\beta}}{|\frac{x_1}{y} - \frac{x_2}{y}|^\gamma} \frac{|f(y)|}{y^{1+\gamma}}\,dy \\
&\qquad\quad + \int_0^1 \frac{\left|k\left(\frac{x_2}{y}\right)\right| \left[\left(\frac{x_2}{y}\right)^\beta - \left(\frac{x_1}{y}\right)^\beta\right]}{|\frac{x_1}{y} - \frac{x_2}{y}|^\gamma \left(\frac{x_1}{y}\right)^\beta} \frac{|f(y)|\,dy}{y^{1+\gamma}} \\
&\quad\le x_1^{-\beta} M_a \int_0^1 \frac{|f(y)|\,dy}{y^{1+\gamma-\beta}} \\
&\qquad\quad + \frac{\left(\frac{x_2}{x_1} - 1\right)^{\beta-\gamma}}{x_2^\beta x_1^\gamma} \int_0^1 \left(\frac{x_2}{y}\right)^\beta \left|k\left(\frac{x_2}{y}\right)\right| \frac{|f(y)|\,dy}{y^{1-\beta}} \\
&\quad\le a^{-\beta} M_a \int_0^1 \frac{|f(y)|\,dy}{y^{1+\gamma-\beta}} + a^{-\beta-\gamma}\left(a^{-1} - 1\right)^{\beta-\gamma} m_a \int_0^1 \frac{|f(y)|\,dy}{y^{1-\beta}} \\
&\quad\le \text{const}\left[\left(\int_0^1 \upsilon^{\left(-1-\gamma+\beta-\frac{\rho}{p}\right)q, -\frac{\sigma}{p-1}}(y)\,dy\right)^{\frac{1}{q}}\right. \\
&\qquad\quad \left. + \left(\int_0^1 \upsilon^{\left(-1+\beta-\frac{\rho}{p}\right)q, -\frac{\sigma}{p-1}}(y)\,dy\right)^{\frac{1}{q}}\right] \|f\|_{\rho,\sigma,p} \\
&\quad\le \text{const}\,\|f\|_{\rho,\sigma,p}\ ,
\end{aligned}$$

where the constant does not depend on x_1 and x_2 and where we took into account that $\left(-1-\gamma+\beta-\frac{\rho}{p}\right)q > -1$ is equivalent to $\frac{1+\rho}{p}+\gamma<\beta$. □

Lemma 5.9 ([7], Theorem in §19) *Let $0 \le a < b \le 1$ and $v \in \mathbf{C}^{0,\gamma}[a,b]$ for some $\gamma \in [0,1)$. Then we have*

$$\mathcal{S}\chi_{[a,b]}v \in \mathbf{H}^{\gamma}(a,b),$$

where $\chi_{[a,b]}$ is the characteristic function of the interval $[a,b]$. Moreover, if $v(a) = 0$ or $v(b) = 0$ are satisfied, then we even get

$$\mathcal{S}\chi_{[a,b]}v \in \mathbf{H}^{\gamma}[a,b) \quad \text{and} \quad \mathcal{S}\chi_{[a,b]}v \in \mathbf{H}^{\gamma}(a,b],$$

respectively.

Corollary 5.10 *Let $0 \le a < b \le 1$ and $v \in \mathbf{H}^{\gamma}(a,b)$ for some $\gamma \in [0,\infty)$ as well as $\chi_{[a,b]}v \in \mathbf{L}^1_{0,0}$. Then $\mathcal{S}\chi_{[a,b]}v \in \mathbf{H}^{\gamma}(a,b)$.*

Proof Without loss of generality, we can assume that $a = 0$ and $b = 1$. Let $[c,d] \subset (0,1)$ and choose $\varepsilon > 0$ such that $\varepsilon < c$ and $d < 1-\varepsilon$. Write

$$\mathcal{S}v = I_1 + I_2 + I_3$$

with

$$I_1(x) = \frac{1}{\pi \mathbf{i}} \int_0^{\varepsilon} \frac{v(y)\,dy}{y-x}, \quad I_2(x) = \frac{1}{\pi \mathbf{i}} \int_{\varepsilon}^{1-\varepsilon} \frac{v(y)\,dy}{y-x},$$

and

$$I_3(x) = \frac{1}{\pi \mathbf{i}} \int_{1-\varepsilon}^{1} \frac{v(y)\,dy}{y-x}.$$

Then $I_1, I_3 \in \mathbf{C}^{\infty}[c,d]$.

Now we consider I_2. If $\gamma \in [0,1)$, then, by our assumptions, $v \in \mathbf{C}^{0,\gamma}[\varepsilon, 1-\varepsilon]$, and Lemma 5.9 delivers $I_2 \in \mathbf{C}^{0,\gamma}[c,d]$. Assume that

$$I_2 \in \mathbf{C}^{[\gamma],\gamma-[\gamma]}[c,d] \quad \text{if} \quad \gamma \in [0,n) \tag{5.5}$$

for some $n \in \mathbb{N}$. We show that then $I_2 \in \mathbf{C}^{[\gamma],\gamma-[\gamma]}[c,d]$ also holds if $\gamma \in [n, n+1)$. Indeed, in that case $v' \in \mathbf{H}^{\gamma-1}(0,1)$ and, by [6, Lemma 6.1],

$$I_2'(x) = \frac{v(\varepsilon)}{\varepsilon - x} - \frac{v(1-\varepsilon)}{1-\varepsilon-x} - \int_\varepsilon^{1-\varepsilon} \frac{v'(y)\,dy}{y-x}\,, \quad \varepsilon < x < 1-\varepsilon\,.$$

By our assumption (5.5),

$$\int_\varepsilon^{1-\varepsilon} \frac{v'(y)\,dy}{y-\cdot} \in \mathbf{C}^{[\gamma]-1,\gamma-[\gamma]}[c,d]\,,$$

which implies $I_2 \in \mathbf{C}^{[\gamma],\gamma-[\gamma]}[c,d]$. The corollary is proved. □

The following corollary can be proved analogously.

Corollary 5.11 *Let $0 \le a < b \le 1$, $c \in \{a,b\}$ and $v \in \mathbf{H}^\gamma\big((a,b) \cup \{c\}\big)$ with $v(c) = v'(c) = \ldots = v^{[\gamma]}(c) = 0$ for some $\gamma \in [0,\infty)$ as well as $\chi_{[a,b]} v \in \mathbf{L}^1_{0,0}$. Then*

$$\mathcal{S}\chi_{[a,b]} v \in \mathbf{H}^\gamma\big((a,b) \cup \{c\}\big).$$

Lemma 5.12 ([7], §29, $3^o, 4^o$, §32) *Let $0 \le \alpha < 1$, $\beta \in \mathbb{R}$, and $\mu = \alpha + \mathbf{i}\beta$ as well as $u \in \mathbf{H}^0(0,1) \cap \mathbf{L}^1_{0,0}$. Moreover, assume that there are $\varepsilon, \delta \in (0,1)$ such that, for $1-\varepsilon < x < 1$,*

$$u(x) = \frac{u^*(x)}{(1-x)^\mu} \quad \text{and} \quad u^* \in \mathbf{C}^{0,\delta}_{[0,1]}(1)\,,$$

where, for $x \in (0,1)$,

$$(1-x)^\mu = (1-x)^\alpha e^{\mathbf{i}\beta \ln(1-x)}\,.$$

Then, for $x \in (1-\varepsilon, 1)$, we have

$$(\mathcal{S}u)(x) = \begin{cases} \dfrac{u^*(1)}{\pi\mathbf{i}} \ln(1-x) + u^{**}(x) : \mu = 0\,, \\ \dfrac{\mathbf{i}\cot(\pi\mu)}{(1-x)^\mu} u^*(1) + u^{**}(x) : \mu \ne 0, \end{cases}$$

where, in case $\alpha = 0$,

$$u^{**} \in \mathbf{C}^{0,\delta}_{[0,1]}(1)$$

and, in case $0 < \alpha < 1$,

$$v^{0,\alpha_0} u^{**} \in \mathbf{C}^{0,\lambda}_{[0,1]}(1)$$

with $\left(v^{0,\alpha_0} u^{**}\right)(1) = 0$ *for*

$$\max\{0, \alpha - \delta\} < \alpha_0 < \alpha \quad \textit{and} \quad 0 < \lambda < \alpha_0 - \max\{0, \alpha - \delta\}\,.$$

Remark 5.13 If (4.3) holds for some $\kappa \in \{-1, 0, 1\}$, then

$$v^{-\nu,\nu+\kappa} \mathbf{L}^p_{\rho,\sigma} \subset \mathbf{L}^1_{0,0}.$$

Proof This follows immediately from

$$\int_0^1 |(v^{-\nu,\nu+\kappa} f)(x)|\, dx \le \left(\int_0^1 |f(x)|^p v^{\rho,\sigma}(x)\, dx\right)^{\frac{1}{p}} \times \left(\int_0^1 v^{\left(-\Re\nu-\frac{\rho}{p}\right)\frac{p}{p-1},\left(\Re\nu+\kappa-\frac{\sigma}{p}\right)\frac{p}{p-1}}(x)\, dx\right)^{\frac{p-1}{p}} < \infty$$

for $f \in \mathbf{L}^p_{\rho,\sigma}$. □

Remark 5.14 ([7], §7) For $\alpha > 0$ and $\beta \in \mathbb{R}$, the function $x \mapsto (1-x)^{\alpha+\mathbf{i}\beta}$ belongs to $\mathbf{C}^{0,\alpha}_{[0,1]}(1)$. Hence, if $f \in \mathbf{C}^{0,\delta}_{[0,1]}(1)$ for some $\delta \in [0, 1)$, then

$$g \in \mathbf{C}^{\min\{\alpha,\delta\}}_{[0,1]}(1)\,,$$

where $g(x) = (1-x)^{\alpha+\mathbf{i}\beta} f(x)$.

Proposition 5.15 *Let* $p \in (1, \infty)$, $\rho, \sigma \in (-1, p-1)$, $0 \le \gamma < \infty$, *and* $a \in \mathbb{C}$, $b \in \mathbb{C} \setminus \{0\}$ *with* $a \pm b \ne 0$. *Moreover, assume that there is a* $\nu \in \mathbb{C}$ *and a* $\kappa \in \{0, \pm 1\}$ *such that* (4.3) *is fulfilled. If* $u \in \mathbf{L}^p_{\rho,\sigma}$ *and*

$$f = au + b\mathcal{S}u \in \mathbf{H}^\gamma(0,1) \cap \mathbf{C}^{0,\delta}_{[0,1]}(1)$$

for some $\delta \in (0,1)$*, then*

$$\begin{cases} \left\{ \begin{array}{c} u \in \mathbf{H}^{\gamma}(0,1) \cap \mathbf{C}^{0,\delta_1}_{[0,1]}(1)\,,\ u(1)=0\,, \\ \delta_1 = \min\{1-\Re\nu,\delta\}-\varepsilon \\ \forall\,\varepsilon \in \big(0,\min\{1-\Re\nu,\delta\}\big) \end{array} \right\} : & \kappa=-1\,, \\ \left\{ \begin{array}{c} u = u_1 + f(1)p_0\upsilon^{\nu,-\nu}\,, \\ u_1 \in \mathbf{H}^{\gamma}(0,1) \cap \mathbf{C}^{0,\delta_1}_{[0,1]}(1)\,,\ u_1(1)=0\,, \\ \delta_1 = \min\{-\Re\nu,\delta\}-\varepsilon \\ \forall\,\varepsilon \in \big(0,\min\{-\Re\nu,\delta\}\big) \end{array} \right\} : & \kappa=0\,,\ -1<\Re\nu<0\,, \\ \left\{ \begin{array}{c} u = u_1 + f(1)p_0\upsilon^{\nu,-\nu}\,, \\ u_1 \in \mathbf{H}^{\gamma}(0,1) \cap \Big(\upsilon^{0,-\nu}\mathbf{C}^{0,\delta}_{[0,1]}(1)\Big) \end{array} \right\} : & \kappa=0\,,\ \Re\nu=0\,, \\ \left\{ \begin{array}{c} u = u_1 + f(1)p_0\upsilon^{\nu,-\nu}\,, \\ u_1 \in \mathbf{H}^{\gamma}(0,1) \cap \Big(\upsilon^{0,-\nu}\mathbf{C}^{0,\min\{\Re\nu,\delta\}}_{[0,1]}(1)\Big) \end{array} \right\} : & \kappa=0\,,\ 0<\Re\nu<1\,, \\ \left\{ \begin{array}{c} u = u_1 + \big[p_0 + f(1)p_1\big]\upsilon^{\nu,-\nu-1}\,, \\ u_1 \in \mathbf{H}^{\gamma}(0,1) \cap \Big(\upsilon^{0,-\nu-1}\mathbf{C}^{0,\min\{\Re\nu+1,\delta\}}_{[0,1]}(1)\Big) \end{array} \right\} : & \kappa=1\,, \end{cases}$$

with certain polynomials $p_j(x)$ *of degree less or equal than* j*.*

Proof Since $\mathcal{S}$ maps $\mathbf{L}^p_{\rho,\sigma}$ into itself, we have $f \in \mathbf{L}^p_{\rho,\sigma}$. First, we consider the case $\kappa=-1$. Hence, $0<\eta<\Re\nu<1-\xi<1$ (cf. (4.3)). Due to Lemma 4.3, the operator $a\mathcal{I}+b\mathcal{S}:\mathbf{L}^p_{\rho,\sigma}\to\mathbf{L}^p_{\rho,\sigma}$ is left-sided invertible and

$$u = \Big(a^2-b^2\Big)^{-1}\Big[af - b\upsilon^{\nu,1-\nu}\mathcal{S}\upsilon^{-\nu,\nu-1}f\Big]\,, \tag{5.6}$$

where $\upsilon^{-\nu,\nu-1}f \in \mathbf{L}^1_{0,0}$ in view of Remark 5.13. Moreover, from (5.6) and (4.6) we deliver

$$u(x) = \frac{1}{a^2-b^2}\left(a[f(x)-f(1)] - \frac{b\upsilon^{\nu,1-\nu}(x)}{\pi\mathbf{i}}\int_0^1 \frac{y^{-\nu}[f(y)-f(1)]}{(1-y)^{1-\nu}}\,\frac{dy}{y-x}\right).$$

In case $\delta > 1-\Re\nu$, Lemma 5.12 delivers that the integral in the last equation belongs to $\upsilon^{0,\alpha_0}\mathbf{C}^{0,\lambda}_{[0,1]}(1)$ for $0<\alpha_0<1-\Re\nu$ and $0<\lambda<\alpha_0$. In particular, $u(1)=0$. Choosing $\alpha_0 = 1-\Re\nu-\frac{\varepsilon}{2}$ and $\lambda = 1-\Re\nu-\varepsilon$, we get, taking into account Remark 5.14, that $u \in \mathbf{C}^{0,\delta_1}_{[0,1]}(1)$ for $\delta_1 = 1-\Re\nu-\varepsilon$ and all $\varepsilon \in (0,1-\Re\nu)$. Analogously, in case $\delta \le 1-\Re\nu$, we obtain $u \in \mathbf{C}^{0,\delta_1}_{[0,1]}(1)$ for $\delta_1 = \delta-\varepsilon$ and all $\varepsilon \in (0,\delta)$. It remains to apply Corollary 5.10 to conclude $u \in \mathbf{H}^{\gamma}(0,1)$.

Now, let us consider the case $\kappa = 0$. Due to Lemma 4.3, the operator $a\mathcal{I} + b\mathcal{S} : \mathbf{L}^p_{\rho,\sigma} \to \mathbf{L}^p_{\rho,\sigma}$ is invertible and

$$u = \left(a^2 - b^2\right)^{-1} \left[af - bv^{\nu,-\nu}\mathcal{S}v^{-\nu,\nu} f\right],$$

where $-1 < \Re\nu < 1$. From (4.7) we get

$$u(x) = \frac{1}{a^2 - b^2}\left(a[f(x) - f(1)] - \frac{bv^{\nu,-\nu}(x)}{\pi \mathbf{i}} \int_0^1 \frac{y^{-\nu}[f(y) - f(1)]}{(1-y)^{-\nu}} \frac{dy}{y - x}\right)$$
$$+ \gamma_0 f(1) v^{\nu,-\nu}(x)\,.$$

with a constant $\gamma_0 \in \mathbb{C}$. In case $-1 < \Re\nu \leq 0$, we get the assertion in the same manner as in the previous case, i.e., $u = u_1 + \gamma_0 v^{\nu,-\nu} f(1)$ with $u_1 \in \mathbf{H}^\gamma(0,1) \cap \mathbf{C}^{0,\delta_1}_{[0,1]}(1)$ and $u_1(1) = 0$. In case $0 < \Re\nu < 1$, we conclude

$$u_1 \in \mathbf{H}^\gamma(0,1) \cap \left(v^{0,-\nu}\mathbf{C}^{0,\min\{\Re\nu,\delta\}}_{[0,1]}(1)\right)$$

by using Corollary 5.11 and Remark 5.14 together with Corollary 5.10. Note that $\nu \neq 0$ due to $b \neq 0$. If $\Re\nu = 0$, then we apply Lemma 5.12 in case $\alpha = 0$.

In case $\kappa = 1$, due to Lemma 4.3, the operator $a\mathcal{I} + b\mathcal{S} : \mathbf{L}^p_{\rho,\sigma} \to \mathbf{L}^p_{\rho,\sigma}$ is invertible from the right and its null space is spanned by $v^{\nu,-\nu-1}$, such that

$$u = \left(a^2 - b^2\right)^{-1} \left[af - bv^{\nu,-\nu-1}\mathcal{S}v^{-\nu,\nu+1} f\right] + \gamma_0 v^{\nu,-\nu-1}$$

with some constant $\gamma_0 \in \mathbb{C}$. Using (4.8) for $1 + \nu$ instead of ν, we deliver

$$u(x) = \frac{1}{a^2 - b^2}\Big(a[f(x) - f(1)]$$
$$- \frac{bv^{\nu,-\nu-1}(x)}{\pi \mathbf{i}} \int_0^1 \frac{y^{-\nu}(1-y)^{\nu+1}[f(y) - 1]}{y - x}\, dy\Big)$$
$$+ \left[\gamma_0 + f(1)(\delta_0 + \delta_1 x)\right] v^{\nu,-\nu-1}(x)\,.$$

Now, we can again make use of Corollary 5.11 and Remark 5.14 together with Corollary 5.10 to get the assertion also for this case. □

For $\rho, \sigma > -1$, we define the space

$$\widetilde{\mathbf{L}}^\infty_{\rho,\sigma} = \bigcap_{1<p<\infty} \mathbf{L}^p_{\rho,\sigma}.$$

Corollary 5.16 *Let $\rho, \sigma > -1$, $0 \le \gamma < \infty$, and $a \in \mathbb{C}$, $b \in \mathbb{C}\setminus\{0\}$ with $a \pm b \neq 0$. If $u \in \widetilde{\mathbf{L}}^{\infty}_{\rho,\sigma}$ and*

$$f = au + b\mathcal{S}u \in \mathbf{H}^{\gamma}(0,1) \cap \mathbf{C}^{0,\delta}_{[0,1]}(1)\,,$$

then

$$u \in \mathbf{H}^{\gamma}(0,1) \cap \mathbf{C}^{0,\delta_1}_{[0,1]}(1) \quad \text{and} \quad u(1) = 0\,,$$

where $\delta_1 = \min\{1 - \Re\nu, \delta\} - \varepsilon$ and $\frac{a+b}{a-b} = e^{2\pi \mathbf{i}\nu}$, $0 \le \Re\nu < 1$, and where $\varepsilon \in \big(0, \min\{1 - \Re\nu, \delta\}\big)$ is arbitrary.

Proof Since, for sufficiently large p, the operator $\mathcal{S}$ maps $\mathbf{L}^{p}_{\rho,\sigma}$ into itself, we have $f \in \widetilde{\mathbf{L}}^{\infty}_{\rho,\sigma}$. Hence, in Lemma 4.3 we can choose $0 \le \Re\nu < 1$ and $p > 1$ sufficiently large such that (4.3) is satisfied for $\kappa = -1$. It remains to apply Proposition 5.15. □

Remark 5.17 Let $p, \rho, \sigma, a, b, \nu$, and κ fulfil the conditions of Proposition 5.15, let $c \in \mathbf{H}^{\gamma}(0,1]$, and let $k(t)$ satisfy the conditions of Lemma 5.8 for some $\gamma \in (0,1)$. If $u \in \widetilde{\mathbf{L}}^{p}_{\rho,\sigma}$ and

$$f = au + b\mathcal{S}u + c\mathcal{B}_k u \in \mathbf{H}^{\gamma}(0,1) \cap \mathbf{C}^{0,\delta}_{[0,1]}(1)$$

for some $\delta \in (0,1)$, then

$$au + b\mathcal{S}u = f - c\mathcal{B}_k u \in \mathbf{H}^{\gamma}(0,1) \cap \mathbf{C}^{0,\min\{\gamma,\delta\}}_{[0,1]}\,,$$

and we can apply Proposition 5.15 with $\min\{\gamma, \delta\}$ instead of δ to deduce a regularity property for u.

Analogously, using Corollary 5.16, we get the following corollary.

Corollary 5.18 *Let ρ, σ, a, and b satisfy the conditions of Corollary 5.16, let $c \in \mathbf{H}^{\gamma}(0,1]$, and let $k(t)$ satisfy the conditions of Lemma 5.8 for some $\gamma \in (0,1)$. If $u \in \widetilde{\mathbf{L}}^{\infty}_{\rho,\sigma}$ and*

$$f = au + b\mathcal{S}u + c\mathcal{B}_k u \in \mathbf{H}^{\gamma}(0,1) \cap \mathbf{C}^{0,\delta}_{[0,1]}(1)$$

for some $\delta \in (0,1)$, then

$$u \in \mathbf{H}^{\gamma}(0,1) \cap \mathbf{C}^{0,\delta_1}_{[0,1]}(1)$$

and $u(1) = 0$, where $\delta_1 = \min\{1 - \Re\nu, \gamma, \delta\}$ and where $\nu \in \mathbb{C}$ is defined as in Corollary 5.16.

6 One-Sided Invertibility

In this section, we are going to investigate the one-sided invertibility of integral operators of the form $a\mathcal{I} + b\mathcal{S} + \mathcal{B}_k$ in the space $\mathbf{L}^p_{\rho,\sigma}$. Under certain conditions regarding the numbers ρ, σ, p, the coefficients a, b, as well as the kernel function k we prove that the homogeneous equation $(a\mathcal{I} + b\mathcal{S} + c\mathcal{B}_k)u = 0$ in the space $\mathbf{L}^p_{\rho,\sigma}$ or the adjoint equation $(a\mathcal{I} + b\mathcal{S} + c\mathcal{B}_k)^* v = 0$ in the space $\mathbf{L}^q_{(1-q)\rho,(1-q)\sigma}$, $\frac{1}{p} + \frac{1}{q} = 1$, have only the trivial solution. In the remaining part of this section, let $\xi = \frac{1+\rho}{p}$ and $\frac{1}{p} + \frac{1}{q} = 1$.

Proposition 6.1 *Let $p \in (1, \infty)$, $\rho, \sigma \in (-1, p-1)$, $a \in \mathbb{C} \setminus \{0\}$, and $k \in \mathbf{C}(\mathbb{R}^+)$ satisfy condition* (A). *Then the homogeneous equations $(a\mathcal{I}+\mathcal{B}_k)u = 0$ in the space $\mathbf{L}^p_{\rho,\sigma}$ or $(a\mathcal{I}+\mathcal{B}_k)^* v = 0$ in the space $\mathbf{L}^q_{(1-q)\rho,(1-q)\sigma}$ have only the trivial solution.*

Proof Let $u \in \mathbf{L}^p_{\rho,\sigma}$, $v \in \mathbf{L}^q_{(1-q)\rho,(1-q)\sigma}$ and

$$(a\mathcal{I} + \mathcal{B}_k)u = 0\,, \quad (a\mathcal{I} + \mathcal{B}_k)^* v \overset{\text{Lemma 4.1}}{=} (\overline{a}\mathcal{I} + \mathcal{B}_{k_1})v = 0\,, \tag{6.1}$$

where $k_1(t) = \overline{k(t^{-1})}t^{-1}$. Due to Corollary 3.3, we can choose $\alpha_0, \beta_0 \in \mathbb{R}$, such that $\alpha < \alpha_0 < \xi < \beta_0 < \beta$ and

$$\lim_{t\to+0} t^{\alpha_0} k(t) = 0 \quad \text{and} \quad \lim_{t\to\infty} t^{\beta_0} k(t) = 0\,.$$

This implies

$$\lim_{t\to 0} t^{1-\beta_0} k_1(t) = 0 \quad \text{and} \quad \lim_{t\to\infty} t^{1-\alpha_0} k_1(t) = 0.$$

Since $\alpha_0 < \frac{1+\rho}{p}$ is equivalent to $\frac{1+(1-q)\rho}{q} < 1 - \alpha_0$, we get, by using Remark 5.2 together with Corollary 5.6 (both in case $n = 0$), that $u \in \mathbf{L}^p_{\rho,0}$ and $v \in \mathbf{L}^q_{(1-q)\rho,0}$. By Lemma 3.8, $\mathcal{B}_k \in \mathcal{L}\big(\mathbf{L}^p_{\rho,\sigma}\big)$ is true for all $\sigma \in (-1, p-1)$. Hence we can consider the equations in (6.1) in the spaces $\mathbf{L}^p_{\rho,0}$ and $\mathbf{L}^q_{(1-q)\rho,0}$, respectively. Because of the relations (4.1) and (4.2), it only remains to apply Lemma 2.2. □

The relations (4.1) and (4.2) together with Lemma 2.2 also immediately deliver the following proposition.

Proposition 6.2 *Let $p \in (1, \infty)$, $\rho \in (-1, p-1)$, $a \in \mathbb{C}$, $b \in \mathbb{C} \setminus \{0\}$, and $k \in \mathbf{C}(\mathbb{R}^+)$ satisfy condition* (A). *Then the homogeneous equations*

$$(a\mathcal{I} + b\mathcal{S} + \mathcal{B}_k)u = 0$$

in the space $\mathbf{L}^p_{\rho,0}$ *or*

$$(a\mathcal{I} + b\mathcal{S} + \mathcal{B}_k)^* v = 0$$

in the space $\mathbf{L}^q_{(1-q)\rho,0}$ *have only the trivial solution.*

Proposition 6.3 *Let* $p \in (1,\infty)$, $\rho, \sigma \in (-1, p-1)$, $a \in \mathbb{C}$, $b \in \mathbb{C} \setminus \{0\}$, *and* $k \in \mathbf{C}(\mathbb{R}^+)$ *satisfy condition* (A). *Moreover, we assume that* $k(t)$ *fulfils*

$$M_{\beta,a,\gamma_1}(k) := \sup\left\{ \frac{\left|s^\beta k(s) - t^\beta k(t)\right|}{|s-t|^{\gamma_1}} : s,t \in [a,\infty)\,, s \neq t \right\} < \infty$$

and

$$N_{\alpha,a,\gamma_2}(k) := \sup\left\{ \frac{|s^\alpha k(s) - t^\alpha k(t)|}{|s-t|^{\gamma_2}} : s,t \in (0,a]\,, s \neq t \right\} < \infty \tag{6.2}$$

for all $a > 0$ *and for some* $\gamma_i \in (0,1)$, $i \in \{1,2\}$, *with* $\xi + \gamma_1 < \beta$ *and* $\alpha < \xi - \gamma_2$. *Let* $a\mathcal{I} + b\mathcal{S}$ *be invertible in* $\mathbf{L}^p_{\rho,\sigma}$, *i.e., there is a* $\nu \in \mathbb{C}$ *satisfying* (4.3) *and* (4.4) *for* $\kappa = 0$. *Moreover, we assume that*

$$\Re\nu < \frac{1}{p} \quad \textit{and} \quad -\Re\nu < \frac{1}{q}\,. \tag{6.3}$$

Then the homogeneous equations

$$(a\mathcal{I} + b\mathcal{S} + \mathcal{B}_k)u = 0$$

in the space $\mathbf{L}^p_{\rho,\sigma}$ *or*

$$(a\mathcal{I} + b\mathcal{S} + \mathcal{B}_k)^* v = 0$$

in the space $\mathbf{L}^q_{(1-q)\rho,(1-q)\sigma}$ *have only the trivial solution.*

Proof Let $u \in \mathbf{L}^p_{\rho,\sigma}$, $v \in \mathbf{L}^q_{(1-q)\rho,(1-q)\sigma}$, and

$$(a\mathcal{I} + b\mathcal{S} + \mathcal{B}_k)u = 0\,,\ (a\mathcal{I} + b\mathcal{S} + \mathcal{B}_k)^* v \overset{\text{Lemma 4.1}}{=} (\overline{a}\mathcal{I} + \overline{b}\mathcal{S} + \mathcal{B}_{k_1})v = 0\,,$$

where $k_1(t) = \overline{k(t^{-1})}t^{-1}$. Since the function $k(t)$ fulfils (6.2), we get that $k_1(t)$ satisfies $M_{1-\alpha,a,\gamma_2}(k_1) < \infty$ for all $a > 0$. Thus, Lemma 5.8 delivers $\mathcal{B}_k u \in \mathbf{H}^{\gamma_1}(0,1]$ and $\mathcal{B}_{k_1} v \in \mathbf{H}^{\gamma_2}(0,1]$. From Proposition 5.15 (in case $\kappa = 0$), we derive $u \in \mathbf{L}^p_{\rho,0}$ and $v \in \mathbf{L}^q_{(1-q)\rho,0}$, where we took into account assumption (6.3). Thus, it only remains to apply (4.1) in combination with Lemma 2.2. □

Proposition 6.4 *Let* $p \in (1,\infty)$, $\rho \in (-1, p-1)$, $a, b \in \mathbb{C}$, *and* $k \in \mathbf{C}(\mathbb{R}^+)$ *satisfy condition (A). If*

$$\inf\left\{\left|a - b\mathbf{i}\cot(\pi\xi + \mathbf{i}t) + \widehat{k}(\xi + \mathbf{i}t)\right| : t \in \mathbb{R}\right\} > 0\,,$$

then the homogeneous equation

$$(a\mathcal{I} + b\mathcal{S} + \mathcal{B}_k)u = 0$$

in the space $\mathbf{L}^p_{\rho,\sigma}$ *or the adjoint equation*

$$(a\mathcal{I} + b\mathcal{S} + \mathcal{B}_k)^* v = 0$$

in the space $\mathbf{L}^q_{(1-q)\rho,(1-q)\sigma}$ *have only the trivial solution.*

Proof Relation (4.1) in combination with Lemma 2.3 immediately delivers the assertion. □

References

1. A. Böttcher, I.M. Spitkovsky, Pseudodifferential operators with heavy spectrum. Integr. Equ. Oper. Theory **19**, 251–269 (1994)
2. R. Duduchava, *Integral Equations with Fixed Singularities* (BSB B. G. Teubner Verlagsgesellschaft, Leipzig, 1979)
3. I. Gohberg, N. Krupnik, *One-Dimensional Linear Singular Integral Equations. Vol I: Introduction*. Operator Theory: Advances and Applications, vol. 53 (Birkhäuser Verlag, Basel, 1992)
4. P. Junghanns, R. Kaiser, A numerical approach for a special crack problem. Dolomites Res. Notes Approx. **10**, 56–67 (2017)
5. P. Junghanns, R. Kaiser, A note on the Fredholm theory of singular integral operators with Cauchy and Mellin kernels, in *In Operator Theory, Analysis and the State Space Approach. In Honor of Rien Kaashoek*. Operator Theory : Advances and Applications, vol. 271 (Birkhäuser, Basel, 2018), 291–325
6. S.G. Mikhlin, S. Prössdorf, *Singular Integral Operators* (Akademie-Verlag, Berlin, 1986)
7. N.I. Muskhelishvili, *Singular Integral Equations. Boundary Problems of Function Theory and Their Application to Mathematical Physics* (P. Noordhoff N. V., Groningen, 1953)
8. R. Schneider, Integral equations with piecewise continuous coefficients in L_p-spaces with weight. J. Integr. Equ. **9**, 135–152 (1985)
9. E.C. Titchmarsh, *Introduction to the Theory of Fourier Integrals*, 3rd edn. (Chelsea Publishing Co., New York, 1986)

A Note on Group Representations, Determinantal Hypersurfaces and Their Quantizations

Igor Klep and Jurij Volčič

Abstract Recently, there have been exciting developments on the interplay between representation theory of finite groups and determinantal hypersurfaces. For example, a finite Coxeter group is determined by the determinantal hypersurface described by its natural generators under the regular representation. This short note solves three problems about extending this result in the negative. On the affirmative side, it is shown that a quantization of a determinantal hypersurface, the so-called free locus, correlates well with representation theory. If $A_1, \dots, A_\ell \in \mathrm{GL}_d(\mathbb{C})$ generate a finite group G, then the family of hypersurfaces $\{X \in \mathrm{M}_n(\mathbb{C})^d : \det(I + A_1 \otimes X_1 + \cdots + A_\ell \otimes X_\ell) = 0\}$ for $n \in \mathbb{N}$ determines G up to isomorphism.

Keywords Linear pencil · Group representation · Determinantal hypersurface · Free locus

Mathematics Subject Classification (2010) Primary 20C15, 15A22; Secondary 47A56, 14J70

The first author was supported by the Slovenian Research Agency grants J1-8132, N1-0057, P1-0222, and partially supported by the Marsden Fund Council of the Royal Society of New Zealand.

I. Klep
Department of Mathematics, Faculty of Mathematics and Physics, University of Ljubljana, Ljubljana, Slovenia
e-mail: igor.klep@fmf.uni-lj.si

J. Volčič (✉)
Department of Mathematics, Texas A&M University, College Station, TX, USA
e-mail: volcic@math.tamu.edu

M. A. Bastos et al. (eds.), *Operator Theory, Functional Analysis and Applications*,
Operator Theory: Advances and Applications 282,
https://doi.org/10.1007/978-3-030-51945-2_19

1 Introduction

To $A_0, \dots, A_\ell \in \mathrm{M}_d(\mathbb{C})$ one assigns the **determinantal hypersurface**

$$\{[\xi_0 : \cdots : \xi_\ell] \in \mathbb{CP}^\ell : \det(\xi_0 A_0 + \cdots + \xi_\ell A_\ell) = 0\}. \tag{1.1}$$

This is a classical object in algebraic geometry [1, 6, 10, 11], where a key question asks which hypersurfaces admit determinantal representations. When A_j are real symmetric matrices, determinantal hypersurfaces pertain to hyperbolic and stable polynomials [2, 15, 18, 23, 24]. The geometry of the hypersurface (1.1) is also explored in multivariate operator theory [3, 4, 26]. If A_j are bounded operators on a Hilbert space and the determinant in (1.1) is replaced with the condition that $\xi_0 A_0 + \cdots + \xi_\ell A_\ell$ is not invertible, then (1.1) is known as the **projective joint spectrum** of $A_0, \dots, A_\ell$ (cf. Taylor spectrum [22] for ensembles of commuting operators).

Through the work of Frobenius [13] and Dedekind [7] on group determinants (see also [9]), determinantal hypersurfaces also pertain to representation theory. Several fascinating developments in this direction [5, 14, 21] have been recently made. This note addresses certain limitations for extensions of these results.

Let G be a finitely generated group. If $T = (g_1, \dots, g_\ell)$ is a finite sequence of generators for G and $\rho : G \to \mathrm{GL}_d(\mathbb{C})$ is a representation of G, then denote

$$\mathscr{Z}_1(T, \rho) = \left\{ \xi \in \mathbb{C}^\ell : \det\left(I_d + \xi_1 \rho(g_1) + \cdots + \xi_\ell \rho(g_\ell)\right) = 0 \right\}. \tag{1.2}$$

It is natural to ask what kind of information the affine hypersurface $\mathscr{Z}_1(T, \rho)$ carries about ρ and G. For example, if G_1, G_2 are finite groups with left regular representations λ_1, λ_2, then $\mathscr{Z}_1(G_1 \backslash \{1\}, \lambda_1) = \mathscr{Z}_1(G_2 \backslash \{1\}, \lambda_2)$ implies that G_1, G_2 are isomorphic [12]. However, one is typically interested in smaller generating sets or in finitely generated groups which are not necessarily finite. In [14], the authors computed the joint spectrum for the infinite dihedral group

$$D_\infty = \langle a, t \mid a^2 = t^2 = 1 \rangle$$

with respect to the generating set $(1, a, t)$, and analyzed its properties through the representation theory of D_∞. Determinantal hypersurfaces also have a strong connection with representation theory in the case of finite Coxeter groups [5]. A Coxeter group is a finitely generated group on generators $g_1, \dots, g_\ell$ satisfying

$$(g_i g_j)^{m_{ij}} = 1$$

where $m_{ii} = 1$ and $m_{ij} \geq 2$ for $i \neq j$. In [5] the authors first showed that if G is a finite Coxeter group, λ is its left regular representation, and $T = (g_1, \dots, g_\ell)$ are the generators as above, then $\mathscr{Z}_1(T, \lambda)$ determines G up to isomorphism. Furthermore,

if G is not of exceptional type (in the Coxeter diagram sense) and ρ is an arbitrary finite-dimensional representation of G, then $\mathscr{Z}_1(T, \rho)$ determines ρ.

These theorems were presented during the Multivariable Spectral Theory and Representation Theory workshop at the Banff International Research Station in April 2019. Several problems about extending these results beyond Coxeter groups were posed by the speakers; among them were the following.

Questions 1.1 Let G be a finite group, T a fixed generating set for G, and ρ_1, ρ_2 irreducible complex representations of G.

(1) Is $\mathscr{Z}_1(T, \rho_1)$ a reduced and irreducible hypersurface?
(2) If $\mathscr{Z}_1(T, \rho_1) = \mathscr{Z}_1(T, \rho_2)$, are ρ_1 and ρ_2 equivalent?

As usual, $\rho_1 : G \to \mathrm{GL}_{d_1}(\mathbb{C})$ and $\rho_2 : G \to \mathrm{GL}_{d_2}(\mathbb{C})$ are equivalent if $d_1 = d_2$ and $\rho_2 = P\rho_2 P^{-1}$ for some $P \in \mathrm{GL}_{d_1}(\mathbb{C})$. A representation ρ_1 is irreducible if its image does not admit a nontrivial common invariant subspace; equivalently, it generates $\mathrm{M}_{d_1}(\mathbb{C})$ as a $\mathbb{C}$-algebra by Burnside's theorem [17, Corollary 1.17]. The hypersurface $\mathscr{Z}_1(T, \rho_1)$ is reduced and irreducible (in the scheme-theoretic sense) if its defining determinant in (1.2) is an irreducible polynomial. The main result of this note is the following.

Theorem 1.2 *Questions* 1.1(1) and (2) *have negative answers in general.*

See Sects. 2.1 and 2.2 for concrete examples. On a more positive side, in Sect. 3 we show that representation theory aligns well with a quantization of the determinantal hypersurface, the free locus; see Theorem 3.1. Furthermore, Proposition 3.4 determines whether a free locus arises from a representation of a finite group, and Proposition 3.7 characterizes finite abelian groups from the perspective of determinantal hypersurfaces. We conclude this note with an open question.

2 Representations Versus Determinants

In this section we give negative answers to Questions 1.1. The representations were found with the help of the computer algebra system GAP and the online repository ATLAS of Finite Group Representations. Verifying equivalence and irreducibility of representations was sometimes done symbolically with the computing system Mathematica.

2.1 Irreducible Representation with Reducible Determinant

The alternating group $G = A_6$ admits a presentation

$$G = \left\langle g_1, g_2 \mid g_1^2, g_2^4, (g_1 g_2)^5, (g_1 g_2^2)^5 \right\rangle.$$

Let

$$A_1 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \end{pmatrix}$$

and

$$A_2 = \begin{pmatrix} 0\,1\,0\,0\,0\,0\,0\,0\,0 \\ 0\,0\,0\,1\,0\,0\,0\,0\,0 \\ 0\,0\,0\,0\,1\,0\,0\,0\,0 \\ 0\,0\,0\,0\,0\,1\,0\,0\,0 \\ 0\,0\,0\,0\,0\,0\,1\,0\,0 \\ 1\,0\,0\,0\,0\,0\,0\,0\,0 \\ 0\,0\,0\,0\,0\,0\,0\,0\,1 \\ 0\,0\,0\,0\,0\,0\,0\,1\,0 \\ 0\,0\,1\,0\,0\,0\,0\,0\,0 \end{pmatrix}.$$

Then $\rho(g_1) = A_1$ and $\rho(g_2) = A_2$ determines a faithful and irreducible representation $\rho : G \to \mathrm{GL}_9(\mathbb{C})$. Indeed, we can directly check that

$$A_1^2 = A_2^4 = (A_1A_2)^5 = (A_1A_2^2)^5 = I,$$

so ρ is a representation of G, and is moreover faithful since it is nontrivial and G is simple. Furthermore, all the possible products of A_1 and A_2 with at most 8 factors span the whole $\mathrm{M}_9(\mathbb{C})$, so ρ is irreducible. However, we claim that the curve $\mathscr{Z}_1((g_1, g_2), \rho)$ in $\mathbb{C}^2$ is not irreducible. We can compute the determinant of $I + x_1\rho(g_1) + x_2\rho(g_2)$,

$$\det \begin{pmatrix} 1+x_1 & x_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & x_1 & x_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & x_1 & 1 & 0 & x_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & x_1 & x_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_1 & 1 & 0 & x_2 & 0 & 0 \\ x_2 & 0 & 0 & 0 & 0 & 1 & 0 & x_1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1+x_1 & 0 & x_2 \\ 0 & 0 & 0 & 0 & 0 & x_1 & 0 & 1+x_2 & 0 \\ -x_1 & -x_1 & x_2-x_1 & -x_1 & -x_1 & -x_1 & -x_1 & -x_1 & 1-x_1 \end{pmatrix}$$

by cofactor expansion along the rows. The reader will have no difficulty verifying that $\det(I + x_1\rho(g_1) + x_2\rho(g_2))$ equals

$$1 + x_1 - 4x_1^2 - 4x_1^3 + 6x_1^4 + 6x_1^5 - 4x_1^6 - 4x_1^7 + x_1^8 + x_1^9 + x_2 + 2x_1x_2 - 2x_1^2x_2$$
$$- 6x_1^3x_2 + 6x_1^5x_2 + 2x_1^6x_2 - 2x_1^7x_2 - x_1^8x_2 + x_1^2x_2^2 + x_1^3x_2^2 - 2x_1^4x_2^2 - 2x_1^5x_2^2$$
$$+ x_1^6x_2^2 + x_1^7x_2^2 - x_1^2x_2^3 + 2x_1^4x_2^3 - x_1^6x_2^3 - 2x_2^4 + x_1^2x_2^4 - x_1^3x_2^4 + x_1^4x_2^4 + x_1^5x_2^4$$
$$- 2x_2^5 - 2x_1x_2^5 - x_1^2x_2^5 + x_1^4x_2^5 - x_1^2x_2^6 - x_1^3x_2^6 + x_1^2x_2^7 + x_2^8 - x_1x_2^8 + x_2^9$$

which is the product of the following two irreducible polynomials:

$$1 + 2x_1 - 2x_1^3 - x_1^4 + x_1x_2 + 2x_1^2x_2 + x_1^3x_2 - x_1x_2^2 - x_1^2x_2^2 + x_1x_2^3 - x_2^4,$$
$$1 - x_1 - 2x_1^2 + 2x_1^3 + x_1^4 - x_1^5 + x_2 - x_1x_2 - x_1^2x_2 + x_1^3x_2 - x_2^4 - x_2^5.$$

Some of the subsequent examples are presented in a more terse way to maintain the focus on their intent.

Note that the above irreducible representation of A_6 has dimension 9, which is not the minimum among nontrivial complex representations of A_6; namely, A_6 admits a representation σ of minimal dimension 5, and $\mathscr{Z}_1((g_1, g_2), \sigma)$ is irreducible. One might thus be tempted to suggest that for a group G generated by a finite set T and its (irreducible) representation σ of minimal dimension, $\mathscr{Z}_1(T, \sigma)$ is irreducible. However, even this weaker conjecture fails. The counterexample is given by the Janko group J_2,

$$J_2 = \left\langle g_1, g_2 \mid g_1^2, g_2^3, (g_1g_2)^7, (g_1g_2g_1^{-1}g_2^{-1})^{12}, (g_1g_2(g_1g_2g_1g_2^{-1})^2)^6 \right\rangle.$$

This sporadic simple group of order 604800 admits two non-isomorphic complex representations σ_1, σ_2 of minimal dimension 14, courtesy of ATLAS of Finite Group Representations. As in the previous example (albeit with slightly longer calculations), one can explicitly check that the curve $\mathscr{Z}_1((g_1, g_2), \sigma_1) = \mathscr{Z}_1((g_1, g_2), \sigma_2)$ has two irreducible components.

2.2 *Non-equivalent Representations with the Same Determinant*

The classical group $G = \mathrm{GL}_2(\mathbb{Z}/3\mathbb{Z})$ admits the presentation

$$G = \left\langle g_1, g_2, g_3 \mid g_1^2, (g_1g_2^{-1})^2, (g_1g_3^{-1})^2, g_2^2g_3g_2^{-1}g_3, g_2g_3^2g_2g_3^{-1} \right\rangle.$$

Let A_1, A_2, A_3 be the matrices

$$\begin{pmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{2}-\frac{i}{2} \\ -\frac{1}{2}+\frac{i}{2} & \frac{1}{\sqrt{2}} \end{pmatrix}, \quad \begin{pmatrix} \frac{1}{2}+\frac{i}{2} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{2}-\frac{i}{2} \end{pmatrix}, \quad \begin{pmatrix} \frac{1}{2}-\frac{i}{2} & \frac{i}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} & \frac{1}{2}+\frac{i}{2} \end{pmatrix}.$$

There are faithful irreducible unitary representations $\rho_+, \rho_- : G \to \mathrm{GL}_2(\mathbb{C})$ given by

$$\rho_\pm(g_1) = \pm A_1, \quad \rho_\pm(g_2) = A_2, \quad \rho_\pm(g_3) = A_3.$$

It is easy to check that ρ_+ and ρ_- are not equivalent. On the other hand,

$$\mathcal{Z}_1((g_1, g_2, g_3), \rho_\pm) = \{(\xi_1, \xi_2, \xi_3) \colon 1 - \xi_1^2 + \xi_2 + \xi_2^2 + \xi_3 + \xi_3^2 = 0\}.$$

3 Free Locus Perspective

In this section we will see how representations of a finitely generated group are determined by a noncommutative relaxation of (1.2). To $A \in \mathrm{M}_d(\mathbb{C})^\ell$ we associate the monic matrix pencil $L_A = I + A_1 x_1 + \cdots + A_\ell x_\ell$ of size d in freely noncommuting variables $x = (x_1, \ldots, x_\ell)$. Thus L is an affine matrix over the free algebra $\mathbb{C}\langle x \rangle$. At a matrix point $X \in \mathrm{M}_n(\mathbb{C})^\ell$ it evaluates as

$$L_A(X) = I_{dn} + A_1 \otimes X_1 + \cdots + A_\ell \otimes X_\ell \in \mathrm{M}_{dn}(\mathbb{C}).$$

The **free locus** [19] of L_A is the disjoint union of determinantal hypersurfaces

$$\mathscr{Z}(L_A) = \bigsqcup_{n \in \mathbb{N}} \mathscr{Z}_n(L_A), \quad \mathscr{Z}_n(L_A) = \left\{ X \in \mathrm{M}_n(\mathbb{C})^\ell \colon \det L_A(X) = 0 \right\}.$$

Given a group G generated by $T = (g_1, \ldots, g_n)$ and a complex representation $\rho : G \to \mathrm{GL}_d(\mathbb{C})$, we write

$$\mathscr{Z}(T, \rho) = \mathscr{Z}\left(L_{\rho(g_1), \ldots, \rho(g_\ell)}\right). \tag{3.1}$$

By the definition of the free locus we see that (3.1) is indeed a quantization of (1.2). The existing results on free loci [16, 19] readily apply to group representations.

Theorem 3.1 *For $i = 1, 2$ let G_i be a group generated by a finite sequence T_i and let ρ_i be a complex representation of G_i. Assume $|T_1| = |T_2|$.*

(1) If ρ_i is irreducible, then there exists $n_0 \in \mathbb{N}$ such that $\mathscr{Z}_n(T_1, \rho_1)$ is a reduced and irreducible hypersurface for all $n \geq n_0$.

(2) If ρ_1 and ρ_2 are irreducible, then $\mathscr{Z}(T_1, \rho_1) = \mathscr{Z}(T_2, \rho_2)$ if and only if $G_1/\ker\rho_1 \cong G_2/\ker\rho_2$ and ρ_1, ρ_2 are equivalent.
(3) For $i = 1, 2$ assume that G_i is finite and ρ_i is a faithful representation. Then $\mathscr{Z}(T_1, \rho_1) = \mathscr{Z}(T_2, \rho_2)$ if and only if $G_1 \cong G_2$ via an isomorphism mapping T_1 to T_2.

Proof

(1) A consequence of [16, Theorem 3.4].
(2) A consequence of [19, Theorem 3.11].
(3) Let $\mathcal{T}_i$ be the $\mathbb{C}$-algebra generated by T_i. Since G_i is finite, its group algebra $\mathbb{C}G_i$ is semisimple by Maschke's theorem [17, Theorem 1.9]. Since $\mathcal{T}_i$ is a quotient of $\mathbb{C}G_i$, it is also semisimple. Then $\mathscr{Z}(T_1, \rho_1) = \mathscr{Z}(T_2, \rho_2)$ if and only if $T_1 \mapsto T_2$ induces an algebra isomorphism $\mathcal{T}_1 \to \mathcal{T}_2$ by [19, Corollary 3.8]. This isomorphism then restricts to the group isomorphism $G_1 \to G_2$. □

Remark 3.2 There is a deterministic bound on n_0 in Theorem 3.1(1) that is exponential in $|T_1|$ and the dimension of ρ_1 by [16, Remark 3.5] (the bound is likely not optimal). Similarly, to verify $\mathscr{Z}(T_1, \rho_1) = \mathscr{Z}(T_2, \rho_2)$ of Theorem 3.1(2,3), it suffices to check $\mathscr{Z}_n(T_1, \rho_1) = \mathscr{Z}_n(T_2, \rho_2)$ for a fixed large enough n, exponential in $|T_i|$ and the dimension of ρ_i by [19, Remark 3.7].

Free loci are defined for monic pencils with arbitrary matrix coefficients; we now describe how the geometry of the free locus $\mathscr{Z}(L_A)$ detects whether the coefficients $A_1, \ldots, A_\ell$ generate a finite group. See also [8] for an efficient algorithm that determines finiteness of a finitely generated linear group.

Definition 3.3 Let $\ell, n \in \mathbb{N}$. Let $C \in \mathrm{GL}_n(\mathbb{Z})$ be the permutation matrix corresponding to the cycle $(1\,2\,\cdots\,n)$. If $\{1, \ldots, n\} = S_1 \sqcup \cdots \sqcup S_\ell$ and P_j is the orthogonal projection onto $\mathrm{span}\{e_k : k \in S_j\}$, then the matrix point

$$X = (P_1 C, \cdots, P_\ell C) \in \mathrm{M}_n(\mathbb{Z})^\ell$$

is called a **cycle partition**. For given ℓ, n we thus have ℓ^n cycle partitions.

Let $\mu_\infty \subset \mathbb{C} \setminus \{0\}$ be the group of all roots of unity. The next proposition shows that if $A_1, \ldots, A_\ell$ generate a finite group, then $\mathscr{Z}(L_A)$ intersects complex lines through cycle partitions only in points from μ_∞.

Proposition 3.4 *Let $A \in \mathrm{M}_d(\mathbb{C})^\ell$. Then $A_1, \ldots, A_\ell$ generate a finite group if and only if the following hold:*

(i) there is a positive definite $P \in \mathrm{M}_d(\mathbb{C})$ such that $A_j^* P A_j = P$ for all j;
(ii) for every cycle partition X and $t \in \mathbb{C}$,

$$tX \in \mathscr{Z}(L_A) \implies t \in \mu_\infty.$$

Proof ($\Leftarrow$) Every A_j is invertible by (i). Let G be a group generated by $A_1, \ldots, A_\ell$. Also by (i), G is a subgroup of the unitary group in $\mathrm{GL}_d(\mathbb{C})$ with respect to the

inner product $\langle u, v\rangle = u^* P v$. Hence every element of G is diagonalizable. By [25, Corollary 4.9], a finitely generated subgroup of $\mathrm{GL}_d(\mathbb{C})$ is finite if and only if it is periodic (or torsion; i.e., every element has finite order). Since a diagonalizable matrix has a finite order if and only if all its eigenvalues lie in μ_∞, it suffices to verify that eigenvalues of every element of G lie in μ_∞.

To $(i_1, \dots, i_n) \in \{1, \dots, \ell\}^n$ we associate the cycle partition $X \in \mathrm{M}_n(\mathbb{Z})^\ell$ by choosing $S_j = \{e_k : i_k = j\}$. We claim that $tX \in \mathscr{Z}(L_A)$ if and only if $(-t)^n$ is an eigenvalue of $A_{i_1} \cdots A_{i_n}$. Indeed, using Schur complements it is easy to check that

$$\begin{aligned}
\det(I - (-1)^n t^n A_{i_1} \cdots A_{i_n}) &= \det \begin{pmatrix} I & tA_{i_1} \\ (-1)^n t^{n-1} A_{i_2} \cdots A_{i_n} & I \end{pmatrix} \\
&= \det \begin{pmatrix} I & & tA_{i_1} & 0 \\ 0 & & I & tA_{i_2} \\ -(-1)^n t^{n-2} A_{i_3} \cdots A_{i_n} & & 0 & I \end{pmatrix} \\
&= \cdots \\
&= \det \begin{pmatrix} I & tA_{i_1} & & \\ & \ddots & \ddots & \\ & & I & tA_{i_{n-1}} \\ tA_{i_n} & & & I \end{pmatrix} \\
&= \det L_A(tX).
\end{aligned}$$

Thus the matrix $A_{i_1} \cdots A_{i_n}$ has finite order if and only if $tX \in \mathscr{Z}(L_A)$ implies $t \in \mu_\infty$, which holds by (ii).

($\Rightarrow$) If $A_1, \dots, A_\ell$ generate a finite group G, then $\mathbb{C}^d$ admits a G-invariant inner product

$$\langle u, v\rangle = \sum_{g \in G} (gu)^* (gv).$$

If P is the positive definite matrix satisfying $\langle u, v\rangle = u^* P v$, then (i) holds. Furthermore, the proof of (ii) is already given in the previous paragraph. □

Remark 3.5 If additional information about $A_1, \dots, A_\ell$ is given, say that their entries generate a number field (finite extension of $\mathbb{Q}$), then the size of the cycle partitions, which have to be tested in Proposition 3.4, can be bounded using Schur's theorem on orders of finite matrix groups [17, Theorem 14.19].

Remark 3.6 Let $p \in \mathbb{N}$ be prime. If μ_∞ in Proposition 3.4 is replaced by the group of power-of-p roots of unity, one obtains a free locus characterization of matrix tuples that generate a finite p-group.

We also show how the free locus certifies whether its defining coefficients generate a finite abelian group. The degree of an affine variety of codimension m is the number of intersection points of the variety with m hyperplanes in general position; in the case of a hypersurface, it is simply the degree of its square-free defining polynomial.

Proposition 3.7 *Let G be a finite group generated by $A_1, \dots, A_\ell \in \mathrm{M}_d(\mathbb{C})$. Then G is abelian if and only if the irreducible components of $\mathscr{Z}_n(L_A)$ have degree n for all $n \in \mathbb{N}$.*

Proof Let $\mathcal{A}$ be the $\mathbb{C}$-algebra generated by $A_1, \dots, A_\ell$. As in the proof of Theorem 3.1(3) we see that $\mathcal{A}$ is semisimple. After a basis change (which does not affect the structure of G or $\mathscr{Z}(L_A)$) we can thus assume that

$$A_j = A_j^{(1)} \oplus \cdots \oplus A_j^{(s)}$$

where $A_1^{(k)}, \dots, A_\ell^{(k)} \in \mathrm{M}_{d_k}(\mathbb{C})$ determine an irreducible representation of G for every $k = 1, \dots, s$. For $X \in \mathrm{M}_n(\mathbb{C})^d$ let us view $\det L_{A^{(k)}}(X)$ as a polynomial in the entries of X. If $d_k = 1$, then $\det L_{A^{(k)}}(X)$ is up to an affine change of coordinates equal to the determinant of a generic $n \times n$ matrix, and hence an irreducible polynomial of degree n. On the other hand, if $d_k > 1$, then $\det L_{A^{(k)}}(X)$ is a polynomial of degree $d_k n > n$ for all n, and irreducible for all large enough n by [16, Theorem 3.4]. Since G is abelian if and only if $d_1 = \cdots = d_s = 1$, and

$$\mathscr{Z}_n(L_A) = \mathscr{Z}_n(L_A^{(1)}) \cup \cdots \cup \mathscr{Z}_n(L_A^{(s)}),$$

it follows that G is abelian if and only if the irreducible components of $\mathscr{Z}_n(L_A)$ are hypersurfaces of degree n. □

Remark 3.8 If $\ell = 2$ and A_1, A_2 are hermitian, then $\mathscr{Z}_1(L_A)$ alone determines whether G is abelian, cf. [20].

The last two propositions offer some directions for future research. Theorem 3.1 implies that the linear group G generated by a tuple A is determined by $\mathscr{Z}(L_A)$. It would be interesting to know which properties of G can be deduced from the geometry of $\mathscr{Z}(L_A)$. For example, intersections of $\mathscr{Z}(L_A)$ with certain lines and hyperplanes determine whether G is finite or abelian. An open problem is how to decide whether a finite group G is nilpotent/solvable/simple (or any other group-theoretic property) by considering the geometry of the hypersurfaces $\mathscr{Z}_n(L_A)$.

Acknowledgments The authors thank Banff International Research Station for the hospitality during the Multivariable Spectral Theory and Representation Theory workshop, and participants for sharing their ideas.

References

1. A. Beauville, Determinantal hypersurfaces. Michigan Math. J. **48**, 39–64 (2000)
2. P. Brändén, Obstructions to determinantal representability. Adv. Math. **226**, 1202–1212 (2011)
3. P. Cade, R. Yang, Projective spectrum and cyclic cohomology. J. Funct. Anal. **265**, 1916–1933 (2013)
4. I. Chagouel, M.I. Stessin, K. Zhu, Geometric spectral theory for compact operators. Trans. Am. Math. Soc. **368**, 1559–1582 (2016)
5. Ž. Čučković, M.I. Stessin, A.B. Tchernev, Determinantal hypersurfaces and representations of Coxeter groups. Preprint. arXiv:1810.12893
6. C. de Concini, D. Eisenbud, C. Procesi, Young diagrams and determinantal varieties. Invent. Math. **56**, 129–165 (1980)
7. R. Dedekind, *Gesammelte mathematische Werke*. Band II (Chelsea Publishing Co., New York, 1968)
8. A.S. Detinko, D.L. Flannery, E.A. O'Brien, Recognizing finite matrix groups over infinite fields. J. Symb. Comput. **50**, 100–109 (2013)
9. L.E. Dickson, An elementary exposition of Frobenius's theory of group-characters and group-determinants. Ann. Math. **4**, 25–49 (1902)
10. L.E. Dickson, Determination of all general homogeneous polynomials expressible as determinants with linear elements. Trans. Am. Math. Soc. **22**, 167–179 (1921)
11. I. Dolgachev, *Classical Algebraic Geometry: a Modern View* (Cambridge University Press, Cambridge, 2012)
12. E. Formanek, D. Sibley, The group determinant determines the group. Proc. Am. Math. Soc. **112**, 649–656 (1991)
13. F.G. Frobenius, *Gesammelte Abhandlungen*. Bände I–III (Springer, Berlin/New York, 1968)
14. R.I. Grigorchuk, R. Yang, Joint spectrum and the infinite dihedral group. Proc. Steklov Inst. Math. **297**, 145–178 (2017)
15. J.W. Helton, V. Vinnikov, Linear matrix inequality representation of sets. Commun. Pure Appl. Math. **60**, 654–674 (2007)
16. J.W. Helton, I. Klep, J. Volčič, Geometry of free loci and factorization of noncommutative polynomials. Adv. Math. **331**, 589–626 (2018)
17. I.M. Isaacs, *Character Theory of Finite Groups*. Pure and Applied Mathematics, vol. 69 (Academic, New York/London, 1976)
18. D. Kerner, V. Vinnikov, Determinantal representations of singular hypersurfaces in $\mathbb{P}^n$. Adv. Math. **231**, 1619–1654 (2012)
19. I. Klep, J. Volčič, Free loci of matrix pencils and domains of noncommutative rational functions. Comment. Math. Helv. **92**, 105–130 (2017)
20. T.S. Motzkin, O. Taussky, Pairs of matrices with property L. Trans. Am. Math. Soc. **73**, 108–114 (1952)
21. M.I. Stessin, A.B. Tchernev, Spectral algebraic curves and decomposable operator tuples. J. Oper. Theory **82**, 75–113 (2019)
22. J.L. Taylor, A joint spectrum for several commuting operators. J. Funct. Anal. **6**, 172–191 (1970)
23. J. Volčič, Stable noncommutative polynomials and their determinantal representations. SIAM J. Appl. Algebra Geom. **3**, 152–171 (2019)
24. D.G. Wagner, Multivariate stable polynomials: theory and applications. Bull. Am. Math. Soc. (N.S.) **48**, 53–84 (2011)
25. B.A.F. Wehrfritz, *Infinite Linear Groups*. Ergebnisse der Matematik und ihrer Grenzgebiete, vol. 76 (Springer, New York/Heidelberg, 1973)
26. R. Yang, Projective spectrum in Banach algebras. J. Topol. Anal. **1**, 289–306 (2009)

Algebras Generated by Toeplitz Operators on the Unit Sphere II: Non Commutative Case

Maribel Loaiza and Nikolai Vasilevski

Abstract In Loaiza and Vasilevski (Commutative algebras generated by Toeplitz operators on the unit sphere. Intgr. Equ. Oper. Theory, v. 92, 25 (2020)), we represented the Hardy space $H^2(S^{2n-1})$, with $n \geq 2$, as a direct sum of weighted Bergman spaces $\mathcal{A}^2_p(\mathbb{B}^{n-1})$, with $p \in \mathbb{Z}_+$, This permitted us to represent Toeplitz operators, whose symbols are invariant under certain $\mathbb{T}$-action, acting on $H^2(S^{2n-1})$, as direct sums of Toeplitz operators, acting on corresponding Bergman spaces. As a benefit we were able to use already known results on Toeplitz operators on Bergman spaces, and a wide variety of commutative algebras, generated by Toeplitz operators on $H^2(S^{2n-1})$, was described.

Following the same approach, in the present paper we pass to the non-commutative case, presenting the detailed description of two non-commutative C^*-algebras generated by Toeplitz operators on the Hardy space $H^2(S^{2n-1})$.

1 Introduction

The paper continues the study of algebras generated by Toeplitz operators, acting on the multidimensional Hardy space $H^2(S^{2n-1})$, which was started in [12]. Recall in this connection the principal difference in algebraic properties of Toeplitz operators acting on the one-dimensional Hardy $H^2(S^1)$ and Bergman $\mathcal{A}^2(\mathbb{D})$ spaces.

The classical result by Brown and Halmos [9] implies that there is no nontrivial commutative C^*-algebra generated by Toeplitz operators acting on the Hardy space $H^2(S^1)$, while there are only two commutative Banach algebras. One of them is generated by Toeplitz operators with analytic symbols, and the other one is generated by Toeplitz operators with conjugate analytic symbols. Of course, such two algebras remain to be commutative for $H^2(S^{2n-1})$, for all $n > 1$.

M. Loaiza · N. Vasilevski (✉)
Departamento de Matemáticas, CINVESTAV, México, Mexico
e-mail: mloaiza@math.cinvestav.mx; nvasilev@math.cinvestav.mx

M. A. Bastos et al. (eds.), *Operator Theory, Functional Analysis and Applications*,
Operator Theory: Advances and Applications 282,
https://doi.org/10.1007/978-3-030-51945-2_20

At the same time, at the turn of this century, it was observed [15, 16] (see also [17]) that there are many nontrivial commutative C^*-algebras, generated by Toeplitz operators acting on the Bergman space $\mathcal{A}^2(\mathbb{D})$ over the unit disk $\mathbb{D} \subset \mathbb{C}$. These results were extended then to the case of Toeplitz operators acting on weighted Bergman spaces on the unit ball $\mathbb{B}^n$, see [13]. Further on many nontrivial Banach algebras generated by Toeplitz operators, that are commutative on each standard weighted Bergman space over $\mathbb{B}^n$, were discovered and studied.

In this connection a rather challenging question appeared: what is the situation with both C^* and Banach algebras generated by Toeplitz operators on the multidimensional Hardy space $H^2(S^{2n-1})$. A first step in this direction was made in [1] (see as well [2]), where Z. Akkar, following all the reasonings of [13], described the commutative C^*-algebras generated by Toeplitz operators on $H^2(S^{2n-1})$.

In [12], we developed an alternative approach to the problem. Therein we represented the Hardy space $H^2(S^{2n-1})$ as a direct sum of weighted Bergman spaces $\mathcal{A}^2_p(\mathbb{B}^{n-1})$, with $p \in \mathbb{Z}_+$, which permitted us to represent Toeplitz operators, acting on $H^2(S^{2n-1})$, as direct sums of Toeplitz operators, acting on corresponding Bergman spaces. The benefit of such an approach is that we can use now all the results on Toeplitz operators on Bergman spaces in their full power. In particular, we showed in [12] how to recover the results of [1] just as simple and straightforward corollaries of [13], and explained how to unhide and describe a wide variety of nontrivial commutative Banach algebras generated by Toeplitz operators on $H^2(S^{2n-1})$.

Following the same approach, in the present paper we pass to the description of non-commutative algebras, and present the detailed description of two non-commutative C^*-algebras generated by Toeplitz operators, acting on the Hardy space $H^2(S^{2n-1})$, based on the already obtained results [4, 6–8] for Toeplitz operators, acting on the Bergman space.

2 Preliminaries: Bergman and Hardy Spaces, Toeplitz Operators

We recall here the results on the representation of the multidimensional Hardy space in terms of the Bergman spaces as well as the corresponding representation of Toeplitz operators, acting on the Hardy space, in terms of Toeplitz operators, acting on Bergman spaces. All proofs and details can be found in [12].

Denote by

$$\mathbb{B}^n = \{z = (z_1, z_2, \dots, z_n) \in \mathbb{C}^n : |z|^2 := |z_1|^2 + \cdots + |z_n|^2 < 1\}$$

the unit ball in $\mathbb{C}^n$ and by

$$S^{2n-1} = \{z = (z_1, z_2, \dots, z_n) \in \mathbb{C}^n : |z_1|^2 + \cdots + |z_n|^2 = 1\}$$

the unit sphere, being the boundary of $\mathbb{B}^n$. The following standard notations will be used throughout the paper. For $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{Z}_+^n$, where $\mathbb{Z}_+ = \{0, 1, 2, \ldots\} \subset \mathbb{Z}$,

$$z^\alpha = z_1^{\alpha_1} \ldots z_n^{\alpha_n},$$
$$\alpha! = \alpha_1! \ldots \alpha_n!,$$
$$|\alpha| = \alpha_1 + \cdots + \alpha_n.$$

The standard Lebesgue measure in $\mathbb{C}^n$ is denoted by $dV(z)$, i.e.

$$dV(z) = dx_1 dy_1 \ldots dx_n dy_n,$$

where $z = (z_1, z_2, \ldots, z_n)$ and $z_k = x_k + iy_k$, $k = 1, \ldots, n$. For each $\lambda \in (-1, \infty)$ introduce the standard normalized weighted measure

$$dv_\lambda(z) = \frac{\Gamma(n + \lambda + 1)}{\pi^n \Gamma(\lambda + 1)} \left(1 - |z|^2\right)^\lambda dV(z).$$

The weighted Bergman space $\mathcal{A}_\lambda^2(\mathbb{B}^n)$ is the closed subspace of $L_2(\mathbb{B}^n, dv_\lambda)$ that consists of all analytic functions. The Hardy space $H^2(S^{2n-1})$ is defined as the set of all holomorphic functions f, defined in $\mathbb{B}^n$, such that

$$\|f\|_2^2 := \sup_{0<r<1} \int_{S^{2n-1}} |f(r\zeta)|^2 d\sigma(\zeta) < \infty,$$

where $d\sigma$ denotes the normalized measure for S^{2n-1}. This space can be defined, alternatively, as the closed subspace of $L_2(S^{2n-1}, d\sigma)$ that consists of all functions f satisfying the tangential Cauchy-Riemann equations:

$$L_{k,j} f = \left(z_k \frac{\partial}{\partial \overline{z_j}} - z_j \frac{\partial}{\partial \overline{z_k}} \right) f = 0, \quad 1 \le k < j \le n. \tag{2.1}$$

Observe that $f \in L_2(S^{2n-1}, d\sigma)$ if and only if

$$\widetilde{f}(z', t_n) := f(z', \sqrt{1 - |z'|^2} t_n) \in L_2(\mathbb{B}^{n-1}, dv(z')) \otimes L_2(S^1, d\mu(t_n)),$$

where

$$z' = (z_1, \ldots, z_{n-1}), \quad z_n = \sqrt{1 - |z'|^2} t_n, \quad dv(z') = \frac{(n-1)!}{\pi^{n-1}} dV(z'),$$

$$dV(z') = dx_1 dy_1 \cdots dx_{n-1} dy_{n-1},$$

and $d\mu(t_n) = \frac{1}{2\pi}\frac{dt_n}{it_n}$ is the normalized arc-length measure on S^1.

Equations (2.1) are not independent. The equations that are independent and equivalent to (2.1) have the form

$$D_k(\widetilde{f}) = \left[\frac{\partial}{\partial \overline{z}_k} + \frac{1}{2}\frac{z_k}{1-|z'|^2} t_n \frac{\partial}{\partial t_n}\right]\widetilde{f} = 0, \quad k = 1, 2, \ldots, n-1.$$

Recall that the discrete Fourier transform $\mathcal{F} : L_2(S^1, d\mu(t_n)) \to \ell_2(\mathbb{Z})$, has the form

$$\mathcal{F} : g \mapsto c_n = \left\{\int_{S^1} g(t_n) t_n^{-p} d\mu(t_n)\right\}_{p\in\mathbb{Z}}.$$

Introduce the unitary operator $U_1 = I \otimes \mathcal{F}$, being an isometric isomorphism of

$$L_2(S^{2n-1}, d\sigma) \cong L_2(\mathbb{B}^{n-1}, dv(z')) \otimes L_2(S^1, d\mu(t_n))$$

onto

$$L_2(\mathbb{B}^{n-1}, dv(z')) \otimes \ell_2(\mathbb{Z}) = \ell_2(\mathbb{Z}, L_2(\mathbb{B}^{n-1}, dv(z')))$$

and acting as follows

$$U_1 \,:\, \widetilde{f}(z', t_n) \;\longmapsto\; \left\{\widetilde{c}_p(z')\right\}_{p\in\mathbb{Z}},$$

where

$$\widetilde{c}_p(z') = \int_{S^1} \widetilde{f}(z', t_n) t_n^{-p} d\mu(t_n).$$

Observe that

$$\ell_2(\mathbb{Z}, L_2(\mathbb{B}^{n-1}, dv(z'))) = \bigoplus_{p\in\mathbb{Z}} L_2(\mathbb{B}^{n-1}, dv(z')).$$

For each $p \in \mathbb{Z}$ introduce then the unitary operator

$$u_p : L_2(\mathbb{B}^{n-1}, dv(z')) \to L_2(\mathbb{B}^{n-1}, dv_p(z')),$$

where $dv_p(z') = \frac{(n+|p|-1)!}{\pi^{n-1}|p|!}(1-|z'|^2)^p dV(z')$, which acts as follows

$$u_p \,:\, \widetilde{c}_p(z') \;\longmapsto\; c_p(z') = \sqrt{\frac{(n-1)!|p|!}{(n+|p|-1)!}}\widetilde{c}_p(z')(1-|z'|^2)^{-\frac{p}{2}}.$$

Finally, introduce the unitary operator

$$U_2 = \bigoplus_{p\in\mathbb{Z}} u_p \,:\, \bigoplus_{p\in\mathbb{Z}} L_2(\mathbb{B}^{n-1}, dv(z')) \;\longrightarrow\; \bigoplus_{p\in\mathbb{Z}} L_2(\mathbb{B}^{n-1}, dv_p(z')).$$

Theorem 2.1 ([12, Theorem 2.1]) *The unitary operator $U = U_2U_1$ is an isometric isomorphism of*

$$L_2(S^{2n-1}, d\sigma) \cong L_2(\mathbb{B}^{n-1}, dv(z')) \otimes L_2(S^1, d\mu(t_n))$$

onto

$$\bigoplus_{p\in\mathbb{Z}} L_2(\mathbb{B}^{n-1}, dv_p(z'))$$

under which $H^2(S^{2n-1})$ is mapped onto $\bigoplus_{p\in\mathbb{Z}_+} \mathcal{A}^2_p(\mathbb{B}^{n-1})$.

For each $p \in \mathbb{Z}_+$, we denote by B_p the Bergman orthogonal projection from $L_2(\mathbb{B}^{n-1}, dv_p(z'))$ onto the weighted Bergman space $\mathcal{A}^2_p(\mathbb{B}^{n-1})$, and let $P_{S^{2n-1}}$ be the Szegő projection from $L_2(S^{2n-1}, d\sigma)$ onto the Hardy space $H^2(S^{2n-1})$.

Corollary 2.2 ([12, Corollary 2.2]) *The Szegő projection and the weighted Bergman projections are connected as follows*

$$U P_{S^{2n-1}} U^{-1} = \bigoplus_{p\in\mathbb{Z}_+} B_p.$$

Let $z = (z', z_n)$ be a point in the unit sphere S^{2n-1} in $\mathbb{C}^n$, where $z' = (z_1, \ldots, z_{n-1})$ and write $z_n = \sqrt{1-|z'|^2}t_n$, where $t_n \in S^1$. Consider a function $\boldsymbol{a}(z', z_n)$ defined in the unit sphere S^{2n-1}. The Toeplitz operator $\boldsymbol{T_a}$, acting on the Hardy space $H^2(S^{2n-1})$, is defined as follows

$$\boldsymbol{T_a} f = P_{S^{2n-1}}(\boldsymbol{a} f).$$

The crucial role of the Fourier transform $\mathcal{F}$ in establishing of Theorem 2.1 suggests us to consider the symbols $\boldsymbol{a}$ of the form

$$\boldsymbol{a} = \boldsymbol{a}(z_1, \ldots, z_{n-1}, |z_n|) = \boldsymbol{a}(z', |z_n|), \tag{2.2}$$

i.e., the symbols that do not depend on t_n, commuting thus with $U_1 = I \otimes \mathcal{F}$.

Given the above $\boldsymbol{a} \in L_\infty(S^{2n-1})$, introduce the associated function

$$\begin{aligned} a &= a(z_1, \ldots, z_{n-1}) = a(z') \\ &= \boldsymbol{a}(z_1, \ldots, z_{n-1}, \sqrt{1-|z_1|^2-\ldots-|z_{n-1}|^2}) = \boldsymbol{a}(z', \sqrt{1-|z'|^2}), \end{aligned} \tag{2.3}$$

where $(z_1, \dots, z_{n-1}) = z' \in \mathbb{B}^{n-1}$. Note that each function $a \in L_\infty(\mathbb{B}^{n-1})$ defines in its turn the function $\boldsymbol{a}$ of the form (2.2) by $\boldsymbol{a}(z', |z_n|) = a(z')$.

We denote then by T_a^p the Toeplitz operator, with symbol a acting on the weighted Bergman space $\mathcal{A}_p^2(\mathbb{B}^{n-1})$, with $p \in \mathbb{Z}_+$.

Theorem 2.3 ([12, Theorem 3.1]) *Given a bounded measurable symbol $\boldsymbol{a}(z', |z_n|)$, defined in S^{2n-1}. Under the isomorphism $U = U_2U_1$, the Toeplitz operator $\boldsymbol{T_a}$, acting on the Hardy space $H^2(S^{2n-1})$, is unitarily equivalent to the operator $\bigoplus_{p\in\mathbb{Z}_+} T_a^p$, acting on*

$$\bigoplus_{p\in\mathbb{Z}_+} \mathcal{A}_p^2(\mathbb{B}^{n-1}),$$

where $a = a(z')$ is of the form (2.3).

In what follows we will consider two classes (algebras) $\boldsymbol{S}$ of functions $\boldsymbol{a}$ of the form (2.2) as well as the corresponding classes (algebras) S of functions a defined by (2.3) for $\boldsymbol{a} \in \boldsymbol{S}$. Of course, symmetrically, each class S of functions $a = a(z')$, $z' \in \mathbb{B}^{n-1}$ defines the class $\boldsymbol{S}$ of functions $\boldsymbol{a}$ of the form (2.2), connected with $a \in S$ by (2.3).

Given a class $\boldsymbol{S}$, we denote by $\boldsymbol{\mathcal{T}}(\boldsymbol{S})$ the closed unital algebra generated by all Toeplitz operators $\boldsymbol{T_a}$, with $\boldsymbol{a} \in \boldsymbol{S}$, acting on the Hardy space $H^2(S^{2n-1})$. For a corresponding class S, we denote by $\mathcal{T}_p(S)$ the closed unital algebra generated by all Toeplitz operators T_a^p, with $a \in S$, acting on the weighted Bergman space $\mathcal{A}_p^2(\mathbb{B}^{n-1})$, with $p \in \mathbb{Z}_+$.

Theorems 2.1 and 2.3 state then that for each generator $\boldsymbol{T_a}$ of the algebra $\boldsymbol{\mathcal{T}}(\boldsymbol{S})$, the operator $U\boldsymbol{T_a}U^{-1}$ leaves invariant each subspace in the direct sum decomposition

$$U(H^2(S^{2n-1})) = \bigoplus_{p\in\mathbb{Z}_+} \mathcal{A}_p^2(\mathbb{B}^{n-1})$$

and

$$U\boldsymbol{T_a}U^{-1} = \bigoplus_{p\in\mathbb{Z}_+} T_a^p. \tag{2.4}$$

Moreover, for each operator $\boldsymbol{T} \in \boldsymbol{\mathcal{T}}(\boldsymbol{S})$ there is a unique sequence of operators $T^p \in \mathcal{T}_p(S)$, $p \in \mathbb{Z}_+$, such that

$$U\boldsymbol{T}U^{-1} = \bigoplus_{p\in\mathbb{Z}_+} T^p.$$

In what follows we will abbreviate the above equation as

$$\boldsymbol{T} \asymp \bigoplus_{p\in\mathbb{Z}_+} T^p. \tag{2.5}$$

3 Compact Semi-commutator C^*-Algebra

We consider the case when $S = \mathrm{VO}_\partial(\mathbb{B}^{n-1})$ with the corresponding class $\boldsymbol{S}_{\mathrm{VO}} \subset L_\infty(S^{2n-1})$. Recall [8], that a bounded continuous function a belongs to $\mathrm{VO}_\partial(\mathbb{B}^{n-1})$ (has a vanishing oscillation at the boundary) if

$$\lim_{z'\to\partial\mathbb{B}^{n-1}} \big[\sup\{|a(z') - a(w')| \,:\, \beta(z', w') \le 1\}\big] = 0,$$

here $\beta(\cdot,\cdot)$ is the Bergman metric in $\mathbb{B}^{n-1}$, and that $\mathrm{VO}_\partial(\mathbb{B}^{n-1})$ is a norm closed C^*-subalgebra of $L_\infty(\mathbb{B}^{n-1})$.

Our choice of $S = \mathrm{VO}_\partial(\mathbb{B}^{n-1})$ is motivated by the following observation. C. Berger, L. Coburn, and K. Zhu described in [8] the largest C^*-subalgebra Q of $L_\infty(\mathbb{B}^{n-1})$ which possesses the compact semi-commutator property, i.e.,

$$[T_a^p, T_b^p) := T_a^p T_b^p - T_{ab}^p \ \text{ is compact for every } \ a,\, b \in Q.$$

Then, $C(\overline{\mathbb{B}}^{n-1}) \subset \mathrm{VO}_\partial(\mathbb{B}^{n-1}) \subset Q$ and the algebra $\mathcal{T}_p(C(\overline{\mathbb{B}}^{n-1}))$ contains all compact operators of $\mathcal{A}_p^2(\mathbb{B}^{n-1})$, which, in particular, implies that the algebra $\mathcal{T}_p(\mathrm{VO}_\partial(\mathbb{B}^{n-1}))$ coincides with $\mathcal{T}_p(Q)$. At the same time it is more pleasant to deal with (bounded uniformly continuous) symbols from $\mathrm{VO}_\partial(\mathbb{B}^{n-1})$ than with (generically discontinuous) symbols from Q.

Note, in this connection, that the C^*-algebra $\mathcal{T}_p(\mathrm{VO}_\partial(\mathbb{B}^{n-1})) = \mathcal{T}_p(Q)$ is irreducible, contains the ideal $\mathcal{K}_p$ of all compact operators, and each its element admits the representation $T^p = T_a^p + K^p$, with $a \in \mathrm{VO}_\partial(\mathbb{B}^{n-1})$ and compact K^p.

We recall as well that the Berezin transform $\mathcal{B}_p : \mathcal{L}(\mathcal{A}_p^2(\mathbb{B}^{n-1})) \to L^\infty(\mathbb{B}^{n-1})$ is given by $\mathcal{B}_p[A](z') := \big\langle Ak_{z'}^p, k_{z'}^p\big\rangle_p$, where $k_{z'}^p$ denotes the normalized reproducing kernel function of $\mathcal{A}_p^2(\mathbb{B}^{n-1})$ defined by

$$k_{z'}^p(w') = \frac{(1-|z'|^2)^{\frac{n+p}{2}}}{(1-w'\cdot\overline{z'})^{n+p}}, \quad z',\ w' \in \mathbb{B}^{n-1}.$$

For the special case of a Toeplitz operator T_a^p,

$$\mathcal{B}_p[T_a^p](z') = \mathcal{B}_p[a](z') = \big\langle a\, k_{z'}^p, k_{z'}^p\big\rangle_p.$$

The aim of this section is to study the algebra $\mathcal{T}(\boldsymbol{S}_{\mathrm{VO}})$ generated by all Toeplitz operators $\boldsymbol{T_a}$ acting on $H^2(S^{2n-1})$ and having symbols $\boldsymbol{a} \in \boldsymbol{S}_{\mathrm{VO}}$, associated to $S = \mathrm{VO}_\partial(\mathbb{B}^{n-1})$.

Lemma 3.1 *For each $\boldsymbol{T} \asymp \bigoplus_{p\in\mathbb{Z}_+} T^p$ we have*

$$\|\boldsymbol{T}\| = \sup_{p\in\mathbb{Z}_+} \|T^p\|,$$

while for the generating operators we have

$$\|\boldsymbol{T_a}\| = \|a(z')\|_\infty.$$

Proof The first equality easily follows by standard arguments from the direct sum representation of the operator $\boldsymbol{T}$. The second follows from $\|T_a^p\| \le \|a(z')\|_\infty$ and [4, Proposition 4.4], implying

$$\|a(z')\|_\infty = \|\lim_{p\to\infty} \mathcal{B}_p\left(T_a^p\right)\|_\infty \le \sup_{p\in\mathbb{Z}} \|T_a^p\| = \|\boldsymbol{T_a}\|. \qquad \square$$

Theorem 3.2 *Each operator $\boldsymbol{T} \in \mathcal{T}(\boldsymbol{S}_{\mathrm{VO}})$, in the direct sum decomposition (2.5), admits the unique representation*

$$\boldsymbol{T} \asymp \bigoplus_{p\in\mathbb{Z}_+} (T_a^p + K^p), \tag{3.1}$$

where $a \in \mathrm{VO}_\partial(\mathbb{B}^{n-1})$, each K^p is compact, and $\|K^p\| \to 0$ as $p \to \infty$.

Proof Follows from the proof of [7, Theorem 4.3] under $\lambda = 0$ and $\ell = m = 1$. For the sake of completeness we sketch it here. Let $\boldsymbol{D}$ be the dense subalgebra of $\mathcal{T}(\boldsymbol{S}_{\mathrm{VO}})$ consisting of all finite sums of finite products of the generators of $\mathcal{T}(\boldsymbol{S}_{\mathrm{VO}})$. By induction, we only need to prove the assertion for a product of just two operators $\boldsymbol{T_a T_b} \asymp \bigoplus_{p\in\mathbb{Z}_+} T_a^p T_b^p$. Since $\mathrm{VO}_\partial(\mathbb{B}^{n-1})$ possesses the semicommutator property, we have that, for each $p \in \mathbb{Z}_+$, the semicommutator

$$T_a^p T_b^p - T_{ab}^p =: K^p$$

is compact. Then [6, Theorem 3.8] implies that $\|K^p\| \to 0$ when $p \to \infty$. Thus

$$\boldsymbol{T_a T_b} \asymp \bigoplus_{p\in\mathbb{Z}_+} T_{ab}^p + K^p,$$

with $\|K^p\| \to 0$ when $p \to \infty$.

Given now $\boldsymbol{T} \in \mathcal{T}(\boldsymbol{S}_{\mathrm{VO}})$, there exists a sequence $\{\boldsymbol{T_m}\}_{m\in\mathbb{N}}$ of elements in $\boldsymbol{D}$ that converges in norm to $\boldsymbol{T}$, where each $\boldsymbol{T_m}$ has the form

$$\boldsymbol{T_m} \asymp \bigoplus_{p\in\mathbb{Z}_+} T^p_{a_m} + K^p_m,$$

with $\|K^p_m\| \to 0$ when $p \to \infty$. Using Lemma 3.1, the convergence of the sequence $\{\boldsymbol{T_m}\}_{m\in\mathbb{N}}$ and that $\|K^p_m\| \to 0$ when $p \to \infty$, it is easy to prove that the sequence $\{a_m\}_{m\in\mathbb{N}}$ is convergent. Denote by a the limit of the last sequence. Then, uniformly on p, $\{T^p_{a_m}\}_{m\in\mathbb{N}}$ converges to T^p_a. This fact implies that the sequence of compact operators $\{K^p_m\}_{m\in\mathbb{N}}$ converges to a compact operator K^p. Then,

$$\boldsymbol{T} \asymp \bigoplus_{p\in\mathbb{Z}_+} T^p_a + K^p.$$

The standard $\frac{\varepsilon}{3}$-trick implies that $\|K^p\| \to 0$ when $p \to \infty$.

To prove its uniqueness of the representation (3.1), we assume that

$$\boldsymbol{T} \asymp \bigoplus_{p\in\mathbb{Z}_+} T^p_a + K^p_1 = \bigoplus_{p\in\mathbb{Z}_+} T^p_b + K^p_2,$$

where $a, b \in \mathrm{VO}_\partial(\mathbb{B}^{n-1})$, K^p_1, K^p_2 are compact operators, for each $p \in \mathbb{Z}_+$, and where $\lim_{p\to\infty} \|K^p_1\| = \lim_{p\to\infty} \|K^p_2\| = 0$. Then,

$$T^p_{a-b} + K^p_1 - K^p_2 = 0, \quad \text{for each } \ p \in \mathbb{Z}_+. \tag{3.2}$$

Last equation and Lemma 3.1 imply that

$$\|a - b\|_\infty = \lim_{p\to\infty} \|T^p_{a-b}\| = \lim_{p\to\infty} \|T^p_{a-b} + K^p_1 - K^p_2\| = 0.$$

Then $a = b$ and, from (3.2), $K^p_1 = K^p_2$. □

Corollary 3.3 *The algebra $\boldsymbol{S}_{\mathrm{VO}}$ possesses the compact semi-commutator property, i.e.,*

$$[\boldsymbol{T_a}, \boldsymbol{T_b}) \ \textit{is compact for every} \ \boldsymbol{a}, \boldsymbol{b} \in \boldsymbol{S}_{\mathrm{VO}}.$$

Each operator $\boldsymbol{T} \in \mathcal{T}(\boldsymbol{S}_{\mathrm{VO}})$ admits a representation

$$\boldsymbol{T} = \boldsymbol{T_a} + \boldsymbol{K},$$

where $\boldsymbol{a} \in \boldsymbol{S}_{\mathrm{VO}}$, $\boldsymbol{K} \in \mathcal{T}(\boldsymbol{S}_{\mathrm{VO}}) \cap \mathcal{K}(H^2(S^{2n-1}))$, and where $\mathcal{K}(H^2(S^{2n-1}))$ is the ideal of all compact operators in $\mathcal{L}(H^2(S^{2n-1}))$.

For $S = L_\infty(\mathbb{B}^{n-1})$, let $\boldsymbol{S}_\infty$ be the corresponding class of symbols defined on S^{2n-1}. We conjecture that the algebra $\boldsymbol{S}_{\mathrm{VO}}$ is the largest C^*-subalgebra among those of $\boldsymbol{S} \subset \boldsymbol{S}_\infty$ that possesses the compact semi-commutator property, i.e.,

$$[\boldsymbol{T}_{\boldsymbol{a}}, \boldsymbol{T}_{\boldsymbol{b}}) \text{ is compact for every } \boldsymbol{a}, \boldsymbol{b} \in \boldsymbol{S} \subset \boldsymbol{S}_\infty \text{ if and only if } \boldsymbol{S} \subseteq \boldsymbol{S}_{\mathrm{VO}}.$$

We list now a family of infinite dimensional irreducible representations of the C^*-algebra $\mathcal{T}(\boldsymbol{S}_{\mathrm{VO}})$.

Proposition 3.4 *For each $p \in \mathbb{Z}_+$ the mapping*

$$\begin{aligned} \iota_p : \mathcal{T}(\boldsymbol{S}_{\mathrm{VO}}) &\longrightarrow \mathcal{T}_p(\mathrm{VO}_\partial(\mathbb{B}^{n-1})) \\ \boldsymbol{T} \asymp \bigoplus_{k\in\mathbb{Z}_+} (T_a^k + K_k) &\longmapsto T_a^p + K^p, \end{aligned}$$

is an irreducible representation of the C^*-algebra $\mathcal{T}(\boldsymbol{S}_{\mathrm{VO}})$.

For different $p \in \mathbb{Z}_+$ the representations ι_p are not unitarily equivalent.

Proof Note that the C^*-algebra, generated by the images under (2.4) of the restrictions of the elements of $\mathcal{T}(\boldsymbol{S}_{\mathrm{VO}})$ onto its invariant subspaces, coincides with the algebra $\mathcal{T}_p(\mathrm{VO}_\partial(\mathbb{B}^{n-1}))$. This is an irreducible C^*-algebra containing the ideal of all compact operators on $\mathcal{A}^2_p(\mathbb{B}^{n-1})$. That is the representation ι_p is irreducible. To prove that, for $p \neq q$, the representations ι_p, ι_q are not unitarily equivalent it suffices to consider the function $\boldsymbol{c}(z', z_n) = 1-|z'|^2$, whose associated function (2.3) is given by $c(z') = 1 - |z'|^2$. Then,

$$\boldsymbol{T}_{\boldsymbol{c}} \asymp \bigoplus_{k\in\mathbb{Z}_+} T_c^k,$$

and $\iota_p(\boldsymbol{T}_{\boldsymbol{c}}) = T_c^p$, which is a diagonal operator with respect to the standard basis of $\mathcal{A}^2_p(\mathbb{B}^{n-1})$. By [13, Theorem 10.1] and [3],

$$\|T^p_{1-|z'|^2}\| = \frac{p+1}{n+p} \neq \frac{q+1}{n+q} = \|T^q_{1-|z'|^2}\|,$$

implying thus that the representations ι_p, ι_q are not unitarily equivalent. □

It is easy to check (see [8] for the unweighted case) the following isomorphism

$$\begin{aligned} \mathcal{T}_p(\mathrm{VO}_\partial(\mathbb{B}^{n-1}))/\mathcal{K}(\mathcal{A}^2_p(\mathbb{B}^{n-1}) &\cong \mathrm{VO}_\partial(\mathbb{B}^{n-1})/C_0(\mathbb{B}^{n-1}) \\ &\cong C(M(\mathrm{VO}_\partial(\mathbb{B}^{n-1})) \setminus \mathbb{B}^{n-1}), \end{aligned}$$

which is defined by

$$T_a^p + \mathcal{K}(\mathcal{A}_p^2(\mathbb{B}^{n-1})) \longmapsto a + C_0(\mathbb{B}^{n-1}) \cong a|_{M(\mathrm{VO}_\partial(\mathbb{B}^{n-1}))\setminus\mathbb{B}^{n-1}},$$

where $M(\mathrm{VO}_\partial(\mathbb{B}^{n-1}))$ denotes the maximal ideal space of $\mathrm{VO}_\partial(\mathbb{B}^{n-1})$.

This leads to the following family of one-dimensional irreducible representations:

Corollary 3.5 *For every $(\eta, p) \in M(\mathrm{VO}_\partial(\mathbb{B}^{n-1})) \setminus \mathbb{B}^{n-1} \times \mathbb{Z}_+$ the map*

$$\pi_{\eta,p} : \boldsymbol{T} \asymp \bigoplus_{k\in\mathbb{Z}_+} (T_a^k + K^k) \overset{\iota_p}{\longmapsto} T_a^p + K^p \longmapsto a(\eta)$$

defines a one-dimensional irreducible representation. Moreover, π_{η_1,p_1} and π_{η_2,p_2} are unitarily equivalent if and only if $\eta_1 = \eta_2$.

Further we have the following variant of [5, Lemma 4.8] and [7, Lemma 5.4].

Lemma 3.6 *The mapping*

$$\nu : \mathcal{T}(\boldsymbol{S}_{\mathrm{VO}}) \longrightarrow \mathrm{VO}_\partial(\mathbb{B}^{n-1}) = C(M(\mathrm{VO}_\partial(\mathbb{B}^{n-1}))),$$

defined by

$$\nu : \boldsymbol{T} \asymp \bigoplus_{p\in\mathbb{Z}_+} T^p \longmapsto \lim_{p\to\infty} \mathcal{B}_p(T^p),$$

where $\mathcal{B}_p$ is the Berezin transform, is a continuous $*$-homomorphism of the C^*-algebra $\mathcal{T}(\boldsymbol{S}_{\mathrm{VO}})$ onto $C(M(\mathrm{VO}_\partial(\mathbb{B}^{n-1})))$.

This result permits us (as in [5, Remark 4.10]) to recover the data of the unique representation (3.1) for each operator $\boldsymbol{T} \asymp \bigoplus_{p\in\mathbb{Z}_+} T^p \in \mathcal{T}(\boldsymbol{S}_{\mathrm{VO}})$:

$$a = \nu(\boldsymbol{T}) = \lim_{p\to\infty} \mathcal{B}_p(T^p) \in \mathrm{VO}_\partial(\mathbb{B}^{n-1}) \quad \text{and} \quad K^p = T^p - T_a^p.$$

Corollary 3.7 *For each $\eta \in M(\mathrm{VO}_\partial(\mathbb{B}^{n-1}))$, the mapping $\nu_\eta : \mathcal{T}(\boldsymbol{S}_{\mathrm{VO}}) \longrightarrow \mathbb{C}$, defined by*

$$\nu_\eta : \boldsymbol{T} \longmapsto \nu(\boldsymbol{T}) = a \longmapsto a(\eta) \in \mathbb{C},$$

is a one-dimensional representation of the C^*-algebra $\mathcal{T}(\boldsymbol{S}_{\mathrm{VO}})$.

We show now that we have listed all (up to the unitary equivalence) irreducible representations of the algebra $\mathcal{T}(\boldsymbol{S}_{\mathrm{VO}})$.

Introduce the set

$$\mathcal{K}_{\mathrm{VO}} := \{\boldsymbol{K} \in \mathcal{T}(\boldsymbol{S}_{\mathrm{VO}}) : \nu(\boldsymbol{K}) = 0\}.$$

In representation (3.1), these are exactly the operators of the form

$$\boldsymbol{K} \asymp \bigoplus_{p\in\mathbb{Z}_+} K^p \tag{3.3}$$

with K^p compact and $\|K^p\| \to 0$ as $p \to \infty$.

Lemma 3.8 *We have that* $\mathcal{T}(\boldsymbol{S}_{\mathrm{VO}}) \cap \mathcal{K}(H^2(S^{2n-1})) = \mathcal{K}_{\mathrm{VO}}$.

Proof Note that each element $\boldsymbol{K} \in \mathcal{K}_{\mathrm{VO}}$ is a compact operator. Indeed, using the representation (3.3) for $\boldsymbol{K}$ we have that

$$\left\| \bigoplus_{p\in\mathbb{Z}_+} K^p - \bigoplus_{p\le k} K^p \right\| = \left\| \bigoplus_{p>k} K^p \right\| = \sup_{p>k} \|K^p\| \to 0, \text{ when } k \to \infty.$$

Then $\bigoplus_{p\in\mathbb{Z}_+} K^p$ is compact and, as a consequence, $\boldsymbol{K}$ is compact. Consider now a compact operator $\boldsymbol{T} \asymp \bigoplus_{p\in\mathbb{Z}_+} T^p$ in $\mathcal{T}(\boldsymbol{S}_{\mathrm{VO}})$. Since $\bigoplus_{p\in\mathbb{Z}_+} T^p$ is compact we have that

$$\lim_{p\to\infty} \|T^p\| = 0.$$

Then

$$|\nu(\boldsymbol{T})| = \lim_{p\to\infty} |\mathcal{B}_p(T^p)| \le \lim_{p\to\infty} \|T^p\| = 0,$$

then $\boldsymbol{T} \in \mathcal{K}_{\mathrm{VO}}$. □

This implies that the Calkin algebra

$$\mathcal{T}(\boldsymbol{S}_{\mathrm{VO}})/(\mathcal{T}(\boldsymbol{S}_{\mathrm{VO}}) \cap \mathcal{K}(H^2(S^{2n-1})))$$

is isomorphic to $\mathrm{VO}_\partial(\mathbb{B}^{n-1})$ via the induced mapping

$$\hat{\nu} : \boldsymbol{T} + \Big(\mathcal{T}(\boldsymbol{S}_{\mathrm{VO}}) \cap \mathcal{K}(H^2(S^{2n-1}))\Big) \longmapsto \lim_{p\to\infty} \mathcal{B}_p(T^p).$$

Theorem 3.9 *The following list of irreducible representations of the C^*-algebra $\mathcal{T}(\boldsymbol{S}_{\mathrm{VO}})$ is complete up to unitary equivalence:*

- $\iota_p : \boldsymbol{T} \asymp \bigoplus_{p\in\mathbb{Z}_+}(T_a^k + K^k) \mapsto T_a^p + K^p$, *for* $p \in \mathbb{Z}_+$,
- $\nu_\eta : \boldsymbol{T} \longmapsto \nu(\boldsymbol{T}) = c \longmapsto c(\eta) \in \mathbb{C}$, *for* $\eta \in M(\mathrm{VO}_\partial(\mathbb{B}^{n-1}))$.

Moreover, the above representations are pairwise not unitarily equivalent.

Proof The proof of this theorem is exactly the same of [7, Theorem 5.7]. We include it here for the sake of completeness. According to [11, Proposition 2.11.2] each irreducible representation of $\mathcal{T}(\boldsymbol{S}_{\mathrm{VO}})$ is either induced by an irreducible representation of

$$\mathcal{T}(\boldsymbol{S}_{\mathrm{VO}})/(\mathcal{T}(\boldsymbol{S}_{\mathrm{VO}}) \cap \mathcal{K}(H^2(S^{2n-1}))) \tag{3.4}$$

or is an extension of an irreducible representation of $\mathcal{T}(\boldsymbol{S}_{\mathrm{VO}}) \cap \mathcal{K}(H^2(S^{2n-1}))$. In the first case, the algebra (3.4) is isomorphic to $\mathrm{VO}_\partial(\mathbb{B}^{n-1})$ and then, its irreducible representations are exactly the representations ν_η. On the other hand, by Lemma 3.8, $\mathcal{T}(\boldsymbol{S}_{\mathrm{VO}}) \cap \mathcal{K}(H^2(S^{2n-1}))$ equals $\mathcal{K}_{\mathrm{VO}}$. Then, each irreducible representation of $\mathcal{K}_{\mathrm{VO}}$, restricted to the different levels p, is either 0 or an irreducible representation of $\mathcal{K}(\mathcal{A}^2_p(\mathbb{B}^{n-1}))$. In the second case, there is only one option: the identity representation, which extends naturally to the identical representation of $\mathcal{T}_p(\mathrm{VO}_\partial(\mathbb{B}^{n-1}))$, and generates thus the representation ι_p. □

As a corollary of Lemmas 3.6 and 3.8, we have also the following result on the Fredholmness of an operator $\boldsymbol{T} \in \mathcal{T}(\boldsymbol{S}_{\mathrm{VO}})$.

Corollary 3.10 *An operator $\boldsymbol{T} \asymp \bigoplus_{p\in\mathbb{Z}_+}(T^p_a + K^p) \in \mathcal{T}(\boldsymbol{S}_{\mathrm{VO}})$ is Fredholm if and only if $a(\eta) \neq 0$ for all $\eta \in M(\mathrm{VO}_\partial(\mathbb{B}^{n-1}))$. In particular,*

$$\textit{ess-sp}\, \boldsymbol{T} = a(M(\mathrm{VO}_\partial(\mathbb{B}^{n-1}))).$$

Using Corollary 3.3 we give the following extended version of Corollary 3.10.

Proposition 3.11 *The Calkin algebra*

$$\widehat{\mathcal{T}}(\boldsymbol{S}_{\mathrm{VO}}) = \mathcal{T}(\boldsymbol{S}_{\mathrm{VO}})/(\mathcal{T}(\boldsymbol{S}_{\mathrm{VO}}) \cap \mathcal{K}(H^2(S^{2n-1})))$$

is isomorphic to $\mathrm{VO}_\partial(\mathbb{B}^{n-1})$. The corresponding homomorphism

$$\pi : \mathcal{T}(\boldsymbol{S}_{\mathrm{VO}}) \longrightarrow \widehat{\mathcal{T}}(\boldsymbol{S}_{\mathrm{VO}}) \cong \mathrm{VO}_\partial(\mathbb{B}^{n-1})$$

is given by

$$\pi : \boldsymbol{T} = \boldsymbol{T}_{\boldsymbol{a}} + \boldsymbol{K} \longmapsto a(\eta),$$

where $\boldsymbol{a} \in \boldsymbol{S}_{\mathrm{VO}}$ and $a \in \mathrm{VO}_\partial(\mathbb{B}^{n-1})$ is connected with $\boldsymbol{a}$ by (2.3).

The Toeplitz operator $\boldsymbol{T}_{\boldsymbol{a}}$, with symbol $\boldsymbol{a} \in \boldsymbol{S}_{\mathrm{VO}}$, is compact if and only if $\boldsymbol{a} \equiv 0$ on S^{2n-1}, being thus a zero operator.

The essential spectrum of each operator $\boldsymbol{T} = \boldsymbol{T}_{\boldsymbol{a}} + \boldsymbol{K}$ of the algebra $\mathcal{T}(\boldsymbol{S}_{\mathrm{VO}})$ and the spectral radius of $\boldsymbol{T}_{\boldsymbol{a}}$ are given by

$$\textit{ess-sp}\, \boldsymbol{T} = \mathrm{clos}\,(\mathrm{Range}\, \boldsymbol{a}) = \mathrm{clos}\,(\mathrm{Range}\, a)\,,$$

$$r(\boldsymbol{T}_{\boldsymbol{a}}) = \sup_{z\in S^{2n-1}} |\boldsymbol{a}(z)| = \sup_{z'\in\mathbb{B}^{n-1}} |a(z')| = \|\boldsymbol{T}_{\boldsymbol{a}}\|.$$

Each common invariant subspace for all the operators from $\mathcal{T}(\boldsymbol{S}_{\mathrm{VO}})$ *is of the form*

$$U^{-1}\left(\bigoplus_{p\in N}\mathcal{A}^2_p(\mathbb{B}^{n-1})\right),$$

where N *is a subset of* $\mathbb{Z}_+$*, and the operator* U *is defined in Theorem* 2.1.

An operator $\boldsymbol{T}=\boldsymbol{T}_{\boldsymbol{a}}+\boldsymbol{K}\in\mathcal{T}(\boldsymbol{S}_{\mathrm{VO}})$ *is Fredholm if and only if there is* $\delta=\delta(\boldsymbol{a})>0$ *such that* $|\boldsymbol{a}(z)|\geq\delta$ *for all* $z\in S^{2n-1}$ *or, equivalently,* $|a(z')|\geq\delta$ *for all* $z'\in\mathbb{B}^{n-1}$.

It is instructive to give the following.

Remark 3.12 In the seminal paper [10] L. Coburn gave a detailed description of the C^*-algebra $\mathcal{T}(C(S^{2n-1}))$ generated by all Toeplitz operators $\boldsymbol{T}_{\boldsymbol{a}}$ with continuous symbols $\boldsymbol{a}\in C(S^{2n-1})$.

The largest possible C^*-algebra that inherits the nice main properties of $\mathcal{T}(C(S^{2n-1}))$ (irreducibility, form of a generic element of the algebra: $\boldsymbol{T}_{\boldsymbol{a}}+\boldsymbol{K}$, essential spectrum formula: clos (Range $\boldsymbol{a}$), etc) has to be generated by Toeplitz operators with symbols from the largest C^*-subalgebra, say $\boldsymbol{Q}_n$, of $L_\infty(S^{2n-1})$ possessing the compact semi-commutator property.

The results of this section show that $\boldsymbol{S}_{\mathrm{VO}}\subset\boldsymbol{Q}_n$. Note that functions from $\boldsymbol{S}_{\mathrm{VO}}$ are not continuous in general, they possess quite sophisticated (of vanishing oscillation type) discontinuities at $S^{2n-1}\cap\{(z',z_n):z_n=0\}$. Further, the algebra $\boldsymbol{S}_{\mathrm{VO}}$ (apart of being a subalgebra of $\boldsymbol{Q}_n$) consists of functions that are invariant under the one-dimensional torus action on S^{2n-1}:

$$\tau\in\mathbb{T}\,:\,(z',z_n)\in S^{2n-1}\quad\longmapsto\quad(z',\tau z_n)\in S^{2n-1}.$$

This $\mathbb{T}$-invariance implies that our C^*-algebra $\mathcal{T}(\boldsymbol{S}_{\mathrm{VO}})$ (contrary to the algebra $\mathcal{T}(\boldsymbol{Q}_n)$) is not anymore irreducible, and we have described all its irreducible representations, as well as, all common invariant subspaces for its elements.

We discuss now briefly the Fredholm index formula for a matrix-valued symbol situation. As shown in [19, Section 2], the index calculation for Fredholm operators with VO_∂ symbols can be reduced to those with continuous up to the boundary symbols in the following way. Given a matrix-valued function $a\in\mathrm{Mat}_q(\mathrm{VO}_\partial(\mathbb{B}^{n-1}))\coloneqq\mathrm{VO}_\partial(\mathbb{B}^{n-1})\otimes\mathrm{Mat}_q(\mathbb{C})$, for each $s\in(0,1)$, we define

$$a_s(r\zeta)=\begin{cases}a(r\zeta), & \text{if } 0\leq r\leq s,\\ a(s\zeta), & \text{if } s<r\leq 1,\end{cases}$$

here $r=|z'|$ and $\zeta\in S^{2n-3}$. Then, obviously, $a_s(r\zeta)\in\mathrm{Mat}_q(C(\overline{\mathbb{B}}^{n-1}))$, for all $s\in(0,1)$.

The matrix-valued version of Corollary 3.10 is as follows.

Corollary 3.13 *Given $a \in \mathrm{Mat}_q(\mathrm{VO}_\partial(\mathbb{B}^{n-1}))$, the operator*

$$\boldsymbol{T} \asymp \bigoplus_{p\in\mathbb{Z}_+} (T_a^p + K^p) \in \boldsymbol{\mathcal{T}}(\boldsymbol{S}_{\mathrm{VO}}) \otimes \mathrm{Mat}_q(\mathbb{C})$$

is Fredholm if and only if the matrix $a(\eta)$ is invertible for all

$$\eta \in M(\mathrm{VO}_\partial(\mathbb{B}^{n-1})).$$

In particular,

$$\textit{ess-sp}\, \boldsymbol{T} = \left\{\lambda : \lambda \in \mathrm{sp}\{a(\eta)\},\ \eta \in M(\mathrm{VO}_\partial(\mathbb{B}^{n-1}))\right\}.$$

In case of being Fredholm,

$$\mathrm{Ind}\, \boldsymbol{T} = \mathrm{Ind}\left[\bigoplus_{p\in\mathbb{Z}_+} (T_a^p + K^p)\right] = 0.$$

Proof Only the index formula needs to be justified. Let $\boldsymbol{T}$ be Fredholm, then the matrix-valued function a is invertible, and

$$\mathrm{Ind}\, \boldsymbol{T} = \sum_{p\in\mathbb{Z}_+} \mathrm{Ind}\,(T_a^p + K^p).$$

For each $p \in \mathbb{Z}_+$, Theorem 2.6 of [19], ensures that there is $s_{\rho,0} \in (0, 1)$ such that for every $s \in (s_{\rho,0}, 1)$, each operator $T_{a_s}^p + K^p$ is Fredholm, and

$$\mathrm{Ind}\,(T_a^p + K^p) = \mathrm{Ind}\,(T_{a_s}^p + K^p).$$

The matrix-valued function a_s is continuous on the closed unit ball $\overline{\mathbb{B}}^{n-1}$. The ball $\overline{\mathbb{B}}^{n-1}$ is retractable to a point, thus the matrix-valued function a_s is homotopic to a constant matrix, say a_p, in a class of invertible continuous on $\overline{\mathbb{B}}^{n-1}$ matrix-functions. This implies that the operator $T_{a_s}^p + K^p$ is homotopic to the scalar-matrix multiplication operator $a_p I$. Thus, for each $p \in \mathbb{Z}_+$,

$$\mathrm{Ind}\,(T_a^p + K^p) = \mathrm{Ind}\, a_p I = 0. \qquad \square$$

Note that, starting from some p_0, all operators $T_a^p + K^p$ are invertible, so that both $\ker \boldsymbol{T}$ and $\mathrm{coker}\, \boldsymbol{T}$ are finite dimensional (see [5, p. 730] for details).

One, of course, easily makes a version of Proposition 3.11 for matrix-valued symbols $a \in \mathrm{Mat}_q(\mathrm{VO}_\partial(\mathbb{B}^{n-1}))$.

4 Full Toeplitz Algebra in Levels

In this section we consider the case when, for each $p \in \mathbb{Z}_+$, the C^*-algebra consisting of all operators T^p of decomposition (2.5) coincides with the full Toeplitz algebra, i.e., with the algebra $\mathcal{T}_p(L_\infty(\mathbb{B}^{n-1}))$, which is generated by all Toeplitz operators T_a^p, with $a \in L_\infty(\mathbb{B}^{n-1})$, acting on the weighted Bergman space $\mathcal{A}_p^2(\mathbb{B}^{n-1})$.

We introduce as well the C^*-algebra $\mathrm{BUC}(\mathbb{B}^{n-1})$ of functions $a(z')$, $z' \in \mathbb{B}^{n-1}$, which are uniformly continuous in $\mathbb{B}^{n-1}$ with respect to the Bergman metric $\beta(\cdot,\cdot)$. Note (see [14, Theorem 7.3]) that each operator T_a^p, with $a \in L_\infty(\mathbb{B}^{n-1})$ can be approximated in norm by Toeplitz operators with symbols from $\mathrm{BUC}(\mathbb{B}^{n-1})$. Moreover, the results of [18] imply that Toeplitz operators with symbols from $\mathrm{BUC}(\mathbb{B}^{n-1})$ form a dense set in $\mathcal{T}_p(L_\infty(\mathbb{B}^{n-1}))$. That is, in particular, $\mathcal{T}_p(\mathrm{BUC}(\mathbb{B}^{n-1})) = \mathcal{T}_p(L_\infty(\mathbb{B}^{n-1}))$.

The above results suggest us to choose $S = \mathrm{BUC}(\mathbb{B}^{n-1})$. Let $\boldsymbol{S}_{\mathrm{BUC}}$ be the corresponding class of functions on S^{2n-1}, connected with functions from $S = \mathrm{BUC}(\mathbb{B}^{n-1})$ by (2.3).

We describe briefly the C^*-algebra $\mathcal{T}(\boldsymbol{S}_{\mathrm{BUC}})$, which is generated by all Toeplitz operators $\boldsymbol{T}_{\boldsymbol{a}}$, with $\boldsymbol{a} \in \boldsymbol{S}_{\mathrm{BUC}}$, acting on the Hardy space $H^2(S^{2n-1})$. Doing this we follow the lines of [7, Section 4], where all details can be found.

We mention first that each operator $\boldsymbol{T}_{\boldsymbol{a}} \in \mathcal{T}(\boldsymbol{S}_{\mathrm{BUC}})$ admits the representation

$$\boldsymbol{T}_{\boldsymbol{a}} \asymp \bigoplus_{p\in\mathbb{Z}_+} T_a^p, \quad \text{with } T_a^p \in \mathcal{T}_p(\mathrm{BUC}(\mathbb{B}^{n-1})).$$

Similarly to Lemma 3.1 we have

$$\|\boldsymbol{T}\| = \sup_{p\in\mathbb{Z}_+} \|T^p\|,$$

while for the generating operators we have

$$\|\boldsymbol{T}_{\boldsymbol{a}}\| = \|a(z')\|_\infty.$$

where the last equality follows from [4, Proposition 4.4].

The proof of the next theorem literally follows the proof of Theorem 3.2 or [7, Theorem 4.3], and thus will be omitted.

Theorem 4.1 *Each operator $\boldsymbol{T}_{\boldsymbol{a}} \in \mathcal{T}(\boldsymbol{S}_{\mathrm{BUC}})$, in the direct sum decomposition (2.5), admits the unique representation*

$$\boldsymbol{T} \asymp \bigoplus_{p\in\mathbb{Z}_+} (T_a^p + N^p), \tag{4.1}$$

where $a \in \mathrm{BUC}(\mathbb{B}^{n-1})$, N^p *belongs to the semi-commutator ideal of the algebra* $\mathcal{T}_p(\mathrm{BUC}(\mathbb{B}^{n-1}))$, *and* $\|N^p\| \to 0$ *as* $p \to \infty$.

Corollary 4.2 *The semi-commutator ideal* $\boldsymbol{\mathcal{N}}(\boldsymbol{S}_{\mathrm{BUC}})$ *of* $\boldsymbol{\mathcal{T}}(\boldsymbol{S}_{\mathrm{BUC}})$ *consists of all its elements*

$$\boldsymbol{N} \asymp \bigoplus_{p\in\mathbb{Z}_+} N^p,$$

coming from the representation (4.1).

Denote by $\mathfrak{T}_p(\mathrm{BUC}(\mathbb{B}^{n-1}))$ the algebra which consists of all operators $T^p = T_a^p + N^p$ that appear on the pth level of decomposition (2.5), or, which is the same, which consists of all operators being the restrictions of elements (4.1) of $\boldsymbol{\mathcal{T}}(\boldsymbol{S}_{\mathrm{BUC}})$ onto its invariant subspace $\mathcal{A}_p^2(\mathbb{B}^{n-1})$.

Lemma 4.3 *For each* $p \in \mathbb{Z}_+$, *the algebra* $\mathfrak{T}_p(\mathrm{BUC}(\mathbb{B}^{n-1}))$ *coincides with the algebra* $\mathcal{T}_p(\mathrm{BUC}(\mathbb{B}^{n-1})) = \mathcal{T}_p(L_\infty(\mathbb{B}^{n-1}))$.

Proof It is easy to see that the mapping

$$\begin{aligned} \iota_p : \boldsymbol{\mathcal{T}}(\boldsymbol{S}_{\mathrm{BUC}}) &\longrightarrow \mathcal{L}(\mathcal{A}_p^2(\mathbb{B}^{n-1})) \\ \boldsymbol{T} \asymp \bigoplus_{k\in\mathbb{Z}_+} T^k &\longmapsto T^p \end{aligned}$$

is a morphism (representation) of the algebra $\boldsymbol{\mathcal{T}}(\boldsymbol{S}_{\mathrm{BUC}})$. Besides, its image is $\mathfrak{T}_p(\mathrm{BUC}(\mathbb{B}^{n-1}))$. Each operator T_a^p with $a \in \mathrm{BUC}(\mathbb{B}^{n-1})$ obviously belongs to $\mathfrak{T}_p(\mathrm{BUC}(\mathbb{B}^{n-1}))$. As it was already mentioned, the set of all such operators is norm dense in $\mathcal{T}_p(L_\infty(\mathbb{B}^{n-1}))$, and the algebra $\mathfrak{T}_p(\mathrm{BUC}(\mathbb{B}^{n-1}))$, being the image of a representation, is norm closed. □

It is instructive to mention that, in spite of the fact that for each $p \in \mathbb{Z}_+$, the algebra $\mathfrak{T}_p(\mathrm{BUC}(\mathbb{B}^{n-1})) = \mathcal{T}_p(\mathrm{BUC}(\mathbb{B}^{n-1}))$ *contains each* Toeplitz operator T_a^p with $a \in L_\infty(\mathbb{B}^{n-1})$, the algebra $\boldsymbol{\mathcal{T}}(\boldsymbol{S}_{\mathrm{BUC}})$ *does not contain* any Toeplitz operator $\boldsymbol{T_a}$ with $\boldsymbol{a}$ from $\boldsymbol{S}_\infty \setminus \boldsymbol{S}_{\mathrm{BUC}}$. That is, the algebra $\boldsymbol{\mathcal{T}}(\boldsymbol{S}_{\mathrm{BUC}})$ is a proper subalgebra of $\boldsymbol{\mathcal{T}}(\boldsymbol{S}_\infty)$. This can be justified literally following the arguments of the proof of [7, Lemma 4.5].

We describe now the irreducible representations of the C^*-algebra $\boldsymbol{\mathcal{T}}(\boldsymbol{S}_{\mathrm{BUC}})$. First, for each $p \in \mathbb{Z}_+$, the representation

$$\begin{aligned} \iota_p : \boldsymbol{\mathcal{T}}(\boldsymbol{S}_{\mathrm{BUC}}) &\longrightarrow \mathcal{L}(\mathcal{A}_p^2(\mathbb{B}^{n-1})) \\ \boldsymbol{T} \asymp \bigoplus_{k\in\mathbb{Z}_+} (T_a^k + N^p) &\longmapsto T_a^p + N^p, \end{aligned} \tag{4.2}$$

whose image is $\mathcal{T}_p(\mathrm{BUC}(\mathbb{B}^{n-1}))$, is irreducible (as the algebra $\mathcal{T}_p(\mathrm{BUC}(\mathbb{B}^{n-1}))$ contains the whole ideal of compact operators). The same reasoning as in case of $\mathcal{T}(\boldsymbol{S}_{\mathrm{VO}})$ implies that for different $p \in \mathbb{Z}_+$ these representations are not unitary equivalent.

Similarly to Lemma 3.6 and [7, Lemma 4.8], we have that the mapping

$$\nu \,:\, \mathcal{T}(\boldsymbol{S}_{\mathrm{BUC}}) \longrightarrow \mathrm{BUC}(\mathbb{B}^{n-1}) = C(M(\mathrm{BUC}(\mathbb{B}^{n-1}))),$$

defined by

$$\nu \,:\, \boldsymbol{T} \asymp \bigoplus_{p\in\mathbb{Z}_+} T^p \longmapsto \lim_{p\to\infty} \mathcal{B}_p(T^p),$$

where $\mathcal{B}_p$ is the Berezin transform, is a continuous $*$-homomorphism of the C^*-algebra $\mathcal{T}(\boldsymbol{S}_{\mathrm{BUC}})$ onto $C(M(\mathrm{BUC}(\mathbb{B}^{n-1})))$. Here $M(\mathrm{BUC}(\mathbb{B}^{n-1}))$ denotes the compact of maximal ideals of the algebra $\mathrm{BUC}(\mathbb{B}^{n-1})$.

As in [7, Remark 4.10], given $\boldsymbol{T} \asymp \bigoplus_{p\in\mathbb{Z}_+} T^p \in \mathcal{T}(\boldsymbol{S}_{\mathrm{BUC}})$, the above result permits us to recover the data of its unique representation (4.1):

$$a = \nu(\boldsymbol{T}^\lambda) = \lim_{p\to\infty} \mathcal{B}_p(T^p) \in \mathrm{BUC}(\mathbb{B}^{n-1}) \quad \text{and} \quad N^p = T^p - T_a^p.$$

Then we have

Corollary 4.4 *For each $\eta \in M(\mathrm{BUC}(\mathbb{B}^{n-1}))$, the mapping*

$$\nu_\eta \,:\, \mathcal{T}(\boldsymbol{S}_{\mathrm{BUC}}) \longrightarrow \mathbb{C},$$

defined by

$$\nu_\eta \,:\, \boldsymbol{T} \longmapsto \nu(\boldsymbol{T}) = a \longmapsto a(\eta) \in \mathbb{C}, \tag{4.3}$$

is a one-dimensional representation of the C^*-algebra $\mathcal{T}(\boldsymbol{S}_{\mathrm{BUC}})$.

We gather all so far obtained information on the C^*-algebra $\mathcal{T}(\boldsymbol{S}_{\mathrm{BUC}})$ in the following proposition.

Proposition 4.5 *The quotient algebra $\widetilde{\mathcal{T}}(\boldsymbol{S}_{\mathrm{BUC}}) = \mathcal{T}(\boldsymbol{S}_{\mathrm{BUC}})/\mathcal{N}(\boldsymbol{S}_{\mathrm{BUC}})$ of $\mathcal{T}(\boldsymbol{S}_{\mathrm{BUC}})$ by its semi-commutator ideal is isomorphic to $\mathrm{BUC}(\mathbb{B}^{n-1})$. The corresponding homomorphism*

$$\widetilde{\pi} \,:\, \mathcal{T}(\boldsymbol{S}_{\mathrm{VO}}) \longrightarrow \widetilde{\mathcal{T}}(\boldsymbol{S}_{\mathrm{VO}}) \cong \mathrm{BUC}(\mathbb{B}^{n-1})$$

is given by

$$\widetilde{\pi} : \boldsymbol{T} \asymp \bigoplus_{p\in\mathbb{Z}_+} (T_a^p + N^p) \longmapsto a(\eta),$$

where $\boldsymbol{a} \in \boldsymbol{S}_{\mathrm{BUC}}$ *and* $a \in \mathrm{BUC}(\mathbb{B}^{n-1})$ *is connected with* $\boldsymbol{a}$ *by* (2.3).

The Toeplitz operator $\boldsymbol{T}_{\boldsymbol{a}}$*, with symbol* $\boldsymbol{a} \in \boldsymbol{S}_{\mathrm{BUC}}$*, is compact if and only if* $\boldsymbol{a} \equiv 0$ *on* S^{2n-1}*, being thus a zero operator.*

The spectrum of each operator $\boldsymbol{T} = \bigoplus_{p\in\mathbb{Z}_+} (T_a^p + N^p)$ *of the algebra* $\mathcal{T}(\boldsymbol{S}_{\mathrm{BUC}})$ *contains* clos (Range a)*, and the spectral radius of* $\boldsymbol{T}_{\boldsymbol{a}}$ *is given by*

$$r(\boldsymbol{T}_{\boldsymbol{a}}) = \sup_{z\in S^{2n-1}} |\boldsymbol{a}(z)| = \sup_{z'\in\mathbb{B}^{n-1}} |a(z')| = \|\boldsymbol{T}_{\boldsymbol{a}}\|.$$

The C^**-algebra* $\mathcal{T}(\boldsymbol{S}_{\mathrm{BUC}})$ *is reducible, its irreducible representations are given by* (4.2) *and* (4.3)*. Each common invariant subspace for all the operators from* $\mathcal{T}(\boldsymbol{S}_{\mathrm{BUC}})$ *is of the form*

$$U^{-1}\left(\bigoplus_{p\in N} \mathcal{A}_p^2(\mathbb{B}^{n-1})\right),$$

where N *is a subset of* $\mathbb{Z}_+$*, and the operator* U *is defined in Theorem* 2.1.

Acknowledgement This work was partially supported by CONACYT Project 238630, México.

References

1. Z. Akkar, Zur Spektraltheorie von Toeplitzoperatoren auf dem Hardyraum $H^2(\mathbb{B}^n)$. Ph.D. Dissertation, Universität des Saarlandes, 2012
2. Z. Akkar, E. Albrecht, Spectral properties of Toeplitz operators on the unit ball and on the unit sphere, in *The Varied Landscape of Operator Theory*. Theta Series in Advanced Mathematics, vol. 17 (Theta, Bucharest, 2014), pp. 1–22
3. H. Alzer, Inequalities for the Beta function of n variables. ANZIAM J. **44**, 609–623 (2003)
4. W. Bauer, L.A. Coburn, Heat flow, weighted Bergman spaces and real analytic Lipschitz approximation. J. Reine Angew. Math. **703**, 225–246 (2015)
5. W. Bauer, N. Vasilevski, On algebras generated by Toeplitz operators and their representations. J. Funct. Anal. **272**, 705–737 (2017)
6. W. Bauer, R. Hagger, N. Vasilevski, Uniform continuity and quantization on bounded symmetric domains. J. Lond. Math. Soc. **96**, 345–366 (2017)
7. W. Bauer, R. Hagger, N. Vasilevski, Algebras of Toeplitz operators on the n-dimensional unit ball. Complex Anal. Oper. Theory **13**, 493–524 (2019)
8. C.A. Berger, L.A. Coburn, K.H. Zhu, Function theory on Cartan domains and the Berezin-Toeplitz symbol calculus. Amer. J. Math. **110**, 921–953 (1988)
9. A. Brown, P.R. Halmos, Algebraic properties of Toeplitz operators. J. Reine Angew. Math. **213**, 89–102 (1964)

10. L. Coburn, Singular integral operators and Toeplitz operators on odd spheres. Indiana Univ. Math. J. **23**, 433–439 (1973)
11. J. Dixmier, *Les C^*-algebres et leurs représensations* (Gauthier-Villars, Paris 1964)
12. M. Loaiza, N. Vasilevski, Commutative algebras generated by Toeplitz operators on the unit sphere. Intgr. Equ. Oper. Theory **92**, 25 (2020). https://doi.org/10.1007/s00020-020-02580-x
13. R. Quiroga-Barranco, N.L. Vasilevski, Commutative algebras of Toeplitz operators on the unit ball I: Bargmann type transforms and spectral representations of Toeplitz operators. Integr. Equ. Oper. Theory **59**, 379–419 (2007)
14. D. Suárez, The essential norm of operators in the Toeplitz algebra $A^p(\mathbb{B}_n)$. Indiana Univ. Math. J. **56**, 2185–2232 (2007)
15. N.L. Vasilevski, Toeplitz operators on the Bergman spaces: inside-the-domain effects. Contemp. Math. **289**, 79–146 (2001)
16. N.L. Vasilevski, Bergman space structure, commutative algebras of Toeplitz operators and hyperbolic geometry. Integr. Equ. Oper. Theory **46**, 235–251 (2003)
17. N.L. Vasilevski, *Commutative Algebras of Toeplitz Operators on the Bergman Space*. Operator Theory: Advances and Applications, vol. 185 (Birkhäuser Verlag, Boston, 2008)
18. J. Xia, Localization and the Toeplitz algebra on the Bergman space. J. Funct. Anal. **269**, 781–814 (2015)
19. J. Xia, D. Zheng, Toeplitz operators and Toeplitz algebra with symbols of vanishing oscillation. J. Operator Theory **76**, 107–131 (2016)

d-Modified Riesz Potentials on Central Campanato Spaces

Katsuo Matsuoka

Dedicated to Professor Lars-Erik Persson in celebration of his 75th birthday

Abstract Recently, we defined the d-modified Riesz potentials $\tilde{I}_{\alpha,d}$ and proved several estimates of boundedness of $\tilde{I}_{\alpha,d}$ on the central Morrey spaces $B^{p,\lambda}(\mathbb{R}^n)$, using the central Campanato spaces $\Lambda^{(d)}_{p,\lambda}(\mathbb{R}^n)$, the generalized σ-Lipschitz spaces $\mathrm{Lip}^{(d)}_{\beta,\sigma}(\mathbb{R}^n)$ and so on. In this paper, we will consider the results of the boundedness for $\tilde{I}_{\alpha,d}$ on the λ-central mean oscillation spaces $\mathrm{CMO}^{p,\lambda}(\mathbb{R}^n)$.

Keywords Central Morrey space · λ-central mean oscillation space · Weak λ-central mean oscillation space · σ-Lipschitz space · σ-BMO space · Central Campanato space · Weak central Campanato space · Generalized σ-Lipschitz space · Generalized σ-BMO space · Riesz potential · Modified Riesz potential · d-modified Riesz potential

Mathematics Subject Classification (2010) Primary 42B35; Secondary 26A33, 46E30, 46E35

This work was supported by Grant-in-Aid for Scientific Research (C) (Grant No. 17K05306), Japan Society for the Promotion of Science.

K. Matsuoka (✉)
College of Economics, Nihon University, Misaki-cho, Kanda, Chiyoda-ku, Tokyo, Japan
e-mail: katsu.m@nihon-u.ac.jp

M. A. Bastos et al. (eds.), *Operator Theory, Functional Analysis and Applications*,
Operator Theory: Advances and Applications 282,
https://doi.org/10.1007/978-3-030-51945-2_21

1 Introduction

Let I_α and $\tilde{I}_\alpha$, $0 < \alpha < n$, be the Riesz potential and the modified Riesz potential, respectively, which are defined for $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ by

$$I_\alpha f(x) = \int_{\mathbb{R}^n} \frac{f(y)}{|x-y|^{n-\alpha}}\, dy$$

and

$$\tilde{I}_\alpha f(x) = \int_{\mathbb{R}^n} f(y) \left(\frac{1}{|x-y|^{n-\alpha}} - \frac{1-\chi_{Q_1}(y)}{|y|^{n-\alpha}} \right) dy$$

(as for the notation χ_{Q_1}, see Sect. 2). Then, the following boundedness results on $L^p(\mathbb{R}^n)$ are well-known: For $0 < \alpha < n$, $1 \le p < n/\alpha$ and $1/q = 1/p - \alpha/n$,

(i) $I_\alpha : L^p(\mathbb{R}^n) \to L^q(\mathbb{R}^n)$, $\quad 1 < p < n/\alpha$ (see [7, 25]);
(ii) $I_\alpha : L^1(\mathbb{R}^n) \to WL^q(\mathbb{R}^n)$, $\quad p = 1$ (see [26]).

And for $0 < \alpha < n$, $n/\alpha \le p < \infty$ and $\beta = \alpha - n/p < 1$,

(iii) $\tilde{I}_\alpha : L^p(\mathbb{R}^n) \to \mathrm{Lip}_\beta(\mathbb{R}^n)$, $\quad n/\alpha < p < \infty$ (cf. [23]);
(iv) $\tilde{I}_\alpha : L^{n/\alpha}(\mathbb{R}^n) \to \mathrm{BMO}(\mathbb{R}^n)$, $\quad p = n/\alpha$, i.e., $\beta = 0$ (cf. [23, 26]).

Here (i) is the so-called Hardy–Littlewood–Sobolev theorem.

After that, in [4], for I_α on the (non-homogeneous) central Morrey space $B^{p,\lambda}(\mathbb{R}^n)$, i.e.,

$$B^{p,\lambda}(\mathbb{R}^n) = \{ f \in L^p_{loc}(\mathbb{R}^n) : \|f\|_{B^{p,\lambda}} < \infty \},$$

where $1 \le p < \infty$, $\lambda \in \mathbb{R}$ and

$$\|f\|_{B^{p,\lambda}} = \sup_{r \ge 1} \frac{1}{r^\lambda} \left(\frac{1}{|Q_r|} \int_{Q_r} |f(y)|^p\, dy \right)^{1/p}$$

(see [1, 4]; cf. [6]), which is the (non-homogeneous) Herz space $K^{-n/p-\lambda}_{p,\infty}(\mathbb{R}^n)$ (cf. [3, 8]), the following boundedness result was obtained: For $0 < \alpha < n$, $1 < p < n/\alpha$, $-n/p + \alpha \le \mu = \lambda + \alpha < 0$ and $1/q = 1/p - \alpha/n$,

(i'-1) $I_\alpha : B^{p,\lambda}(\mathbb{R}^n) \to B^{q,\mu}(\mathbb{R}^n)$.

On the other hand, in [11, 19] (cf. [10]), from the B_σ-Morrey–Campanato estimates for I_α and $\tilde{I}_\alpha$, the following boundedness results on $B^{p,\lambda}(\mathbb{R}^n)$ were obtained as the corollaries: For $0 < \alpha < n$, $-n + \alpha \le \mu = \lambda + \alpha < 0$ and $1/q = 1 - \alpha/n$,

(ii'-1) $I_\alpha : B^{1,\lambda}(\mathbb{R}^n) \to WB^{q,\mu}(\mathbb{R}^n)$.

Here $WB^{p,\lambda}(\mathbb{R}^n)$ is a weak central Morrey space, i.e.,

$$WB^{p,\lambda}(\mathbb{R}^n) = \{f \in L^p_{loc}(\mathbb{R}^n) : \|f\|_{WB^{p,\lambda}} < \infty\},$$

where $1 \le p < \infty$, $\lambda \in \mathbb{R}$ and

$$\|f\|_{WB^{p,\lambda}} = \sup_{r\ge 1} \frac{1}{r^\lambda} \left(\frac{1}{|Q_r|} \sup_{t>0} t^p \, |\{y \in Q_r : |f(y)| > t\}| \right)^{1/p}$$

(see [11]). Also for $0 < \alpha < n$, $1 \le p < n/\alpha$, $-n/p + \alpha \le \mu = \lambda + \alpha < 1$ and $1/q = 1/p - \alpha/n$,

(i'-2) $\tilde{I}_\alpha : B^{p,\lambda}(\mathbb{R}^n) \to \mathrm{CMO}^{q,\mu}(\mathbb{R}^n)$, $\quad 1 < p < n/\alpha$;
(ii'-2) $\tilde{I}_\alpha : B^{1,\lambda}(\mathbb{R}^n) \to W\mathrm{CMO}^{q,\mu}(\mathbb{R}^n)$, $\quad p = 1$.

And for $0 < \alpha < n$, $n/\alpha \le p < \infty$, $-n/p + \alpha \le \lambda + \alpha < 1$ and $\beta = \alpha - n/p$,

(iii') $\tilde{I}_\alpha : B^{p,\lambda}(\mathbb{R}^n) \to \mathrm{Lip}_{\beta,\lambda+n/p}(\mathbb{R}^n)$, $\quad n/\alpha < p < \infty$;
(iv') $\tilde{I}_\alpha : B^{n/\alpha,\lambda}(\mathbb{R}^n) \to \mathrm{BMO}_{\lambda+n/p}(\mathbb{R}^n)$, $\quad p = n/\alpha$, i.e., $\beta = 0$.

As for the spaces $\mathrm{CMO}^{p,\lambda}(\mathbb{R}^n)$, $W\mathrm{CMO}^{p,\lambda}(\mathbb{R}^n)$ and $\mathrm{Lip}_{\beta,\sigma}(\mathbb{R}^n)$, $\mathrm{BMO}_\sigma(\mathbb{R}^n)$, refer to Definitions 2.1, 2.2 and 2.11. Furthermore, the following boundedness results on $\mathrm{CMO}^{p,\lambda}(\mathbb{R}^n)$ were also gotten excepting (ii") (as for (ii"), see Corollary 3.5): For $0 < \alpha < 1$, $1 \le p < n/\alpha$, $-n/p + \alpha \le \lambda + \alpha = \mu < 1$ and $1/q = 1/p - \alpha/n$ (see [19]; cf. [16]),

(i") $\tilde{I}_\alpha : \mathrm{CMO}^{p,\lambda}(\mathbb{R}^n) \to \mathrm{CMO}^{q,\mu}(\mathbb{R}^n)$, $\quad 1 < p < n/\alpha$;
(ii") $\tilde{I}_\alpha : \mathrm{CMO}^{1,\lambda}(\mathbb{R}^n) \to W\mathrm{CMO}^{q,\mu}(\mathbb{R}^n)$, $\quad p = 1$.

And for $0 < \alpha < 1$, $n/\alpha \le p < \infty$, $-n/p + \alpha \le \mu = \lambda + \alpha < 1$ and $\beta = \alpha - n/p$,

(iii") $\tilde{I}_\alpha : \mathrm{CMO}^{p,\lambda}(\mathbb{R}^n) \to \mathrm{Lip}_{\beta,\lambda+n/p}(\mathbb{R}^n)$, $\quad n/\alpha < p < \infty$;
(iv") $\tilde{I}_\alpha : \mathrm{CMO}^{n/\alpha,\lambda}(\mathbb{R}^n) \to \mathrm{BMO}_{\lambda+n/p}(\mathbb{R}^n)$, $\quad p = n/\alpha$, i.e., $\beta = 0$.

Recentry, for the whole of $\mu = \lambda + \alpha$ such that $1 \le \mu < \infty$, we extended the boundedness of $\tilde{I}_\alpha$ for $B^{p,\lambda}(\mathbb{R}^n)$. In order to do so, in [17, 18], we introduced the central Campanato spaces $\Lambda^{(d)}_{q,\mu}(\mathbb{R}^n)$ and the generalized σ-Lip spaces $\mathrm{Lip}^{(d)}_{\beta,\sigma}(\mathbb{R}^n)$ (see Definitions 2.3 and 2.11), and also we defined the "higher-degree" modification of I_α, i.e., the d-modified Riesz potentials $\tilde{I}_{\alpha,d}$, $0 < \alpha < n$ and $d \in \mathbb{N} \cup \{0\}$. Then, the following boundedness results were shown: For $0 < \alpha < n$, $1 < p < n/\alpha$, $d \in \mathbb{N} \cup \{0\}$, $-n/p + \alpha \le \mu = \lambda + \alpha < d + 1$ and $1/q = 1/p - \alpha/n$,

(I) $\tilde{I}_{\alpha,d} : B^{p,\lambda}(\mathbb{R}^n) \to \Lambda^{(d)}_{q,\mu}(\mathbb{R}^n)$, $\quad 1 < p < n/\alpha$;
(II) $\tilde{I}_{\alpha,d} : B^{1,\lambda}(\mathbb{R}^n) \to W\Lambda^{(d)}_{q,\mu}(\mathbb{R}^n)$, $\quad p = 1$.

And for $0 < \alpha < n$, $n/\alpha \le p < \infty$, $d \in \mathbb{N} \cup \{0\}$, $-n/p + \alpha + d \le \lambda + \alpha < d + 1$ and $\beta = \alpha - n/p$,

(III) $\tilde{I}_{\alpha,d} : B^{p,\lambda}(\mathbb{R}^n) \to \mathrm{Lip}^{(d)}_{\beta,\lambda+n/p}(\mathbb{R}^n), \quad n/\alpha < p < \infty;$

(IV) $\tilde{I}_{\alpha,d} : B^{n/\alpha,\lambda}(\mathbb{R}^n) \to \mathrm{BMO}^{(d)}_{\lambda+n/p}(\mathbb{R}^n), \quad p = n/\alpha$, i.e., $\beta = 0$.

In this paper, therefore, we will investigate the extension of the above boundedness results of $\tilde{I}_{\alpha,d}$ on $B^{p,\lambda}(\mathbb{R}^n)$, i.e., the boundedness of $\tilde{I}_{\alpha,d}$ for $\mathrm{CMO}^{p,\lambda}(\mathbb{R}^n)$. As a by-product, we can obtain the estimates of $\tilde{I}_\alpha$ for $B^{p,\lambda}(\mathbb{R}^n)$.

We note that the same results in this paper still hold for the homogeneous versions of the function spaces.

2 Central Campanato Spaces and Generalized σ-Lip Spaces

We start by explaining the notation used in the present paper. The symbol $A \lesssim B$ stands for $A \le CB$ for some constant $C > 0$. If $A \lesssim B$ and $B \lesssim A$, we then write $A \sim B$. For $r > 0$, by Q_r, we mean the following: $Q_r = \{y \in \mathbb{R}^n : |y| < r\}$ or $Q_r = \{y = (y_1, y_2, \cdots, y_n) \in \mathbb{R}^n : \max_{1\le i\le n} |y_i| < r\}$. Further, for $x \in \mathbb{R}^n$, we set $Q(x,r) = x + Q_r = \{x + y : y \in Q_r\}$. For a measurable set $G \subset \mathbb{R}^n$, we denote by $|G|$ and χ_G the Lebesgue measure of G and the characteristic function of G, respectively. And also, for a function $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ and a measurable set $G \subset \mathbb{R}^n$ with $0 < |G| < \infty$, let

$$f_G = \fint_G f(y)\,dy = \frac{1}{|G|}\int_G f(y)\,dy,$$

and let $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$.

First of all, we define the λ-central mean oscillation (λ-CMO) spaces $\mathrm{CMO}^{p,\lambda}(\mathbb{R}^n)$ and the weak λ-CMO spaces $W\mathrm{CMO}^{p,\lambda}(\mathbb{R}^n)$ (see [1, 15, 16]; cf. [9, 10]) and introduce the central Campanato spaces $\Lambda^{(d)}_{p,\lambda}(\mathbb{R}^n)$ and the weak central Campanato spaces $W\Lambda^{(d)}_{p,\lambda}(\mathbb{R}^n)$ (see [6, 12, 17]; cf. [22]).

Definition 2.1 For $1 \le p < \infty$ and $\lambda \in \mathbb{R}$,

$$\mathrm{CMO}^{p,\lambda}(\mathbb{R}^n) = \{f \in L^p_{loc}(\mathbb{R}^n) : \|f\|_{\mathrm{CMO}^{p,\lambda}} < \infty\},$$

where

$$\|f\|_{\mathrm{CMO}^{p,\lambda}} = \sup_{r\ge 1} \frac{1}{r^\lambda}\left(\fint_{Q_r} |f(y) - f_{Q_r}|^p\,dy\right)^{1/p}.$$

Particularly

$$\mathrm{CMO}^{p,0}(\mathbb{R}^n) = \mathrm{CMO}^p(\mathbb{R}^n).$$

Definition 2.2 For $1 \le p < \infty$ and $\lambda \in \mathbb{R}$,

$$W\mathrm{CMO}^{p,\lambda}(\mathbb{R}^n) = \{f \in L^p_{loc}(\mathbb{R}^n) : \|f\|_{W\mathrm{CMO}^{p,\lambda}} < \infty\},$$

where

$$\|f\|_{W\mathrm{CMO}^{p,\lambda}} = \sup_{r\ge 1} \frac{1}{r^\lambda} \left(\frac{1}{|Q_r|} \sup_{t>0} t^p |\{y \in Q_r : |f(y) - f_{Q_r}| > t\}| \right)^{1/p},$$

Particularly

$$W\mathrm{CMO}^{p,0}(\mathbb{R}^n) = W\mathrm{CMO}^p(\mathbb{R}^n).$$

Here remark that $\mathrm{CMO}^p(\mathbb{R}^n)$ and $W\mathrm{CMO}^p(\mathbb{R}^n)$, so-called the central mean oscillation (CMO) space and the weak CMO space, are defined by

$$\mathrm{CMO}^p(\mathbb{R}^n) = \{f \in L^p_{loc}(\mathbb{R}^n) : \|f\|_{\mathrm{CMO}^p} < \infty\},$$

where

$$\|f\|_{\mathrm{CMO}^p} = \sup_{r\ge 1} \left(\fint_{Q_r} |f(y) - f_{Q_r}|^p \, dy \right)^{1/p},$$

and

$$W\mathrm{CMO}^p(\mathbb{R}^n) = \{f \in L^p_{loc}(\mathbb{R}^n) : \|f\|_{W\mathrm{CMO}^p} < \infty\},$$

where

$$\|f\|_{W\mathrm{CMO}^p} = \sup_{r\ge 1} \left(\frac{1}{|Q_r|} \sup_{t>0} t^p |\{y \in Q_r : |f(y) - f_{Q_r}| > t\}| \right)^{1/p},$$

respectively (see [2, 5]).

Definition 2.3 For $1 \le p < \infty$, $d \in \mathbb{N}_0$ and $-n/p \le \lambda < d+1$, a function $f \in L^p_{loc}(\mathbb{R}^n)$ will be said to belong to the central Campanato space $\Lambda^{(d)}_{p,\lambda}(\mathbb{R}^n)$ if and only if for every $r \ge 1$, there is a polynomial $P^d_r f$ of degree at most d such that

$$\|f\|_{\Lambda^{(d)}_{p,\lambda}} = \sup_{r\ge 1} \frac{1}{r^\lambda} \left(\fint_{Q_r} |f(y) - P^d_r f(y)|^p \, dy \right)^{1/p} < \infty.$$

Particularly

$$\Lambda^{(0)}_{p,\lambda}(\mathbb{R}^n) = \mathrm{CMO}^{p,\lambda}(\mathbb{R}^n).$$

Definition 2.4 For $1 \le p < \infty$, $d \in \mathbb{N}_0$ and $-n/p \le \lambda < d+1$, a function $f \in L^p_{loc}(\mathbb{R}^n)$ will be said to belong to the weak central Campanato space $W\Lambda^{p,\lambda}(\mathbb{R}^n)$ if and only if for every $r \ge 1$, there is a polynomial $P^d_r f$ of degree at most d such that

$$\|f\|_{W\Lambda^{p,\lambda}} = \sup_{r\ge 1} \frac{1}{r^\lambda} \left(\frac{1}{|Q_r|} \sup_{t>0} t^p |\{y \in Q_r : |f(y) - P^d_r f(y)| > t\}| \right)^{1/p} < \infty.$$

Particularly

$$W\Lambda^{(0)}_{p,\lambda}(\mathbb{R}^n) = W\mathrm{CMO}^{p,\lambda}(\mathbb{R}^n).$$

Here we note that identifying functions which differ by a polynomial of degree at most d, a.e., we see that $\Lambda^{(d)}_{p,\lambda}(\mathbb{R}^n)$ and $W\Lambda^{(d)}_{p,\lambda}(\mathbb{R}^n)$ are Banach and quasi-Banach spaces, respectively.

Remark 2.5 (Remark 6.2 of [22]) For $1 \le p < \infty$, $d \in \mathbb{N}_0$, $-n/p \le \lambda < d+1$ and $f \in L^p_{loc}(\mathbb{R}^n)$, we have

$$\|f\|_{\Lambda^{(d)}_{p,\lambda}} \sim \sup_{r\ge 1} \inf_{P\in\mathcal{P}^d(\mathbb{R}^n)} \frac{1}{r^\lambda} \left(\fint_{Q_r} |f(y) - P(y)|^p \, dy \right)^{1/p}$$

and

$$\|f\|_{W\Lambda^{p,\lambda}} \sim \sup_{r\ge 1} \inf_{P\in\mathcal{P}^d(\mathbb{R}^n)} \frac{1}{r^\lambda} \left(\frac{1}{|Q_r|} \sup_{t>0} t^p |\{y \in Q_r : |f(y) - P(y)| > t\}| \right)^{1/p},$$

where $\mathcal{P}^d(\mathbb{R}^n)$ is the set of all polynomials having degree at most d.

Next we define the generalized σ-Lipschitz spaces $\mathrm{Lip}^{(d)}_{\beta,\sigma}(\mathbb{R}^n)$.

Definition 2.6 (Definition 8.1 of [22]; cf. [12, 18]) Let $U = \mathbb{R}^n$ or $U = Q_r$ with $r > 0$. For $d \in \mathbb{N}_0$ and $0 \le \beta \le 1$, a continuous function f will be said to belong to the generalized Lipschitz (Lip) space on U, i.e., $\mathrm{Lip}^{(d)}_\beta(U)$ if and only if

$$\|f\|_{\mathrm{Lip}^{(d)}_\beta(U)} = \sup_{x, x+h\in U, h\ne 0} \frac{1}{|h|^\beta} |\triangle^{d+1}_h f(x)| < \infty,$$

where $\triangle_h^k$ is a difference operator, which is defined inductively by

$$\triangle_h^0 f = f, \quad \triangle_h^1 f = \triangle_h f = f(\cdot + h) - f(\cdot),$$

$$\triangle_h^k f = \triangle_h^{k-1} f(\cdot + h) - \triangle_h^{k-1} f(\cdot), \quad k = 2, 3, \cdots.$$

In particular, we define

$$\mathrm{BMO}^{(d)}(U) = \mathrm{Lip}_0^{(d)}(U),$$

which we call the generalized BMO space on U.

Remark 2.7 Let $U = \mathbb{R}^n$ or $U = Q_r$ with $r > 0$. For $0 < \beta < 1$, $d \in \mathbb{N}_0$ and $\beta = 1$, $d \in \mathbb{N}$, the spaces $\mathrm{Lip}_\beta^{(d)}(U)$ coincide with $\mathrm{Lip}_\beta^{(0)}(U) = \mathrm{Lip}_\beta(U)$ (see Remark 2.12) and $\mathrm{Lip}_1^{(1)}(U)$, respectively, which are the so-called Nikol'skiĭ spaces (cf. Remark 2.4 of [18, 24]).

Remark 2.8 (Theorem 8.3 of [22]) Let $U = \mathbb{R}^n$ or $U = Q_r$ with $r > 0$. For $1 \le p < \infty$, $d \in \mathbb{N}_0$, $0 \le \beta \le 1$ and $f \in L_{loc}^p(\mathbb{R}^n)$, we have

$$\|f\|_{\mathrm{Lip}_\beta^{(d)}(U)} \sim \|f\|_{\mathcal{L}_{p,\beta}^{(d)}(U)},$$

where $\mathcal{L}_{p,\beta}^{(d)}(U)$ is the Campanato space on U as defined below.

Definition 2.9 Let $U = \mathbb{R}^n$ or $U = Q_r$ with $r > 0$. For $1 \le p < \infty$, $d \in \mathbb{N}_0$ and $-n/p \le \lambda < d + 1$, a function $f \in L_{loc}^p(U)$ will be said to belong to the Campanato space on U, i.e., $\mathcal{L}_{p,\lambda}^{(d)}(U)$ if and only if for every $Q(x, s) \subset U$, there is a polynomial $P_{Q(x,s)}^d f$ of degree at most d such that

$$\|f\|_{\mathcal{L}_{p,\lambda}^{(d)}(U)} = \sup_{Q(x,s)\subset U} \frac{1}{s^\lambda} \left(\fint_{Q(x,s)} |f(y) - P_{Q(x,s)}^d f(y)|^p \, dy \right)^{1/p} < \infty.$$

Particularly, when $d = 0$, we use the well-known notation

$$\mathcal{L}_{p,\lambda}(U) = \mathcal{L}_{p,\lambda}^{(0)}(U).$$

Remark 2.10 (Remark 6.2 of [22]) Let $U = \mathbb{R}^n$ or $U = Q_r$ with $r > 0$. For $1 \le p < \infty$, $d \in \mathbb{N}_0$, $-n/p \le \lambda < d + 1$ and $f \in L_{loc}^p(U)$, we have

$$\|f\|_{\mathcal{L}_{p,\lambda}^{(d)}(U)} \sim \sup_{Q(x,s)\subset U} \inf_{P\in\mathcal{P}^d(U)} \frac{1}{s^\lambda} \left(\fint_{Q(x,s)} |f(y) - P(y)|^p \, dy \right)^{1/p}.$$

Definition 2.11 (Definition 2.8 of [18]; cf. Definition 11 of [12]) For $d \in \mathbb{N}_0$, $0 \le \beta \le 1$ and $0 \le \sigma < \infty$, the continuous function f will be said to belong to the generalized σ-Lipschitz (σ-Lip) space, i.e., $\mathrm{Lip}^{(d)}_{\beta,\sigma}(\mathbb{R}^n)$ if and only if

$$\|f\|_{\mathrm{Lip}^{(d)}_{\beta,\sigma}} = \sup_{r \ge 1} \frac{1}{r^\sigma} \|f\|_{\mathrm{Lip}^{(d)}_{\beta}(Q_r)} < \infty.$$

In particular, we define

$$\mathrm{BMO}^{(d)}_{\sigma}(\mathbb{R}^n) = \mathrm{Lip}^{(d)}_{0,\sigma}(\mathbb{R}^n)$$

and

$$\mathrm{Lip}_{\beta,\sigma}(\mathbb{R}^n) = \mathrm{Lip}^{(0)}_{\beta,\sigma}(\mathbb{R}^n), \quad \mathrm{BMO}_{\sigma}(\mathbb{R}^n) = \mathrm{BMO}^{(0)}_{\sigma}(\mathbb{R}^n),$$

which we call the generalized σ-BMO space and the σ-Lip space, the σ-BMO space, respectively.

Identifying functions which differ by a polynomial of degree at most d, a.e., we see that $\mathrm{Lip}^{(d)}_{\beta}(\mathbb{R}^n)$ and $\mathrm{Lip}^{(d)}_{\beta,\sigma}(\mathbb{R}^n)$ are Banach spaces (see [12] and [22]).

Remark 2.12 ([17]; cf. [10]) We note that particularly

$$\mathrm{Lip}^{(d)}_{\beta,0}(\mathbb{R}^n) = \mathrm{Lip}^{(d)}_{\beta}(\mathbb{R}^n), \quad \mathrm{BMO}^{(d)}_{0}(\mathbb{R}^n) = \mathrm{BMO}^{(d)}(\mathbb{R}^n),$$

and

$$\mathrm{Lip}^{(0)}_{\beta}(\mathbb{R}^n) = \mathrm{Lip}_{\beta}(\mathbb{R}^n), \quad \mathrm{BMO}^{(0)}(\mathbb{R}^n) = \mathrm{BMO}(\mathbb{R}^n).$$

3 d-Modified Riesz Potentials

Now, as we stated in Sect. 1, under the condition $\mu = \lambda + \alpha \ge 1$ we consider the boundedness of modified Riesz potentials $\tilde{I}_\alpha$ on $\mathrm{CMO}^{p,\lambda}(\mathbb{R}^n)$. Then, we use the definition of "higher-degree" modification of Riesz potentials I_α, i.e., the following definition of d-modified Riesz potentials $\tilde{I}_{\alpha,d}$, $0 < \alpha < n$ and $d \in \mathbb{N}_0$ (see [13, 14, 20]; cf. Definition 3.1 of [17]).

Definition 3.1 For $0 < \alpha < n$ and $d \in \mathbb{N}_0$, we define the modified Riesz potential of order α and degree d, i.e., d-modified Riesz potential $\tilde{I}_{\alpha,d}$, as follows: For any $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ and $x \in \mathbb{R}^n$,

$$\tilde{I}_{\alpha,d} f(x) = \int_{\mathbb{R}^n} f(y) \left\{ K_\alpha(x-y) - \left(\sum_{\{l:|l|\le d\}} \frac{x^l}{l!} (D^l K_\alpha)(-y) \right) (1 - \chi_{Q_1}(y)) \right\} dy,$$

where for any $x \in \mathbb{R}^n \setminus \{0\}$,

$$K_\alpha(x) = \frac{1}{|x|^{n-\alpha}}$$

and for $x = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n$ and $l = (l_1, l_2, \cdots, l_n) \in \mathbb{N}_0^n$, $|l| = l_1 + l_2 + \cdots + l_n$, $x^l = x_1^{l_1} x_2^{l_2} \cdots x_n^{l_n}$, $l! = l_1! l_2! \cdots l_n!$ and D^l is the partial derivative of order l, i.e.,

$$D^l = (\partial/\partial x_1)^{l_1} (\partial/\partial x_2)^{l_2} \cdots (\partial/\partial x_n)^{l_n}.$$

Note that in particular

$$\tilde{I}_{\alpha,0} = \tilde{I}_\alpha$$

and that $\tilde{I}_{\alpha,d}(|f|) \not\equiv \infty$ on $\mathbb{R}^n$, if

$$\int_{\mathbb{R}^n} \frac{|f(y)|}{(1+|y|)^{n-\alpha+d+1}}\, dy < \infty$$

(cf. [21]). If $I_\alpha f$ is well-defined, then $\tilde{I}_{\alpha,d} f$ is also well-defined and $I_\alpha f - \tilde{I}_{\alpha,d} f$ is a polynomial of degree at most d.

Then our first results for a d-modified Riesz potential $\tilde{I}_{\alpha,d}$ are the following strong and weak estimates on $\mathrm{CMO}^{p,\lambda}(\mathbb{R}^n)$, where $1 \le p < n/\alpha$.

Theorem 3.2 *Let $0 < \alpha < 1$, $1 < p < n/\alpha$, $d \in \mathbb{N}_0$, $-n/p + \alpha \le \mu = \lambda + \alpha < d + 1$ and $q = pn/(n - p\alpha)$, i.e., $1/q = 1/p - \alpha/n$. Then $\tilde{I}_{\alpha,d}$ is bounded from $\mathrm{CMO}^{p,\lambda}(\mathbb{R}^n)$ to $\Lambda^{(d)}_{q,\mu}(\mathbb{R}^n)$, that is, there exists a constant $C > 0$ such that*

$$\|\tilde{I}_{\alpha,d} f\|_{\Lambda^{(d)}_{q,\mu}} \le C \|f\|_{\mathrm{CMO}^{p,\lambda}}, \quad f \in \mathrm{CMO}^{p,\lambda}(\mathbb{R}^n).$$

Theorem 3.3 *Let $0 < \alpha < 1$, $d \in \mathbb{N}_0$, $-n + \alpha \le \mu = \lambda + \alpha < d + 1$ and $q = n/(n - \alpha)$, i.e., $1/q = 1 - \alpha/n$. Then $\tilde{I}_{\alpha,d}$ is bounded from $\mathrm{CMO}^{1,\lambda}(\mathbb{R}^n)$ to $W\Lambda^{(d)}_{q,\mu}(\mathbb{R}^n)$, that is, there exists a constant $C > 0$ such that*

$$\|\tilde{I}_{\alpha,d} f\|_{W\Lambda^{(d)}_{q,\mu}} \le C \|f\|_{\mathrm{CMO}^{1,\lambda}}, \quad f \in \mathrm{CMO}^{1,\lambda}(\mathbb{R}^n).$$

In the above theorems, if $d = 0$, then we can get the following strong and weak estimates for a modified Riesz potential $\tilde{I}_{\alpha,d}$ on $\mathrm{CMO}^{p,\lambda}(\mathbb{R}^n)$.

Corollary 3.4 (Corollary 2.7 of [19]; cf. [17]) *Let $0 < \alpha < 1$, $1 < p < n/\alpha$, $-n/p + \alpha \le \mu = \lambda + \alpha < 1$ and $q = pn/(n - p\alpha)$, i.e., $1/q = 1/p - \alpha/n$. Then*

$\tilde{I}_\alpha$ is bounded from $\mathrm{CMO}^{p,\lambda}(\mathbb{R}^n)$ *to* $\mathrm{CMO}^{q,\mu}(\mathbb{R}^n)$*, that is, there exists a constant* $C > 0$ *such that*

$$\|\tilde{I}_\alpha f\|_{\mathrm{CMO}^{q,\mu}} \le C\|f\|_{\mathrm{CMO}^{p,\lambda}}, \quad f \in \mathrm{CMO}^{p,\lambda}(\mathbb{R}^n).$$

Corollary 3.5 *Let* $0 < \alpha < 1$, $-n + \alpha \le \mu = \lambda + \alpha < 1$ *and* $q = n/(n-\alpha)$*, i.e.,* $1/q = 1 - \alpha/n$*. Then* $\tilde{I}_\alpha$ *is bounded from* $\mathrm{CMO}^{1,\lambda}(\mathbb{R}^n)$ *to* $W\mathrm{CMO}^{q,\mu}(\mathbb{R}^n)$*, that is, there exists a constant* $C > 0$ *such that*

$$\|\tilde{I}_\alpha f\|_{W\mathrm{CMO}^{q,\mu}} \le C\|f\|_{\mathrm{CMO}^{1,\lambda}}, \quad f \in \mathrm{CMO}^{1,\lambda}(\mathbb{R}^n).$$

Next, for a d-modified Riesz potential $\tilde{I}_{\alpha,d}$, we prove our second result, i.e., the following estimates on $\mathrm{CMO}^{p,\lambda}(\mathbb{R}^n)$, where $n/\alpha \le p < \infty$.

Theorem 3.6 *Let* $0 < \alpha < 1$, $n/\alpha \le p < \infty$, $d \in \mathbb{N}_0$, $-n/p + \alpha + d \le \lambda + \alpha < d+1$ *and* $\beta = \alpha - n/p$*. Then* $\tilde{I}_{\alpha,d}$ *is bounded from* $\mathrm{CMO}^{p,\lambda}(\mathbb{R}^n)$ *to* $\mathrm{Lip}^{(d)}_{\beta,\lambda+n/p}(\mathbb{R}^n)$*, that is, there exists a constant* $C > 0$ *such that*

(i) *when* $n/\alpha < p < \infty$*,*

$$\|\tilde{I}_{\alpha,d} f\|_{\mathrm{Lip}^{(d)}_{\beta,\lambda+n/p}} \le C\|f\|_{\mathrm{CMO}^{p,\lambda}}, \quad f \in \mathrm{CMO}^{p,\lambda}(\mathbb{R}^n);$$

(ii) *when* $p = n/\alpha$*, i.e.,* $\beta = 0$*,*

$$\|\tilde{I}_{\alpha,d} f\|_{\mathrm{BMO}^{(d)}_{\lambda+\alpha}} \le C\|f\|_{\mathrm{CMO}^{n/\alpha,\lambda}}, \quad f \in \mathrm{CMO}^{n/\alpha,\lambda}(\mathbb{R}^n).$$

In Theorem 3.6, if $d = 0$, then we can obtain the following estimates for a modified Riesz potential $\tilde{I}_\alpha$ on $\mathrm{CMO}^{p,\lambda}(\mathbb{R}^n)$.

Corollary 3.7 *Let* $0 < \alpha < 1$, $n/\alpha \le p < \infty$, $-n/p + \alpha \le \lambda + \alpha < 1$ *and* $\beta = \alpha - n/p$*. Then* $\tilde{I}_{\alpha,d}$ *is bounded from* $\mathrm{CMO}^{p,\lambda}(\mathbb{R}^n)$ *to* $\mathrm{Lip}_{\beta,\lambda+n/p}(\mathbb{R}^n)$*, that is, there exists a constant* $C > 0$ *such that*

(i) *when* $n/\alpha < p < \infty$*,*

$$\|\tilde{I}_{\alpha,d} f\|_{\mathrm{Lip}_{\beta,\lambda+n/p}} \le C\|f\|_{\mathrm{CMO}^{p,\lambda}}, \quad f \in \mathrm{CMO}^{p,\lambda}(\mathbb{R}^n);$$

(ii) *when* $p = n/\alpha$*, i.e.,* $\beta = 0$*,*

$$\|\tilde{I}_{\alpha,d} f\|_{\mathrm{BMO}_{\lambda+\alpha}} \le C\|f\|_{\mathrm{CMO}^{n/\alpha,\lambda}}, \quad f \in \mathrm{CMO}^{n/\alpha,\lambda}(\mathbb{R}^n).$$

4 Proofs of Theorems

First of all, we state the following well-definedness of $\tilde{I}_{\alpha,d}$ for $\mathrm{CMO}^{p,\lambda}(\mathbb{R}^n)$, which is shown by the same argument as in the proof of Theorem 3.6 of [16].

Lemma 4.1 *Let $0 < \alpha < 1$, $1 \le p < \infty$, $d \in \mathbb{N}_0$ and $-n/p + \alpha \le \lambda + \alpha < d + 1$. Then for $f \in \mathrm{CMO}^{p,\lambda}(\mathbb{R}^n)$, $\tilde{I}_{\alpha,d} f$ is well-defined.*

Proof Let $f \in \mathrm{CMO}^{p,\lambda}(\mathbb{R}^n)$, $r \ge 1$ and $x \in Q_r$, and let

$$\begin{aligned}\tilde{I}_{\alpha,d} f(x) &= \tilde{I}_{\alpha,d}(\tilde{f}\chi_{Q_{2r}})(x) + \tilde{I}_{\alpha,d}(\tilde{f}(1-\chi_{Q_{2r}}))(x) + f_{Q_{4r}}\tilde{I}_{\alpha,d}1(x)\\ &= I_\alpha(\tilde{f}\chi_{Q_{2r}})(x) - \sum_{\{l:|l|\le d\}} \frac{x^l}{l!}\int_{Q_{2r}\setminus Q_1} \tilde{f}(y)(D^l K_\alpha)(-y)\,dy\\ &\quad + \int_{\mathbb{R}^n\setminus Q_{2r}} \tilde{f}(y)\left(K_\alpha(x-y) - \sum_{\{l:|l|\le d\}} \frac{x^l}{l!}(D^l K_\alpha)(-y)\right) dy\\ &\quad + f_{Q_{4r}}\tilde{I}_{\alpha,d}1(x), \end{aligned} \tag{4.1}$$

where $\tilde{f} = f - f_{Q_{4r}}$. Then, since $\tilde{f}\chi_{Q_{2r}} \in L^p(\mathbb{R}^n)$, the first term is well-defined. The second term is also well-defined, because $(D^l K_\alpha)(\chi_{Q_{2r}} - \chi_{Q_1}) \in L^{p'}(\mathbb{R}^n)$. Here we note that the second term is a polynomial of degree at most d. For the third term, the integral converges absolutely in virtue of Lemmas 4.2 and 4.4, which are shown later, and so the present term is well-defined. As $\int_{\mathbb{R}^n\setminus Q_1}(D^l K_\alpha)(-y)dy$ converges absolutely under the assumption $0 < \alpha < 1$, $\tilde{I}_{\alpha,d}1 \in \mathcal{P}^d(Q_r)$, and then the forth term is well-defined.

Further, if we let for $1 \le s < r$,

$$\begin{aligned} f_{Q_{4r}} - f_{Q_{4s}} &= (f - f_{Q_{4s}})\chi_{Q_{2s}} - (f - f_{Q_{4r}})\chi_{Q_{2r}}\\ &\quad + (f - f_{Q_{4s}})(1-\chi_{Q_{2s}}) - (f - f_{Q_{4r}})(1-\chi_{Q_{2r}}), \end{aligned}$$

then it follows that for $x \in Q_s \subset Q_r$,

$$\begin{aligned} 0 &= \tilde{I}_{\alpha,d}(f_{Q_{4r}} - f_{Q_{4s}})(x)\\ &= \tilde{I}_{\alpha,d}((f - f_{Q_{4s}})\chi_{Q_{2s}})(x) - \tilde{I}_{\alpha,d}((f - f_{Q_{4r}})\chi_{Q_{2r}})(x)\\ &\quad + \tilde{I}_{\alpha,d}((f - f_{Q_{4s}})(1-\chi_{Q_{2s}}))(x) - \tilde{I}_{\alpha,d}((f - f_{Q_{4r}})(1-\chi_{Q_{2r}}))(x). \end{aligned}$$

Therefore, $\tilde{I}_{\alpha,d} f$ is independent of Q_r containing x, and so $\tilde{I}_{\alpha,d} f$ is well-defined on $\mathbb{R}^n$. □

In order to prove Theorems 3.2, 3.3 and 3.6, it is necessary to use the following three lemmas.

Lemma 4.2 (Lemma 7.3 of [21]) *Let $x \in \mathbb{R}^n$, $0 < \alpha < n$ and $d \in \mathbb{N}_0$. If $y \in \mathbb{R}^n \setminus Q_{2|x|}$, then*

$$\left| K_\alpha(x-y) - \sum_{\{l:|l|\le d\}} \frac{x^l}{l!}(D^l K_\alpha)(-y) \right| \le C \frac{|x|^{d+1}}{|y|^{n-\alpha+d+1}}. \tag{4.2}$$

Here note that let $Q_{2|x|} = \{0\}$, when $x = 0$.

Lemma 4.3 (Lemma 4.2 of [18]) *Let $x \in \mathbb{R}^n$, $0 < \alpha < n$ and $d \in \mathbb{N}_0$. If $y \in \mathbb{R}^n \setminus Q_{2|x|}$, then*

$$\left| \triangle_h^{d+1} \left(K_\alpha(x-y) - \sum_{\{l:|l|\le d\}} \frac{x^l}{l!}(D^l K_\alpha)(-y) \right) \right| \le C \frac{|h|^{d+1}}{|y|^{n-\alpha+d+1}}. \tag{4.3}$$

Lemma 4.4 (Lemma 4.1 of [16]; cf. Lemma 4.2 of [19]) *Let $1 \le p < \infty$ and $\lambda \in \mathbb{R}$. If $\beta < 0$ and $\beta + \lambda < 0$, then there exists a positive constant C such that*

$$\int_{\mathbb{R}^n \setminus Q_r} \frac{|f(y) - f_{Q_{2r}}|}{|y|^{n-\beta}}\, dy \le C r^{\beta+\lambda} \|f\|_{\mathrm{CMO}^{p,\lambda}}$$

for all $f \in \mathrm{CMO}^{p,\lambda}(\mathbb{R}^n)$ and $r \ge 1$.

Proof of Theorem 3.2 Let $f \in \mathrm{CMO}^{p,\lambda}(\mathbb{R}^n)$, $r \ge 1$ and $x \in Q_r$. As $\tilde{I}_{\alpha,d} f$ is well-defined by Lemma 4.1, hence we prove only that

$$\|\tilde{I}_{\alpha,d} f\|_{\Lambda^{(d)}_{q,\mu}} \lesssim \|f\|_{\mathrm{CMO}^{p,\lambda}},$$

by the same argument as in the proof of Theorem 3.1 in [17].

Now, in (4.1), let us put

$$\tilde{R}^d_r f(x) = R^d_r \tilde{f}(x) + f_{Q_{4r}} \tilde{I}_{\alpha,d} 1(x) \in \mathcal{P}^d(Q_r),$$

where

$$R^d_r \tilde{f}(x) = -\sum_{\{l:|l|\le d\}} \frac{x^l}{l!} \int_{Q_{2r}\setminus Q_1} \tilde{f}(y)(D^l K_\alpha)(-y)\, dy,$$

and

$$\tilde{J}_{\alpha,d,r} f(x) = \int_{\mathbb{R}^n \setminus Q_{2r}} \tilde{f}(y) \left(K_\alpha(x-y) - \sum_{\{l:|l|\le d\}} \frac{x^l}{l!}(D^l K_\alpha)(-y) \right) dy.$$

Then we have

$$\left(\int_{Q_r} |\tilde{I}_{\alpha,d} f(x) - \tilde{R}_r^d f(x)|^q \, dx\right)^{1/q}$$
$$\leq \left(\int_{Q_r} |I_\alpha(\tilde{f}\chi_{Q_{2r}})(x)|^q \, dx\right)^{1/q} + \left(\int_{Q_r} |\tilde{J}_{\alpha,d,r} f(x)|^q \, dx\right)^{1/q}$$
$$=: I_1 + I_2. \tag{4.4}$$

First we estimate I_1. By applying the strong (p, q) boundedness of I_α, we have

$$I_1 \leq \|I_\alpha(\tilde{f}\chi_{Q_{2r}})\|_{L^q} \lesssim \|\tilde{f}\chi_{Q_{2r}}\|_{L^p} \lesssim r^\lambda |Q_{2r}|^{1/p} \|\tilde{f}\|_{B^{p,\lambda}}$$
$$\sim r^{\lambda+n/p} \|f\|_{\mathrm{CMO}^{p,\lambda}} = r^{\mu+n/q} \|f\|_{\mathrm{CMO}^{p,\lambda}}.$$

Next, in order to estimate I_2, using (4.2), it follows that for $x \in Q_r$ and $y \in \mathbb{R}^n \setminus Q_{2r}$,

$$\left| K_\alpha(x-y) - \sum_{\{l:|l|\leq d\}} \frac{x^l}{l!} (D^l K_\alpha)(-y) \right| \lesssim \frac{|x|^{d+1}}{|y|^{n-\alpha+d+1}} \leq \frac{r^{d+1}}{|y|^{n-\alpha+d+1}}.$$

Consequently, we obtain by Lemma 4.4 and the assumptions $0 < \alpha < 1$ and $\mu+\sigma < d+1$,

$$|\tilde{J}_{\alpha,d,r} f(x)| \lesssim r^{d+1} \int_{\mathbb{R}^n \setminus Q_{2r}} \frac{|\tilde{f}(y)|}{|y|^{n-\alpha+d+1}} \, dy \lesssim r^{\lambda+\alpha} \|f\|_{\mathrm{CMO}^{p,\lambda}}$$
$$= r^\mu \|f\|_{\mathrm{CMO}^{p,\lambda}}, \tag{4.5}$$

which shows that $\tilde{J}_{\alpha,d,r} f$ is well-defined for all $x \in Q_r$. Thus we get

$$I_2 = \|\tilde{J}_{\alpha,d,r} f\|_{L^q(Q_r)} \lesssim r^\mu \|f\|_{\mathrm{CMO}^{p,\lambda}} \cdot |Q_r|^{1/q} \sim r^{\mu+n/q} \|f\|_{\mathrm{CMO}^{p,\lambda}},$$

and then

$$\|\tilde{I}_{\alpha,d} f\|_{\Lambda^{(d)}_{q,\mu}} \lesssim \sup_{r\geq 1} \frac{1}{r^\mu} \left(\fint_{Q_r} \left|\tilde{I}_{\alpha,d} f(y) - \tilde{R}_r^d f(y)\right|^q \, dy\right)^{1/q}$$
$$\lesssim \sup_{r\geq 1} \frac{1}{r^\mu} \left(\frac{1}{|Q_r|}\right)^{1/q} \cdot r^{\mu+n/q} \|f\|_{\mathrm{CMO}^{p,\lambda}} \sim \|f\|_{\mathrm{CMO}^{p,\lambda}}.$$

This concludes the proof. □

Proof of Theorem 3.3 This proof is similar to that of Theorem 3.6 of [16]. Hence, in the same way as (4.4), we have that for $f \in \mathrm{CMO}^{1,\lambda}(\mathbb{R}^n)$, $r \geq 1$ and $x \in Q_r$,

$$\begin{aligned}
&\sup_{t>0}(2t)^q \left|\left\{x \in Q_r : |\tilde{I}_{\alpha,d} f(x) - \tilde{R}_r^d f(x)| > 2t\right\}\right| \\
&\qquad \leq 2^q \left\{\sup_{t>0} t^q \left|\left\{x \in Q_r : |I_\alpha(\tilde{f}\chi_{Q_{2r}})(x)| > t\right\}\right| \right. \\
&\qquad\qquad \left. + \sup_{t>0} t^q \left|\{x \in Q_r : |\tilde{J}_{\alpha,d,r} f(x)| > t\}\right|\right\} \\
&\qquad =: 2^q \left(I_3 + I_4\right).
\end{aligned}$$

Then it follows from the weak $(1, q)$ boundedness of I_α that

$$I_3^{1/q} \lesssim r^{\lambda+n} \|f\|_{\mathrm{CMO}^{1,\lambda}} = r^{\mu+n/q} \|f\|_{\mathrm{CMO}^{1,\lambda}},$$

and by applying (4.5) with $p = 1$,

$$I_4^{1/q} \lesssim r^{\mu+n/q} \|f\|_{\mathrm{CMO}^{1,\lambda}}.$$

Thus, we obtain

$$\|\tilde{I}_{\alpha,d} f\|_{W\Lambda^{(d)}_{q,\mu}} \lesssim \|f\|_{\mathrm{CMO}^{1,\lambda}},$$

which completes the proof. □

Proof of Theorem 3.6 Let $f \in \mathrm{CMO}^{p,\lambda}(\mathbb{R}^n)$, $r \geq 1$ and $x \in Q_r$. Similarly to the proof of Theorem 3.2, we prove only that for $n/\alpha \leq p < \infty$,

$$\|\tilde{I}_{\alpha,d} f\|_{\mathrm{Lip}^{(d)}_\beta(Q_r)} \lesssim r^{\lambda+n/p} \|f\|_{\mathrm{CMO}^{p,\lambda}},$$

following similar arguments to the proof of Theorem 3.2 in [18].

Now, using the notation in the proof of Theorem 3.2, it follows from Remarks 2.8, 2.10 and (4.1) with

$$R_r^d \tilde{f}(x) = -\sum_{\{l:1\leq|l|\leq d\}} \frac{x^l}{l!} \int_{Q_{2r}\setminus Q_1} \tilde{f}(y)(D^l K_\alpha)(-y)\, dy$$

that

$$\begin{aligned}\|\tilde{I}_{\alpha,d} f\|_{\mathrm{Lip}_\beta^{(d)}(Q_r)} &\sim \|\tilde{I}_{\alpha,d} f\|_{\mathcal{L}_{1,\beta}^{(d)}(Q_r)} \\ &\sim \sup_{Q(x,s)\subset Q_r} \inf_{P\in\mathcal{P}^d(Q_r)} \frac{1}{s^\beta}\left(\fint_{Q(x,s)} |\tilde{I}_{\alpha,d} f(y) - P(y)|\,dy\right) \\ &\le \sup_{Q(x,s)\subset Q_r} \frac{1}{s^\beta}\left(\fint_{Q(x,s)} |\tilde{I}_{\alpha,d} f(y) - \tilde{R}_r^d f(y)|\,dy\right) \\ &\le \|\tilde{I}_\alpha(\tilde{f}\chi_{Q_{2r}})\|_{\mathrm{Lip}_\beta(Q_r)} + \|\tilde{J}_{\alpha,d,r} f\|_{\mathcal{L}_{1,\beta}^{(d)}(Q_r)} \\ &=: I_5 + I_6.\end{aligned}$$

We firstly estimate I_5. When $n/\alpha < p < \infty$, which implies $0 < \beta < 1$, if we apply the $(L^p, \mathrm{Lip}_\beta)$ boundedness of $\tilde{I}_\alpha$, then we get

$$\begin{aligned}I_5 &\le \|\tilde{I}_\alpha(\tilde{f}\chi_{Q_{2r}})\|_{\mathrm{Lip}_\beta} \lesssim \|\tilde{f}\chi_{Q_{2r}}\|_{L^p} \lesssim r^{\lambda+n/p}\|\tilde{f}\|_{B^{p,\lambda}} \\ &= r^{\lambda+n/p}\|f\|_{\mathrm{CMO}^{p,\lambda}}.\end{aligned}$$

Similarly, when $p = n/\alpha$, by using the $(L^{n/\alpha}, \mathrm{BMO})$ boundedness of $\tilde{I}_\alpha$, we obtain

$$I_5 \le \|\tilde{I}_\alpha(\tilde{f}\chi_{Q_{2r}})\|_{\mathrm{BMO}} \lesssim r^{\lambda+\alpha}\|f\|_{\mathrm{CMO}^{n/\alpha,\lambda}}.$$

Next in order to estimate I_6, since by Remark 2.8,

$$I_6 \sim \|\tilde{J}_{\alpha,d,r} f\|_{\mathrm{Lip}_\beta^{(d)}(Q_r)} = \sup_{x,x+h\in Q_r, h\neq 0} \frac{1}{|h|^\beta}\,|\triangle_h^{d+1}\tilde{J}_{\alpha,d,r} f(x)|,$$

we estimate $\triangle_h^{d+1}\tilde{J}_{\alpha,d,r} f(x)$. To do so, if we use (4.3), Lemma 4.4 and the assumptions $0 < \alpha < 1$ and $\lambda + \alpha < d + 1$, then we have for $x \in Q_r$ and $y \in \mathbb{R}^n \setminus Q_{2r}$,

$$\begin{aligned}&|\triangle_h^{d+1}\tilde{J}_{\alpha,d,r} f(x)| \\ &\quad= \left|\int_{\mathbb{R}^n\setminus Q_{2r}} \tilde{f}(y)\left\{\triangle_h^{d+1}\left(K_\alpha(x-y) - \sum_{\{l:|l|\le d\}} \frac{x^l}{l!}(D^l K_\alpha)(-y)\right)\right\} dy\right| \\ &\quad\lesssim |h|^{d+1}\int_{\mathbb{R}^n\setminus Q_{2r}} \frac{|\tilde{f}(y)|}{|y|^{n-\alpha+d+1}}\,dy \lesssim |h|^{d+1} r^{\alpha-d-1+\lambda}\|f\|_{\mathrm{CMO}^{p,\lambda}}.\end{aligned}$$

Hence

$$I_6 \lesssim \sup_{x,x+h\in Q_r,h\neq 0} \frac{1}{|h|^\beta} \cdot |h|^{d+1} r^{\alpha-d-1+\lambda} \|f\|_{\mathrm{CMO}^{p,\lambda}}$$
$$\leq r^{\lambda+\alpha-\beta} \|f\|_{\mathrm{CMO}^{p,\lambda}} = r^{\lambda+n/p} \|f\|_{\mathrm{CMO}^{p,\lambda}}.$$

Thus it follows that for $n/\alpha \leq p < \infty$,

$$\|\tilde{I}_{\alpha,d} f\|_{\mathrm{Lip}^{(d)}_{\beta,\lambda+n/p}} = \sup_{r\geq 1} \frac{1}{r^{\lambda+n/p}} \|\tilde{I}_{\alpha,d} f\|_{\mathrm{Lip}^{(d)}_{\beta}(Q_r)} \lesssim \|f\|_{\mathrm{CMO}^{p,\lambda}}.$$

This shows the conclusion. □

Acknowledgments The author would like to express his deep gratitude to the anonymous referees for their careful reading and fruitful comments.

References

1. J. Alvarez, M. Guzmán-Partida, J. Lakey, Spaces of bounded λ-central mean oscillation, Morrey spaces, and λ-central Carleson measures. Collect. Math. **51**, 1–47 (2000)
2. Y. Chen, K. Lau, Some new classes of Hardy spaces. J. Funct. Anal. **84**, 255–278 (1989)
3. H. Feichtinger, An elementary approach to Wiener's third Tauberian theorem on Euclidean n-space, in *Proceedings, Conference at Cortona 1984, Symposia Mathematica*, vol. 29 (Academic Press, New York, 1987), pp. 267–301
4. Z. Fu, Y. Lin, S. Lu, λ-central BMO estimates for commutators of singular integral operators with rough kernels. Acta Math. Sin. (Engl. Ser.) **24**, 373–386 (2008)
5. J. García-Cuerva, Hardy spaces and Beurling algebras. J. Lond. Math. Soc. **39**, 499–513 (1989)
6. J. García-Cuerva, M.J.L. Herrero, A theory of Hardy spaces associated to the Herz spaces. Proc. Lon. Math. Soc. **69**, 605–628 (1994)
7. G.H. Hardy, J.E. Littlewood, Some properties of fractional integrals. I. Math. Z. **27**, 565–606 (1928); II, ibid. **34**, 403–439 (1932)
8. C. Herz, Lipschitz spaces and Bernstein's theorem on absolutely convergent Fourier transforms. J. Math. Mech. **18**, 283–324 (1968)
9. Y. Komori-Furuya, K. Matsuoka, Some weak-type estimates for singular integral operators on CMO spaces. Hokkaido Math. J. **39**, 115–126 (2010)
10. Y. Komori-Furuya, K. Matsuoka, Strong and weak estimates for fractional integral operators on some Herz-type function spaces, in *Proceedings of the Maratea Conference FAAT 2009, Rendiconti del Circolo Mathematico di Palermo, Serie II, Supplementary*, vol. 82 (2010), pp. 375–385
11. Y. Komori-Furuya, K. Matsuoka, E. Nakai, Y. Sawano, Integral operators on B_σ-Morrey–Campanato spaces. Rev. Mat. Complut. **26**, 1–32 (2013)
12. Y. Komori-Furuya, K. Matsuoka, E. Nakai, Y. Sawano, Applications of Littlewood–Paley theory for $\dot{B}_\sigma$-Morrey spaces to the boundedness of integral operators. J. Funct. Spaces Appl. **2013**, 859402 (2013)
13. T. Kurokawa, Riesz potentials, higher Riesz transforms and Beppo Levi spaces. Hiroshima Math. J. **18**, 541–597 (1988)

14. T. Kurokawa, Weighted norm inequalities for Riesz potentials. Jpn. J. Math. **14**, 261–274 (1988)
15. S. Lu, D. Yang, The central BMO spaces and Littlewood–Paley operators. Approx. Theory Appl. (N.S.) **11**, 72–94 (1995)
16. K. Matsuoka, B_σ-Morrey–Campanato estimates and some estimates for singular integrals on central Morrey spaces and λ-CMO spaces, in *Banach and Function Spaces IV (Kitakyushu 2012)* (Yokohama Publishers, Yokohama, 2014), pp. 325–335
17. K. Matsuoka, Generalized fractional integrals on central Morrey spaces and generalized λ-CMO spaces, in *Function Spaces X, Banach Center Publications*, vol. 102 (Institute of Mathematics, Polish Academy of Sciences, Warsawa, 2014), pp. 181–188
18. K. Matsuoka, Generalized fractional integrals on central Morrey spaces and generalized σ-Lipschitz spaces, in *Current Trends in Analysis and its Applications: Proceedings of the 9th ISAAC Congress, Kraków 2013*. Springer Proceedings in Mathematics and Statistics (Birkhäuser, Basel, 2015), pp. 179–189
19. K. Matsuoka, E. Nakai, Fractional integral operators on $B^{p,\lambda}$ with Morrey–Campanato norms, in *Function Spaces IX, Banach Center Publishers*, vol. 92 (Institute of Mathematics, Polish Academy of Sciences, Warsawa, 2011), pp. 249–264
20. Y. Mizuta, On the behaviour at infinity of superharmonic functions. J. Lond. Math. Soc. **27**, 97–105 (1983)
21. Y. Mizuta, *Potential Theory in Euclidean Spaces* (Gakkōtosho, Tokyo, 1996)
22. E. Nakai, Y. Sawano, Hardy spaces with variable exponents and generalized Campanato spaces. J. Funct. Anal. **262**, 3665–3748 (2012)
23. J. Peetre, On the theory of $\mathcal{L}_{p,\lambda}$ spaces. J. Funct. Anal. **4**, 71–87 (1969)
24. A. Pietsch, *History of Banach Spaces and Linear Operators* (Birkhäuser, Boston, 2007)
25. S.L. Sobolev, On a theorem in functional analysis. Mat. Sbornik **4**, 471–497 (1938) (in Russian)
26. A. Zygmund, On a theorem of Marcinkiewicz concerning interpolation of operations. J. Math. Pures Appl. **35**, 223–248 (1956)

On Some Consequences of the Solvability of the Caffarelli–Silvestre Extension Problem

Jan Meichsner and Christian Seifert

Abstract We consider the Caffarelli–Silvestre extension problem, i.e., a Bessel type ODE in a Banach space X with a closed and typically unbounded operator A as right-hand side and point out a couple of consequences arising from the assumption of the well-posedness of the problem. In the end a conjecture is stated concerning the implications of analyticity of the solution of the extension problem.

Keywords Fractional powers · Non-negative operator · Dirichlet-to-Neumann operator

Mathematics Subject Classification (2010) Primary 47A05; Secondary 47D06, 47A60

1 Introduction

Since the well-known Caffarelli–Silvestre paper [4] from 2007 various authors considered an abstract version of the there presented problem by having a look on the abstract (incomplete) ODE problem

$$u''(t) + \frac{1-2\alpha}{t}u'(t) = Au(t) \quad (t > 0), \quad u(0) = x \tag{1.1}$$

J. Meichsner (✉)
Technische Universität Hamburg, Institut für Mathematik, Hamburg, Germany
e-mail: jan.meichsner@tuhh.de

C. Seifert
Technische Universität Hamburg, Institut für Mathematik, Hamburg, Germany

Technische Universität Clausthal, Institut für Mathematik, Clausthal-Zellerfeld, Germany
e-mail: christian.seifert@tuhh.de; christian.seifert@tu-clausthal.de

M. A. Bastos et al. (eds.), *Operator Theory, Functional Analysis and Applications*,
Operator Theory: Advances and Applications 282,
https://doi.org/10.1007/978-3-030-51945-2_22

in some Hilbert or more general Banach space X, see [2, 5, 8, 9]. The abstract setting made it necessary to introduce new tools and viewpoints completely different from the techniques used in [4] for the rather concrete problem considered there. This development was first started by Stinga–Torrea in [9] by the introduction of explicit formulas for solutions of (1.1) and further continued by Galé–Miana–Stinga in [5]. Typically, the operator A appearing on the right-hand side of (1.1) is assumed to be sectorial (with or without dense domain). The basic question of interest is to establish existence and uniqueness results. For the special choice of $\alpha = 1/2$ the singular first order term in (1.1) vanishes and the problem reduces to an incomplete second order Cauchy problem as already considered by Balakrishnan in [3, Theorem 6.1], where the well-posedness for a sectorial operator A is shown under the additional assumption of a globally bounded solution u of (1.1). Concerning generalisations of Balakrishnan's result, in [9] the authors considered the situation for a self-adjoint, positive, so in particular sectorial, operator on a Hilbert space. The authors derived two explicit formulas for a solution to (1.1) including the Stinga-Torrea formula, a Poisson formula for the solution. The uniqueness of the solution is also shown under the assumption of A having a pure point spectrum. The corresponding Banach space setting was considered in [5] where the authors constructed explicitly solutions for both, generators of tempered integrated semigroups and tempered integrated cosine families. These results particularly cover generators of bounded C_0-semigroups. Combined with the uniqueness result from [8], (1.1) is a well-posed problem and this fact actually extends to arbitrary sectorial right-hand sides. Given an initial condition for the first derivative one can also tackle the problem by cosine functions, see e.g. [1, Chapter 3]. The connection between the various results was explored a bit in [8], where one can see that the boundedness assumption appearing in Balakrishnan's paper 'couples' the two initial values and thus is suitable to replace one of them.

In [6], the authors consider the problem from a slightly different point of view. They assume the ODE (1.1) to be given for $\alpha = 1/2$ but merely assume A to be closed with non-empty resolvent set. As a consequence of the well-posedness of the problem, the existence of a generator B of a C_0-semigroup is deduced which satisfies $B^2 = A$. Further, analyticity of the solution u of (1.1) on a sector is shown to imply sectoriality of A and then $B = \sqrt{A}$ in the functional calculus sense.

This is the starting point of this paper. More precisely, we attempt to state the right setting for proving similar results as explained above in the more general case $\alpha \in (0, 1)$. Making use of similar strategies as in [6, Theorem 6.3.3], generalisations of the first result on factorizing A (as B^2) can be obtained; cf. Corollary 2.13. For the second result that analyticity of the solution u of (1.1) on a sector yields sectoriality of the operator A, we formulate a conjecture.

2 The Solution Operator to the Caffarelli–Silvestre Extension Problem

In this paper X will always denote a complex Banach space with norm $\|\cdot\|$. Further let A be a closed linear operator in X with non-empty resolvent set $\rho(A)$ and with dense domain $\mathcal{D}(A)$. We will interpret $\mathcal{D}(A)$ as a Banach space equipped with the graph norm

$$\|x\|_{\mathcal{D}(A)} := \|x\| + \|Ax\| \quad (x \in \mathcal{D}(A)).$$

Moreover, let $\alpha \in (0, 1)$ be a parameter.

For $x \in X$ let $P_\alpha(x)$ be the problem of finding $u \in C_b\big([0,\infty); \mathcal{D}(A)\big) \cap C^2\big((0,\infty); X\big)$ such that $u(0) = x$ and

$$u''(t) + \frac{1-2\alpha}{t} u'(t) = Au(t) \quad (t > 0).$$

For the whole paper, we may assume as a standing assumption that $P_\alpha(x)$ has a unique solution, denoted by u_x, for every $x \in \mathcal{D}(A)$. This fact will not be mentioned explicitly anymore but the reader should be always aware of it.

Remark 2.1 As noted in the introduction, there are various cases where $P_\alpha(x)$ has a unique solution for all $x \in \mathcal{D}(A)$, e.g., when A is the generator of a bounded C_0-semigroup ([5, Theorem 2.1]) or when A is the generator of a tempered integrated cosine family ([5, Theorem 1.3] combined with [8, Theorem 5.8]).

For $x \in \mathcal{D}(A)$ and $t \geq 0$ we may then define $Ux(t) := u_x(t)$.

Lemma 2.2 *We have* $U \in \mathcal{L}\big(\mathcal{D}(A), C_b\big([0,\infty); \mathcal{D}(A)\big)\big)$.

Proof The proof essentially works as the one for the special case $\alpha = 1/2$ as presented in [6, Theorem 6.3.3] with the help of the closed graph theorem. So let (x_n) be a sequence in $\mathcal{D}(A)$ which converges w.r.t. $\|\cdot\|_{\mathcal{D}(A)}$ towards $x \in \mathcal{D}(A)$. Further, assume that (Ux_n) is convergent in $C_b\big([0,\infty); \mathcal{D}(A)\big)$ towards some element v. Let $n, m \in \mathbb{N}$. From the fact that Ux_n and Ux_m solve $P_\alpha(x_n)$ and $P_\alpha(x_m)$, respectively, we get (for $t \geq 0$)

$$\begin{aligned} &\left\| t^{2\alpha-1} \frac{\mathrm{d}}{\mathrm{d}t} t^{1-2\alpha} \frac{\mathrm{d}}{\mathrm{d}t} Ux_n(t) - t^{2\alpha-1} \frac{\mathrm{d}}{\mathrm{d}t} t^{1-2\alpha} \frac{\mathrm{d}}{\mathrm{d}t} Ux_m(t) \right\| \\ &\quad = \|AUx_n(t) - AUx_m(t)\| \leq \|Ux_n - Ux_m\|_{C_b([0,\infty);\mathcal{D}(A))}, \end{aligned}$$

which implies the existence of some function $u_2 \in C_b\big([0,\infty); X\big)$ such that $\Big(t \mapsto t^{2\alpha-1} \frac{\mathrm{d}}{\mathrm{d}t} t^{1-2\alpha} \frac{\mathrm{d}}{\mathrm{d}t} Ux_n(t)\Big)$ converges in $C_b\big([0,\infty); X\big)$ towards u_2. Multiplying the ODE (at $t_1 > 0$) for the initial condition x_k by $t_1^{1-2\alpha}$, and integrating the

resulting term from t_1 to $t_2 > 0$ results in

$$t_1^{1-2\alpha}\frac{\mathrm{d}}{\mathrm{d}t}Ux_k(t_1) - t_2^{1-2\alpha}\frac{\mathrm{d}}{\mathrm{d}t}Ux_k(t_2) = \int_{t_1}^{t_2} s^{1-2\alpha}AUx_k(s)\mathrm{d}s.$$

By the boundedness of $s \mapsto AUx_k(s)$ and the integrability of $s \mapsto s^{1-2\alpha}$ we can send t_1 to 0. Hence, for $k \in \mathbb{N}$, we can define

$$y_k := \lim_{t\to 0+} t^{1-2\alpha}\frac{\mathrm{d}}{\mathrm{d}t}Ux_k(t).$$

Let $t > 0$. Multiplying this time the difference of the ODE's (at $s > 0$) for initial conditions x_n and x_m by $s^{1-2\alpha}$, and integrating the term from 0 to t results in

$$\begin{aligned} &t^{1-2\alpha}\frac{\mathrm{d}}{\mathrm{d}t}Ux_n(t) - t^{1-2\alpha}\frac{\mathrm{d}}{\mathrm{d}t}Ux_m(t) \\ &\quad = y_n - y_m + \int_0^t s^{1-2\alpha}A(Ux_n - Ux_m)(s)\mathrm{d}s. \end{aligned} \tag{2.1}$$

Now, we multiply (2.1) by $t^{2\alpha-1}$, integrate again, take norms and solve for $\|y_n - y_m\|$. Then we obtain

$$\begin{aligned} &\frac{t^{2\alpha}}{2\alpha}\|y_n - y_m\| \\ &= \left\| U(x_n - x_m)(t) - (x_n - x_m) - \int_0^t s^{2\alpha-1}\int_0^s r^{1-2\alpha}AU(x_n - x_m)(r)\mathrm{d}r\mathrm{d}s \right\| \\ &\le \left(2 + \frac{t^2}{4-4\alpha}\right)\|Ux_n - Ux_m\|_{C_b([0,\infty);\mathcal{D}(A))}. \end{aligned}$$

Hence, (y_n) is a Cauchy sequence in X.

By (2.1), we estimate

$$\begin{aligned} &\left\| t^{1-2\alpha}\frac{\mathrm{d}}{\mathrm{d}t}Ux_n(t) - t^{1-2\alpha}\frac{\mathrm{d}}{\mathrm{d}t}Ux_m(t) \right\| \\ &\quad = \left\| y_n - y_m + \int_0^t s^{1-2\alpha}A(Ux_n - Ux_m)(s)\mathrm{d}s \right\| \\ &\quad \le \|y_n - y_m\| + \frac{t^{2-2\alpha}}{2-2\alpha}\|Ux_n - Ux_m\|_{C_b([0,\infty);\mathcal{D}(A))}. \end{aligned}$$

Therefore, there exists $u_1 \in C\big((0,\infty); X\big)$ such that

$$\big(t \mapsto t^{1-2\alpha}\frac{\mathrm{d}}{\mathrm{d}t}Ux_n(t)\big) \to \big(t \mapsto t^{1-2\alpha}u_1(t)\big)$$

uniformly on compacts in $[0,\infty)$. Thus, $v \in C^1\big((0,\infty); X\big)$.

Furthermore, for $t > 0$ we have

$$\begin{aligned}
&\left\|\frac{\mathrm{d}^2}{\mathrm{d}t^2}Ux_n(t) - \frac{\mathrm{d}^2}{\mathrm{d}t^2}Ux_m(t)\right\| \\
&\quad\le \left\|t^{2\alpha-1}\frac{\mathrm{d}}{\mathrm{d}t}t^{1-2\alpha}\frac{\mathrm{d}}{\mathrm{d}t}Ux_n(t) - t^{2\alpha-1}\frac{\mathrm{d}}{\mathrm{d}t}t^{1-2\alpha}\frac{\mathrm{d}}{\mathrm{d}t}Ux_m(t)\right\| \\
&\qquad + \frac{|1-2\alpha|}{t}\left\|\frac{\mathrm{d}}{\mathrm{d}t}Ux_n(t) - \frac{\mathrm{d}}{\mathrm{d}t}Ux_m(t)\right\| \\
&\quad\to 0,
\end{aligned}$$

uniformly for t in a compact set $K \subset (0,\infty)$. Therefore, $v \in C^2\big((0,\infty); X\big)$. Since

$$v(0) = \lim_{n\to\infty}(Ux_n)(0) = \lim_{n\to\infty} x_n = x,$$

by uniqueness of the solution of $P_\alpha(x)$ we obtain $v = Ux$. Now, the closed graph theorem yields $U \in \mathcal{L}\big(\mathcal{D}(A), C_b\big([0,\infty); \mathcal{D}(A)\big)\big)$. □

As a corollary of the previous lemma we obtain continuity properties of certain derivatives of Ux.

Corollary 2.3 *For $x \in \mathcal{D}(A)$ we have*

$$\Big(t \mapsto t^{1-2\alpha}\frac{\mathrm{d}}{\mathrm{d}t}Ux(t)\Big) \in C\big([0,\infty); X\big)$$

and

$$\Big(t \mapsto t^{2\alpha-1}\frac{\mathrm{d}}{\mathrm{d}t}t^{1-2\alpha}\frac{\mathrm{d}}{\mathrm{d}t}Ux(t)\Big) \in C_b\big([0,\infty); X\big).$$

Let $\lambda \in \rho(A)$. As a consequence of the assumed uniqueness one can prove the equations $U(\lambda - A)^{-1}x = (\lambda - A)^{-1}Ux$ for $x \in \mathcal{D}(A)$ and $UAx = AUx$ for $x \in \mathcal{D}(A^2)$.

For $x \in X$ let $\widetilde{P_\alpha}(x)$ be the problem of finding $u \in C_b\big([0,\infty); X\big)$ with $u(0) = x$ and such that for all $\phi \in C_c^\infty\big((0,\infty)\big)$ we have

$$\int_0^\infty u(t)\phi(t)\mathrm{d}t \in \mathcal{D}(A) \quad \text{and} \quad \int_0^\infty u(t)\frac{\mathrm{d}}{\mathrm{d}t}t^{1-2\alpha}\frac{\mathrm{d}}{\mathrm{d}t}t^{2\alpha-1}\phi(t)\mathrm{d}t = A\int_0^\infty u(t)\phi(t)\mathrm{d}t.$$

We can interpret $\widetilde{P_\alpha}(x)$ as a weak version of $P_\alpha(x)$.

Lemma 2.4 *The mapping U interpreted as a function from $\mathcal{D}(A) \subseteq X$ to $C_b\big([0,\infty); \mathcal{D}(A)\big) \subseteq C_b\big([0,\infty); X\big)$ admits a continuous extension to all of X, again denoted by U. For $x \in X$ the function $Ux \in C_b\big([0,\infty); X\big)$ is the unique solution of $\widetilde{P_\alpha}(x)$.*

Proof Let $x \in \mathcal{D}(A)$. By Lemma 2.2 there exist $C_1 \geq 0$ such that

$$\begin{aligned}
&\|Ux\|_{C_b([0,\infty);X)} \\
&\quad = \left\|(\lambda - A)U(\lambda - A)^{-1}x\right\|_{C_b([0,\infty);X)} \\
&\quad \leq (|\lambda| + 1)\left(\left\|U(\lambda - A)^{-1}x\right\|_{C_b([0,\infty);X)} + \left\|AU(\lambda - A)^{-1}x\right\|_{C_b([0,\infty);X)}\right) \\
&\quad \leq C_1\left\|(\lambda - A)^{-1}x\right\|_{\mathcal{D}(A)} \leq C_2\|x\|
\end{aligned}$$

for some $C_2 > 0$. Hence, by density of $\mathcal{D}(A)$ in X, U can be continuously extended to the whole of X. We denote the extension again by U. For given $x \in X$ let (x_n) be a sequence in $\mathcal{D}(A)$ convergent towards x in X and $\phi \in C_c^\infty\big((0,\infty)\big)$. Then $Ux_n \to Ux$ in $C_b\big([0,\infty); X\big)$. Hence,

$$\begin{aligned}
\int_0^\infty Ux(t)\frac{\mathrm{d}}{\mathrm{d}t}t^{1-2\alpha}\frac{\mathrm{d}}{\mathrm{d}t}t^{2\alpha-1}\phi(t)\mathrm{d}t &= \int_0^\infty \lim_{n\to\infty} Ux_n(t)\frac{\mathrm{d}}{\mathrm{d}t}t^{1-2\alpha}\frac{\mathrm{d}}{\mathrm{d}t}t^{2\alpha-1}\phi(t)\mathrm{d}t \\
&= \lim_{n\to\infty}\int_0^\infty t^{2\alpha-1}\frac{\mathrm{d}}{\mathrm{d}t}t^{1-2\alpha}\frac{\mathrm{d}}{\mathrm{d}t}Ux_n(t)\phi(t)\mathrm{d}t \\
&= \lim_{n\to\infty}\int_0^\infty AUx_n(t)\phi(t)\mathrm{d}t \\
&= \lim_{n\to\infty} A\int_0^\infty Ux_n(t)\phi(t)\mathrm{d}t,
\end{aligned}$$

by Hille's theorem. Since

$$\int_0^\infty Ux_n(t)\phi(t)\mathrm{d}t \to \int_0^\infty Ux(t)\phi(t)\mathrm{d}t,$$

by closedness of A we observe

$$\int_0^\infty Ux(t)\phi(t)\mathrm{d}t \in \mathcal{D}(A)$$

and

$$\int_0^\infty Ux(t)\frac{\mathrm{d}}{\mathrm{d}t}t^{1-2\alpha}\frac{\mathrm{d}}{\mathrm{d}t}t^{2\alpha-1}\phi(t)\mathrm{d}t = A\int_0^\infty Ux(t)\phi(t)\mathrm{d}t.$$

Also

$$Ux(0) = \lim_{n\to\infty} Ux_n(0) = \lim_{n\to\infty} x_n = x.$$

So Ux is a solution to $\widetilde{P_\alpha}(x)$ as claimed.

Note that the solution of $\widetilde{P_\alpha}(x)$ is unique. Indeed, if u is a solution to $\widetilde{P_\alpha}(0)$ then $v := (\lambda - A)^{-1}u$ is a solution to $P_\alpha(0)$, hence $v = 0$ and, therefore, $u = 0$. □

The uniqueness of a solution to $\widetilde{P_\alpha}(x)$ yields the extension of the already found relationships between U, A and its resolvent $(\lambda - A)^{-1}$ to $U(\lambda - A)^{-1} = (\lambda - A)^{-1}U$ and $AUx = UAx$ for $x \in \mathcal{D}(A)$.

Remark 2.5 In view of Lemma 2.4, the authors came to the conclusion that defining a solution to the considered problem in [8] would have been better in the way as in $\widetilde{P_\alpha}(x)$.

Next we show that the smoothness of u and its scaled derivatives tells us precisely about the smoothness of the initial datum x.

Definition 2.6 We define the operator D_α in $C_b\big([0,\infty);\, X\big)$ by

$$\mathcal{D}(D_\alpha) := \big\{u \in C_b\big([0,\infty);\, X\big) \cap C^1\big((0,\infty);\, X\big) \mid \big(t \mapsto t^{1-2\alpha}u'(t)\big) \in C_b\big([0,\infty);\, X\big)\big\},$$

$$D_\alpha u := \big(t \mapsto t^{1-2\alpha}u'(t)\big).$$

The operator D_α acts as a scaled first derivative. Note that D_α is closed.

Lemma 2.7 *Let $k \in \mathbb{N}_0$, $x \in X$. Then $x \in \mathcal{D}(A^k)$ if and only if*

$$Ux \in \mathcal{D}\big((D_{1-\alpha}D_\alpha)^k\big).$$

Proof For $k = 0$ there is nothing to prove. Therefore, let $k \geq 1$ and $x \in \mathcal{D}(A^k)$. Consider first $k = 1$. Then $Ux \in \mathcal{D}(D_{1-\alpha}D_\alpha)$ by Corollary 2.3. If $k \geq 2$ we note that

$$D_{1-\alpha}D_\alpha Ux = AUx = UAx$$

by (1.1) and commutativity of A and U on $\mathcal{D}(A)$. Since $Ax \in \mathcal{D}(A)$, arguing inductively shows $Ux \in \mathcal{D}\big((D_{1-\alpha}D_\alpha)^k\big)$ and

$$(D_{1-\alpha}D_\alpha)^k Ux = UA^k x.$$

Conversely, let $k \geq 1$ and $Ux \in \mathcal{D}\big((D_{1-\alpha}D_\alpha)^k\big)$. Consider again first the case $k = 1$ and choose for $s > 0$ a sequence (ϕ_k) in $C_c^\infty\big((0,\infty)\big)$ which converges towards δ_s in $C_c\big((0,\infty)\big)'$. The function Ux solves $\widetilde{P_\alpha}(x)$. Hence,

$$\int_0^\infty D_{1-\alpha}D_\alpha Ux(t)\phi_k(t)\mathrm{d}t = A\int_0^\infty Ux(t)\phi_k(t)\mathrm{d}t$$

which, by closedness of A, gives $D_{1-\alpha}D_\alpha Ux(s) = AUx(s)$ as $k \to \infty$. Further, $Ux(s) \to x$ as $s \to 0+$. Using the closedness of A again yields $x \in \mathcal{D}(A)$ and $D_{1-\alpha}D_\alpha Ux = AUx = UAx$. Now one can argue again inductively. □

Let us define $B_\alpha : \mathcal{D}(A) \to X$ by

$$B_\alpha x := \lim_{t\to 0+} t^{1-2\alpha}\frac{\mathrm{d}}{\mathrm{d}t}Ux(t).$$

Note that B_α is the 'generalized Dirichlet-to-Neumann operator' associated to the problem $P_\alpha(x)$. Since $(\lambda - A)^{-1}U = U(\lambda - A)^{-1}$, we obtain

$$B_\alpha(\lambda - A)^{-1}x = (\lambda - A)^{-1}B_\alpha x$$

for $x \in \mathcal{D}(A)$. From the proof of Lemma 2.2 one can see the existence of a constant $C > 0$, depending on a fixed parameter $T \in [0,\infty)$ (let us take $T = 1$, say), such that $\|B_\alpha x\| \leq C\,\|x\|_{\mathcal{D}(A)}$. Commutativity with the resolvent of A and continuity with respect to the graph norm imply that B_α is closable as an operator in X with domain $\mathcal{D}(A) \subseteq X$. To see this, consider (x_n) in $\mathcal{D}(A)$ convergent towards 0 in X

and such that $B_\alpha x_n \to y$ in X. Then

$$\left\|(\lambda - A)^{-1} y\right\| = \lim_{n\to\infty} \left\|B_\alpha(\lambda - A)^{-1} x_n\right\| \le C \cdot \lim_{n\to\infty} \left\|(\lambda - A)^{-1} x_n\right\|_{\mathcal{D}(A)} = 0,$$

which shows the closability since the above inequality implies $y = 0$. Let us denote the closure of B_α in X again by B_α. The next statement will be an auxiliary lemma for complex-valued functions.

Lemma 2.8 *Let* $r \in [0, 1)$, $u \in C_b\big([0,\infty)\big) \cap C^2\big((0,\infty)\big)$ *such that* $\big(t \mapsto t^r \frac{\mathrm{d}}{\mathrm{d}t} t^{-r} u'(t)\big) \in C_b\big([0,\infty)\big)$. *Then also* $\big(t \mapsto t^{-r} u'(t)\big) \in C_b\big([0,\infty)\big)$.

Proof For $r = 0$ this is a standard result. The integrability of u'' near $t = 0$ implies the continuity of u' at $t = 0$ while the boundedness of u' for large values of t follows from Taylor's theorem. For the general case consider the function v given by $v(s) := u\big(s^{\frac{1}{1+r}}\big)$. By assumption $v \in C_b\big([0,\infty)\big) \cap C^2\big((0,\infty)\big)$ and with the relationship $t := s^{\frac{1}{1+r}}$ we can write $v'(s) = u'(t)t^{-r}$ and

$$v''(s) = t^{-r} \frac{\mathrm{d}}{\mathrm{d}t} t^{-r} u'(t).$$

The claim follows now by applying the result for $r = 0$ to the function v. □

Corollary 2.9 *Let* $x \in \mathcal{D}(A)$ *and* $\alpha \in [1/2, 1)$. *Then*

$$\big(t \mapsto t^{1-2\alpha} \frac{\mathrm{d}}{\mathrm{d}t} U x(t)\big) \in C_b\big([0,\infty); X\big).$$

Proof Let $r := 2\alpha - 1$ and $x' \in X'$. Then Lemma 2.8 applied to $t \mapsto \langle x', U x(t)\rangle_{X'\times X}$ yields boundedness of this function. Hence, Ux is weakly bounded and therefore bounded. □

Let us now assume that, additionally to unique solvability of $P_\alpha(x)$ for $x \in \mathcal{D}(A)$, the 'conjugated problem' $P_{1-\alpha}(x)$ is also uniquely solvable for all $x \in \mathcal{D}(A)$. We adapt all so far introduced notations by simply changing α to $1 - \alpha$ and specify in the following the solution operators by indices, i.e., $U_\alpha x$ will be the solution to $P_\alpha(x)$ while $U_{1-\alpha} x$ solves $P_{1-\alpha}(x)$. With this notation we can now state and prove the main theorem of this paper.

Theorem 2.10 *Let* $x \in \mathcal{D}(A^2)$. *Then* $t^{1-2\alpha} \frac{\mathrm{d}}{\mathrm{d}t} U_\alpha x(t) = U_{1-\alpha} B_\alpha x(t)$ *for all* $t \in [0,\infty)$.

For $t = 0$ we interpret the left-hand side of the equation in Theorem 2.10 as the limit from the right.

Proof For the proof we will distinguish two cases for α.

(i) Let us begin with $\alpha \in [1/2, 1)$. Define

$$v := \big(t \mapsto t^{1-2\alpha} \frac{\mathrm{d}}{\mathrm{d}t} U_\alpha x(t)\big).$$

By Corollary 2.9 we have $v \in C_b\big([0,\infty); X\big)$ and by the definition of B_α we have $v(0) = B_\alpha x$. Let us finally check that v solves the ODE for $P_{1-\alpha}(B_\alpha x)$. Then $v = U_{1-\alpha} B_\alpha x$. Let $t > 0$. Then

$$t^{2\alpha-1} \frac{\mathrm{d}}{\mathrm{d}t} v(t) = t^{2\alpha-1} \frac{\mathrm{d}}{\mathrm{d}t} t^{1-2\alpha} \frac{\mathrm{d}}{\mathrm{d}t} U_\alpha x(t) = A U_\alpha x(t) = U_\alpha A x(t)$$

and

$$t^{1-2\alpha} \frac{\mathrm{d}}{\mathrm{d}t} t^{2\alpha-1} \frac{\mathrm{d}}{\mathrm{d}t} v(t) = t^{1-2\alpha} \frac{\mathrm{d}}{\mathrm{d}t} U_\alpha A x(t) = A v(t),$$

where we used the closedness of A for the last equality.

(ii) Let us turn to the case $\alpha \in (0, 1/2)$. In this case, consider the function $\tilde{v}$ given by

$$\tilde{v}(t) := t^{2\alpha-1} v'(t) = U_\alpha A x(t) \quad (t \geq 0),$$

which belongs to $C_b\big([0,\infty); \mathcal{D}(A)\big) \subseteq C_b\big([0,\infty); X\big)$ and which solves $P_\alpha(Ax)$. In particular, $\tilde{v}$ solves $\widetilde{P_\alpha}(Ax)$. Further let w be defined by

$$w(t) := U_{1-\alpha} B_\alpha x(t) \quad (t \geq 0).$$

Let $\lambda \in \rho(A)$. Since $x \in \mathcal{D}(A^2)$ there exists $y \in \mathcal{D}(A)$ such that $x = (\lambda - A)^{-1} y$. It follows that

$$B_\alpha x = B_\alpha (\lambda - A)^{-1} y = (\lambda - A)^{-1} B_\alpha y \in \mathcal{D}(A).$$

This implies that

$$\tilde{w} := \big(t \mapsto t^{2\alpha-1} w'(t)\big) \in C_b\big([0,\infty); X\big),$$

where for the boundedness of $\tilde{w}$ we used Lemma 2.8 with $r := 1 - 2\alpha$. Note that $\tilde{w}(0) = B_{1-\alpha} B_\alpha x$. If we apply (i) to $1 - \alpha$ instead of α we find

$$t^{1-2\alpha} \frac{\mathrm{d}}{\mathrm{d}t} t^{2\alpha-1} \frac{\mathrm{d}}{\mathrm{d}t} U_{1-\alpha} x(t) = U_{1-\alpha} A x(t) = t^{1-2\alpha} \frac{\mathrm{d}}{\mathrm{d}t} U_\alpha B_{1-\alpha} x(t).$$

Sending $t \to 0+$ we conclude $Ax = B_\alpha B_{1-\alpha} x$. Let us show that the operators B_α and $B_{1-\alpha}$ commute on $\mathcal{D}(A^2)$. Note that for $y \in \mathcal{D}(A)$ we have $B_{1-\alpha}(\lambda - A)^{-1} y = (\lambda - A)^{-1} B_{1-\alpha} y$ (by uniqueness of solutions). Hence, we get $B_{1-\alpha} Ax = AB_{1-\alpha} x$ for $x \in \mathcal{D}(A^2)$. This implies that the bounded operator $B_{1-\alpha}(\lambda - A)^{-1}$ commutes on $\mathcal{D}(A)$ with A and this in turn implies the commutativity with B_α. So finally

$$\begin{aligned} B_\alpha B_{1-\alpha}(\lambda - A)^{-1} y &= B_{1-\alpha}(\lambda - A)^{-1} B_\alpha y \\ &= B_{1-\alpha} B_\alpha (\lambda - A)^{-1} y = B_{1-\alpha} B_\alpha x. \end{aligned}$$

Therefore, we observe $\tilde{w}(0) = Ax$. Finally, we show that $\tilde{w}$ solves $\widetilde{P_\alpha}(Ax)$. Let $\phi \in C_c^\infty\big((0,\infty)\big)$. Since w solves the ODE in (1.1) we obtain

$$\int_0^\infty \tilde{w}(t) \frac{\mathrm{d}}{\mathrm{d}t} t^{1-2\alpha} \frac{\mathrm{d}}{\mathrm{d}t} t^{2\alpha-1} \phi(t) \mathrm{d}t = -\int_0^\infty AU_{1-\alpha} B_\alpha x(t) \frac{\mathrm{d}}{\mathrm{d}t} t^{2\alpha-1} \phi(t) \mathrm{d}t.$$

If we apply Hille's theorem to the operator A it follows that

$$\int_0^\infty \tilde{w}(t)\phi(t)\mathrm{d}t = -\int_0^\infty U_{1-\alpha} B_\alpha x(t) \frac{\mathrm{d}}{\mathrm{d}t} t^{2\alpha-1} \phi(t) \mathrm{d}t \in \mathcal{D}(A)$$

and

$$-\int_0^\infty AU_{1-\alpha} B_\alpha x(t) \frac{\mathrm{d}}{\mathrm{d}t} t^{2\alpha-1} \phi(t) \mathrm{d}t = A \int_0^\infty \tilde{w}(t)\phi(t)\mathrm{d}t.$$

We conclude $\tilde{w} = \tilde{v}$ by uniqueness of the solution to $\widetilde{P_\alpha}(Ax)$, and therefore $w' = v'$. Since $w(0) = v(0)$ it follows that $w = v$. □

Remark 2.11 On the first glance it seems awkward that the existence of a solution to $P_\alpha(x)$ may not imply the existence of a solution to $P_{1-\alpha}(x)$, but meanwhile the authors are convinced that this is may indeed be the case, although a concrete example for this is missing. Note that in the case $\alpha = 1/2$ both problems are identical which is why in this case only solvability of one problem is needed to prove the result.

As a corollary to Theorem 2.10 we obtain a result similar to the one of Lemma 2.7.

Corollary 2.12 *Let $x \in X$. Then $x \in \mathcal{D}(B_\alpha)$ if and only if $U_\alpha x \in \mathcal{D}(D_\alpha)$. In either case, $D_\alpha U_\alpha x = U_{1-\alpha} B_\alpha x$.*

Proof Let $x \in \mathcal{D}(B_\alpha)$. Then there exists a sequence (x_n) in $\mathcal{D}(A)$ which is convergent towards x in X such that $B_\alpha x_n \to B_\alpha x$. By Lemma 2.4, we have $U_\alpha x_n \to U_\alpha x$ and by Theorem 2.10 we obtain

$$D_\alpha U_\alpha x_n = U_{1-\alpha} B_\alpha x_n \to U_{1-\alpha} B_\alpha x.$$

Since D_α is closed we get $U_\alpha x \in \mathcal{D}(D_\alpha)$ and $D_\alpha U_\alpha x = U_{1-\alpha} B_\alpha x$.

Conversely, let $U_\alpha x \in \mathcal{D}(D_\alpha)$. Choose a sequence (x_n) in $\mathcal{D}(A^2)$ such that $x_n \to x$ in X (since A is densely defined with non-empty resolvent set all its powers A^k are densely defined as well). Let $s > 0$. Choose a delta sequence (ϕ_k) in $C_c^\infty\big((0,\infty)\big)$ which converges towards δ_s in $C\big((0,\infty)\big)'$, i.e.,

$$\forall f \in C\big((0,\infty)\big) : \lim_{k\to\infty} \int_0^\infty f(t) t \phi_k(t) \mathrm{d}t = f(s).$$

By Theorem 2.10, for $t \geq 0$ we have

$$t^{1-2\alpha} \frac{\mathrm{d}}{\mathrm{d}t} U_\alpha x_n(t) = U_{1-\alpha} B_\alpha x_n(t) = B_\alpha U_{1-\alpha} x_n(t).$$

Using the closedness of B_α, by Hille's theorem we obtain

$$\int_0^\infty t^{1-2\alpha} \frac{\mathrm{d}}{\mathrm{d}t} U_\alpha(t) x_n \, \phi_k(t) \mathrm{d}t = B_\alpha \int_0^\infty U_{1-\alpha}(t) x_n \, \phi_k(t) \mathrm{d}t.$$

Sending $n \to \infty$ we get

$$\int_0^\infty t^{1-2\alpha} \frac{\mathrm{d}}{\mathrm{d}t} U_\alpha(t) x \, \phi_k(t) \mathrm{d}t = B_\alpha \int_0^\infty U_{1-\alpha}(t) x \, \phi_k(t) \mathrm{d}t.$$

Now, the limit $k \to \infty$ yields by closedness of B_α that

$$s^{1-2\alpha} \frac{\mathrm{d}}{\mathrm{d}s} U_\alpha(s) x = B_\alpha U_{1-\alpha} x(s)$$

Sending $s \to 0+$ and using a last time the closedness of B_α finally yields $x \in \mathcal{D}(B_\alpha)$. □

Corollary 2.13 *We have* $B_\alpha B_{1-\alpha} = B_{1-\alpha} B_\alpha = A$.

Proof Note, that we already showed $B_\alpha B_{1-\alpha} x = B_{1-\alpha} B_\alpha x = Ax$ for $x \in \mathcal{D}(A^2)$. So, let now $x \in \mathcal{D}(A)$. Then $x \in \mathcal{D}(B_\alpha)$, essentially by definition and the fact that

$D_\alpha U_\alpha x = U_{1-\alpha} B_\alpha x$ by Corollary 2.12. But since $x \in \mathcal{D}(A)$ we have $D_\alpha U_\alpha x = U_{1-\alpha} B_\alpha x \in \mathcal{D}(D_{1-\alpha})$ so applying Corollary 2.12 once more with $1-\alpha$ instead of α we obtain $B_\alpha x \in \mathcal{D}(B_{1-\alpha})$ and

$$U_\alpha B_{1-\alpha} B_\alpha x = D_{1-\alpha} D_\alpha U_\alpha x = A U_\alpha x = U_\alpha A x$$

in $C_b\big([0,\infty); X\big)$. Evaluating at 0 yields $B_{1-\alpha} B_\alpha x = Ax$. Conversely, let $x \in \mathcal{D}(B_{1-\alpha} B_\alpha)$. By two applications of Corollary 2.12 we then obtain that $U_\alpha x \in \mathcal{D}(D_{1-\alpha} D_\alpha)$. By Lemma 2.7, this implies $x \in \mathcal{D}(A)$. The equality $B_{1-\alpha} B_\alpha x = Ax$ is now verified in the same manner as before. The missing part of the Corollary follows by exchanging α and $1-\alpha$. □

We end this paper by a conjecture, formulated as an open problem. Assume that there exists an open sector $S_\phi \subseteq \mathbb{C}$ of (half-)opening angle $\phi \in (0, \pi/2)$ around the positive real axis such that for every $\delta \in (0, \phi)$ the function $U_\alpha x$ is analytic on S_δ and continuous and bounded on $\overline{S_\delta}$ for all $x \in \mathcal{D}(A)$ w.r.t. the norm $\|\cdot\|_{\mathcal{D}(A)}$. Note that this implies that the same holds for $U_{1-\alpha} x$. Actually, it is well known that this even implies analyticity of the mappings U_α and $U_{1-\alpha}$ w.r.t. the topology of $\mathcal{L}\big(\mathcal{D}(A)\big)$. We conjecture that A is then sectorial of angle $\omega := \pi - 2\phi$, and $B_\alpha = A^\alpha$ and $B_{1-\alpha} = A^{1-\alpha}$. Put differently, does analyticity of the solutions of $P_\alpha(x)$ for all $x \in \mathcal{D}(A)$ yield sectoriality of the operator A? An affirmative answer to this question would yield the converse of [7, Theorem 2.2] which states that sectoriality of A yields analyticity of the solution of $P_\alpha(x)$.

References

1. W. Arendt, C.J.K. Batty, M. Hieber, F. Neubrander, *Vector-valued Laplace Transforms and Cauchy Problems*, 2nd edn. (Birkhäuser, Basel, 2011)
2. W. Arendt, A.F.M. ter Elst, M. Warma, Fractional powers of sectorial operators via the Dirichlet-to-Neumann operator. Comm. Partial Differential Equations **43**, 1–24 (2018)
3. A.V. Balakrishnan, Fractional powers of closed operators and the semigroups generated by them. Pacific J. Math. **10**, 419–437 (1960)
4. L. Caffarelli, L. Silvestre, An extension problem related to the fractional Laplacian. Comm. Partial Differential Equations **32**(8), 1245–1260 (2007)
5. J.E. Galé, P.J. Miana, P.R. Stinga, Extension problem and fractional operators: semigroups and wave equations. J. Evol. Equ. **13**, 343–368 (2013)
6. C. Martinez, M. Sanz, *The Theory of Fractional Powers of Operators*. North-Holland Mathematics Studies, vol. 187 (Elsevier Science, Amsterdam, 2001)
7. J. Meichsner, C. Seifert, A Note on the harmonic extension approach to fractional powers of non-densely defined operators. Proc. Appl. Math. Mech. **19**, e201900296 (2019)
8. J. Meichsner, C. Seifert, On the harmonic extension approach to fractional powers in Banach spaces. Arxiv preprint (2019). https://arxiv.org/abs/1905.06779
9. P.R. Stinga, J.L. Torrea, Extension problem and Harnack's inequality for some fractional operators. Comm. Partial Differential Equations **35**, 2092–2122 (2010)

Time-Dependent Approach to Uniqueness of the Sommerfeld Solution to a Problem of Diffraction by a Half-Plane

A. Merzon, P. Zhevandrov, J. E. De la Paz Méndez, and T. J. Villalba Vega

Abstract We consider the Sommerfeld problem of diffraction by an opaque half-plane interpreting it as the limiting case as $t \to \infty$ of the corresponding non-stationary diffraction problem. We prove that the Sommerfeld formula for the solution is the limiting amplitude of the solution of this non-stationary problem which belongs to a certain functional class and is unique in it. For the proof of the uniqueness of solution to the non-stationary problem we reduce this problem, after the Fourier–Laplace transform in t, to a stationary diffraction problem with a complex wave number. This permits us to use the proof of the uniqueness in the Sobolev space H^1 as in (Castro and Kapanadze, J Math Anal Appl 421(2):1295–1314, 2015). Thus we avoid imposing the radiation condition from the beginning and instead obtain it in a natural way.

Keywords Diffraction · Uniqueness

Mathematics Subject Classification (2010) Primary 35Q60; Secondary 78A45

Supported by CONACYT and CIC (UMSNH).

A. Merzon (✉)
Instituto de Física y Matemáticas, Universidad Michoacana de San Nicolás de Hidalgo, Morelia, Michoacán, México

P. Zhevandrov
Facultad de Ciencias Físico-Matemáticas, Universidad Michoacana de San Nicolás de Hidalgo, Morelia, Michoacán, México

J. E. De la Paz Méndez · T. J. Villalba Vega
Universidad Autonoma de Guerrero, Chilpancingo, Guerrero, México

M. A. Bastos et al. (eds.), *Operator Theory, Functional Analysis and Applications*,
Operator Theory: Advances and Applications 282,
https://doi.org/10.1007/978-3-030-51945-2_23

1 Introduction

The main goal of this paper is to prove the uniqueness of a solution to the Sommerfeld half-plane problem [23, 32, 33] with a real wave number, proceeding from the uniqueness of the corresponding time-dependent problem in a certain functional class. The existence and uniqueness of solutions to this problem was considered in many papers, for example in [8, 12, 25]. However, in our opinion, the problem of uniqueness is still not solved in a satisfactory form from the point of view of the boundary value problems (BVPs). The fact is that this problem is a homogeneous BVP boundary value problem which admits various nontrivial solutions. Usually the "correct" solutions are chosen by physical reasoning [23, 25, 32, 33], for example, using the Sommerfeld radiation conditions and regularity conditions at the edge.

The question is: from where do the radiation and regularity conditions arise, from the mathematical point of view?

Our goal is to show that they arise automatically from the *non-stationary* problem. This means the following: we prove that the Sommerfeld solution is a limiting amplitude of a solution to the corresponding non-stationary problem which is unique in an appropriate functional class. Since the Sommerfeld solution, as is well-known, satisfies the radiation and regularity conditions, our limiting amplitude also satisfies them. Of course, the limiting amplitude principle (LAP) is very well-known for the diffraction by smooth obstacles, see e.g. [28, 29], but we are unaware of its rigorous proof in the case of diffraction by a half-plane.

The literature devoted to diffraction by wedges including the Sommerfeld problem is enormous (see e.g. the review in [20]), and we will only indicate some papers where the uniqueness is treated. In paper [25] a uniqueness theorem was proven for the Helmholtz equation $(\Delta + 1)u = 0$ in two-dimensional regions D of half-plane type. These regions can have a finite number of bounded obstacles with singularities on their boundaries. In particular, the uniqueness of solution u to the Sommerfeld problem was proven by means of the decomposition of the solution into the sum $u = g + h$, where g describes the geometrical optics incoming and reflected waves and h satisfies the Sommerfeld radiation condition (clearly, u should also satisfy the regularity conditions at the edge).

In paper [8] exact conditions were found for the uniqueness in the case of complex wave number. The problem was considered in Sobolev spaces for a wide class of generalized incident waves, and for DD and NN boundary conditions. In paper [12] the same problem was considered also for the complex wave number and for DN boundary conditions. In both papers the Wiener-Hopf method has been used. Time-dependent scattering by wedges was considered in many papers although their number is not so large as the number of papers devoted to the stationary scattering by wedges. We indicate here the following papers: [1–4, 13, 14, 24, 26–31]. The detailed description of these papers is given in [19].

In [6, 7, 10, 17–20, 22], the diffraction by a wedge of magnitude ϕ (which can be a half-plane in the case $\phi = 0$ as in [20]) with real wavenumber was considered as a stationary problem which is the "limiting case" of a non-stationary one. More

precisely, we seek the solutions of the classical diffraction problems as *limiting amplitudes* of solutions to corresponding non-stationary problems, which are unique in some appropriate functional class. We also, like in [25], decomposed the solution of non-stationary problem separating a "bad" incident wave, so that the other part of solution belongs to a certain appropriate functional class. Thus we avoided the a priori use of the radiation and regularity conditions and instead obtained them in a natural way. In papers [10, 17, 19] we considered the time-dependent scattering with DD, DN and NN boundary conditions and proved the uniqueness of solution in an appropriate functional class. But these results were obtained only for $\phi \neq 0$ because in the proof of uniqueness we used the Method of Complex Characteristics [15, 16, 21] which "works" only for $\phi \neq 0$.

For $\phi = 0$ we need to use other methods, namely, the reduction of the uniqueness problem for the stationary diffraction to the uniqueness problem for the corresponding non-stationary diffraction, which, in turn, is reduced to the proof of uniqueness of solution of the stationary problem but with a *complex wavenumber*, see e.g. [5].

Note that in [18] we proved the LAP for $\phi \neq 0$ and for the DD boundary conditions. Similar results for the NN and DN boundary conditions were obtained in [6, 7, 10]. A generalization of these results to the case of generalized incident wave (cf. [8]) was given in [19]. This approach (stationary diffraction as the limit of time-dependent one) permits us to justify all the classical explicit formulas [13, 14, 20, 28–31] and to prove their coincidence with the explicit formulas given in [17, 19, 22]. In other words, all the classical formulas are limiting amplitudes of solutions to non-stationary problems as $t \to \infty$. For the Sommerfeld problem, this was proven in [20], except for the proof of the uniqueness of the solution to the non-stationary problem in an appropriate class. This paper makes up for this omission.

Our plan is as follows. The non-stationary diffraction problem is reduced by means of the Fourier–Laplace transform with respect to time t to a stationary one with a complex wave number. For this problem the uniqueness theorems can be proven more easily in Sobolev classes (see an important paper [5]) and do not use the radiation conditions. Then we prove that the Fourier–Laplace transforms of solutions to non-stationary diffraction half-plane problem, whose amplitude tends to the Sommerfeld solution, also belong to a Sobolev space for a rather wide class of incident waves. This permits us to reduce the problem to the case of [5].

Let us pass to the problem setting. We consider the two-dimensional time-dependent scattering of a plane wave by the half-plane

$$W^0 := \left\{(x_1, x_2) \in \mathbb{R}^2 : x_2 = 0, x_1 \geq 0\right\}.$$

(Obviously, W^0 is a half-line in $\mathbb{R}^2$, but if one recalls that the initial problem is three-dimensional, W^0 becomes a half-plane; the third coordinate is suppressed in

all what follows.) The non-stationary incident plane wave in the absence of obstacles reads

$$u_i(x,t) = e^{-i\omega_0(t-\mathbf{n}\cdot x)} f(t-\mathbf{n}\cdot x), \quad x \in \mathbb{R}^2, \quad t \in \mathbb{R}, \tag{1.1}$$

where

$$\omega_0 > 0, \quad \mathbf{n} = (n_1, n_2) = (\cos(\pi+\alpha), \sin(\pi+\alpha)), \tag{1.2}$$

and f is "a profile function", such that $f \in L^1_{loc}(\mathbb{R})$, and

$$f(s) = 0,\ s < 0,\ \sup(1+|s|)^p |f(s)| < \infty \text{ for some } p \in \mathbb{R},\ \lim_{s\to+\infty} f(s) = 1. \tag{1.3}$$

Remark 1.1 Obviously, these functions satisfy the D'Alembert equation $\Box u_i(x,t) = 0$ in the sense of distributions.

For definiteness, we assume that

$$\frac{\pi}{2} < \alpha < \pi. \tag{1.4}$$

In this case the front of the incident wave u_i reaches the half-plane W^0 for the first time at the moment $t = 0$ and at this moment the reflected wave $u_r(x,t)$ is born (see Fig. 1). Thus

$$u_r(x,t) \equiv 0, \quad t < 0.$$

Note that for $t \to \infty$ the limiting amplitude of u_i is exactly equal to the Sommerfeld incident wave [33] by (1.3), cf. also (2.1) below.

The time-dependent scattering with the Dirichlet boundary conditions is described by the mixed problem

$$\left.\begin{cases} \Box u(x,t) := (\partial_t^2 - \Delta)u(x,t) = 0,\ x \in Q \\ u(x_1, \pm 0, t) = 0,\ x_1 > 0 \end{cases}\right| \ t \in \mathbb{R}, \tag{1.5}$$

where $Q := \mathbb{R}^2 \setminus W^0$. The "initial condition" reads

$$u(x,t) = u_i(x,t), \quad x \in Q, \quad t < 0, \tag{1.6}$$

where u_i is the incident plane wave (1.1).

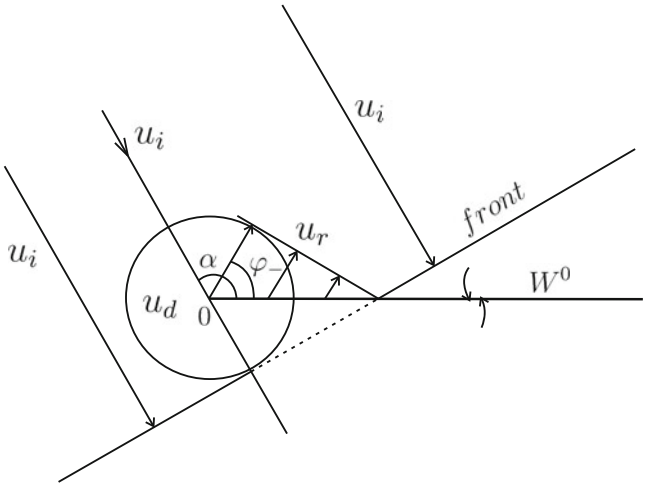


Fig. 1 Time-dependent diffraction by a half-plane

Introduce the non-stationary "scattered" wave u_s as the difference between u and u_i,

$$u_s(x,t) := u(x,t) - u_i(x,t), \quad x \in Q, \quad t \in \mathbb{R}. \tag{1.7}$$

Since $\Box u_i(x,t) = 0,\ (x,t) \in Q \times \mathbb{R}$, we get from (1.6), (1.5) that

$$\Box u_s(x,t) = 0, \quad (x,t) \in Q \times \mathbb{R}, \tag{1.8}$$

$$u_s(x,t) = 0, \quad x \in Q, \quad t < 0, \tag{1.9}$$

$$u_s(x_1, \pm 0, t) = -u_i(x_1, 0, t), \quad x_1 > 0, \quad t > 0. \tag{1.10}$$

Denote

$$\varphi_\pm := \pi \pm \alpha. \tag{1.11}$$

Everywhere below we assume that

$$x_1 = r\cos\varphi, \quad x_2 = r\sin\varphi, \quad 0 \le \varphi < 2\pi. \tag{1.12}$$

Let us define the nonstationary incident wave in the presence of the obstacle W^0, which is the opaque screen,

$$u_i^0(\rho,\varphi,t) := \begin{cases} u_i(\rho,\varphi,t), & 0 < \varphi < \varphi_+, \\ 0, & \varphi_+ < \varphi < 2\pi. \end{cases} \tag{1.13}$$

Remark 1.2 The function u_s has no physical sense, since $u_i \neq u_i^0$. The wave u_s coincides with the scattered wave $u_s^0 := u - u_i^0$ in the zone $\{(\rho, \varphi) : 0 < \varphi < \varphi_+\}$, but in the zone $\{(\rho, \varphi) : \varphi_+ < \varphi \leq 2\pi\}$ we have $u_s^0 = u_s + u_i$.

The goal of the paper is to prove that the Sommerfeld solution of half-plane diffraction problem is the limiting amplitude of the solution to time-dependent problem (1.5), (1.6) (with any f satisfying (1.3)) and this solution is unique in an appropriate functional class.

The paper is organized as follows. In Sect. 2 we recall the Sommerfeld solution. In Sect. 3 we reduce the time-dependent diffraction problem to a "stationary" one and define a functional class of solutions. In Sect. 4 we give an explicit formula for the solution of time-dependent problem and prove that the Sommerfeld solution is its limiting amplitude. In Sect. 5 we prove that the solution belongs to a certain functional class. Finally, in Sect. 6 we prove the uniqueness.

2 Sommerfeld's Diffraction

Let us recall the Sommerfeld solution [23, 33]. The stationary incident wave (rather, the incident wave limiting amplitude) in the presence of the obstacle is

$$\mathcal{A}_i^0(\rho, \varphi) = \begin{cases} e^{-i\omega_0 \rho \cos(\varphi - \alpha)}, & \varphi \in (0, \varphi_+), \\ 0, & \varphi \in (\varphi_+, 2\pi). \end{cases} \tag{2.1}$$

We denote this incident wave as $\mathcal{A}_i^0$ since it is the limiting amplitude of the non-stationary incident wave u_i^0 given by (1.13):

$$\mathcal{A}_i^0(\rho, \varphi) = \lim_{t \to \infty} e^{i\omega_0 t} u_i^0(x, t),$$

in view of formula (1.1), see Remark 1.2. The Sommerfeld half-plane diffraction problem can be formulated as follows: find a function $\mathcal{A}(x)$, $x \in \overline{Q}$, such that

$$\begin{cases} (\Delta + \omega_0^2)\mathcal{A}(x) = 0, & x \in Q, \\ \mathcal{A}(x_1, \pm 0) = 0, & x_1 > 0, \end{cases} \tag{2.2}$$

$$\mathcal{A}(x) = \mathcal{A}_i^0(x) + \mathcal{A}_r(x) + \mathcal{A}_d(x), \quad x \in Q, \tag{2.3}$$

where $\mathcal{A}_r(x)$ is the reflected wave,

$$\mathcal{A}_r(x) = \begin{cases} -e^{-i\omega_0 \rho \cos(\varphi + \alpha)}, & \varphi \in (0, \varphi_-), \\ 0, & \varphi \in (\varphi_-, 2\pi), \end{cases} \tag{2.4}$$

and $\mathcal{A}_d(x)$ is the wave diffracted by the edge,

$$\mathcal{A}_d(x) \to 0, \quad |x| \to \infty. \tag{2.5}$$

A. Sommerfeld [33] found the solution of this problem in the form

$$\mathcal{A}(\rho, \varphi) = \frac{1}{4\pi} \int_{\mathcal{C}} \zeta(\gamma, \varphi) e^{-i\omega\rho\cos\gamma} d\gamma, \quad \rho \geq 0, \quad \varphi \in [0, 2\pi],$$

where

$$\zeta(\gamma, \varphi) := \left(1 - e^{i(-\gamma+\varphi-\alpha)/2}\right)^{-1} - \left(1 - e^{i(-\gamma+\varphi+\alpha)/2}\right)^{-1}, \quad \gamma \in \mathbb{C} \tag{2.6}$$

and $\mathcal{C}$ is the Sommerfeld contour (see [20, formula (1.1) and Fig. 3]).

In the rest of the paper we prove that this solution is the limiting amplitude of the solution of time-dependent problem (1.5) and is unique in an appropriate functional class.

The Sommerfeld diffraction problem can also be considered for NN and DN half-plane. The corresponding formulas for the solution can be found in [19].

Sommerfeld obtained his solution using an original method of solutions of the Helmholtz equation on a Riemann surface. Note that a similar approach was used for the diffraction by a wedge of rational angle [9], where well-posedness in suitable Sobolev space was proved.

3 Reduction to a "Stationary" Problem: Fourier–Laplace Transform

Let $\widehat{h}(\omega)$, $\omega \in \mathbb{C}^+$, denote the Fourier–Laplace transform $\mathcal{F}_{t\to\omega}$ of $h(t)$,

$$\widehat{h}(\omega) = \mathcal{F}_{t\to\omega}[h(t)] = \int_0^\infty e^{i\omega t} h(t)\, dt, \quad h \in L_1(\mathbb{R}^+); \tag{3.1}$$

$\mathcal{F}_{t\to\omega}$ is extended by continuity to $S'(\overline{\mathbb{R}^+})$. Assuming that $u_s(x,t)$ belongs to $S'(\mathbb{R}^2 \times \overline{\mathbb{R}^+})$ (see (1.9) and Definition 3.1), we apply this transform to system (1.8)–(1.10), and obtain

$$\left\{ \begin{array}{ll} (\Delta + \omega^2)\widehat{u}_s(x,\omega) = 0, & x \in Q, \\ \widehat{u}_s(x_1, \pm 0, \omega) = -\widehat{u}_i(x_1, \pm 0, \omega), & x_1 > 0 \end{array} \right| \quad \omega \in \mathbb{C}^+. \tag{3.2}$$

Let us calculate $\widehat{u}_i(x,\omega)$. Changing the variable $t-\mathbf{n}\cdot x=\tau$, and using the fact that $\operatorname{supp} f\subset\overline{\mathbb{R}^+}$ we obtain from (1.1) and (1.2) that

$$\widehat{u}_i(x,\omega)=e^{i\omega\mathbf{n}\cdot x}\,\widehat{f}(\omega-\omega_0). \tag{3.3}$$

Hence,

$$\widehat{u}_i(x_1,0,\omega)=e^{i\omega n_1 x_1}\,\widehat{f}(\omega-\omega_0),\quad x_1>0,$$

and the boundary condition in (3.2) is $\widehat{u}_s(x_1,0,\omega)=-g(\omega)e^{i\omega n_1 x_1}$. Therefore we come to the following family of BVPs depending on $\omega\in\mathbb{C}^+$: find $\widehat{u}_s(x,\omega)$ such that

$$\begin{cases}(\Delta+\omega^2)\widehat{u}_s(x,\omega)=0, & x\in Q,\\ \widehat{u}_s(x_1,\pm 0,\omega)=-g(\omega)e^{i\omega n_1 x_1}, & x_1>0.\end{cases} \tag{3.4}$$

We are going to prove the existence and uniqueness of solution to problem (1.5), (1.6) such that u_s given by (1.7) belongs to the space $\mathcal{M}$, which is defined as follows:

Definition 3.1 $\mathcal{M}$ is the space of functions $u(x,t)\in S'(\mathbb{R}^2\times\overline{\mathbb{R}^+})$ such that its Fourier–Laplace transform $\widehat{u}(x,\omega)$ is a holomorphic function on $\omega\in\mathbb{C}^+$ with values in $C^2(Q)$ and

$$\widehat{u}(\cdot,\cdot,\omega)\in H^1(Q) \tag{3.5}$$

for any $\omega\in\mathbb{C}^+$.

Remark 3.2 We use the classical definition [11] of the space $H^1(Q)$ as the completion of the space of smooth functions on Q with respect to the corresponding norm. This definition does not coincide with the frequently used definition of $H^1(Q)$ as the space restrictions of distributions from $H^1(\mathbb{R})$ to Q. In our case these definitions lead to different spaces; in particular, the latter definition does not allow for functions which are discontinuous across W^0. In [34], another space allowing for the same class of functions was introduced; the proof of uniqueness of the solution to our problem in that space is an open question.

Remark 3.3 Note that $u_i(x,t)\big|_{\mathbb{R}^2\times\overline{\mathbb{R}^+}}\notin\mathcal{M}$, where for $\varphi\in D(\mathbb{R}^2)$,

$$\left\langle u_i(x,t)\big|_{\mathbb{R}^2\times\overline{\mathbb{R}^+}},\varphi\right\rangle:=\int\limits_{\mathbb{R}^2\times\mathbb{R}^+}u(x,t)\varphi(x,t)\,dx\,dt.$$

In fact, $\left|e^{i\omega\mathbf{n}\cdot x}\right|=e^{\omega_2\rho\cos(\varphi-\alpha)}$ and, for $\alpha-\pi/2<\varphi<\alpha+\pi/2$, $\omega\in\mathbb{C}^+$ it grows exponentially as $\rho\to\infty$, and hence does not satisfy (3.5); because of this we use system (1.8)–(1.10) instead of (1.5) (they are equivalent by (1.6)) since (1.8)–(1.10)

involves only the values of u_i on the boundary and the latter possess the Fourier–Laplace transforms which do not grow exponentially.

Remark 3.4 Since for a (weak) solution of the Helmholtz equation $u_s \in H^1(Q)$ the Dirichlet and Neumann data exist in the trace sense and in the distributional sense, respectively (see, e.g., [5]), problem (3.4) is well-posed. Hence, problem (1.8)–(1.10) is well-posed too.

4 Connection Between the Non-stationary Diffraction Problem (1.5) and (1.6) and the Sommerfeld Half-Plane Problem

In paper [20] we solved problem (1.5) and (1.6). Let us recall the corresponding construction. First we define the non-stationary reflected wave [20, formula (26)]:

$$u_r(x,t) = \left\{ \begin{array}{cc} -e^{-i\omega_0(t-\overline{\mathbf{n}}\cdot x)} f(t-\overline{\mathbf{n}}\cdot x), & \varphi \in (0,\varphi_-) \\ 0, & \varphi \in (\varphi_-, 2\pi) \end{array} \right| \quad t \geq 0, \tag{4.1}$$

where $\overline{\mathbf{n}} := (n_1, -n_2) = (-\cos\alpha, \sin\alpha)$ (see Fig. 1).

Note that its limiting amplitude coincides with (2.4) similarly to the incident wave.

Second, we define the non-stationary diffracted wave (cf. [20, formula (31) for $\phi = 0$]). Let

$$\mathcal{Z}(\beta,\varphi) := Z(\beta + 2\pi i - i\varphi), \tag{4.2}$$

and

$$u_d(\rho,\varphi,t) = \frac{i}{8\pi}\int\limits_{\mathbb{R}} \mathcal{Z}(\beta,\varphi) F(t-\rho\cosh\beta)\, d\beta, \tag{4.3}$$

where $\varphi \in (0,2\pi)$, $\varphi \neq \varphi_\pm$; $t \geq 0$,

$$F(s) = f(s)e^{-i\omega_0 s}, \tag{4.4}$$

$$Z(z) = -U\Big(-\frac{i\pi}{2} + z\Big) + U\Big(-\frac{5i\pi}{2} + z\Big), \tag{4.5}$$

$$U(\zeta) = \coth\Big(q(\zeta - i\frac{\pi}{2} + i\alpha)\Big) - \coth\Big(q(\zeta - i\frac{\pi}{2} - i\alpha)\Big), \quad q = \frac{1}{4} \tag{4.6}$$

for the Dirichlet boundary conditions. Below in Lemma 8.1 we give the necessary properties of the function $\mathcal{Z}$, from which the convergence of integral (4.3) follows.

Obviously, the condition supp $F \subset [0, \infty)$ (see (3.1)) implies that supp $u_d(\cdot, \cdot, t) \subset [0, +\infty)$.

Remark 4.1 The function $U(\gamma + \varphi)$ essentially coincides with the Sommerfeld kernel (2.6). This is for a reason. In paper [17] it was proven that the solution to the corresponding time-dependent diffraction problem by an arbitrary angle $\phi \in (0, \pi]$ belonging to a certain class similar to $\mathcal{M}$ necessarily has the form of the Sommerfeld type integral with the Sommerfeld type kernel.

Finally, we proved [20, Th. 3.2, Th 4.1] the following.

Theorem 4.2

(i) *For $f \in L^1_{loc}(\mathbb{R})$ the function*

$$u(\rho, \varphi, t) := u_i^0(\rho, \varphi, t) + u_r(\rho, \varphi, t) + u_d(\rho, \varphi, t), \quad \varphi \neq \varphi_\pm \tag{4.7}$$

belongs to $L^1_{loc}(Q \times \mathbb{R}^+)$. It is continuous up to $\partial Q \times \mathbb{R}$ and satisfies the boundary and initial conditions (1.5), (1.6). The D'Alembert equation in (1.5) holds in the sense of distributions.

(ii) *The LAP holds for Sommerfeld's diffraction by a half-plane:*

$$\lim_{t \to \infty} e^{i\omega_0 t} u(\rho, \varphi, t) = \mathcal{A}(\rho, \varphi), \quad \varphi \neq \varphi_\pm$$

(the limit here and everywhere else is pointwise).

Since the main object of our consideration will be the "scattered" wave $u_s(x, t)$ given by (1.7), we clarify the connection between u_s and the Sommerfeld solution $\mathcal{A}$.

Corollary 4.3 *Define $\mathcal{A}_i(x) = e^{-i\omega_0 \rho \cos(\varphi + \alpha)}$, which is the limiting amplitude of $u_i(x, t)$ given by (1.1). The limiting amplitude of $u_s(x, t)$ is the function*

$$\mathcal{A}_s(x) = \mathcal{A}(x) - \mathcal{A}_i(x), \tag{4.8}$$

i.e. $\lim_{t \to \infty} e^{i\omega_0 t} u_s(x, t) = \mathcal{A}_s(x)$.

Proof The statement follows from (1.7). □

Remark 4.4 The function $\mathcal{A}_s$ is the limiting amplitude of the scattered non-stationary wave $u_s(x, t)$ and $\mathcal{A}_s$ satisfies the following nonhomogeneous BVP:

$$\begin{cases} (\Delta + \omega_0^2)\mathcal{A}_s(x) = 0, & x \in Q, \\ \mathcal{A}_s(x_1, \pm 0) = -\mathcal{A}_i(x_1, 0), & x_1 > 0. \end{cases} \tag{4.9}$$

This BVP (as well as (2.2)) is ill-posed since the homogeneous problem admits many solutions (i.e., the solution is nonunique).

Remark 4.5 $\mathcal{A}_s$ can be decomposed similarly to (2.3). Namely, by (4.8) and (2.3), we have

$$\mathcal{A}_s = \mathcal{A}_i^0 + \mathcal{A}_r(x) + \mathcal{A}_d(x) - \mathcal{A}_i(x) = \mathcal{A}_r(x) + \mathcal{A}_d(x) - \mathcal{A}_i^1(x), \tag{4.10}$$

where $\mathcal{A}_i^1(x) = \mathcal{A}_i(x) - \mathcal{A}_i^0(x)$. Obviously, problems (4.9), (4.10) and (2.2), (2.3) with condition (2.5) are equivalent, but the first problem is more convenient as we will see later.

5 Solution of the "Stationary" Problem

In this section we will obtain an explicit formula for the solution of (3.4) and prove that it belongs to $H^1(Q)$ for all $\omega \in \mathbb{C}^+$.

Let $\mathcal{Z}(\beta, \varphi)$ be given by (4.2). First, we will need the Fourier–Laplace transforms of the reflected and diffracted waves (4.1), (4.3).

Lemma 5.1 *The Fourier–Laplace transforms of u_r and u_d are*

$$\widehat{u}_r(x,\omega) = \begin{cases} -\widehat{f}(\omega-\omega_0)e^{-i\omega\rho\cos(\varphi+\alpha)}, & \varphi \in (0,\varphi_-), \\ 0, & \varphi \in (\varphi_-, 2\pi), \end{cases} \tag{5.1}$$

$$\widehat{u}_d(\rho,\varphi,\omega) = \frac{i}{8\pi}\widehat{f}(\omega-\omega_0)\int_{\mathbb{R}} \mathcal{Z}(\beta,\varphi)\, e^{i\omega\rho\cosh\beta}\, d\beta, \ \omega \in \mathbb{C}^+, \ \varphi \neq \varphi_\pm. \tag{5.2}$$

Proof From (4.1) we have

$$\widehat{u}_r(x,\omega) = \begin{cases} -\mathcal{F}_{t\to\omega}\left[e^{-i\omega_0(t-\overline{\mathbf{n}}\cdot x)} f(t-\overline{\mathbf{n}}\cdot x)\right], & \varphi \in (0,\varphi_-), \\ 0, & \varphi \in (\varphi_-, 2\pi). \end{cases}$$

Further,

$$-\mathcal{F}_{t\to\omega}\left[e^{-i\omega_0(t-\overline{\mathbf{n}}\cdot x)} f(t-\overline{\mathbf{n}}\cdot x)\right] = -e^{i\omega_0(\overline{\mathbf{n}}\cdot x)}\int_0^\infty e^{i(\omega-\omega_0)t} f(t-\overline{\mathbf{n}}\cdot x)\, dt.$$

Changing the variable $t - \overline{\mathbf{n}}\cdot x = \tau$, we obtain

$$\widehat{u}_r(x,\omega) = -e^{i\omega\,\overline{\mathbf{n}}\cdot x}\int_{-\overline{\mathbf{n}}\cdot x}^\infty e^{i(\omega-\omega_0)\tau} f(\tau)\, d\tau, \quad \varphi \in (0,\varphi_-).$$

Moreover, by (4.1),

$$-\overline{\mathbf{n}} \cdot x = \rho \cos(\varphi - \alpha) \le c < 0, \quad \varphi \in (0, \varphi_-),$$

since $\pi/2 < \alpha < \varphi + \alpha < \pi$ by (1.4) and (1.11). Hence, we obtain (5.1), since $\operatorname{supp} f \subset \overline{\mathbb{R}^+}$. The second formula in (5.1) follows from definition (4.1) of u_r.

Let us prove (5.2). Everywhere below we put $\omega = \omega_1 + i\omega_2$, $\omega_{1,2} \in \mathbb{R}$, $\omega_2 > 0$, for $\omega \in \mathbb{C}^+$. By Lemma 8.1(i), (1.3) and (4.4) we have

$$\left| e^{i\omega t} \mathcal{Z}(\beta, \varphi) F(t - \rho \cosh \beta) \right| \le C e^{-\omega_2 t} e^{-\beta/2} (1+t)^{-p}, \quad \rho < 0, \varphi \ne \varphi_\pm, \beta \in \mathbb{R}.$$

Hence, by the Fubini Theorem there exists the Fourier–Laplace transform of $u_d(\cdot, \cdot, t)$ and

$$\widehat{u}_d(\rho, \varphi, \omega) = \frac{i}{8\pi} \int_{\mathbb{R}} \mathcal{Z}(\beta, \varphi) \mathcal{F}_{t \to \omega} \Big[F(t - \rho \cosh \beta) \Big] \, d\beta, \quad \varphi \ne \varphi_\pm. \tag{5.3}$$

We have

$$G(\rho, \beta, \omega) := \mathcal{F}_{t \to \omega} \Big[F(t - \rho \cosh \beta) \Big] = \int_0^\infty e^{i\omega t} F(t - \rho \cosh \beta) \, dt, \quad \omega \in \mathbb{C}^+.$$

Making the change of the variable $\tau = t - \rho \cosh \beta$ in the last integral and using the fact that $\operatorname{supp} F \subset [0, \infty)$ and $\widehat{F}(\omega) = \widehat{f}(\omega - \omega_0)$ by (4.4), we get $G(\rho, \beta, \omega) = e^{i\omega\rho \cosh \beta} \widehat{f}(\omega - \omega_0)$. Substituting this expression into (5.3) we obtain (5.2). Lemma 5.1 is proven. □

5.1 Estimates for $\widehat{u}_r$, $\partial_\rho \widehat{u}_r$, $\partial_\varphi \widehat{u}_r$

Lemma 5.2 *For any $\omega \in \mathbb{C}$, there exist $C(\omega), c(\omega) > 0$, such that both functions $\widehat{u}_r$ and $\partial_\rho \widehat{u}_r$ admit the same estimate*

$$\left. \begin{aligned} \left| \widehat{u}_r(\rho, \varphi, \omega) \right| &\le C(\omega) e^{-c(\omega)\rho} \\ \left| \partial_\rho \widehat{u}_r(\rho, \varphi, \omega) \right| &\le C(\omega) e^{-c(\omega)\rho} \end{aligned} \right| \quad \rho > 0, \ \varphi \in (0, 2\pi), \ \varphi \ne \varphi_\pm. \tag{5.4}$$

and $\partial_\varphi \widehat{u}_r(\rho, \varphi, \omega)$ admits the estimate

$$|\partial_\varphi \widehat{u}_r(\rho, \varphi, \omega)| \le C(\omega) \rho \, e^{-c(\omega)\rho}, \quad \rho > 0. \tag{5.5}$$

Proof By (1.4) there exits $c(\omega) > 0$ such that

$$\left|e^{-i\omega\rho\cos(\varphi+\alpha)}\right| = e^{\omega_2\rho\cos(\varphi+\alpha)} \le e^{-c(\omega)\rho}, \quad 0 < \varphi < \varphi_-$$

by (1.4). Therefore (5.4) holds for $\widehat{u}_r$. Hence, differentiating (5.1) we obtain (5.4) for $\partial_\rho\widehat{u}_r$ and (5.5) for $\partial_\varphi\widehat{u}_r$, for $\varphi \neq \varphi_-$. □

5.2 *Estimates for $\widehat{u}_d$*

Proposition 5.3 *There exist $C(\omega), c(\omega) > 0$ such that the function $\widehat{u}_d$, and $\partial_\rho\widehat{u}_d$, $\partial_\varphi\widehat{u}_d$ admit the estimates*

$$\begin{aligned}
\left|\widehat{u}_d(\rho,\varphi,\omega)\right| &\le C(\omega)e^{-c(\omega)\rho},\\
\left|\partial_\rho\widehat{u}_d(\rho,\varphi,\omega)\right| &\le C(\omega)e^{-c(\omega)\rho}(1+\rho^{-1/2}),\\
\left|\partial_\varphi\widehat{u}_d(\rho,\varphi,\omega)\right| &\le C(\omega)e^{-c(\omega)\rho}\rho(1+\rho^{-1/2})
\end{aligned} \tag{5.6}$$

for $\rho > 0$, $\varphi \in (0, 2\pi)$, $\varphi \neq \varphi_\pm$.

Proof

(I) By (5.2), in order to prove (5.6) for $\widehat{u}_d$ it suffices to prove that

$$|A(\rho,\varphi,\omega)| \le C(\omega)e^{-c(\omega)\rho}, \tag{5.7}$$

where

$$A(\rho,\varphi,\omega) := \int_{\mathbb{R}} \mathcal{Z}(\beta,\varphi)e^{i\omega\rho\cosh\beta}\,d\beta, \quad \varphi \neq \varphi_\pm. \tag{5.8}$$

Represent A as $A = A_1 + A_2$, where

$$\left.\begin{aligned}
A_1(\rho,\varphi,\omega) &:= \int_{-1}^{1} \mathcal{Z}(\beta,\varphi)e^{i\omega\rho\cosh\beta}\,d\beta\\
A_2(\rho,\varphi,\omega) &:= \int_{|\beta|\ge 1} \mathcal{Z}(\beta,\varphi)e^{i\omega\rho\cosh\beta}\,d\beta
\end{aligned}\right| \quad \varphi \in (0,2\pi), \quad \varphi \neq \varphi_\pm. \tag{5.9}$$

The estimate (5.7) for A_2 follows from (8.1) (see Appendix 1). It remains to prove the same estimate for the function A_1. Let

$$\varepsilon_\pm := \varphi_\pm - \varphi. \tag{5.10}$$

Representing A_1 as

$$A_1(\rho,\varphi,\omega) = -4\mathcal{K}_0(\rho, w, \varepsilon_+) + 4\mathcal{K}_0(\rho, w, \varepsilon_-) + \int_{-1}^{1} \check{\mathcal{Z}}(\beta,\varphi) e^{i\omega\rho\cosh\beta}\, d\beta,$$

where $\mathcal{K}_0$ is defined by (8.7), we obtain (5.7) for A_1 from Lemma 8.2 (i) and (8.3).

(II) Let us prove (5.6) for $\partial_\rho \widehat{u}_d$. By (5.2) it suffices to prove that

$$|B(\rho,\varphi,\omega)| \le C(\omega) e^{-c(\omega)\rho}(1+\rho^{1/2}), \quad \varphi \ne \varphi_\pm, \tag{5.11}$$

where

$$B(\rho,\varphi,\omega) := \int_{\mathbb{R}} \mathcal{Z}(\beta,\varphi)\cosh\beta\; e^{i\omega\rho\cosh\beta}\, d\beta.$$

Represent B as $B_1 + B_2$, where $B_{1,2}(\rho,\varphi,\omega)$ are defined similarly to (5.9),

$$B_1(\rho,\varphi,\omega) := \int_{-1}^{1} \mathcal{Z}(\beta,\varphi)\cosh\beta e^{i\omega\rho\cosh\beta}\, d\beta,$$

$$B_2(\rho,\varphi,\omega) := \int_{|\beta|\ge 1} \mathcal{Z}(\beta,\varphi)\cosh\beta e^{i\omega\rho\cosh\beta}\, d\beta, \quad \varphi \ne \varphi_\pm.$$

From (8.1) for $\mathcal{Z}$ we have

$$|B_2(\rho,\varphi,\omega)| \le C_1 \int_1^\infty e^{\beta/2} e^{-\frac{1}{2}\omega_2\rho e^\beta}\, d\beta.$$

Making the change of the variable $\xi := \rho e^\beta$, we get

$$|B_2(\rho,\varphi,\omega)| \le \begin{cases} C_1(\omega)\rho^{-1/2}, & \rho \le 1, \\ \displaystyle\int_\rho^\infty \frac{e^{-\omega_2\xi/2}}{\xi^{1/2}} d\xi, & \rho \ge 1. \end{cases}$$

Since for $\rho \geq 1$,

$$\int_{\rho}^{\infty} \frac{e^{-\omega_2 \xi/2}}{\xi^{1/2}} d\xi \leq \frac{2}{\omega_2} e^{-\omega_2 \rho/2},$$

Equation (5.11) is proved for B_2.

It remains to prove estimate (5.11) for B_1. Using (8.2) and (8.8) we write

$$B_1(\rho, \varphi, \omega) = -4\mathcal{K}_1(\rho, \omega, \varepsilon_+) + 4\mathcal{K}_1(\rho, \omega, \varepsilon_-) + \int_{-1}^{1} \check{\mathcal{Z}}(\beta, \varphi) \cdot \cos\beta \; e^{i\omega\rho\cosh\beta} \, d\beta.$$

Hence, B_1 satisfies (5.7) (and, therefore, (5.11)) by Lemma 8.2 (i) and (8.3).

(III) Let us prove (5.6) for $\partial_\varphi \widehat{u}_d$. By (5.2) it suffices to prove this estimate for $\partial_\varphi A$, where A is given by (5.8). From (9.3) we have

$$\begin{aligned} \partial_\varphi A(\rho, \varphi, \omega) &= -\omega\rho A_3(\rho, \varphi, \omega), \\ A_3(\rho, \varphi, \omega) &= \int_{\mathbb{R}} \mathcal{Z}(\beta, \varphi) \sinh\beta \; e^{i\omega\rho\cosh\beta} \, d\beta, \quad \varphi \neq \varphi_\pm. \end{aligned} \tag{5.12}$$

Similarly to the proof of estimate (5.11) for B, we obtain the same estimate for A_3, so, by (5.12), the estimate (5.6) follows. Proposition 5.3 is proven. □

Now define the function

$$u_s^0(\rho, \varphi, t) = u(\rho, \varphi, t) - u_i^0(\rho, \varphi, t), \quad \varphi \neq \varphi_+, \quad t > 0, \tag{5.13}$$

where u_i^0 is given by (1.13). Then by (4.7),

$$u_s^0(\rho, \varphi, t) = u_r(\rho, \varphi, t) + u_d(\rho, \varphi, t), \quad \varphi \neq \varphi_\pm, \quad t > 0, \tag{5.14}$$

where u_r is given by (4.1) and u_d is given by (4.3).

Corollary 5.4 *Let $\widehat{u}_s^0(\rho, \varphi, \omega)$ be the Fourier–Laplace transform of the function $u_s^0(\rho, \varphi, t)$. Then the functions $\widehat{u}_s^0$, $\partial_\rho \widehat{u}_s^0$ and $\partial_\varphi \widehat{u}_s^0$ satisfy (5.6).*

Proof From (5.14) we have

$$\widehat{u}_s^0(\rho, \varphi, \omega) = \widehat{u}_r(\rho, \varphi, \omega) + \widehat{u}_d(\rho, \varphi, \omega), \quad \varphi \neq \varphi_\pm, \quad \omega \in \mathbb{C}^+, \tag{5.15}$$

where $\widehat{u}_r$ and $\widehat{u}_d$ are defined by (5.1) and (5.2), respectively. Hence the statement follows from Lemma 5.2 and Proposition 5.3. □

5.3 Estimates for $\widehat{u}_s(x,\omega)$

To estimate $\widehat{u}_s$ it is convenient to introduce one more "part" u_i^1 of the non-stationary incident wave u_i, namely the difference between u_i and u_i^0.

From (1.7) and (5.13) it follows that

$$u_s(\rho,\varphi,t) = u_s^0(\rho,\varphi,t) - u_i^1(\rho,\varphi,t), \quad \varphi \neq \varphi_\pm \tag{5.16}$$

where $u_i^1(\rho,\varphi,t) := u_i(\rho,\varphi,t) - u_i^0(\rho,\varphi,t)$. From (1.1) and (1.13) it follows that

$$u_i^1(\rho,\varphi,t) = \begin{cases} 0, & 0 < \varphi < \varphi_+, \\ -u_i(\rho,\varphi,t), & \varphi_+ < \varphi < 2\pi. \end{cases} \tag{5.17}$$

By (3.3),

$$\widehat{u}_i^1(\rho,\varphi,\omega) = \begin{cases} 0, & 0 < \varphi < \varphi_+, \\ -\widehat{f}(\omega-\omega_0)\, e^{i\omega\mathbf{n}\cdot x}, & \varphi_+ < \varphi < 2\pi. \end{cases} \tag{5.18}$$

Lemma 5.5 *There exist $C(\omega), c(\omega) > 0$ such that $\widehat{u}_i^1$, $\partial_\rho \widehat{u}_i^1$ satisfy (5.4) and $\partial_\varphi \widehat{u}_i^1$ satisfies (5.5) for $\varphi \in (0,2\pi)$, $\varphi \neq \varphi_\pm$.*

Proof By (3.3) it suffices to prove the statement for $e^{i\omega\mathbf{n}\cdot x}$ when $\varphi \in (\varphi_+, 2\pi)$. Since $|e^{i\omega\mathbf{n}\cdot x}| = e^{\omega_2\rho\cos(\varphi-\alpha)}$, $\varphi \in (\varphi_+, 2\pi)$ we have

$$\begin{aligned} \partial_\rho e^{\omega_2\rho\cos(\varphi-\alpha)} &= \omega_2\cos(\varphi-\alpha)e^{\omega_2\rho\cos(\varphi-\alpha)}, \\ \partial_\varphi e^{\omega_2\rho\cos(\varphi-\alpha)} &= -\omega_2\rho\sin(\varphi-\alpha)e^{\omega_2\rho\cos(\varphi-\alpha)}, \end{aligned} \tag{5.19}$$

and for $\varphi \in (\varphi_+, 2\pi)$, we have $|e^{\omega_2\rho\cos(\varphi-\alpha)}| \le e^{-c\omega_2\rho}$, $c > 0$, $\varphi \in (\varphi_+, 2\pi)$, because $\cos(\varphi-\alpha) \le -c < 0$ by (1.4). Hence the statement follows from (5.19). □

Corollary 5.6 *The functions $\widehat{u}_s$, $\partial_\rho \widehat{u}_s$ and $\partial_\varphi \widehat{u}_s$ satisfy (5.6) for $\varphi \in (0,2\pi)$, $\varphi \neq \varphi_\pm$.*

Proof From (5.16) it follows that

$$\widehat{u}_s(\rho,\varphi,\omega) = \widehat{u}_s^0(\rho,\varphi,\omega) - \widehat{u}_i^1(\rho,\varphi,\omega). \tag{5.20}$$

Thus the statement follows from Corollary 5.4 and Lemma 5.5. □

It is possible to get rid of the restriction $\varphi \neq \varphi_\pm$ in Corollary 5.6.

Let $l_\pm = \{(\rho,\varphi) : \rho > 0,\ \varphi = \varphi_\pm\}$.

Proposition 5.7 *The functions* $\widehat{u}_s(\cdot,\cdot,\omega)$, $\partial_\rho \widehat{u}_s(\cdot,\cdot,\omega)$ *and* $\partial_\varphi \widehat{u}_s(\cdot,\cdot,\omega)$ *belong to* $C^2(Q)$, *and satisfy* (5.6) *in* Q *(including* $l_+ \cup l_-$*), and*

$$(\Delta + \omega^2)\widehat{u}_s(\rho,\varphi,\omega) = 0, \quad (\rho,\varphi) \in Q, \quad \omega \in \mathbb{C}^+. \tag{5.21}$$

Proof The function $\widehat{u}_s(\rho,\varphi,\omega)$ satisfies (5.21) in $Q \setminus (l_+ \cup l_-)$. This follows directly from the explicit formulas (5.20). In fact, (5.20) and (5.15) imply

$$\widehat{u}_s = \widehat{u}_r + \widehat{u}_d - \widehat{u}_i^1. \tag{5.22}$$

The function $\widehat{u}_r$ satisfies (5.21) for $\varphi \neq \varphi_\pm$, $\widehat{u}_i^1$ satisfies (5.21) for $\varphi \neq \varphi_\pm$ by (5.17) and (3.3) and $\widehat{u}_d$ satisfies (5.21) for $\varphi \neq \varphi_\pm$ by (5.2), see Appendix 2. It remains only to prove that $\widehat{u}_s \in C^2(Q)$, because this will mean that (5.6) holds by Corollary 5.6 (and continuity) and (5.21) holds in Q including $l_\pm$.

Let us prove this for φ close to φ_-. The case of φ close to φ_+ is analyzed similarly.

Let $h(s)$ be defined in $(\mathbb{C}\setminus\mathbb{R}) \cap B(s^*)$, where $B(s^*)$ is a neighborhood of $s^* \in \mathbb{R}$. Define the jump of h at the point s^* as

$$\mathcal{J}(h,s^*) := \lim_{\varepsilon\to 0+} h(s^* + i\varepsilon) - \lim_{\varepsilon\to 0+} h(s^* - i\varepsilon).$$

We have $\mathcal{J}\left(\widehat{u}_r(\rho,\varphi,\omega),\varphi_-\right) = \widehat{f}(\omega-\omega_0)e^{-i\omega\rho}$ by (5.1).

Similarly,

$$\mathcal{J}\left(\partial_\varphi \widehat{u}_r(\rho,\varphi,\omega),\varphi_-\right) = 0, \quad \mathcal{J}\left(\partial_{\varphi\varphi} \widehat{u}_r(\rho,\varphi,\omega),\varphi_-\right) = -\widehat{f}(\omega-\omega_0)(i\omega\rho)e^{i\omega\rho}.$$

From (5.2), (5.10), (8.2), and (8.3) we have

$$\begin{aligned}\mathcal{J}\left(\widehat{u}_d(\rho,\varphi,\omega),\varphi_-\right) &= \frac{i}{8\pi}\widehat{f}(\omega-\omega_0)\int_{-1}^{1}\frac{4}{\beta+i\varepsilon}e^{i\omega\rho\cosh\beta}\,d\beta\Big|_{\varepsilon_-=-0}^{\varepsilon_-=+0}\\ &= -\mathcal{J}\left(\widehat{u}_r(\rho,\varphi,\omega),\varphi_-\right).\end{aligned} \tag{5.23}$$

Further, by (8.4),

$$\mathcal{J}\left(\partial_\varphi \widehat{u}_d(\rho,\varphi,\omega),\varphi_-\right) = 0 = -\mathcal{J}\left(\partial_\varphi \widehat{u}_r(\rho,\varphi,\omega),\varphi_-\right).$$

Finally, consider

$$M := \mathcal{J}\left(\partial_{\varphi\varphi} \widehat{u}_d(\rho,\varphi,\omega),\varphi_-\right).$$

Similarly to (5.23), expanding $e^{i\omega\rho\cos\beta}$ in the Taylor series in β (at 0) and noting that all the terms $\int \frac{\beta^k\, d\beta}{(\beta+i\varepsilon_-)^3}$, $k \neq 2$, are continuous, we obtain

$$M = -\frac{i}{\pi}\widehat{f}(\omega-\omega_0)\int_{-1}^{1}\frac{e^{i\omega\rho\cos\beta}}{(\beta+i\varepsilon_-)^3}d\beta\Big|_{\varepsilon_-=+0}^{\varepsilon_-=-0}$$

$$= \frac{-i\,\widehat{f}(\omega-\omega_0)(i\omega\rho)e^{i\omega\rho}}{2\pi}\int_{-1}^{1}\frac{\beta^2}{(\beta+i\varepsilon_-)^3}d\beta\Big|_{\varepsilon_-=+0}^{\varepsilon_-=-0}.$$

Hence,

$$M = \widehat{f}(\omega-\omega_0)(i\omega\rho)e^{i\omega\rho} = -\mathcal{J}\,(\widehat{u}_r(\rho,\varphi,\omega),\varphi_-)\,.$$

Since $\widehat{u}_i^i(\rho,\varphi,\omega)$ is smooth on l_- by (5.18), we obtain from (5.22) that $\widehat{u}_s \in C^2(l_-)$.

Similarly using (5.1), (5.17) and (1.1) we obtain $\widehat{u}_s \in C^2(l_+)$. So $\widehat{u}_s \in C^2(Q)$. Proposition 5.7 is proven. □

Corollary 5.8

(i) *The function $\widehat{u}_s(\cdot,\cdot,\omega)$ belongs to the space $H^1(Q)$ for any $\omega \in \mathbb{C}^+$.*
(ii) *The function $u_s(x,t) \in \mathcal{M}$.*

Proof

(i) Everywhere below $x = (\rho,\varphi) \in Q \setminus (l_1 \cup l_2)$. It suffices to prove that

$$u_s(\cdot,\cdot,\omega),\ \partial_{x_k}u_s(\cdot,\cdot,\omega) \in L_2(Q), \quad k=1,2, \quad \omega\in\mathbb{C}^+. \tag{5.24}$$

First, by Proposition 5.7, $\widehat{u}_s(x,\omega)$ satisfies (5.6). Hence, $\widehat{u}_s(\cdot,\omega) \in L_2(Q)$ for any $\omega \in \mathbb{C}^+$. Further, using (1.12), we have

$$|\partial_{x_1}u_s(\cdot,\cdot,\omega)|^2 \le |\cos\varphi|^2|\partial_\rho u_s(\cdot,\cdot,\omega)|^2 + \frac{|\sin\varphi|^2}{\rho^2}|\partial_\varphi u_s(\cdot,\cdot,\omega)|^2.$$

Hence, by Proposition 5.7,

$$|\partial_{x_1}u_s(\cdot,\cdot,\omega)|^2 \le C(\omega)e^{-2c(\omega)}\left(1+\frac{1}{\rho}\right).$$

This implies that $\partial_{x_1}u_s \in L_2(Q)$, since $c(\omega) > 0$. Similarly, $\partial_{x_2}u_s(\cdot,\cdot,\omega) \in L_2(Q)$. (5.24) is proven.

(ii) The statement follows from Definition 3.1. □

6 Uniqueness

In Sect. 5 we proved the existence of solution to (1.8)–(1.10) belonging to $\mathcal{M}$. In this section prove the uniqueness of this solution in the same space.

Recall that we understand the uniqueness of the time-dependent Sommerfeld problem (1.5)–(1.6) as the uniqueness of the solution u_s given by (1.7) of the mixed problem (1.8)–(1.10) in the space $\mathcal{M}$.

The following theorem is the main result of the paper.

Theorem 6.1

(i) *Problem* (1.8)–(1.10) *admits a solution belonging to the space* $\mathcal{M}$. *Its limiting amplitude exists and is the solution of problem* (4.9). *The connection between this limiting amplitude and the Sommerfeld solution is given by* (4.8).
(ii) *Problem* (1.8)–(1.10) *admits a unique solution in the space* $\mathcal{M}$.

Proof The statements contained in item (i) follow from Corollary 5.8, Corollary 4.3, and Remark 4.4.

(ii) Let us prove the uniqueness. We follow closely the proof of Theorem 2.1 from [5]. Suppose that there exist two solutions $u_s(x,t)$ and $v_s(x,t)$ of system (1.8)–(1.10) belonging to $\mathcal{M}$. Consider $w_s(x,t) := u_s(x,t) - v_s(x,t)$.

Then $\widehat{w}_s(\cdot,\cdot,\omega) = \widehat{u}_s(\cdot,\cdot,\omega) - \widehat{v}_s(\cdot,\cdot,\omega)$, where $\widehat{u}_s$, $\widehat{v}_s$ (and, therefore, $\widehat{w}_s$) satisfy all the conditions of Proposition 5.7 and $\widehat{w}_s|_{W^0} = 0$ by (3.4).

Let us prove that $\widehat{w}_s(\cdot,\cdot,\omega) \equiv 0$. Let R be a sufficiently large positive number and $B(R)$ be the open disk centered at the origin with radius R. Set $Q_R := Q \cap B(R)$. Note that Q_R has a piecewise smooth boundary S_R and denote by $n(x)$ the outward unit normal vector at the non-singular points $x \in S_R$.

The first Green identity for $w_s(\rho,\varphi,\cdot)$ and its complex conjugate $\overline{w}_s$ in the domain Q_R, together with zero boundary conditions on S_R, yield

$$\int\limits_{Q_R} \Big[|\nabla \widehat{w}_s|^2 - \omega^2 |\widehat{w}_s|^2 \Big]\, dx = \int\limits_{\partial B(R)\cap Q} \Big(\partial_n \widehat{w}_s\Big) \cdot \Big(\overline{w}_s\Big)\, dS_R.$$

From the real and imaginary parts of the last identity, we obtain

$$\int\limits_{Q_R} \Big[|\nabla \widehat{w}_s|^2 + \big(\operatorname{Im}\omega\big)^2 |\widehat{w}_s|^2 \Big]\, dx = \operatorname{Re} \int\limits_{\partial B(R)\cap Q} \big(\partial_n \widehat{w}_s\big)\big(\overline{\widehat{w}_s}\big)\, dS_R \tag{6.1}$$

for $\operatorname{Re}\omega = 0$ and

$$-2\big(\operatorname{Re}\omega\big)\big(\operatorname{Im}\omega\big) \int\limits_{Q_R} |\widehat{w}_s|^2\, dx = \operatorname{Im} \int\limits_{\partial B(R)\cap Q} \big(\partial_n w_s\big)\big(\overline{w}_s\big)\, dS_R \tag{6.2}$$

for Re $\omega \neq 0$. Recall that we consider the case Im $k \neq 0$. Now, note that since $\widehat{\omega}_s \in H^1(Q)$, there exist a monotonic sequence of positive numbers $\{R_j\}$ such that $R_j \to \infty$ as $j \to \infty$ and

$$\lim_{j\to\infty} \int\limits_{\partial B(R_j)\cap Q} \left[\partial_n \widehat{w}_s\right]\left[\overline{\widehat{w}_s}\right] dS_{R_j} = 0. \tag{6.3}$$

Indeed, in polar coordinates (ρ, φ), we have that the integrals

$$\int\limits_0^\infty \left(R \int\limits_0^{2\pi} |\widehat{w}_s(\rho,\varphi)|^2 d\varphi \right) dR \quad \text{and} \quad \int\limits_0^\infty \left(R \int\limits_0^{2\pi} |\partial_n \widehat{w}_s(\rho,\varphi)|^2 d\varphi \right) dR$$

are finite. This fact, in particular, implies that there exist a monotonic sequence of positive numbers R_j such that $R_j \to \infty$ as $j \to \infty$ and

$$\int\limits_0^{2\pi} |\widehat{w}_s(R_j,\varphi)|^2 d\varphi = o(R_j^{-1}), \quad \int\limits_0^{2\pi} |\partial_n \widehat{w}_s(R_j,\varphi)|^2 d\varphi = o(R_j^{-1}) \text{ as } j \to \infty.$$

Further, applying the Cauchy-Schwarz inequality for every R_j, we get

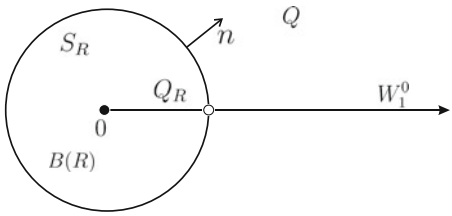
$$\begin{aligned}
\left| \int\limits_0^{2\pi} \partial_n \widehat{w}_s(R_i,\varphi)\overline{\widehat{w}_s}(R_i,\varphi) d\varphi \right| &\leq \int\limits_0^{2\pi} |\partial_n \widehat{w}_s(R_i,\varphi)\widehat{w}_s(R_i,\varphi)| d\varphi \\
&\leq \left(\int\limits_0^{2\pi} |\partial_n \widehat{w}_s(R_i,\varphi)|^2 d\varphi \right)^{1/2} \left(\int\limits_0^{2\pi} |\widehat{w}_s(R_i,\varphi)|^2 d\varphi \right)^{1/2} \\
&= o(R_j^{-1}) \quad \text{as} \quad j \to \infty,
\end{aligned}$$

and therefore we obtain (6.3).

Since the expressions under the integral sign in the left hand sides of equalities (6.1) and (6.2) are non-negative, we have that these integrals are monotonic with respect to R. This observation together with (6.3) implies

$$\int\limits_Q \left[|\nabla \widehat{w}_s|^2 + \left(\operatorname{Im}\omega\right)^2 |\widehat{w}_s|^2 \right] d\varphi = \lim_{R\to\infty} \int\limits_{Q_R} \left[|\nabla \widehat{w}_s|^2 + \left(\operatorname{Im}\omega\right)^2 |\widehat{w}_s|^2 \right] d\varphi = 0$$

Fig. 2 Uniqueness



for Re $\omega = 0$ and

$$\int_Q |\widehat{w}_s|^2 \, d\varphi = \lim_{R\to\infty} \int_{Q_R} |\widehat{w}_s|^2 \, d\varphi = 0$$

for Re $\omega \neq 0$. Thus, it follows from the last two identities that $\widehat{w}_s = 0$ in Q (Fig. 2). □

7 Conclusion

We proved that the Sommerfeld solution to the half-plane diffraction problem for a wide class of incident waves is the limiting amplitude of the solution of the corresponding time-dependent problem in a functional class of generalized solutions. The solution of the time-dependent problem is shown to be unique in this class. It is also shown that the limiting amplitude automatically satisfies the Sommerfeld radiation condition and the regularity condition at the edge.

Acknowledgments The authors are grateful to CONACYT (México) and CIC (UMSNH) for partial financial support. We are also grateful to anonymous referees for valuable comments.

8 Appendix 1

Lemma 8.1

(i) *The functions* $\mathcal{Z}$ *(given by* (4.2)*) and* $\partial_\varphi \mathcal{Z}$ *admit uniform with respect to* $\varphi \in [0, 2\pi]$ *estimates*

$$|\mathcal{Z}(\beta, \varphi)| \leq Ce^{-|\beta|/2}, \quad |\partial_\varphi \mathcal{Z}(\beta, \varphi)| \leq Ce^{-|\beta|/2}, \quad |\beta| \geq 1. \tag{8.1}$$

(ii) *The function* $\mathcal{Z}$ *admits the representation*

$$\mathcal{Z}(\beta, \varphi) = -\frac{4}{\beta + i\varepsilon_+} + \frac{4}{\beta + i\varepsilon_-} + \breve{\mathcal{Z}}(\beta, \varphi), \quad \varepsilon_\pm \neq 0 \tag{8.2}$$

with

$$\check{\mathcal{Z}}(\beta,\varphi)\in C^\infty(\mathbb{R}\times[0,2\pi]),\quad |\check{\mathcal{Z}}(\beta,\varphi)|\le C,\quad \beta\in\mathbb{R}\times[0,2\pi]. \tag{8.3}$$

(iii) *The function* $\partial_\varphi\mathcal{Z}$ *admits the representation*

$$\partial_\varphi\mathcal{Z}=-\frac{4i}{(\beta+i\varepsilon_+)^2}+\frac{4i}{(\beta+i\varepsilon_-)^2}+\check{Z}_1(\beta,\varphi),\quad \varepsilon_\pm\ne 0, \tag{8.4}$$

with

$$\check{Z}_1(\beta,\varphi)\in C^\infty(\mathbb{R}\times[0,2\pi]),\quad |\check{Z}_1(\beta,\varphi)|\le C,\quad \beta\in\mathbb{R}\times[0,2\pi]. \tag{8.5}$$

Proof

(i) For $a=im$, $b=in$, we have

$$\coth a-\coth b=\frac{-\sinh(\alpha/2)}{\sinh(b)\sinh(a)}.$$

Hence for $m=-\pi/8+a/4$ and $n=-\pi/8-a/4$ we obtain the estimate (8.1) for $U(\zeta)$ given by (4.6) with respect to ζ. So (8.1) for $\mathcal{Z}$ follows from (4.5) and (4.2).

(ii) From (4.5) and (4.6) it follows that the function $\mathcal{Z}$ admits the representation

$$\mathcal{Z}(\beta,\varphi)=Z_+(\beta,\varphi)+Z_-(\beta,\varphi)+Z^+(\beta,\varphi)+Z^-(\beta,\varphi),$$

where

$$\begin{aligned} Z_\pm(\beta,\varphi)&=\pm\coth\left(\frac{\beta+i(\varphi_\pm-\varphi)}{4}\right),\\ Z^\pm(\beta,\varphi)&=\pm\coth\left(\frac{\beta-i(\varphi_\pm+\varphi)}{4}\right).\end{aligned} \tag{8.6}$$

Further, since $|\coth z-1/z|\le C$, $|\operatorname{Im} z|\le\pi$, $z\ne 0$, we have

$$Z_\pm(\beta,\varphi)=\pm\frac{4}{\beta+i\varepsilon_\pm}+\check{Z}_\pm(\beta,\varphi),\quad \varphi\ne\varphi_\pm,$$

and

$$\check{Z}_\pm(\beta,\varphi)\in C^\infty(\mathbb{R}\times[0,2\pi]),\quad |\check{Z}_\pm(\beta,\varphi)|\le C,\quad (\beta,\varphi)\in\mathbb{R}\times[0,2\pi].$$

Finally, by (1.4),

$$Z^{\pm}(\beta,\varphi)\in C^{\infty}(\mathbb{R}\times[0,2\pi]),\quad |Z^{\pm}(\beta,\varphi)|\le C,\quad (\beta,\varphi)\in\mathbb{R}\times[0,2\pi].$$

Therefore, (8.2) and (8.3) are proven.

(iii) From (8.2) and (5.10) we get (8.4). Finally, by (8.6),

$$\partial_{\varphi}Z^{\pm}(\beta,\varphi)\in C^{\infty}(\mathbb{R}\times[0,2\pi]),\quad |\partial_{\varphi}Z^{\pm}(\beta,\varphi)|\le C,\quad (\beta,\varphi)\in\mathbb{R}\times[0,2\pi].$$

Moreover, since

$$\partial_{\varphi}Z_{\pm}(\beta,\varphi)\pm[4i/(\beta+\varepsilon_{\pm})^2]\in C^{\infty}([\mathbb{R}\times[0,2\pi]),$$

and is bounded in the same region, (8.5) holds.

□

For $\varepsilon,\beta\in\mathbb{R}$, $\varepsilon\neq 0$, $\rho>0$, $\omega\in\mathbb{C}^{+}$, let

$$K_0(\beta,\rho,\omega,\varepsilon):=\frac{e^{i\omega\rho\cosh\beta}}{\beta+i\varepsilon},\quad \mathcal{K}_0(\rho,\omega,\varepsilon):=\int_{-1}^{1}K(\beta,\rho,\omega,\varepsilon)\,d\beta, \tag{8.7}$$

$$\begin{aligned}&K_1(\beta,\rho,\omega,\varepsilon):=\cosh\beta\cdot e^{i\omega\rho\cosh\beta},\quad \mathcal{K}_1(\rho,\omega,\varepsilon):=\int_{-1}^{1}K_1(\beta,\rho,\omega,\varepsilon)\,d\beta,\\ &K_2(\beta,\rho,\varphi,\varepsilon):=\frac{e^{i\omega\rho\cosh\beta}}{(\beta+i\varepsilon)^2},\quad \mathcal{K}_2(\rho,\omega,\varepsilon):=\int_{-1}^{1}K_2(\beta,\rho,\omega,\varepsilon)\,d\beta\,d\beta.\end{aligned} \tag{8.8}$$

Lemma 8.2 *There exist $C(\omega)>0$, $c(\omega)>0$ such that the functions $\mathcal{K}_0$, $\mathcal{K}_1$, and $\mathcal{K}_2$ satisfy the estimates*

$$|\mathcal{K}_{0,1,2}(\rho,\omega,\varepsilon)|\le C(\omega)e^{-c(\omega)\rho},\quad \rho>0,\ \varphi\in(0,2\pi),\ \varepsilon\neq 0. \tag{8.9}$$

Proof It suffices to prove (8.9) for $0<\varepsilon<\varepsilon_0$, since the functions $\mathcal{K}_0,\mathcal{K}_1,\mathcal{K}_2$ are odd with respect to ε, and for $\varepsilon\ge\varepsilon_0>0$ they satisfy the estimate

$$\left|\mathcal{K}_{0,1,2}(\beta,\rho,\omega,\varepsilon)\right|\le C(\varepsilon_0)\int_{-1}^{1}e^{-\omega_2\rho}\,d\beta\le 2C(\varepsilon_0)e^{-\omega_2\rho}.$$

(I) Let us prove (8.9) for $\mathcal{K}_0$. Let

$$\cosh\beta := 1 + h(\beta), \quad \beta \in \mathbb{C}. \tag{8.10}$$

Define $\varepsilon_0 = \varepsilon_0(\omega)$ such that

$$|h(\beta)| < \frac{1}{4}, \quad |\omega_1||h(\beta)| \le \frac{\omega_2}{4} \quad \text{for} \quad |\beta| \le 2\varepsilon_0 := r, \tag{8.11}$$

and define the contour

$$\gamma_r := \{\beta = re^{i\theta}, \quad -\pi < \theta < 0\}. \tag{8.12}$$

Then we have by the Cauchy Theorem

$$\mathcal{K}_0(\rho,\omega,\varepsilon) = I_1(\rho,\omega,\varepsilon) + I_2(\rho,\omega,\varepsilon) - 2\pi i \operatorname{Res}_{\beta=-i\varepsilon} K_0(\beta,\rho,\omega,\varepsilon),$$

where

$$I_1(\rho,\omega,\varepsilon) = \int\limits_{\gamma_r} K_0(\beta,\rho,\omega,\varepsilon)\, d\beta, \quad I_2(\rho,\omega,\varepsilon) = \left(\int\limits_{-1}^{-r} + \int\limits_{r}^{1}\right) K_0(\beta,\rho,\omega,\varepsilon)\, d\beta$$

and $0 < \varepsilon < \varepsilon_0$. First,

$$|\operatorname{Res}_{\beta=-i\varepsilon} K_0(\beta,\rho,\omega,\varepsilon)| = e^{-\omega_2\rho\cos\varepsilon} \le e^{-\frac{1}{2}\omega_2\rho}, \quad 0 < \varepsilon < \varepsilon_0, \tag{8.13}$$

by (8.11). Further, from (8.10) we have

$$\begin{aligned} \left|I_1(\rho,\omega,\varepsilon)\right| &\le \int\limits_{\gamma_r} \frac{\left|e^{-\omega_2\rho\left(1+h(\beta)\right)} e^{i\omega_1\rho\left(1+h(\beta)\right)}\right|}{|\beta + i\varepsilon|} |d\beta| \\ &\le \frac{1}{\varepsilon_0}\, e^{-\omega_2\rho} \int\limits_{\gamma_r} \left|e^{-\omega_2\rho\, h(\beta) + i\omega_1\rho\, h(\beta)}\right| \left|d\beta\right|, \end{aligned} \tag{8.14}$$

since for $\beta \in \gamma_r$ we have $|\beta + i\varepsilon| \ge |\beta| - \varepsilon = 2\varepsilon_0 - \varepsilon > \varepsilon_0$, see Fig. 3.

Let $h(\beta) := h_1(\beta) + ih_2(\beta)$. Then

$$|I_1(\rho,\omega,\varepsilon)| \le \frac{1}{\varepsilon_0}\, e^{-\omega_2\rho} \int\limits_{\gamma_r} e^{\omega_2\rho\,|h_1(\beta)|}\, e^{|\omega_1|\rho\,|h_2(\beta)|}\, d\beta \le 2\pi e^{-\omega_2\rho/2}, \tag{8.15}$$

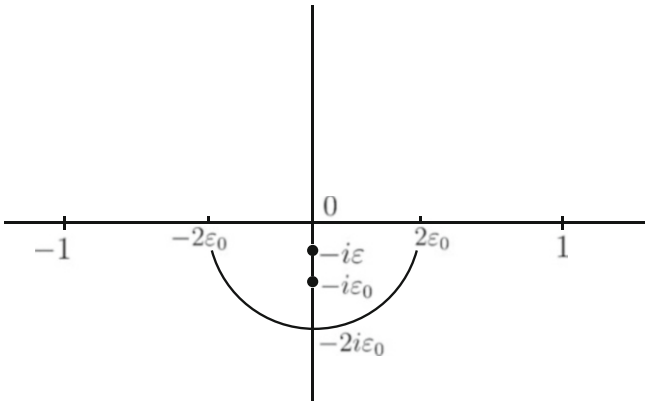


Fig. 3 Contour γ_r

by (8.11). Finally,

$$|I_2(\rho,\omega,\varepsilon)| \le \int\limits_{[-1,-r]\cup[r,1]} \left| \frac{e^{-\omega_2\rho\cosh\beta + i\omega_1\rho\cosh\beta}}{\beta+i\varepsilon} \right| d\beta \le \frac{e^{-\omega_2\rho}}{2\varepsilon_0(\omega)}, \tag{8.16}$$

since $|\beta + i\varepsilon| \ge 2\varepsilon_0,\ \beta \in [-1,-r]\cup[r,1]$. From (8.14)–(8.16), we obtain (8.9) for $\mathcal{K}_0$.

(II) Let us prove (8.9) for $\mathcal{K}_1$. Let $h(\beta), \varepsilon_0(\omega), \gamma_r$ be defined by (8.10)–(8.12). Then we have by the Cauchy Theorem

$$\mathcal{K}_1(\rho,\omega,\varepsilon) := \int\limits_{\gamma_r\cup[-1,r]\cup[r,1]} K_1(\beta,\rho,\omega,\varepsilon)\,d\beta - 2\pi i\ \mathrm{Res}_{\beta=-i\varepsilon}\ K_1(\beta,\rho,\omega,\varepsilon), \quad 0<\varepsilon<\varepsilon_0. \tag{8.17}$$

First, similarly to (8.13), we obtain

$$\left|\mathrm{Res}_{\beta=-i\varepsilon}\ K_1(\beta,\rho,\omega,\varepsilon)\right| \le |\omega| e^{-\frac{\omega_2\rho}{2}},$$

by (8.11). Further, by (8.11) similarly to the proof of (8.14), (8.15), and using (8.10), we get

$$\left|\int\limits_{\gamma_r} K_1(\beta,\rho,\omega,\varepsilon)\,d\beta\right| \le \frac{|\omega|}{\varepsilon_0}\cdot\frac{5}{4}\,e^{-\omega_2\rho}\int\limits_{\gamma_r} |e^{-\omega_2\rho\,h(\beta)}\ e^{i\omega_1\rho\,h(\beta)}|\ |d\beta| \le C(\omega)e^{-\frac{\omega_2\rho}{2}}. \tag{8.18}$$

Finally, similarly to the proof of (8.16) we get the estimate

$$\left| \int\limits_{[-1,-r]\cup[r,1]} K_1(\beta,\rho,\omega,\varepsilon)\, d\beta \right| \le C(\omega)\, e^{-\omega_2 \rho}. \tag{8.19}$$

From (8.17)–(8.19), we obtain (8.9) for $\mathcal{K}_1$.

(III) Estimate (8.9) for $\mathcal{K}_2$ is proved similarly to the same estimate for $\mathcal{K}_0$ and $\mathcal{K}_1$ with obvious changes. Lemma 8.2 is proven. □

9 Appendix 2

Lemma 9.1 *We have*

$$\left(\Delta + \omega^2\right) u_d(\rho,\varphi,\omega) = 0, \quad \varphi \ne \varphi_\pm, \quad \omega \in \mathbb{C}^+. \tag{9.1}$$

Proof By (5.2) it suffices to prove (9.1) for

$$A_d(\rho,\varphi,\omega) := \int\limits_{\mathbb{R}} \mathcal{Z}(\beta,\varphi) e^{i\omega\rho\cosh\beta}\, d\beta. \tag{9.2}$$

Since $\omega \in \mathbb{C}^+$ the integral (9.2) converges after differentiation with respect to ρ and φ. We have

$$\partial_\rho A_d(\rho,\varphi,\omega) = (i\omega) \int\limits_{\mathbb{R}} \mathcal{Z}(\beta,\varphi) \cosh\beta\, e^{i\omega\rho\cosh\beta}\, d\beta,$$

$$\partial_\rho^2 A_d(\rho,\varphi,\omega) = -\omega^2 \int\limits_{\mathbb{R}} \mathcal{Z}(\beta,\varphi) \cosh^2\beta\, e^{i\omega\rho\cosh\beta}\, d\beta.$$

Integrating by parts, we have by (4.2) and (8.1)

$$\begin{aligned} \partial_\varphi A_d(\rho,\varphi,\omega) &= \int\limits_{\mathbb{R}} \partial_\varphi \left(Z^0(\beta + 2\pi i - i\varphi) \right) e^{i\omega\rho\cosh\beta}\, d\beta \\ &= -\omega\rho \int\limits_{\mathbb{R}} \mathcal{Z}(\beta,\varphi) \sinh\beta\, e^{i\omega\rho\cosh\beta}\, d\beta, \quad \varphi \ne \varphi_\pm. \end{aligned} \tag{9.3}$$

Hence, similarly to (9.3)

$$\partial^2_{\varphi\varphi} A_d(\rho,\varphi,\omega) = -i\omega\rho \int\limits_{\mathbb{R}} \mathcal{Z}(\beta,\varphi)\Big[\cosh\beta + i\omega\rho\sinh^2\beta\Big] e^{i\omega\rho\cosh\beta}\, d\beta,$$

and

$$\begin{aligned}(\Delta+\omega^2)u_d(\rho,\varphi,\omega) =& \partial^2_\rho A_d(\rho,\varphi,\omega) + \frac{1}{\rho}\partial_\rho A_d(\rho,\varphi\omega) \\ &+ \frac{1}{\rho^2}\partial^2_\varphi A_d(\rho,\varphi,\omega) + \omega^2 A_d(\rho,\varphi,\omega) = 0. \qquad \square\end{aligned}$$

References

1. J. Bernard, Progresses on the diffraction by a wedge: transient solution for line source illumination, single face contribution to scattered field, anew consequence of reciprocity on the spectral function. Rev. Tech. Thomson **25**(4), 1209–1220 (1993)
2. J. Bernard, On the time domain scattering by a passive classical frequency dependent-shaped region in a lossy dispersive medium. Ann. Telecommun. **49**(11–12), 673–683 (1994)
3. J. Bernard, G. Pelosi, G. Manara, A. Freni, Time domains scattering by an impedance wedge for skew incidence, in *Proceeding of the Conference ICEAA* (1991), pp. 11–14
4. V. Borovikov, *Diffraction at Polygons and Polyhedrons* (Nauka, Moscow, 1996)
5. L.P. Castro, D. Kapanadze, Wave diffraction by wedges having arbitrary aperture angle. J. Math. Anal. Appl. **421**(2), 1295–1314 (2015)
6. J. De la Paz Méndez, A. Merzon, DN-Scattering of a plane wave by wedges. Math. Methods Appl. Sci. **34**(15), 1843–1872 (2011)
7. J. De la Paz Méndez, A. Merzon, Scattering of a plane wave by "hard-soft" wedges, in *Recent Progress in Operator Theory and Its Applications*. Operator Theory: Advances and Applications, vol. 220 (2012), pp. 207–227
8. A. Dos Santos, F. Teixeira, The Sommerfeld problem revisited: solution spaces and the edges conditions. Math. Anal. Appl. **143**, 341–357 (1989)
9. T. Ehrhardt, A. Nolasco, F.-O. Speck, A Riemann surface approach for diffraction from rational wedges. Oper. Matrices **8**(2), 301–355 (2014)
10. A. Esquivel Navarrete, A. Merzon, An explicit formula for the nonstationary diffracted wave scattered on a NN-wedge. Acta Appl. Math. **136**(1), 119–145 (2015)
11. P. Grisvard, *Elliptic Problems in Nonsmooth Domains* (Pitman Publishing, London, 1985)
12. A. Heins, The Sommerfeld half-plane problem revisited I: the solution of a pair of complex Wiener–Hopf integral equations. Math. Methods Appl. Sci. **4**, 74–90 (1982)
13. I. Kay, The diffraction of an arbitrary pulse by a wedge. Commun. Pure Appl. Math. **6**, 521–546 (1953)
14. J. Keller, A. Blank, Diffraction and reflection of pulses by wedges and corners. Commun. Pure Appl. Math. **4**(1), 75–95 (1951)
15. A. Komech, Elliptic boundary value problems on manifolds with piecewise smooth boundary. Math. USSR-Sb. **21**(1), 91–135 (1973)
16. A. Komech, Elliptic differential equations with constant coefficients in a cone. Mosc. Univ. Math. Bull. **29**(2), 140–145 (1974)

17. A. Komech, N. Mauser, A. Merzon, On Sommerfeld representation and uniqueness in scattering by wedges. Math. Methods Appl. Sci. **28**(2), 147–183 (2005)
18. A. Komech, A. Merzon, Limiting amplitude principle in the scattering by wedges. Math. Methods Appl. Sci. **29**, 1147–1185 (2006)
19. A. Komech, A. Merzon, J. De la Paz Mendez, Time-dependent scattering of generalized plane waves by wedges. Math. Methods Appl. Sci. **38**, 4774–4785 (2015)
20. A. Komech, A. Merzon, A. Esquivel Navarrete, J. De La Paz Méndez, T. Villalba Vega, Sommerfeld's solution as the limiting amplitude and asymptotics for narrow wedges. Math. Methods Appl. Sci. **42**(15), 4957–4970 (2019)
21. A. Komech, A. Merzon, P. Zhevandrov, A method of complex characteristics for elliptic problems in angles and its applications. Am. Math. Soc. Transl. **206**(2), 125–159 (2002)
22. A. Merzon, A. Komech, J. De la Paz Méndez, T. Villalba Vega, On the Keller–Blank solution to the scattering problem of pulses by wedges. Math. Methods Appl. Sci. **38**, 2035–2040 (2015)
23. R. Nagem, M. Zampolli, G. Sandri, A. Sommerfeld (eds.), in *Mathematical Theory of Diffraction*. Progress in Mathematical Physics, vol. 35 (Birkhäuser, Boston, 2004)
24. F. Oberhettinger, On the diffraction and reflection of waves and pulses by wedges and corners. J. Res. Natl. Bur. Stand. **61**(2), 343–365 (1958)
25. A. Peters, J. Stoker, A uniqueness and a new solution for Sommerfeld's and other diffraction problems. Commun. Pure Appl. Math. **7**(3), 565–585 (1954)
26. K. Rottbrand, Time-dependent plane wave diffraction by a half-plane: explicit solution for Rawlins' mixed initial boundary value problem. Z. Angew. Math. Mech. **78**(5), 321–335 (1998)
27. K. Rottbrand, Exact solution for time-dependent diffraction of plane waves by semi-infinite soft/hard wedges and half-planes. Z. Angew. Math. Mech. **79**, 763–774 (1999)
28. V. Smirnov, S. Sobolev, Sur une méthode nouvelle dans le probléme plan des vibrations élastiques. Trudy Seismol. Inst. Acad. Nauk SSSR **20**, 1–37 (1932)
29. S. Sobolev, Theory of diffraction of plane waves. Trudy Seismol. Inst. Acad. Nauk SSSR **41**(1), 75–95 (1934)
30. S. Sobolev, General theory of diffraction of waves on Riemann surfaces. Tr. Fiz.-Mat. Inst. Steklova **9**, 39–105 (1935) [Russian]. English translation: S.L. Sobolev, General theory of diffraction of waves on Riemann surfaces, in *Selected Works of S.L. Sobolev*, vol. I (Springer, New York, 2006), pp. 201–262
31. S. Sobolev, Some questions in the theory of propagations of oscillations Chap. XII, in *Differential and Integral Equations of Mathematical Physics*, ed. by F. Frank, P. Mizes (Moscow, Leningrad, 1937), pp. 468–617 [Russian]
32. A. Sommerfeld, Mathematische theorie der diffraction. Math. Ann. **47**, 317–374 (1896)
33. A. Sommerfeld, *Optics (Lectures on Theoretical Physics)*, vol. 4 (Academic Press, New York, 1954)
34. F.-O. Speck, From Sommerfeld diffraction problems to operator factorisation. Constr. Math. Anal. **2**, 183–216 (2019)

On the Operator Jensen-Mercer Inequality

H. R. Moradi, S. Furuichi, and M. Sababheh

Abstract Mercer's inequality for convex functions is a variant of Jensen's inequality, with an operator version that is still valid without operator convexity. This paper is two folded. First, we present a Mercer-type inequality for operators without assuming convexity nor operator convexity. Yet, this form refines the known inequalities in the literature. Second, we present a log-convex version for operators. We then use these results to refine some inequalities related to quasi-arithmetic means of Mercer's type for operators.

Keywords Jensen-Mercer operator inequality · Log-convex functions · Operator quasi-arithmetic mean

Mathematics Subject Classification (2010) Primary 47A63; Secondary 47A64, 46L05, 47A60

The author (S.F.) was partially supported by JSPS KAKENHI Grant Number 16K05257.

H. R. Moradi
Young Researchers and Elite Club, Mashhad Branch, Islamic Azad University, Mashhad, Iran
e-mail: hrmoradi@mshdiau.ac.ir

S. Furuichi (✉)
Department of Information Science, College of Humanities and Sciences, Nihon University, Setagaya-ku, Tokyo, Japan
e-mail: furuichi@chs.nihon-u.ac.jp

M. Sababheh
Department of Basic Sciences, Princess Sumaya University for Technology, Amman, Jordan
e-mail: sababheh@psut.edu.jo

M. A. Bastos et al. (eds.), *Operator Theory, Functional Analysis and Applications*,
Operator Theory: Advances and Applications 282,
https://doi.org/10.1007/978-3-030-51945-2_24

1 Introduction

Recall that a function $f : I \subseteq \mathbb{R} \to \mathbb{R}$ is said to be convex on the interval I, if it satisfies the Jensen inequality

$$f\left(\sum_{i=1}^{n} w_i x_i\right) \leq \sum_{i=1}^{n} w_i f(x_i), \tag{1.1}$$

for all choices of positive scalars $w_1, \ldots, w_n$ with $\sum_{i=1}^{n} w_i = 1$ and $x_i \in I$. It is well known that this general form is equivalent to the same inequality when $n = 2$.

In 2003, Mercer found a variant of (1.1), which reads as follows.

Theorem 1.1 ([7, Theorem 1.2]) *If f is a convex function on $[m, M]$, then*

$$f\left(M + m - \sum_{i=1}^{n} w_i x_i\right) \leq f(M) + f(m) - \sum_{i=1}^{n} w_i f(x_i), \tag{1.2}$$

for all $x_i \in [m, M]$ and all $w_i \in [0, 1]$ $(i = 1, \ldots, n)$ with $\sum_{i=1}^{n} w_i = 1$.

There are many versions, variants and generalizations for the inequality (1.2); see for example [1, 2, 9].

It is customary in the field of Mathematical inequalities to extend scalar inequalities, like (1.1) and (1.2), to operators on Hilbert spaces. For this end, we adopt the following notations. Let $\mathscr{H}$ and $\mathscr{K}$ be Hilbert spaces, $\mathbb{B}(\mathscr{H})$ and $\mathbb{B}(\mathscr{K})$ be the C^*-algebras of all bounded operators on the appropriate Hilbert space. An operator $A \in \mathscr{H}$ is called self-adjoint if $A = A^*$, where A^* denotes the adjoint operator of A. If $A \in \mathscr{H}$, the notation $A \geq 0$ will be used to declare that A is positive, in the sense that $\langle Ax, x\rangle \geq 0$ for all $x \in \mathscr{H}$. If $\langle Ax, x\rangle > 0$ for all non zero $x \in \mathscr{H}$, we write $A > 0$, and we say than that A is positive definite. On the class of self-adjoint operators, the $\leq$ partial order relation is well known, where we write $A \leq B$ if $B - A \geq 0$, when A, B are self-adjoint.

In studying operator inequalities, the notion of spectrum cannot be avoided. If $A \in \mathscr{H}$, the spectrum of A is defined by

$$\sigma(A) = \{\lambda \in \mathbb{C} : A - \lambda \mathbf{1}_{\mathscr{H}} \text{ is not invertible}\},$$

where $\mathbf{1}_{\mathscr{H}}$ denotes the identity operator on $\mathscr{H}$. Finally, in these terminologies, a linear map $\Phi : \mathbb{B}(\mathscr{H}) \to \mathbb{B}(\mathscr{K})$ is said to be positive if $\Phi(A) \geq 0$ whenever $A \geq 0$ and Φ is called unital if $\Phi(\mathbf{1}_{\mathscr{H}}) = \mathbf{1}_{\mathscr{K}}$.

Recall that a continuous function $f : I \to \mathbb{R}$ is said to be operator convex if

$$f\left(\frac{A + B}{2}\right) \leq \frac{f(A) + f(B)}{2}$$

for all self-adjoint $A, B \in \mathbb{B}(\mathscr{H})$ and $\sigma(A), \sigma(B) \subset I$. This is equivalent to the Jensen operator inequality, valid for the self-adjoint operators A_i whose spectra are in the interval I,

$$f\left(\sum_{i=1}^{n} w_i A_i\right) \leq \sum_{i=1}^{n} w_i f(A_i), \quad w_i > 0, \quad \sum_{i=1}^{n} w_i = 1. \tag{1.3}$$

It is evident that a convex function is not necessarily operator convex, and the function $f(x) = x^4$ provides such an example. Thus, a convex function does not necessarily satisfy the operator Jensen inequality (1.3). However, it turns out that a convex function satisfies the following operator version of the Mercer inequality (1.2).

Theorem 1.2 ([5, Theorem 1]) *Let $A_1, \ldots, A_n \in \mathbb{B}(\mathscr{H})$ be self-adjoint operators with spectra in $[m, M]$ and let $\Phi_1, \ldots, \Phi_n : \mathbb{B}(\mathscr{H}) \to \mathbb{B}(\mathscr{K})$ be positive linear maps with $\sum_{i=1}^{n} \Phi_i(\mathbf{1}_{\mathscr{H}}) = \mathbf{1}_{\mathscr{K}}$. If $f : [m, M] \subseteq \mathbb{R} \to \mathbb{R}$ is a convex function, then*

$$f\left((M+m)\,\mathbf{1}_{\mathscr{K}} - \sum_{i=1}^{n} \Phi_i(A_i)\right) \leq (f(M) + f(m))\,\mathbf{1}_{\mathscr{K}} - \sum_{i=1}^{n} \Phi_i(f(A_i)). \tag{1.4}$$

Further, in the same reference, the following series of inequalities was proved

$$\begin{aligned} f\left((M+m)\mathbf{1}_{\mathscr{K}} - \sum_{i=1}^{n} \Phi_i(A_i)\right) &\leq (f(M) + f(m))\,\mathbf{1}_{\mathscr{K}} \\ &\quad + \frac{\sum_{i=1}^{n} \Phi_i(A_i) - M\mathbf{1}_{\mathscr{K}}}{M-m} f(m) + \frac{m\mathbf{1}_{\mathscr{K}} - \sum_{i=1}^{n} \Phi_i(A_i)}{M-m} f(M) \\ &\leq (f(M) + f(m))\mathbf{1}_{\mathscr{K}} - \sum_{i=1}^{n} \Phi_i(f(A_i)). \end{aligned}$$

Later, related and analogous results have been established in [3, 4, 6].

Our main goal of this article is to present a refinement of the operator inequality (1.4) without using convexity of f. Rather, using the idea by Mićić et al. [8], we assume a boundedness condition on f''. Then a discussion of log-convex version of Mercer's operator inequality will be presented.

2 Main Results

In this section we present our main results in two parts. In the first part, we discuss the twice differentiable case, then we discuss the log-convex case.

2.1 Twice Differentiable Functions

We begin with the non-convex version of Theorem 1.2. We use the following symbol in this paper.

(i) $\mathbf{A}^{sa} = (A_1, \ldots, A_n)$, where $A_i \in \mathbb{B}(\mathscr{H})$ are self-adjoint operators with $\sigma(A_i) \subseteq [m, M]$ for some scalars $0 < m < M$.
(ii) $\mathbf{\Phi}^{+} = (\Phi_1, \ldots, \Phi_n)$, where $\Phi_i : \mathbb{B}(\mathscr{H}) \to \mathbb{B}(\mathscr{K})$ are positive linear maps.

Theorem 2.1 *Let $A_1, \ldots, A_n \in \mathbb{B}(\mathscr{H})$ be self-adjoint operators with spectra in $[m, M]$ and let $\Phi_1, \ldots, \Phi_n : \mathbb{B}(\mathscr{H}) \to \mathbb{B}(\mathscr{K})$ be positive linear maps with $\sum_{i=1}^{n} \Phi_i(\mathbf{1}_{\mathscr{H}}) = \mathbf{1}_{\mathscr{K}}$. If $f : [m, M] \subseteq \mathbb{R} \to \mathbb{R}$ is a continuous twice differentiable function such that $\alpha \leq f'' \leq \beta$ with $\alpha, \beta \in \mathbb{R}$, then*

$$\begin{aligned}
&(f(M) + f(m))\,\mathbf{1}_{\mathscr{K}} - \sum_{i=1}^{n} \Phi_i(f(A_i)) - \beta J(m, M, \mathbf{A}^{sa}, \mathbf{\Phi}^{+}) \\
&\leq f\left((M+m)\mathbf{1}_{\mathscr{K}} - \sum_{i=1}^{n} \Phi_i(A_i)\right) \qquad (2.1) \\
&\leq (f(M) + f(m))\,\mathbf{1}_{\mathscr{K}} - \sum_{i=1}^{n} \Phi_i(f(A_i)) - \alpha J(m, M, \mathbf{A}^{sa}, \mathbf{\Phi}^{+}), \qquad (2.2)
\end{aligned}$$

where

$$\begin{aligned}
J(m, M, \mathbf{A}^{sa}, \mathbf{\Phi}^{+}) &:= (M+m) \sum_{i=1}^{n} \Phi_i(A_i) - Mm\mathbf{1}_{\mathscr{K}} \\
&\quad - \frac{1}{2}\left(\left(\sum_{i=1}^{n} \Phi_i(A_i)\right)^2 + \sum_{i=1}^{n} \Phi_i\left(A_i^2\right)\right) \geq 0.
\end{aligned}$$

Proof Notice that for any convex function f and $m \leq t \leq M$, we have

$$f(t) = f\left(\frac{M-t}{M-m}m + \frac{t-m}{M-m}M\right) \leq L_f(t), \qquad (2.3)$$

where

$$L_f(t) := \frac{M-t}{M-m} f(m) + \frac{t-m}{M-m} f(M). \tag{2.4}$$

Letting

$$g_\alpha(t) := f(t) - \frac{\alpha}{2} t^2 \quad (m \le t \le M),$$

we observe that g is convex noting the assumption $\alpha \le f''$. Applying (2.3) to the function g, we have $g(t) \le L_g(t)$, which leads to

$$f(t) \le L_f(t) - \frac{\alpha}{2} \left\{ (M+m)\, t - Mm - t^2 \right\}. \tag{2.5}$$

Since $m \le M + m - t \le M$, we can replace t in (2.5) with $M + m - t$, to get

$$f(M+m-t) \le L_0(t) - \frac{\alpha}{2} \left\{ (M+m)\, t - Mm - t^2 \right\},$$

where

$$L_0(t) := L(M+m-t) = f(M) + f(m) - L_f(t).$$

Using functional calculus for the operator

$$m\mathbf{1}_{\mathscr{K}} \le \sum_{i=1}^{n} \Phi_i(A_i) \le M\mathbf{1}_{\mathscr{K}},$$

we infer that

$$\begin{aligned} & f\left((M+m)\mathbf{1}_{\mathscr{K}} - \sum_{i=1}^{n} \Phi_i(A_i) \right) \\ & \le L_0\left(\sum_{i=1}^{n} \Phi_i(A_i) \right) \\ & \quad - \frac{\alpha}{2} \left\{ (M+m) \sum_{i=1}^{n} \Phi_i(A_i) - Mm\mathbf{1}_{\mathscr{K}} - \left(\sum_{i=1}^{n} \Phi_i(A_i) \right)^2 \right\}. \end{aligned} \tag{2.6}$$

On the other hand, by applying functional calculus for the operator

$$m\mathbf{1}_{\mathscr{H}} \le A_i \le M\mathbf{1}_{\mathscr{H}}$$

in (2.5), we get

$$f(A_i) \leq L_f(A_i) - \frac{\alpha}{2}\left\{(M+m)\,A_i - Mm\mathbf{1}_{\mathscr{H}} - A_i^2\right\}.$$

Applying the positive linear maps Φ_i and adding in the last inequality yield

$$\begin{aligned}
&\sum_{i=1}^{n} \Phi_i(f(A_i)) \\
&\quad \leq L_0\left(\sum_{i=1}^{n} \Phi_i(A_i)\right) \\
&\qquad - \frac{\alpha}{2}\left\{(M+m)\sum_{i=1}^{n} \Phi_i(A_i) - Mm\mathbf{1}_{\mathscr{K}} - \sum_{i=1}^{n} \Phi_i(A_i^2)\right\}. \qquad (2.7)
\end{aligned}$$

Combining the two inequalities (2.6) and (2.7), we get (2.2).

Finally we give the proof of $J(m, M, \Phi_i, A_i) \geq 0$. Since

$$m\mathbf{1}_{\mathscr{H}} \leq A_i \leq M\mathbf{1}_{\mathscr{H}},$$

we have

$$(M\mathbf{1}_{\mathscr{H}} - A_i)(A_i - m\mathbf{1}_{\mathscr{H}}) \geq 0$$

which implies

$$(M+m)A_i - mM\mathbf{1}_{\mathscr{H}} - A_i^2 \geq 0.$$

Thus we have

$$(M+m)\Phi(A_i) - mM\Phi(\mathbf{1}_{\mathscr{H}}) - \Phi(A_i^2) \geq 0.$$

Taking a summation on $i = 1, \cdots, n$ of this inequality with taking an account for $\sum_{i=1}^{n} \Phi_i(\mathbf{1}_{\mathscr{H}}) = \mathbf{1}_{\mathscr{K}}$, we obtain

$$(M+m)\sum_{i=1}^{n} \Phi_i(A_i) - Mm\mathbf{1}_{\mathscr{K}} - \sum_{i=1}^{n} \Phi(A_i^2) \geq 0. \qquad (2.8)$$

Further, noting that $m \leq \sum_{i=1}^{n} \Phi_i(A_i) \leq M$, we also have

$$(M+m)\sum_{i=1}^{n} \Phi_i(A_i) - mM\mathbf{1}_{\mathscr{K}} - \left(\sum_{i=1}^{n} \Phi_i(A_i)\right)^2 \geq 0. \qquad (2.9)$$

Adding (2.8) and (2.9) and dividing by 2, we obtain $J(m, M, \mathbf{A}^{sa}, \mathbf{\Phi}^{+}) \geq 0$.

The inequality (2.1) follows similarly by taking into account that

$$L_f(t) - \frac{\beta}{2}\left\{(M+m)t - Mm - t^2\right\} \leq f(t), \quad m \leq t \leq M.$$

The details are left to the reader. This completes the proof. □

In the following example, we present the advantage of using twice differentiable functions in Theorem 2.1.

Example 2.2 Let $f(t) = \sin t \ (0 \leq t \leq 2\pi)$,

$$A = \begin{pmatrix} \frac{\pi}{4} & 0 \\ 0 & \frac{\pi}{2} \end{pmatrix}$$

and $\Phi(A) = \frac{1}{2}Tr[A]$. Actually the function $f(t) = \sin t$ is concave on $[0, \pi]$. Letting $m = \frac{\pi}{4}$ and $M = \frac{\pi}{2}$, we obtain

$$0.9238 \approx f((M+m) - \Phi(A)) \not\leq f(M) + f(m) - \Phi(f(A)) \approx 0.8535.$$

That is, (1.4) may fail without the convexity assumption. However, by considering the weaker assumptions assumed in Theorem 2.1, we get

$$\begin{aligned} 0.9238 \approx & f((M+m) - \Phi(A)) \\ \lneqq & f(M) + f(m) - \Phi(f(A)) \\ & - \alpha\left\{(M+m)\Phi(A) - Mm - \frac{1}{2}\left\{\Phi(A)^2 + \Phi(A^2)\right\}\right\} \\ \approx & 0.9306, \end{aligned}$$

since $f''(t) = -\sin t$ which gives $\alpha = -1$.

To better understand the relation between Theorems 1.2 and 2.1, we present the following remark, where we clarify how the first theorem is retrieved from the second.

Remark 2.3 The inequality (2.2) in Theorem 2.1 with an assumption on a twice differentiable function f such that $\alpha \leq f'' \leq \beta$ for $\alpha, \beta \in \mathbb{R}$ gives a better upper bound of

$$f\left((M+m)\mathbf{1}_{\mathscr{K}} - \sum_{i=1}^{n} \Phi_i(A_i)\right)$$

than that in (1.4), since $J(m, M, \mathbf{A}^{sa}, \mathbf{\Phi}^{+}) \geq 0$, if we take $\alpha \geq 0$. Additionally to this result, we obtained a reverse type inequality (2.1) which gives a lower bound of

$$f\left((M+m)\,\mathbf{1}_{\mathscr{K}} - \sum_{i=1}^{n} \Phi_i\,(A_i)\right).$$

2.2 Log-Convex Functions

We conclude this section by presenting Mercer-type operator inequalities for log-convex functions. Recall that a positive function defined on an interval I (or, more generally, on a convex subset of some vector space) is called *log-convex* if $\log f\,(x)$ is a convex function of x. We observe that such functions satisfy the elementary inequality

$$f\,((1-v)\,a + vb) \leq [f\,(a)]^{1-v}\,[f\,(b)]^{v}, \quad 0 \leq v \leq 1$$

for any $a, b \in I$. f is called *log-concave* if the inequality above is reversed (that is, when $\frac{1}{f}$ is log-convex). By virtue of the arithmetic-geometric mean inequality, we have

$$f((1-v)a + vb) \leq [f(a)]^{1-v}\,[f(b)]^{v} \leq (1-v)f(a) + vf(b), \tag{2.10}$$

which implies convexity of log-convex functions. This double inequality is of special interest since (2.10) can be written as

$$f\,(t) \leq [f\,(m)]^{\frac{M-t}{M-m}}[f\,(M)]^{\frac{t-m}{M-m}} \leq L_f\,(t)\,, \qquad m \leq t \leq M \tag{2.11}$$

where $L_f\,(t)$ is as in (2.4).

Manipulating the inequality (2.11), we have the following extension of Theorem 1.2 to the context of log-convex functions. The proof is left to the reader.

Theorem 2.4 *Let all the assumptions of Theorem 1.2 hold except that* $f : [m, M] \to (0, \infty)$ *is log-convex. Then*

$$\begin{aligned} &f\left((M+m)\mathbf{1}_{\mathscr{K}} - \sum_{i=1}^{n} \Phi_i(A_i)\right) \\ &\quad \leq [f(m)]^{\frac{\sum_{i=1}^{n}\Phi_i(A_i) - m\mathbf{1}_{\mathscr{K}}}{M-m}}[f(M)]^{\frac{M\mathbf{1}_{\mathscr{K}} - \sum_{i=1}^{n}\Phi_i(A_i)}{M-m}} \\ &\quad \leq (f(M) + f(m))\,\mathbf{1}_{\mathscr{K}} - \sum_{i=1}^{n} \Phi_i(f(A_i)). \end{aligned}$$

3 Applications

In this section, we present some applications of the main results that we have shown so far. First, we review and introduce the notations.

(i) $\mathbf{A}^+ = (A_1, \ldots, A_n)$, where $A_i \in \mathbb{B}(\mathscr{H})$ are positive invertible operators with $\sigma(A_i) \subseteq [m, M]$ for some scalars $0 < m < M$.
(ii) $\mathbf{\Phi}^+ = (\Phi_1, \ldots, \Phi_n)$, where $\Phi_i : \mathbb{B}(\mathscr{H}) \to \mathbb{B}(\mathscr{K})$ are positive linear maps.
(iii) $C([m, M])$ is the set of all real valued continuous functions on an interval $[m, M]$.

We also need to remind the reader that a function $f \in C([m, M])$ is called operator monotone increasing (or operator increasing for short) if $f(A) \leq f(B)$ whenever A, B are self-adjoint operators with spectra in $[m, M]$ and such that $A \leq B$. That is, when f preserves the order of self-adjoint operator. A function $f \in C([m, M])$ is said to be operator decreasing if $-f$ is operator monotone.

The so called operator quasi-arithmetic mean of Mercer's type was defined in [5] as follows:

$$\widetilde{M}_{\varphi}\left(\mathbf{A}^+, \mathbf{\Phi}^+\right) := \varphi^{-1}\left(\left(\varphi(M) + \varphi(m)\right)\mathbf{1}_{\mathscr{K}} - \sum_{i=1}^{n} \Phi_i\left(\varphi\left(A_i\right)\right)\right).$$

In this reference, the following result was shown.

Theorem 3.1 *Let $\varphi, \psi \in C([m, M])$ be two strictly monotonic functions.*

(i) *If either $\psi \circ \varphi^{-1}$ is convex and ψ^{-1} is operator increasing, or $\psi \circ \varphi^{-1}$ is concave and ψ^{-1} is operator decreasing, then*

$$\widetilde{M}_{\varphi}\left(\mathbf{A}^+, \mathbf{\Phi}^+\right) \leq \widetilde{M}_{\psi}\left(\mathbf{A}^+, \mathbf{\Phi}^+\right). \tag{3.1}$$

(ii) *If either $\psi \circ \varphi^{-1}$ is concave and ψ^{-1} is operator increasing, or $\psi \circ \varphi^{-1}$ is convex and ψ^{-1} is operator decreasing, then the inequality in (3.1) is reversed.*

By virtue of Theorem 2.1, we have the following extension of this result.

Theorem 3.2 *Let $\varphi, \psi \in C([m, M])$ be two strictly monotonic functions and $\psi \circ \varphi^{-1}$ is twice differentiable function.*

(i) *If $\alpha \leq \left(\psi \circ \varphi^{-1}\right)''$ with $\alpha \in \mathbb{R}$ and ψ^{-1} is operator monotone, then*

$$\begin{aligned}&\widetilde{M}_{\varphi}\left(\mathbf{A}^+, \mathbf{\Phi}^+\right)\\ &\quad \leq \psi^{-1}\left\{\psi\left(\widetilde{M}_{\psi}\left(\mathbf{A}^+, \mathbf{\Phi}^+\right)\right) - \alpha K\left(m, M, \varphi, \mathbf{A}^+, \mathbf{\Phi}^+\right)\right\},\end{aligned} \tag{3.2}$$

where

$$K\left(m, M, \varphi, \mathbf{A}^{+}, \mathbf{\Phi}^{+}\right)$$
$$:= (\varphi(M) + \varphi(m)) \sum_{i=1}^{n} \Phi_i(\varphi(A_i)) - \varphi(M)\varphi(m)\mathbf{1}_{\mathscr{K}}$$
$$- \frac{1}{2}\left(\left(\sum_{i=1}^{n} \Phi_i(\varphi(A_i))\right)^2 + \sum_{i=1}^{n} \Phi_i\left(\varphi(A_i)^2\right)\right).$$

(ii) *If* $\left(\psi \circ \varphi^{-1}\right)'' \leq \beta$ *with* $\beta \in \mathbb{R}$ *and* ψ^{-1} *is operator monotone, then the reverse inequality is valid in* (3.2) *with* β *instead of* α.

Proof Let $f = \psi \circ \varphi^{-1}$ in (2.2) and replace A_i, m and M with $\varphi(A_i)$, $\varphi(m)$ and $\varphi(M)$ respectively. This implies

$$\psi\left(\widetilde{M}_{\varphi}\left(\mathbf{A}^{+}, \mathbf{\Phi}^{+}\right)\right) \leq \psi\left(\widetilde{M}_{\psi}\left(\mathbf{A}^{+}, \mathbf{\Phi}^{+}\right)\right) - \alpha K\left(m, M, \varphi, \mathbf{A}^{+}, \mathbf{\Phi}^{+}\right).$$

Since ψ^{-1} is operator monotone, the first conclusion follows immediately. The other case follows in a similar manner from (2.1). □

Similarly, Theorem 2.4 implies the following version.

Theorem 3.3 *Let* $\varphi, \psi \in C([m, M])$ *be two strictly monotonic functions. If* $\psi \circ \varphi^{-1}$ *is log-convex function and* ψ^{-1} *is operator increasing, then*

$$\widetilde{M}_{\varphi}(\mathbf{A}^{+}, \mathbf{\Phi}^{+})$$
$$\leq \psi^{-1}\left\{[\psi(m)]^{\frac{\sum_{i=1}^{n}\Phi_i(\varphi(A_i))-\varphi(m)\mathbf{1}_{\mathscr{K}}}{\varphi(M)-\varphi(m)}}[\psi(M)]^{\frac{\varphi(M)\mathbf{1}_{\mathscr{K}}-\sum_{i=1}^{n}\Phi_i(\varphi(A_i))}{\varphi(M)-\varphi(m)}}\right\}$$
$$\leq \widetilde{M}_{\psi}\left(\mathbf{A}^{+}, \mathbf{\Phi}^{+}\right).$$

Remark 3.4 By choosing appropriate functions φ and ψ, and making suitable substitutions, the above results imply some improvements of certain inequalities governing operator power mean of Mercer's type. We leave the details of this idea to the interested reader as an application of our main results.

In the end of the article, we show the example such that there is no relationship between inequalities in Theorems 3.2 and 3.3. Here, we restrict ourselves to the power function $f(t) = t^p$ with $p < 0$.

Example 3.5 It is sufficient to compare (2.5) and the first inequality of (2.11). We take $m = 1$ and $M = 3$. Setting

$$g(t) = \frac{M-t}{M-m}m^p + \frac{t-m}{M-m}M^p - \frac{p(p-1)M^{p-2}}{2}\left\{(M+m)t - Mm - t^2\right\} - \left(m^{\frac{M-t}{M-m}}M^{\frac{t-m}{M-m}}\right)^p.$$

Some calculations show that $g(2) \approx -0.0052909$ when $p = -0.2$, while $g(2) \approx 0.0522794$ when $p = -1$. We thus conclude that there is no ordering between the RHS of inequality in (2.5) and the RHS of first inequality of (2.11).

Acknowledgments The authors would like to thank the referees for their careful and insightful comments to improve our manuscript. The authors are also grateful to Dr. Trung Hoa Dinh for fruitful discussion and revising the manuscript.

References

1. S. Abramovich, J. Barić, J. Pečarić, A variant of Jessen's inequality of Mercer's type for superquadratic functions. J. Inequal. Pure Appl. Math. **9**(3) (2008), Article 62
2. E. Anjidani, M.R. Changalvaiy, Reverse Jensen-Mercer type operator inequalities. Electron J. Linear Algebra, **31**, 87–99 (2016)
3. J. Barić, A. Matković, J. Pečarić, A variant of the Jensen-Mercer operator inequality for superquadratic functions. Math. Comput. Modelling **51**, 1230–1239 (2010)
4. S. Ivelić, A. Matković, J. Pečarić, On a Jensen-Mercer operator inequality. Banach J. Math. Anal. **5**, 19–28 (2011)
5. A. Matković, J. Pečarić, I. Perić, A variant of Jensen's inequality of Mercer's type for operators with applications. Linear Algebra Appl. **418**, 551–564 (2006)
6. A. Matković, J. Pečarić, I. Perić, Refinements of Jensen's inequality of Mercer's type for operator convex functions. Math. Ineq. Appl. **11**, 113–126 (2008)
7. A.McD. Mercer, A variant of Jensen's inequality. J. Inequal. Pure Appl. Math. **4**(4) (2003), Article 73
8. J. Mićić, H.R. Moradi, S. Furuichi, Choi-Davis-Jensen's inequality without convexity. J. Math. Inequal. **12**, 1075–1085 (2018)
9. H.R. Moradi, M.E. Omidvar, M. Adil Khan, K. Nikodem, *Around Jensen's inequality for strongly convex functions*. Aequationes Math. **92**, 25–37 (2018)

A Numerical Approach for Approximating Variable-Order Fractional Integral Operator

Somayeh Nemati

Abstract In this work, we propose a numerical method to find an approximation of the variable-order integral of a given function using a generalization of the modified hat functions. First, an operational matrix of the basis functions corresponding to the variable-order integral operator is introduced. Then, using this matrix and an approximation of the given function, we find an approximation of the variable-order integral operator of the function. An error estimate is proved. Two test examples are included to show the efficiency and accuracy of our new technique. Finally, this new technique is used to solve the variable-order differential equations numerically and some illustrative problems are provided to validate the applicability and accuracy of this new scheme.

Keywords Variable-order integral operator · Modified hat functions · Operational matrix

Mathematics Subject Classification (2010) 34A08; 65M70

1 Introduction

A recent generalization of the theory of fractional calculus is to let the fractional order of the derivatives to be of variable order. In [16], authors have investigated operators when the order of fractional derivative is a variable on time. The non-local properties of systems are more visible with variable-order fractional calculus, and

S. Nemati (✉)
Department of Applied Mathematics, Faculty of Mathematical Sciences, University of Mazandaran, Babolsar, Iran
Center for Research and Development in Mathematics and Applications (CIDMA), Department of Mathematics, University of Aveiro, Aveiro, Portugal
e-mail: s.nemati@umz.ac.ir; s.nemati@ua.pt

M. A. Bastos et al. (eds.), *Operator Theory, Functional Analysis and Applications*, Operator Theory: Advances and Applications 282,
https://doi.org/10.1007/978-3-030-51945-2_25

many real world phenomena in physics, mechanics, control, and signal processing have been described by this approach [2, 13–15].

Hat functions are defined on the interval [0, 1] and are continuous with shape hats [17]. In the last few years, a modification of hat functions has been introduced and used for solving some different kinds of problems. For instance: two-dimensional linear Fredholm integral equations [5], integral equations of Stratonovich–Volterra [6] and Volterra–Fredholm type [7], fractional integro-differential equations [8], fractional pantograph differential equations [9], fractional optimal control problems [10] and systems of fractional differential equations [11]. These functions are defined on the interval $[0, T]$, with $T > 0$. In this paper, we consider a general definition of these basis functions on $[a, b]$, with $b > a$, called generalized modified hat functions (GMHFs). These basis functions are utilized to present a numerical method to approximate the variable-order fractional integral of a given function. For $y : [a, b] \to \mathbb{R}$ and a real function $\alpha(t)$ such that $0 < \alpha(t) \leq 1$ for all $t \in [a, b]$, the left Riemann–Liouville fractional integral operator of order $\alpha(t)$ of y is defined by (see, e.g., [1])

$$ {}_aI_t^{\alpha(t)} y(t) = \frac{1}{\Gamma(\alpha(t))} \int_a^t (t-s)^{\alpha(t)-1} y(s) ds, \quad t > a, $$

where Γ is the Euler gamma function.

The organization of this paper is as follows. In Sect. 2, we give a new general definition of the modified hat functions and introduce some of their main properties. Section 3 is devoted to introducing a numerical method based on the GMHFs to compute the variable-order fractional integral of a given function. The error analysis of this new approximation is provided and the integral of two test functions are computed to illustrate the accuracy of the method. An application of this new technique for solving variable-order fractional differential equations is proposed in Sect. 4. Finally, we conclude the paper in Sect. 5.

2 Definition and Function Approximation

In this section, we introduce a generalization of the modified hat functions and present some of their main properties. To this aim, the interval $[a, b]$ is divided into m subintervals $[a + ih, a + (i + 1)h]$, $i = 0, 1, 2, \ldots, m - 1$, of equal lengths h, where $h = \frac{b-a}{m}$ and $m \geq 2$ is an even integer number. A $(m + 1)$-set of GMHFs are defined on the interval $[a, b]$ as follows

$$ \psi_0(t) = \begin{cases} \frac{1}{2h^2} (t - (a + h)) (t - (a + 2h)), \\ \qquad \text{if} \quad a \leq t \leq a + 2h, \\ 0, \qquad \text{otherwise}, \end{cases} \tag{2.1} $$

if i is odd and $1 \leq i \leq m-1$:

$$\psi_i(t) = \begin{cases} \frac{-1}{h^2}\left(t-(a+(i-1)h)\right)\left(t-(a+(i+1)h)\right), \\ \qquad \text{if} \quad a+(i-1)h \leq t \leq a+(i+1)h, \\ 0, \qquad \text{otherwise}, \end{cases} \tag{2.2}$$

if i is even and $2 \leq i \leq m-2$:

$$\psi_i(t) = \begin{cases} \frac{1}{2h^2}\left(t-(a+(i-1)h)\right)\left(t-(a+(i-2)h)\right), \\ \qquad \text{it} \quad a+(i-2)h \leq t \leq a+ih, \\ \frac{1}{2h^2}\left(t-(a+(i+1)h)\right)\left(t-(a+(i+2)h)\right), \\ \qquad \text{if} \quad a+ih \leq t \leq a+(i+2)h, \\ 0, \qquad \text{otherwise}, \end{cases} \tag{2.3}$$

and

$$\psi_m(t) = \begin{cases} \frac{1}{2h^2}\left(t-(b-h)\right)\left(t-(b-2h)\right), \\ \qquad \text{if} \quad b-2h \leq t \leq b, \\ 0, \qquad \text{otherwise}. \end{cases} \tag{2.4}$$

The GMHFs build a set of linearly independent functions in the space $L^2[a,b]$. Figure 1 displays a set of the GMHFs obtained by $m=4$ and defined on the interval $[-1,1]$.

The following properties can be easily investigated using the definition of the GMHFs:

$$\psi_i(a+jh) = \begin{cases} 1, & i=j, \\ 0, & i \neq j, \end{cases} \qquad \sum_{i=0}^{m} \psi_i(t) = 1, \quad t \in [a,b],$$

$$\psi_i(t)\psi_j(t) = \begin{cases} 0, \text{ if } i \text{ is even and } |i-j| \geq 3, \\ 0, \text{ if } i \text{ is odd and } |i-j| \geq 2. \end{cases}$$

Any arbitrary function $y \in L^2[a,b]$ may be approximated in terms of the GMHFs as [5]

$$y(t) \simeq y_m(t) = \sum_{i=0}^{m} y_i \psi_i(t) = Y^T \Psi(t), \tag{2.5}$$

where

$$\Psi(t) = [\psi_0(t), \psi_1(t), \ldots, \psi_m(t)]^T, \tag{2.6}$$

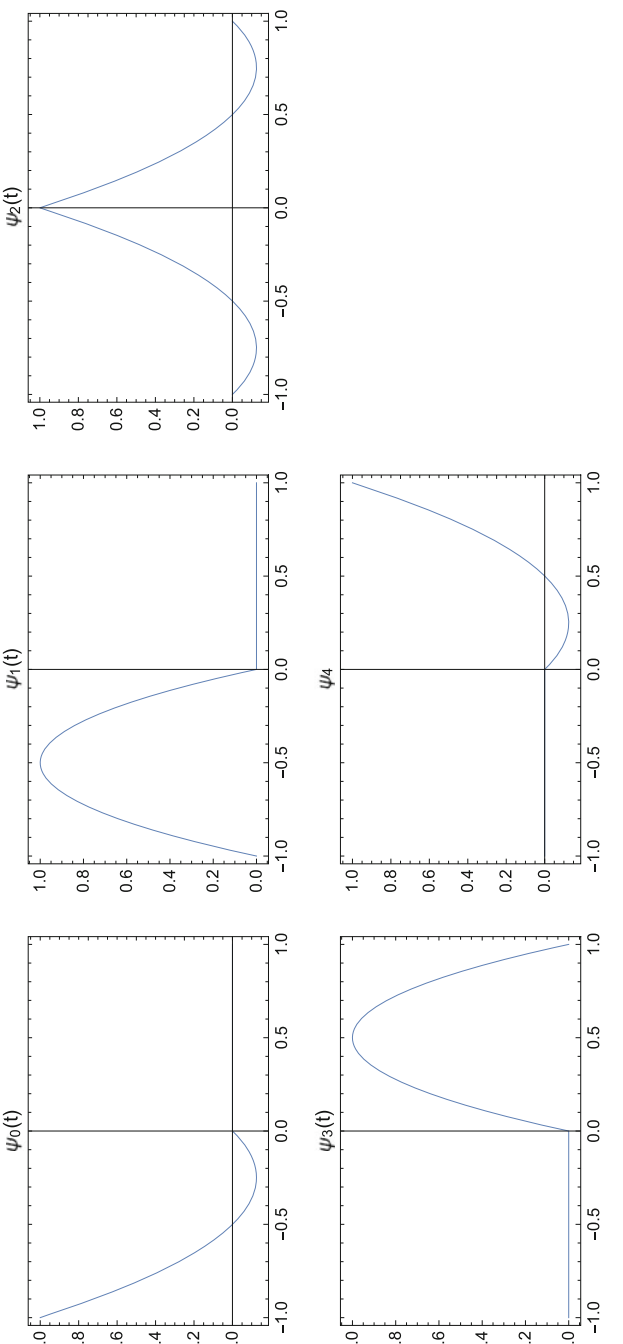


Fig. 1 Plot of the GMHFs with $m = 4$ on $[-1, 1]$

and

$$Y = [y_0, y_1, \ldots, y_m]^T,$$

with

$$y_i = y(a + ih). \tag{2.7}$$

We have the following theorem for the error of the function approximation using the GMHFs.

Theorem 2.1 *If a function $y \in C^3([a, b])$ is approximated by the family of first $(m + 1)$ GMHFs as defined by (2.1)–(2.4), then*

$$|y(t) - y_m(t)| = O(h^3),$$

where $h = \frac{b-a}{m}$, and y_m is defined by (2.5).

Proof It can be proved in a same way as done in [8, Theorem 4.1]. □

3 Numerical Method

In this section, we propose a numerical method based on the GMHFs for computing the left Riemann–Liouville variable-order integral of a given function.

3.1 Variable-Order Integration Rule

In order to introduce the method, we first obtain an operational matrix of variable-order integration for the GMHFs basis vector. To do this, we apply the Riemann–Liouville fractional integral operator of order $0 < \alpha(t) \leq 1$, for all $t \in [a, b]$, to the function $\psi_i(t)$, and obtain

$${}_aI_t^{\alpha(t)}\psi_i(t) = \frac{1}{\Gamma(\alpha(t))}\int_a^t (t-s)^{\alpha(t)-1}\psi_i(s)ds, \quad i = 0, 1, \ldots, m.$$

Then, we may expand these functions using the GMHFs as follows:

$${}_aI_t^{\alpha(t)}\psi_i(t) \simeq \sum_{j=0}^{m} {}_aP_{ij}^{\alpha}\psi_j(t), \quad i = 0, 1, \ldots, m,$$

wherein according to (2.7), ${}_aP_{ij}^{\alpha}$ is the value of ${}_aI_t^{\alpha(t)}\psi_i(t)$ at $a+jh$. That is

$$ {}_aP_{ij}^{\alpha} = \frac{1}{\Gamma(\alpha(a+jh))}\int_a^{a+jh}(a+jh-s)^{\alpha(a+jh)-1}\psi_i(s)ds, \quad i,j=0,1,\ldots,m. \tag{3.1}$$

By substituting the definitions of the functions $\psi_i(t)$ given by (2.1)–(2.4), we compute the integral part of (3.1). For $i=0$, we obtain

$$ {}_aP_{00}^{\alpha} = 0, \tag{3.2}$$

$$ {}_aP_{01}^{\alpha} = \frac{h^{\alpha(a+h)}}{2\Gamma(\alpha(a+h)+3)}\left[\alpha(a+h)\left(3+2\alpha(a+h)\right)\right], \tag{3.3}$$

and for $j>1$,

$$\begin{aligned} {}_aP_{0j}^{\alpha} =& \frac{h^{\alpha(a+jh)}}{2\Gamma(\alpha(a+jh)+3)} \\ &\times\Big[j^{\alpha(a+jh)+1}\left(2j-6-3\alpha(a+jh)\right) \\ &+2j^{\alpha(a+jh)}\left(1+\alpha(a+jh)\right)\left(2+\alpha(a+jh)\right) \\ &-(j-2)^{\alpha(a+jh)+1}\left(2j-2+\alpha(a+jh)\right)\Big]. \end{aligned} \tag{3.4}$$

If i is odd, we get for $j<i$,

$$ {}_aP_{ij}^{\alpha} = 0, \tag{3.5}$$

$$ {}_aP_{ii}^{\alpha} = \frac{2h^{\alpha(a+ih)}}{\Gamma(\alpha(a+ih)+3)}(1+\alpha(a+ih)), \tag{3.6}$$

and for $j>i$,

$$\begin{aligned} {}_aP_{ij}^{\alpha} =& \frac{2h^{\alpha(a+jh)}}{\Gamma(\alpha(a+jh)+3)} \\ &\times\Big[(j-i-1)^{\alpha(a+jh)+1}(j-i+1+\alpha(a+jh)) \\ &-(j-i+1)^{\alpha(a+jh)+1}(j-i-1-\alpha(a+jh))\Big]. \end{aligned} \tag{3.7}$$

Finally, if i is even, we have for $j < i-1$,

$$
\begin{aligned}
{}_aP_{ij}^{\alpha} &= 0, && (3.8)\\
{}_aP_{i,i-1}^{\alpha} &= \frac{h^{\alpha(a+(i-1)h)}}{2\Gamma(\alpha(a+(i-1)h)+3)}\left[-\alpha(a+(i-1)h)\right], && (3.9)\\
{}_aP_{ii}^{\alpha} &= \frac{h^{\alpha(a+ih)}}{2\Gamma(\alpha(a+ih)+3)}\left[2^{\alpha(a+ih)+1}(2-\alpha(a+ih))\right], && (3.10)\\
{}_aP_{i,i+1}^{\alpha} &= \frac{h^{\alpha(a+(i+1)h)}}{2\Gamma(\alpha(a+(i+1)h)+3)}\\
&\quad\times\Big[3^{\alpha(a+(i+1)h)+1}(4-\alpha(a+8i+1)h))\\
&\quad-6(2+\alpha(a+(i+1)h))\Big], && (3.11)
\end{aligned}
$$

and for $j > i+1$,

$$
\begin{aligned}
{}_aP_{ij}^{\alpha} &= \frac{h^{\alpha(a+jh)}}{2\Gamma(\alpha(a+jh)+3)}\\
&\quad\times\Big[(j-i+2)^{\alpha(a+jh)+1}(2j-2i+2-\alpha(a+jh))\\
&\quad-6(j-i)^{\alpha(a+jh)+1}(2+\alpha(a+jh))\\
&\quad-(j-i-2)^{\alpha(a+jh)+1}(2j-2i-2+\alpha(a+jh))\Big]. && (3.12)
\end{aligned}
$$

As a result, we can write

$${}_aI_t^{\alpha(t)}\Psi(t) \simeq P_a^{\alpha(t)}\Psi(t), \tag{3.13}$$

where $P_a^{\alpha(t)}$ is an upper Hessenberg matrix of dimension $(m+1)\times(m+1)$ called operational matrix of variable-order integration of order $\alpha(t)$ in the Riemann-Liouville sense, and is given using (3.2)–(3.12) as follows:

$$
P_a^{\alpha(t)} = \begin{bmatrix}
0 & {}_aP_{01}^{\alpha} & {}_aP_{02}^{\alpha} & {}_aP_{03}^{\alpha} & {}_aP_{04}^{\alpha} & \cdots & {}_aP_{0(m-1)}^{\alpha} & {}_aP_{0m}^{\alpha}\\
0 & {}_aP_{11}^{\alpha} & {}_aP_{12}^{\alpha} & {}_aP_{13}^{\alpha} & {}_aP_{14}^{\alpha} & \cdots & {}_aP_{1(m-1)}^{\alpha} & {}_aP_{1m}^{\alpha}\\
0 & {}_aP_{21}^{\alpha} & {}_aP_{22}^{\alpha} & {}_aP_{23}^{\alpha} & {}_aP_{24}^{\alpha} & \cdots & {}_aP_{2(m-1)}^{\alpha} & {}_aP_{2m}^{\alpha}\\
0 & 0 & 0 & {}_aP_{33}^{\alpha} & {}_aP_{34}^{\alpha} & \cdots & {}_aP_{3(m-1)}^{\alpha} & {}_aP_{3m}^{\alpha}\\
0 & 0 & 0 & {}_aP_{43}^{\alpha} & {}_aP_{44}^{\alpha} & \cdots & {}_aP_{4(m-1)}^{\alpha} & {}_aP_{4m}^{\alpha}\\
\vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots\\
0 & 0 & 0 & 0 & 0 & \cdots & {}_aP_{(m-1)(m-1)}^{\alpha} & {}_aP_{(m-1)m}^{\alpha}\\
0 & 0 & 0 & 0 & 0 & \cdots & {}_aP_{m(m-1)}^{\alpha} & {}_aP_{mm}^{\alpha}
\end{bmatrix}. \tag{3.14}
$$

In order to compute the left Riemann–Liouville variable-order integral operator of an arbitrary function $y \in L^2[a,b]$, first, we approximate this function using the GMHFs as (2.5). Then, using (3.13), we have

$$ {}_aI_t^{\alpha(t)} y(t) \simeq {}_aI_t^{\alpha(t)} y_m(t) = Y^T {}_aI_t^{\alpha(t)} \Psi(t) \simeq Y^T P_a^{\alpha(t)} \Psi(t). $$

3.2 Error Bound

The aim of this section is to give an error bound for the numerical estimation of the left Riemann–Liouville variable-order integral operator of an arbitrary function in the sense of L^2-norm. To this aim, we consider the following L^2-norm for a function $f \in L^2[a,b]$:

$$ \|f\|_2 = \left(\int_a^b |f(t)|^2 dt \right)^{\frac{1}{2}}. $$

With the help of Theorem 2.1 we obtain the following result.

Theorem 3.1 *Assume that $y \in C^3([a,b])$. Suppose that the left Riemann–Liouville variable-order integral operator of y is approximated by $Y^T P_a^{\alpha(t)} \Psi(t)$ where $\Psi(t)$ and $P_a^{\alpha(t)}$ are given by (2.6) and (3.14), respectively, and the elements of Y are $y_i = y(a+ih)$, $i = 0, 1, \ldots, m$, with $h = \frac{b-a}{m}$. Then,*

$$ \left\| {}_aI_t^{\alpha(t)} y - Y^T P_a^{\alpha(t)} \Psi \right\|_2 = O(h^3). $$

Proof By assuming $0 < \alpha(t) \leq 1$, for all $t \in [a,b]$, we consider the following definition of the norm for the operator ${}_aI_t^{\alpha(t)}$,

$$ \|{}_aI_t^{\alpha(t)}\|_2 = \sup_{\|f\|_2=1} \|{}_aI_t^{\alpha(t)} f\|_2. $$

We show that this operator is bounded. Using Schwarz's inequality, we obtain

$$ \begin{aligned} \left\| {}_aI_t^{\alpha(t)} f \right\|_2 &= \left\| \frac{1}{\Gamma(\alpha(t))} \int_a^t (t-s)^{\alpha(t)-1} f(s) ds \right\|_2 \\ &\leq \|f\|_2 \left\| \frac{1}{\Gamma(\alpha(t))} \int_a^t (t-s)^{\alpha(t)-1} ds \right\|_2 \\ &= \left\| \frac{(t-a)^{\alpha(t)}}{\Gamma(\alpha(t)+1)} \right\|_2, \end{aligned} $$

where we have used $\|f\|_2 = 1$. Since we have $\Gamma(t) > 0.8$ for all $t > 0$, we can write $\Gamma(\alpha(t)+1) > \frac{4}{5}$. Therefore, we get

$$\left\| \frac{(t-a)^{\alpha(t)}}{\Gamma(\alpha(t)+1)} \right\|_2 < \frac{5}{4} \left\| (t-a)^{\alpha(t)} \right\|_2,$$

which together with $0 < \alpha(t) \leq 1$ help us to have one of the following statements:

1. If $0 < t - a < 1$, then

$$\left\| \frac{(t-a)^{\alpha(t)}}{\Gamma(\alpha(t)+1)} \right\|_2 < \frac{5}{4} \|1\|_2 = \frac{5}{4}(b-a)^{\frac{1}{2}}.$$

2. If $1 \leq t - a \leq b - a$, then

$$\left\| \frac{(t-a)^{\alpha(t)}}{\Gamma(\alpha(t)+1)} \right\|_2 < \frac{5}{4} \|t-a\|_2 = \frac{5}{4\sqrt{3}}(b-a)^{\frac{3}{2}}.$$

Hence there is a positive constant dependent on a and b so that

$$\|{}_aI_t^{\alpha(t)}\|_2 < C.$$

We use the boundedness property of ${}_aI_t^{\alpha(t)}$ to continue the proof. Using Theorem 2.1, we have

$$\|y - y_m\|_2 = O(h^3) \tag{3.15}$$

and

$$\left\| Y^T {}_aI_t^{\alpha(t)} \Psi - Y^T P_a^{\alpha(t)} \Psi \right\|_2 = O(h^3). \tag{3.16}$$

Finally, using (3.15) and (3.16), we obtain

$$\begin{aligned}
\left\| {}_aI_t^{\alpha(t)} y - Y^T P_a^{\alpha(t)} \Psi \right\|_2 &= \left\| {}_aI_t^{\alpha(t)} y - {}_aI_t^{\alpha(t)} y_m + {}_aI_t^{\alpha(t)} y_m - Y^T P_a^{\alpha(t)} \Psi \right\|_2 \\
&\leq \left\| {}_aI_t^{\alpha(t)} y - {}_aI_t^{\alpha(t)} y_m \right\|_2 + \left\| {}_aI_t^{\alpha(t)} y_m - Y^T P_a^{\alpha(t)} \Psi \right\|_2 \\
&\leq \left\| {}_aI_t^{\alpha(t)} \right\|_2 \|y - y_m\|_2 + \left\| Y^T {}_aI_t^{\alpha(t)} \Psi - Y^T P_a^{\alpha(t)} \Psi \right\|_2 \\
&= O(h^3),
\end{aligned}$$

which completes the proof. □

3.3 Test Examples

In this section, in order to illustrate the accuracy and applicability of the proposed method, we consider two functions and employ the method for computing the left Riemann–Liouville variable-order integral of these functions. We note that the method was carried out using Mathematica 12. To show the accuracy, the l^2 norm of the error and the convergence order are defined, respectively, by

$$E_m = \left(\frac{1}{m} \sum_{i=1}^{m} (J(t_i) - J_m(t_i))^2 \right)^{\frac{1}{2}}, \quad \epsilon_m = \log_2 (E_m / E_{2m}),$$

where $t_i = a + ih$. Furthermore, in computation of E_m, we have

$$J(t) = {}_a I_t^{\alpha(t)} y(t), \quad J_m(t) = Y^T P_a^{\alpha(t)} \Psi(t)$$

as the exact and approximate integrals of the function y, respectively.

Example 3.1 Suppose that $y(t) = \sin(t)$, $\alpha(t) = t$, $a = 0$ and $b = 1$. The exact left Riemann–Liouville variable-order integral of y is

$$J(t) = \frac{t^t}{(1+t)\Gamma(t)} {}_1F_2 \left(1; 1 + \frac{t}{2}, \frac{3}{2} + \frac{t}{2}; -\frac{t^2}{4} \right),$$

where ${}_pF_q$ is the generalized hypergeometric function defined by

$${}_pF_q \left(a_1, \ldots, a_p; b_1, \ldots, b_q; z \right) = \sum_{k=0}^{\infty} \frac{(a_1)_k \ldots (a_p)_k}{(b_1)_k \ldots (b_q)_k} \frac{z^k}{k!},$$

and $(a)_k = a(a+1)\ldots(a+k-1)$ is Pochhammer symbol.

Example 3.2 Consider $y(t) = e^t$, $\alpha(t) = \sin(t)$, $a = 1$ and $b = 3$. In this case, we have

$$J(t) = e^t \left(1 - \frac{\Gamma(\sin(t), t-1)}{\Gamma(\sin(t))} \right),$$

where $\Gamma(c, t)$ is the incomplete gamma function defined by

$$\Gamma(\beta, t) = \int_t^{\infty} s^{\beta-1} e^{-s} ds.$$

We have employed the present method to obtain approximations of the left Riemann–Liouville variable-order integral operator by considering the information given in Examples 3.1 and 3.2. Numerical results are displayed in Table 1 and Fig. 2.

Table 1 Numerical results for the l^2 norm of the error and the convergence order

	Example 3.1		Example 3.2	
m	E_m	ϵ_m	E_m	ϵ_m
2	$1.99e{-3}$	3.63	$2.23e{-1}$	3.42
4	$1.61e{-4}$	3.45	$2.08e{-2}$	3.46
8	$1.47e{-5}$	3.36	$1.89e{-3}$	3.39
16	$1.43e{-6}$	3.30	$1.80e{-4}$	3.35
32	$1.45e{-7}$	3.26	$1.76e{-5}$	3.33
64	$1.51e{-8}$	3.24	$1.75e{-6}$	3.31
128	$1.60e{-9}$	3.22	$1.77e{-7}$	3.28
256	$1.72e{-10}$	3.18	$1.82e{-8}$	2.51
512	$1.90e{-11}$	–	$3.19e{-9}$	–

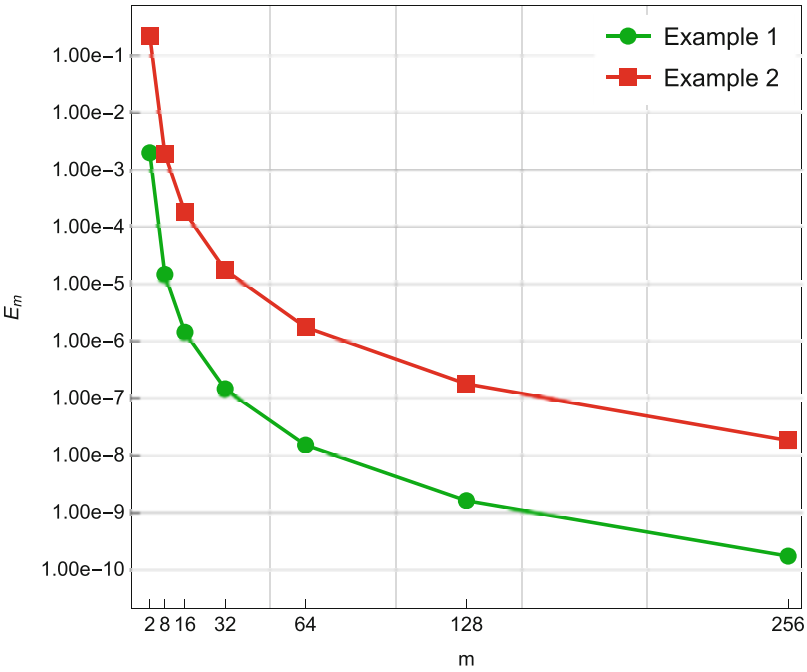


Fig. 2 The l^2 norm of the error for some selected values of m in logarithmic scale

In Table 1, the l^2 norm of the error and the convergence order are presented which confirm the $O(h^3)$ accuracy of this method. Furthermore, in Fig. 2, the results for the l^2 norm of the error are plotted in a logarithmic scale.

4 Application to Variable-Order Fractional Differential Equations

In this section, a numerical method based on the use of operational matrix of variable-order integration for the GMHFs basis vector is introduced for solving variable-order fractional differential equations. Moreover, some examples are provided to show the accuracy of this method.

4.1 Method of Solution

Consider the variable-order fractional initial value problem

$$\begin{cases} {}_a^C D_t^{\alpha(t)} y(t) = f(t, y(t)),\ 0 < \alpha(t) \le 1, \quad a \le t \le b, \\ y(a) = y_0, \end{cases} \tag{4.1}$$

where y is the unknown function, $f : [a, b] \times \mathbb{R} \longrightarrow \mathbb{R}$ is a given function, y_0 is a real constant and ${}_a^C D_t^{\alpha(t)}$ denotes the left Caputo fractional derivative of order $\alpha(t)$ defined by Almeida et al. [1]:

$${}_a^C D_t^{\alpha(t)} y(t) = \frac{1}{\Gamma(1-\alpha(t))} \int_a^t (t-s)^{-\alpha(t)} y'(s) ds, \quad t > a.$$

Let $y(t) \in C^1[a, b]$ be the exact solution of (4.1), then an approximation of $y'(t)$ based on the GMHFs can be considered as follows:

$$y'(t) = C^T \Psi(t), \tag{4.2}$$

where C is a vector with unknown parameters $c_i, i = 0, 1, \ldots, m$, and $\Psi(t)$ is given by (2.6). Therefore, an approximation of y can be given using (3.13), (4.2) and the initial condition as

$$\begin{aligned} y(t) &= \int_a^t y'(s) ds + y(a) \simeq C^T {}_a I_t^1 \Psi(t) + y_0 \\ &\simeq (C^T P_a^1 + Y_a^T) \Psi(t) = Y^T \Psi(t), \end{aligned} \tag{4.3}$$

where

$$Y_a = [y_0, y_0, \ldots, y_0]^T, \quad Y = (C^T P_a^1 + Y_a^T)^T = [y_0, y_1, \ldots, y_m]^T.$$

In a similar way, we have

$${}_a^C D_t^{\alpha(t)} y(t) = {}_a I_t^{1-\alpha(t)} y'(t) \simeq C^T P_a^{1-\alpha(t)} \Psi(t). \tag{4.4}$$

On the other hand, an approximation of the function $f : ([a, b] \times \mathbb{R}) \to \mathbb{R}$ based on the GMHFs is given by

$$f(t, y(t)) \simeq \sum_{i=0}^{n} f(t_i, y(t_i))\psi_i(t) = \sum_{i=0}^{n} f(t_i, y_i)\psi_i(t) = F(\Theta, Y)\Psi(t), \tag{4.5}$$

with $t_i = a + ih$, and

$$F(\Theta, Y) = [f(a, y_0), f(a + h, y_1), \ldots, f(b, y_m)].$$

Now, by substituting (4.4) and (4.5) into (4.1), we get

$$C^T P_a^{1-\alpha(t)} = F(\Theta, Y), \tag{4.6}$$

which is a system of nonlinear algebraic equations in terms of the unknown parameters of the vector C. The first equation of the system (4.6) is a degenerate equation. By letting $\alpha(t) \to 1$, we impose the condition $y'(a) = f(a, y_0)$ for its solvability. With this condition, we obtain $c_0 = f(a, y_0)$. In order to determine the remaining unknown entries of C, we substitute $c_0 = f(a, y_0)$ into equations of (4.6). By solving the resulting system, an approximation of the function y is given by (4.3).

4.2 Numerical Simulations

Here, we employ the proposed method for solving three examples to demonstrate the applicability and accuracy of our new method.

Example 4.3 As the first example, consider the problem (4.1) with $0 \le t \le 1$ and [4]

$$f(t, y) = t^3 + \frac{\Gamma(4)}{\Gamma(4 - \alpha(t))} t^{3-\alpha(t)} + t^2 + \frac{\Gamma(3)}{\Gamma(3 - \alpha(t))} t^{2-\alpha(t)} - y.$$

The exact solution of this problem is $y(t) = t^3 + t^2$. By considering $\alpha(t) = \sin t$, this problem has been solved by different values of m. Plots of the error for $m = 4, 8, 16, 32$ are displayed in Fig. 3. In [4], this problem has been solved by considering different fixed values of $\alpha \in (0, 1)$ using a high order numerical methods based on a second-degree compound quadrature formula and the convergence order of the error at $t = 1$ has been reported. Since the proposed method in this paper gives the exact solution at $t = 1$, we report the L^2-norm of the error with different m for $\alpha = 0.75$, and compare the convergence order with the method of [4] in Table 2. Also, we can see in this table the CPU times (in seconds), which have been obtained on a 2.5 GHz Core i7 personal computer with 16 GB of

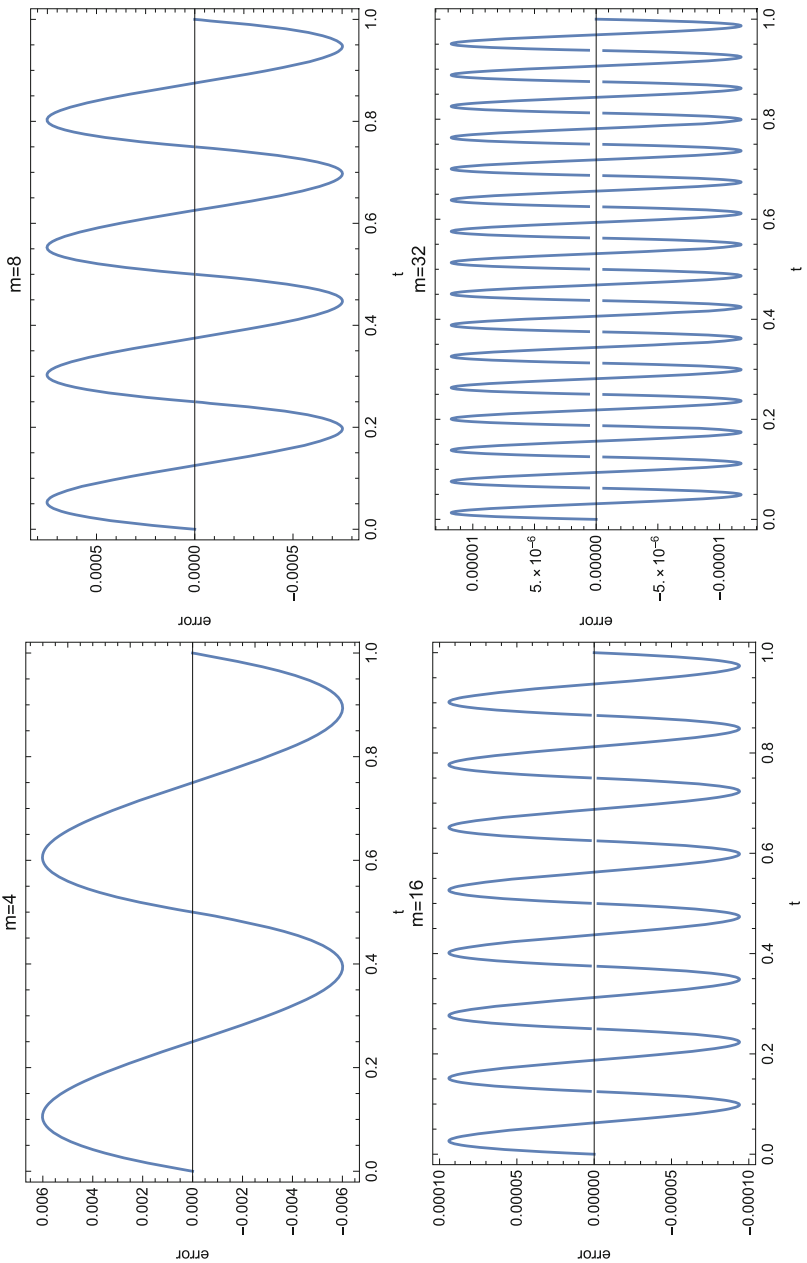


Fig. 3 Plots of the error with $m = 4, 8, 16, 32$ for Example 4.3 with $\alpha(t) = \sin t$

Table 2 Numerical results for the L^2-error and convergence order for Example 4.3 with $\alpha = 0.75$

	Present method			Method of [4]
m	L^2-error	Convergence order	CPU time	Convergence order
10	$2.76e{-4}$	3.00	0.000	1.79
20	$3.45e{-5}$	3.00	0.016	2.07
40	$4.31e{-6}$	3.00	0.016	2.17
80	$5.39e{-7}$	2.99	0.062	2.21
160	$6.74e{-8}$	3.00	0.234	2.23
320	$8.42e{-9}$	3.00	0.906	2.24
640	$1.05e{-9}$	–	2.828	–

RAM using Mathematica 12. The Mathematica function FindRoot was used for solving the resulting systems.

Example 4.4 Consider the problem (4.1) with $0 \le t \le 1$ and

$$f(t, y) = \frac{\Gamma\left(\frac{9}{2}\right) t^{\frac{7}{2}-\alpha(t)}}{\Gamma\left(\frac{9}{2} - \alpha(t)\right)} + t^7 \sin(t) - y^2 \sin(t).$$

The exact solution of this problem is $t^{7/2}$. By setting $\alpha(t) = 1 - 0.5\exp(-t)$, the numerical results of employing the proposed method for solving this example are given in Fig. 4 and Table 3.

Example 4.5 As the last example, consider the following Riccati fractional differential equation [3, 12]

$${}_0^C D_t^{\alpha(t)} y(t) = 2y(t) - y^2(t) + 1, \quad 0 < \alpha \le 1, \quad 0 \le t \le 1,$$

with initial condition $y(0) = 0$. When $\alpha(t) = 1$, the exact solution is

$$y(t) = 1 + \sqrt{2}\tanh\left(\sqrt{2}t + \frac{1}{2}\ln\left(\frac{\sqrt{2}-1}{\sqrt{2}+1}\right)\right).$$

By taking $\alpha(t) = 1$, a comparison between the numerical results obtained by the present method with $m = 50$, the modified homotopy perturbation method [12] using the fourth-order term and Chebyshev wavelet method [3] using $\hat{m} = 192$ is provided at some selected points in Table 4. Furthermore, the numerical results for $y(t)$ with $\alpha = 0.65, 0.75, 0.85, 0.95$ and $m = 10$ together with the exact solution for $\alpha = 1$ are plotted in Fig. 5. As it could be expected, when α is close to 1, the numerical solution is close to the exact solution of the case $\alpha = 1$.

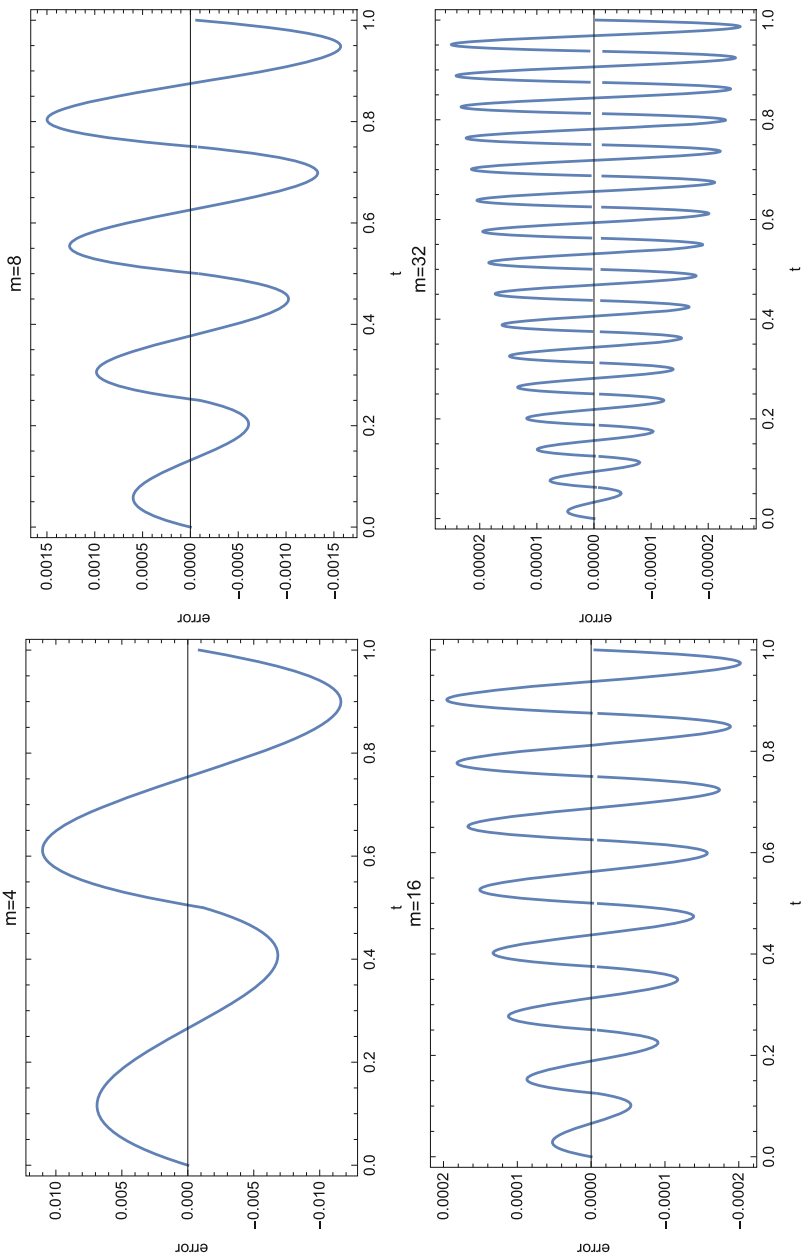


Fig. 4 Plots of the error with $m = 4, 8, 16, 32$ for Example 4.4 with $\alpha(t) = 1 - 0.5\exp(-t)$

Table 3 Numerical results for Example 4.4 with $\alpha(t) = 1 - 0.5\exp(-t)$

m	L^2-error	Convergence order	CPU time
2	$5.53e{-2}$	3.04	0.000
4	$6.72e{-3}$	3.01	0.000
8	$8.34e{-4}$	3.00	0.016
16	$1.04e{-4}$	3.00	0.031
32	$1.30e{-5}$	3.00	0.078
64	$1.63e{-6}$	3.00	0.687
128	$2.04e{-7}$	–	3.953

Table 4 Numerical results for Example 4.5 with $\alpha(t) = 1$

t	Method of [12]	Method of [3]	Present method	Exact solution
0	0	0.000001	0	0
0.1	0.110294	0.110311	0.110295	0.110295
0.2	0.241965	0.241995	0.241977	0.241977
0.3	0.395106	0.395123	0.395105	0.395105
0.4	0.568115	0.567829	0.567812	0.567812
0.5	0.757564	0.756029	0.756014	0.756014
0.6	0.958259	0.953576	0.953566	0.953566
0.7	1.163459	1.152955	1.152949	1.152949
0.8	1.365240	1.346365	1.346364	1.346364
0.9	1.554960	1.526909	1.526911	1.526911
1	1.723810	1.689494	1.689498	1.689498

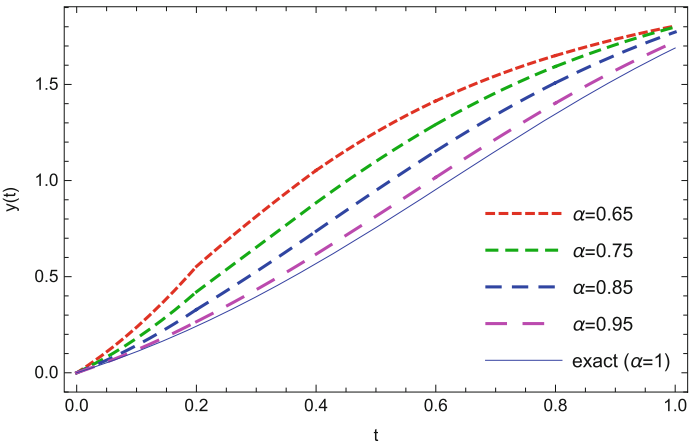


Fig. 5 Numerical solutions with $m = 10$ and different values of α together with the exact solution with $\alpha = 1$ for Example 4.5

5 Concluding Remarks

A new numerical technique has been introduced for computing the left Riemann–Liouville variable-order integral of a given function. A generalized class of the modified hat functions (GMHFs) has been considered and used to suggest our new scheme. The given function is easily expanded using the GMHFs. An operational matrix of variable-order integral of the basis vector is computed and used to approximate the integral of the function under consideration. The convergence order of our numerical method is proved and confirmed by the results of two illustrative examples. Finally, this new technique is employed to solve variable-order differential equations and the numerical results demonstrate the efficiency of the method.

Acknowledgement The author is grateful to two anonymous referees for several positive and constructive comments, which helped her to improve the manuscript.

References

1. R. Almeida, D. Tavares, D.F.M. Torres, *The Variable-Order Fractional Calculus of Variations* (Springer, Berlin, 2019)
2. C.F.M. Coimbra, C.M. Soon, M.H. Kobayashi, The variable viscoelasticity operator. Ann. Phys. **14**, 378–389 (2005)
3. Y. Li, Solving a nonlinear fractional differential equation using Chebyshev wavelets. Commun. Nonlinear Sci. Numer. Simul. **15**, 2284–2292 (2010)
4. Z. Li, Y. Yan, N.J. Ford, Error estimates of a high order numerical method for solving linear fractional differential equations. Appl. Numer. Math. **114**, 201–220 (2017)
5. F. Mirzaee, E. Hadadiyan, Numerical solution of linear Fredholm integral equations via two dimensional modification of hat functions. Appl. Math. Comput. **250**, 805–816 (2015)
6. F. Mirzaee, E. Hadadiyan, Approximation solution of nonlinear Stratonovich Volterra integral equations by applying modification of hat functions. J. Comput. Appl. Math. **302**, 272–284 (2016)
7. F. Mirzaee, E. Hadadiyan, Numerical solution of Volterra–Fredholm integral equations via modification of hat functions. Appl. Math. Comput. **280**, 110–123 (2016)
8. S. Nemati, P.M. Lima, Numerical solution of nonlinear fractional integro-differential equations with weakly singular kernels via a modification of hat functions. Appl. Math. Comput. **327**, 79–92 (2018)
9. S. Nemati, P. Lima, S. Sedaghat, An effective numerical method for solving fractional pantograph differential equations using modification of hat functions. Appl. Numer. Math. **131**, 174–189 (2018)
10. S. Nemati, P. Lima, D.F.M Torres, A numerical approach for solving fractional optimal control problems using modified hat functions. Commun. Nonlinear Sci. Numer. Simulat. **78**, 104849 (2019)
11. S. Nemati, D.F.M Torres, A new spectral method based on two classes of hat functions for solving systems of fractional differential equations and an application to respiratory syncytial virus infection. Soft Comput. (2020). https://doi.org/10.1007/s00500-019-04645-5
12. Z. Odibat, S. Momani, Modified homotopy perturbation method: application to quadratic Riccati differential equation of fractional order. Chaos Solitons Fractals **36**, 167–174 (2008)

13. T. Odzijewicz, A.B. Malinowska, D.F.M. Torres, Fractional variational calculus of variable order, in *Advances in Harmonic Analysis and Operator Theory*. Operator Theory: Advances and Applications, vol. 229 (Birkhäuser, Basel, 2013)
14. P.W. Ostalczyk, P. Duch, D.W. Brzeziński, D. Sankowski, Order functions selection in the variable, fractional-order PID controller, in *Advances in Modelling and Control of Non-integer-Order Systems*. Lecture Notes in Electrical Engineering, vol. 320 (2015), pp. 159–170.
15. M.R. Rapaić, A. Pisano, Variable-order fractional operators for adaptive order and parameter estimation. IEEE Trans. Autom. Control **59(3)**, 798–803 (2014)
16. S.G. Samko, B. Ross, Integration and differentiation to a variable fractional order. Integral Transform. Spec. Funct. **1**, 277–300 (1993)
17. M.P. Tripathi, V.K. Baranwal, R.K. Pandey, O.P. Singh, A new numerical algorithm to solve fractional differential equations based on operational matrix of generalized hat functions. Commun. Nonlinear Sci. Numer. Simul. **18**, 1327–1340 (2013)

Langlands Reciprocity for C^*-Algebras

Igor V. Nikolaev

Abstract We introduce a C^*-algebra $\mathcal{A}_V$ of a variety V over the number field K and a C^*-algebra $\mathcal{A}_G$ of a reductive group G over the ring of adeles of K. Using Pimsner's Theorem, we construct an embedding $\mathcal{A}_V \hookrightarrow \mathcal{A}_G$, where V is a G-coherent variety, e.g. the Shimura variety of G. The embedding is an analog of the Langlands reciprocity for C^*-algebras. It follows from the K-theory of the inclusion $\mathcal{A}_V \subset \mathcal{A}_G$ that the Hasse-Weil L-function of V is a product of the automorphic L-functions corresponding to irreducible representations of the group G.

Keywords Langlands program · Serre C^*-algebra

Mathematics Subject Classification (2010) Primary 11F70; Secondary 46L85

1 Introduction

The Langlands conjectures say that all zeta functions are automorphic [9]. In this note we study (one of) the conjectures in terms of the C^*-algebras [5]. Namely, denote by $G(\mathbf{A}_K)$ a reductive group G over the ring of adeles $\mathbf{A}_K$ of a number field K and by $G(K)$ its discrete subgroup over K. The Banach algebra $L^1(G(K)\backslash G(\mathbf{A}_K))$ consists of the integrable complex-valued functions endowed with the operator norm. The addition of functions $f_1, f_2 \in L^1(G(K)\backslash G(\mathbf{A}_K))$ is defined pointwise and the multiplication is given by the convolution product:

$$(f_1 * f_2)(g) = \int_{G(K)\backslash G(\mathbf{A}_K)} f_1(gh^{-1}) f_2(h) dh. \tag{1.1}$$

I. V. Nikolaev (✉)
Department of Mathematics, St. John's University, New York, NY, USA

M. A. Bastos et al. (eds.), *Operator Theory, Functional Analysis and Applications*,
Operator Theory: Advances and Applications 282,
https://doi.org/10.1007/978-3-030-51945-2_26

Consider the enveloping C^*-algebra, $\mathcal{A}_G$, of the algebra $L^1(G(K)\backslash G(\mathbf{A}_K))$; we refer the reader to [5, Section 13.9] for details of this construction. The algebra $\mathcal{A}_G$ encodes all unitary irreducible representations of the locally compact group $G(\mathbf{A}_K)$ induced by $G(K)$. Such representations are related to the automorphic cusp forms and non-abelian class field theory [8]. The algebra $\mathcal{A}_G$ has an amazingly simple structure. Namely, let us assume $G \cong GL_n$. Then $\mathcal{A}_G$ is a stationary *AF-algebra*, see Lemma 3.1; such an algebra is defined by a positive integer matrix $B \in SL_n(\mathbf{Z})$ [2] and its K-theory is well understood [6].

Let V be a complex projective variety. For an automorphism $\sigma : V \to V$ and an invertible sheaf $\mathcal{L}$ of the linear forms on V, one can construct a twisted homogeneous coordinate ring $B(V, \mathcal{L}, \sigma)$ of the variety V, i.e. a non-commutative r ing such that:

$$Mod\ (B(V, \mathcal{L}, \sigma))\ /\ Tors \cong\ Coh\ (V),$$

where Mod is the category of graded left modules over the graded ring $B(V, \mathcal{L}, \sigma)$, $Tors$ the full subcategory of Mod of the torsion modules and Coh the category of quasi-coherent sheaves on the variety V [18, p. 180]. The norm-closure of a self-adjoint representation of the ring $B(V, \mathcal{L}, \sigma)$ by the linear operators on a Hilbert space is called the *Serre C^*-algebra* of V [10]. In what follows, we shall focus on the case when V is defined over a number field K, i.e. V is an arithmetic variety. The corresponding Serre C^*-algebra is denoted by $\mathcal{A}_V$. The Hasse-Weil L-function of V was calculated in [10] in terms of the K-theory of algebra $\mathcal{A}_V$.

It is known that the Langlands philosophy does not distinguish between the arithmetic and automorphic objects [9]. Therefore one can expect a regular map between the C^*-algebras $\mathcal{A}_V$ and $\mathcal{A}_G$, provided V is a G-coherent variety, see Definition 1.1. We prove that such a map is an embedding $\mathcal{A}_V \hookrightarrow \mathcal{A}_G$. To give an exact statement, we shall need the following notions. The i-th *trace cohomology* $\{H^i_{tr}(V) \mid 0 \leq i \leq 2\dim_{\mathbf{C}} V\}$ of an arithmetic variety V is an additive abelian subgroup of $\mathbf{R}$ obtained from a canonical trace on the Serre C^*-algebra of V [10]. Likewise, the group $K_0(\mathcal{A}_G)$ of the stationary AF-algebra $\mathcal{A}_G$ is an additive abelian subgroup of $\mathbf{R}$ [6, Chapter 6].

Definition 1.1 The arithmetic variety V is called G-coherent, if

$$H^i_{tr}(V) \subseteq K_0(\mathcal{A}_G) \quad \text{for all} \quad 0 \leq i \leq 2\dim_{\mathbf{C}} V. \tag{1.2}$$

Remark 1.2 If $V \cong Sh\ (G, X)$ is the Shimura variety corresponding to the Shimura datum (G, X) [4], then V is a G-coherent variety. This remark follows from an adaption of the argument for the Shimura curves considered in Sect. 4. To put it simple, the arithmetic variety V is G-coherent if for all $0 \leq i \leq 2\dim_{\mathbf{C}} V$ the number fields $\mathbf{k}_i := H^i_{tr}(V) \otimes \mathbf{Q}$ are subfields of (or coincide with) a number field $\mathbf{K} := K_0(\mathcal{A}_G) \otimes \mathbf{Q}$. A quick example are elliptic curves with complex multiplication, see Proposition 4.2.

Theorem 1.3 *There exists a canonical embedding $\mathcal{A}_V \hookrightarrow \mathcal{A}_G$, where V is a G-coherent variety.*

Remark 1.4 Theorem 1.3 can be viewed as an analog of the Langlands reciprocity for C^*-algebras. In other words, the coordinate ring $\mathcal{A}_V$ of a G-coherent variety V is a sub-algebra of the algebra $\mathcal{A}_G$.

An application of Theorem 1.3 is as follows. Recall that to each arithmetic variety V one can attach the Hasse-Weil (motivic) L-function. Likewise, to each irreducible representation of the group $G(\mathbf{A}_K)$ one can attach an automorphic (standard) L-function, see [8] and [9]. Theorem 1.3 implies one of the conjectures of [9].

Corollary 1.5 *The Hasse-Weil L-function of a G-coherent variety V is a product of the automorphic L-functions.*

The paper is organized as follows. The definitions and preliminary results can be found in Sect. 2. Theorem 1.3 and Corollary 1.5 are proved in Sect. 3. An example is constructed in Sect. 4.

2 Preliminaries

This section is a brief account of preliminary facts involved in our paper; we refer the reader to [2, 5, 9, 18].

2.1 AF-Algebras

A *C^*-algebra* is an algebra A over $\mathbf{C}$ with a norm $a \mapsto \|a\|$ and an involution $a \mapsto a^*$ such that it is complete with respect to the norm and $\|ab\| \leq \|a\| \, \|b\|$ and $\|a^*a\| = \|a^2\|$ for all $a, b \in A$. Any commutative C^*-algebra is isomorphic to the algebra $C_0(X)$ of continuous complex-valued functions on some locally compact Hausdorff space X; otherwise, A represents a noncommutative topological space.

An *AF-algebra* (Approximately Finite C^*-algebra) is defined to be the norm closure of an ascending sequence of finite dimensional C^*-algebras M_n, where M_n is the C^*-algebra of the $n \times n$ matrices with entries in $\mathbf{C}$. Here the index $n = (n_1, \ldots, n_k)$ represents the semi-simple matrix algebra $M_n = M_{n_1} \oplus \cdots \oplus M_{n_k}$. The ascending sequence mentioned above can be written as

$$M_1 \xrightarrow{\varphi_1} M_2 \xrightarrow{\varphi_2} \ldots,$$

where M_i are the finite dimensional C^*-algebras and φ_i the homomorphisms between such algebras. If $\varphi_i = Const$, then the AF-algebra $\mathcal{A}$ is called *stationary*; such an algebra defines and is defined by a *shift automorphism* $\sigma_\varphi : \mathcal{A} \to \mathcal{A}$

corresponding to a map $i \mapsto i+1$ on φ_i [6, p. 37]. The homomorphisms φ_i can be arranged into a graph as follows. Let $M_i = M_{i_1} \oplus \cdots \oplus M_{i_k}$ and $M_{i'} = M_{i'_1} \oplus \cdots \oplus M_{i'_k}$ be the semi-simple C^*-algebras and $\varphi_i : M_i \to M_{i'}$ the homomorphism. One has two sets of vertices $V_{i_1}, \ldots, V_{i_k}$ and $V_{i'_1}, \ldots, V_{i'_k}$ joined by b_{rs} edges whenever the summand M_{i_r} contains b_{rs} copies of the summand $M_{i'_s}$ under the embedding φ_i. As i varies, one obtains an infinite graph called the *Bratteli diagram* of the AF-algebra. The matrix $B = (b_{rs})$ is known as a *partial multiplicity* matrix; an infinite sequence of B_i defines a unique AF-algebra.

For a unital C^*-algebra A, let $V(A)$ be the union (over n) of projections in the $n \times n$ matrix C^*-algebra with entries in A; projections $p, q \in V(A)$ are *equivalent* if there exists a partial isometry u such that $p = u^*u$ and $q = uu^*$. The equivalence class of projection p is denoted by $[p]$; the equivalence classes of orthogonal projections can be made to a semigroup by putting $[p]+[q] = [p+q]$. The Grothendieck completion of this semigroup to an abelian group is called the K_0-group of the algebra A. The functor $A \to K_0(A)$ maps the category of unital C^*-algebras into the category of abelian groups, so that projections in the algebra A correspond to a positive cone $K_0^+ \subset K_0(A)$ and the unit element $1 \in A$ corresponds to an order unit $u \in K_0(A)$. The ordered abelian group (K_0, K_0^+, u) with an order unit is called a *dimension group*; an order-isomorphism class of the latter we denote by (G, G^+).

If $\mathcal{A}$ is an AF-algebra, then its dimension group $(K_0(\mathcal{A}), K_0^+(\mathcal{A}), u)$ is a complete isomorphism invariant of algebra $\mathcal{A}$ [7]. The order-isomorphism class $(K_0(\mathcal{A}), K_0^+(\mathcal{A}))$ is an invariant of the *Morita equivalence* of algebra $\mathcal{A}$, i.e. an isomorphism class in the category of finitely generated projective modules over $\mathcal{A}$.

2.2 Trace Cohomology

Let V be an n-dimensional complex projective variety endowed with an automorphism $\sigma : V \to V$ and denote by $B(V, \mathcal{L}, \sigma)$ its twisted homogeneous coordinate ring, see [18]. Let R be a commutative graded ring, such that $V = Spec\,(R)$. Denote by $R[t, t^{-1}; \sigma]$ the ring of skew Laurent polynomials defined by the commutation relation $b^\sigma t = tb$ for all $b \in R$, where b^σ is the image of b under automorphism σ. It is known, that $R[t, t^{-1}; \sigma] \cong B(V, \mathcal{L}, \sigma)$.

Let $\mathcal{H}$ be a Hilbert space and $\mathcal{B}(\mathcal{H})$ the algebra of all bounded linear operators on $\mathcal{H}$. For a ring of skew Laurent polynomials $R[t, t^{-1}; \sigma]$, consider a homomorphism:

$$\rho : R[t, t^{-1}; \sigma] \longrightarrow \mathcal{B}(\mathcal{H}). \tag{2.1}$$

Recall that $\mathcal{B}(\mathcal{H})$ is endowed with a $*$-involution; the involution comes from the scalar product on the Hilbert space $\mathcal{H}$. We shall call representation (2.1) $*$-coherent, if (i) $\rho(t)$ and $\rho(t^{-1})$ are unitary operators, such that $\rho^*(t) = \rho(t^{-1})$ and (ii) for all $b \in R$ it holds $(\rho^*(b))^{\sigma(\rho)} = \rho^*(b^\sigma)$, where $\sigma(\rho)$ is an automorphism of $\rho(R)$ induced by σ. Whenever $B = R[t, t^{-1}; \sigma]$ admits a $*$-coherent representation, $\rho(B)$

is a $*$-algebra; the norm closure of $\rho(B)$ is a C^*-algebra [5]. We shall denote it by $\mathcal{A}_V$ and refer to $\mathcal{A}_V$ as the *Serre* C^**-algebra* of variety V.

Let $\mathcal{K}$ be the C^*-algebra of all compact operators on $\mathcal{H}$. We shall write $\tau : \mathcal{A}_V \otimes \mathcal{K} \to \mathbf{R}$ to denote the canonical normalized trace on $\mathcal{A}_V \otimes \mathcal{K}$, i.e. a positive linear functional of norm 1 such that $\tau(yx) = \tau(xy)$ for all $x, y \in \mathcal{A}_V \otimes \mathcal{K}$, see [1, p. 31]. Denote by $C(V)$ the C^*-algebra of complex-valued functions on the Hausdorff space V. Because $\mathcal{A}_V$ is a crossed product C^*-algebra of the form $\mathcal{A}_V \cong C(V) \rtimes \mathbf{Z}$ [12, Lemma 5.3.2], one can use the Pimsner-Voiculescu six term exact sequence for the crossed products, see e.g. [1, p. 83] for the details. Thus one gets the short exact sequence of the algebraic K-groups: $0 \to K_0(C(V)) \xrightarrow{i_*} K_0(\mathcal{A}_V) \to K_1(C(V)) \to 0$, where the map i_* is induced by the natural embedding of $C(V)$ into $\mathcal{A}_V$. We have $K_0(C(V)) \cong K^0(V)$ and $K_1(C(V)) \cong K^{-1}(V)$, where K^0 and K^{-1} are the topological K-groups of V, see [1, p.80]. By the Chern character formula, one gets $K^0(V) \otimes \mathbf{Q} \cong H^{even}(V; \mathbf{Q})$ and $K^{-1}(V) \otimes \mathbf{Q} \cong H^{odd}(V; \mathbf{Q})$, where H^{even} (H^{odd}) is the direct sum of even (odd, resp.) cohomology groups of V. Notice that $K_0(\mathcal{A}_V \otimes \mathcal{K}) \cong K_0(\mathcal{A}_V)$ because of a stability of the K_0-group with respect to tensor products by the algebra $\mathcal{K}$, see e.g. [1, p. 32]. One gets the commutative diagram in Fig. 1, where τ_* denotes a homomorphism induced on K_0 by the canonical trace τ on the C^*-algebra $\mathcal{A}_V \otimes \mathcal{K}$. Since $H^{even}(V) := \oplus_{i=0}^{n} H^{2i}(V)$ and $H^{odd}(V) := \oplus_{i=1}^{n} H^{2i-1}(V)$, one gets for each $0 \leq i \leq 2n$ an injective homomorphism $\tau_* : H^i(V) \longrightarrow \mathbf{R}$.

Definition 2.1 By an i-th trace cohomology group $H^i_{tr}(V)$ of variety V one understands the abelian subgroup of $\mathbf{R}$ defined by the map τ_*.

2.3 *Langlands Reciprocity*

Let V be an n-dimensional complex projective variety over a number field K; consider its reduction $V(\mathbf{F}_p)$ modulo the prime ideal $\mathfrak{P} \subset K$ over a non-ramified prime p. Recall that the *Weil* zeta function is defined as:

$$Z_p(t) = \exp\left(\sum_{r=1}^{\infty} |V(\mathbf{F}_{p^r})| \frac{t^r}{r}\right), \quad r \in \mathbf{C}, \tag{2.2}$$

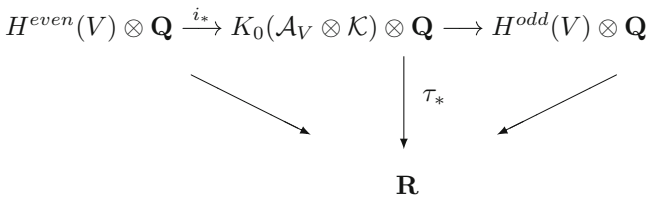


Fig. 1 The trace cohomology

where $|V(\mathbf{F}_{p^r})|$ is the number of points of variety $V(\mathbf{F}_{p^r})$ defined over the field with p^r elements. It is known that:

$$Z_p(t) = \frac{P_1(t)\dots P_{2n-1}(t)}{P_0(t)\dots P_{2n}(t)}, \tag{2.3}$$

where $P_0(t) = 1 - t$, $P_{2n} = 1 - p^n t$ and each $P_i(t)$ for $1 \leq i \leq 2n-1$ is a polynomial with integer coefficients, such that $P_i(t) = \prod(1 - \alpha_{ij}t)$ for some algebraic integers α_{ij} of the absolute value $p^{\frac{i}{2}}$. Consider an infinite product:

$$L(s,V) := \prod_p Z_p(p^{-s}) = \frac{L^1(s,V)\dots L^{2n-1}(s,V)}{L^0(s,V)\dots L^{2n}(s,V)}, \tag{2.4}$$

where $L^i(s,V) = \prod_p P_i(p^{-s})$; the $L(s,V)$ is called the *Hasse-Weil* (or motivic) L-function of V.

On the other hand, if K is a number field then the *adele ring* $\mathbf{A}_K$ of K is a locally compact subring of the direct product $\prod K_v$ taken over all places v of K; the $\mathbf{A}_K$ is endowed with a canonical topology. Consider a reductive group $G(\mathbf{A}_K)$ over $\mathbf{A}_K$; the latter is a topological group with a canonical discrete subgroup $G(K)$. Denote by $L^2(G(K)\backslash G(\mathbf{A}_K))$ the Hilbert space of all square-integrable complex-valued functions on the homogeneous space $G(K) \setminus G(\mathbf{A}_K)$ and consider the right regular representation $\mathcal{R}$ of the locally compact group $G(\mathbf{A}_K)$ by linear operators on the space $L^2(G(K)\backslash G(\mathbf{A}_K))$ given by formula (1.1). It is well known, that each irreducible component π of the unitary representation $\mathcal{R}$ can be written in the form $\pi = \otimes \pi_v$, where v are all unramified places of K. Using the *spherical functions*, one gets an injection $\pi_v \mapsto [A_v]$, where $[A_v]$ is a conjugacy class of matrices in the group $GL_n(\mathbf{C})$. The *automorphic* L-function is given by the formula:

$$L(s,\pi) = \prod_v \left(\det\left[I_n - [A_v](Nv)^{-s}\right]\right)^{-1}, \quad s \in \mathbf{C}, \tag{2.5}$$

where Nv is the norm of place v; we refer the reader to [9, p. 170] and [8, p. 201] for details of this construction.

The following conjecture relates the Hasse-Weil and automorphic L-functions.

Conjecture 2.2 ([9]) *For each $0 \leq i \leq 2n$ there exists an irreducible representation π_i of the group $G(\mathbf{A}_K)$, such that $L^i(s,V) \equiv L(s,\pi_i)$.*

3 Proofs

3.1 Proof of Theorem 1.3

We shall split the proof in two lemmas.

Lemma 3.1 *The algebra $\mathcal{A}_G$ is isomorphic to a stationary AF-algebra.*

Proof Let $\mathbf{A}_K^\times$ be the idele group, i.e. a group of invertible elements of the adele ring $\mathbf{A}_K$. Denote by $Gal\ (K^{ab}|K)$ the Galois group of the maximal abelian extension K^{ab} of the number field K. The Artin reciprocity says that there exists a continuous isomorphism:

$$K^\times\backslash\mathbf{A}_K^\times/C_K \longrightarrow Gal\ (K^{ab}|K), \tag{3.1}$$

where C_K is the closure of the image in $K^\times\backslash\mathbf{A}_K^\times$ of the identity connected component of the archimedean part K^∞ of the $\mathbf{A}_K^\times$.

Recall that $Gal\ (K^{ab}|K)$ is a profinite abelian group, i.e. a topological group isomorphic to the inverse limit of finite abelian groups. It follows from the Artin reciprocity (3.1), that $K^\times\backslash\mathbf{A}_K^\times/C_K$ is also a profinite abelian group. Since every finite abelian group is a product of the cyclic groups $\mathbf{Z}/p_i^{k_i}\mathbf{Z}$, we can write the group $K^\times\backslash\mathbf{A}_K^\times/C_K$ in the form:

$$K^\times\backslash\mathbf{A}_K^\times/C_K \cong \varprojlim \prod_{i=1}^{l_m}\left(\mathbf{Z}/p_i^{k_i}\mathbf{Z}\right), \tag{3.2}$$

where $m \to \infty$. Notice that the cyclic group $\mathbf{Z}/p_i^{k_i}\mathbf{Z}$ can be embedded into the finite field $\mathbf{F}_{q_i}$, where $q_i = p_i^{k_i}$. Thus the group $GL_n(\mathbf{Z}/p_i^{k_i}\mathbf{Z})$ is correctly defined and from (3.2) one gets an isomorphism

$$GL_n(K^\times\backslash\mathbf{A}_K^\times/C_K) \cong \varprojlim \prod_{i=1}^{l_m} GL_n(\mathbf{F}_{q_i}), \tag{3.3}$$

where $GL_n(\mathbf{F}_{q_i})$ is a finite group of order $\prod_{j=0}^{n-1}(q_i^n - q_i^j)$ and such a group is no longer abelian. In particular, it follows from (3.3) that the $GL_n(K^\times\backslash\mathbf{A}_K^\times/C_K)$ is a profinite group.

(i) Let us show that the group $GL_n(K^\times\backslash\mathbf{A}_K^\times/C_K)$ being profinite implies that the $\mathcal{A}_G$ is an AF-algebra. Indeed, if G is a finite group then the group algebra $\mathbf{C}[G]$ has the form

$$\mathbf{C}[G] \cong M_{n_1}(\mathbf{C}) \oplus \cdots \oplus M_{n_h}(\mathbf{C}),$$

where n_i are degrees of the irreducible representations of G and h is the total number of such representations [16, Proposition 10]. In view of (3.3), we have

$$GL_n(K^\times \backslash \mathbf{A}_K^\times / C_K) \cong \varprojlim G_i, \tag{3.4}$$

where G_i is a finite group. Consider a group algebra

$$\mathbf{C}[G_i] \cong M_{n_1}^{(i)}(\mathbf{C}) \oplus \cdots \oplus M_{n_h}^{(i)}(\mathbf{C}) \tag{3.5}$$

corresponding to G_i. Notice that the $\mathbf{C}[G_i]$ is a finite-dimensional C^*-algebra. The inverse limit (3.4) defines an ascending sequence of the finite-dimensional C^*-algebras of the form

$$\varprojlim M_{n_1}^{(i)}(\mathbf{C}) \oplus \cdots \oplus M_{n_h}^{(i)}(\mathbf{C}). \tag{3.6}$$

Since $\mathcal{A}_G$ is the norm closure of the group algebra $\mathbf{C}[GL_n(K^\times \backslash \mathbf{A}_K^\times)] \cong \mathbf{C}[GL_n(K^\times) \backslash GL_n(\mathbf{A}_K^\times)]$ [5, Section 13.9], we conclude that there exists a C^*-homomorphism $h : \mathcal{A}_G \to \mathbb{A}_G$, where $\mathbb{A}_G$ is an AF-algebra defined by the limit (3.6). To calculate the kernel of h, recall that $C_K \cong \varprojlim U_i$, where U_i are open subgroups of the group $K^\times \backslash \mathbf{A}_K^\times$. We repeat the construction of (3.4)–(3.6) and obtain an AF-algebra $\mathbb{A}_U$. One gets an exact sequence of the C^*-algebras $1 \to \mathbb{A}_U \to \mathcal{A}_G \to \mathbb{A}_G \to 1$. In other words, the $\mathcal{A}_G$ is an extension of the AF-algebra $\mathbb{A}_U$ by the AF-algebra $\mathbb{A}_G$. But any such an extension must be an AF-algebra itself [3]. Item (i) is proved.

(ii) It remains to prove that the $\mathcal{A}_G$ is a stationary AF-algebra. Indeed, denote by Fr_q the Frobenius map, i.e. an endomorphism of the finite field $\mathbf{F}_q$ acting by the formula $x \mapsto x^q$. The map Fr_{q_i} induces an automorphism of the group $GL_n(\mathbf{F}_{q_i})$. Using formula (3.3), one gets an automorphism of the group $GL_n(\mathbf{A}_K)$ and the corresponding group algebra $\mathbf{C}[GL_n(\mathbf{A}_K)]$. Taking the norm closure of the algebra $\mathbf{C}[GL_n(\mathbf{A}_K)]$, we conclude that there exists a non-trivial automorphism ϕ of the AF-algebra $\mathcal{A}_G$. But the AF-algebra admits an automorphism $\phi \neq \pm Id$ if and only if it is a stationary AF-algebra [6, p. 37]. Thus the algebra $\mathcal{A}_G$ is a stationary AF-algebra. Lemma 3.1 is proved. □

Remark 3.2 It follows from formula (3.4) that the AF-algebra $\mathcal{A}_G$ is determined by a partial multiplicity matrix B of rank n, i.e. $B \in SL_n(\mathbf{Z})$. Consider an isomorphism

$$\mathcal{A}_G \rtimes \mathbf{Z} \cong \mathcal{O}_B \otimes \mathcal{K},$$

where the crossed product is taken by the shift automorphism of $\mathcal{A}_G$, $\mathcal{O}_B$ is the *Cuntz-Krieger* algebra defined by matrix B and $\mathcal{K}$ is the C^*-algebra of compact operators [1, Exercise 10.11.9]. Consider a continuous group of modular automorphisms $\{\sigma^t : \mathcal{O}_B \to \mathcal{O}_B \mid t \in \mathbf{R}\}$ acting on the generators $s_1, \ldots, s_n$ of the algebra $\mathcal{O}_B$ by the formula $s_k \mapsto e^{it} s_k$. Then a pull back of σ^t corresponds

to the action of continuous symmetry group $GL_n(\mathbf{A}_K)$ on the homogeneous space $GL_n(K)\backslash GL_n(\mathbf{A}_K)$. This observation can be applied to prove Weil's conjecture on the Tamagawa numbers.

Lemma 3.3 *The algebra $\mathcal{A}_V$ embeds into the AF-algebra $\mathcal{A}_G$, where V is a G-coherent variety.*

Proof We shall use the *Pimsner's Theorem* [13, Theorem 7] about an embedding of the crossed product algebra $\mathcal{A}_V$ into an AF-algebra. It will develop that the G-coherence of V implies that the AF-algebra is Morita equivalent to the algebra $\mathcal{A}_G$ of Lemma 3.1. We pass to a detailed argument.

Let V be a complex projective variety. Following [13] we shall think of V as a compact metrizable topological space X. Recall that for a homeomorphism $\varphi : X \to X$ the point $x \in X$ is called *non-wandering* if for each neighborhood U of x and every $N > 0$ there exists $n > N$ such that

$$\varphi^n(U) \cap U \neq \emptyset.$$

(In other words, the point x does not "wander" too far from its initial position in the space X.) If each point $x \in X$ is a non-wandering point, then the homeomorphism φ is called non-wandering.

Let $\sigma : V \to V$ be an automorphism of finite order of the G-coherent variety V, such that the representation (2.1) is $*$-coherent. Then the crossed product

$$\mathcal{A}_V = C(V) \rtimes_\sigma \mathbf{Z}$$

is the Serre C^*-algebra of V. Since σ is of a finite order, it is a non-wandering homeomorphism of X. In particular, the σ is a pseudo-non-wandering homeomorphism [13, Definition 2]. Then there exists a unital (dense) embedding

$$\mathcal{A}_V \hookrightarrow \mathcal{A}, \tag{3.7}$$

where $\mathcal{A}$ is an AF-algebra defined by the homeomorphism φ [13, Theorem 7].

Let us show that the algebra $\mathcal{A}$ is Morita equivalent to the AF-algebra $\mathcal{A}_G$. Indeed, the embedding (3.7) induces an injective homomorphism of the K_0-groups

$$K_0(\mathcal{A}_V) \hookrightarrow K_0(\mathcal{A}). \tag{3.8}$$

As explained in Sect. 2.2, the map (3.8) defines an inclusion

$$H^i_{tr}(V) \subseteq K_0(\mathcal{A}). \tag{3.9}$$

On the other hand, the trace cohomology of a G-coherent variety V must satisfy an inclusion

$$H^i_{tr}(V) \subseteq K_0(\mathcal{A}_G). \tag{3.10}$$

Let $b^* = \max_{0\le i\le 2n} b_i$ be the maximal Betti number of variety V. Then in formulas (3.9) and (3.10) the inclusion is an isomorphism, i.e. $H^*_{tr}(V) \cong K_0(\mathcal{A})$ and $H^*_{tr}(V) \cong K_0(\mathcal{A}_G)$. One concludes that

$$K_0(\mathcal{A}) \cong K_0(\mathcal{A}_G). \tag{3.11}$$

In other words, the AF-algebras $\mathcal{A}$ and $\mathcal{A}_G$ are Morita equivalent. The embedding $\mathcal{A}_V \hookrightarrow \mathcal{A}_G$ follows from formulas (3.7) and (3.11). Lemma 3.3 is proved. □

Theorem 1.3 follows from Lemma 3.3.

3.2 *Proof of Corollary 1.5*

Corollary 1.5 follows from an observation that the Frobenius action $\sigma(Fr_p^i) : H^i_{tr}(V) \to H^i_{tr}(V)$ extends to a Hecke operator $T_p : K_0(\mathcal{A}_G) \to K_0(\mathcal{A}_G)$, whenever $H^i_{tr}(\mathcal{A}_V) \subseteq K_0(\mathcal{A}_G)$. Let us pass to a detailed argument.

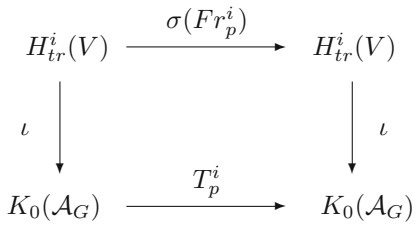
Recall that the Frobenius map on the i-th trace cohomology of variety V is given by an integer matrix $\sigma(Fr_p^i) \in GL_{b_i}(\mathbf{Z})$, where b_i is the i-th Betti number of V; moreover,

$$|V(\mathbf{F}_p)| = \sum_{i=0}^{2n} (-1)^i \, tr\, \sigma(Fr_p^i), \tag{3.12}$$

where $V(\mathbf{F}_p)$ is the reduction of V modulo a good prime p [10]. (Notice that (3.12) is sufficient to calculate the Hasse-Weil L-function $L(s, V)$ of variety V via Eq. (2.2); hence the map $\sigma(Fr_p^i) : H^i_{tr}(V) \to H^i_{tr}(V)$ is motivic.)

Definition 3.4 Denote by T_p^i an endomorphism of $K_0(\mathcal{A}_G)$, such that the diagram in Fig. 2 is commutative, where ι is the embedding (1.2). By $\mathfrak{H}_i$ we understand the algebra over $\mathbf{Z}$ generated by the $T_p^i \in End\ (K_0(\mathcal{A}_G))$, where p runs through all but a finite set of primes.

Fig. 2 Hecke operator T_p^i



Remark 3.5 The algebra $\mathfrak{H}_i$ is commutative. Indeed, the endomorphisms T_p^i correspond to multiplication of the group $K_0(\mathcal{A}_G)$ by the real numbers; the latter commute with each other. We shall call the $\{\mathfrak{H}_i \mid 0 \leq i \leq 2n\}$ an i-th *Hecke algebra*.

Lemma 3.6 *The algebra $\mathfrak{H}_i$ defines an irreducible representations π_i of the group $G(\mathbf{A}_K)$.*

Proof Let $f \in L^2(G(K)\backslash G(\mathbf{A}_K))$ be an eigenfunction of the Hecke operators T_p^i; in other words, the Fourier coefficients c_p of the function f coincide with the eigenvalues of the Hecke operators T_p up to a scalar multiple. Such an eigenfunction is defined uniquely by the algebra $\mathfrak{H}_i$.

Let $\mathcal{L}_f \subset L^2(G(K)\backslash G(\mathbf{A}_K))$ be a subspace generated by the right translates of f by the elements of the locally compact group $G(\mathbf{A}_K)$. It is immediate (see e.g. [8, Example on p. 197]), that $\mathcal{L}_f$ is an irreducible subspace of the space $L^2(G(K)\backslash G(\mathbf{A}_K))$; therefore it gives rise to an irreducible representation π_i of the locally compact group $G(\mathbf{A}_K)$. Lemma 3.6 follows. □

Lemma 3.7 $L(s, \pi_i) \equiv L^i(s, V)$.

Proof Recall that the function $L^i(s, V)$ can be written as

$$L^i(s, V) = \prod_p \left(\det\left[I_n - \sigma(Fr_p^i)p^{-s}\right]\right)^{-1}, \tag{3.13}$$

where $\sigma(Fr_p^i) \in GL_{b_i}(\mathbf{Z})$ is a matrix form of the action of Fr_p^i on the trace cohomology $H_{tr}^i(V)$.

On the other hand, from (2.5) one gets

$$L(s, \pi_i) = \prod_p \left(\det\left[I_n - [A_p^i]p^{-s}\right]\right)^{-1}, \tag{3.14}$$

where $[A_p^i] \subset GL_n(\mathbf{C})$ is a conjugacy class of matrices corresponding to the irreducible representation π_i of the group $G(\mathbf{A}_K)$. As explained, for such a representation we have an inclusion $T_p^i \in [A_p^i]$. But the action of the Hecke operator T_p^i is an extension of the action of $\sigma(Fr_p^i)$ on $H_{tr}^i(V)$, see Fig. 2. Therefore

$$\sigma(Fr_p^i) = [A_p^i] \tag{3.15}$$

for all but a finite set of primes p. Comparing formulas (3.13)–(3.15), we get that $L(s, \pi_i) \equiv L^i(s, V)$. Lemma 3.7 follows. □

Corollary 1.5 follows from Lemma 3.7 and formula (2.4).

4 Example

We shall illustrate Theorem 1.3 and Corollary 1.5 for the group

$$G \cong GL_2(\mathbf{A}_K),$$

where $K = \mathbf{Q}(\sqrt{D})$ is a real quadratic field.

Proposition 4.1 *$K_0(\mathcal{A}_G) \cong \mathbf{Z} + \mathbf{Z}\omega$, where*

$$\omega = \begin{cases} \frac{1+\sqrt{D}}{2}, & \text{if } D \equiv 1 \ mod\ 4, \\ \sqrt{D}, & \text{if } D \equiv 2, 3 \ mod\ 4. \end{cases} \tag{4.1}$$

Proof By Lemma 3.1 and Remark 3.2, the $\mathcal{A}_G$ is a stationary AF-algebra given by partial multiplicity matrix $B \in SL_2(\mathbf{Z})$. In particular, $K_0(\mathcal{A}_G) \cong \mathbf{Z} + \mathbf{Z}\omega$, where $\omega \in \mathbf{Q}(\lambda_B)$, where λ_B is the Perron-Frobenius eigenvalue of matrix B. Moreover, by the construction $End\ (K) \cong End\ (K_0(\mathcal{A}_G))$, where End is the endomorphism ring of the corresponding $\mathbf{Z}$-module. But $End\ (K) \cong O_K$, where O_K is the ring of integers of K. Thus, $\lambda_B \in K$ and ω is given by formula (4.1). Proposition 4.1 follows. □

Proposition 4.2 *Let $\mathcal{E}_{CM} \cong \mathbf{C}/O_k$ be an elliptic curve with complex multiplication by the ring of integers of the imaginary quadratic field $k = \mathbf{Q}(\sqrt{-D})$. Then $\mathcal{E}_{CM}$ is a G-coherent variety of the group $G \cong GL_2(\mathbf{A}_K)$.*

Proof The noncommutative torus $\mathcal{A}_\theta$ is a C^*-algebra generated by the unitary operators u and v satisfying the commutation relation $vu = e^{2\pi i\theta}uv$ for a constant $\theta \in \mathbf{R}$ [15]. The Serre C^*-algebra of an elliptic curve $\mathcal{E}_\tau \cong \mathbf{C}/(\mathbf{Z} + \mathbf{Z}\tau)$ is isomorphic to $\mathcal{A}_\theta$ for any $\{\tau \mid Im\ \tau > 0\}$, see [12, Theorem 1.3.1]. In particular [11], if $\tau \in O_k$ then

$$\begin{cases} H^0_{tr}(\mathcal{E}_{CM}) & = H^2_{tr}(\mathcal{E}_{CM}) \cong \mathbf{Z}, \\ H^1_{tr}(\mathcal{E}_{CM}) & \cong \mathbf{Z} + \mathbf{Z}\omega. \end{cases} \tag{4.2}$$

Comparing formulas (4.1) and (4.2), one concludes that

$$H^i_{tr}(\mathcal{E}_{CM}) \subseteq K_0(\mathcal{A}_G),$$

i.e. the $\mathcal{E}_{CM}$ is a G-coherent variety of the group $G \cong GL_2(\mathbf{A}_K)$. Proposition 4.2 is proved. □

Remark 4.3 The embedding of $\mathcal{A}_\theta$ into an AF-algebra was initially constructed in [14].

Proposition 4.4 *$L(s, \mathcal{E}_{CM}) \equiv \frac{L(s,\ \pi_1)}{L(s,\ \pi_0)L(s,\ \pi_2)}$, where π_i are irreducible representations of the locally compact group $GL_2(\mathbf{A}_K)$.*

Proof The Hasse-Weil L-function of the $\mathcal{E}_{CM}$ has the form:

$$L(s, \mathcal{E}_{CM}) = \frac{\prod_p \left[\det\ (I_2 - \sigma(Fr_p^1)p^{-s}\right]^{-1}}{\zeta(s)\zeta(s-1)}, \qquad s \in \mathbf{C}, \tag{4.3}$$

where $\zeta(s)$ is the Riemann zeta function and the product is taken over the set of good primes; we refer the reader to formula (3.13). It is immediate that

$$\begin{cases} L(s, \pi_0) = \zeta(s), \\ L(s, \pi_2) = \zeta(s-1), \end{cases}$$

where $L(s, \pi_0)$ and $L(s, \pi_2)$ are the automorphic L-functions corresponding to the irreducible representations π_0 and π_2 of the group $GL_2(\mathbf{A}_K)$. An irreducible representation π_1 gives rise to an automorphic L-function

$$L(s, \pi_1) = \prod_p \left(\det\left[I_2 - [A_p^1]p^{-s}\right]\right)^{-1}.$$

But formula (3.15) says that $[A_p^1] = \sigma(Fr_p^1)$ and therefore the numerator of (4.3) coincides with the $L(s, \pi_1)$. Proposition 4.4 is proved. □

Remark 4.5 Proposition 4.4 can be proved in terms of the Grössencharacters [17, Chapter II, §10].

Acknowledgments I thank the referees for their interest and helpful comments on the draft of this paper.

References

1. B. Blackadar, *K-Theory for Operator Algebras* (MSRI Publications, Springer, 1986)
2. O. Bratteli, Inductive limits of finite dimensional C^*-algebras. Trans. Amer. Math. Soc. **171**, 195–234 (1972)
3. L.G. Brown, Extensions of AF algebras: the projection lifting problem, in *Operator Algebras and Applications, Proceedings of Symposia in Pure Mathematics*, vol. 38 (1982), pp. 175–176
4. P. Deligne, *Travaux de Shimura*, vol. 244. Séminaire Bourbaki, Lecture Notes in Mathematics (Springer, Berlin, 1971), pp. 123–165
5. J. Dixmier, *C^*-Algebras* (North-Holland Publishing Company, Amsterdam, 1977)
6. E.G. Effros, Dimensions and C^*-algebras, in *Board of the Mathematical Sciences, Regional Conference Series in Mathematics*, vol. 46 (AMS, Providence, 1981)
7. G.A. Elliott, On the classification of inductive limits of sequences of semisimple finite-dimensional algebras. J. Algebra **38**, 29–44 (1976)
8. S. Gelbart, An elementary introduction to the Langlands program. Bull. Amer. Math. Soc. **10**, 177–219 (1984)

9. R.P. Langlands, L-functions and automorphic representations, in *Proceedings of the ICM 1978, Helsinki* (1978), pp. 165–175
10. I.V. Nikolaev, On traces of Frobenius endomorphisms. Finite Fields Appl. **25**, 270–279 (2014)
11. I.V. Nikolaev, On a symmetry of complex and real multiplication. Hokkaido Math. J. **45**, 43–51 (2016)
12. I.V. Nikolaev, *Noncommutative Geometry*. De Gruyter Studies in Mathematics, vol. 66 (De Gruyter, Berlin, 2017)
13. M.V. Pimsner, Embedding some transformation group C^*-algebras into AF-algebras. Ergodic Theory Dyn. Syst. **3**, 613–626 (1983)
14. M.V. Pimsner, D.V. Voiculescu, Imbedding the irrational rotation C^*-algebra into an AF-algebra. J. Oper. Theory **4**, 201–210 (1980)
15. M.A. Rieffel, Non-commutative tori – a case study of non-commutative differentiable manifolds. Contemp. Math. **105**, 191–211 (1990)
16. J.-P. Serre, *Représentations Linéaires des Groupes Finis* (Hermann, Paris, 1967)
17. J.H. Silverman, *Advanced Topics in the Arithmetic of Elliptic Curves*. GTM, vol. 151 (Springer, Berlin, 1994)
18. J.T. Stafford, M. van den Bergh, Noncommutative curves and noncommutative surfaces. Bull. Amer. Math. Soc. **38**, 171–216 (2001)

Compact Sequences in Quasifractal Algebras

Steffen Roch

Abstract The paper is devoted to the study of compact sequences in quasifractal algebras. We are particularly interested in the relations between the essential ranks of fractally restricted sequences $(A_n)|_{\mathbb{M}}$ and the essential rank of the full sequence. We will also ask whether the dependence of the essential ranks of the fractally restricted sequences $(A_n)|_{\mathbb{M}}$ on (a coset of) $\mathbb{M}$ is continuous, which requires to provide the fractal variety of the algebra with a suitable topology.

Keywords Fractal restriction · Quasifractal algebras · Continuous fields · Continuous trace algebras

Mathematics Subject Classification (2010) Primary 47N40; Secondary 65J10, 46L99

1 Introduction

Let H be a Hilbert space and $(P_n)_{n\in\mathbb{N}}$ a sequence of orthogonal projections of finite rank which converges strongly to the identity operator on H. Let $\mathcal{F}$ denote the set of all bounded sequences $(A_n)_{n\geq 1}$ of operators $A_n \in L(\operatorname{im} P_n)$ and $\mathcal{G}$ the set of all sequences $(A_n) \in \mathcal{F}$ with $\|A_n\| \to 0$. Provided with the operations

$$(A_n) + (B_n) := (A_n + B_n), \quad (A_n)(B_n) := (A_n B_n), \quad (A_n)^* := (A_n^*)$$

and the norm $\|(A_n)\| := \sup \|A_n\|$, $\mathcal{F}$ becomes a unital C^*-algebra and $\mathcal{G}$ a closed ideal of $\mathcal{F}$. Let $\delta(n)$ denote the rank of P_n. Then one can identify $L(\operatorname{im} P_n)$ with the C^*-algebra $\mathcal{C}_n := M_{\delta(n)}(\mathbb{C})$ of the complex $\delta(n) \times \delta(n)$ matrices, and $\mathcal{F}$ and $\mathcal{G}$ can be identified with the direct product and the direct sum of the sequence $(\mathcal{C}_n)$,

S. Roch (✉)
Department of Mathematics, Technical University of Darmstadt, Darmstadt, Germany
e-mail: roch@mathematik.tu-darmstadt.de

M. A. Bastos et al. (eds.), *Operator Theory, Functional Analysis and Applications*,
Operator Theory: Advances and Applications 282,
https://doi.org/10.1007/978-3-030-51945-2_27

respectively. In what follows we will think of $\mathcal{F}$ and $\mathcal{G}$ in this way (although some results hold for products and sums of sequences of general C^*-algebras as well).

The quotient algebra $\mathcal{F}/\mathcal{G}$ plays a significant role in numerical analysis, which stems from the observation that asymptotic properties of the sequence $(A_n) \in \mathcal{F}$ can often be rephrased as a property of the coset $(A_n) + \mathcal{G}$. To mention only two examples, it is

$$\|(A_n) + \mathcal{G}\|_{\mathcal{F}/\mathcal{G}} = \limsup_{n\to\infty} \|A_n\|$$

for every sequence $(A_n) \in \mathcal{F}$, and the coset $(A_n)+\mathcal{G}$ is invertible in $\mathcal{F}/\mathcal{G}$ if and only if the A_n are invertible for all sufficiently large n and if the norms of their inverses are uniformly bounded, which is equivalent to saying that (A_n) is a stable sequence.

Fractal and quasifractal algebras The correspondence between asymptotic properties of the sequence $(A_n) \in \mathcal{F}$ and properties of the coset $(A_n) + \mathcal{G}$ is particularly close when the sequence (A_n) belongs to a fractal subalgebra of $\mathcal{F}$. For example, if (A_n) is a sequence in a fractal subalgebra of $\mathcal{F}$, then the limit $\lim \|A_n\|$ exists, and the sequence (A_n) is stable if one of its (infinite) subsequences is stable.

The perhaps simplest way to define fractal algebras is the following (which is equivalent to the original definition in [15]):

> A C^*-subalgebra $\mathcal{A}$ of $\mathcal{F}$ is *fractal* if every partial zero sequence in $\mathcal{A}$ is a zero sequence, i.e., if every sequence $(A_n) \in \mathcal{A}$ with $\liminf \|A_n\| = 0$ satisfies $\lim \|A_n\| = 0$.

Not every subalgebra of $\mathcal{F}$ is fractal, but it turns out that every *separable* subalgebra of $\mathcal{F}$ has a *fractal restriction*. Here is the precise statement of the *fractal restriction theorem* from [9] (see [12] for a shorter proof):

> If $\mathcal{A}$ is a separable C^*-subalgebra of $\mathcal{F}$, then there is an infinite subset $\mathbb{M}$ of $\mathbb{N}$ such that the algebra $\mathcal{A}|_{\mathbb{M}}$ of all restricted sequences $(A_n)_{n\in\mathbb{M}}$ is a fractal subalgebra of $\mathcal{F}|_{\mathbb{M}}$.

Since every restriction of a separable subalgebra of $\mathcal{F}$ is separable again, the separable subalgebras of $\mathcal{F}$ are *quasifractal* in the sense of the following definition:

> A C^*-subalgebra $\mathcal{A}$ of $\mathcal{F}$ is called *quasifractal* if every restriction of $\mathcal{A}$ has a fractal restriction.

We denote by $\operatorname{fr}\mathcal{A}$ the set of all infinite subsets $\mathbb{M}$ of $\mathbb{N}$ for which $\mathcal{A}|_{\mathbb{M}}$ is a fractal algebra. Two elements $\mathbb{M}_1$, $\mathbb{M}_2$ of $\operatorname{fr}\mathcal{A}$ are called *equivalent*, symbolically written as $\mathbb{M}_1 \sim \mathbb{M}_2$, if $\mathbb{M}_1 \cup \mathbb{M}_2 \in \operatorname{fr}\mathcal{A}$. The relation $\sim$ is an equivalence relation (Lemma 4.2 in [13]). We denote the set of all equivalence classes of this relation by $(\operatorname{fr}\mathcal{A})^{\sim}$ and call $(\operatorname{fr}\mathcal{A})^{\sim}$ the *fractal variety* of the algebra $\mathcal{A}$. For example, if $\mathcal{A}$ is fractal, then $\operatorname{fr}\mathcal{A}$ consists of all infinite subsets of $\mathbb{N}$ and $(\operatorname{fr}\mathcal{A})^{\sim}$ is a singleton. A standard example for a fractal algebra is the algebra of the finite sections method for Toeplitz operators with continuous generating function (see [2] for Toeplitz operators and their finite sections and Corollary 1.10 in [5] for the fractality result). Examples for quasifractal algebras can be found in [13].

Compact sequences We will need several notions of the compactness and of the rank of an element in a general C^*-algebra and, in particular, of a sequence in $\mathcal{F}$. These notions can be subsumed under the following simple scheme.

Let $\mathcal{A}$ be a unital C^*-algebra and $\mathcal{J}$ a non-empty self-adjoint subset of $\mathcal{A}$ such that $\mathcal{A}\mathcal{J} \subseteq \mathcal{J}$ and $\mathcal{J}\mathcal{A} \subseteq \mathcal{J}$. We call the subsets with these properties the *semi-ideals* of $\mathcal{A}$. Clearly, $0 \in \mathcal{J}$. Every semi-ideal $\mathcal{J}$ induces a *rank function* on $\mathcal{A}$, i.e., a function $r : \mathcal{A} \to \mathbb{N} \cup \{\infty\}$ which satisfies

(a) $r(a) = 0$ if and only if $a = 0$,
(b) $r(a + b) \leq r(a) + r(b)$,
(c) $r(ab) \leq \min\{r(a),\ r(b)\}$,
(d) $r(a) = r(a^*)$

for all $a,\ b \in \mathcal{A}$, as follows: Set $r(0) := 0$. If a nonzero $a \in \mathcal{A}$ is a finite sum of elements of $\mathcal{J}$, then $r(a)$ is the smallest positive integer such that a can be written as a sum of $r(a)$ elements from $\mathcal{J}$. Finally, $r(a) := \infty$ if a is not a finite sum of elements of $\mathcal{J}$. The closure of the set of the elements with finite rank is a closed ideal of $\mathcal{A}$, called the ideal of the compact elements (relative to $\mathcal{J}$). We will reify this scheme in several settings:

1. Let $\mathcal{A}$ be a unital C^*-algebra and $\mathcal{J}$ the set of all elements $k \in \mathcal{A}$ with the property that for every $a \in \mathcal{A}$ there is a number $\alpha \in \mathbb{C}$ such that $kak = \alpha k$. Then $\mathcal{J}$ is a semi-ideal; the associated rank function is called the *algebraic rank* and denoted by $\operatorname{alg\,rank} a$, and the associated ideal of the compact elements is denoted by $\mathcal{C}(\mathcal{A})$.
2. For $\mathcal{A} = \mathcal{F}$, consider the set $\mathcal{J}$ of all sequences $(K_n) \in \mathcal{F}$ such that $\operatorname{rank} K_n \leq 1$ for all $n \in \mathbb{N}$. Then $\mathcal{J}$ is a semi-ideal of $\mathcal{F}$; the associated rank function is called the *sequential rank* and denoted by $\operatorname{seq\,rank}(A_n)$, and the associated ideal of the compact elements is denoted by $\mathcal{K}$. The elements of $\mathcal{K}$ are called the compact sequences. It is not hard to show that $\mathcal{G} \subseteq \mathcal{K}$ and that $\operatorname{seq\,rank}(A_n) = \sup_n \operatorname{rank} A_n$ for every sequence $(A_n) \in \mathcal{F}$.
3. If $\mathcal{A}$ is a C^*-subalgebra of $\mathcal{F}$ and $\mathcal{J}$ is the set of all sequences (K_n) *in* $\mathcal{A}$ such that $\operatorname{rank} K_n \leq 1$ for all $n \in \mathbb{N}$, then we denote the corresponding ideal of the compact sequences by $\mathcal{K}(\mathcal{A})$. Clearly, $\mathcal{K} = \mathcal{K}(\mathcal{F})$.
4. Let $\mathcal{J}$ be the set of all cosets $(K_n) + \mathcal{G}$ of sequences $(K_n) \in \mathcal{F}$ with $\operatorname{seq\,rank}(K_n) \leq 1$. Then $\mathcal{J}$ is a semi-ideal of $\mathcal{F}/\mathcal{G}$ and the corresponding ideal of the compact elements is nothing but $\mathcal{K}/\mathcal{G}$. If r denotes the rank function associated with $\mathcal{J}$ then we call $\operatorname{ess\,rank}(A_n) := r((A_n)+\mathcal{G})$ the *essential rank* of $(A_n) \in \mathcal{F}$. In particular, the sequences of essential rank 0 are just the sequences in $\mathcal{G}$. The essential rank of a sequence $(A_n) \in \mathcal{F}$ can be characterized as the smallest integer $r \geq 0$ such that (A_n) can be written as $(G_n) + (K_n)$ with $(G_n) \in \mathcal{G}$ and $\sup_n \operatorname{rank} K_n = r$. The advantage of the essential rank of a sequence (A_n) over its sequential rank is that it depends on the coset of (A_n) modulo $\mathcal{G}$ only.

See Sections 4.1–4.4 in [11] for some basic facts on compact sequences.

Compact Sequences in Fractal Algebras Let $\mathcal{A}$ be a unital and fractal C^*-algebra of $\mathcal{F}$. A crucial result ([10]) on compact sequences in fractal algebras is that then

$$(\mathcal{A} \cap \mathcal{K})/\mathcal{G} = \mathcal{C}\big(\mathcal{A}/\mathcal{G}\big).$$

Consequently, then $(\mathcal{A} \cap \mathcal{K})/\mathcal{G}$ is a dual algebra, i.e., it is *-isomorphic to a direct sum of ideals $K(H_t)$ of compact operators on a Hilbert space H_t. There is also a relation between the algebraic rank of the coset $(A_n) + \mathcal{G}$ and the essential rank of (A_n) when (A_n) is a compact sequence in a fractal algebra $\mathcal{A}$. The general form of this relation involves the local weights of $\mathcal{A}$ which we are not going to introduce here. This relation takes a particular simple form when $\mathcal{A}$ has local weight 1, in which case

$$\operatorname{ess\,rank}(A_n) = \operatorname{alg\,rank}\big((A_n) + \mathcal{G}\big) \tag{1.1}$$

for every sequence $(A_n) \in \mathcal{A}$ of finite essential rank. One can show that a fractal algebra $\mathcal{A}$ has local weight 1 if and only if

$$\mathcal{A} \cap \mathcal{K} = \mathcal{K}(\mathcal{A}). \tag{1.2}$$

We therefore refer to (1.2) as the *local weight 1 condition* in what follows.

The contents of this paper The goal of this paper is to study compact sequences in quasifractal algebras. We are particularly interested in the relations of the essential ranks of the restricted sequences $(A_n)|_{\mathbb{M}}$ with $\mathbb{M} \in \operatorname{fr}\mathcal{A}$ to the essential rank of (A_n) on the one side and to the algebraic ranks of the cosets $(A_n)|_{\mathcal{M}} + \mathcal{G}|_{\mathbb{M}}$ on the other side. We will also ask whether the dependence of the essential ranks of $(A_n)|_{\mathbb{M}}$ on $\mathbb{M}^\sim$ is continuous with respect to the topology on $(\operatorname{fr}\mathcal{A})^\sim$ introduced in [13]. We will embed the latter question into a broader context and discuss to what extend $(\mathcal{A} \cap \mathcal{K})/\mathcal{G}$ can be viewed of as an algebra with continuous trace.

2 Restrictions of Compact Sequences

In what follows we will often use boldface letters to denote sequences in $\mathcal{F}$. We start with picking up the first question raised at the end of the introduction: to what extend do the compactness properties of the fractal restrictions of a sequence determine the compactness properties of the full sequence? Clearly, if $\mathbf{K} \in \mathcal{F}$ is a compact sequence and $\mathbb{M}$ an infinite subset of $\mathbb{N}$, then the restriction $\mathbf{K}|_{\mathbb{M}}$ is a compact sequence in $\mathcal{F}|_{\mathbb{M}}$, and

$$\operatorname{ess\,rank}(\mathbf{K}|_{\mathbb{M}}) \le \operatorname{ess\,rank}\mathbf{K}. \tag{2.1}$$

Trivially, the converse is also true: If *every* restriction of a sequence $\mathbf{K}$ is compact, then $\mathbf{K}$ is compact. Remarkably, already the *fractal* restrictions of $\mathbf{K}$ will do the job.

Proposition 2.1 *Let $\mathcal{A}$ be a quasifractal C^*-subalgebra of $\mathcal{F}$ and $\mathbf{K} \in \mathcal{A}$. Then*

(*a*) $\mathbf{K}$ *is compact if and only if every restriction* $\mathbf{K}|_{\mathbb{M}}$ *with* $\mathbb{M} \in \operatorname{fr} \mathcal{A}$ *is compact.*
(*b*) *if* $\mathbf{K}$ *is of finite essential rank,*

$$\operatorname{ess\,rank} \mathbf{K} = \max_{\mathbb{M} \in \operatorname{fr} \mathcal{A}} \operatorname{ess\,rank} (\mathbf{K}|_{\mathbb{M}}). \tag{2.2}$$

Proof We shall employ a criterion from [11] which characterizes the compactness of a sequence (K_n) in terms of the singular values of the matrices K_n. Thus, let $\Sigma_1(A) \geq \ldots \geq \Sigma_n(A) \geq 0$ denote the singular values of a matrix $A \in M_n(\mathbb{C})$. If $\mathbf{K} = (K_n)$ is a compact sequence, then every (fractal or not) restriction of $\mathbf{K}$ is compact. If $\mathbf{K}$ fails to be compact, then the negation of property (*a*) in Theorem 4.5 in [11] yields a positive constant C and strictly increasing sequences (n_r) and (k_r) with $n_r \geq k_r$ such that

$$\Sigma_{k_r}(K_{n_r}) \geq C \quad \text{for all } r.$$

The same theorem then implies that no restriction of the subsequence (K_{n_r}) of (K_n) is compact. This proves assertion (*a*).

Let now $\mathbf{K}$ be a sequence of finite essential rank r. The estimate $\geq$ in (2.2) comes from (2.1). For the reverse estimate, we have to show that there is a fractal restriction $\mathbf{K}|_{\mathbb{M}}$ of $\mathbf{K}$ of essential rank r. Let $C := \limsup \Sigma_r(K_n)$. By Corollary 4.6 in [11], C is positive, and the set

$$\mathbb{M}' := \{n \in \mathbb{N} : \Sigma_r(K_n) \geq C/2\}$$

is infinite. Since $\mathcal{A}$ is quasifractal, there is an infinite subset $\mathbb{M}$ of $\mathbb{M}'$ such that $\mathbf{K}|_{\mathbb{M}}$ is fractal. Moreover, by construction,

$$\limsup_{\mathbb{M} \ni m \to \infty} \Sigma_r(K_m) \geq C/2 > 0 \quad \text{and} \quad \lim_{\mathbb{M} \ni m \to \infty} \Sigma_{r+1}(K_m) = 0.$$

Thus, $\mathbf{K}|_{\mathbb{M}}$ is a sequence of finite essential rank r. □

It would be desirable to replace the essential ranks of the restrictions $\mathbf{K}|_{\mathbb{M}}$ on the right-hand side of (2.2) by a more intrinsic quantity such as the algebraic rank of the coset $\mathbf{K}|_{\mathbb{M}} + \mathcal{G}|_{\mathbb{M}}$. Since, in general, the equality $\operatorname{ess\,rank} \mathbf{K} = \operatorname{alg\,rank} (\mathbf{K} + \mathcal{G})$ does not even hold for a sequence $\mathbf{K}$ in a fractal algebra, we will need additional conditions to guarantee the equality

$$\operatorname{ess\,rank} \mathbf{K}|_{\mathbb{M}} = \operatorname{alg\,rank} (\mathbf{K}|_{\mathbb{M}} + \mathcal{G}|_{\mathbb{M}}) \tag{2.3}$$

for all $\mathbb{M} \in \operatorname{fr} \mathcal{A}$.

There are several conditions which ensure the equality (2.3). Our goal is to show (2.3) for sequences $\mathbf{K}$ which are stably regularizable. Recall that an element a of a C^*-algebra $\mathcal{A}$ is *Moore-Penrose invertible* if there is a $b \in \mathcal{A}$ such that $aba = a$ and $bab = b$ and such that ab and ba are self-adjoint. The element b is uniquely determined by these conditions; it is called the Moore-Penrose inverse of a and denoted by $a^\dagger$. All we need on Moore-Penrose invertibility is in Section 2.2.1 in [5]. A sequence $\mathbf{K} \in \mathcal{F}$ is then called *stably regularizable* if it is the sum of a Moore-Penrose invertible sequence and a sequence in $\mathcal{G}$. Stable regularizability of a sequence $\mathbf{K}$ is equivalent to the Moore-Penrose invertibility of its coset $\mathbf{K} + \mathcal{G}$ in $\mathcal{F}/\mathcal{G}$.

So let $\mathbf{K} = (K_n) \in \mathcal{A}$ be a sequence of finite essential rank r which is stably regularizable. Then, by definition, there are sequences $(L_n) \in \mathcal{F}$ and $(G_n) \in \mathcal{G}$ such that $(K_n) = (L_n) + (G_n)$ and (L_n) is Moore-Penrose invertible in $\mathcal{F}$ (and, if $\mathcal{G} \subseteq \mathcal{A}$, even in $\mathcal{A}$ by inverse closedness). In particular, every L_n is Moore-Penrose invertible and $\sup \|L_n^\dagger\| < \infty$, such that $(L_n)^\dagger = (L_n^\dagger)$. Clearly, (L_n) is a compact sequence and

$$\operatorname{ess\,rank}(K_n) = \operatorname{ess\,rank}(L_n).$$

To continue we need two simple lemmas.

Lemma 2.2 *Let $\mathcal{A}$ be a C^*-algebra, $\mathcal{J}$ a semi-ideal in $\mathcal{A}$ and r the associated rank function. If $k \in \mathcal{A}$ is Moore-Penrose invertible in $\mathcal{F}$ and of finite essential rank, then*

$$r(k) = r(k^*) = r(k^\dagger) = r(k^\dagger k) = r(k^* k).$$

Proof From $k = kk^\dagger k$ and the properties of the rank function we conclude that

$$r(k) \le r(k^\dagger k) \le \min\{r(k),\, r(k^\dagger)\},$$

which implies that $r(k) = r(k^\dagger k)$ and $r(k) \le r(k^\dagger)$. Since $(k^\dagger)^\dagger = k$, we obtain $r(k) = r(k^\dagger)$. Further we infer from Theorem 2.15 in [5] that $k^\dagger k = (k^*k+q)^{-1}k^*k$, where q is the Moore-Penrose projection of k. Hence,

$$r(k) = r(k^\dagger k) \le r(k^* k) \le r(k)$$

which finally implies that $r(k) = r(k^*k)$. □

Lemma 2.3 *Let $\mathcal{A}$ be a C^*-algebra, $p \in \mathcal{A}$ a projection, and $k \in p\mathcal{A}p$. Then*

$$\operatorname{alg\,rank}_{p\mathcal{A}p} k = \operatorname{alg\,rank}_{\mathcal{A}} k.$$

Proof Let $k \in p\mathcal{A}p$ be of algebraic rank 1 in $p\mathcal{A}p$ and $a \in \mathcal{A}$. Then $k = pkp$, and from $kak = pkp\, pap\, pkp$ we conclude that k has algebraic rank 1 in $\mathcal{A}$.

Set $r := \operatorname{alg\,rank}_{p\mathcal{A}p} k$ and $s := \operatorname{alg\,rank}_{\mathcal{A}} k$. Thus, k is a sum of r elements of algebraic rank 1 in $p\mathcal{A}p$. As we have just seen, these r elements are of algebraic rank 1 in $\mathcal{A}$. Hence, $s \le r$.

Conversely, let k be the sum $k_1 + \ldots + k_s$ of elements k_i of algebraic rank 1 in $\mathcal{A}$. Then, since $k \in p\mathcal{A}p$, the element $k = pkp$ is the sum $pk_1p + \ldots + pk_mp$ of m elements on algebraic rank at most 1 in $p\mathcal{A}p$. Hence, $r \le s$. □

From Lemma 2.2 we infer that $\mathbf{P} = (\Pi_n) := (L_n)^\dagger(L_n)$ is a projection with

$$\operatorname{ess\,rank}(\Pi_n) = \operatorname{ess\,rank}(L_n) = \operatorname{ess\,rank}(K_n) = r.$$

Proposition 2.4 *Let $\mathcal{A}$ be a quasifractal C^*-subalgebra of $\mathcal{F}$ which contains $\mathcal{G}$ and satisfies the local weight* 1 *condition $\mathcal{K}(\mathcal{A}) = \mathcal{A} \cap \mathcal{K}$. Further, let $\mathbf{P} \in \mathcal{A}$ be a projection of finite essential rank and $\mathbb{M} \in \operatorname{fr}\mathcal{A}$. Then*

$$\mathcal{K}\big((\mathbf{P}\mathcal{A}\mathbf{P})|_{\mathbb{M}}\big) = (\mathbf{P}\mathcal{A}\mathbf{P})|_{\mathbb{M}} \cap \mathcal{K}|_{\mathbb{M}}. \tag{2.4}$$

Proof Let $\operatorname{ess\,rank}\mathbf{P} = r$. Since $\mathbf{P} \in \mathcal{K} \cap \mathcal{A} = \mathcal{K}(\mathcal{A})$, there are sequences $(K_n^i) \in \mathcal{A}$ of spatial rank one and $(G_n) \in \mathcal{G}$ such that

$$\mathbf{P} = (K_n^1) + \ldots + (K_n^r) + (G_n). \tag{2.5}$$

We may even assume that $(K_n^i) \in \mathbf{P}\mathcal{A}\mathbf{P}$ (otherwise multiply (2.5) from both sides by $\mathbf{P}$ to get a decomposition of $\mathbf{P}$ into r sequences $\mathbf{P}(K_n^i)\mathbf{P} \in \mathbf{P}\mathcal{A}\mathbf{P}$ of essential rank one and a sequence in $\mathbf{P}\mathcal{G}\mathbf{P}$). Let $\mathbb{M} \in \operatorname{fr}\mathcal{A}$ and $\mathbf{A} \in \mathbf{P}\mathcal{A}\mathbf{P}$. Then

$$\begin{aligned}\mathbf{A}|_{\mathbb{M}} = (\mathbf{P}\mathbf{A})|_{\mathbb{M}} &= \Big(\big((K_n^1) + \ldots + (K_n^r) + (G_n)\big)\mathbf{A}\Big)\,|_{\mathbb{M}} \\ &= \big((K_n^1)\mathbf{A}\big)|_{\mathbb{M}} + \ldots + \big((K_n^r)\mathbf{A}\big)|_{\mathbb{M}} + \big((G_n)\mathbf{A}\big)|_{\mathbb{M}}\end{aligned}$$

with sequences $\big((K_n^1)\mathbf{A}\big)|_{\mathbb{M}}$ of spatial rank one in $(\mathbf{P}\mathcal{A}\mathbf{P})|_{\mathbb{M}}$, which implies that

$$(\mathbf{P}\mathcal{A}\mathbf{P})|_{\mathbb{M}} \subseteq \mathcal{K}\big((\mathbf{P}\mathcal{A}\mathbf{P})|_{\mathbb{M}}\big).$$

Since $\mathbf{P}$ is compact, we also have $(\mathbf{P}\mathcal{A}\mathbf{P})|_{\mathbb{M}} \subseteq \mathcal{K}|_{\mathbb{M}}$, which results in the inclusion $\supseteq$ in (2.4). The reverse inclusion is evident. □

Thus, if the algebra $\mathcal{A}$ satisfies the local weight one condition, then so does the algebra $(\mathbf{P}\mathcal{A}\mathbf{P})|_{\mathbb{M}}$. Since $\mathbb{M} \in \operatorname{fr}\mathcal{A}$, the latter algebra is also fractal, and we conclude from (1.1) that

$$\operatorname{ess\,rank}\mathbf{K}|_{\mathbb{M}} = \operatorname{alg\,rank}_{(\mathbf{P}\mathcal{A}\mathbf{P})|_{\mathbb{M}}/\mathcal{G}|_{\mathbb{M}}}\big(\mathbf{K}|_{\mathbb{M}} + \mathcal{G}|_{\mathbb{M}}\big) \tag{2.6}$$

for all sequences $\mathbf{K} \in \mathbf{P}\mathcal{A}\mathbf{P}$. To get rid of the restriction to the algebra $\mathbf{P}\mathcal{A}\mathbf{P}$ on the right-hand side of (2.6), we employ Lemma 2.3 which states in the present context

that

$$\operatorname{alg\,rank}_{(\mathbf{P}\mathcal{A}\mathbf{P})|_{\mathbb{M}}/\mathcal{G}|_{\mathbb{M}}}\big(\mathbf{K}|_{\mathbb{M}}+\mathcal{G}|_{\mathbb{M}}\big)=\operatorname{alg\,rank}_{\mathcal{A}|_{\mathbb{M}}/\mathcal{G}|_{\mathbb{M}}}\big(\mathbf{K}|_{\mathbb{M}}+\mathcal{G}|_{\mathbb{M}}\big).$$

The following theorem summarizes what we obtained.

Theorem 2.5 *Let $\mathcal{A}$ be a quasifractal C^*-subalgebra of $\mathcal{F}$ which contains $\mathcal{G}$ and satisfies $\mathcal{K}(\mathcal{A})=\mathcal{A}\cap\mathcal{K}$. Further, let $\mathbf{P}\in\mathcal{A}$ be a projection of finite essential rank. Then, for all $\mathbf{K}\in\mathbf{P}\mathcal{A}\mathbf{P}$,*

$$\operatorname{ess\,rank}\mathbf{K}=\max_{\mathbb{M}\in\operatorname{fr}\mathcal{A}}\operatorname{ess\,rank}\mathbf{K}|_{\mathbb{M}}$$

and, for all $\mathbb{M}\in\operatorname{fr}\mathcal{A}$,

$$\operatorname{ess\,rank}\mathbf{K}|_{\mathbb{M}}=\operatorname{alg\,rank}_{\mathcal{A}|_{\mathbb{M}}/\mathcal{G}|_{\mathbb{M}}}\big(\mathbf{K}|_{\mathbb{M}}+\mathcal{G}|_{\mathbb{M}}\big).$$

Remark 2.6 Let $\mathbf{K}$ be a sequence of finite essential rank in a quasifractal algebra $\mathcal{A}$. Does the essential rank of a restriction $\mathbf{K}|_{\mathbb{M}}$ with $\mathbb{M}\in\operatorname{fr}\mathcal{A}$ depend on the coset $\mathbb{M}^{\sim}$ only? In general the answer is NO, even when $\mathcal{A}$ is a fractal algebra (examples can be easily found among sequences of projections). The crucial ingredient for an affirmative answer is the local weight 1 condition.

Let $\mathcal{A}$ and $\mathbf{P}$ be as in Theorem 2.5, and let $\mathbf{K}=(K_n)\in\mathbf{P}\mathcal{A}\mathbf{P}$ be a sequence of finite essential rank. Further, let $\mathbb{M}_1,\ \mathbb{M}_2\in\operatorname{fr}\mathcal{A}$ and $\mathbb{M}_1\sim\mathbb{M}_2$. Then the algebra $(\mathbf{P}\mathcal{A}\mathbf{P})|_{\mathbb{M}_1\cup\mathbb{M}_2}$ is fractal and has local weight 1 by equality (2.4) in Proposition 2.4. Now it follows as in Theorem 4.4 in [10] and its corollaries that the matrices K_n have the same spatial rank for all sufficiently large $n\in\mathbb{M}_1\cup\mathbb{M}_2$. (Just to mention the point: The cited theorem deals with Fredholm sequences, i.e. with sequences that are invertible modulo $\mathcal{K}$. As in operator theory, where there is a close relation between Fredholm operators and their kernel dimension and compact operators and their rank, there is a close relation between the α_n-numbers dealt with in Theorem 4.4 in [10] and the ranks of K_n.) Hence, $\operatorname{ess\,rank}\mathbf{K}|_{\mathbb{M}_1}=\operatorname{ess\,rank}\mathbf{K}|_{\mathbb{M}_2}$.

3 Algebras with Continuous Trace

There are several (equivalent) definitions of continuous trace algebras in the literature; see, e.g. [7] and [8, Definition 5.13]. (The latter reference is an excellent introduction into the field.) Here is Fell's original definition from [4, IV.4.1].

Definition 3.1 A C^*-algebra $\mathcal{A}$ is called an *algebra with continuous trace* if

(a) $\mathcal{A}$ is liminal,
(b) $\operatorname{Prim}\mathcal{A}$ is a Hausdorff space (with respect to the hull-kernel topology),
(c) for every $L_0\in\operatorname{Prim}\mathcal{A}$, there are a neighborhood U of L_0 in $\operatorname{Prim}\mathcal{A}$ and an $a\in\mathcal{A}$ such that $a+L$ is a projection of rank 1 for all $L\in U$.

The rank condition in (c) has to be understood as follows: Since $\mathcal{A}$ is liminal, there is, for every $L \in \operatorname{Prim}\mathcal{A}$, a unique (up to unitary equivalence) irreducible representation (H, π) of $\mathcal{A}$ with $\ker\pi = L$. Then the requirement is that $\pi(a)$ is a projection on H with (spatial) rank 1. Condition (c) is also referred to as *Fell's condition*.

Example

(a) The following examples taken from [8] are instructive. Of the 'dimension-drop-algebras'

$$\mathcal{C}_1 := \{f \in C([0,\,1],\,M_2(\mathbb{C})) : f(0) = \operatorname{diag}(\alpha,\,\alpha) \text{ with } \alpha \in \mathbb{C}\},$$

$$\mathcal{C}_2 := \{f \in C([0,\,1],\,M_2(\mathbb{C})) : f(0) = \operatorname{diag}(\alpha,\,\beta) \text{ with } \alpha,\,\beta \in \mathbb{C}\},$$

$$\mathcal{C}_3 := \{f \in C([0,\,1],\,M_2(\mathbb{C})) : f(0) = \operatorname{diag}(\alpha,\,0) \text{ with } \alpha \in \mathbb{C}\},$$

$\mathcal{C}_3$ is the only one with continuous trace. For $\mathcal{C}_1$, the identity matrix $\operatorname{diag}(1,\,1)$ is a projection of algebraic rank one at 0, but there is no rank one projection close to it in a neighborhood of 0, whereas for $\mathcal{C}_2$, the Hausdorff property is violated.

(b) Let (s_n) be a dense subsequence of $[0,\,1]$. For $i \in \{1,\,2,\,3\}$, let $\mathcal{S}(\mathcal{C}_i)$ stand for the smallest closed subalgebra of $\mathcal{F}$ which contains the ideal $\mathcal{G}$ and all sequences $(f(s_n))$ with $f \in \mathcal{C}_i$ where $\mathcal{C}_i$ is as in (a). (Here we assume that $\delta \geq 2$ and identify a 2×2-matrix A with the $\delta(n) \times \delta(n)$-matrix $\left(\begin{smallmatrix} A & 0 \\ 0 & 0 \end{smallmatrix}\right)$.) Each of the algebras $\mathcal{S}(\mathcal{C}_i)$ is quasifractal, $\mathcal{S}(\mathcal{C}_i) \subseteq \mathcal{K}$, and the mapping

$$\mathcal{S}(\mathcal{C}_i)/\mathcal{G} \to \mathcal{C}_i, \quad (f(s_n)) + \mathcal{G} \to f$$

is a *-isomorphism from $\mathcal{S}(\mathcal{C}_i)/\mathcal{G}$ onto $\mathcal{C}_i$. In particular, $\operatorname{Prim}(\mathcal{S}(\mathcal{C}_i)/\mathcal{G}) \cong \operatorname{Prim}\mathcal{C}_i$. Thus, also assumptions such as compactness and quasifractality cannot change the picture sketched in part (a): among the algebras $\mathcal{S}(\mathcal{C}_i)/\mathcal{G}$, only $\mathcal{S}(\mathcal{C}_2)/\mathcal{G}$ i s an algebra with continuous trace. □

Let us now go back to the question from the introduction whether $(\mathcal{A} \cap \mathcal{K})/\mathcal{G}$ is an algebra with continuous trace when $\mathcal{A}$ is quasifractal. One can show that the ideal $\mathcal{K}$ is a liminal algebra. Since C^*-subalgebras and quotients of liminal C^*-algebras are liminal (see [3], 4.2.4, 4.3.4 and 4.3.5), condition (a) in Definition 3.1 is satisfied for every algebra $(\mathcal{A} \cap \mathcal{K})/\mathcal{G}$. On the other side, the Examples 3 show that we cannot expect the Hausdorff property of $\operatorname{Prim}(\mathcal{A} \cap \mathcal{K})/\mathcal{G}$ in condition (b) to hold, even for very simple examples of quasifractal algebras. So our focus will be on Fell's condition (c). The following three sections will prepare a natural additional condition for the ideal of the compact sequences in $\mathcal{A}$ which will allow us to overcome the problems with Fell's condition (b).

4 A Topology on $(\mathrm{fr}\,\mathcal{A})^\sim$

Here we recall from [13] the definition of a topology on $(\mathrm{fr}\,\mathcal{A})^\sim$ which makes $(\mathrm{fr}\,\mathcal{A})^\sim$ to a compact Hausdorff space. For $\mathcal{A}$ as C^*-subalgebra of $\mathcal{F}$, let $\mathcal{L}_{\min}(\mathcal{A})$ denote the smallest closed complex subalgebra[1] of $l^\infty := l^\infty(\mathbb{N})$ which contains all sequences $(\|A_n\|)$ where $(A_n) \in \mathcal{A}$. Clearly, $\mathcal{L}_{\min}(\mathcal{A})$ is a commutative C^*-algebra, which is unital if $\mathcal{A}$ is unital.

For a C^*-subalgebra $\mathcal{L}$ of l^∞, we let $\mathrm{cr}\,\mathcal{L}$ stand for the set of all infinite subsets $\mathbb{M}$ of $\mathbb{N}$ such that all sequences in the restriction $\mathcal{L}|_{\mathbb{M}}$ converge. The algebra $\mathcal{L}$ is called *quasiconvergent* if every infinite subset of $\mathbb{N}$ has an infinite subset in $\mathrm{cr}\,\mathcal{L}$. Then, for every C^*-subalgebra $\mathcal{A}$ of $\mathcal{F}$, $\mathrm{fr}\,\mathcal{A} = \mathrm{cr}\,\mathcal{L}_{\min}(\mathcal{A})$, and $\mathcal{A}$ is quasifractal if and only if $\mathcal{L}_{\min}(\mathcal{A})$ is quasiconvergent.

Let $\mathcal{L}$ be a unital C^*-subalgebra of l^∞. Then, for every $\mathbb{M} \in \mathrm{cr}\,\mathcal{L}$, the mapping

$$\varphi_{\mathbb{M}} : \mathcal{L} \to \mathbb{C}, \quad a \mapsto \lim (a|_{\mathbb{M}}) \tag{4.1}$$

is a character (i.e., a non-zero continuous linear functional) on $\mathcal{L}$. Since $\mathcal{L} \cap c_0$ is in the kernel of the mapping (4.1), the quotient mapping

$$\varphi_{\mathbb{M}} : \mathcal{L}/(\mathcal{L} \cap c_0) \to \mathbb{C}, \quad a + (\mathcal{L} \cap c_0) \mapsto \lim (a|_{\mathbb{M}})$$

is well defined, and this mapping is a character of $\mathcal{L}/(\mathcal{L} \cap c_0)$.

Proposition 4.1 (Proposition 4.5 in [13]) *Let $\mathcal{L}$ be a unital and quasiconvergent C^*-subalgebra of l^∞. Then the set $\{\varphi_{\mathbb{M}} : \mathbb{M} \in \mathrm{cr}\,\mathcal{L}\}$ is strictly spectral for $\mathcal{L}/(\mathcal{L} \cap c_0)$, i.e., if $b \in \mathcal{L}/(\mathcal{L} \cap c_0)$ and $\varphi_{\mathbb{M}}(b)$ is invertible for all $\mathbb{M} \in \mathrm{cr}\,\mathcal{L}$, then b is invertible.*

In order to conclude that $\{\varphi_{\mathbb{M}} : \mathbb{M} \in \mathrm{cr}\,\mathcal{L}\}$ is all of the maximal ideal space $\mathrm{Max}\,\mathcal{L}/(\mathcal{L} \cap c_0)$ of $\mathcal{L}/(\mathcal{L} \cap c_0)$ we need a further property of $\mathcal{L}$: separability,[2] and in order to make the mapping $\mathbb{M} \mapsto \varphi_{\mathbb{M}}$ injective, we introduce an equivalence relation $\sim$ on $\mathrm{cr}\,\mathcal{L}$ by calling $\mathcal{M}_1$, $\mathbb{M}_2$ of $\mathrm{cr}\,\mathcal{L}$ *equivalent* if $\mathbb{M}_1 \cup \mathbb{M}_2 \in \mathrm{cr}\,\mathcal{L}$. We denote the equivalence class of $\mathbb{M} \in \mathrm{cr}\,\mathcal{L}$ by $\mathbb{M}^\sim$ and write $(\mathrm{cr}\,\mathcal{L})^\sim$ for the set of all equivalence classes.

Theorem 4.2 (Corollary 4.7 in [13]) *Let $\mathcal{L}$ be a unital, separable and quasiconvergent C^*-subalgebra of l^∞. Then the mapping $\mathbb{M}^\sim \mapsto \varphi_{\mathbb{M}}$ is a bijection from $(\mathrm{cr}\,\mathcal{L})^\sim$ onto $\mathrm{Max}\,(\mathcal{L}/(\mathcal{L} \cap c_0))$.*

If now $\mathcal{A}$ is a unital and quasifractal C^*-subalgebra of $\mathcal{F}$, then $\mathcal{L}_{\min}(\mathcal{A})$ is a unital and quasiconvergent C^*-subalgebra of l^∞.

[1] The algebra $\mathcal{L}_{\min}(\mathcal{A})$ is denoted by $\mathcal{L}(\mathcal{A})$ in [13].

[2] Actually, one only needs that $\mathrm{Max}\,\mathcal{L}/(\mathcal{L} \cap c_0)$ is first countable; see [6].

Corollary 4.3 *Let $\mathcal{A}$ be a unital and quasifractal C^*-subalgebra of $\mathcal{F}$ such that $\mathcal{L}_{\min}(\mathcal{A})$ is separable. Then the mapping $\mathbb{M}^\sim \mapsto \varphi_{\mathbb{M}}$ is a bijection from $(\mathrm{fr}\,\mathcal{A})^\sim$ onto $\mathrm{Max}\,(\mathcal{L}_{\min}(\mathcal{A})/(\mathcal{L}_{\min}(\mathcal{A}) \cap c_0))$.*

Under the conditions of the corollary, one thus can transfer the Gelfand topology from $\mathrm{Max}\,(\mathcal{L}_{\min}(\mathcal{A})/(\mathcal{L}_{\min}(\mathcal{A}) \cap c_0))$ onto $(\mathrm{fr}\,\mathcal{A})^\sim$ making the latter a compact Hausdorff space.

Besides $\mathcal{L}_{\min}(\mathcal{A})$, there is another way to associate a commutative C^*-algebra with a quasifractal algebra $\mathcal{A}$ which induces a Hausdorff topology on $(\mathrm{fr}\,\mathcal{A})^\sim$. Let $\mathcal{L}_{\max}(\mathcal{A})$ stand for the set of all sequences $(\alpha_n) \in l^\infty$ with the property that the restricted sequence $(\alpha_n)|_{\mathbb{M}}$ converges for every $\mathbb{M} \in \mathrm{fr}\,A$. Clearly, $\mathcal{L}_{\min}(\mathcal{A}) \subseteq \mathcal{L}_{\max}(\mathcal{A})$ (hence the notation). Moreover, the algebra $\mathcal{L}_{\max}(\mathcal{A})$ is unital and quasiconvergent whenever $\mathcal{A}$ is quasifractal, and $\mathrm{fr}\,\mathcal{A} = \mathrm{cr}\,\mathcal{L}_{\max}(\mathcal{A})$. So one can also generate a topology on $(\mathrm{fr}\,\mathcal{A})^\sim$ using $\mathcal{L}_{\max}(\mathcal{A})$. It turns out that these topologies coincide in important cases. For example, using ideas from the proof of Theorem 4.2 = Corollary 4.7 in [13], one can show that $\mathcal{L}_{\min}(\mathcal{A}) = \mathcal{L}_{\max}(\mathcal{A})$ whenever $\mathcal{A}$ is a unital, separable and quasifractal C^*-subalgebra of $\mathcal{F}$.

On the other side, there are contexts in which the use of $\mathcal{L}_{\max}(\mathcal{A})$ is of advantage. For example, it is often useful to have $\mathcal{L}_{\max}(\mathcal{A})$ as a subalgebra inside $\mathcal{A}$ (in the sense that a sequence $(\alpha_n) \in l^\infty$ is identified with the sequence $(\alpha_n P_n) \in \mathcal{F}$). We call $\mathcal{A}$ an $L_{\max}$*-algebra* in this case. The requirement to be an $L_{\max}$-algebra can be easily fulfilled: For a C^*-subalgebra $\mathcal{A}$ of $\mathcal{F}$, let $\mathcal{A}_{\max}$ stand for the smallest closed C^*-subalgebra of $\mathcal{F}$ which contains $\mathcal{A}$ and $\mathcal{L}_{\max}(\mathcal{A})$. It is elementary to check that

$$\mathcal{L}_{\max}(\mathcal{A}_{\max}) = \mathcal{L}_{\max}(\mathcal{A}),$$

thus $\mathcal{A}_{\max}$ is an $L_{\max}$-algebra containing $\mathcal{A}$.

Theorem 4.4 *Let $\mathcal{A}$ be a unital C^*-subalgebra of $\mathcal{F}$ with $\mathcal{G} \in \mathcal{A}$, and assume that $\mathcal{A}$ is an $\mathcal{L}_{\max}$-algebra. Then*

$$\mathcal{L}_{\min}(\mathcal{A}) = \mathcal{L}_{\max}(\mathcal{A}) = \mathcal{A} \cap l^\infty. \tag{4.2}$$

Proof We first show that $\mathcal{L}_{\min}(\mathcal{L}) = \mathcal{L}$ for every C^*-subalgebra $\mathcal{L}$ of l^∞. If $(\alpha_n) \in \mathcal{L}$, then $(|\alpha_n|) = ((\alpha_n)(\alpha_n)^*)^{1/2} \in \mathcal{L}$; hence, $\mathcal{L}_{\min}(\mathcal{L}) \subseteq \mathcal{L}$. Conversely, every sequence in $\mathcal{L}$ can be written as a linear combination of four non-negative sequences. Since $(\beta_n) = (|\beta_n|) \in \mathcal{L}_{\min}(\mathcal{L})$ for each of these sequences, the reverse inclusion $\mathcal{L} \subseteq \mathcal{L}_{\min}(\mathcal{L})$ follows.

Now to the proof of (4.2). The inclusion $\mathcal{L}_{\min}(\mathcal{A}) \subseteq \mathcal{L}_{\max}(\mathcal{A})$ holds trivially, as already mentioned. To get the reverse inclusion, note that

$$\mathcal{L}_{\max}(\mathcal{A}) = \mathcal{L}_{\min}(\mathcal{L}_{\max}(\mathcal{A})) \subseteq \mathcal{L}_{\min}(\mathcal{A}),$$

where the equality comes from the remark at the beginning of the proof and the inclusion holds because $\mathcal{L}_{\max}(\mathcal{A}) \subseteq \mathcal{A}$ and the functor $\mathcal{L}_{\min}$ is increasing. This settles the first equality in (4.2).

For the second equality, set $\mathcal{L} := \mathcal{A} \cap l^\infty$, considered as a subalgebra of l^∞. Evidently, $\mathcal{L}_{\max}(\mathcal{A}) \subseteq \mathcal{L}$. For the reverse inclusion, note that $\mathcal{L} \subseteq \mathcal{L}_{\max}(\mathcal{L})$, again by the above remark. Since $\mathcal{L} \subseteq \mathcal{A}$ and the functor $\mathcal{L}_{\max}$ is increasing, the assertion follows. □

The following section shows another instance where $\mathcal{L}_{\max}(\mathcal{A})$ naturally occurs.

5 Sequences with Sequential Rank 1

Let H be a Hilbert space and $K \in L(H)$ an operator with $\operatorname{rank} K \le 1$. Then, for every $A \in L(H)$, there is an $\alpha \in \mathbb{C}$ such that $KAK = \alpha K$. The number α is uniquely determined and satisfies $|\alpha| \le \|A\|\,\|K\|$ whenever $K \ne 0$. If $K = 0$, then $KAK = \alpha K$ holds for every $\alpha \in \mathbb{C}$, whereas the condition $|\alpha| \le \|A\|\,\|K\|$ only holds when $\alpha = 0$. Thus, both conditions together determine α uniquely.

Let now $\mathcal{A}$ be a C^*-subalgebra of $\mathcal{F}$ and let $(K_n) \in \mathcal{A}$ be a sequence with $\max_n \operatorname{rank} K_n \le 1$. Then, as we have just seen, for every sequence $(A_n) \in \mathcal{A}$, there is a *unique* sequence $(\alpha_n) \in l^\infty$ such that

$$K_n A_n K_n = \alpha_n K_n \quad \text{and} \quad |\alpha_n| \le \|A_n\|\,\|K_n\| \quad \text{for every } n \in \mathbb{N}. \tag{5.1}$$

Theorem 5.1 *Let $\mathcal{A}$ be a unital C^*-subalgebra of $\mathcal{F}$ with $\mathcal{G} \subseteq \mathcal{A}$ and let $(A_n) \in \mathcal{A}$. Then the sequence (α_n) determined by (5.1) belongs to $\mathcal{L}_{\max}(\mathcal{A})$.*

Proof It is helpful to define the *support* of a sequence $(A_n) \in \mathcal{A}$ by

$$\operatorname{supp}(A_n) := \{\mathbb{M} \in \operatorname{fr}\mathcal{A} : (A_n)|_{\mathbb{M}} \notin \mathcal{G}|_{\mathbb{M}}\}.$$

If $\mathbb{M} \in \operatorname{supp}(A_n)$, then the restriction $\mathcal{A}|_{\mathbb{M}}$ is fractal and $(A_n)|_{\mathbb{M}}$ is not a partial zero sequence. Thus,

$$\mathbb{M} \in \operatorname{supp}(A_n) \quad \text{if and only if} \quad \lim_{n\in\mathbb{M}} \|A_n\| > 0. \tag{5.2}$$

Note also that if one of two equivalent sets in $\operatorname{fr}\mathcal{A}$ belongs to $\operatorname{supp}(A_n)$, then so does the other. In this sense, the support of a sequence is compatible with the equivalence relation $\sim$.

Let now (A_n), $(K_n) \in \mathcal{A}$ with $\max_n \operatorname{rank} K_n \le 1$, and let (α_n) be the sequence uniquely determined by (5.1). Further let $\mathbb{M} \in \operatorname{fr}\mathcal{A}$. We distinguish two cases.

Case A $\mathbb{M} \notin \operatorname{supp}(A_n)$. Then $(A_n)|_{\mathbb{M}} \in \mathcal{G}|_{\mathbb{M}}$ by (5.2). The second condition in (5.1) then implies that $(\alpha_n)|_{\mathbb{M}}$ is a zero sequence; hence, convergent.

Case B $\mathbb{M} \in \operatorname{supp}(A_n)$. Suppose $(\alpha_n)|_{\mathbb{M}}$ is not convergent. Then, since (α_n) is a bounded sequence, there are complex numbers $\alpha \ne \beta$ and disjoint infinite subsets $\mathbb{M}_\alpha$, $\mathbb{M}_\beta$ of $\mathbb{M}$ such that $\alpha_n \to \alpha$ on $\mathbb{M}_\alpha$ and $\alpha_n \to \beta$ on $\mathbb{M}_\beta$.

Consider the sequence $(\alpha_n K_n - \alpha K_n)$ which is equal to $(K_n A_n K_n) - \alpha(K_n)$ and, hence, belongs to $\mathcal{A}$. Its restriction onto $\mathbb{M}_\alpha$ is a zero sequence, whereas its restriction to $\mathbb{M}_\beta$ is not. The latter follows from the estimate

$$\|\alpha_n K_n - \alpha K_n\| \geq |\alpha - \beta| \, \|K_n\| - |\beta - \alpha_n| \, \|K_n\| \tag{5.3}$$

and the fact that the sequence $(\|K_n\|)|_{\mathbb{M}}$ has a positive limit, k, by (5.2), which implies that the right hand side of (5.3) converges to $k|\alpha - \beta| > 0$.

Hence, $(\alpha_n K_n - \alpha K_n)|_{\mathbb{M}}$ is a partial zero sequence but not a zero sequence, which contradicts the fractality of $\mathcal{A}|_{\mathbb{M}}$. This contradiction shows that the sequence $(\alpha_n)|_{\mathbb{M}}$ converges for every $\mathbb{M} \in \operatorname{fr} \mathcal{A}$, whence $(\alpha_n) \in \mathcal{L}_{\max}(\mathcal{A})$. □

Let $\mathcal{A}$ be a unital C^*-algebra and $\mathcal{C}$ a unital C^*-subalgebra in the center of $\mathcal{A}$, i.e., the elements of $\mathcal{C}$ commute which each element of $\mathcal{A}$. Then $\mathcal{C}$ is a unital commutative C^*-subalgebra of $\mathcal{A}$. Let $\mathcal{J}$ stand for the set of all elements $k \in \mathcal{A}$ with the following property: For every $a \in \mathcal{A}$ there is a $c \in \mathcal{C}$ such that $kak = ck$. Then $\mathcal{J}$ forms a semi-ideal of $\mathcal{A}$. The associated rank function is called the $\mathcal{C}$-rank, and the associated ideal of the $\mathcal{C}$-compact elements is denoted by $\mathcal{J}_{\mathcal{C}}(\mathcal{A})$. In case $\mathcal{C}$ is the complete center of $\mathcal{A}$, the $\mathcal{C}$-rank is also called the *central rank* and denoted by cen rank a, and the associated ideal of the centrally compact elements is denoted by $\mathcal{J}_{cen}(\mathcal{A})$. Note also that $\mathcal{J}_{\mathbb{C}}(\mathcal{A}) = \mathcal{C}(\mathcal{A})$, with the notation introduced in the introduction.

Theorem 5.2 *Let $\mathcal{A}$ be a unital C^*-subalgebra of $\mathcal{F}$ with $\mathcal{G} \subseteq \mathcal{A}$, and assume $\mathcal{A}$ is an $\mathcal{L}_{\max}$-algebra. Then*

$$\mathcal{K}(\mathcal{A}) = \mathcal{J}_{cen}(\mathcal{A}).$$

The inclusion $\supseteq$ holds without the $\mathcal{L}_{\max}$-assumption.

Proof Let (K_n) be a sequence of central rank 1 in $\mathcal{A}$. Since $\mathcal{G} \subseteq \mathcal{A}$ by assumption, the center of $\mathcal{A}$ is a subalgebra of l^∞, and rank $K_n \leq 1$ for every $n \in \mathbb{N}$. Hence, $(K_n) \in \mathcal{K}(\mathcal{A})$ and $\mathcal{J}_{cen}(\mathcal{A}) \subseteq \mathcal{K}(\mathcal{A})$.

Conversely, let $(K_n) \in \mathcal{A}$ be a sequence with $\max_n \operatorname{rank} K_n \leq 1$. By (5.1), for every $(A_n) \in \mathcal{A}$, there is a unique sequence $(\alpha_n) \in l^\infty$ such that $K_n A_n K_n = \alpha_n K_n$ for every n. The sequence (α_n) belongs to $\mathcal{L}_{\max}(\mathcal{A})$ by Theorem 5.1. Since $\mathcal{A}$ is an $\mathcal{L}_{\max}$-algebra, the sequence $(\alpha_n I_n)$ lies in $\mathcal{A}$ and, hence in the center of $\mathcal{A}$. Thus, (K_n) is of central rank 1, whence $\mathcal{K}(\mathcal{A}) \subseteq \mathcal{J}_{cen}(\mathcal{A})$. □

6 Fell's Condition: Transition to $\mathcal{C}$-Compact Elements

The theory of compact sequences in fractal algebras $\mathcal{A}$ is most satisfying when the algebra $\mathcal{A}$ is of local weight 1, which happens if and only if $\mathcal{A} \cap \mathcal{K} = \mathcal{K}(\mathcal{A})$. It is only natural to impose an analogous condition when studying compact sequences in

quasifractal algebras. Thus, in what follows, we let $\mathcal{A}$ be a unital quasifractal C^*-subalgebra of $\mathcal{F}$ with $\mathcal{G} \subseteq \mathcal{A}$ which is an $\mathcal{L}_{\max}$-algebra, and we suppose further that $\mathcal{L}_{\max}(\mathcal{A})$ is separable and that

$$\mathcal{A} \cap \mathcal{K} = \mathcal{K}(\mathcal{A}).$$

Then, by Theorem 5.2, $\mathcal{A} \cap \mathcal{K} = \mathcal{K}(\mathcal{A}) = \mathcal{J}_{cen}(\mathcal{A})$ where $\operatorname{Cen}\mathcal{A} = \mathcal{L}_{\max}(\mathcal{A}) = \mathcal{A} \cap l^\infty(\mathbb{N})$ by Theorem 4.4.

We are going to localize the algebra $\mathcal{B} := \mathcal{A}/\mathcal{G}$ over its central subalgebra $\mathcal{C} := (\mathcal{L}_{\max}(\mathcal{A}) + \mathcal{G})/\mathcal{G} \cong \mathcal{L}_{\max}(\mathcal{A})/c_0$ via the local principle by Allan/Douglas (see Theorems 2.2.2 and 2.2.11 in [14]). By Theorem 4.2,

$$\operatorname{Max}\mathcal{C} = \operatorname{Max}(\mathcal{L}_{\max}(\mathcal{A})/c_0) \cong (\operatorname{fr}\mathcal{A})^\sim.$$

Let $\mathcal{I}_x$ stand for the smallest closed ideal of $\mathcal{B}$ which contains the maximal ideal x of $\mathcal{C}$ and let Φ_x denote the quotient map $\mathcal{B} \to \mathcal{B}/\mathcal{I}_x$. Then the local principle states (among other things) that an element $b \in \mathcal{B}$ is invertible in $\mathcal{B}$ if and only if $\Phi_x(b)$ is invertible in $\mathcal{B}/\mathcal{I}_x$ for every $x \in \operatorname{Max}\mathcal{B}$ and that the function

$$\operatorname{Max}\mathcal{C} \to \mathbb{R}^+, \quad x \mapsto \|\Phi_x(b)\| \tag{6.1}$$

is upper semi-continuous for every $b \in \mathcal{B}$. In the present context on can say more: If $x = \mathbb{M}^\sim \in (\operatorname{fr}\mathcal{A})^\sim$ then, by Theorem 4.9 in [13] (note that $\mathcal{L}_{\max}(\mathcal{A}) = \mathcal{L}_{\min}(\mathcal{A})$ by Theorem 4.4), the local algebra $\mathcal{A}/\mathcal{I}_x$ is *-isomorphic to the fiber $\mathcal{A}(\mathbb{M}^\sim)$, and the function (6.1) is continuous for every $b \in \mathcal{B}$. (Roughly speaking, the continuity results from including $\mathcal{L}_{\max}(\mathcal{A}) = \mathcal{L}_{\min}(\mathcal{A})$ into $\mathcal{A}$.)

If $(K_n) \in \mathcal{A}$ is a sequence of central rank one, then its coset $k := (A_n) + \mathcal{G}$ has the property that, for every $b \in \mathcal{B}$, there is a $c \in \mathcal{C}$ such that $kbk = ck$. Thus, $k \in \mathcal{B}$ has $\mathcal{C}$-rank 1. The ideal $\mathcal{J}_{cen}(\mathcal{A})/\mathcal{G}$ of $\mathcal{B}$ we are interested in will be denoted by $\mathcal{J}$. As we have just seen, $\mathcal{J}$ is generated by certain elements of $\mathcal{C}$-rank 1. Hence, $\mathcal{J} \subseteq \mathcal{J}_{\mathcal{C}}(\mathcal{B})$. Note also that $\mathcal{C}$ is a subalgebra of the center of $\mathcal{B}$, hence $\mathcal{J}_{\mathcal{C}}(\mathcal{B}) \subseteq \mathcal{J}_{cen}(\mathcal{B})$. We cannot claim that equality holds in either of these inclusions.

7 Fell's Condition for $\mathcal{C}$-Compact Elements

In this section, we are going to show that Fell's condition is satisfied for the algebras $\mathcal{J}$ and $\mathcal{J}_{\mathcal{C}}(\mathcal{B})$. We prepare the next steps with a few lemmas. Since now $\mathcal{C}$ stands for a central subalgebra, we adopt the notation introduced in the previous section and write $\mathcal{J}_{\mathbb{C}}(\mathcal{A})$ instead of $\mathcal{C}(\mathcal{A})$ for the closed ideal of a C^*-algebra $\mathcal{A}$ generated by its elements of $\mathbb{C}$-rank = algebraic rank 1 to avoid confusion.

Lemma 7.1 *Let $\mathcal{B}$ be a unital C^*-algebra, $\mathcal{C}$ a central C^*-subalgebra of $\mathcal{B}$ which contains the unit element, $x \in \operatorname{Max}\mathcal{C}$, and $\mathcal{J} \subseteq \mathcal{J}_{\mathcal{C}}(\mathcal{B})$ a closed ideal of $\mathcal{B}$ which is generated by elements of $\mathcal{C}$-rank one. Then*

(*a*) *if $k \in \mathcal{B}$ is of $\mathcal{C}$-rank one, then $\Phi_x(k)$ is of algebraic rank ≤ 1 in $\mathcal{B}/\mathcal{I}_x$,*
(*b*) *$\Phi_x(\mathcal{J}) \subseteq \mathcal{J}_{\mathbb{C}}(\mathcal{B}/\mathcal{I}_x)$,*
(*c*) *$\Phi_x(\mathcal{J})$ is a dual algebra.*

Proof (*a*) For every $b \in \mathcal{B}$, there is a $c \in \mathcal{C}$ such that $kbk = ck$. Thus, $\Phi_x(k)\Phi_x(b)\Phi_x(k) = \Phi_x(c)\Phi_x(k)$. Since $\Phi_x(c)$ is the value of the Gelfand transform of c at x, hence a complex number, the assertion follows. Since $\mathcal{J}$ is generated as an ideal by sequences of $\mathcal{C}$-rank 1, (*b*) is a consequence of (*a*).

Finally, we know from [10] that $\mathcal{J}_{\mathbb{C}}(\mathcal{A})$ is a dual algebra for every C^*-algebra $\mathcal{A}$, and it is well known (see, e.g., [1]) that dual algebras are *-isomorphic to C^*-subalgebras of $K(H)$ for some Hilbert space H and that, hence, C^*-subalgebras of dual algebras are dual again. Therefore $\Phi_x(\mathcal{J})$, which is a C^*-subalgebra of a dual algebra by (*b*), is dual. □

Lemma 7.2

(*a*) *Let $\mathcal{B}$ be a C^*-algebra with identity e, $\mathcal{J}$ a proper closed ideal of $\mathcal{B}$, $j \in \mathcal{J}$ self-adjoint, and f a continuous function on $\sigma(j)$ with $f(0) = 0$. Then $f(j) \in \mathcal{J}$.*
(*b*) *In addition, let $\mathcal{C}$ be a central C^*-subalgebra of $\mathcal{B}$ which contains e, let j have $\mathcal{C}$-rank one, and assume that f is continuously differentiable on some interval containing $\sigma(j)$ and that $f(0) = f'(0) = 0$. Then $f(j)$ has $\mathcal{C}$-rank ≤ 1.*

Proof We will prove part (*b*) only. Let $[a, b]$ be a closed interval which contains $\sigma(j)$. Since $\mathcal{J}$ is proper, $0 \in \sigma(j) \subseteq [a, b]$. By the Weierstrass' approximation theorem, there is a sequence of polynomials p_n such that $p_n \to f$ and $p_n' \to f'$ uniformly on $[a, b]$ and $p_n(0) = p_n'(0) = 0$. Define polynomials q_n such that $p_n(x) = xq_n(x)$ for all x.

Since f is differentiable and $f'(0) = 0$, the function g defined by $g(x) := f(x)/x$ if $x \neq 0$ and $g(0) := 0$ is continuous on $[a, b]$ and differentiable on $(a, 0)$ and $(0, b)$. Let $x \in (0, b]$. Then, by the mean value theorem, there is a $\xi \in (0, x)$ such that

$$q_n(x) - g(x) = \frac{p_n(x) - f(x)}{x} = p_n'(\xi) - f'(\xi),$$

hence $|q_n(x) - g(x)| \leq \|p_n' - f'\|_\infty$. This estimate holds for $x \in [a, 0)$ and $x = 0$ as well. Thus,

$$\|q_n - g\|_\infty \leq \|p_n' - f'\|_\infty,$$

implying that the q_n converge uniformly to g. By the functional calculus again, $p_n(j) \to f(j)$ and $p_n(j) = jq_n(j) \to jg(j)$ in the norm of $\mathcal{B}$. Hence, $f(j) = jg(j)$, and $f(j)$ has $\mathcal{C}$-rank ≤ 1 because j has $\mathcal{C}$-rank 1. □

The next results state that projections of algebraic rank one in $\Phi_x(\mathcal{J})$ can be lifted both to elements of $\mathcal{C}$-rank one in $\mathcal{J}$ and to local projections.

Proposition 7.3 *Let $\mathcal{B}$ be a unital C^*-algebra, $\mathcal{C}$ a central C^*-subalgebra of $\mathcal{B}$ which contains the unit element, and $\mathcal{J} \subseteq \mathcal{J}_{\mathcal{C}}(\mathcal{B})$ a closed ideal of $\mathcal{B}$ which is generated by elements of $\mathcal{C}$-rank one. Let $a \in \mathcal{J}$, $x \in \operatorname{Max}\mathcal{C}$, and suppose that $\Phi_x(a)$ is a projection of algebraic rank one in $\mathcal{B}/\mathcal{I}_x$. Then there is an element $k \in \mathcal{J}$ of $\mathcal{C}$-rank one such that $\Phi_x(k) = \Phi_x(a)$.*

Proof Set $p := \Phi_x(a)$. Being of algebraic rank one, $p \neq 0$. Since $a \in \mathcal{J}$, a is a limit of sums of elements $r_{in} \in \mathcal{J}$ with $\mathcal{C}$-rank one:

$$a = \lim_{n\to\infty} \sum_{i=1}^{n} r_{in}.$$

Then

$$p = \Phi_x(a) = \lim_{n\to\infty} \sum_{i=1}^{n} \Phi_x(r_{in})$$

and, because p is a projection,

$$p = \lim_{n\to\infty} \sum_{i=1}^{n} p\Phi_x(r_{in})p. \tag{7.1}$$

Since p is of algebraic rank one, there are numbers $\lambda_{in} \in \mathbb{C}$ such that $p\Phi_x(r_{in})p = \lambda_{in}p$. Not all of these numbers can be zero (otherwise (7.1) would imply that $p = 0$). Let $\lambda := \lambda_{i_0 n_0} \neq 0$. Then, with $r := r_{i_0 n_0}$, $p\Phi_x(r)p = \lambda p$, whence

$$p = \frac{1}{\lambda} p\Phi_x(r)p = \frac{1}{\lambda}\Phi_x(a)\Phi_x(r)\Phi_x(a) = \Phi_x(\lambda^{-1}ara).$$

Then $k := \lambda^{-1}ara$ belongs to $\mathcal{J}$ and has the desired properties. □

Proposition 7.4 *Let $\mathcal{B}$ be a unital C^*-algebra, $\mathcal{C}$ a central C^*-subalgebra of $\mathcal{B}$ which contains the unit element, and $\mathcal{J}$ a closed ideal of $\mathcal{B}$. For $x \in \operatorname{Max}\mathcal{C}$, let $p \in \Phi_x(\mathcal{J})$ be a projection. Then there are an open neighborhood U of x and a positive element $a \in \mathcal{J}$ such that $\Phi_x(a) = p$ and the $\Phi_y(a)$ is a projection for every $y \in U$. Moreover, if the function*

$$\operatorname{Max}\mathcal{C} \ni y \mapsto \|\Phi_y(j)\| \tag{7.2}$$

is continuous for every $j \in \mathcal{J}$, then U can be chosen such that $\Phi_y(a) \neq 0$ for all $y \in U$.

Proof Let $b \in \mathcal{J}$ be such that $\Phi_x(b) = p$. Without loss we may assume that b is positive (otherwise first replace b by $(b + b^*)/2$ and then by $\max(0, b)$, using the continuous functional calculus and Lemma 7.2 (a)). From

$$0 = \|p - p^2\| = \|\Phi_x(b) - \Phi_x(b)^2\|$$

and the upper semi-continuity of the function (7.2) we conclude that, given $\varepsilon > 0$, there is an open neighborhood U_ε of x such that

$$\|\Phi_y(b) - \Phi_y(b)^2\| \leq \varepsilon - \varepsilon^2 \quad \text{for all } y \in U_\varepsilon.$$

Since $|s - s^2| \leq \varepsilon - \varepsilon^2$ for $s \in \mathbb{R}$ implies $s \in [-\varepsilon, \varepsilon] \cup [1 - \varepsilon,\ 1 + \varepsilon]$, we conclude that

$$\sigma(\Phi_y(b)) \subseteq [-\varepsilon, \varepsilon] \cup [1 - \varepsilon,\ 1 + \varepsilon] \quad \text{for all } y \in U_\varepsilon.$$

Let f be a bounded continuous function on $\mathbb{R}$ which is 0 on $[-\varepsilon, \varepsilon]$ and 1 on $[1 - \varepsilon,\ 1 + \varepsilon]$, and set $a := f(b)$. Then $a \in \mathcal{J}$ by Lemma 7.2 (a), and $f(\Phi_y(b)) = \Phi_y(f(b)) = \Phi_y(a)$. Thus, $\Phi_y(a)$ is a projection for all $y \in U_\varepsilon$, and $\Phi_x(a) = f(\Phi_x(b)) = f(p) = p$ (note that f is the identical mapping on $\sigma(p) \subseteq \{0,\ 1\}$).

Up to here we only used the upper semi-continuity of the functions (7.2). If these functions are continuous, then $\|\Phi_x(a)\| = \|p\| = 1$ implies $\|\Phi_y(a)\| \geq 1/2$ for all y in an open neighborhood $U \subseteq U_\varepsilon$ of x, thus giving the final assertion. □

Theorem 7.5 *Let $\mathcal{B}$ be a unital C^*-algebra, $\mathcal{C}$ a central C^*-subalgebra of $\mathcal{B}$ which contains the unit element, and $\mathcal{J} \subseteq \mathcal{J}_{\mathcal{C}}(\mathcal{B})$ a closed ideal of $\mathcal{B}$ which is generated by elements of $\mathcal{C}$-rank* 1. *We assume that the function* (7.2) *is continuous for every $j \in \mathcal{J}$. Finally, for $x \in \operatorname{Max}\mathcal{C}$, let $p \in \Phi_x(\mathcal{J})$ be a projection of algebraic rank* 1. *Then there are an open neighborhood U of x and a positive element $a \in \mathcal{J}$ with $\mathcal{C}$-rank* 1 *such that $\Phi_x(a) = p$ and that $\Phi_y(a)$ is a projection of algebraic rank* 1 *for all $y \in U$.*

Proof Let $p \in \Phi_x(\mathcal{J})$ be a projection of algebraic rank 1. By Proposition 7.3, there is an element $k \in \mathcal{J}$ of $\mathcal{C}$-rank 1 such that $\Phi_x(k) = p$. Further, as in the proof of Proposition 7.4, we find a positive element $j \in \mathcal{J}$ such that $\Phi_x(j) = p$. Then $b := k^* j k$ is a positive element of $\mathcal{J}$ with $\Phi_x(b) = p$, and b has $\mathcal{C}$-rank 1 (note that $b = k^* j k \neq 0$ because $p \neq 0$).

We proceed as in the proof of Proposition 7.4, but now with the function f defined there being chosen continuously differentiable. Then $a := f(b)$ belongs to $\mathcal{J}$ by Lemma 7.2 (a) and has $\mathcal{C}$-rank ≤ 1 by Lemma 7.2 (b). It follows as in the proof of Proposition 7.4 that $\Phi_x(a) = p$ and that $\Phi_y(a)$ is a projection for all y in a certain neighborhood U of x, and this projection has algebraic rank ≤ 1 by Lemma 7.1 (a).

Finally, since $p \neq 0$, a has $\mathcal{C}$-rank 1, and the continuity of the functions (7.2) implies that $\Phi_y(a)$ is a projection of algebraic rank 1 for all y in a certain open neighborhood of x (possibly smaller than U). □

One may consider Theorem 7.5 as stating Fell's condition for $\mathcal{J}$ with respect to the base space $\operatorname{Max}\mathcal{C}$. To get from here Fell's condition for $\mathcal{J}$ with respect to the primitive ideal space $\operatorname{Prim}\mathcal{J}$, we still need a 'transfer lemma' which relates the two settings. For, we need some facts on primitive ideal spaces. First, the mapping

$$\varphi : \operatorname{Prim}\mathcal{A} \to \operatorname{Max}\mathcal{C}, \quad L \mapsto L \cap \mathcal{C}$$

is onto, and it is continuous with respect to the hull-kernel topology on $\operatorname{Prim}\mathcal{A}$ (and it is open if and only if the function $\operatorname{Max}\mathcal{C} \ni y \mapsto \|\Phi_y(j)\|$ is continuous for every $a \in \mathcal{A}$) (Section C1 in [16]) and, second, the mapping

$$\lambda_{\mathcal{J}} : \operatorname{Prim}^{\mathcal{J}}\mathcal{A} = \{L \in \operatorname{Prim}\mathcal{A} : \mathcal{J} \not\subseteq L\} \to \operatorname{Prim}\mathcal{J}, \quad L \mapsto L \cap \mathcal{J}$$

is a homeomorphism (Proposition A27 in [8]). Define

$$\varphi_{\mathcal{J}} : \operatorname{Prim}\mathcal{J} \to \operatorname{Max}\mathcal{C}, \quad \varphi_{\mathcal{J}} := \varphi \circ \lambda_{\mathcal{J}}^{-1}.$$

Let $L_0 \in \operatorname{Prim}\mathcal{J}$ and $x_0 := \varphi_{\mathcal{J}}(L_0)$. Then $x_0 = \varphi(\chi_{\mathcal{J}}^{-1}(L_0)) = \chi_{\mathcal{J}}^{-1}(L_0) \cap \mathcal{C}$; hence, $x_0 \subseteq \chi_{\mathcal{J}}^{-1}(L_0)$. Since $\chi_{\mathcal{J}}^{-1}(L_0)$ is a closed ideal of $\mathcal{A}$, this implies $I_{x_0} \subseteq \chi_{\mathcal{J}}^{-1}(L_0)$, from which we conclude that

$$I_{x_0} \cap \mathcal{J} \subseteq \lambda_{\mathcal{J}}^{-1}(L_0) \cap \mathcal{J} = \lambda_{\mathcal{J}}(\lambda_{\mathcal{J}}^{-1}(L_0)) = L_0.$$

Summarizing, we obtain

$$I_{x_0} \cap \mathcal{J} \subseteq L_0 \quad \text{for } x_0 = \varphi_{\mathcal{J}}(L_0). \tag{7.3}$$

Lemma 7.6 *Let $\mathcal{B}$, $\mathcal{C}$ and $\mathcal{J} \subseteq \mathcal{J}_{\mathcal{C}}(\mathcal{B})$ be as in Theorem 7.5.*

(*a*) *Let $L_0 \in \operatorname{Prim}\mathcal{J}$ and p a projection of algebraic rank 1 in $\mathcal{J}/L_0$. Set $x_0 := \varphi_{\mathcal{J}}(L_0)$. Then there is a $a \in \mathcal{J}$ of $\mathcal{C}$-rank 1 such that $a + L_0 = p$ and $a + I_{x_0}$ is a projection of algebraic rank 1 in $\Phi_{x_0}(\mathcal{J})$.*

(*b*) *Let $x_0 \in \operatorname{Max}\mathcal{C}$ and $a \in \mathcal{J}$ of $\mathcal{C}$-rank 1 such that $a + I_{x_0}$ is a projection of algebraic rank 1 in $\Phi_{x_0}(\mathcal{J})$. Then there is an $L_0 \in \operatorname{Prim}\mathcal{J}$ with $\varphi_{\mathcal{J}}(L_0) = x_0$ such that $a + L_0$ is a projection of algebraic rank 1 in $\mathcal{J}/L_0$. This L_0 is unique.*

Proof

(*a*) Given a projection p of algebraic rank 1 in $\mathcal{J}/L_0$, choose a positive $b \in \mathcal{J}$ such that $b + L_0 = p$. Since $\mathcal{J}$ is generated by sequences of $\mathcal{C}$-rank 1, there are elements $j_{kn} \in \mathcal{J}$ of $\mathcal{C}$-rank 1 and numbers $k_n \in \mathbb{N}$ such that

$$b = \lim_{n\to\infty} \sum_{k=1}^{k_n} j_{kn}.$$

Passing to cosets modulo L_0 and multiplying by the projection $b + L_0 = p$ yields

$$b + L_0 = \lim_{n\to\infty} \sum_{k=1}^{k_n} (b + L_0)\,(j_{kn} + L_0)\,(b + L_0). \tag{7.4}$$

Since $b + L_0 = p$ is of algebraic rank 1, there are $\alpha_{kn} \in \mathbb{C}$ such that

$$(b + L_0)\,(j_{kn} + L_0)\,(b + L_0) = \alpha_{kn}(b + L_0).$$

By (7.4) and since $p \neq 0$, at least one of the α_{kn} is not zero; say $\alpha_{k'n'}$. Then $j := j_{k'n'}/\alpha_{k'n'}$ is in $\mathcal{J}$, is of $\mathcal{C}$-rank 1, and satisfies

$$(b + L_0)\,(j + L_0)\,(b + L_0) = b + L_0.$$

We may moreover assume that $j \geq 0$ (otherwise replace j by j^*bj). Then $a := bjb \in \mathcal{J}$ is positive, of $\mathcal{C}$-rank 1, and $a + L_0 = b + L_0 = p$.

Since a has $\mathcal{C}$-rank 1, the coset $a + I_{x_0}$ has algebraic rank ≤ 1. Actually, it has algebraic rank 1 since $a + I_{x_0} = 0$ would imply $p = a + L_0 = 0$ by (7.3), which is impossible. Since $b \neq 0$ is positive and has $\mathcal{C}$-rank 1, $\Phi_{x_0}(b/\|\Phi_{x_0}(b)\|)$ is a projection of algebraic rank 1. Thus,

$$\big(b/\|\Phi_{x_0}(b)\|\big)^2 - b/\|\Phi_{x_0}(b)\| \in I_{x_0}$$

whence, by (7.3),

$$\big(b/\|\Phi_{x_0}(b)\|\big)^2 - b/\|\Phi_{x_0}(b)\| \in L_0.$$

Thus, $b/\|\Phi_{x_0}(b)\| + L_0$ is a projection. Since already $b + L_0 = p$ is a projection, we conclude that $\|\Phi_{x_0}(b)\| = 1$, i.e., $b + L_0$ is a projection.

(b) Since $a^2 - a,\ a^* - a \in I_{x_0}$, we conclude from (7.3) that $a^2 - a,\ a^* - a \in L$ for each $L \in \operatorname{Prim}\mathcal{J}$ with $\varphi_{\mathcal{J}}(L) = x_0$. Thus $a + L$ is a projection for each such L. We show next that $a + L$ has algebraic rank ≤ 1. Given $k \in \mathcal{A}$, there is a $c \in \mathcal{C}$ such that $aka = ca$. Then $aka - c(x_0)k \in I_{x_0}$, whence $aka - c(x_0)k \in L$, again by (7.3).

Further, $\Phi_{x_0}(J)$ is a dual ideal which therefore is generated by projections of algebraic rank 1, i.e., by the $a + L$. So there is at least one $L_0 \in \operatorname{Prim}\mathcal{J}$ with $\varphi_{\mathcal{J}}(L_0) = x_0$ such that $a + L_0$ is a projection of algebraic rank 1. Finally, this L_0 is unique: Each projection of algebraic rank 1 in $\Phi_{x_0}(J)$ sits in a unique elementary component of $\Phi_{x_0}(J)$, and these components are in a one-to-one correspondence with the primitive ideals L of $\mathcal{J}$ with $\varphi_{\mathcal{J}}(L) = x_0$. □

The following theorem settles Fell's condition (= condition (c) in Definition 3.1) for ideals $\mathcal{J}$ which are generated by elements of $\mathcal{C}$-rank 1.

Theorem 7.7 *Let $\mathcal{B}$ be a unital C^*-algebra, $\mathcal{C}$ a central C^*-subalgebra of $\mathcal{B}$ which contains the unit element, and $\mathcal{J} \subseteq \mathcal{J}_{\mathcal{C}}(\mathcal{B})$ a closed ideal of $\mathcal{B}$ which is generated by elements of $\mathcal{C}$-rank 1. Assume that the function (7.2) is continuous for every $j \in \mathcal{J}$. Then $\mathcal{J}$ satisfies Fell's condition, i.e, for every $L_0 \in \operatorname{Prim}\mathcal{J}$, there are a neighborhood U of L_0 in $\operatorname{Prim}\mathcal{J}$ and an $a \in \mathcal{J}$ such that $a + L$ is a projection of algebraic rank 1 for all $L \in U$.*

Proof Let $L_0 \in \operatorname{Prim}\mathcal{J}$ and $x_0 := \varphi_{\mathcal{J}}(L_0)$. Let $p \in \mathcal{J}/L_0$ be a projection of algebraic rank 1. Then, by part (a) of the Transfer Lemma 7.6, there is a $a \in \mathcal{J}$ of $\mathcal{C}$-rank 1 such that $a + L_0 = p$ and $a + I_{x_0}$ is a projection of algebraic rank one in $\Phi_{x_0}(\mathcal{J})$. By Theorem 7.5, we can choose the element a positive and of $\mathcal{C}$-rank 1 and such that $\Phi_x(a)$ is a projection of algebraic rank one for all x in an open neighborhood $V \subseteq \operatorname{Max}\mathcal{C}$ of x_0. Part (b) of the transfer lemma then implies that, for every $x \in V$, there is a unique $L_x \in \operatorname{Prim}\mathcal{J}$ such that $\varphi_{\mathcal{J}}(L_x) = x$ and that $a + L_x$ is a projection of algebraic rank 1 in $\mathcal{J}/L_x$.

Since $\varphi_{\mathcal{J}}$ is continuous, $U_1 := \varphi_{\mathcal{J}}^{-1}(V)$ is open in $\operatorname{Prim}\mathcal{J}$. Further, by Lemma A.30 in [8], $U_2 := \{L \in \operatorname{Prim}\mathcal{J} : \|a + L\| > 1/2\}$ is open (note that $\operatorname{Spec}\mathcal{J}$ and $\operatorname{Prim}\mathcal{J}$ are naturally homeomorphic). Thus, $U := U_1 \cap U_2$ is open in $\operatorname{Prim}\mathcal{J}$. The assertion will follow once we have shown that $U = \{L_x : x \in V\}$. The inclusion $\supseteq$ follows by the definition of U.

The reverse inclusion $\subseteq$ is a consequence of the transfer lemma again. Indeed, let $L \in U$. Since $L \in U_1$, $\varphi_{\mathcal{J}}(L) = x_0 \in V$. By Lemma 7.6, $a + L$ is a projection of algebraic rank ≤ 1, implying that the norm of $a + L$ is either 0 or 1. Since $L \in U_2$, we conclude that $\|a + L\| = 1$. Hence, $a + L$ is a projection of rank 1, whence $L = L_{x_0}$ by the uniqueness assertion in the transfer lemma. □

8 Continuity of Local Ranks

If $\mathcal{J}$ is as in Theorem 7.7 and $x \in \operatorname{Max}\mathcal{C}$, then $\mathcal{J}/(I_x \cap \mathcal{J})$ is a dual algebra. Hence, every projection in $\mathcal{J}/(I_x \cap \mathcal{J})$ has a finite algebraic rank. We are interested in the dependence of these 'local ranks' on x.

Theorem 8.1 *Let $\mathcal{B}$ be a unital C^*-algebra, $\mathcal{C}$ a central C^*-subalgebra of $\mathcal{B}$ which contains the unit element, and $\mathcal{J} \subseteq \mathcal{J}_{\mathcal{C}}(\mathcal{B})$ a closed ideal of $\mathcal{B}$ which is generated by elements of $\mathcal{C}$-rank one. Assume that the function (7.2) is continuous for every $j \in \mathcal{J}$. Further let $a \in \mathcal{J}$ be such that $a + (I_x \cap \mathcal{J})$ is a projection for all x in an open set $U \subseteq \operatorname{Max}\mathcal{C}$. Then the function*

$$\operatorname{Max}\mathcal{C} \to \mathbb{Z}^+, \quad x \mapsto \operatorname{alg\,rank}(a + (I_x \cap \mathcal{J}))$$

is continuous on U.

Proof Let $x_0 \in U$ and $s := \operatorname{alg\,rank}(a + (I_{x_0} \cap \mathcal{J}))$. Then there are finitely many elementary components $\mathcal{E}_1, \ldots, \mathcal{E}_r$ of the dual algebra $\mathcal{J}/(I_{x_0} \cap \mathcal{J})$ and s mutually orthogonal projections $\pi_j^i \in \mathcal{E}_i$, $j = 1, \ldots, n_i$, of algebraic rank 1 such that

$$a + (I_{x_0} \cap \mathcal{J}) = \sum_{i=1}^{r} \sum_{j=1}^{n_i} \pi_j^i.$$

In each component $\mathcal{E}_i$, we fix a projection β_i of algebraic rank 1. Since $\mathcal{E}_i$ is elementary, there are partial isometries $\alpha_j^i \in \mathcal{E}_i$ such that

$$(\alpha_j^i)^* \alpha_j^i = \pi_j^i \quad \text{and} \quad \alpha_j^i (\alpha_j^i)^* = \beta_i \quad \text{for } j = 1, \ldots, n_i.$$

Now we continue the locally defined elements π_j^i, α_j^i and β_i to elements of $\mathcal{J}$. By Theorem 7.5, there are positive elements $b_i \in \mathcal{J}$ of central rank 1 such that $b_i + (I_{x_0} \cap \mathcal{J}) = \beta_i$ and $b_i + (I_x \cap \mathcal{J})$ is a projection of algebraic rank 1 for all x in a neighborhood W_i of x_0.

Next, it follows as in Lemmas 3.1 and 3.2 of [4], there are elements p_j^i, a_j^i, $b_i' \in \mathcal{J}$ and a neighborhood $V \subseteq \cap_i W_i$ of x_0 such that

$$p_j^i + (I_{x_0} \cap \mathcal{J}) = \pi_j^i, \quad a_j^i + (I_{x_0} \cap \mathcal{J}) = \alpha_j^i, \quad b_i' + (I_{x_0} \cap \mathcal{J}) = \beta_i$$

and that, for all $x \in V$, $b_i' + (I_x \cap \mathcal{J})$ is a projection and the $p_j^i + (I_x \cap \mathcal{J})$ are mutually orthogonal projections with

$$(a_j^i)^* a_j^i + (I_x \cap \mathcal{J}) = p_j^i + (I_x \cap \mathcal{J}), \quad a_j^i (a_j^i)^* + (I_x \cap \mathcal{J}) = b_i' + (I_x \cap \mathcal{J}). \tag{8.1}$$

Since $b_i + (I_{x_0} \cap \mathcal{J}) = b_i' + (I_{x_0} \cap \mathcal{J}) = \beta_i$ and the function (7.2) is upper semi-continuous, we may assume that

$$\|(b_i + (I_x \cap \mathcal{J})) - (b_i' + (I_x \cap \mathcal{J}))\| < 1 \quad \text{for all } x \in V$$

(otherwise we lessen V accordingly). Since $b_i + (I_x \cap \mathcal{J})$ is a projection of algebraic rank 1, we conclude via a Neumann series argument that $b_i' + (I_x \cap \mathcal{J})$ is of algebraic rank 1, too. But then (8.1) implies that $p_j^i + (I_x \cap \mathcal{J})$ is of algebraic rank 1 and that $p := \sum_{i=1}^{r} \sum_{j=1}^{n_i} p_j^i + (I_x \cap \mathcal{J})$ is a projection of algebraic rank s for all x in the neighborhood V of x_0. □

References

1. M.C.F. Berglund, Ideal C^*-algebras. Duke Math. J. **40**, 241–257 (1973)
2. A. Böttcher, B. Silbermann, *Introduction to Large Truncated Toeplitz Matrices* (Springer, Berlin, 1999)
3. J. Dixmier, *C^*-Algebras* (North Holland Publishing Company, Amsterdam, 1982)
4. J.M.G. Fell, The structure of algebras of operator fields. Acta Math. **106**, 233–280 (1961)

5. R. Hagen, S. Roch, B. Silbermann, *C^*-Algebras and Numerical Analysis* (Dekker, New York, 2001)
6. V. Nistor, N. Prudhon, Exhaustive families of representations and spectra of pseudodifferential operators. J. Oper. Theory **78**, 247–279 (2017)
7. G.K. Pedersen, *C^*-Algebras and Their Automorphism Groups* (Academic, London, 1979)
8. I. Raeburn, D.P. Williams, *Morita Equivalence and Continuous Trace C^*-Algebras* (American Mathematical Society, Providence, 1998)
9. S. Roch, Algebras of approximation sequences: Fractality, in in *Problems and Methods in Mathematical Physics*. Operator Theory: Advances and Applications, vol. 121 (Birkhäuser, Basel, 2001), pp. 471–497
10. S. Roch, Algebras of approximation sequences: fredholm theory in fractal algebras. Studia Math. **150**, 53–77 (2002)
11. S. Roch, *Finite Sections of Band-Dominated Operators*, vol. 191. (Memoirs of the American Mathematical Societ, Providence, 2008), p. 895
12. S. Roch, Extension-restriction theorems for algebras of approximation sequences, in *Operator Theory, Operator Algebras, ans Matrix Theory*. Operator Theory: Advances and Applications, vol. 267 (Birkhäuser, Basel 2018), pp. 261–284
13. S. Roch, Beyond fractality: Piecewise fractal and quasifractal algebras, in *The Diversity and Beauty of Applied Operator Theory*. Operator Theory: Advances and Applications, vol. 268 (Birkhäuser, Basel 2018), pp. 413–428
14. S. Roch, P.A. Santos, B. Silbermann, *Non-commutative Gelfand Theories. A Tool-kit for Operator Theorists and Numerical Analysts* (Universitext, Springer, London, 2011)
15. S. Roch, B. Silbermann, C^*-algebra techniques in numerical analysis. J. Oper. Theory **35**, 241–280 (1996)
16. D. Williams, *Crossed Products of C^*-Algebras* (American Mathematical Society, Providence, 2007)

Dilation Theory: A Guided Tour

Orr Moshe Shalit

Abstract Dilation theory is a paradigm for studying operators by way of exhibiting an operator as a compression of another operator which is in some sense well behaved. For example, every contraction can be dilated to (i.e., is a compression of) a unitary operator, and on this simple fact a penetrating theory of non-normal operators has been developed. In the first part of this survey, I will leisurely review key classical results on dilation theory for a single operator or for several commuting operators, and sample applications of dilation theory in operator theory and in function theory. Then, in the second part, I will give a rapid account of a plethora of variants of dilation theory and their applications. In particular, I will discuss dilation theory of completely positive maps and semigroups, as well as the operator algebraic approach to dilation theory. In the last part, I will present relatively new dilation problems in the noncommutative setting which are related to the study of matrix convex sets and operator systems, and are motivated by applications in control theory. These problems include dilating tuples of noncommuting operators to tuples of commuting normal operators with a specified joint spectrum. I will also describe the recently studied problem of determining the optimal constant $c = c_{\theta,\theta'}$, such that every pair of unitaries U, V satisfying $VU = e^{i\theta}UV$ can be dilated to a pair of cU', cV', where U', V' are unitaries that satisfy the commutation relation $V'U' = e^{i\theta'}U'V'$. The solution of this problem gives rise to a new and surprising application of dilation theory to the continuity of the spectrum of the almost Mathieu operator from mathematical physics.

Keywords Dilations · Isometric dilation · Unitary dilation · Matrix convex sets · q-commuting unitaries · Completely positive maps · CP-semigroups

Mathematics Subject Classification (2010) 47A20, 46L07, 46L55, 47A13, 47B32, 47L25

O. M. Shalit (✉)
Faculty of Mathematics, Technion - Israel Institute of Technology, Haifa, Israel
e-mail: oshalit@technion.ac.il

M. A. Bastos et al. (eds.), *Operator Theory, Functional Analysis and Applications*,
Operator Theory: Advances and Applications 282,
https://doi.org/10.1007/978-3-030-51945-2_28

Dilation theory is a collection of results, tools, techniques, tricks, and points of view in operator theory and operator algebras, that fall under the unifying idea that one can learn a lot about an operator (or family of operators, or a map, etc.) by viewing it as "a part of" another, well understood operator. This survey on dilation theory consists of three parts. The first part is a stand-alone exposition aimed at giving an idea of what dilation theory is about by describing several representative results and applications that are, in my opinion, particularly interesting. The climax of the first part is in Sect. 4, where as an application of the material in the first three sections, we see how to prove the Pick interpolation theorem using the commutant lifting theorem. Anyone who took a course in operator theory can read Part 1.

Out of the theory described in the first part, several different research directions have developed. The second part of this survey is an attempt to give a quick account of some of these directions. In particular, we will cover Stinespring's dilation theorem, and the operator algebraic approach to dilation theory that was invented by Arveson. This survey up to Sect. 7 contains what everyone working in dilation theory and/or nonselfadjoint operator algebras should know. I will also cover a part of the dilation theory of CP-semigroups, and take the opportunity to report on my work with Michael Skeide, which provides our current general outlook on the subject.

In the third and last part I will survey some recent dilation results in the noncommutative setting, in particular those that have been motivated by the study of matrix convex sets. Then I will focus on my recent joint work with Malte Gerhold, where we study the problem of dilating q-commuting unitaries. Experts on dilation theory can read the last three sections in this survey independently.

I made an effort to include in this survey many applications of dilation theory. The theory is interesting and elegant in itself, but the applications give it its vitality. I believe that anyone, including experts in dilation theory, will be able to find in this survey an interesting application which they have not seen before.

Some results are proved and others are not. For some results, only an idea of the proof is given. The guiding principle is to include proofs that somehow together convey the essence or philosphy of the field, so that the reader will be able to get the core of the theory from this survey, and then be able to follow the references for more.

As for giving references: this issue has given me a lot of headaches. On the one hand, I would like to give a historically precise picture, and to give credit where credit is due. On the other hand, making too big of a fuss about this might result in an unreadable report, that looks more like a legal document than the inviting survey that I want this to be. Some results have been rediscovered and refined several times before reaching their final form. Who should I cite? My solution was to always prefer the benefit of the reader. For "classical" results, I am very happy to point the reader to an excellent textbook or monograph, that contains a proof, as well as detailed references and sometimes also historical remarks. I attach a specific paper to a theorem only when it is clear-cut and useful to do so. In the case of recent results, I sometimes give all relevant references and an account of the historical development, since this appears nowhere else.

There are other ways to present dilation theory, and by the end of the first section the reader will find references to several alternative sources. Either because of my ignorance, or because I had to make choices, some things were left out. I have not been able to cover all topics that could fall under the title, nor did I do full justice to the topics covered. After all this is just a survey, and that is the inevitable nature of the genre.

Part 1: An Exposition of Classical Dilation Theory

1 The Concept of Dilations

The purpose of this introductory section is to present the notion of *dilation*, and to give a first indication that this notion is interesting and can be useful.

Let $\mathcal{H}$ be a Hilbert space, and $T \in B(\mathcal{H})$ be a **contraction**, that is, T is an operator such that $\|T\| \leq 1$. Then $I - T^*T \geq 0$, and so we can define $D_T = \sqrt{I - T^*T}$. Halmos [72] observed that the simple construction

$$U = \begin{pmatrix} T & D_{T^*} \\ D_T & -T^* \end{pmatrix}$$

gives rise to a unitary operator on $\mathcal{H} \oplus \mathcal{H}$. Thus, every contraction T is a **compression** of a unitary U, meaning that

$$T = P_{\mathcal{H}} U\big|_{\mathcal{H}}, \tag{1.1}$$

where $P_{\mathcal{H}}$ denotes the orthogonal projection of $\mathcal{H} \oplus \mathcal{H}$ onto $\mathcal{H} \oplus \{0\} \cong \mathcal{H}$. In this situation we say that U is a **dilation** of T, and below we shall write $T \prec U$ to abbreviate that U is a dilation of T.

This idea can be pushed further. Let $\mathcal{K} = \mathcal{H}^{N+1} = \mathcal{H} \oplus \cdots \oplus \mathcal{H}$ be the $(N+1)$th-fold direct sum of $\mathcal{H}$ with itself, and consider the following $(N+1) \times (N+1)$ operator matrix

$$U = \begin{pmatrix} T & 0 & 0 & \cdots & 0 & D_{T^*} \\ D_T & 0 & 0 & \cdots & 0 & -T^* \\ 0 & I & 0 & \cdots & 0 & 0 \\ 0 & 0 & I & & & 0 \\ \vdots & \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & 0 & I & 0 \end{pmatrix}. \tag{1.2}$$

Egerváry [58] observed that U is unitary on $B(\mathcal{K})$, and moreover, that

$$U^k = \begin{pmatrix} T^k & * \\ * & * \end{pmatrix}, \quad k = 1, 2, \ldots, N, \tag{1.3}$$

in other words, if we identify $\mathcal{H}$ with the first summand of $\mathcal{K}$, then

$$p(T) = P_{\mathcal{H}} p(U)\big|_{\mathcal{H}} \tag{1.4}$$

for every polynomial p of degree at most N. Such a dilation was called an N-**dilation** in [95]. Thus, an operator U satisfying (1.1) might be referred to as a 1-**dilation**, however, the recent ubiquity of 1-dilations has led me to refer to it simply as a **dilation**.

We see that, in a sense, every contraction is a "part of" a unitary operator. Unitaries are a very well understood class of operators, and contractions are as general a class as one can hope to study. Can we learn anything interesting from the dilation picture?

Theorem 1.1 (von Neumann's Inequality [168]) *Let T be a contraction on some Hilbert space $\mathcal{H}$. Then, for every polynomial $p \in \mathbb{C}[z]$,*

$$\|p(T)\| \leq \sup_{|z|=1} |p(z)|. \tag{1.5}$$

Proof Suppose that the degree of p is N. Construct U on $\mathcal{K} = \mathcal{H} \oplus \cdots \oplus \mathcal{H}$ (direct sum $N + 1$ times) as in (1.2). Using (1.4), we find that

$$\|p(T)\| = \|P_{\mathcal{H}} p(U)\big|_{\mathcal{H}}\| \leq \|p(U)\| = \sup_{z \in \sigma(U)} |p(z)|,$$

by the spectral theorem, where $\sigma(U)$ denotes the spectrum of U. Since for every unitary U, the spectrum $\sigma(U)$ is contained in the unit circle $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$, the proof is complete. ∎

Remark 1.2 The above proof is a minor simplification of the proof of von Neumann's inequality due to Sz.-Nagy [163], which uses the existence of a unitary **power dilation** (see the next section). The simplification becomes significant when $\dim \mathcal{H} < \infty$, because then U is a unitary which acts on a finite dimensional space, and the spectral theorem is then a truly elementary matter. Note that even the case $\dim \mathcal{H} = 1$ is not entirely trivial: in this case von Neumann's inequality is basically the maximum modulus principle in the unit disc; as was observed in [145], this fundamental results can be proved using linear algebra!

Remark 1.3 A **matrix valued polynomial** is a function $z \mapsto p(z) \in M_n$ where $p(z) = \sum_{k=0}^{N} A_k z^k$, and $A_1, \ldots, A_k \in M_n = M_n(\mathbb{C})$ (the $n \times n$ matrices over $\mathbb{C}$). Equivalently, a matrix valued polynomial is an $n \times n$ matrix of polynomials $p = (p_{ij})$, where the ijth entry is given by $p_{ij}(z) = \sum_{k=0}^{N} (A_k)_{ij} z^k$. If $T \in B(\mathcal{H})$, then

we may evaluate a matrix valued polynomial p at T by setting $p(T) = \sum_{k=0}^{N} A_k \otimes T^k \in M_n \otimes B(\mathcal{H})$, or, equivalently, $p(T)$ is the $n \times n$ matrix over $B(\mathcal{H})$ with ijth entry equal to $p_{ij}(T)$ (the operator $p(T)$ acts on the direct sum of $\mathcal{H}$ with itself n times). It is not hard to see that if p is a matrix valued polynomial with values in M_n and T is a contraction, then the inequality (1.5) still holds but with $|p(z)|$ replaced by $\|p(z)\|_{M_n}$: first one notes that it holds for unitary operators, and then one obtains it for a general contraction by the dilation argument that we gave.

The construction (1.2) together with Theorem 1.1 illustrate what *dilation theory* is about and how it can be used: every object in a general class of objects (here, contractions) is shown to be "a part of" an object in a smaller, better behaved class (here, unitaries); the objects in the better behaved class are well understood (here, by the spectral theorem), and thus the objects in the general class inherit some good properties. The example we have just seen is an excellent one, since proving von Neumann's inequality for non-normal contractions is not trivial. The simple construction (1.2) and its multivariable generalizations have other applications, for example they lead to concrete cubature formulas and operator valued cubature formulas [95, Section 4.3]. In Sect. 4 we will examine in detail a deeper application of dilation theory—an operator theoretic proof of Pick's interpolation theorem. Additional applications are scattered throughout this survey.

Before continuing I wish to emphasize that "dilation theory" and even the word "dilation" itself mean different things to different people. As we shall see further down the survey, the definition changes, as do the goals and the applications. Besides the expository essay [95], the subject is presented nicely in the surveys [8] and [13], and certain aspects are covered in books, e.g., [2, 118], and [121] (the forthcoming book [26] will surely be valuable when it appears). Finally, the monograph [23] is an indispensable reference for anyone who is seriously interested in dilation theory of contractions.

2 Classical Dilation Theory of a Single Contraction

2.1 Dilations of a Single Contraction

It is quite natural to ask whether one can modify the construction (1.2) so that (1.3) holds for all $k \in \mathbb{N}$, and not just up to some power N. We will soon see that the answer is affirmative. Let us say that an operator $U \in B(\mathcal{K})$ is a **power dilation** of $T \in B(\mathcal{H})$, if $\mathcal{H}$ is a subspace of $\mathcal{K}$ and if $T^k = P_{\mathcal{H}} U^k\big|_{\mathcal{H}}$ for all $k \in \mathbb{N} = \{0, 1, 2, \ldots\}$. The reader should be warned that this is not the standard terminology used, in the older literature one usually finds the word **dilation** used to describe what we just called a power dilation, whereas the concept of N-dilation does not appear much (while in the older-older literature one can again find *power dilation*).

In this and the next sections I will present what I and many others refer to as *classical dilation theory*. By and large, this means the theory that has been pushed

and organized by Sz.-Nagy and Foias (though there are many other contributors), and appears in the first chapter of [23]. The book [23] is the chief reference for classical dilation theory. The proofs of most of the results in this and the next section, as well as references, further comments and historical remarks can be found there.

Theorem 2.1 (Sz.-Nagy's Unitary Dilation Theorem [163]) *Let T be a contraction on a Hilbert space $\mathcal{H}$. Then there exists a Hilbert space $\mathcal{K}$ containing $\mathcal{H}$ and a unitary U on $\mathcal{K}$, such that*

$$T^k = P_{\mathcal{H}} U^k\big|_{\mathcal{H}} \quad , \quad \text{for all } k = 0, 1, 2, \ldots \tag{2.1}$$

*Moreover, $\mathcal{K}$ can be chosen to be **minimal** in the sense that the smallest reducing subspace for U that contains $\mathcal{H}$ is $\mathcal{K}$. If $U_i \in B(\mathcal{K}_i)$ $(i = 1, 2)$ are two minimal unitary dilations, then there exists a unitary $W : \mathcal{K}_1 \to \mathcal{K}_2$ acting as the identity on $\mathcal{H}$ such that $U_2 W = W U_1$.*

One can give a direct proof of existence of the unitary dilation by just writing down an infinite operator matrix U acting on $\ell^2(\mathbb{Z}, \mathcal{H}) = \oplus_{n\in\mathbb{Z}}\mathcal{H}$, similarly to (1.2). A minimal dilation is then obtained by restricting U to the reducing subspace $\bigvee_{n\in\mathbb{Z}} U^n\mathcal{H}$ (the notation means the closure of the span of the subspaces $U^n\mathcal{H}$). The uniqueness of the minimal unitary dilation is then a routine matter. We will follow a different path, that requires us to introduce another very important notion: **the minimal isometric dilation**.

Before presenting the isometric dilation theorem, it is natural to ask whether it is expected to be of any use. A unitary dilation can be useful because unitaries are "completely understood" thanks to the spectral theorem. Are isometries well understood? The following theorem shows that, in a sense, isometries are indeed very well understood.

For a Hilbert space $\mathcal{G}$, we write $\ell^2(\mathbb{N}, \mathcal{G})$ for the direct sum $\oplus_{n\in\mathbb{N}}\mathcal{G}$. The **unilateral shift of multiplicity** $\dim \mathcal{G}$ (or simply the shift) is the operator $S : \ell^2(\mathbb{N}, \mathcal{G}) \to \ell^2(\mathbb{N}, \mathcal{G})$ given by

$$S(g_0, g_1, g_2, \ldots) = (0, g_0, g_1, \ldots).$$

The space $\mathcal{G}$ is called the **multiplicity space**. Clearly, the shift is an isometry. Similarly, the **bilateral shift** on $\ell^2(\mathbb{Z}, \mathcal{G})$ is defined to be the operator

$$U(\ldots, g_{-2}, g_{-1}, \boxed{g_0}, g_1, g_2, \ldots) = (\ldots g_{-3}, g_{-2}, \boxed{g_{-1}}, g_0, g_1, \ldots)$$

where we indicate with a box the element at the 0th summand of $\ell^2(\mathbb{Z}, \mathcal{G})$.

Theorem 2.2 (Wold Decomposition) *Let V be an isometry on a Hilbert space $\mathcal{H}$. Then there exists a (unique) direct sum decomposition $\mathcal{H} = \mathcal{H}_s \oplus \mathcal{H}_u$ such that $\mathcal{H}_s$ and $\mathcal{H}_u$ are reducing for V, and such that $V\big|_{\mathcal{H}_s}$ is unitarily equivalent to a unilateral shift and $V\big|_{\mathcal{H}_u}$ is unitary.*

For the proof, the reader has no choice but to define $\mathcal{H}_u = \cap_{n\geq 0} V^n\mathcal{H}$. Once one shows that $\mathcal{H}_u$ is reducing, it remains to show that $V\big|_{\mathcal{H}_u^\perp}$ is a unilateral shift. Hint: the multiplicity space is $\mathcal{H} \ominus V\mathcal{H}$, (this suggestive notation is commonly used in the theory; it means "the orthogonal complement of $V\mathcal{H}$ inside $\mathcal{H}$", which is in this case just $(V\mathcal{H})^\perp$).

Theorem 2.3 (Sz.-Nagy's Isometric Dilation Theorem) *Let T be a contraction on a Hilbert space $\mathcal{H}$. Then there exists a Hilbert space $\mathcal{K}$ containing $\mathcal{H}$ and an isometry V on $\mathcal{K}$, such that*

$$T^* = V^*\big|_{\mathcal{H}} \tag{2.2}$$

and in particular, V is a power dilation of T. Moreover, $\mathcal{K}$ can be chosen to be **minimal** *in the sense that the minimal invariant subspace for V that contains $\mathcal{H}$ is $\mathcal{K}$. If $V_i \in B(\mathcal{K}_i)$ $(i = 1, 2)$ are two minimal isometric dilations, then there exists a unitary $W : \mathcal{K}_1 \to \mathcal{K}_2$ acting as the identity on $\mathcal{H}$ such that $V_2W = WV_1$.*

Proof Set $D_T = (I - T^*T)^{1/2}$—this is the so called **defect operator**, which measures how far T is from being an isometry—and let $\mathcal{D} = \overline{D_T(\mathcal{H})}$. Construct $\mathcal{K} = \mathcal{H} \oplus \mathcal{D} \oplus \mathcal{D} \oplus \ldots$, in which $\mathcal{H}$ is identified as the first summand. Now we define, with respect to the above decomposition of $\mathcal{K}$, the block operator matrix

$$V = \begin{pmatrix} T & 0 & 0 & 0 & 0 & \cdots \\ D_T & 0 & 0 & 0 & 0 & \cdots \\ 0 & I_{\mathcal{D}} & 0 & 0 & 0 & \cdots \\ 0 & 0 & I_{\mathcal{D}} & 0 & 0 & \\ 0 & 0 & 0 & I_{\mathcal{D}} & 0 & \\ \vdots & \vdots & & & \ddots & \ddots \end{pmatrix}$$

which is readily seen to satisfy $V^*\big|_{\mathcal{H}} = T^*$. That V is an isometry, and the fact that $\mathcal{K} = \bigvee_{n\in\mathbb{N}} V^n\mathcal{H}$, can be proved directly and without any pain.

The uniqueness is routine, but let's walk through it for once. If $V_i \in B(\mathcal{K}_i)$ are two minimal isometric dilations, then we can define a map W on $\operatorname{span}\{V_1^n h : n \in \mathbb{N}, h \in \mathcal{H}\} \subseteq \mathcal{K}_1$ by first prescribing

$$WV_1^n h = V_2^n h \in \mathcal{K}_2.$$

This map preserves inner products: assuming that $m \leq n$, and using the fact that V_i is an isometric dilation of T, we see

$$\langle WV_1^m g, WV_1^n h\rangle = \langle V_2^m g, V_2^n h\rangle = \langle g, V_2^{n-m} h\rangle = \langle g, T^{n-m} h\rangle = \langle V_1^m g, V_1^n h\rangle.$$

The map W therefore well defines an isometry from the dense subspace $\text{span}\{V_1^n h : n \in \mathbb{N}, h \in \mathcal{H}\} \subseteq \mathcal{K}_1$ onto the dense subspace $\text{span}\{V_2^n h : n \in \mathbb{N}, h \in \mathcal{H}\} \subseteq \mathcal{K}_2$, and therefore extends to a unitary which, by definition, intertwines V_1 and V_2. ■

Note that the minimal isometric dilation is actually a **coextension**: $T^* = V\big|_{\mathcal{H}}^*$. A coextension is always a power dilation, so $T^n = P_{\mathcal{H}} V^n\big|_{\mathcal{H}}$ for all $n \in \mathbb{N}$, but, of course, the converse is not true (see Theorem 3.10 and (3.6) below for the general form of a power dilation). Note also that the minimality requirement is more stringent than the minimality required from the minimal unitary dilation. The above proof of existence together with the uniqueness assertion actually show that every isometry which is a power dilation of T and is minimal in the sense of the theorem, is in fact a coextension.

Because the minimal isometric dilation V is actually a coextension of T, the adjoint V^* is a coisometric extension of T^*. Some people prefer to speak about **coisometric extensions** instead of isometric coextensions.

Once the existence of an isometric dilation is known, the existence of a unitary dilation follows immediately, by the Wold decomposition. Indeed, given a contraction T, we can dilate it to an isometry V. Since $V \cong S \oplus V_u$, where S is a unilateral shift and V_u is a unitary, we can define a unitary dilation of T by $U \oplus V_u$, where U is the bilateral shift of the same multiplicity as S. This proves the existence part of Theorem 2.1; the uniqueness of the minimal unitary dilation is proved as for the minimal isometric one.

2.2 *A Glimpse at Some Applications of Single Operator Dilation Theory*

The minimal unitary dilation of a single contraction can serve as the basis of the development of operator theory for non-normal operators. This idea was developed to a high degree by Sz.-Nagy and Foias and others; see the monograph [23] ([23] also contains references to alternative approaches to non-normal operators, in particular the theories of de Branges–Rovnyak, Lax–Phillips, and Livsič and his school). The minimal unitary dilation can be used to define a refined functional calculus on contractions, it can be employed to analyze one-parameter semigroups of operators, it provides a "functional model" by which to analyze contractions and by which they can be classified, and it has led to considerable progress in the study of invariant subspaces.

To sketch just one of the above applications, let us briefly consider the functional calculus (the following discussion might be a bit difficult for readers with little background in function theory and measure theory; they may skip to the beginning of Sect. 3 without much loss). We let $H^\infty = H^\infty(\mathbb{D})$ denote the algebra of bounded analytic functions on the open unit disc $\mathbb{D}$. Given an operator $T \in B(\mathcal{H})$, we wish to define a functional calculus $f \mapsto f(T)$ for all $f \in H^\infty$. If the spectrum of T is contained in $\mathbb{D}$, then we can apply the holomorphic functional calculus to T to

define a homomorphism $f \mapsto f(T)$ from the algebra $\mathcal{O}(\mathbb{D})$ of analytic functions on $\mathbb{D}$ into $B(\mathcal{H})$. In fact, if $f \in \mathcal{O}(\mathbb{D})$ and $\sigma(T) \subset \mathbb{D}$, then we can simply plug T into the power series of f. Thus, in this case we know how to define $f(T)$ for all $f \in H^\infty$.

Now, suppose that T is a contraction, but that $\sigma(T)$ is not contained in the *open* disc $\mathbb{D}$. Given a bounded analytic function $f \in H^\infty$, how can we define $f(T)$? Note that the holomorphic functional calculus cannot be used, because $\sigma(T)$ contains points on the circle $\mathbb{T} = \partial\mathbb{D}$, while not every $f \in H^\infty$ extends to a holomorphic function on a neighborhood of the closed disc.

The first case that we can treat easily is the case when f belongs to the **disc algebra** $A(\mathbb{D}) \subset H^\infty$, which is the algebra of all bounded analytic functions on the open unit disc that extend continuously to the closure $\overline{\mathbb{D}}$. This case can be handled using von Neumann's inequality (Theorem 1.5), which is an immediate consequence of Theorem 2.1. Indeed, it is not very hard to show that $A(\mathbb{D})$ is the closure of the polynomials with respect to the supremum norm $\|f\|_\infty = \sup_{|z|\leq 1} |f(z)|$. If p_n is a sequence of polynomials that converges uniformly on $\overline{\mathbb{D}}$ to f, and T is a contraction, then von Neumann's inequality implies that $p_n(T)$ is a Cauchy sequence, so we can define $f(T)$ to be $\lim_n p_n(T)$. It is not hard to show that the functional calculus $A(\mathbb{D}) \ni f \mapsto f(T)$ has all the properties one wishes for: it is a homomorphism extending the polynomial functional calculus, it is continuous, and it agrees with the continuous functional calculus if T is normal.

Defining a functional calculus for H^∞ is a more delicate matter, but here again the unitary dilation leads to a resolution. The rough idea is that we can look at the minimal unitary dilation U of T, and use spectral theory to analyze what can be done for U. If the spectral measure of U is absolutely continuous with respect to Lebesgue measure on the unit circle, then it turns out that we can define $f(U)$ for all $f \in H^\infty$, and then we can simply define $f(T)$ to be the compression of $f(U)$ to $\mathcal{H}$. In this case the functional calculus $f \mapsto f(T)$ is a homomorphism that extends the polynomial and holomorphic functional calculi, it is continuous in the appropriate sense, and it agrees with the Borel functional calculus when T is normal. (Of course, this only becomes useful if one can find conditions that guarantee that the minimal unitary dilation of T has absolutely continuous spectral measure. A contraction T is said to be **completely nonunitary** (c.n.u.) if is has no reducing subspace $\mathcal{M}$ such that the restriction $T\big|_{\mathcal{M}}$ is unitary. Every contraction splits as a direct sum $T = T_0 \oplus T_1$, where T_0 is unitary and T_1 is c.n.u. Sz.-Nagy and Foias have shown that if T is c.n.u., then the spectral measure of its minimal unitary dilation is absolutely continuous.)

If the spectral measure of U is not absolutely continuous with respect to Lebesgue measure, then there exits a subalgebra H^∞_U of H^∞ for which there exists a functional calculus $f \mapsto f(U)$, and then one can compress to get $f(T)$; it was shown that H^∞_U is precisely the subalgebra of functions in H^∞ on which $f \mapsto f(T)$ is a well defined homomorphism. See [23, Chapter III] for precise details.

There are two interesting aspects to note. First, an intrinsic property of the minimal dilation—the absolute continuity of its spectral measure—provides us with

nontrivial information about T (whether or not it has an H^∞-functional calculus). The second interesting aspect is that there do exist interesting functions $f \in H^\infty$ that are not holomorphic on a neighborhood of the closed disc (and are not even continuous up to the boundary) for which we would like to evaluate $f(T)$. This technical tool has real applications. See [23, Section III.8], for example.

3 Classical Dilation Theory of Commuting Contractions

3.1 Dilations of Several Commuting Contractions

The manifold applications of the unitary dilation of a contraction on a Hilbert space motivated the question (which is appealing and natural in itself, we must admit) whether the theory can be extended in a sensible manner to families of operators. The basic problem is: given commuting contractions $T_1, \ldots, T_d \in B(\mathcal{H})$, to determine whether there exist commuting isometries/unitaries $U_1, \ldots, U_d$ on a larger Hilbert space $\mathcal{K} \supseteq \mathcal{H}$ such that

$$T_1^{n_1} \cdots T_d^{n_d} = P_{\mathcal{H}} U_1^{n_1} \cdots U_d^{n_d}\big|_{\mathcal{H}} \tag{3.1}$$

for all $(n_1, \ldots, n_d) \in \mathbb{N}^d$. Such a family $U_1, \ldots, U_d$ is said to be an **isometric/unitary dilation** (I warned you that the word is used differently in different situations!).

Clearly, it would be nice to have a unitary dilation, because commuting unitaries are completely understood by spectral theory. On the other hand, isometric dilations might be easier to construct. Luckily, we can have the best of both worlds, according to the following theorem of Itô and Brehmer (see [23, Proposition I.6.2]).

Theorem 3.1 *Every family of commuting isometries has a commuting unitary extension.*

Proof The main idea of the proof is that given commuting isometries $V_1, \ldots, V_d$ on a Hilbert space $\mathcal{H}$, one may extend them to commuting isometries $W_1, \ldots, W_d$ such that (a) W_1 is a unitary, and (b) if V_i is unitary then W_i is a unitary. Given that this is possible, one my repeat the above process d times to obtain a unitary extension. The details are left to the reader. ■

In particular, every family of commuting isometries has a commuting unitary dilation. Thus, a family of commuting contractions $T_1, \ldots, T_d \in B(\mathcal{H})$ has an isometric dilation if and only if it has a unitary dilation.

Theorem 3.2 (Andô's Isometric Dilation Theorem [9]) *Let T_1, T_2 be two commuting contractions on a Hilbert space $\mathcal{H}$. Then there exists a Hilbert space $\mathcal{K} \supseteq \mathcal{H}$ and two commuting isometries V_1, V_2 on $\mathcal{K}$ such that*

$$T_1^{n_1} T_2^{n_2} = P_{\mathcal{H}} V_1^{n_2} V_2^{n_2}\big|_{\mathcal{H}} \quad \text{for all } n_1, n_2 \in \mathbb{N}. \tag{3.2}$$

In fact, V_1, V_2 can be chosen such that $V_i^*\big|_{\mathcal{H}} = T_i^*$ for $i = 1, 2$.

In other words, every pair of contractions has an isometric dilation, and in fact, an isometric coextension.

Proof The proof begins similarly to the proof of Theorem 2.3: we define the Hilbert space $\mathcal{K} = \oplus_{n\in\mathbb{N}}\mathcal{H} = \mathcal{H} \oplus \mathcal{H} \oplus \cdots$, and we define two isometries W_1, W_2 by

$$W_i(h_0, h_1, h_2, \ldots) = (T_i h_0, D_{T_i} h_0, 0, h_1, h_2, \ldots)$$

for $i = 1, 2$. These are clearly isometric coextensions, but they do not commute:

$$\begin{aligned} W_j W_i(h_0, h_1, h_2, \ldots) &= W_j(T_i h_0, D_{T_i} h_0, 0, h_1, h_2, \ldots) \\ &= (T_j T_i h_0, D_{T_j} T_i h_0, 0, D_{T_i} h_0, 0, h_1, h_2, \ldots). \end{aligned}$$

Of course, in the zeroth entry we have equality $T_1 T_2 h_0 = T_2 T_1 h_0$ and from the fifth entry on we have $(h_1, h_2, \ldots) = (h_1, h_2, \ldots)$. The problem is that usually $D_{T_1} T_2 h_0 \neq D_{T_2} T_1 h_0$ and $D_{T_1} h_0 \neq D_{T_2} h_0$. However ,

$$\begin{aligned} \|D_{T_1} T_2 h_0\|^2 + \|D_{T_2} h_0\|^2 &= \langle T_2^*(I - T_1^* T_1) T_2 h_0, h_0\rangle + \langle (I - T_2^* T_2) h_0, h_0\rangle \\ &= \langle (I - T_2^* T_1^* T_1 T_2) h_0, h_0\rangle \\ &= \|D_{T_2} T_1 h_0\|^2 + \|D_{T_1} h_0\|^2, \end{aligned}$$

and this allows us to define a unitary operator $U_0 : \mathcal{G} := \mathcal{H} \oplus \mathcal{H} \oplus \mathcal{H} \oplus \mathcal{H} \to \mathcal{G}$ that satisfies

$$U_0(D_{T_1} T_2 h_0, 0, D_{T_2} h_0, 0) = (D_{T_2} T_1 h_0, 0, D_{T_1} h_0, 0).$$

Regrouping $\mathcal{K} = \mathcal{H} \oplus \mathcal{G} \oplus \mathcal{G} \oplus \cdots$, we put $U = I_{\mathcal{H}} \oplus U_0 \oplus U_0 \oplus \cdots$, and now we define $V_1 = U W_1$ and $V_2 = W_2 U^{-1}$. The isometries V_1 and V_2 are isometric coextensions—multiplying by U and U^{-1} did not spoil this property of W_1, W_2. The upshot is that V_1 and V_2 commute; we leave this for the reader to check. ■

As a consequence (by Theorem 3.1),

Theorem 3.3 (Andô's Unitary Dilation Theorem [9]) *Every pair of contractions has a unitary dilation.*

One can also get minimal dilations, but it turns out that in the multivariable setting minimal dilations are not unique, so they are not canonical and don't play a prominent role. Once the existence of the unitary dilation is known, the following two-variable version of von Neumann's inequality follows just as in the proof of Theorem 1.1.

Theorem 3.4 *Let T_1, T_2 be two commuting contractions on a Hilbert space $\mathcal{H}$. Then for every complex two-variable polynomial p,*

$$\|p(T)\| \leq \sup_{z \in \mathbb{T}^2} |p(z)|.$$

Here and below we use the shorthand notation $p(T) = p(T_1, \ldots, T_d)$ whenever p is a polynomial in d variables and $T = (T_1, \ldots, T_d)$ is a d-tuple of operators. The proof of the above theorem (which is implicit in the lines preceding it) gives rise to an interesting principle: whenever we have a unitary or a normal dilation then we have a von Neumann type inequality. This principle can be used in reverse, to show that for three or more commuting contractions there might be no unitary dilation, in general.

Example 3.5 There exist three contractions T_1, T_2, T_3 on a Hilbert space $\mathcal{H}$ and a complex polynomial p such that

$$\|p(T)\| > \|p\|_\infty := \sup_{z \in \mathbb{T}^3} |p(z)|.$$

Consequently, T_1, T_2, T_3 have no unitary, and hence also no isometric, dilation. There are several concrete examples. The easiest to explain, in my opinion, is the one presented by Crabb and Davie [39]. One takes a Hilbert space $\mathcal{H}$ of dimension 8 with orthonormal basis $e, f_1, f_2, f_3, g_1, g_2, g_3, h$, and defines partial isometries T_1, T_2, T_3 by

$$\begin{aligned} T_i e &= f_i \\ T_i f_i &= -g_i \\ T_i f_j &= g_k \ , \ k \neq i, j \\ T_i g_j &= \delta_{ij} h \\ T_i h &= 0 \end{aligned}$$

for $i, j, k = 1, 2, 3$. Obviously, these are contractions, and checking that $T_i T_j = T_j T_i$ is probably easier than you guess. Now let $p(z_1, z_2, z_3) = z_1 z_2 z_3 - z_1^3 - z_2^3 - z_3^3$. Directly evaluating we see that $p(T_1, T_2, T_3)e = 4h$, so that $\|p(T_1, T_2, T_3)\| \geq 4$. On the other hand, it is elementary to show that $|p(z)| < 4$ for all $z \in \mathbb{T}^3$, so by compactness of $\mathbb{T}^3$ we get $\|p\|_\infty < 4$, as required.

Remark 3.6 At more or less the same time that the above example appeared, Kaijser and Varopoulos discovered three 5×5 commuting contractive matrices that do not satisfy von Neumann's inequality [165]. On the other hand, it was known (see [56, p. 21]) that von Neumann's inequality holds for any d-tuple of 2×2 matrices, in fact, every such d-tuple has a commuting unitary dilation. It was therefore begged of operator theorists to decide whether or not von Neumann's inequality holds for

3-tuples of 3×3 and 4×4 commuting contractive matrices. Holbrook found a 4×4 counter example in 2001 [78], and the question of whether von Neumann's inequality holds for three 3×3 contractions remained outrageously open until finally, only very recently, Knese [87] showed how results of Kosiński on the three point Pick interpolation problem in the polydisc [88] imply that in the 3×3 case the inequality holds (it is still an open problem whether or not every three commuting 3×3 contractions have a commuting unitary dilation; the case of four 3×3 contractions was settled negatively in [35]).

It is interesting to note that the first example of three contractions that do not admit a unitary dilation did not involve a violation of a von Neumann type inequality. Parrott [110] showed that if U and V are two noncommuting unitaries, then the operators

$$T_1 = \begin{pmatrix} 0 & 0 \\ I & 0 \end{pmatrix} \ , \ T_2 = \begin{pmatrix} 0 & 0 \\ U & 0 \end{pmatrix} \ , \ T_3 = \begin{pmatrix} 0 & 0 \\ V & 0 \end{pmatrix} \tag{3.3}$$

are three commuting contractions that have no isometric dilation. However, these operators can be shown to satisfy von Neumann's inequality.

What is it exactly that lies behind this dramatic difference between $d = 2$ and $d = 3$? Some people consider this to be an intriguing mystery, and there has been effort made into trying to understand which d-tuples are the ones that admit a unitary dilation (see, e.g., [160] and the references therein), or at least finding sufficient conditions for the existence of a nice dilation.

A particularly nice notion of dilation is that of *regular dilation*. For a d-tuple $T = (T_1, \ldots, T_d) \in B(\mathcal{H})^d$ and $n = (n_1, \ldots, n_d) \in \mathbb{N}^d$, we write $T^n = T_1^{n_1} \cdots T_d^{n_d}$. If $n = (n_1, \ldots, n_d) \in \mathbb{Z}^d$, then we define

$$T(n) = (T^{n_-})^* T^{n_+}$$

where $n_+ = (\max\{n_1, 0\}, \ldots, \max\{n_d, 0\})$ and $n_- = n_+ - n$. For a commuting unitary tuple U and $n \in \mathbb{Z}^d$, we have $U(n) = U^n := U_1^{n_1} \cdots U_d^{n_d}$. Now, if $\mathcal{K}$ contains $\mathcal{H}$ and $U = (U_1, \ldots, U_d) \in B(\mathcal{K})^d$ a d-tuple of unitaries, say that U is a **regular dilation** of T if

$$T(n) = P_{\mathcal{H}} U^n \big|_{\mathcal{H}} \quad \text{for all } n \in \mathbb{Z}^d. \tag{3.4}$$

Note that a unitary (power) dilation of a single contraction is automatically a regular dilation, because applying the adjoint to (1.4) gives $T(k) = P_{\mathcal{H}} U^k \big|_{\mathcal{H}}$ for all $k \in \mathbb{Z}$. However, a given unitary dilation of a pair of contractions need not satisfy (3.4), and in fact there are pairs of commuting contractions that have no regular unitary dilation.

In contrast with the situation of unitary dilations, the tuples of contractions that admit a regular unitary dilation can be completely characterized.

Theorem 3.7 (Regular Unitary Dilation) *A d-tuple $T = (T_1, \ldots, T_d)$ of commuting contractions on a Hilbert space has a regular unitary dilation if and only if,*

$$\sum_{\{i_1,\ldots,i_m\}\subseteq S} (-1)^m T_{i_1}^* \cdots T_{i_m}^* T_{i_1} \cdots T_{i_m} \geq 0 \quad , \quad \text{for all } S \subseteq \{1, \ldots, d\}. \tag{3.5}$$

The conditions (3.5) are sometimes called *Brehmer's conditions*. For the proof, one shows that the function $n \mapsto T(n)$ is a **positive definite function on the group** $\mathbb{Z}^d$ (see Section I.9 in [23]), and uses the fact that every positive definite function on a group has a unitary dilation [23, Section I.7].

Corollary 3.8 *The following are sufficient conditions for a d-tuple $T = (T_1, \ldots, T_d)$ of commuting contractions on a Hilbert space to have a regular unitary dilation:*

1. $\sum_{i=1}^d \|T_i\|^2 \leq 1$.
2. *$T_1, \ldots, T_d$ are all isometries.*
3. *$T_1, \ldots, T_d$ doubly commute, in the sense that $T_i T_j^* = T_j^* T_i$ for all $i \neq j$ (in addition to $T_i T_j = T_j T_i$ for all i, j).*

Proof It is not hard to show that the conditions listed in the corollary are sufficient for Brehmer's conditions (3.5) to hold. ■

3.2 Commutant Lifting

We return to the case of two commuting contractions. The following innocuous looking theorem, called the *commutant lifting theorem*, has deep applications (as we shall see in Sect. 4) and is the prototype for numerous generalizations. It originated in the work of Sarason [137], was refined by Sz.-Nagy and Foias (see [23]), and has become a really big deal (see [63]).

Theorem 3.9 (Commutant Lifting Theorem) *Let A be a contraction on a Hilbert space $\mathcal{H}$, and let $V \in B(\mathcal{K})$ be the minimal isometric dilation of A. For every contraction B that commutes with A, there exists an operator $R \in B(\mathcal{K})$ such that*

1. *R commutes with V,*
2. *$B = R^*\big|_{\mathcal{H}}$,*
3. *$\|R\| \leq 1$.*

In other words, every operator commuting with A can be "lifted" to an operator commuting with its minimal dilation, without increasing its norm.

Proof Let $U, W \in B(\mathcal{L})$ be the commuting isometric coextension of A, B, where $\mathcal{L}$ is a Hilbert space that contains $\mathcal{H}$ (the coextension exists by Andô's isometric

dilation theorem, Theorem 3.2). The restriction of the isometry U to the subspace $\bigvee_{n\in\mathbb{N}} U^n\mathcal{H}$ is clearly

1. an isometry,
2. a dilation of A,
3. a minimal dilation,

and therefore (by uniqueness of the minimal isometric dilation), the restriction of U to $\bigvee_{n\in\mathbb{N}} U^n\mathcal{H}$ is unitarily equivalent to *the* minimal isometric dilation V on $\mathcal{H}$, so we identify $\mathcal{K} = \bigvee_{n\in\mathbb{N}} U^n\mathcal{H}$ and $V = U\big|_{\mathcal{H}}$. It follows (either from our knowledge on the minimal dilation, or simply from the fact that U is a coextension) that V is a coextension of A. With respect to the decomposition $\mathcal{L} = \mathcal{K} \oplus \mathcal{K}^\perp$,

$$U = \begin{pmatrix} V & X \\ 0 & Z \end{pmatrix} \quad , \quad W = \begin{pmatrix} R & Q \\ P & N \end{pmatrix}$$

It is evident that $\|R\| \leq 1$ and that R is a coextension of B. We wish to show that $RV = VR$.

From $UW = WU$ we find that $VR + XP = RV$. Thus, the proof will be complete if we show that $X = 0$. Equivalently, we have to show that $\mathcal{K}$ is invariant under U^*. To see this, consider $U^*U^n h$ for some $h \in \mathcal{H}$ and $n \in \mathbb{N}$. If $n \geq 1$ we get $U^{n-1}h$ which is in $\mathcal{K}$. If $n = 0$ then we get $U^*h = A^*h \in \mathcal{H} \subseteq \mathcal{K}$, because U is a coextension of A. That completes the proof. ■

3.3 *Dilations of Semigroups and Semi-Invariant Subspaces*

Above we treated the case of a single operator or a tuple of commuting operators. However, dilation theory can also be developed, or at least examined, in the context of operator semigroups.

Let $T = \{T_s\}_{s\in\mathcal{S}} \subset B(\mathcal{H})$ be a family of operators parametrized by a semigroup $\mathcal{S}$ with unit e. Then T is a said to be a **semigroup of operators over** $\mathcal{S}$ if

1. $T_e = I$,
2. $T_{st} = T_s T_t$ for all $s, t \in \mathcal{S}$.

If $\mathcal{S}$ is a topological semigroup, then one usually requires the semigroup T to be continuous in some sense. A semigroup $V = \{V_s\}_{s\in\mathcal{S}} \subset B(\mathcal{K})$ is said to be a **dilation** of T if $\mathcal{K} \supset \mathcal{H}$ and if

$$T_s = P_{\mathcal{H}} V_s\big|_{\mathcal{H}} \quad , \quad \text{for all } s \in \mathcal{S}.$$

Note that Sz.-Nagy's unitary dilation theorem can be rephrased by saying that every semigroup of contractions over $\mathcal{S} = \mathbb{N}$ has a unitary dilation, in the above sense. Similarly, there are notions of **extension** and **coextension** of a semigroup of operators. Some positive results have been obtained for various semigroups. Sz.-

Nagy proved that every semigroup $T = \{T_t\}_{t\in\mathbb{R}_+}$ of contractions that is point-strong continuous (in the sense that $t \mapsto T_t h$ is continuous for all $h \in \mathcal{H}$) has isometric and unitary dilations, which are also point-strong continuous (see [23, Section I.10]). This result was extended to the two parameter case by Słociński [151] and Ptak [130]; the latter also obtained the existence of regular dilations for certain types of multi-parameter semigroups. Douglas proved that every commutative semigroup of isometries has a unitary extension [52]. Letting the commutative semigroup be $\mathcal{S} = \mathbb{N}^d$, we recover Theorem 3.1. Douglas's result was generalized by Laca to semigroups of isometries parametrized by an Ore semigroup [92], and in fact to "twisted" representations.

A result that somewhat sheds light on the question, which tuples of operators have a unitary dilation and which don't, is due to Opela. If $T = \{T_i\}_{i\in I} \subset B(\mathcal{H})$ is a family of operators, we say that T **commutes according to the graph** $\mathcal{G}$ with vertex set I, if $T_i T_j = T_j T_i$ whenever $\{i, j\}$ is an edge in the (undirected) graph $\mathcal{G}$. We can consider T as a semigroup parameterized by a certain quotient of the free semigroup over I. Opela proved the following compelling result: given a graph $\mathcal{G}$, every family $T = \{T_i\}_{i\in I}$ of contractions commuting according to $\mathcal{G}$ has a unitary dilation that commutes according to $\mathcal{G}$, if and only if $\mathcal{G}$ contains no cycles [108].

It is interesting that in the general setting of semigroups of operators, one can say something about the structure of dilations.

Theorem 3.10 (Sarason's Lemma [136]) *Let $V = \{V_s\}_{s\in\mathcal{S}} \subset B(\mathcal{K})$ be a semigroup of operators over a semigroup with unit $\mathcal{S}$, and let $\mathcal{H}$ be a subspace of $\mathcal{K}$. Then the family $T = \{T_s := P_{\mathcal{H}} V_s\big|_{\mathcal{H}}\}_{s\in\mathcal{S}}$ is a semigroup over $\mathcal{S}$ if and only if there exist two subspaces $\mathcal{M} \subseteq \mathcal{N} \subseteq \mathcal{K}$, invariant under V_s for all s, such that $\mathcal{H} = \mathcal{N} \ominus \mathcal{M} := \mathcal{N} \cap \mathcal{M}^\perp$.*

Proof The sufficiency of the condition is easy to see, if one writes the elements of the semigroup V as 3×3 block operator matrices with respect to the decomposition $\mathcal{K} = \mathcal{M} \oplus \mathcal{H} \oplus \mathcal{N}^\perp$.

For the converse, one has no choice but to define $\mathcal{N} = \vee_{s\in\mathcal{S}} V_s\mathcal{H}$ (clearly an invariant space containing $\mathcal{H}$), and then it remains to prove that $\mathcal{M} := \mathcal{N} \ominus \mathcal{H}$ is invariant for V, or—what is the same—that $P_{\mathcal{H}} V_t \mathcal{M} = 0$ for all $t \in \mathcal{S}$. Fixing t, we know that for all s,

$$P_{\mathcal{H}} V_t P_{\mathcal{H}} V_s P_{\mathcal{H}} = T_t T_s = T_{ts} = P_{\mathcal{H}} V_{ts} P_{\mathcal{H}} = P_{\mathcal{H}} V_t V_s P_{\mathcal{H}}.$$

It follows that $P_{\mathcal{H}} V_t P_{\mathcal{H}} = P_{\mathcal{H}} V_t$ on $\vee_s V_s\mathcal{H} = \mathcal{N}$. In particular, $P_{\mathcal{H}} V_t \mathcal{M} = P_{\mathcal{H}} V_t P_{\mathcal{H}} \mathcal{M} = 0$ (since $\mathcal{M} \perp \mathcal{H}$), as required. ∎

The theorem describes how a general dilation looks like. A subspace $\mathcal{H}$ as above, which is the difference of two invariant subspaces, is said to be **semi-invariant** for the family V. In the extreme case where $\mathcal{M} = \{0\}$, the space $\mathcal{H} = \mathcal{N}$ is just an invariant subspace for V, and V is an extension of T. In the other extreme case when $\mathcal{N} = \mathcal{K}$, the space $\mathcal{H}$ is a coinvariant subspace for V, and V is a coextension of T. In general, the situation is more complicated, but still enjoys some structure.

In the special $\mathcal{S} = \mathbb{N}$, case Sarason's lemma implies that V is a (power) dilation of an operator T if and only if it has the following block form:

$$V = \begin{pmatrix} * & * & * \\ 0 & T & * \\ 0 & 0 & * \end{pmatrix}. \tag{3.6}$$

Sarason's lemma is interesting and useful also in the case of dilations of a single contraction.

Remark 3.11 Up to this point in the survey, rather than attempting to present a general framework that encapsulates as much of the theory as possible, I chose to sew the different parts together with a thin thread. There are, of course, also "high level" approaches. In Sect. 7 we will see how the theory fits in the framework of operator algebras, which is one unifying viewpoint (see also [41, 118, 121]). There are other viewpoints. A notable one is due to Sz.-Nagy—very soon after he proved his unitary dilation theorem for a single contraction, he found a far-reaching generalization in terms of dilations of positive functions on $*$-*semigroups*; see [164], which contains a theorem from which a multitude of dilation theorems can be deduced (see also [162] for a more recent discussion with some perspective). Another brief but high level look on dilation theory can be found in Arveson's survey [13].

4 An Application: Pick Interpolation

The purpose of this section is to illustrate how classical dilation theory can be applied in a nontrivial way to prove theorems in complex function theory. The example we choose is classical—Pick's interpolation theorem—and originates in the work of Sarason [137]. Sarason's idea to use commutant lifting to solve the Pick interpolation problem works for a variety of other interpolation problems as well, including Carathéodory interpolation, matrix valued interpolation, and mixed problems. It can also be applied in different function spaces and multivariable settings. Here we will focus on the simplest case. Good references for operator theoretic methods and interpolation are [2] and [63], and the reader is referred to these sources for details and references.

4.1 The Problem

Recall that H^∞ denotes the algebra of bounded analytic functions on the unit disc $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$. For $f \in H^\infty$ we define

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

This norm turns H^∞ into a Banach algebra.

The *Pick interpolation problem* is the following: given n points $z_1, \ldots, z_n$ in the unit disc and n target points $w_1, \ldots, w_n \in \mathbb{C}$, determine whether or not there exists a function $f \in H^\infty$ such that

$$f(z_i) = w_i \quad , \quad \text{for all } i = 1, \ldots, n, \tag{4.1}$$

and

$$\|f\|_\infty \leq 1. \tag{4.2}$$

It is common knowledge that one can always find a polynomial (unique, if we take it to be of degree less than or equal to $n-1$) that *interpolates the data*, in the sense that (4.1) holds. The whole point is that we require (4.2) to hold as well. Clearly, this problem is closely related to the problem of finding the H^∞ function of minimal norm that interpolates the points.

Recall that an $n \times n$ matrix $A = (a_{ij})_{i,j=1}^n$ is said to be **positive semidefinite** if for every $v = (v_1, \ldots, v_n)^T \in \mathbb{C}^n$

$$\langle Av, v \rangle = \sum_{i,j=1}^n a_{ij} v_j \overline{v}_i \geq 0.$$

If A is positive semidefinite, then we write $A \geq 0$.

Theorem 4.1 (Pick's Interpolation Theorem) *Given points $z_1, \ldots, z_n$ and $w_1, \ldots, w_n$ as above, there exists a function $f \in H^\infty$ satisfying (4.1)-(4.2), if and only if the following matrix inequality holds:*

$$\left(\frac{1 - w_i \overline{w_j}}{1 - z_i \overline{z_j}} \right)_{i,j=1}^n \geq 0. \tag{4.3}$$

The $n \times n$ matrix on the left hand side of (4.3) is called the *Pick matrix*. What is remarkable about this theorem is that it gives an explicit and practical necessary and sufficient condition for the solvability of the interpolation problem: that the Pick matrix be positive semidefinite.

At this point it is not entirely clear how this problem is related to operator theory on Hilbert spaces, since there is currently no Hilbert space in sight. To relate this problem to operator theory we will represent H^∞ as an operator algebra. The space on which H^∞ acts is an interesting object in itself, and to this space we devote the next subsection. Some important properties of H^∞ functions as operators will be studied in Sect. 4.3, and then, in Sect. 4.4, we will prove Theorem 4.1.

4.2 The Hilbert Function Space H^2

The *Hardy space* $H^2 = H^2(\mathbb{D})$ is the space of analytic functions $h(z) = \sum_{n=0}^{\infty} a_n z^n$ on the unit disc $\mathbb{D}$ that satisfy $\sum |a_n|^2 < \infty$. It is not hard to see that H^2 is a linear space, and that

$$\left\langle \sum a_n z^n, \sum b_n z^n \right\rangle = \sum a_n \overline{b}_n$$

is an inner product which makes H^2 into a Hilbert space, with norm

$$\|h\|_{H^2}^2 = \sum |a_n|^2.$$

In fact, after noting that every $(a_n)_{n=0}^{\infty} \in \ell^2 := \ell^2(\mathbb{N}, \mathbb{C})$ gives rise to a power series that converges (at least) in $\mathbb{D}$, it is evident that the map

$$(a_n)_{n=0}^{\infty} \mapsto \sum_{n=0}^{\infty} a_n z^n$$

is a unitary isomorphism of ℓ^2 onto $H^2(\mathbb{D})$, so the Hardy space is a Hilbert space, for free. The utility of representing a Hilbert space in this way will speak for itself soon.

For $w \in \mathbb{D}$, consider the element $k_w \in H^2$ given by

$$k_w(z) = \sum_{n=0}^{\infty} \overline{w}^n z^n = \frac{1}{1 - z\overline{w}}.$$

Then for $h(z) = \sum a_n z^n$, we calculate

$$\langle h, k_w \rangle = \left\langle \sum a_n z^n, \sum \overline{w}^n z^n \right\rangle = \sum a_n w^n = h(w).$$

We learn that the linear functional $h \mapsto h(w)$ is a bounded functional, and that the element of H^2 that implements this functional is k_w. The functions k_w are called **kernel functions**, and the function $k : \mathbb{D} \times \mathbb{D} \to \mathbb{C}$ given by $k(z, w) = k_w(z)$ is called the **reproducing kernel** of H^2. The fact that point evaluation in H^2 is a bounded linear functional lies at the root of a deep connection between function theory on the one hand, and operator theory, on the other.

The property of H^2 observed in the last paragraph is so useful and important that it is worth a general definition. A Hilbert space $\mathcal{H} \subseteq \mathbb{C}^X$ consisting of functions on a set X, in which point evaluation $h \mapsto h(x)$ is bounded for all $x \in X$, is said to be a **Hilbert function space** or a **reproducing kernel Hilbert space**. See [119] for a general introduction to this subject, and [2] for an introduction geared towards

Pick interpolation (for readers that are in a hurry, Chapter 6 in [147] contains an elementary introduction to H^2 as a Hilbert function space). If $\mathcal{H}$ is a Hilbert function space on X, then by the Riesz representation theorem, for every $x \in X$ there is an element $k_x \in \mathcal{H}$ such that $h(x) = \langle h, k_x \rangle$ for all $h \in \mathcal{H}$, and one may define the **reproducing kernel** of $\mathcal{H}$ by $k(x, y) = k_y(x)$.

The **multiplier algebra** of a Hilbert function space $\mathcal{H}$ on a set X is defined to be

$$\mathrm{Mult}(\mathcal{H}) = \{f : X \to \mathbb{C} : fh \in \mathcal{H} \ \text{ for all } h \in \mathcal{H}\}.$$

Every $f \in \mathrm{Mult}(\mathcal{H})$ gives rise to a linear **multiplication operator** $M_f : \mathcal{H} \to \mathcal{H}$ that acts as $M_f h = fh$, for all $h \in \mathcal{H}$. By the closed graph theorem, multiplication operators are bounded. The following characterization of multiplication operators is key to some applications.

Proposition 4.2 *Let $\mathcal{H}$ be a Hilbert function space on a set X. If $f \in \mathrm{Mult}(\mathcal{H})$, then $M_f^* k_x = \overline{f(x)} k_x$ for all $x \in X$. Conversely, if $T \in B(\mathcal{H})$ is such that for all $x \in X$ there is some $\lambda_x \in \mathbb{C}$ such that $Tk_x = \lambda_x k_x$, then there exists $f \in \mathrm{Mult}(\mathcal{H})$ such that $T = M_f^*$.*

Proof For all $h \in H$ and $x \in X$,

$$\langle h, M_f^* k_x \rangle = \langle fh, k_x \rangle = f(x)h(x) = f(x)\langle h, k_x \rangle = \langle h, \overline{f(x)} k_x \rangle,$$

so $M_f^* k_x = \overline{f(x)} k_x$. The converse is similar. ∎

Corollary 4.3 *Every $f \in \mathrm{Mult}(\mathcal{H})$ is a bounded function, and*

$$\sup_{x \in X} |f(x)| \leq \|M_f\|.$$

Proposition 4.4 $\mathrm{Mult}(H^2) = H^\infty$ *and $\|M_f\| = \|f\|_\infty$ for every multiplier.*

Proof Since $1 \in H^2$, every multiplier $f = M_f 1$ is in H^2. In particular, every multiplier is an analytic function. By the above corollary, $\mathrm{Mult}(H^2) \subseteq H^\infty$, and $\|f\|_\infty \leq \|M_f\|$ for every multiplier f.

Conversely, if $p(z) = \sum_{n=0}^N a_n z^n$ is a polynomial, then it is straightforward to check that

$$\|p\|_{H^2}^2 = \frac{1}{2\pi} \int_0^{2\pi} |p(e^{it})|^2 dt.$$

An approximation argument then gives

$$\|h\|_{H^2}^2 = \lim_{r \nearrow 1} \frac{1}{2\pi} \int_0^{2\pi} |h(re^{it})|^2 dt$$

for all $h \in H^2$. This formula for the norm in H^2 implies that $H^\infty \subseteq \mathrm{Mult}(H^2)$, and that $\|M_f\| \leq \|f\|_\infty$. ∎

We will henceforth identify f with M_f, and we will think of H^∞ as a subalgebra of $B(H^2)$.

4.3 The Shift M_z

We learned that every bounded analytic function $f \in H^\infty$ defines a bounded multiplication operator $M_f : H^2 \to H^2$, but there is one that stands out as the most important. If we abuse notation a bit and denote the identity function $\mathbf{id} : \mathbb{D} \to \mathbb{D}$ simply as z, then we obtain the multiplier M_z, defined by

$$(M_z h)(z) = zh(z).$$

It is quite clear that M_z is an isometry, and in fact it is unitarily equivalent to the unilateral shift of multiplicity one on ℓ^2 (defined before Theorem 2.2). We will collect a couple of important results regarding this operator, before getting back to the proof of Pick's theorem.

Recall that the **commutant** of a set of operators $\mathcal{S} \subset B(\mathcal{H})$ is the algebra

$$\mathcal{S}' = \{T \in B(\mathcal{H}) : ST = TS \ \text{ for all } S \in \mathcal{S}\}.$$

Proposition 4.5 $\{M_z\}' = (H^\infty)' = H^\infty$.

Proof Clearly $H^\infty \subseteq (H^\infty)' \subseteq \{M_z\}'$. Now suppose that $T \in \{M_z\}'$. We claim that $T = M_f$ for $f = T1$. Indeed, if $p(z) = \sum_{n=0}^N a_n z^n$ is a polynomial, then

$$Tp = T \sum_{n=0}^N a_n M_z^n 1 = \sum_{n=0}^N a_n M_z^n T1 = M_p f = fp.$$

An easy approximation argument would show that $T = M_f$, if we knew that $f \in H^\infty$; but we still don't. To finesse this subtlety, we find for an arbitrary $h \in H^2$ a sequence of polynomials p_n converging in norm h, and evaluate at all points $w \in \mathbb{D}$, to obtain:

$$f(w)p_n(w) = (Tp_n)(w) \xrightarrow{n\to\infty} (Th)(w),$$

while $f(w)p_n(w) \to f(w)h(w)$, on the other hand. This means that $Th = fh$ for all h and therefore $f \in \mathrm{Mult}(H^2) = H^\infty$, as required. ∎

Let $z_1, \ldots, z_n \in \mathbb{D}$. It is not hard to see that $k_{z_1}, \ldots, k_{z_n}$ are linearly independent. Let $\mathcal{G} = \mathrm{span}\{k_{z_1}, \ldots, k_{z_n}\}$, and let $A = P_{\mathcal{G}} M_z\big|_{\mathcal{G}}$. By Proposition 4.2, $\mathcal{G}$ is coinvariant for M_z, i.e., $M_z^* \mathcal{G} \subseteq \mathcal{G}$, and $A^* = M_z^*\big|_{\mathcal{G}}$ is the diagonal operator given by

$$A^* : k_{z_i} \mapsto \overline{z_i} k_{z_i}. \tag{4.4}$$

We claim that M_z is the minimal isometric dilation of A. Well, it's clearly an isometric dilation, we just need to show that it is minimal. But $k_{z_i}(z) = \frac{1}{1-z\overline{z_i}}$, so $k_{z_i} - \overline{z_i} M_z k_{z_i} = 1 \in \bigvee_{n\in\mathbb{N}} M_z^n \mathcal{G}$. It follows that all the polynomials are in $\bigvee_{n\in\mathbb{N}} M_z^n \mathcal{G}$, whence $H^2 = \bigvee_{n\in\mathbb{N}} M_z^n \mathcal{G}$.

More generally, if we have a multiplier f, and we define $B = P_{\mathcal{G}} M_f\big|_{\mathcal{G}}$, then we have that $B^* = M_f^*\big|_{\mathcal{G}}$ and that

$$B^* : k_{z_i} \mapsto \overline{f(z_i)} k_{z_i}. \tag{4.5}$$

4.4 Proof of Pick's Interpolation Theorem

We can now prove Theorem 4.1. We first show that (4.3) is a necessary condition. Suppose that $f \in H^\infty$ satisfies $f(z_i) = w_i$ for all $i = 1, \ldots, n$ and that $\|f\|_\infty \leq 1$. Define $B = P_{\mathcal{G}} M_f\big|_{\mathcal{G}}$, where $\mathcal{G} = \operatorname{span}\{k_{z_1}, \ldots, k_{z_n}\}$ as in the previous subsection. Then, by (4.5) B^* is the diagonal operator given by $B^* k_{z_i} = \overline{w_i} k_{z_i}$. Since $\|M_f\| \leq 1$, also $\|B^*\| \leq 1$, thus for all $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$,

$$\begin{aligned} 0 \leq \left\| \sum_{i=1}^n \alpha_i k_{z_i} \right\|^2 - \left\| B^* \sum_{i=1}^n \alpha_i k_{z_i} \right\|^2 &= \left\| \sum_{i=1}^n \alpha_i k_{z_i} \right\|^2 - \left\| \sum_{i=1}^n \overline{w_i} \alpha_i k_{z_i} \right\|^2 \\ &= \sum_{i,j=1}^n \alpha_j \left(1 - \overline{w_j} w_i\right) \overline{\alpha_i} \left\langle k_{z_j}, k_{z_i} \right\rangle \\ &= \sum_{i,j=1}^n \alpha_j \left(\frac{1 - w_i \overline{w_j}}{1 - z_i \overline{z_j}} \right) \overline{\alpha_i}. \end{aligned}$$

That is, the Pick matrix is positive semidefinite, and (4.3) holds.

Conversely, suppose that (4.3) holds. Define a diagonal operator $D : \mathcal{G} \to \mathcal{G}$ by $D k_{z_i} = \overline{w_i} k_{z_i}$ for $i = 1, \ldots, n$, and let $B = D^*$. Then the above computation can be rearranged to show that $\|B\| = \|B^*\| \leq 1$. Now, the diagonal operator B^* clearly commutes with the diagonal operator $A^* = M_z^*\big|\mathcal{G}$, so B commutes with A. Since M_z is the minimal isometric dilation of A, the commutant lifting theorem (Theorem 3.9) implies that B has a coextension to an operator T that commutes with M_z and has $\|T\| \leq 1$. By Proposition 4.5, $T = M_f$ for some $f \in H^\infty$, and by Proposition 4.2, $f(z_i) = w_i$ for all $i = 1, \ldots, n$. Since $\|f\|_\infty = \|T\| \leq 1$, the proof is complete.

5 Spectral Sets, Complete Spectral Sets, and Normal Dilations

Classical dilation theory does not end with dilating commuting contractions to commuting unitaries. Let us say that a d-tuple $N = (N_1, \ldots, N_d)$ is a **normal** tuple if $N_1, \ldots, N_d$ are all normal operators and, in addition, they all commute with one another. Recall that the **joint spectrum** $\sigma(N)$ of a normal tuple is the set

$$\sigma(N) = \{(\rho(N_1), \ldots, \rho(N_d)) : \rho \in \mathcal{M}(C^*(N))\} \subset \mathbb{C}^d,$$

where $\mathcal{M}(C^*(N))$ is the space of all nonzero complex homomorphisms from the unital C*-algebra $C^*(N)$ generated by N to $\mathbb{C}$. If N acts on a finite dimensional space, then the joint spectrum is the set of joint eigenvalues, belonging to an orthogonal set of joint eigenvectors that simultaneously diagonalize $N_1, \ldots, N_d$. A commuting tuple of unitaries $U = (U_1, \ldots, U_d)$ is the same thing as a normal tuple with joint spectrum contained in the torus $\mathbb{T}^d$. Since normal tuples are in a sense "completely understood", it is natural to ask which operator tuples T have a normal dilation N (where the definition of dilation is as in (3.1)) with the spectrum $\sigma(N)$ prescribed to be contained in some set $X \subset \mathbb{C}^d$.

Suppose that $T = (T_1, \ldots, T_d)$ is a commuting tuple of operators and that $N = (N_1, \ldots, N_d)$ is a normal dilation with $\sigma(N) = X \subset \mathbb{C}^d$. Then we immediately find that

$$\|p(T)\| \leq \|p(N)\| = \sup_{z \in X} |p(z)|$$

for every polynomial p in d variables. In fact, it is not too hard to see that the above inequality persists when p is taken to be a rational function that is regular on X. This motivates the following definition: a subset $X \subseteq \mathbb{C}^d$ is said to be a **K-spectral set** for T if X contains the joint spectrum $\sigma(T)$ of T, and if for every rational function f that is regular on X,

$$\|f(T)\| \leq K \|f\|_{X,\infty}, \tag{5.1}$$

where $\|f\|_{X,\infty} = \sup_{z \in X} |f(z)|$. If X is a K-spectral set for T with $K = 1$, then it is simply said to be a **spectral set** for T.

I do not wish to define the joint spectrum of a non-normal commuting tuple, nor to go into how to evaluate a rational function on a tuple of operators, so I will be somewhat sloppy in what follows (see [11, Section 1.1]; for a textbook treatment, I recommend also [118]). Two simplifying comments are in order:

1. In the case $d = 1$, i.e., just one operator T, the spectrum $\sigma(T)$ is the usual spectrum, and the evaluation $f(T)$ of a rational function on T can be done naturally, and this is the same as using the holomorphic functional calculus.

2. One may also discuss **polynomial spectral sets**, in which (5.1) is required to hold only for polynomials [38]. If X is polynomially convex (and in particular, if X is convex), then considering polynomials instead of rational functions leads to the same notion.

Thus, with the terminology introduced above, we can rephrase Theorem 3.4 by saying that the bidisc $\overline{\mathbb{D}}^2$ is a spectral set for every pair $T = (T_1, T_2)$ of commuting contractions, and Example 3.5 shows that there exists three commuting contractions for which the tridisc $\overline{\mathbb{D}}^3$ is not a spectral set.

The notion of a spectral set for a single operator is due to von Neumann [168]. A nice presentation of von Neumann's theory can be found in Sections 153–155 of [133]. The reader is referred to [19] for a rather recent survey with a certain emphasis on the single variable case. To give just a specimen of the kind of result that one can encounter, which is quite of a different nature than what I am covering in this survey, let me mention the result of Crouzeix [40], which says that for every $T \in B(\mathcal{H})$, the numerical range $W(T) := \{\langle Th, h\rangle : \|h\| = 1\}$ of T is a K-spectral set for some $K \geq 2$ (it is easy to see that one cannot have a constant smaller than 2; Crouzeix conjectured that $K = 2$, and this conjecture is still open at the time of me writing this survey).

It is plain to see that if T has a commuting normal dilation N with spectrum $\sigma(N) \subseteq X$, then X is a polynomial spectral set for T, and it is true that in fact X is a spectral set. It is natural to ask whether the converse implication holds, that is, whether the assumption that a set X is a spectral set for a tuple T implies that there exists a normal dilation with spectrum constrained to X (or even to the *Shilov boundary* ∂X). There are cases when this is true (see [19]), but in general the answer is *no*. For example, we already mentioned that Parrott's example [110] of three commuting contractions that have no unitary dilation (hence also no normal dilation with spectrum contained in $\overline{\mathbb{D}}^3$) does not involve a violation of von Neumann's inequality, in other words the tuple T from (3.3) has $\overline{\mathbb{D}}^3$ as a spectral set but has no unitary dilation.

The situation was clarified by Arveson's work [11], where the notion of *complete spectral set* was introduced. To explain this notion, we need matrix valued polynomials and rational functions. Matrix valued polynomials in several commuting (or noncommuting) variables, and the prescriptions for evaluating them at d-tuples of commuting (or noncommuting) operators, are defined in a similar manner to their definition in the one variable case in Remark 1.3. Once one knows how to evaluate a rational function in several variables at a commuting tuple, the passage to matrix valued rational functions is done similarly.

Given a tuple $T \in B(\mathcal{H})^d$ of commuting contractions, we say that a set $X \subset \mathbb{C}^d$ is a **complete K-spectral set** for T, if $\sigma(T) \subseteq X$ and if for every matrix valued rational function f that is regular on X, (5.1) holds, where now for an $n \times n$ matrix valued rational function $\|f\|_{X,\infty} = \sup_{z \in X} \|f(z)\|_{M_n}$. If X is a complete K-spectral set for T with $K = 1$, then it is simply said to be a **complete spectral set** for T.

Theorem 5.1 (Arveson's Dilation Theorem [11]) *Let $T = (T_1, \ldots, T_d)$ be a tuple of commuting operators on a Hilbert space $\mathcal{H}$. Let $X \subset \mathbb{C}^d$ be a compact set and let ∂X be the Shilov boundary of X with respect to the algebra $rat(X) \subseteq C(X)$ of rational functions that are regular on X. Then X is a complete spectral set for T if and only if only if there exists a normal tuple $N = (N_1, \ldots, N_d)$ acting on a Hilbert space $\mathcal{K} \supseteq \mathcal{H}$, such that $\sigma(N) \subseteq \partial X$ and for every matrix valued rational function f that is regular on X,*

$$f(T) = P_{\mathcal{H}} f(N)\big|_{\mathcal{H}}.$$

Putting Arveson's dilation theorem together with some comments made above, we see that $\overline{\mathbb{D}}^3$ is a spectral set for the triple T from (3.3), but it is not a complete spectral set. On the other hand, we know that for every pair of commuting contractions $T = (T_1, T_2)$, the bidisc $\overline{\mathbb{D}}^2$ is a complete spectral set. Agler and McCarthy proved a sharper result: if $T = (T_1, T_2)$ acts on a finite dimensional space, and $\|T_1, \|, \|T_2\| < 1$, then there exists a one dimensional complex algebraic subvariety $V \subseteq \mathbb{D}^2$ (in fact, a so-called *distinguished variety*, which means that $\overline{V} \cap \partial(\mathbb{D}^2) = \overline{V} \cap \mathbb{T}^2$), such that V is a complete spectral set for T [3].

If $X \subset \mathbb{C}$ is a spectral set for an operator T, one may ask whether or not it is a complete spectral set. We close this section by mentioning some notable results in this direction. It is known that if $X \subset \mathbb{C}$ is a compact spectral set for T such that $rat(X) + \overline{rat(X)}$ is dense in $C(\partial X)$, then X is a complete spectral set, and T has a normal dilation with spectrum in ∂X. The condition is satisfied, for example, when X is the closure of a bounded and simply connected open set (this result is due to Berger, Foias and Lebow (independently); see [118, Theorem 4.4]). The same is true if X is an annulus (Agler [1]), but false if X is triply connected (Agler et al. [5] and Dritschel and McCullough [54]).

If a pair of commuting operators $T = (T_1, T_2)$ has the **symmetrized bidisc** $\Gamma := \{(z_1 + z_2, z_1 z_2) : z_1, z_2 \in \overline{\mathbb{D}}\}$ as a spectral set, then in fact Γ is a complete spectral set for T (Agler and Young [4]). Pairs of operators having Γ as a spectral set have a well developed model theory (see, e.g., Sarkar [139] and the references therein). Building on earlier work of Bhattacharyya et al. [29], and inspired by Agler and McCarthy's distinguished varieties result mentioned above, Pal and Shalit showed that if Γ is a spectral set for a pair $T = (T_1, T_2)$ of commuting operators acting on a finite dimensional space, then there exists a *distinguished* one dimensional algebraic variety $V \subseteq \Gamma$ which is a complete spectral set for T [109].

Part 2: A Rapid Overview of Dilation Theories

6 Additional Results and Generalizations of Dilation Theory

6.1 Some Further Remarks on N-Dilations

The notion of a 1-dilation of a single operator, which is usually referred to simply as *dilation*, has appeared through the years and found applications in operator theory; see e.g. [22, 36, 72] (the reader should be warned that the terminology is not universally accepted; for example, as we already mentioned, a power dilation is usually simply referred to as *dilation*. Even more confusingly, in [22], a *unitary N-dilation of T* means what we call here a unitary 1-dilation of T that acts on $\mathcal{H} \oplus \mathbb{C}^N$).

Egerváry's simple construction (1.2) of an N-dilation, and with it the concept of N-dilations, have been largely forgotten until [95] seemed to revive some interest in it (see also [106]). The motivation was that the well-known Sz.-Nagy unitary (power) dilation of a contraction T (given by Theorem 2.1) always acts on an infinite dimensional space whenever T is nonunitary, even if T acts on a finite dimensional space. Arguably, one cannot say that an infinite dimensional object is better understood than a matrix. That's what led to the rediscovery of (1.2) and thence to the dilation-theoretic proof of von Neumann's inequality that we presented, which has the conceptual advantage of never leaving the realm of finite dimensional spaces, in the case where T acts on a finite dimensional space to begin with.

Let $T = (T_1, \ldots, T_d)$ be a d-tuple of commuting operators acting on a Hilbert space $\mathcal{H}$, and let $U = (U_1, \ldots, U_d)$ be a d-tuple of commuting operators acting on a Hilbert space $\mathcal{K} \supseteq \mathcal{H}$. We say that U is a an **N-dilation** of T if

$$p(T) = P_{\mathcal{H}} p(U)\big|_{\mathcal{H}}$$

for every polynomial in d complex variables of degree less than or equal to N. We say that this dilation is a **unitary/normal dilation** if every U_i $(i = 1, \ldots, d)$ is unitary/normal. The construction (1.2) shows that every contraction has a unitary N-dilation acting on $\mathcal{H}^{N+1}$. In particular, it shows that every contraction acting on a finite dimensional space has a unitary N-dilation acting on a finite dimensional space, for all N.

Curiously, it appears that the proof of Theorem 3.2 cannot be modified to show that every pair of commuting contractions on a finite dimensional space has a commuting unitary N-dilation on a finite dimensional space, for all N. It was shown by McCarthy and Shalit that indeed such a finitary version of Andô's dilation theorem holds [99]. Interestingly, the proof made use of Andô's dilation theorem. So, if one uses this finitary dilation theorem to prove von Neumann's inequality for pairs of matrices, one does not truly avoid infinite dimensional spaces. It is an

open problem to come up with an explicit construction of a unitary N-dilation for commuting matrices.

In fact, in [99] it was also proved that a d-tuple of contractions acting on a finite dimensional space has a unitary dilation if and only if for all N it has a unitary N-dilation acting on a finite dimensional space. Likewise, it was shown that for such a tuple, the existence of a regular unitary dilation is equivalent to the existence of a *regular unitary N-dilation* (you can guess what that means) acting on a finite dimensional space, for all N. Additional finitary dilation results appeared, first in the setting of normal dilations of commuting tuples [38], and then in the setting of 1-dilations of noncommuting operators [46, Section 7.1]. A similar phenomenon was also observed in [69]. At last, Hartz and Lupini found a finite dimensional version of Stinespring's dilation theorem (see Sect. 7.1), which provides a general principle by which one can deduce finite dimensional dilation theorems from their infinite dimensional counterparts [73].

It is interesting to note that N-dilations found an application in simulating open quantum systems on a quantum computer [79], and they also appeared in the context of quantum information theory [94]. The notion of N-dilations also appeared in the dilation theory in general Banach spaces (about which will say a few words below), see [62].

6.2 Models

Another direction in which dilation theory for commuting d-tuples has been developed is that of *operator models*. Roughly, the idea is that certain classes of d-tuples of operators can be exhibited as the compressions of a particular "model" d-tuple of operators. We will demonstrate this with a representative example; for a broader point of view see [105], Chapter 14 in [2], or the surveys [8] and [138].

Our example is the *d-shift* on the *Drury-Arveson space* H^2_d [12, 55] (see also the survey [146]). For a fixed d, we let H^2_d denote the space of all analytic functions $f(z) = \sum_\alpha c_\alpha z^\alpha$ on the unit ball $\mathbb{B}_d$ such that (with standard multi-index notation)

$$\|f\|^2_{H^2_d} := \sum_\alpha |c_\alpha|^2 \frac{\alpha!}{|\alpha|!} < \infty.$$

This norm turns the space H^2_d into a Hilbert space of analytic functions on $\mathbb{B}_d$, such that point evaluation is bounded. In fact, H^2_d is the reproducing kernel Hilbert space determined by the kernel $k(z, w) = \frac{1}{1-\langle z,w\rangle}$. Some readers might jump to their feet and object that this space is nothing but the good old *symmetric Fock space*, but it is fruitful and enlightening to consider it as a space of analytic functions (so please, sit down).

For the record, let the reader know that the possibility $d = \infty$ is allowed, but we do not dwell upon this point.

On H^2_d there is a d-tuple of operators $S = (S_1, \ldots, S_d)$, called the *d-shift*, and defined by

$$S_i f(z) = z_i f(z) \quad , \quad i = 1, \ldots, d,$$

where $z = (z_1, \ldots, z_d)$ is the complex variable, and so S_i is multiplication by the ith coordinate function z_i. The tuple S is plainly a commuting tuple: $S_i S_j = S_j S_i$ (multiplication of functions is commutative). A short combinatorial exercise shows that $\sum S_i S_i^*$ is equal to the orthogonal projection onto the orthogonal complement of the constant functions, and in particular $\sum S_i S_i^* \leq I$. Thus S is a **row contraction**, meaning that the row operator $[S_1 \ S_2 \ \cdots \ S_d] : H^2_d \oplus \cdots \oplus H^2_d \to H^2_d$ is a contraction. Another calculation reveals that S is **pure**, in the sense that $\sum_{|\alpha|=n} S^\alpha (S^\alpha)^* \xrightarrow{n\to\infty} 0$ in the strong operator topology.

The remarkable fact is that H^2_d is a *universal model* for pure commuting row contractions. I will now explain what these words mean. If $\mathcal{G}$ is a Hilbert space, we can consider the space $H^2_d \otimes \mathcal{G}$ (which can be considered as a Hilbert space of analytic $\mathcal{G}$-valued functions), and the d-shift promotes to a shift $S \otimes I_\mathcal{G}$ on $H^2_d \otimes \mathcal{G}$, which is called a **multiple of the d-shift**. A subspace $\mathcal{M} \subseteq H^2_d \otimes \mathcal{G}$ is said to be **coinvariant** if it is invariant for $S_i^* \otimes I_\mathcal{G}$ for all $i = 1, \ldots, d$.

Theorem 6.1 (Universality of the d-Shift) *Let $T = (T_1, \ldots, T_d) \in B(\mathcal{H})^d$ be a pure, commuting row contraction. Then there exists a Hilbert space $\mathcal{G}$ and a coinvariant subspace $\mathcal{M} \subseteq H^2_d \otimes \mathcal{G}$ such that T is unitarily equivalent to the compression of $S \otimes I_\mathcal{G}$ to $\mathcal{M}$.*

Thus, every row contraction T is unitarily equivalent to the corestriction of a multiple of the d-shift to a coinvariant subspace. In particular, for every polynomial p in d variables,

$$\|p(T)\| = \left\| P_\mathcal{M} \left(p(S) \otimes I_\mathcal{G} \right) \big|_\mathcal{M} \right\| \leq \|p(S)\|, \tag{6.1}$$

and this inequality replaces von Neumann's inequality in this setting (and this was Drury's motivation [55]). It can be shown [55] (see also [12, 44]) that there exists no constant C such that $\|P(S)\| \leq C \sup_{z \in \overline{\mathbb{B}}_d} |p(z)|$, and in particular, commuting row contractions in general do not have normal dilations with spectrum contained in $\overline{\mathbb{B}}_d$.

6.3 Dilation Theory for Noncommutative Operator Tuples

Dilation theory also plays a role in the analysis of tuples of noncommuting operators. Recall that a **row contraction** is a tuple $T = (T_1, \ldots, T_d)$ such that $\sum T_i T_i^* \leq I$ (as in Sect. 6.2, we allow, but do not belabor, the case $d = \infty$, in which case the sum is understood in the strong-operator topology sense). A **row isometry**

is a tuple $V = (V_1, \ldots, V_d)$ such that $V_i^* V_j = \delta_{ij} I$, for all i, j. Thus, the operators $V_1, \ldots, V_d$ are all isometries which have mutually orthogonal ranges, and this is equivalent to the condition that the *row operator* $[V_1\ V_2\ \cdots\ V_d]$ is an isometry. The Sz.-Nagy isometric dilation theorem extends to the setting of (noncommuting) row contractions. The following theorem is due to Frazho [64] (the case $d = 2$), Bunce [33] (the case $d \in \mathbb{N} \cup \{\infty\}$) and Popescu [123] (who proved the existence of dilation in the case $d \in \mathbb{N} \cup \{\infty\}$, and later developed a far reaching generalization of Sz-Nagy's and Foias's theory for noncommuting tuples and more).

Theorem 6.2 (Row Isometric Dilation of Row Contractions) *Let $T \in B(\mathcal{H})^d$ be a row contraction. Then there exists a Hilbert space $\mathcal{K}$ containing $\mathcal{H}$ and a row isometry $V = (V_1, \ldots, V_d) \in B(\mathcal{K})^d$ such that $V_i^* \big| \mathcal{H} = T_i^*$ for all i.*

There is also a very closely related dilation result, that shows that the shift $L = (L_1, \ldots, L_d)$ on the full Fock space is a universal model for *pure* row contractions, which reads similarly to Theorem 6.1, with the free shift L replacing the commutative shift S. Correspondingly, there is a von Neumann type inequality $\|p(T)\| \leq \|p(L)\|$ which holds for every row contraction T and every polynomial p in noncommuting variables [123, 124].

Popescu has a large body of work in which this dilation/model theory is developed, applied, and generalized. In particular, the theory can be modified to accommodate tuples satisfying certain polynomial relations [127] (see also [150, Section 8]) or tuples in certain *noncommutative polydomains* [128].

The isometric dilation of a row contractions lies at the heart of the free functional calculus for row contractions (see, e.g., [126]), and is important for understanding the algebraic structure of *noncommutative Hardy algebras* (also called *analytic Toeplitz algebras*, see [45]), as well as for the study and classification of algebras of bounded *nc analytic functions* on the *nc unit ball* and its subvarieties [134, 135].

6.4 Dilations in Banach Spaces

Until now, we have only considered operators on Hilbert spaces. But there are other kinds of interesting spaces, and the concept of dilations has appeared and been used in various settings. In the setting of Banach spaces, one may hope to dilate a contraction to an invertible isometry (that is, a surjective isometry); more generally one may wish to dilate a semigroup of operators to a group representation. Results along these lines, including a direct analogoue of Sz.-Nagy's unitary dilation theorem, were obtained by Stroescu; see [161].

However, Banach spaces form a huge class of spaces, and the dilation theory in the context of general Banach spaces contains the additional aspect that one might like to ensure that the dilating space shares some properties with the original space. For example, if T is a contraction on an L^p-space, one might wish to dilate to an invertible isometry acting on an L^p-space. Moreover, if T is **positive**, in the sense that $Tf \geq 0$ (almost everywhere) whenever $f \geq 0$ (almost everywhere), then one

might hope to dilate to a positive invertible isometry. The following theorem is an example of the kind of result one can look for.

Theorem 6.3 (Akcoglu–Sucheston [6]) *Let $T : X \to X$ be a positive contraction on an L^p-space $X = L^p(\mu)$ ($1 \leq p < \infty$). Then there exists another L^p-space $Y = L^p(\nu)$, a positive invertible isometry $U : Y \to Y$, a positive isometry $J : X \to Y$, and a positive projection $Q : Y \to Y$ such that*

$$JT^n = QU^n J \quad , \quad \text{for all } n \in \mathbb{N}.$$

Note that even in the case $p = 2$, this is not exactly Sz.-Nagy's dilation theorem: the assumptions are stronger, but so is the conclusion. For a modern approach to dilations in Banach spaces, generalizations, and also an overview of the history of the theory and its applications, see [62]. Operator algebras are another class of spaces in which dilation theory was developed and applied; we will discuss this setting in Sects. 7 and 8 below.

6.5 Dilations of Representations of C*-Correspondences

A **Hilbert C*-module** is a complex linear space E which is a right module over a C*-algebra $\mathcal{A}$, which carries an "$\mathcal{A}$-valued inner product" $\langle \cdot, \cdot \rangle : E \times E \to \mathcal{A}$, that satisfies the following conditions:

1. $\langle x, x \rangle \geq 0$ for all $x \in E$,
2. $\langle x, ya \rangle = \langle x, y \rangle a$ for all $x, y \in E$ and $a \in \mathcal{A}$,
3. $\langle x, y \rangle = \langle y, x \rangle^*$ for all $x, y \in E$,
4. $\langle x, \alpha y + \beta z \rangle = \alpha \langle x, y \rangle + \beta \langle x, z \rangle$ for all $x, y, z \in E$ and $\alpha, \beta \in \mathbb{C}$,
5. $\|x\| := \|\langle x, x \rangle\|^{1/2}$ is a norm on E which makes E into a Banach space.

The notion was introduced by Kaplansky [83] for the case where the C*-algebra $\mathcal{A}$ is commutative, and then developed further by Paschke [112] and Rieffel [131] for general C*-algebras. It is now a standard tool in some fields in operator algebras; see [93] or Part I of [154] for an introduction.

Hilbert modules evolved into a more refined notion, called *Hilbert correspondences*, that involves a left action. A linear operator $T : E \to E$ is said to be **adjointable** if there exists a linear operator $S : E \to E$ so that $\langle Tx, y \rangle = \langle x, Sy \rangle$ for all $x, y \in E$. One can show that every adjointable operator is a bounded right module map, but the converse is not true. The set of all adjointable operators on a Hilbert C*-correspondence E is denoted $\mathscr{B}^a(E)$ or $\mathcal{L}(E)$; it is a C*-algebra. A Hilbert **C*-correspondence from $\mathcal{A}$ to $\mathcal{B}$** is a Hilbert $\mathcal{B}$-module E which also carries a left action of $\mathcal{A}$ by adjointable operators. If $\mathcal{A} = \mathcal{B}$ then we say **C*-correspondence over $\mathcal{A}$**.

Given a Hilbert C*-correspondence E over the C*-algebra $\mathcal{A}$, a **covariant representation** of E on a Hilbert space $\mathcal{H}$ is a pair (T, σ) where T is linear map

$T : E \to B(\mathcal{H})$ and $\sigma : \mathcal{A} \to B(\mathcal{H})$ is a nondegenerate $*$-representation such that $T(a \cdot x \cdot b) = \sigma(a)T(x)\sigma(b)$ for all $a, b \in \mathcal{A}$ and $x \in E$. A covariant representation is said to be **contractive/completely contractive/bounded**, etc., if T is contractive/completely contractive/bounded, etc; it is said to be **isometric** if $T(x)^*T(y) = \sigma(\langle x, y\rangle)$ for all $x, y \in E$.

Muhly and Solel proved that every completely contractive covariant representation (T, σ) of E on $\mathcal{H}$ has an **isometric dilation** (V, π) of E on $\mathcal{K} \supseteq \mathcal{H}$, [100, Theorem 3.3]. By this, we mean an isometric covariant representation (V, π) on a Hilbert space $\mathcal{K}$ that contains $\mathcal{H}$, such that

1. $\mathcal{H}$ is reducing for π, and $P_{\mathcal{H}}\pi(a)\big|_{\mathcal{H}} = \sigma(a)$ for all $a \in \mathcal{A}$,
2. $P_{\mathcal{H}}V(x)\big|_{\mathcal{H}} = T(x)$ for all $x \in E$,
3. $P_{\mathcal{H}}V(x)\big|_{\mathcal{H}^\perp} = 0$ for all $x \in E$.

Moreover, they proved that such an isometric dilation can be chosen to be minimal in a certain sense, and that the minimal isometric dilation is unique up to unitary equivalence (the third condition that an isometric dilation is required to satisfy looks more like something that should be called a *coextension*, it is actually a consequence of minimality; sometimes it is not required). The isometric dilation theorem was used in [100] to analyze the representation theory of the **tensor algebra** $\mathcal{T}_+(E)$, which is a particular nonselfadjoint operator algebra, formed from the C*-correspondence in a way which we shall not go into. This has shed light on problems regarding an enormous class of operator algebras. Remarkably, Muhly and Solel's minimal isometric dilation enjoys also a commutant lifting theorem, and this, in turn, can lead to a Nevanlinna-Pick type interpolation theorem for so-called *noncommutative Hardy algebras*, with a proof reminiscent to the one we gave in Sect. 4 (see [103]).

This dilation theorem of Muhly and Solel is a far reaching generalization of Sz.-Nagy's isometric dilation theorem (Theorem 2.3). In fact, the latter is obtained from the simplest case $E = \mathcal{A} = \mathbb{C}$ of Muhly and Solel's theorem. The row-isometric dilation of a row contraction (Theorem 6.2) is obtained as the "second simplest" case $E = \mathbb{C}^d$ and $\mathcal{A} = \mathbb{C}$. Muhly and Solel's isometric dilation theorem also reduces to some known dilation results in the context of crossed product and semi-crossed product operator algebras, as well in graph C*-algebras.

On the other hand, Andô's theorem, for example, is not a special case of Muhly and Solel's isometric dilation theorem—a single C*-correspondence is not sufficient to encode a pair of commuting contractions. The missing ingredient is the notion of *product systems*. A **product system** over a monoid (i.e., a semigroup with unit e) $\mathcal{S}$ is a family $E = \{E_s\}_{s\in\mathcal{S}}$ of C*-correspondences over a C*-algebra $\mathcal{A}$, such that for every $s, t \in \mathcal{S}$ there exists an isomorphism of correspondences (i.e., an adjointable surjective isometry which is a bimodule map) $u_{s,t} : E_s \odot E_t \to E_{st}$ such that the multiplication $x_s y_t := u_{s,t}(x_s \odot y_t)$ is associative

$$(w_r x_s)y_t = w_r(x_s y_t).$$

(Here $E_s \odot E_t$, denotes the *internal* (or *interior*) *tensor product* of E_s and E_t, sometimes also denoted $E_s \otimes E_t$; see [93, Chapter 4].) A **covariant representation** of a product system $E = \{E_s\}_{s\in\mathcal{S}}$ on $\mathcal{H}$ is a family $T = \{T_s\}_{s\in\mathcal{S}}$ such that for all $s \in \mathcal{S}$, the pair (T_s, T_e) is a covariant representation of E_s on $\mathcal{H}$, which satisfies in addition

$$T_{st}(x_s \odot y_t) = T_s(x_s)T_t(y_t)$$

for all $s, t \in \mathcal{S}$ and all $x_s \in E_s$ and $y_t \in E_t$. An **isometric** representation of E on $\mathcal{K}$ is a covariant representation $V = \{V_s\}_{s\in\mathcal{S}}$ of E such that for all $s \in \mathcal{S}$, the pair (V_s, V_e) is an isometric representation. One then says that V is an **isometric dilation** of T if

1. $\mathcal{H}$ is reducing for π, and $P_{\mathcal{H}}T_e(a)\big|_{\mathcal{H}} = V_e(a)$ for all $a \in \mathcal{A}$,
2. $P_{\mathcal{H}}V(x)\big|_{\mathcal{H}} = T(x)$ for all $x \in E_s$.

The theory of isometric dilations of completely contractive representations of product systems, is analogous to the theory of isometric dilations of semigroups of contractions. Moreover, some of the proofs rely on the same ideas and approaches, albeit at a technical sophistication level that is one order of magnitude higher. In fact, several results (but not all) can be *reduced* to the case of operator semigroups (see [143]). We will see in Sect. 8 that the dilation theory of covariant representations is important for the dilation theory of CP-semigroups.

Here are some sample results. Solel proved a version of Andô's theorem in this setting: every completely contractive covariant representation of a product system over $\mathbb{N}^2$ has an isometric dilation [157, Theorem 4.4]. Solel also proved an analogue of Theorem 3.7 (regular dilations) for product systems over $\mathbb{N}^d$ using a direct proof [158] (see also [152, 153]). Shalit later found another proof by reducing to the case of operator semigroups [143]. The method of [143] was later used in [141, Section 5] to prove a counterpart to Theorem 3.1 (see also [66]). Vernik generalized Opela's result on dilations of contractions commuting according to a graph (see Sect. 3.3) to the setting of product system representations [166]. All of the above results reduce to their counterparts that we discussed in earlier sections, when one considers the special case $\mathcal{A} = E_s = \mathbb{C}$ for all $s \in \mathcal{S}$, where $\mathcal{S}$ is the appropriate monoid.

7 The Operator Algebraic Perspective

The operator algebraic outlook on dilation theory began with Arveson's visionary papers [10, 11]. Arveson sought to develop a systematic study of nonselfadjoint operator algebras, which is based on studying the relations between an operator algebra and the C*-algebras that it generates. From the outset, the approach was general and powerful enough to cover also certain operator spaces. On the one hand, this approach opened the door by which operator algebraic techniques entered into

operator theory: these techniques have shed light on classical dilation theory, and they also created a powerful framework by which new dilation results could be obtained. On the other hand, the general philosophy of dilation theory found its way into operator algebras, and has led to remarkable developments.

In this section I will present Stinespring's dilation theorem, and how Arveson's extension theorem and his notion of *C*-dilation* have made Stinespring's theorem into a "dilation machine" that produces and explains dilation results in operator theory. Then I will briefly discuss how dilation theory is related to the notions of boundary representations and the C*-envelope, which lie at the heart of the above mentioned analysis of the relationship between on operator algebra/space and the C*-algebras that it generates.

I will not attempt to cover all the manifold ways in which dilation theory appears in the theory of operator algebras, and I'll just mention a few (of my favorite) recent examples: [51, 82, 84, 85]. The reader is referred to the survey [48] or the paper [41] in order to get an idea of the role it plays, in particular in operator algebras related to dynamical systems and semicrossed products.

7.1 Completely Positive Maps and Stinespring's Theorem

An **operator space** is a subspace $\mathcal{M} \subseteq B(\mathcal{H})$ of the bounded operators on some Hilbert space $\mathcal{H}$. We say that $\mathcal{M}$ is **unital** if $1 = I_{\mathcal{H}} \in \mathcal{M}$. If $\mathcal{M}$ is a subalgebra of $B(\mathcal{H})$, then it is called an **operator algebra** (note that operator algebras are not assumed to be closed under the adjoint). A unital operator space $\mathcal{M}$ is said to be an **operator system** if it is closed under the adjoint operation. Since every C*-algebra can be represented faithfully on a Hilbert space, we can consider a subspace of a C*-algebra as an operator space (and likewise for unital operator spaces, operator algebras and operator systems). C*-algebras are operator algebras, and unital C*-algebras are operator systems, of course.

An operator space $\mathcal{M} \subseteq B(\mathcal{H})$ inherits from $B(\mathcal{H})$ a norm and a notion of positivity: an element $a \in \mathcal{M}$ is said to be **positive**, if it is positive as an operator on $\mathcal{H}$, i.e., $\langle ah, h\rangle \geq 0$ for all $h \in \mathcal{H}$. Operator systems are spanned by their positive elements, indeed, if $a \in \mathcal{M}$ then its real and imaginary parts are also in $\mathcal{M}$, and if a is selfadjoint then $\frac{1}{2}(\|a\| \cdot 1 \pm a) \geq 0$ and the difference of these two positive elements is a.

As a consequence, it makes sense to speak of positive maps. If $\mathcal{M}$ and $\mathcal{N}$ are operator systems, a linear map $\phi : \mathcal{M} \to \mathcal{N}$ is said to be **positive** if it takes positive elements to positive elements. The matrix spaces $M_n(\mathcal{M}) \subseteq M_n(B(\mathcal{H})) = B(\mathcal{H}^n)$ and $M_n(\mathcal{N})$ are also operator systems, and then ϕ induces a linear map $\phi^{(n)} : M_n(\mathcal{M}) \to M_n(\mathcal{N})$

$$\phi^{(n)} = \phi \otimes \mathbf{id}_{M_n} : M_n(\mathcal{M}) = \mathcal{M} \otimes M_n \to M_n(\mathcal{N}) = \mathcal{N} \otimes M_n$$

acting elementwise as

$$\phi^{(n)} : (a_{ij})_{i,j=1}^n \mapsto (\phi(a_{ij}))_{i,j=1}^n \in M_n(\mathcal{N}).$$

The map ϕ is said be **completely positive** (or **CP** for short) if $\phi^{(n)}$ is positive for all n. Likewise, ϕ is said to be **completely contractive** (or **CC** for short) if $\phi^{(n)}$ is contractive for all n. A map is **UCP** if it is a unital CP map, and **UCC** if it is a unital CC map.

Completely positive maps were introduced by Stinespring [159], but it was Arveson who observed how important they are and opened the door to their becoming an indispensable tool in operator theory and operator algebras [10]. There are several excellent sources to learn about operator spaces/systems and completely positive (and bounded) maps; see for example [118] and [122].

Completely positive maps arise also in mathematical physics in a natural way [89]. The evolution of an open quantum system is described by a semigroup of completely positive maps [49], and noisy channels in quantum information theory are modelled as trace preserving completely positive maps [107]. In quantum probability [111], semigroups of unit preserving completely positive maps play the role of Markov semigroups.

The simplest examples of completely positive maps are $*$-homomorphisms between C*-algebras. Next, a map of the form $B(\mathcal{K}) \ni T \mapsto V^*TV \in B(\mathcal{H})$, where V is some fixed operator in $B(\mathcal{H}, \mathcal{K})$, is readily seen to be completely positive. Since compositions of CP maps are evidently CP, we see that whenever $\mathcal{A}$ is a C*-algebra, $\pi : \mathcal{A} \to B(\mathcal{K})$ is a $*$-homomorphism, and $V \in B(\mathcal{H}, \mathcal{K})$, then the map $a \mapsto V^*\pi(a)V$ is a CP map. The following fundamental theorem shows that essentially all CP maps on C*-algebras are of this form.

Theorem 7.1 (Stinespring's Theorem [159]) *Let $\mathcal{A}$ be a unital C*-algebra and let $\phi : \mathcal{A} \to B(\mathcal{H})$ be a CP map. Then there exists a Hilbert space $\mathcal{K}$, an operator $V \in B(\mathcal{H}, \mathcal{K})$, and a $*$-representation $\pi : \mathcal{A} \to B(\mathcal{K})$ such that*

$$\phi(a) = V^*\pi(a)V \quad , \quad \text{for all } a \in \mathcal{A}.$$

The tuple $(\pi, \mathcal{K}, V)$ can be chosen such that $\mathcal{K} = [\pi(\mathcal{A})\mathcal{H}]$—the smallest closed subspace containing $\pi(a)h$ for all $a \in \mathcal{A}$ and $h \in \mathcal{H}$—and in this case the triple $(\pi, \mathcal{K}, V)$ is unique up to unitary equivalence.

Proof On the algebraic tensor product $\mathcal{A} \otimes \mathcal{H}$, we define a semi-inner product by setting $\langle a \otimes g, b \otimes h \rangle = \langle g, \phi(a^*b)h \rangle_{\mathcal{H}}$ and extending sesquilinearly (the complete positivity guarantees that this is a positive semidefinite form). Quotienting out the kernel and then completing gives rise to the Hilbert space $\mathcal{K}$. The image of all the elementary tensors $b \otimes h \in \mathcal{A} \otimes \mathcal{H}$ in $\mathcal{K}$ form a total set, and we continue to denote these images as $b \otimes h$. One needs to check that for every $a \in \mathcal{A}$, the map $\pi(a) : b \otimes h \mapsto ab \otimes h$ extends to a well defined, bounded linear operator on $\mathcal{K}$. Once this is done, it is easy to verify that the map $a \mapsto \pi(a)$ is a $*$-homomorphism.

To recover ϕ, we define $V : \mathcal{H} \to \mathcal{K}$ by $V(h) = 1 \otimes h$, and then all that remains to do is to compute $\langle V^*(a \otimes h), g \rangle = \langle a \otimes h, V(g) \rangle = \langle a \otimes h, 1 \otimes g \rangle = \langle h, \phi(a^*)g \rangle$,

so $V^*(a \otimes h) = \phi(a)h$, and thus

$$V^*\pi(a)Vh = V^*(a \otimes h) = \phi(a)h,$$

as required. If $[\pi(\mathcal{A})\mathcal{H}] \subsetneq \mathcal{K}$, then we replace $\mathcal{K}$ with $[\pi(\mathcal{A})\mathcal{H}]$, and obtain a minimal representation. The uniqueness is a standard matter, and is left to the reader. ■

If $\mathcal{K} = [\pi(\mathcal{A})\mathcal{H}]$, then $(\pi, \mathcal{K}, V)$ (or just π sometimes) is called the **minimal Stinespring representation of** ϕ.

Remark 7.2 If ϕ is unital, then $1 = \phi(1) = V^*\pi(1)V = V^*V$, so V is an isometry. In this case it is convenient to identify $\mathcal{H}$ with $V\mathcal{H} \subseteq \mathcal{K}$, and the Stinespring representation manifests itself as a dilation

$$\phi(a) = P_{\mathcal{H}}\pi(a)\big|_{\mathcal{H}}.$$

In this situation, the minimal Stinespring representation is referred to as the **minimal Stinespring dilation** of ϕ.

7.2 Arveson's Extension Theorem and C*-Dilations

The utility of completely positive maps comes from the following extension theorem of Arveson. For a proof, see Arveson's paper or Paulsen's book [118, Chapter 7].

Theorem 7.3 (Arveson's Extension Theorem [10]) *Let $\mathcal{M}$ be an operator system in a C*-algebra $\mathcal{A}$, and let $\phi : \mathcal{M} \to B(\mathcal{H})$ be a CP map. Then there exists a CP map $\hat{\phi} : \mathcal{A} \to B(\mathcal{H})$ such that $\|\hat{\phi}\| = \|\phi\|$ and which extends ϕ, i.e., $\hat{\phi}(a) = \phi(a)$ for all $a \in \mathcal{M}$.*

We will now see how the combination of Stinespring's dilation theorem and Arveson's extension theorem serve as kind of all purpose "dilation machine", that produces dilation theorems in varied settings.

Let $1 \in \mathcal{M} \subseteq \mathcal{A}$ be a unital operator space. A linear map $\phi : \mathcal{M} \to B(\mathcal{H})$ is said to have a **C*-dilation** to $\mathcal{A}$ if there exists a $*$-representation $\pi : \mathcal{A} \to B(\mathcal{K})$, $\mathcal{K} \supseteq \mathcal{H}$, such that

$$\phi(a) = P_{\mathcal{H}}\pi(a)\big|_{\mathcal{H}}\ , \quad \text{for all } a \in \mathcal{M}.$$

Arveson showed that every UCP map is UCC, and that, conversely, every UCC map as above extends to a UCP map $\widetilde{\phi} : \widetilde{\mathcal{M}} := \mathcal{M} + \mathcal{M}^* \to B(\mathcal{H})$ given by $\widetilde{\phi}(a + b^*) = \phi(a) + \phi(b)^*$. Combining this basic fact with Theorems 7.1 and 7.3 we obtain the following versatile dilation theorem.

Theorem 7.4 *Every UCC or UCP map has a C*-dilation.*

Arveson's dilation theorem[1] (Theorem 5.1) follows from the above theorem, once one carefully works through the delicate issues of joint spectrum, Shilov boundary and functional calculus (see [11]). We shall illustrate the use of the dilation machine by proving Arveson's dilation theorem for the simple but representative example of the polydisc $\overline{\mathbb{D}}^d$.

Theorem 7.5 (Arveson's Dilation Theorem for the Polydisc) *A d-tuple of commuting contractions $T = (T_1, \ldots, T_d)$ has a unitary dilation if and only if the polydisc $\overline{\mathbb{D}}^d$ is a complete spectral set for T.*

Proof To relate the statement of the theorem to the language of Sect. 5, we note that the Shilov boundary of $\overline{\mathbb{D}}^d$ is just the torus $(\partial\overline{\mathbb{D}})^d = \mathbb{T}^d$, and therefore a unitary dilation is nothing but a normal dilation with joint spectrum contained in $\mathbb{T}^d$. Recall that $\overline{\mathbb{D}}^d$ being a complete spectral set is equivalent to that

$$\|p(T)\| \leq \|p\|_\infty := \sup_{z \in \overline{\mathbb{D}}^d} \|p(z)\| \tag{7.1}$$

for every matrix valued polynomial p (since $\overline{\mathbb{D}}^d$ is convex it suffices to consider matrix valued polynomials, and there is no need to worry about matrix valued rational functions).

If $U = (U_1, \ldots, U_d)$ is a tuple of commuting unitaries and p is a matrix valued polynomial, then, using the spectral theorem, it is not hard to see that $\|p(U)\| = \sup_{z \in \sigma(U)} \|p(z)\| \leq \|p\|_\infty$. Now, if U is a dilation of T then $\|p(T)\| \leq \|p(U)\|$, and so the inequality (7.1) holds. That was the easy direction.

Conversely, suppose that $\overline{\mathbb{D}}^d$ is a complete spectral set for a commuting tuple $T \in B(\mathcal{H})^d$, that is, suppose that (7.1) holds for every matrix valued polynomial p. Let $\mathcal{M} = \mathbb{C}[z_1, \ldots, z_d]$ be the space of polynomials in d variables, considered as a unital subspace of the C*-algebra $C(\mathbb{T}^d)$, equipped with the usual supremum norm. It is useful to note that

$$\sup_{z \in \overline{\mathbb{D}}^d} \|p(z)\| = \sup_{z \in \mathbb{T}^d} \|p(z)\|,$$

by applying the maximum modulus principle in several variables. The fact that $\overline{\mathbb{D}}^d$ is a complete spectral set for T implies that the unital map $\phi : \mathcal{M} \to B(\mathcal{H})$ given by $\phi(p) = p(T)$, is completely contractive. By Theorem 7.4, ϕ has a C*-dilation $\pi : C(\mathbb{T}^d) \to B(\mathcal{K})$, such that

$$p(T) = \phi(p) = P_{\mathcal{H}} \pi(p)\big|_{\mathcal{H}} \ , \ \ p \in \mathcal{M}.$$

[1]The reader should be aware that Theorem 7.4 is sometimes referred to as *Arveson's dilation theorem*, whereas I used this name already for the more specific Theorem 5.1.

Now, π is a $*$-representation, and the coordinate functions $z_1, \ldots, z_d \in C(\mathbb{T}^d)$ are all unitary, so $U_i = \pi(z_i)$, $i = 1, \ldots, d$, are commuting unitaries. Since $\pi(p) = p(\pi(z_1), \ldots, \pi(z_d))$, we find that

$$p(T) = P_{\mathcal{H}} p(U)\big|_{\mathcal{H}}$$

for all $p \in \mathcal{M}$, that is, U is a unitary dilation of T, as required. ■

Following Arveson, the above method has been used extensively for proving the existence of dilations in certain situations. The burden is then shifted from the task of *constructing* a dilation, to that of showing that certain naturally defined maps are UCP or UCC. In other words, by proving an inequality one obtains an existence proof—a good bargain from which analysts have profited for a century. Of course, "good bargain" does not mean that we cheat, we still need to prove something. Let me give an example of how this works.

Example 7.6 We now prove that for every contraction $T \in B(\mathcal{H})$, the map $\Phi : \mathbb{C}[z] \to B(\mathcal{H})$ given by $\Phi(p) = p(T)$ is UCC. Combining this with Arveson's dilation theorem for the disc (the case $d = 1$ in Theorem 7.5), we obtain a genuinely new proof of Sz.-Nagy's unitary dilation theorem (Theorem 2.1). The reader should be able to adapt the details of this proof to get a proof that every row contraction has a row-isometric dilation, and that every pure commuting row contraction can be modelled by the d-shift (for hints, see the introduction of [125] or [8], respectively).

Let S be the unilateral shift on ℓ^2, and let e_n denote the nth standard basis vector in ℓ^2. If T is a contraction and $r \in (0, 1)$, we let $D_{rT} = (I - r^2 TT^*)^{1/2}$, and define $K_r(T) : \mathcal{H} \to \ell^2 \otimes \mathcal{H}$ by

$$K_r(T)h = \sum_n e_n \otimes \left(r^n D_{rT} T^{n*} h\right),$$

for all $h \in \mathcal{H}$. We compute:

$$\begin{aligned} K_r(T)^* K_r(T)h &= \sum r^{2n} T^n D_{rT}^2 T^{n*} h \\ &= \sum r^{2n} T^n (I - r^2 TT^*) T^{n*} h \\ &= \sum r^{2n} T^n T^{n*} h - \sum r^{2(n+1)} T^{n+1} T^{(n+1)*} h = h \end{aligned}$$

so that $K_r(T)$ is an isometry. On $C^*(S)$ we define a UCP map

$$\phi_r(a) = K_r(T)^* (a \otimes I) K_r(T).$$

We compute that

$$\begin{aligned}\phi_r(S) &= K_r(T)^*(S \otimes I)K_r(T)h \\ &= K_r(T)^* \sum e_{n+1} \otimes \left(r^n D_{rT} T^{n*} h\right) \\ &= \sum r^{2n+1} T^{n+1} D_{rT}^2 T^{n*} h = rTh.\end{aligned}$$

Likewise, $\phi_r(S^n) = r^n T^n$ for all $n \in \mathbb{N}$.

Now we define a UCP map $\Phi := \lim_{r \nearrow 1} \phi_r$. Then $\Phi(S^n) = T^n$ for all n. We see that the map $p(S) \mapsto p(T)$ is UCC. But S is unitarily equivalent to the multiplication operator M_z on H^2 (see Sect. 4.3). Thus, for every matrix valued polynomial p

$$\|p(T)\| \leq \|p(S)\| = \|p(M_z)\| = \sup_{|z|=1} \|p(z)\|,$$

so $\overline{\mathbb{D}}$ is a complete spectral set for T. By Theorem 7.5, T has a unitary dilation.

7.3 Boundary Representations and the C*-Envelope

The ideas in this section are best motivated by the following classical example.

Example 7.7 Consider the disc algebra $A(\mathbb{D})$, which is equal to the closure of the polynomials with respect to the norm $\|p\|_\infty = \sup_{z \in \overline{\mathbb{D}}} |p(z)|$. The disc algebra is an operator algebra, being a subalgebra of the C*-algebra $C(\overline{\mathbb{D}})$ of continuous functions on the disc $\overline{\mathbb{D}}$. Moreover, $C^*(A(\overline{\mathbb{D}})) = C(\overline{\mathbb{D}})$, that is, the C*-subalgebra generated by $A(\mathbb{D}) \subseteq C(\overline{\mathbb{D}})$ is equal to $C(\overline{\mathbb{D}})$. However, $C(\overline{\mathbb{D}})$ is not determined uniquely by being "the C*-algebra generated by the disc algebra". In fact, by the maximum modulus principle, $A(\overline{\mathbb{D}})$ is also isometrically isomorphic to the closed subalgebra of $C(\mathbb{T})$ generated by all polynomials, and the C*-subalgebra of $C(\mathbb{T})$ generated by the polynomials is equal to $C(\mathbb{T})$.

Now, $C(\mathbb{T})$ is the quotient of $C(\overline{\mathbb{D}})$ by the ideal of all continuous functions vanishing on the circle $\mathbb{T}$. If $\pi : C(\overline{\mathbb{D}}) \to C(\mathbb{T})$ denotes the quotient map, then $\pi(f) = f\big|_{\mathbb{T}}$, and we note, using the maximum principle again, that π is isometric on $A(\mathbb{D})$. It turns out that $\mathbb{T}$ is the minimal subset $E \subseteq \overline{\mathbb{D}}$ such that the map $f \mapsto f\big|_E$ is isometric on $A(\mathbb{D})$.

The above phenomenon arises in all *uniform algebras*, that is, in all unital subalgebras $\mathcal{A}$ of $C(X)$ that separate the points of X, where X is some compact Hausdorff space. For every such algebra there exists a set $\partial_{\mathcal{A}} \subseteq X$—called the **Shilov boundary** of $\mathcal{A}$—which is the unique minimal closed subset $E \subseteq X$ such that $f \mapsto f\big|_E$ is isometric (see [68] for the theory of uniform algebras).

In the spirit of noncommutative analysis, Arveson sought to generalize the Shilov boundary to the case where the commutative C*-algebra $C(X)$ is replaced by a noncommutative C*-algebra $\mathcal{B} = C^*(\mathcal{A})$ generated by a unital operator algebra $\mathcal{A}$. An ideal $\mathcal{I} \lhd \mathcal{B}$ is said to be a **boundary ideal** for $\mathcal{A}$ in $\mathcal{B}$, if the restriction of the quotient map $\pi : \mathcal{B} \to \mathcal{B}/\mathcal{I}$ to $\mathcal{A}$ is completely isometric. The **Shilov ideal** of $\mathcal{A}$ in $\mathcal{B}$ is the unique largest boundary ideal for $\mathcal{A}$ in $\mathcal{B}$. If $\mathcal{J}$ is the Shilov ideal of $\mathcal{A}$ in $\mathcal{B}$, then **C*-envelope** of $\mathcal{A}$ is defined to be the C*-algebra $C_e^*(\mathcal{A}) = \mathcal{B}/\mathcal{J}$. (The above notions were introduced in [10, 11], but it took some time until the terminology settled down. A good place for the beginner to start learning this stuff is [118].)

Example 7.8 If $\mathcal{B} = C(X)$ is a commutative C*-algebra generated by the uniform algebra $\mathcal{A}$, then the Shilov ideal of $\mathcal{A}$ is just the ideal $\mathcal{I}_{\partial_\mathcal{A}}$ of functions vanishing on the Shilov boundary $\partial_\mathcal{A}$. In this case $C_e^*(\mathcal{A}) = C(\partial_\mathcal{A})$.

The C*-envelope $C_e^*(\mathcal{A}) = \mathcal{B}/\mathcal{J} = \pi(\mathcal{B})$ has the following universal property: if $i : \mathcal{A} \to \mathcal{B}'$ is a completely isometric homomorphism such that $\mathcal{B}' = C^*(i(\mathcal{A}))$, then there exists a unique surjective $*$-homomorphism $\rho : \mathcal{B}' \to C_e^*(\mathcal{A})$ such that $\pi(a) = \rho(i(a))$ for all $a \in \mathcal{A}$. It follows that the C*-envelope depends only on the structure of $\mathcal{A}$ as an operator algebra, not on the concrete realization $\mathcal{A} \subseteq \mathcal{B}$ with which we started. Thus, if $\mathcal{A}_i \subseteq \mathcal{B}_i = C^*(\mathcal{A}_i)$ for $i = 1, 2$ have trivial Shilov ideals, then every completely isometric homomorphism $\phi : \mathcal{A}_1 \to \mathcal{A}_2$ extends to $*$-isomorphism $\rho : \mathcal{B}_1 \to \mathcal{B}_2$.

In fact, the algebraic structure is not essential here, and the above notions also make sense for unital operator spaces. For some purposes, it is most convenient to work with operator systems, and focus is then shifted to this case (as in the next subsection). For example, the Shilov ideal of an operator system $\mathcal{S} \subseteq \mathcal{B} = C^*(\mathcal{S})$ is the largest ideal $\mathcal{I} \lhd \mathcal{B}$ such that the quotient map $\pi : \mathcal{B} \mapsto \mathcal{B}/\mathcal{I}$ restricts to a complete isometry on $\mathcal{S}$, etc.

How does one find the Shilov ideal? Let us return to Example 7.7 (recalling also Example 7.8). If $A(\overline{\mathbb{D}})$ is given as a subalgebra of $C(\overline{\mathbb{D}})$, how could one characterize its Shilov boundary? A little bit of function theory shows that a point $z \in \overline{\mathbb{D}}$ is in the unit circle $\mathbb{T}$ if and only if it has a unique representing measure. Recall that when we have a uniform algebra $\mathcal{A} \subseteq C(X)$, a probability measure μ is said to be a **representing measure** for x if

$$f(x) = \int_X f d\mu$$

for all $f \in \mathcal{A}$. For example, the Lebesgue measure on the circle is a representing measure for the point 0, because

$$f(0) = \frac{1}{2\pi} \int_0^{2\pi} f(e^{it}) dt$$

for every $f \in A(\overline{\mathbb{D}})$, as an easy calculation shows. Every point can be represented by the delta measure δ_z; the points on the circle are singled out by being those with

a unique representing measure, that is, they can be represented *only* by the delta measure. In the general case of a uniform algebra $\mathcal{A} \subseteq C(X)$, the points in X that have a unique representing measure are referred to as the **Choquet boundary** of $\mathcal{A}$. It is not hard to show that the Choquet boundary of $A(\overline{\mathbb{D}})$ is $\mathbb{T}$. In general, the Choquet boundary of a uniform algebra is dense in the Shilov boundary.

Let return to the noncommutative case, so let $\mathcal{A} \subseteq \mathcal{B} = C^*(\mathcal{A})$ be again a unital operator algebra generating a C*-algebra. Point evaluations correspond to the irreducible representations of a commutative C*-algebra, and probability measures correspond to states, that is, positive maps into $\mathbb{C}$. With this in mind, the reader will hopefully agree that the following generalization is potentially useful: an irreducible representation $\pi : \mathcal{B} \to B(\mathcal{H})$ is said to be a **boundary representation** if the only UCP map $\Phi : \mathcal{B} \to B(\mathcal{H})$ that extends $\pi\big|_{\mathcal{A}}$ is π itself.

Arveson proved in [10] that if an operator algebra $\mathcal{A} \subseteq \mathcal{B} = C^*(\mathcal{A})$ has **sufficiently many boundary representations**, in the sense that

$$\|A\| = \sup\{\|\pi^{(n)}(A)\| : \pi : \mathcal{B} \to B(\mathcal{H}_\pi) \text{ is a boundary representation}\}$$

for all $A \in M_n(\mathcal{A})$, then the Shilov ideal exists, and is equal to the intersection of all boundary ideals. For some important operator algebras, the existence of sufficiently many boundary representations was obtained (see also [11]), but the problem of existence of boundary representations in general remained open almost 45 years.[2] Following a sequence of important developments [15, 53, 101], Davidson and Kennedy proved that every operator system has sufficiently many boundary representations [42]. Their proof implies that every unital operator space, and in particular every unital operator algebra, has sufficiently many boundary representations as well.

7.4 Boundary Representations and Dilations

It is interesting that the solution to the existence problem of boundary representations was obtained through dilations. Davidson and Kennedy worked in the setting of operator systems, so let us follow them in this subsection.

If $\mathcal{S} \subseteq \mathcal{B} = C^*(\mathcal{S})$ is an operator system inside the C*-algebra that it generates, an irreducible representation $\pi : \mathcal{B} \to B(\mathcal{H})$ is a boundary representation if the only UCP map $\Phi : \mathcal{B} \to B(\mathcal{H})$ that extends $\pi\big|_{\mathcal{S}}$ is π itself. This leads to the following definition: a UCP map $\phi : \mathcal{S} \to B(\mathcal{H})$ is said to have the **unique extension property** if there exists a unique UCP map $\Phi : \mathcal{B} \to B(\mathcal{H})$ that extends ϕ, and, moreover, *this* Φ *is a* $*$*-representation*. Thus, an irreducible representation π is a

[2] The existence of the C*-envelope was obtained much earlier, without making use of boundary representations; see [118].

boundary representation if and only if the restriction $\pi\big|_{\mathcal{S}}$ has the unique extension property.

Now, the unique extension property nicely captures the idea that $\mathcal{S}$ is in some sense rigid in $\mathcal{B}$, but it is hard to verify it in practice. The following notion is more wieldy. A UCP map $\psi : \mathcal{S} \to B(\mathcal{K})$ is said to be a **dilation** of a UCP map $\phi : \mathcal{S} \to B(\mathcal{H})$ if $\phi(a) = P_{\mathcal{H}}\psi(a)\big|_{\mathcal{H}}$ for all $a \in \mathcal{S}$. The dilation ψ is said to be a **trivial** dilation if $\mathcal{H}$ is reducing for $\psi(\mathcal{S})$, that is, $\psi = \phi \oplus \rho$ for some UCP map ρ. A UCP map ϕ is said to be **maximal** if it has only trivial dilations.

Penetrating observations of Muhly and Solel [101], and consequently Dritschel and McCullough [53], can be reformulated as the following theorem. The beauty is that the notion of maximality is intrinsic to the operator system $\mathcal{S}$, and does not take the containing C*-algebra $\mathcal{B}$ into account (similar reformulations exist for the categories of unital operator spaces and operator algebras).

Theorem 7.9 *A UCP map $\phi : \mathcal{S} \to B(\mathcal{H})$ has the unique extension property if and only if it is maximal.*

Following Dritschel and McCullough's proof of the existence of the C*-envelope [53] and Arveson's consequent work [15], Davidson and Kennedy proved the following theorem (as above, similar reformulations exist for the categories of unital operator spaces and operator algebras).

Theorem 7.10 *Every UCP map can be dilated to a maximal UCP map, and every pure UCP map can be dilated to a pure maximal UCP map.*

Davidson and Kennedy proved that *pureness* guarantees that the $*$-representation, which is the unique UCP extension of the maximal dilation, is in fact irreducible. Moreover, they showed that pure UCP maps completely norm an operator space. Thus, by dilating sufficiently many pure UCP maps, and making use of the above theorems, they concluded that there exist sufficiently many boundary representations [42].

Example 7.11 Let us see what are the maximal dilations in the case of the disc algebra $A(\overline{\mathbb{D}}) \subseteq C(\overline{\mathbb{D}})$ (we switch back from the category of operator systems to the category of unital operator algebras). A representation π of $C(\overline{\mathbb{D}})$ is determined uniquely by a normal operator N with spectrum in $\overline{\mathbb{D}}$ by the relation $N = \pi(z)$. A UCC representation $\phi : A(\overline{\mathbb{D}}) \to B(\mathcal{H})$ is determined uniquely by the image of the coordinate function z, which is a contraction $T = \phi(z) \in B(\mathcal{H})$. Conversely, by the $A(\overline{\mathbb{D}})$ functional calculus (see Sect. 2.2), every contraction $T \in B(\mathcal{H})$ gives rise to a UCC homomorphism of $A(\overline{\mathbb{D}})$ into $B(\mathcal{H})$. In this context, a dilation of a UCC map ϕ into $B(\mathcal{H})$ is simply a representation $\rho : A(\overline{\mathbb{D}}) \to B(\mathcal{K})$ such that

$$f(T) = \phi(f) = P_{\mathcal{H}}\rho(f)\big|_{\mathcal{H}} = P_{\mathcal{H}}f(V)\big|_{\mathcal{H}},$$

for all $f \in A(\overline{\mathbb{D}})$, where $V = \rho(z) \in B(\mathcal{K})$. So ρ is a dilation of ϕ if and only if $V = \rho(z)$ is a (power) dilation of $T = \phi(z)$ in the sense of Sect. 2.

With the above notation, it is not hard to see the following two equivalent statements: *(i) a dilation $\rho : A(\overline{\mathbb{D}}) \to B(\mathcal{K})$ is maximal if and only if V is a unitary*, and *(ii) a representation $\pi : C(\overline{\mathbb{D}}) \to B(\mathcal{K})$ is such that $\rho = \pi\big|_{A(\overline{\mathbb{D}})}$ has the unique extension property if and only if V is a unitary*. The fact that every UCC representation of $A(\overline{\mathbb{D}})$ has a maximal dilation, is equivalent to the fact that every contraction has a unitary dilation.

What are the boundary representations of the disc algebra? The irreducible representations of $C(\overline{\mathbb{D}})$ are just point evaluations δ_z for $z \in \overline{\mathbb{D}}$. The boundary representations are those point evaluations δ_z whose restriction to $A(\overline{\mathbb{D}})$ have a unique extension to a UCP map $C(\overline{\mathbb{D}}) \to \mathbb{C}$. But UCP maps into the scalars are just states, and states of $C(\overline{\mathbb{D}})$ are given by probability measures. Hence, boundary representations are point evaluations δ_z such that z has a unique representing measure, so that $z \in \mathbb{T}$.

The Shilov ideal can be obtained as the intersection of the kernels of the boundary representations, and so it is the ideal of functions vanishing on $\mathbb{T}$. The C*-envelope is the quotient of $C(\overline{\mathbb{D}})$ by this ideal, thus it is $C(\mathbb{T})$, as we noted before.

Note that to find the boundary representations of the disc algebra we did not need to invoke the machinery of maximal dilations. In the commutative case, the existence of sufficiently many boundary representations is no mystery: all of them are obtained as extensions of evaluation at a boundary point. The machinery of maximal dilations allows us to find the boundary representations in the noncommutative case, where there are no function theoretic tools at our disposal.

8 Dilations of Completely Positive Semigroups

The dilation theory of semigroups of completely positive maps can be considered as a kind of "quantization" of classical isometric dilation theory of contractions on a Hilbert space. The original motivation comes from mathematical physics [49, 60]. The theory is also very interesting and appealing in itself, having connections and analogies (and also surprising differences) with classical dilation theory. Studying dilations of CP-semigroups has led to the discovery of results and some structures that are interesting in themselves.

In this section, I will only briefly review some results in dilation theory of CP-semigroups from the last two decades, of the kind that I am interested in. There are formidable subtleties and technicalities that I will either ignore, or only gently hint at. For a comprehensive and up-to-date account, including many references (also to other kinds of dilations), see [149]. All of the facts that we state without proof or reference below have either a proof or a reference in [149]. The reader is also referred to the monographs [14] and [111] for different takes on quantum dynamics and quantum probability.

8.1 CP-Semigroups, E-Semigroups and Dilations

Let $\mathcal{S}$ be a commutative monoid. By a **CP-semigroup** we mean family $\Theta = \{\Theta_s\}_{s\in\mathcal{S}}$ of contractive CP-maps on a unital C*-algebra $\mathcal{B}$ such that

1. $\Theta_0 = \mathbf{id}_{\mathcal{B}}$,
2. $\Theta_s \circ \Theta_t = \Theta_{s+t}$ for all $s, t \in \mathcal{S}$.

If $\mathcal{S}$ carries some topology, then we usually require $s \mapsto \Theta_s$ to be continuous in some sense. A CP-semigroup Θ is said to be a **Markov semigroup** (or a **CP$_0$-semigroup**) if every Θ_s is a unital map. A CP-semigroup Θ is called an **E-semigroup** if every element Θ_s is a $*$-endomorphism. Finally, an **E$_0$-semigroup** is a Markov semigroup which is also an E-semigroup.

One-parameter semigroups of $*$-automorphisms model the time evolution in a closed (or a reversible) quantum mechanical system, and one-parameter Markov semigroups model the time evolution in an open (or irreversible) quantum mechanical system [49].

The prototypical example of a CP-semigroup is given by

$$\Theta_s(b) = T_s b T_s^* \ , \ \ b \in \mathcal{B} \tag{8.1}$$

where $T = \{T_s\}_{s\in\mathcal{S}}$ is a semigroup of contractions in $\mathcal{B}$. We call such a semigroup **elementary**. Of course, not all CP-semigroups are elementary.

For us, a **dilation** of a CP-semigroup is a triplet $(\mathcal{A}, \alpha, p)$, where $\mathcal{A}$ is a C*-algebra, $p \in \mathcal{A}$ is a projection such that $\mathcal{B} = p\mathcal{A}p$, and $\alpha = \{\alpha_s\}_{s\in\mathcal{S}}$ is an E-semigroup on $\mathcal{A}$, such that

$$\Theta_s(b) = p\alpha_s(b)p$$

for all $b \in \mathcal{B}$ and $s \in \mathcal{S}$. A **strong dilation** is a dilation in which the stronger condition

$$\Theta_s(pap) = p\alpha_s(a)p$$

holds for all $a \in \mathcal{A}$ and $s \in \mathcal{S}$. It is a fact, not hard to show, that if Θ is a Markov semigroup then every dilation is strong. Examples show that this is not true for general CP-semigroups. Sometimes, to lighten the terminology a bit, we just say that α is a (strong) dilation.

Remark 8.1 It is worth pausing to emphasize that the dilation defined above is entirely different from Stinespring's dilation: the Stinespring dilations of Θ_s and $\Theta_{s'}$ cannot be composed. The reader should also be aware that there are other notions of dilations, for example in which the "small" algebra $\mathcal{B}$ is embedded as a *unital* subalgebra of the "large" algebra $\mathcal{A}$ (see, e.g., [67] or [167] and the references therein), or where additional restrictions are imposed (see, e.g. [91], and the papers that cite it).

The most important is the one-parameter case, where $\mathcal{S} = \mathbb{N}$ or $\mathbb{R}_+$. The following result was proved first by Bhat in the case $\mathcal{B} = B(\mathcal{H})$ [24] (slightly later SeLegue gave a different proof in the case $\mathcal{B} = B(\mathcal{H})$ [140]), then it was proved by Bhat and Skeide for general unital C*-algebras [28], and then by Muhly and Solel in the case of unital semigroups on von Neumann algebras [102] (slightly later this case was also proved by Arveson [14, Chapter 8]).

Theorem 8.2 *Every CP-semigroup $\Theta = \{\Theta_s\}_{s\in\mathcal{S}}$ on $\mathcal{B}$ over $\mathcal{S} = \mathbb{N}$ or $\mathcal{S} = \mathbb{R}_+$ has a strong dilation $(\mathcal{A}, \alpha, p)$.*

Moreover, if Θ is unital then α can be chosen unital; if $\mathcal{B}$ is a von Neumann algebra and Θ is normal, then $\mathcal{A}$ can be taken to be a von Neumann algebra and α normal; and if further $\mathcal{S} = \mathbb{R}_+$ and Θ is point weak-$*$ continuous, in the sense that $t \mapsto \rho(\Theta_t(b))$ is continuous for all $\rho \in \mathcal{B}_*$, then α can also be chosen to be point weak-$*$ continuous.

Proof Let us illustrate the proof for the case where $\mathcal{B} = B(\mathcal{H})$, $\mathcal{S} = \mathbb{N}$, θ is a normal contractive CP map on $B(\mathcal{H})$, and $\Theta_n = \theta^n$ for all n. Then it is well known that θ must have the form $\theta(b) = \sum_i T_i b T_i^*$ for a row contraction $T = (T_i)$ (see, e.g., [89, Theorem 1]). By Theorem 6.2, T has a row isometric coextension $V = (V_i)$ on a Hilbert space $\mathcal{K}$. Letting $\mathcal{A} = B(\mathcal{K})$, $p = P_{\mathcal{H}}$ and

$$\alpha(a) = \sum_i V_i a V_i^*$$

we obtain a strong dilation (as the reader will easily verify). ■

The above proof suggests that there might be strong connections between operator dilation theory and the dilation theory of CP-semigroups. This is true, but there are some subtleties. Consider an elementary CP-semigroup (8.1) acting on $\mathcal{B} = B(\mathcal{H})$, where T is a semigroup of contractions on $\mathcal{H}$. If $V = \{V_s\}_{s\in\mathcal{S}}$ is an isometric dilation of T, then $\alpha(a) = V_s a V_s^*$ is a dilation of Θ. If V is a coextension, then α is a strong dilation. Thus, we know that we can find (strong) dilations for elementary semigroups when the semigroup is such that isometric dilations (coextensions) exist for every contractive semigroup; for example, when $\mathcal{S} = \mathbb{N}, \mathbb{N}^2, \mathbb{R}_+, \mathbb{R}_+^2$. Moreover, we expect that we won't always be able to dilate CP-semigroups over monoids for which isometric dilations don't always exist (for example $\mathcal{S} = \mathbb{N}^3$). This analogy based intuition is almost correct, and usually helpful.

Remark 8.3 As in classical dilation theory, there is also a notion of *minimal dilation*. However, it turns out that there are several reasonable notions of minimality. In the setting of normal continuous semigroups on von Neumann algebras, the most natural notions of minimality turn out to be equivalent in the one-parameter case, but in the multi-parameter case they are not equivalent. See [14, Chapter 8] and [149, Section 21] for more on this subject.

Theorem 8.2 has the following interesting interpretation. A one-parameter CP-semigroup models the time evolution in an open quantum dynamical system, and a one-parameter automorphism semigroup models time evolution in a closed one. In many cases, E-semigroups can be extended to automorphism semigroups, and so Theorem 8.2 can be interpreted as saying that every open quantum dynamical system can be embedded in a closed (reversible) one (this interpretation was the theoretical motivation for the first dilation theorems, see [49, 60]).

8.2 Main Approaches and Results

There are two general approaches by which strong dilations of CP-semigroups can be constructed.

The Muhly–Solel Approach One approach, due to Muhly and Solel [102], seeks to represent a CP-semigroup in a form similar to (8.1), and then to import ideas from classical dilation theory. To give little more detail, if $\Theta = \{\Theta_s\}_{s\in\mathcal{S}}$ is a CP-semigroup on a von Neumann algebra $\mathcal{B} \subseteq B(\mathcal{H})$, then one tries to find a product system $E^{\odot} = \{E_s\}_{s\in\mathcal{S}}$ of $\mathcal{B}'$-correspondences over $\mathcal{S}$ (see Sect. 6.5) and a completely contractive covariant representation $T = \{T_s\}_{s\in\mathcal{S}}$ such that

$$\Theta_s(b) = \widetilde{T}_s(\mathbf{id}_{E_s} \odot b)\widetilde{T}_s^* \tag{8.2}$$

for every $s \in \mathcal{S}$ and $b \in \mathcal{B}$. Here, $\widetilde{T}_s : E_s \odot \mathcal{H} \to \mathcal{H}$ is given by $\widetilde{T}_s(x \odot h) = T_s(x)h$. This form is reminiscent of (8.1), and it is begging to try to dilate Θ by constructing an isometric dilation V for T, and then defining

$$\alpha_s(a) = \widetilde{V}_s(\mathbf{id}_{E_s} \odot a)\widetilde{V}_s^*$$

for $a \in \mathcal{A} := V_0(\mathcal{B}')'$. This is a direct generalization of the apporach to dilating elementary semigroups discussed in the paragraph following the proof of Theorem 8.2, and in fact it also generalizes the proof we gave for that theorem. This approach was used successfully to construct and analyze dilations in the discrete and continuous one-parameter cases (Muhly and Solel, [102, 104]), in the discrete two-parameter case (Solel, [157]; this case was solved earlier by Bhat for $\mathcal{B} = B(\mathcal{H})$ [25]), and in the "strongly commuting" two-parameter case (Shalit, [141, 142, 144]).

However, it turns out that finding a product system and representation giving back Θ as in (8.2) is not always possible, and one needs a new notion to proceed. A **subproduct system** is a family $\mathcal{E}^{\ominus} = \{\mathcal{E}_s\}_{s\in\mathcal{S}}$ of C*-correspondences such that, roughly, $\mathcal{E}_{s+t} \subseteq \mathcal{E}_s \odot \mathcal{E}_t$ (up to certain identifications that iterate associatively). Following earlier works [14, 102], it was shown that for every CP-semigroup there is a *subproduct system* and a representation, called the **Arveson–Stinespring subproduct system and representation**, satisfying (8.2) (Shalit and Solel, [150]). Subproduct systems have appeared implicitly in the theory in several places, and in

[150] they were finally formally introduced (at the same time, subproduct systems of Hilbert spaces were introduced in Bhat and Mukherjee's paper [27]).

The approach to dilations introduced in [150] consists of two parts: first, embed the Arveson–Stinespring subproduct system associated with a CP-semigroup into a product system, and then dilate the representation to an isometric dilation. This approach was used to find necessary and sufficient conditions for the existence of dilations. In particular, it was used to prove that a Markov semigroup has a (certain kind of) minimal dilation if and only if the Arveson–Stinespring subproduct system can be embedded into a product system. Moreover, the framework was used to show that there exist CP-semigroups over $\mathbb{N}^3$ that have no *minimal* strong dilations, as was suggested from experience with classical dilation theory. Vernik later used these methods to prove an analogue of Opela's theorem (see 3.3) for completely positive maps commuting according to a graph [166].

The reader is referred to [150] for more details. The main drawback of that approach is that it works only for CP-semigroups of normal maps on von Neumann algebras.

The Bhat–Skeide Approach The second main approach to dilations of CP-semigroups is due to Bhat and Skeide [28]. It has several advantages, one of which is that it works for semigroups on unital C*-algebras (rather than von Neumann algebras). The Bhat–Skeide approach is based on a fundamental and useful representation theorem for CP maps called Paschke's *GNS representation* [112], which I will now describe.

For every CP map $\phi : \mathcal{A} \to \mathcal{B}$ between two unital C*-algebras, there exists a C*-correspondence E from $\mathcal{A}$ to $\mathcal{B}$ (see Sect. 6.5) and a vector $\xi \in E$ such that $\phi(a) = \langle \xi, a\xi \rangle$ for all $a \in \mathcal{A}$. The existence of such a representation follows from a construction: one defines E to be the completion of the algebraic tensor product $\mathcal{A} \otimes \mathcal{B}$ with respect to the $\mathcal{B}$-valued inner product

$$\langle a \otimes b, a' \otimes b' \rangle = b^* \phi(a^* a') b',$$

equipped with the natural left and right actions. Letting $\xi = 1_{\mathcal{A}} \otimes 1_{\mathcal{B}}$, it is immediate that $\phi(a) = \langle \xi, a\xi \rangle$ for all $a \in \mathcal{A}$. Moreover, ξ is **cyclic**, in the sense that it generates E as a C*-correspondence. The pair (E, ξ) is referred to as the **GNS representation** of ϕ. The GNS representation is unique in the sense that whenever F is a C*-correspondence from $\mathcal{A}$ to $\mathcal{B}$ and $\eta \in F$ is a cyclic vector such that $\phi(a) = \langle \eta, a\eta \rangle$ for all $a \in \mathcal{A}$, then there is an isomorphism of C*-correspondences from E onto F that maps ξ to η.

In the Bhat–Skeide approach to dilations, the idea is to find a product system $F^{\odot} = \{F_s\}_{s \in \mathcal{S}}$ of $\mathcal{B}$-correspondences and a unit, i.e., a family $\xi^{\odot} = \{\xi_s \in F_s\}_{s \in \mathcal{S}}$ satisfying $\xi_{s+t} = \xi_s \odot \xi_t$, such that Θ is recovered as

$$\Theta_s(b) = \langle \xi_s, b\xi_s \rangle \tag{8.3}$$

for all $s \in \mathcal{S}$ and all $b \in \mathcal{B}$. If Θ is a Markov semigroup, the dilation is obtained via a direct limit construction. For non Markov semigroups, a dilation can be obtained via a unitalization procedure. In [28], dilations were constructed this way in the continuous and discrete one-parameter cases. This strategy bypasses product system representations, but, interestingly, it can also be used to prove the existence of an isometric dilation for any completely contractive covariant representation of a one-parameter product system [155].

Again, it turns out that constructing such a product system is not always possible. However, if one lets $(\mathcal{F}_s, \xi_s)$ be the GNS representation of the CP map Θ_s, then it is not hard to see that $\mathcal{F}^{\widehat{\Theta}} = \{\mathcal{F}_s\}_{s\in\mathcal{S}}$ is a subproduct system (called the **GNS subproduct system**) and $\xi^{\odot} = \{\xi_s\}$ is a unit.

The above observation was used by Shalit and Skeide to study the existence of dilations of CP-semigroups in a very general setting [149]. If one can embed the GNS subproduct system into a product system, then one has (8.3), and can invoke the Bhat–Skeide approach to obtain a dilation. The paper [149] develops this framework to give a unified treatment of dilation theory of CP-semigroups over a large class of monoids $\mathcal{S}$, including noncommutative ones. One of the main results in [149], is the following, which generalizes a result obtained earlier in [150].

Theorem 8.4 *A Markov semigroup over an Ore monoid admits a full strict dilation if and only if its GNS subproduct system embeds into a product system.*

This theorem essentially enables to recover almost all the other known dilation theorems and counter examples. It is used in [149] to show that every Markov semigroup over $\mathcal{S} = \mathbb{N}^2$ on a von Neumann algebra has a unital dilation, and also that for certain multi-parameter semigroups (the so called *quantized convolution semigroups*) there is always a dilation. The theorem is also used in the converse direction, to construct a large class of examples that have no dilation whatsoever.

In the setting of normal semigroups on von Neumann algebras, the Bhat–Skeide and Muhly–Solel approaches to dilations are connected to each other by a functor called *commutant*; see [149, Appendix A(iv)] for details.

New Phenomena We noted above some similarities between the theory of isometric dilations of contractions, and the dilation theory of CP-semigroups. In particular, in both theories, there always exists a dilation when the semigroup is parameterized by $\mathcal{S} = \mathbb{N}, \mathbb{N}^2, \mathbb{R}_+$, and the results support the possibility that this is true for $\mathcal{S} = \mathbb{R}_+^2$ as well. Moreover, in both settings, there exist semigroups over $\mathcal{S} = \mathbb{N}^3$ for which there is no dilation.

But there are also some surprises. By Corollary 3.8, if a tuple of commuting contractions $T = (T_1, \ldots, T_d)$ satisfies $\sum_{i=1}^d \|T_i\|^2 \leq 1$, then T has a regular unitary dilation. Therefore, one might think that if a commuting d-tuple of CP maps $\Theta_1, \ldots, \Theta_d$ are such that $\sum_{i=1}^d \|\Theta_i\|$ is sufficiently small, then this tuple has a dilation. This is false, at least in a certain sense (that is, if one requires a strong and *minimal* dilation); see [150, Section 5.3].

Moreover, by Corollary 3.8, a tuple of commuting isometries always has a unitary dilation, and it follows that every tuple of commuting coisometries has an isometric

(in fact, unitary) coextension. In (8.1) coisometries correspond to unital maps. Hence, one might expect that commuting unital CP maps always have dilations. Again, this is false [149, Section 18] (see also [148]). The reason for the failure of these expectations is that not every subproduct system over $\mathbb{N}^d$ (when $d \geq 3$) can be embedded into a product system (in fact, there are subproduct systems that cannot even be embedded in a *superproduct system*). However, the product system of an elementary CP-semigroup is a trivial product system, so the obstruction to embeddability does not arise in that case.

8.3 Applications of Dilations of CP-Semigroups

Besides the interesting interpretation of dilations given at the end of Sect. 8.1, dilations have some deep applications in noncommutative dynamics. In this section, we will use the term CP-semigroup to mean a one-parameter semigroup $\Theta = \{\Theta_s\}_{s\in\mathcal{S}}$ (where $\mathcal{S} = \mathbb{N}$ or $\mathcal{S} = \mathbb{R}_+$) of normal CP maps acting on a von Neumann algebra $\mathcal{B}$. In case of continuous time (i.e., $\mathcal{S} = \mathbb{R}_+$), we will also assume that Θ is *point weak-$*$ continuous*, in the sense that $t \mapsto \rho(\Theta_t(b))$ is continuous for every $\rho \in \mathcal{B}_*$. We will use the same convention for E- or E_0-semigroups.

The Noncommutative Poisson Boundary Let Θ be a normal UCP map on a von Neumann algebra $\mathcal{B}$. Then one can show that the fixed point set $\{b \in \mathcal{B} : \Theta(b) = b\}$ is an operator system, and moreover that it is the image of a completely positive projection $E : \mathcal{B} \to \mathcal{B}$. Hence, the *Choi-Effros product* $x \circ y = E(xy)$ turns the fixed point set into a von Neumann algebra $H^\infty(\mathcal{B}, \Theta)$, called the **noncommutative Poisson boundary** of Θ. The projection E and the concrete structure on $H^\infty(\mathcal{B}, \Theta)$ are hard to get a grip with.

Arveson observed that if $(\mathcal{A}, \alpha, p)$ is the minimal dilation of Θ, and if $\mathcal{A}^\alpha$ is the fixed point algebra of α, then the compression $a \mapsto pap$ is a unital, completely positive order isomorphism between $\mathcal{A}^\alpha$ and $H^\infty(\mathcal{B}, \Theta)$. Hence $\mathcal{A}^\alpha$ is a concrete realization of the noncommutative Poisson boundary. See the survey [81] for details (it has been observed that this result holds true also for dilations of abelian CP-semigroups [129]).

Continuity of CP-Semigroups Recall that one-parameter CP-semigroups on a von Neumann algebra $\mathcal{B} \subseteq B(\mathcal{H})$ are assumed to be point weak-$*$ continuous. Since CP-semigroups are bounded, this condition is equivalent to *point weak-operator* continuity, i.e., that $t \mapsto \langle \Theta_t(b)g, h\rangle$ is continuous for all $b \in \mathcal{B}$ and all $g, h \in \mathcal{H}$. Another natural kind of continuity to consider is *point strong-operator continuity*, which means that $t \mapsto \Theta_t(b)h$ is continuous in norm for all $b \in \mathcal{B}$ and $h \in \mathcal{H}$. For brevity, below we shall say a semigroup is **weakly continuous** if it is point weak-operator continuous, and **strongly continuous** if it is point strong-operator continuous.

Strong continuity is in some ways easier to work with and hence it is desirable, but it is natural to use the weak-$*$ topology, because it is independent of the representation of the von Neumann algebra. Happily, it turns out that weak (and hence point weak-$*$) continuity implies strong continuity.

One possible approach to prove the above statement is via dilation theory. First, one notices that the implication is easy for E-semigroups. Indeed, if α is an E-semigroup on $\mathcal{A} \subseteq B(\mathcal{K})$, then

$$\|\alpha_t(a)k - \alpha_s(a)k\|^2 = \langle \alpha_t(a^*a)k, k\rangle + \langle \alpha_s(a^*a)k, k\rangle - 2\,\mathrm{Re}\langle \alpha_t(a)k, \alpha_s(a)k\rangle.$$

Assuming that s tends to t, the expression on the right hand side tends to zero if α is weakly continuous. Now, if Θ is a weakly continuous CP-semigroup, then its dilation α given by Theorem 8.2 is also weakly continuous. By the above argument, α is strongly continuous, and this continuity is obviously inherited by $\Theta_t(\cdot) = p\alpha_t(\cdot)p$. Hence, by dilation theory, weak continuity implies strong continuity.

The above argument is a half cheat, because, for a long time, the known proofs that started from assuming weak continuity and ended with a weakly continuous dilation, actually assumed implicitly, somewhere along the way, that CP-semigroups are strongly continuous [14, 28, 102]. This gap was pointed out and fixed by Markiewicz and Shalit [98], who proved directly that a weakly continuous CP-semigroup is strongly continuous. Later, Skeide proved that the minimal dilation of a weakly continuous semigroup of CP maps is strongly continuous, independently of [98], thereby recovering the result "weakly continuous $\Rightarrow$ strongly continuous" with a proof that truly goes through the construction of a dilation; see [156, Appendix A.2].

Existence of E_0-Semigroups As we have seen above, dilations can be used to study CP-semigroups. We will now see an example, where dilations are used in the theory of E_0-semigroups.

The fundamental classification theory E_0-semigroups on $\mathcal{B} = B(\mathcal{H})$ was developed Arveson, Powers, and others about two decades ago; see the monograph [14] for the theory, and in particular for the results stated below (the classification theory of E_0-semigroups on arbitrary C* or von Neumann algebras is due to Skeide; see [156]). For such E_0-semigroups there exists a crude grouping into *type I, type II, and type III* semigroups. However, it is not at all obvious that there exist any E_0-semigroups of every type. Given a semigroup of isometries on a Hilbert space H, one may use second quantization to construct E_0-semigroups on the symmetric and anti-symmetric Fock spaces over H, called the CCR and CAR flows, respectively. CAR and CCR flows are classified in term of their *index*. These E_0-semigroups are of type I, and, conversely, every type I E_0-semigroup is cocycle conjugate to a CCR flow, which is, in turn, conjugate to a CAR flow.

It is much more difficult to construct an E_0-semigroup that is not type I. How does one construct a non-trivial E_0-semigroup? Theorem 8.2 provides a possible way: construct a Markov semigroup, and then take its minimal dilation. This procedure

has been applied successfully to provide examples of non type I E_0-semigroups, even with prescribed *index*; see [14, 81].

Part 3: Recent Results in the Dilation Theory of Noncommuting Operators

9 Matrix Convexity and Dilations

In recent years, dilation theory has found a new role in operator theory, through the framework of matrix convexity. In this section I will quickly introduce matrix convex sets in general, special examples, minimal and maximal matrix convex over a given convex set, and the connection to dilation theory. Then I will survey the connection to the UCP interpolation problem, some dilation results, and finally an application to spectrahedral inclusion problems.

9.1 Matrix Convex Sets

Fix $d \in \mathbb{N}$. The "noncommutative universe" $\mathbb{M}^d$ is the set of all d-tuples of $n \times n$ matrices, of all sizes n, that is,

$$\mathbb{M}^d = \bigsqcup_{n=1}^{\infty} M_n^d.$$

Sometimes it is useful to restrict attention to the subset $\mathbb{M}_{\text{sa}}^d$, which consists of all tuples of selfadjoint matrices. We will refer to a subset $\mathcal{S} \subseteq \mathbb{M}^d$ as a **noncommutative (nc) set**, and we will denote by $\mathcal{S}_n$ or $\mathcal{S}(n)$ the n**th level of** $\mathcal{S}$, by which we mean $\mathcal{S}_n = \mathcal{S}(n) := \mathcal{S} \cap M_n^d$. Let us endow tuples with the row norm $\|A\| := \|\sum_i A_i A_i^*\|$; this induces a metric on $B(\mathcal{H})^d$ for every d, and in particular on M_n^d for every n. We will say that a nc set $\mathcal{S}$ is **closed** if $\mathcal{S}_n$ is closed in M_n^d for all n. We will say that $\mathcal{S}$ is **bounded** if there exists some $C > 0$ such that $\|X\| \leq C$ for all $X \in \mathcal{S}$.

For a tuple $X = (X_1, \ldots, X_n) \in M_n^d$ and a linear map $\phi : M_n \to M_k$, we write $\phi(X) = (\phi(X_1), \ldots, \phi(X_d)) \in M_k^d$. In particular, if A and B are $n \times k$ matrices, then we write $A^*XB = (A^*X_1B, \ldots, A^*X_dB)$. Another operation that we can preform on tuples is the **direct sum**, that is, if $X \in M_m^d$ and $Y \in M_n^d$, then we let $X \oplus Y = (X_1 \oplus Y_1, \ldots, X_d \oplus Y_d) \in M_{m+n}^d$.

A **matrix convex set** $\mathcal{S}$ is a nc set $\mathcal{S} = \sqcup_{n=1}^{\infty} \mathcal{S}_n$ which is invariant under direct sums and the application of UCP maps:

$$X \in \mathcal{S}_m, Y \in \mathcal{S}_n \Longrightarrow X \oplus Y \in \mathcal{S}_{m+n}$$

and

$$X \in \mathcal{S}_n \text{ and } \phi \in \mathrm{UCP}(M_n, M_k) \Longrightarrow \phi(X) \in \mathcal{S}_k.$$

It is not hard to check that a nc set $\mathcal{S} \subseteq \mathbb{M}^d$ is matrix convex if and only if it is closed under **matrix convex combinations** in the following sense: whenever $X^{(j)} \in \mathcal{S}_{n_j}$ and $V_j \in M_{n_j,n}$ for $j = 1, \ldots, k$ are such that $\sum_{j=1}^{k} V_j^* V_j = I_n$, then $\sum V_j^* X^{(j)} V_j \in \mathcal{S}_n$.

Remark 9.1 The above notion of matrix convexity is due to Effros and Winkler [57]. Other variants appeared before and after. A very general take on matrix convexity that I will not discuss here has recently been initiated by Davidson and Kennedy [43]. I will follow a more pedestrian point of view, in the spirit of [47] (note: the arxiv version [47] is a corrected version of the published version [46]. The latter contains several incorrect statements in Section 6, which result from a missing hypothesis; the problem and its solution are explained in [47]). We refer to the first four chapters of [47] for explanations and/or references to the some of the facts that will be mentioned below without proof. The papers [61, 113, 115] make a connection between the geometry of matrix convex sets, in particular various kinds of extreme points, and dilation theory. For a comprehensive and up-to-date account of matrix convex sets the reader can consult [90].

Example 9.2 Let $A \in B(\mathcal{H})^d$. The **matrix range** of A is the nc set $\mathcal{W}(A) = \sqcup_{n=1}^{\infty} \mathcal{W}_n(A)$ given by

$$\mathcal{W}_n(A) = \{\phi(A) : \phi : B(\mathcal{H}) \to M_n \text{ is UCP}\}.$$

The matrix range is a closed and bounded matrix convex set. Conversely, every closed and bounded matrix convex set is the matrix range of some operator tuple.

If $d = 1$, then the first level $\mathcal{W}_1(A)$ of the matrix range of an operator A coincides with the closure of the **numerical range** $W(A)$

$$W(A) = \{\langle Ah, h\rangle : \|h\| = 1\}.$$

We note, however, that for $d \geq 2$, the first level $\mathcal{W}_1(A)$ does not, in general, coincide with the closure of what is sometimes referred to as the **joint numerical range** of a tuple [96].

Matrix ranges of single operators were introduced by Arveson [11], and have been picked up again rather recently. The matrix range of an operator tuple A is a complete invariant of the operator system generated by A, and—as we shall

see below—it is useful when considering interpolation problems for UCP maps. Moreover, in the case of a *fully compressed* tuple A of compact operators or normal operators, the matrix range determines A up to unitary equivalence [115]. The importance of matrix ranges has led to the investigation of random matrix ranges, see [70].

Example 9.3 Let $A \in B(\mathcal{H})^d$. The **free spectrahedron** determined by A is the nc set $\mathcal{D}_A = \sqcup_{n=1}^{\infty} \mathcal{D}_A(n)$ given by

$$\mathcal{D}_A(n) = \left\{ X \in M_n^d : \mathrm{Re} \sum_{j=1}^{d} X_j \otimes A_j \leq I \right\}.$$

A free spectrahedron is always a closed matrix convex set, that contains the origin in its interior. Conversely, every closed matrix convex set with 0 in its interior is a free spectrahedron. In some contexts it is more natural to work with just selfadjoint matrices. For $A \in B(\mathcal{H})^d_{\mathrm{sa}}$ one defines

$$\mathcal{D}_A^{\mathrm{sa}} = \left\{ X \in \mathbb{M}^d_{\mathrm{sa}} : \sum_{j=1}^{d} X_j \otimes A_j \leq I \right\}.$$

The first level $\mathcal{D}_A(1)$ is called a **spectrahedron**. Most authors use the word *spectrahedron* to describe only sets of the form $\mathcal{D}_A(1)$ where A is a tuple of *matrices*; and likewise for the term *free spectrahedron*. This distinction is important for applications of the theory, since spectrahedra determined by tuples of matrices form a class of reasonably tractable convex sets that arise in applications, and not every convex set with 0 in its interior can be represented as $\mathcal{D}_A(1)$ for a tuple A acting on a finite dimensional space.

For a matrix convex set $\mathcal{S} \subseteq \mathbb{M}^d$ we define its **polar dual** to be

$$\mathcal{S}^\circ = \left\{ X \in \mathbb{M}^d : \mathrm{Re}\left(\sum X_j \otimes A_j \right) \leq I \ \text{ for all } A \in \mathcal{S} \right\}.$$

If $\mathcal{S} \subseteq \mathbb{M}^d_{\mathrm{sa}}$, then it is more convenient to use the following variant

$$\mathcal{S}^\bullet = \left\{ X \in \mathbb{M}^d_{\mathrm{sa}} : \sum X_j \otimes A_j \leq I \ \text{ for all } A \in \mathcal{S} \right\}.$$

By the Effros–Winkler Hahn–Banach type separation theorem [57], $\mathcal{S}^{\circ\circ} = \mathcal{S}$ whenever $\mathcal{S}$ is a matrix convex set containing 0 (if $0 \notin \mathcal{S}$, then $\mathcal{S}^{\circ\circ}$ is equal to the *matrix convex hull* of $\mathcal{S}$ and 0). It is not hard to see that $\mathcal{D}_A = \mathcal{W}(A)^\circ$, and that when $0 \in \mathcal{W}(A)$, we also have $\mathcal{W}(A) = \mathcal{D}_A^\circ$.

Example 9.4 Another natural and important way in which matrix convex sets arise, is as positivity cones in operator systems. In [65] it was observed that a finite dimensional abstract operator system $\mathcal{M}$ (see [118, Chapter 13]) generated by d

linearly independent elements $A_1, \ldots, A_d \in \mathcal{M}$, corresponds to a matrix convex set $\mathcal{C} \subseteq \mathbb{M}_{\text{sa}}^d$ where every $\mathcal{C}_n$ is the cone in $(M_n^d)_{\text{sa}}$ consisting of the matrix tuples $X = (X_1, \ldots, X_d)$ such that $\sum X_j \otimes A_j$ is positive in $M_n(\mathcal{M})$. Such matrix convex sets can be described by a slight modification of the notion of free spectrahedron:

$$\mathcal{C} = \left\{ X \in \mathbb{M}_{\text{sa}}^d : \sum X_j \otimes A_j \geq 0 \right\}.$$

9.2 The UCP Interpolation Problem

Suppose we are given two d-tuples of operators $A = (A_1, \ldots, A_d) \in B(\mathcal{H})^d$ and $B = (B_1, \ldots, B_d) \in B(\mathcal{K})^d$. A very natural question to ask is whether there exists a completely positive map $\phi : B(\mathcal{H}) \to B(\mathcal{K})$ such that $\phi(A_i) = B_i$. This is the *CP interpolation problem*. In the realm of operator algebras it is sometimes more useful to ask about the existence of a UCP map that interpolates between the operators, and in quantum information theory it makes sense to ask whether there exists a completely positive trace preserving (CPTP) interpolating map. A model result is the following.

Theorem 9.5 *Let $A \in \mathcal{B}(H)^d$ and $B \in \mathcal{B}(K)^d$ be d-tuples of operators.*

1. *There exists a UCP map $\phi : B(\mathcal{H}) \to B(\mathcal{K})$ such that $\phi(A_i) = B_i$ for all $i = 1, \ldots, d$ if and only if $\mathcal{W}(B) \subseteq \mathcal{W}(A)$.*
2. *There exists a unital completely isometric map $\phi : B(\mathcal{H}) \to B(\mathcal{K})$ such that $\phi(A_i) = B_i$ for all $i = 1, \ldots, d$ if and only if $\mathcal{W}(B) = \mathcal{W}(A)$.*

This result was obtained by Davidson, Dor-On, Shalit and Solel in [47, Theorem 5.1]. An earlier result was obtained by Helton, Klep and McCullough in the case where $\mathcal{H}$ and $\mathcal{K}$ are finite dimensional, and the condition $\mathcal{W}(B) \subseteq \mathcal{W}(A)$ is replaced by the *dual* condition $\mathcal{D}_A \subseteq \mathcal{D}_B$, under the blanket assumption that $\mathcal{D}_A$ is bounded [74] (see a somewhat different approach in [7]). Later, Zalar showed that the condition $\mathcal{D}_A \subseteq \mathcal{D}_B$ is equivalent to the existence of an interpolating UCP map without the assumption that $\mathcal{D}_A$ is bounded, and also in the case of operators on an infinite dimensional space [171]. Variants of the above theorem were of interest to mathematical physicists for some time, see the references in the above papers.

From Theorem 9.5 one can deduce also necessary and sufficient conditions for the existence of contractive CP (CCP) or completely contractive (CC) maps sending one family of operators to another, as well as approximate versions (see [47, Section 5]). The theorem also leads to more effective conditions under additional assumptions, for example when dealing with normal tuples (recall, that $A = (A_1, \ldots, A_d)$ is said to be normal if A_i is normal and $A_i A_j = A_j A_i$ for all i, j).

Corollary 9.6 *Let $A \in \mathcal{B}(H)^d$ and $B \in \mathcal{B}(K)^d$ be two normal d-tuples of operators. Then there exists a UCP map $\phi : B(\mathcal{H}) \to B(\mathcal{K})$ such that*

$$\phi(A_i) = B_i \ , \quad \text{for all } i = 1, \ldots, d,$$

if and only if

$$\sigma(B) \subseteq \operatorname{conv} \sigma(A).$$

This result was first obtained by Li and Poon [97], in the special case where A and B each consist of commuting selfadjoint matrices. It was later recovered in [47], in the above generality, as a consequence of Theorem 9.6 together with the fact that for a normal tuple N, the matrix range $\mathcal{W}(N)$ is the *minimal matrix convex set that contains the joint spectrum* $\sigma(N)$ *in its first level* (see [47, Corollary 4.4]). The next section is dedicated to explaining what are the minimal and maximal matrix convex sets over a convex set, and how these notions are related to dilation theory.

9.3 Minimal and Maximal Matrix Convex Sets

Every level $\mathcal{S}_n$ of a matrix convex set $\mathcal{S}$ is a convex subset of M_n^d. In particular, the first level $\mathcal{S}_1$ is a convex subset of $\mathbb{C}^d$. Conversely, given a convex set $K \subseteq \mathbb{C}^d$ (or $K \subseteq \mathbb{R}^d$), we may ask whether there exists a matrix convex set $\mathcal{S} \subseteq \mathbb{M}^d$ (or $\mathcal{S} \subseteq \mathbb{M}_{\text{sa}}^d$) such that $\mathcal{S}_1 = K$. The next question to ask is, to what extent does the first level $\mathcal{S}_1 = K$ determine the matrix convex set $\mathcal{S}$?

In order to approach the above questions, and also as part of a general effort to understand inclusions between matrix convex sets (motivated by results as Theorem 9.5), notions of minimal and maximal matrix convex sets have been introduced by various authors [46, 65, 75]. These are very closely related (via Example 9.4) to the notion of minimal and maximal operator systems that was introduced earlier [120].

For brevity, we shall work in the selfadjoint setting. Let $K \subseteq \mathbb{R}^d$ be a convex set. By the Hahn–Banach theorem, K can be expressed as the intersection of a family of half spaces:

$$K = \{x \in \mathbb{R}^d : f_i(x) \leq c_i \ \text{ for all } i \in \mathcal{I}\}$$

where $\{f_i\}_{i\in\mathcal{I}}$ is a family of linear functionals and $\{c_i\}_{i\in\mathcal{I}}$ is a family of scalars. Writing $f_i(x) = \sum_j a_j^i x_j$, we define

$$\mathcal{W}_n^{\max}(K) = \{X \in (M_n^d)_{\text{sa}} : \sum_j a_j^i X_j \leq c_i I_n \ \text{ for all } i \in \mathcal{I}\}$$

and $\mathcal{W}^{\max}(K) = \sqcup_{n=1}^{\infty} \mathcal{W}_n^{\max}(K)$. In other words, $\mathcal{W}^{\max}(K)$ is the nc set determined by the linear inequalities that determine K. It is clear that $\mathcal{W}^{\max}(K)$ is matrix convex, and a moment's thought reveals that it contains every matrix convex set that has K as its first level.

That settles the question, of whether or not there exists a matrix convex set with first level equal to K. It follows, that there has to exist a minimal matrix convex set that has K as its first level—simply intersect over all such matrix convex sets. There is a useful description of this minimal matrix convex set. We define

$$\mathcal{W}^{\min}(K) = \left\{ X \in \mathbb{M}_{\mathrm{sa}}^d : \exists \text{ normal } T \text{ with } \sigma(T) \subseteq K \text{ s.t. } X \prec T \right\}. \tag{9.1}$$

Recall that $X \prec T$ means that T is a dilation of X.

$\mathcal{W}^{\min}(K)$ is clearly invariant under direct sums. To see that it is invariant also under the application of UCP maps, one may use Stinespring's theorem as follows. If $X \prec T$, T is normal, and ϕ is UCP, then the map $T \mapsto X \mapsto \phi(X)$ is UCP. By Stinespring's theorem there is a $*$-representation π such that $\phi(X) \prec \pi(T)$, and $\pi(T)$ is a normal tuple with $\sigma(\pi(T)) \subseteq \sigma(T)$ (alternatively, one may use the dilation guaranteed by Theorem 8.2).

We see that the set defined in (9.1) is matrix convex. On the other hand, any matrix convex set containing K in the first level must contain all unitary conjugates of tuples formed from direct sums of points in K, as well as their compressions, therefore the minimal matrix convex set over K contains all X that have a normal dilation T *acting on a finite dimensional space* such that $\sigma(T) \subseteq K$. But for $X \in \mathbb{M}_{\mathrm{sa}}^d$, the existence of a normal dilation $X \prec T$ with $\sigma(T) \subseteq K$ implies the existence of a normal dilation *acting on a finite dimensional space* (see [46, Theorem 7.1]), thus the nc set $\mathcal{W}^{\min}(K)$ that we defined above is indeed the minimal matrix convex set over K.

Example 9.7 Let $\overline{\mathbb{D}}$ be the closed unit disc in $\mathbb{C}$. Let us compute $\mathcal{W}^{\min}(\overline{\mathbb{D}})$ and $\mathcal{W}^{\max}(\overline{\mathbb{D}})$. We can consider $\overline{\mathbb{D}}$ as a subset of $\mathbb{R}^2$, and pass to the selfadjoint setting (and back) by identifying $T = \operatorname{Re} T + i \operatorname{Im} T \in \mathbb{M}^1$ with the selfadjoint tuple $(\operatorname{Re} T, \operatorname{Im} R) \in \mathbb{M}_{\mathrm{sa}}^2$. The minimal matrix convex set is just

$$\mathcal{W}^{\min}(\overline{\mathbb{D}}) = \{X \in \mathbb{M}^1 : \|X\| \leq 1\},$$

because by Theorem 2.1, every contraction has a unitary dilation. Since the set of real linear inequalities determining the disc is

$$\overline{\mathbb{D}} = \{z \in \mathbb{C} : \operatorname{Re}\left(e^{i\theta} z\right) \leq 1 \quad \text{for all } \theta \in \mathbb{R}\},$$

it follows that

$$\mathcal{W}^{\max}(\overline{\mathbb{D}}) = \{X : \operatorname{Re}\left(e^{i\theta} X\right) \leq I \quad \text{for all } \theta \in \mathbb{R}\},$$

which equals the set of all matrices with numerical range contained in the disc.

Given a convex set $K \subseteq \mathbb{C}^d$, Passer, Shalit and Solel introduced a constant $\theta(K)$ that quantifies the difference between the minimal and maximal matrix convex sets over K [116, Section 3]. For two convex sets K, L, we define

$$\theta(K, L) = \inf\{C : \mathcal{W}^{\max}(K) \subseteq C\mathcal{W}^{\min}(L)\},$$

and $\theta(K) = \theta(K, K)$. Note that $C\mathcal{W}^{\min}(L) = \mathcal{W}^{\min}(CL)$.

Remark 9.8 In the theory of operator spaces, there are the notions of minimal and maximal operator spaces over a normed space V, and there is a constant $\alpha(V)$ that quantifies the difference between the minimal and maximal operator space structures [117] (see also [118, Chapter 14] and [122, Chapter 3]). These notions are analogous to the above notions of minimal and maximal matrix convex sets, but one should not confuse them.

By the characterization of the minimal and maximal matrix convex sets, the inclusion $\mathcal{W}^{\max}(K) \subseteq \mathcal{W}^{\min}(L)$ is a very general kind of dilation result: it means that *every d-tuple X satisfying the linear inequalities defining K, has a normal dilation $X \prec N$ such that $\sigma(N) \subseteq L$*. Let us now review a few results obtained regarding this dilation problem.

Theorem 9.9 (Theorem 6.9, [116]) *For $p \in [1, \infty]$, let $\overline{\mathbb{B}}_{p,d}$ denote the unit ball in $\mathbb{R}^d$ with respect to the ℓ^p norm, and let $\overline{\mathbb{B}}_{p,d}(\mathbb{C})$ denote the unit ball in $\mathbb{C}^d$ with respect to the ℓ^p norm. Then*

$$\theta(\overline{\mathbb{B}}_{p,d}) = d^{1-|1/2-1/p|}$$

and

$$\theta(\overline{\mathbb{B}}_{p,d}(\mathbb{C})) = 2d^{1-|1/2-1/p|}.$$

See [116] for many other (sharp) inclusions $\mathcal{W}^{\max}(K) \subseteq \mathcal{W}^{\min}(L)$. Interestingly, the fact that $\theta(\overline{\mathbb{B}}_{1,d}) = \sqrt{d}$ has implications in quantum information theory—it allows to find a quantitative measure of how much *noise* one needs to add to a d-tuple of quantum effects to guarantee that they become *jointly measurable*; see [30].

The case $d = 2$ in the above theorem was first obtained in [76, Section 14] and [47, Section 7] using other methods. It also follows from the following result.

Theorem 9.10 (Theorem 5.8, [65]) *Let $K \subseteq \mathbb{R}^d$ be a symmetric convex set, i.e. $K = -K$. Then*

$$\mathcal{W}^{\max}(K) \subseteq d\mathcal{W}^{\min}(K).$$

The above result was originally proved by Fritz, Netzer and Thom [65] for cones *with a symmetric base*; to pass between the language of convex bodies and that of cones, one may use the gadget developed in [116, Section 7]. In [116, Theorem 4.5]

it was observed that Theorem 9.10 is also a consequence of the methods of [47, Section 7] together with some classical results in convex geometry.

Already in [57, Lemma 3.1] it was observed that there is only one matrix convex $\mathcal{S}$ with $\mathcal{S}_1 = [a, b] \subset \mathbb{R}$, namely the **matrix interval** given by $\mathcal{S}_n = \{X \in (M_n)_{\text{sa}} : aI_n \leq X \leq bI_n\}$. Said differently, $\mathcal{W}^{\max}([a, b]) = \mathcal{W}^{\min}([a, b])$. It is natural to ask whether there exists any other convex body (i.e., a compact convex set) K with the property that $\mathcal{W}^{\min}(K) = \mathcal{W}^{\max}(K)$.

Theorem 9.11 *Let $K \subseteq \mathbb{R}^d$ be a convex body. Then $\mathcal{W}^{\max}(K) = \mathcal{W}^{\min}(K)$ if and only if K is a simplex, that is, if K is the convex hull of a set of affinely independent points. In fact, $\mathcal{W}_2^{\max}(K) = \mathcal{W}_2^{\min}(K)$ already implies that K is a simplex.*

The result that the equality $\mathcal{W}^{\max}(K) = \mathcal{W}^{\min}(K)$ is equivalent to K being a simplex was first obtained by Fritz, Netzer and Thom [65, Corollary 5.3] for polyhedral cones. In [116, Theorem 4.1] it was proved for general convex bodies, and it was also shown that one does not need to check equality $\mathcal{W}_n^{\max}(K) = \mathcal{W}_n^{\min}(K)$ for all n in order to deduce that K is a simplex—it suffices to check this for some $n \geq 2^{d-1}$. For *simplex pointed* convex bodies, it was shown that $\mathcal{W}_2^{\max}(K) = \mathcal{W}_2^{\min}(K)$ already implies that K is a simplex [116, Theorem 8.8]. Huber and Netzer later obtained this for all polyhedral cones [80], and finally Aubrun, Lami, Palazuelos and Plavala proved the result for all cones [16, Corollary 2].

Remark 9.12 The minimal matrix convex set over a "commutative" convex set $K \subseteq \mathbb{R}^d$ can be considered as the *matrix convex hull* of K. There are some variations on this theme. Helton, Klep and McCullough studied the matrix convex hull of *free semialgebraic sets* [75]. Instead of $\mathcal{W}^{\min}(K)$ and $\mathcal{W}^{\max}(K)$, which are the minimal and maximal matrix convex sets with prescribed first level, one can also discuss the minimal and maximal matrix convex sets with a prescribed kth level (see [90], or [169, 170] for the version of this notion in the framework of operator systems). In the recent paper [114], Passer and Paulsen define, given a matrix convex set $\mathcal{S}$, the minimal and maximal matrix convex sets $\mathcal{W}^{\min\text{-}k}(\mathcal{S})$ and $\mathcal{W}^{\max\text{-}k}(\mathcal{S})$ such that $\mathcal{W}_k^{\min\text{-}k}(\mathcal{S}) = \mathcal{W}_k^{\max\text{-}k}(\mathcal{S}) = \mathcal{S}_k$, and they utilize quantitative measures of discrepancy between $\mathcal{W}^{\min\text{-}k}(\mathcal{S})$, $\mathcal{W}^{\max\text{-}k}(\mathcal{S})$ and $\mathcal{S}$ to glean information on the operator system corresponding to $\mathcal{S}$; unfortunately, these results are beyond the scope of this survey. The paper [114] also ties together some of the earlier work in this direction, so it is a good place to start if one is interested in this problem.

9.4 Further Dilation Results

There are many other interesting dilation results in [46, 50, 65, 76, 113, 116]. In this section I will review a few more.

Problem 9.13 Fix $d \in \mathbb{N}$. What is the smallest constant C_d such that for every d-tuple of contractions A, there exists a d-tuple of commuting normal operators B, such that $A \prec B$ holds with $\|B_i\| \leq C_d$ for all i?

First, we note that the sharp dilation constant $\theta(\overline{\mathbb{B}}_{\infty,d}) = \sqrt{d}$ obtained in Theorem 9.10 implies the following result, which is a solution to Problem 9.13 in the selfadjoint setting.

Theorem 9.14 (Theorem 6.7, [116]) *For every d-tuple $A = (A_1, \ldots, A_d)$ of selfadjoint contractions, there exists a d-tuple of commuting selfadjoints $N = (N_1, \ldots, N_d)$ with $\|N_i\| \leq \sqrt{d}$ for $i = 1, \ldots, d$, such that $A \prec N$. Moreover, $\sqrt{d}$ is the optimal constant for selfadjoints.*

It is interesting to note that one of the proofs of the above theorem goes through a concrete construction of the dilation. The nonselfadjoint version of Problem 9.13 is more difficult, and it does not correspond to an inclusion problem of some $\mathcal{W}^{\max}$ in some $\mathcal{W}^{\min}$. The best general result in the nonselfadjoint case is the following theorem obtained by Passer.

Theorem 9.15 (Theorem 4.4, [113]) *For every d-tuple $A = (A_1, \ldots, A_d)$ of contractions, there exists a d-tuple of commuting normal operators $N = (N_1, \ldots, N_d)$ with $\|N_i\| \leq \sqrt{2d}$ for $i = 1, \ldots, d$, such that $A \prec N$.*

Thus

$$\sqrt{d} \leq C_d \leq \sqrt{2d}.$$

In the next section we will improve the lower bound in the case $d = 2$.

Helton, Klep, McCullough and Schweighofer obtained a remarkable result, which is analogous to Theorem 9.14, but in which the dilation constant is *independent of the number of operators d* [76]. Following Ben-Tal and Nemirovski [21], Helton et al. defined a constant $\vartheta(n)$ as follows:

$$\frac{1}{\vartheta(n)} = \min\left\{ \int_{\partial \mathbb{B}_n} \left| \sum_{i=1}^n a_i x_i^2 \right| d\mu(x) \ : \ \sum_{i=1}^n |a_i| = 1 \right\}$$

where μ is the uniform probability measure on the unit sphere $\partial \mathbb{B}_n \subset \mathbb{R}^n$.

Theorem 9.16 (Theorem 1.1, [76]) *Fix n and a real n-dimensional Hilbert space $\mathcal{H}$. Let $\mathcal{F} \subseteq B(\mathcal{H})_{sa}$ be a family of selfadjoint contractions. Then there exists a real Hilbert space $\mathcal{K}$, an isometry $V : \mathcal{H} \to \mathcal{K}$, and a commuting family $\mathcal{C}$ in the unit ball of $B(\mathcal{K})_{sa}$ such that for every contraction $A \in \mathcal{F}$, there exists $N \in \mathcal{C}$ such that*

$$\frac{1}{\vartheta(n)} A = V^* N V.$$

Moreover, $\vartheta(n)$ is the smallest constant such that the above holds for all finite sets of contractive selfadjoints $\mathcal{F} \subseteq B(\mathcal{H})_{sa}$.

Note the difference from Theorem 9.14: the dimension of matrices is fixed at $n \times n$, but the number of matrices being simultaneously dilated is **not** fixed. In other words, the constant $\vartheta(n)$ depends only on the size of the matrices being dilated (in fact, it is shown that n can be replaced with the maximal rank of the matrices being dilated). It is also shown that

$$\vartheta(n) \sim \frac{\sqrt{\pi n}}{2}.$$

In the next subsection I will explain the motivation for obtaining this result.

9.5 An Application: Matricial Relaxation of Spectrahedral Inclusion Problems

Any dilation result, such as Theorem 9.14 or 9.16, leads to a von Neumann type inequality. For example, if A is a d-tuple of selfadjoint contractions, then by Theorem 9.14, for every matrix valued polynomial p of degree at most one, we have the following inequality:

$$\|p(A)\| \leq \sup\left\{\|p(z)\| : z \in [-\sqrt{d}, \sqrt{d}]^d\right\}.$$

This result is by no means trivial, but it is the kind of application of dilation theory that we have already seen above several times.

We will now see a deep application of Helton, Klep, McCullough and Schweighofer's theorem (Theorem 9.16) that is of a different nature from the applications that we have seen hitherto, and is the main motivation for the extraordinary paper [76]. The application builds on earlier work of Ben-Tal and Nemirovski [21] in control theory and optimization, related to what is sometimes called *the matrix cube problem*. I will give a brief account; the reader who seeks a deeper understanding should start with the introductions of [21] and [76].

In the analysis of a linear controlled dynamical system (as in [21]), one is led to the problem of deciding whether the cube $[-1, 1]^d$ is contained in the spectrahedron $\mathcal{D}_A^{\mathrm{sa}}(1)$, for a given a d-tuple of selfadjoint $n \times n$ matrices $A_1, \ldots, A_d$; this is called *the matrix cube problem*. More generally, given another d-tuple of selfadjoint matrices $B_1, \ldots, B_d$, it is of practical interest to solve the spectrahedral inclusion problem, that is, to be able to decide whether

$$\mathcal{D}_B^{\mathrm{sa}}(1) \subseteq \mathcal{D}_A^{\mathrm{sa}}(1).$$

Note that the matrix cube problem is a special case of the spectrahedral inclusion problem, since $[-1,1]^d = \mathcal{D}_C^{\mathrm{sa}}(1)$ for the d-tuple of $2d \times 2d$ diagonal matrices $C_1 = \mathrm{diag}(1,-1,0,\ldots,0)$, $C_2 = \mathrm{diag}(0,0,1,-1,0,\ldots,0)$, $\ldots$, $C_d = (0,\ldots,0,1,-1)$. The free spectrahedron $\mathcal{D}_C^{\mathrm{sa}}$ determined by C is nothing but the nc set consisting of all d-tuples of selfadjoint contractions.

The problem of deciding whether one spectrahedron is contained in another is a *hard* problem. In fact, deciding whether or not $[-1,1]^d \subseteq \mathcal{D}_A^{\mathrm{sa}}(1)$ has been shown to be NP hard (note that the naive solution of checking whether all the vertices of the cube are in $\mathcal{D}_A^{\mathrm{sa}}(1)$ requires one to test the positive semidefiniteness of 2^d matrices). However, Ben-Tal and Nemirovski introduced a tractable *relaxation* of this problem [21]. In [74], Helton, Klep and McCullough showed that the relaxation from [21] is equivalent to the *free relaxation* $\mathcal{D}_C^{\mathrm{sa}} \subseteq \mathcal{D}_A^{\mathrm{sa}}$, and the subsequent work in [76] gives a full understanding of this relaxation, including sharp estimates of the error bound.

Let's take a step back. Fix two d-tuples of selfadjoint matrices A and B. We mentioned that the problem of determining whether $\mathcal{D}_B^{\mathrm{sa}}(1) \subseteq \mathcal{D}_A^{\mathrm{sa}}(1)$ is hard. In [74], it was observed that the *free relaxation*, that is, the problem $\mathcal{D}_B^{\mathrm{sa}} \subseteq \mathcal{D}_A^{\mathrm{sa}}$ is tractable. Indeed, as explained after Theorem 9.5, the inclusion $\mathcal{D}_B^{\mathrm{sa}} \subseteq \mathcal{D}_A^{\mathrm{sa}}$ is equivalent to the UCP interpolation problem, that is, to the existence of a UCP map sending B_i to A_i for all $i = 1,\ldots,d$ [74, Theorem 3.5]. Now, the UCP interpolation problem can be shown to be equivalent to the solution of a certain *semidefinite program* [74, Section 4]. In practice, there are numerical software packages that can solve such problems efficiently.

So we see that instead of solving the matrix cube problem $[-1,1]^d \subseteq \mathcal{D}_A^{\mathrm{sa}}(1)$, one can solve the free relaxation $\mathcal{D}_C^{\mathrm{sa}} \subseteq \mathcal{D}_A^{\mathrm{sa}}$. Now, the whole point of the sharp results in [76] is that they give a tight estimate of how well the tractable free relaxation approximates the hard matrix cube problem. To explain this, we need the following lemma.

Lemma 9.17 *Suppose that A is a d-tuple of selfadjoint $n \times n$ matrices. Then,*

$$[-1,1]^d \subseteq \mathcal{D}_A^{\mathrm{sa}}(1) \Rightarrow \mathcal{D}_C^{\mathrm{sa}} \subseteq \vartheta(n)\mathcal{D}_A^{\mathrm{sa}}.$$

Proof Suppose that $[-1,1]^d \subseteq \mathcal{D}_A^{\mathrm{sa}}(1)$. If $X \in \mathcal{D}_C^{\mathrm{sa}}(n)$, then by Theorem 9.16, $X \prec \vartheta(n)N$, where N is a normal tuples and $\sigma(N) \subseteq [-1,1]^d \subseteq \mathcal{D}_A^{\mathrm{sa}}(1)$. So

$$\sum X_j \otimes A_j \prec \vartheta(n) \sum N_j \otimes A_j \leq \vartheta(n) I,$$

where the last inequality follows easily by the spectral theorem and the assumption $[-1,1]^d \subseteq \mathcal{D}_A^{\mathrm{sa}}(1)$. ∎

Finally, we can now understand how to give an approximate solution to the matrix cube problem. Simply, one tests whether $\mathcal{D}_C^{\mathrm{sa}} \subseteq \vartheta(n)\mathcal{D}_A^{\mathrm{sa}}$, which is a tractable problem. If the inclusion holds, then it holds at every level and in particular $[-1,1]^d \subseteq \vartheta(n)\mathcal{D}_A^{\mathrm{sa}}(1)$. If not, then, using the lemma, we conclude that $[-1,1]^d \subsetneq \mathcal{D}_A^{\mathrm{sa}}(1)$. Thus, we are able to determine the containment of $[-1,1]^d$ in $\mathcal{D}_A^{\mathrm{sa}}(1)$, up to

a multiplicative error of $\vartheta(n)$, which is known to high precision, and independent of d.

10 Dilation of q-Commuting Unitaries

This section is dedicated to presenting the results Gerhold and Shalit from [69], on dilations of q-commuting unitaries.

Let $\theta \in \mathbb{R}$ and write $q = e^{i\theta}$. If u and v are two unitaries that satisfy $vu = quv$, then we say that u and v are **q-commuting**. We denote by $\mathcal{A}_\theta$ the universal C*-algebra generated by a pair of q-commuting unitaries, and we call $\mathcal{A}_\theta$ a **rational/irrational rotation C*-algebra** if $\frac{\theta}{2\pi}$ is rational/irrational respectively. We shall write u_θ, v_θ for the generators of $\mathcal{A}_\theta$. The rotation C*-algebras have been of widespread interest ever since they were introduced by Rieffel [132]. A good reference for this subject is Boca's book [31].

In an attempt to make some progress in our understanding of the general constant C_d from Problem 9.13, Malte Gerhold and I studied a certain refinement of that problem which is of independent interest. Instead of dilating arbitrary tuples of contractions, we considered the task of dilating pairs of unitaries u, v that satisfy the q-commutation relation $vu = quv$, and studied the dependence of the dilation constant on the parameter q. In the context of Problem 9.13, it is worth noting that, by a result of Buske and Peters [34] (see also [86]), every pair of q-commuting contractions has a q-commuting unitary power dilation; therefore, this work has implications to all pairs of q-commuting operators. Surprisingly, our dilation results also have implications for the continuity of the norm and the spectrum of the almost Mathieu operator from mathematical physics (this application will be discussed in the final section).

For every $\theta \in \mathbb{R}$ we define the optimal dilation constant

$$c_\theta := \inf\{c > 1 \mid (u_\theta, v_\theta) \prec c(U, V) \text{ where } U, V \text{ are commuting unitaries}\}.$$

We note that the infimum is actually a minimum, and that it is equal to the infimum of the constants c that satisfy: for every q-commuting pair of unitaries U, V there exists a commuting normal dilation M, N such that $\|M\|, \|N\| \leq c$ (see [69, Proposition 2.3]). Thus, c_θ is a lower bound for the constant C_2 from Problem 9.13.

10.1 Continuity of the Dilation Scale

Theorem 10.1 (Theorem 3.2, [69]) *Let $\theta, \theta' \in \mathbb{R}$, set $q = e^{i\theta}, q' = e^{i\theta'}$, and put $c = e^{\frac{1}{4}|\theta - \theta'|}$. Then for any pair of q-commuting unitaries U, V there exists a pair of q'-commuting unitaries U', V' such that cU', cV' dilates U, V.*

Proof The proof makes use of the Weyl operators on symmetric Fock space (see [111, Section 20]). For a Hilbert space H let $H^{\otimes_s k}$ be the k-fold symmetric tensor product of H, and let

$$\Gamma(H) := \bigoplus_{k=0}^{\infty} H^{\otimes_s k}$$

be the symmetric Fock space over H. The **exponential vectors**

$$e(x) := \sum_{k=0}^{\infty} \frac{1}{\sqrt{k!}} x^{\otimes k} \quad , \quad x \in H,$$

form a linearly independent and total subset of $\Gamma(H)$. For $z \in H$ we define the **Weyl unitary** $W(z) \in B(\Gamma(H))$ which is determined by

$$W(z)e(x) = e(z+x) \exp\left(-\frac{\|z\|^2}{2} - \langle x, z\rangle\right)$$

for all exponential vectors $e(x)$.

Consider Hilbert spaces $H \subset K$ with p the projection onto H, and the symmetric Fock spaces $\Gamma(H) \subset \Gamma(K)$ with P the projection onto $\Gamma(H)$. We write $p^\perp$ for the projection onto the orthogonal complement $H^\perp$. Note that for exponential vectors we have $Pe(x) = e(px)$. For every $y, z \in K$, the Weyl unitaries $W(y), W(z)$ satisfy:

1. $W(z)$ and $W(y)$ commute up to the phase factor $e^{2i \operatorname{Im}\langle y,z\rangle}$.
2. $PW(z)\big|_{\Gamma(H)} = e^{-\frac{\|p^\perp z\|^2}{2}} W(pz)$, so it is a scalar multiple of a unitary on $\Gamma(H)$.
3. $PW(z)\big|_{\Gamma(H)}$ and $PW(y)\big|_{\Gamma(H)}$ commute up to a phase factor $e^{2i \operatorname{Im}\langle py,z\rangle}$.

In [69] it is shown that, assuming without loss that $\theta > \theta'$, things can be arranged so that there are two linearly independent vectors z, y so that pz and py are linearly independent, and such that

1. $p^\perp y = -ip^\perp z$,
2. $\theta' = 2\operatorname{Im}\langle y, z\rangle$,
3. $\theta = 2\operatorname{Im}\langle py, z\rangle$.

Then we get q'-commutation of $W(z)$ and $W(y)$, q-commutation of the operators $PW(z)\big|_{\Gamma(H)}$ and $PW(y)\big|_{\Gamma(H)}$, and

$$\theta - \theta' = -2\operatorname{Im}\langle p^\perp y, z\rangle = 2\|p^\perp z\|^2 = 2\|p^\perp y\|^2,$$

so

$$\left\| PW(z)\big|_{\Gamma(H)} \right\| = \left\| PW(y)\big|_{\Gamma(H)} \right\| = e^{-\frac{\|p^{\perp}y\|^2}{2}} = e^{-\frac{|\theta-\theta'|}{4}}.$$

Now if we put

$$U = e^{\frac{|\theta-\theta'|}{4}} PW(z)\big|_{\Gamma(H)} \ , \ V = e^{\frac{|\theta-\theta'|}{4}} PW(y)\big|_{\Gamma(H)},$$

and

$$U' = W(z) \ , \ V' = W(y)$$

then we get the statement for this particular q-commuting pair U, V. Since the Weyl unitaries give rise to a universal representation of $\mathcal{A}_\theta$, the general result follows (see [69, Proposition 2.3]). ■

From the above result we obtained continuity of the dilation scale.

Corollary 10.2 (Corollary 3.4, [69]) *The optimal dilation scale c_θ depends Lipschitz continuously on θ. More precisely, for all $\theta, \theta' \in \mathbb{R}$ we have*

$$|c_\theta - c_{\theta'}| \leq 0.39 \left|\theta - \theta'\right|.$$

10.2 The Optimal Dilation Scale

The main result of [69] is the following theorem.

Theorem 10.3 (Theorems 6.3 and 6.4, [69]) *Let $\theta, \theta' \in \mathbb{R}$, $q = e^{i\theta}$, $q' = e^{i\theta'}$, and put $\gamma = \theta' - \theta$. The smallest constant $c_{\theta,\theta'}$ such that every pair of q-commuting unitaries can be dilated to $c_{\theta,\theta'}$ times a pair of q'-commuting unitaries is given by*

$$c_{\theta,\theta'} = \frac{4}{\|u_\gamma + u_\gamma^* + v_\gamma + v_\gamma^*\|}.$$

In particular, for every $\theta \in \mathbb{R}$,

$$c_\theta = \frac{4}{\|u_\theta + u_\theta^* + v_\theta + v_\theta^*\|}.$$

Proof Since it is a nice construction that we have not yet seen, let us show just that the value of $c_{\theta,\theta'}$ is no bigger than $\frac{4}{\|u_\gamma + u_\gamma^* + v_\gamma + v_\gamma^*\|}$; for the optimality of the dilation constant we refer the reader to [69] (the formula for $c_\theta = c_{\theta,0}$ follows, since it is not hard to see that $c_\theta = c_{-\theta}$).

Represent $C^*(U, V)$ concretely on a Hilbert space $\mathcal{H}$. Let u_γ, v_γ be the universal generators of $\mathcal{A}_\gamma$ and put $h_\gamma := u_\gamma + u_\gamma^* + v_\gamma + v_\gamma^*$. We claim that there exists a state φ on $\mathcal{A}_\gamma$ such that $|\varphi(u_\gamma)| = |\varphi(v_\gamma)| = \frac{\|h_\gamma\|}{4}$ (for the existence of such a state, see [69]). Assuming the existence of such a state, we define

$$U' = U \otimes \frac{\pi(u_\gamma)}{\varphi(u_\gamma)} \quad , \quad V' = V \otimes \frac{\pi(v_\gamma)}{\varphi(v_\gamma)}.$$

on $\mathcal{K} = \mathcal{H} \otimes \mathcal{L}$, where $\pi : A_\gamma \to B(\mathcal{L})$ is the GNS representation of φ. These are q'-commuting scalar multiples of unitaries, and they have norm $\frac{4}{\|h_\gamma\|}$. By construction, there exists a unit vector $x \in \mathcal{L}$ such that $\varphi(a) = \langle \pi(a)x, x\rangle$ for all $a \in A_\gamma$. Consider the isometry $W : \mathcal{H} \to \mathcal{H} \otimes \mathcal{L}$ defined by

$$Wh = h \otimes x \quad , \quad h \in \mathcal{H}.$$

Then

$$W^* U' W = \frac{1}{\varphi(u_\gamma)} \langle \pi(u_\gamma)x, x\rangle U = U$$

and

$$W^* V' W = \frac{1}{\varphi(v_\gamma)} \langle \pi(v_\gamma)x, x\rangle V = V,$$

and the proof of the existence of a dilation is complete. ∎

The operator $h_\theta = u_\theta + u_\theta^* + v_\theta + v_\theta^*$ is called the *almost Mathieu operator*, and it has been intensively studied by mathematical physicists, before and especially after Hofstadter's influential paper [77] (we will return to it in the next section). However, the precise behaviour of the norm $\|h_\theta\|$ as a function of θ is still not completely understood. We believe that the most detailed analysis is contained in the paper [32].

In [69, Section 7] we obtained numerical values for $c_\theta = 4/\|h_\theta\|$ for various θ. We calculated by hand $c_{\frac{4}{5}\pi} \approx 1.5279$, allowing us to push the lower bound $C_2 \geq 1.41...$ to $C_2 \geq 1.52$. We also made some numerical computations, which lead to an improved estimate $C_2 \geq \max_\theta c_\theta \geq 1.5437$. The latter value is an approximation of the constant c_{θ_s} attained at the *silver mean* $\theta_s = \frac{2\pi}{\gamma_s} = 2\pi(\sqrt{2} - 1)$ (where $\gamma_s = \sqrt{2} + 1$ is the *silver ratio*) which we conjecture to be the angle where the maximum is attained. However, we do not expect that the maximal value of c_θ will give a tight lower approximation for C_2. Determining the value of C_2 remains an open problem.

10.3 An Application: Continuity of the Spectrum of Almost Mathieu Operators

The almost Mathieu operator $h_\theta = u_\theta + u_\theta^* + v_\theta + v_\theta^*$, which appears in the formula $c_\theta = \frac{4}{\|h_\theta\|}$, arises as the Hamiltonian in a certain mathematical model describing an electron in a lattice under the influence of a magnetic field; see Hofstadter [77]. This operator has been keeping mathematicians and physicists busy for more than a generation. Hofstadter's paper included a picture that depicts the spectrum (computed numerically) of h_θ for various values of θ, famously known as the *Hofstadter butterfly* (please go ahead and google it). From observing the Hofstadter butterfly, one is led to making several conjectures.

First and foremost, it appears that the spectrum of h_θ varies continuously with θ; since θ is a physical parameter of the system studied, and the spectrum is supposed to describe possible energy levels, any other possibility is unreasonable. There are other natural conjectures to make, suggested just by looking at the picture. The most famous one is perhaps what Barry Simon dubbed as the *Ten Martini Problem*, which asks whether the spectrum is a Cantor set for irrational angles. This problem was settled (in greater generality) by Avila and Jitomirskaya (see [17] for the conclusive work as well as for references to earlier work).

The continuity of the spectrum $\sigma(h_\theta)$ is a delicate problem that attracted a lot of attention. For example, in [37] Choi, Elliott, and Yui showed that the spectrum $\sigma(h_\theta)$ of h_θ depends Hölder continuously (in the Hausdorff metric) on θ, with Hölder exponent $1/3$. This was soon improved by Avron, Mouche, and Simon to Hölder continuity with exponent $1/2$ [18]. The $1/2$-Hölder continuity of the spectrum also follows from a result of Haagerup and Rørdam, who showed that there exist $1/2$-Hölder *norm* continuous paths $\theta \mapsto u_\theta \in B(\mathcal{H})$, $\theta \mapsto v_\theta \in B(\mathcal{H})$ [71, Corollary 5.5].

As an application of our dilation techniques, we are able to recover the best possible continuity result regarding the spectrum of the operator h_θ. This result is not new, but our proof is new and simple, and I believe that it is a beautiful and exciting application of dilation theory with which to close this survey. The following theorem also implies that the rotation C*-algebras form a continuous field of C*-algebras, a result due to Elliott [59]. Our dilation methods can also be used to recover the result of Bellisard [20], that the norm of h_θ is a Lipschitz continuous function of θ.

Theorem 10.4 *Let p be a selfadjoint $*$-polynomial in two noncommuting variables. Then the spectrum $\sigma(p(u_\theta, v_\theta))$ of $p(u_\theta, v_\theta)$ is $\frac{1}{2}$-Hölder continuous in θ with respect to the Hausdorff distance for compact subsets of $\mathbb{R}$.*

Proof Let us present the idea of the proof for the most important case

$$p(u_\theta, v_\theta) = h_\theta = u_\theta + u_\theta^* + v_\theta + v_\theta^*,$$

without going into the details of Hölder continuity. The idea is that, due to Theorem 10.1, when $\theta \approx \theta'$ we have the dilation $(u_\theta, v_\theta) \prec (cu_{\theta'}, cv_{\theta'})$ with $c = e^{\frac{1}{4}|\theta-\theta'|} \approx 1$. Thus,

$$cu_{\theta'} = \begin{pmatrix} u_\theta & x \\ y & z \end{pmatrix}$$

and so x and y must be small, to be precise,

$$\|x\|, \|y\| \leq \sqrt{c^2 - 1} \approx 0.$$

A similar estimate holds for the off diagonal block of $cv_{\theta'}$ which dilates v_θ. By a basic lemma in operator theory, for any selfadjoint operators a and b, the Hausdorff distance between their spectra is bounded as follows:

$$d(\sigma(a), \sigma(b)) \leq \|a - b\|.$$

We have $h_\theta = u_\theta + u_\theta^* + v_\theta + v_\theta^*$, and so

$$ch_{\theta'} = \begin{pmatrix} h_\theta & * \\ * & * \end{pmatrix} \approx \begin{pmatrix} h_\theta & 0 \\ 0 & * \end{pmatrix},$$

because the off diagonal blocks have small norm, and therefore

$$\sigma(h_\theta) \subseteq \sigma\left(\begin{pmatrix} h_\theta & 0 \\ 0 & * \end{pmatrix}\right) \approx \sigma(ch_{\theta'}) \approx \sigma(h_{\theta'}).$$

In the same way one shows that $\sigma(h_{\theta'})$ is approximately contained in $\sigma(h_\theta)$, and therefore the Hausdorff distance between the spectra is small. ■

It is interesting to note that the above proof generalizes very easily to higher dimensional noncommutative tori. Determining the precise dilation scales for higher dimensional noncommutative tori remains an open problem.

Acknowledgments This survey paper grew out of the talk that I gave at the International Workshop on Operator Theory and its Applications (IWOTA) that took place in the Instituto Superior Técnico, Lisbon, Portugal, in July 2019. I am grateful to the organizers of IWOTA 2019 for inviting me to speak in this incredibly successful workshop, and especially to Amélia Bastos, for inviting me to contribute to these proceedings. I used a preliminary version of this survey as lecture notes for a mini-course that I gave in the workshop Noncommutative Geometry and its Applications, which took place in January 2020, in NISER, Bhubaneswar, India. I am grateful to the organizers Bata Krishna Das, Sutanu Roy and Jaydeb Sarkar, for the wonderful hospitality and the opportunity to speak and organize my thoughts on dilation theory. I also owe thanks to Michael Skeide and to Fanciszek Szafraniec, for helpful feedback on preliminary versions. Finally, I wish to thank an anonymous referee for several useful comments and corrections.

This project was partially supported by ISF Grant no. 195/16.

References

1. J. Agler, Rational dilation on an annulus. Ann. Math. **121**, 537–563 (1985)
2. J. Agler, J.E. McCarthy, *Pick Interpolation and Hilbert Function Spaces*. Graduate Studies in Mathematics, vol. 44 (American Mathematical Society, Providence, 2002)
3. J. Agler, J.E. McCarthy, Distinguished varieties. Acta Math. **194**, 133–153 (2005)
4. J. Agler, N.J. Young, Operators having the symmetrized bidisc as a spectral set. Proc. Edinb. Math. Soc. **43**, 195–210 (2000)
5. J. Agler, J. Harland, B.J. Raphael, *Classical Function Theory, Operator Dilation Theory and Machine Computation on Multiply-Connected Domains*. Memoirs of the American Mathematical Society (American Mathematical Society, Providence, 2008)
6. M.A. Akcoglu, L. Sucheston, Dilations of positive contractions on L_p spaces. Can. Math. Bull. **20**, 285–292 (1977)
7. C.G. Ambrozie, A. Gheondea, An interpolation problem for completely positive maps on matrix algebras: solvability and parametrization. Linear Multilinear Algebra **63**, 826–851 (2015)
8. C. Ambrozie, V. Muller, Commutative dilation theory, in *Operator Theory*, ed. by D. Alpay (Springer, Berlin, 2014)
9. T. Andô, On a pair of commutative contractions. Acta Sci. Math. **24**, 88–90 (1963)
10. W.B. Arveson, Subalgebras of C^*-algebras. Acta Math. **123**, 141–224 (1969)
11. W.B. Arveson, Subalgebras of C*-algebras II. Acta Math. **128**, 271–308 (1972)
12. W.B. Arveson, Subalgebras of C^*-algebras III: multivariable operator theory. Acta Math. **181**, 159–228 (1998)
13. W.B. Arveson, Dilation theory yesterday and today, in *A Glimpse of Hilbert Space Operators: Paul R. Halmos in Memoriam*, ed. by S. Axler, P. Rosenthal, D. Sarason (Birkhäuser, Basel, 2010)
14. W.B. Arveson, *Non-commutative Dynamics and E-semigroups*. Springer Monographs in Mathematics (Springer, Berlin, 2003)
15. W.B. Arveson, The noncommutative Choquet boundary. J. Am. Math. Soc. **21**, 1065–1084 (2008)
16. G. Aubrun, L. Lami, C. Palazuelos, M. Plavala, Entangleability of cones (2020). arXiv:1911.09663
17. A. Avila, S. Jitomirskaya, The ten martini problem. Ann. Math. **170**, 303–342 (2009)
18. J. Avron, P.H.M.v. Mouche, B. Simon, On the measure of the spectrum for the almost Mathieu operator. Commun. Math. Phys. **132**, 103–118 (1990)
19. C. Badea, B. Beckermann, Spectral sets, in *Handbook of Linear Algebra*, ed. by L. Hogben (Chapman and Hall/CRC, Boca Raton, 2014)
20. J. Bellisard, Lipshitz continuity of gap boundaries for Hofstadter-like spectra. Commun. Math. Phys. **160**, 599–613 (1994)
21. A. Ben-Tal, A. Nemirovski, On tractable approximations of uncertain linear matrix inequalities affected by interval uncertainty. SIAM J. Optim. **12**, 811–833 (2002)
22. H. Bercovici, D. Timotin, The numerical range of a contraction with finite defect numbers. J. Math. Anal. Appl. **417**, 42–56 (2014)
23. H. Bercovici, C. Foias, L. Kerchy, B. Sz.-Nagy, *Harmonic Analysis of Operators on Hilbert Space*. Universitext (Springer, Berlin, 2010)
24. B.V.R. Bhat, An index theory for quantum dynamical semigroups. Trans. Am. Math. Soc. **348**, 561–583 (1996)
25. B.V.R. Bhat, A generalized intertwining lifting theorem, in *Operator Algebras and Applications, II, Waterloo, ON, 1994-1995*, Fields Institute Communications, vol. 20 (American Mathemtical Society, Providence, 1998), pp. 1–10
26. B.V.R. Bhat, T. Bhattacharyya, Dilations, completely positive maps and geometry (forthcoming book)

27. B.V.R. Bhat, M. Mukherjee, Inclusion systems and amalgamated products of product systems. Infin. Dimens. Anal. Quant. Probab. Relat. Top. **13**, 1–26 (2010)
28. B.V.R. Bhat, M. Skeide, Tensor product systems of Hilbert modules and dilations of completely positive semigroups. Infin. Dimens. Anal. Quant. Probab. Relat. Top. **3**, 519–575 (2000)
29. T. Bhattacharyya, S. Pal, S. Shyam Roy, Dilations of - contractions by solving operator equations. Adv. Math. **230**, 577–606 (2012)
30. A. Bluhm, I. Nechita, Joint measurability of quantum effects and the matrix diamond. J. Math. Phys. **59**, 112202 (2018)
31. F.P. Boca, *Rotation C*-Algebras and Almost Mathieu Operators* (The Theta Foundation, Bucharest, 2001)
32. F.P. Boca, A. Zaharescu, Norm estimates of almost Mathieu operators. J. Funct. Anal. **220**, 76–96 (2005)
33. J.W. Bunce, Models for n-tuples of noncommuting operators. J. Funct. Anal. **57**, 21–30 (1984)
34. D.R. Buske, J.R. Peters, Semicrossed products of the disk algebra: contractive representations and maximal ideals. Pac. J. Math. **185**, 97–113 (1998)
35. M.D. Choi, K.R. Davidson, A 3×3 dilation counterexample. Bull. Lond. Math. Soc. **45**, 511–519 (2013)
36. M.D. Choi, C.K. Li, Constrained unitary dilations and numerical ranges. J. Operator Theory **46**, 435–447 (2001)
37. M.-D. Choi, G.A. Elliott, N. Yui, Gauss polynomials and the rotation algebra. Invent. Math. **99**, 225–246 (1990)
38. D. Cohen, Dilations of matrices. Thesis (M.Sc.), Ben-Gurion University (2015). arXiv:1503.07334
39. M.J. Crabb, A.M. Davie, Von Neumann's inequality for hilbert space operators. Bull. Lond. Math. Soc. **7**, 49–50 (1975)
40. M. Crouzeix, Numerical range and functional calculus in Hilbert space. J. Funct. Anal. **244**, 668–990 (2007)
41. K.R. Davidson, E.G. Katsoulis, Dilation theory, commutant lifting and semicrossed products. Doc. Math. **16**, 781–868 (2011)
42. K.R. Davidson, M. Kennedy, The Choquet boundary of an operator system. Duke Math. J. **164**, 2989–3004 (2015)
43. K.R. Davidson, M. Kennedy, Noncommutative Choquet theory. arXiv:1905.08436
44. K.R. Davidson, D.R. Pitts, Nevanlinna-Pick interpolation for non-commutative analytic Toeplitz algebras. Integr. Equ. Operator Theory **31**, 321–337 (1998)
45. K.R. Davidson, D.R. Pitts, The algebraic structure of non-commutative analytic Toeplitz algebras. Math. Ann. **311**, 275–303 (1998)
46. K.R. Davidson, A. Dor-On, O.M. Shalit, B. Solel, Dilations, inclusions of matrix convex sets, and completely positive maps. Int. Math. Res. Not. **2017**, 4069–4130 (2017)
47. K.R. Davidson, A. Dor-On, O.M. Shalit, B. Solel, Dilations, inclusions of matrix convex sets, and completely positive maps (2018). arXiv:1601.07993v3 [math.OA]
48. K.R. Davidson, A.H. Fuller, E.T.A. Kakariadis, Semicrossed products of operator algebras: a survey. New York J. Math. **24a**, 56–86 (2018)
49. E.B. Davies, *Quantum Theory of Open Systems* (Academic, Cambridge, 1976)
50. A. Dor-On, Techniques in operator algebras: classification, dilation and non-commutative boundary theory. Thesis (Ph.D.) University of Waterloo (2017)
51. A. Dor-On, G. Salomon, Full Cuntz-Krieger dilations via non-commutative boundaries. J. Lond. Math. Soc. **98**, 416–438 (2018)
52. R. Douglas, On extending commutative ssemigroupsemigroups of isometries. Bull. Lond. Math. Soc. **1**, 157–159 (1969)
53. M. Dritschel, S. McCullough, Boundary representations for families of representations of operator algebras and spaces. J. Operator Theory **53**, 159–167 (2005)
54. M.A. Dritschel, S. McCullough, The failure of rational dilation on a triply connected domain. J. Am. Math. Soc. **18**, 873–918 (2005)

55. S.W. Drury, A generalization of von Neumann's inequality to the complex ball. Proc. Am. Math. Soc. **68**, 300–304 (1978)
56. S.W. Drury, *Remarks on von Neumann's Inequality*. Lecture Notes in Mathematics, vol. 995 (Springer, Berlin, 1983)
57. E.G. Effros, S. Winkler, Matrix convexity: operator analogues of the bipolar and Hahn-Banach theorems. J. Funct. Anal. **144**, 117–152 (1997)
58. E. Egerváry, On the contractive linear transformations of n-dimensional vector space. Acta Sci. Math. Szeged **15**, 178–182 (1954)
59. G.A. Elliott, Gaps in the spectrum of an almost periodic Schrödinger operator. C.R. Math. Rep. Acad. Sci. Can. **4**, 255–299 (1982)
60. D.E. Evans, J.T. Lewis, Dilations of dynamical semi-groups. Commun. Math. Phys. **50**, 219–227 (1976)
61. E. Evert, J.W. Helton, I. Klep, S. McCullough, Extreme points of matrix convex sets, free spectrahedra, and dilation theory. J. Geom. Anal. **28**, 1373–1408 (2018)
62. S. Fackler, J. Glück, A toolkit for constructing dilations on Banach spaces. Proc. Lond. Math. Soc. **118**, 416–440 (2019)
63. C. Foias, A.E. Frazho, *The Commutant Lifting Approach to Interpolation Problems* (Birkhäuser, Basel, 1990)
64. A.E. Frazho, Models for noncommuting operators. J. Funct. Anal. **48**, 1–11 (1982)
65. T. Fritz, T. Netzer, A. Thom, Spectrahedral containment and operator systems with finite-dimensional realization. SIAM J. Appl. Algebra Geom. **1**, 556–574 (2017)
66. A.H. Fuller, Finitely correlated representations of product systems of C*-correspondences over $\mathbb{N}^k$. J. Funct. Anal. **260**, 574–611 (2011)
67. D.J. Gaebler, Continuous unital dilations of completely positive semigroups. J. Funct. Anal. **269**, 998–1027 (2015)
68. T.W. Gamelin, *Uniform Algebras*, vol. 311 (American Mathematical Society, Providence, 2005)
69. M. Gerhold, O.M. Shalit, Dilations of q-commuting unitaries (2019). arXiv:1902.10362
70. M. Gerhold, O.M. Shalit, On the matrix range of random matrices (2020). arXiv:1911.12102
71. U. Haagerup, M. Rørdam, Perturbations of the rotation C*-algebras and of the Heisenberg commutation relations. Duke Math. J. **77**, 227–256 (1995)
72. P.R. Halmos, Normal dilations and extensions of operators. Summa Brasil. Math. **2**, 125–134 (1950)
73. M. Hartz, M. Lupini, Dilation theory in finite dimensions and matrix convexity (2019). arXiv:1910.03549
74. J.W. Helton, I. Klep, S. McCullough, The matricial relaxation of a linear matrix inequality. Math. Program. **138**, 401–445 (2013)
75. J.W. Helton, I. Klep, S. McCullough, Matrix convex hulls of free semialgebraic sets. Trans. Am. Math. Soc. **368**, 3105–3139 (2016)
76. J.W. Helton, I. Klep, S. McCullough, M. Schweighofer, *Dilations, Linear Matrix Inequalities, the Matrix Cube Problem and Beta Distributions*. Memoirs of the American Mathematical Society (American Mathematical Society, Providence, 2019)
77. D.R. Hofstadter, Energy levels and wave functions of Bloch electrons in rational and irrational magnetic fields. Phys. Rev. B. **14**, 2239–2249 (1976)
78. J.A. Holbrook, Schur norms and the multivariate von Neumann inequality, in *Operator Theory: Advances and Applications*, vol. 127 (Birkhäuser, Basel, 2001)
79. Z. Hu, R. Xia, S. Kais, A quantum algorithm for evolving open quantum dynamics on quantum computing devices (2019). arXiv:1904.00910
80. B. Huber, T. Netzer, A note on non-commutative polytopes and polyhedra (2019). arXiv:1809.00476
81. M. Izumi, E_0-semigroups: around and beyond Arveson's work. J. Oper. Theory **68**, 335–363 (2012)
82. E.T.A. Kakariadis, O.M. Shalit, Operator algebras of monomial ideals in noncommuting variables. J. Math. Anal. Appl. **472**, 738–813 (2019)

83. I. Kaplansky, Modules over operator algebras. Am. J. Math. **75**, 839–858 (1953)
84. E. Katsoulis, C. Ramsey, *Crossed Products of Operator Algebras*. Memoirs of the American Mathematical Society (American Mathematical Society, Providence, 2019)
85. M. Kennedy, O.M. Shalit, Essential normality, essential norms and hyperrigidity. J. Funct. Anal. **270**, 2812–2815 (2016)
86. D. Keshari, N. Mallick, q-commuting dilation. Proc. Am. Math. Soc. **147**, 655–669 (2019)
87. G. Knese, The von Neumann inequality for 3×3 matrices. Bull. Lond. Math. Soc. **48**, 53–57 (2016)
88. Ł. Kosiński, Three-point Nevanlinna-Pick problem in the polydisc. Proc. Lond. Math. Soc. **111**, 887–910 (2015)
89. K. Kraus, *States, Effects, and Operations: Fundamental Notions of Quantum Theory*. Lecture Notes in Physics, vol. 190 (Springer, Berlin, 1983)
90. T.-L. Kriel, An introduction to matrix convex sets and free spectrahedra (2016). arXiv:1611.03103v6
91. B. Kümmerer, Markov dilations and W*-algberas. J. Funct. Anal. **63**, 139–177 (1985)
92. M. Laca, From endomorphisms to automorphisms and back: dilations and full corners. J. Lond. Math. Soc. **61**, 893–904 (2000)
93. E.C. Lance, *Hilbert C*-modules: A Toolkit for Operator Algebraists*, vol. 210 (Cambridge University Press, Cambridge, 1995)
94. J. Levick, R.T.W. Martin, Matrix N-dilations of quantum channels. Oper. Matrices **12**, 977–995 (2018)
95. E. Levy, O.M. Shalit, Dilation theory in finite dimensions: the possible, the impossible and the unknown. Rocky Mountain J. Math. **44**, 203–221 (2014)
96. C.K. Li, Y.T. Poon, Convexity of the joint numerical range. SIAM J. Matrix Anal. Appl. **21**, 668–678 (2000)
97. C.K. Li, Y.T. Poon, Interpolation by completely positive maps. Linear Multilinear Algebra **59**, 1159–1170 (2011)
98. D. Markiewicz, O.M. Shalit, Continuity of CP-semigroups in the point-strong topology. J. Operator Theory **64**, 149–154 (2010)
99. J.E. McCarthy, O.M. Shalit, Unitary N-dilations for tuples of commuting matrices. Proc. Am. Math. Soc. **14**, 563–571 (2013)
100. P.S. Muhly, B. Solel, Tensor algebras over C^*-correspondences: representations, dilations, and C^*-envelopes. J. Funct. Anal. **158**, 389–457 (1998)
101. P.S. Muhly, B. Solel, An algebraic characterization of boundary representations, in *Operator Theory: Advances and Applications*, vol. 104 (Birkhäuser, Basel, 1998), pp. 189–196
102. P.S. Muhly, B. Solel, Quantum Markov processes (correspondences and dilations). Int. J. Math. **13**, 863–906 (2002)
103. P.S. Muhly, B. Solel, Hardy algebras, W*-correspondences and interpolation theory. Math. Ann. **330**, 353–415 (2004)
104. P. Muhly, B. Solel, Quantum Markov semigroups (product systems and subordination). Int. J. Math. **18**, 633–669 (2007)
105. V. Müller, F-H. Vasilescu, Standard models for some commuting multioperators. Proc. Am. Math. Soc. **117**, 979–989 (1993)
106. B. Nagy, On contractions in Hilbert space. Acta Sci. Math **15**, 87–92 (2013)
107. I.L. Nielsen, M.A. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000)
108. D. Opela, A generalization of Andô's theorem and Parrott's example. Proc. Am Math. Soc. **134**, 2703–2710 (2006)
109. S. Pal, O.M. Shalit, Spectral sets and distinguished varieties in the symmetrized bidisc. J. Funct. Anal. **266**, 5779–5800 (2014)
110. S. Parrott, Unitary dilations for commuting contractions. Pac. J. Math. **34**, 481–490 (1970)
111. K.R. Parthasarathy, *An Introduction to Quantum Stochastic Calculus*. Monographs in Mathematics, vol. 85 (Birkhäuser, Basel, 2012)
112. W.L. Paschke, Inner product modules over B^*-algebras. Trans. Am. Math. Soc. **182**, 443–468 (1973)

113. B. Passer, Shape, scale, and minimality of matrix ranges. Trans. Am. Math. Soc. **372**, 1451–1484 (2019)
114. B. Passer, V.I. Paulsen, Matrix range characterization of operator system properties (preprint). arXiv:1912.06279
115. B. Passer, O.M. Shalit, Compressions of compact tuples. Linear Algebra Appl. **564**, 264–283 (2019)
116. B. Passer, O.M. Shalit, B. Solel, Minimal and maximal matrix convex sets. J. Funct. Anal. **274**, 3197–3253 (2018)
117. V.I. Paulsen, Representations of function algebras, abstract operator spaces, and Banach space geometry. J. Funct. Anal. **109**(1), 113–129 (1992)
118. V.I. Paulsen, *Completely Bounded Maps and Operator Algebras* (Cambridge University Press, Cambridge, 2002)
119. V.I. Paulsen, M. Raghupathi, *An Introduction to the Theory of Reproducing Kernel Hilbert Spaces*. Cambridge Studies in Advanced Mathematics, vol. 152 (Cambridge University Press, Cambridge, 2016)
120. V.I. Paulsen, I.G. Todorov, M. Tomforde, Operator system structures on ordered spaces. Proc. Lond. Math. Soc. **102**, 25–49 (2011)
121. G. Pisier, *Similarity Problems and Completely Bounded Maps*. Lecture Notes of Mathematics, vol. 1618 (Springer, Berlin, 1996)
122. G. Pisier, *Introduction to Operator Space Theory*, vol. 294 (Cambridge University Press, Cambridge, 2003)
123. G. Popescu, Isometric dilations for infinite sequences of noncommuting operators. Trans. Am. Math. Soc. **316**, 523–536 (1989)
124. G. Popescu, Von Neumann inequality for $(B(\mathcal{H})^n)_1$. Math. Scand. **68**, 292–304 (1991)
125. G. Popescu, Poisson transforms on some C*-algebras generated by isometries. J. Funct. Anal. **161**, 27–61 (1999)
126. G. Popescu, Free holomorphic functions on the unit ball of $B(\mathcal{H})^n$. J. Funct. Anal. **241**, 268–333 (2006)
127. G. Popescu, Operator theory on noncommutative varieties. Ind. Univ. Math. J. **56**, 389–442 (2006)
128. G. Popescu, Berezin transforms on noncommutative polydomains. Trans. Am. Math. Soc. **368**, 4357–4416 (2016)
129. B. Prunaru, Lifting fixed points of completely positive semigroups. Integr. Equ. Oper. Theory **72**, 219–222 (2012)
130. M. Ptak, Unitary dilations of multiparameter semigroups of operators. Ann. Polon. Math. **45**, 237–243 (1985)
131. M. Rieffel, Induced representations of C*-algebras. Adv. Math. **13**, 176–257 (1974)
132. M. Rieffel, C*-algebras associated with irrational rotations. Pac. J. Math. **93**, 415–429 (1981)
133. F. Riesz, B. Sz.-Nagy, *Functional Analysis* (Dover, New York, 1990) (first published in 1955)
134. G. Salomon, O.M. Shalit, E. Shamovich, Algebras of bounded noncommutative analytic functions on subvarieties of the noncommutative unit ball. Trans. Am. Math. Soc. **370**, 8639–8690 (2018)
135. G. Salomon, O.M. Shalit, E. Shamovich, Algebras of noncommutative functions on subvarieties of the noncommutative ball: the bounded and completely bounded isomorphism problem. J. Funct. Anal. (2019). arXiv:1806.00410
136. D. Sarason, On spectral sets having connected complement. Acta Sci. Math. **26**, 289–299 (1965)
137. D. Sarason, Generalized interpolation in H^∞. Trans. Am. Math. Soc. **127**, 179–203 (1967)
138. J. Sarkar, Applications of Hilbert module approach to multivariable operator theory, in *Operator Theory*, ed. by D. Alpay (Springer, Berlin, 2014)
139. J. Sarkar, Operator theory on symmetrized bidisc. Ind. Univ. Math. J. **64**, 847–873 (2015)
140. D. SeLegue, Minimal Dilations of CP maps and C*-extension of the Szegö limit theorem. Ph.D Dissertation, University of California, Berkeley (1997)

141. O.M. Shalit, E_0-dilation of strongly commuting CP_0-semigroups. J. Funct. Anal. **255**, 46–89 (2008)
142. O.M. Shalit, What type of dynamics arise in E_0-dilations of commuting Quantum Markov Semigroups?. Infin. Dimens. Anal. Quant. Probab. Relat. Top. **11**, 393–403 (2008)
143. O.M. Shalit, Representing a product system representation as a contractive semigroup and applications to regular isometric dilations. Can. Math. Bull. **53**, 550–563 (2010)
144. O.M. Shalit, E-dilations of strongly commuting CP-semigroups (the nonunital case). Houston J. Math. **35**(1), 203–232 (2011)
145. O.M. Shalit, A sneaky proof of the maximum modulus principle. Am. Math. Month. **120**(4), 359–362 (2013)
146. O.M. Shalit, Operator theory and function theory in Drury-Arveson space and its quotients, in *Operator Theory*, ed. by D. Alpay (Springer, Berlin, 2014)
147. O.M. Shalit, *A First Course in Functional Analysis* (Chapman and Hall/CRC, Boca Raton, 2017)
148. O.M. Shalit, M. Skeide, Three commuting, unital, completely positive maps that have no minimal dilation. Integr. Equ. Oper. Theory **71**, 55–63 (2011)
149. O.M. Shalit, M. Skeide, CP-semigroups and dilations, subproduct systems and superproduct systems: the multi-parameter case and beyond** (in preparation)
150. O.M. Shalit, B. Solel, Subproduct systems. Doc. Math. **14**, 801–868 (2009)
151. M. Słociński, Unitary dilation of two-parameter semi-groups of contractions. Bull. Acad. Polon. Sci. Ser. Sci. Math. Astronom. Phys. **22**, 1011–1014 (1974)
152. A. Skalski, On isometric dilations of product systems of C*-correspondences and applications to families of contractions associated to higher-rank graphs. Ind. Univ. Math. J. **58**, 2227–2252 (2009)
153. A. Skalski, J. Zacharias, Wold decomposition for representations of product systems of C*-correspondences. Int. J. Math. **19**, 455–479 (2008)
154. M. Skeide, Hilbert modules and applications in quantum probability. Habilitationsschrift, Cottbus (2001). http://web.unimol.it/skeide/
155. M. Skeide, Isometric dilations of representations of product systems via commutants. Int. J. Math. **19**, 521–539 (2008)
156. M. Skeide, Classification of E_0-semigroups by product systems. Mem. Am. Math. Soc. **240**, 1137 (2016)
157. B. Solel, Representations of product systems over semigroups an dilations of commuting CP maps. J. Funct. Anal. **235**, 593–618 (2006)
158. B. Solel, Regular dilations of representations of product systems. Math. Proc. R. Ir. Acad. **108**, 89–110 (2008)
159. W.F. Stinespring, Positive functions on C*-algebras. Proc. Am. Math. Soc. **6**, 211–216 (1955)
160. J. Stochel, F.H. Szafraniec, Unitary dilation of several contractions, in *Operator Theory: Advances and Applications*, vol. 127 (Birkhäuser, Basel, 2001), pp. 585–598
161. E. Stroescu, Isometric dilations of contractions on Banach spaces. Pac. J. Math. **47**, 257–262 (1973)
162. F.H. Szafraniec, Murphy's Positive definite kernels and Hilbert C*-modules sreorganized. Banach Center Publ. **89**, 275–295 (2010)
163. B. Sz.-Nagy, Sur les contractions de l'espace de Hilbert. Acta Sci. Math. **15**, 87–92 (1953)
164. B. Sz.-Nagy, Extensions of linear transformations in Hilbert space which extend beyond this space. Appendix to F. Riesz and B. Sz.-Nagy, Functional Analysis, Ungar, New York, 1960. Translation of "Prolongements des transformations de l'espace de Hilbert qui sortent de cet espace", Budapest (1955)
165. N.T. Varopoulos, On an inequality of von Neumann and an application of the metric theory of tensor products to operators theory. J. Funct. Anal. **16**, 83–100 (1974)
166. A. Vernik, Dilations of CP-maps commuting according to a graph. Houston J. Math. **42**, 1291–1329 (2016)
167. F. vom Ende, G. Dirr, Unitary dilations of discrete-time quantum-dynamical semigroups. J. Math. Phys. **60**, 1–17 (2019)

168. J. von Neumann, Eine Spektraltheorie für allgemeine Operatoren eines unitären Raumes. Math. Nachr. **4**, 258–281 (1951)
169. B. Xhabli, Universal operator system structures on ordered spaces and their applications. Thesis (Ph.D.) University of Houston (2009)
170. B. Xhabli, The super operator system structures and their applications in quantum entanglement theory. J. Funct. Anal. **262**, 1466–1497 (2012)
171. A. Zalar, Operator positivestellensatze for noncommutative polynomials positive on matrix convex sets. J. Math. Anal. Appl. **445**, 32–80 (2017)

Riesz-Fischer Maps, Semi-frames and Frames in Rigged Hilbert Spaces

Francesco Tschinke

Abstract In this note we present a review, some considerations and new results about maps with values in a distribution space and domain in a σ-finite measure space X. Namely, this is a survey about Bessel maps, frames and bases (in particular Riesz and Gel'fand bases) in a distribution space. In this setting, the Riesz-Fischer maps and semi-frames are defined and new results about them are obtained. Some examples in tempered distributions space are examined.

Keywords Frames · Bases · Distributions · Rigged Hilbert space

Mathematics Subject Classification (2010) Primary 42C15; Secondary 47A70, 46F05

1 Introduction

Given a Hilbert space $\mathcal{H}$ with inner product $\langle\cdot|\cdot\rangle$ and norm $\|\cdot\|$, a frame is a sequence of vectors $\{f_n\}$ in $\mathcal{H}$ if there exist $A, B > 0$ such that:

$$A\|f\|^2 \leq \sum_{k=1}^{\infty} |\langle f\,|\,f_n\rangle|^2 \leq B\|f\|^2, \quad \forall f \in \mathcal{H}.$$

As known, this notion generalizes orthonormal bases, and has reached an increasing level of popularity in many fields of interests, such as signal theory, image processing, etc., but it is also an important tool in pure mathematics: in fact it plays key roles in wavelet theory, time-frequency analysis, the theory of shift-invariant spaces, sampling theory and many other areas (see [10, 11, 19, 20, 24]).

F. Tschinke (✉)
Università di Palermo, Palermo, Italy
e-mail: francesco.tschinke@unipa.it

M. A. Bastos et al. (eds.), *Operator Theory, Functional Analysis and Applications*,
Operator Theory: Advances and Applications 282,
https://doi.org/10.1007/978-3-030-51945-2_29

A generalization of frame, the *continuous* frame, was proposed by Kaiser [24] and by Alí et al. [1, 2]: if (X, μ) is a measure space with μ as σ-finite positive measure, a function $F : x \in X \mapsto F_x \in \mathcal{H}$ is a continuous frame with respect to (X, μ) if:

(i) F is weakly measurable, i.e. the map $x \in X \mapsto \langle f|F_x\rangle \in \mathbb{C}$ is μ-measurable for all $f \in \mathcal{H}$;
(ii) there exist $A, B > 0$ such that, for all $f \in \mathcal{H}$:

$$A\|f\|^2 \leq \int_X |\langle f|F_x\rangle|^2 d\mu \leq B\|f\|^2, \quad \forall f \in \mathcal{H}.$$

Today, the notion of continuous frames in Hilbert spaces and their link with the theory of coherent states is well-known in the literature.

With the collaboration of C. Trapani and T. Triolo [29], the author introduced bases and frames in distributional spaces. To illustrate the motivations for this study, we have to consider the *rigged Hilbert space* (or Gel'fand triple) [16, 17]: that is, if $\mathcal{H}$ is a Hilbert space, the triple:

$$\mathcal{D}[t] \subset \mathcal{H} \subset \mathcal{D}^\times[t^\times],$$

where $\mathcal{D}[t]$ is a dense subspace of $\mathcal{H}$ endowed with a locally convex topology t stronger than the Hilbert norm and $\mathcal{D}^\times[t^\times]$ is the conjugate dual space of $\mathcal{D}$ endowed with the strong dual topology $t^\times$. If $\mathcal{D}$ is reflexive, the inclusions are dense and continuous.

In this setting, let us consider the *generalized eigenvectors* of an operator, i.e. eigenvectors that do not belong to $\mathcal{H}$. More precisely: if A is an essentially self-adjoint operator in $\mathcal{D}$ which maps $\mathcal{D}[t]$ into $\mathcal{D}[t]$ continuously, then A has a continuous extension $\hat{A}$ given by its adjoint, (i.e. $\hat{A} = A^\dagger$) from $\mathcal{D}^\times$ into itself. A *generalized eigenvector* of A, with eigenvalue $\lambda \in \mathbb{C}$, is an eigenvector of $\hat{A}$; that is, a conjugate linear functional $\omega_\lambda \in \mathcal{D}^\times$ such that:

$$\langle Af|\omega_\lambda\rangle = \lambda \langle f|\omega_\lambda\rangle, \quad \forall f \in \mathcal{D}.$$

The above equality can be read as $\hat{A}\omega_\lambda = A^\dagger\omega_\lambda = \lambda\omega_\lambda$.

A simple and explicative example is given by the derivative operator: $A := i\frac{d}{dx} : \mathcal{S}(\mathbb{R}) \to \mathcal{S}(\mathbb{R})$ where $\mathcal{S}(\mathbb{R})$ is the Schwartz space (i.e. infinitely differentiable rapidly decreasing functions). The rigged Hilbert space is:

$$\mathcal{S}(\mathbb{R}) \subset L^2(\mathbb{R}) \subset \mathcal{S}^\times(\mathbb{R}), \tag{1.1}$$

where the set $\mathcal{S}^\times(\mathbb{R})$ is known as space of *tempered distributions*. Then $\omega_\lambda(x) = \frac{1}{\sqrt{2\pi}}e^{-i\lambda x}$ -that does not belong to $L^2(\mathbb{R})$—is a generalized eigenvector of A with λ as eigenvalue.

Each function ω_λ can be viewed as a regular distribution of $\mathcal{S}^\times(\mathbb{R})$ through the following integral representation:

$$\langle \phi | \omega_\lambda \rangle = \int_{\mathbb{R}} \phi(x)\overline{\omega_\lambda(x)}dx = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \phi(x)e^{i\lambda x}dx := \check{\phi}(\lambda)$$

and the linear functional $\phi \mapsto \check{\phi}(\lambda)$ defined on $\mathcal{S}(\mathbb{R})$ is continuous. Furthermore by the Fourier-Plancherel theorem, one has:

$$\|\check{\phi}\|_2^2 = \int_{\mathbb{R}} |\langle \phi | \omega_\lambda \rangle|^2 dx = \|\phi\|_2^2.$$

With a limiting procedure, the Fourier transform can be extended to $L^2(\mathbb{R})$. Since a function $f \in L^2(\mathbb{R})$ defines a regular tempered distribution, we have, for all $\phi \in \mathcal{S}(\mathbb{R})$:

$$\begin{aligned} \langle \phi | f \rangle :&= \int_{\mathbb{R}} f(x)\phi(x)dx = \int_{\mathbb{R}} \left(\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{f}(\lambda)e^{i\lambda x}d\lambda \right) \phi(x)dx \\ &= \int_{\mathbb{R}} \hat{f}(\lambda)\check{\phi}(\lambda)d\lambda = \int_{\mathbb{R}} \hat{f}(\lambda)\langle \phi | \omega_\lambda \rangle \, d\lambda. \end{aligned}$$

That is:

$$f = \int_{\mathbb{R}} \hat{f}(\lambda)\omega_\lambda d\lambda. \tag{1.2}$$

in weak sense. The family $\{\omega_\lambda;\ \lambda \in \mathbb{R}\}$ of the previous example can be considered as the range of a weakly measurable function $\omega : \mathbb{R} \to \mathcal{S}^\times(\mathbb{R})$ which allows a representation as in (1.2) of any $f \in L^2(\mathbb{R})$ in terms of generalized eigenvectors of A. This is an example of a *distribution basis*. More precisely, since the Fourier-Plancherel theorem corresponds to the Parseval identity, this is an example of *Gel'fand distribution basis* [29, Subsec. 3.4], that plays, in $\mathcal{S}^\times(\mathbb{R})$, the role of an orthonormal basis in a Hilbert space.

The example above is a particular case of the Gel'fand-Maurin theorem (see [16, 18] for details), which states that, if $\mathcal{D}$ is a domain in a Hilbert space $\mathcal{H}$ which is a nuclear space under a certain topology τ, and A is an essentially self-adjoint operator on $\mathcal{D}$ which maps $\mathcal{D}[t]$ into $\mathcal{D}[t]$ continuously, then A admits a *complete set of generalized eigenvectors*.

If $\sigma(\overline{A})$ is the spectrum of the closure of the operator A, the completeness of the set $\{\omega_\lambda; \lambda \in \sigma(\overline{A})\}$ is understood in the sense that the Parseval identity holds, that is:

$$\|f\| = \left(\int_{\sigma(\overline{A})} |\langle f | \omega_\lambda \rangle|^2 d\lambda \right)^{1/2}, \quad \forall f \in \mathcal{D}. \tag{1.3}$$

To each λ there corresponds the subspace $\mathcal{D}_\lambda^\times \subset \mathcal{D}^\times$ of all generalized eigenvectors whose eigenvalue is λ. For all $f \in \mathcal{D}$ it is possible to define a linear functional $\tilde{f}_\lambda$ on $\mathcal{D}_\lambda^\times$ by $\tilde{f}_\lambda(\omega_\lambda) := \langle \omega_\lambda | f \rangle$ for all $\omega_\lambda \in \mathcal{D}_\lambda^\times$. Following [16, 17], the correspondence $\mathcal{D} \to \mathcal{D}_\lambda^{\times\times}$ defined by $f \mapsto \tilde{f}_\lambda$ is called the *spectral decomposition of the element f corresponding to A*. If $\tilde{f}_\lambda \equiv 0$ implies $f = 0$ (i.e. the map $f \mapsto \tilde{f}_\lambda$ is injective) then A is said to have *a complete system of generalized eigenvectors*.

The completeness and condition (1.3) may be interpreted as a kind of *orthonormality* of the ω_λ's: the family $\{\omega_\lambda\}_{\lambda \in \sigma(\overline{A})}$ in [29] is called a *Gel'fand basis*.

Another meaningful situation comes from quantum mechanics. Let us consider the rigged Hilbert space (1.1) corresponding to the one-dimensional case. The Hamiltonian operator H is an essentially self-adjoint operator on $\mathcal{S}(\mathbb{R})$, with self-adjoint extension $\overline{H}$ on the domain $\mathcal{D}(\overline{H})$, dense in $L^2(\mathbb{R})$. According to the *spectral expansion theorem* in the case of non-degenerate spectrum, for all $f \in L^2(\mathbb{R})$, the following decomposition holds:

$$f = \sum_{n \in J} c_n u_n + \int_{\sigma_c} c(\alpha) u_\alpha d\mu(\alpha).$$

The set $\{u_n\}_{n \in J}$, $J \subset \mathbb{N}$, is an orthonormal system of eigenvectors of H; the measure μ is a continuous measure on $\sigma_c \subset \mathbb{R}$ and $\{u_\alpha\}_{\alpha \in \sigma_c}$ are *generalized eigenvectors* of H in $\mathcal{S}^\times(\mathbb{R})$. This decomposition is *unique*. Furthermore:

$$\|f\|^2 = \sum_{n \in J} |c_n|^2 + \int_{\sigma_c} |c(\alpha)|^2 d\mu(\alpha).$$

The subset σ_c, corresponding to the continuous spectrum, is a union of intervals of $\mathbb{R}$, i.e. the index α is continuous. The generalized eigenvectors u_α are distributions: they do not belong to $L^2(\mathbb{R})$, therefore the "orthonormality" between generalized eigenvectors is not defined. Nevertheless, it is often denoted by the physicists with the Dirac delta: "$\langle u_\alpha | u_{\alpha'} \rangle$"$=\delta_{\alpha - \alpha'}$.

Frames, semi-frames, Riesz bases, etc. are families of sequences that generalize orthonormal bases in Hilbert space maintaining the possibility to *reconstruct* vectors of the space as superposition of more 'elementary' vectors renouncing often to the uniqueness of the representation, but gaining in versatility.

In this sense, they have been considered in literature in various spaces of functions and distributions: see for example the following (not exhaustive) list: [4, 8, 12, 14, 15, 25, 26].

It is remarkable that in a separable Hilbert space, orthonormal bases and Riesz bases are countable and notions corresponding to Riesz basis have been formulated in the continuous setting, but it is known that they exist only if the space given by the index set is discrete [7, 22, 23]. On the other hand, in the distributions and rigged Hilbert space setting the corresponding objects can be continuous.

Revisiting some results of [29] about Bessel maps, frames and (Gel'fand and Riesz) bases in distribution set-up, in this paper the notions of Riesz-Fischer map and of semi-frames in a space of distributions are proposed.

After some preliminaries and notations (Sect. 2), in Sect. 3 distribution Bessel maps are considered and the notion of distribution Riesz-Fischer maps is proposed, showing some new results about them (such as bounds and duality properties). Since distribution Bessel maps are not, in general, bounded by a Hilbert norm, we consider appropriate to define in Sect. 4 the *distribution semi-frames*, notion already introduced in a Hilbert space by J.-P. Antoine and P. Balasz [3]. Finally, distribution frames, distribution bases, Gel'fand and Riesz bases, considered in [29], are revisited in Sect. 5 with some additional examples.

2 Preliminary Definitions and Facts

2.1 Rigged Hilbert Space

Let $\mathcal{D}$ be a dense subspace of a Hilbert space $\mathcal{H}$ endowed with a locally convex topology t finer than the topology induced by a Hilbert norm. Denote as $\mathcal{D}^\times$ the vector space of all continuous conjugate linear functionals on $\mathcal{D}[t]$, i.e., the conjugate dual of $\mathcal{D}[t]$, endowed with the *strong dual topology* $t^\times = \beta(\mathcal{D}^\times, \mathcal{D})$, which can be defined by the seminorms:

$$q_{\mathcal{M}}(F) = \sup_{g \in \mathcal{M}} |\langle F|g\rangle|, \quad F \in \mathcal{D}^\times,$$

where $\mathcal{M}$ is a bounded subset of $\mathcal{D}[t]$. In this way, a *rigged Hilbert space* is defined in a standard fashion:

$$\mathcal{D}[t] \hookrightarrow \mathcal{H} \hookrightarrow \mathcal{D}^\times[t^\times], \tag{2.1}$$

where $\hookrightarrow$ denotes a continuous injection. Since the Hilbert space $\mathcal{H}$ can be identified with a subspace of $\mathcal{D}^\times[t^\times]$, we will systematically read (2.1) as a chain of topological inclusions: $\mathcal{D}[t] \subset \mathcal{H} \subset \mathcal{D}^\times[t^\times]$. These identifications imply that the sesquilinear form $B(\cdot,\cdot)$ which puts $\mathcal{D}$ and $\mathcal{D}^\times$ in duality is an extension of the inner product of $\mathcal{H}$; i.e. $B(\xi,\eta) = \langle\xi|\eta\rangle$, for every $\xi, \eta \in \mathcal{D}$ (to simplify notations we adopt the symbol $\langle\cdot|\cdot\rangle$ for both of them) and also the embedding map $I_{\mathcal{D},\mathcal{D}^\times} : \mathcal{D} \to \mathcal{D}^\times$ can be taken to act on $\mathcal{D}$ as $I_{\mathcal{D},\mathcal{D}^\times} f = f$ for every $f \in \mathcal{D}$. For more insights, besides to [16, 17], see also [21]. In this paper, if is not otherwise specified, we will work with a rigged Hilbert space $\mathcal{D}[t] \hookrightarrow \mathcal{H} \hookrightarrow \mathcal{D}^\times[t^\times]$ with $\mathcal{D}[t]$ reflexive, in this way the embedding $\hookrightarrow$ is continuous and dense.

2.2 The Space $\mathcal{L}(\mathcal{D}, \mathcal{D}^\times)$

If $\mathcal{D}[t] \subset \mathcal{H} \subset \mathcal{D}^\times[t^\times]$ is a rigged Hilbert space, let us denote by $\mathcal{L}(\mathcal{D}, \mathcal{D}^\times)$ the vector space of all continuous linear maps from $\mathcal{D}[t]$ into $\mathcal{D}^\times[t^\times]$. If $\mathcal{D}[t]$ is barreled (e.g., reflexive), an involution $X \mapsto X^\dagger$ can be introduced in $\mathcal{L}(\mathcal{D}, \mathcal{D}^\times)$ by the identity:

$$\left\langle X^\dagger \eta | \xi \right\rangle = \overline{\langle X\xi | \eta \rangle}, \quad \forall \xi, \eta \in \mathcal{D}.$$

Hence, in this case, $\mathcal{L}(\mathcal{D}, \mathcal{D}^\times)$ is a †-invariant vector space. We also denote by $\mathcal{L}(\mathcal{D})$ the algebra of all continuous linear operators $Y : \mathcal{D}[t] \to \mathcal{D}[t]$ and by $\mathcal{L}(\mathcal{D}^\times)$ the algebra of all continuous linear operators $Z : \mathcal{D}^\times[t^\times] \to \mathcal{D}^\times[t^\times]$. If $\mathcal{D}[t]$ is reflexive, for every $Y \in \mathcal{L}(\mathcal{D})$ there exists a unique operator $Y^\times \in \mathcal{L}(\mathcal{D}^\times)$, the adjoint of Y, such that

$$\langle \Phi | Y g \rangle = \left\langle Y^\times \Phi | g \right\rangle, \quad \forall \Phi \in \mathcal{D}^\times, g \in \mathcal{D}.$$

In similar way an operator $Z \in \mathcal{L}(\mathcal{D}^\times)$ has an adjoint $Z^\times \in \mathcal{L}(\mathcal{D})$ such that $(Z^\times)^\times = Z$. In the monograph [5] the topic is treated more deeply.

2.3 Weakly Measurable Maps

In this paper a *weakly measurable map* is considered as a subset of $\mathcal{D}^\times$: if (X, μ) is a measure space with μ a σ-finite positive measure, $\omega : x \in X \mapsto \omega_x \in \mathcal{D}^\times$ is a weakly measurable map if, for every $f \in \mathcal{D}$, the complex valued function $x \mapsto \langle f | \omega_x \rangle$ is μ-measurable. Since the form which puts $\mathcal{D}$ and $\mathcal{D}^\times$ in conjugate duality is an extension of the inner product of $\mathcal{H}$, we write $\langle f | \omega_x \rangle$ for $\overline{\langle \omega_x | f \rangle}$, $f \in \mathcal{D}$. If not otherwise specified, throughout the paper we will work with a measure space (X, μ) above described.

Definition 2.1 Let $\mathcal{D}[t]$ be a locally convex space, $\mathcal{D}^\times$ its conjugate dual and $\omega : x \in X \to \omega_x \in \mathcal{D}^\times$ a weakly measurable map, then:

(i) ω is *total* if, $f \in \mathcal{D}$ and $\langle f | \omega_x \rangle = 0$ μ-a.e. $x \in X$ implies $f = 0$;
(ii) ω is μ-*independent* if the unique measurable function ξ on (X, μ) such that, if $\int_X \xi(x) \langle g | \omega_x \rangle \, d\mu = 0$ for every $g \in \mathcal{D}$, then $\xi(x) = 0$ μ-a.e.

3 Bessel and Riesz-Fischer Distribution Maps

3.1 Bessel Distribution Maps

Definition 3.1 Let $\mathcal{D}[t]$ be a locally convex space. A weakly measurable map ω is a *Bessel distribution map* (briefly: Bessel map) if for every $f \in \mathcal{D}$,

$$\int_X |\langle f|\omega_x\rangle|^2 d\mu < \infty.$$

The following Proposition is the analogue of Proposition 2 and Theorem 3 in [30, Section 2, Chapter 4].

Proposition 3.2 ([29, Proposition 3.1]) *If $\mathcal{D}[t]$ a Fréchet space, and $\omega : x \in X \mapsto \omega_x \in \mathcal{D}^\times$ a weakly measurable map. The following statements are equivalent:*

(i) *ω is a Bessel map;*
(ii) *there exists a continuous seminorm p on $\mathcal{D}[t]$ such that:*

$$\left(\int_X |\langle f|\omega_x\rangle|^2 d\mu\right)^{1/2} \leq p(f), \quad \forall f \in \mathcal{D};$$

(iii) *for every bounded subset $\mathcal{M}$ of $\mathcal{D}$ there exists $C_\mathcal{M} > 0$ such that:*

$$\sup_{f\in\mathcal{M}} \left| \int_X \xi(x)\, \langle \omega_x|f\rangle\, d\mu \right| \leq C_\mathcal{M} \|\xi\|_2, \quad \forall \xi \in L^2(X,\mu). \tag{3.1}$$

We have also the following:

Lemma 3.3 ([29, Lemma 3.4]) *If $\mathcal{D}$ is a Fréchet space and ω a Bessel distribution map, then:*

$$\int_X \langle f|\omega_x\rangle\, \omega_x d\mu$$

converges for every $f \in \mathcal{D}$ to an element of $\mathcal{D}^\times$. Moreover, the map

$$\mathcal{D} \ni f \mapsto \int_X \langle f|\omega_x\rangle\, \omega_x d\mu \in \mathcal{D}^\times$$

is continuous.

Let ω be a Bessel map: the previous lemma allows to define on $\mathcal{D} \times \mathcal{D}$ the sesquilinear form Ω:

$$\Omega(f,g) = \int_X \langle f|\omega_x\rangle\, \langle \omega_x|g\rangle\, d\mu.$$

By Proposition 3.2, one has for all $f, g \in \mathcal{D}$,

$$|\Omega(f,g)| = \left| \int_X \langle f|\omega_x\rangle \langle g|\omega_x\rangle \, d\mu \right| \leq \| \langle f|\omega_x\rangle \|_2 \, \| \langle g|\omega_x\rangle \|_2 \leq p(f)p(g).$$

This means that Ω is jointly continuous on $\mathcal{D}[t]$. Hence there exists an operator $S_\omega \in \mathcal{L}(\mathcal{D}, \mathcal{D}^\times)$, with $S_\omega = S_\omega^\dagger$, $S_\omega \geq 0$, such that:

$$\Omega(f,g) = \langle S_\omega f|g\rangle = \int_X \langle f|\omega_x\rangle \langle \omega_x|g\rangle \, d\mu, \quad \forall f, g \in \mathcal{D} \tag{3.2}$$

that is,

$$S_\omega f = \int_X \langle f|\omega_x\rangle \, \omega_x d\mu, \quad \forall f \in \mathcal{D}.$$

Since ω is a Bessel distribution map and $\xi \in L^2(X,\mu)$, we put for all $g \in \mathcal{D}$:

$$\Lambda_\omega^\xi(g) := \int_X \xi(x) \langle \omega_x|g\rangle \, d\mu. \tag{3.3}$$

Then Λ_ω^ξ is a continuous conjugate linear functional on $\mathcal{D}$, i.e. $\Lambda_\omega^\xi \in \mathcal{D}^\times$. We write:

$$\Lambda_\omega^\xi := \int_X \xi(x)\omega_x d\mu$$

in weak sense. Therefore we can define a linear map $T_\omega : L^2(X,\mu) \to \mathcal{D}^\times[t^\times]$, which will be called the *synthesis operator*, by:

$$T_\omega : \xi \mapsto \Lambda_\omega^\xi.$$

By (3.1), it follows that T_ω is continuous from $L^2(X,\mu)$, endowed with its natural norm, into $\mathcal{D}^\times[t^\times]$. Hence, it possesses a continuous adjoint $T_\omega^\times : \mathcal{D}[t] \to L^2(X,\mu)$, which is called the *analysis operator*, acting as follows:

$$T_\omega^\times : f \in \mathcal{D}[t] \mapsto \xi_f \in L^2(X,\mu), \text{ where } \xi_f(x) = \langle f|\omega_x\rangle, \ x \in X.$$

One has that $S_\omega = T_\omega T_\omega^\times$.

3.2 Riesz-Fischer Distribution Map

Definition 3.4 Let $\mathcal{D}[t]$ be a locally convex space. A weakly measurable map $\omega : x \in X \mapsto \omega_x \in \mathcal{D}^\times$ is called a *Riesz-Fischer distribution map* (briefly: Riesz-Fischer map) if, for every $h \in L^2(X, \mu)$, there exists $f \in \mathcal{D}$ such that:

$$\langle f|\omega_x\rangle = h(x) \quad \mu\text{-a.e.} \tag{3.4}$$

In this case, we say that f is a solution of equation $\langle f|\omega_x\rangle = h(x)$.

Clearly, if f_1 and f_2 are solutions of (3.4), then

$$f_1 - f_2 \in \omega^\perp := \{g \in \mathcal{D} : \langle g|\omega_x\rangle = 0, \quad \mu - a.e.\}.$$

If ω is total, the solution is unique. We prove the following:

Lemma 3.5 *Let $\mathcal{D}$ be a reflexive locally convex space, $h(x)$ be a measurable function and $x \in X \mapsto \omega_x \in \mathcal{D}^\times[t^\times]$ a weakly measurable map. Then the equation:*

$$\langle f|\omega_x\rangle = h(x)$$

admits a solution $f \in \mathcal{D}$ if, and only if, there exists a bounded subset $\mathcal{M}$ of $\mathcal{D}$ such that

$$|h(x)| \leq \sup_{f\in\mathcal{M}} |\langle f|\omega_x\rangle| \quad \mu\text{-a.e.}$$

Proof Necessity is obvious. Conversely, let $x \in X$ be a point such that $\langle f|\omega_x\rangle = h(x) \neq 0$. Let us consider the subspace V_x of $\mathcal{D}^\times$ given by $V_x := \{\alpha\omega_x\}_{\alpha\in\mathbb{C}}$, and let us define the functional μ on V_x by: $\mu(\alpha\omega_x) := \alpha h(x)$. We have that

$$|\mu(\alpha\omega_x)| = |\alpha h(x)| \leq |\alpha| \sup_{f\in\mathcal{M}} |\langle f|\omega_x\rangle| = \sup_{f\in\mathcal{M}} |\langle f|\alpha\omega_x\rangle|,$$

in other words:

$$|\mu(F_x)| \leq \sup_{f\in\mathcal{M}} |\langle f|F_x\rangle|$$

for all $F_x \in V_x$. By the Hahn-Banach theorem, there exists an extension $\tilde{\mu}$ to $\mathcal{D}^\times$ such that

$$|\tilde{\mu}(F)| \leq \sup_{f\in\mathcal{M}} |\langle f|F\rangle|,$$

for every $F \in \mathcal{D}^\times$. Since $\mathcal{D}$ is reflexive, there exists $\bar{f} \in \mathcal{D}$ such that $\tilde{\mu}(F) = \langle \bar{f} | F \rangle$. The statement follows from the fact that $\mu(\omega_x) = h(x)$. □

If M is a subspace of $\mathcal{D}$ and the topology of $\mathcal{D}$ is generated by the family of seminorms $\{p_\alpha\}_{\alpha \in I}$, then the topology on the quotient space $\mathcal{D}/M$ is defined, as usual, by the seminorms $\{\tilde{p}_\alpha\}_{\alpha \in I}$, where

$$\tilde{p}_\alpha(\tilde{f}) := \inf\{p_\alpha(g) : g \in f + M\}.$$

The following proposition can be compared to the case of Riesz-Fischer sequences: see [30, Chapter 4, Section 2, Proposition 2].

Proposition 3.6 *Assume that $\mathcal{D}[t]$ is a Fréchet space. If the map $\omega : x \in X \to \omega_x \in \mathcal{D}^\times$ is a Riesz-Fischer map, then for every continuous seminorm p on $\mathcal{D}$, there exists a constant $C > 0$ such that, for every solution f of (3.4),*

$$\tilde{p}(\tilde{f}) := \inf\{p(g) : g \in f + \omega^\perp\} \leq C \| \langle f | \omega_x \rangle \|_2.$$

Proof Since $\omega^\perp$ is closed, it follows that the quotient $\mathcal{D}/\omega^\perp := \mathcal{D}_{\omega^\perp}$ is a Fréchet space. If $f \in \mathcal{D}$, we put $\tilde{f} := f + \omega^\perp$. Let $h \in L^2(X, \mu)$ and f a solution of (3.4) corresponding to h; then, we can define an operator $S : L^2(X, \mu) \to \mathcal{D}_{\omega^\perp}$ by $h \mapsto \tilde{f}$. Let us consider a sequence $h_n \in L^2(X, \mu)$ such that $h_n \to 0$ and, for each $n \in \mathbb{N}$, let f_n be a corresponding solution of (3.4). One has that

$$\int_X |h_n(x)|^2 d\mu \to 0, \quad \text{i.e.} \quad \int_X |\langle f_n | \omega_x \rangle|^2 d\mu \to 0.$$

This implies that $\langle f_n | \omega_x \rangle \to 0$ in measure, so there exists a subsequence such that $\langle f_{n_k} | \omega_x \rangle \to 0$ a.e. (see [13]). On the other hand, if $Sh_n = \tilde{f}_n$ is a sequence convergent to $\tilde{f}$ in $\mathcal{D}_{\omega^\perp}$ w.r. to the quotient topology defined by the seminorms $\tilde{p}(\cdot)$, it follows that the sequence is convergent in the *weak* topology of $\mathcal{D}_{\omega^\perp}$, i.e.:

$$\left\langle \tilde{f}_n | \tilde{F} \right\rangle \to \left\langle \tilde{f} | \tilde{F} \right\rangle \quad \forall \tilde{F} \in \mathcal{D}^\times_{\omega^\perp}.$$

Let us consider the canonical surjection $\rho : \mathcal{D} \to \mathcal{D}_{\omega^\perp}$, $\rho : f \mapsto \tilde{f} = f + \omega^\perp$. Its transpose map (adjoint) $\rho^\dagger : \mathcal{D}^\times_{\omega^\perp} \to \mathcal{D}^\times$ is injective (see [21], p. 263) and $\rho^\dagger[\mathcal{D}^\times_{\omega^\perp}] = \omega^{\perp\perp}$. Then $\rho^\dagger : \mathcal{D}^\times_{\omega^\perp} \to \omega^{\perp\perp}$ is invertible. Hence,

$$\left\langle \tilde{f}_n | \tilde{F} \right\rangle = \left\langle \rho(f_n) | (\rho^\dagger)^{-1}(F) \right\rangle = \left\langle f_n | \rho^\dagger((\rho^\dagger)^{-1}(F)) \right\rangle = \langle f_n | F \rangle, \quad \forall F \in \omega^{\perp\perp}.$$

Thus, if $\tilde{f}_n \to \tilde{f}$ in the topology of $\mathcal{D}_{\omega^\perp}$, then $\langle f_n | F \rangle \to \langle f | F \rangle$, for all $F \in \omega^{\perp\perp}$, and, in particular, since $\omega \subset \omega^{\perp\perp}$, one has $\langle f_n | \omega_x \rangle \to \langle f | \omega_x \rangle$. Since $\langle f_n | \omega_x \rangle$ has a subsequence convergent to 0, one has $f \in \omega^\perp$. From the closed graph theorem, it follows that the map S is continuous, i.e. for all continuous seminorms $\tilde{p}$ on $\mathcal{D}_{\omega^\perp}$

there exists $C > 0$ such that: $\tilde{p}(Sh) \leq C\|h\|_2$, for all $h \in L^2(X, \mu)$. The statement follows from the definition of Riesz-Fischer map. □

Corollary 3.7 *Assume that $\mathcal{D}[t]$ is a Fréchet space. If the map $\omega : x \in X \to \omega_x \in \mathcal{D}^\times$ is a total Riesz-Fischer map, then for every continuous seminorm p on $\mathcal{D}$, there exists a constant $C > 0$ such that, for the solution f of (3.4),*

$$p(f) \leq C\| \langle f|\omega_x\rangle \|_2.$$

Remark 3.8 For an arbitrary weakly measurable map ω, we define the subset of $\mathcal{D}[t]$: $D(V_\omega) := \{f \in \mathcal{D} : \langle f|\omega_x\rangle \in L^2(X, \mu)\}$ and the analysis operator $V_\omega : f \in D(V_\omega) \mapsto \langle f|\omega_x\rangle \in L^2(X, \mu)$. Clearly, ω is a Riesz-Fischer map if and only if $V_\omega : D(V_\omega) \to L^2(X, \mu)$ is surjective. If ω is total, it is injective too, so V_ω is invertible. A consequence of Corollary 3.7 is that $V_\omega^{-1} : L^2(X, \mu) \to D(V_\omega)$ is continuous.

3.3 Duality

Definition 3.9 Let $\mathcal{D} \hookrightarrow \mathcal{H} \hookrightarrow \mathcal{D}^\times$ be a rigged Hilbert space and ω a weakly measurable map. We call *dual map of* ω, if it exists, a weakly measurable map θ such that for all $f, g \in \mathcal{D}$:

$$\left| \int_X \langle f|\theta_x\rangle \langle \omega_x|g\rangle \, d\mu \right| < \infty$$

and

$$\langle f|g\rangle = \int_X \langle f|\theta_x\rangle \langle \omega_x|g\rangle \, d\mu, \quad \forall f, g \in \mathcal{D}.$$

Proposition 3.10 *Suppose that ω is a Riesz-Fischer map. Then the map θ is a Bessel map.*

Proof For all $h \in L^2(X, \mu)$ there exists $\bar{f} \in \mathcal{D}$ such that $\langle \bar{f}|\omega_x\rangle = h(x)$ μ-a.e. Since θ is a dual map, one has that:

$$\left| \int_X h(x) \langle \theta_x|g\rangle \, d\mu \right| < \infty$$

for all $h \in L^2(X, \mu)$. It follows that $\langle \theta_x|g\rangle \in L^2(X, \mu)$ (see [28, Chapter 6, Exercise 4]). □

Proposition 3.11 *Let $\mathcal{D}$ be reflexive and let ω be a μ-independent Bessel map. Furthermore, suppose that for all $h \in L^2(X, \mu)$ there exists a bounded subset $\mathcal{M} \subset \mathcal{D}$ such that:*

$$\left| \int_X h(x) \, \langle \omega_x | g \rangle \, d\mu \right| \leq \sup_{f \in \mathcal{M}} | \langle f | g \rangle |, \quad \forall g \in \mathcal{D},$$

then the dual map θ is a Riesz-Fischer map.

Proof Since $h \in L^2(X, \mu)$, and since ω is a Bessel map, one has:

$$\left| \int_X h(x) \, \langle \omega_x | g \rangle \, d\mu \right| < \infty.$$

Let us consider

$$g = \int_X \langle g | \omega_x \rangle \, \theta_x d\mu$$

as element of $\mathcal{D}^\times$. We define the following functional on $\mathcal{D}$ (as subspace of $\mathcal{D}^\times$):

$$\mu(g) := \int_X h(x) \, \langle \omega_x | g \rangle \, d\mu.$$

By hypothesis, one has:

$$|\mu(g)| \leq \sup_{f \in \mathcal{M}} | \langle f | g \rangle |.$$

By the Hahn-Banach theorem, there exists an extension $\tilde{\mu}$ to $\mathcal{D}^\times$ such that:

$$|\tilde{\mu}(G)| \leq \sup_{f \in \mathcal{M}} | \langle f | G \rangle |, \quad \forall G \in \mathcal{D}^\times.$$

Since $\mathcal{D}$ is reflexive, there exists $\tilde{f} \in \mathcal{D}^{\times\times} = \mathcal{D}$ such that $\tilde{\mu}(G) = \left\langle \tilde{f} | G \right\rangle$. In particular

$$\left\langle \tilde{f} | g \right\rangle = \int_X h(x) \, \langle \omega_x | g \rangle \, d\mu.$$

Since θ is dual of ω, we have too:

$$\left\langle \tilde{f} | g \right\rangle = \int_X \left\langle \tilde{f} | \theta_x \right\rangle \langle \omega_x | g \rangle \, d\mu.$$

But ω is μ-independent, then it follows that $h(x) = \left\langle \tilde{f} | \theta_x \right\rangle$ μ-a.e. □

4 Semi-frames and Frames

4.1 Distribution Semi-frames

Definition 4.1 Given a rigged Hilbert space $\mathcal{D} \hookrightarrow \mathcal{H} \hookrightarrow \mathcal{D}^\times$, a Bessel map ω is a *distribution upper semi-frame* if it is complete (total) and if there exists $B > 0$:

$$0 < \int_X |\langle f|\omega_x\rangle|^2 d\mu \leq B\|f\|^2, \quad \forall f \in \mathcal{D}, \quad f \neq 0.$$

Since the injection $\mathcal{D} \hookrightarrow \mathcal{H}$ is continuous, it follows that there exists a continuous seminorm p on $\mathcal{D}$ such that $\|f\| \leq p(f)$ for all $f \in \mathcal{D}$. If $\xi \in L^2(X, \mu)$, then the continuous conjugate functional Λ_ω^ξ on $\mathcal{D}$ defined in (3.3) is bounded in $\mathcal{D}[\|\cdot\|]$; it follows that it has a bounded extension $\tilde{\Lambda}_\omega^\xi$ to $\mathcal{H}$, defined, as usual, by a limiting procedure. Therefore, there exists a unique vector $h_\xi \in \mathcal{H}$ such that:

$$\tilde{\Lambda}_\omega^\xi(g) = \langle h_\xi|g\rangle, \quad \forall g \in \mathcal{H}.$$

This implies that the synthesis operator T_ω takes values in $\mathcal{H}$, it is bounded and $\|T_\omega\| \leq B^{1/2}$; its hilbertian adjoint $C_\omega := T_\omega^*$ extends the analysis operator $T_\omega^\times$.

The action of C_ω can be easily described: if $g \in \mathcal{H}$ and $\{g_n\}$ is a sequence of elements of $\mathcal{D}$, norm converging to g, then the sequence $\{\eta_n\}$, where $\eta_n(x) = \langle g_n|\omega_x\rangle$, is convergent in $L^2(X, \mu)$. Put $\eta = \lim_{n\to\infty} \eta_n$. Then,

$$\langle T_\omega \xi|g\rangle = \lim_{n\to\infty} \int_X \xi(x)\,\langle \omega_x|g_n\rangle\, d\mu = \int_X \xi(x)\overline{\eta(x)} d\mu.$$

Hence $T_\omega^* g = \eta$.

The function $\eta \in L^2(X, \mu)$ depends linearly on g, for each $x \in X$. Thus we can define a linear functional $\check{\omega}_x$ by

$$\langle g|\check{\omega}_x\rangle = \lim_{n\to\infty} \langle g_n|\omega_x\rangle, \quad g \in \mathcal{H};\ g_n \to g. \tag{4.1}$$

Of course, for each $x \in X$, $\check{\omega}_x$ extends ω_x; however $\check{\omega}_x$ need not be continuous, as a functional on $\mathcal{H}$. We conclude that:

$$T_\omega^* : g \mapsto \langle g|\check{\omega}_x\rangle \in L^2(X, \mu).$$

Moreover, in this case, the sesquilinear form Ω in (3.2), which is well defined on $\mathcal{D} \times \mathcal{D}$, is bounded with respect to $\|\cdot\|$ and possesses a bounded extension $\hat{\Omega}$ to $\mathcal{H}$. Hence there exists a bounded operator $\hat{S}_\omega$ in $\mathcal{H}$, such that:

$$\hat{\Omega}(f, g) = \langle \hat{S}_\omega f|g\rangle, \quad \forall f, g \in \mathcal{H}. \tag{4.2}$$

Since

$$\left\langle \hat{S}_\omega f | g \right\rangle = \int_X \langle f|\omega_x\rangle \langle \omega_x|g\rangle \, d\mu, \quad \forall f, g \in \mathcal{D},$$

$\hat{S}_\omega$ extends the frame operator S_ω and $S_\omega : \mathcal{D} \to \mathcal{H}$. It is easily seen that $\hat{S}_\omega = \hat{S}_\omega^*$ and $\hat{S}_\omega = T_\omega T_\omega^*$. By definition, we have:

$$0 < \|\hat{S}_\omega f\| \leq B\|f\|, \quad \forall f \in \mathcal{H}, \quad f \neq 0.$$

Then $\hat{S}_\omega$ is bounded, self-adjoint and injective too. This means that $\mathsf{Ran}\, S_\omega$ is dense in $\mathcal{H}$, and $\hat{S}_\omega^{-1}$ is densely defined. If ω is not a frame, $\hat{S}_\omega^{-1}$ is an unbounded, self-adjoint operator (see [3]).

Remark 4.2 If $\{\omega_x\}_{x\in X}$ is an upper semi-frame, then there exists a continuous seminorm p on $\mathcal{D}$ such that $\| \langle f|\omega_x\rangle \|_2 \leq p(f)$ for all $f \in \mathcal{D}$. In fact, the injection $\mathcal{D} \hookrightarrow \mathcal{H}$ is continuous, i.e. $\|f\| \leq p(f)$ for all $f \in \mathcal{D}$. The converse is not true: let us consider the rigged Hilbert space $\mathcal{S}(\mathbb{R}) \hookrightarrow L^2(\mathbb{R}) \hookrightarrow \mathcal{S}^\times(\mathbb{R})$; the system of derivative of Dirac's deltas $\{\delta'_x\}_{x\in\mathbb{R}}$ is total. Since $\mathcal{S}(\mathbb{R})$ is a Fréchet space, (ii) of Proposition 3.2 it holds. However $\{\delta'_x\}_{x\in\mathbb{R}}$ is not a distribution upper semi-frame; in fact:

$$\int_{\mathbb{R}} |\left\langle \phi|\delta'_x\right\rangle|^2 dx = \|\phi'\|_2^2 \quad \forall \phi \in \mathcal{S}(\mathbb{R}),$$

but the derivative operator $\frac{d}{dx} : \mathcal{S}(\mathbb{R}) \to L^2(\mathbb{R})$ is unbounded (clearly with respect to the topology of the Hilbert norm).

Remark 4.3 In [29] it is defined the notion of *bounded Bessel map*, that is a Bessel map in rigged Hilbert space such that:

$$\int_X |\langle f|\omega_x\rangle|^2 d\mu \leq B\|f\|^2, \quad \forall f \in \mathcal{D}.$$

It is a more general notion than upper bounded semi-frame. In fact, we can consider, as example, the distribution $\omega_x := \eta_K(x)\delta_x$ where $\eta_K(x)$ is a C^∞-function with compact support K and $M := \max_{x\in K} |\eta_K(x)|$:

$$\begin{aligned}\int_{\mathbb{R}} |\langle \phi|\omega_x\rangle|^2 dx &= \int_{\mathbb{R}} |\langle \phi|\eta_K(x)\delta_x\rangle|^2 dx \\ &= \int_{\mathbb{R}} |\overline{\eta_K(x)}\phi(x)|^2 dx \leq M^2 \int_K |\phi(x)|^2 dx \leq M^2\|\phi\|_2^2.\end{aligned}$$

Therefore ω is a bounded Bessel map, but it is not total, then it is not an upper semi-frame.

Definition 4.4 Given a rigged Hilbert space $\mathcal{D} \hookrightarrow \mathcal{H} \hookrightarrow \mathcal{D}^\times$, a Bessel map ω is a *distribution lower semi-frame* if there exists $A > 0$ such that:

$$A\|f\|^2 \leq \int_X |\langle f|\omega_x\rangle|^2 d\mu, \quad \forall f \in \mathcal{D}.$$

By definition, it follows that ω is total. If $\mathcal{D}$ is a Fréchet space, by Proposition 3.2 one has $S_\omega \in \mathcal{L}(\mathcal{D}, \mathcal{D}^\times)$ and, if ω is not a frame, S_ω is unbounded. Furthermore, S_ω is injective, and S_ω^{-1} is bounded.

Example Let us consider the space $\mathcal{O}_M$, known (see [27]) as the set of infinitely differentiable functions on $\mathbb{R}$ that are polynomially bounded together with their derivatives. Let us consider $g(x) \in \mathcal{O}_M$ such that $0 < m < |g(x)|$. If we define $\omega_x := g(x)\delta_x$, then $\{\omega_x\}_{x\in\mathbb{R}}$ is a distribution lower semi-frame with $A = m^2$.

The proof of the following lemma is analogous to that of [3, Lemma 2.5]:

Lemma 4.5 *Let ω be an upper semi-frame with upper frame bound M and θ a total family dual to ω. Then θ is a lower semi-frame, with lower frame bound M^{-1}.*

4.2 Distribution Frames

This section is devoted to distribution frames, with main results already shown in [29].

Definition 4.6 ([29, Definition 3.6]) Let $\mathcal{D}[t] \subset \mathcal{H} \subset \mathcal{D}^\times[t^\times]$ be a rigged Hilbert space, with $\mathcal{D}[t]$ a reflexive space and ω a Bessel map. We say that ω is a *distribution frame* if there exist $A, B > 0$ such that:

$$A\|f\|^2 \leq \int_X |\langle f|\omega_x\rangle|^2 d\mu \leq B\|f\|^2, \quad \forall f \in \mathcal{D}.$$

A distribution frame ω is clearly, in particular, an upper bounded semi-frame. Thus, we can consider the operator $\hat{S}_\omega$ defined in (4.2). It is easily seen that, in this case,

$$A\|f\| \leq \|\hat{S}_\omega f\| \leq B\|f\|, \quad \forall f \in \mathcal{H}.$$

This inequality, together with the fact that $\hat{S}_\omega$ is symmetric, implies that $\hat{S}_\omega$ has a bounded inverse $\hat{S}_\omega^{-1}$ everywhere defined in $\mathcal{H}$.

Remark 4.7 It is worth noticing that the fact that ω and S_ω extend to $\mathcal{H}$ does not mean that ω a frame in the Hilbert space $\mathcal{H}$, because we do not know if the extension of S_ω has the form of (3.2) with $f, g \in \mathcal{H}$.

To conclude this section, we recall a list of properties of frames proved in [29].

Lemma 4.8 ([29, Lemma 3.8]) *Let ω be a distribution frame. Then, there exists $R_\omega \in \mathcal{L}(\mathcal{D})$ such that $S_\omega R_\omega f = R_\omega^\times S_\omega f = f$, for every $f \in \mathcal{D}$.*

As a consequence, the reconstruction formulas for distribution frames hold for all $f \in \mathcal{D}$:

$$f = R_\omega^\times S_\omega f = \int_X \langle f|\omega_x\rangle \, R_\omega^\times \omega_x d\mu;$$
$$f = S_\omega R_\omega f = \int_X \langle R_\omega f|\omega_x\rangle \, \omega_x d\mu.$$

These representations have to be interpreted in the weak sense.

Remark 4.9 The operator R_ω acts as an inverse of S_ω. On the other hand the operator $\hat{S}_\omega$ has a bounded inverse $\hat{S}_\omega^{-1}$ everywhere defined in $\mathcal{H}$. It results that [29, Remark 3.7]: $\hat{S}_\omega^{-1}\mathcal{D} \subset \mathcal{D}$ and $R_\omega = \hat{S}_\omega^{-1} \restriction_{\mathcal{D}}$.

There exists the dual frame:

Proposition 4.10 ([29, Lemma 3.10]) *Let ω be a distribution frame. Then there exists a weakly measurable function θ such that:*

$$\langle f|g\rangle = \int_X \langle f|\theta_x\rangle \, \langle \omega_x|g\rangle \, d\mu, \quad \forall f, g \in \mathcal{D}.$$

Where $\theta_x := R_\omega^\times \omega_x$. The frame operator S_θ for θ is well defined and we have: $S_\theta = I_{\mathcal{D},\mathcal{D}^\times} R_\omega$.

The distribution function θ, constructed in Proposition 4.10, is also a distribution frame, called the *canonical dual frame* of ω. Indeed, it results that [29]:

$$B^{-1}\|f\|^2 \leq \langle S_\theta f|f\rangle \leq A^{-1}\|f\|^2, \quad \forall f \in \mathcal{D}.$$

4.3 Parseval Distribution Frames

Definition 4.11 If ω is a distribution frame, then we say that:

(a) ω is a *tight* distribution frame if we can choose $A = B$ as frame bounds. In this case, we usually refer to A as a frame bound for ω;
(b) ω is a *Parseval* distribution frame if $A = B = 1$ are frame bounds.

More explicitly a weakly measurable distribution function ω is called a *Parseval distribution frame* if [29, Definition 3.13]:

$$\int_X |\langle f|\omega_x\rangle|^2 d\mu = \|f\|^2, \quad f \in \mathcal{D}.$$

It is clear that a Parseval distribution frame is a frame in the sense of Definition 4.6 with $S_\omega = I_{\mathcal{D}}$, the identity operator of $\mathcal{D}$.

Lemma 4.12 ([29, Lemma 3.14]) *Let $\mathcal{D} \subset \mathcal{H} \subset \mathcal{D}^\times$ be a rigged Hilbert space and $\omega : x \in X \mapsto \omega_x \in \mathcal{D}^\times$ a weakly measurable map. The following statements are equivalent.*

(i) *ω is a Parseval distribution frame;*
(ii) $\langle f|g\rangle = \int_X \langle f|\omega_x\rangle \langle \omega_x|g\rangle \, d\mu, \quad \forall f, g \in \mathcal{D}$;
(iii) *$f = \int_X \langle f|\omega_x\rangle \, \omega_x d\mu$, the integral on the r.h.s. is understood as a continuous conjugate linear functional on $\mathcal{D}$, that is an element of $\mathcal{D}^\times$.*

The representation in (iii) of Lemma 4.12 is not necessarily unique.

5 Distribution Basis

Definition 5.1 ([29, Definition 2.3]) Let $\mathcal{D}[t]$ be a locally convex space, $\mathcal{D}^\times$ its conjugate dual and $\omega : x \in X \mapsto \omega_x \in \mathcal{D}^\times$ a weakly measurable map. Then ω is a *distribution basis* for $\mathcal{D}$ if, for every $f \in \mathcal{D}$, there exists a *unique* measurable function ξ_f such that:

$$\langle f|g\rangle = \int_X \xi_f(x) \, \langle \omega_x|g\rangle \, d\mu, \quad \forall f, g \in \mathcal{D}$$

and, for every $x \in X$, the linear functional $f \in \mathcal{D} \to \xi_f(x) \in \mathbb{C}$ is continuous in $\mathcal{D}[t]$.

The above formula can be represented by:

$$f = \int_X \xi_f(x) \omega_x d\mu$$

in weak sense.

Remark 5.2 Clearly, if ω is a distribution basis, then it is μ-independent. Furthermore, since $f \in \mathcal{D} \mapsto \xi_f(x) \in \mathbb{C}$ continuously, there exists a unique weakly μ-measurable map $\theta : X \to \mathcal{D}^\times$ such that: $\xi_f(x) = \langle f|\theta_x\rangle$ for every $f \in \mathcal{D}$. We call θ *dual* map of ω. If θ is μ-independent, then it is a distribution basis too.

5.1 Gel'fand Distribution Basis

The Gel'fand distribution basis, introduced in [29], is a good substitute for the notion of an *orthonormal basis* which is meaningless in the present framework.

Definition 5.3 A weakly measurable map ζ is *Gel'fand distribution basis* if it is a μ-independent Parseval distribution frame.

By definition and Lemma 4.12, this means that, for every $f \in \mathcal{D}$ there exists a unique function $\xi_f \in L^2(X, \mu)$ such that:

$$f = \int_X \xi_f(x)\zeta_x d\mu \tag{5.1}$$

with $\xi_f(x) = \langle f|\zeta_x\rangle$ μ-a.e. Furthermore $\|f\|^2 = \int_X |\langle f|\zeta_x\rangle|^2 d\mu$ and ζ is total too.

For every $x \in X$, the map $f \in \mathcal{H} \mapsto \xi_f(x) \in \mathbb{C}$ defines as in (4.1) a linear functional $\check{\zeta}_x$ on $\mathcal{H}$, then for all $f \in \mathcal{H}$:

$$f = \int_X \left\langle f|\check{\zeta}_x\right\rangle \zeta_x d\mu.$$

We have the following characterization result [29]:

Proposition 5.4 ([29, Proposition 3.15]) *Let $\mathcal{D} \subset \mathcal{H} \subset \mathcal{D}^\times$ be a rigged Hilbert space and let $\zeta : x \in X \mapsto \zeta_x \in \mathcal{D}^\times$ be a Bessel distribution map. Then the following statements are equivalent.*

(a) *ζ is a Gel'fand distribution basis.*
(b) *The synthesis operator T_ζ is an isometry of $L^2(X, \mu)$ onto $\mathcal{H}$.*

Example ([29, Example 3.17]) Given the rigged Hilbert space:

$$\mathcal{S}(\mathbb{R}) \hookrightarrow L^2(\mathbb{R}) \hookrightarrow \mathcal{S}^\times(\mathbb{R}),$$

for $x \in \mathbb{R}$ the function $\zeta_x(y) = \frac{1}{\sqrt{2\pi}} e^{-ixy}$, defines a (regular) tempered distribution: in fact, denoting as usual by $\hat{g}$, $\check{g}$, respectively, the Fourier transform and the inverse Fourier transform of $g \in L^2(\mathbb{R})$, one has that:

$$\mathcal{S}(\mathbb{R}) \ni \phi \mapsto \langle \phi|\zeta_x\rangle = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \phi(y) e^{-ixy} dy = \hat{\phi}(x) \in \mathbb{C}.$$

For all $x \in \mathbb{R}$ the set of functions $\zeta := \{\zeta_x(y)\}_{x\in\mathbb{R}}$ is a Gel'fand distribution basis, because the synthesis operator $T_\zeta : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ defined by:

$$(T_\zeta \xi)(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \xi(y) e^{-ixy} dy = \hat{\xi}(x), \quad \forall \xi \in L^2(\mathbb{R})$$

is an isometry onto $L^2(\mathbb{R})$ by Plancherel theorem. The analysis operator is: $T_\zeta^* f = \check{f}$, for all $f \in L^2(\mathbb{R})$.

Example ([29, Example 3.18]) Let us consider again $\mathcal{S}(\mathbb{R}) \hookrightarrow L^2(\mathbb{R}) \hookrightarrow \mathcal{S}^\times(\mathbb{R})$. For $x \in \mathbb{R}$, let us consider the Dirac delta $\delta_x : \mathcal{S}(\mathbb{R}) \to \mathbb{C}$, $\phi \mapsto \langle \phi | \delta_x \rangle := \phi(x)$. The set of Dirac deltas $\delta := \{\delta_x\}_{x \in \mathbb{R}}$ is a Gel'fand distribution basis. In fact, the Parseval identity holds:

$$\int_{\mathbb{R}} |\langle \delta_x | \phi \rangle|^2 dx = \int_{\mathbb{R}} |\phi(x)|^2 dx = \|\phi\|_2^2, \quad \forall \phi \in \mathcal{S}(\mathbb{R}).$$

The synthesis operator: $T_\delta : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ is:

$$\langle T_\delta \xi | \phi \rangle = \int_{\mathbb{R}} \xi(x) \langle \delta_x | \phi \rangle \, dx = \int_{\mathbb{R}} \xi(x) \overline{\phi(x)} dx = \langle \xi | \phi \rangle, \quad \forall \phi \in \mathcal{S}(\mathbb{R}).$$

Then $T_\delta \xi = \xi$ for all $\xi \in L^2(\mathbb{R})$. Since T_δ is an identity, it is an isometry onto $L^2(\mathbb{R})$.

5.2 *Riesz Distribution Basis*

Proposition 5.4 and (5.1) suggest a more general class of bases that will play the same role as Riesz bases in the ordinary Hilbert space framework.

Definition 5.5 Let $\mathcal{D} \subset \mathcal{H} \subset \mathcal{D}^\times$ be a rigged Hilbert space. A weakly measurable map $\omega : x \in X \mapsto \omega_x \in \mathcal{D}^\times$ is a *Riesz distribution basis* if ω is a μ-independent distribution frame.

One has the following:

Proposition 5.6 ([29, Proposition 3.19]) *Let $\mathcal{D} \subset \mathcal{H} \subset \mathcal{D}^\times$ be a rigged Hilbert space and let $\omega : x \in X \mapsto \omega_x \in \mathcal{D}^\times$ be a Bessel distribution map. Then the following statements are equivalent:*

(a) *ω is a Riesz distribution basis;*
(b) *If ζ is a Gel'fand distribution basis, then the operator W defined, for $f \in \mathcal{H}$, by:*

$$f = \int_X \xi_f(x) \zeta_x d\mu \mapsto Wf = \int_X \xi_f(x) \omega_x d\mu$$

is continuous and has bounded inverse;
(c) *the synthesis operator T_ω is a topological isomorphism of $L^2(X, \mu)$ onto $\mathcal{H}$.*

Proposition 5.7 *If ω is a Riesz distribution basis then ω possesses a unique dual frame θ which is also a Riesz distribution basis.*

Example Let us consider $f \in C^\infty(\mathbb{R})$: $0 < m < |f(x)| < M$. Let us define $\omega_x := f(x)\delta_x$: then $\{\omega_x\}_{x\in\mathbb{R}}$ is a distribution frame, in fact:

$$\int_{\mathbb{R}} |\langle \omega_x|\phi\rangle|^2 dx = \int_{\mathbb{R}} |\overline{f(x)}\phi(x)|^2 dx \leq M^2\|\phi\|_2^2, \quad \forall\phi \in \mathcal{S}(\mathbb{R}),$$

and

$$m^2\|\phi\|_2^2 \leq \int_{\mathbb{R}} |\overline{f(x)}\phi(x)|^2 dx \leq M^2\|\phi\|_2^2, \quad \forall\phi \in \mathcal{S}(\mathbb{R}).$$

Furthermore, $\{\omega_x\}_{x\in\mathbb{R}}$ is μ-independent. In fact, putting:

$$\int_{\mathbb{R}} \xi(x)\,\langle \omega_x|g\rangle\, dx = 0, \quad \forall g \in \mathcal{S}(\mathbb{R}),$$

one has:

$$\int_{\mathbb{R}} \xi(x)\,\langle \omega_x|g\rangle\, dx = \int_{\mathbb{R}} \xi(x)\overline{f(x)}\,\langle \delta_x|g\rangle dx = 0, \quad \forall g \in \mathcal{S}(\mathbb{R}).$$

Since $\{\delta_x\}_{x\in\mathbb{R}}$ is μ-independent, it follows that $\xi(x)\overline{f(x)} = 0$ a.e., then $\xi(x) = 0$ a.e.. By definition, $\{\omega_x\}_{x\in\mathbb{R}}$ is a Riesz distribution basis.

6 Concluding Remarks

In a Hilbert space, frames, semi-frames, Bessel, Riesz-Fischer sequences, and Riesz bases are related through the action of a linear operator on elements of an orthonormal basis (see also [29, Remark 3.22]). On the other hand, in literature some studies on bounds (upper and lower) of these sequences have been already considered and their links with the linear operators related to them have been studied (see [3, 6, 9]). For that, it is desirable to continue an analogous study in rigged Hilbert spaces by considering linear operators in $\mathcal{L}(\mathcal{D}, \mathcal{D}^\times)$.

Acknowledgement The author would like to thank R. Corso for some valuable comments and suggestions.

References

1. S.T. Ali, J.P. Antoine, J.P. Gazeau, Continuous frames in Hilbert spaces. Ann. Phys. **222**, 1–37 (1993)
2. S.T. Ali, J.P. Antoine, J.P. Gazeau, *Coherent States, Wavelets and Their Generalizations*, 2nd edn. (Springer, Berlin, 2014)

3. J.-P. Antoine, P. Balazs, Frames and semi-frames. J. Phys. A Math. Theor. **44**, 205201 (2011)
4. J.-P. Antoine, C. Trapani, Reproducing pairs of measurable functions and partial inner product spaces. Adv. Oper. Theory **2**, 126–146 (2017)
5. J.-P. Antoine, A. Inoue, C. Trapani, *Partial *-Algebras and Their Operator Realizations* (Kluwer, Dordrecht, 2002)
6. P. Balazs, D.T. Stoeva, J.-P. Antoine, Classification of general sequences by frame-related operators. Sampl. Theory Signal Image Process. **10**, 151–170 (2011)
7. P. Balasz, M. Speckbacher, Frames, their relatives and reproducing kernel Hilbert spaces. J. Phys. A Math. Theor. **53**, 015204 (2020)
8. G. Bellomonte, C. Trapani, Riesz-like bases in rigged Hilbert spaces. Zeitschr. Anal. Anwen. **35**, 243–265 (2016)
9. P. Casazza, O. Christensen, S. Li, A. Lindner, Riesz-Fischer sequences and lower frame bounds. Zeitschr. Anal. Anwen. **21**, 305–314 (2002)
10. O. Christensen, *Frames and Bases: An Introductory Course* (Birkhäuser, Boston, 2008)
11. O. Christensen, *An Introduction to Frames and Riesz Bases* (Birkhaäuser, Boston, 2016)
12. E. Cordero, H. Feichtinger, F. Luef, Banach Gelfand triples for Gabor analysis, in *Pseudod-ifferential Operators*. Lecture Notes in Mathematics, vol. 1949 (Springer, Berlin, 2008), pp. 1–33
13. G. de Barra, *Measure Theory and Integration* (New Age International (P) limited Publishers, Darya Ganj, 1981)
14. H.G. Feichtinger, K. Gröchenig, Gabor frames and time-frequency analysis of Distributions. J. Funct. Anal. **146**, 464–495 (1997)
15. H.G. Feichtinger, G. Zimmermann, A Banach space of test functions for Gabor analysis, in *Gabor Analysis and Algorithms: Theory and Applications* (Birkhäuser, Boston, 1998)
16. I.M. Gel'fand, N.Ya. Vilenkin, *Generalized Functions*, vol. IV (Academic, New York, 1964)
17. I.M. Gel'fand, G.E. Shilov, E. Saletan, *Generalized Functions*, vol. III (Academic, New York, 1967)
18. G.G. Gould, The spectral representation of normal operators on a rigged Hilbert space. J. London Math. Soc. **43**, 745–754 (1968)
19. K. Gröchenig, *Foundations of Time-Frequency Analysis* (Birkhäuser, Basel, 2001)
20. C. Heil, *A Basis Theory Primer*. Expanded Edition (Birkhäuser/Springer, New York, 2011)
21. J. Horvath, *Topological Vector Spaces and Distributions* (Addison-Wesley, Boston, 1966)
22. H. Hosseini Giv, M. Radjabalipour, On the structure and properties of lower bounded analytic frames. Iran. J. Sci. Technol. **37**, 227–230 (2013)
23. M.S. Jakobsen, J. Lemvig, Density and duality theorems for regular Gabor frames. J. Funct. Anal. **270**, 229–263 (2016)
24. G. Kaiser, *A Friendly Guide to Wavelets* (Birkhäuser, Boston, 1994)
25. G. Kyriazis, P. Petrushev, On the construction of frames for spaces of distributions. J. Funct. Anal. **257**, 2159–2187 (2009)
26. S. Pilipovic, D.T. Stoeva, Fréchet frames, general definition and expansion. Anal. Appl. **12**, 195–208 (2014)
27. M. Reed, B. Simon, *Methods of Modern Mathematical Physics*, vols. I and II (Academic, New York, 1980)
28. W. Rudin, *Real and Complex analysis* (McGraw-Hill, New York, 1987)
29. C. Trapani, S. Triolo, F. Tschinke, Distribution frames and bases. J. Fourier Anal. and Appl. **25**, 2109–2140 (2019)
30. R.M. Young, *An Introduction to Nonharmonic Fourier Series*, 2nd edn. (Academic, Cambridge, 2001)

Periodic Coherent States Decomposition and Quantum Dynamics on the Flat Torus

Lorenzo Zanelli

Abstract We provide a result on the coherent states decomposition for functions in $L^2(\mathbb{T}^n)$ where $\mathbb{T}^n := (\mathbb{R}/2\pi\mathbb{Z})^n$. We study such a decomposition with respect to the quantum dynamics related to semiclassical elliptic Pseudodifferential operators, and we prove a related invariance result.

Keywords Coherent states · Toroidal Pdo · Quantum dynamics

Mathematics Subject Classification (2010) 81R30, 58J40, 58C40

1 Introduction

Let us introduce the usual class of semiclassical coherent states on $\mathbb{R}^n$:

$$\phi_{(x,\xi)}(y) := \alpha_h \, e^{\frac{i}{h}(x-y)\cdot\xi} e^{-\frac{|x-y|^2}{2h}}, \ (x,\xi) \in \mathbb{R}^{2n}, \ y \in \mathbb{R}^n, \ 0 < h \leq 1 \tag{1.1}$$

with the $L^2(\mathbb{R}^n)$—normalization constant $\alpha_h := 2^{-\frac{n}{2}}(\pi h)^{-3n/4}$, and where h is a 'semiclassical parameter'. For any $\psi \in \mathcal{S}'(\mathbb{R}^n)$ the coherent state decomposition reads, in the distributional sense, as

$$\psi(x_0) = \int_{\mathbb{R}^{2n}} \phi^{\star}_{(x,\xi)}(x_0) \Big(\int_{\mathbb{R}^n} \phi_{(x,\xi)}(y)\psi(y)\, dy \Big) dx\, d\xi \tag{1.2}$$

as shown for example in [10, Proposition 3.1.6].

We now observe that for the flat torus $\mathbb{T}^n := (\mathbb{R}/2\pi\mathbb{Z})^n$ the well known inclusion $L^2(\mathbb{T}^n) \subset \mathcal{S}'(\mathbb{R}^n)$ implies that distributional equality (1.2) makes sense also for functions in $L^2(\mathbb{T}^n)$.

L. Zanelli (✉)
Department of Mathematics "Tullio Levi-Civita", University of Padova, Padova, Italy
e-mail: lzanelli@math.unipd.it

M. A. Bastos et al. (eds.), *Operator Theory, Functional Analysis and Applications*,
Operator Theory: Advances and Applications 282,
https://doi.org/10.1007/978-3-030-51945-2_30

The first aim of our paper is to prove the decomposition of any $\varphi \in L^2(\mathbb{T}^n)$ with respect to the family of periodic coherent states Φ given by the periodization of (1.1). In view of this target, we recall that the periodization operator

$$\Pi(\phi)(y) := \sum_{k\in\mathbb{Z}^n} \phi(y - 2\pi k)$$

maps $\mathcal{S}(\mathbb{R}^n)$ into $C^\infty(\mathbb{T}^n)$, as is shown for example in [14, Theorem 6.2]. Thus, we can define for all $0 < h \leq 1$,

$$\Phi_{(x,\xi)}(y) := \sum_{k\in\mathbb{Z}^n} \phi_{(x,\xi)}(y - 2\pi k), \qquad (x,\xi) \in \mathbb{T}^n \times h\,\mathbb{Z}^n, \qquad y \in \mathbb{T}^n. \tag{1.3}$$

Notice that the family of coherent states in (1.3) is well posed also for $\xi \in \mathbb{R}^n$ and the related phase space is $\mathbb{T}^n \times \mathbb{R}^n$. However, our target is to show that the decomposition of periodic functions can be done with respect to the minimal set of coherent states in (1.3) for $\xi \in h\,\mathbb{Z}^n \subset \mathbb{R}^n$. Furthermore, we notice that the phase space $\mathbb{T}^n \times h\,\mathbb{Z}^n$ is necessary in order to deal with a well defined setting of toroidal Weyl operators acting on $L^2(\mathbb{T}^n)$ and more in general with semiclassical toroidal Pseudodifferential operators (see Sect. 2).

The first result of the paper is the following.

Theorem 1.1 *Let $\varphi_h \in C^\infty(\mathbb{T}^n)$ be such that $\|\Delta_x \varphi_h\|_{L^2} \leq c\,h^{-M}$ for some $c > 0$, $M \in \mathbb{N}$, $\|\varphi_h\|_{L^2} = 1$ with $0 < h \leq 1$, $h^{-1} \in \mathbb{N}$ and let $\Phi_{(x,\xi)}$ be as in (1.3). Then*

$$\varphi_h = \sum_{\xi\in h\,\mathbb{Z}^n} \int_{\mathbb{T}^n} \langle \Phi_{(x,\xi)}, \varphi_h\rangle_{L^2}\, \Phi^\star_{(x,\xi)}\, dx + \mathcal{O}_{L^2}(h^\infty). \tag{1.4}$$

Moreover, there exists $f(h) > 0$ depending on φ_h such that

$$\varphi_h = \sum_{\xi\in h\,\mathbb{Z}^n,\ |\xi|\leq f(h)} \int_{\mathbb{T}^n} \langle \Phi_{(x,\xi)}, \varphi_h\rangle_{L^2}\, \Phi^\star_{(x,\xi)}\, dx + \mathcal{O}_{L^2}(h^\infty). \tag{1.5}$$

The following inclusion involving the set of frequencies $\xi \in h\,\mathbb{Z}^n \subset \mathbb{R}^n$ allows to consider decomposition (1.4) minimal with respect to (1.2). The above result shows also that the sum over the frequencies can be taken in the bounded region $|\xi| \leq f(h)$, i.e. we can consider a finite sum by taking into account an $\mathcal{O}(h^\infty)$ remainder in $L^2(\mathbb{T}^n)$.

An analogous result of (1.4) in the two dimensional setting is shown in [4, Proposition 60] by the use of a different periodization operator. Same construction of coherent states as in [4] for $\mathbb{T}^2$ is used in [2, 6] for the study of quantum cat maps and equipartition of the eigenfunctions of quantized ergodic maps. In the paper [8], covariant integral quantization using coherent states for semi-direct product groups is implemented for the motion of a particle on the circle and in particular the resolution of the identity formula is proved. Another class of coherent states

on the torus are defined also in [7], with a related resolution of the identity, in the understanding of the Quantum Hall effect. We also recall [5] where coherent states and Bargmann Transform are studied on $L^2(\mathbb{S}^n)$. The literature on coherent states are quite rich, and thus we address the reader to [1].

We now devote our attention to the periodic coherent states decomposition for eigenfunctions of elliptic semiclassical toroidal Pseudodifferential operators (see Sect. 2). We will see that the formula (1.4) can be reduced in view of a phase-space localization of eigenfunctions.

This is the content of the second main result of the paper.

Theorem 1.2 *Let* $\mathrm{Op}_h(b)$ *be an elliptic semiclassical* Ψ*do as in* (2.1) *and* $h^{-1} \in \mathbb{N}$. *Let* $E \in \mathbb{R}$, *and let* $\psi_h \in C^\infty(\mathbb{T}^n)$ *be such that* $\|\psi_h\|_{L^2} = 1$ *and* $\|\Delta_x \psi_h\|_{L^2} \leq c\, h^{-M}$, *and which is eigenfunction of the eigenvalue problem on* $\mathbb{T}^n$ *given by*

$$\mathrm{Op}_h(b)\psi_h = E_h \psi_h$$

where $E_h \leq E$ *for any* $0 < h \leq 1$. *Then, there exists* $g(h, E) \in \mathbb{R}_+$ *such that*

$$\psi_h = \sum_{\xi \in h\,\mathbb{Z}^n,\ |\xi| \leq g(h,E)} \int_{\mathbb{T}^n} \langle \Phi_{(x,\xi)}, \psi_h \rangle_{L^2}\, \Phi^\star_{(x,\xi)}\, dx + \mathcal{O}_{L^2}(h^\infty). \tag{1.6}$$

We notice that for the operators $-h^2\Delta_x + V(x)$ all the eigenfunctions with eigenvalues $E_h \leq E$ fulfill $\|\Delta_x \psi_h\|_{L^2} \leq c\, h^{-2}$. In particular, we have the asymptotics $g(h, E) \to +\infty$ as $h \to 0^+$. We also underline that the function $g(E, h)$ and the estimate on remainder $\mathcal{O}_{L^2}(h^\infty)$ do not depend on the particular choice of ψ_h. This implies that all these eigenfunctions take the form (1.6) and therefore also any finite linear combination of eigenfunctions of kind $\sum_{1 \leq \alpha \leq N} c_\alpha\, \psi_{h,\alpha}$ where $|c_\alpha| \leq 1$. We remind that Weyl Law on the number $\mathcal{N}(h)$ of eigenvalues $E_{h,\alpha} \leq E$ (with their multiplicity) for semiclassical elliptic operators (see for example [9]) reads $\mathcal{N}(h) \simeq (2\pi h)^{-n}(\mathrm{vol}(U(E)) + \mathcal{O}(1))$.

The proof of the above result is mainly based on a uniform estimate for our toroidal version of the Fourier-Bros-Iagolnitzer (FBI) transform

$$(\mathcal{T}\psi_h)(x, \xi) := \langle \Phi_{(x,\xi)}, \psi_h \rangle_{L^2}$$

on the unbounded region given by all $x \in \mathbb{T}^n$ and $\xi \in h\mathbb{Z}^n$ such that $|\xi| > g(h, E)$. The FBI transform on any compact manifold has already been defined and studied in the literature, see for example [16].

We remind that, in the euclidean setting of $\mathbb{R}^{2n}$, the function

$$T_h(\psi_h)(x, \xi) := \langle \phi_{(x,\xi)}, \psi_h \rangle_{L^2(\mathbb{R}^n)}$$

is the usual version of the FBI transform, which is well posed for any $\psi_h \in \mathcal{S}'(\mathbb{R}^n)$. This is used to study the phase space localization by the Microsupport of ψ_h (see for example [10]), namely $MS(\psi_h)$ the complement of the set of points (x_0, ξ_0) such

that $T_h(\psi_h)(x,\xi) \simeq \mathcal{O}(e^{-\delta/h})$ uniformly in a neighborhood of (x_0,ξ_0). In the case of the weaker estimate $T_h(\psi_h)(x,\xi) \simeq \mathcal{O}(h^\infty)$ one can define the semiclassical Wave Front Set $WF(\psi_h)$. Is well known (see [10]) that the Microsupport (or the semiclassical Wave Front Set) of eigenfunctions for elliptic operators is localized in the sublevel sets

$$U(E) := \{(x,\xi) \in \mathbb{R}^{2n} \mid b(x,\xi) \leq E\},$$

i.e. $MS(\psi_h) \subseteq U(E)$. The well posedness of $WF(\psi_h)$ and $MS(\psi_h)$ in the periodic setting can be seen starting from the euclidean setting and thanks to distributional inclusion $L^2(\mathbb{T}^n) \subset \mathcal{S}'(\mathbb{R}^n)$, (see for example section 3.1 of [3]). The semiclassical study in the phase space for eigenfunctions in the periodic setting has also been studied in [18] with respect to weak KAM theory.

In our Theorem 1.2 we are interested to show another kind of semiclassical localization, namely to localize the bounded region

$$\Omega(E,h) := \{(x,\xi) \in \mathbb{T}^n \times \mathbb{R}^n \mid x \in \mathbb{T}^n, |\xi| \leq g(E,h)\}$$

which will be bigger than $MS(\psi_h)$, h—dependent and such that the coherent state decomposition of ψ_h can be done up to a remainder $\mathcal{O}_{L^2}(h^\infty)$.

We now focus our attention to the decompositon (1.6) under the time evolution.

Theorem 1.3 *Let $\varphi_h \in C^\infty(\mathbb{T}^n)$, L^2—normalized such that*

$$\varphi_h = \sum_{1 \leq j \leq J(h)} c_j \, \psi_{h,j} \tag{1.7}$$

where $\psi_{h,j}$ are given in Theorem 1.2 and $J(h) \leq J_0 h^{-Q}$ for some $J_0, Q > 0$. Let $\mathrm{Op}_h(b)$ be an elliptic semiclassical Ψdo as in (2.1) and

$$U_h(t) := \exp\{(-i\mathrm{Op}_h(b)t)/h\}.$$

Then, there exists $\ell(h) > 0$ such that for any $t \in \mathbb{R}$,

$$U_h(t)\varphi_h = \sum_{\xi \in h\,\mathbb{Z}^n,\ |\xi| \leq \ell(h)} \int_{\mathbb{T}^n} \langle \Phi_{(x,\xi)}, U_h(t)\varphi_h \rangle_{L^2} \, \Phi^\star_{(x,\xi)} \, dx + \mathcal{O}_{L^2}(h^\infty). \tag{1.8}$$

The equality (1.8) shows that time evolution under the L^2—unitary map $U_h(t)$ does not change such a decomposition, since $\ell(h)$ does not depend on time. The function $\ell(h)$ is not necessarily the same as the function $f(h)$ contained in Theorem 1.1 but we have that $\ell(h) \geq f(h)$. In other words, this quantum dynamics preserves the coherent state decomposition (1.5). The same result holds for any eigenfunctions in Theorem 1.2 since in this case $U_h(t)\psi_h = \exp\{(-iE_h t)/h\}\psi_h$. Notice that here we can assume that $Q > n$, namely the linear combination

(1.7) can be done with more eigenfunctions than the ones that have eigenvalues $E_h \leq E$ with fixed energy $E > 0$. Notice also that we have $\langle \Phi_{(x,\xi)}, U_h(t)\varphi_h \rangle_{L^2} = \langle U_h(-t)\Phi_{(x,\xi)}, \varphi_h \rangle_{L^2}$ for any $t \in \mathbb{R}$ and that the time evolution of the periodization of coherent states has been used in [17] in the context of optimal transport theory.

2 Semiclassical Toroidal Pseudodifferential Operators

Let us define the flat torus $\mathbb{T}^n := (\mathbb{R}/2\pi\mathbb{Z})^n$ and introduce the class of symbols $b \in S^m_{\rho,\delta}(\mathbb{T}^n \times \mathbb{R}^n)$, $m \in \mathbb{R}$, $0 \leq \delta$, $\rho \leq 1$, given by functions in $C^\infty(\mathbb{T}^n \times \mathbb{R}^n; \mathbb{R})$ which are 2π-periodic in each variable x_j, $1 \leq j \leq n$ and for which for all $\alpha, \beta \in \mathbb{Z}^n_+$ there exists $C_{\alpha\beta} > 0$ such that for all $(x, \xi) \in \mathbb{T}^n \times \mathbb{R}^n$,

$$|\partial_x^\beta \partial_\xi^\alpha b(x, \xi)| \leq C_{\alpha\beta m} \langle \xi \rangle^{m-\rho|\alpha|+\delta|\beta|}$$

where $\langle \xi \rangle := (1+|\xi|^2)^{1/2}$. In particular, the set $S^m_{1,0}(\mathbb{T}^n \times \mathbb{R}^n)$ is denoted by $S^m(\mathbb{T}^n \times \mathbb{R}^n)$.

We introduce the semiclassical toroidal Pseudodifferential Operators by the following.

Definition 2.1 Let $\psi \in C^\infty(\mathbb{T}^n; \mathbb{C})$ and $0 < h \leq 1$,

$$\mathrm{Op}_h(b)\psi(x) := (2\pi)^{-n} \sum_{\kappa \in \mathbb{Z}^n} \int_{\mathbb{T}^n} e^{i\langle x-y,\kappa \rangle} b(x, h\kappa)\psi(y)\, dy. \tag{2.1}$$

This is the semiclassical version (see [12, 13]) of the quantization by Pseudodifferential Operators on the torus developed in [14] and [15]. See also [11] for the notion of vector valued Pseudodifferential Operators on the torus.

We now notice that we have a map $\mathrm{Op}_h(b) : C^\infty(\mathbb{T}^n) \longrightarrow \mathcal{D}'(\mathbb{T}^n)$. Indeed, remind that $u \in \mathcal{D}'(\mathbb{T}^n)$ are the linear maps $u : C^\infty(\mathbb{T}^n) \longrightarrow \mathbb{C}$ such that there exist $C > 0$ and $k \in \mathbb{N}$, for which $|u(\phi)| \leq C \sum_{|\alpha| \leq k} \|\partial_x^\alpha \phi\|_\infty$ for all $\phi \in C^\infty(\mathbb{T}^n)$.

Given a symbol $b \in S^m(\mathbb{T}^n \times \mathbb{R}^n)$, the toroidal Weyl quantization reads (see [12, 13])

$$\mathrm{Op}^w_h(b)\psi(x) := (2\pi)^{-n} \sum_{\kappa \in \mathbb{Z}^n} \int_{\mathbb{T}^n} e^{i\langle x-y,\kappa \rangle} b\Big(y, \frac{h}{2}\kappa\Big)\psi(2y - x)\, dy.$$

In particular, it holds

$$\mathrm{Op}^w_h(b)\psi(x) = (\mathrm{Op}_h(\sigma) \circ T_x\, \psi)(x)$$

where $T_x : C^\infty(\mathbb{T}^n) \to C^\infty(\mathbb{T}^n)$ defined as $(T_x\psi)(y) := \psi(2y - x)$ is linear, invertible and L^2-norm preserving, and σ is a suitable toroidal symbol related to b, i.e.

$$\sigma \sim \sum_{\alpha\geq 0} \frac{1}{\alpha!} \triangle_\xi^\alpha D_y^{(\alpha)} b(y, h\xi/2)\big|_{y=x},$$

where $\triangle_{\xi_j} f(\xi + e_j) - f(\xi)$ is the difference operator (see [14, Theorem 4.2]).

The typical example is given by

$$\begin{aligned} \mathrm{Op}_h(H) &= \Big(-\frac{1}{2}h^2\Delta_x + V(x)\Big)\psi(x) \\ &= (2\pi)^{-n} \sum_{\kappa\in\mathbb{Z}^n} \int_{\mathbb{T}^n} e^{i\langle x-y,\kappa\rangle} \Big(\frac{1}{2}|h\kappa|^2 + V(x)\Big)\psi(y)\, dy \end{aligned}$$

namely the related symbol is the mechanical type Hamiltonian $H(x,\xi) = \frac{1}{2}|\xi|^2 + V(x)$. Also in the case of the Weyl operators we have

$$-\frac{1}{2}h^2\Delta_x + V(x) = \mathrm{Op}_h^w(H)$$

for the same symbol (see for example [13]).

In our paper we are interested in uniform elliptic operators, namely such that the symbol $b \in S^m(\mathbb{T}^n \times \mathbb{R}^n)$ fulfills for some constants $C, c > 0$ the lower bound

$$|b(x,\xi)| \geq C\,\langle\xi\rangle^m$$

for any $x \in \mathbb{T}^n$ and $|\xi| \geq c$. This property guarantees bounded sublevels sets for b and discrete spectrum for the operator $\mathrm{Op}_h(b)$ for any fixed $0 < h \leq 1$. As we see in Theorem 1.2, this assumption permits also to prove the semiclassical localization of all the eigenfunctions within these sublevels sets, and this localization can be studied by our semiclassical coherent states (1.3).

3 Proofs of the Main Results

3.1 *Proof of Theorem 1.1*

We remind that $\Phi_{(x,\xi)}(y) := \Pi(\phi_{(x,\xi)})(y)$ and $\Pi(\phi)(y) := \sum_{k\in\mathbb{Z}^n} \phi(y - 2\pi k)$. Thus,

$$\begin{aligned} \Phi_{(x+2\pi\beta,\xi)}(y) &= \sum_{k\in\mathbb{Z}^n} \phi_{(x+2\pi\beta,\xi)}(y - 2\pi k) = \sum_{k\in\mathbb{Z}^n} \phi_{(x,\xi)}(y - 2\pi k - 2\pi\beta) \\ &= \Phi_{(x,\xi)}(y). \end{aligned}$$

We mainly adapt, in our toroidal setting, the proof of [10, Proposition 3.1.6] written for the euclidean setting. Thus, we define the operator $\mathcal{T}^\star$ on functions $\Psi \in L^2(\mathbb{T}^n \times h\mathbb{Z}^n)$ as

$$(\mathcal{T}^\star \Psi)(y) := \sum_{\xi \in h\,\mathbb{Z}^n} \int_{\mathbb{T}^n} \Psi(x,\xi) \Phi^\star_{(x,\xi)}(y)\, dx.$$

It can be easily seen that $\mathcal{T}^\star$ equals the adjoint of the operator $(\mathcal{T}\psi)(x,\xi) := \langle \Phi_{(x,\xi)}, \psi \rangle_{L^2(\mathbb{T}^n)}$, i.e.

$$\langle \mathcal{T}^\star \Psi, \psi \rangle_{L^2(\mathbb{T}^n)} = \langle \Psi, \mathcal{T}\psi \rangle_{L^2(\mathbb{T}^n \times h\mathbb{Z}^n)}.$$

Thus, for all $\psi_1, \psi_2 \in C^\infty(\mathbb{T}^n) \subset L^2(\mathbb{T}^n)$ we have

$$\langle \mathcal{T}^\star \circ \mathcal{T}\psi_1, \psi_2 \rangle_{L^2(\mathbb{T}^n)} = \langle \mathcal{T}\psi_1, \mathcal{T}\psi_2 \rangle_{L^2(\mathbb{T}^n \times h\mathbb{Z}^n)}.$$

It remains to prove that

$$\langle \mathcal{T}\psi_1, \mathcal{T}\psi_2 \rangle_{L^2(\mathbb{T}^n \times h\mathbb{Z}^n)} = \langle \psi_1, \psi_2 \rangle_{L^2(\mathbb{T}^n)} + \mathcal{O}(h^\infty) \tag{3.1}$$

which implies

$$\mathcal{T}^\star \circ \mathcal{T} = \mathrm{Id} \bmod \mathcal{O}(h^\infty) \tag{3.2}$$

on $L^2(\mathbb{T}^n)$, and equality (3.2) is exactly the statement (1.4).

In order to prove (3.1), we recall that the periodization operator Π can be rewritten in the form (see [14, Theorem 6.2]):

$$\Pi(\phi) = \mathcal{F}^{-1}_{\mathbb{T}^n}\Big(\mathcal{F}_{\mathbb{R}^n}\phi\Big|_{\mathbb{Z}^n}\Big). \tag{3.3}$$

where $\mathcal{F}^{-1}_{\mathbb{T}^n}$ stands for the inverse toroidal Fourier Transform, and $\mathcal{F}_{\mathbb{R}^n}$ is the usual euclidean version. In view of (3.3) it follows

$$\begin{aligned}(\mathcal{T}\psi)(x,\xi) &:= \langle \Phi_{(x,\xi)}, \psi \rangle_{L^2(\mathbb{T}^n)} = \langle \mathcal{F}_{\mathbb{R}^n}\phi_{x,\xi}|_{\mathbb{Z}^n}, \mathcal{F}_{\mathbb{T}^n}\psi \rangle_{L^2(\mathbb{Z}^n)} \\ &= \sum_{k \in \mathbb{Z}^n} \widehat{\phi}_{x,\xi}(k)^\star \widehat{\psi}(k),\end{aligned}$$

where $\widehat{\phi}_{x,\xi}(k) := \mathcal{F}_{\mathbb{R}^n}\phi_{x,\xi}(k)$ and $\widehat{\psi}(k) := \mathcal{F}_{\mathbb{T}^n}\psi(k)$. Thus,

$$\begin{aligned}&\langle \mathcal{T}\psi_1, \mathcal{T}\psi_2 \rangle_{L^2(\mathbb{T}^n \times h\mathbb{Z}^n)} \\ &\quad = \sum_{\xi \in h\,\mathbb{Z}^n} \int_{\mathbb{T}^n} \Big(\sum_{k \in \mathbb{Z}^n} \widehat{\phi}_{x,\xi}(k)^\star \widehat{\psi}(k) \Big)^\star \Big(\sum_{\mu \in \mathbb{Z}^n} \widehat{\phi}_{x,\xi}(\mu)^\star \widehat{\psi}(\mu) \Big) dx.\end{aligned}$$

We can rewrite this equality, in the distributional sense, as

$$\langle \mathcal{T}\psi_1, \mathcal{T}\psi_2\rangle_{L^2(\mathbb{T}^n\times h\mathbb{Z}^n)} = \sum_{k,\mu\in\mathbb{Z}^n} \widehat{\psi}_1(k)^\star \widehat{\psi}_2(\mu) \sum_{\xi\in h\,\mathbb{Z}^n} \int_Q \widehat{\phi}_{x,\xi}(k)\widehat{\phi}_{x,\xi}(\mu)^\star dx$$

where $Q := [0, 2\pi]^n$ and $\psi_1, \psi_2 \in C^\infty(\mathbb{T}^n)$. Now let $\xi = h\alpha$ with $\alpha \in \mathbb{Z}^n$, so that

$$\langle \mathcal{T}\psi_1, \mathcal{T}\psi_2\rangle_{L^2(\mathbb{T}^n\times h\mathbb{Z}^n)} = \sum_{k,\mu\in\mathbb{Z}^n} \widehat{\psi}_1(k)^\star \widehat{\psi}_2(\mu) \sum_{\alpha\in\mathbb{Z}^n} \int_Q \widehat{\phi}_{x,h\alpha}(k)\widehat{\phi}_{x,h\alpha}(\mu)^\star dx$$

By using the explicit form of $\widehat{\phi}_{x,h\alpha}$ and the condition $h^{-1} \in \mathbb{N}$, a direct computation shows that

$$\langle \mathcal{T}\psi_1, \mathcal{T}\psi_2\rangle_{L^2(\mathbb{T}^n)} = \sum_{k,\mu\in\mathbb{Z}^n} \widehat{\psi}_1(k)^\star \widehat{\psi}_2(\mu)\Big[\Big(\sum_{\alpha\in\mathbb{Z}^n} e^{i\alpha(k-\mu)}\Big) + \mathcal{O}(h^\infty)\Big] \tag{3.4}$$

where $\mathcal{O}(h^\infty)$ does not depend on the functions ψ_1, ψ_2. We now use the assumption $\|\Delta_x\psi\|_{L^2} \le c\,h^{-M}$ for some fixed $c, M > 0$ so that Fourier components fulfill $|\widehat{\psi}_k| \le |k|^{-2}(2\pi)^{n/2}c\,h^{-M}$, and $\|\psi\|_{L^2} = 1$ gives $|\widehat{\psi}_0| \le (2\pi)^{n/2}$. Consequently,

$$\sum_{k\in\mathbb{Z}^n} |\widehat{\psi}_1(k)| \le (2\pi)^{n/2} + (2\pi)^{n/2}c \sum_{k\in\mathbb{Z}^n\setminus\{0\}} |k|^{-2}\,h^{-M}, \tag{3.5}$$

and

$$\Big|\sum_{k,\mu\in\mathbb{Z}^n} \widehat{\psi}_1(k)^\star \widehat{\psi}_2(\mu)\Big| \le \sum_{k\in\mathbb{Z}^n} |\widehat{\psi}_1(k)| \sum_{\mu\in\mathbb{Z}^n} |\widehat{\psi}_2(\mu)|. \tag{3.6}$$

To conclude, since

$$\delta(k-\mu) = \sum_{\alpha\in\mathbb{Z}^n} e^{i\alpha(k-\mu)},$$

we get

$$\begin{aligned}\langle \mathcal{T}\psi_1, \mathcal{T}\psi_2\rangle_{L^2(\mathbb{T}^n\times h\mathbb{Z}^n)} &= \sum_{k\in\mathbb{Z}^n} \widehat{\psi}_1(k)^\star \widehat{\psi}_2(k) + \mathcal{O}(h^\infty)\\ &= \langle \psi_1, \psi_2\rangle_{L^2(\mathbb{T}^n)} + \mathcal{O}(h^\infty).\end{aligned} \tag{3.7}$$

The estimates (3.5)–(3.6) together with (3.4) ensure that the remainder in (3.7) has order $\mathcal{O}(h^\infty)$.

In order to prove (1.5), we observe that

$$\varphi_h = \sum_{\xi \in h\,\mathbb{Z}^n} \int_{\mathbb{T}^n} \langle \Phi_{(x,\xi)}, \varphi_h \rangle_{L^2}\, \Phi^{\star}_{(x,\xi)}\, dx + \mathcal{O}(h^{\infty}).$$

is given by an L^2-convergent series. Thus, for any fixed φ_h we can say that there exists $f(h) > 0$ such that

$$\varphi_h = \sum_{\xi \in h\,\mathbb{Z}^n, |\xi| < f(h)} \int_{\mathbb{T}^n} \langle \Phi_{(x,\xi)}, \varphi_h \rangle_{L^2}\, \Phi^{\star}_{(x,\xi)}\, dx + \mathcal{O}(h^{\infty}).$$

□

3.2 *Proof of Theorem 1.2*

We apply the statement of Theorem 1.1, for a set of linearly independent eigenfunctions $\psi_{h,i}$ generating all the eigenspaces linked to eigenvalues $E_h \leq E$ and $f_i(h) > 0$ given by Theorem 1.1:

$$\psi_{h,i} = \sum_{\xi \in h\,\mathbb{Z}^n, |\xi| < f_i(h)} \int_{\mathbb{T}^n} \langle \Phi_{(x,\xi)}, \psi_{h,i} \rangle_{L^2}\, \Phi^{\star}_{(x,\xi)}\, dx + R_{h,i}$$

where $\|R_{h,i}\|_{L^2} = \mathcal{O}(h^{\infty})$.

Moreover, we recall that the Weyl Law on the number $\mathcal{N}(h)$ of eigenvalues $E_h \leq E$ (counted with their multiplicity) for semiclassical elliptic operators (see for example [9]) reads

$$\mathcal{N}(E,h) \simeq (2\pi h)^{-n}(\mathrm{vol}(U(E)) + \mathcal{O}(1)).$$

We define:

$$g(E,h) := \max_{1 \leq i \leq \mathcal{N}(E,h)} f_i(h).$$

Since any eigenfunction ψ_h linked to $E_h \leq E$ will be written as $\psi_h = \sum_i \langle \psi_{h,i}, \psi_h \rangle \psi_{h,i}$ then the linearity of decomposition (1.4) ensures also the decomposition (1.6) for such ψ_h. Namely,

$$\psi_h = \sum_{\xi \in h\,\mathbb{Z}^n,\ |\xi| \leq g(h,E)} \int_{\mathbb{T}^n} \langle \Phi_{(x,\xi)}, \psi_h \rangle_{L^2}\, \Phi^{\star}_{(x,\xi)}\, dx + R_h$$

where $R_h = \sum_{1\le i\le \mathcal{N}(E,h)} R_{i,h}$. To conclude:

$$\|R_h\|_{L^2} \le \sum_{1\le i\le \mathcal{N}(E,h)} \|R_{i,h}\|_{L^2} \le \mathcal{N}(E,h) \max_{1\le i\le \mathcal{N}(E,h)} \|R_{i,h}\|_{L^2}$$
$$= \mathcal{N}(E,h)\cdot \mathcal{O}(h^\infty) = \mathcal{O}(h^\infty).$$

□

3.3 *Proof of Theorem 1.3*

We assume that $\varphi_h \in C^\infty(\mathbb{T}^n)$ is L^2-normalized and

$$\varphi_h = \sum_{1\le j\le J(h)} c_j\, \psi_{h,j}$$

where the L^2-normalized eigenfunctions $\psi_{h,j}$ of $\mathrm{Op}_h(b)$ are given in Theorem 1.2 and we assume $J(h) \le J_0 h^{-Q}$ for some $J_0, Q > 0$ that are independent on $0 < h \le 1$.

Define

$$\ell(h) := \max_{1\le j\le J(h)} f_j(h)$$

where $f_i(h)$ are associated to the functions $\psi_{h,i}$ and given by Theorem 1.1.

We now observe that if $U_h(t) := \exp\{(-i\mathrm{Op}_h(b)t)/h\}$ then

$$U(t)\varphi_h = \sum_{1\le j\le J(h)} c_j\, e^{-\frac{i}{h}E_{j,h}}\psi_{h,j}$$

for any $t \in \mathbb{R}$.

We can now apply the decomposition formula (1.4) with the condition on the frequencies $|\xi| \le \ell(h)$ and for the wave function $U(t)\varphi_h$ and get the expected result, namely

$$U_h(t)\varphi_h = \sum_{\xi\in h\,\mathbb{Z}^n,\ |\xi|\le \ell(h)} \int_{\mathbb{T}^n} \langle \Phi_{(x,\xi)}, U_h(t)\varphi_h\rangle_{L^2}\, \Phi^\star_{(x,\xi)}\, dx + \sum_{1\le j\le J} R_{j,h}$$

for any $t \in \mathbb{R}$. The remainder $R_h := \sum_{1\le j\le J} R_{j,h}$ can be estimated as in the previous Theorem, namely

$$\|R_h\|_{L^2} \le \sum_{1\le i\le J} \|R_{j,h}\|_{L^2} \le J_0\, h^{-Q} \max_{1\le j\le J} \|R_{j,h}\|_{L^2}$$
$$= J_0\, h^{-Q}\cdot \mathcal{O}(h^\infty) = \mathcal{O}(h^\infty).$$

□

References

1. J.P. Antoine, F. Bagarello, J.P. Gazeau (eds.), *Coherent States and Their Applications. A Contemporary Panorama*. Springer Proceedings in Physics, vol. 205 (Springer, Berlin, 2018)
2. A. Bouzouina, S. De Bievre, Equipartition of the eigenfunctions of quantized ergodic maps on the torus. Commun. Math. Phys. **178**, 83–105 (1996)
3. F. Cardin, L. Zanelli, The geometry of the semiclassical wave front set for Schrödinger eigenfunctions on the torus. Math. Phys. Anal. Geom. **20**(2), Art. 10, 20 (2017)
4. M. Combescure, D. Robert, *Coherent States and Applications in Mathematical Physics* (Springer, Berlin, 2012)
5. E. Diaz-Ortiz, C. Villegas-Blas, On a Bargmann transform and coherent states for the n-sphere. J. Math. Phys. **53**(6), 062103, 25 (2012)
6. F. Faure, S. Nonnenmacher, S. De Bievre, Scarred eigenstates for quantum cat maps of minimal periods. Commun. Math. Phys. **239**, 449–492 (2003)
7. M. Fremling, Coherent state wave functions on a torus with a constant magnetic field. J. Phys. A **46**(27), 275302, 23 (2013)
8. R. Fresneda, J.P. Gazeau, D. Noguera, Quantum localisation on the circle. J. Math. Phys. **59**(5), 052105, 19 (2018)
9. V. Guillemin, S. Sternberg, *Semi-Classical Analysis* (International Press, Boston, 2013)
10. A. Martinez, *Introduction to Semiclassical and Microlocal Analysis* (Springer, New York, 2002)
11. B.B. Martinez, R. Denk, J.H. Monzón, T. Nau, Generation of semigroups for vector-valued pseudodifferential operators on the torus. J. Fourier Anal. Appl. **22**, 823–853 (2016)
12. A. Parmeggiani, L. Zanelli, Wigner measures supported on weak KAM tori. J. Anal. Math. **123**, 107–137 (2014)
13. T. Paul, L. Zanelli, On the dynamics of WKB wave functions whose phase are weak KAM solutions of H-J equation. J. Fourier Anal. Appl. **20**, 1291–1327 (2014)
14. M. Ruzhansky, V. Turunen, Quantization of pseudo-differential operators on the torus. J. Fourier Anal. Appl. **16**, 943–982 (2010)
15. M. Ruzhansky, V. Turunen, *Pseudo-Differential Operators and Symmetries: Background Analysis and Advanced Topics* (Birkhäuser, Basel, 2010)
16. J. Wunsch, M. Zworski, The FBI transform on compact C^∞-manifolds. Trans. Amer. Math. Soc. **353**, 1151–1167 (2001)
17. L. Zanelli, On the optimal transport of semiclassical measures. Appl. Math. Optim. **74**, 325–342 (2016)
18. L. Zanelli, Schrödinger spectra and the effective Hamiltonian of the weak KAM theory on the flat torus. J. Math. Phys. **57**(8), 081507, 12 (2016)