W. Utschick · H. Boche · R. Ma

FOUNDATIONS IN
COMMUNICATION

VOLUME 4

Advanced To
System and

Foundations in Signal Processing, Communications and Networking

Series Editors: W. Utschick, H. Boche, R. Mathar

Foundations in Signal Processing, Communications and Networking

Series Editors: W. Utschick, H. Boche, R. Mathar

Vol. 1. Dietl, G. K. E.
Linear Estimation and Detection in Krylov Subspaces, 2007
ISBN 978-3-540-68478-7

Vol. 2. Dietrich, F. A.
Robust Signal Processing for Wireless Communications, 2008
ISBN 978-3-540-74246-3

Vol. 3. Stanczak, S., Wiczanowski, M. and Boche, H.
Fundamentals of Resource Allocation in Wireless Networks, 2009
ISBN 978-3-540-79385-4

Vol. 4. Pohl, V., Boche, H.
Advanced Topics in System and Signal Theory, 2009
ISBN 978-3-642-03639-2

Volker Pohl · Holger Boche

Advanced Topics in System and Signal Theory

A Mathematical Approach

With 5 Figures

Springer

Series Editors:
Wolfgang Utschick
TU Munich
Associate Institute for Signal Processing
Arcisstrasse 21
80290 Munich, Germany

Rudolf Mathar
RWTH Aachen University
Institute of Theoretical Information Technology
52056 Aachen, Germany

Prof. Dr.-Ing. Dr.rer.nat. Holger Boche
Technische Universität Berlin
Heinrich Hertz Chair for Mobile Communications
and
Fraunhofer German-Sino Lab for Mobile Communications
and
Fraunhofer Institute for Telecommunications
Heinrich Hertz Institute
Einsteinufer 37, 10587 Berlin
Germany
boche@hhi.fhg.de

Author:
Dr. Volker Pohl
Fraunhofer-Institut für Nachrichtentechnik
Heinrich-Hertz-Inst.
Einsteinufer 37
10587 Berlin
Germany
volker.pohl@mk.tu-berlin.de

Prof. Dr.-Ing. Dr.rer.nat. Holger Boche
Technische Universität Berlin
Heinrich Hertz Chair for Mobile Communications
and
Fraunhofer German-Sino Lab for Mobile Communications
and
Fraunhofer Institute for Telecommunications
Heinrich Hertz Institute
Einsteinufer 37, 10587 Berlin
Germany
boche@hhi.fhg.de

ISSN 1863-8538 e-ISSN 1863-8546
ISBN 978-3-642-03638-5 e-ISBN 978-3-642-03639-2
DOI 10.1007/978-3-642-03639-2
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2009934791

Cover design: eStudio Calamar S.L.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

The requirement of causality in system theory is inevitably accompanied by the appearance of certain mathematical operations, namely the Riesz projection, the Hilbert transform, and the spectral factorization mapping. A classical example illustrating this is the determination of the so-called Wiener filter (the linear, minimum means square error estimation filter for stationary stochastic sequences [88]). If the filter is not required to be causal, the transfer function of the Wiener filter is simply given by $H(\omega) = \Phi_{xy}(\omega)/\Phi_{xx}(\omega)$, where $\Phi_{xx}(\omega)$ and $\Phi_{xy}(\omega)$ are certain given functions. However, if one requires that the estimation filter is causal, the transfer function of the optimal filter is given by

$$H(\omega) = \frac{1}{[\Phi_{xx}]_{+}(\omega)} \mathfrak{P}_{+} \left(\frac{\Phi_{xy}(\omega)}{[\Phi_{xx}]_{-}(\omega)} \right) , \qquad \omega \in (-\pi, \pi] .$$

Here $[\Phi_{xx}]_{+}$ and $[\Phi_{xx}]_{-}$ represent the so called spectral factors of Φ_{xx}, and $\mathfrak{P}_{+}$ is the so called Riesz projection. Thus, compared to the non-causal filter, two additional operations are necessary for the determination of the causal filter, namely the *spectral factorization mapping* $\Phi_{xx} \mapsto ([\Phi_{xx}]_{+}, [\Phi_{xx}]_{-})$, and the *Riesz projection* $\mathfrak{P}_{+}$.

In applications the two functions $\Phi_{xx}(\omega)$ and $\Phi_{xy}(\omega)$ are usually not perfectly known but disturbed by measurement errors, or their values are only given at a finite number of sampling points $\{\omega_k\}_{k=1}^{N}$. The question arises, how these errors in the given data influence the calculation of the optimal filter $H(\omega)$. The answer will depend strongly on the metric in which the errors are measured, i.e. on the function spaces on which these problems are considered, and an answer requires the investigation of the continuity and boundedness of the involved operations (Riesz projection and spectral factorization) on the desired function spaces.

This monograph is intended primarily for engineers working on such robustness problems under a causality constraint. At the beginning, it presents the mathematical methods, necessary to approaching these problems. Then some related classical results concerning the boundedness and continuity of the Hilbert transform and Riesz projection are presented. Finally, these methods and results are applied to selected topics from signal processing.

The first part of the this monograph gives a very brief introduction to the main mathematical methods used later in the book. This part serves primarily as a review of results needed in later chapters, so that the work becomes essentially self-contained. The different topics are only covered as far as they will be needed and proofs are sometimes omitted. Appropriate reference are given for those who want a more detailed introduction to the different topics. This work presupposes a working knowledge of real and complex analysis (roughly as contained in [70]) as well as some basic elements of functional analysis (e.g. as in [54]).

The second part collects the basic abstract results concerning the continuity and the boundedness of the Hilbert transform and the Riesz projection on different Banach spaces. These results are the basis for the applications discussed in the third part of this monograph. Here four applications from signal processing are investigated in some detail, namely the expansions of transfer functions in orthonormal bases in Chapter 7, the linear approximation from measured data in Chapter 8, the calculation of the Hilbert transform in Chapter 9, and the spectral factorization in Chapter 10. All these topics are essentially problems of recovery or approximation of causal function from measured data, which are generally corrupted by small errors. It is investigated how these errors influence the possibility of recovering the desired signal from the measurements under the restriction that this recovery is based only on past and present measurements, i.e. under the requirement of causality.

The authors are very grateful to Brendan Farrell and Sander Wahls of the Technical University Berlin for reading the manuscript and for many helpful comments and suggestions on it. Parts of the book were written when the first author was staying at the Department of Electrical Engineering of the Technion in Haifa. He is very thankful to his host, Prof. Yonina Eldar and her whole group for their hospitality and help.

Haifa and Berlin
June 2009

Volker Pohl
Holger Boche

Contents

Part I

Mathematical Preliminaries

1

Function Spaces and Operators

This first chapter recalls shortly the most basic facts from analysis, and introduces the basic notations used throughout the whole book. The proofs of these standard results are almost always omitted and can be found in numerous textbooks.

1.1 Banach and Hilbert spaces

Let $\mathcal{X}$ be a complex vector space. A non-negative, real valued functional $\|\cdot\|_{\mathcal{X}}$ on $\mathcal{X}$ is said to be a *norm* on $\mathcal{X}$ if

1. $\|x\|_{\mathcal{X}} = 0$ if and only if $x = 0$
2. $\|\alpha x\|_{\mathcal{X}} = |\alpha| \, \|x\|_{\mathcal{X}}$
3. $\|x_1 + x_2\|_{\mathcal{X}} \leq \|x_1\|_{\mathcal{X}} + \|x_2\|_{\mathcal{X}}$

for all $x_1, x_2 \in \mathcal{X}$ and all scalars $\alpha \in \mathbb{C}$. A complex vector space $\mathcal{X}$ together with a certain norm $\|\cdot\|_{\mathcal{X}}$ is called a *normed vector space*.

A subset $\mathcal{Y} \subset \mathcal{X}$ is called a *subspace* of $\mathcal{X}$ if $\mathcal{Y}$ is itself a vector space. This is the case if and only if

$$0 \in \mathcal{Y} \qquad \text{and} \qquad \alpha\, y_1 + \beta\, y_2 \in \mathcal{Y}$$

for all scalars $\alpha, \beta \in \mathbb{C}$ and every $y_1, y_2 \in \mathcal{Y}$.

A vector space $\mathcal{Y}$ is said to be of *dimension* N if every $y \in \mathcal{Y}$ has a unique representation of the form $y = \alpha_1 \phi_1 + \cdots + \alpha_N \phi_N$ for a fixed collection $\{\phi_k\}_{k=1}^{N}$ of elements in $\mathcal{Y}$ and with certain $\alpha_1, \ldots, \alpha_N \in \mathbb{C}$. The dimension of a vector space may be finite or infinite.

Let $\{x_k\}_{k=1}^{\infty}$ be a sequence of elements of $\mathcal{X}$. If

$$\lim_{m,k\to\infty} \|x_m - x_k\|_{\mathcal{X}} = 0$$

then $\{x_k\}_{k=1}^{\infty}$ is said to be a *Cauchy sequence* in $\mathcal{X}$. The sequence $\{x_k\}_{k=1}^{\infty}$ *converges in* $\mathcal{X}$, if there exists an element $x \in \mathcal{X}$ such that

V. Pohl, H. Boche, *Advanced Topics in System and Signal Theory*, Foundations in Signal Processing, Communications and Networking 4,
DOI 10.1007/978-3-642-03639-2_1,

$$\lim_{k\to\infty} \|x - x_k\|_{\mathcal{X}} = 0 \; .$$

A normed linear space $\mathcal{X}$ is said to be *complete* if every Cauchy sequence converges in $\mathcal{X}$. A complete normed vector space is called a *Banach space.*

Let $\mathcal{H}$ be a complex vector space. A mapping $\langle \cdot, \cdot \rangle_{\mathcal{H}} : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$, which assigns to every pair of vectors in $\mathcal{H}$ a scalar, is called an *inner product* on $\mathcal{H}$ if for all $x_1, x_2, x_3 \in \mathcal{H}$ and all scalars $\alpha, \beta \in \mathbb{C}$ the following conditions are satisfied

1. $\langle x_1, x_2 \rangle_{\mathcal{H}} = \overline{\langle x_2, x_1 \rangle}_{\mathcal{H}}$
2. $\langle \alpha\, x_1 + \beta\, x_2, x_3 \rangle_{\mathcal{H}} = \alpha \, \langle x_1, x_3 \rangle_{\mathcal{H}} + \beta \, \langle x_2, x_3 \rangle_{\mathcal{H}}$
3. $\langle x_1, x_1 \rangle_{\mathcal{H}} \geq 0$ and $\langle x_1, x_1 \rangle_{\mathcal{H}} = 0$ if and only if $x_1 = 0$.

Such a vector space $\mathcal{H}$ together with a specific inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ is called an *inner product space.* The inner product on such a space induces a norm on $\mathcal{H}$ by

$$\|x\|_{\mathcal{H}} := \sqrt{\langle x, x \rangle_{\mathcal{H}}} \; .$$

If $\mathcal{H}$ is complete with respect to this norm, one says that $\mathcal{H}$ is a *Hilbert space.*

Example 1.1 (Euclidean Vector Spaces). Let $\mathbb{C}^N$ be the N-dimensional Euclidean vector space of N-tuples $x = (x_1, x_2, \cdots, x_N)$ of complex numbers and define by

$$\|x\|_{\mathbb{C}^N} := \sqrt{|x_1|^2 + |x_2|^2 + \cdots + |x_N|^2}$$

a norm on $\mathbb{C}^N$. Since $\mathbb{C}^N$ is finite dimensional, it is clear that $\mathbb{C}^N$ is a Banach space. On this Banach space one can also define an inner product by

$$\langle x, y \rangle_{\mathbb{C}^N} := x_1 \, \overline{y_1} + x_2 \, \overline{y_2} + \cdots + x_N \, \overline{y_N} \; .$$

for all $x, y \in \mathbb{C}^N$. This inner product is compatible with the above defined norm, which means that $\|x\|_{\mathbb{C}^N} = \sqrt{\langle x, x \rangle_{\mathbb{C}^N}}$. Therefore $\mathbb{C}^N$ is also a Hilbert space.

Example 1.2 (ℓ^p-Spaces). Let $1 \leq p \leq \infty$ be a real number and let $x = \{x_k\}_{k=-\infty}^{\infty}$ be a double infinite sequence of complex numbers and define

$$\|x\|_{\ell^p} := \begin{cases} \left(\sum_{k=-\infty}^{\infty} |x_k|^p\right)^{1/p} & \text{if } 1 \leq p < \infty \\ \sup_{k\in\mathbb{Z}} |x_k| & \text{if } p = \infty \; . \end{cases}$$

Then $\ell^p = \ell^p(\mathbb{C})$ denotes the Banach space of all double infinite complex sequences $x = \{x_k\}_{k=-\infty}^{\infty}$ for which the norm $\|x\|_{\ell^p}$ is finite, and $\ell^p_+ = \ell^p_+(\mathbb{C})$ is the Banach space of all infinite sequences $x = \{x_k\}_{k=0}^{\infty}$ with $\|x\|_{\ell^p} < \infty$.

On the particular spaces $\ell^2(\mathbb{C})$ and $\ell^2_+(\mathbb{C})$ one can define the inner product

$$\langle x, y \rangle_{\ell^2} := \textstyle\sum_{k=-\infty}^{\infty} x_k \, \overline{y_k}$$

which is compatible with the norm $\|\cdot\|_{\ell^p}$ and which makes $\ell^2(\mathbb{C})$ and $\ell^2_+(\mathbb{C})$ to Hilbert spaces.

Example 1.3 (L^p-Spaces). Let $\mathbb{T} := \{z \in \mathbb{C} : |z| = 1\}$ be the unit circle in the complex plane. Choose a number $1 \leq p < \infty$ and denote by L^p the set of all Lebesgue measurable functions f defined on $\mathbb{T}$ with

$$\|f\|_p := \left(\tfrac{1}{2\pi} \int_{-\pi}^{\pi} \left|f(\mathrm{e}^{\mathrm{i}\theta})\right|^p \,\mathrm{d}\theta\right)^{1/p} < \infty \,.$$

Together with this norm, L^p is a Banach space. The space L^∞ is the set of all essentially bounded Lebesgue measurable functions on $\mathbb{T}$ with the norm

$$\|f\|_\infty := \operatorname{ess\,sup}_{\theta \in [-\pi,\pi]} \left|f(\mathrm{e}^{\mathrm{i}\theta})\right| \,.$$

Or more generally, let μ be an arbitrary finite positive measure on the unit circle $\mathbb{T}$. Then $L^p(\mu)$ with $1 \leq p < \infty$ denotes the set of all measurable functions on $\mathbb{T}$ with $\|f\|_p := \int_{\mathbb{T}} |f(\zeta)|^p \,\mathrm{d}\mu(\zeta) < \infty$, and $L^\infty(\mu)$ is the set of all essentially bounded (with respect to μ) functions.

For the particular case $p = 2$ one defines for arbitrary $f, g \in L^2$ an inner product in L^2 by

$$\langle f, g\rangle_{L^2} := \left(\tfrac{1}{2\pi} \int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\theta})\, \overline{g(\mathrm{e}^{\mathrm{i}\theta})} \,\mathrm{d}\theta\right)^{1/2} \tag{1.1}$$

which is compatible with the norm on L^2. With this inner product L^2 becomes a Hilbert space.

Example 1.4. The set of all continuous functions f on the unit circle $\mathbb{T}$ equipped with the supremum (uniform) norm

$$\|f\|_\infty = \sup_{\zeta \in \mathbb{T}} |f(\zeta)|$$

is a Banach space and will be denoted by $\mathcal{C}(\mathbb{T})$.

Compared to Banach spaces, Hilbert spaces possess some nice geometrical properties which are very similar to the finite dimensional Euclidean spaces. A particularly useful concept is orthonormal bases.

Definition 1.5. *Let $\mathcal{H}$ be an inner product space. Two vectors $x, y \in \mathcal{H}$ are called* orthogonal *if $\langle x, y\rangle_{\mathcal{H}} = 0$. A family $S = \{x_k\}$ of non-zero vectors in $\mathcal{H}$ is called an* orthogonal set *if any two distinct vectors in this family are orthogonal. If in addition $\|x_k\|_{\mathcal{H}} = 1$ for all $x_k \in S$, the family S is called an* orthonormal set.

Let $\mathcal{M}$ be a subset of the Hilbert space $\mathcal{H}$. Then $\mathcal{M}^\perp$ denotes the set of all $y \in \mathcal{H}$ which are orthogonal to every $x \in \mathcal{M}$:

$$\mathcal{M}^\perp := \{\, y \in \mathcal{H} : \langle y, x\rangle_{\mathcal{H}} = 0 \text{ for every } x \in \mathcal{M} \,\} \,.$$

Note that an orthogonal set may contain a finite or an infinite number of elements. The notation $x \perp y$ will be used to indicate that the vectors x and y are orthogonal.

Theorem 1.6 (Bessel's inequality and equality). *Let $\{x_k\}_{k=1}^N$ be an orthonormal set in an inner product space $\mathcal{H}$ where N may be infinity. Then for every $x \in \mathcal{H}$ holds*

$$\sum_{k=1}^N |\langle x, x_k\rangle|_{\mathcal{H}}^2 \leq \|x\|_{\mathcal{H}}^2 \tag{1.2}$$

with equality for all x in the subspace spanned by $x_1, \cdots, x_N$. Moreover

$$\left\| x - \sum_{k=1}^N \langle x, x_k\rangle_{\mathcal{H}}\, x_k \right\|_{\mathcal{H}}^2 = \|x\|_{\mathcal{H}}^2 - \sum_{k=1}^N \left|\langle x, x_k\rangle_{\mathcal{H}}\right|^2 \tag{1.3}$$

for all $x \in \mathcal{H}$.

Assume that $S = \{x_k\}_{k=1}^\infty$ is an orthogonal sequence in an inner product space $\mathcal{H}$. For any $x \in \mathcal{H}$, we define its *generalized Fourier coefficients* (with respect to S) by $c_k := \langle x, x_k\rangle_{\mathcal{H}}$. Then we consider the truncated *generalized Fourier series* of x with respect to the orthonormal system S given by

$$s_N := \sum_{k=1}^N \langle x, x_k\rangle_{\mathcal{H}}\, x_k = \sum_{k=1}^N c_k\, x_k \ .$$

It is clear that $s_N \in \text{span}\{x_1, \cdots, x_N\}$. Therefore Theorem 1.6 shows that $\|s_N\|_{\mathcal{H}} = \sum_{k=1}^N |c_k|_{\mathcal{H}}^2$. Moreover, Bessel's inequality (1.2) implies that the infinite series $\sum_{k=1}^\infty |\langle x, x_k\rangle|_{\mathcal{H}}^2$ converges for every $x \in \mathcal{H}$ and its sum is upper bounded by $\|x\|_{\mathcal{H}}^2$. Thus the sequence $c = \{c_k\}_{k=1}^\infty$ is an element of ℓ^2 and consequently for all $N > M$, one has that

$$\|s_N - s_M\|_{\mathcal{H}}^2 = \sum_{k=M+1}^N |\langle x, x_k\rangle|_{\mathcal{H}}^2$$

which shows that s_N is a Cauchy sequence in $\mathcal{H}$. Therefore, if one additionally assumes that $\mathcal{H}$ is complete, i.e. if $\mathcal{H}$ is a Hilbert space, the partial sum s_N converges in $\mathcal{H}$ as $N \to \infty$. However s_N needs not to converge to x, in general. This only happens if the orthonormal system S is complete.

Definition 1.7 (Orthonormal base). *An orthonormal sequence $S = \{x_k\}_{k=1}^\infty$ in a Hilbert space $\mathcal{H}$ is called a* complete orthonormal system *or an* orthonormal basis *of $\mathcal{H}$ if the generalized Fourier series $\sum_{k=1}^\infty \langle x, x_k\rangle_{\mathcal{H}}\, x_k$ converges to x for every $x \in \mathcal{H}$.*

For a complete orthonormal sequence $S = \{x_k\}_{k=1}^\infty$, the left hand side of Bessel's equality (1.3) converges to zero as $N \to \infty$. This gives immediately the following result.

Theorem 1.8 (Parseval's identity). *Let $\{x_k\}_{k=1}^\infty$ be a complete orthonormal basis of a Hilbert space $\mathcal{H}$. Then*

$$\|x\|_{\mathcal{H}}^2 = \sum_{k=1}^\infty \left|\langle x, x_k\rangle_{\mathcal{H}}\right|^2 \tag{1.4}$$

for every $x \in \mathcal{H}$.

A Hilbert space is called *separable* if it contains a complete orthonormal sequence and it can be shown that every basis contains only a countable number of elements. Parsval's identity implies that every (infinite dimensional) separable Hilbert space $\mathcal{H}$ is *isometricly isomorphic* to the Hilbert space ℓ^2_+. To see this let $\{x_k\}_{k=1}^\infty$ be a complete orthonormal basis of $\mathcal{H}$ and define the linear operator $\mathfrak{T}\, x = \{c_k\}_{k=1}^\infty$, where $c_k = \langle x, x_k\rangle_{\mathcal{H}}$, for all $x \in \mathcal{H}$. Then Parseval's identity shows that $\mathfrak{T}$ maps $\mathcal{H}$ to ℓ^2_+ with $\|\mathfrak{T}\, x\|_{\ell^2} = \|x\|_{\mathcal{H}}$. Conversely every sequence $c = \{c_k\}_{k=1}^\infty \in \ell^2_+$ defines by $x_N = \sum_{k=1}^N c_k\, x_k$ a Cauchy sequence in $\mathcal{H}$. Since the basis $\{x_k\}_{k=1}^\infty$ is complete, x_N converges to an $x \in \mathcal{H}$ as $N \to \infty$, and again by Parseval's identity one has $\|x\|_{\mathcal{H}} = \|c\|_{\ell^2}$

The following example of an orthonormal system is the case we are most interested in throughout this book. This example will be discussed in more detail in Sec. 2.

Example 1.9 (Fourier series). Consider the Hilbert space $\mathcal{H} = L^2$ of Lebesque square-integrable functions on the unit circle $\mathbb{T}$ with the inner product (1.1). Define for $k = 0, \pm 1, \pm 2, \cdots$ the functions $\varphi_k(\mathrm{e}^{\mathrm{i}\theta}) := \mathrm{e}^{\mathrm{i}k\theta}$ in L^2. Then it is easy to verify that the set $\mathcal{S} = \{\varphi_k\}_{k=-\infty}^\infty$ is an orthonormal set in L^2. Moreover $\mathcal{S}$ is complete in L^2. In this case, for every $f \in L^2$ the coefficients

$$\hat{f}(k) := \langle f, \varphi_k\rangle_{L^2} = \tfrac{1}{2\pi} \int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\tau})\, \mathrm{e}^{-\mathrm{i}k\tau}\, \mathrm{d}\tau$$

are called *Fourier coefficients* of f and the formal series

$$\textstyle\sum_{k=-\infty}^\infty \hat{f}(k)\, \varphi_k = \sum_{k=-\infty}^\infty \langle f, \varphi_k\rangle_{L^2}\ \varphi_k$$

is the *Fourier series* of f, which converges in L^2 to f by Bessel's equality. If $\hat{f} = \{\hat{f}(k)\}_{k=-\infty}^\infty$ denotes the set of all Fourier coefficients of f, then Parseval's identity shows that $\|\hat{f}\|_{\ell^2} = \|f\|_{L^2}$, and the previous two equations establish an isometric isomorphism between L^2 and ℓ^2.

1.2 Operators on Banach spaces

Let $\mathcal{X}$ and $\mathcal{Y}$ be normed vector spaces. We consider mappings $\mathfrak{T} : x \mapsto y$, defined on a linear subspace $\mathscr{D}(\mathfrak{T})$ of $\mathcal{X}$, which assign to each $x \in \mathscr{D}(\mathfrak{T})$ an element $y = \mathfrak{T}(x) \in \mathcal{Y}$. The subspace $\mathscr{D}(\mathfrak{T}) \subset \mathcal{X}$ is called the *domain* of $\mathfrak{T}$ and

$$\begin{aligned} \mathscr{R}(\mathfrak{T}) &:= \{\ y \in \mathcal{Y}\ :\ y = \mathfrak{T}(x) \text{ for some } x \in \mathscr{D}(\mathfrak{T})\ \} \\ \mathscr{N}(\mathfrak{T}) &:= \{\ x \in \mathcal{X}\ :\ \mathfrak{T}(x) = 0\ \} \end{aligned}$$

denote the *range* and the *null space* (the *kernel*) of a mapping $\mathfrak{T} : \mathscr{D}(\mathfrak{T}) \to \mathcal{Y}$, respectively. It is easily verified that $\mathscr{R}(\mathfrak{T})$ and $\mathscr{N}(\mathfrak{T})$ are subspaces of $\mathcal{Y}$ and $\mathcal{X}$, respectively. If not mentioned otherwise, all mappings are always defined on the whole space $\mathcal{X}$ in the following. For us, the boundedness and continuity of such mappings will be of particular interest.

Definition 1.10 (Boundedness). *A mapping $\mathfrak{T} : \mathcal{X} \to \mathcal{Y}$ is said to be* bounded *if there exists a constant C such that*

$$\|\mathfrak{T}(x)\|_{\mathcal{Y}} \leq C \qquad \text{for all } x \in \mathcal{X} \text{ with } \|x\|_{\mathcal{X}} \leq 1 \,.$$

Definition 1.11 (Continuity). *A mapping $\mathfrak{T} : \mathcal{X} \to \mathcal{Y}$ is said to be* continuous at the point $x_0 \in \mathcal{X}$ *if to every $\epsilon > 0$ there exists a $\delta > 0$ such that*

$$\|\mathfrak{T}(x) - \mathfrak{T}(x_0)\|_{\mathcal{Y}} < \epsilon \qquad \text{for all } x \in \mathcal{X} \text{ with } \|x - x_0\|_{\mathcal{X}} < \delta \,.$$

$\mathfrak{T}$ is said to be continuous*, if it is continuous at every point $x \in \mathcal{X}$.*

The continuity of a mapping $\mathfrak{T} : \mathcal{X} \to \mathcal{Y}$ is in general a local property, i.e. it may be continuous at certain points $x \in \mathcal{X}$ but discontinuous at other points. Moreover the boundedness and the continuity of a mapping are in general two completely different properties which depend strongly on the vector spaces $\mathcal{X}$, $\mathcal{Y}$ and especially on the norm on these spaces. This general behavior will become evident in Section 10 where the spectral factorization mapping is investigated on different Banach spaces.

Definition 1.12. *A mapping $\mathfrak{T} : \mathcal{X} \to \mathcal{Y}$ is said to be* linear*, if*

$$\mathfrak{T}(\alpha\, x_1 + \beta\, x_2) = \alpha\, \mathfrak{T}(x_1) + \beta\, \mathfrak{T}(x_2)$$

for all $x_1, x_2 \in \mathcal{X}$ and all scalars α and β.

If the mapping $\mathfrak{T}$ is linear, we will also write $\mathfrak{T}\, x$ instated of $\mathfrak{T}(x)$. For linear mappings there is a direct relation between the boundedness and continuity of these mapping:

Theorem 1.13. *Let $\mathfrak{T} : \mathcal{X} \to \mathcal{Y}$ be a mapping between the normed vector spaces $\mathcal{X}$ and $\mathcal{Y}$. If $\mathfrak{T}$ is linear then the following statements are equivalent:*

(*a*) *$\mathfrak{T}$ is continuous at* 0.
(*b*) *$\mathfrak{T}$ is continuous.*
(*c*) *$\mathfrak{T}$ is bounded.*
(*d*) *There exists a constant C such that*

$$\|\mathfrak{T}x\|_{\mathcal{Y}} \leq C\, \|x\|_{\mathcal{X}} \qquad \text{for all } x \in \mathcal{X} \,. \tag{1.5}$$

Statement (d) of the previous theorem motivates the definition of the norm of a linear continuous mapping $\mathfrak{T} : \mathcal{X} \to \mathcal{Y}$ as the minimal possible constant C for which (1.5) holds.

Definition 1.14. *Let $\mathfrak{T} : \mathcal{X} \to \mathcal{Y}$ be a linear mapping. Then*

$$\|\mathfrak{T}\|_{\mathcal{X} \to \mathcal{Y}} := \sup_{\substack{x \in \mathcal{X} \\ x \neq 0}} \frac{\|\mathfrak{T}\, x\|_{\mathcal{Y}}}{\|x\|_{\mathcal{X}}} = \sup_{\substack{x \in \mathcal{X} \\ \|x\|_{\mathcal{X}} \leq 1}} \|\mathfrak{T}\, x\|_{\mathcal{Y}} \,. \tag{1.6}$$

is the called the operator norm *of $\mathfrak{T}$.*

The set of all bounded linear operators from $\mathcal{X}$ to $\mathcal{Y}$ together with the norm (1.6) *is again a normed linear space and will be denoted by $\mathscr{B}(\mathcal{X}, \mathcal{Y})$.*

One easily verifies that (1.6) indeed defines a norm on the set of all linear mappings from $\mathcal{X}$ to $\mathcal{Y}$. Moreover, if $\mathcal{Y}$ is a complete normed space then $\mathscr{B}(\mathcal{X},\mathcal{Y})$ is a Banach space as well.

Particularly important is the case where $\mathcal{Y} = \mathbb{C}$. Then $\mathscr{B}(\mathcal{X},\mathbb{C})$ contains all bounded linear functionals on $\mathcal{X}$. This space is called the *dual space* of $\mathcal{X}$ and it is usually denoted by $\mathcal{X}^*$. In the case of a Hilbert space $\mathcal{H}$, the dual $\mathcal{H}^*$ can be identified with the space $\mathcal{H}$ itself, since *Riesz representation theorem* states that to every bounded linear functional $c \in \mathcal{H}^*$ there exists a unique $f \in \mathcal{H}$ such that $c(h) = \langle h, f\rangle_{\mathcal{H}}$ for all $h \in \mathcal{H}$.

It is also worth to notice that the null space of a bounded linear mapping is a closed subspace of its domain.

Theorem 1.15. *Let $\mathfrak{T} : \mathcal{X} \to \mathcal{Y}$ be a bounded linear operator on a Banach space $\mathcal{X}$. Then $\mathscr{N}(\mathfrak{T})$ is a closed subspace of $\mathcal{X}$.*

Proof. Let $x_1, x_2 \in \mathscr{N}(\mathfrak{T})$, then $\mathfrak{T}(\alpha\, x_1 + \beta\, x_2) = \alpha\, \mathfrak{T}x_1 + \beta\, \mathfrak{T}x_2 = 0$ and since $0 \in \mathscr{N}(\mathfrak{T})$, $\mathscr{N}(\mathfrak{T})$ is a subspace of $\mathcal{X}$. It remains to show that $\mathscr{N}(\mathfrak{T})$ is closed. Let $\{x_n\}_{n\in\mathbb{N}}$ with $x_n \in \mathscr{N}(\mathfrak{T})$ for all $n \in \mathbb{N}$ and with $x_n \to x$ in $\mathcal{X}$. Then

$$\|\mathfrak{T}x\|_{\mathcal{Y}} = \|\mathfrak{T}x_n - \mathfrak{T}x\|_{\mathcal{Y}} \le \|\mathfrak{T}\|_{\mathcal{X}\to\mathcal{Y}}\, \|x_n - x\|_{\mathcal{X}} \;.$$

for every $n \in \mathbb{N}$. Since $\|x_n - x\|_{\mathcal{X}} \to 0$ and $\|\mathfrak{T}\|_{\mathcal{X}\to\mathcal{Y}} < \infty$ this implies $\mathfrak{T}x = 0$ and consequently $x \in \mathscr{N}(\mathfrak{T})$. □

Let $\mathcal{H}$ and $\mathcal{K}$ be two Hilbert spaces and let $\mathfrak{T} \in \mathscr{B}(\mathcal{H},\mathcal{K})$ be a bounded linear operator between $\mathcal{H}$ and $\mathcal{K}$. Then the *adjoint* $\mathfrak{T}^* : \mathcal{K} \to \mathcal{H}$ is defined by the relation

$$\langle \mathfrak{T}\, h, k\rangle_{\mathcal{K}} = \langle h, \mathfrak{T}^*\, k\rangle_{\mathcal{H}} \qquad \text{for all } h \in \mathcal{H} \text{ and } k \in \mathcal{K} \;.$$

It is not hard to prove that the adjoint $\mathfrak{T}^*$ is also a linear operator with $\|\mathfrak{T}^*\|_{\mathcal{K}\to\mathcal{H}} = \|\mathfrak{T}\|_{\mathcal{H}\to\mathcal{K}}$. The operator $\mathfrak{T}$ is called an *isometry*, if

$$\langle \mathfrak{T}\, h_1, \mathfrak{T}\, h_2\rangle_{\mathcal{K}} = \langle h_1, h_2\rangle_{\mathcal{H}} \quad \text{for all } h_1, h_2 \in \mathcal{H}$$

which is equivalent to $\mathfrak{T}^*\mathfrak{T} = \mathfrak{I}_{\mathcal{H}}$, where $\mathfrak{I}_{\mathcal{H}}$ denotes the identity mapping on $\mathcal{H}$. The operator $\mathfrak{T}$ is called *unitary* if $\mathfrak{T}$ is an isometry and if $\mathscr{R}(\mathfrak{T}) = \mathcal{K}$, which is equivalent to

$$\mathfrak{T}^*\, \mathfrak{T} = \mathfrak{I}_{\mathcal{H}} \quad \text{and} \quad \mathfrak{T}\, \mathfrak{T}^* = \mathfrak{I}_{\mathcal{K}} \;,$$

which means $\mathfrak{T}^* = \mathfrak{T}^{-1}$.

We will also need the notation of a derivative of a mapping $\mathfrak{T} : \mathcal{X} \to \mathcal{Y}$ from an arbitrary normed vector space $\mathcal{X}$ into an other normed vector space $\mathcal{Y}$. This derivative is defined in a way that is similar as for usual functions.

Definition 1.16 (Fréchet Derivative). *Let $\mathcal{X}$ and $\mathcal{Y}$ be two normed linear spaces, and let $\mathcal{V}$ be an open subset of $\mathcal{X}$. A mapping $\mathfrak{T} : \mathcal{V} \to \mathcal{Y}$ is said to* *be (Fréchet) differentiable at $x \in \mathcal{V}$ if there exists a bounded linear mapping $\mathfrak{A} : \mathcal{X} \to \mathcal{Y}$ such that*

$$\lim_{h \to 0} \frac{\|\mathfrak{T}(x+h) - \mathfrak{T}(x) - \mathfrak{A}\, h\|_{\mathcal{Y}}}{\|h\|_{\mathcal{X}}} = 0 \,. \tag{1.7}$$

The linear operator $\mathfrak{A}$ is called the (Fréchet) derivative of $\mathfrak{T}$ at x *and its value at h will also be denoted by $\mathfrak{A}\, h = \mathfrak{T}'(x)\, h$.*

Thus the (Fréchet) derivative of a mapping $\mathfrak{T} : \mathcal{X} \to \mathcal{Y}$ is a bounded linear operator $\mathfrak{T}'(x) : \mathcal{X} \to \mathcal{Y}$. In usual calculus, the derivative of a function f is a number $f'(x)$ which represents the slope of the tangent on the graph of f at a point x. Thus, it gives the best linear approximation of f in a neighborhood of x. Similarly, the Fréchet derivative gives the best linear approximation $\mathfrak{T}'(x) : \mathcal{X} \to \mathcal{Y}$ of the operator $\mathfrak{T} : \mathcal{X} \to \mathcal{Y}$ in a neighborhood of x. If $\mathfrak{T}$ is linear, it follows from the above definition that $\mathfrak{T}'(x)\, h = \mathfrak{T}\, h$, i.e. the Fréchet derivative of $\mathfrak{T}$ is equal to $\mathfrak{T}$. Thus, the best linear approximation of a linear operator is the operator itself. For the special case $\mathcal{X} = \mathcal{Y} = \mathbb{C}$, the above definition consides with the usual derivate of complex functions. It is not hard to see that if the Fréchet derivative of $\mathfrak{T}$ exists, it is unique. If $\mathfrak{T}$ is Fréchet differentiable at some point x, $\mathfrak{T}$ is continuous at x.

Proposition 1.17. *If a mapping $\mathfrak{T} : \mathcal{X} \to \mathcal{Y}$, defined in an open subset $\mathcal{V}$ of $\mathcal{X}$, has a Fréchet derivative at a point $x \in \mathcal{V}$, then it is continuous at x.*

Proof. Since $\mathcal{V}$ is open and $x \in \mathcal{V}$, the triangle inequality gives

$$\|\mathfrak{T}(x+h) - \mathfrak{T}(x)\|_{\mathcal{Y}} \le \|\mathfrak{T}(x+h) - \mathfrak{T}(x) - \mathfrak{T}'(x)\, h\|_{\mathcal{Y}} + \|\mathfrak{T}'(x)\|_{\mathcal{Y}}\, \|h\|_{\mathcal{X}}$$

for all h in a certain neighborhood of zero, where we have also used that the Fréchet derivative $\mathfrak{T}'(x)$ is a bounded operator. Since $\mathfrak{T}$ has a Fréchet derivative at x, for every $\epsilon > 0$ there exists a $\delta > 0$ such that $\|\mathfrak{T}(x+h) - \mathfrak{T}(x) - \mathfrak{T}'(x)\, h\|_{\mathcal{Y}} \le \epsilon\, \|h\|_{\mathcal{X}}$ for all $\|h\|_{\mathcal{X}} \le \delta$. Therewith, the above inequality becomes

$$\|\mathfrak{T}(x+h) - \mathfrak{T}(x)\|_{\mathcal{Y}} \le (\epsilon + \|\mathfrak{T}'(x)\|_{\mathcal{Y}})\, \|h\|_{\mathcal{X}}$$

which shows that $\mathfrak{T}$ is continuous at x. □

Subsequently, we will be mainly interested in the case where the mapping $\mathfrak{T}$ is defined on subsets Ω of the complex numbers $\mathbb{C}$, i.e. where $\mathcal{X} = \mathbb{C}$. Thus, we will consider functions over $\mathbb{C}$ with values in a certain normed vector space $\mathcal{Y}$. In this case, all linear mappings $\mathfrak{A} : \mathbb{C} \to \mathcal{Y}$ have the form $\mathfrak{A}\, h = A\, h$ with a certain vector $A \in \mathcal{Y}$. Therefore, the Fréchet derivative of $\mathfrak{T}$ can be identified with a corresponding vector $\mathfrak{T}'(x) = A(x) \in \mathcal{Y}$, and this vector will be called the (Fréchet) derivative of $\mathfrak{T}$. Moreover, if the function $\mathfrak{T} : \mathbb{C} \to \mathcal{Y}$ is (Fréchet) differentiable at every point $z \in \Omega$, one says that $\mathfrak{T}$ is *holomorphic* (complex differentiable) in Ω.

1.3 Spaces of Smooth Functions

Let $\Omega \subset \mathbb{C}$ be a compact set in the complex plane and let $f : \Omega \to \mathbb{C}$ be a complex function defined on Ω. Then the real function $\omega_f(\delta)$ defined by

$$\omega_f(\delta) := \sup_{|z_1 - z_2| \le \delta} |f(z_1) - f(z_2)| \quad \text{for all } z_1, z_2 \in \Omega$$

is called the *modulus of continuity* of f. As we will see later, this modulus of continuity of a function plays an important roll in analysis since it characterizes certain properties of f. For example, ω_f determines how fast the Fourier coefficients $\hat{f}(k)$ of f decrease as $|k|$ increases and in turn how well f can be approximated by polynomials. We state some almost obvious properties of the modulus of continuity.

(a) $\omega_f(\delta_1) \le \omega_f(\delta_2)$ whenever $\delta_1 < \delta_2$.
(b) A function f is uniformly continuous if and only if $\omega_f(\delta) \to 0$ as $\delta \to 0$.
(c) Let $N \in \mathbb{N}$ be a natural number, then $\omega_f(N\,\delta) \le N\,\omega_f(\delta)$. Indeed, assume that $|z_1 - z_2| \le N\,\delta$, and divide the line connecting z_1 and z_2 into N equal intervals $(z^{(k-1)}, z^{(k)})$ with $z^{(k)} = z_1 + k\,(z_2 - z_1)/N$ and with $k = 1, 2, \ldots, N$. Then $|z^{(k)} - z^{(k-1)}| \le \delta$ for all k and consequently

$$|f(z_1) - f(z_2)| \le \sum_{k=1}^{N} \left| f(z^{(k)}) - f(z^{(k-1)}) \right| \le N\,\omega(\delta) \ ,$$

from what follows that

$$\omega_f(N\,\delta) = \sup_{|z_1 - z_2| \le N\,\delta} |f(z_1) - f(z_2)| \le N\,\omega(\delta) \ .$$

(d) Let α be an arbitrary positive real number. Then $\omega_f(\alpha\,\delta) \le (\alpha+1)\,\omega_f(\delta)$. Indeed let N be the largest integer not exceeding α, i.e. $N \le \alpha < (N+1)$. Then it follows from properties (a) and (b) that

$$\omega_f(\alpha\,\delta) \le \omega_f([N+1]\,\delta) \le (N+1)\,\omega_f(\delta) \le (\alpha+1)\,\omega_f(\delta) \ .$$

These properties of the modulus of continuity motivate the following definition:

Definition 1.18. *A continuous, increasing, real valued function $\omega(\tau)$ defined on the interval $[0, \pi]$ is called a* majorant *if $\omega(0) = 0$ and if the function $\omega(\tau)/\tau$ is non increasing.*

Given a majorant ω and a bounded domain $\Omega \subset \mathbb{C}$, one denotes by $\mathcal{C}_\omega(\Omega)$ the set of all functions $f : \Omega \to \mathbb{C}$ for which

$$\|f\|_\omega := \sup_{z_1 \ne z_2} \frac{|f(z_1) - f(z_2)|}{\omega\left(|z_1 - z_2|\right)} < \infty \ . \tag{1.8}$$

It is clear from this definition that every $f \in \mathcal{C}_\omega(\Omega)$ is continuous on Ω because there exists a constant C such that $|f(z_1) - f(z_2)| \le C\,\omega(|z_1 - z_2|)$ for all $z_1, z_2 \in \Omega$, and since $\omega(\tau) \to 0$ as $\tau \to 0$ it follows that f is continuous. Moreover there exists a constant $C(\omega)$, dependent only on the majorant ω, such that $\|f\|_\infty \le C(\omega)\,\|f\|_\omega$ for all $f \in \mathcal{C}_\omega(\Omega)$. Indeed, let $f \in \mathcal{C}_\omega(\Omega)$ and define for some arbitrary $z_0 \in \Omega$ the function $f_0(z) := f(z) - f(z_0)$. For this function holds obviously $f_0(z_0) = 0$, $\|f_0\|_\omega = \|f\|_\omega$, and $\|f\|_\infty \le 2\|f_0\|_\infty$. Therewith, one obtains for an arbitrary $z \in \Omega$

$$|f_0(z)| = |f_0(z) - f_0(z_0)| \le \|f_0\|_\omega\,\omega(|z - z_0|) \le \omega(|\Omega|)\,\|f_0\|_\omega$$

where $|\Omega| = \sup_{z_1,z_2\in\Omega} |z_1 - z_2|$ which is finite since Ω is compact. Altogether one obtains $\|f\|_\infty \le C(\omega)\,\|f\|_\omega$. However, the functional (1.8) defines not a norm on $\mathcal{C}_\omega(\Omega)$ since for every constant function f, one has $\|f\|_\omega = 0$. For this reason, one defines the norm on $\mathcal{C}_\omega(\Omega)$ by

$$\|f\|_{\mathcal{C}_\omega(\Omega)} := \|f\|_\infty + \|f\|_\omega = \|f\|_\infty + \sup_{z_1 \ne z_2} \frac{|f(z_1) - f(z_2)|}{\omega\left(|z_1 - z_2|\right)} \tag{1.9}$$

which is finite for every $f \in \mathcal{C}_\omega(\Omega)$ since $\|f\|_{\mathcal{C}_\omega(\Omega)} \le [1 + C(\omega)]\,\|f\|_\omega$. It follows in particular that $\mathcal{C}_\omega(\Omega)$ equipped with the norm (1.9) is continuously embedded into the space $\mathcal{C}(\Omega)$ of all continuous functions on Ω with $\|f\|_\infty \le \|f\|_{\mathcal{C}_\omega(\Omega)}$. Moreover it follows from (1.9) that $\|f\|_\omega \le \|f\|_{\mathcal{C}_\omega(\Omega)}$ such that for every $f \in \mathcal{C}_\omega(\Omega)$ the definition (1.8) gives

$$|f(z_1) - f(z_2)| \le \|f\|_{\mathcal{C}_\omega(\Omega)}\,\omega\left(|z_1 - z_2|\right) . \tag{1.10}$$

This relation shows that the modulus of continuity ω_f of every $f \in \mathcal{C}_\omega(\Omega)$ is upper bounded by the majorant ω:

$$\omega_f(\delta) \le \|f\|_{\mathcal{C}_\omega(\Omega)}\,\omega(\delta) . \tag{1.11}$$

Thus the majorant ω characterizes the smoothness of the functions in $\mathcal{C}_\omega(\Omega)$.

For a given majorant ω, the space $\mathcal{C}_\omega(\Omega)$ will be not separable, in general. Therefore, we will also need the separable subspace $\mathcal{C}_{\omega,0}(\Omega)$ which is defined as the closure of all polynomials defined in Ω with respect to the corresponding $\mathcal{C}_\omega$-norm (1.9).

We will distinguish between different classes of majorants.

Definition 1.19. *A majorant ω is called* regular *if there exists a constant C such that*

$$\int_0^x \frac{\omega(\tau)}{\tau}\,\mathrm{d}\tau + x \int_x^\pi \frac{\omega(\tau)}{\tau^2}\,\mathrm{d}\tau \le C\,\omega(x), \quad 0 < x < 1 .$$

Moreover, we say that a majorant ω is weak regular of type 1 *(or* fast*) if there exists a constant C such that*

$$\int_0^x \frac{\omega(\tau)}{\tau}\, \mathrm{d}\tau \le C\,\omega(x), \quad 0 < x < 1 \tag{1.12}$$

and ω is said to be weak regular of type 2 *(or* slow*) if there exists a constant C such that*

$$x\int_x^\pi \frac{\omega(\tau)}{\tau^2}\, \mathrm{d}\tau \le C\,\omega(x), \quad 0 < x < 1\ . \tag{1.13}$$

Thus a majorant ω is regular if it is weak regular of type 1 and weak regular of type 2. Conversely, it is clear that every regular majorant is also weak regular of type 1 and type 2.

Example 1.20 (Hölder continuous functions Λ_α). Consider the majorant $\omega(\tau) = \tau^\alpha$ with $0 < \alpha < 1$. It is easily verified that

$$\int_0^x \frac{\omega(\tau)}{\tau}\, \mathrm{d}\tau = \frac{1}{\alpha}\,\omega(x) \quad \text{and} \quad x\int_x^\pi \frac{\omega(\tau)}{\tau^2}\, \mathrm{d}\tau \le \frac{1}{1-\alpha}\,\omega(x)\ .$$

The first equality shows that ω is weak regular of type 1, whereas the second inequality proves that ω is also weak regular of type 2. Therefore ω is a regular majorant. The space $\mathcal{C}_\omega(\Omega)$ with this special majorant will be denoted by $\Lambda_\alpha(\Omega)$. It is the set of all *Hölder continuous functions of exponent* α on Ω.

Example 1.21 (Lipschitz continuous functions Lip_K). Consider now the majorant $\omega(\tau) = \tau$. Exactly as in the previous example, one verifies that ω is weak regular of type 1. However, now it holds that

$$x\int_x^\pi \frac{\omega(\tau)}{\tau^2}\, \mathrm{d}\tau = [\log(1/x) + \log\pi]\ \omega(x)$$

which shows that there exists no constant C such that (1.13) is satisfied. Therefore ω is not weak regular of type 2. With this particular majorant $\mathcal{C}_\omega(\Omega)$ is denoted by $\Lambda_1(\Omega)$ and the elements of $\Lambda_1(\Omega)$ are *Lipschitz continuous functions* on Ω.

In the case of Lipschitz continuous functions, the proportionality constant (cf. (1.11)) between the majorant ω and the actual modulus of continuity ω_f is sometimes of importance. Whereas $\Lambda_1(\Omega)$ contains all functions f for which there exits a constant $C < \infty$ such that

$$|f(z_1) - f(z_2)| \le C\ |z_1 - z_2|\ , \qquad z_1, z_2 \in \Omega \tag{1.14}$$

we define the set $\mathrm{Lip}_K(\Omega)$ as the set of all functions f which satisfy (1.14) with $C = K$. Finally, we note that every function f differentiable on Ω belongs to $\Lambda_1(\Omega)$ and every differentiable function f whose first derivative $|f'(z)|$ is upper bounded for all $z \in \Omega$ by M belongs to $\mathrm{Lip}_M(\Omega)$.

We will also need the following generalization of the class Λ_α of Hölder continuous functions. Let Ω be an open set in $\mathbb{C}$, and let $\alpha = k + \beta$ be a positive real number with an integer $k \geq 0$ and with $0 < \beta \leq 1$. Then the *Hölder-Zygmund class* $\Lambda_\alpha(\Omega)$ is defined as the set of all k-times continuously differentiable functions $f : \Omega \to \mathbb{C}$ for which the k-th derivative $f^{(k)}$ satisfies a *Hölder (-Zygmund) condition* of order β, i.e. for which

$$
\begin{aligned}
|f^{(k)}(\theta+\tau) - f^{(k)}(\theta)| &\leq C\,|\tau|^\beta \quad \text{for} \quad 0<\beta<1 \\
|f^{(k)}(\theta+\tau) + f^{(k)}(\theta-\tau) - 2\,f^{(k)}(\tau)| &\leq C\,|\tau| \quad\;\; \text{for} \quad \beta = 1\ .
\end{aligned}
$$

and for all $\theta \in \Omega$.

Notes

The material in this section can be found in any introductory analysis or functional analysis textbook, e.g. [28, 70] The classes of smooth function characterized by regular majorants as described in Sect. 1.3 where extensively used by Dyakonov in [32] and the distinction between fast and slow majorants (cf. Def. 1.19) follows [31].

2

Fourier Analysis and Analytic Functions

2.1 Trigonometric Series

One of the most important tools for the investigation of linear systems is Fourier analysis. Let $f \in L^1$ be a complex-valued Lebesgue-integrable function on the unit circle $\mathbb{T}$. Then the *Fourier coefficients* of f are the complex numbers

$$\hat{f}(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\theta})\, \mathrm{e}^{-\mathrm{i}n\theta}\, \mathrm{d}\theta\,, \qquad n = 0, \pm 1, \pm 2, \dots\,. \tag{2.1}$$

Let $\mathcal{B}$ be an arbitrary subspace of L^1. Then we define by

$$\mathcal{B}_+ := \{\, f \in \mathcal{B} : \hat{f}(n) = 0 \text{ for all } n < 0 \,\}$$

the subspace of all functions in $\mathcal{B}$ for which all Fourier coefficients with negative index are equal to zero.

Definition 2.1 (Conjugate function). *Let* $\mathcal{B} \subset L^1$ *and* $f \in \mathcal{B}$. *The function* $\widetilde{f} \in \mathcal{B}$ *is called the* conjugate function of f in $\mathcal{B}$ *if it satisfies the conditions*

$$f + \mathrm{i}\widetilde{f} \in \mathcal{B}_+ \qquad \textit{and} \qquad \tfrac{1}{2\pi} \int_{-\pi}^{\pi} \widetilde{f}(\mathrm{e}^{\mathrm{i}\theta})\, \mathrm{d}\theta = 0\,.$$

In general, a function $f \in \mathcal{B}$ need not possess a conjugate function according to the above definition. However, for every $f \in \mathcal{B} \subset L^1$ it is always possible to define a conjugate functions $\widetilde{f}$ which exists almost everywhere on $\mathbb{T}$ but which does not necessarily belong to $\mathcal{B}$ (see e.g. [41] and also the discussion in Section 5). The second condition on $\widetilde{f}$ in the above definition is only a normalization of $\widetilde{f}$ ensuring that $\widetilde{f}$ is unique if it exists.

The importance of the Fourier coefficients (2.1) originates from the fact that they determine the function f uniquely. In saying this, f is considered as an element of L^1 and functions which differ only on a set of Lebesgue measure zero are identified as equivalent. The interesting question is now, how can we

V. Pohl, H. Boche, *Advanced Topics in System and Signal Theory*, Foundations in Signal Processing, Communications and Networking 4,
DOI 10.1007/978-3-642-03639-2_2, © Springer-Verlag Berlin Heidelberg 2009

recapture f and its conjugate function $\widetilde{f}$ from the Fourier coefficients $\hat{f}(n)$. This is usually done by partial sums of the form

$$(S_N^{(w)} f)(\mathrm{e}^{\mathrm{i}\theta}) = \sum_{n=-N}^{N} w(n)\, \hat{f}(n)\, \mathrm{e}^{\mathrm{i}n\theta} \tag{2.2}$$

in which the sequence $\{w(n)\}_{n=-N}^{N}$ of complex numbers is an arbitrary *window function* which weights the Fourier coefficients. The series

$$(\widetilde{S}_N^{(w)} f)(\mathrm{e}^{\mathrm{i}\theta}) = \sum_{n=-N}^{N} -\mathrm{i}\ \mathrm{sgn}(n)\, w(n)\, \hat{f}(n)\, \mathrm{e}^{\mathrm{i}n\theta} \tag{2.3}$$

with the sign function

$$\mathrm{sgn}(k) = \begin{cases} 1\,, & k > 0 \\ 0\,, & k = 0 \\ -1\,, & k < 0 \end{cases}$$

and with the same window $\{w(n)\}_{n=-N}^{N}$ and the same degree N is called the *series conjugate* to $S_N^{(w)} f$. The usage of the name conjugate series is justified by Definition 2.1 and by the fact that series

$$(S_N^{(w)} f)(\mathrm{e}^{\mathrm{i}\theta}) + \mathrm{i}\,(\widetilde{S}_N^{(w)} f)(\mathrm{e}^{\mathrm{i}\theta}) = w(0)\, \hat{f}(0) + 2\sum_{k=1}^{N} w(n)\, \hat{f}(n)\, \mathrm{e}^{\mathrm{i}n\theta} \tag{2.4}$$

has only nonnegative Fourier coefficients.

The question is whether the series (2.2) and (2.3) converge as N tends to infinity, and if they do converge, do they converge to f and $\widetilde{f}$, respectively? The answer depends on the window $\{w(n)\}_{n=-N}^{N}$ and on the actual topology, i.e. one may ask whether (2.2) and (2.3) converge pointwise, uniformly, or in some type of norm to f and $\widetilde{f}$, respectively. Here, we discuss only some of the most important windows $\{w(n)\}_{n=-N}^{N}$ which are needed in this book. All of these windows will be symmetric, in the sense that $w(-n) = w(n)$ for all n.

To investigate the convergence behavior of (2.2) and (2.3), it is sometimes more convenient to write those series in closed form by inserting the Fourier coefficients (2.1) into (2.2) and (2.3). This gives the following integral representations

$$(S_N^{(w)} f)(\mathrm{e}^{\mathrm{i}\theta}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\tau})\, K_N^{(w)}(\theta - \tau)\, \mathrm{d}\tau \tag{2.5}$$

and

$$(\widetilde{S}_N^{(w)} f)(\mathrm{e}^{\mathrm{i}\theta}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\tau})\, \widetilde{K}_N^{(w)}(\theta - \tau)\, \mathrm{d}\tau\ . \tag{2.6}$$

with the kernels

$$K_N^{(w)}(\tau) = w(0) + 2\sum_{n=1}^{N} w(n)\cos(n\tau) \tag{2.7}$$

$$\widetilde{K}_N^{(w)}(\tau) = 2\sum_{n=1}^{N} w(n)\sin(n\tau). \tag{2.8}$$

Of course, these kernels depend on the (symmetric) window $\{w(n)\}_{n=-N}^{N}$ which is indicated by the superscript (w). It is immediately clear that the kernels (2.7) and (2.8) are 2π-periodic. Next, three special windows and the corresponding series are discussed.

2.1.1 Fourier Series

The most simple and best known window is the rectangular window given by

$$w(n) = 1\,, \qquad n = 0, \pm 1, \pm 2, \ldots, \pm N\ . \tag{2.9}$$

In this case, the partial sums (2.2) and (2.3) are just the truncated *Fourier* and the *conjugate Fourier series* of f, and will be denoted by s_N and $\widetilde{s}_N$ respectively. The kernel (2.7) of this particular series is called the *Dirichlet kernel* given by

$$\mathcal{D}_N(\tau) = 1 + 2\sum_{n=1}^{N}\cos(n\tau) = \frac{\sin\left([N+\frac{1}{2}]\tau\right)}{\sin(\tau/2)}\ . \tag{2.10}$$

Similarly, the kernel (2.8) of the conjugate series is called *conjugate Dirichlet kernel* given by

$$\widetilde{\mathcal{D}}_N(\tau) = 2\sum_{n=1}^{N}\sin(n\tau) = \frac{1}{\tan(\tau/2)} - \frac{\cos\left([N+\frac{1}{2}]\tau\right)}{\sin(\tau/2)}\ .$$

Therewith the partial Fourier series are given by

$$s_N(f;\mathrm{e}^{\mathrm{i}\theta}) = \sum_{n=-N}^{N} \hat{f}(n)\,\mathrm{e}^{\mathrm{i}n\theta} = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\tau})\,\mathcal{D}_N(\theta-\tau)\,\mathrm{d}\tau \tag{2.11}$$

and likewise for $\widetilde{s}_N$.

If the function f belongs to L^2 then it is well known that the partial sums s_N and $\widetilde{s}_N$ of the Fourier series converge to f and $\widetilde{f}$ in the L^2-norm, respectively (cf. Example 1.9). However, if f belongs only to L^1, the partial sums s_N and $\widetilde{s}_N$ do not converge to f and $\widetilde{f}$ in the L^1-norm, in general. Moreover if f is a continuous function on $\mathbb{T}$, one might require that s_N and $\widetilde{s}_N$ converge uniformly to f and $\widetilde{f}$, respectively, i.e. that

$$\lim_{N\to\infty}\|f - s_N(f;\cdot)\|_\infty = 0 \qquad \text{and} \qquad \lim_{N\to\infty}\|\widetilde{f} - \widetilde{s}_N(f;\cdot)\|_\infty = 0$$

for all $f \in \mathcal{C}(\mathbb{T})$. However, the partial sums s_N and $\widetilde{s}_N$ do not show such nice behavior. It even happens that for some continuous functions f, the truncated Fourier series $s_n(f;\cdot)$ does not even converge pointwise. More precisely, one has the following result.

Theorem 2.2. *To every $\theta \in [-\pi, \pi)$, there corresponds a set $E(\theta) \subset \mathcal{C}(\mathbb{T})$ of second category which is dense in $\mathcal{C}(\mathbb{T})$ such that*

$$\sup_{N \in \mathbb{N}} |s_N(f; \mathrm{e}^{\mathrm{i}\theta})| = \infty$$

for every $f \in E(\theta)$.

Proof. Let $\theta \in [-\pi, \pi)$ be arbitrary but fixed. Then for every N (2.11) defines a linear functional $s_N(f; \mathrm{e}^{\mathrm{i}\theta}) = \mathfrak{D}_N f$ on $\mathcal{C}(\mathbb{T})$, for which holds

$$|\mathfrak{D}_N f| \;\leq\; \|f\|_\infty \tfrac{1}{2\pi} \textstyle\int_{-\pi}^{\pi} |\mathcal{D}_N(\tau)| \, \mathrm{d}\tau \;=\; \|f\|_\infty \, \|\mathcal{D}_N\|_1$$

which shows that $\|\mathfrak{D}_N\| \leq \|D_N\|_1$. Actually, equality holds. To see this, we define the function

$$g_N(\mathrm{e}^{\mathrm{i}\tau}) := \begin{cases} 1 & \text{for all } \tau \text{ for which } \mathcal{D}_N(\theta - \tau) \geq 0 \\ -1 & \text{for all } \tau \text{ for which } \mathcal{D}_N(\theta - \tau) < 0 \end{cases}$$

for which certainly holds $\|g_N\|_\infty = 1$ and $|\mathfrak{D}_N \, g_N| = \|\mathcal{D}_N\|_1$ for all $N \in \mathbb{N}$. Moreover, by Lusin's Theorem (see e.g. [70, §2.24]), for every $\epsilon > 0$ there exists an $f_N \in \mathcal{C}(\mathbb{T})$ with $\|f_N\|_\infty \leq 1$ such that

$$\|f_N - g_N\|_\infty = \operatorname*{ess\,sup}_{\zeta \in \mathbb{T}} |f_N(\zeta) - g_N(\zeta)| < \epsilon \, .$$

Therewith, one obtains

$$\begin{aligned} |\mathfrak{D}_N \, f_N| &= |\mathfrak{D}_N \, g_N - \mathfrak{D}_N(g_N - f_N)| \geq |\mathfrak{D}_N \, g_N| - |\mathfrak{D}_N(g_N - f_N)| \\ &\geq \|\mathcal{D}_N\|_1 - \epsilon \, \|\mathcal{D}_N\|_1 \end{aligned}$$

such that

$$\|\mathfrak{D}_N\| \;=\; \sup_{\substack{f \in \mathcal{C}(\mathbb{T}) \\ \|f\|_\infty \leq 1}} |\mathfrak{D}_N f| \;\geq\; |\mathfrak{D}_N f_N| \;\geq\; (1 - \epsilon) \, \|\mathcal{D}_N\|_1$$

which shows that $\|\mathfrak{D}_N\| = \|\mathcal{D}_N\|_1$. Next, it is shown that $\|\mathcal{D}_N\|_1$ diverges as $N \to \infty$. Since the kernel $\mathcal{D}_N$, given by (2.10), is an even function and since $\sin(\tau/2) \leq \tau/2$ for all $0 \leq \tau < \pi$, one obtains

$$\begin{aligned} \|\mathcal{D}_N\|_1 &\geq \frac{2}{\pi} \int_0^{\pi} |\sin([N + 1/2]\tau)| \, \frac{\mathrm{d}\tau}{\tau} = \frac{2}{\pi} \int_0^{[N+1/2]\pi} |\sin(\tau)| \, \frac{\mathrm{d}\tau}{\tau} \\ &> \frac{2}{\pi} \sum_{k=1}^{N} \int_{[k-1]\pi}^{k\pi} \frac{|\sin(\tau)|}{\tau} \, \mathrm{d}\tau \geq \frac{2}{\pi} \sum_{k=1}^{N} \frac{2}{k\,\pi} \geq \frac{4}{\pi^2} \, \log(N+1) \quad (2.12) \end{aligned}$$

which shows that $\|\mathfrak{D}_N\| = \|\mathcal{D}_N\|_1 \to \infty$ as $N \to \infty$. This divergence of the operator norm $\|\mathfrak{D}_N\|$ implies by the Banach-Steinhaus theorem (see e.g. [70, §5.8]) that there exists a dense subset $E(\theta) \subset \mathcal{C}(\mathbb{T})$ of second category such that $\sup_{N\in\mathbb{N}} |\mathfrak{D}_N f| = \infty$ for all $f \in E(\theta)$. □

The previous proof showed that the divergence of the Fourier series $s_N(f;\mathrm{e}^{\mathrm{i}\theta})$ at some points $\theta \in [-\pi,\pi)$ is caused by the slow decay of the envelope of the Dirichlet kernel $\mathcal{D}_N(\tau)$ as $\tau \to \infty$. It decreases only proportional to $1/\tau$ which causes the divergence of $\|\mathcal{D}_N\|_1$ as $N \to \infty$ (cf. (2.12) and Fig. 2.2). Thus, to obtain an approximation series (2.2) which converges uniformly, one needs a window $w(n)$ such that the corresponding kernel $K_N^{(w)}(\tau)$ (2.7) decreases faster than $1/\tau$ as $\tau \to \infty$. Two such methods will be discussed in the next subsections.

2.1.2 First arithmetic means – Fejér series

The unfavorable convergence behavior of the partial sums (2.11) is resolved if one considers the *(first) arithmetic means* of the partial sums (2.11)

$$\sigma_N(f;\mathrm{e}^{\mathrm{i}\theta}) := \frac{1}{N}\sum_{k=0}^{N-1} s_k(f;\mathrm{e}^{\mathrm{i}\theta}) \,. \tag{2.13}$$

If (2.11) is inserted into this arithmetic mean, a straight forward calculation gives the representations

$$\sigma_N(f;\mathrm{e}^{\mathrm{i}\theta}) = \sum_{n=-N}^{N} w(n)\,\hat{f}(n)\,\mathrm{e}^{\mathrm{i}n\theta} = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\tau})\,\mathcal{F}_N(\theta-\tau)\,\mathrm{d}\tau \tag{2.14}$$

with the window function

$$w(n) = 1 - \tfrac{|n|}{N}\,, \qquad n = 0, \pm 1, \pm 2, \cdots \pm N\,. \tag{2.15}$$

and with the kernel

$$\mathcal{F}_N(\tau) = \frac{1}{N}\sum_{k=0}^{N-1}\mathcal{D}_k(\tau) = \frac{1}{N}\,\frac{\sin^2(N\,\tau/2)}{\sin^2(\tau/2)}\,. \tag{2.16}$$

This first arithmetic mean (2.14) will be called the *Fejér series* of f with the *Fejér kernel* (2.16). For illustration, the window function (2.15) and the corresponding kernel (2.16) are plotted in Fig. 2.1 and 2.2, respectively. Similarly, forming the arithmetic mean of the conjugate partial sums $s_N(f;\mathrm{e}^{\mathrm{i}\theta})$, one obtains the conjugate Fejér mean $\widetilde{\sigma}_N(f;\mathrm{e}^{\mathrm{i}\theta})$ with the same window (2.15) and with the *conjugate Fejér kernel*

$$\widetilde{\mathcal{F}}_N(\tau) = \frac{1}{N}\sum_{k=0}^{N-1}\widetilde{\mathcal{D}}_k(\tau) = \frac{1}{\tan(\tau/2)} - \frac{1}{N}\,\frac{\sin(N\,\tau)}{2\,\sin^2(\tau/2)}\,.$$

It turns out that the Fejér series posses a much better convergence behavior with respect to continuous functions than the Fourier series. This will follow immediately from the observation that the Fejér kernel is a so called *approximate identity*:

Proposition 2.3. *The Fejér kernel* (2.16) *is an* approximate identity, *i.e. it satisfies for all* $N \in \mathbb{N}$ *the following three properties*

(*a*) $\mathcal{F}_N(\tau) \geq 0$ *for all* $\tau \in [-\pi, \pi)$

(*b*) $\frac{1}{2\pi} \int_{-\pi}^{\pi} \mathcal{F}_N(\tau)\, \mathrm{d}\tau = 1$

(*c*) $\lim_{N\to\infty} \mathcal{F}_N(\tau) = 0$ *for all* $0 < |\tau| \leq \pi$

Proof. Property (*a*) is obvious from (2.16) and (*b*) follows at once from (2.16) and (2.10). To verify (*c*) choose an arbitrary but fixed $0 < \epsilon < \pi$. Then one obtains from (2.16) that

$$|\mathcal{F}_N(\tau)| \leq \frac{1}{N} \frac{1}{|\sin^2(\tau/2)|} \leq \frac{1}{N} \frac{1}{|\sin^2(\epsilon/2)|} \quad \text{for all } \epsilon \leq |\tau| \leq \pi$$

where the right had side goes to zero as $N \to \infty$. □

Remark. Sometimes, kernels with the three properties (*a*), (*b*), and (*c*) are also called *positive kernels*. However, the proof of the following theorem will show that the definition of an approximate identity seems to be more appropriate. Moreover, we will call a kernel *positive* if it satisfies only property (*a*). With this definitions, a positive kernel need not be an approximate identity, in general, but an approximate identity is always positive.

Theorem 2.4. *Let* $f \in \mathcal{C}(\mathbb{T})$ *be a continuous function on* $\mathbb{T}$. *Then its Fejér series* (2.14) *converges uniformly to* f, *i.e.*

$$\lim_{N\to\infty} \|\sigma_N(f;\cdot) - f\|_\infty = 0\ .$$

Proof. Since f is given on the unit circle, the integral representation on the right hand side of (2.14) may be written as

$$\sigma_N(f; \mathrm{e}^{\mathrm{i}\theta}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}(\theta-\tau)})\, \mathcal{F}_N(\tau)\, \mathrm{d}\tau\ .$$

Therewith and using Properties (a) and (b) of the Fejér kernel, one gets

$$|\sigma_N(f; \mathrm{e}^{\mathrm{i}\theta}) - f(\mathrm{e}^{\mathrm{i}\theta})| \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(\mathrm{e}^{\mathrm{i}(\theta-\tau)}) - f(\mathrm{e}^{\mathrm{i}\theta})|\, \mathcal{F}_N(\tau)\, \mathrm{d}\tau\ . \tag{2.17}$$

We have to show that to every $\epsilon > 0$ there exists an $N_0 \in \mathbb{N}$ such that $|\sigma_N(f; \mathrm{e}^{\mathrm{i}\theta}) - f(\mathrm{e}^{\mathrm{i}\theta})| < \epsilon$ for all $N \geq N_0$ and for all $\theta \in [-\pi, \pi)$. To this end, we fix $\epsilon > 0$ and choose $\delta > 0$ such that

$$\sup_{-\delta\leq\tau\leq\delta} |f(e^{i(\theta-\tau)}) - f(e^{i\theta})| \leq \tfrac{\epsilon}{2} \qquad \text{for each } \theta \in [-\pi,\pi) \,. \tag{2.18}$$

This is always possible since f is continuous at every θ. Next we split up the integral in (2.17) into an integration over the interval $[-\delta,\delta]$ and over the interval $\delta \leq |\tau| \leq \pi$. Using (2.18) to upper bound the integral over $[-\delta,\delta]$, one obtains for an arbitrary $\theta \in [-\pi,\pi)$

$$\begin{aligned}&|\sigma_N(f;e^{i\theta}) - f(e^{i\theta})|\\ &\qquad\leq \frac{\epsilon}{2}\frac{1}{2\pi}\int_{-\delta}^{\delta} \mathcal{F}_N(\tau)\,d\tau \;+\; 2\,\|f\|_\infty \frac{1}{2\pi}\int_{|\tau|\geq\delta} \mathcal{F}_N(\tau)\,d\tau \,.\end{aligned} \tag{2.19}$$

By properties (a) and (b) of the Fejér kernel, the first term on the right hand side is certainly smaller than $\epsilon/2$. Since $\|f\|_\infty < \infty$, Property (c) of the Fejér kernel shows that there exists an N_0 such that $\mathcal{F}_N(\tau) < \epsilon/(4\,\|f\|_\infty)$ for all $N > N_0$ and all $\delta \leq |\tau| < \pi$. Therewith also the second term in (2.19) is upper bounded by $\epsilon/2$ such that

$$|\sigma_N(f;e^{i\theta}) - f(e^{i\theta})| \leq \epsilon \qquad \text{for all } N \geq N_0 \text{ and all } \theta \in [-\pi,\pi) \,.$$

This is what we wanted to show. □

Since $\mathcal{C}(\mathbb{T})$ is dense in every L^p with $1 \leq p < \infty$, the previous theorem implies that for every $f \in L^p$ the Fejér series $\sigma_N(f;\cdot)$ converges to f in L^p.

2.1.3 Delayed first arithmetic means – de-la-Vallée-Poussin Series

As we will see, it is favorable in some sense, to introduce a delay K in the first arithmetic mean (2.13) and to take the mean of the partial sums $s_K, s_{K+1}, \ldots, s_{K+N-1}$. This gives the so-called *delayed first arithmetic mean*

$$\sigma_{N,K}(f;e^{i\theta}) := \frac{1}{N}\sum_{k=K}^{K+N-1} s_k(f;e^{i\theta}) \,. \tag{2.20}$$

wherein $s_k(f;\cdot)$ is again the partial Fourier series (2.11) of f. For $K = 0$, one obtains again the first arithmetic mean (2.13), i.e. $\sigma_{N,0} = \sigma_N$. It is clear that the delayed arithmetic mean can be expressed as the difference of the two (not-delayed) arithmetic means σ_{K+N} and σ_K. Since $N\,\sigma_{K,N} = (K + N)\,\sigma_{K+N} - K\,\sigma_K$ one has the representation

$$\sigma_{N,K}(f;e^{i\theta}) = (1 + \tfrac{K}{N})\,\sigma_{K+N}(f;e^{i\theta}) - \tfrac{K}{N}\,\sigma_K(f;e^{i\theta}) \tag{2.21}$$

for the delayed arithmetic means. Subsequently we consider mainly the particular case where $K = N$. Then the delayed arithmetic mean becomes

$$\sigma_{N,N}(f;e^{i\theta}) = 2\,\sigma_{2N}(f;e^{i\theta}) - \sigma_N(f;e^{i\theta}) \,. \tag{2.22}$$

For the particular case ($K = N$), we insert the partial sums (2.11) into (2.20). A straight forward calculation gives the representations

$$\sigma_{N,N}(f;\mathrm{e}^{\mathrm{i}\theta}) = \sum_{n=-2N}^{2N} w(n)\,\hat{f}(n)\,\mathrm{e}^{\mathrm{i}n\theta} = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\tau})\,\mathcal{V}_N(\theta-\tau)\,\mathrm{d}\tau \quad (2.23)$$

with the trapezoid window function

$$w(n) = \begin{cases} 1 & 0 \le |n| \le N \\ 2\,(1-\frac{n}{2N}) & N+1 < |n| \le 2N \end{cases} \quad (2.24)$$

and with the kernel

$$\mathcal{V}_N(\tau) = 2\,\mathcal{F}_{2N}(\tau) - \mathcal{F}_N(\tau) = \frac{1}{N}\,\frac{\cos(N\tau)-\cos(2N\tau)}{2\,\sin^2(\tau/2)}\;. \quad (2.25)$$

This particular ($K = N$) delayed first arithmetic mean (2.23) is called the *de-la-Vallée-Poussin series* of f and (2.25) is the *de-la-Vallée-Poussin kernel.*

Similarly, forming the delayed arithmetic mean of the conjugate partial sums $\widetilde{s}_N(f;\mathrm{e}^{\mathrm{i}\theta})$, one obtains the *conjugate de-la-Vallée-Poussin series* $\widetilde{\sigma}_{N,N}(f;\mathrm{e}^{\mathrm{i}\theta})$ with the same window (2.24) and with the *conjugate de-la-Vallée-Poussin kernel*

$$\widetilde{\mathcal{V}}_N(\tau) = 2\,\widetilde{\mathcal{F}}_{2N}(\tau) - \widetilde{\mathcal{F}}_N(\tau) = \frac{1}{\tan(\tau/2)} - \frac{1}{N}\,\frac{\sin(N\tau)-\sin(2N\tau)}{2\,\sin^2(\tau/2)}\;.$$

It should be noted, that the de-la-Vallée-Poussin series $\sigma_{N,N}(f;\cdot)$ of a function f involves $4N-1$ Fourier coefficients of f whereas the Fejér series $\sigma_N(f;\cdot)$ uses only $2N-1$ Fourier coefficients of f.

The Fejér series $\sigma_N(f;\cdot)$ of a continuous function f converges uniformly to f. Since by (2.22) the de-la-Vallée-Possin series $\sigma_{N,N}(f;\cdot)$ is just the difference of two Fejér series one has immediately the following corollary of Theorem 2.4.

Corollary 2.5. *Let $f \in \mathcal{C}(\mathbb{T})$ be a continuous function on $\mathbb{T}$. Then its de-la-Vallée-Poussin series* (2.23) *converges uniformly to f, i.e.*

$$\lim_{N\to\infty} \|\sigma_{N,N}(f;\cdot) - f\|_\infty = 0\;.$$

Proof. By Theorem 2.4, the Fejér series $\sigma_N(f;\mathrm{e}^{\mathrm{i}\theta})$ converges to $f(\mathrm{e}^{\mathrm{i}\theta})$, uniformly in θ and therefore, it is a Cauchy sequence (uniformly in θ). Using the representation (2.22) of the de-la-Vallée-Poussin series, one obtains

$$|\sigma_{N,N}(f;\mathrm{e}^{\mathrm{i}\theta}) - f(\mathrm{e}^{\mathrm{i}\theta})| \le |\sigma_{2N}(f;\mathrm{e}^{\mathrm{i}\theta}) - f(\mathrm{e}^{\mathrm{i}\theta})| + |\sigma_{2N}(f;\mathrm{e}^{\mathrm{i}\theta}) - \sigma_N(f;\mathrm{e}^{\mathrm{i}\theta})|$$

where the right hand side goes to zero as $N \to \infty$, uniformly in θ. □

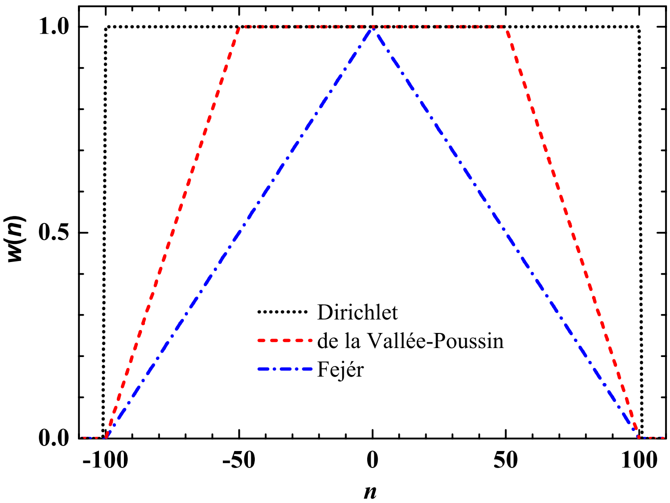

Fig. 2.1. Window functions corresponding to the Fourier series, the first arithmetic means, and the delayed first arithmetic mean for the order $N = 100$.

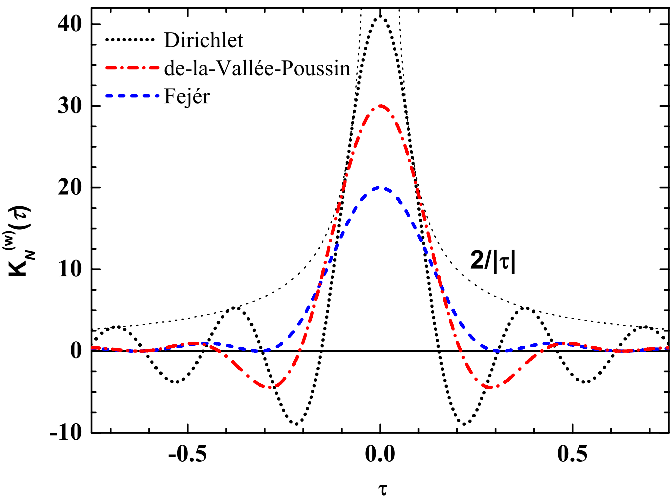

Fig. 2.2. Dirichlet, Fejér, and de-la-Vallée-Poussin kernel for the order $N = 20$.

Figures 2.1 and 2.2 show the window functions and the kernels, respectively, of the three weighted trigonometric series discussed above. The window function $w(n)$ determines primarily the convergence behavior of the partial sums (2.2) and (2.3) as $N \to \infty$. In the next subsection, for example, we will show that the de-la-Vallée-Poussin mean possesses the property that the approximation error decreases almost as fast as possible as $N \to \infty$. However, in applications other properties of the approximation series may also be of some importance. Therefore there does not exist an "optimal" window function, in general, but for different applications, different window functions may be favorable. Consequently, there exists many more possible window functions. In Section 10.5 we will discuss the optimal kernel for the approximation of spectral densities in some detail.

2.1.4 Best approximation by trigonometric polynomials

In what follows, $\mathcal{P}(N)$ denotes the set of all trigonometric polynomials with degree at most N, i.e. the set of all functions of the form

$$f(\mathrm{e}^{\mathrm{i}\theta}) = \frac{a_0}{2} + \sum_{k=1}^{N} a_k \cos(k\,\theta) + b_k \sin(k\,\theta)\,, \qquad \theta \in [-\pi, \pi)$$

with real coefficients $\{a_k\}_{k=0}^N$ and $\{b_k\}_{k=1}^N$. The subset of all $f \in \mathcal{P}(N)$ with a zero constant term a_0 i.e. all $f \in \mathcal{P}(N)$ for which $\int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\theta})\,\mathrm{d}\theta = 0$ is denoted by $\mathcal{P}_0(N)$.

Theorem 2.4 and Corollary 2.5 imply that every continuous function $f \in \mathcal{C}(\mathbb{T})$ can be uniformly approximated by a trigonometric polynomial, e.g. by the Fejér or de-la-Vallée-Poussin mean. Thus given an $\epsilon > 0$ one always finds a polynomial $p \in \mathcal{P}(N)$ of sufficiently large degree N such that $\|f - p\|_\infty < \epsilon$. Of course, for practical reasons, it is desirable to find the polynomial with the smallest degree N which satisfies the error requirement. We want to show that in a sense the de-la-Vallée-Poussin means are such approximation polynomials with an almost minimal degree.

Given a continuous function $f \in \mathcal{C}(\mathbb{T})$ and a fixed degree $N \geq 0$, the *best approximation* of f of degree N is defined as the number

$$B_N[f] := \inf_{p_N \in \mathcal{P}(N)} \|f - p_N\|_\infty\,. \tag{2.26}$$

It is clear that $B_N[f]$ tends to zero as $N \to \infty$. This follows from letting $\sigma_N(f;\cdot)$ be the Fejér series of f. Then $B_N[f] \leq \|f - \sigma_N(f;\cdot)\|_\infty$ and the right hand side converges to zero by Theorem 2.4. Next we observe that the infimum in (2.26) is attained in $\mathcal{P}(N)$.

Proposition 2.6. *To every $f \in \mathcal{C}(\mathbb{T})$ there exists a polynomial $p_N^* \in \mathcal{P}(N)$ such that $\|f - p_N^*\|_\infty = B_N[f]$.*

Proof. Let $f \in \mathcal{C}(\mathbb{T})$ be arbitrary and fix a degree N. Then by definition (2.26) of $B_N[f]$, there exists a sequence $\{p_N^{(k)}\}_{k=1}^{\infty}$ of polynomials in $\mathcal{P}(N)$ such that to every $\epsilon > 0$ there exists a K_0 such that

$$\|f - p_N^{(k)}\|_\infty \leq B_N[f] + \epsilon \,. \tag{2.27}$$

for all $k \geq K_0$. In particular, all trigonometric polynomials $p_N^{(k)}$ are uniformly bounded which implies that all Fourier coefficients $\hat{p}_N^{(k)}(n)$, $n = 0, \pm 1, \pm 2, \ldots$ of these polynomials are uniformly bounded. By the theorem of Bolzano-Weierstrass, for every n there exists a subsequence of $\hat{p}_N^{(k_i)}(n)$ which converges to a limit $\hat{p}_N^*(n)$. The corresponding subsequence of trigonometric polynomials converges uniformly to the polynomial

$$p_N^*(\mathrm{e}^{\mathrm{i}\theta}) = \textstyle\sum_{n=-N}^{N} \hat{p}_N^*(n)\, \mathrm{e}^{\mathrm{i}n\theta} \,, \qquad \theta \in [-\pi, \pi)$$

for which (2.27) gives $\|f - p_N^*\|_\infty \leq B_N[f]$. Since the reverse inequality is obvious, one gets the desired statement. □

Even though we know that the best approximation is attained, it will be difficult, in general, to determine the optimal polynomial p_N^*. However, the next theorem will show that the de-la-Vallée-Poussin mean $\sigma_{N,N}(f;\cdot)$ of f is always near the optimal polynomial, in the sense that the approximation error $\|f - \sigma_{N,N}(f;\cdot)\|_\infty$ is upper bounded by four times the best approximation.

Theorem 2.7. *Let $f \in \mathcal{C}(\mathbb{T})$ and let $\sigma_{N,K}(f;\cdot)$ with $K \geq N \geq 0$ be its delayed arithmetic mean. Then*

$$\|f - \sigma_{N,K}(f;\cdot)\|_\infty \leq 2\,(1 + \tfrac{K}{N})\, B_N[f]$$

and in particular $\|f - \sigma_{N,N}(f;\cdot)\|_\infty \leq 4\, B_N[f]$.

Proof. Fix the degree N and denote by $p_N \in \mathcal{P}(N)$ the trigonometric polynomial which attains the best approximation according to Proposition 2.6. Write f as

$$f(\mathrm{e}^{\mathrm{i}\theta}) = p_N(\mathrm{e}^{\mathrm{i}\theta}) + r(\mathrm{e}^{\mathrm{i}\theta}) \,, \qquad \theta \in [-\pi, \pi) \,. \tag{2.28}$$

It follows for the rest term that

$$|r(\mathrm{e}^{\mathrm{i}\theta})| = |f(\mathrm{e}^{\mathrm{i}\theta}) - p_N(\mathrm{e}^{\mathrm{i}\theta})| \leq \|f - p_N\|_\infty = B_N[f]$$

for all $\theta \in [-\pi, \pi)$, which implies for the Fejér mean of r that

$$|\sigma_k(r;\mathrm{e}^{\mathrm{i}\theta})| \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |r(\mathrm{e}^{\mathrm{i}\tau})|\, \mathcal{F}_k(\theta - \tau)\, \mathrm{d}\tau \leq B_N[f]$$

for all $\theta \in [-\pi, \pi)$ and for every arbitrary $k \geq 0$, using that $\mathcal{F}_k$ is an approximate identity. Applying (2.21) one obtains for the delayed arithmetic mean of r that

$$\begin{aligned} |\sigma_{N,K}(r;\mathrm{e}^{\mathrm{i}\theta})| &\leq (1+\tfrac{K}{N})\,|\sigma_{K+N}(f;\mathrm{e}^{\mathrm{i}\theta})| + \tfrac{K}{N}\,|\sigma_K(f;\mathrm{e}^{\mathrm{i}\theta})| \\ &\leq (1+2\,\tfrac{K}{N})\,B_N[f] \end{aligned} \tag{2.29}$$

for all $\theta \in [-\pi,\pi)$ and arbitrary $K, N \geq 0$.

Consider the partial Fourier series (2.11) of the polynomial p_N and assume that $k \geq N$. Then $s_k(p_N;\mathrm{e}^{\mathrm{i}\theta}) = p_N(\mathrm{e}^{\mathrm{i}\theta})$ for all θ. Therefore, it follows from (2.28) for the partial Fourier series of f that

$$s_k(f;\mathrm{e}^{\mathrm{i}\theta}) = p_N(\mathrm{e}^{\mathrm{i}\theta}) + s_k(r;\mathrm{e}^{\mathrm{i}\theta})\;, \qquad \theta \in [-\pi,\pi)$$

for all $k \geq N$, and for an arbitrary $K \geq N$. The delayed first arithmetic mean of f becomes

$$\frac{1}{N}\sum_{k=K}^{K+N-1} s_k(f;\mathrm{e}^{\mathrm{i}\theta}) = p_N(\mathrm{e}^{\mathrm{i}\theta}) + \frac{1}{N}\sum_{k=K}^{K+N-1} s_k(r;\mathrm{e}^{\mathrm{i}\theta})\;,$$

which is equivalent to

$$\sigma_{N,K}(f;\mathrm{e}^{\mathrm{i}\theta}) = p_N(\mathrm{e}^{\mathrm{i}\theta}) + \sigma_{N,K}(r;\mathrm{e}^{\mathrm{i}\theta})\;, \qquad \theta \in [-\pi,\pi)\;. \tag{2.30}$$

Finally, one obtains for all $K \geq N$

$$\begin{aligned} |f(\mathrm{e}^{\mathrm{i}\theta}) - \sigma_{N,K}(f;\mathrm{e}^{\mathrm{i}\theta})| &\leq |f(\mathrm{e}^{\mathrm{i}\theta}) - p_N(\mathrm{e}^{\mathrm{i}\theta})| + |p_N(\mathrm{e}^{\mathrm{i}\theta}) - \sigma_{N,K}(f;\mathrm{e}^{\mathrm{i}\theta})| \\ &\leq B_N[f] + |\sigma_{N,K}(r;\mathrm{e}^{\mathrm{i}\theta})| \\ &\leq 2\,(1+\tfrac{K}{N})\,B_N[f] \end{aligned}$$

where the second inequality follows from the definition of p_N and from (2.30), whereas the last inequality is a consequence of (2.29). □

We already saw that the best approximation $B_N[f]$ of every continuous function $f \in \mathcal{C}(\mathbb{T})$ converges to zero as $N \to \infty$. However, one would expect that it is easier to approximate a "simple" function than a "complicated" functions. Thus, the best approximation $B_N[f]$ of a "simple" function should converge more rapidly than for a complicated function. The following theorem will show that this is indeed the case, and that in this context a "simple function" is a smooth function, i.e. the smoother the function f, the faster $B_N[f]$ converges to zero.

Theorem 2.8. *Let f be a k-times differentiable function on the unit circle $\mathbb{T}$ whose k-th derivative $f^{(k)}$ has a modulus of continuity of ω. Then there exists a constant C_k, which depends only on k, such that*

$$B_N[f] \leq C_k\,\omega\left(\tfrac{1}{N}\right)\,N^{-k}\;.$$

For functions from the Hölder-Zygmund class $\Lambda_\alpha(\mathbb{T})$ (cf. Section 1.3), we obtain immediately the following corollary as a special case of Theorem 2.8.

Corollary 2.9. *If $f \in \Lambda_\alpha(\mathbb{T})$ for $\alpha > 0$, then there exists a constant C_α such that*

$$B_N[f] \le C_\alpha\, N^{-\alpha}\,.$$

Proof (Theorem 2.8). The proof consists of several steps. Each considers a special case of Theorem 2.8.
1) First it is shown that if $f \in \mathcal{C}(\mathbb{T})$ and has modulus of continuity ω, then $B_N[f] < 12\,\omega(1/N)$. To this end, we use the following approximation

$$f_N(\mathrm{e}^{\mathrm{i}\theta}) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\tau})\,\mathcal{J}_N(\theta-\tau)\,\mathrm{d}\tau \tag{2.31}$$

with the so-called *Jackson kernel*

$$\mathcal{J}_N(\tau) = \frac{3}{N\,(2\,N^2+1)}\,\left(\frac{\sin(N\,t/2)}{\sin(t/2)}\right)^4\,. \tag{2.32}$$

This approximation method has the following three properties[1] which are used throughout the rest of the proof.

(a) $f_N \in \mathcal{P}(2\,N-2)$, i.e. f_N is a trigonometric polynomial of degree $2\,N-2$.
(b) $f_N \in \mathcal{P}_0(2\,N-2)$ whenever $\int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\theta})\,\mathrm{d}\theta = 0$.
(c) The kernel (2.32) satisfies the relation

$$\tfrac{1}{2\,\pi}\int_{-\pi}^{\pi}\mathcal{J}_N(\tau)\,\mathrm{d}\tau = 1\,. \tag{2.33}$$

Replacing $\theta-\tau$ by s and splitting up the integral in (2.31) gives

$$\begin{aligned} f_N(\mathrm{e}^{\mathrm{i}\theta}) &= \frac{1}{2\pi}\int_{-\pi}^{0} f(\mathrm{e}^{\mathrm{i}(\theta-s)})\,\mathcal{J}_N(s)\,\mathrm{d}s + \frac{1}{2\pi}\int_{0}^{\pi} f(\mathrm{e}^{\mathrm{i}(\theta-s)})\,\mathcal{J}_N(s)\,\mathrm{d}s \\ &= \frac{1}{2\pi}\int_0^{\pi}\Big[f(\mathrm{e}^{\mathrm{i}(\theta+s)}) + f(\mathrm{e}^{\mathrm{i}(\theta-s)})\Big]\,\mathcal{J}_N(s)\,\mathrm{d}s \end{aligned}$$

where the second line follows after the change variable $s \mapsto -s$ in the first integral. Next we use property (2.33) of the kernel $\mathcal{J}_N$. This yields

$$\begin{aligned} \big|f(\mathrm{e}^{\mathrm{i}\theta}) - f_N(\mathrm{e}^{\mathrm{i}\theta})\big| &= \left|\frac{1}{2\pi}\int_0^{\pi}\Big[2\,f(\mathrm{e}^{\mathrm{i}\theta}) - f(\mathrm{e}^{\mathrm{i}(\theta+s)}) - f(\mathrm{e}^{\mathrm{i}(\theta-s)})\Big]\,\mathcal{J}_N(s)\,\mathrm{d}s\right| \\ &\le \frac{1}{2\pi}\int_0^{\pi}\Big(\Big|f(\mathrm{e}^{\mathrm{i}\theta}) - f(\mathrm{e}^{\mathrm{i}(\theta+s)})\Big| + \Big|f(\mathrm{e}^{\mathrm{i}\theta}) - f(\mathrm{e}^{\mathrm{i}(\theta-s)})\Big|\Big)\,\mathcal{J}_N(s)\,\mathrm{d}s \\ &\le \frac{1}{\pi}\int_0^{\pi}\omega(s)\,\mathcal{J}_N(s)\,\mathrm{d}s \qquad \text{for every } \theta \in [-\pi,\pi)\,. \end{aligned}$$

[1] For a proof of these properties, we refer e.g. to [61, vol. 1, Chap. IV], or just note that $\mathcal{J}_N$ is a normalized square of the Fejér kernel (2.16).

By the properties of the modulus of continuity (cf. Section 1.3), it holds $\omega(s) = \omega(N\,s\,\frac{1}{N}) \leq (N\,s+1)\,\omega(1/N)$ such that

$$\left|f(\mathrm{e}^{\mathrm{i}\theta}) - f_N(\mathrm{e}^{\mathrm{i}\theta})\right| \leq \omega\left(\tfrac{1}{N}\right) \left[\frac{N}{\pi}\int_0^{\pi} s\,\mathcal{J}_N(s)\,\mathrm{d}s + 1\right] \tag{2.34}$$

for every $\theta \in [-\pi, \pi)$. Next, we derive an upper bound for the integral on the right hand side. To do this, the integration interval is divided into two intervals as follows

$$\int_0^{\pi} s\,\mathcal{J}_N(s)\,\mathrm{d}s = \int_0^{\pi/N} s\,\mathcal{J}_N(s)\,\mathrm{d}s + \int_{\pi/N}^{\pi} s\,\mathcal{J}_N(s)\,\mathrm{d}s\;.$$

Now, in the first integral we use that $|\sin(N\,s/2)| \leq N\,\sin(s/2)$ and in the second integral we apply the inequalities $|\sin(x)| \leq 1$ and $\sin(s/2) \geq s/\pi$ for all $s \in [0, \pi)$. Therewith, one obtains

$$\begin{aligned}\int_0^{\pi} s\,\mathcal{J}_N(s)\,\mathrm{d}s &\leq \frac{3}{N\,(2\,N^2+1)}\left\{N^4\int_0^{\pi/N} s\,\mathrm{d}s + \pi^4\int_{\pi/N}^{\pi} s^{-3}\,\mathrm{d}s\right\}\\ &\leq \frac{3}{N\,(2\,N^2+1)}\left\{\frac{\pi^2N^2}{2} + \frac{\pi^2N^2}{2}\right\} = \frac{3\,\pi^2 N}{2\,N^2+1}\;.\end{aligned}$$

Using this upper bound in (2.34), one obtains that

$$\left|f(\mathrm{e}^{\mathrm{i}\theta}) - f_N(\mathrm{e}^{\mathrm{i}\theta})\right| \leq \left[\frac{3\,\pi}{2} + 1\right]\,\omega\left(\tfrac{1}{N}\right) < 6\,\omega\left(\tfrac{1}{N}\right)$$

for every $\theta \in [-\pi, \pi)$. Since f_N is a trigonometric polynomial of degree $2N-2$, it follows that $B_{2N-2}[f] < 6\,\omega(1/N)$. Assume first that $N = 2\,M$ is an even natural number. Then

$$B_N[f] = B_{2\,M}[f] \leq B_{2\,M-2}[f] < 6\,\omega(1/M) = 6\,\omega(2/N) \leq 12\,\omega(1/N)\;.$$

Similarly, if $N = 2\,M-1$ is an odd natural number, one obtains

$$B_N[f] = B_{2\,M-1}[f] \leq B_{2\,M-2}[f] < 6\,\omega(1/M) = 6\,\omega(2/[N+1]) \leq 12\,\omega(1/N)\;.$$

This is what we wanted to show. Moreover, this already proves the theorem for the case $k = 0$.

2) As a consequence of the first part, one obtains for the special case of Lipschitz continuous functions $f \in \mathrm{Lip}_K$ that $B_N[f] < 12\,K/N$.

3) Assume now that the given function f satisfies the condition

$$\textstyle\int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\theta})\,\mathrm{d}\theta = 0\;. \tag{2.35}$$

Then $f_N \in \mathcal{P}_0(2\,N-2)$ by property (b) of the approximation polynomial f_N. Denote by $b_N[f]$ the best approximation of f by polynomials in $\mathcal{P}_0(N)$, i.e.

$$b_N[f] := \inf_{p_N \in \mathcal{P}_0(N)} \|f - p_N\|_\infty \ .$$

Following the derivation under point 1) of this proof, one obtains that $b_N[f] \leq 12\,\omega(1/N)$ for all $f \in \mathcal{C}(\mathbb{T})$ which satisfy (2.35). Moreover, as under point 2), for all $f \in \mathrm{Lip}_K$ which satisfy (2.35), one obtains

$$b_N[f] \leq 12\,K/N \ . \tag{2.36}$$

4) Assume now that f possesses a bounded derivative f' and denote by $b'_N = b_N[f']$ the best approximation of f' in the class $\mathcal{P}_0(N)$. Then there exists a trigonometric polynomial $u \in \mathcal{P}_0(N)$ such that

$$|f'(\mathrm{e}^{\mathrm{i}\theta}) - u(\mathrm{e}^{\mathrm{i}\theta})| \leq b'_N \ , \qquad \text{for all } \theta \in [-\pi, \pi) \ . \tag{2.37}$$

Integrating u gives a trigonometric polynomial $v \in \mathcal{P}_0(N)$ such that $v'(\mathrm{e}^{\mathrm{i}\theta}) = u(\mathrm{e}^{\mathrm{i}\theta})$. With the definition $\varphi(\mathrm{e}^{\mathrm{i}\theta}) := f(\mathrm{e}^{\mathrm{i}\theta}) - v(\mathrm{e}^{\mathrm{i}\theta})$ relation (2.37) can be written as $|\varphi'(\mathrm{e}^{\mathrm{i}\theta})| \leq b'_N$ for all $\theta \in [-\pi, \pi)$. This shows in particular that $\varphi \in \mathrm{Lip}_{b'_N}$. Using point 2) of this proof, we get $B_N[\varphi] \leq 12\,b'_N[f]/N$. Therefore, by the definition of the $B_N[\varphi]$, there exists a $w \in \mathcal{P}(N)$ such that

$$\big|\varphi(\mathrm{e}^{\mathrm{i}\theta}) - w(\mathrm{e}^{\mathrm{i}\theta})\big| = \big|f(\mathrm{e}^{\mathrm{i}\theta}) - \big[v(\mathrm{e}^{\mathrm{i}\theta}) + w(\mathrm{e}^{\mathrm{i}\theta})\big]\big| \leq \frac{12}{N}\,b'_N$$

for all $\theta \in [-\pi, \pi)$. Setting $u_N = v + w \in \mathcal{P}(N)$, the last inequality shows that

$$B_N[f] \leq \frac{12}{N}\,b'_N \ . \tag{2.38}$$

5) Assume now that f satisfies the conditions of the theorem. Then according to 4), relation (2.38) holds. Moreover, the first derivative f' has again a continuous and bounded derivative f''. Thus $f' \in \mathrm{Lip}_{b''_N}$ where b''_N denotes the best approximation of f'' in $\mathcal{P}_0(N)$. Moreover, since f is continuous on the unit circle $\mathbb{T}$, we have that $\int_{-\pi}^{\pi} f'(\mathrm{e}^{\mathrm{i}\theta})\,\mathrm{d}\theta = f(\pi) - f(-\pi) = 0$. Thus, f' satisfies also the assumption (2.35) under point 3) of this proof. Applying (2.36) to f' one obtains $b'_N = b_N[f''] \leq 12\,b''_N/N$. Now one applies the same arguments to f'', $f^{(3)}$, and so forth, up to $f^{(k-1)}$. This gives the relations

$$b_N^{(n)} \leq \frac{12}{N}\,b_N^{(n+1)} \qquad n = 1, 2, \ldots, k-1 \ . \tag{2.39}$$

Finally, it follows from point 3) of this proof that $b_N^{(k)} = b_N[f^{(k)}] \leq 12\,\omega(1/N)$. Combining this with the inequalities (2.39) and with (2.38), one obtains the statement of the theorem with $C_k = 12^{k+1}$. □

2.2 Hardy Spaces on the Unit Disk

This section gives a short introduction to a class of function spaces which contain analytic functions in the unit disk. Their importance for system theory results from the fact that their elements can be interpreted as causal transfer functions of linear systems which are bounded in a certain L^p-norm. After a short introduction of these so called Hardy spaces, we will present some results which will be needed in later parts of this book.

2.2.1 Basic definitions

Denote by H the set of all functions that are *analytic* (i.e. *holomorphic*) inside the unit disk $\mathbb{D}$. Then, it is clear that every $f(z)$ in H is bounded for all $z \in \mathbb{D}$. However, as $|z|$ approaches 1, the modulus $|f(z)|$ may go to infinity. Hardy spaces are subsets of H whose elements satisfy a certain growth condition toward the boundary of the unit disk.

Definition 2.10 (Hardy spaces). *Let $f \in H$ be an analytic function inside the unit disk $\mathbb{D}$. For $0 < p < \infty$, we set*

$$\|f\|_p := \sup_{0<r<1} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} |f(r\mathrm{e}^{\mathrm{i}\theta})|^p \, \mathrm{d}\theta \right)^{1/p}$$

and for $p = \infty$, we define

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

Then for $0 < p \leq \infty$ the Hardy class H^p *is defined as the set of all functions f analytic in $\mathbb{D}$ for which $\|f\|_p < \infty$.*

It is not hard to verify that $H^\infty \subset H^q \subset H^p$ for all $0 < p < q < \infty$. The Hardy spaces were defined by their behavior inside the unit disk $\mathbb{D}$. The following theorem[2] characterizes the radial limits of functions in H^p.

Theorem 2.11. *Let $f \in H^p$ with $1 \leq p \leq \infty$. Then the radial limit*

$$\widetilde{f}(\mathrm{e}^{\mathrm{i}\theta}) := \lim_{r \nearrow 1} f(r\mathrm{e}^{\mathrm{i}\theta})$$

exists for almost all $\theta \in [-\pi, \pi)$. Moreover $\widetilde{f} \in L^p$ with $\|\widetilde{f}\|_{L^p} = \|f\|_{H^p}$.

Thus, the radial limit of every $f \in H^p$ exists. From now on, this radial limit will also be denoted by f. As a consequence H^p can be considered as a closed subspace of L^p and therefore every H^p with $1 \leq p \leq \infty$ is a Banach space by itself. Moreover, since H^2 is a closed subspace of the Hilbert space L^2, it is also Hilbert space with the inner product

$$\langle f, g \rangle = \lim_{r \nearrow 1} \frac{1}{2\pi} \int_{-\pi}^{\pi} f(r\mathrm{e}^{\mathrm{i}\theta}) \, \overline{g(r\mathrm{e}^{\mathrm{i}\theta})} \, \mathrm{d}\theta = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\theta}) \, \overline{g(\mathrm{e}^{\mathrm{i}\theta})} \, \mathrm{d}\theta \,.$$

[2] For a proof, we refer e.g. to [70, § 17.11], [30, § 3.2]

Conversely, let $f \in H^p$. Then Cauchy's theorem implies that all Fourier coefficients $\hat{f}(n)$ with negative index n vanish, because f is analytic inside the unit disk $\mathbb{D}$ and therefore

$$\hat{f}(n) = \frac{1}{2\pi \mathrm{i}} \oint_{\mathbb{T}} f(\zeta)\, \zeta^{-(n+1)}\, \mathrm{d}\zeta = 0 \quad \text{for all}\, n < 0$$

where the integration over $\mathbb{T}$ has to be done counter-clockwise. Consequently, H^p could be defined as the subspace of those L^p functions for which all negative Fourier coefficients are equal to zero:

$$H^p = \{f \in L^p : \hat{f}(n) = 0 \text{ for all } n < 0\} \,.$$

Therefore every $f \in L^p$ with the Fourier series $f(\mathrm{e}^{\mathrm{i}\theta}) = \sum_{n=0}^{\infty} \hat{f}(n)\, \mathrm{e}^{\mathrm{i}n\theta}$ can be identified with the H^p-function

$$f(z) = \textstyle\sum_{n=0}^{\infty} \hat{f}(n)\, z^n$$

which is analytic for every $z \in \mathbb{D}$.

As explained in Example 1.9, there is a natural isometric isomorphism between L^2 and $\ell^2(\mathbb{Z})$, given by associating every $f \in L^2$ with the sequence $\{\hat{f}(n)\}_{n=-\infty}^{\infty}$ of its Fourier coefficients. In the same way, the Hardy space H^2 is isometricly isomorphic to the sequence space $\ell^2(\mathbb{Z}_+)$ because $\ell^2(\mathbb{Z}_+)$ may be considered as the subspace of all sequences $\{\alpha_k\}_{k=-\infty}^{\infty}$ in $\ell^2(\mathbb{Z})$ for which $\alpha_k = 0$ for $k < 0$.

The next theorem gives a useful characterization of the Fourier coefficients of H^1 functions, which will be used frequently in the following. The proof is omitted but may be found in [30] or [41], for example.

Theorem 2.12 (Hardy's inequality). *Let $f \in H^1$ with Fourier expansion $f(z) = \sum_{k=0}^{\infty} \hat{f}(k)\, z^k$. Then its Fourier coefficients satisfy the inequality*

$$\|f\|_1 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(\mathrm{e}^{\mathrm{i}\theta})|\, \mathrm{d}\theta \geq \frac{1}{\pi} \sum_{k=0}^{\infty} \frac{|\hat{f}(k)|}{k+1} \,. \tag{2.40}$$

2.2.2 Zeros of H^p-functions

Let $f \in H^p$ be an arbitrary function in a certain Hardy space H^p. Define by

$$\mathcal{Z}(f) := \{z \in \mathbb{T} : f(z) = 0\}$$

the *zero set* of f, i.e. the set of all those points in $\mathbb{D}$ where f is zero. For every function f holomorhic in the unit disk $\mathbb{D}$, it is well known that either $\mathcal{Z}(f) = \mathbb{D}$, or $\mathcal{Z}(f)$ has no limit point in $\mathbb{D}$. In the first case f is identically zero which is of little interest. Thus, the zeros of a non-zero holomorphic function f in $\mathbb{D}$ are isolated points in $\mathbb{T}$, and if the number of zeros is infinite, the limit points of the zeros have to lie outside of $\mathbb{D}$, i.e. on the boundary $\mathbb{T}$ of $\mathbb{D}$. If we only assume that f is holomorphic in $\mathbb{D}$, this is all we can say about the zeros of f, by the Theorem of Weierstrass (see e.g. [70, Chapter 15]). However, if we consider functions in the Hardy spaces H^p which are not only holomorphic in $\mathbb{D}$ but satisfy a certain growth behavior toward the boundary of $\mathbb{D}$, more can be said about the distribution of the zeros in $\mathbb{D}$, namely that the zeros have to converge with a certain rate toward the limit points on $\mathbb{T}$ (if they exist). The basis of deriving these conditions on the zeros of H^p functions is the following *Jensen's formula.*

Theorem 2.13 (Jensen's formula). *Let $f \in H(\mathbb{D})$ be a holomorphic function in $\mathbb{D}$ with $f(0) \neq 0$, let $0 < r < 1$, and let $\alpha_1, \dots, \alpha_N$ be the zeros of f in the closed disk $\overline{\mathbb{D}}_r(0) = \{z \in \mathbb{C} : |z| \leq r\}$ listed according to their multiplicities. Then*

$$|f(0)| \prod_{n=1}^{N} \frac{r}{|\alpha_n|} = \exp\left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \log\left|f(r\mathrm{e}^{\mathrm{i}\theta})\right| \, \mathrm{d}\theta\right) . \tag{2.41}$$

Remark 2.14. Since f is considered in the disk $\overline{\mathbb{D}}_r(0) \subset \mathbb{D}$ with $r < 1$ and since f is holomorphic in $\mathbb{D}$, the zeros of f have no limit point in $\overline{\mathbb{D}}_r(0)$. Consequently, the number N of zeros in $\overline{\mathbb{D}}_r(0)$ is finite.

Remark 2.15. The assumption $f(0) \neq 0$ is no real limitation. Because if f has a zero of order m at 0, one can apply Jensen's formula to the function $f(z)/z^m$.

Proof. One orders the zeros $\{\alpha_n\}_{n=1}^{N}$ of f in $\overline{\mathbb{D}}_r(0)$ according to their location in $\mathbb{D}_r(0)$ and on the boundary of $\mathbb{D}_r(0)$, i.e. such that $|\alpha_1| \leq \dots \leq |\alpha_M| < r$ and $|\alpha_{M+1}| = \dots = |\alpha_N| = r$. Define the function

$$g(z) := f(z) \prod_{n=1}^{M} \frac{r^2 - \overline{\alpha}_n z}{r(\alpha_n - z)} \prod_{n=M+1}^{N} \frac{\alpha_n}{\alpha_n - z} , \qquad z \in \mathbb{D} . \tag{2.42}$$

It is clear that g has no zeros in $\overline{\mathbb{D}}_r(0)$ and since $f \in H(\mathbb{D})$, also $g \in H(\mathbb{D})$. Thus, there exists a $\rho > r$ such that g has no zeros in the open disk $\mathbb{D}_\rho(0)$ and is holomorphic in $\mathbb{D}_\rho(0)$. It follows that $\log|g|$ is a harmonic function in $\mathbb{D}_\rho(0)$ (see e.g. [70, § 13.12]). Consequently, $\log|g|$ possesses the mean value property

$$\log|g(0)| = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log\left|g(r\mathrm{e}^{\mathrm{i}\theta})\right| \, \mathrm{d}\theta . \tag{2.43}$$

To determine the right hand side of (2.43), one easily verifies that the factors in (2.42) for $1 \leq n \leq M$ have modulus 1. For the remaining factors with $M + 1 \leq n \leq N$ holds

$$\frac{\alpha_n}{\alpha_n - z} = \frac{1}{1 - \mathrm{e}^{\mathrm{i}(\theta - \tau_n)}}$$

if we write $\alpha_n = r\mathrm{e}^{\mathrm{i}\tau_n}$ and $z = r\mathrm{e}^{\mathrm{i}\theta}$. Therewith (2.43) becomes

$$\log|g(0)| = \frac{1}{2\pi}\int_{-\pi}^{\pi} \log\left|f(r\mathrm{e}^{\mathrm{i}\theta})\right| \,\mathrm{d}\theta - \sum_{n=M+1}^{N} \frac{1}{2\pi}\int_{-\pi}^{\pi} \log\left|1-\mathrm{e}^{\mathrm{i}(\theta-\tau_n)}\right| \,\mathrm{d}\theta \tag{2.44}$$

It is a consequence of Cauchy's theorem that

$$\tfrac{1}{2\pi}\textstyle\int_{-\pi}^{\pi} \log\left|1-\mathrm{e}^{\mathrm{i}\theta}\right| \,\mathrm{d}\theta = 0$$

(see e.g. [70, § 15.17]) which implies that the second term on the right hand side of (2.44) is zero. The definition (2.42) of the function g gives at once

$$|g(0)| = |f(0)| \textstyle\prod_{n=1}^{M} \tfrac{r}{\alpha_n} \; .$$

Taking the exponential function of (2.43) shows finally (2.41). □

The next theorem proves a necessary condition on the zeros of a function f in order that $f \in H^p$ for some $1 \le p \le \infty$.

Theorem 2.16. *Let $f \in H^p$ with $1 \le p \le \infty$ be an analytic function in $\mathbb{D}$ with $f(0) \neq 0$, and let $\{\alpha_n\}_{n=1}^{\infty}$ be the zeros of f, listed according to their multiplicities. Then these zeros satisfy the* Blaschke condition

$$\sum_{n=1}^{\infty} (1-|\alpha_n|) < \infty \; . \tag{2.45}$$

Proof. If f has only finitely many zeros condition (2.45) is satisfied. Therefore, we assume that there are infinitely many zeros. Denote by $N(r)$ the number of zeros of f in the closed disk $\overline{\mathbb{D}}_r(0)$ for a radius $r < 1$. Fix $K \in \mathbb{N}$ and choose $r < 1$ such that $N(r) > K$. Then Jensen's formula gives

$$|f(0)| \prod_{n=1}^{K} \frac{r}{|\alpha_n|} \le |f(0)| \prod_{n=1}^{N(r)} \frac{r}{|\alpha_n|} = \exp\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \log\left|f(r\mathrm{e}^{\mathrm{i}\theta})\right| \,\mathrm{d}\theta\right) < \infty$$

where the right hand side is bounded since $f \in H^p \subset H^1$. Thus there exists a constant $C_0 < \infty$ such that $\prod_{n=1}^{K} |\alpha_n| \ge r^K \,|f(0)|/C_0$. Since this inequality holds for arbitrary K, it is still valid for $K \to \infty$, i.e.

$$\prod_{n=1}^{\infty} |\alpha_n| \ge \frac{|f(0)|}{C_0} > 0 \; .$$

Now we define $u_n := 1 - |\alpha_n|$ for all $n \in \mathbb{N}$ and notice that the power series expansion of the exponential function implies $1 - x \le \exp(-x)$ for all $x \in \mathbb{R}$. Replacing x by u_n and multiplying the resulting inequalities gives

$$0 < \prod_{n=1}^{\infty} |\alpha_n| = \prod_{n=1}^{\infty} (1 - u_n) \le \prod_{n=1}^{\infty} \exp(-u_n)$$
$$\le \exp\left(-\sum_{n=1}^{\infty} u_n\right) = \exp\left(-\sum_{n=1}^{\infty}(1 - |\alpha_n|)\right)$$

which implies (2.45). □

Remark 2.17. Let $\{\alpha_n\}$ be a sequence in $\mathbb{D}$ with $|\alpha_n| = (n-1)/n$. This sequence does not satisfy the Blaschke condition (2.45) and therefore there exists no function f in any H^p-space $(1 \le p \le \infty)$ with zeros at α_n. However, by the Weierstrass factorization theorem there exists a holomorphic function $f \in H(\mathbb{D})$ with zeros at α_n, but with $\|f\|_p = \infty$ for all $1 \le p \le \infty$. Conversely, if a function $f \in H^p$ is known to have zeros at $\{\alpha_n\}_{n\in\mathbb{N}}$, then this function has to be identically zero in $\mathbb{D}$.

So the Blaschke condition (2.45) is a necessary condition on the zeros of a holomorphic function f in order that f belongs to a Hardy space H^p. It even turns out that (2.45) is also a sufficient condition for the existence of a function $f \in H^p$, $1 \le p \le \infty$ which has zeros only at the prescribed points $\{\alpha_n\}_{n=1}^{\infty}$. The form of such a function is characterized in the following theorem.

Theorem 2.18 (Blaschke product). *Let $\{\alpha_n\}_{n=1}^{\infty}$ be a sequence of nonzero complex numbers in $\mathbb{D}$ such that $\{\alpha_n\}_{n=1}^{\infty}$ satisfies the Blaschke condition* (2.45). *Let $k \ge 0$ be a nonnegative integer, and define the* Blaschke product

$$B(z) := z^k \prod_{n=1}^{\infty} \frac{|\alpha_n|}{\alpha_n} \frac{\alpha_n - z}{1 - \overline{\alpha}_n z}\,, \qquad z \in \mathbb{D}\,. \tag{2.46}$$

Then $B \in H^\infty \subset H^p$, $1 \le p < \infty$, and B has zeros only at the points α_n and a zero of order k at 0*. Moreover, $|B(z)| < 1$ for all $z \in \mathbb{D}$ and $|B(\mathrm{e}^{\mathrm{i}\theta})| = 1$ almost everywhere.*

Remark 2.19. The term "Blaschke product" will be used for all functions of the form (2.46) even if the product contains only a finite number of factors and even if it contains no factor, i.e. even if $B(z) = 1$ for all $z \in \mathbb{D}$.

The function $B(z)$ in the above theorem is the product of z^k and of the factors

$$b_n(z) := \frac{|\alpha_n|}{\alpha_n} \frac{\alpha_n - z}{1 - \overline{\alpha}_n z}\,, \quad z \in \mathbb{D}\,. \tag{2.47}$$

Each factor b_n has a zero at $z = \alpha_n$ inside the unit disk $\mathbb{D}$, and a pole at $z = \overline{\alpha}_n^{-1}$ outside the closed unit disk $\overline{\mathbb{D}}$. Thus, each factor $b_n \in H(\mathbb{D})$ is a holomorphic function in $\mathbb{D}$ with precisely one zero at α_n. Moreover, it is easily verified that each factor has the properties that $|b_n(z)| < 1$ for all $z \in \mathbb{D}$ and that $|b_n| = 1$ for all $|z| = 1$. The Blaschke product is given by the infinite product $B(z) = z^k \prod_{n=1}^{\infty} b_n(z)$ of holomorphic functions. To prove Theorem 2.18, we basically have to show that this product converges uniformly to a holomorphic function. Therefore, as a preparation, the following Lemma studies conditions for the uniform convergence of infinite products.

Lemma 2.20. *Let $\{u_n\}_{n=1}^{\infty}$ be a sequence of bounded complex functions on a subset $S \subset \mathbb{D}$ of the unit disk such that the sum $\sum_{n=1}^{\infty} |u_n(z)|$ converges uniformly on S. Then the product*

$$f(z) = \prod_{n=1}^{\infty} [1 + u_n(z)]$$

converges uniformly on S. Moreover $f(z) = 0$ at some $z \in S$ if and only if $u_n(z) = -1$ for some $n \in \mathbb{N}$.

Proof. The assumption on $\{u_n\}$ implies that there exists a constant $C_0 < \infty$ such that $\sum_{n=1}^{\infty} |u_n(z)| \le C_0$ for all $z \in S$. The power series expansion of the exponential function shows that $1 + x \le \exp(x)$. Replacing x by $|u_1(z)|, |u_2(z)|, \ldots$ and multiplying the inequalities yields

$$\prod_{n=1}^{\infty} (1 + |u_n(z)|) \le \exp\left(\sum_{n=1}^{\infty} |u_n(z)|\right) \le \exp(C_0) =: C_1 < \infty \tag{2.48}$$

for all $z \in S$. Next, we define the partial products

$$p_N(z) := \prod_{n=1}^{N} [1 + u_n(z)] \qquad \text{and} \qquad q_N(z) := \prod_{n=1}^{N} [1 + |u_n(z)|]$$

and show that for every $N \in \mathbb{N}$

$$|p_N(z) - 1| \le q_N(z) - 1 \,. \tag{2.49}$$

For $N = 1$ this inequality is certainly satisfied. For $N > 1$, the statement is proved by induction. For $p_{N+1}(z)$ holds obviously

$$\begin{aligned} p_{N+1}(z) - 1 &= p_N(z)\ [1 + u_{N+1}(z)] - 1 \\ &= [p_N(z) - 1]\ [1 + u_{N+1}(z)] + u_{N+1}(z) \,. \end{aligned}$$

Therewith, it follows for the modulus

$$\begin{aligned} |p_{N+1}(z) - 1| &\le |p_N(z) - 1|\ |1 + u_{N+1}(z)| + |u_{N+1}(z)| \\ &\le (q_N(z) - 1)\ (1 + |u_{N+1}(z)|) + |u_{N+1}(z)| \\ &= q_{N+1}(z) - 1 \end{aligned}$$

where for the second line we used that (2.49) holds for p_N.

Now we apply (2.49) to (2.48). This shows that $|p_N(z) - 1| \le q_N(z) - 1 \le C_1 - 1$ for all $N \in \mathbb{N}$ and all $z \in S$. Taking $C_2 = C_1$, one obtains $|p_N(z)| \le C_2$ for all $N \in \mathbb{N}$ and $z \in S$. Since $\sum_{n=1}^{\infty} |u_n(z)|$ is assumed to converge uniformly, for every $0 < \epsilon < 1/2$ there exists an $N_0 \in \mathbb{N}$ such that

$$\textstyle\sum_{n=N+1}^{\infty} |u_n(z)| < \epsilon \tag{2.50}$$

for all $N > N_0$ and for all $z \in S$. Now, for $M > N > N_0$ holds

$$\begin{aligned} |p_M(z) - p_N(z)| &= |p_N(z)| \left| \prod_{n=N+1}^{M} [1 + u_n(z)] - 1 \right| \\ &\le |p_N(z)| \left(\prod_{n=N+1}^{M} [1 + |u_n(z)|] - 1 \right) \\ &\le |p_N(z)| \, (\mathrm{e}^{\epsilon} - 1) \le 2 \, |p_N(z)| \, \epsilon \le 2 \, C_2 \, \epsilon \end{aligned} \tag{2.51}$$

where the first inequality follows from (2.49), whereas the last line from (2.48), (2.50), and from the uniform boundedness of p_N. This last result shows that p_N is a Cauchy sequence in S which converges uniformly to a limit function f on S.

Finally (2.51) implies $|p_M(z)| \ge (1 - 2\epsilon)|p_{N_0}(z)|$ for all $M > N_0$. It follows for the limit function f that

$$|f(z)| \ge (1 - 2\epsilon)|p_{N_0}(z)|$$

for all $z \in S$. This shows that $f(z) = 0$ if and only if $p_{N_0}(z) = 0$, i.e. if and only if $u_n(z) = -1$ for some n. □

With this we are able to prove Theorem 2.18.

Proof (Theorem 2.18). Without loss of generality, we assume that $k = 0$. Again, we define the individual factors b_n of the Blaschke product by (2.47), and consider the term $1 - b_n$ inside the unit disk. Adding the term $1/|\alpha_n| - 1/|\alpha_n|$ followed by a straight forward rearranging yields

$$\begin{aligned} 1 - b_n(z) &= \frac{1}{|\alpha_n|} \left(1 - \frac{|\alpha_n|^2 - \overline{\alpha}_n z}{1 - \overline{\alpha}_n z} \right) - \frac{1 - |\alpha_n|}{|\alpha_n|} \\ &= (1 - |\alpha_n|) \, \frac{|\alpha_n| + \overline{\alpha}_n z}{(1 - \overline{\alpha}_n z) \, |\alpha_n|} \, . \end{aligned}$$

Remembering that all zeros α_n are inside the unit disk $\mathbb{D}$ gives

$$|1 - b_n(z)| \le (1 - |\alpha_n|) \, \frac{1 + r}{1 - r} \, , \quad \text{for all } |z| \le r < 1$$

and consequently

$$\sum_{n=1}^{\infty} |1 - b_n(z)| \leq \frac{1+r}{1-r} \sum_{n=1}^{\infty} (1 - |\alpha_n|) \, .$$

This shows that $\sum_{n=1}^{\infty} |1 - b_n(z)|$ converges uniformly on compact subsets of $\mathbb{D}$ since the zeros α_n satisfy the Blaschke condition (2.45).

Setting $u_n(z) = b_n(z) - 1$, Lemma 2.20 implies that $B(z) = \prod_{n=1}^{\infty} b_n(z)$ converges uniformly on compact subsets of $\mathbb{D}$ and that $B(z) = 0$ if and only if $B_n(z) = 0$ for some $n \in \mathbb{N}$. Since every b_n is holomorphic in $\mathbb{D}$ and since B converges uniformly on compact subsets of $\mathbb{D}$, $B(z)$ is also holomorphic in $\mathbb{D}$ (see e.g. [70, Theorem 10.28]). Moreover, since each factor b_n has absolute value less than 1 in $\mathbb{D}$, it follows that $|B(z)| < 1$ for all $z \in \mathbb{D}$, and consequently that $B \in H^\infty$ with $\|B\|_\infty \leq 1$.

Since $B \in H^\infty$, the boundary function $B(\mathrm{e}^{\mathrm{i}\theta})$ exists almost everywhere for $\theta \in [-\pi, \pi)$, and since $B(z)$ has absolute value smaller that 1 for all $z \in \mathbb{D}$, the boundary function has to satisfy $|B(\mathrm{e}^{\mathrm{i}\theta})| \leq 1$ almost everywhere. Now, let $B_N(z) = \prod_{n=1}^{N} b_n(z)$ the partial product of B. Then $B(z)/B_N(z)$ is again a Blaschke product, and thus holomorphic in $\mathbb{D}$. Consequently, it satisfies the mean value property. Together with the triangle inequality, one has

$$\frac{B(0)}{B_N(0)} \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \frac{B(\mathrm{e}^{\mathrm{i}\theta})}{B_N(\mathrm{e}^{\mathrm{i}\theta})} \right| \mathrm{d}\theta = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| B(\mathrm{e}^{\mathrm{i}\theta}) \right| \mathrm{d}\theta$$

where the last equality is a consequence of $|B_N(\mathrm{e}^{\mathrm{i}\theta})| = 1$ for all $\theta \in [-\pi, \pi)$. Now, letting $N \to \infty$, we obtain that $1 \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| B(\mathrm{e}^{\mathrm{i}\theta}) \right| \mathrm{d}\theta$. Consequently, since $|B(\mathrm{e}^{\mathrm{i}\theta})| \leq 1$, one gets $|B(\mathrm{e}^{\mathrm{i}\theta})| = 1$ almost everywhere. □

Thus, given a function $f \in H^p$ with $1 \leq p \leq \infty$ with zeros at $\{\alpha_n\}_{n \in \mathbb{N}}$ (which will satisfy the Blaschke condition (2.45)), we can form the Blaschke product $B \in H^\infty$ with the zeros of f. Now we can try to divide out the zeros of f by dividing f by the corresponding Blaschke product B. Of course, the resulting quotient $g := f/B$ is again a holomorphic function in $\mathbb{D}$, and since B has absolute value 1 almost everywhere on the unit circle, we even expect that g may have the same H^p-norm as the original f. That this reasoning is indeed true is shown by the next theorem.

Theorem 2.21. *Let $f \in H^p$ with $1 \leq p \leq \infty$, let B be the Blaschke product (2.46) formed with the zeros of f, and set $g(z) := f(z)/B(z)$, $z \in \mathbb{D}$. Then $g \in H^p$ with $\|g\|_p = \|f\|_p$.*

Proof. Let $\{\alpha_n\}_{n=1}^{\infty}$ be the sequence of zeros of f in $\mathbb{D}$, and let $b_n(z)$ be the factor of the Blaschke product corresponding to the zero α_n as defined in (2.47). Moreover, let

$$B_N(z) = \prod_{n=1}^{N} b_n(z) \, , \qquad z \in \mathbb{D}$$

be the partial Blaschke product formed by the first N zeros of f, and let $g_N = f/B_N$. For every fixed N, $B_N(r\,\mathrm{e}^{\mathrm{i}\theta}) \to 1$ uniformly as $r \to 1$. It follows that $g_N(r\,\mathrm{e}^{\mathrm{i}\theta}) \to f(\mathrm{e}^{\mathrm{i}\theta})$ and consequently that

$$\|g_N\|_p = \|f\|_p \, . \tag{2.52}$$

Since $|b_n(z)| < 1$ for all $z \in \mathbb{D}$ and all n, we have that

$$0 \le |g_1(z)| \le |g_2(z)| \le \cdots \le \infty \quad \text{and} \quad |g_n(z)| \to |g(z)|$$

for every $z \in \mathbb{D}$. Fixing $0 < r < 1$, set $g_r(z) := g(r\,z)$ and $(g_N)_r(z) := g_N(r\,z)$, and applying Lebesgue's monotone convergence theorem, one gets

$$\lim_{N\to\infty} \|(g_N)_r\|_p^p = \lim_{N\to\infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} |g_N(r\,\mathrm{e}^{\mathrm{i}\theta})|^p \,\mathrm{d}\theta = \frac{1}{2\pi} \int_{-\pi}^{\pi} |g(r\,\mathrm{e}^{\mathrm{i}\theta})|^p \,\mathrm{d}\theta = \|g_r\|_p^p \, .$$

Since g_N is analytic in $\mathbb{D}$ and because of (2.52), the left hand side is upper bounded by $\|f\|_p^p$ for every $0 < r < 1$. Leting $r \to 1$, we obtain $\|g\|_p \le \|f\|_p$. However, since $|B(z)| \le 1$ for all $z \in \mathbb{D}$, we also have that $|g(z)| \ge |f(z)|$ for all $z \in \mathbb{D}$, which shows that we even have equality, i.e. that $\|g\|_p = \|f\|_p$. □

2.2.3 Inner-Outer factorization

The last theorem showed that every function $f \in H^p$ can be factorized into a Blaschke product and a function without zeros in the unit disk. This section considers a somewhat different factorization of functions in H^p into so called inner and outer functions.

Definition 2.22 (Inner and Outer functions). *An* inner function *is an analytic function $f \in H^\infty$ such that $|f(z)| \le 1$ in the unit disk and such that $|f(\mathrm{e}^{\mathrm{i}\theta})| = 1$ almost everywhere on the unit circle.*
An outer function *is an analytic function O in the unit disk $\mathbb{D}$ of the form*

$$O(z) = c\,\exp\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \log\phi(\mathrm{e}^{\mathrm{i}\tau})\,\frac{\mathrm{e}^{\mathrm{i}\tau}+z}{\mathrm{e}^{\mathrm{i}\tau}-z}\,\mathrm{d}\tau\right) \, . \tag{2.53}$$

Here c is a constant with $|c| = 1$, and ϕ is a positive measurable function on $\mathbb{T}$ such that $\log\phi \in L^1$.

Every Blaschke product is an inner function by Theorem 2.18. However, there exist other inner functions. The following theorem characterizes all inner functions as the product of a Blaschke product and a so-called singular function.

Theorem 2.23. *Let $f \in H^p$ be an inner function and let B be the Blaschke product formed with the zeros of f. Then there exists a positive Borel measure μ on $\mathbb{T}$ which is singular with respect to Lebesgue measure and a complex constant c with $|c| = 1$ such that*

$$f(z) = B(z)\,S(z) \, , \qquad z \in \mathbb{D} \tag{2.54}$$

with

$$S(z) = c\,\exp\left(-\int_{-\pi}^{\pi} \frac{\mathrm{e}^{\mathrm{i}\tau}+z}{\mathrm{e}^{\mathrm{i}\tau}-z}\,\mathrm{d}\mu(\tau)\right) \, , \qquad z \in \mathbb{D} \, . \tag{2.55}$$

We will call the function S in the factorization (2.54) a *singular function*. Thus, an inner factor is the product of a Blaschke product B and of a singular function S, in general. Nevertheless, it may happen that B, or S, or even both factors are identically 1.

Example 2.24. Probably the simplest singular function is obtained by taking μ in (2.55) to be the unit mass at $\tau = 0$ and by letting $c = 1$. This yields the singular function

$$S(z) = \exp\left(\frac{z+1}{z-1}\right), \qquad z \in \mathbb{D}.$$

It is a holomorphic function in $\mathbb{D}$ with an essential singularity at $z = 1$.

Proof (Theorem 2.23). Let $g := f/B$, then g is a holomorphic function without any zeros in $\mathbb{D}$, from which follows that $\log|g|$ is a harmonic function in $\mathbb{D}$ (see e.g. [70, Theorem 13.12]). By Theorem 2.21 it follows that $|g(z)| \leq 1$ for $z \in \mathbb{D}$ and that $|g(\mathrm{e}^{\mathrm{i}\theta})| = 1$ almost everywhere, which implies that $\log|g| \leq 0$ in $\mathbb{D}$ and $\log|g(\mathrm{e}^{\mathrm{i}\theta})| = 0$ a.e. on $\mathbb{T}$. It is known that every bounded harmonic function in $\mathbb{D}$ can be represented by the Poisson integral of a unique Borel measure on $\mathbb{T}$ (see e.g. [70, Theorem 11.30]). We conclude for our case that $\log|g|$ is the Poisson integral of $-\mathrm{d}\mu$ with some positive Borel measure μ on $\mathbb{D}$. However, since $\log|g(\mathrm{e}^{\mathrm{i}\theta})| = 0$ a.e. on $\mathbb{T}$ the measure μ has to be singular (with respect to Lebesgue measure). Now $\log|g|$, as the Poisson integral of $-\mathrm{d}\mu$, is the real part of the function

$$G(z) = -\int_{-\pi}^{\pi} \frac{\mathrm{e}^{\mathrm{i}\tau} + z}{\mathrm{e}^{\mathrm{i}\tau} - z}\, \mathrm{d}\mu(\tau)$$

(see e.g. Section 5.1) which implies that S has the form (2.55). □

The previous theorem clarified the general form of an inner function whereas the general form of an outer function is given by (2.53). The next theorem studies basic properties of outer functions needed frequently throughout this book.

Theorem 2.25. *Let O_ϕ be an outer function related to a positive (real valued) measurable function ϕ as in Definition 2.22. Then*

(*a*) $\log|O_\phi|$ *is the Poisson integral of* $\log\phi$.
(*b*) $\lim_{r\to 1}|O_\phi(r\mathrm{e}^{\mathrm{i}\theta})| = \phi(\mathrm{e}^{\mathrm{i}\theta})$ *for almost all* $\theta \in [-\pi, \pi)$.
(*c*) $O_\phi \in H^p$ *if and only if* $\phi \in L^p$ *and* $\|O_\phi\|_p = \|\phi\|_p$.

Proof. Statement (*a*) follows from the definition of the outer function (2.53) since the exponent of O_ϕ can be written as

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \log\phi(\mathrm{e}^{\mathrm{i}\tau})\, \frac{\mathrm{e}^{\mathrm{i}\tau} + r\mathrm{e}^{\mathrm{i}\theta}}{\mathrm{e}^{\mathrm{i}\tau} - r\mathrm{e}^{\mathrm{i}\theta}}\, \mathrm{d}\tau = (\mathfrak{P}\log\phi)(r\mathrm{e}^{\mathrm{i}\theta}) + \mathrm{i}\,(\mathfrak{Q}\log\phi)(r\mathrm{e}^{\mathrm{i}\theta})$$

with the Poisson integral $\mathfrak{P}\log\phi$ and the conjugate Poisson integral $\mathfrak{Q}\log\phi$ of $\log\phi$ (cf. Section 5.1). It follows that $|O_\phi| = \exp(\mathfrak{P}\log\phi)$ which proves (a).

The Poisson integral $(\mathfrak{P}f)(r\mathrm{e}^{\mathrm{i}\theta})$ of a function $f \in L^1$ converges to f in the L^1-norm as $r \to 1$ (see e.g. Theorem 5.3 for a proof). By this property of the Poisson integral (a) implies (b).

Applying statement (b), we have

$$\begin{aligned}\|\phi\|_p^p &= \frac{1}{2\pi}\int_{-\pi}^{\pi} |\phi(\mathrm{e}^{\mathrm{i}\theta})|^p \,\mathrm{d}\theta = \frac{1}{2\pi}\int_{-\pi}^{\pi} \lim_{r\to 1} |O_\phi(r\mathrm{e}^{\mathrm{i}\theta})|^p \,\mathrm{d}\theta \\ &\le \lim_{r\to 1} \frac{1}{2\pi}\int_{-\pi}^{\pi} |O_\phi(r\mathrm{e}^{\mathrm{i}\theta})|^p \,\mathrm{d}\theta = \|O_\phi\|_p^p\end{aligned}$$

where the inequality follows from Fatou's Lemma (see e.g. [70, § 1.28]). Thus $\|\phi\|_p \le \|O\phi\|_p$. For the converse assume that $\phi \in L^p$. Then Jensen's inequality (cf. [70, §3.3]) gives

$$\begin{aligned}|O_\phi(r\mathrm{e}^{\mathrm{i}\theta})|^p &= \exp\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \log\phi^p(\mathrm{e}^{\mathrm{i}\tau})\,\mathcal{P}_r(\theta-\tau)\,\mathrm{d}\tau\right) \\ &\le \frac{1}{2\pi}\int_{-\pi}^{\pi} \phi^p(\mathrm{e}^{\mathrm{i}\tau})\,\mathcal{P}_r(\theta-\tau)\,\mathrm{d}\tau\end{aligned}$$

in which the left hand side of the inequality is just the Poisson integral $\mathfrak{P}\log\phi^p$ with the Poisson kernel $\mathcal{P}_r$ (cf. also (5.5) and (5.4) for the definition). Integration of the last inequality with respect to θ and using that the Poisson kernel $\mathcal{P}_r$ satisfies $\int_{-\pi}^{\pi} \mathcal{P}_r(\theta)\,\mathrm{d}\theta = 1$ (cf. Section 5.2) gives $\|O_\phi\|_p \le \|\phi\|_p$. This finishes the proof of (c). □

Finally, the following theorem will give the desired factorization result under point (c). It shows that every function $f \in H^p$ can always be factorized into an inner and an outer function.

Theorem 2.26. *For $1 \le p \le \infty$ let $f \in H^p$ be a nonzero function. Then*

(a) $\log|f| \in L^1$.

(b) *the outer function defined by*

$$O_f(z) = \exp\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \log\left|f(\mathrm{e}^{\mathrm{i}\tau})\right| \frac{\mathrm{e}^{\mathrm{i}\tau}+z}{\mathrm{e}^{\mathrm{i}\tau}-z}\,\mathrm{d}\tau\right), \qquad z \in \mathbb{D}$$

is an element of H^p.

(c) *there exists an inner function I_f such that $f = O_f\, I_f$.*

Proof. We consider first the case $p = 1$. Assume that $f \in H^1$, let B be the Blaschke product (2.46) formed with the zeros of f, and set $g = f/B$. By Theorem 2.21 $g \in H^1$ and $|g(\mathrm{e}^{\mathrm{i}\theta})| = |f(\mathrm{e}^{\mathrm{i}\theta})|$ for almost all $\theta \in [-\pi, \pi)$. Therefore, it is sufficient to prove the theorem for g instead of f. Since g is holomorphic without any zero in $\mathbb{D}$ the function $\log|g|$ is harmonic in $\mathbb{D}$ (see e.g. [70, § 13.12]) and therefore it satisfies (2.43). Since $g(0) \neq 0$ we assume, without loss of generality, that $g(0) = 1$ and define the two functions

$$\log^+ x := \begin{cases} 0 & , \quad x < 1 \\ \log x \, , & \quad x \geq 1 \end{cases} \qquad \text{and} \qquad \log^- x := \begin{cases} \log(1/x) \, , & \quad x < 1 \\ 0 & , \quad x \geq 1 \end{cases}$$

on the positive real axis, such that obviously $\log x = \log^+ x - \log^- x$. Therewith, it follows from (2.43) that

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \log^- |g(\mathrm{e}^{\mathrm{i}\theta})| \,\mathrm{d}\theta = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log^+ |g(\mathrm{e}^{\mathrm{i}\theta})| \,\mathrm{d}\theta \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |g(\mathrm{e}^{\mathrm{i}\theta})| \,\mathrm{d}\theta = \|g\|_1 \, .$$

This shows that $\log^+|g|$ and $\log^-|g|$ are in L^1, such that $\log|g| \in L^1$, which proves (*a*). It follows that O_f is a well defined outer function, and Theorem 2.25 (c) implies that $O_f \in H^1$, proving (*b*).

It remains to show (*c*). To this end we show next that $|g(z)| \leq |O_g(z)|$ for all $z \in \mathbb{D}$. Since we know from Theorem 2.25 that $\log|O_g|$ is equal to the Poisson integral of $\log|g|$, we have to show

$$\log|g(z)| \leq \log|O_g(z)| = (\mathfrak{P} \log|g|)(z) \, . \tag{2.56}$$

For $0 < r < 1$ and $z \in \mathbb{D}$ we define $g_r(z) := g(r\,z)$. Since g is a holomorphic function without zeros in $\mathbb{D}$, $\log|g_r|$ is harmonic in $\mathbb{D}$ (see e.g. [70, Theorem 13.12]) and can be represented as a Poisson integral. We therefore have

$$\log|g_r(z)| = \mathfrak{P}\left[\log|g_r|\right](z) = \mathfrak{P}\left[\log^+|g_r|\right](z) - \mathfrak{P}\left[\log^-|g_r|\right](z) \, . \tag{2.57}$$

We know from Theorem 2.11 that $g_r(\mathrm{e}^{\mathrm{i}\theta}) \to g(\mathrm{e}^{\mathrm{i}\theta})$ as $r \to 1$. It follows that the left hand side of (2.57) converges to $\log|g|$ and that the first term on the right hand side converges to $\mathfrak{P}\left[\log^+|g|\right](z)$ as $r \to 1$. This last statement follows from

$$\begin{aligned} &\left|\mathfrak{P}\left[\log^+|g_r|\right](\rho\,\mathrm{e}^{\mathrm{i}\theta}) - \mathfrak{P}\left[\log^+|g|\right](\rho\,\mathrm{e}^{\mathrm{i}\theta})\right| \\ &\qquad \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} \left|\log^+|g_r(\mathrm{e}^{\mathrm{i}\tau})| - \log^+|g(\mathrm{e}^{\mathrm{i}\tau})|\right| \mathcal{P}_\rho(\theta - \tau) \,\mathrm{d}\tau \\ &\qquad\qquad \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} \left||g_r(\mathrm{e}^{\mathrm{i}\tau})| - |g(\mathrm{e}^{\mathrm{i}\tau})|\right| \mathcal{P}_\rho(\theta - \tau) \,\mathrm{d}\tau \end{aligned}$$

and from the fact that $g_r(\mathrm{e}^{\mathrm{i}\theta}) \to g(\mathrm{e}^{\mathrm{i}\theta})$. Here $\mathcal{P}_\rho$ denotes the Poisson kernel (see (5.5) and (5.4) for the definition), and the last line was obtained using the relation $|\log^+ u - \log^+ v| \leq |u - v|$ for all real numbers u, v, which may easily be verified. Therewith, letting $r \to 1$ in (2.57), one obtains

$$\mathfrak{P}\left[\log^-|g|\right](z) \leq \lim_{r \to 1} \mathfrak{P}\left[\log^-|g_r|\right](z) = \mathfrak{P}\left[\log^+|g_r|\right](z) - \log|g(z)|$$

where the first inequality follows from Fatou's Lemma (see e.g. [70, § 1.28]). Combining $\log^+$ and $\log^-$ to log, one obtains the desired relation (2.56), which proves that $|g(z)| \leq |O_g(z)|$ for all $z \in \mathbb{D}$.

Now we define the function

$$I_g(z) := \frac{g(z)}{O_g(z)} \qquad z \in \mathbb{D} .$$

Obviously, I_g is analytic in $\mathbb{D}$, $|I_g(z)| \leq 1$ for all $z \in \mathbb{D}$ and $|I_g(\mathrm{e}^{\mathrm{i}\theta})| = 1$ almost everywhere. Thus I_g is an inner function. □

By Theorem 2.23, every inner function can be written as the product of a Blaschke product and an singular function. Consequently, it follows form point (c) of the previous theorem that every $f \in H^p$ may be written as

$$f(z) = O_f(z)\, B_f(z)\, S_f(z) \;, \qquad z \in \mathbb{D}$$

with an outer function O_f, the Blaschke product B_f formed with the zeros of f, and a singular function S_f.

2.3 Vector-valued Hardy Spaces

The previous section introduced the Hardy space of complex valued functions. In general, it is possible to extend the concept of Hardy spaces to functions taking values in arbitrary Banach spaces. To give a completely satisfactory definition of such spaces, one needs some results from the integration theory of functions with values in Banach spaces. Although this is a straight forward generalization of the standard integration of complex valued Lebesgue measurable functions, it would be out of the scope of our intentions here. However, for the case of functions with values in a separable or even a finite dimensional Hilbert spaces, almost the whole theory can be led back to the scalar case of the previous section. Therefore, we shall restrict ourselves to these cases. Later, we will be especially interested in the finite dimensional case, since this is the suitable framework for modeling linear systems with a finite number of inputs and outputs. Nevertheless, the basic definitions are given for the slightly more general case of separable Hilbert spaces.

We start with a formal extension of ℓ^p and L^p spaces to the case of vector valued functions. To emphasize the difference to the scalar case, vector valued functions will be denoted by bold face letters.

Definition 2.27 (Vector-valued ℓ^p spaces). *Let $\mathcal{H}$ be a separable Hilbert space and let $\hat{\boldsymbol{f}} = \{\hat{\boldsymbol{f}}(k)\}_{k=-\infty}^{\infty}$ be a double infinite sequence of elements from $\mathcal{H}$. For $1 \leq p < \infty$ and $p = \infty$ define*

$$\|\hat{\boldsymbol{f}}\|_{\ell^p} := \left(\sum_{k=-\infty}^{\infty} \|\hat{\boldsymbol{f}}(k)\|_{\mathcal{H}}^p \right)^{1/p} \qquad \textit{and} \quad \|\hat{\boldsymbol{f}}\|_{\ell^\infty} := \sup_{k\in\mathbb{Z}} \|\hat{\boldsymbol{f}}(k)\|_{\mathcal{H}}$$

respectively.
Then, for $1 \leq p \leq \infty$ *the set* $\ell^p(\mathcal{H})$ *denotes the set of all double infinite sequences* $\hat{\boldsymbol{f}} = \{\hat{\boldsymbol{f}}(k)\}_{k=-\infty}^{\infty}$ *with values in* $\mathcal{H}$ *for which* $\|\hat{\boldsymbol{f}}\|_{\ell^p} < \infty$ *and* $\ell^p_+(\mathcal{H})$ *denotes the set of all infinite sequences* $\hat{\boldsymbol{f}} = \{\hat{\boldsymbol{f}}(k)\}_{k=0}^{\infty}$ *in* $\mathcal{H}$ *with* $\|\hat{\boldsymbol{f}}\|_{\ell^p} < \infty$*.*

Of course, $\ell^p_+(\mathcal{H})$ my be considered as the subspace of $\ell^p(\mathcal{H})$ in which for all elements $\hat{\boldsymbol{f}}$ holds that $\hat{\boldsymbol{f}}(k) = 0$ for all $k < 0$. Moreover, it is clear that for the special case $\mathcal{H} = \mathbb{C}$, one obtains again the usual ℓ^p spaces. As in the scalar case, $\ell^2(\mathcal{H})$ is a Hilbert space with the inner product

$$\langle \hat{\boldsymbol{f}}, \hat{\boldsymbol{g}} \rangle_{\ell^2(\mathcal{H})} = \sum_{k=-\infty}^{\infty} \langle \hat{\boldsymbol{f}}(k), \hat{\boldsymbol{g}}(k) \rangle_{\mathcal{H}} \,.$$

Definition 2.28 (Vector-valued L^p spaces). *Let* $\mathcal{H}$ *be a separable Hilbert space and let* f *be a measurable function with values in* $\mathcal{H}$*. For* $1 \leq p < \infty$ *define*

$$\|\boldsymbol{f}\|_p := \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \|\boldsymbol{f}(\mathrm{e}^{\mathrm{i}\theta})\|_{\mathcal{H}}^p \,\mathrm{d}\theta \right)^{1/p}$$

and for $p = \infty$ *define*

$$\|\boldsymbol{f}\|_\infty := \operatorname*{ess\,sup}_{\zeta \in \mathbb{T}} \|\boldsymbol{f}(\zeta)\|_{\mathcal{H}} \,.$$

Then for $1 \leq p \leq \infty$ *the set* $L^p(\mathcal{H})$ *denotes the set of all measurable functions* $\boldsymbol{f}$ *with values in* $\mathcal{H}$ *for which* $\|\boldsymbol{f}\|_p < \infty$*.*

Of course, if the dimension of the Hilbert space $\mathcal{H}$ is one, the above definition of $L^p(\mathcal{H})$ coincides with the usual L^p-spaces on the unit circle, and $L^2(\mathcal{H})$ is a Hilbert space with the inner product

$$\langle \boldsymbol{f}, \boldsymbol{g} \rangle_{L^2(\mathcal{H})} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \langle \boldsymbol{f}(\mathrm{e}^{\mathrm{i}\theta}), \boldsymbol{g}(\mathrm{e}^{\mathrm{i}\theta}) \rangle_{\mathcal{H}} \,\mathrm{d}\theta \,.$$

Since $\mathcal{H}$ is assumed to be separable there exists a complete orthonormal basis $\{\boldsymbol{e}_n\}_{n=1}^{\infty}$ in $\mathcal{H}$ such that every $\boldsymbol{f} \in \mathcal{H}$ can be written as $\boldsymbol{f} = f_1\,\boldsymbol{e}_1 + f_2\,\boldsymbol{e}_2 + \cdots + f_n\,\boldsymbol{e}_n + \cdots$ where $f_n := \langle \boldsymbol{f}, \boldsymbol{e}_n \rangle_{\mathcal{H}}$ are the components of $\boldsymbol{f}$ with respect to the basis $\{\boldsymbol{e}_n\}_{n=1}^{\infty}$ and the norm of $\boldsymbol{f}$ in $\mathcal{H}$ is just the ℓ^2-norm of the sequence $(f_1, f_2, \cdots)$ of its components: $\|\boldsymbol{f}\|_{\mathcal{H}}^2 = \sum_{n=1}^{\infty} |f_n|^2$. Moreover, given a sequence $\hat{\boldsymbol{f}} = \{\hat{\boldsymbol{f}}(k)\}_{k=-\infty}^{\infty}$ with values in the Hilbert space $\mathcal{H}$, we can define its *coordinate sequences* $\hat{f}_n = \{\hat{f}_n(k)\}_{k=-\infty}^{\infty}$ with $n = 1, 2, \ldots$ by

$$\hat{f}_n(k) = \langle \hat{\boldsymbol{f}}(k), \boldsymbol{e}_n \rangle_{\mathcal{H}} \,, \qquad k \in \mathbb{Z} \,, n \in \mathbb{N} \,.$$

Similarly, given a function $\boldsymbol{f}$ on the unit circle with values in $\mathcal{H}$, its *coordinate functions* f_n, $n = 1, 2, \ldots$ are defined by

$$f_n(\zeta) = \langle \boldsymbol{f}(\zeta), \boldsymbol{e}_n \rangle_{\mathcal{H}}\,, \qquad \zeta \in \mathbb{T}\,,\, n \in \mathbb{N}\,.$$

The following proposition gives a characterization of the spaces $\ell^p(\mathcal{H})$ and $L^p(\mathcal{H})$ in terms of the individual coordinates. It will be show that $\ell^p(\mathcal{H})$ is equivalent to the set of all sequences $\hat{\boldsymbol{f}} = \{\hat{f}_n\}_{n=1}^{\infty}$ whose individual entries $\hat{f}_n = \{\hat{f}_n(k)\}_{k=-\infty}^{\infty}$ belong to ℓ^p and it will be show that $L^p(\mathcal{H})$ is precisely the set of all sequences $\boldsymbol{f} = \{f_n\}_{n=1}^{\infty}$ whose individual components f_n are elements of L^p.

Proposition 2.29. *Let $\mathcal{H}$ be a separable Hilbert space with an arbitrary orthonormal basis $\{\boldsymbol{e}_n\}_{n=1}^{\infty}$ and let $1 \le p \le \infty$.*
A sequence $\hat{\boldsymbol{f}}$ of elements in $\mathcal{H}$ belongs to $\ell^p(\mathcal{H})$ if and only if all coordinate sequences $\hat{f}_n = \langle \hat{\boldsymbol{f}}, \boldsymbol{e}_n \rangle_{\mathcal{H}}$, $n \in \mathbb{N}$ belong to ℓ^p.
A function $\boldsymbol{f}$ on the unit circle $\mathbb{T}$ and with values in $\mathcal{H}$ belongs to $L^p(\mathcal{H})$ if and only if all coordinate functions $f_n = \langle \boldsymbol{f}, \boldsymbol{e}_n \rangle$, $n \in \mathbb{N}$ belong to L^p.

Proof. We prove the statement for $L^p(\mathcal{H})$. By the identification of $\mathcal{H}$ with ℓ^2 and with the triangle inequality, one has

$$\|\boldsymbol{f}(\zeta)\|_{\mathcal{H}} = \left(\textstyle\sum_{n=1}^{\infty} |f_n(\zeta)|^2\right)^{1/2} \le \textstyle\sum_{n=1}^{N} |f_n(\zeta)| \qquad \text{for every } \zeta \in \mathbb{T}$$

and provided that the right hand side exists. Therewith one gets at once $\|\boldsymbol{f}\|_{\infty} \le \sum_{n=1}^{\infty} \|f_n\|_{\infty}$. For $p < \infty$ we take both sides to the power p, integrate over the unit circle $\mathbb{T}$, and apply Minkowski's inequality to the right hand side integral. This gives $\|\boldsymbol{f}\|_p \le \sum_{n=1}^{\infty} \|f_n\|_p$ which proves the “if” part of the proposition.
To verify the “only if” part, note that by the identification of $\mathcal{H}$ with ℓ^2 one has that $\|\boldsymbol{f}(\zeta)\|_{\mathcal{H}}^p \ge |f_n(\zeta)|^p$ for all $\zeta \in \mathbb{T}$ and for all $n = \mathbb{N}$. This gives immediately $\|\boldsymbol{f}\|_{\infty} \ge \|f_n\|_{\infty}$ and for $p < \infty$, the integration of both sides, gives $\|\boldsymbol{f}\|_p \ge \|f_n\|_p$ for every $1 \le n \le N$.
The analogous proof for $\ell^p(\mathcal{H})$ is left as an exercise. □

As in the scalar case, we want to consider the Fourier series expansion of functions in $L^p(\mathcal{H})$. To avoid the introduction of the integration over functions with values in Hilbert spaces, the Fourier series are introduced in the weak sense, as follows: Let $\boldsymbol{f} \in L^1(\mathcal{H})$, we want to write $\boldsymbol{f}$ in the form

$$\boldsymbol{f}(\mathrm{e}^{\mathrm{i}\theta}) = \sum_{k=-\infty}^{\infty} \hat{\boldsymbol{f}}(k)\, \mathrm{e}^{\mathrm{i}k\theta} \tag{2.58}$$

where $\{\hat{\boldsymbol{f}}(k)\}_{k=-\infty}^{\infty}$ is a sequence of elements from the Hilbert space $\mathcal{H}$. We call this expansion Fourier series of $\boldsymbol{f}$ if for every $\boldsymbol{g} \in \mathcal{H}$

$$\left\langle \boldsymbol{f}(\mathrm{e}^{\mathrm{i}\theta}), \boldsymbol{g} \right\rangle_{\mathcal{H}} = \sum_{k=-\infty}^{\infty} \left\langle \hat{\boldsymbol{f}}(k), \boldsymbol{g} \right\rangle_{\mathcal{H}} \mathrm{e}^{\mathrm{i}k\theta}$$

is an ordinary Fourier series of the complex valued function $\langle \boldsymbol{f}(\mathrm{e}^{\mathrm{i}\theta}), \boldsymbol{g} \rangle_{\mathcal{H}}$. The coefficients of the scalar Fourier series are of course given by (2.1), thus

$$\langle \hat{\boldsymbol{f}}(k), \boldsymbol{g} \rangle_{\mathcal{H}} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \langle \boldsymbol{f}(\mathrm{e}^{\mathrm{i}\theta}), \boldsymbol{g} \rangle_{\mathcal{H}} \mathrm{e}^{-\mathrm{i}k\theta} \, \mathrm{d}\theta \, . \tag{2.59}$$

It may not be immediately clear whether there exists such a sequence $\{\hat{\boldsymbol{f}}(k)\}_{k=-\infty}^{\infty}$ of vectors in $\mathcal{H}$ such that (2.59) is satisfied for all $\boldsymbol{g} \in \mathcal{H}$. However, for a fixed $\boldsymbol{f} \in L^1(\mathcal{H})$ and $k \in \mathbb{Z}$ the right hand side of (2.59) defines a conjugate-linear functional Φ_k on $\mathcal{H}$ which is also continuous, since by the Cauchy-Schwarz inequality

$$\begin{aligned} |\Phi_k(\boldsymbol{g})| &\leq \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \langle \boldsymbol{f}(\mathrm{e}^{\mathrm{i}\tau}), \boldsymbol{g} \rangle_{\mathcal{H}} \right| \, \mathrm{d}\tau \\ &\leq \frac{1}{2\pi} \int_{-\pi}^{\pi} \|\boldsymbol{f}(\mathrm{e}^{\mathrm{i}\tau})\|_{\mathcal{H}} \, \|\boldsymbol{g}\|_{\mathcal{H}} \, \mathrm{d}\tau = \|\boldsymbol{f}\|_1 \, \|\boldsymbol{g}\|_{\mathcal{H}} \, . \end{aligned}$$

But this implies by the Riesz representation theorem (for Hilbert spaces) that there exists a unique $\hat{\boldsymbol{f}}(k) \in \mathcal{H}$ such that (2.59) holds for all $\boldsymbol{g} \in \mathcal{H}$. To determine the coefficient vectors $\hat{\boldsymbol{f}}(k)$ in the Fourier series (2.58), one may choose an orthonormal basis $\{\boldsymbol{e}_n\}$ in $\mathcal{H}$ and write every coefficient vector in this basis as $\hat{\boldsymbol{f}}(k) = \sum_{l \in \mathbb{N}} \hat{f}_l(k) \, \boldsymbol{e}_l$. Plugging this representation into (2.59) together with a $\boldsymbol{g} = \boldsymbol{e}_n$ gives

$$\hat{f}_n(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f_n(\mathrm{e}^{\mathrm{i}\tau}) \, \mathrm{e}^{-\mathrm{i}k\tau} \, \mathrm{d}\tau$$

where $f_n(\mathrm{e}^{\mathrm{i}\tau}) = \langle \boldsymbol{f}(\mathrm{e}^{\mathrm{i}\tau}), \boldsymbol{e}_n \rangle_{\mathcal{H}}$ is the n-th coordinate of $\boldsymbol{f}(\mathrm{e}^{\mathrm{i}\theta})$ with respect to the orthonormal basis $\{\boldsymbol{e}_l\}$. Thus, if an orthonormal basis in $\mathcal{H}$ is fixed, the Fourier coefficients $\hat{\boldsymbol{f}}(k)$ of the series (2.58) can be determined componentwise. Therefore, we will simply write

$$\hat{\boldsymbol{f}}(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \boldsymbol{f}(\mathrm{e}^{\mathrm{i}\tau}) \, \mathrm{e}^{-\mathrm{i}k\tau} \, \mathrm{d}\tau \, , \qquad k = 0, \pm 1, \pm 2, \ldots \tag{2.60}$$

where the integration on the right hand side means a component wise integration of every coordinate with respect to the chosen orthonormal basis. The above discussion holds in particular for the case $\mathcal{H} = \mathbb{C}^N$ with the usual orthonormal basis $\boldsymbol{e}_1 = \{1, 0, 0, \cdots, 0\}$, $\boldsymbol{e}_2 = \{0, 1, 0, \cdots, 0\}$, ..., $\boldsymbol{e}_N = \{0, 0, 0, \cdots, 1\}$.

Due to the separation of the Fourier series into its components, it is not hard to verify that the *Parseval theorem*

$$\|\boldsymbol{f}\|_2^2 = \sum_{k=-\infty}^{\infty} \|\hat{\boldsymbol{f}}(k)\|_{\mathcal{H}}^2$$

holds for $L^2(\mathcal{H})$ because Parseval's equality holds for every single component f_n.

With these preparations, we can introduce vector valued Hardy spaces.

Definition 2.30 (Vector valued Hardy spaces). *Let $\mathcal{H}$ be a separable Hilbert space, and let $1 \le p \le \infty$. Then $H^p(\mathcal{H})$ denotes the subset of $L^p(\mathcal{H})$ of all $\boldsymbol{f} \in L^p(\mathcal{H})$ whose Fourier coefficients* (2.60) *with negative indices vanish, thus*

$$H^p(\mathcal{H}) := \{\boldsymbol{f} \in L^p(\mathcal{H}) : \hat{\boldsymbol{f}}(k) = 0 \text{ for all } k < 0\} .$$

Let $\{\boldsymbol{e}_n\}_{n=1}^{\infty}$ be an arbitrary orthonormal basis in $\mathcal{H}$, then the above discussions on the Fourier series expansion make it clear that the definition of the Hardy spaces $H^p(\mathcal{H})$ is equivalent to the following statement

Proposition 2.31. *A function $\boldsymbol{f} \in L^p(\mathcal{H})$ belongs to $H^p(\mathcal{H})$ if and only if all of its coordinate functions $f_n = \langle \boldsymbol{f}, \boldsymbol{e}_n \rangle_{\mathcal{H}}$ are elements of H^p.*

Similar to the case of scalar functions, every $\boldsymbol{f} \in H^p(\mathcal{H})$ can be associated with a function

$$\boldsymbol{F}(z) := \sum_{k=0}^{\infty} \hat{\boldsymbol{f}}(k)\, z^k \qquad (z \in \mathbb{D}) \tag{2.61}$$

which is analytic for all $z \in \mathbb{D}$, where $\hat{\boldsymbol{f}}(k)$ are the Fourier coefficients (2.60) of $\boldsymbol{f}$. As in the case of scalar functions, we need to show that the function $\boldsymbol{F}(r\mathrm{e}^{\mathrm{i}\theta})$ converges to $\boldsymbol{f}(\mathrm{e}^{\mathrm{i}\theta})$ in $L^p(\mathcal{H})$ as $r \to 1$. However, since the general proof can be simply reduced to the scalar case, we just state the result, which is completely analog to the scalar case, but omit the lengthy and technical proof.

Theorem 2.32. *Let $\mathcal{H}$ be a separable Hilbert space, let $1 \le p \le \infty$ and let $\boldsymbol{f} \in H^p(\mathcal{H})$. Define $\boldsymbol{F}_r(\mathrm{e}^{\mathrm{i}\theta}) = \boldsymbol{F}(r\mathrm{e}^{\mathrm{i}\theta})$ with $\boldsymbol{F}$ given by* (2.61)*. Then it holds that*

$$\lim_{r \to 1} \|\boldsymbol{F}_r - \boldsymbol{f}\|_p = 0 .$$

2.4 Operator-valued Analytic Functions

Next we consider analytic functions with values in the space of bounded linear operators. Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be separable Hilbert spaces and denote by $\mathscr{B}(\mathcal{H}_1, \mathcal{H}_2)$ the set of all bounded linear operators $\mathbf{H}$ from $\mathcal{H}_1$ to $\mathcal{H}_2$. It is known that $\mathscr{B}(\mathcal{H}_1, \mathcal{H}_2)$ is a Banach space with respect to the usual operator norm

$$\|\mathbf{H}\|_{\mathcal{H}_1 \to \mathcal{H}_2} := \sup_{\boldsymbol{f} \in \mathcal{H}_1, \|\boldsymbol{f}\|_{\mathcal{H}_1} \le 1} \|\mathbf{H}\boldsymbol{f}\|_{\mathcal{H}_2} . \tag{2.62}$$

Therewith, we define operator valued bounded analytic functions.

Definition 2.33 (Bounded analytic functions). *Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two separable Hilbert spaces, and let $\{\hat{\mathbf{H}}(k)\}_{k=0}^{\infty}$ be a sequence of elements in $\mathscr{B}(\mathcal{H}_1, \mathcal{H}_2)$. Therewith, we define the function*

$$\mathbf{H}(z) = \sum_{k=0}^{\infty} \hat{\mathbf{H}}(k)\, z^k , \qquad z \in \mathbb{D} . \tag{2.63}$$

Now $H^\infty(\mathcal{H}_1, \mathcal{H}_2)$ *denotes the set of all* bounded analytic functions with values in $\mathscr{B}(\mathcal{H}_1, \mathcal{H}_2)$*, that is the set of all functions of the form* (2.63) *which are uniformly bounded in* $\mathbb{D}$*, i.e. for which*

$$\|\mathbf{H}\|_\infty := \sup_{z\in\mathbb{D}} \|\mathbf{H}(z)\|_{\mathcal{H}_1\to\mathcal{H}_2} < \infty .$$

Equation (2.63) means that the power series on the right hand side is assumed to converge in $\mathscr{B}(\mathcal{H}_1, \mathcal{H}_2)$ for every $z \in \mathbb{D}$. It shows that $\mathbf{H}$ is holomorphic in $\mathbb{D}$. Since $\mathscr{B}(\mathcal{H}_1, \mathcal{H}_2)$ is a Banach space, the usual differentiation is defined on $H^\infty(\mathcal{H}_1, \mathcal{H}_2)$ (cf. Def. 1.16). Then, as in the scalar case, it follows from the power series representation (2.63) that $\mathbf{H}$ is analytic (complex differentiable) for all $z \in \mathbb{D}$. For the particular case $\mathcal{H}_1 = \mathcal{H}_2 = \mathbb{C}$, one obtains the usual Hardy space H^∞.

Given a bounded analytic function $\mathbf{H}$ with values in $\mathscr{B}(\mathcal{H}_1, \mathcal{H}_2)$, we will be particularly interested in multiplication operators $\mathbf{O_H} : L^p(\mathcal{H}_1) \to L^p(\mathcal{H}_2)$, with a certain $p \in (1, \infty)$, defined by

$$(\mathbf{O_H}\boldsymbol{f})(\zeta) := \mathbf{H}(\zeta)\, \boldsymbol{f}(\zeta) , \qquad \zeta \in \mathbb{T}$$

and $\mathbf{O}^+_\mathbf{H} : H^p(\mathcal{H}_1) \to H^p(\mathcal{H}_2)$ given by

$$(\mathbf{O}^+_\mathbf{H}\boldsymbol{f})(z) := \mathbf{H}(z)\, \boldsymbol{f}(z) , \qquad z \in \mathbb{D} .$$

The bounded analytic function $\mathbf{H} \in H^\infty(\mathcal{H}_1, \mathcal{H}_2)$ will be called the *symbol* of $\mathbf{O_H}$ and $\mathbf{O}^+_\mathbf{H}$. The norm of these operators is defined as usual by

$$\|\mathbf{O_H}\| = \sup_{\boldsymbol{f}\in L^p(\mathcal{H}_1)} \frac{\|\mathbf{O_H}\boldsymbol{f}\|_{L^p(\mathcal{H}_2)}}{\|\boldsymbol{f}\|_{L^p(\mathcal{H}_1)}} \quad \text{and} \quad \|\mathbf{O}^+_\mathbf{H}\| = \sup_{\boldsymbol{f}\in H^p(\mathcal{H}_1)} \frac{\|\mathbf{O}^+_\mathbf{H}\boldsymbol{f}\|_{H^p(\mathcal{H}_p)}}{\|\boldsymbol{f}\|_{H^p(\mathcal{H}_1)}} .$$

The following proposition will formally prove that the norm of these two operators are given by the norm $\|\mathbf{H}\|_\infty$ of the symbol $\mathbf{H}$.

Proposition 2.34. *Let* $\mathbf{H} \in H^\infty(\mathcal{H}_1, \mathcal{H}_2)$ *be a bounded analytic function. Then for the norms of the multiplication operators* $\mathbf{O_H}$ *and* $\mathbf{O}^+_\mathbf{H}$ *with symbol* $\mathbf{H}$ *it holds that*

$$\|\mathbf{O_H}\| = \operatorname*{ess\,sup}_{\zeta\in\mathbb{T}} \|\mathbf{H}(\zeta)\|_{\mathcal{H}_1\to\mathcal{H}_2} = \sup_{z\in\mathbb{D}} \|\mathbf{H}(z)\|_{\mathcal{H}_1\to\mathcal{H}_2} = \|\mathbf{O}^+_\mathbf{H}\| .$$

Proof. First consider $\mathbf{O_H}$

$$\begin{aligned}
\|\mathbf{O_H}\boldsymbol{f}\|_{L^p(\mathcal{H}_2)} &= \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \|\mathbf{H}(\mathrm{e}^{\mathrm{i}\theta})\boldsymbol{f}(\mathrm{e}^{\mathrm{i}\theta})\|^p_{\mathcal{H}_2}\, \mathrm{d}\theta\right)^{1/p} \\
&\le \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \|\mathbf{H}(\mathrm{e}^{\mathrm{i}\theta})\|^p_{\mathcal{H}_1\to\mathcal{H}_2}\, \|\boldsymbol{f}(\mathrm{e}^{\mathrm{i}\theta})\|^p_{\mathcal{H}_1}\, \mathrm{d}\theta\right)^{1/p} \\
&\le \operatorname*{ess\,sup}_{\zeta\in\mathbb{T}} \|\mathbf{H}(\zeta)\|_{\mathcal{H}_1\to\mathcal{H}_2}\, \|\boldsymbol{f}\|_{L^p(\mathcal{H}_1)}
\end{aligned}$$

which shows that $\|\mathbf{O_H}\| \le \|\mathbf{H}\|_\infty$. To prove the reverse inequality, we choose an arbitrary $\epsilon > 0$ and define the set

$$M(\epsilon) := \left\{\theta \in [-\pi,\pi) \;:\; \|\mathbf{H}(\mathrm{e}^{\mathrm{i}\theta})\|_{\mathcal{H}_1\to\mathcal{H}_2} \ge \|\mathbf{H}\|_\infty - \epsilon/2\right\}$$

and denote by $\chi_{M(\epsilon)}$ the indicator function of $M(\epsilon)$. The Lebesgue measure of $M(\epsilon)$ will be denoted by $\mu_\epsilon := \frac{1}{2\pi}\int_{-\pi}^{\pi}\chi_{M(\epsilon)}(\mathrm{e}^{\mathrm{i}\theta})\,\mathrm{d}\theta$. Moreover, to every $\theta \in M(\epsilon)$ there exists a $\boldsymbol{g}(\mathrm{e}^{\mathrm{i}\theta}) \in \mathcal{H}_1$ such that

$$\|\mathbf{H}(\mathrm{e}^{\mathrm{i}\theta})\,\boldsymbol{g}(\mathrm{e}^{\mathrm{i}\theta})\|_{\mathcal{H}_2} \ge (\|\mathbf{H}\|_\infty - \epsilon)\,\|\boldsymbol{g}(\mathrm{e}^{\mathrm{i}\theta})\|_{\mathcal{H}_1}\,. \tag{2.64}$$

Now, we define the function

$$\boldsymbol{f}_\epsilon(\mathrm{e}^{\mathrm{i}\theta}) := \tfrac{1}{\mu_\epsilon}\,\chi_{M(\epsilon)}(\mathrm{e}^{\mathrm{i}\theta})\,\boldsymbol{g}(\mathrm{e}^{\mathrm{i}\theta})$$

where every $\boldsymbol{g}(\mathrm{e}^{\mathrm{i}\theta})$ is chosen such that (2.64) holds for all $\theta \in M(\epsilon)$. For this function, one obtains

$$\begin{aligned}\|\mathbf{O_H}\,\boldsymbol{f}_\epsilon\|_{L^p(\mathcal{H}_2)} &= \left(\frac{1}{2\pi}\int_{M(\epsilon)} \tfrac{1}{\mu_\epsilon^p}\,\|\mathbf{H}(\mathrm{e}^{\mathrm{i}\theta})\,\boldsymbol{g}(\mathrm{e}^{\mathrm{i}\theta})\|_{\mathcal{H}_2}^p\,\mathrm{d}\theta\right)^{1/p}\\ &\ge (\|\mathbf{H}\|_\infty - \epsilon)\,\frac{1}{\mu_\epsilon}\left(\frac{1}{2\pi}\int_{M(\epsilon)} \|\boldsymbol{g}(\mathrm{e}^{\mathrm{i}\theta})\|_{\mathcal{H}_1}^p\,\mathrm{d}\theta\right)^{1/p}\\ &= (\|\mathbf{H}\|_\infty - \epsilon)\,\|\boldsymbol{f}_\epsilon\|_{L^p(\mathcal{H}_1)}\,.\end{aligned}$$

Since ϵ was chosen arbitrary, this shows that $\|\mathbf{O_H}\| \ge \|\mathbf{H}\|_\infty$ and together with the first part of this proof, one has $\|\mathbf{O_H}\| = \|\mathbf{H}\|_\infty$.

It remains to show that $\|\mathbf{O}^+_\mathbf{H}\| = \|\mathbf{O_H}\|$. Since $H^p(\mathcal{H}_1) \subset L^p(\mathcal{H}_1)$, it is clear that $\|\mathbf{O}^+_\mathbf{H}\| \le \|\mathbf{O_H}\|$. To prove the reverse inequality, we consider polynomials in $L^p(\mathcal{H}_1)$ of the form

$$\boldsymbol{p}(\mathrm{e}^{\mathrm{i}\theta}) = \sum_{k=-N_1}^{N_2} \hat{\boldsymbol{p}}(k)\,\mathrm{e}^{\mathrm{i}k\theta} = \mathrm{e}^{-\mathrm{i}N_1\theta}\sum_{k=0}^{N_1+N_2} \hat{\boldsymbol{p}}(k-N_1)\,\mathrm{e}^{\mathrm{i}k\theta} = \mathrm{e}^{-\mathrm{i}N_1\theta}\,\boldsymbol{p}_c(\mathrm{e}^{\mathrm{i}\theta})$$

with $\hat{\boldsymbol{p}}(k) \in \mathcal{H}_1$ and $N_1, N_2 \ge 0$. However, the polynomial $\boldsymbol{p}_c(\mathrm{e}^{\mathrm{i}\theta})$, obtained from $\boldsymbol{p}(\mathrm{e}^{\mathrm{i}\theta})$ by factoring out $\mathrm{e}^{-\mathrm{i}N_1\theta}$ belongs to $H^p(\mathcal{H}_1)$, and it is easily verified that $\|\mathbf{O}^+_\mathbf{H}\,\boldsymbol{p}_c\|_{H^p(\mathcal{H}_1)} = \|\mathbf{O_H}\,\boldsymbol{p}\|_{L^p(\mathcal{H}_1)}$. Moreover, the polynomials $\mathcal{P}(\mathcal{H}_1)$ of the above form are dense in $L^2(\mathcal{H}_1)$. Therefore, to every $\mathbf{H} \in H^\infty(\mathcal{H}_1,\mathcal{H}_2)$ there exist a polynomials $\boldsymbol{p} \in L^p(\mathcal{H}_1)$ with $\|\boldsymbol{p}\|_{L^p(\mathcal{H}_1)} = 1$ so that

$$\|\mathbf{O}^+_\mathbf{H}\,\boldsymbol{p}_c\|_{H^p(\mathcal{H}_1)} = \|\mathbf{O_H}\,\boldsymbol{p}\|_{L^p(\mathcal{H}_2)} \ge \|\mathbf{O_H}\| - \epsilon\,. \tag{2.65}$$

In turn this implies that

$$\|\mathbf{O}^+_\mathbf{H}\| = \sup_{\substack{\boldsymbol{f}\in H^p(\mathcal{H}_1)\\ \|\boldsymbol{f}\|_{H^p(\mathcal{H}_1)}\le 1}} \|\mathbf{O}^+_\mathbf{H}\boldsymbol{f}\|_{H^p(\mathcal{H}_2)} \ge \|\mathbf{O}^+_\mathbf{H}\,\boldsymbol{p}_c\|_{H^p(\mathcal{H}_1)} \ge \|\mathbf{O_H}\| - \epsilon$$

which shows that $\|\mathbf{O}^+_\mathbf{H}\| \ge \|\mathbf{H}\|_\infty$ and altogether that $\|\mathbf{O}^+_\mathbf{H}\| = \|\mathbf{H}\|_\infty$. □

Since every symbol $\mathbf{H} \in H^\infty(\mathcal{H}_1, \mathcal{H}_2)$ is analytic in $\mathbb{D}$ it is clear that $(\mathbf{O_H} \boldsymbol{f})(z) = \mathbf{H}(z)\, \boldsymbol{f}(z)$ belongs to $H^p(\mathcal{H}_2)$ provided that $\boldsymbol{f} \in H^p(\mathcal{H}_1)$. For this reason and because $H^p(\mathcal{H}_1)$ and $H^p(\mathcal{H}_2)$ are subspaces of $L^p(\mathcal{H}_1)$ and $L^p(\mathcal{H}_2)$, respectively, the operator $\mathbf{O}^+_{\mathbf{H}}$ can be considered as the restriction of $\mathbf{O_H}$ to the subspace $H^p(\mathcal{H}_1)$ of $L^p(\mathcal{H}_1)$.

The most important case is $p = 2$. Then the operators $\mathbf{O_H}$ and $\mathbf{O}^+_{\mathbf{H}}$ are mappings from the Hilbert space $L^2(\mathcal{H}_1)$ into the Hilbert space $L^2(\mathcal{H}_2)$. In this case, it is easily verified that the adjoint of the operator $\mathbf{O_H}$ is given by $\mathbf{O}^*_{\mathbf{H}} = \mathbf{O}_{\mathbf{H}^*}$ where $\mathbf{H}^*(z) = [\mathbf{H}(z)]^*$ is the adjoint of $\mathbf{H}(z) : \mathcal{H}_1 \to \mathcal{H}_2$ for every $z \in \mathbb{D}$.

Example 2.35. Assume that the Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ both have finite dimension. Then, without any loss of generality, we may assume that $\mathcal{H}_1 = \mathbb{C}^N$ and $\mathcal{H}_2 = \mathbb{C}^M$ with the dimensions $N, M \geq 1$ and with the usual Euclidean norm in $\mathbb{C}^N$ and $\mathbb{C}^M$. It is well known that every bounded linear operator $\mathbf{H} \in \mathscr{B}(\mathbb{C}^N, \mathbb{C}^M)$ can be identified with a complex $M \times N$ matrix $\mathbf{H}$ with M rows and N columns. Therefore, $\mathscr{B}(\mathbb{C}^N, \mathbb{C}^M)$ can be identified with the set $\mathbb{C}^{M\times N}$ of all complex $M \times N$ matrices, and the norm of any matrix $\mathbf{H} \in \mathbb{C}^{M\times N}$, induced by the Euclidean vector norm in $\mathbb{C}^N$ and $\mathbb{C}^M$, is known to be

$$\|\mathbf{H}\|_{\mathbb{C}^{M\times N}} = \sup_{\boldsymbol{f}\in\mathbb{C}^N} \frac{\|\mathbf{H}\boldsymbol{f}\|_{\mathbb{C}^M}}{\|\boldsymbol{f}\|_{\mathbb{C}^N}} = \sqrt{\lambda_{\max}\{\mathbf{H}^*\mathbf{H}\}}$$

wherein $\lambda_{\max}(\mathbf{H}^*\mathbf{H})$ is the largest singular value of the matrix $\mathbf{H}$. This norm is also known as the *spectral norm* of $\mathbf{H}$.

Moreover, every $\mathbf{H} \in H^\infty(\mathbb{C}^N, \mathbb{C}^M)$ has the general form (2.63) in which all $\hat{\mathbf{H}}(k) \in \mathbb{C}^{M\times N}$ are complex $M \times N$ matrices. To shorten the notation, the space $H^\infty(\mathbb{C}^N, \mathbb{C}^M)$ of all matrix valued bounded analytic functions will be denoted by $H^\infty(\mathbb{C}^{M\times N})$, and the norm in this space is given by

$$\|\mathbf{H}\|_\infty = \operatorname*{ess\,sup}_{\zeta\in\mathbb{T}} \lambda_{\max}\{\mathbf{H}^*(\zeta)\mathbf{H}(\zeta)\} = \sup_{z\in\mathbb{D}} \lambda_{\max}\{\mathbf{H}^*(z)\mathbf{H}(z)\}\ .$$

Notes

Still the classical reference for trigonometric series is the volume of Zygmund [92]. Theorem 2.8 is due to Jackson [51]. Detailed proofs can also be found in [92, Chap. III] or [61, vol. 1, Chap. IV]. There are numerous text books containing the basic theory of Hardy spaces in different detail and various forms. The scalar case can be found for example in [30, 41, 45, 48, 70]. The exposition here and the given proofs are primarily taken from [70] where also most of the omitted proofs and auxiliary results can be found. The vector valued case is considered in detail in [44, 62, 83]. The notion of inner and outer functions was introduced by Beurling in his seminal paper on shift-invariant subspaces [7]. It seems be worthwhile to consider this original approach to the inner-outer factorization since it gives somewhat more descriptive i.e geometrical proofs of the theorems in Section 2.2.3. We also refer the reader to [43, 44, 45, 57, 62, 83].

3

Banach Algebras

Let $\mathcal{L}_1$ be a linear system with transfer function f_1. An important task in applications is to find the transfer function f_2 of a second linear system $\mathcal{L}_2$ such that the series connection of $\mathcal{L}_1$ and $\mathcal{L}_2$ has a prescribed transfer function f, i.e. such that $f = f_1 f_2$. If, for example, the desired transfer function $f \equiv 1$ is the identity then $\mathcal{L}_2$ is the linear inverse system of $\mathcal{L}_1$.

To investigate such problems, the notion of vector spaces is not sufficient, since in these spaces only the operation of addition is defined, but not multiplication. For these reasons we have to consider space on which also a multiplication operation is defined. Such spaces are known as *algebras*. Additionally, we still require that our spaces are complete normed linear spaces. Spaces with these properties are called Banach algebras.

Definition 3.1 (Banach algebra). *A* complex algebra *is a vector space* $\mathcal{A}$ *over the complex field* $\mathbb{C}$ *in which a multiplication is defined satisfying for all* $f, g, h \in \mathcal{A}$

1) $f\,(g\,h) = (f\,g)\,h$ *(associative law)*

2) $(f+g)\,h = f\,h + g\,h$ *and* $f\,(g+h) = f\,g + f\,h$ *(distributive law)*

and for which

$$\alpha\,(f\,g) = f\,(\alpha\,g) = (\alpha\,f)\,g$$

for every scalar $\alpha \in \mathbb{C}$.

If $\mathcal{A}$ *is a normed linear space which satisfies the* submultiplicative *condition*

$$\|f\,g\|_{\mathcal{A}} \leq \|f\|_{\mathcal{A}}\,\|g\|_{\mathcal{A}} \tag{3.1}$$

for all $f, g \in \mathcal{A}$*, then* $\mathcal{A}$ *is called a* normed complex algebra*. If, in addition,* $\mathcal{A}$ *is complete with respect to its norm* $\|\cdot\|_{\mathcal{A}}$ *and if* $\mathcal{A}$ *contains a* unit element e *such that for all* $f \in \mathcal{A}$

$$fe = ef = f \qquad \textit{and} \qquad \|e\|_{\mathcal{A}} = 1$$

V. Pohl, H. Boche, *Advanced Topics in System and Signal Theory*, Foundations in Signal Processing, Communications and Networking 4,
DOI 10.1007/978-3-642-03639-2_3,

then $\mathcal{A}$ *is called a* Banach algebra.

If the multiplication on $\mathcal{A}$ *is also commutative, i.e. if* $fg = gf$ *for all* $f, g \in \mathcal{A}$*, then* $\mathcal{A}$ *is a* commutative Banach algebra.

Note that the required submultiplicativity (3.1) of the norm makes multiplication a continuous operation on the Banach space $\mathcal{A}$. This continuity implies that if $f_n \to f$ and $g_n \to g$ in $\mathcal{A}$ then $f_n g_n \to fg$ in $\mathcal{A}$, since by (3.1) one has

$$\|f_n g_n - fg\|_{\mathcal{A}} \le \|f_n - f\|_{\mathcal{A}} \, \|g_n\|_{\mathcal{A}} + \|f\|_{\mathcal{A}} \, \|g_n - g\|_{\mathcal{A}} \, .$$

It should be noted that the presence of a unit element is omitted in most of the definitions of a Banach algebra. However, for us, one of the main motivations is the investigation of the inverse of an element from $\mathcal{A}$. Of course, this only makes sense if there exists a unit in the algebra. Moreover, the most important algebras, for our intentions, have a unit and all other algebras can be supplied with a unit in a canonical way (see e.g. [72]).

If not mentioned otherwise, we always assume that the Banach algebras are commutative. Even though this condition is not necessary in some of the following results, this assumption will slightly simplifies some proofs and it is no strong restriction for us since most of the algebras considered here are commutative anyway.

Finally, we give some examples of Banach algebras.

Example 3.2 (L^∞ and H^∞). The space L^∞ of all essentially bounded functions on the unit circle is a Banach algebra if the multiplication of two elements $f, g \in L^\infty$ is defined pointwise by

$$(f\,g)(\zeta) := f(\zeta)\,g(\zeta)\,, \qquad \zeta \in \mathbb{T}\,. \tag{3.2}$$

It is clear that the so defined multiplication satisfies the associative, distributive, and commutative law and one easily verifies that also the submultiplicative condition (3.1) holds for all $f, g \in L^\infty$. The unit element of L^∞ is the constant function $e(\zeta) = 1$ for all $\zeta \in \mathbb{T}$. With obvious modifications, the Hardy space $H^\infty = (L^\infty)_+$ is also a commutative Banach algebra with unit under pointwise multiplication.

Notice that the spaces L^p and H^p with $p < \infty$ are not Banach algebras, in general. In the cases L^2 and H^2 this follows at once from the well known fact that every $f \in L^1$ can be written uniquely as the (pointwise) product of two functions $g, h \in L^2$: $f(\zeta) = g(\zeta)\,h(\zeta)$, $\zeta \in \mathbb{T}$. Thus the product of two L^2 functions belongs no longer to L^2, in general. Therefore the multiplication is not submultiplicative.

Example 3.3 (Continuous functions $\mathcal{C}(\mathbb{T})$ and the disk algebra $A(\mathbb{D})$). Let $\mathcal{C}(\mathbb{T})$ be the set of all continuous functions on the unit circle $\mathbb{T}$ equipped with the supremum norm. Together with the pointwise multiplication (3.2),

$\mathcal{C}(\mathbb{T})$ becomes a commutative Banach algebra in which the constant function $e(\zeta) \equiv 1$ is the unit element. The subspace

$$A(\mathbb{D}) := \mathcal{C}(\mathbb{T})_+ = \{ f \in \mathcal{C}(\mathbb{T}) : \hat{f}(n) = 0 \text{ for all } n < 0 \}$$

together with the same norm and multiplication is also Banach algebra and it is known as the *disk algebra.* It contains all $f \in H^\infty$ which are continuous in the closure $\overline{\mathbb{D}}$ of the unit disk. Equivalently, $A(\mathbb{D})$ is the closure of all polynomials $f(z) = \sum_{k\geq 0} \hat{f}(k)\, z^k$ in the infinity norm.

Example 3.4 (ℓ^1). We consider the space ℓ^1 of absolutely summable series $x = \{x_k\}_{k\in\mathbb{Z}}$ with the norm $\|x\|_{\ell^1} = \sum_{k\in\mathbb{Z}} |x_k|$ (cf. also Example 1.2). On this space, we define the multiplication of two elements $x, y \in \ell^1$ by the convolution

$$z_n = (x\, y)_n := \textstyle\sum_{k=-\infty}^{\infty} x_k\, y_{n-k}\ , \qquad n \in \mathbb{Z}\ .$$

One easily verifies that the usual laws of multiplication are satisfied by this definition. Moreover, since

$$\begin{aligned} \|z\|_{\ell^1} &= \sum_{n=-\infty}^{\infty} |z_n| = \sum_{n=-\infty}^{\infty} \left| \sum_{k=-\infty}^{\infty} x_k\, y_{n-k} \right| \\ &\leq \sum_{n=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} |x_k|\, |y_{n-k}| = \sum_{k=-\infty}^{\infty} |x_k| \sum_{n=-\infty}^{\infty} |y_{n-k}| = \|x\|_{\ell^1}\, \|y\|_{\ell^1} \end{aligned}$$

this multiplication satisfies the submultiplicativity relation (3.1). Therefore l^1 equipped with this multiplication is a commutative Banach algebra. The unit element is obviously $e = \{\ldots, 0, 0, 1, 0, 0, \ldots\}$, the sequence which has only zero entries except for the unit element at position zero.

Example 3.5 (The Wiener algebras $\mathcal{W}$ and $\mathcal{W}_+$). The Wiener algebra $\mathcal{W}$ is the set of all absolutely convergent Fourier series

$$f(\mathrm{e}^{\mathrm{i}\theta}) = \sum_{k=-\infty}^{\infty} \hat{f}(k)\, \mathrm{e}^{\mathrm{i}k\theta}\ , \qquad \theta \in [-\pi, \pi)$$

equipped with the norm

$$\|f\|_{\mathcal{W}} = \sum_{k=-\infty}^{\infty} |\hat{f}(k)| \ < \ \infty$$

and with the pointwise multiplication $(f\, g)(\mathrm{e}^{\mathrm{i}\theta}) = f(\mathrm{e}^{\mathrm{i}\theta})\, g(\mathrm{e}^{\mathrm{i}\theta})$. One easily verifies that $\mathcal{W}$ is isometrically isomorphic to the previous example by identifying each $f \in \mathcal{W}$ with the sequence $\{\cdots, \hat{f}(-1), \hat{f}(0), \hat{f}(1), \cdots\} \in \ell^1$ of its Fourier coefficients. This shows that $\mathcal{W}$ is a commutative Banach algebra with the unit element $e(\mathrm{e}^{\mathrm{i}\theta}) \equiv 1$. As in Examples 3.2 and 3.3 the subspace $\mathcal{W}_+$ of all series of the form $f(\mathrm{e}^{\mathrm{i}\theta}) = \sum_{k=0}^{\infty} \hat{f}(k)\, \mathrm{e}^{\mathrm{i}k\theta}$ together with the same norm $\|\cdot\|_{\mathcal{W}}$ is also a Banach algebra.

Example 3.6 (Linear bounded operators). Let $\mathcal{X}$ be an arbitrary Banach space and let $\mathscr{B}(\mathcal{X})$ be the Banach space of all bounded linear operators on $\mathcal{X}$ with respect to the usual operator norm (1.6). If one defines for all $\mathfrak{T}, \mathfrak{U} \in \mathscr{B}(\mathcal{X})$ a multiplication in the obvious way by $(\mathfrak{T}\mathfrak{U})(x) := \mathfrak{T}(\mathfrak{U}x)$ for all $x \in \mathcal{X}$, $\mathscr{B}(\mathcal{X})$ becomes a Banach algebra. The identity operator $\mathfrak{I}$ is the unit element in this algebra. Note that $\mathscr{B}(\mathcal{X})$ is not commutative, in general.

3.1 The Invertible Elements

Let $\mathcal{A}$ be a commutative Banach algebra. An element $f \in \mathcal{A}$ is called *invertible* in $\mathcal{A}$ if there exists an element $f^{-1} \in \mathcal{A}$ such that

$$f^{-1}f = e = ff^{-1} .$$

If $f \in \mathcal{A}$ is invertible, the element $f^{-1} \in \mathcal{A}$ is called the *inverse* of f and it is easily seen that any $f \in \mathcal{A}$ has at most one inverse in $\mathcal{A}$. The set of all invertible elements of a Banach algebra $\mathcal{A}$ will be denoted by $\mathcal{G}(\mathcal{A})$. If $f, g \in \mathcal{G}(\mathcal{A})$ are two invertible elements of $\mathcal{A}$, then $g^{-1}f^{-1} \in \mathcal{G}(\mathcal{A})$ is the inverse of $fg \in \mathcal{A}$.

Next we want to characterize the set $\mathcal{G}(\mathcal{A})$ of all invertible elements more closely. We start with the observation that the unit element e of $\mathcal{A}$ is invertible in $\mathcal{A}$ with $e^{-1} = e$. The following theorem will show that all elements in $\mathcal{A}$, which are close enough to e, are invertible as well.

Theorem 3.7. *Let $\mathcal{A}$ be a Banach algebra and let $f \in \mathcal{A}$ with $\|f\|_{\mathcal{A}} < 1$, then $e + f \in \mathcal{G}(\mathcal{A})$ and $e - f \in \mathcal{G}(\mathcal{A})$ and it holds*

$$(e+f)^{-1} = \textstyle\sum_{k=0}^{\infty}(-1)^k f^k \qquad \text{and} \qquad (e-f)^{-1} = \textstyle\sum_{k=0}^{\infty} f^k . \tag{3.3}$$

Proof. Consider the elements $g_n \in \mathcal{A}$ given by

$$g_n = \textstyle\sum_{k=0}^{n}(-1)^k f^k = e - f + f^2 - \cdots + (-1)^n f^n , \qquad n \in \mathbb{N} .$$

Since $\|f\|_{\mathcal{A}} < 1$ and by the submultiplicity (3.1) of the norm, it holds that $\|f^n\|_{\mathcal{A}} \le \|f\|_{\mathcal{A}}^n < 1$, which shows that $f^n \to 0$ in $\mathcal{A}$. This implies that g_n is a Cauchy sequence in $\mathcal{A}$. Because $\mathcal{A}$ is complete, there exists an $g \in \mathcal{A}$ such that $g_n \to g$ in $\mathcal{A}$. Moreover, the identity

$$g_n(e+f) = e + (-1)^n f^{n+1} = (e+f)\, g_n$$

and the continuity of the multiplication shows that $g_n(e+f) \to e$ and that $(e+f)g_n \to e$ as $n \to \infty$. This implies the left hand side of (3.3). The proof of the statement for $(e-f)^{-1}$ is completely analogous. □

Thus, all elements in the neighborhood of e are invertible. This result is generalized in the following theorem: If $f \in \mathcal{A}$ is known to be invertible, then all elements in a neighborhood of f are also invertible and the inverse f^{-1} depends continuously on f. Moreover, the mapping $f \mapsto f^{-1}$ is differentiable at any $f \in \mathcal{G}(\mathcal{A})$.

Theorem 3.8. *Let $\mathcal{A}$ be a Banach algebra.*

(a) If $f \in \mathcal{G}(\mathcal{A})$ and $h \in \mathcal{A}$ with $\|h\|_{\mathcal{A}} < \frac{1}{\|f^{-1}\|_{\mathcal{A}}}$ then $f + h \in \mathcal{G}(\mathcal{A})$.
(b) $\mathcal{G}(\mathcal{A})$ is an open subset of $\mathcal{A}$.
(c) The inversion $\mathfrak{I} : f \mapsto f^{-1}$ is a homomorphism of $\mathcal{G}(\mathcal{A})$ onto $\mathcal{G}(\mathcal{A})$ which is differentiable at any $f \in \mathcal{G}(\mathcal{A})$. Its derivative at $f \in \mathcal{G}(\mathcal{A})$ is given by the linear mapping $\mathfrak{I}'[f] : h \mapsto -f^{-1}\, h\, f^{-1}$.

Proof. Since $f+h = f(e+f^{-1}\,h)$, since f is invertible, and since $\|f^{-1}\,h\|_{\mathcal{A}} \leq \|f^{-1}\|_{\mathcal{A}}\,\|h\|_{\mathcal{A}} < 1$, Theorem 3.7 implies that $f+h$ is invertible in $\mathcal{A}$ with

$$\begin{aligned}(f+h)^{-1} &= (e+f^{-1}\,h)^{-1}\,f^{-1}\\ &= \sum_{k=0}^{\infty}(-1)^k(f^{-1}\,h)^k\,f^{-1} = f^{-1} - f^{-1}\,h\,f^{-1} + \cdots\end{aligned}$$

and that $\mathcal{G}(\mathcal{A})$ is an open set. Moreover, the series for $(f+h)^{-1}$ implies

$$\begin{aligned}\|(f+h)^{-1} - f^{-1} + f^{-1}\,h\,f^{-1}\|_{\mathcal{A}} &\leq \sum_{k=2}^{\infty}\|(f^{-1}\,h)^k\|_{\mathcal{A}}\,\|f^{-1}\|_{\mathcal{A}}\\ &\leq \left(\sum_{k=0}^{\infty}\|f^{-1}\|_{\mathcal{A}}^k\,\|h\|_{\mathcal{A}}^k\right)\|f^{-1}\|_{\mathcal{A}}^3\,\|h\|_{\mathcal{A}}^2 \leq \frac{\|f^{-1}\|_{\mathcal{A}}^3}{1-\|f^{-1}\|_{\mathcal{A}}\,\|h\|_{\mathcal{A}}}\,\|h\|_{\mathcal{A}}^2\,. \qquad (3.4)\end{aligned}$$

Dividing both sides by $\|h\|_{\mathcal{A}}$ and let $\|h\|_{\mathcal{A}} \to 0$, the right hand side converges to zero, which implies that the mapping $h \mapsto -f^{-1}\,h\,f^{-1}$ is the derivative of the mapping $f \mapsto f^{-1}$ at f (cf. Def. 1.16). □

3.1.1 Basic properties of spectra

Definition 3.9 (Spectrum). *Let $\mathcal{A}$ be a Banach algebra and let $f \in \mathcal{G}(\mathcal{A})$. The* spectrum *of f is the set $\sigma(f)$ of all $\lambda \in \mathbb{C}$ such that $f - \lambda e$ is* not *invertible and the* spectral radius

$$r_\sigma(f) = \sup_{\lambda\in\sigma(f)}|\lambda|$$

of f is the radius of the smallest closed disc in $\mathbb{C}$ with center at 0, which contains $\sigma(f)$. The complement $\rho(f) = \mathbb{C}\backslash\sigma(f)$ of the spectrum $\sigma(f)$ in $\mathbb{C}$ is the resolvent set *of f.*

Let f be an arbitrary element of the Banach algebra $\mathcal{A}$. If we choose $\lambda \in \mathbb{C}$ satisfying $|\lambda| > \|f\|_{\mathcal{A}}$, then the element $\frac{1}{\lambda}\,f - e$ will be in a small neighborhood of e. Therefore $f - \lambda e = \lambda(\frac{1}{\lambda}\,f - e)$ will be invertible in $\mathcal{A}$ by Theorem 3.7 and λ will not be an element of the spectrum. This indicates that the spectrum of any $f \in \mathcal{A}$ is contained in a bounded set around zero. However, if $f \in \mathcal{G}(\mathcal{A})$ then $\lambda = 0 \notin \sigma(f)$ and therefore it is not obvious whether the spectrum of f is empty. The next theorem proves the non-emptiness and boundedness of the spectrum for every $f \in \mathcal{A}$.

Theorem 3.10. *Let $\mathcal{A}$ be a Banach algebra and $f \in \mathcal{A}$. Then*

(*a*) *the spectrum $\sigma(f)$ is compact and nonempty*
(*b*) *the spectral radius of f is given by*

$$r_\sigma(f) = \lim_{n\to\infty} \|f^n\|_{\mathcal{A}}^{1/n} .$$

The second part of this theorem is also known as the *spectral radius formula.* It belongs to the key results of the whole theory of Banach algebras since the spectral radius formula links the purely algebraic property of invertibility on an algebra to the metric properties of the associated Banach space. It expresses the spectral radius of an element $f \in \mathcal{A}$ in terms of a limit of the norms of powers of f. Theorem 3.11 below will give a consequence of the first part of the previous theorem, i.e. of the nonemptiness of the spectrum of every element of the algebra. It will give a characterization of those Banach algebras in which every nonzero element is invertible.

Proof. Assume $|\lambda| > \|f\|_{\mathcal{A}}$ then $f - \lambda\, e = -\lambda(e - \lambda^{-1} f)$ is invertible in $\mathcal{A}$ by Theorem 3.7 because $\|\lambda^{-1} f\|_{\mathcal{A}} = |\lambda|^{-1}\|f\|_{\mathcal{A}} < 1$. This shows that $\lambda \notin \sigma(f)$ and proves that $\sigma(f)$ is bounded and that $r_\sigma(f) \le \|f\|_{\mathcal{A}}$. It shows also that

$$|\lambda| \le \|f\|_{\mathcal{A}} \qquad \text{if} \qquad \lambda \in \sigma(f) . \tag{3.5}$$

To prove that $\sigma(f)$ is closed, we show that its complement $\rho(f)$ is open. To this end, define the mapping $\Psi : \mathbb{C} \to \mathcal{A}$ by $\Psi(\lambda) = f - \lambda\, e$. Since $\|\Psi(\lambda) - \Psi(\mu)\|_{\mathcal{A}} \le |\lambda - \mu|\, \|e\|_{\mathcal{A}}$, the mapping Ψ is continuous which implies that the pre-image of every open set is open. The resolvent set of f is the pre-image of $\mathcal{G}(\mathcal{A})$: $\rho(f) = \Psi^{-1}(\mathcal{G}(\mathcal{A}))$. Since $\mathcal{G}(\mathcal{A})$ is open (Theorem 3.8), $\rho(f)$ is open as well. Therefore $\sigma(f)$ is closed and consequently compact.

Define the mapping $F : \rho(f) \to \mathcal{A}$ by $F(\lambda) := (f - \lambda\, e)^{-1}$. Then Theorem 3.8 (*c*) implies that F is complex differentiable (analytic). To see this, choose $\lambda, \mu \in \rho(f)$ and set $g := \lambda e - f$ and $g + h := \mu e - f$. Therewith, we have $F(\lambda) = g^{-1}$, $F(\mu) = (g + h)^{-1}$, and $h = (\mu - \lambda)e$. Thereby, λ and μ are chosen such that $\|h\|_{\mathcal{A}} = |\mu - \lambda|\, \|e\|_{\mathcal{A}} < \frac{1}{2}\, \|g^{-1}\|_{\mathcal{A}}^{-1} = \frac{1}{2}\, \|F(\lambda)\|_{\mathcal{A}}^{-1}$, which is always possible, since $\rho(f)$ is open. As in the proof of Theorem 3.8, we consider the series expansion of $F(\mu) = (g + h)^{-1}$. This implies an equation similar to (3.4) wherein f^{-1} has to be replaced by $g^{-1} = F(\lambda)$ and $(f + h)^{-1}$ has to be replaced by $(g + h)^{-1} = F(\mu)$. Dividing the resulting equation by $|\lambda - \mu|$ yields

$$\begin{aligned}\frac{\|F(\mu) - F(\lambda) + F(\lambda)^2(\mu - \lambda)\|_{\mathcal{A}}}{|\mu - \lambda|} &\le \frac{\|F(\lambda)\|_{\mathcal{A}}^3\, \|e\|^2\, |\lambda - \mu|}{1 - \|F(\lambda)\|_{\mathcal{A}}\, \|e\|_{\mathcal{A}}\, |\lambda - \mu|} \\ &\le 2\, \|F(\lambda)\|_{\mathcal{A}}^3\, \|e\|_{\mathcal{A}}^2\, |\lambda - \mu| \, . \end{aligned} \tag{3.6}$$

The right hand side converges to zero as $\mu \to \lambda$ which shows that F is differentiable at every $\lambda \in \rho(f)$ with the derivative $F'(\lambda) = -F(\lambda)^2$ (cf. Def. 1.16).

Let $\varphi \in \mathcal{A}^*$ be an arbitrary bounded linear functional on $\mathcal{A}$ and define

$$\Phi(\lambda) := \varphi[F(\lambda)] = \varphi[(f - \lambda\, e)^{-1}]\,, \qquad \lambda \in \rho(f)\,. \tag{3.7}$$

Then (3.6) and the boundedness and linearity of φ shows that $\Phi(\lambda)$ is an analytic function at every $\lambda \in \rho(f)$ with $\Phi'(\lambda) = -\varphi[(f - \lambda\, e)^{-2}]$. For $|\lambda| > \|f\|_{\mathcal{A}}$ write Φ as $\Phi(\lambda) = \varphi[-(e - \lambda^{-1} f)^{-1}\lambda^{-1}]$. Then Theorem 3.7 implies that the power series

$$\Phi(\lambda) \;=\; -\tfrac{1}{\lambda}\,\phi\left[\textstyle\sum_{k=0}^{\infty} \frac{1}{\lambda^k}\, f^k\right] \;=\; -\textstyle\sum_{k=0}^{\infty} \lambda^{-(k+1)} \varphi(f^k)\,.$$

converges in $\mathcal{A}$. Since Φ is analytic in $\rho(f)$, the Cauchy integral formula implies

$$\varphi(f^k) = \frac{1}{2\pi \mathrm{i}} \int_{\partial B(0,r)} \lambda^k \Phi(\lambda)\,\mathrm{d}\lambda\,, \qquad k = 0, 1, 2, \ldots \tag{3.8}$$

where the integration is taken counter-clockwise along the circle $\partial B(0,r)$ around zero with radius $r > \|f\|_{\mathcal{A}}$. If $\sigma(f)$ were empty then Φ would be analytic in $\mathbb{C}$ and the Cauchy theorem would imply that all integrals in (3.8) are zero, and in particular for $k = 0$ that $\varphi(e) = 0$. However, by the Hahn-Banach theorem there exists a bounded linear functional with $\varphi \in \mathcal{A}^*$ with $\varphi(e) = \|e\|_{\mathcal{A}} \neq 0$. Since φ was arbitrary, this is a contradiction which shows that the spectrum $\sigma(f)$ is not empty.

It remains to prove the spectral radius formula (*b*). We consider again the function Φ defined in (3.7). As it was shown, this function is analytic for every $\lambda \in \rho(f)$, which implies that

$$M(r) := \sup_{\theta \in [-\pi, \pi)} |\Phi(r\mathrm{e}^{\mathrm{i}\theta})| < \infty$$

for each $r > r_\sigma(f)$. Therewith (3.8) gives

$$\left|\varphi(f^k)\right| = \left| \frac{r^{k+1}}{2\pi} \int_{-\pi}^{\pi} \Phi(r\mathrm{e}^{\mathrm{i}\theta})\, \mathrm{e}^{\mathrm{i}(k+1)\theta}\,\mathrm{d}\theta \right| \;\leq\; r^{k+1} M(r)\,. \tag{3.9}$$

Now, every $f \in \mathcal{A}$ defines a linear functional on $\mathcal{A}^*$ by $\Psi_f(\varphi) := \varphi(f)$, $\varphi \in \mathcal{A}^*$. It is a consequence of the Hahn-Banach theorem (see e.g. [54]) that the norm of f is the same as the norm of Ψ_f, i.e.

$$\|f\|_{\mathcal{A}} = \|\Psi_f\|_{\mathcal{A}^*} = \sup_{\varphi \in \mathcal{A}^*,\ \|\varphi\|_{\mathcal{A}^*} \leq 1} |\Psi_f(\varphi)|\,. \tag{3.10}$$

Then (3.9) shows that $|\Psi_{f^k}(\varphi)| = |\varphi(f^k)| \leq r^{k+1} M(r)$ for every $\varphi \in \mathcal{A}^*$, and therefore (3.10) gives $\|f^k\|_{\mathcal{A}} \leq r^{k+1} M(r)$ for $k = 1, 2, 3, \ldots$ and for $r > r_\sigma(f)$. Consequently, we obtain

$$\limsup_{k \to \infty} \|f^k\|_{\mathcal{A}}^{1/k} \leq r_\sigma(f)\,. \tag{3.11}$$

Conversely, let $\lambda \in \sigma(f)$. Then $f - \lambda\, e$ is not invertible and the relation

$$f^k - \lambda^k e = (f - \lambda e)(f^{k-1} + \lambda f^{k-2} + \lambda^2 f^{k-3} + \cdots + \lambda^{k-1} e)$$

shows that also $f^k - \lambda^k e$ is not invertible such that $\lambda^k \in \sigma(f^k)$ which then implies by (3.5) that $|\lambda^k| \leq \|f^k\|$ for all $k = 1, 2, 3, \ldots$. Since $\lambda \in \sigma(f)$, one has $|\lambda| \leq r_\sigma(f)$ such that the previous inequality implies

$$r_\sigma(f) \leq \inf_{k \geq 1} \|f^k\|^{1/k} \tag{3.12}$$

The spectral radius formula (b) follows from (3.11) and (3.12). □

An algebra $\mathcal{A}$ is called a *division algebra* if every nonzero element of $\mathcal{A}$ is invertible. The following result will show that the complex field $\mathbb{C}$ is essentially the only division algebra. This observation is an easy consequence of the fact that the spectrum $\sigma(f)$ of every $f \in \mathcal{A}$ is always nonempty.

Theorem 3.11 (Gelfand-Mazur). *Let $\mathcal{A}$ be a Banach algebra in which every non-zero element is invertible, then $\mathcal{A}$ is isomorphic to $\mathbb{C}$.*

Proof. By Theorem 3.10 (a), for every $f \in \mathcal{A}$ there exixts a $\lambda \in \mathbb{C}$ such that $f - \lambda e$ is not invertible in $\mathcal{A}$. By the assumption of the theorem, zero is the only noninvertible element of $\mathcal{A}$. So $f = \lambda e$ and $\|f\|_\mathcal{A} = |\lambda|\, \|e\|_\mathcal{A}$. This shows that $f \mapsto \lambda$ is a bijective mapping from $\mathcal{A}$ onto $\mathbb{C}$. Moreover, since obviously $f + g \mapsto \lambda(f) + \lambda(g)$ and $f\, g \mapsto \lambda(f)\, \lambda(g)$, the mapping $f \mapsto \lambda(f)$ defines an algebra-isomorphism of $\mathcal{A}$ onto $\mathbb{C}$. □

Note that the isomorphism between $\mathcal{A}$ and $\mathbb{C}$ in the previous theorem is even an isometry if $\mathcal{A}$ has the additional property that $\|e\|_\mathcal{A} = 1$.

3.1.2 Exponential and logarithm on Banach algebras

Let $\mathcal{A}$ be a Banach algebra and $f \in \mathcal{A}$. Then $f^k \in \mathcal{A}$ for every $k \in \mathbb{N}$ and since the multiplication is continuous in $\mathcal{A}$, one has $\|f^k\|_\mathcal{A} \leq \|f\|_\mathcal{A}^k$. Because of these properties of Banach algebras, the usual functions known from analysis (e.g. exp, log, cos, sin), which are defined by means of a power series of real or complex numbers, can be defined also on every Banach algebra $\mathcal{A}$ as the power series of an $f \in \mathcal{A}$. This power series will converge for every $f \in \mathcal{A}$ as long as the common series of $\|f\|_\mathcal{A}$ converges in $\mathbb{R}$. Here, we only consider shortly the exponential and the logarithm functions, since they will be needed frequently in later chapters.

Let $f \in \mathcal{A}$ be an arbitrary element of a Banach algebra $\mathcal{A}$. We define the exponential of f by

$$\mathrm{e}^f = \exp(f) := \textstyle\sum_{k=0}^\infty \frac{1}{k!}\, f^k\ , \qquad f \in \mathcal{A}\ . \tag{3.13}$$

Since $\|\exp(f)\|_\mathcal{A} \leq \sum_{k=0}^\infty \frac{1}{k!} \|f\|_\mathcal{A}^k$ this exponential series converges for every $f \in \mathcal{A}$ by the convergence of the usual exponential series and one has

$$\|\exp f\|_{\mathcal{A}} \le \exp \|f\|_{\mathcal{A}} \qquad \text{for every } f \in \mathcal{A}\,. \tag{3.14}$$

It is not hard to verify that all the well known properties of the scalar exponential function like the relation $\exp(x+y) = \exp(x)\,\exp(y)$ are still true for the exponential function on Banach algebras. Also, similar to the scalar case, the exponential function is continuous at every point $f \in \mathcal{A}$. This is shown in the following lemma which will be needed in later chapters

Lemma 3.12. *Let $\mathcal{A}$ be a Banach algebra, let $f \in \mathcal{A}$ and $h \in \mathcal{A}$ with $\|h\|_{\mathcal{A}} < 1$, then*

$$\|\exp(f+h) - \exp(f)\|_{\mathcal{A}} \le 2\,\|\exp f\|_{\mathcal{A}}\,\|h\|_{\mathcal{A}}\,.$$

Proof. Since $\exp(f+h) - \exp(f) = \exp(f)[\exp(h) - e]$ the power series (3.13) gives

$$\begin{aligned}\|\exp(f+h) - \exp(f)\|_{\mathcal{A}} &\le \|\exp(f)\|_{\mathcal{A}}\,\|\exp(h) - e\|_{\mathcal{A}} \\ &\le \|\exp(f)\|_{\mathcal{A}}\,\|h\|_{\mathcal{A}} \sum_{k=0}^{\infty} \frac{\|h\|_{\mathcal{A}}^k}{(k+1)!} \le \|\exp(f)\|_{\mathcal{A}}\,\|h\|_{\mathcal{A}} \sum_{k=0}^{\infty} \left(\tfrac{\|h\|_{\mathcal{A}}}{2}\right)^k \\ &\le 2\,\|\exp(f)\|_{\mathcal{A}}\,\|h\|_{\mathcal{A}} \end{aligned} \tag{3.15}$$

for all $f \in \mathcal{A}$ and all $h \in \mathcal{A}$ with $\|h\|_{\mathcal{A}} < 1$. □

We define the set

$$\exp(\mathcal{A}) := \{\,\exp(f) \;:\; f \in \mathcal{A}\,\}\,.$$

Every element of $\exp(\mathcal{A})$ is invertible in $\mathcal{A}$, i.e. $\exp(\mathcal{A}) \subset \mathcal{G}(\mathcal{A})$. Indeed, if $f = \exp(g)$ arbitrary with some $g \in \mathcal{A}$, then $h = \exp(-g)$ is also an element of $\mathcal{A}$ and $f\,h = \exp(g)\,\exp(-g) = \exp(0) = e$.

Let $g \in \mathcal{A}$. If $g = \exp(f)$ for some $f \in \mathcal{A}$, then f is said to be a *logarithm* of g. From the definition of the set $\exp(\mathcal{A})$ it is clear that each $g \in \exp(\mathcal{A})$ possesses a logarithm in $\mathcal{A}$. As in the scalar case, the logarithm f of g is not unique, in general, i.e. for an $g \in \mathcal{A}$ there may exist several different $f \in \mathcal{A}$ such that $g = \exp(f)$. However, as in the scalar case, the logarithm on $\mathcal{A}$ can be defined by means of the power series

$$\log(e - f) = -\textstyle\sum_{k=1}^{\infty} \frac{1}{k}\, f^k\,, \qquad f \in \mathcal{A},\ \|f\|_{\mathcal{A}} < 1\,. \tag{3.16}$$

The convergence of this series follows from the convergence of the scalar series for $\|f\|_{\mathcal{A}}$, and one verifies the common properties of the logarithm similar as in the scalar case. In particular, one verifies that $e - g = \exp[\log(e-g)]$. We will frequently need the following continuity result for the logarithm.

Lemma 3.13. *Let $\mathcal{A}$ be a Banach algebra, let $f \in \exp(\mathcal{A})$ and $h \in \mathcal{A}$ with $\|h\|_{\mathcal{A}} < \frac{1}{2}\|f^{-1}\|_{\mathcal{A}}^{-1}$. Then $\log(f-h) \in \exp(\mathcal{A})$ and*

$$\|\log(f-h) - \log(f)\|_{\mathcal{A}} \le 2\,\|f^{-1}\|_{\mathcal{A}}\,\|h\|_{\mathcal{A}}\,.$$

Proof. Of course, $\log(f-h) = \log f + \log(e - f^{-1}h)$. Both terms on the right hand side exist since $f \in \exp(\mathcal{A})$ and $\|f^{-1}h\|_{\mathcal{A}} < 1$. Using the power series (3.16), one obtains

$$\begin{aligned}\|\log(f-h) - \log(f)\|_{\mathcal{A}} &\le \|f^{-1}h\|_{\mathcal{A}} \sum_{k=0}^{\infty} \frac{\|f^{-1}h\|_{\mathcal{A}}^k}{k+1} \\ &\le \|f^{-1}h\|_{\mathcal{A}} \frac{1}{1-\|f^{-1}h\|_{\mathcal{A}}} \le 2\,\|f^{-1}\|_{\mathcal{A}}\|h\|_{\mathcal{A}}\end{aligned}$$

using that $\|f^{-1}h\|_{\mathcal{A}} \le \|f^{-1}\|_{\mathcal{A}}\|h\|_{\mathcal{A}} < 1/2$. □

Note that the previous lemma shows in particular that $\exp(\mathcal{A})$ is an open set.

3.2 Complex Homomorphisms and Ideals

Linear functionals which preserve the multiplication operation of the algebra play a very important roll in the theory of Banach algebras. Linear functionals with this property are known as homomorphisms.

Definition 3.14 (Complex homomorphisms). *Let $\mathcal{A}$ be a complex algebra and let $\gamma : \mathcal{A} \to \mathbb{C}$ be a linear functional on $\mathcal{A}$ which is not identical to zero. Then γ is called a* complex homomorphism *on $\mathcal{A}$ if it is a* multiplicative complex linear functional, *i.e. if it is linear and satisfies*

$$\gamma(fg) = \gamma(f)\,\gamma(g) \qquad \text{for all } f, g \in \mathcal{A}\,.$$

The set of all complex homomorphisms on $\mathcal{A}$ will be denoted by $\Gamma(\mathcal{A})$.

It should be noted that the above definition does not require that the multiplicative linear functionals on $\mathcal{A}$ are continuous. The next theorem gives basic properties of homomorphisms on Banach algebras, in particular it will show that every multiplicative linear functional on an algebra $\mathcal{A}$ is continuous.

Theorem 3.15. *Let $\mathcal{A}$ be a Banach algebra and let $\gamma \in \Gamma(\mathcal{A})$ be an arbitrary complex homomorphism on $\mathcal{A}$. Then*

$$\begin{aligned}&(a) \quad \gamma(e) = 1 \\ &(b) \quad \gamma(f) \neq 0 \qquad \text{for all } f \in \mathcal{G}(\mathcal{A}) \\ &(c) \quad \|\gamma\| = \sup_{f\in\mathcal{A},\ \|f\|_{\mathcal{A}}\le 1} |\gamma(f)| = 1\,.\end{aligned}$$

For part (a) and (b) of this theorem, $\mathcal{A}$ needs only to be a complex algebra, but not a normed algebra. Statement (b) shows that for every invertible element $f \in \mathcal{A}$, $\gamma(f)$ is non-zero for all complex homomorphisms on $\mathcal{A}$. We will see later that this property actually uniquely characterizes all invertible elements of $\mathcal{A}$. Statement (c) implies that all complex homomorphisms of a Banach algebra are bounded and therefore continuous.

Proof. Let $g \in \mathcal{A}$ with $\gamma(g) \neq 0$. Then

$$\gamma(g) = \gamma(ge) = \gamma(g)\gamma(e) ,$$

which proves (a). Assume now that $f \in \mathcal{G}(\mathcal{A})$. Then (a) gives the identity

$$1 = \gamma(e) = \gamma(ff^{-1}) = \gamma(f)\gamma(f^{-1})$$

so that $\gamma(f) \neq 0$.

Now, let $f \in \mathcal{A}$ with $\|f\|_{\mathcal{A}} \leq 1$ arbitrary and let $\lambda \in \mathbb{C}$ with $|\lambda| > 1$. Then $\|\lambda^{-1} f\|_{\mathcal{A}} < 1$ and Theorem 3.7 shows that $e - \lambda^{-1} f$ is invertible in $\mathcal{A}$ such that parts (b) and (a) give

$$\gamma(e - \lambda^{-1} f) = 1 - \lambda^{-1}\gamma(f) \neq 0 .$$

Hence $\gamma(f) \neq \lambda$ for $|\lambda| > 1$. This shows that $|\gamma(f)| \leq 1$ where equality holds for $f = e$. This finally gives statement (c). □

Closely related to the complex homomorphisms of a Banach algebra $\mathcal{A}$ are the maximal ideals of $\mathcal{A}$.

Definition 3.16 (Ideals). *A subset $\mathcal{I}$ of a commutative complex algebra $\mathcal{A}$ is said to be an* ideal *if*

(a) *$\mathcal{I}$ is a subspace of $\mathcal{A}$*
(b) *$f\,g \in \mathcal{I}$ whenever $f \in \mathcal{I}$ and $g \in \mathcal{A}$.*

If $\mathcal{I} \neq \mathcal{A}$ then $\mathcal{I}$ is said to be a proper ideal*. A proper ideal which is not contained in any larger proper ideal is called a* maximal ideal*.*

Without proof, we give some basic properties of ideals of a commutative Banach algebra $\mathcal{A}$.

Proposition 3.17. *Let $\mathcal{A}$ be a commutative Banach algebra.*

(a) *A proper ideal $\mathcal{I}$ of $\mathcal{A}$ does not contain any invertible element of $\mathcal{A}$.*
(b) *Let $\mathcal{I}$ be an ideal of $\mathcal{A}$, then the closure $\overline{\mathcal{I}}$ is also an ideal of $\mathcal{A}$.*
(c) *Every proper ideal of $\mathcal{A}$ is contained in a maximal ideal of $\mathcal{A}$.*
(d) *Every maximal ideal of $\mathcal{A}$ is closed.*

The verification of (a) and (b) is almost trivial, whereas (c) follows immediately from the Hausdorf maximality principle. To verify (d) one uses additionally that $\mathcal{G}(\mathcal{A})$ is open together with (a).

To show the close relation between the homomorphisms on a Banach algebra $\mathcal{A}$ and the maximal ideals on $\mathcal{A}$, we need to recall shortly the notion of quotient spaces and quotient algebras. To this end, assume at the beginning that $\mathcal{A}$ is an arbitrary vector space over $\mathbb{C}$ and $\mathcal{M}$ is a subspace of $\mathcal{A}$. Then with every $f \in \mathcal{A}$ one associates the *coset*

$$\chi(f) := f + \mathcal{M} = \{f + g : g \in \mathcal{M}\} .$$

This definition implies that $\chi(f) = \chi(g)$ whenever $f - g \in \mathcal{M}$ and that $\chi(f) \cap \chi(g) = \emptyset$ whenever $f - g \notin \mathcal{M}$. The set of all cosets of $\mathcal{M}$ is usually denoted by $\mathcal{A}/\mathcal{M}$. If one defines the addition and the scalar multiplication on $\mathcal{A}/\mathcal{M}$, respectively by

$$\chi(f) + \chi(g) = \chi(f+g) \qquad \text{and} \qquad \alpha\,\chi(f) = \chi(\alpha\, f)$$

for all $f, g \in \mathcal{A}$ and $\alpha \in \mathbb{C}$, then the set $\mathcal{A}/\mathcal{M}$ becomes a vector space with the zero element $\chi(0) = \mathcal{M}$. If $\mathcal{A}$ is even a normed space and if $\mathcal{M}$ is a *closed* subspace of $\mathcal{A}$, then one defines the so called *quotient norm* on $\mathcal{A}/\mathcal{M}$

$$\|\chi(f)\| := \inf_{g \in \mathcal{M}} \|f + g\|_{\mathcal{A}} \,. \tag{3.17}$$

With this norm, $\mathcal{A}/\mathcal{M}$ becomes a normed linear space and if $\mathcal{A}$ is a Banach space, so will be $\mathcal{A}/\mathcal{M}$:

Lemma 3.18. *Let $\mathcal{A}$ be a commutative Banach algebra and let $\mathcal{M}$ be a closed subspace of $\mathcal{A}$. Then $\mathcal{A}/\mathcal{M}$ is a Banach space.*

Proof. We first show that $\mathcal{A}/\mathcal{M}$ with the norm (3.17) is indeed a normed linear space . Let $f \in \mathcal{M}$ (i.e. f belongs to the zero element $\chi(0)$ of $\mathcal{A}/\mathcal{M}$), then $\|\chi(f)\| = 0$ because $-f \in \mathcal{M}$. Conversely, let $f \notin \mathcal{M}$. Since $\mathcal{M}$ is closed $\mathcal{A}\backslash\mathcal{M}$ is open and there exists an $\epsilon > 0$ and neighborhood $B_\epsilon(f) := \{g \in \mathcal{A} : \|f - g\|_{\mathcal{A}} < \epsilon\}$ of f with $B_\epsilon(f) \cap \mathcal{M} = \emptyset$, which implies that $\|\chi(f)\| \geq \epsilon > 0$. To verify the triangle inequality, let $f_1, f_2 \in \mathcal{A}$, and let $\epsilon > 0$ arbitrary. Then there exist $h_1, h_2 \in \mathcal{M}$ such that

$$\|f_1 + h_1\|_{\mathcal{A}} \leq \|\chi(f_1)\| + \epsilon/2 \quad \text{and} \quad \|f_2 + h_2\|_{\mathcal{A}} \leq \|\chi(f_2)\| + \epsilon/2 \,. \tag{3.18}$$

Therewith, one gets

$$\begin{aligned} \|\chi(f_1 + f_2)\| &= \inf_{h \in \mathcal{M}} \|f_1 + f_2 + h\|_{\mathcal{A}} \leq \|f_1 + h_1 + f_2 + h_2\|_{\mathcal{A}} \\ &\leq \|f_1 + h_1\|_{\mathcal{A}} + \|f_2 + h_2\|_{\mathcal{A}} \leq \|\chi(f_1)\| + \|\chi(f_2)\| + \epsilon \,, \end{aligned}$$

which proves the triangle inequality since ϵ was arbitrary.

Let $\chi(f_n)$ be a Cauchy sequence in $\mathcal{A}/\mathcal{M}$. Then to every $\epsilon > 0$ there exists an $N_0 \in \mathbb{N}$ such that

$$\|\chi(f_n) - \chi(f_m)\| = \inf_{h \in \mathcal{M}} \|f_n - f_m + h\|_{\mathcal{A}} < \epsilon \qquad \text{for all } n, m \geq N_0$$

which shows that there exist $g_n, g_m \in \mathcal{A}$ with $f_n - g_n \in \mathcal{M}$ and $f_m - g_m \in \mathcal{M}$ such that $\|g_n - g_m\|_{\mathcal{A}} < 2\epsilon$ for all $n, m \geq N_0$. Thus, $\{g_n\}$ is a Cauchy sequence in the Banach space $\mathcal{A}$ which converges to a $g \in \mathcal{A}$. It follows that

$$\|\chi(g_n) - \chi(g)\| = \inf_{h \in \mathcal{M}} \|g_n - g + h\|_{\mathcal{A}} \leq \|g_n - g\|_{\mathcal{A}} \leq 2\epsilon$$

which implies that $\chi(g_n)$ converges to $\chi(g)$ in $\mathcal{A}/\mathcal{M}$. Consequently $\mathcal{A}/\mathcal{M}$ is complete. □

Assume next that $\mathcal{A}$ is even a commutative Banach algebra and that $\mathcal{M}$ is a closed ideal of $\mathcal{A}$. Then one can define a multiplication on $\mathcal{A}/\mathcal{M}$ by

$$\chi(f)\,\chi(g) = \chi(f\,g) \qquad \text{for all } f, g \in \mathcal{A}\,. \tag{3.19}$$

In this way also $\mathcal{A}/\mathcal{M}$ becomes a commutative algebra with unit element $\chi(e)$. This is proved in the following lemma.

Lemma 3.19. *Let $\mathcal{A}$ be a commutative Banach algebra. If $\mathcal{M}$ is a proper closed ideal then $\mathcal{A}/\mathcal{M}$ is a commutative Banach algebra.*

Proof. First it is shown that the multiplication (3.19) is well defined. Assume that $\chi(f_1) = \chi(f_2)$ and $\chi(g_1) = \chi(g_2)$ then there exist $h_1, h_2 \in \mathcal{M}$ such that $f_1 = f_2 + h_1$ and $g_1 = g_2 + h_2$. Therewith

$$f_1\,g_1 = f_2\,g_2 + h_1\,g_2 + f_2\,h_2 + h_1\,h_2 = f_2\,g_2 + h$$

with $h = h_1\,g_2 + f_2\,h_2 + h_1\,h_2$. Since $h_1, h_2 \in \mathcal{M}$ and because $\mathcal{M}$ is an ideal, it follows that $h \in \mathcal{M}$ which shows that $\chi(f_1\,g_1) = \chi(f_2\,g_2)$.

Lemma 3.18 implies that $\mathcal{A}/\mathcal{M}$ is a Banach space, and it remains to show that the multiplication (3.19) on $\mathcal{A}/\mathcal{M}$ satisfies the submultiplicity relation (3.1). To this end, let $f_1, f_2 \in \mathcal{A}$ and $\epsilon > 0$ arbitrary. As in the proof of Lemma 3.18, there exist $h_1, h_2 \in \mathcal{M}$ such that (3.18) holds. Note that $(f_1 + h_1)(f_2 + h_2) = f_1\,f_2 + h_0$ with $h_0 := h_1\,f_2 + f_1\,h_2 + h_1\,h_2 \in \mathcal{M}$, since $\mathcal{M}$ is an ideal. Then

$$\begin{aligned}
\|\chi(f_1)\,\chi(f_1)\| = \|\chi(f_1\,f_2)\| &= \inf_{h\in\mathcal{M}} \|f_1\,f_2 + h\|_{\mathcal{A}} \le \|f_1\,f_2 + h_0\|_{\mathcal{A}} \\
&= \|(f_1 + h_1)(f_2 + h_2)\|_{\mathcal{A}} \le \|f_1 + h_1\|_{\mathcal{A}}\,\|f_2 + h_2\|_{\mathcal{A}} \\
&\le (\,\|\chi(f_1)\| + \tfrac{\epsilon}{2}\,)(\,\|\chi(f_2)\| + \tfrac{\epsilon}{2}\,) \\
&= \|\chi(f_1)\|\,\|\chi(f_2)\| + \tfrac{\epsilon}{2}(\,\|\chi(f_1)\| + \|\chi(f_2)\| + \tfrac{\epsilon}{2}\,)\,.
\end{aligned}$$

For $\epsilon \to 0$ the right hand side converges to $\|\chi(f_1)\|\,\|\chi(f_2)\|$, but this is what we had to prove. Finally, from the definition (3.19) follows that $\chi(f)\chi(e) = \chi(fe) = \chi(f)$ for every $f \in \mathcal{A}$, which shows that $\chi(e)$ is the multiplicative unity in $\mathcal{A}/\mathcal{M}$. □

It should be noted that one necessarily needs that $\mathcal{M}$ is closed in order for χ to be a homomorphism of $\mathcal{A}$ onto $\mathcal{A}/\mathcal{M}$.

The next theorem reveals the close relation between the maximal ideals and the homomorphisms of a commutative Banach algebra.

Theorem 3.20. *Let $\mathcal{A}$ be a commutative Banach algebra and let $\Gamma(\mathcal{A})$ be the set of all complex homomorphisms on $\mathcal{A}$. A subset $\mathcal{M}$ of $\mathcal{A}$ is a maximal ideal of $\mathcal{A}$ if and only if there exists a $\gamma \in \Gamma(\mathcal{A})$ such that $\mathcal{M}$ is the null space of γ, i.e. $\mathcal{M} = \mathscr{N}(\gamma)$.*

Proof. If $\mathcal{M}$ is a maximal ideal of $\mathcal{A}$, then $\mathcal{M}$ is closed (Proposition 3.17) and $\mathcal{A}/\mathcal{M}$ is a Banach algebra (Lemma 3.19). Next, it is shown that $\mathcal{A}/\mathcal{M}$ is even a division algebra (cf. Theorem 3.11). To this end, choose $f \in \mathcal{A}$ but with $f \notin \mathcal{M}$ and consider the set $\mathcal{K} = \{g\,f + h : g \in \mathcal{A}, h \in \mathcal{M}\}$. It is easily seen that $\mathcal{K}$ is an ideal of $\mathcal{A}$ which contains $\mathcal{M}$ ($\mathcal{M}$ is the subset of $\mathcal{K}$, obtained by setting $g = 0$). Since $\mathcal{M}$ is assumed to be maximal, this shows that $\mathcal{K} = \mathcal{A}$. Consequently there exists a certain $g \in \mathcal{A}$ and $h \in \mathcal{M}$ such that $g\,f + h = e$. If $\chi : \mathcal{A} \to \mathcal{A}/\mathcal{M}$ denotes again the quotient map, then the last equation gives $\chi(g)\,\chi(f) = \chi(e)$. Since $\chi(e)$ is the unit element of $\mathcal{A}/\mathcal{M}$ this shows that every nonzero element $f \in \mathcal{A}$ is invertible in $\mathcal{A}/\mathcal{M}$ and by Theorem 3.11 there exists an isomorphism $k : \mathcal{A}/\mathcal{M} \to \mathbb{C}$. Finally, define the linear functional $\gamma : \mathcal{A} \to \mathbb{C}$ by $\gamma(f) = k(\chi(f))$. It is easily seen that $\gamma \in \Gamma(\mathcal{A})$ and that $\mathcal{M} = \mathscr{N}(\gamma) = \{f \in \mathcal{A} : k(\chi(f)) = 0\}$.

Conversely, let $\gamma_1 \in \Gamma(\mathcal{A})$ be an arbitrary homomorphism on $\mathcal{A}$ and let $\mathcal{M} = \mathscr{N}(\gamma_1) = \{f \in \mathcal{A} : \gamma_1(f) = 0\}$ be its null space. By Theorem 1.15, $\mathcal{M}$ is a closed subspace of $\mathcal{A}$. Since $\gamma_1(f\,g) = \gamma_1(f)\,\gamma_1(g) = 0$ for all $g \in \mathcal{A}$, $f\,g \in \mathcal{M}$ which shows that $\mathcal{M}$ is a closed ideal. It is proper because $\mathcal{M}$ contains no invertible element of $\mathcal{A}$, by Theorem 3.15–(b). □

As an important application of the previous theorem, we have the following characterization of all invertible elements of a Banach algebra.

Theorem 3.21. *Let $\mathcal{A}$ be a commutative Banach algebra, and let $\Gamma(\mathcal{A})$ be the set of all complex homomorphisms of $\mathcal{A}$. An element $f \in \mathcal{A}$ is invertible in $\mathcal{A}$ if and only if $\gamma(f) \neq 0$ for all $\gamma \in \Gamma(\mathcal{A})$.*

Proof. If $f \in \mathcal{A}$ is invertible, Theorem 3.15–(b) implies that $\gamma(f) \neq 0$ for every $\gamma \in \Gamma(\mathcal{A})$. Conversely, assume that f is not invertible and define the set $\mathcal{K} = \{f\,g : g \in \mathcal{A}\}$. This $\mathcal{K}$ is a proper ideal since it contains no invertible element. Therefore $\mathcal{K}$ is contained in a maximal ideal $\mathcal{M}$ (Proposition 3.17) and Theorem 3.20 shows that there exists a $\gamma \in \Gamma(\mathcal{A})$ such that $\gamma(g) = 0$ for all $g \in \mathcal{M}$, and in particular $\gamma(f) = 0$. □

Definition 3.22 (Gelfand Transform). *Let $\mathcal{A}$ be a commutative Banach algebra and $\Gamma(\mathcal{A})$ the set of all complex homomorphisms of $\mathcal{A}$. Then to every $f \in \mathcal{A}$ one assigns a function $\check{f} : \Gamma(\mathcal{A}) \to \mathbb{C}$ by the formula*

$$\check{f}(\gamma) := \gamma(f)\,, \qquad \gamma \in \Gamma(\mathcal{A})\,.$$

The so defined function $\check{f}$ is called the Gelfand transform *of f and the set of all Gelfand transforms $\check{f}$ of elements $f \in \mathcal{A}$ will be denoted by $\check{\mathcal{A}}$.*

Thus the Gelfand transform of an $f \in \mathcal{A}$ is defined on the set $\Gamma(\mathcal{A})$ of all homomorphisms of $\mathcal{A}$ and has values in the complex field $\mathbb{C}$. Sometimes, the term "Gelfand transform" is also applied to the mapping $f \mapsto \check{f}$.

3.3 Involutions

Definition 3.23 (Involution). *A mapping $f \mapsto f^*$ of a complex algebra $\mathcal{A}$ into $\mathcal{A}$ which satisfies the following four properties, for all $f, g \in \mathcal{A}$, and $\lambda \in \mathbb{C}$*

(*a*) $(f+g)^* = f^* + g^*$

(*b*) $(\lambda f)^* = \overline{\lambda} f^*$

(*c*) $(fg)^* = g^* f^*$

(*d*) $f^{**} = f$.

is called an involution *on $\mathcal{A}$.*

Example 3.24. Consider the algebras L^∞, $\mathcal{C}(\mathbb{T})$, or $\mathcal{W}$ of functions on $\mathbb{T}$, introduces in the Examples 3.2-3.5. Then the operation $f \mapsto \overline{f}$ which is the pointwise complex conjugate, i.e. $\overline{f}(\zeta) = \overline{f(\zeta)}$ for all $\zeta \in \mathbb{T}$, defines an involution on these algebras.

The previous example of an involution, is the involution we will be most concerned with in later sections. However, there are other examples of involutions:

Example 3.25. Let $\mathcal{H}$ be a Hilbert space with scalar product $\langle\cdot,\cdot\rangle_{\mathcal{H}}$ and denote by $\mathscr{B}(\mathcal{H})$ the Banach algebra of all bounded linear operators on $\mathcal{H}$. Then, it can be shown that to every $\mathfrak{T} \in \mathscr{B}(\mathcal{H})$ there exists a unique $\mathfrak{T}^* \in \mathscr{B}(\mathcal{H})$, the *adjoint* of $\mathfrak{T}$, such that

$$\langle \mathfrak{T}\, x, y\rangle_{\mathcal{H}} = \langle x, \mathfrak{T}^*\, y\rangle_{\mathcal{H}} \qquad \text{for all } x, y \in \mathcal{H}\ .$$

The operation $\mathfrak{T} \mapsto \mathfrak{T}^*$ is an involution on $\mathscr{B}(\mathcal{H})$.

For us, it will be important that the involution $f \mapsto f^*$ is a continuous mapping. The following theorem gives sufficient conditions on the algebra $\mathcal{A}$ such that every involution on $\mathcal{A}$ is continuous.

Theorem 3.26. *Let $\mathcal{A}$ be a commutative Banach algebra and let $\Gamma(\mathcal{A})$ be the set of all complex homomorphisms on $\mathcal{A}$. If the intersection of the null spaces of all $\gamma \in \Gamma(\mathcal{A})$ contains only the zero element of $\mathcal{A}$, i.e. if*

$$\bigcap_{\gamma \in \Gamma(\mathcal{A})} \mathscr{N}(\gamma) = \{0\} \tag{3.20}$$

then every involution on $\mathcal{A}$ is continuous.

An algebra which satisfies (3.20) is called *semisimple.*

Proof. The theorem is proved using the closed graph theorem, which states that every linear and closed operator between two Banach spaces is continuous. Since the involution mapping is linear, we have to show that the set $G = \{(f, f^*) : f \in \mathcal{A}\}$ is closed in $\mathcal{A} \times \mathcal{A}$.

To this end, let $\{f_n\}$ be a sequence in $\mathcal{A}$ such that $f_n \to f$ and such that $f_n^* \to g$ in $\mathcal{A}$. We have to show that $g = f^*$. Let $\gamma \in \varGamma(\mathcal{A})$ be an arbitrary complex homomorphism on $\mathcal{A}$, and define $\eta(f) := \overline{\gamma(f^*)}$. Using the properties of the involution in Def. 3.23, it is easily verified that η is linear and multiplicative, i.e. that η is a complex homomorphism on $\mathcal{A}$. Because of part c) of Theorem 3.15, γ and η are continuous, and therefore one has the following identity

$$\overline{\gamma(f^*)} = \eta(f) = \lim_{n\to\infty} \eta(f_n) = \lim_{n\to\infty} \overline{\gamma(f_n^*)} = \overline{\gamma(g)} \,.$$

This implies that $g = f^* + h_\gamma$ with $h_\gamma \in \mathscr{N}(\gamma)$. However, since γ was chosen arbitrary and because of (3.20), $h_\gamma = 0$ so that $g = f^*$. $\square$

Notes

We refer to standard textbooks (e.g. [40, 60, 68] or [70, Chap. 18],[72, Chap. 10 and 11]) for a detailed introduction to Banach algebras.

4

Signal Models and Linear Systems

Nowadays, signal processing uses almost exclusively digital techniques. The majority of applications such as communications, television, speech and image processing, radar, sonar, to name just some, are based on digital signal processing. The major advantage of digital methods over analog ones is the availability of high-speed digital computers at low costs which allow for very efficient and flexible implementation of advanced signal processing algorithms along with decreasing implementation effort and costs. Therefore, we will discuss only digital systems in the present and in all of the following sections.

4.1 Signal Models

We consider digital systems with N input ports and M outputs. The input and output signal of such a system is a sequence $\hat{\boldsymbol{x}} = \{\hat{\boldsymbol{x}}(k)\}_{k=-\infty}^{\infty}$ and $\hat{\boldsymbol{y}} = \{\hat{\boldsymbol{y}}(k)\}_{k=-\infty}^{\infty}$ of vectors in $\mathbb{C}^N$ and $\mathbb{C}^M$, respectively. Thus, at a certain time instant k the signals $\hat{\boldsymbol{x}}(k) = [\,\hat{x}_1(k), \hat{x}_2(k), \cdots, \hat{x}_N(k)\,]$ are applied to the N inputs of the digital system $\mathcal{L}$ and the signals $\hat{\boldsymbol{y}}(k) = [\,\hat{y}_1(k), \hat{y}_2(k), \cdots, \hat{y}_M(k)\,]$ can be observed on the output ports of $\mathcal{L}$. If $N = M = 1$, the system $\mathcal{L}$ will said to be a single-input single-output (SISO) system.

Apart from this *time-domain* description of the input and output signals of the linear system $\mathcal{L}$, it is often advantageous to also consider the signals in the so-called *frequency domain*. We recall from Section 2 that given a sequence $\hat{x} = \{\hat{x}(k)\}_{k=-\infty}^{\infty}$, the *Fourier series* of $\hat{x}$ is given by

$$x(\mathrm{e}^{\mathrm{i}\theta}) = \sum_{k=-\infty}^{\infty} \hat{x}(k)\,\mathrm{e}^{\mathrm{i}k\theta}\;, \qquad \theta \in [-\pi, \pi) \tag{4.1}$$

provided that the series converges in a certain sense. This Fourier series of the time domain sequence $\hat{x}$ is sometimes called its *(frequency) spectrum.*

Conversely, given the spectrum $x(\mathrm{e}^{\mathrm{i}\theta})$, $\theta \in [-\pi, \pi)$ of a signal, the time domain symbols are obtained as the Fourier coefficients of x:

V. Pohl, H. Boche, *Advanced Topics in System and Signal Theory*, Foundations in Signal Processing, Communications and Networking 4,
DOI 10.1007/978-3-642-03639-2_4,

$$\hat{x}(k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} x(\mathrm{e}^{\mathrm{i}\theta})\,\mathrm{e}^{-\mathrm{i}k\theta}\,\mathrm{d}\theta\ , \qquad k = 0, \pm 1, \pm 2, \dots\ . \tag{4.2}$$

If $x \in L^1(\mathbb{T})$, this integral exists for every $k \in \mathbb{Z}$.

Besides the frequency spectrum of a time domain sequence $\hat{x} = \{\hat{x}(k)\}_{k=-\infty}^{\infty}$ one considers also the so called *$\mathscr{D}$-transform* of $\hat{x}$, which is given by

$$x(z) = \sum_{k=-\infty}^{\infty} \hat{x}(k)\, z^k\ , \qquad z \in \mathbb{C}\ . \tag{4.3}$$

Setting $z = \mathrm{e}^{\mathrm{i}\theta}$, $\theta \in [-\pi, \pi)$ in the $\mathscr{D}$-transform $x(z)$, one obtains again the frequency spectrum $x(\mathrm{e}^{\mathrm{i}\theta})$. Of course, the sum in (4.3) does not exit for all $z \in \mathbb{C}$. However, a minimal requirement on a signal $\hat{\boldsymbol{x}}$ will always be that its frequency spectrum (4.1) exists for almost all $\theta \in [-\pi, \pi)$. This implies that the $\mathscr{D}$-transform exists at least for almost all $z \in \mathbb{T}$. However, in many cases we will consider so called causal signals (see below). Then the $\mathscr{D}$-transform in (4.3) converges for all $z \in \mathbb{D}$.

Remark 4.1 (Notation). Note that the notations used here differ from the common notations in system theory. Usually, the spectrum of signal is defined as the (discrete) Fourier transform of its time domain representation, and the reverse transformation from the frequency to the time domain is usually done by the inverse Fourier transform. These transformations are obtained from (4.1) and (4.2) by replacing θ with $-\theta$. Moreover, usually the frequency representation of the signal is denoted by a "hat", or something similar, and not the time domain signal as it is done here.

4.1.1 Causal Signals

A digital signal $\hat{\boldsymbol{x}} = \{\hat{\boldsymbol{x}}(k)\}_{k=-\infty}^{\infty}$ with values in $\mathbb{C}^N$ is said to be *causal* if $\hat{\boldsymbol{x}}(k) = 0$ for all $k < 0$. The spectrum of such a causal signal is given by $\boldsymbol{x}(\mathrm{e}^{\mathrm{i}\theta}) = \sum_{k=0}^{\infty} \hat{\boldsymbol{x}}(k)\,\mathrm{e}^{\mathrm{i}k\theta}$ and its $\mathscr{D}$-transform is given by a power series

$$\boldsymbol{x}(z) = \sum_{k=0}^{\infty} \hat{\boldsymbol{x}}(k)\, z^k\ , \qquad z \in \mathbb{D} \tag{4.4}$$

which converges for every $z \in \mathbb{D}$ (provided that the spectrum exist) and which is analytic for all $z \in \mathbb{D}$. Thus, the $\mathscr{D}$-transform of every causal signal is an analytic function inside the unit disk $\mathbb{D}$, i.e. $\boldsymbol{x} \in H(\mathbb{C}^N)$.

4.1.2 Bounded Signals

A minimal requirement on useful signals is certainly that every symbol $\hat{\boldsymbol{x}}(k)$, $k = \dots, -2, -1, 0, 1, 2, \dots$ has to be finite. Thus, a digital signal

$\hat{\boldsymbol{x}} = \{\hat{\boldsymbol{x}}(k)\}_{k=-\infty}^{\infty}$ with values in $\mathbb{C}^N$ is called a *bounded signal* if there exists a constant $C < \infty$ so that

$$\|\hat{\boldsymbol{x}}(k)\|_{\mathbb{C}^N} \le C \qquad \text{for all } k \in \mathbb{Z} .$$

Thus, $\ell^\infty(\mathbb{C}^N)$ can be interpreted as the set of all bounded signals with values in $\mathbb{C}^N$ and $\ell^\infty_+(\mathbb{C}^N)$ can be identified with the set of all *causal and bounded signals*. Sometimes bounded signals are called *finite-power signals*.

4.1.3 Energy Signals

The requirement that the signals should be bounded is too weak in some applications. Instead, one requires that the energy of the signal remains finite. This results in the notation of finite-energy signals: A digital signal $\hat{\boldsymbol{x}} = \{\hat{\boldsymbol{x}}(k)\}_{k=-\infty}^{\infty}$ with values in $\mathbb{C}^N$ is called an *energy signal* if

$$\|\hat{\boldsymbol{x}}\|_{\ell^2} = \left(\sum_{k=-\infty}^{\infty} \|\hat{\boldsymbol{x}}(k)\|^2_{\mathbb{C}^N} \right)^{1/2} < \infty . \tag{4.5}$$

In other words, every energy signal belongs to $\ell^2(\mathbb{C}^N)$ and the set of all *causal energy signals* can be identified with $\ell^2_+(\mathbb{C}^N)$. Since $\ell^2(\mathbb{C}^N) \subset \ell^\infty(\mathbb{C}^N)$, it is clear that every energy signal has to be bounded. Additionally, the symbols $\hat{\boldsymbol{x}}(k)$ of an energy signal have to vanish as $|k| \to \infty$. Because the convergence of the series (4.5) implies that $\sum_{k=0}^{n} \|\hat{\boldsymbol{x}}(k)\|^2_{\mathbb{C}^N}$ is a Cauchy sequence in $\mathbb{C}^N$. Consequently, to every $\epsilon > 0$ there exists an $N \in \mathbb{N}$ such that $\sum_{k=n}^{m} \|\hat{\boldsymbol{x}}(k)\|^2_{\mathbb{C}^N} < \epsilon$ for all $m \ge n \ge N$. It follows in particular for $m = n$ that $\|\hat{\boldsymbol{x}}(k)\|^2_{\mathbb{C}^N} < \epsilon$ for all $k \ge N$, and it is clear that a similar relation holds for the anti-causal part $(k < 0)$ of the series. All this implies that

$$\lim_{|k|\to\infty} \|\hat{\boldsymbol{x}}(k)\|^2_{\mathbb{C}^N} = 0$$

for every energy signal. So energy signals are concentrated in time (in the above sense) since the signal components $\hat{\boldsymbol{x}}(k)$ die away as $|k| \to \infty$.

By Parseval's identity and in view of Example 1.9, it holds that $\|\hat{\boldsymbol{x}}\|_{\ell^2} = \|\boldsymbol{x}\|_2$. Thus, $\hat{\boldsymbol{x}}$ is an energy signal if and only its frequency spectrum $\boldsymbol{x}$ satisfies

$$\|\boldsymbol{x}\|_2 = \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \|\boldsymbol{x}(\mathrm{e}^{\mathrm{i}\theta})\|^2_{\mathbb{C}^N} \, \mathrm{d}\theta \right)^{1/2} < \infty .$$

Therefore $L^2(\mathbb{C}^N)$ can be considered as the set of all spectra of energy signals, whereas the Hardy space $H^2(\mathbb{C}^N)$ contains the spectra of all causal energy signals. It is worth noting that $\ell^2(\mathbb{C}^N)$ and $\ell^2_+(\mathbb{C}^N)$, as well as $L^2(\mathbb{C}^N)$ and $H^2(\mathbb{C}^N)$, are Hilbert spaces, which makes working with energy signals in some respects easier than with other signal models.

4.2 Linear Systems – Properties and Representation

A digital system $\mathcal{L}$ is a transformation that takes any digital input signal $\hat{\boldsymbol{x}} = \{\hat{\boldsymbol{x}}_k\}_{k=-\infty}^{\infty}$ to a digital signal $\hat{\boldsymbol{y}} = \{\hat{\boldsymbol{y}}_k\}_{k=-\infty}^{\infty} = \mathcal{L}[\hat{\boldsymbol{x}}]$ at the output of the system. Similarly, one may consider the signals in the frequency domain. Then the system $\mathcal{L}$ takes the spectrum $\boldsymbol{x}$ of the input signal to the spectrum $\boldsymbol{y} = \mathcal{L}[\boldsymbol{x}]$ of the output signal.

Usually the input and output signals have certain specified or required properties (causality, boundedness, finite energy, etc.). These properties are characterized by the Banach space from which these signals are taken. Thus, a (digital) system is a mapping

$$\mathcal{L} : \mathcal{B}_1 \to \mathcal{B}_2$$

which maps all signals in a certain Banach space $\mathcal{B}_1$ onto another Banach space $\mathcal{B}_2$. In the following we consider in particular systems $\mathcal{L} : \ell_+^\infty(\mathbb{C}^N) \to \ell_+^\infty(\mathbb{C}^M)$ mapping causal bounded signals at the input onto causal bounded signals at the output of $\mathcal{L}$, and we consider systems $\mathcal{L} : H^2(\mathbb{C}^N) \to H^2(\mathbb{C}^M)$ mapping causal energy signals onto causal energy signals.

4.2.1 Basic System Properties

The above definition of a digital system is still quite general. In applications, additional assumptions on the properties of $\mathcal{L}$ are usually made. Some of these properties will be discussed next.

Definition 4.2 (Linearity). *A digital system $\mathcal{L}$ is said to be* linear *if*

$$\mathcal{L}[\alpha\,\hat{\boldsymbol{x}}_1 + \beta\,\hat{\boldsymbol{x}}_2] = \alpha\,\mathcal{L}[\hat{\boldsymbol{x}}_1] + \beta\,\mathcal{L}[\hat{\boldsymbol{x}}_2]$$

for all input signal $\hat{\boldsymbol{x}}_1$ and $\hat{\boldsymbol{x}}_2$ and arbitrary complex numbers α and β.

Of course, this linearity condition can be formulated equivalently in the frequency domain as $\mathcal{L}[\alpha\,\boldsymbol{x}_1 + \beta\,\boldsymbol{x}_2] = \alpha\,\mathcal{L}[\boldsymbol{x}_1] + \beta\,\mathcal{L}[\boldsymbol{x}_2]$. If not mentioned otherwise, all systems considered in the following are assumed to be linear digital systems.

Definition 4.3 (Stability). *Let $\mathcal{L} : \mathcal{B}_1 \to \mathcal{B}_2$ be a linear system mapping signals from a Banach space $\mathcal{B}_1$ to signals in a Banach space $\mathcal{B}_2$. The linear system $\mathcal{L}$ is called stable, if the linear mapping $\mathcal{L} : \mathcal{B}_1 \to \mathcal{B}_2$ is bounded, i.e. if*

$$\|\mathcal{L}\|_{\mathcal{B}_1\to\mathcal{B}_2} = \sup_{\boldsymbol{x}\in\mathcal{B}_1} \frac{\|\mathcal{L}\boldsymbol{x}\|_{\mathcal{B}_2}}{\|\boldsymbol{x}\|_{\mathcal{B}_1}} < \infty\,.$$

The stability of $\mathcal{L}$ implies that there exists a constant $C = \|\mathcal{L}\|_{\mathcal{B}_1\to\mathcal{B}_2}$ such that

$$\|\mathcal{L}\boldsymbol{x}\|_{\mathcal{B}_2} \le C\,\|\boldsymbol{x}\|_{\mathcal{B}_1} \qquad \text{for all} \qquad \boldsymbol{x} \in \mathcal{B}_1\,.$$

Thus, if the system is stable, then the norm of the output signal $\mathcal{L}\boldsymbol{x}$ can always be controlled by the norm $\|\boldsymbol{x}\|_{\mathcal{B}_1}$ of the input signal. The above definition shows that the stability of a linear system depends strongly on the Banach spaces $\mathcal{B}_1$ and $\mathcal{B}_2$ of the input and output signals. In many cases, the signal spaces of the input and output signal are equal. We consider in particular the following two particular cases:

(a) If $\mathcal{B}_1 = H^2(\mathbb{C}^N)$ and $\mathcal{B}_2 = H^2(\mathbb{C}^M)$ for some dimensions $N, M \in \mathbb{N}$, i.e. if the input and output of the linear system are signals of finite energy, and if $\mathcal{L} : H^2(\mathbb{C}^N) \to H^2(\mathbb{C}^M)$ is bounded, then the linear system $\mathcal{L}$ is called *energy stable.*
(b) If $\mathcal{B}_1 = \ell^\infty(\mathbb{C}^N)$ and $\mathcal{B}_2 = \ell^\infty(\mathbb{C}^M)$ for some dimensions $N, M \in \mathbb{N}$, i.e. if the input and and output of $\mathcal{L}$ are bounded signals, and if $\mathcal{L} : \ell^\infty(\mathbb{C}^N) \to \ell^\infty(\mathbb{C}^M)$ is bounded, then the linear system $\mathcal{L}$ is called *bounded input bounded output (BIBO) stable.*

Definition 4.4 (Causality). *A linear system $\mathcal{L}$ is said to be* causal *if for every input signal $\hat{\boldsymbol{x}}$ with $\hat{\boldsymbol{x}}(k) = 0$ for all $k < 0$, $(\mathcal{L}\hat{\boldsymbol{x}})(k) = 0$ for all $k < 0$.*

In other words, we cannot influence the past of the output by present or future inputs. Thus, a causal linear system maps causal input signals onto causal output signals. Since the $\mathscr{D}$-transform of every causal signal is an analytic function in $\mathbb{D}$, every causal linear system with N inputs and M-outputs can be considered as a mapping $\mathcal{L} : H(\mathbb{C}^N) \to H(\mathbb{C}^M)$ which maps analytic functions $\boldsymbol{x} \in H(\mathbb{C}^N)$ with values in $\mathbb{C}^N$ onto analytic functions $\mathcal{L}\boldsymbol{x} \in H(\mathbb{C}^M)$.

A linear system $\mathcal{L}$ is called *time-invariant*, if the output of the system does not depend on the absolute time, but only on the input signal. Thus, whether one applies an input signal $\hat{\boldsymbol{x}}$ to the system $\mathcal{L}$ now or K time instances later, the output $\hat{\boldsymbol{x}} = \mathcal{L}\hat{\boldsymbol{x}}$ will be identical, except for a time delay of the K time slots. To get a proper formal definition, we introduce the (right) shift operator $\mathcal{R} : \ell^\infty_+(\mathbb{C}^N) \to \ell^\infty_+(\mathbb{C}^N)$ on the space of all bounded and causal input signals. It is defined by

$$\mathcal{R}\{\hat{\boldsymbol{x}}(0), \hat{\boldsymbol{x}}(1), \hat{\boldsymbol{x}}(2), \dots\} = \{0, \hat{\boldsymbol{x}}(0), \hat{\boldsymbol{x}}(1), \dots\} , \qquad \{\hat{\boldsymbol{x}}(k)\}_{k=0}^\infty \in \ell^\infty(\mathbb{C}^N) .$$

Applying the right shift operator $n - \textit{times}$ consecutively on the sequence $\hat{\boldsymbol{x}} \in \ell^\infty(\mathbb{C}^N)$ yields obviously

$$\mathcal{R}^n\{\hat{\boldsymbol{x}}(0), \hat{\boldsymbol{x}}(1), \hat{\boldsymbol{x}}(2), \dots\} = \{\underbrace{0, \dots, 0,}_{n-\text{times}} \hat{\boldsymbol{x}}(0), \hat{\boldsymbol{x}}(1), \dots\} .$$

This right shift operator $\mathcal{R}$ has also a natural $\mathscr{D}$-domain description. Let $\boldsymbol{x}(z) = \sum_{k=0}^\infty \hat{\boldsymbol{x}}(k)\, z^k$ be a function in a certain $H^p(\mathbb{C}^N)$-space. Then the right shift of $\boldsymbol{x}$ is given by

$$(\mathcal{R}\boldsymbol{x})(z) = \sum_{k=1}^\infty \hat{\boldsymbol{x}}(k)\, z^k = z\, \boldsymbol{x}(z) , \qquad z \in \mathbb{D}$$

and $(\mathcal{R}^n\boldsymbol{x})(z) = z^n\, \boldsymbol{x}(z)$.

Example 4.5 (Unit pulses). Let $\hat{\delta}_0 = \{1, 0, 0, \cdots\}$ be the sequence that contains a one at the first position and is zero elsewhere. We call this particular sequence *unit pulse*. Then $\mathcal{R}^n \hat{\delta}_0$ is the sequence which contains a one at position n and zeros elsewhere, and in the $\mathscr{D}$-domain one has $(\mathcal{R}^n \delta_0)(z) = z^n$.

Using the right shift operator we define now the time-invariance of a linear system.

Definition 4.6 (Time-invariance). *A linear system $\mathcal{L}$ is called* time-invariant*, if it commutes with the right shift operator $\mathcal{R}$, i.e. if*

$$\mathcal{R}\mathcal{L}\hat{\boldsymbol{x}} = \mathcal{L}\mathcal{R}\hat{\boldsymbol{x}}, \quad \text{for all } \hat{\boldsymbol{x}} \in \mathscr{D}(\mathcal{L})$$

wherein $\mathscr{D}(\mathcal{L}) \subset \ell^\infty_+(\mathbb{C}^N)$ is the domain of $\mathcal{L}$, i.e. the set of all possible input signals.

4.2.2 Linear Time-invariant Systems

In future sections, we consider almost exclusively linear and time-invariant (LTI) systems $\mathcal{L}$. It is well known that the input-output relation of an LTI systems can be characterized by the so called *transfer function* or equivalently by the *impulse response* of the system. However, to obtain such a simple description of the system $\mathcal{L}$ it is generally necessary to require that the system $\mathcal{L}$ is stable. The following theorem establishes these two characterizations of the input-output relation of an LTI systems. This theorem is derived for causal systems, since such systems are considered primarily in the following.

Theorem 4.7. *Let $1 < p < \infty$ and $\mathcal{L} : H^p(\mathbb{C}^N) \to H^p(\mathbb{C}^M)$ be a system with input $\boldsymbol{f} \in H^p(\mathbb{C}^N)$ and output $\boldsymbol{g} = \mathcal{L}\,\boldsymbol{f} \in H^p(\mathbb{C}^M)$. Then $\mathcal{L}$ is linear, stable, causal, and time-invariant if and only if there exists a bounded analytic matrix function $\mathbf{H}(z) = \sum_{k=1}^{\infty} \hat{\mathbf{H}}(k)\, z^k$ in $H^\infty(\mathbb{C}^{M\times N})$ such that*

$$(\mathcal{L}\boldsymbol{f})(z) = (\mathbf{O}^+_{\mathbf{H}}\, \boldsymbol{f})(z) = \mathbf{H}(z)\, \boldsymbol{f}(z)\ , \qquad z \in \mathbb{D}\ . \tag{4.6}$$

for all $\boldsymbol{f} \in H^p(\mathbb{C}^N)$, or equivalently that

$$\hat{\boldsymbol{g}}(k) = \sum_{n=0}^{\infty} \hat{\mathbf{H}}(n)\, \hat{\boldsymbol{f}}(k-n)\ , \qquad k = 0, 1, 2, \ldots \tag{4.7}$$

for all $\boldsymbol{f} \in H^p(\mathbb{C}^N)$. Moreover $\|\mathcal{L}\| = \|\mathbf{O}_H\| = \|\mathbf{H}\|_\infty$.

Proof. First we show that (4.6) and (4.7) are equivalent if the statement of the theorem is true. Thus, assume that there exists $\mathbf{H} \in H^\infty(\mathbb{C}^{M\times N})$ such that $\boldsymbol{g}(z) = \mathbf{H}(z)\, \boldsymbol{f}(z)$ for all $z \in \mathbb{D}$. Let $\{\boldsymbol{e}_m\}_{m=1}^M$ be a basis of the Hilbert space $\mathbb{C}^M$. We consider the individual components $g_m(z) = \langle \boldsymbol{g}(z), \boldsymbol{e}_m\rangle_{\mathbb{C}^M}$. Their Fourier coefficients are given by

$$\hat{g}_m(k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \left\langle \mathbf{H}(e^{i\tau})\,\boldsymbol{f}(e^{i\tau})\,,\,\boldsymbol{e}_m\right\rangle_{\mathbb{C}^M}\, e^{-ik\tau}\,d\tau\;,$$

which exist for all k, since $\mathbf{O}_H\boldsymbol{f} \in H^p(\mathbb{C}^M)$. Now, we insert the power series expansion of $\mathbf{H}$. Since this series converges absolutely, one can interchange the integration with the summation and obtain

$$\begin{aligned}
\hat{g}_m(k) &= \sum_{n=0}^{\infty}\frac{1}{2\pi}\int_{-\pi}^{\pi}\left\langle \hat{\mathbf{H}}(n)\,\boldsymbol{f}(e^{i\tau})\,,\,\boldsymbol{e}_m\right\rangle_{\mathbb{C}^M} e^{-i(k-n)\tau}\,d\tau\\
&= \sum_{n=0}^{\infty}\frac{1}{2\pi}\int_{-\pi}^{\pi}\left\langle \boldsymbol{f}(e^{i\tau})\,,\,\hat{\mathbf{H}}^*(n)\,\boldsymbol{e}_m\right\rangle_{\mathbb{C}^M} e^{-i(k-n)\tau}\,d\tau\\
&\overset{(a)}{=} \sum_{n=0}^{\infty}\left\langle \hat{\boldsymbol{f}}(k-n)\,,\,\hat{\mathbf{H}}^*(n)\,\boldsymbol{e}_m\right\rangle_{\mathbb{C}^M} = \left\langle \sum_{n=0}^{\infty}\hat{\mathbf{H}}(n)\,\hat{\boldsymbol{f}}(k-n)\,,\,\boldsymbol{e}_m\right\rangle_{\mathbb{C}^M}
\end{aligned}$$

where $\hat{\mathbf{H}}^*(n)$ denotes the adjoint of the complex matrix $\hat{\mathbf{H}}(n)$ and where (a) follows from the definition of the Fourier series for vector valued functions, see (2.59). Collecting all equations for the individual components $\hat{g}_m(k)$, $m = 1, 2, \ldots, M$ into one vector, one obtains (4.7). This shows that (4.6) and (4.7) are equivalent.

Assume now that there exists an $\mathbf{H} \in H^\infty(\mathbb{C}^{M\times N})$ such that (4.6) respectively (4.7) describes the input-output relation of $\mathcal{L}$. Then it follows at once from (4.6) and (4.7) that $\mathcal{L}$ is linear, causal, and time-invariant. Moreover, Proposition 2.34 implies that $\|\mathbf{O}_{\mathbf{H}}^{+}\| = \|\mathbf{H}\|_\infty < \infty$.

For the "only if part", we have to show that if $\mathcal{L}$ is linear, stable, causal, and time-invariant, then there exists an $\mathbf{H} \in H^p(\mathbb{C}^{M\times N})$ with the prescribed properties. To this end we consider first the SISO case $N = M = 1$: Let δ_0 be the unit pulse (cf. Example 4.5) and set $H(z) := (\mathcal{L}\,\delta_0)(z)$. It is clear that $H \in H^p$ and in particular that H is an analytic function in $\mathbb{D}$. Now, let $\delta_k = \mathcal{R}^k\delta_0$ with $k = 0, 1, 2, \ldots$ be the unit pulse at time instance k. Since $\mathcal{L}$ is time-invariant, we have

$$(\mathcal{L}\,\delta_k)(z) = (\mathcal{L}\,\mathcal{R}^k\,\delta_0)(z) = (\mathcal{R}^k\,\mathcal{L}\,\delta_0)(z) = z^k\,H(z)\;.$$

The span of $\{\delta_k\}_{k=0}^{\infty}$ is dense in H^p. Therefore, the partial sum $f_K = \sum_{k=0}^{K}\hat{f}(k)\,\delta_k$ converges to f as $K \to \infty$ for every $f \in H^p$. Applying $\mathcal{L}$ to f_K gives

$$(\mathcal{L}\,f_K)(z) = \sum_{k=0}^{K}\hat{f}(k)\,(\mathcal{L}\delta_k)(z) = H(z)\sum_{k=0}^{K}\hat{f}(k)\,z^k = H(z)\,f_K(z)\;.$$

Since $\|\mathcal{L}\,f_K - \mathcal{L}\,f\|_{H^p(\mathbb{C}^M)} \le \|\mathcal{L}\|\,\|f_K - f\|_{H^p(\mathbb{C}^N)}$ and because of the boundedness of $\mathcal{L}$ follows that $\mathcal{L}\,f_K$ converges to $\mathcal{L}\,f$ as $K \to \infty$. Since $\mathcal{L}$ is assumed to be bounded, Proposition 2.34 implies that $h \in H^\infty$. Thus in the SISO case, the so constructed transfer function H has all the desired properties.

The vector valued case is readily obtained from the scalar case by considering the individual components of the signals in $H^p(\mathbb{C}^N)$ and $H^p(\mathbb{C}^M)$ and the individual operators $[\mathcal{L}]_{j,i}$ between the i-th input and the j-th output of the system $\mathcal{L}$. Clearly, every $[\mathcal{L}]_{j,i}$ is a bounded time-invariant causal linear operator which is characterized by an analytic transfer function $H_{j,i}(z)$. Thus, by the linearity of the system $\mathcal{L}$, we have that $(\mathcal{L}\boldsymbol{f})(z) = \mathbf{H}(z)\,\boldsymbol{f}(z)$ for every $\boldsymbol{f} \in H^p(\mathbb{C}^N)$ in which $\mathbf{H}$ is an $M \times N$ matrix with the individual entries $[\mathbf{H}(z)]_{j,i} = H_{j,i}(z) \in H^\infty$. Proposition 2.34 implies that $\mathbf{H}$ is an element of $H^\infty(\mathbb{C}^{M\times N})$, because $\|\mathcal{L}\| = \|\mathbf{O_H}\| = \|\mathbf{H}\|_\infty$ and $\|\mathcal{L}\| < \infty$ by the assumption of the theorem. □

The previous theorem showed that every stable time-invariant causal linear system $\mathcal{L} : H^p(\mathbb{C}^N) \to H^p(\mathbb{C}^M)$ is uniquely determined by a matrix valued bounded analytic function $\mathbf{H} \in H^\infty(\mathbb{C}^{M\times N})$ with power series expansion (2.63) in such a way that the input-output relation of $\mathcal{L}$ can be written as (4.6) in the $\mathscr{D}$-domain or as (4.7) in the time domain. Thereby, the bounded analytic function $\mathbf{H}$ is called the *(matrix) transfer function* of the linear system $\mathcal{L}$ and the corresponding sequence $\hat{\mathbf{H}} = \{\hat{\mathbf{H}}(k)\}_{k=0}^{\infty}$ of Fourier coefficients is called the *(matrix) impulse response* of $\mathcal{L}$.

Especially in later chapters, we will always identify a matrix transfer function $\mathbf{H} \in L^\infty(\mathbb{C}^{M\times N})$ with the multiplication operator $\mathbf{O_H} : L^2(\mathbb{C}^N) \to L^2(\mathbb{C}^M)$ defined by $(\mathbf{O_H}\,\boldsymbol{f})(\zeta) = \mathbf{H}(\zeta)\,\boldsymbol{f}(\zeta)$ for all $\zeta \in \mathbb{T}$. Thus, speaking for example of the range or null space of $\mathbf{H}$, we always mean the range or null space of $\mathbf{O_H}$.

Of particular interest are linear systems $\mathcal{L} : H^2(\mathbb{C}^N) \to H^2(\mathbb{C}^M)$ which map energy signals onto energy signals. For these particular systems, Theorem 4.7 contains the following well known result as a special case:

Corollary 4.8. *A causal, linear, and time-invariant system $\mathcal{L}$ is energy stable if and only if its transfer function $\mathbf{H}$ is a bounded analytic function $\mathbf{H} \in H^\infty(\mathbb{C}^{M\times N})$.*

Theorems 4.7 holds for linear systems $\mathcal{L} : H^p(\mathbb{C}^N) \to H^p(\mathbb{C}^M)$ with $1 < p < \infty$. However, it cannot be applied for systems $\mathcal{L}$ mapping bounded signals $\hat{\boldsymbol{f}} \in \ell_+^\infty(\mathbb{C}^N)$ onto bounded signals $\mathcal{L}\,\hat{\boldsymbol{f}} \in \ell_+^\infty(\mathbb{C}^M)$ since $\hat{\boldsymbol{f}} \in \ell_+^\infty(\mathbb{C}^N)$ does not imply that $\boldsymbol{f} \in H^p(\mathbb{C}^N)$, in general. To see this, assume that $\hat{f} = \{\hat{f}(k)\}_{k=0}^{\infty}$ is an element of ℓ_+^∞ (which means that $\sup_{k\geq 0} |\hat{f}(k)| = \|\hat{f}\|_\infty < \infty$) and let $f(z) = \sum_{k=0}^{\infty} \hat{f}(k)\, z^k$ be the corresponding $\mathscr{D}$-transform. Then for an arbitrary $z \in \mathbb{D}$, we have

$$|f(z)| \leq \sum_{k=0}^{\infty} |\hat{f}(k)|\,|z|^k \leq \|\hat{f}\|_{\ell^\infty} \sum_{k=0}^{\infty} |z|^k = \frac{1}{1-|z|}\,\|\hat{f}\|_{\ell^\infty}$$

where equality holds for some $\hat{f} \in \ell_+^\infty$ and $z \in \mathbb{D}$. Since the right hand side diverges as $|z| \to 1$, it is clear that f need not be an element of any H^p space.

However, for causal linear systems $\mathcal{L} : \ell_+^\infty(\mathbb{C}^N) \to \ell_+^\infty(\mathbb{C}^M)$ the following result is obtained.

Theorem 4.9. *Let $\mathcal{L} : \ell_+^\infty(\mathbb{C}^N) \to \ell_+^\infty(\mathbb{C}^M)$ be a system with input $\hat{\boldsymbol{f}} \in \ell_+^\infty(\mathbb{C}^N)$ and output $\hat{\boldsymbol{g}} = \mathcal{L}\,\hat{\boldsymbol{f}} \in \ell_+^\infty(\mathbb{C}^M)$. Then $\mathcal{L}$ is linear, stable, causal, and time-invariant if and only if there exists an $M \times N$ matrix $\mathbf{H}(z)$ whose individual entries $H_{i,j}(z) = [\mathbf{H}(z)]_{i,j}$ have the form*

$$H_{i,j}(z) = \sum_{k=0}^{\infty} \hat{H}_{i,j}(k)\, z^k \qquad \textit{with} \qquad \|\hat{H}_{i,j}\|_{\ell^1} = \sum_{k=0}^{\infty} \left|\hat{H}_{i,j}(k)\right| < \infty$$

for all $i = 1, \ldots, M$ and $j = 1, \ldots, N$, and such that (4.6) *and* (4.7) *hold for every $\hat{\boldsymbol{f}} \in \ell_+^\infty(\mathbb{C}^N)$.*

Proof. As in the proof of Theorem 4.7, it is clear that (4.6) and (4.7) are equivalent. Given $\mathbf{H}$ with the specified properties, (4.6) and (4.7) show that $\mathcal{L}$ is linear, causal, and time-invariant. It remains to show that $\mathcal{L}$ is also bounded. To this end, we consider an arbitrary component $\hat{g}_i$ of the system output $\hat{\boldsymbol{g}} = \mathcal{L}\,\hat{\boldsymbol{f}}$. According to (4.7) its is given by

$$\hat{g}_i(k) = \sum_{j=1}^{N} \sum_{n=0}^{\infty} \hat{H}_{i,j}(n)\, f_j(k-n)\ , \quad i = 1, \ldots, M\ ; \quad k = 1, 2, \ldots\ .$$

For its modulus, one finds the upper bound

$$|\hat{g}_i(k)| \le \sum_{j=1}^{N} \sum_{n=0}^{\infty} \left|\hat{H}_{i,j}(n)\right|\ |f_j(k-n)| \le N \max_{i,j} \|\hat{H}_{i,j}\|_{\ell_+^1}\ \|\hat{\boldsymbol{f}}\|_{\ell_+^\infty(\mathbb{C}^N)}$$

for all $i = 1, \ldots, M$ and all $k \in \mathbb{N}$. Define the constant $C_0 := N \max_{i,j} \|\hat{H}_{i,j}\|_{\ell_+^1}$ which depends on the impulse response matrix $\hat{\mathbf{H}}$ and which is finite, one obtains for the norm of the output signal

$$\|\hat{\boldsymbol{g}}(k)\|_{\mathbb{C}^M}^2 = \sum_{i=1}^{M} |g_i(k)|^2 \le M\, C_0^2\, \|\hat{\boldsymbol{f}}\|_{\ell_+^\infty(\mathbb{C}^N)}^2\ .$$

Since this last bound holds for all $k \in \mathbb{N}$, it shows that $\mathcal{L}$ is bounded.

For the "only if part" we start again with the SISO case ($N = M = 1$) and define $\hat{H} := \mathcal{L}\,\hat{\delta}_0$, in which $\hat{\delta}_0 = \{1, 0, 0, \cdots\} \in \ell_+^\infty$ is the unit pulse (cf. Example 4.5) and where $\hat{H} = \{\hat{H}(0), \hat{H}(1), \hat{H}(2), \cdots\}$ is the *impulse response* of the (SISO) system $\mathcal{L}$. Applying the unit pulse at time instant k to $\mathcal{L}$ gives the output signal

$$\mathcal{L}\,\hat{\delta}_k = \mathcal{L}\,\mathcal{R}^k\,\hat{\delta}_0 = \mathcal{R}^k\,\mathcal{L}\,\hat{\delta}_0 = \mathcal{R}^k\,\hat{H} = \{\underbrace{0, \cdots, 0}_{k-\text{times}}, \hat{H}(0), \hat{H}(1), \cdots\}\ ,$$

using that $\mathcal{L}$ is time-invariant. Of course, every $\hat{f} \in \ell_+^\infty$ can be written as $\hat{f} = \sum_{k=0}^{\infty} \hat{f}(k)\,\hat{\delta}_k$. Applying the operator $\mathcal{L}$ to this function gives

$$\mathcal{L}\hat{f} = \mathcal{L}\left(\sum_{k=0}^{\infty} \hat{f}(k)\,\hat{\delta}_k\right) = \sum_{k=0}^{\infty} \hat{f}(k)\,\mathcal{L}\hat{\delta}_k = \sum_{k=0}^{\infty} \hat{f}(k)\,\mathcal{R}^k \hat{H}\ . \tag{4.8}$$

Next, we consider the n-th symbol in the output sequence $\mathcal{L}\hat{f}$, which is equal to

$$(\mathcal{L}\hat{f})(n) = \sum_{k=0}^{n} \hat{f}(k)\,\hat{H}(n-k) = \sum_{m=0}^{\infty} \hat{H}(m)\,\hat{f}(n-m)\ ,$$

using for the last equation that $\hat{f}(k) = 0$ for all $k < 0$. Moreover, for the modulus of the output symbols hold

$$\left|(\mathcal{L}\hat{f})(n)\right| \leq \sum_{m=0}^{\infty} |\hat{H}(m)|\,|\hat{f}(n-m)| \leq \|\hat{f}\|_{\ell^\infty} \sum_{m=0}^{\infty} |\hat{H}(m)| < \infty\ , \tag{4.9}$$

since $\mathcal{L}$ was assumed to be bounded. This shows that $\|\hat{H}\|_{\ell^1} < \infty$, and that the right hand side of (4.8) converges uniformly for all time instances n. This justifies the interchange of $\mathcal{L}$ with the infinite sum in (4.8). Thus, it was shown that if the SISO system $\mathcal{L}$ is linear, causal, time-invariant, and bounded then the output $\hat{g}$ of $\mathcal{L}$ is obtained by the convolution of the input $\hat{f}$ with the impulse response $\hat{H}$ of $\mathcal{L}$ (4.7), and $\hat{H} \in \ell^1$. If $\mathcal{L}$ has several inputs and outputs, the statement of the theorem follows from the SISO case by the linearity of $\mathcal{L}$. □

Unlike Theorem 4.7 for energy stable systems, the previous Theorem 4.9 makes no statement on the norm $\|\mathcal{L}\|$ of the linear system $\mathcal{L}$ in terms of the transfer function $\mathbf{H}$ or in terms of the impulse response $\hat{\mathbf{H}}$ of $\mathcal{L}$. In the case of BIBO stable systems only the obvious upper bound

$$\|\mathcal{L}\|_{\ell_+^\infty(\mathbb{C}^N)\to\ell_+^\infty(\mathbb{C}^M)} \leq \sum_{n=0}^{\infty} \|\hat{\mathbf{H}}(n)\|_{\mathbb{C}^N\to\mathbb{C}^M} \tag{4.10}$$

can be given, in general. It follows at once from the representation (4.7) of $\mathcal{L}$, since

$$\begin{aligned}\|(\mathcal{L}\hat{\boldsymbol{f}})(k)\|_{\mathbb{C}^M} &\leq \sum_{n=0}^{\infty} \|\hat{\mathbf{H}}(n)\|_{\mathbb{C}^N\to\mathbb{C}^M}\,\|\hat{\boldsymbol{f}}(k-n)\|_{\mathbb{C}^N} \\ &\leq \|\hat{\boldsymbol{f}}\|_{\ell^\infty(\mathbb{C}^N)} \sum_{n=0}^{\infty} \|\hat{\mathbf{H}}(n)\|_{\mathbb{C}^N\to\mathbb{C}^M}\ .\end{aligned}$$

Therein $\|\hat{\mathbf{H}}(n)\|_{\mathbb{C}^N\to\mathbb{C}^M}$ is the operator norm of the matrices $\hat{\mathbf{H}}(n) : \mathbb{C}^N \to \mathbb{C}^M$, which is known to be equal to the *spectral norm*, i.e. largest singular value of $\hat{\mathbf{H}}(n)$:

$$\|\hat{\mathbf{H}}(n)\|_{\mathbb{C}^N \to \mathbb{C}^M} = \sqrt{\lambda_{\max}\left\{\hat{\mathbf{H}}^*(n)\,\hat{\mathbf{H}}(n)\right\}}\,.$$

However, in the case of a SISO system ($N = M = 1$) even equality holds in (4.10). This is proved in the following theorem.

Theorem 4.10. *Let $\mathcal{L} : \ell_+^\infty \to \ell_+^\infty$ be a linear, stable, causal, and time-invariant SISO system with impulse response $\hat{H} = \{\hat{H}(k)\}_{k=0}^\infty$. Then the norm of $\mathcal{L}$ is given by*

$$\|\mathcal{L}\| = \|\hat{H}\|_{\ell^1} = \sum_{k=0}^{\infty} \left|\hat{H}(k)\right|\,.$$

Proof. Equation (4.9) implies that $\|\mathcal{L}\,\hat{f}\|_{\ell^\infty} \leq \|\hat{H}\|_{\ell^1}\|\hat{f}\|_{\ell^\infty}$ which shows that $\|\mathcal{L}\| \leq \|\hat{H}\|_{\ell^1}$. To show that also the inverse inequality holds, we write every individual element $\hat{H}(k)$ of the impulse response as $\hat{H}(k) = |\hat{H}(k)|\,\mathrm{e}^{\mathrm{i}\Phi_H(k)}$ and define a sequence of input signals by

$$\hat{f}_n(k) := \begin{cases} \mathrm{e}^{-\mathrm{i}\Phi_H(n-k)} & k = 0, 1, 2, \dots, n \\ 0 & \text{otherwise} \end{cases}, \qquad \text{for } n = 1, 2, \dots\,.$$

For these functions, it obviously holds that $\|\hat{f}_n\|_{\ell^\infty} = 1$ and that

$$(\mathcal{L}\,\hat{f}_n)(n) = \sum_{m=0}^{n} |\hat{H}(m)|\,\mathrm{e}^{\mathrm{i}\Phi_H(m)}\,\hat{f}_n(n-m) = \sum_{m=0}^{n} |\hat{H}(m)|\,.$$

Therefore $\sup_{n\in\mathbb{N}} |(\mathcal{L}\,\hat{f}_n)(n)| = \|\hat{H}\|_{\ell^1}$, which gives

$$\|\mathcal{L}\| = \sup_{\hat{f}\in\ell_+^\infty,\, \|\hat{f}\|_{\ell^\infty}\leq 1} \|\mathcal{L}\hat{f}\|_{\ell^\infty} \geq \sup_{n\in\mathbb{N}} \|\mathcal{L}\,\hat{f}_n\|_{\ell^\infty} = \|\hat{H}\|_{\ell^1}$$

and which proves that $\|\mathcal{L}\| = \|\hat{H}\|_{\ell^1}$. □

Theorem 4.7 and 4.9 establish that the input-output relation of a linear system can be represented by a convolution of the input signal with the impulse response of the system, or equivalently in the frequency domain by a multiplication of the spectrum of the input signal with the transfer function of the system. This result is considered as one of the cornerstones of liner system theory and consequently discussed in depth in almost any textbook on this topic (e.g. [63, 67] among many others). It should be noted however, that this statement is not true in general, i.e. there exists linear systems whose input-output map is not entirely characterized by its impulse response. For a discussion of such cases, we refer to a series of papers published by Irvine Sandberg [73, 75, 76], and to [26, 74], in which the time continuous systems are considered.

Notes

[65]

Part II

Fundamental Operators

5

Poisson Integral and Hilbert Transformation

Given the transfer function $f(\mathrm{e}^{\mathrm{i}\theta}) = \sum_{k=-\infty}^{\infty} \hat{f}(k)\,\mathrm{e}^{\mathrm{i}k\theta}$ with $\theta \in [-\pi, \pi)$ of a non causal system. Then the operation $\mathfrak{P}_+ : f \mapsto \sum_{k=0}^{\infty} \hat{f}(k)\,\mathrm{e}^{\mathrm{i}k\theta}$, which cuts off the anti-causal part of the transfer function, is called the Riesz projection. This operation plays a prominent roll in system theory, as soon as the causality of certain system has to be enforced. For example, in estimation and detection problems, the determination of the causal linear filter which minimizes the means square error criterion (the so called Wiener filter) involves the Riesz projection, and also the so called spectral factorization comprises a Riesz projection (cf. Section 10). In Section 6 we will investigate the analytic behavior of the Riesz projection on different Banach spaces in some detail. However, in the present section, we first investigate the behavior of the Poisson and the conjugate Poisson integrals, since the real and the imaginary part of the Riesz projection are essentially given by these two integrals, respectively. However, these results are also of considerable interest by themselves since these integral transforms are very important in many different areas of physics and engineering.

5.1 Definitions

Suppose $f \in L^1$ is a function on the unit circle $\mathbb{T}$ with Fourier coefficients $\hat{f}(k)$, $k \in \mathbb{Z}$ given by (2.1) and consider the function defined by

$$(\mathfrak{R}f)(z) := \hat{f}(0) + 2\sum_{k=1}^{\infty} \hat{f}(k)\, z^k\ , \qquad z \in \mathbb{D}\ . \tag{5.1}$$

Using the definition of the Fourier coefficients (2.1) in the above series of $\mathfrak{R}f$ and interchanging the order of integration and summation, one obtains

$$(\mathfrak{R}f)(r\mathrm{e}^{\mathrm{i}\theta}) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\tau})\,\frac{1 + r\mathrm{e}^{\mathrm{i}(\theta-\tau)}}{1 - r\mathrm{e}^{\mathrm{i}(\theta-\tau)}}\,\mathrm{d}\tau \tag{5.2}$$

V. Pohl, H. Boche, *Advanced Topics in System and Signal Theory*, Foundations in Signal Processing, Communications and Networking 4,
DOI 10.1007/978-3-642-03639-2_5, © Springer-Verlag Berlin Heidelberg 2009

with $0 \le r < 1$ and $-\pi \le \theta < \pi$. Taking the modulus of (5.2) and applying the triangle inequality one obtains

$$|(\mathfrak{R}f)(r\mathrm{e}^{\mathrm{i}\theta})| \le \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(\mathrm{e}^{\mathrm{i}\tau})|\,\frac{1+r}{1-r}\,\mathrm{d}\tau = \frac{1+r}{1-r}\,\|f\|_1\,. \tag{5.3}$$

This shows that (5.1), and equivalently (5.2), defines for every $f \in L^1$ a function $(\mathfrak{R}f)(z)$ in the unit disk $z \in \mathbb{D}$. Moreover (5.1) shows that $(\mathfrak{R}f)(z)$ is analytic at every $z \in \mathbb{D}$. The mapping $\mathfrak{R} : L^1(\mathbb{T}) \to H(\mathbb{D})$ is called the *Herglotz-Riesz transform.*

The kernel

$$\mathcal{H}_r(\tau) := \frac{1 + r\mathrm{e}^{\mathrm{i}\tau}}{1 - r\mathrm{e}^{\mathrm{i}\tau}}$$

of the Herglotz-Riesz transform (5.2) is a complex function of r and τ. It can be written as $\mathcal{H}_r(\tau) = \mathcal{P}_r(\tau) + \mathrm{i}\,\mathcal{Q}_r(\tau)$ wherein the real part $\mathcal{P}_r(\tau)$ is known as the *Poisson kernel* and the imaginary part $\mathcal{Q}_r(\tau)$ is called the *conjugate Poisson kernel.* It is easily verified that they are given by

$$\mathcal{P}_r(\tau) := \frac{1-r^2}{1-2r\cos\tau + r^2} \qquad \text{and} \qquad \mathcal{Q}_r(\tau) := \frac{2\,r\,\sin\tau}{1-2\,r\,\cos\tau + r^2} \tag{5.4}$$

respectively. With these one defines the *Poisson integral*

$$(\mathfrak{P}f)(r\mathrm{e}^{\mathrm{i}\theta}) := \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\tau})\,\mathcal{P}_r(\theta-\tau)\,\mathrm{d}\tau \tag{5.5}$$

and the *conjugate Poisson integral*

$$(\mathfrak{Q}f)(r\mathrm{e}^{\mathrm{i}\theta}) := \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\tau})\,\mathcal{Q}_r(\theta-\tau)\,\mathrm{d}\tau\,. \tag{5.6}$$

Consequently, the Herglotz-Riesz transform can be written as

$$(\mathfrak{R}f)(r\mathrm{e}^{\mathrm{i}\theta}) = (\mathfrak{P}f)(r\mathrm{e}^{\mathrm{i}\theta}) + \mathrm{i}\,(\mathfrak{Q}f)(r\mathrm{e}^{\mathrm{i}\theta})\,. \tag{5.7}$$

Since $\mathfrak{R}f$ is an analytic function in $\mathbb{D}$, the Poisson and the conjugate Poisson integral $\mathfrak{P}f$ and $\mathfrak{Q}f$ are harmonic functions in $\mathbb{D}$ and $\mathfrak{Q}f$ is the harmonic conjugate of $\mathfrak{P}f$.

Given $f \in L^1$, it follows from (5.3) that $(\mathfrak{R}f)(z)$, $(\mathfrak{P}f)(z)$, and $(\mathfrak{Q}f)(z)$ are bounded at every $z \in \mathbb{D}$. However, these functions need not to exist on the unit circle $\mathbb{T}$, in general. The boundary behavior of $\mathfrak{R}f$, $\mathfrak{P}f$, and $\mathfrak{Q}f$ will be investigated in the following subsections separately for these three operators. More precisely, given a function $f \in \mathcal{B}$ from a certain Banach space $\mathcal{B} \subset L^1$ of functions on the unit circle, we ask whether the boundary functions $(\mathfrak{P}f)(\mathrm{e}^{\mathrm{i}\theta}) := \lim_{r\to 1}(\mathfrak{P}f)(r\mathrm{e}^{\mathrm{i}\theta})$, $(\mathfrak{Q}f)(\mathrm{e}^{\mathrm{i}\theta}) := \lim_{r\to 1}(\mathfrak{Q}f)(r\mathrm{e}^{\mathrm{i}\theta})$, and $(\mathfrak{R}f)(\mathrm{e}^{\mathrm{i}\theta}) := \lim_{r\to 1}(\mathfrak{R}f)(r\mathrm{e}^{\mathrm{i}\theta})$ exist and whether they belong again to $\mathcal{B}$.

5.2 The Poisson Integral

Assume at the moment that $f \in L^2$ is a real valued function. Then its Fourier series and the Herglotz-Riesz transform (5.1) converge (in L^2). Moreover, since f is real, its Fourier coefficients satisfy $\hat{f}(-k) = \overline{\hat{f}(k)}$ for all $k \in \mathbb{N}$. Therefore, it follows immediately from the definition of the Herglotz-Riesz transform that the Poisson integral $(\mathfrak{P} f)(r\mathrm{e}^{\mathrm{i}\theta})$ converges to f for $r \to 1$. We will first show that this behavior of the Poisson integral also holds in a much wider sense. As a preparation, we note that the Poisson kernel $\mathcal{P}_r$ is an *approximate identity*, which means that:

(a) $\mathcal{P}_r(\tau) > 0$ for all $\tau \in [-\pi, \pi)$ and $0 \le r < 1$

(b) $\displaystyle\frac{1}{2\pi}\int_{-\pi}^{\pi} \mathcal{P}_r(\tau)\,\mathrm{d}\tau = 1$ for all $0 \le r < 1$

(c) $\displaystyle\lim_{r \to 1} \mathcal{P}_r(\tau) = 0$ for all $0 < |\tau| \le \pi$

Moreover, $\mathcal{P}_r$ is a continuous function for all $0 \le r < 1$, and even, i.e. $\mathcal{P}_r(-\tau) = \mathcal{P}_r(\tau)$. Properties (a) and (b) are easily verified, whereas property (c) results from the inequality $\cos(\tau) \le 1 - \frac{2}{\pi^2}\tau^2$ for all $\tau \in [-\pi, \pi]$. Then, one gets for an arbitrary $\epsilon > 0$ that

$$\mathcal{P}_r(\tau) \le \frac{1 - r^2}{(1-r)^2 + \frac{4r}{\pi^2}\,\epsilon^2} \qquad \text{for all } \epsilon \le |\tau| \le \pi$$

where the right hand side converges to zero as $r \to 1$.

The first theorem investigates the boundary behavior of the Poisson integral for all continuous functions on the unit circle.

Theorem 5.1. *Let $f \in \mathcal{C}(\mathbb{T})$ be arbitrary and set $F_r(\mathrm{e}^{\mathrm{i}\theta}) := (\mathfrak{P} f)(r\mathrm{e}^{\mathrm{i}\theta})$. Then*

$$\|F_r\|_\infty \le \|f\|_\infty \qquad \textit{for all } 0 \le r < 1 \tag{5.8}$$

and

$$\lim_{r \to 1} \|F_r - f\|_\infty = 0\,. \tag{5.9}$$

Proof. The first statement (5.8) follows from Property (b) of the Poisson kernel, because by the definition of the Poisson integral one has

$$\left|F_r(\mathrm{e}^{\mathrm{i}\theta})\right| \le \tfrac{1}{2\pi}\int_{-\pi}^{\pi} |f(\mathrm{e}^{\mathrm{i}\theta})|\,\mathcal{P}_r(\theta-\tau)\,\mathrm{d}\tau \le \|f\|_\infty\,\tfrac{1}{2\pi}\int_{-\pi}^{\pi} \mathcal{P}_r(\theta-\tau)\,\mathrm{d}\tau = \|f\|_\infty$$

for all $\theta \in [-\pi, \pi]$.

The second statement, i.e. the uniform convergence of the Poisson integral, can be proven exactly as it was done for the Fejér series (Theorem 2.4) using that the Poisson kernel $\mathcal{P}_r$ is an approximate identity. Therefore, this part of the proof is left as an exercise. □

Let $f \in \mathcal{C}(\mathbb{T})$ be an arbitrary continuous function on the unit circle $\mathbb{T}$ and define the function $F(r\mathrm{e}^{\mathrm{i}\theta}) := (\mathfrak{P}f)(r\mathrm{e}^{\mathrm{i}\theta})$ in the unit disk $\mathbb{D}$, i.e for all $0 \le r < 1$ and $\theta \in [-\pi, \pi)$. This function F is a harmonic function in the open unit disk $\mathbb{D}$, and the second statement of Theorem 5.1 implies that

$$\lim_{r \to 1} F(r\mathrm{e}^{\mathrm{i}\theta}) = f(\mathrm{e}^{\mathrm{i}\theta}) \qquad \text{for each } \theta \in [-\pi, \pi)\ .$$

Therefore, F is called the *harmonic extension* of f into the unit disk $\mathbb{D}$, and we can define the function

$$f(r\mathrm{e}^{\mathrm{i}\theta}) := \begin{cases} f(\mathrm{e}^{\mathrm{i}\theta}) & \text{for } r = 1 \\ (\mathfrak{P}f)(r\mathrm{e}^{\mathrm{i}\theta}) & \text{for } 0 \le r < 1 \end{cases}$$

which is harmonic in the open unit disk $\mathbb{D}$ and continuous in the closed unit disk $\overline{\mathbb{D}}$. The first statement (5.8) of Theorem 5.1 shows that the Poisson integral $\mathfrak{P}$ is a bounded operator on $\mathcal{C}(\mathbb{T})$ and that the maximum modulus of the extended function f is attained on the unit circle $\mathbb{T}$.

Remark 5.2. The Poisson integral plays a very important role in a variety of areas of engineering and physics since the above theorem provides a solution to the important boundary value problem: Let Ω be an open set in the complex plane with the boundary $\partial\Omega$. One looks for a function $F(z)$ with $z = x + \mathrm{i}\,y \in \Omega$ which satisfies Laplace's equation

$$\Delta F = \frac{\partial^2 F}{\partial x^2} + \frac{\partial^2 F}{\partial y^2} = 0$$

at every point $z \in \Omega$ and whose boundary values on $\partial\Omega$ are equal to a given continuous function $f \in \mathcal{C}(\partial\Omega)$. For $\Omega = \mathbb{D}$, the solution to this problem is given by Theorem 5.1. For arbitrary sets Ω, a solution can also be obtained from Theorem 5.1 using a conformal mapping from Ω onto $\mathbb{D}$.

Assume now that the given boundary function f is not continuous but belongs to L^p for $1 \le p \le \infty$. Then the Poisson integral $F(z) = (\mathfrak{P}f)(z)$ still defines a harmonic and continuous function for all $z \in \mathbb{D}$. But since f is not continuous, the Poisson integral $F = \mathfrak{P}f$ does not converge pointwise to the boundary values f, in general. Nevertheless, the pointwise convergence in the case of $f \in \mathcal{C}(\mathbb{T})$ was just a consequence of the norm-convergence (5.9). The next theorem will show that the Poisson integral $(\mathfrak{P}f)(z)$ shows the same behavior on L^p with $1 \le p \le \infty$. Namely, it converges to f in L^p as $|z| \to 1$.

Theorem 5.3. *Let $f \in L^p$ with $1 \le p \le \infty$ and set $F_r(\mathrm{e}^{\mathrm{i}\theta}) := (\mathfrak{P}f)(r\mathrm{e}^{\mathrm{i}\theta})$, then*

$$\|F_r\|_p \le \|f\|_p \qquad \text{for all } 0 \le r < 1\,. \tag{5.10}$$

Moreover, if $1 \le p < \infty$ then

$$\lim_{r \to 1} \|F_r - f\|_p = 0\,. \tag{5.11}$$

Proof. For a fixed r and θ, the Poisson kernel defines a positive measure $\mathrm{d}\mu_{r,\theta}(\tau) := \frac{1}{2\pi}\,\mathcal{P}_r(\theta-\tau)\,\mathrm{d}\tau$ on $[-\pi,\pi)$ with $\int_{-\pi}^{\pi}\mathrm{d}\mu_{r,\theta}(\tau) = 1$ by the second property of the Poisson kernel. Therewith, one sees that

$$|F_r(\mathrm{e}^{\mathrm{i}\theta})| \;\le\; \int_{-\pi}^{\pi}|f(\mathrm{e}^{\mathrm{i}\tau})|\,\mathrm{d}\mu_{r,\theta}(\tau) \;\le\; \Big(\int_{-\pi}^{\pi}|f(\mathrm{e}^{\mathrm{i}\tau})|^p\,\mathrm{d}\mu_{r,\theta}(\tau)\Big)^{1/p}$$

where the second inequality was obtained by applying Hölder's inequality with an arbitrary $1 \le p < \infty$. One obtains

$$\begin{aligned}\|F_r\|_p^p &= \frac{1}{2\pi}\int_{-\pi}^{\pi}|F_r(\mathrm{e}^{\mathrm{i}\theta})|^p\,\mathrm{d}\theta \le \frac{1}{2\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}|f(\mathrm{e}^{\mathrm{i}\tau})|^p\,\mathrm{d}\mu_{r,\theta}(\tau)\,\mathrm{d}\theta\\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi}|f(\mathrm{e}^{\mathrm{i}\theta})|^p\,\mathrm{d}\theta = \|f\|_p^p\end{aligned}$$

using Fubini's theorem (to exchange the order of integration) and $\int_{-\pi}^{\pi}\mathrm{d}\mu_{r,\theta}(\tau) = 1$. This proves (5.10) for $p < \infty$. The statement for $p = \infty$ follows directly from

$$|F_r(\mathrm{e}^{\mathrm{i}\theta})| \;\le\; \int_{-\pi}^{\pi}|f(\mathrm{e}^{\mathrm{i}\tau})|\,\mathrm{d}\mu_{r,\theta}(\tau) \;\le\; \|f\|_\infty\int_{-\pi}^{\pi}\mathrm{d}\mu_{r,\theta}(\tau) = \|f\|_\infty\,.$$

To prove (5.11), we use that $\mathcal{C}(\mathbb{T})$ is dense in L^p and apply Theorem 5.1. Let $\epsilon > 0$ arbitrary and choose $g \in \mathcal{C}(\mathbb{T})$ such that $\|g-f\|_p < \epsilon/3$ and set $G_r(\mathrm{e}^{\mathrm{i}\theta}) := (\mathfrak{P}g)(r\mathrm{e}^{\mathrm{i}\theta})$. Then

$$\|F_r - f\|_p \le \|F_r - G_r\|_p + \|G_r - g\|_p + \|g - f\|_p$$

Since $F_r(\mathrm{e}^{\mathrm{i}\theta}) - G_r(\mathrm{e}^{\mathrm{i}\theta}) = (\mathfrak{P}[f-g])(r\mathrm{e}^{\mathrm{i}\theta})$, it follows from (5.10) that $\|F_r - G_r\|_p \le \|f-g\|_p < \epsilon/3$. Moreover $\|G_r-g\|_p \le \|G_r-g\|_\infty$ and by Theorem 5.1, there exists an $R_0 < 1$ such that $\|G_r - g\|_\infty < \epsilon/3$ for all $r > R_0$. Altogether this shows that $\|F_r - f\|_p < \epsilon$ for all $r > R_0$ which proves (5.11). □

Finally, we consider the Poisson integral on spaces $\mathcal{C}_\omega(\mathbb{T})$ of smooth functions on the unit circle, as introduced in Section 1.3. Here, it will depend on the majorant ω whether the Poisson integral is bounded or not. The following theorem gives necessary and sufficient conditions on the majorant ω such that $\mathfrak{P} : \mathcal{C}_\omega(\mathbb{T}) \to \mathcal{C}_\omega(\overline{\mathbb{D}})$ is bounded.

Theorem 5.4. *Let ω be a majorant. Then the Poisson integral $\mathfrak{P} : \mathcal{C}_\omega(\mathbb{T}) \to \mathcal{C}_\omega(\overline{\mathbb{D}})$ is bounded if and only if ω is a weak regular majorant of type 2. In other words, there exists a constant $C = C(\omega)$ such that*

$$\|\mathfrak{P}f\|_{\mathcal{C}_\omega(\overline{\mathbb{D}})} \le C(\omega)\;\|f\|_{\mathcal{C}_\omega(\mathbb{T})} \tag{5.12}$$

for all $f \in \mathcal{C}_\omega(\mathbb{T})$ if and only if ω is a weak regular majorant of type 2.

This theorem states that if $f \in \mathcal{C}_\omega(\overline{\mathbb{D}})$ is given with a regular majorant of type 2 then the modulus of continuity of the Poisson integral $\mathfrak{P}f$ is always upper bounded by the same majorant ω in the closed unit disk $\overline{\mathbb{D}}$. At all points $z \in \mathbb{D}$, the Poisson integral $(\mathfrak{P}f)(z)$ is holomorphic. It easily follows (see proof) that $\mathfrak{P}f$ belongs to all spaces $\mathcal{C}_\omega(\mathbb{D})$ with an arbitrary majorant ω. Thus, the non-trivial statement of this theorem is that the boundary function $(\mathfrak{P}f)(\mathrm{e}^{\mathrm{i}\theta})$, $\theta \in [-\pi, \pi)$ again belongs to $\mathcal{C}_\omega(\mathbb{T})$ as long as ω is weak regular of type 2. For the proof of this theorem, we will need the following auxiliary lemma.

Lemma 5.5. *Let ω be a weak regular majorant of type 2. Then there exists a constant C such that*

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\omega(|\mathrm{e}^{\mathrm{i}\tau}-\mathrm{e}^{\mathrm{i}t}|)(1-r^2)}{1-2r\cos(\tau-t)+r^2}\,\mathrm{d}\tau \le C\,\frac{\omega(1-r)}{r} \tag{5.13}$$

for $0 < r \le 1$ and for all $t \in [-\pi, \pi)$.

Proof. It is not hard to see that $|\mathrm{e}^{\mathrm{i}\tau} - \mathrm{e}^{\mathrm{i}t}| \le |\tau - t|$ and that there exists a constant $C_1 = 4/\pi^2$ such that $1-2r\cos(\tau-t)+r^2 \ge (1-r)^2 + C_1\, r\, (\tau-t)^2$. Therewith, the following upper bound for the left-hand side of (5.13), denoted by $L(t)$, is obtained

$$\begin{aligned} L(t) &\le \frac{1-r^2}{2\,\pi}\int_{-\pi}^{\pi}\frac{\omega(|\tau-t|)}{(1-r)^2+C_1\,(\tau-t)^2}\,\mathrm{d}\tau \\ &\le \frac{1-r^2}{\pi}\int_{0}^{\pi}\frac{\omega(s)}{(1-r)^2+C_1\,s^2}\,\mathrm{d}s \end{aligned} \tag{5.14}$$

for all $t \in [-\pi, \pi)$. The last inequality was obtained by the substitution $s = \tau - t$ and using that the integrand is a positive function. Now, the right hand side of (5.14) is split up into a sum of an integral from 0 to $1-r$, denoted by L_1, and an integral from $1-r$ to π, denoted by L_2. To obtain an upper bound for L_1, it is used that $\omega(s)$ is a monotone increasing function and that $C_1\, s^2 \ge 0$:

$$L_1 \le \frac{(1-r^2)\,\omega(1-r)}{\pi\,(1-r)^2}\int_0^{1-r}\mathrm{d}s \;\le\; \frac{(1+r)\,r}{\pi}\,\frac{\omega(1-r)}{r} \le \frac{2}{\pi}\,\frac{\omega(1-r)}{r}$$

Similarly, the following upper bound for L_2 is obtained

$$L_2 \le \frac{1+r}{\pi\, C_1\, r}\,(1-r)\int_{1-r}^{\pi}\frac{\omega(s)}{s^2}\,\mathrm{d}s \le \frac{2}{\pi}\,\frac{C_2}{C_1}\,\frac{\omega(1-r)}{r}$$

using that ω is a weak regular majorant of type 2. These two upper bounds together with (5.14) give statement (5.13) of the lemma. □

Proof (Theorem 5.4). 1) Sufficiency: In the first part, it is shown that if ω is weak regular of type 2 then $\mathfrak{P}$ is continuous. To this end, we have show that there exists a constant $C(\omega)$, which depends only on the majorant ω such that

$|(\mathfrak{P}f)(z_1) - (\mathfrak{P}f)(z_2)| \leq C(\omega)\,\omega(|z_1 - z_2|)$ for all z_1, z_2 in the closed unit disk $\overline{\mathbb{D}}$ if and only if ω is weak regular of type 2.

First, we note that for all z_1, z_2 strictly inside the unit disk such a constant $C(\omega)$ exists for every arbitrary majorant ω. This follows easily from the fact that $(\mathfrak{P}f)(z)$ is harmonic in $\mathbb{D}$. Indeed, we know that $\mathfrak{P}f$ is the real part of the Herglotz-Riesz transform $\mathfrak{R}f$ (cf. (5.7)), which is an analytic function inside the unit disk. Moreover, by the maximum modulus principle for analytic function $|(\mathfrak{R}f)(z)| \leq \|f\|_\infty \leq C_1(\omega)\,\|f\|_{\mathcal{C}_\omega(\mathbb{T})}$ for all $z \in \mathbb{D}$, where the last inequality follows from the properties of functions in $\mathcal{C}_\omega(\mathbb{T})$ with a certain constant $C_1(\omega)$, which depends only on the majorant ω (cf. Section 1.3). By a simple application of Schwarz Lemma (see e.g. [70, § 12.5]), the modulus of the the first derivative of $\mathfrak{R}f$ can be upper bounded as follows

$$|(\mathfrak{R}f)'(z)| \leq \frac{2\,\|f\|_\infty}{1-|z|^2} \leq \frac{2\,C_1(\omega)}{1-|z|^2}\,\|f\|_{\mathcal{C}_\omega(\mathbb{T})}\,, \qquad \text{for all } z \in \mathbb{D}\,.$$

Consequently, for all points z_1, z_2 with $|z_1|, |z_2| \leq 1/2$, we obtain

$$|(\mathfrak{R}f)(z_1) - (\mathfrak{R}f)(z_2)| \leq 8\,C_1(\omega)\,\|f\|_{\mathcal{C}_\omega(\mathbb{T})}\,|z_1 - z_2|\,.$$

Finally, we use that $\omega(\tau)/\tau$ is non-increasing and that $|z_1 - z_2| \leq 1$. This implies that $\omega(|z_1 - z_2|)/\|z_1 - z_2| \geq \omega(1)$. Noting again that $\mathfrak{P}f$ is the real part of $\mathfrak{R}f$, we get

$$|(\mathfrak{P}f)(z_1) - (\mathfrak{P}f)(z_2)| \leq \frac{8\,C_1(\omega)}{\omega(1)}\,\|f\|_{\mathcal{C}_\omega(\mathbb{T})}\,\omega(|z_1 - z_2|)$$

for all z_1, z_2 with $|z_1|, |z_2| \leq 1/2$, and for every arbitrary majorant ω.

It remains to study the behavior at points z_1, z_2 lying close to the boundary of the unit disk. Thus, without loss of generality, we consider three point $z_1 = r_1 \mathrm{e}^{\mathrm{i}t_1}$, $z_2 = r_2 \mathrm{e}^{\mathrm{i}t_2}$, and $z = r_1 \mathrm{e}^{\mathrm{i}t_2}$ with $1/2 \leq r_1, r_2 \leq 1$ and with $t_1, t_2 \in [-\pi, \pi)$. The triangle inequality gives

$$\begin{aligned} &|(\mathfrak{P}f)\,(z_1) - (\mathfrak{P}f)\,(z_2)| \\ &\qquad \leq |(\mathfrak{P}f)\,(z_1) - (\mathfrak{P}f)\,(z)| + |(\mathfrak{P}f)\,(z) - (\mathfrak{P}f)\,(z_2)|\;. \end{aligned} \tag{5.15}$$

The first term on the right hand side becomes

$$\begin{aligned} \left|(\mathfrak{P}f)\,(r_1\mathrm{e}^{\mathrm{i}t_1}) - (\mathfrak{P}f)\,(r_1\mathrm{e}^{\mathrm{i}t_2})\right| &\leq \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|f(\mathrm{e}^{\mathrm{i}(\tau+t_1)}) - f(\mathrm{e}^{\mathrm{i}(\tau+t_2)})\right|\mathcal{P}_{r_1}(\tau)\,\mathrm{d}\tau \\ &\leq \|f\|_{\mathcal{C}_\omega(\mathbb{T})}\,\omega\left(|\mathrm{e}^{\mathrm{i}t_1} - \mathrm{e}^{\mathrm{i}t_2}|\right) \leq 2\,\|f\|_{\mathcal{C}_\omega(\mathbb{T})}\,\omega\left(|z_1 - z|\right)\;. \end{aligned} \tag{5.16}$$

Here, the last inequality follows from the obvious relation $|z_1 - z| \leq |\mathrm{e}^{\mathrm{i}t_1} - \mathrm{e}^{\mathrm{i}t_2}|$ and by using that $\omega(t)/t$ is non-increasing. Note that this result implies that the function defined by $g(\mathrm{e}^{\mathrm{i}t}) := (\mathfrak{P}f)(r_1\mathrm{e}^{\mathrm{i}t})$ is an element of $\mathcal{C}_\omega(\mathbb{T})$.
Next, the second term on the right hand side of (5.15) is investigated. Without

loss of generality, it is assumed that $r_2 \geq r_1$. First, we consider the case that z_2 lies on the unit circle, i.e. that $r_2 = 1$. Using that $\lim_{r\to 1} (\mathfrak{P}f)(r\mathrm{e}^{\mathrm{i}t_2}) = f(\mathrm{e}^{\mathrm{i}t_2})$ and that $\int_{-\pi}^{\pi} \mathcal{P}_r(\tau)\,\mathrm{d}\tau = 2\pi$ for all $0 \leq r \leq 1$, the second term on the right hand side of (5.15) becomes

$$\begin{aligned}\left|(\mathfrak{P}f)(r_1\mathrm{e}^{\mathrm{i}t_2}) - f(\mathrm{e}^{\mathrm{i}t_2})\right| &= \left|\frac{1}{2\pi}\int_{-\pi}^{\pi}\left[f(\mathrm{e}^{\mathrm{i}\tau}) - f(\mathrm{e}^{\mathrm{i}t_2})\right]\,\mathcal{P}_{r_1}(t_2-\tau)\,\mathrm{d}\tau\right| \\ &\leq \|f\|_{\mathcal{C}_\omega(\mathbb{T})}\,\frac{1}{2\pi}\int_{-\pi}^{\pi}\omega\left(|\mathrm{e}^{\mathrm{i}\tau} - \mathrm{e}^{\mathrm{i}t_2}|\right)\mathcal{P}_{r_1}(t_2-\tau)\,\mathrm{d}\tau \\ &\leq 2\,C_2\,\|f\|_{\mathcal{C}_\omega(\mathbb{T})}\,\omega\,(1-r_1) \leq 2\,C_2\,\|f\|_{\mathcal{C}_\omega(\mathbb{T})}\,\omega\,(|z-z_1|)\ .\end{aligned} \tag{5.17}$$

Here, the first inequality is a consequence of $f \in \mathcal{C}_\omega(\mathbb{T})$ whereas the second inequality follows from Lemma 5.5 with a certain constant C_2, and using that $r_1 \geq 1/2$. Now we consider the case $r_2 < 1$ and define the function $g(\mathrm{e}^{\mathrm{i}t}) := (\mathfrak{P}f)(r_2\mathrm{e}^{\mathrm{i}t})$. As a consequence of the first part of this proof, $g \in \mathcal{C}_\omega(\mathbb{T})$ with $\|g\|_{\mathcal{C}_\omega(\mathbb{T})} \leq 2\,\|f\|_{\mathcal{C}_\omega(\mathbb{T})}$. Therewith, the second term on the right hand side of (5.15) can be written as

$$\begin{aligned}\left|(\mathfrak{P}f)\,(z) - (\mathfrak{P}f)\,(z_2)\right| &= \left|(\mathfrak{P}g)\,(\tfrac{r_1}{r_2}\,\mathrm{e}^{\mathrm{i}t_2}) - (\mathfrak{P}g)\,(\mathrm{e}^{\mathrm{i}t_2})\right| \\ &\leq 4\,C_2\,\|f\|_{\mathcal{C}_\omega(\mathbb{T})}\,\omega(1-\tfrac{r_1}{r_2})\ ,\end{aligned}$$

using (5.17) for the last inequality. Since $r_2 \geq 1/2$ and because the majorant $\omega(t)$ is an increasing function, it holds that $\omega(\frac{r_2-r_1}{r_2}) \leq \omega(2\,[r_2-r_1]) \leq 2\,\omega(r_2-r_1)$, whereas the last inequality follows from the fact that $\omega(t)/t$ is non-increasing. Because $|z-z_2| = r_2 - r_1$ we therefore have

$$\left|(\mathfrak{P}f)\,(z) - (\mathfrak{P}f)\,(z_2)\right| \leq 8\,C_2\|f\|_{\mathcal{C}_\omega(\mathbb{T})}\,\omega(|z-z_2|)\ . \tag{5.18}$$

Using (5.16) and (5.18) in (5.15) gives

$$\left|(\mathfrak{P}f)\,(z_1) - (\mathfrak{P}f)\,(z_2)\right| \leq 2\,\|f\|_{\mathcal{C}_\omega(\mathbb{T})}\,\omega\,(|z_1-z|) + 8\,C_2\,\|f\|_{\mathcal{C}_\omega(\mathbb{T})}\,\omega(|z-z_2|)\ .$$

Since $|z_1-z| \leq |z_1-z_2|$ and $|z-z_2| \leq |z_1-z_2|$ and because $\omega(t)$ is an increasing function, this gives $|(\mathfrak{P}f)\,(z_1) - (\mathfrak{P}f)\,(z_2)| \leq (2+8\,C_2)\,\|f\|_{\mathcal{C}_\omega(\mathbb{T})}\,\omega(|z_1-z_2|)$, and this is what we wanted to show.

2) Necessity: It will be shown that if $\mathfrak{P} : \mathcal{C}_\omega(\mathbb{T}) \to \mathcal{C}_\omega(\mathbb{T})$ is continuous then ω is weak regular of type 2. We assume that there exists a constant C such that (5.12) holds. This implies in particular that $\mathfrak{P}f \in \mathcal{C}_\omega(\mathbb{T})$ for all $f \in \mathcal{C}_\omega(\mathbb{T})$. Next, the special function $f_1(\mathrm{e}^{\mathrm{i}\tau}) := \omega(|\tau|)$ is considered. Using the properties of a majorant, it can be verified that $f_1 \in \mathcal{C}_\omega(\mathbb{T})$ and therefore that $\mathfrak{P}f_1 \in \mathcal{C}_\omega(\mathbb{T})$. This implies that there exists a constant C_3 such that for all $0 \leq r \leq 1$ and all $-\pi \leq \tau < \pi$

$$\left|(\mathfrak{P}f_1)(r\mathrm{e}^{\mathrm{i}\tau}) - (\mathfrak{P}f_1)(\mathrm{e}^{\mathrm{i}\tau})\right| \leq C_3\,\omega(1-r)$$

holds. This last inequality holds in particular for $\tau = 0$, i.e. $|(\mathfrak{P}f_1)(r) - (\mathfrak{P}f_1)(1)| \leq C_3\,\omega(1-r)$. Because $(\mathfrak{P}f_1)(1) = f_1(1) = \omega(0) = 0$ and because f_1 is non-negative and even, this last inequality becomes

$$\frac{1}{\pi}\int_0^{\pi} f_1(\mathrm{e}^{\mathrm{i}\tau})\,\frac{1-r^2}{1-2r\cos\tau + r^2}\,\mathrm{d}\tau \leq C_3\,\omega(1-r)\;.$$

Next, we derive a lower bound for the Poisson integral on the left hand side, which will finally show that ω is weak regular of type 2. To this end, the denominator of the kernel is written as $(1-r)^2 + 2r\,[1-\cos\tau]$ and upper bounded using the inequality $1-\cos\tau \leq \tau^2/2$. The nominator of the kernel is lower bounded by $1-r \leq 1-r^2$ and by increasing the lower integration limit the value of the integral decreases since the integrand is non negative. All this gives finally

$$\frac{1-r}{\pi}\int_{1-r}^{\pi} f_1(\mathrm{e}^{\mathrm{i}\tau})\,\frac{\mathrm{d}\tau}{(1-r)^2+\tau^2} \leq C_5\,\omega(1-r)\;.$$

Since the integration variable τ is greater than $(1-r)$, it follows that $(1-r)^2 + \tau^2 \leq 2\tau^2$. Using this in the above inequality and recalling that $f_1(\mathrm{e}^{\mathrm{i}\tau}) = \omega(\tau)$ for all $\tau \geq 0$ gives

$$(1-r)\int_{1-r}^{\pi}\frac{\omega(\tau)}{\tau^2}\,\mathrm{d}\tau \leq 2\pi\,C_3\,\omega(1-r)$$

which shows that ω is weak regular of type 2. $\square$

5.3 The Conjugate Poisson Integral

The Poisson integral $(\mathfrak{P}f)(z)$ converges (in norm) to f as $|z| \to 1$ for all functions f in the spaces L^p, $\mathcal{C}(\mathbb{T})$. This result was obtained in the previous section easily from the fact that the kernel of the Poisson integral is an approximate identity. The kernel of the conjugate Poisson integral does not behave like an approximate identity. In particular, it holds that $\frac{1}{2\pi}\int_{-\pi}^{\pi}|\mathcal{Q}_r(\tau)|\,\mathrm{d}\tau = \frac{2}{\pi}\log[(r+1)/(r-1)]$, which goes to infinity as $r \to 1$. However, one can show that for every $f \in L^1$ the conjugate Poisson integral $(\mathfrak{P}f)(r\mathrm{e}^{\mathrm{i}t})$ has non-tangential limits

$$\widetilde{f}(\mathrm{e}^{\mathrm{i}t}) = \lim_{r\to 1}(\mathfrak{Q}f)(r\mathrm{e}^{\mathrm{i}t}) \quad \text{for almost all } t \in [-\pi,\pi)\;.$$

The linear mapping $\mathfrak{H} : f \mapsto \widetilde{f}$ is called the *conjugate mapping* or the *Hilbert transform* of f. Starting with (5.6), one can show (see e.g. [41, Chapter III]) that $\widetilde{f}$ is given by the principal value integral

$$\widetilde{f}(\mathrm{e}^{\mathrm{i}t}) = (\mathfrak{H}f)(\mathrm{e}^{\mathrm{i}t}) := \lim_{\epsilon\to 0}\frac{1}{2\pi}\int_{\epsilon<|\tau|\leq\pi}\frac{f(\mathrm{e}^{\mathrm{i}(t+\tau)})}{\tan(\tau/2)}\,\mathrm{d}\tau \qquad (5.19)$$

for almost all $t \in [-\pi, \pi)$. Since the kernel of the Hilbert transform

$$\mathcal{H}(\tau) = \frac{1}{\tan(\tau/2)} = \frac{\sin\tau}{1-\cos\tau} \tag{5.20}$$

has a singularity at $\tau = 0$, the *truncated Hilbert transform*

$$(\mathfrak{H}_\epsilon f)(\mathrm{e}^{\mathrm{i}t}) := \frac{1}{2\pi}\int_{\epsilon<|\tau|\le\pi} \frac{f(\mathrm{e}^{\mathrm{i}(\tau+t)})}{\tan(\tau/2)}\,\mathrm{d}\tau \tag{5.21}$$

is used in the following to investigate the existence and boundedness of the Hilbert transform. Here, the integration over the singularity is left out, and the Hilbert transform is obtained by letting $\epsilon \to 0$, i.e.

$$\widetilde{f}(\mathrm{e}^{\mathrm{i}t}) = (\mathfrak{H}f)(\mathrm{e}^{\mathrm{i}t}) = \lim_{\epsilon\to 0}(\mathfrak{H}_\epsilon f)(\mathrm{e}^{\mathrm{i}t}) \ .$$

5.3.1 Boundedness of the Hilbert transform

As in the case of the Poisson integral, we consider the boundedness behavior of the Hilbert transform on the spaces L^p, $\mathcal{C}(\mathbb{T})$, and $\mathcal{C}_\omega(\mathbb{T})$ and start again with L^p. Suppose that $f \in L^p$ for some $1 \le p \le \infty$. Does it follow that the conjugate function $\widetilde{f} = \mathfrak{H}f$ belongs also to L^p? The answer is affirmative if $1 < p < \infty$.

Theorem 5.6 (M. Riesz). *Let $1 < p < \infty$. Then there exists a constant $C(p)$, which depends only on p, such that*

$$\|\widetilde{f}\|_p = \|\mathfrak{H}f\|_p \le C(p)\,\|f\|_p \qquad \textit{for every } f \in L^p(\mathbb{T}) \ .$$

The proof of this classical theorem, due to Marcel Riesz, can be found in many standard textbooks (e.g. [41, 48, 70, 92]). Therefore it is not presented here. However it is important to note that for $p = 1$ and $p = \infty$ the above theorem does not hold. This can be verified by considering some counter examples.

Example 5.7 (counter example for $p = \infty$). In the case of $p = \infty$, we consider the partial sums f_N and its conjugate $\widetilde{f}_N$ given by

$$f_N(\mathrm{e}^{\mathrm{i}\tau}) = -\tfrac{2}{\pi}\sum_{k=1}^{N}\tfrac{\sin(k\tau)}{k} \qquad \text{and} \qquad \widetilde{f}_N(\mathrm{e}^{\mathrm{i}\tau}) = \tfrac{2}{\pi}\sum_{k=1}^{N}\tfrac{\cos(k\tau)}{k}$$

respectively. The function f_N is just the partial Fourier series of the function

$$f(t) := \begin{cases} \frac{1}{\pi}\tau + 1\ , & \tau \in [-\pi, 0) \\ 0\ , & \tau = 0 \\ \frac{1}{\pi}\tau - 1\ , & \tau \in (0, \pi] \end{cases}$$

which is known to be uniformly bounded by $\|f_N\|_\infty \le 1.117$ for all $N \in \mathbb{N}$. Consequently, the limit $f = \lim_{N\to\infty} f_N$ is an element of L^∞. However, for the partial sums of the conjugate functions $\widetilde{f}(\mathrm{e}^{\mathrm{i}\tau})$ at $\tau = 0$ holds

$$|\widetilde{f}_N(1)| = \sum_{k=1}^{N} \frac{1}{k} \geq \int_1^{N+1} \frac{\mathrm{d}\,x}{x} = \log(N+1) \xrightarrow[N\to\infty]{} \infty\,.$$

Therefore, the conjugate function $\mathfrak{H}f = \widetilde{f} = \lim_{N\to\infty} \widetilde{f}_N$ of f cannot be an element of L^∞.

Example 5.8 (counter example for $p = 1$). For the case $p = 1$, one considers the functions f_N with the conjugate functions $\widetilde{f}_N$ given by

$$f_N(\mathrm{e}^{\mathrm{i}\tau}) = \sum_{k=2}^{N} \frac{\cos(k\tau)}{\log k} \qquad \text{and} \quad \widetilde{f}_N(\mathrm{e}^{\mathrm{i}\tau}) = \sum_{k=2}^{N} \frac{\sin(k\tau)}{\log k}$$

respectively. The series $\{1/\log k\}_{k=2}^{\infty}$ is convex and converges to zero as $k \to \infty$. Therefore, Proposition 5.15 in the appendix implies that f_N converges to a nonnegative function in L^1 as $N \to \infty$. However, the limit of $\widetilde{f}_N$ is not an integrable function. To see this, we consider the functions

$$g_N(\mathrm{e}^{\mathrm{i}\tau}) = f_N(\mathrm{e}^{\mathrm{i}\tau}) + \mathrm{i}\,\widetilde{f}_N(\mathrm{e}^{\mathrm{i}\tau}) = \sum_{k=2}^{N} \frac{\mathrm{e}^{\mathrm{i}k\tau}}{\log k}\,.$$

By this construction $g_N(z)$ is an analytic function for every $z \in \mathbb{D}$ and all $N \in \mathbb{N}$, but Hardy's inequality (Theorem 2.12) shows that the limit function $\lim_{N\to\infty} g_N$ is not an element of H^1 because

$$\sum_{k=2}^{N} \frac{1}{(k+1)\,\log k} \geq \sum_{k=3}^{N+1} \frac{1}{k\,\log k} \geq \int_3^{N+1} \frac{\mathrm{d}x}{x\,\log x} \geq \log\log(N+2) \xrightarrow[N\to\infty]{} \infty\,.$$

This implies that $\|g_N\|_1 \to \infty$ as $N \to \infty$, and since $\|g_N\|_1 \leq \|f_N\|_1 + \|\widetilde{f}_N\|_1$, it follows that $\|\widetilde{f}_N\|_1 \to \infty$ as $N \to \infty$ and $\lim_{N\to\infty} \widetilde{f}_N$ is not in L^1.

The above result of M. Riesz also does not hold for the space of continuous function. This is again verified by a counter example.

Example 5.9 (counter example for $\mathcal{C}(\mathbb{T})$). Consider the functions f_N with the corresponding conjugate functions $\widetilde{f}_N$ given by

$$f_N(\mathrm{e}^{\mathrm{i}\tau}) = \sum_{k=2}^{N} \frac{\sin(k\tau)}{k\,\log k} \qquad \text{and} \quad \widetilde{f}_N(\mathrm{e}^{\mathrm{i}\tau}) = -\sum_{k=2}^{N} \frac{\cos(k\tau)}{k\,\log k}\,.$$

The series f_N converges uniformly for all $\theta \in [-\pi, \pi)$ as $N \to \infty$ (by Theorem 1.3 in Chapter V of [92]). Therefore the function $f(\mathrm{e}^{\mathrm{i}\tau}) = \sum_{k=2}^{\infty} \frac{\sin(k\tau)}{k\,\log k}$ is a continuous function on $\mathbb{T}$.

However, the corresponding conjugate function $\widetilde{f}$ is not continuous on $\mathcal{C}(\mathbb{T})$. To see this, consider the partial sums

$$g_N(\mathrm{e}^{\mathrm{i}\tau}) = f_N(\mathrm{e}^{\mathrm{i}\tau}) + \mathrm{i}\,\widetilde{f}_N(\mathrm{e}^{\mathrm{i}\tau}) = \frac{1}{\mathrm{i}} \sum_{k=2}^{N} \frac{\mathrm{e}^{\mathrm{i}k\tau}}{k\,\log k}\,.$$

Then the same calculation as in Example 5.8 shows that

$$|g_N(1)| = \sum_{k=2}^{N} \frac{1}{k\,\log k} \xrightarrow[N\to\infty]{} \infty\,,$$

which implies that g_N, and consequently $\widetilde{f}_N$, does not converge to a continuous function as $N \to \infty$.

Thus if the function f is continuous, the Hilbert transform $\mathfrak{H}f$ need not be continuous and may even be unbounded at some points. By restricting the domain of the Hilbert transform to a subset of $\mathcal{C}(\mathbb{T})$ one may achieve that the Hilbert transform of all functions from this subset is a continuous function. In the following we consider the subspaces $\mathcal{C}_\omega(\mathbb{T}) \subset \mathcal{C}(\mathbb{T})$ of smooth functions, and the subsequent lemma gives sufficient conditions on the majorant ω such that the Hilbert transform of every $f \in \mathcal{C}_\omega(\mathbb{T})$ is a continuous function.

Lemma 5.10. *If ω is a weak regular majorant of type 1 and if $f \in \mathcal{C}_\omega(\mathbb{T})$ then the Hilbert transform $\widetilde{f}(\mathrm{e}^{\mathrm{i}t}) = (\mathfrak{H}f)(\mathrm{e}^{\mathrm{i}t})$ exists for all $t \in [-\pi, \pi)$ and is continuous.*

Proof. Let $\epsilon > 0$ and consider the truncated Hilbert transform $\mathfrak{H}_\epsilon f$ (5.21) which may be written as

$$(\mathfrak{H}_\epsilon f)(\mathrm{e}^{\mathrm{i}t}) = \frac{1}{2\pi} \int_\epsilon^\pi \frac{f(\mathrm{e}^{\mathrm{i}(t+\tau)}) - f(\mathrm{e}^{\mathrm{i}(t-\tau)})}{\tan(\tau/2)} \, \mathrm{d}\tau \, .$$

Using the assumption that $f \in \mathcal{C}_\omega(\mathbb{T})$, one obtains an upper bound for the modulus of $\mathfrak{H}_\epsilon f$ by

$$\begin{aligned} \left|(\mathfrak{H}_\epsilon f)(\mathrm{e}^{\mathrm{i}t})\right| &\le \frac{1}{2\pi} \int_\epsilon^\pi \frac{\left|f(\mathrm{e}^{\mathrm{i}(t+\tau)}) - f(\mathrm{e}^{\mathrm{i}t})\right| + \left|f(\mathrm{e}^{\mathrm{i}t}) - f(\mathrm{e}^{\mathrm{i}(t-\tau)})\right|}{\tan(\tau/2)} \, \mathrm{d}\tau \\ &\le \|f\|_{\mathcal{C}_\omega(\mathbb{T})} \, \frac{1}{\pi} \int_\epsilon^\pi \frac{\omega(|\mathrm{e}^{\mathrm{i}(t+\tau)} - \mathrm{e}^{\mathrm{i}t}|)}{\tan(\tau/2)} \, \mathrm{d}\tau \end{aligned}$$

using that $|\mathrm{e}^{\mathrm{i}\tau} - 1| \le \tau$ and that $\omega(\tau)$ is an increasing function. With $\tan(\tau/2) \ge \tau/2$ for all $0 \le \tau \le \pi$, the upper bound becomes finally

$$\left|(\mathfrak{H}_\epsilon f)(\mathrm{e}^{\mathrm{i}t})\right| \le \|f\|_{\mathcal{C}_\omega(\mathbb{T})} \, \frac{2}{\pi} \int_0^\pi \frac{\omega(\tau)}{\tau} \, \mathrm{d}\tau \; \le \|f\|_{\mathcal{C}_\omega(\mathbb{T})} \, \frac{2}{\pi} \, C\, \omega(\pi) \, .$$

The last integral always exists, since ω is a weak regular majorant of type 1. This result shows that $\left|(\mathfrak{H}_\epsilon f)(\mathrm{e}^{\mathrm{i}t})\right|$ is uniformly bounded for all t and ϵ. Therefore, $\mathfrak{H}_\epsilon f$ converges to the Hilbert transform $\mathfrak{H}f$ as $\epsilon \to 0$.

The continuity of $\mathfrak{H}f$ follows from the fact that f is Dini continuous. Let ω_f be the modulus of continuity of f. Then f is said to be *Dini continuous* if there exists an $a > 0$ such that

$$\int_0^a \frac{\omega_f(\tau)}{\tau} \, \mathrm{d}\tau < \infty \, .$$

Since $f \in \mathcal{C}_\omega(\mathbb{T})$, we have that $\omega_f(\tau) \le \|f\|_{\mathcal{C}_\omega(\mathbb{T})} \, \omega(\tau)$ (cf. Section 1.3). Therefore the Dini continuity of f follows from the weak regularity (of type 1) of the majorant ω. The proof that the Hilbert transform of every Dini continuous function is continuous omitted here. But a proof can be found in [41, Chapt. III, Theorem 1.3]. □

Thus, for every $f \in \mathcal{C}_\omega(\mathbb{T})$ with a weak regular majorant ω of type 1, the Hilbert transform $\widetilde{f} = \mathfrak{H} f$ will always exist and $\widetilde{f}$ will be a continuous function. However, the conjugate function $\widetilde{f}$ does not belong to $\mathcal{C}_\omega(\mathbb{T})$, in general, under these conditions on ω. Nevertheless, the next theorem will show that if ω is additionally weak regular of type 2, the Hilbert transform $\mathfrak{H} f$ will always belong to $\mathcal{C}_\omega(\mathbb{T})$, and $\mathfrak{H} : \mathcal{C}_\omega(\mathbb{T}) \to \mathcal{C}_\omega(\mathbb{T})$ will even be a continuous mapping.

Theorem 5.11. *Let $\mathfrak{H}$ be the Hilbert transform on the domain $\mathcal{C}_\omega(\mathbb{T})$. If ω is a regular majorant, then there exists a constant $C(\omega)$, which depends only on the majorant ω, such that*

$$\|\mathfrak{H} f\|_{\mathcal{C}_\omega(\mathbb{T})} \le C(\omega)\, \|f\|_{\mathcal{C}_\omega(\mathbb{T})} \qquad \text{for every } f \in \mathcal{C}_\omega(\mathbb{T}) \ .$$

For the proof of this theorem we need an auxiliary lemma which characterizes the smoothness of the truncated Hilbert transform (5.20).

Lemma 5.12. *Let ω be a regular majorant and let $f \in \mathcal{C}_\omega(\mathbb{T})$. Then there exists a constant C such that*

$$\left|(\mathfrak{H}_\epsilon f)\,(\mathrm{e}^{\mathrm{i}t_1}) - (\mathfrak{H}_\epsilon f)\,(\mathrm{e}^{\mathrm{i}t_2})\right| \le C\, \|f\|_{\mathcal{C}_\omega(\mathbb{T})}\, \omega(\epsilon) \ .$$

Proof. Since the kernel $\mathcal{H}(\tau)$, given in (5.20), has a singularity at $\tau = 0$. It will be advantageous to separate the integration of the Hilbert transform into a regular part and a singular part. Therefore, the kernel (5.20) of the Hilbert transform is written as

$$\frac{1}{\tan(\tau/2)} = \frac{2}{\tau} + \mathcal{K}(\tau) \quad \text{with} \quad \mathcal{K}(\tau) = \frac{\tau - 2\tan(\tau/2)}{\tau \tan(\tau/2)} \ ,$$

and it is easily verified that $\mathcal{K}(\tau)$ is continuous and bounded by $|\mathcal{K}(\tau)| < \frac{2}{\pi}$ for all $\tau \in [-\pi, \pi]$. With this separated kernel one obtains

$$\begin{aligned}\left|(\mathfrak{H}_\epsilon f)\,(\mathrm{e}^{\mathrm{i}t_1}) - (\mathfrak{H}_\epsilon f)\,(\mathrm{e}^{\mathrm{i}t_2})\right| \le &\left|\frac{1}{\pi}\int_{\epsilon<|\tau|\le\pi} \frac{f(\mathrm{e}^{\mathrm{i}(\tau+t_1)}) - f(\mathrm{e}^{\mathrm{i}(\tau+t_2)})}{\tau}\,\mathrm{d}\tau\right| + \\ &+ \left|\frac{1}{2\pi}\int_{\epsilon<|\tau|\le\pi} \left[f(\mathrm{e}^{\mathrm{i}(\tau+t_1)}) - f(\mathrm{e}^{\mathrm{i}(\tau+t_2)})\right] \mathcal{K}(\tau)\,\mathrm{d}\tau\right| \ .\end{aligned}$$

The first and the second term on the right hand side of this inequality are denoted by $|T_1|$ and $|T_2|$, respectively. For $|T_2|$ an upper bound is immediately found using that $f \in \mathcal{C}_\omega(\mathbb{T})$ and that $\mathcal{K}(\tau)$ is integrable

$$|T_2| \le \tfrac{2}{\pi}\, \|f\|_{\mathcal{C}_\omega(\mathbb{T})}\, \omega\left(|t_1 - t_2|\right) \int_0^\pi |\mathcal{K}(\tau)|\,\mathrm{d}\tau = C_{T_2}\, \|f\|_{\mathcal{C}_\omega(\mathbb{T})}\, \omega\left(|t_1 - t_2|\right) \ .$$

Now ϵ is chosen to be $\epsilon = |t_1 - t_2|/2$. Therewith the previous bound becomes

$$|T_2| \le 2\ C_{T_2}\, \|f\|_{\mathcal{C}_\omega(\mathbb{T})}\, \omega\left(\epsilon\right) \ . \tag{5.22}$$

The first term $|T_1|$ is written as the difference of two integrals. After a variable substitution in both integrals, $|T_1|$ becomes

$$|T_1| \le \left| \frac{1}{\pi} \int_{I_0} \left[f(\mathrm{e}^{\mathrm{i}\tau}) - f(\mathrm{e}^{\mathrm{i}t_1})\right] \left(\tfrac{1}{t_1-\tau} - \tfrac{1}{t_2-\tau}\right) \mathrm{d}\tau \right| + \\ + \left| \frac{1}{\pi} \int_{I_\epsilon(t_2)} \frac{f(\mathrm{e}^{\mathrm{i}\tau}) - f(\mathrm{e}^{\mathrm{i}t_1})}{t_1-\tau} \mathrm{d}\tau \right| + \left| \frac{1}{\pi} \int_{I_\epsilon(t_1)} \frac{f(\mathrm{e}^{\mathrm{i}\tau}) - f(\mathrm{e}^{\mathrm{i}t_1})}{t_2-\tau} \mathrm{d}\tau \right| .$$

The three terms on the right hand side of the last inequality are denoted by $|L_1|$, $|L_2|$ and $|L_3|$, respectively. The integration intervals in these integrals are defined as $I_\epsilon(t_i) := \{\tau : t_i - \epsilon \le \tau \le t_i + \epsilon\}$ and $I_0 = \{\tau \in [-\pi,\pi) : \tau \notin I_\epsilon(t_1), \tau \notin I_\epsilon(t_2)$. Now upper bounds are derived for all three terms separately. First, $|L_1|$ is considered. Because of the special choice for ϵ, it is $|t_1 - t_2| = 2\epsilon$ and it holds that $|\tau - t_2| \ge |\tau - t_1|/3$ for all $\tau \in I_0$. Therewith, the following upper bound for $|L_1|$ is obtained

$$|L_1| \le \frac{2\epsilon}{\pi} \int_{I_0} \frac{\left|f(\mathrm{e}^{\mathrm{i}\tau}) - f(\mathrm{e}^{\mathrm{i}t_1})\right|}{|(t_1-\tau)(t_2-\tau)|} \mathrm{d}\tau \le \tfrac{6\epsilon}{\pi} \|f\|_{\mathcal{C}_\omega(\mathbb{T})} \int_{\epsilon \le |\tau - t_1| \le \pi} \frac{\omega(|\tau - t_1|)}{|\tau - t_1|^2} \mathrm{d}\tau \tag{5.23}$$

using the assumption that $f \in \mathcal{C}_\omega(\mathbb{T})$. After the substitution $s := \tau - t_1$ this bound becomes

$$|L_1| \le \tfrac{12}{\pi} \|f\|_{\mathcal{C}_\omega(\mathbb{T})} \, \epsilon \int_\epsilon^\pi \frac{\omega(s)}{s^2} \mathrm{d}s \le \tfrac{12}{\pi} \|f\|_{\mathcal{C}_\omega(\mathbb{T})} \, C \, \omega(\epsilon) \tag{5.24}$$

using that ω is weak regular of type 2 and in particular that ω satisfies (1.13). For the term $|L_2|$ the upper bound

$$|L_2| \le \tfrac{1}{\pi} \|f\|_{\mathcal{C}_\omega(\mathbb{T})} \int_{t_2-\epsilon}^{t_2+\epsilon} \frac{\omega(|\tau - t_1|)}{|\tau - t_1|} \mathrm{d}\tau \le \tfrac{1}{\pi} \|f\|_{\mathcal{C}_\omega(\mathbb{T})} \int_\epsilon^{3\epsilon} \frac{\omega(s)}{s} \mathrm{d}s$$

is obtained using again that $f \in \mathcal{C}_\omega(\mathbb{T})$ and that $|t_1 - t_2| = 2\epsilon$. Since ω is a regular majorant, it holds that $\omega(3\epsilon)/3\epsilon \le \omega(\epsilon)/\epsilon$ and there exists a constant C such that (1.12) is fulfilled. Therewith the upper bound becomes

$$|L_2| \le \tfrac{3}{\pi} \|f\|_{\mathcal{C}_\omega(\mathbb{T})} \, C \, \omega(\epsilon) \, . \tag{5.25}$$

With similar arguments and using again that $|\tau - t_2| \ge |\tau - t_1|/3$, an upper bound for the last term $|L_3|$ is obtained

$$|L_3| \le \tfrac{6}{\pi} \|f\|_{\mathcal{C}_\omega(\mathbb{T})} \, C \, \omega(\epsilon) \, . \tag{5.26}$$

The three bounds (5.24),(5.25),(5.26) give an upper bound for $|T_1|$:

$$|T_1| \le C_{T_1} \|f\|_{\mathcal{C}_\omega(\mathbb{T})} \, \omega(\epsilon) \qquad \text{with} \qquad C_{T_1} = \tfrac{21}{\pi} C \, .$$

Together with (5.22) one obtains the desired statement with $C = C_{T_1} + 2\, C_{T_2}$.
□

Proof (Theorem 5.11). It has to be shown that there exists a constant C_1 such that for an arbitrary $f \in \mathcal{C}_\omega(\mathbb{T})$

$$|\widetilde{f}(\mathrm{e}^{\mathrm{i}t_1}) - \widetilde{f}(\mathrm{e}^{\mathrm{i}t_2})| \leq C_1 \, \|f\|_{\mathcal{C}_\omega(\mathbb{T})} \, \omega(|\mathrm{e}^{\mathrm{i}t_1} - \mathrm{e}^{\mathrm{i}t_2}|) \quad \text{for all } t_1, t_2 \in [-\pi, \pi) \ .$$

By the triangle inequality, it holds that

$$\begin{aligned} \left|\widetilde{f}(\mathrm{e}^{\mathrm{i}t_1}) - \widetilde{f}(\mathrm{e}^{\mathrm{i}t_2})\right| &\leq \left|\widetilde{f}(\mathrm{e}^{\mathrm{i}t_1}) - (\mathfrak{H}_\epsilon f)(\mathrm{e}^{\mathrm{i}t_1})\right| \\ &\quad + \left|(\mathfrak{H}_\epsilon f)(\mathrm{e}^{\mathrm{i}t_1}) - (\mathfrak{H}_\epsilon f)(\mathrm{e}^{\mathrm{i}t_2})\right| + \left|(\mathfrak{H}_\epsilon f)(\mathrm{e}^{\mathrm{i}t_2}) - \widetilde{f}(\mathrm{e}^{\mathrm{i}t_2})\right| \ . \end{aligned} \tag{5.27}$$

At the beginning, we consider the first and the third term on the right hand side of this inequality. By Lemma 5.10 the Hilbert transform $\widetilde{f}$ exists under the above conditions. This means in particular that the integral in (5.21) exists for $\epsilon \to 0$ and therefore it holds that

$$\widetilde{f}(\mathrm{e}^{\mathrm{i}t}) - (\mathfrak{H}_\epsilon f)(\mathrm{e}^{\mathrm{i}t}) = \frac{1}{2\pi} \int_{-\epsilon}^{\epsilon} \frac{f(\mathrm{e}^{\mathrm{i}(t+\tau)}) - f(\mathrm{e}^{\mathrm{i}t})}{\tan(\tau/2)} \, \mathrm{d}\tau$$

using that $\tan(\tau/2)$ is an odd function. Applying similar arguments as in the proof of Lemma 5.10, and using that ω is weak regular of type 1, one easily obtains

$$\left|\widetilde{f}(\mathrm{e}^{\mathrm{i}t}) - (\mathfrak{H}_\epsilon f)(\mathrm{e}^{\mathrm{i}t})\right| \leq \|f\|_{\mathcal{C}_\omega(\mathbb{T})} \, \frac{2}{\pi} \int_0^{\epsilon} \frac{\omega(\tau)}{\tau} \, \mathrm{d}\tau \leq \tfrac{2}{\pi} \, \|f\|_{\mathcal{C}_\omega(\mathbb{T})} \, \omega(\epsilon)$$

for all $t \in [-\pi, \pi)$. This is an upper bound for the first and the third term on the right hand side of (5.27). Moreover, by Lemma 5.12, the second term on the right hand side of (5.27) is upper bounded by $C_3 \, \|f\|_{\mathcal{C}_\omega(\mathbb{T})} \, \omega(\epsilon)$ with a certain constant C_3. By choosing $\epsilon = |\mathrm{e}^{\mathrm{i}t_1} - \mathrm{e}^{\mathrm{i}t_2}|/2$, (5.27) becomes

$$\left|\widetilde{f}(\mathrm{e}^{\mathrm{i}t_1}) - \widetilde{f}(\mathrm{e}^{\mathrm{i}t_2})\right| \leq (C_2 + C_3 + C_2) \ \|f\|_{\mathcal{C}_\omega(\mathbb{T})} \, \omega(\epsilon) \ .$$

This is what we wanted to show. □

5.3.2 Boundedness of the Conjugate Poisson Integral

To every function f on the unit circle $\mathbb{T}$ the Poisson integral $\mathfrak{P}f$ defines a harmonic function in $\mathbb{D}$. If f belongs to $\mathcal{C}(\mathbb{T})$, L^p with $1 \leq p \leq \infty$, or $\mathcal{C}_\omega(\mathbb{T})$ with a weak regular majorant of type 2, then, as it was shown in Section 5.2, the Poisson integral $(\mathfrak{P}f)(r\mathrm{e}^{\mathrm{i}t})$ converges to the boundary function f as $r \to 1$ in the norm of the actual Banach space. Moreover, the conjugate Poisson integral $\mathfrak{Q}f$ is the harmonic conjugate of $\mathfrak{P}f$ in the unit disk. Also this harmonic function is uniquely defined by the Poisson integral of its boundary values on the unit circle, and since the Hilbert transform was defined as the limit of $(\mathfrak{Q}f)(r\mathrm{e}^{\mathrm{i}\theta})$ as $r \to 1$, the conjugate Poisson integral of f may be written as the Poisson integral $\mathfrak{P}$ of the conjugate function $\widetilde{f} = \mathfrak{H}f$, i.e.

$$(\mathfrak{Q}f)(z) = (\mathfrak{P}\widetilde{f})(z) = (\mathfrak{P}\mathfrak{H}f)(z) \,, \qquad z \in \mathbb{D} \,.$$

Since we already characterized the boundedness behavior of the Poisson integral $\mathfrak{P}$ and of the Hilbert transform $\mathfrak{H}$, we get immediately the following two results on the boundedness of the conjugate Poisson integral. We start again with the classical result on L^p-spaces:

Theorem 5.13. *Let $f \in L^p$ with $1 < p < \infty$ and set $\widetilde{F}_r(\mathrm{e}^{\mathrm{i}\theta}) := (\mathfrak{Q}f)(r\mathrm{e}^{\mathrm{i}\theta})$. Then there exists a constant $C(p)$ which only depends on p such that*

$$\|\widetilde{F}_r\|_p \leq C(p)\,\|f\|_p \qquad \textit{for all } 0 \leq r < 1 \,.$$

Proof. Since $\widetilde{F}_r(\mathrm{e}^{\mathrm{i}\theta}) = (\mathfrak{P}\widetilde{f})(r\mathrm{e}^{\mathrm{i}\theta})$, Theorem 5.3 and 5.6 imply that

$$\|\widetilde{F}_r\|_p \leq \|\widetilde{f}\|_p \leq C(p)\,\|f\|_p$$

for every $f \in L^p$ and all $0 \leq r < 1$. □

Since the Hilbert transform is unbounded in the case of $p = 1$ and $p = \infty$ also the conjugate Poisson integral is unbounded in these cases. For the same reason, $\mathfrak{Q}$ is an unbounded operator on $\mathcal{C}(\mathbb{T})$. However, for the subspace $\mathcal{C}_\omega(\mathbb{T})$ of smooth functions on the unit circle, we have sufficient conditions for the boundedness of the conjugate Poisson integral.

Corollary 5.14. *Let ω be a regular majorant. Then the conjugate Poisson integral $\mathfrak{Q} : \mathcal{C}_\omega(\mathbb{T}) \to \mathcal{C}_\omega(\overline{\mathbb{D}})$ is bounded, i.e. there exists a constant $C = C(\omega)$ such that*

$$\|\mathfrak{Q}f\|_{\mathcal{C}_\omega(\overline{\mathbb{D}})} \leq C(\omega)\;\|f\|_{\mathcal{C}_\omega(\mathbb{T})} \qquad \textit{for all } f \in \mathcal{C}_\omega(\mathbb{T})\,. \tag{5.28}$$

Proof. The statement follows from Theorem 5.4 and 5.11. Because if ω is weak regular of type 1 and type 2 then there exist two constants $C_4(\omega)$ and $C_5(\omega)$ which depend only on the majorant ω such that

$$\|\mathfrak{Q}f\|_{\mathcal{C}_\omega(\overline{\mathbb{D}})} = \|\mathfrak{P}\widetilde{f}\|_{\mathcal{C}_\omega(\overline{\mathbb{D}})} \leq C_4(\omega)\,\|\widetilde{f}\|_{\mathcal{C}_\omega(\mathbb{T})} \leq C_4(\omega)\,C_5(\omega)\,\|f\|_{\mathcal{C}_\omega(\mathbb{T})}$$

for all $f \in \mathcal{C}_\omega(\mathbb{T})$. □

Thus, if ω is weak regular of type 1 and 2, the conjugate Poisson integral is bounded on $\mathcal{C}_\omega(\mathbb{T})$. Thereby, the weak regularity of type 2 of the majorant ω guarantees the boundedness of the Poisson integral, whereas the weak regularity of type 1 is additionally necessary to guarantee the boundedness of the Hilbert transform.

Notes

Textbooks related to this section include [1, 41, 48, 70, 92], among many others. Some of the results concerning the $\mathcal{C}_\omega$ spaces appeared in [20].

Appendix – An Auxiliary Result

Proposition 5.15. *Let $\{a(k)\}_{k=0}^{\infty}$ be a real and convex sequence with $\lim_{k\to\infty} a(k) = 0$. Then the series*

$$\frac{a_0}{2} + \sum_{k=1}^{\infty} a(k)\,\cos(k\tau) \tag{5.29}$$

converges for all $\tau \in [-\pi, \pi)$ except $\tau = 0$ to a nonnegative function $\phi \in L^1$ with Fourier series (5.29)*.*

Remark 5.16. The convexity of the series $\{a(k)\}_{k=0}^{\infty}$ means that for all k the following always holds

$$a(k+1) \le \tfrac{1}{2}\left[a(k) + a(k+2)\right] \,.$$

Proof. Let $\Delta a_k = a(k+1) - a(k)$ and $\Delta^2 a_k = \Delta(\Delta a_k)$. Then the convexity of $\{a(k)\}$ implies that $\Delta^2 a_k = \Delta a_{k+1} - \Delta a_k \ge 0$ which shows that Δa_k is a monotone increasing sequence. Since $a(k) \to 0$ also $\Delta a_k \to 0$. Therefore $\Delta a_k \le 0$ for all k.

Let $S_N(\tau) = \frac{a_0}{2} + \sum_{k=1}^{N} a(k)\,\cos(k\tau)$ be the N-th partial sum of (5.29). Two successive summations by parts give

$$S_N(\tau) = \frac{1}{2}\sum_{k=0}^{N-2} \Delta^2 a_k\,(k+1)\,\mathcal{F}_n(\tau) - \tfrac{1}{2}\,\Delta a_{N-1}\,N\,\mathcal{F}_{N-1}(\tau) + \tfrac{1}{2}\,a(N)\,\mathcal{D}_N(\tau) \tag{5.30}$$

in which $\mathcal{D}_N$ and $\mathcal{F}_N$ are the Dirichlet and Fejér kernel given by (2.10) and (2.16), respectively. Let $\epsilon > 0$ arbitrary and set $I_\epsilon := (-\epsilon, \epsilon)$. The $\mathcal{D}_N(\tau)$ as well as $N\,\mathcal{F}_{N-1}(\tau)$ are obviously uniformly bounded for every $\tau \notin I_\epsilon$, and since $a(k) \to 0$ as well as $\Delta a_k \to 0$ as $k \to \infty$, the last two terms in (5.30) vanish as $N \to \infty$. Consequently $S_N(\tau)$ tends to

$$s(\tau) = \tfrac{1}{2}\textstyle\sum_{k=0}^{\infty} \Delta^2 a_k\,(k+1)\,\mathcal{F}_n(\tau)\,.$$

Since $(k+1)\,\mathcal{F}_n(\tau)$ is uniformly bounded for every $\tau \notin I_\epsilon$ and since $\sum_{k=0}^{N} \Delta^2 a_k = \Delta a_{N+1} - \Delta a_0 \to -\Delta a_0 \ge 0$, the series of $s(\tau)$ converges uniformly for every $\tau \notin I_\epsilon$. Therefore, $s(\tau)$ represents a non-negative continuous function for every $\tau \ne 0$. Moreover, s is integrable on $[-\pi, \pi]$, because

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |s(\tau)|\,\mathrm{d}\tau = \frac{1}{4\pi}\sum_{k=0}^{\infty} \Delta^2 a_k\,(k+1)\int_{-\pi}^{\pi} \mathcal{F}_n(\tau)\,\mathrm{d}\tau = \frac{1}{2}\sum_{k=0}^{\infty} \Delta^2 a_k\,(k+1)$$

using the approximate identity property of the Fejér kernel (cf. Proposition 2.3), and it remains to show that the last sum converges. To this end, we consider the equation

$$\begin{aligned} a(n+1) - a(0) &= \Delta a_0 + \Delta a_1 + \cdots + \Delta a_n \\ &= \sum_{k=0}^{n+1} (n+1) \Delta^2 a_k + (n+1)\, \Delta a_n \end{aligned}$$

where the second line was obtained by partial summation. Now we let $n \to \infty$ and obtain $\sum_{k=0}^{n+1}(n+1)\Delta^2 a_k = a(0) < \infty$ using that $a(k) \to 0$, that $\Delta a_n \to 0$, and that $n\Delta a_n \to 0$. The later statement follows from Abel's classical theorem and from the fact that the sequence Δa_k is non-positive and monotone increasing. □

6

Causal Projections

Let $\mathcal{B} \subset L^1$ be a Banach space. Then to every $f \in \mathcal{B}$ the Fourier coefficients $\{\hat{f}(k)\}_{k=-\infty}^{\infty}$ exist, and one can define the causal subspace

$$\mathcal{B}_+ := \{f \in \mathcal{B} : \hat{f}(k) = 0 \text{ for all } k < 0\} \tag{6.1}$$

and the anticausal subspace

$$\mathcal{B}_- := \{f \in \mathcal{B} : \hat{f}(k) = 0 \text{ for all } k \geq 0\}$$

of $\mathcal{B}$. In this chapter, we consider the decomposition of a transfer function $f \in \mathcal{B}$ into its causal part f_c and its anti-causal part f_{ac}, given by

$$f_c = \textstyle\sum_{k=0}^{\infty} \hat{f}(k)\,\mathrm{e}^{\mathrm{i}k\theta} \qquad \text{and} \qquad f_{ac} = \textstyle\sum_{k=-\infty}^{-1} \hat{f}(k)\,\mathrm{e}^{\mathrm{i}k\theta} \tag{6.2}$$

respectively. It is clear that f is the algebraic sum of its causal and anti-causal part: $f = f_c + f_{ac}$. However, it is not clear at the outset whether this decomposition defines two functions f_c and f_{ac} which belong again to $\mathcal{B}$, i.e. it is not clear whether the sums in (6.2) converge in the norm of $\mathcal{B}$ for each $f \in \mathcal{B}$. This depends strongly on the space $\mathcal{B}$. For illustration, we consider two examples.

Example 6.1. Let $\mathcal{B} = L^2$, the space of all square integrable functions on the unit circle $\mathbb{T}$. Let $f \in L^2$ be an arbitrary element of L^2 with its Fourier series $f(\mathrm{e}^{\mathrm{i}\theta}) = \sum_{k=-\infty}^{\infty} \hat{f}(k)\,\mathrm{e}^{\mathrm{i}k\theta}$. By Parseval's identity it follows immediately that

$$\|f_c\|_2 = \sqrt{\sum_{k=0}^{\infty} |\hat{f}(k)|^2} \leq \|f\|_2 \quad \text{and} \quad \|f_{ac}\|_2 = \sqrt{\sum_{k=-\infty}^{-1} |\hat{f}(k)|^2} \leq \|f\|_2$$

which shows that $f \in L^2$ always implies that $f_c \in L^2$ and $f_{ac} \in L^2$.

Example 6.2. Let $\mathcal{B} = \mathcal{C}(\mathbb{T})$ the set of all continuous functions on the unit circle and consider the function

V. Pohl, H. Boche, *Advanced Topics in System and Signal Theory*, Foundations in Signal Processing, Communications and Networking 4,
DOI 10.1007/978-3-642-03639-2_6,

$$f(e^{i\theta}) = \sum_{k=2}^{\infty} \frac{\sin(k\theta)}{k \log k}, \qquad \theta \in [-\pi, \pi) .$$

This series converges uniformly for all $\theta \in [-\pi, \pi)$ (by Theorem 1.3 in Chapter V of [92]), and therefore $f \in \mathcal{C}(\mathbb{T})$. It causal part is obviously given by

$$f_c(e^{i\theta}) = \frac{1}{2i} \sum_{k=2}^{\infty} \frac{1}{k \log k} e^{ik\theta}, \qquad \theta \in [-\pi, \pi) .$$

As it was already shown in Example 5.9, this series diverges at $\theta = 0$, and therefore, the causal part f_c of a continuous function f is not necessarily a continuous functions on $\mathbb{T}$.

This chapter characterizes Banach spaces $\mathcal{B}$ on which every $f \in \mathcal{B}$ can be decomposed into its causal and anticausal part such that $f_c \in \mathcal{B}_+$ and $f_{ac} \in \mathcal{B}_-$. The question whether or not such a decomposition is always possible is equivalent to the question whether the projection $f \mapsto f_c$ is bounded. This relation is discussed at the beginning in Section 6.1. Then Section 6.2 considers the decomposition of L^p spaces with $1 \leq p \leq \infty$. The situation we shall discuss in more detail is that in which $\mathcal{B}$ is a subset of $\mathcal{C}(\mathbb{T})$. Example 6.2 already showed that $\mathcal{C}(\mathbb{T})$ itself cannot be decomposed into a direct sum of $\mathcal{C}(\mathbb{T})_+$ and $\mathcal{C}(\mathbb{T})_-$. Section 6.3 will give necessary and sufficient conditions on the smoothness of the functions in $\mathcal{C}(\mathbb{T})$ such that such a decomposition is possible.

6.1 Complemented Subspaces and Projections

Definition 6.3. *Let $\mathcal{B}$ be a Banach space and let $\mathcal{M}$ be a closed subspace of $\mathcal{B}$. A closed subspace $\mathcal{N}$ of $\mathcal{B}$ is called the* direct complement *of $\mathcal{M}$ in $\mathcal{B}$ if*

$$\mathcal{B} = \mathcal{M} + \mathcal{N} \qquad \text{and} \qquad \mathcal{M} \cap \mathcal{N} = \{0\} .$$

In this case, we say that $\mathcal{M}$ is complemented *in $\mathcal{B}$ and that $\mathcal{B}$ is the* direct sum *of $\mathcal{M}$ and $\mathcal{N}$, which is written as*

$$\mathcal{B} = \mathcal{M} \oplus \mathcal{N} .$$

In the above definition, it is important that the two subspaces $\mathcal{M}$ and $\mathcal{N}$ are closed in $\mathcal{B}$. It should be noted that the direct complement of a closed subspace is not uniquely defined, in general. Moreover, not every closed subspace of a Banach spaces possesses a direct complement. Lindenstrauss and Tzafiri [58] even showed that in every Banach space which is not isomorphic to a Hilbert space there always exist closed subspaces which have no direct complement. Later we will see several examples of closed subspaces which can not be complemented. At the moment, we start with an important example of subspaces which always have a direct complement.

Theorem 6.4. *Let $\mathcal{B}$ be a Banach space and let $\mathcal{M}$ be a finite-dimensional subspace of $\mathcal{B}$. Then $\mathcal{M}$ is complemented in $\mathcal{B}$.*

Proof. Assume that the dimension of $\mathcal{M}$ is N. Then there exists a basis $\{\phi_k\}_{k=1}^N \in \mathcal{M}$ and a system of biorthonormal bounded linear functionals $\{c_k\}_{k=1}^N \in \mathcal{B}^*$ such that every $f \in \mathcal{M}$ can be written as

$$f = \sum_{k=1}^N c_k(f)\,\phi_k \ . \tag{6.3}$$

Define the subspace

$$\mathcal{N} := \{f \in \mathcal{B} : c_k(f) = 0 \text{ for all } k = 1, 2, \ldots, N\} \ .$$

It is easily verified that $\mathcal{N}$ is closed in $\mathcal{B}$ and that $\mathcal{M} \cap \mathcal{N} = \{0\}$. Indeed, assume that $f \in \mathcal{M} \cap \mathcal{N}$ is given by (6.3), then

$$c_n(f) = \sum_{k=1}^N c_k(f)\,c_n(\phi_k) = c_n(f) = 0 \qquad \text{for every } n = 1, 2, \ldots, N$$

by the biorthonormality of $\{\phi_k\}_{k=1}^N$ and $\{c_k\}_{k=1}^N$. But this implies $f = 0$. Finally, for an arbitrary $f \in \mathcal{B}$, define

$$f_1 := \sum_{k=1}^N c_k(f)\,\phi_k \qquad \text{and} \qquad f_2 := f - \sum_{k=1}^N c_k(f)\,\phi_k$$

It is clear that $f_1 \in \mathcal{M}$, and since $c_n(f_2) = c_n(f) - \sum_{k=1}^N c_k(f)\,c_n(\phi_k) = 0$ for all $n = 1, 2, \cdots, N$, it follows that $f_2 \in \mathcal{N}$. This proves that $\mathcal{B} = \mathcal{M} \oplus \mathcal{N}$. □

There exists a close relation between complemented subspaces and projections. This relation is considered next:

Definition 6.5. *Let $\mathcal{X}$ be a vector space. A linear mapping $\mathfrak{P} : \mathcal{X} \to \mathcal{X}$ is called a* projection *if*

$$\mathfrak{P}(\mathfrak{P}f) = \mathfrak{P}f \qquad \textit{for all } f \in \mathcal{X} \ .$$

It is easily verified that if $\mathfrak{P}$ is a projection on $\mathcal{X}$ then the operator $\mathfrak{Q} = \mathfrak{I} - \mathfrak{P}$ is also a projection on $\mathcal{X}$. The projection $\mathfrak{Q}$ is called the *complementary projection* of $\mathfrak{P}$. The following lemma collects some properties of the range and null space of projections and complementary projections.

Lemma 6.6. *Let $\mathcal{X}$ be a vector space, let $\mathfrak{P}$ be a projection on $\mathcal{X}$, and let $\mathfrak{Q} = \mathfrak{I} - \mathfrak{P}$ be the complementary projection. Then*

(a) $\mathscr{R}(\mathfrak{P}) = \mathscr{N}(\mathfrak{Q}) = \{f \in \mathcal{X} : \mathfrak{P}f = f\}$

(b) $\mathscr{N}(\mathfrak{P}) = \mathscr{R}(\mathfrak{Q})$

(c) $\mathscr{R}(\mathfrak{P}) \cap \mathscr{N}(\mathfrak{P}) = \{0\}$ *and* $\mathscr{R}(\mathfrak{P}) + \mathscr{N}(\mathfrak{P}) = \mathcal{X}$

The simple verification of the statements is left as an exercise. Note that the above lemma only assumes that $\mathcal{X}$ is a vector space which need not be normed. Point (c) of Lemma 6.6 states that $\mathcal{X}$ is the algebraic sum of the range and the null space of the projection $\mathfrak{P}$. However, if one considers projections on a Banach space $\mathcal{B}$, the subspaces $\mathscr{R}(\mathfrak{P})$ and $\mathscr{N}(\mathfrak{P})$ need not be closed, in general, such that $\mathcal{B}$ is not necessarily the direct sum of $\mathscr{R}(\mathfrak{P})$ and $\mathscr{N}(\mathfrak{P})$. The next theorem gives a sufficient condition on the projection $\mathfrak{P}$ in order for the range and kernel to be closed.

Theorem 6.7. *Let $\mathfrak{P}$ be a projection on a Banach space $\mathcal{B}$. If $\mathfrak{P}$ is continuous on $\mathcal{B}$ then*

$$\mathcal{B} = \mathscr{R}(\mathfrak{P}) \oplus \mathscr{N}(\mathfrak{P}) \, .$$

Proof. The algebraic part of the statement is equivalent to part (c) of Lemma 6.6. The null space $\mathscr{N}(\mathfrak{P})$ is closed by Theorem 1.15 and since by part (a) of Lemma 6.6 $\mathscr{R}(\mathfrak{P}) = \mathscr{N}(\mathfrak{I} - \mathfrak{P})$, and since the operator $\mathfrak{I} - \mathfrak{P}$ is linear and bounded whenever $\mathfrak{P}$ is linear and bounded, Theorem 1.15 implies also that $\mathscr{R}(\mathfrak{P})$ is closed. □

Conversely, assume that the Banach space $\mathcal{B}$ is the direct sum of two closed subspaces $\mathcal{M}$ and $\mathcal{N}$, and let $f = f_1 + f_2$ with $f_1 \in \mathcal{M}$ and $f_2 \in \mathcal{N}$ be an arbitrary element of $\mathcal{B}$. Define the operator $\mathfrak{P}$ on $\mathcal{B}$ by the equation $\mathfrak{P}f = f_1$ for all $f \in \mathcal{B}$. It is easy to see that $\mathscr{R}(\mathfrak{P}) = \mathcal{M}$, that $\mathscr{N}(\mathfrak{P}) = \mathcal{N}$ and that $\mathfrak{P}(\mathfrak{P}f) = \mathfrak{P}f$ for every $f \in \mathcal{B}$. Thus, $\mathfrak{P}$ is a projection of $\mathcal{B}$ onto $\mathcal{M}$ with null space $\mathcal{N}$. Moreover, since the direct sum of $\mathscr{R}(\mathfrak{P})$ and $\mathscr{N}(\mathfrak{P})$ is the whole $\mathcal{B}$, it follows that the so defined projection $\mathfrak{P}$ is even continuous. This is shown by the next theorem.

Theorem 6.8. *Let $\mathcal{B} = \mathcal{M} \oplus \mathcal{N}$ be a Banach space which is the direct sum of two closed subspaces $\mathcal{M}$ and $\mathcal{N}$. Then the projection $\mathfrak{P}$ with range $\mathcal{M}$ and null space $\mathcal{N}$ is continuous.*

Proof. The theorem is proved with the aid of the closed graph theorem. Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence in $\mathcal{B}$ with $f_n \to f$ and $\mathfrak{P}f_n \to g$. Since $\mathfrak{P}f_n \in \mathcal{M}$ for all $n \in \mathbb{N}$ and since $\mathcal{M}$ is assumed to be closed, one has that $g \in \mathcal{M}$. Similarly, since $f_n - \mathfrak{P}f_n = (\mathfrak{I} - \mathfrak{P})f_n \in \mathcal{N}$ for all $n \in \mathbb{N}$ and since $\mathcal{N}$ is closed, we have that $f - g \in \mathcal{N}$. It follows that $\mathfrak{P}f = \mathfrak{P}g = g$. Therefore, the graph of $\mathfrak{P}$ is closed in $\mathcal{B}$ and by the closed graph theorem $\mathfrak{P}$ is continuous. □

As a consequence of the previous two theorems, we have

Corollary 6.9. *A closed subspace $\mathcal{M}$ of a Banach space $\mathcal{B}$ is complemented in $\mathcal{B}$ if and only if $\mathcal{M}$ is the range of a bounded linear projection $\mathfrak{P} : \mathcal{X} \to \mathcal{M}$.*

As a first application of this corollary we prove that in a Hilbert space every closed subspace is complemented.

Theorem 6.10. *Let $\mathcal{H}$ be a Hilbert space and $\mathcal{M}$ be a closed subspace of $\mathcal{H}$. Then $\mathcal{M}$ is complemented in $\mathcal{H}$ with $\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^{\perp}$.*

Proof. Since $\mathcal{M}$ is a closed subspace of $\mathcal{H}$, there exists an orthonormal basis $\Phi = \{\phi_k\}_{k=1}^{\infty}$ of $\mathcal{M}$ and there exists an orthonormal basis S of $\mathcal{H}$ which contains Φ as a subset: $S = \Phi \cup \{\psi_k\}_{k=1}^{\infty}$. Thus, every $f \in \mathcal{H}$ can be written as

$$f = \textstyle\sum_{k=1}^{\infty} a_k \, \phi_k + \sum_{k=1}^{\infty} b_k \, \psi_k \, .$$

By Parseval's identity, $\|f\|_{\mathcal{H}}^2 = \sum_{k=1}^{\infty} |a_k|^2 + \sum_{k=1}^{\infty} |b_k|^2$ and the linear operator $\mathfrak{P} : \mathcal{H} \to \mathcal{M}$ defined by $\mathfrak{P} : f \mapsto \sum_{k=1}^{\infty} a_k \phi_k$ is obviously a projection with $\mathscr{R}(\mathfrak{P}) = \mathcal{M}$ and $\mathscr{N}(\mathfrak{P}) = \mathcal{M}^\perp$. Since $\|\mathfrak{P}f\|_{\mathcal{H}}^2 = \sum_{k=1}^{\infty} |a_k|^2 \leq \|f\|_{\mathcal{H}}^2$, the projection $\mathfrak{P}$ is bounded and the statement of the theorem follows from Corollary 6.9. □

6.2 Projections from L^p to H^p

The Banach spaces L^p with $1 \leq p \leq \infty$, $p \neq 2$ are not Hilbert spaces. By a result of Lindenstrauss and Tzafiri [58] there always exist closed subspaces of L^p, $p \neq 2$ which are not complemented in L^p. For our applications, the causal subspace $H^p = L_+^p$ is of particular interest. Therefore, we ask in the following whether this particular subspace is complemented in L^p. By Corollary 6.9, this is equivalent to the question whether there is a bounded projection $\mathfrak{T} : L^p \to H^p$.

We start by deriving a result, which allows us to consider only the "natural" projection from $L^p \to H^p$. For $f \in L^p$, one possible projection $L^p \to H^p$ is the natural projection $\mathfrak{P}_+$ which is defined formally by

$$\mathfrak{P}_+ : \sum_{k=-\infty}^{\infty} \hat{f}(k)\, \mathrm{e}^{\mathrm{i}k\theta} \mapsto \sum_{k=0}^{\infty} \hat{f}(k)\, \mathrm{e}^{\mathrm{i}k\theta} \tag{6.4}$$

and which is called *Riesz projection*. Inserting the Fourier coefficients (2.1) of f into the sum of the right hand side of (6.4) gives a closed form expression for the Riesz projection

$$(\mathfrak{P}_+ f)(z) = \frac{1}{2\pi\mathrm{i}} \int_{\mathbb{T}} \frac{f(\zeta)}{\zeta - z}\, \mathrm{d}\zeta = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\tau})\, \frac{\mathrm{e}^{\mathrm{i}\tau}}{\mathrm{e}^{\mathrm{i}\tau} - z}\, \mathrm{d}\tau\,, \qquad z \in \mathbb{D}\,. \tag{6.5}$$

Among all bounded projections $L^p \to H^p$, the Riesz projection has the remarkable property that it is the projection with the least operator norm. This will be proved in the next theorem. Therein, we use the following notation. If f is an arbitrary function on the unit circle $\mathbb{T}$ and $\lambda \in \mathbb{R}$, then f_λ will denote the right-shifted (or rotated) function defined by $f_\lambda(\mathrm{e}^{\mathrm{i}t}) = f(\mathrm{e}^{\mathrm{i}(t-\lambda)})$.

Theorem 6.11. *Let $1 \leq p < \infty$ and let $\mathfrak{T} : L^p \to H^p$ be an arbitrary bounded projection, i.e. $\mathfrak{T}f = f$ for all $f \in H^p$ and $\|\mathfrak{T}f\|_p \leq C\, \|f\|_p$ for all $f \in L^p$ with a positive constant C. Then*

$$(\mathfrak{P}_+ f)(\rho \mathrm{e}^{\mathrm{i}\tau}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} (\mathfrak{T} f_\lambda)(\rho \mathrm{e}^{\mathrm{i}(\tau+\lambda)})\, \mathrm{d}\lambda \tag{6.6}$$

and

$$\|\mathfrak{P}_+\|_{L^p \to H^p} \leq \|\mathfrak{T}\|_{L^p \to H^p}\ . \tag{6.7}$$

Proof. First, consider for a fixed $f \in L^p$ the mapping $\lambda \mapsto f_\lambda$ from the unit circle $\mathbb{T}$ into L^p. This mapping is continuous. To see this, we use that $\mathcal{C}(\mathbb{T})$ is dense in L^p for all $1 \leq p < \infty$, i.e. to every $\epsilon > 0$ there exists a $g \in \mathcal{C}(\mathbb{T})$ such that $\|f - g\|_p < \epsilon/2$. Therewith, one obtains

$$\|f - f_\lambda\|_p \leq \|f - g\|_p + \|g - g_\lambda\|_p + \|g_\lambda - f_\lambda\|_p \leq \|g - g_\lambda\|_\infty + \epsilon \,.$$

Since g is continuous, there exists a constant C such that $\|g - g_\lambda\|_\infty \leq C\,\lambda$ which finally proves the continuity of the mapping $\lambda \mapsto f_\lambda$ on L^p. Moreover, since $\mathfrak{T}$ is assumed to be continuous, the mapping $\lambda \mapsto (\mathfrak{T} f_\lambda)_{-\lambda}$ is continuous as well. This shows that the integral on the right hand side of (6.6) is well defined.

Next, we prove the identity (6.6). Let $\phi_k(\zeta) = \zeta^k$, $\zeta \in \mathbb{T}$. Since $\mathfrak{T}$ is a projection $L^p \to H^p$, it holds that $\mathfrak{T}\phi_k = \phi_k$ for all $k \geq 0$. For $k < 0$ we set $g_k := \mathfrak{T}\,\phi_k$. Of course $g_k \in H^p$. Consider the function $f \in L^p$ given by

$$f(\zeta) = \textstyle\sum_{k=-\infty}^{\infty} c_k\,\zeta^k = \sum_{k=-\infty}^{\infty} c_k\,\phi_k(\zeta)\,, \quad z \in \mathbb{T}\,.$$

Applying the projection $\mathfrak{T}$ onto the function $f_\lambda(\mathrm{e}^{\mathrm{i}t}) := f(\mathrm{e}^{\mathrm{i}(t-\lambda)})$, one obtains at $z = \rho\,\mathrm{e}^{\mathrm{i}(\tau+\lambda)}$, $0 \leq \rho < 1$ the H^p-function

$$(\mathfrak{T} f_\lambda)(\rho\,\mathrm{e}^{\mathrm{i}(\tau+\lambda)}) = \sum_{k=-\infty}^{-1} c_k\,\mathrm{e}^{-\mathrm{i}k\lambda}\,g_k(\rho\,\mathrm{e}^{\mathrm{i}(\tau+\lambda)}) + \sum_{k=0}^{\infty} c_k\,\rho\,\mathrm{e}^{\mathrm{i}k\tau}\,.$$

This term is integrated with respect to λ, which gives

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} (\mathfrak{T} f_\lambda)(\rho\,\mathrm{e}^{\mathrm{i}(\tau+\lambda)})\,\mathrm{d}\lambda = \sum_{k=0}^{\infty} c_k\,\rho\,\mathrm{e}^{\mathrm{i}k\tau} + \sum_{k=-\infty}^{-1} c_k\,\frac{1}{2\pi}\int_{-\pi}^{\pi} g_{k,\tau}(\rho\,\mathrm{e}^{\mathrm{i}\lambda})\,\mathrm{e}^{-\mathrm{i}k\lambda}\,\mathrm{d}\lambda$$

where $g_{k,\tau}(\rho\mathrm{e}^{\mathrm{i}\lambda}) := g_k(\rho\mathrm{e}^{\mathrm{i}(\tau+\lambda)})$ with $g_{k,\tau} \in H^p$ and $\|g_{k,\tau}\|_p = \|g_k\|_p$. Since the functions $g_{k,\tau}$ are analytic in $\mathbb{D}$, the Cauchy integral theorem shows that the integrals in the last sum are equal to zero for all $k < 0$. But the first term is the Riesz projection of f such that (6.6) is obtained.

Relation (6.7) follows immediately from representation (6.6), noting that $\mathfrak{T}$ is bounded and that the rotation $f \mapsto f_\lambda$ does not change the L^p norm. Therefore, for every fixed λ, one has $\|(\mathfrak{T} f_\lambda)_{-\lambda}\|_p \leq \|\mathfrak{T}\|_{L^p \to H^p}\,\|f\|_p$. Using this in (6.6), one obtains (6.7). □

Thus, the Riesz projection has the least operator norm among all bounded projection $L^p \to H^p$. Therefore, to investigate whether H^p is a complemented subspace in L^p, we only have to investigate the boundedness of the Riesz projection $\mathfrak{P}_+$. Because, if $\mathfrak{P}_+$ is bounded, then Corollary 6.9 implies that H^p is a complemented subspace. If on the other hand, $\mathfrak{P}_+$ is unbounded, the previous theorem shows that there exist no other bounded projector $L^p \to H^p$, and so H^p would not complemented in this case.

Thus, we have to investigate the Riesz projection $\mathfrak{P}_+$ on the spaces L^p. We note at the beginning that $\mathfrak{P}_+$ is closely related to the Herglotz-Riesz transform $\mathfrak{R}$ defined in Section 5.1. Comparing the definition (5.1) of $\mathfrak{R}f$ with the definition of the Riesz projections (6.4) shows that

$$\begin{aligned}(\mathfrak{P}_+ f)(z) &= \tfrac{1}{2}\left[(\mathfrak{R}f)(z) + \hat{f}(0)\right] \\ &= \tfrac{1}{2}\left[(\mathfrak{P}f)(z) + \mathrm{i}\,(\mathfrak{Q}f)(z) + \hat{f}(0)\right], \qquad z \in \mathbb{D}. \end{aligned} \tag{6.8}$$

Theorem 6.12 (M. Riesz). *Let $f \in L^p$ for $1 < p < \infty$ be a function on the unit circle. Then there exists a constant $K(p)$ such that*

$$\|\mathfrak{P}_+ f\|_{H^p} \leq K(p)\,\|f\|_{L^p}$$

where the constant $K(p)$ depends only on p, but not on f.

Proof. For an arbitrary $0 \leq r < 1$ define the function $F_r(\mathrm{e}^{\mathrm{i}\theta}) := (\mathfrak{P}f)(r\mathrm{e}^{\mathrm{i}\theta})$, $\widetilde{F}_r(\mathrm{e}^{\mathrm{i}\theta}) := (\mathfrak{Q}f)(r\mathrm{e}^{\mathrm{i}\theta})$, and $G_r(\mathrm{e}^{\mathrm{i}\theta}) := (\mathfrak{P}_+ f)(r\mathrm{e}^{\mathrm{i}\theta})$. Therewith relation (6.8) becomes

$$G_r(\mathrm{e}^{\mathrm{i}\theta}) = \tfrac{1}{2}\left[F_r(\mathrm{e}^{\mathrm{i}\theta}) + \mathrm{i}\,\widetilde{F}_r(\mathrm{e}^{\mathrm{i}\theta}) + \hat{f}(0)\right], \qquad \theta \in [-\pi, \pi).$$

By the definition (2.1) of the Fourier coefficient $\hat{f}(0)$, it is clear that $|\hat{f}(0)| \leq \|f\|_1 \leq \|f\|_p$ and since, according to Theorem 5.4 and 5.13, the Poisson and the conjugate Poisson integrals are bounded on L^p, one obtains

$$\begin{aligned}\|G_r\|_{L^p} &\leq \tfrac{1}{2}\left[\,\|F_r\|_{L^p} + \|\widetilde{F}_r\|_{L^p} + |\hat{f}(0)|\,\right] \\ &\leq \tfrac{1}{2}\left[\,\|f\|_p + C(p)\,\|f\|_p + \|f\|_p\,\right]\end{aligned}$$

with the boundedness constant $C(p)$ of the conjugate Poisson integral, which only depends on p. Therewith, one obtains finally

$$\|\mathfrak{P}_+ f\|_{H^p} = \sup_{0 \leq r < 1} \|G_r\|_{L^p} \leq K(p)\,\|f\|_p$$

with the constant $K(p) = 1 + \frac{1}{2}\,C(p)$. □

We still have to consider the cases $p = 1$ and $p = \infty$. In both cases, it is no longer true the the Riesz projection $\mathfrak{P}_+$ is bounded. Consequently, we have

Theorem 6.13. *There is no bounded causal projection of L^1 onto H^1 and there is no bounded causal projection of L^∞ onto H^∞.*

Proof. The unboundedness of $\mathfrak{P}_+$ is proved with the same counter examples as used in Example 5.7 and 5.8: In the case $p = 1$ consider the function $f(\mathrm{e}^{\mathrm{i}\tau}) = \sum_{k=2}^{\infty} \frac{\cos(k\tau)}{\log k}$ and its Riesz projection $(\mathfrak{P}_+ f)(\mathrm{e}^{\mathrm{i}\tau}) = \sum_{k=2}^{\infty} \frac{\exp(\mathrm{i}k\tau)}{\log k}$. It was shown in Example 5.8 that $f \in L^1$ but that $\mathfrak{P}_+ f \notin H^1$.
In the case $p = \infty$ one considers $g(\mathrm{e}^{\mathrm{i}\tau}) = \frac{2}{\pi} \sum_{k=2}^{\infty} \frac{\sin(k\tau)}{k}$ with its Riesz projection $(\mathfrak{P}_+ g)(\mathrm{e}^{\mathrm{i}\tau}) = \frac{1}{\mathrm{i}\pi} \sum_{k=2}^{\infty} \frac{\exp(\mathrm{i}k\tau)}{k}$, and its was shown in Example 5.7 that $\|g\|_\infty < \infty$ whereas $(\mathfrak{P}_+ g)(\mathrm{e}^{\mathrm{i}\tau})$ is unbounded at $\tau = 0$. □

6.3 Projections in Spaces of Smooth Functions

We saw that there exists no bounded projection of L^∞ onto H^∞. In other words, given an energy stable (non-causal) transfer function $f \in L^\infty$ of the form

$$f(e^{i\tau}) = \sum_{k=-\infty}^{\infty} \hat{f}(k)\, e^{ik\tau}\,, \qquad \tau \in [-\pi, \pi)$$

with impulse response $\{\hat{f}(k)\}_{k=-\infty}^{\infty}$, then a simple truncation of the noncausal part of f by $f_c = \mathfrak{P}_+ f$ yields a causal but in general not a stable transfer function f_c.

The space L^∞, and in particular the causal subspace H^∞, contains quite complicated functions, such that it may not be surprising at first glance that there exist some problems with the boundedness of the Riesz projection on L^∞ – maybe just these complicated functions cause the unboundedness of the Riesz projection. In the following, we investigate the Riesz projection on subspaces of L^∞ with smooth function. The first subspace which we will consider is the set of all continuous functions $\mathcal{C}(\mathbb{T}) \subset H^\infty$ on the unit circle.

Theorem 6.14. *There exists no bounded causal projection $\mathcal{C}(\mathbb{T})$ onto $A(\mathbb{D})$. Equivalently, $A(\mathbb{D})$ is not complemented in $\mathcal{C}(\mathbb{T})$.*

Proof. It is clear that Theorem 6.11 holds also for projections from $\mathcal{C}(\mathbb{T})$ to $A(\mathbb{D})$. The proof is even much more direct. Therefore, it remains to show that the Riesz projection $\mathfrak{P}_+$ is unbounded on $\mathcal{C}(\mathbb{T})$, but this follows already from the counter example given in Example 5.9. □

Next, we consider causal projections on the spaces $\mathcal{C}_\omega(\mathbb{T}) \subset \mathcal{C}(\mathbb{T})$ of smooth functions and derive necessary and sufficient conditions on the majorant ω such that there exists a bounded causal projection of $\mathcal{C}_\omega(\mathbb{T})$ onto $A_\omega(\mathbb{D})$. Our approach is similar as for the L^p spaces. At the beginning, it is shown that the Riesz projection $\mathfrak{P}_+$ has the least operator norm among all bounded causal projections $\mathfrak{T}$ from $\mathcal{C}_\omega(\mathbb{T})$ onto $A_\omega(\mathbb{D})$. Afterwards, necessary and sufficient condition on ω are derived such that $\mathfrak{P}_+$ is a bounded projection $\mathcal{C}_\omega(\mathbb{T}) \to A_\omega(\mathbb{D})$.

Theorem 6.15. *Let ω be a majorant and let $\mathfrak{T}$ be an arbitrary bounded projection $\mathfrak{T} : \mathcal{C}_\omega(\mathbb{T}) \to A_\omega(\mathbb{D})$. Then $\mathfrak{P}_+$ can be written as in* (6.6) *and*

$$\|\mathfrak{P}_+\|_{\mathcal{C}_\omega(\mathbb{T}) \to A_\omega(\mathbb{D})} \le \|\mathfrak{T}\|_{\mathcal{C}_\omega(\mathbb{T}) \to A_\omega(\mathbb{D})}\ . \tag{6.9}$$

Proof. The representation (6.6) of the Riesz projection is verified in the same way as it was done in the proof of Theorem 6.11.

Let $f \in \mathcal{C}_\omega(\mathbb{T})$, then by (6.6) and by the linearity of $\mathfrak{T}$, one obtains

$$\begin{aligned}\left|(\mathfrak{P}_+f)(\mathrm{e}^{\mathrm{i}\tau_1}) - (\mathfrak{P}_+f)(\mathrm{e}^{\mathrm{i}\tau_2})\right| &\leq \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\mathfrak{T}f_\lambda(\mathrm{e}^{\mathrm{i}(\tau_1+\lambda)}) - \mathfrak{T}f_\lambda(\mathrm{e}^{\mathrm{i}(\tau_2+\lambda)})\right|\,\mathrm{d}\lambda \\ &\leq \frac{1}{2\pi}\int_{-\pi}^{\pi}\|\mathfrak{T}f_\lambda\|_{A_\omega(\mathbb{D})}\omega(|\tau_1-\tau_2|)\,\mathrm{d}\lambda \\ &\leq \|\mathfrak{T}\|_{\mathcal{C}_\omega(\mathbb{T})\to A_\omega(\mathbb{D})}\,\|f\|_{\mathcal{C}_\omega(\mathbb{T})}\,\omega(|\tau_1-\tau_2|)\;.\end{aligned}$$

Therein f_λ is defined by $f_\lambda(\mathrm{e}^{\mathrm{i}\tau}) = f(\mathrm{e}^{\mathrm{i}(\tau-\lambda)})$, and for the last line it was used that $f \in \mathcal{C}_\omega(\mathbb{T})$ such that $f_\lambda \in \mathcal{C}_\omega(\mathbb{T})$ with $\|f\|_{\mathcal{C}_\omega(\mathbb{T})} = \|f_\lambda\|_{\mathcal{C}_\omega(\mathbb{T})}$. The above inequality shows that $\|\mathfrak{P}_+f\|_{\mathcal{C}_\omega(\mathbb{T})} \leq \|\mathfrak{S}\|_{\mathcal{C}_\omega(\mathbb{T})\to A_\omega(\mathbb{D})}\,\|f\|_{\mathcal{C}_\omega(\mathbb{T})}$ for all $f \in \mathcal{C}_\omega(\mathbb{T})$ which is equivalent to (6.9). □

It remains to investigate the boundedness of the Riesz projection $\mathfrak{P}_+$. By Theorem 6.14, $\mathfrak{P}_+$ is unbounded on $\mathcal{C}(\mathbb{T})$. However, by restricting the domain of the operator to a certain subset of $\mathcal{C}(\mathbb{T})$, $\mathfrak{P}_+$ may become continuous on this subset. This is shown by the following theorem. It gives necessary and sufficient condition on the smoothness of the functions in $\mathcal{C}(\mathbb{T})$ such that the Riesz projection is always bounded and consequently continuous.

Theorem 6.16 (Boundedness of the Riesz projection). *Let ω be a majorant. Then $\mathfrak{P}_+$ is a bounded projection from $C_\omega(\mathbb{T})$ onto $A_\omega(\mathbb{D})$ if and only if ω is a regular majorant. Thus, there exists a constant $C = C(\omega)$ such that $\|\mathfrak{P}_+f\|_{\mathcal{C}_\omega(\overline{\mathbb{D}})} \leq C(\omega)\,\|f\|_{\mathcal{C}_\omega(\mathbb{T})}$ for all $f \in \mathcal{C}_\omega(\mathbb{T})$ if and only if ω is a regular majorant.*

Proof. According to (6.8), the Riesz projection $\mathfrak{P}_+$ can be expressed in terms of the Herglotz-Riesz transform $\mathfrak{R}$ given by (5.1). By the definition of the norm in the $\mathcal{C}_\omega$-spaces (1.9), it is clear that $|\widehat{f}_0| \leq \|f\|_\infty \leq \|f\|_{\mathcal{C}_\omega(\mathbb{T})}$. Therefore, the Riesz projection $\mathfrak{P}_+$ is continuous if and only if the HR-transform $\mathfrak{R}$ is continuous, because

$$\|\mathfrak{P}_+f\|_{\mathcal{C}_\omega(\overline{\mathbb{D}})} \leq \tfrac{1}{2}\left(|\widehat{f}_0| + \|\mathfrak{R}f\|_{\mathcal{C}_\omega(\overline{\mathbb{D}})}\right) \leq \tfrac{1}{2}\,(1+C_1)\,\|f\|_{\mathcal{C}_\omega(\mathbb{T})}$$

and conversely

$$\|\mathfrak{R}f\|_{\mathcal{C}_\omega(\overline{\mathbb{D}})} \leq |\widehat{f}_0| + 2\,\|\mathfrak{P}_+f\|_{\mathcal{C}_\omega(\overline{\mathbb{D}})} \leq (1+2\,C_2)\,\|f\|_{\mathcal{C}_\omega(\mathbb{T})}\;.$$

Therein, C_1 and C_2 are the continuity constants of $\mathfrak{R}$ and $\mathfrak{P}_+$, respectively. Therefore, it remains to prove that $\mathfrak{R} : \mathcal{C}_\omega(\mathbb{T}) \to \mathcal{C}_\omega(\overline{\mathbb{D}})$ is bounded if and only if ω is a regular majorant.

1) Sufficiency: The first part will show that if ω is regular then $\mathfrak{R}$ is continuous. To this end, let $\mathfrak{R}f = \mathfrak{P}f + \mathrm{i}\,\mathfrak{Q}f$ be the decomposition of $\mathfrak{R}$ into its real and imaginary parts as in (5.7). By Theorem 5.4 there exists a constant C_3 such that (5.12) holds for all $f \in \mathcal{C}_\omega(\mathbb{T})$, and by Theorem 5.14 there exists a constant C_4 such that (5.28) holds for all $f \in \mathcal{C}_\omega(\mathbb{T})$. Consequently

$$\|\mathfrak{R}f\|_{\mathcal{C}_\omega(\overline{\mathbb{D}})} \leq \|\mathfrak{P}f\|_{\mathcal{C}_\omega(\overline{\mathbb{D}})} + \|\mathfrak{Q}f\|_{\mathcal{C}_\omega(\overline{\mathbb{D}})} \leq (C_3+C_4)\,\|f\|_{\mathcal{C}_\omega(\mathbb{T})}$$

for all $f \in \mathcal{C}_\omega(\mathbb{T})$ which shows that $\mathfrak{R}$ is continuous.

2) Necessity: We show that if $\mathfrak{R} : \mathcal{C}_\omega(\mathbb{T}) \to \mathcal{C}_\omega(\overline{\mathbb{D}})$ is continuous then ω is a regular majorant. Assume that there exists a constant C_5 such that $\|\mathfrak{R}f\|_{\mathcal{C}_\omega(\overline{\mathbb{D}})} \leq C_5 \, \|f\|_{\mathcal{C}_\omega(\mathbb{T})}$ for all $f \in \mathcal{C}_\omega(\mathbb{T})$. Since

$$\|\mathfrak{P}f\|_{\mathcal{C}_\omega(\overline{\mathbb{D}})} \leq \|\mathfrak{P}f + \mathrm{i}\,\mathfrak{Q}f\|_{\mathcal{C}_\omega(\overline{\mathbb{D}})} = \|\mathfrak{R}f\|_{\mathcal{C}_\omega(\overline{\mathbb{D}})} \leq C_5 \, \|f\|_{\mathcal{C}_\omega(\mathbb{T})}$$

it follows that also $\mathfrak{P}$ is continuous, and from a similar inequality follows that $\mathfrak{Q}$ is continuous. Theorem 5.4 shows then that ω is weak regular of type 2.

It remains to show that ω is also weak regular of type 1. To this end, we consider the odd function defined by

$$f_2(\mathrm{e}^{\mathrm{i}\tau}) = \begin{cases} -\omega(\pi/2)\,\sin(\tau) & \pi/2 < \tau \leq \pi \\ -\omega(\tau) & 0 < \tau \leq \pi/2 \\ \ \ \omega(-\tau) & -\pi/2 < \tau \leq 0 \\ \ \ \omega(\pi/2)\,\sin(-\tau) & -\pi < \tau \leq -\pi/2 \end{cases}$$

and study the conjugate Poisson integral $(\mathfrak{Q}f_2)(r\mathrm{e}^{\mathrm{i}t}) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f_2(\mathrm{e}^{\mathrm{i}\tau})\, \mathcal{Q}_r(t-\tau)\,\mathrm{d}\tau$. Since f_2 and $\mathcal{Q}_r$ are odd functions, $\mathfrak{Q}f_2$ can be written at $t = 0$ as

$$(\mathfrak{Q}f_2)(r) = \frac{1}{\pi}\int_0^{\pi} f_2(\mathrm{e}^{\mathrm{i}\tau})\, \mathcal{Q}_r(-\tau)\,\mathrm{d}\tau \ . \tag{6.10}$$

Using the properties of a majorant, it can be verified that the function f_2 belongs to $\mathcal{C}_\omega(\mathbb{T})$. Since $\mathfrak{Q}$ is also continuous it follows that there exists a constant C_6 such that

$$|(\mathfrak{Q}f_2)(r) - (\mathfrak{Q}f_2)(1)| \leq C_6\, \omega(1-r) \ . \tag{6.11}$$

Next we consider $(\mathfrak{Q}f_2)(1)$. For this term holds

$$(\mathfrak{Q}f_2)(1) \geq \frac{1}{\pi}\int_0^{1-r} f_2(\mathrm{e}^{\mathrm{i}\tau}) \frac{-\sin\tau}{1-\cos\tau}\,\mathrm{d}\tau \geq \frac{4}{\pi^2}\int_0^{1-r} \frac{\omega(\tau)}{\tau}\,\mathrm{d}\tau \tag{6.12}$$

where the first inequality is a consequence of the positivity of the integrand and the second inequality follows from the relations $\sin\tau \geq \frac{2}{\pi}\,\tau$ and $1-\cos\tau \leq \frac{1}{2}\,\tau^2$ for all $0 \leq \tau \leq \pi$ and from the definition of f_2. Note that the term on the right hand side is equivalent to the definition of a weak regular majorant of type 1. Therewith, it remains to show that there exists a constant C_7 such that the middle term in (6.12) is upper bounded by $C_7\, \omega(1-r)$. To this end, the middle term in (6.12) is written as

$$\left|\frac{1}{\pi}\int_0^{1-r} f_2(\mathrm{e}^{\mathrm{i}\tau})\, \mathcal{Q}_1(-\tau)\,\mathrm{d}\tau\right| \leq \left|\frac{1}{\pi}\int_0^{\pi} f_2(\mathrm{e}^{\mathrm{i}\tau})\, \mathcal{Q}_1(-\tau)\,\mathrm{d}\tau - (\mathfrak{Q}f_2)(r)\right| + \\ + \left|(\mathfrak{Q}f_2)(r) - \frac{1}{\pi}\int_{1-r}^{\pi} f_2(\mathrm{e}^{\mathrm{i}\tau})\, \mathcal{Q}_1(-\tau)\,\mathrm{d}\tau\right| \ . \tag{6.13}$$

Because of (6.10), the first term on the right hand side is equal to $|(\mathfrak{Q}f_2)(1) - (\mathfrak{Q}f_2)(r)|$ which is upper bounded by $C_6\,\omega(1-r)$ as (6.11) shows. It remains to investigate the second term on the right hand side, which will be denoted by T_0. In it, the integral which belongs to $(\mathfrak{Q}f_2)(r)$ is split up into an integration over $[0, 1-r]$ and an integration over $[1-r, \pi]$. This gives

$$|T_0| \leq \left|\frac{1}{\pi}\int_0^{1-r} f_2(\mathrm{e}^{\mathrm{i}\tau})\, \mathcal{Q}_r(-\tau)\,\mathrm{d}\tau\right| + \left|\frac{1}{\pi}\int_{1-r}^{\pi} f_2(\mathrm{e}^{\mathrm{i}\tau})\, [\mathcal{Q}_r(-\tau) - \mathcal{Q}_1(-\tau)]\,\mathrm{d}\tau\right| . \tag{6.14}$$

The two terms on the right hand side are denoted by T_1 and T_2 and investigated separately. First, T_1 is considered. It should be noted that it can be written as $\frac{1}{2\pi}\int_{|\tau|\leq 1-r} f_2(\mathrm{e}^{\mathrm{i}\tau})\,\mathcal{Q}_r(-\tau)\,\mathrm{d}\tau$. Since $\mathcal{Q}_r(\tau)$ is an odd function, an arbitrary constant can be added to f_2 without changing the value of the integral. Therefore, the following chain of upper bounds for T_1 is obtained

$$\begin{aligned}
|T_1| &\overset{(a)}{\leq} \frac{1}{2\pi}\int_{|\tau|\leq 1-r} \left|f_2(\mathrm{e}^{\mathrm{i}\tau}) - f_2(\mathrm{e}^{\mathrm{i}0})\right|\, |\mathcal{Q}_r(-\tau)|\,\mathrm{d}\tau \\
&\overset{(b)}{\leq} \|f_2\|_{\mathcal{C}_\omega(\mathbb{T})}\,\frac{1}{\pi}\int_0^{1-r} \omega(|\mathrm{e}^{\mathrm{i}\tau} - \mathrm{e}^{\mathrm{i}0}|)\, |\mathcal{Q}_r(-\tau)|\,\mathrm{d}\tau \\
&\overset{(c)}{\leq} \|f_2\|_{\mathcal{C}_\omega(\mathbb{T})}\,\omega(1-r)\,\frac{2}{\pi}\int_0^{1-r} \frac{r\sin\tau}{1 - 2r\cos\tau + r^2}\,\mathrm{d}\tau \\
&\overset{(d)}{\leq} 4\,\|f_2\|_{\mathcal{C}_\omega(\mathbb{T})}\,\omega(1-r)
\end{aligned}$$

In it, the triangle inequality was used to obtain (a), and (b) follows from $f_2 \in \mathcal{C}_\omega(\mathbb{T})$. For (c), it was used that $\left|\mathrm{e}^{\mathrm{i}\tau} - \mathrm{e}^{\mathrm{i}0}\right| \leq |\tau|$, that ω is non-decreasing, and that $\tau \leq 1-r$. To obtain (d), observe that $\sin\tau$ is a monotonic increasing function on $[0, 1-r]$ such that $\sin(\tau) \leq \sin(1-r) \leq 1-r$. Therewith, the numerator of the integrand can be upper bounded by $r\sin\tau \leq r(1-r) \leq 1-r^2$, such that the integrand becomes the Poisson kernel. This shows that the integral is always smaller that 2π.
Next, we consider the second term $|T_2|$ on the right hand side of (6.14). Plugging in the kernels $\mathcal{Q}_r$ and $\mathcal{Q}_1$ and applying triangle inequality gives

$$|T_2| = \frac{1}{\pi}\int_{1-r}^{\pi} |f_2(\mathrm{e}^{\mathrm{i}\tau})| \frac{(1-r)^2\sin\tau}{[1-\cos\tau]\,[1-2r\cos\tau + r^2]}\,\mathrm{d}\tau\ .$$

The second bracket in the denominator may be written as $(1-r)^2 + 2r(1-\cos\tau)$. Next, the inequality $\frac{1}{2}\,\tau^2(1 - \frac{1}{12}\,\tau^2) \leq 1 - \cos\tau$ is used. Since $\tau \leq \pi$, it follows that $\frac{c_0}{2}\tau^2 \leq 1 - \cos\tau$ with the constant $c_0 = 1 - \frac{\pi^2}{12}$. Using this inequality in the above equation for $|T_2|$ gives

$$|T_2| \leq \frac{2}{c_0\,\pi}\int_{1-r}^{\pi} |f_2(\mathrm{e}^{\mathrm{i}\tau})| \frac{(1-r)^2\,\sin\tau}{\tau^2[(1-r)^2 + c_0\,r\,\tau^2]}\,\mathrm{d}\tau\ . \tag{6.15}$$

Now, we distinguish two cases. First we assume that $r \geq R_0 > 0$ for an arbitrary $R_0 < 1$. Then we use that $(1-r)^2 + c_0\, r\, \tau^2 \geq c_0\, r\, \tau^2$, that $\sin\tau \leq \tau$, and that $\tau \geq 1-r$. Therewith, $|T_2|$ can be upper bounded by

$$|T_2| \leq \frac{2}{c_0^2\, R_0\, \pi}(1-r) \int_{1-r}^{\pi} \frac{|f_2(\mathrm{e}^{\mathrm{i}\tau})|}{\tau^2}\, \mathrm{d}\tau\, .$$

Now, we use the definition of f_2 and split up the integral into an integration from $1-r$ to $\pi/2$ and an integration from $\pi/2$ to π. By the definition of f_2, the integration over $[-\pi/2, \pi]$ will only give a certain constant C_8. Therefore, we obtain

$$|T_2| \leq \frac{2}{c_0^2\, R_0\, \pi}\left[(1-r)\int_{1-r}^{\pi/2} \frac{\omega(\tau)}{\tau^2}\, \mathrm{d}\tau + C_8\,(1-r)\right]\, .$$

It was already shown in the first step of this proof that the majorant ω is weak regular of type 2. Therefore, the first term in the brackets is upper bounded by $C_9\, \omega(1-r)$. Moreover, since $\omega(\tau)/\tau$ is non-increasing it holds that $(1-r) \leq \omega(1-r)/\omega(1)$. Altogether, we get that $|T_2| \leq C_{10}\, \omega(1-r)$ with a certain constant C_{10}. It remains to consider the case that $r \leq R_0 < 1$. To this end we use the following obvious relations in (6.15): $\sin\tau \leq 1$ and $(1-r)^2 + c_0\, r\, \tau^2 \geq (1-r)^2$. This yields the upper bound

$$|T_2| \leq \frac{2}{c_0\, \pi}\, \frac{1-r}{1-R_0} \int_{1-r}^{\pi} \frac{|f_2(\mathrm{e}^{\mathrm{i}\tau})|}{\tau^2}\, \mathrm{d}\tau\, .$$

Now we can proceed as in the case $r \geq R_0 > 0$ and also obtain $|T_2| \leq C_{11}\, \omega(1-r)$ with a certain constant C_{11}.

Collecting the upper bounds for $|T_2|$, $|T_1|$, and $|T_0|$ and substitute them in (6.13) and (6.12) shows that there exists a constant C_{12} such that

$$\int_0^{1-r} \frac{\omega(\tau)}{\tau}\, \mathrm{d}\tau \leq C_{12}\, \omega(1-r)$$

which proves that ω is weak regular of type 1. □

Since the Riesz projection simply truncates the left side of the Fourier series, the theorem implies that the truncation error $\|f - \mathfrak{P}_+ f\|_{\mathcal{C}_\omega(\mathbb{T})}$ is bounded for all $f \in \mathcal{C}_\omega(\mathbb{T})$, because

$$\|f - \mathfrak{P}_+ f\|_{\mathcal{C}_\omega(\mathbb{T})} \leq \|f\|_{\mathcal{C}_\omega(\mathbb{T})} + \|\mathfrak{P}_+ f\|_{\mathcal{C}_\omega(\mathbb{T})} \leq \left(1 + \|\mathfrak{P}_+\|_{\mathcal{C}_\omega(\mathbb{T}) \to A_\omega(\mathbb{D})}\right)\, \|f\|_{\mathcal{C}_\omega(\mathbb{T})}$$

where the last inequality is a consequence of the continuity of $\mathfrak{P}_+$.

Note that Theorem 6.16 does not say that $f_+ = \mathfrak{P}_+ f$ becomes necessarily unbounded if $f \in \mathcal{C}_\omega(\mathbb{T})$ with a non-regular majorant ω. If ω is not regular, it is only no longer possible to control the modulus of continuity of f_+. For instance, if $f \in \mathcal{C}_{\omega,0}(\mathbb{T})$ with a weak regular majorant ω of type 1, then an analysis of the proof of Theorem 6.16 shows that f_+ is always bounded ($\|f\|_\infty < \infty$) and therefore an element of $A(\mathbb{D})$, but f_+ does not belong to $A_\omega(\mathbb{D})$, in general. However, if f is assumed to be only continuous then f_+ may become unbounded. It can be shown [8] that to every set $E \subset [-\pi, \pi)$ of measure zero there exists a function $f \in \mathcal{C}(\mathbb{T})$ such that

$$\lim_{r \to 1} |f_+(r\mathrm{e}^{\mathrm{i}\tau})| = \infty \quad \text{for all } \tau \in E\ . \tag{6.16}$$

According to a classical result of Kolmogorov [55], it holds for every $f \in L^1(\mathbb{T})$ that the set $E \subset \mathbb{T}$ for which (6.16) holds is of measure zero. Thus, with regard to the divergence behavior toward $\mathbb{T}$, the Riesz projection of continuous functions behaves as bad as for L^1-functions. Even if f is absolute continuous, no improvement of this divergence behavior can be observed [11]. This bad convergence behavior of f_+ has important consequences for several applications. In the determination of the Wiener filter, for instance, this divergence of f_+ may result in an unbounded (i.e. unstable) Wiener filter if the given spectrum is only continuous [11]. All this emphasizes again the importance of the spaces $\mathcal{C}_\omega(\mathbb{T})$ of smooth functions in which such a complicated behavior of the causal projections does not occur.

6.4 Inner-Outer Factorization on Subspaces of H^∞

As an application of the previous results, this section considers the inner-outer factorization in subspaces of H^∞. Every $f \in H^\infty$ possesses a unique (up to a unitary constant) inner-outer factorization $f = f_I\, f_O$ into an inner function f_I and an outer function f_O which both belong again to H^∞ (cf. Sec. 2.2). However, if one considers subspaces of $\mathcal{A} \subset H^\infty$, this result may no longer hold. Of course any $f \in \mathcal{A} \subset H^\infty$ has an inner-outer factorization in H^∞ but the factors f_I and f_O do not belong to $\mathcal{A}$, in general. For example, if $f \in A(\mathbb{D}) \subset H^\infty$, its outer functions f_O may not belong to the disk algebra $A(\mathbb{D})$. In this section, we characterize subspaces $\mathcal{A} \subset H^\infty$ such that the inner and outer factors of every $f \in \mathcal{A}$ belong again to $\mathcal{A}$.

6.4.1 Scalar Case

The subspaces of H^∞ on which the inner-outer factorization exists, will be characterized in terms of two properties of certain Banach algebras which are formulated using Toeplitz operators. Given a function $\varphi \in L^\infty$, the *Toeplitz operator* $\mathcal{T}_\varphi$ associated with φ is defined by

$$(\mathcal{T}_\varphi f)(z) = \frac{1}{2\pi\mathrm{i}} \int_{\mathbb{T}} \frac{\varphi(\zeta) f(\zeta)}{\zeta - z}\, \mathrm{d}\zeta = (\mathfrak{P}_+ \varphi f)(z) \tag{6.17}$$

for all $f \in H^1$ and $z \in \mathbb{D}$, where $\mathfrak{P}_+$ is again the orthogonal Riesz-projection (6.5) from L^2 to H^2.

Definition 6.17 (I-property). *A subalgebra $\mathcal{A} \in H^\infty$ is said to have the* I-property *if all $f \in \mathcal{A}$ which are invertible in H^∞ are also invertible in $\mathcal{A}$, i.e. if from $f^{-1} \in H^\infty$ always follows that $f^{-1} \in \mathcal{A}$.*

Definition 6.18 (K-property). *A subalgebra $\mathcal{A} \subset H^\infty$ is said to have the* K-property *if for all $\varphi \in H^\infty$ hold that $\mathcal{T}_{\overline{\varphi}}(\mathcal{A}) \subset \mathcal{A}$. It is said that $\mathcal{A}$ has the* strong K-property *if in addition there exists a constant C such that*

$$\|\mathcal{T}_{\overline{\varphi}} f\|_{\mathcal{A}} \leq C \, \|\varphi\|_\infty \, \|f\|_{\mathcal{A}} \ . \tag{6.18}$$

The next two theorems show under which conditions on the majorant ω, the spaces $A_\omega(\mathbb{D})$ posses the I- and the K-property, respectively.

Theorem 6.19. *Let ω be an arbitrary majorant. Then $A_\omega(\mathbb{D})$ has the I-property.*

Theorem 6.20. *Let ω be a regular majorant. Then $A_\omega(\mathbb{D})$ has the strong K-property, i.e. for all $\varphi \in H^\infty$ and $f \in A_\omega(\mathbb{D})$ holds*

$$\|\mathcal{T}_{\overline{\varphi}} f\|_{A_\omega(\mathbb{D})} \leq C \, \|\varphi\|_{H^\infty} \, \|f\|_{A_\omega(\mathbb{D})} \ . \tag{6.19}$$

Proof (Theorem 6.19). We assume $f \in A_\omega(\mathbb{D})$ and $f^{-1} \in H^\infty$. This means that $\sup_{|z|<1} \left|f^{-1}(z)\right| < \infty$, which shows that there exists a constant $\delta > 0$ such that $|f(z)| \geq \delta$ for all $z \in \mathbb{D}$. It follows that there exists a constant C such that for all $z_1, z_2 \in \mathbb{D}$

$$\left|\frac{1}{f(z_1)} - \frac{1}{f(z_2)}\right| \leq \frac{1}{\delta^2} \, |f(z_1) - f(z_2)| \leq C\omega\left(|z_1 - z_2|\right)$$

using that $f \in A_\omega(\mathbb{D})$. This shows that indeed $f^{-1} \in A_\omega(\mathbb{D})$. □

Theorem 6.20, concerning the K-property of $A_\omega(\mathbb{D})$, will be proved using a technique which is called pseudoanalytic extension. To this end, we need some further notations. In what follows, $\mathbb{D}_- = \mathbb{C} \setminus \overline{\mathbb{D}}$ denotes the outside of the closed unit disk. Furthermore, let $z = x + \mathrm{i}y$ be a complex number, then we write $z^* := 1/\overline{z}$ for all $z \in \mathbb{C} \setminus \{0\}$. The Cauchy-Riemann differential operator is defined, as usual, by $\overline{\partial} := \frac{1}{2}(\partial/\partial x + \mathrm{i}\ \partial/\partial y)$, and $\mathcal{C}^1(\mathbb{D}_-)$ is the set of all continuously differentiable functions on $\mathbb{D}_-$.

The following lemma gives a characterization of the functions in $A_\omega(\mathbb{D})$, in terms of the existence of a so called pseudoanalytic extension of f onto the outside of the unit disk $\mathbb{D}_-$. The proof of this lemma is omitted here but it can be found in [32, Lemma 7].

Lemma 6.21 (Pseudoanalytic extension). *Let ω be a regular majorant and $f \in H^\infty$. Then $f \in A_\omega(\mathbb{D})$ if and only if there exists a function $F \in \mathcal{C}^1(\mathbb{D}_-) \cap L^\infty$ and a constant C such that*

$$\lim_{z \in \mathbb{D}_-, \ z \to \zeta} F(z) = f(\zeta) \ , \qquad a.e.\ \zeta \in \mathbb{T} \tag{6.20}$$

$$\left|\overline{\partial} F(z)\right| \leq C \, \frac{\omega\left(|z| - 1\right)}{|z| - 1} \ , \qquad \textit{for all } z \in \mathbb{D}_- \ . \tag{6.21}$$

Remark 6.22. By (6.20) F is an extension of the function f given on $\mathbb{T}$ to the outside of the unit disk. However, by (6.21) this extended function F is not (quite) analytic in $\mathbb{D}_-$ since $\overline{\partial} F \neq 0$. This is the reason for the notion of "‘pseudoanalytic extension"’. Such pseudoanalytic extensions also exist for other classes of smooth functions (see e.g. [34]).

With the help of the pseudoanalytic extension, we are able to prove the strong K-property of $A_\omega(\mathbb{D})$.

Proof (Theorem 6.20). Set $g := \mathcal{T}_{\overline{\varphi}} f$. It has to be shown that $g \in A_\omega(\mathbb{D})$ for any $f \in A_\omega(\mathbb{D})$. Since g is the orthogonal projection of $\overline{\varphi} f$ onto H^2, g can be written as

$$g = \overline{\varphi} f - \overline{h} \;, \qquad \text{a.e. on } \mathbb{T} \tag{6.22}$$

where $h \in H^2$ with $h(0) = 0$. Now let F be the pseudoanalytic extension of f according to Lemma 6.21, and define by $\Phi(z) := \overline{\varphi(z^*)}$ and $H(z) := \overline{h(z^*)}$ two analytic functions for $z \in \mathbb{D}_-$. Finally set

$$G(z) := F(z)\,\Phi(z) - H(z) \quad \text{for all } z \in \mathbb{D}_- \;.$$

We show that G is a pseudoanalytic extension of g according to Lemma 6.21. Clearly, G is continuously differentiable in $\mathbb{D}_-$ and by the properties of the Riesz-projector $\mathfrak{P}_+$, G is bounded in $\mathbb{D}_-(R) = \{z \in \mathbb{C} : |z| > R\}$ for all $R > 1$. Thus, $G \in \mathcal{C}^1(\mathbb{D}_-) \cap L^\infty(\mathbb{D}_-(R))$ and because of (6.22) and the definitions of F, Φ, and H it holds that $G|_{\mathbb{T}} = g|_{\mathbb{T}}$. Since Φ and H are analytic in $\mathbb{D}_-$, it follows that $\overline{\partial} G = \Phi \cdot \overline{\partial} F$ and therefore

$$\left|\overline{\partial} G(z)\right| \leq \|\Phi\|_{H^\infty(\mathbb{D}_-)} \cdot \left|\overline{\partial} F(z)\right| \leq c_1 \frac{\omega\left(|z| - 1\right)}{|z| - 1} \tag{6.23}$$

for all $z \in \mathbb{D}_-$ and with a certain constant c_1, where it was used that F satisfies (6.21). Thus, G satisfies the conditions of Lemma 6.21 which shows that $g \in A_\omega(\mathbb{D})$.

It remains to show that for $g = \mathcal{T}_{\overline{\varphi}} f$ the upper bound (6.19) exists. It follows from (6.23) that $g'(z) \leq \|\varphi\|_{H^\infty} |f'(z)|$ for all $z \in \mathbb{D}$, using that $\left|\overline{\partial} G(z)\right| = |g'(z^*)| \cdot |z^*|^2$. Now, we define

$$\|f\|_{\mathcal{PE}} := \sup_{|z|<1} |f'(z)| \frac{1 - |z|}{\omega(1 - |z|)} \;.$$

Therewith, it is clear that $\|g\|_{\mathcal{PE}} \leq \|\varphi\|_{H^\infty} \|f\|_{\mathcal{PE}}$. Moreover, it is clear that $|g(0)| \leq \|\varphi\|_{H^\infty} |f(0)|$. Therefore, it is sufficient to consider the case $|g(0)| = |f(0)| = 0$. It remains to show that for this case $\|f\|_{A_\omega(\mathbb{D})} \leq c_0 \|f\|_{\mathcal{PE}}$ for all $f \in A_\omega(\mathbb{D})$ and with a constant c_0 which is independent on f. Assume that $f \in A_\omega(\mathbb{D})$ is given with $\|f\|_{\mathcal{PE}} < \infty$, then it holds

$$|f'(z)| \leq \|f\|_{\mathcal{PE}} \frac{\omega(1 - |z|)}{1 - |z|} \;, \qquad \text{for all } z \in \mathbb{D} \;. \tag{6.24}$$

Now, we consider two points $z_1 = \mathrm{e}^{\mathrm{i}\theta_1}$ and $z_2 = \mathrm{e}^{\mathrm{i}\theta_2}$ on the unit circle with $|z_1 - z_2| \leq \delta$ and define $r_\delta := 1 - \delta$. For these two points holds

$$\begin{aligned}\left|f(\mathrm{e}^{\mathrm{i}\theta_1}) - f(\mathrm{e}^{\mathrm{i}\theta_2})\right| \leq \left|f(\mathrm{e}^{\mathrm{i}\theta_1}) - f(r_\delta \mathrm{e}^{\mathrm{i}\theta_1})\right| + \\ + \left|f(r_\delta \mathrm{e}^{\mathrm{i}\theta_1}) - f(r_\delta \mathrm{e}^{\mathrm{i}\theta_2})\right| + \left|f(r_\delta \mathrm{e}^{\mathrm{i}\theta_2}) - f(\mathrm{e}^{\mathrm{i}\theta_2})\right| \ . \end{aligned} \tag{6.25}$$

Using that f satisfies (6.24) one gets for the first term on the right hand side

$$\begin{aligned}\left|f(\mathrm{e}^{\mathrm{i}\theta_1}) - f(r_\delta \mathrm{e}^{\mathrm{i}\theta_1})\right| &\leq \int_{r_\delta}^1 \left|f'(\rho \mathrm{e}^{\mathrm{i}\theta_1})\right| \mathrm{d}\rho \leq \|f\|_{\mathcal{PE}} \int_{r_\delta}^1 \frac{\omega(1-\rho)}{1-\rho}\, \mathrm{d}\rho \\ &\leq c_2 \|f\|_{\mathcal{PE}}\, \omega(1 - r_\delta) = c_2 \|f\|_{\mathcal{PE}}\, \omega(\delta)\end{aligned}$$

where for the second line the weak regularity (1.12) of ω was used. The constant c_2 depends only on the majorant ω. Of course, the same relations holds also for the last term in (6.25). Similarly, for the middle term on the right side of (6.25) one gets

$$\begin{aligned}\left|f(r_\delta \mathrm{e}^{\mathrm{i}\theta_1}) - f(r_\delta \mathrm{e}^{\mathrm{i}\theta_2})\right| \leq \int_{\theta_2}^{\theta_1} \left|f'(r_\delta \mathrm{e}^{\mathrm{i}\theta})\right|\, \mathrm{d}\theta \leq \|f\|_{\mathcal{PE}} \frac{\omega\,(1-r_\delta)}{1-r_\delta} |\theta_2 - \theta_1| \\ \leq c_3\, \|f\|_{\mathcal{PE}}\, \omega(\delta)\ .\end{aligned}$$

using (6.24) and the inequality $|\theta_2 - \theta_1| \leq \frac{\pi}{\sqrt{2}} \left|\mathrm{e}^{\mathrm{i}\theta_2} - \mathrm{e}^{\mathrm{i}\theta_1}\right| \leq \frac{\pi}{\sqrt{2}}\, \delta$. Altogether, we get $\left|f(\mathrm{e}^{\mathrm{i}\theta_1}) - f(\mathrm{e}^{\mathrm{i}\theta_2})\right| \leq c_0\, \|f\|_{\mathcal{PE}}\, \omega(\delta)$ with a constant c_0 independent of f, which shows that indeed $\|f\|_{\mathcal{PE}} \leq c_0 \|f\|_{A_\omega(\mathbb{D})}$. □

The next theorem will show that the I- and the K-property are sufficient conditions for the existens of an inner-outer factorization on a certain Banach algebra $\mathcal{A}$.

Theorem 6.23. *Let $\mathcal{A} \subset H^\infty$ be a Banach algebra which has the I- and K-properties and let $f \in \mathcal{A}$. Then f has an inner-outer factorization $f = f_I\, f_O$ with $f_O, f_I \in \mathcal{A}$. If $\mathcal{A}$ also has the strong K-property, then there exists a constant C such that $\|f_O\|_{\mathcal{A}} \leq C\, \|f\|_{\mathcal{A}}$.*

Proof. Since $\mathcal{A} \subset H^\infty$, there exists an inner-outer factorization with $f_I, f_O \in H^\infty$. From the properties of these factors follows that $f_O(\zeta) = \overline{f_I}(\zeta)\, f(\zeta)$ for almost all $\zeta \in \mathbb{T}$ and that $\mathfrak{P}_+(f_O) = f_O$. Therewith, we get

$$f_O = \mathfrak{P}_+\left(\overline{f_I}\, f\right) = \mathcal{T}_{\overline{f_I}} f\ . \tag{6.26}$$

Since $f \in \mathcal{A}$ and because of the K-property of $\mathcal{A}$ this shows that $f_O \in \mathcal{A}$. The I-property shows that $f_O^{-1} \in \mathcal{A}$ and since $f_I = f_O^{-1} \cdot f$, it follows that $f_I \in \mathcal{A}$. The upper bound for $\|f_O\|_{\mathcal{A}}$ follows directly from (6.26) using that $\mathcal{A}$ has the strong K-property (6.18) and that $\|f_I\|_\infty = 1$. □

This result shows in particular that the mapping $f \mapsto f_O$ is bounded on every Banach algebra with the I- and the K-property.

6.4.2 Matrix functions

In the previous subsection we characterized sub-algebras $\mathcal{A}$ of H^∞ such that the inner and outer factor of an arbitrary function $f \in \mathcal{A}$ belongs again to $\mathcal{A}$. These investigation should now be extended to matrix valued functions $\mathbf{H} \in H^\infty(\mathbb{C}^{M\times N})$. Throughout this section, we use the notations introduced in Section 2.3 and 2.4.

First, we recall the general notion of inner- and outer functions for bounded analytic functions with values in the space $\mathscr{B}(\mathcal{H}_1, \mathcal{H}_2)$ of bounded operators from a Hilbert space $\mathcal{H}_1$ into a Hilbert space $\mathcal{H}_2$, as introduced in the previous section.

Definition 6.24 (Inner and outer functions). *Let $\mathcal{H}_1, \mathcal{H}_2$ be separable Hilbert spaces, and let $\mathbf{H} \in H^\infty(\mathcal{H}_1, \mathcal{H}_2)$ be a bounded analytic function with values in $\mathscr{B}(\mathcal{H}_1, \mathcal{H}_2)$. Then $\mathbf{H}$ is called*

(*a*) inner, *if $\mathbf{O}^+_{\mathbf{H}}$ is an isometry from $H^2(\mathcal{H}_1)$ into $H^2(\mathcal{H}_2)$.*
(*b*) outer, *if the image of $\mathbf{O}^+_{\mathbf{H}}$ on $H^2(\mathcal{H}_1)$ is dense in $H^2(\mathcal{H}_2)$, i.e. if*

$$\overline{\mathbf{O}^+_{\mathbf{H}}\, H^2(\mathcal{H}_1)} = H^2(\mathcal{H}_2)\,.$$

The definition of an inner function can be based directly on the analytic function $\mathbf{H}$ rather than on $\mathbf{O}_{\mathbf{H}}$.

Proposition 6.25. *The bounded analytic function $\mathbf{H} \in H^\infty(\mathcal{H}_1, \mathcal{H}_2)$ is an inner function if and only if $\mathbf{H}(\zeta)$ is an isometry from $\mathcal{H}_1$ into $\mathcal{H}_2$, i.e.*

$$\mathbf{H}^*(\zeta)\,\mathbf{H}(\zeta) = \mathbf{I}_{\mathcal{H}_1} \qquad \text{for almost all } \zeta \in \mathbb{T}\,. \tag{6.27}$$

Proof. We have to show that $\mathbf{O}^+_{\mathbf{H}}$ is an isometry $H^2(\mathcal{H}_1) \to H^2(\mathcal{H}_1)$ if and only if (6.27) holds. The "if part" is trivial. To prove the "only if part", we note first that if $\mathbf{O}^+_{\mathbf{H}}$ is an isometry, also $\mathbf{O}_{\mathbf{H}}$ will be an isometry. To see this, let

$$\boldsymbol{p}(\mathrm{e}^{\mathrm{i}\theta}) = \textstyle\sum_{k=N_1}^{N_2} \hat{\boldsymbol{p}}(k)\,\mathrm{e}^{\mathrm{i}k\theta} = \mathrm{e}^{-\mathrm{i}N_1\theta} \sum_{k=0}^{N_1+N_2} \hat{\boldsymbol{p}}(k-N_1)\,\mathrm{e}^{\mathrm{i}k\theta} = \mathrm{e}^{-\mathrm{i}N_1\theta}\,\boldsymbol{p}_c(\mathrm{e}^{\mathrm{i}\theta})$$

with $\hat{\boldsymbol{p}}(k) \in \mathcal{H}_1$ for all k, be a trigonometric polynomial in $L^2(\mathcal{H}_1)$. Then it is clear that $\boldsymbol{p}_c \in H^2(\mathcal{H}_1)$ and that $\|\boldsymbol{p}_c\|_{H^2(\mathcal{H}_1)} = \|\boldsymbol{p}\|_{L^2(\mathcal{H}_1)}$. An obvious calculation gives

$$\|\mathbf{O}_{\mathbf{H}}\boldsymbol{p}\|_{L^2(\mathcal{H}_2)} = \|\mathbf{O}_{\mathbf{H}}\boldsymbol{p}_c\|_{L^2(\mathcal{H}_2)} = \left\|\mathbf{O}^+_{\mathbf{H}}\boldsymbol{p}_c\right\|_{H^2(\mathcal{H}_2)} = \|\boldsymbol{p}\|_{L^2(\mathcal{H}_1)}\,.$$

Since the trigonometric polynomials are dense in $L^2(\mathcal{H}_1)$ this holds for all $\boldsymbol{f} \in L^2(\mathcal{H}_1)$.
Next, we choose the special function $\boldsymbol{f} = \chi_{[\tau,\tau+\delta]}(\theta)\,\boldsymbol{g}$ with $\boldsymbol{g} \in \mathcal{H}_1$ arbitrary, and where $\chi_{[\tau,\tau+\delta]}$ is the indicator function of the interval $[\tau, \tau+\delta]$. For this function, one obtains

$$\tfrac{1}{\delta}\int_{\tau}^{\tau+\delta} \|\mathbf{H}(\mathrm{e}^{\mathrm{i}\theta})\boldsymbol{g}\|_{\mathcal{H}_2}^2 \,\mathrm{d}\theta = \|\boldsymbol{g}\|_{\mathcal{H}_1}^2$$

which implies for $\delta \to 0$ that $\|\mathbf{H}(\mathrm{e}^{\mathrm{i}\theta})\boldsymbol{g}\|_{\mathcal{H}_2}^2 = \|\boldsymbol{g}\|_{\mathcal{H}_1}^2$ for almost all $\theta \in [-\pi, \pi)$. Therefore, $\mathbf{H}(\zeta)$ is an isometry for almost all $\zeta \in \mathbb{T}$. □

We start with the inner-outer factorization result in $H^\infty(\mathbb{C}^{M\times N})$, i.e. in the space of $M \times N$ matrices with entries from H^∞. For this space, it follows in particular from the above definition that every $\mathbf{H}_I \in H^\infty(\mathbb{C}^{N\times N})$ is an *inner* function if

$$\mathbf{H}_I^*(\zeta)\,\mathbf{H}_I(\zeta) = \mathbf{I}_N \qquad \text{for almost all } \zeta \in \mathbb{T}\ .$$

Moreover, it is clear that every outer function $\mathbf{H}_O \in H^\infty(\mathbb{C}^{N\times N})$ is invertible in $H^\infty(\mathbb{C}^{N\times N})$, which means that there exists an $\mathbf{H}_O^{-1} \in H^\infty(\mathbb{C}^{N\times N})$ such that $\mathbf{H}_O^{-1}(z)\,\mathbf{H}_O(z) = \mathbf{H}_O(z)\,\mathbf{H}_O^{-1}(z) = \mathbf{I}_N$ for all $z \in \mathbb{D}$.

Theorem 6.26. *Let* $\mathbf{H} \in H^\infty(\mathbb{C}^{M\times N})$ *and assume that there exists a* $\delta > 0$ *such that*

$$\mathbf{H}^*(z)\mathbf{H}(z) \geq \delta^2 \mathbf{I}_N \qquad \text{for all } z \in \mathbb{D}\ . \tag{6.28}$$

Then there exists an inner function $\mathbf{H}_I \in H^\infty(\mathbb{C}^{M\times N})$ *and an outer function* $\mathbf{H}_O \in H^\infty(\mathbb{C}^{N\times N})$ *such that*

$$\mathbf{H}(z) = \mathbf{H}_I(z)\,\mathbf{H}_O(z) \qquad \text{for all } z \in \mathbb{D}\ .$$

Moreover, let $\mathbf{H} \in H^\infty(\mathbb{C}^{M\times N})$ *and assume that there exists a* $\delta > 0$ *such that*

$$\mathbf{H}(z)\mathbf{H}^*(z) \geq \delta^2 \mathbf{I}_N \qquad \text{for all } z \in \mathbb{D}\ . \tag{6.29}$$

Then there exists an inner function $\mathbf{H}_I \in H^\infty(\mathbb{C}^{N\times M})$ *and an outer function* $\mathbf{H}_O \in H^\infty(\mathbb{C}^{M\times M})$ *such that*

$$\mathbf{H}(z) = \mathbf{H}_O(z)\,\mathbf{H}_I(z) \qquad \text{for all } z \in \mathbb{D}\ .$$

These factorizations are unique up to a constant unitary matrix.

Note that $M \geq N$ is a necessary requirement for (6.28) to hold. Similarly, (6.29) can be satisfied only if $N \leq M$. Subsequently, we consider only the case that $M \geq N$, but it will always be clear how the argumentation has to be changed for the case $M \leq N$.

What happens, if the entries of the matrix $\mathbf{H}$ belongs to a certain subalgebra $\mathcal{A} \subset H^\infty$? By the previous theorem, it is clear that there exists an inner-outer factorization of $\mathbf{H}$ with $\mathbf{H}_I \in H^\infty(\mathbb{C}^{M\times N})$ and with $\mathbf{H}_O \in H^\infty(\mathbb{C}^{N\times N})$. But do the entries of the inner and outer functions belong again to $\mathcal{A} \in H^\infty$? In general, the answer will be negative. However, there exist subspaces $\mathcal{A}$ of H^∞ such that all entries of the inner and outer factors belong again to $\mathcal{A}$. We will show in the following that Banach algebras $\mathcal{A}$ with the I and the K properties (cf. Def. 6.17 and 6.18) are such subspaces of H^∞. Matrices of the size $M \times N$ with entries of a certain algebra $\mathcal{A}$ will be denoted in the following by $\mathcal{A}^{M\times N}$. With this notations, it is clear that $(H^\infty)^{M\times N} = H^\infty(\mathbb{C}^{M\times N})$.

As a preparation, we present an auxiliary lemma, which shows that for subalgebras $\mathcal{A} \subset H^\infty$ with the I-property every quadratic matrix $\mathbf{H} \in \mathcal{A}^{M\times M}$ which is invertible in H^∞ is also invertible in $\mathcal{A}$.

Lemma 6.27. *Let $\mathcal{A}$ be a subalgebra of H^∞ which has the I-property and let $\mathbf{H} \in \mathcal{A}^{N\times N}$. If $\mathbf{H}^{-1} \in (H^\infty)^{N\times N}$ then $\mathbf{H}^{-1} \in \mathcal{A}^{N\times N}$ for all $N \geq 1$.*

Proof. Assume that the matrix $\mathbf{H} \in \mathcal{A}^{N\times N}$ is invertible in $(H^\infty)^{N\times N}$ then its inverse can be written as $\mathbf{H}^{-1} = (\det \mathbf{H})^{-1}\, \mathbf{H}_{\text{adj}}$. It is clear that for the adjunct matrix holds $\mathbf{H}_{\text{adj}} \in \mathcal{A}^{N\times N}$ since the entries of $\mathbf{H}_{\text{adj}}$ are products and sums of functions from $\mathcal{A}$. Because $\mathbf{H}^{-1} \in (H^\infty)^{N\times N}$, we have $(\det \mathbf{H})^{-1} \in H^\infty$ and from the I-property of $\mathcal{A}$ follows that $(\det \mathbf{H})^{-1} \in \mathcal{A}$. Altogether we have $\mathbf{H}^{-1} \in \mathcal{A}^{N\times N}$. □

Next we state the result on the inner-outer factorization in subalgebras of H^∞ with the I- and the (strong) K-property.

Theorem 6.28. *Let $\mathcal{A}$ be a subalgebra of H^∞ which has the I- and the K-property. Then for all $\mathbf{H} \in \mathcal{A}^{M\times N}$ for which there exists an $\delta > 0$ such that*

$$\mathbf{H}^*(z)\mathbf{H}(z) \geq \delta^2 \mathbf{I}_{\mathbb{C}^N}\ , \quad \text{for all } z \in \mathbb{D} \tag{6.30}$$

there exists a (left) inner-outer factorization $\mathbf{H} = \mathbf{H}_I\, \mathbf{H}_O$ in $\mathcal{A}$ which is unique up to a constant unitary matrix, with the inner factor $\mathbf{H}_I \in \mathcal{A}^{M\times N}$ and the outer factor $\mathbf{H}_O \in \mathcal{A}^{N\times N}$. If $\mathcal{A}$ has additionally the strong K-property, then there exists a constant C such that

$$\left\| [\mathbf{H}_O]_{m,n} \right\|_{\mathcal{A}} \leq C \cdot \textstyle\sum_{k=1}^{M} \left\| [\mathbf{H}]_{k,n} \right\|_{\mathcal{A}} \tag{6.31}$$

for all $m = 1, 2, \ldots, M$ and all $n = 1, 2, \ldots, N$.

Proof. Since $\mathcal{A} \subset H^\infty$ the matrix $\mathbf{H}$ belongs to $(H^\infty)^{M\times N}$ and therefore there exists an inner-outer factorization $\mathbf{H} = \mathbf{H}_I\, \mathbf{H}_O$ with $\mathbf{H}_I \in (H^\infty)^{M\times N}$, for which $\mathbf{H}_I^*(\zeta)\mathbf{H}_I(\zeta) = \mathbf{I}_{\mathbb{C}^N}$ a.e. $\zeta \in \mathbb{T}$, and with the matrix $\mathbf{H}_O \in (H^\infty)^{N\times N}$ which is invertible in H^∞. Since this factorization is unique in H^∞ (up to a constant unitary matrix), it is unique (if it exists) in every subalgebra of H^∞. For $\zeta \in \mathbb{T}$ follows $\mathbf{H}_O(\zeta) = \mathbf{H}_I^*(\zeta)\,\mathbf{H}(\zeta)$, or in more detail

$$[\mathbf{H}_O]_{m,n} = \textstyle\sum_{k=1}^{N} [\mathbf{H}_I^*]_{m,k}\, [\mathbf{H}]_{k,n}\ , \quad 1 \leq m, n \leq N\ .$$

By the properties of the outer function holds $[\mathbf{H}_O]_{m,n} = \mathfrak{P}_+([\mathbf{H}_O]_{m,n})$ and therefore follows that

$$\begin{aligned}[\mathbf{H}_O]_{m,n} &= \textstyle\sum_{k=1}^{N} \mathfrak{P}_+\left([\mathbf{H}_I^*]_{m,k}\, [\mathbf{H}]_{k,n}\right) \\ &= \textstyle\sum_{k=1}^{N} \mathcal{T}_{[\mathbf{H}_I^*]_{m,k}}\left([\mathbf{H}]_{k,n}\right)\ .\end{aligned} \tag{6.32}$$

Since all $[\mathbf{H}]_{k,n} \in \mathcal{A}$ and since $\mathcal{A}$ has the K-property, it follows that $[\mathbf{H}_O]_{m,n} \in \mathcal{A}$ and therefore $\mathbf{H}_O \in \mathcal{A}^{N\times N}$. Since $\mathbf{H}_O$ is invertible in H^∞, Lemma 6.27 shows that $\mathbf{H}_O^{-1} \in \mathcal{A}^{N\times N}$. Since $\mathbf{H}_I = \mathbf{H}\,\mathbf{H}_O^{-1}$ it follows finally that $\mathbf{H}_I \in \mathcal{A}^{M\times N}$.
Using that $\|[\mathbf{H}_I]_{m,n}\|_\infty = 1$ for all m, n and applying Def. 6.18 of the strong K-property, the upper bound (6.31) follows at once from (6.32). □

Note that the previous theorem states in particular that the mapping $\mathbf{H} \mapsto \mathbf{H}_O$ is bounded (if it exists) in every Banach algebra with the I- and the strong $K-$property. It was proved in Theorem 6.19 and Theorem 6.19 that the spaces $A_\omega(\mathbb{D})$ of smooth and analytic functions in $\mathbb{D}$ possesses the I as well as the strong K-property, provided that the majorant ω is regular. Therefore, Theorem 6.28 holds in particular for $\mathcal{A} = A_\omega(\mathbb{D})$ with a regular majorant ω.

Notes

The result of M. Riesz on the boundedness of the Riesz projection on L^p can be found e.g. in [48, 92]. Theorem 6.11 is a variant of results for polynomial operations (see e.g. [54, Chapter VII]). Equation (6.6) is a variant of the *Zygmund-Martsinkevich-Berman identity* for polynomial operations which was proved in [6] and the related result to (6.7) was established in [59].

More on the pseudoanalytic extension, used in Subsection 6.4 can be found in [34]. Lemma 6.21 is due to Dyakonov [32].

Part III

Causality Aspects in Signal and System Theory

7

Disk Algebra Bases

This chapter considers the approximation of the transfer function of a linear system in terms of a causal filter bank. Assume that $f(\mathrm{e}^{\mathrm{i}\theta})$ with $\theta \in [-\pi, \pi)$ is the transfer function of an arbitrary discrete-time linear system $\mathcal{L}$, and let $\Phi = \{\varphi_k\}_{k=1}^{\infty}$ be a set of transfer functions of an orthonormal filterbank. It is assumed that f as well as all φ_k are elements of a certain Banach algebra $\mathcal{B}$ which characterizes the system theoretical properties of $\mathcal{L}$ and of the filterbank Φ. Moreover, since f as well as $\{\varphi_k\}_{k=1}^{\infty}$ should represent causal systems, these transfer functions have to belong to the causal subspace $\mathcal{B}_+$ of $\mathcal{B}$. Then, it is desirable to obtain an approximation of f in this filterbank of the form

$$(\mathfrak{A}_N f)(\mathrm{e}^{\mathrm{i}\theta}) = \sum_{k=1}^{N} c_k(f)\, \varphi_k(\mathrm{e}^{\mathrm{i}\theta}) \,, \qquad \theta \in [-\pi, \pi) \tag{7.1}$$

with constants $c_k(f)$ which are uniquely determined by f. From this approximation, we require that $\mathfrak{A}_N f \in \mathcal{B}_+$ represents the transfer function of a causal and stable linear system, and we wish that every causal and stable transfer function $f \in \mathcal{B}_+$ can be approximated arbitrarily well in the filterbank Φ, i.e.

$$\|f - \mathfrak{A}_N f\|_{\mathcal{B}} \to 0 \quad \text{as} \quad N \to \infty \,. \tag{7.2}$$

If such a system $\Phi = \{\varphi_k\}_{k=1}^{\infty}$ of functions in $\mathcal{B}$ exists such that (7.2) holds for every $f \in \mathcal{B}_+$, one says that Φ is a basis in $\mathcal{B}_+$.

Here, we consider almost exclusively the approximation of causal and energy stable linear systems, i.e. the case $\mathcal{B} = H^\infty$. However, since H^∞ is not a separable space it is clear that no basis exists in H^∞. Therefore we consider the disk algebra $A(\mathbb{D}) \subset H^\infty$ which is equal to the closure of all polynomials in H^∞-norm and which is separable.

This chapter considers three problems connected with bases in the disk algebra. First, the question of the existence of a disk algebra basis is discussed, then the robustness of such basis expansions with respect to errors in the given data is analyzed, and finally the existence of uniformly stable bases is discussed.

V. Pohl, H. Boche, *Advanced Topics in System and Signal Theory*, Foundations in Signal Processing, Communications and Networking 4,
DOI 10.1007/978-3-642-03639-2_7, © Springer-Verlag Berlin Heidelberg 2009

7.1 On the Existence of Disk Algebra Bases

Definition 7.1 (Schauder Base). *Let $\mathcal{B}$ be an arbitrary Banach space. A sequence $\{\varphi_k\}_{k=1}^{\infty}$ of elements of $\mathcal{B}$ is called a (*Schauder*) *basis *for $\mathcal{B}$ if for every $f \in \mathcal{B}$ there exists a unique sequence $\{\eta_k\}_{k=1}^{\infty}$ of complex numbers such that*

$$f = \sum_{k=1}^{\infty} \eta_k \, \varphi_k \tag{7.3}$$

where the equality sign means that the sum converges to f in the norm of the Banach space $\mathcal{B}$.

The coefficients η_k in the basis representation (7.3) are given by bounded linear functionals $c_k : f \mapsto \eta_k$. This important property of Schauder bases is proved in the following proposition.

Proposition 7.2. *Let $\{\varphi_k\}_{k=1}^{\infty}$ be a Schauder basis in a Banach space $\mathcal{B}$. Then the coefficients η_k in the basis representation* (7.3) *are given by*

$$\eta_k = c_k(f) \ , \quad k = 1, 2, \dots \quad \textit{and for all } f \in \mathcal{B}$$

where $c_k \in \mathcal{B}^*$, $k = 1, 2, \dots$ are bounded linear functionals on $\mathcal{B}$. Moreover the sequences $\{c_k\}_{k=1}^{\infty}$ and $\{\varphi_k\}_{k=1}^{\infty}$ are biorthogonal, *which means that*

$$c_k(\varphi_n) = \begin{cases} 1 & k = n \\ 0 & k \neq n \end{cases} . \tag{7.4}$$

Proof. Denote by $\mathcal{S}$ the set of all sequences $\eta = \{\eta_k\}_{k=1}^{\infty}$ for which the series (7.3) converges in $\mathcal{B}$, and define by

$$\|\eta\|_{\mathcal{S}} := \sup\nolimits_{N \geq 1} \| \textstyle\sum_{k=1}^{N} \eta_k \, \varphi_k \|_{\mathcal{B}}$$

a norm in $\mathcal{S}$. It is easily verified that $\mathcal{S}$ is complete with respect to this norm. Consider now the linear mapping $M : \mathcal{S} \to \mathcal{B}$ defined by $M\eta = \sum_{k=1}^{\infty} \eta_k \varphi_k$ for every $\eta \in \mathcal{S}$. By the definition of a Schauder base and the set $\mathcal{S}$, the mapping M is one-to-one and onto, and by the definition of the norm $\| \cdot \|_{\mathcal{S}}$, it is clear that $\|M\eta\|_{\mathcal{B}} \leq \|\eta\|_{\mathcal{S}}$, i.e. M is bounded. Consequently, the open mapping theorem shows that also the inverse $M^{-1} : \mathcal{B} \to \mathcal{S}$ is a bounded linear operator.

Now for an arbitrary $k \in \mathbb{N}$ the linear functional c_k on $\mathcal{B}$ defined by

$$c_k(f) := \eta_k \quad \text{where } f = \textstyle\sum_{k=1}^{\infty} \eta_k \, \varphi_k$$

is considered. Since $\eta_n \, \varphi_n = f - \sum_{k \neq n} \eta_k \, \varphi_k$ it follows easily that $\|\eta_n \, \varphi_n\|_{\mathcal{B}} \leq 2 \, \|f\|_{\mathcal{B}} \leq 2 \, \|\eta\|_{\mathcal{S}}$. Therewith, one gets

$$|c_n(f)| = |\eta_n| = \frac{\|\eta_n \, \varphi_n\|_{\mathcal{B}}}{\|\varphi_n\|_{\mathcal{B}}} \leq \frac{2\|\eta\|_{\mathcal{S}}}{\|\varphi_n\|_{\mathcal{B}}} \leq \frac{2\|M^{-1}\|}{\|\varphi_n\|_{\mathcal{B}}} \, \|f\|_{\mathcal{B}} \tag{7.5}$$

which shows that the linear functionals $\{c_k\}_{k=1}^{\infty}$ are bounded. Thus, we have

$$f = \sum_{k=1}^{\infty} c_k(f)\,\varphi_k \quad \text{for all } f \in \mathcal{B} \tag{7.6}$$

where the sequence $\{c_k\}_{k=1}^{\infty}$ of bounded linear functionals of $\mathcal{B}$ is obviously uniquely determined. Because of the uniqueness of the representation (7.6), it follows from $c_n(f) = \sum_{k=1}^{\infty} c_k(f)\,c_n(\varphi_k)$ that (7.4) holds. □

Example 7.3 (Classical Fourier series). Let $\mathcal{B} = L^p$ with $1 < p < \infty$ the set of all Lebesgue integrable functions on the unit circle. The sequence of functions $\varphi_k(\mathrm{e}^{\mathrm{i}\theta}) = \frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{\mathrm{i}k\theta}$ with $k = 0, \pm 1, \pm 2, \dots$ forms a basis of $\mathcal{B}$ and the biorthogonal coefficient functionals $c_n \in \mathcal{B}^*$ are given by

$$c_k(f) = \tfrac{1}{2\pi} \int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\theta})\,\mathrm{e}^{-\mathrm{i}k\theta}\,\mathrm{d}\theta\ .$$

Whereas the classical Fourier series of the previous example is a basis for many function spaces, it is not a basis for some other important spaces, e.g. L^1. The Fourier series is also not a basis for $\mathcal{C}(\mathbb{T})$, because it is well known that there exist continuous functions on $\mathbb{T}$ for which the classical Fourier series fails to converge at some points on the unit circle.

Example 7.4 (Faber-Schauder system). Define the continuous functions

$$\phi_0(\tau) := \begin{cases} 1 + \frac{\tau}{\pi}\,, & \tau \in [-\pi, 0) \\ 1 - \frac{\tau}{\pi}\,, & \tau \in [0, \pi] \\ 0\,, & \text{otherwise} \end{cases}$$

$$\phi_n(\tau) := \phi_0(2^n\,\tau)\,, \qquad n = 1, 2, \dots\ .$$

Therewith, one defines *Faber-Schauder system* $\{\xi_n\}_{n=0}^{\infty}$

$$\xi_0(\mathrm{e}^{\mathrm{i}\theta}) := 1$$

$$\xi_n(\mathrm{e}^{\mathrm{i}\theta}) := \phi_k\left(\theta - \pi + \tfrac{2l+1}{2^k}\,\pi\right) \quad \text{with } n = 2^k + l$$

where $k = 0, 1, 2, \dots$ and $0 \le l < 2^k$. It is well known that this Faber-Schauder system $\{\xi_n\}_{n=0}^{\infty}$ is a basis for $\mathcal{C}(\mathbb{T})$. In fact, the Faber-Schauder system was the first example of a basis for the space of continuous functions [36, 79]. Note that the definition given above is just one possible way to define a Faber-Schauder system.

Every function $\varphi \in \mathcal{C}(\mathbb{T})$ defines by

$$\langle f, \varphi \rangle := \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\theta})\,\overline{\varphi(\mathrm{e}^{\mathrm{i}\theta})}\,\mathrm{d}\theta \qquad \text{for all } f \in \mathcal{C}(\mathbb{T}) \tag{7.7}$$

a linear functional on $\mathcal{C}(\mathbb{T})$. A basis $\{\varphi_n\}_{n=1}^{\infty}$ in $\mathcal{C}(\mathbb{T})$ is called an *orthogonal basis* in $\mathcal{C}(\mathbb{T})$ if the basis functions $\{\varphi_n\}$ define by (7.7) a sequence of coefficient functionals $c_n(f) = \langle f, \varphi_n \rangle$ such that

$$f = \sum_{n=1}^{\infty} \langle f, \varphi_n \rangle \, \varphi_n \qquad \text{for all } f \in \mathcal{C}(\mathbb{T}) \; . \tag{7.8}$$

Since the sequence of coefficient functionals $\{c_n\}_{n\in\mathbb{N}}$ is biorthogonal to the sequence of basis functions $\{\varphi_n\}_{n\in\mathbb{N}}$, it holds for every orthonormal basis in $\mathcal{C}(\mathbb{T})$ that

$$\langle \varphi_n, \varphi_k \rangle = \begin{cases} 1 & n = k \\ 0 & n \neq k \end{cases} \; .$$

Example 7.5 (Franklin system). Take the Faber-Schauder system $\{\xi_n\}_{n=0}^{\infty}$ defined in Example 7.4 and apply the Gram-Schmidt orthogonalization to this system. This gives the orthogonal system $\{\psi_n\}_{n=0}^{\infty}$, where

$$\begin{aligned} \psi_0(\mathrm{e}^{\mathrm{i}\theta}) &= \xi_0(\mathrm{e}^{\mathrm{i}\theta}) \\ \psi_{n+1}(\mathrm{e}^{\mathrm{i}\theta}) &= \left[\xi_{n+1} - \sum_{k=0}^{n} \langle \xi_{n+1}, \psi_k \rangle \, \psi_k \right] / \sqrt{\langle \xi_{n+1}, \xi_{n+1} \rangle} \; . \end{aligned}$$

The so defined *Franklin system* $\Psi = \{\psi_n\}_{n=0}^{\infty}$ is an orthogonal basis for $\mathcal{C}(\mathbb{T})$ [39], and it is the basis for the following considerations in this section.

For an arbitrary Banach space $\mathcal{B}$, it is generally not easy to determine whether or not there exists a basis for $\mathcal{B}$. Here, we consider the special case $\mathcal{B} = A(\mathbb{D})$, i.e. we consider bases in the disk algebra $A(\mathbb{D})$. The question whether there exists a basis in the disk algebra was posed by Banach [5] and an affirmative answer was given by Bockarev [23] forty years later. Since $A(\mathbb{D}) \subset \mathcal{C}(\mathbb{T})$, it is not surprising that the basis in $A(\mathbb{D})$ can be constructed from the basis in $\mathcal{C}(\mathbb{T})$ by an analytic continuation inside the unit disk.

7.2 Robust Approximation in Disk Algebra Bases

As mentioned at the end of the previous section, there exist orthogonal bases in the disk algebra $A(\mathbb{D})$. Given such a basis $\Phi = \{\varphi_k\}_{k=1}^{\infty}$, it is possible to approximate every $f \in A(\mathbb{D})$ by the partial sum

$$(\mathfrak{A}_N f)(z) = \sum_{k=1}^{N} \langle f, \varphi_k \rangle \, \varphi_k(z) \, , \qquad z \in \overline{\mathbb{D}} \tag{7.9}$$

such that the approximation error $\|\mathfrak{A}_N f - f\|_\infty$ can be made arbitrary small by increasing the approximation degree N. In this section, we investigate the robustness behavior of such basis expansions in $A(\mathbb{D})$. However, first we will discuss some general properties of such basis approximations, which will be needed subsequently.

Properties of basis expansions in $A(\mathbb{D})$

1. Let $\Phi = \{\varphi_k\}_{k=1}^{\infty}$ be an arbitrary orthonormal basis in $A(\mathbb{D})$. Since for every $f \in A(\mathbb{D})$ it always holds that $\|f\|_2 \leq \|f\|_\infty$. It follows that $\{\varphi_k\}_{k=1}^{\infty}$ is orthogonal sequence in H^2. Moreover, since $A(\mathbb{D})$ is dense in H^2 it follows that Φ is a complete orthonormal sequence in H^2. Thus Φ is also a base for H^2.
2. Let f be an arbitrary function in L^2 and let $f = f_+ + f_-$ with $f_+ = \mathfrak{P}_+ f$ its decomposition into its causal and anticausal part. We already know that the Riesz projection $\mathfrak{P}_+ : f \mapsto f_+$ is a bounded operator from L^2 onto H^2 (cf. Theorem 6.12). Moreover, a simple calculation shows that $\mathfrak{P}_+$ is a selfadjoint operator on L^2, i.e. $\langle \mathfrak{P}_+ f, g\rangle_2 = \langle f, \mathfrak{P}_+ g\rangle_2$ for all $f, g \in L^2$.
3. Let $\Phi = \{\varphi_k\}_{k=1}^{\infty}$ be an orthogonal base in $A(\mathbb{D})$, then the approximation operator $\mathfrak{A}_N$ is an orthogonal projection from L^2 onto the subspace of $A(\mathbb{D})$ which is spanned by $\varphi_1, \cdots, \varphi_N$. Since this subspace is contained in the range of $\mathfrak{P}_+$ (which is equal to H^2), the relation $\mathfrak{P}_+ \mathfrak{A}_N f = \mathfrak{A}_N f$ holds for all $f \in L^2$. Moreover, because $\mathfrak{P}_+$ is selfadjoint, one gets
$$\langle f_+, \varphi_k\rangle_2 = \langle \mathfrak{P}_+ f, \varphi_k\rangle_2 = \langle f, \mathfrak{P}_+ \varphi_k\rangle_2 = \langle f, \varphi_k\rangle_2 \ .$$
Using this relation in the representation (7.9) of $\mathfrak{A}_N f$ and $\mathfrak{A}_N f_+$ one obtains also that $\mathfrak{A}_N f = \mathfrak{A}_N \mathfrak{P}_+ f$ for all $f \in L^2$. Thus, altogether
$$\mathfrak{P}_+ \mathfrak{A}_N f = \mathfrak{A}_N \mathfrak{P}_+ f = \mathfrak{A}_N f \qquad \text{for all } f \in L^2 \ . \tag{7.10}$$
Note that this relation implies immediately that $\mathfrak{A}_N f_- = 0$ for all $f \in L^2$.
4. A function in H^2 need not belong to H^∞, in general. However, all functions in H^2 are bounded on any disk with a radius smaller than 1. This follows from the following inequality which holds for all $f \in H^2$.
$$|f(z)| \leq \frac{\|f\|_2}{\sqrt{1 - |z|^2}} \qquad \text{for } |z| < 1 \ . \tag{7.11}$$
To see this, write f as its Fourier series $f(z) = \sum_{k=0}^{\infty} \hat{f}(k)\, z^k$. Applying the Cauchy-Schwartz inequality and the Parseval's identity yields
$$|f(z)|^2 \leq \left(\sum_{k=0}^{\infty} \left|\hat{f}_k\right|^2\right) \left(\sum_{k=0}^{\infty} |z|^{2k}\right) = \|f\|_2^2 \, \frac{1}{1 - |z|^2} \qquad \text{for all } |z| < 1 \ .$$

Robustness behavior

Assume that the actual function $f \in A(\mathbb{D})$ is disturbed by an error $g \in \mathcal{C}(\mathbb{T})$ such that only the function $\widetilde{f} = f + g$ is known. Then for the approximation error holds
$$\|\mathfrak{A}_N \widetilde{f} - f\|_\infty \geq \big|\, \|\mathfrak{A}_N g\|_\infty - \|\mathfrak{A}_N f - f\|_\infty \big| \ . \tag{7.12}$$

The second term can be made as small as desired since $\Phi = \{\varphi_n\}_{n=1}^{\infty}$ is a basis in $A(\mathbb{D})$ whereas the first term, due to the disturbance g, will remain larger than zero as long as $g \neq 0$. The question, we want to investigate in the following is whether the residual error $\|\mathfrak{A}_N\, g\|_\infty$ remains bounded for every $g \in \mathcal{C}(\mathbb{T})$, i.e. whether $\mathfrak{A}_N$ satisfies

$$\sup_{N\in\mathbb{N}} \|\mathfrak{A}_N\, g\|_\infty < \infty \qquad \text{for all } g \in \mathcal{C}(\mathbb{T})\,. \tag{7.13}$$

If this condition is satisfied, we will say that the basis Φ allows a robust approximation of any $f \in A(\mathbb{D})$. Otherwise the approximation error (7.12) might get unbounded for some disturbances $g \in \mathcal{C}(\mathbb{T})$. By the *theorem of Banach-Steinhaus* condition (7.13) is satisfied if and only if the operators $\mathfrak{A}_N : \mathcal{C}(\mathbb{T}) \to A(\mathbb{D})$ are uniformly bounded, i.e. if there is a constant $C < \infty$ such that

$$\|\mathfrak{A}_N\|_{\mathcal{C}(\mathbb{T})\to A(\mathbb{D})} = \sup_{\substack{g\in\mathcal{C}(\mathbb{T})\\ \|g\|_\infty\leq 1}} \|\mathfrak{A}_N\, g\|_\infty \leq C \qquad \text{for all } N \in \mathbb{N}\,.$$

Thus we have to analyze the norm of the approximation operators $\mathfrak{A}_N$. To this end, the coefficient functionals (7.7) are inserted in the definition (7.9) of the approximation operators. This gives the following integral representation of the approximation operators

$$(\mathfrak{A}_N\, f)(z) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\theta})\ K_N(\mathrm{e}^{\mathrm{i}\theta}, z)\,\mathrm{d}\theta \tag{7.14}$$

with the kernel

$$K_N(\zeta, z) = \sum_{k=1}^{N} \varphi_k(z)\,\overline{\varphi_k(\zeta)} \tag{7.15}$$

corresponding to the given disk algebra basis $\Phi = \{\varphi_k\}_{k=1}^{\infty}$. The numbers

$$L_N := \sup_{|z|<1}\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \left|K_N(\mathrm{e}^{\mathrm{i}\theta}, z)\right|\,\mathrm{d}\theta\right) = \sup_{|z|<1} \|K_N(\cdot, z)\|_1\,, \qquad N \in \mathbb{N}\,.$$

are the so-called *Lebesgue constants* of the basis Φ. These Lebesgue constants play an important role in our considerations because of the following result.

Proposition 7.6. *The operator norm* $\|\mathfrak{A}_N\|_{\mathcal{C}(\mathbb{T})\to A(\mathbb{D})}$ *of the basis expansion* (7.9) *is given by the Lebesgue constant of the basis* Φ, *i.e.*

$$\|\mathfrak{A}_N\|_{\mathcal{C}(\mathbb{T})\to A(\mathbb{D})} = \sup_{|z|<1} \|K_N(\cdot, z)\|_1\,.$$

Proof. From the integral representation (7.14) of the approximation operator follows at once that

$$
\begin{aligned}
\|\mathfrak{A}_N\|_{\mathcal{C}(\mathbb{T})\to A(\mathbb{D})} &= \sup_{|z|<1} \sup_{\substack{f\in\mathcal{C}(\mathbb{T})\\ \|f\|_\infty\le 1}} |(\mathfrak{A}_N f)(z)| \\
&\le \sup_{|z|<1} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \left|K_N(\mathrm{e}^{\mathrm{i}\theta}, z)\right| \mathrm{d}\theta \right) \qquad (7.16)\\
&= \sup_{|z|<1} \|K_N(\cdot, z)\|_1 \,.
\end{aligned}
$$

To see that there exists an $f \in \mathcal{C}(\mathbb{T})$ with $\|f\|_1 \le 1$ for which equality holds, write the kernel as $K_N(\mathrm{e}^{\mathrm{i}\theta}, z) = \left|K_N(\mathrm{e}^{\mathrm{i}\theta}, z)\right| \mathrm{e}^{\mathrm{i}\phi(\theta,z)}$ and set $f(\mathrm{e}^{\mathrm{i}\theta}) = \mathrm{e}^{-\mathrm{i}\phi(\theta,z)}$. Since all basis function φ_k are continuous, also the kernel K_N and the above defined f are continuous in θ for every $z \in \mathbb{D}$. Using this f one gets $|(\mathfrak{A}_N f)(z)| = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left|K_N(\mathrm{e}^{\mathrm{i}\theta}, z)\right| \mathrm{d}\theta$ which shows that equality holds in (7.16). □

Thus to assess whether a certain basis $\Phi = \{\varphi_k\}_{k=1}^{\infty}$ allows a robust approximation in the form (7.9), one has to investigate its Lebesgue constants. The next theorem will show that the Lebesgue constants L_N of every arbitrary orthogonal basis in $A(\mathbb{D})$ diverge as N tends to infinity.

Theorem 7.7. *Let $\Phi = \{\varphi_k\}_{k=1}^{\infty}$ be an arbitrary orthogonal basis in $A(\mathbb{D})$ with the kernel K_N given by* (7.15)*, then*

$$\lim_{N\to\infty} \sup_{|z|<1} \|K_N(\cdot, z)\|_1 = \infty \,. \qquad (7.17)$$

Proof. The theorem is proved by contradiction in three steps.

1) Contrary to the statement of the theorem, it is assumed that (7.17) does not hold. Thus, by going to a subsequence if necessary, we may suppose without loss of generality that there exists a constant C_1 such that

$$\sup_{|z|<1} \frac{1}{2\pi} \int_{-\pi}^{\pi} \left|K_N(\mathrm{e}^{\mathrm{i}\theta}, z)\right| \mathrm{d}\theta \le C_1 \qquad (7.18)$$

for all $N \in \mathbb{N}$. Then Proposition 7.6 implies that the norms of the operators $\mathfrak{A}_N : \mathcal{C}(\mathbb{T}) \to A(\mathbb{D})$ are uniformly bounded by C_1: $\|\mathfrak{A}_N\|_{\mathcal{C}(\mathbb{T})\to A(\mathbb{D})} \le C_1$.

2) Let f be an arbitrary function in L^2. Since $\{\varphi_k\}_{k\in\mathbb{N}}$ is a basis in H^2 the sequence $\mathfrak{A}_N f_+ = \mathfrak{A}_N \mathfrak{P}_+ f$ converges to $f_+ = \mathfrak{P}_+ f$ in H^2. Because of (7.10), the relation $\mathfrak{A}_N f = \mathfrak{A}_N f_+$ holds such that

$$\lim_{N\to\infty} \|\mathfrak{A}_N f - f_+\|_2 = 0 \,.$$

Define the sequence $g_N := \mathfrak{A}_N f - f_+$ in H^2 which converges to zero (in H^2) as N tends to infinity. Let $0 < R < 1$, then (7.11) shows that

$$\sup_{|z|\le R} |g_N(z)| \le \frac{\|g_N\|_2}{\sqrt{1 - R^2}}$$

and since $\|g_N\|_2$ converges to zero, it follows $\lim_{N\to\infty} \sup_{|z|<R} |g_N(z)| = 0$ which finally shows that

$$\lim_{N\to\infty} (\mathfrak{A}_N f)(z) = f_+(z) \quad \text{for all } |z| < 1 \ . \tag{7.19}$$

Thus for every $f \in L^2$, the approximation $(\mathfrak{A}_N f)(z)$ converges to the causal part $f_+(z)$ at all points z inside the unit disk.

3) In particular (7.19) holds for all $f \in \mathcal{C}(\mathbb{T}) \subset L^2$ and implies for every $z \in \mathbb{D}$ that

$$\begin{aligned} |f_+(z)| &= \lim_{N\to\infty} |(\mathfrak{A}_N f)(z)| \\ &\leq \limsup_{N\to\infty} \|\mathfrak{A}_N\|_{\mathcal{C}(\mathbb{T})\to A(\mathbb{D})} \|f\|_\infty \leq C_1 \|f\|_\infty \end{aligned} \tag{7.20}$$

using that the norms $\|\mathfrak{A}_N\|_{\mathcal{C}(\mathbb{T})\to A(\mathbb{D})}$ are upper bounded by C_1, because of assumption (7.18). Inequality (7.20) shows that the causal part $f_+(z)$ of every $f \in \mathcal{C}(\mathbb{T})$ is upper bounded by $C_1 \|f\|_\infty$ inside the unit disk. However Theorem 6.14 implies that there exist functions $f \in \mathcal{C}(\mathbb{T})$ with $\|f\|_\infty \leq 1$ such that the causal part $f_+ = \mathfrak{P}_+ f$ is unbounded, i.e. to every constant $C_2 > 0$ there exists a function $f \in \mathcal{C}(\mathbb{T})$ with $\|f\|_\infty \leq 1$ such that $\sup_{|z|<1} |f_+(z)| > C_2$. This is contradictory to (7.20), and it follows that the initial assumption (7.18) was wrong. □

Taking into account Proposition 7.6, the previous theorem implies that the norms $\|\mathfrak{A}_N\|_{\mathcal{C}(\mathbb{T})\to A(\mathbb{D})}$ are unbounded as $N \to \infty$ for *every* orthogonal basis Φ in $A(\mathbb{D})$. As an immediate consequence, we have

Corollary 7.8. *Let $\Phi = \{\varphi_k\}_{k=1}^\infty$ be an arbitrary orthogonal basis in $A(\mathbb{D})$. Then there are continuous functions $g \in \mathcal{C}(\mathbb{T})$ with $\|g\|_\infty \leq 1$ such that*

$$\lim_{N\to\infty} \|\mathfrak{A}_N g\|_\infty = \infty \ .$$

This result is a consequence of the theorem of Banach Steinhaus, which implies also that the set of continuous functions for which this corollary holds is of second category (a "nonmeager set", see e.g. [72]). In turn, this corollary implies that there is no orthogonal basis in $A(\mathbb{D})$ for which the approximation operator (7.9) is robust against disturbances of continuous functions.

Corollary 7.9. *Let $\Phi = \{\varphi_k\}_{k=1}^\infty$ be an arbitrary orthogonal basis in $A(\mathbb{D})$. To every $f \in A(\mathbb{D})$ there are disturbances $g \in \mathcal{C}(\mathbb{T})$ with $\|g\|_\infty \leq 1$ such that for the approximation error of $\widetilde{f} = f + \delta \cdot g$ holds*

$$\lim_{N\to\infty} \|\mathfrak{A}_N \widetilde{f} - f\|_\infty = \infty$$

for every number $\delta > 0$.

Of course, this corollary follows immediately from inequality (7.12) and Corollary 7.8. It shows that it is not possible to control the approximation error by means of the degree N of the approximation operator $\mathfrak{A}_N$, because even though an increase in N reduces the second term of the approximation error in (7.12), there always exist disturbances $g \in \mathcal{C}(\mathbb{T})$ such that an increasing of N increases the whole approximation error $\|\mathfrak{A}_N \widetilde{f} - f\|_\infty$.

7.3 Bases in Spaces of Smooth Functions

We saw in the previous section that there exists no basis in $A(\mathbb{D})$ such that the approximation (7.9) is robust against disturbances of continuous functions. The question arises for which kind of disturbances it will be possible to control the approximation error by means of the approximation degree N? Of course, since $\Phi = \{\varphi_k\}_{k=1}^\infty$ is assumed to be a basis in $A(\mathbb{D})$, it would be sufficient that the disturbance g is an element of $A(\mathbb{D})$ in order that the term $\|\mathfrak{A}_N g\|_\infty$ in (7.12) remains always bounded. However, this would imply that the disturbance g has to be an analytic function in $\mathbb{D}$.

In this section we characterize subspaces of $\mathcal{C}(\mathbb{T})$ such that for all disturbances from these subspaces, the approximation error can be controlled by the approximation degree N. This characterization is done in terms of the smoothness of the functions. It will turn out that it is sufficient that the disturbance f belongs to $\mathcal{C}_\omega(\mathbb{T})$ with a regular majorant ω (cf. Section 1.3).

Bases in $\mathcal{C}_\omega(\mathbb{T})$

Throughout this section we consider a special set of orthonormal functions on $\mathbb{T}$, the so-called *Franklin system*. The functions of this system are always denoted as ψ_k with $k = 0, 1, 2, \ldots$ and the construction of ψ_k is given in the following example.

Example 7.10 (Franklin system). For an integer $k = 2^n + l$, with $n = 0, 1, 2, \ldots$ and $0 \le l < 2^n$ define $\tau_k = 2\pi \frac{2l+1}{2^{n+1}}$ and put $\tau_0 = 0$. Then, the *Franklin system* on $\mathbb{T}$ is the orthonormal set of real valued, continuous, and piecewise linear functions $\{\psi_k\}_{k=0}^\infty$ such that ψ_k has nodes at $\{e^{i\tau_l}\}_{l \le k}$.

At the beginning, it will be shown that this Franklin system is a basis for the space $\mathcal{C}_\omega(\mathbb{T})$ afterward it will be shown that there exists a basis in $A_\omega(\mathbb{D})$ which can easily be obtained from the Franklin system. We consider functions f given on the unit circle $\mathbb{T}$. The generalized Fourier coefficients with respect to the Franklin system $\{\psi_n\}_{n=0}^\infty$ are always denoted by $c_k(f) = \langle f, \psi_k \rangle$. The first theorem gives necessary and sufficient conditions on the coefficients $c_k(f)$ such that f belongs to the space $\mathcal{C}_\omega(\mathbb{T})$ with a regular majorant ω.

Theorem 7.11. *Let ω be a regular majorant. A function f given on $\mathbb{T}$ belongs to $\mathcal{C}_\omega(\mathbb{T})$ if and only if there exists a constant C such that*

$$|c_k(f)| \le \tfrac{C}{\sqrt{k+1}}\, \omega(\tfrac{1}{k+1}) \ . \tag{7.21}$$

for all $k = 0, 1, 2, \ldots$. *A function* f *belongs to* $\mathcal{C}_{\omega,0}(\mathbb{T})$ *if and only if*

$$\lim_{k\to\infty} |c_k(f)|\, \frac{\sqrt{k+1}}{\omega(\frac{1}{k+1})} = 0 \ . \tag{7.22}$$

This theorem gives a characterization of all functions $f \in \mathcal{C}_\omega(\mathbb{T})$ with a regular majorant ω. If the coefficients $c_k(f) = \langle f, \psi_k\rangle$ of a function f with respect to the Franklin system are known then it can be ascertained by the above theorem whether f belongs to $\mathcal{C}_\omega(\mathbb{T})$ or $\mathcal{C}_{\omega,0}(\mathbb{T})$. Note that such a simple characterization of f is in general not possible if f belongs only to $\mathcal{C}(\mathbb{T})$.

The following proof of Theorem 7.11 follows closely the proof of Theorem 27 in [90, III.D]. There spaces of Lipschitz continuous functions were considered whereas Theorem 7.11 makes a statement on general smoothness classes $\mathcal{C}_\omega(\mathbb{T})$. However, we will need the following auxiliary lemma.

Lemma 7.12. *Let* ω *be a regular majorant. Then there exist two constants* C_1 *and* C_2 *such that*

$$\frac{1}{2^n}\sum_{k=0}^{n} \frac{\omega(1/2^k)}{1/2^k} \le C_1\, \omega(\tfrac{1}{2^n}) \quad \textit{and} \quad \sum_{k=n}^{\infty} \omega(\tfrac{1}{2^{k+1}}) \le C_2\, \omega(\tfrac{1}{2^n}) \ . \tag{7.23}$$

Proof. We start with the inequality on the left hand side of (7.23). Writing the sum on the left hand side in a slightly different form, one obtains

$$\begin{aligned} \frac{1}{2^n}\sum_{k=0}^{n} \frac{\omega(1/2^k)}{1/2^k} &= \frac{\omega(1)}{2^n} + 2\frac{1}{2^n}\sum_{k=1}^{n}\frac{\omega(1/2^k)}{1/2^k}\,\frac{1}{2} \\ &\le \frac{\omega(1)}{2^n} + \frac{1}{2^{n-1}}\sum_{k=1}^{n}\frac{\omega(1/2^k)}{1/2^k}\int_{1/2^{k+1}}^{1/2^k}\frac{\mathrm{d}\tau}{\tau} \ . \end{aligned} \tag{7.24}$$

The second line follows from the fact that the last integral is equal to $\log 2$ which is larger than $1/2$. Recall that the function $\omega(\tau)/\tau$ is non-increasing. Therefore it follows

$$\frac{\omega(b)}{b}\int_a^b \frac{\mathrm{d}\tau}{\tau} \le \int_a^b \frac{\omega(\tau)}{\tau^2}\,\mathrm{d}\tau \qquad \text{for all} \qquad 0 \le a \le b \le 1 \ .$$

Using this inequality in (7.24), one gets

$$\begin{aligned} \frac{1}{2^n}\sum_{k=0}^{n} \frac{\omega(1/2^k)}{1/2^k} &\le \frac{\omega(1)}{2^n} + \frac{1}{2^{n-1}}\sum_{k=1}^{n}\int_{1/2^{k+1}}^{1/2^k}\frac{\omega(\tau)}{\tau^2}\,\mathrm{d}\tau \\ &= \frac{\omega(1)}{2^n} + 4\,\frac{1}{2^{n+1}}\int_{1/2^{n+1}}^{1/2}\frac{\omega(\tau)}{\tau^2}\,\mathrm{d}\tau \le \frac{\omega(1)}{2^n} + 6\,C\,\omega(1/2^n) \ . \end{aligned}$$

For the last inequality it was used that ω is a regular majorant and satisfies (1.13) with a certain constant C and that $\omega(\tau/2) \le (3/2)\,\omega(\tau)$ (cf. Sec. 1.3).

Next the right inequality in (7.23) is proved. One easily verifies that

$$\sum_{k=n}^{\infty} \omega(1/2^{k+1}) = \sum_{k=n}^{\infty} \frac{\omega(1/2^{k+1})}{1/2^{k+1}} \left(\frac{1}{2^k} - \frac{1}{2^{k+1}} \right) . \tag{7.25}$$

Since $\omega(\tau)/\tau$ is a non-increasing function, it follows that

$$\int_a^b \frac{\omega(\tau)}{\tau}\, d\tau \ge \frac{\omega(a)}{a} \int_a^b d\tau = \frac{\omega(a)}{a}\,(b-a) \qquad \text{for all } 0 \le a \le b \le 1 .$$

Set $a = 1/2^{k+1}$ and $b = 1/2^k$ and using it in (7.25), one gets

$$\sum_{k=n}^{\infty} \omega(1/2^{k+1}) \le \sum_{k=n}^{\infty} \int_{1/2^{k+1}}^{1/2^k} \frac{\omega(\tau)}{\tau}\, d\tau = \int_0^{1/2^n} \frac{\omega(\tau)}{\tau}\, d\tau .$$

Since ω is a regular majorant and satisfies (1.12), one obtains finally that the last term is smaller or equal than $C_2\,\omega(1/2^n)$ with a certain constant C_2. □

Proof (Theorem 7.11). For the "only if" part, it can be shown as in [90, III.D] that

$$\left| \frac{1}{2\pi} \int_{-\pi}^{\pi} f(e^{i\tau})\, \overline{\psi_k(e^{i\tau})}\, d\tau \right| \le C_0\, \|f\|_{\mathcal{C}_\omega(\mathbb{T})}\, \frac{1}{\sqrt{k+1}}\, \omega(\tfrac{1}{k+1})$$

with a certain constant C_0. But this is equivalent to (7.21). The second statement (7.22) follows from (7.21) by standard arguments using that the polynomials are dense $\mathcal{C}_{\omega,0}(\mathbb{T})$.

It remains to show that for every sequence of coefficients $\{c_k(f)\}_{k=0}^{\infty}$ with property (7.21) the corresponding function f belongs to $\mathcal{C}_\omega(\mathbb{T})$. To this end we write f as

$$f(e^{it}) = c_0\, \psi_0(e^{it}) + \sum_{k=0}^{\infty} F_k(e^{it}) \tag{7.26}$$

with the functions $F_k(e^{it}) = \sum_{n=2^k}^{2^{k+1}-1} c_n(f)\, \psi_n(e^{it})$. Moreover, as in [90], there exists a constant C such that

$$\left|F_k(e^{it})\right| \le \sum_{n=2^k}^{2^{k+1}-1} |c_n(f)|\, \left|\psi_n(e^{it})\right| \le C\, \omega(\tfrac{1}{2^k}) \tag{7.27}$$

and such that

$$\left|F_k(e^{it_1}) - F_k(e^{it_2})\right| \le C\, \omega(\tfrac{1}{2^k})\, 2^{k+1} |e^{it_1} - e^{it_2}| . \tag{7.28}$$

Let $t_1, t_2 \in [-\pi, \pi)$ and choose N such that $1/2^{N+1} < |e^{it_1} - e^{it_2}| \le 1/2^N$. Now one gets from (7.26)

$$\big|f(\mathrm{e}^{\mathrm{i}t_1}) - f(\mathrm{e}^{\mathrm{i}t_2})\big| \leq |c_0|\,|\mathrm{e}^{\mathrm{i}t_1} - \mathrm{e}^{\mathrm{i}t_2}| + \sum_{k=0}^{N} \big|F_k(\mathrm{e}^{\mathrm{i}t_1}) - F_k(\mathrm{e}^{\mathrm{i}t_2})\big| + 2 \sum_{k=N+1}^{\infty} \|F_k\|_\infty$$

where the first term on the right hand side follows from the special properties of the functions ψ_k of the Franklin system. Next the relations (7.27) and (7.28) are used in the above inequality. This gives

$$\begin{aligned}\big|f(\mathrm{e}^{\mathrm{i}t_1}) - f(\mathrm{e}^{\mathrm{i}t_2})\big| &\leq |c_0|\,|\mathrm{e}^{\mathrm{i}t_1} - \mathrm{e}^{\mathrm{i}t_2}| \\ &\quad + C\sum_{k=0}^{N} \omega\left(\tfrac{1}{2^k}\right) 2^{k+1} |\mathrm{e}^{\mathrm{i}t_1} - \mathrm{e}^{\mathrm{i}t_2}| + 2\,C \sum_{k=N+1}^{\infty} \omega\left(\tfrac{1}{2^k}\right)\end{aligned}$$

To get an upper bound for the second and third term, Lemma 7.12 is used. This finally gives

$$\big|f(\mathrm{e}^{\mathrm{i}t_1}) - f(\mathrm{e}^{\mathrm{i}t_2})\big| \leq |c_0|\,|\mathrm{e}^{\mathrm{i}t_1} - \mathrm{e}^{\mathrm{i}t_2}| + (C_1 + C_2)\,\omega(1/2^N)$$

which shows that there exists a constant C_3 such that $\big|f(\mathrm{e}^{\mathrm{i}t_1}) - f(\mathrm{e}^{\mathrm{i}t_2})\big| \leq C_3\,\omega(|\mathrm{e}^{\mathrm{i}t_1} - \mathrm{e}^{\mathrm{i}t_2}|)$. This proves that $f \in \mathcal{C}_\omega(\mathbb{T})$. □

Based on these preparations, we will show in the next theorem that for every regular majorant ω all function $f \in \mathcal{C}_{\omega,0}(\mathbb{T})$ can be approximated uniformly in the Franklin system $\{\psi_k\}_{k=0}^\infty$. It follows in particular that the Franklin system is a basis for $\mathcal{C}_{\omega,0}(\mathbb{T})$.

Theorem 7.13. *Let ω be a regular majorant. Then there exists a constant C such that*

$$\Big\|\textstyle\sum_{k=0}^N c_k(f)\,\psi_k\Big\|_{\mathcal{C}_\omega(\mathbb{T})} \leq C(\omega)\ \|f\|_{\mathcal{C}_\omega(\mathbb{T})}$$

for all $N \in \mathbb{N}$. For all $f \in \mathcal{C}_{\omega,0}(\mathbb{T})$ holds

$$\lim_{N\to\infty} \Big\|f - \textstyle\sum_{k=0}^N c_k(f)\,\psi_k\Big\|_{\mathcal{C}_\omega(\mathbb{T})} = 0\ .$$

Corollary 7.14. *Let ω be a regular majorant, then the Franklin system $\{\psi_k\}_{k=0}^\infty$ is a basis in $\mathcal{C}_{\omega,0}(\mathbb{T})$.*

Proof (Theorem 7.13). From the proof of Theorem 7.11 we already know that

$$|c_k(f)| \leq C_0\ \|f\|_{\mathcal{C}_\omega(\mathbb{T})}\ \tfrac{1}{\sqrt{k+1}}\,\omega(\tfrac{1}{k+1})\ .$$

Therewith (7.27) and (7.28) show that

$$\begin{aligned}|F_k(t)| &\leq C_1\ \|f\|_{\mathcal{C}_\omega(\mathbb{T})}\ \omega(\tfrac{1}{2^k}) \\ |F_k(t_1) - F_k(t_2)| &\leq C_2\ \|f\|_{\mathcal{C}_\omega(\mathbb{T})}\ \omega(\tfrac{1}{2^k})\,2^{k+1}|\mathrm{e}^{\mathrm{i}t_1} - \mathrm{e}^{\mathrm{i}t_2}|\ .\end{aligned}$$

Therewith, one gets similar to the proof of [90, III.D.27] that

$$\left\| \sum_{l=0}^{2^k} c_l(f)\, \psi_l \right\|_{\mathcal{C}_\omega(\mathbb{T})} \leq C_3 \, \|f\|_{\mathcal{C}_\omega(\mathbb{T})} \, .$$

This is the statement of the theorem only for a subsequence of the generalized Fourier sequence. For the general case $n = 2^k + m$ a longer extension is needed, which is left as an exercise. □

Bases in $A_\omega(\mathbb{D})$

Next we investigate the existence of a basis in the spaces $A_\omega(\mathbb{D}) = [\mathcal{C}_\omega(\mathbb{T})]_+$ and $A_{\omega,0}(\mathbb{D}) = [\mathcal{C}_{\omega,0}(\mathbb{T})]_+$ of all smooth and analytic functions inside the unit disk $\mathbb{D}$. It turns out that a basis in $A_{\omega,0}(\mathbb{D})$ can be obtained by an analytic extension of the Franklin system $\{\psi_k\}_{k=0}^\infty$ inside the unit disk by means of the Heglotz-Riesz transform, $\mathfrak{R}$, defined in (5.1). To this end we set for $k = 0, 1, 2 \ldots$

$$\varphi_k(z) := (\mathfrak{R}\, \psi_k)(z) \, , \qquad z \in \mathbb{D} \, .$$

Then it can be shown that $\Phi = \{\varphi_k\}_{k=0}^\infty$ is a basis in $A_{\omega,0}(\mathbb{D})$. This follows easily from the result of the previous paragraph and from the fact that the Riesz projection, and consequently the Herglotz-Riesz transform, is a continuous operator on $\mathcal{C}_\omega(\mathbb{T}) \to A_\omega(\mathbb{D})$ provided that ω is a regular majorant (cf. Theorem 6.16). For this reason, we left the technical details as an exercise and state only the results.

Similar to the previous paragraph, we give first a characterization of the functions in $A_\omega(\mathbb{D})$ in terms of the generalized Fourier coefficients $c_k(f) = \langle f, \varphi_k \rangle$ with respect to the orthonormal system $\Phi = \{\varphi_k\}_{k=0}^\infty$.

Theorem 7.15. *Let ω be a regular majorant. The sequence $\{c_k\}_{k=0}^\infty$ of complex numbers is a series of Fourier coefficients of a function $f \in A_\omega(\mathbb{D})$ with respect to the orthonormal system $\{\varphi_k\}_{k=0}^\infty$ if and only if there exists a constant C such that*

$$|c_k| \leq \frac{C}{\sqrt{k+1}} \, \omega\left(\frac{1}{k+1}\right) \, .$$

The sequence $\{c_k\}_{k=0}^\infty$ represents the Fourier coefficients of a function $f \in A_{\omega,0}(\mathbb{D})$ with respect to the orthonormal system $\{\varphi_k\}_{k=0}^\infty$ if and only if

$$\lim_{k\to\infty} |c_k| \, \frac{\sqrt{k+1}}{\omega(\frac{1}{k+1})} = 0 \, .$$

This theorem is equivalent to Theorem 7.11 for $\mathcal{C}_\omega(\mathbb{T})$. Based on the properties of the coefficients $c_k(f) = \langle f, \varphi_k \rangle$, the above theorem allows us to decide whether f belongs to $A_\omega(\mathbb{D})$ or $A_{\omega,0}(\mathbb{D})$. Note that without the requirement that ω is regular, this will not be possible. Based on this characterization of the functions in $A_{\omega,0}(\mathbb{D})$, it can be shown that $\{\varphi_k\}_{k=0}^\infty$ is a basis in $A_{\omega,0}(\mathbb{D})$.

Theorem 7.16. *Let ω be a regular majorant, then the sequence $\Phi = \{\varphi_k\}_{k=0}^{\infty}$ is a (Schauder) basis for the space $A_{\omega,0}(\mathbb{D})$, i.e. for all $f \in A_{\omega,0}(\mathbb{D})$ holds*

$$\lim_{N\to\infty} \left\| f - \sum_{k=0}^{N} c_k(f)\, \varphi_k \right\|_{A_\omega(\mathbb{D})} = 0$$

with $c_k(f) = \langle f, \varphi_k \rangle$, and the representation

$$f(z) = \sum_{k=0}^{\infty} c_k(f)\, \varphi_k(z)\ , \qquad z \in \mathbb{D}$$

of f in the basis Φ is unique.

Remark 7.17. By a simple normalization $\widetilde{\varphi}_k := \sqrt{k+1}\,\varphi_k$, it can be achieved that all $\widetilde{\varphi}_k$ are uniformly bounded by a common constant $\|\widetilde{\varphi}_k\|_\infty \leq C$ for all k. Then $\{\widetilde{\varphi}_k\}_{k=0}^{\infty}$ is said to be a uniformly stable basis in $A_{\omega,0}(\mathbb{D})$. Such bases are considered in more detail in the next chapter.

7.4 Uniformly Stable Basis Representations

We saw in Section 7.1 that there exists bases $\Phi = \{\varphi_k\}_{k=1}^{\infty}$ in $A(\mathbb{D})$ such that the partial sums $\sum_{k=1}^{N} \langle f, \varphi_k \rangle\, \varphi_k$ converges uniformly to f for all $f \in A(\mathbb{D})$. Since Φ is a basis in $A(\mathbb{D})$, it is clear that $\varphi_k \in A(\mathbb{D})$ for all k and consequently that $\|\varphi_k\|_\infty < \infty$ for all basis functions φ_k. Thus, the stability norm of all individual filters φ_k is bounded such that every φ_k represents the transfer function of a causal and stable filter by itself. Next, we ask whether or not it is possible to upper bound the stability norm of the individual filters φ_k by a universal constant C_0, i.e. does there exists a constant $C_0 < \infty$ such that $\|\varphi_k\|_\infty \leq C_0$ for all $k \in \mathbb{N}$? If such a constant C_0 exists, we will speak of a *uniformly stable basis* Φ. However, the next theorem shows that there exists no uniformly stable basis in $A(\mathbb{D})$.

Theorem 7.18. *Let $\Phi = \{\varphi_k\}_{k\in\mathbb{N}}$ be a system of orthogonal functions in $A(\mathbb{D})$ and assume that there exists a constant $C_0 < \infty$ such that $\|\varphi_k\|_\infty \leq C_0$ for all $k \in \mathbb{N}$. Then there always exists a function $f \in A(\mathbb{D})$ such that*

$$\limsup_{N\to\infty} \left\| f - \sum_{k=0}^{N} \langle f, \varphi_k \rangle\, \varphi_k \right\|_\infty = \infty\ . \tag{7.29}$$

Thus, if the elements of the orthonormal system Φ are uniformly bounded then Φ can not be a basis in $A(\mathbb{D})$ since the above theorem shows that there always exists an $f \in A(\mathbb{D})$ such that its generalized Fourier series (with respect to Φ) does not converge to f in $A(\mathbb{D})$.

Proof. Let $\Phi = \{\varphi_k\}_{k=1}^{\infty}$ be a fixed orthogonal system in $A(\mathbb{D})$ and consider the partial sum (7.9) of an arbitrary function $f \in A(\mathbb{D})$ in its integral representation (7.14) with the kernel (7.15). For a fixed $z \in \mathbb{D}$, the approximation operator $\mathfrak{A}_N$ defines a linear functional on $A(\mathbb{D})$ with norm given by

$$\sup_{\|f\|_\infty \le 1} |(\mathfrak{A}_N f)(z)| = \|K_N(\cdot, z)\|_{A(\mathbb{D})^*} \tag{7.30}$$

and wherein $A(\mathbb{D})^*$ is the dual space of the disk algebra $A(\mathbb{D})$.

Now a result of Bourgain [24, Theorem in Section 0] is used. It states that for a fixed $N \in \mathbb{N}$ there exists a $\widehat{z} \in \overline{\mathbb{D}}$ such that

$$\frac{1}{N} \sum_{n=1}^{N} \|K_n(\cdot, \widehat{z})\|_{A(\mathbb{D})^*} > c(C_0) \log N .$$

with a constant $c(C_0)$ which depends only on the upper bound C_0 of the stability norms of the basis functions φ_k. It follows that for every $N \in \mathbb{N}$ there exists an n such that

$$\|K_n(\cdot, \widehat{z})\|_{A(\mathbb{D})^*} > c(C_0) \log N .$$

Using (7.30), it follows that there exists an $f_* \in A(\mathbb{D})$ with $\|f_*\|_\infty \le 1$ such that

$$\max_{1 \le n \le N} \|\mathfrak{A}_n f_*\|_\infty > c(C_0) \log N . \tag{7.31}$$

Let $\{\epsilon_n\}_{n \in \mathbb{N}}$ be a monotone decreasing sequence with $\lim_{n\to\infty} \epsilon_n = 0$ and $\lim_{n\to\infty} \epsilon_n \log n = \infty$. Then (7.31) shows that

$$\limsup_{N\to\infty} \frac{1}{\epsilon_N \log N} \max_{1 \le n \le N} \|\mathfrak{A}_n f_*\|_\infty = \infty \tag{7.32}$$

and in particular that $\|\mathfrak{A}_N f_*\|_\infty \to \infty$ as $N \to \infty$.
Since $\|f_* - \mathfrak{A}_N f_*\|_\infty \ge |\, \|\mathfrak{A}_N f_*\|_\infty - \|f_*\|_\infty |$ and because $\|f_*\|_\infty \le 1$, it follows from (7.32) that f_* satiesfies (7.29). □

The previous theorem proved that there exists no uniformly stable basis in $A(\mathbb{D})$. However, if subsets of $A(\mathbb{D})$ are considered, a uniformly stable basis may exist. As an example, we consider again bases in the space $A_\omega(\mathbb{D})$ of analytic functions in $\mathbb{D}$ of which the modulus of continuity is upper bounded by a regular majorant ω. For this subset the following result is obtained.

Theorem 7.19. *Consider the orthonormal system $\Phi = \{\varphi_k(z) = z^k\}_{k=0}^{\infty}$ in $A(\mathbb{D})$ and let ω be a regular majorant. Then*

$$\lim_{N\to\infty} \left\| f - \sum_{k=0}^{N} \langle f, \varphi_k \rangle \varphi_k \right\|_\infty = 0 \qquad \textit{for all } f \in A_{\omega,0}(\mathbb{D}) .$$

It is clear that the basis functions $\varphi_k(z) = z^k$ in this theorem are uniformly bounded by $\|\varphi_k\|_\infty \le 1$ for all k and the above theorem shows that every sufficiently smooth function in $A(\mathbb{D})$ can be approximated arbitrary well (with respect to the stability norm $\|\cdot\|_\infty$) in this system Φ of orthogonal functions. However, Φ is not a basis of $A(\mathbb{D})$.

A proof of this result for the special case of Hölder continuous functions, i.e. for majorants of the form $w(\tau) = \tau^\alpha$ with $0 < \alpha < 1$ can be found in [92, Chapter II, Sect. 10]. The extension to regular majorants is merely a technical exercise. Therefore the proof of Theorem 7.19 is left out.

Notes

Textbooks related to this section include [5, 90, 91, 92]. The proofs in Section 7.3 and 7.4 use mainly ideas from Wojtaszczyk's book [90]. The existence of a basis in the disk algebra was first shown by Bockarev [23]. The classical references to the Faber-Schauder and Franklin system, used extensive in this chapter, are [36, 79] and [39], respectively. The interest of the authors [9, 18] for this topic was inspired by some works of Akcay e.g. [2, 3].

8

Causal Approximations

The previous chapter considered the approximation of causal transfer functions in bases of the disk algebra in the form (7.1). There, it was important that the coefficients $\eta_k = c_k(f)$ depend only on the given function f which should be approximated, but not on the approximation degree N.

The present chapter considers the *causal approximations* of causal transfer functions by means of a sequence $\{\varphi_k\}_{k=1}^{\infty}$ of the form

$$(\mathfrak{A}_N f)(z) = \sum_{n=1}^{N} c_{n,N}(f)\, \varphi_n(z)\,, \qquad z \in \mathbb{D} \tag{8.1}$$

with certain numbers $c_{n,N}(f)$ which are uniquely determined by the given function f. In contrast to a basis expansion, the coefficients $c_{n,N}$ depend now, in general, on the approximation degree N. Thus, the approximation method (8.1) is more general than the basis expansion. However, we still will require that the coefficient functionals depend *linearly* on the given function f.

In Section 7.1, it was shown that there exist bases in the disk algebra $A(\mathbb{D})$, which implies of course, that there exist sequences $\Phi = \{\varphi_n\}_{n=1}^{\infty}$ in $A(\mathbb{D})$ such that the approximation (8.1) converges uniformly to f for all $f \in A(\mathbb{D})$. However, it was shown in Section 7.2 that there exists no disk algebra basis which is robust against errors in the given data f. In this section, we want to investigate whether the more general approximation method (8.1) allows a *robust* approximation of functions in $A(\mathbb{D})$. We require that the approximation method (8.1) has the following three natural properties.

(A) *Robustness*: The approximation error $\|g - \mathfrak{A}_N g\|_\infty$ should decrease as the approximation degree N increases, and $\sup_{N\in\mathbb{N}} \|\mathfrak{A}_N g\|_\infty$ ought to be bounded for all $g \in \mathcal{C}(\mathbb{T})$.
(B) *Causality*: The approximation $\mathfrak{A}_N f$ should represent a causal transfer function. This is certainly achieved if all individual transfer functions φ_n are causal.
(C) *Linearity*: The calculation of the coefficients $c_n(f, N)$ should be sufficiently simple. Therefore, we require that the coefficients depend linearly on f.

V. Pohl, H. Boche, *Advanced Topics in System and Signal Theory*, Foundations in Signal Processing, Communications and Networking 4,
DOI 10.1007/978-3-642-03639-2_8, © Springer-Verlag Berlin Heidelberg 2009

Since the functions f are assumed to be continuous, the *Riesz representation theorem* (see e.g. [70]) implies that the coefficient functionals c_n have the following general form

$$c_n(f,N) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\theta})\,\mathrm{d}\mu_{n,N}(\mathrm{e}^{\mathrm{i}\theta}) \tag{8.2}$$

with unique Borel measures $\mu_{n,N}$ which depend on the approximation degree N, in general. We will distinguish between two cases. In the first place, it is assumed that all measures $\mu_{n,N}$ are *absolute continuous* (with respect to the Lebesgue measure on $\mathbb{T}$). In this case, there exist functions $\gamma_{n,N} \in L^1$ such that the coefficient functionals (8.2) can be written as $c_n(f,N) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\theta})\,\gamma_{n,N}(\mathrm{e}^{\mathrm{i}\theta})\,\mathrm{d}\theta$. This case is of interest for filterbank and wavelet applications. There, the two systems of functions $\{\varphi\}_{n=1}^{\infty}$ and $\{\gamma\}_{n=1}^{\infty}$ are biorthonormal to each other and they are usually called synthesis and analysis filterbank, respectively.
Secondly, we consider the case where the measures $\mu_{n,N}$ are concentrated on finite sets $\mathcal{T}_{n,N} = \{\mathrm{e}^{\mathrm{i}\theta_1(n,N)}, \mathrm{e}^{\mathrm{i}\theta_2(n,N)}, \cdots, \mathrm{e}^{\mathrm{i}\theta_M(n,N)}\}$ of discrete points on the unit circle. This is certainly the practically relevant case since in applications the function f is only known at discrete points (samples) and the calculation of the coefficients $c_n(f,N)$ has to be based on these samples only.

Of course, for any causal approximation $\mathfrak{A}_N$ in $A(\mathbb{D})$, one wants to have that

$$\lim_{N\to\infty} \|f - \mathfrak{A}_N f\|_\infty = 0 \qquad \text{for all } f \in A(\mathbb{D})\ . \tag{8.3}$$

Thus, it should be possible to approximate every causal function f arbitrarily close by $\mathfrak{A}_N f$. A necessary condition for (8.3) to hold is the following property of the approximation method.

Definition 8.1 (M–property). *We say that an approximation method $\mathfrak{A}_N$ possesses the M–*property *if for all polynomials of the form $p_m(z) = z^m$ holds*

$$\lim_{N\to\infty} \|p_m - \mathfrak{A}_N\, p_m\|_\infty = 0 \qquad \textit{for all}\ \ m = 0,1,2,\dots\ .$$

In the following, we only require that the approximation methods possess this M–property but we do not require explicitly (8.3). Thus the class of approximation methods under consideration is somewhat larger than necessary. Nevertheless, we are going to show that even in this larger class there exists no approximation operator on $A(\mathbb{D})$ which possesses all three properties (A), (B), and (C). Consequently, there is also no robust, causal, and linear approximation method of the form (8.1) which satisfies (8.3). Before we are going to prove this, we show that there exists approximation methods with only two out of the three properties (A), (B), (C).

8.1 Non-linear and Causal Approximations

For $f \in \mathcal{C}(\mathbb{T})$, we look for a causal and stable transfer function $g \in H^\infty$ which approximates f as closely as possible. Moreover, if we require that the approximative transfer function g is also continuous, we have to look for an optimal g in the disk algebra $A(\mathbb{D})$. It is known [41] that

$$E[f, A(\mathbb{D})] = \inf_{g \in A(\mathbb{D})} \|f - g\|_\infty = \inf_{g \in H^\infty} \|f - g\|_\infty = E[f, H^\infty]$$

in which $E[f, A(\mathbb{D})]$ and $E[f, H^\infty]$ are called the *best approximation* of f in $A(\mathbb{D})$ and H^∞, respectively [92]. Thus, the best approximation is equal in $A(\mathbb{D})$ and H^∞, which means that the residual approximation error $\|f - g\|_\infty$ coincides in both spaces. However, it was shown in Chapter 6 that $A(\mathbb{D})$ is not complemented in $\mathcal{C}(\mathbb{T})$ (cf. Theorem 6.14). This implies that the optimal g, for which the best approximation is attained, belongs only to H^∞ but not to $A(\mathbb{D})$, in general. Thus, to every $f \in \mathcal{C}(\mathbb{T})$ there exists a unique function $g^{opt} \in H^\infty$ such that

$$E[f, A(\mathbb{D})] = E[f, H^\infty] = \|f - g^{opt}\|_\infty \ .$$

But this g^{opt} is not a continuous function on $\mathbb{T}$, in general. To ascertain that the best causal approximation g^{opt} is again continuous on $\mathbb{T}$, one has to consider the approximation problem on subspaces of $\mathcal{C}(\mathbb{T})$. For example, it follows from Theorem 6.16 that $A(\mathbb{D})$ is complemented in the space of smooth functions $\mathcal{C}_\omega(\mathbb{T})$ if ω is a regular majorant. Consequently, if f is assumed to belong to $\mathcal{C}_\omega(\mathbb{T})$, the best causal approximation g^{opt} will belong to $A(\mathbb{D})$ (actually, it even belongs to $A_\omega(\mathbb{D})$).

In the following, we proceed from the assumption that the given function f belongs to such a subset of $\mathcal{C}_\omega(\mathbb{T}) \subset \mathcal{C}(\mathbb{T})$ for which the optimal approximation g^{opt} belongs to $A(\mathbb{D})$. How does g^{opt} depend on the given f and does there exist a linear mapping $\mathcal{M}_+^{opt} : f \mapsto g^{opt}$ which gives to every $f \in \mathcal{C}_\omega(\mathbb{T})$ the optimal approximation $g^{opt} \in A(\mathbb{D})$? The general relation between f and g^{opt} is quite complicated. However, it is known [41, Section IV] that to every $f \in \mathcal{C}(\mathbb{T}) \subset L^\infty$ there exists an $F \in H_0^1$ such that $f - g^{opt} = E[f, A(\mathbb{D})] \frac{\overline{F}}{|F|}$. Consequently, the causal and stable transfer function, which approximates the given f best, can be written as

$$g^{opt} = \mathcal{M}_+^{opt} f = f - E[f, A(\mathbb{D}))] \, \mathrm{e}^{-\mathrm{i} \arg(F)}$$

with a certain function $F \in H_0^1$. This last relation shows that 1) $\mathcal{M}_+^{opt} f = f$ whenever $f \in A(\mathbb{D})$ and 2) that the mapping $\mathcal{M}_+^{opt}$ is *non-linear*, in general.

Similarly, one can consider the approximation of $f \in \mathcal{C}(\mathbb{T})$ by polynomials of a certain maximal degree N. Then the best approximation by a (causal) polynomial in $\mathcal{P}_N$ is given by

$$E_+[f, N] = \inf_{p \in \mathcal{P}_N} \|f - p\|_\infty = \|f - p_N^{opt}\|_\infty \tag{8.4}$$

where the best approximation $E_+[f,N]$ is always attained by a polynomial $p_N^{opt} \in \mathcal{P}_N$. The mapping $\mathcal{M}_{+,N}^{opt} : f \mapsto p_N^{opt}$ is unique but again $\mathcal{M}_{+,N}^{opt}$ is a non-linear operator, in general. We summarize this without a formal proof in the following lemma.

Lemma 8.2. *There exists an approximation method* $\mathcal{M}_{+,N}^{opt} : \mathcal{C}(\mathbb{T}) \to \mathcal{P}_N$ *which maps every* $f \in \mathcal{C}(\mathbb{T})$ *onto a unique causal polynomial* $p_N^{opt} \in \mathcal{P}_N$ *which satisfies* (8.4). *This method has the properties (A) and (B) but it is non-linear.*

Consequently, there exist non-linear approximation methods which have the desired properties (A) and (B). In particular it holds for the FIR approximation (8.4) that the residual approximation error decreases with increasing approximation degree N and that $\lim_{N\to\infty} E_+[f,N] = E[f,A(\mathbb{D})]$. However, we note again that the optimal causal polynomial approximation p_N^{opt} only converges to a function $g^{opt} \in H^\infty$ as $N \to \infty$, in general.

8.2 Non-causal, Linear Approximations

Next we consider linear approximations of the form

$$(\mathfrak{A}_N^{(w)} f)(z) = \sum_{n=-N}^{N} w(n/N)\, \hat{f}(n)\, z^n \ , \qquad z \in \overline{\mathbb{D}} \tag{8.5}$$

where $\hat{f}(n)$ are the usual Fourier coefficients (2.1) of f, and $w(x)$ is a window function defined for $-1 \le x \le 1$ and with $w(x) = 0$ for all $|x| > 1$. Since the above approximation $\mathfrak{A}_N^{(w)} f$ contains non-zero negative Fourier coefficients, it is clear that this approximation is *non-causal*. It is obvious that the coefficients $c_n(f,N) = w(n/N)\, \hat{f}(n)$ depend linearly on f. Therefore, the approximation (8.5) has property (C).

As discussed in Section 2.1, there exist several window functions w such that $\mathfrak{A}_N^{(w)} f$ converges uniformly to f as $N \to \infty$ for all $z \in \overline{\mathbb{D}}$, i.e. there exists windows such that $\lim_{N\to\infty} \|f - \mathfrak{A}_N^{(w)} f\|_\infty = 0$ for all $f \in \mathcal{C}(\mathbb{T})$. Examples are the Fejér or the de la Vallée-Poussin window. This shows that the approximation method (8.5) also possesses property (A). We summarize this in the following lemma.

Lemma 8.3. *There exist approximation methods* $\mathfrak{A}_N^{(w)} : \mathcal{C}(\mathbb{T}) \to A(\mathbb{D})$ *which are stable (A) and linear (C) but which are non-causal.*

Given an approximation method (8.5) with property (A) and (C), it seems to be reasonable to cut off the non-causal part of the series to obtain a method which also has property (B) (i.e. causality). However, if this is done, the approximation series will no longer be stable. This is shown in the next section.

8.3 Behavior of Causal Approximations

This section investigates approximation methods of the form (8.1) with properties (B) and (C). Thus, we assume that the functions $\{\varphi_n\}_{n=1}^{\infty}$ in (8.1) represent the transfer functions of causal and stable linear systems ($\varphi_n \in A(\mathbb{D})$ for all $n \in \mathbb{N}$) and that the coefficients $c_n(f,N)$ are of the form (8.2). Moreover, if not mentioned otherwise, it is always assume that the approximation methods (8.1) have the M–property defined Def. 8.1.

Absolute continuous measures

In the first part of this section, we investigate the case that all measures $\mu_{n,N}$ in the coefficient functionals (8.2) are *absolute continuous*. Then, by the *Radon-Nikodym theorem*, the coefficients $c_n(f,N)$ can always be written as

$$c_n(f,N) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\theta})\,\overline{\gamma_{n,N}(\mathrm{e}^{\mathrm{i}\theta})}\,\mathrm{d}\theta\ . \tag{8.6}$$

with functions $\gamma_{n,N} \in L^1$ which are uniquely defined by the measures $\mu_{n,N}$. If these coefficients are plugged into the definition (8.1) of the approximation operator, an integral representation of $\mathfrak{A}_N$ is obtained:

$$(\mathfrak{A}_N f)(z) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\theta})\,K_N(\mathrm{e}^{\mathrm{i}\theta},z)\,\mathrm{d}\theta\ , \qquad z \in \mathbb{D} \tag{8.7}$$

with the *reproducing kernel*

$$K_N(\mathrm{e}^{\mathrm{i}\theta},z) = \sum_{n=1}^{N} \varphi_n(z)\,\overline{\gamma_n(\mathrm{e}^{\mathrm{i}\theta})}\ . \tag{8.8}$$

Example 8.4. Approximation methods with coefficients of the form (8.6) and with property (B) and (C), can easily be derived from the non-causal linear approximation methods considered in Section 8.2, simply by truncating the anti-causal part in these series (8.5). As a concrete example, we consider the *causal Fejér means*: Let the basis function φ_k and the functions $\gamma_{n,N}$ in the coefficient functionals (8.6) be given by

$$\varphi_n(z) = z^n \qquad \text{and} \qquad \gamma_{n,N}(\mathrm{e}^{\mathrm{i}\theta}) = (1-\tfrac{n}{N})\,\mathrm{e}^{\mathrm{i}n\theta}$$

respectively, for $n = 0,1,2,\dots,N-1$. It is clear that all $\varphi_k \in A(\mathbb{D})$ and all $\gamma_{n,N} \in L^1$. Therewith, the integral representation (8.7) of the approximation operator has the kernel

$$K_N(\mathrm{e}^{\mathrm{i}\theta},z) = \textstyle\sum_{n=0}^{N-1}\left(1-\frac{n}{N}\right)\,\left(z\,\mathrm{e}^{-\mathrm{i}\theta}\right)^n\ .$$

Since $A(\mathbb{D}) \subset \mathcal{C}(\mathbb{T})$, it is clear from the behavior of the non-causal Fejér mean as discussed in Section 8.2 that $\lim_{N\to\infty}\|f-\mathfrak{A}_N f\|_\infty = 0$ for all $f \in A(\mathbb{D})$.

This shows in particular that this approximation method has property M. Thus, the causal Fejér mean is a perfect approximation method for all $f \in A(\mathbb{D})$. However, in order that $\mathfrak{A}_N$ also has property (A) (i.e. stability) this should hold for all $f \in \mathcal{C}(\mathbb{T})$.

By the above definition of the approximation method, it possesses the property (B) and (C) and it remains to investigate whether this method has also property (A), i.e. whether it is robust. To do this, we follow the approach of Section 7.2 and investigate the Lebesgue constant of the approximation method $\mathfrak{A}_N$. The question is whether $\sup_{N \in \mathbb{N}} \|\mathfrak{A}_N f\|_\infty < \infty$ for all $f \in \mathcal{C}(\mathbb{T})$. By the *uniform boundedness principle* (Theorem of Banach-Steinhaus) this will be satisfied if and only if the operators $\mathfrak{A}_N : \mathcal{C}(\mathbb{T}) \to A(\mathbb{D})$ are uniformly bounded (see e.g. [70]), i.e. if there exists a constant $C < \infty$ such that

$$\|\mathfrak{A}_N\|_{\mathcal{C}(\mathbb{T}) \to A(\mathbb{D})} = \sup_{\substack{f \in \mathcal{C}(\mathbb{T}) \\ \|f\|_\infty \leq 1}} \|\mathfrak{A}_N f\|_\infty \leq C \quad \text{for all } N \in \mathbb{N} \, .$$

As in the case of the basis expansion (cf. Proposition 7.6), the norm of the approximation operators $\mathfrak{A}_N : \mathcal{C}(\mathbb{T}) \to A(\mathbb{D})$ is equal to its Lebesgue constant:

$$\|\mathfrak{A}_N\|_{\mathcal{C}(\mathbb{T}) \to A(\mathbb{D})} = \sup_{|z|<1} \underbrace{\left(\frac{1}{2\pi} \int_{-\pi}^{\pi} |K_N(\mathrm{e}^{\mathrm{i}\theta}, z)| \, \mathrm{d}\theta \right)}_{=: L_N(z)} = \sup_{|z|<1} L_N(z) \, . \tag{8.9}$$

In order to decide whether the approximation method $\mathfrak{A}_N : \mathcal{C}(\mathbb{T}) \to A(\mathbb{D})$ is robust, we investigate the uniform boundedness of the Lebesgue constant. As a preparation, we give an auxiliary proposition on a relation between the Fourier coefficients $\hat{\varphi}_n(k)$ of the basis functions φ_n and the Fourier coefficients $\hat{\gamma}_{n,N}(k)$ of the functions $\gamma_{n,N}$ in the coefficient functionals (8.6). This relation is a direct consequence of the required property M of the approximation method, and will be needed for the proof of Theorem 8.7 below.

Proposition 8.5. *Let $\mathfrak{A}_N$ be an approximation method of the form* (8.1) *with $\varphi_n \in A(\mathbb{D})$ for all n and in which the coefficients $c_n(f, N)$ are given by the functionals* (8.6) *with $\gamma_{n,N} \in L^1$ for all $n = 1, 2, \ldots, N$. If $\mathfrak{A}_N$ has property M, then it holds that*

$$\lim_{N \to \infty} \sum_{n=1}^{N} \hat{\varphi}_n(l) \, \overline{\hat{\gamma}_{n,N}(m)} = \begin{cases} 1 & \text{for } l = m \\ 0 & \text{for } l \neq m \end{cases} \, . \tag{8.10}$$

Proof. For a fixed integer $m \geq 0$, we consider the approximation $(\mathfrak{A}_N p_m)(z) = \sum_{n=1}^{N} c_n(p_m) \, \varphi_n(z)$ of the polynomial $p_m(z) = z^m$. According to (8.6), the coefficients $c_n(p_m, N)$ are given by

$$c_n(p_m, N) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}m\theta} \, \overline{\gamma_{n,N}(\mathrm{e}^{\mathrm{i}\theta})} \, \mathrm{d}\theta = \overline{\hat{\gamma}_{n,N}(m)} \tag{8.11}$$

where the right hand side is the conjugate complex of the m-th Fourier coefficient of $\gamma_{n,N}$. Since all basis functions φ_n belong to $A(\mathbb{D})$, they can be written as $\varphi_n(z) = \sum_{l=0}^{\infty} \hat{\varphi}_n(l)\, z^l$. Using this representation in (8.1) and taking into account (8.11), one gets

$$(\mathfrak{A}_N p_m)(z) = \sum_{n=1}^{N} c_n(p_m, N)\, \varphi_n(z) = \sum_{l=0}^{\infty} \left(\sum_{n=1}^{N} \hat{\varphi}_n(l)\, \overline{\hat{\gamma}_{n,N}(m)} \right) z^l \ .$$

Since $\mathfrak{A}_N$ possesses the M–property, it holds that $\lim_{N\to\infty}(\mathfrak{A}_N p_m)(z) = z^m$, which gives (8.10). $\square$

Example 8.6. Consider the causal Fejér means given in Example 8.4. This approximation method possesses property M and consequently, it has to satisfy (8.10). Indeed, it is easy to verify that for this example holds

$$\lim_{N\to\infty} \sum_{n=1}^{N} \hat{\varphi}_n(l)\, \overline{\hat{\gamma}_{n,N}(m)} = \lim_{N\to\infty} \hat{\varphi}_l(l)\, \overline{\hat{\gamma}_{m,N}(m)} = \lim_{N\to\infty} (1 - \tfrac{n}{N}) = 1 \ .$$

After these preparations, we are able to show that every approximation method (8.1) with coefficient functionals $c_n(f, N)$ of the form (8.6) will be non-robust against continuous errors in the given data.

Theorem 8.7. *For every approximation method* $\mathfrak{A}_N : \mathcal{C}(\mathbb{T}) \to A(\mathbb{D})$ *of the form* (8.1) *with* $\varphi_n \in A(\mathbb{D})$ *and with coefficients* $c_n(f, N)$ *of the form* (8.6) *there exist functions* $f \in \mathcal{C}(\mathbb{T})$ *such that* $\sup_{N\in\mathbb{N}} \|\mathfrak{A}_N f\|_\infty = \infty$ *and such that* $\limsup_{N\to\infty} \|f - \mathfrak{A}_N f\|_\infty = \infty$.

Proof. We have to show that the operator norms $\|\mathfrak{A}_N\|_{\mathcal{C}(\mathbb{T})\to A(\mathbb{D})}$ are not uniformly bounded. To this end, the Lebesgue constant $\sup_{|z|<1} L_N(z)$, with $L_N(z) = \frac{1}{2\pi} \int_{-\pi}^{\pi} K_N(\mathrm{e}^{\mathrm{i}\theta}, z)\, \mathrm{d}\theta$ and with the kernel (8.8) is considered. But first, the expression

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |L_N(\rho \mathrm{e}^{\mathrm{i}\tau})|\, \mathrm{d}\tau = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} |K_N(\mathrm{e}^{\mathrm{i}\theta}, \rho \mathrm{e}^{\mathrm{i}\tau})|\, \mathrm{d}\tau\, \mathrm{d}\theta \tag{8.12}$$

with $0 < \rho < 1$ is investigated, where in the last term the order of integration has been interchanged. First, the inner integral is analysed. Inserting the expression (8.8) for the kernel and replacing the basis functions by its Taylor series $\varphi_n(z) = \sum_{l=0}^{\infty} \hat{\varphi}_n(l)\, z^l$ gives

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |K_N(\mathrm{e}^{\mathrm{i}\theta}, \rho \mathrm{e}^{\mathrm{i}\tau})|\, \mathrm{d}\tau = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \sum_{l=0}^{\infty} \Big[\sum_{n=1}^{N} \overline{\gamma_{n,N}(\mathrm{e}^{\mathrm{i}\theta})}\, \hat{\varphi}_n(l) \Big] \rho^l\, \mathrm{e}^{\mathrm{i}l\tau} \right| \mathrm{d}\tau \ .$$

The integrand of the last expression is an H^1-function. Therefore, we can apply Hardy's inequality (2.40) and obtain

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}|K_N(\mathrm{e}^{\mathrm{i}\theta},\rho\mathrm{e}^{\mathrm{i}\tau})|\,\mathrm{d}\tau \geq \frac{1}{\pi}\sum_{l=0}^{\infty}\frac{\rho^l}{l+1}\Big|\sum_{n=1}^{N}\overline{\gamma_{n,N}(\mathrm{e}^{\mathrm{i}\theta})}\,\hat{\varphi}_n(l)\Big| .$$

Using this result in (8.12) and interchanging integration and summation, one obtains

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}|L_N(\rho\mathrm{e}^{\mathrm{i}\tau})|\,\mathrm{d}\tau \geq \frac{1}{\pi}\sum_{l=0}^{\infty}\frac{\rho^l}{l+1}\,\frac{1}{2\pi}\int_{-\pi}^{\pi}\Big|\sum_{n=1}^{N}\overline{\gamma_{n,N}(\mathrm{e}^{\mathrm{i}\theta})}\,\hat{\varphi}_n(l)\,\mathrm{e}^{\mathrm{i}l\theta}\Big|\,\mathrm{d}\theta$$

where we introduced an additional term $\mathrm{e}^{\mathrm{i}l\theta}$ in the integrand, which certainly does not change the value of the integral. Next, the operation of taking the modulus is pulled out of the integral. In this way, the value of the integration becomes smaller or remains unchanged. Afterwards, the inner summation is interchanged with the integration, which finally yields

$$\begin{aligned}\frac{1}{2\pi}\int_{-\pi}^{\pi}|L_N(\rho\mathrm{e}^{\mathrm{i}\tau})|\,\mathrm{d}\tau &\geq \frac{1}{\pi}\sum_{l=0}^{\infty}\frac{\rho^l}{l+1}\left|\sum_{n=1}^{N}\hat{\varphi}_n(l)\,\frac{1}{2\pi}\int_{-\pi}^{\pi}\overline{\gamma_{n,N}(\mathrm{e}^{\mathrm{i}\theta})}\,\mathrm{e}^{\mathrm{i}l\theta}\,\mathrm{d}\theta\right| \\ &= \frac{1}{\pi}\sum_{l=0}^{\infty}\frac{\rho^l}{l+1}\left|\sum_{n=1}^{N}\hat{\varphi}_n(l)\,\overline{\hat{\gamma}_{n,N}(l)}\right| .\end{aligned}$$

Now, we take the limit inferior of the last expression and apply Proposition 8.5. This gives

$$\liminf_{N\to\infty}\frac{1}{2\pi}\int_{-\pi}^{\pi}|L_N(\rho\mathrm{e}^{\mathrm{i}\tau})|\,\mathrm{d}\tau \geq \frac{1}{\pi}\sum_{l=0}^{\infty}\frac{\rho^l}{l+1} = \frac{1}{\pi\,\rho}\log\frac{1}{1-\rho}$$

for all $0<\rho<1$. It is clear that

$$\sup_{|z|<1}|L_N(z)| \geq \sup_{|z|<\rho}|L_N(z)| \geq \frac{1}{2\pi}\int_{-\pi}^{\pi}|L_N(\rho\mathrm{e}^{\mathrm{i}\tau})|\,\mathrm{d}\tau . \tag{8.13}$$

Together with (8.9) one obtains therefore

$$\liminf_{N\to\infty}\|\mathfrak{A}_N\|_{\mathcal{C}(\mathbb{T})\to A(\mathbb{D})} = \liminf_{N\to\infty}\sup_{|z|<1}\left|\frac{1}{2\pi}\int_{-\pi}^{\pi}K_N(\mathrm{e}^{\mathrm{i}\theta},z)\,\mathrm{d}\theta\right| = \infty$$

which shows that the operators $\mathfrak{A}_N : \mathcal{C}(\mathbb{T}) \to A(\mathbb{D})$ are not uniformly bounded, and by the uniform boundedness principle this implies that $\sup_{N\in\mathbb{N}}\|\mathfrak{A}_N f\|_\infty = \infty$ for all f belonging to a dense subset of $\mathcal{C}(\mathbb{T})$. □

It follows in particular, that to every such approximation operator $\mathfrak{A}_N$, characterized in Theorem 8.7, there exist continuous functions $f\in\mathcal{C}(\mathbb{T})$ such that the approximation error $\|f-\mathfrak{A}_N f\|_\infty$ increases as the approximation degree N is increased. In conclusion, Theorem 8.7 shows that there exists no robust, causal and linear approximation method on $\mathcal{C}(\mathbb{T})$.

Discrete measures

As in the previous paragraph, we consider approximation methods (8.1) with property (B), (C), and with the M–property. But now, it is assumed that every measure $\mu_{n,N}$ in the coefficient functionals (8.2) is discrete, i.e. every $\mu_{n,N}$ is concentrated on a finite sampling set

$$\mathcal{T}_{n,N} = \{\mathrm{e}^{\mathrm{i}\theta_1(n,N)}, \mathrm{e}^{\mathrm{i}\theta_2(n,N)}, \cdots, \mathrm{e}^{\mathrm{i}\theta_M(n,N)}\}$$

of points on the unit circle. Note that both the sampling points $\theta_m(n,N)$ as well as the number of sampling points $M(n,N)$ may be different for every measure $\mu_{n,N}$ and for different degrees N. With these assumptions, the coefficient functionals (8.2) become

$$c_n(f,N) = \sum_{m=0}^{M(n,N)} f(\mathrm{e}^{\mathrm{i}\theta_m(n,N)})\, \nu_{n,N}(m) \tag{8.14}$$

in which $\nu_m(n,N) = \mu_{n,N}(\mathrm{e}^{\mathrm{i}\theta_m(n,N)})$ denotes the measure of the point $\mathrm{e}^{\mathrm{i}\theta_m(n,N)}$.

First, we derive a representation of the approximation operator $\mathfrak{A}_N$ which is similar to the integral representation (8.7) in the case of absolute continuous measures. To this end, we plug the coefficients (8.14) into the approximation (8.1). This gives

$$(\mathfrak{A}_N f)(z) = \sum_{n=1}^{N} \sum_{m=1}^{M(n,N)} f(\mathrm{e}^{\mathrm{i}\theta_m(n,N)})\, \nu_{n,N}(m)\, \varphi_n(z) \ . \tag{8.15}$$

In this representation, there may exist sampling points which appear several times in the above double sum. Thus, there may exist indices $m_1 \neq m_2 \neq \cdots \neq m_q$ and indices $n_1 \neq n_2 \neq \cdots \neq n_q$ but such that $\theta_{m_1}(n_1,N) = \theta_{m_2}(n_2,N) = \cdots = \theta_{m_q}(n_q,N)$ for some $q \geq 1$. Of course, also the individual measures $\nu_{n,N}(m)$ are different in this case for different indices, i.e. $\nu_{n_1,N}(m_1) \neq \nu_{n_2,N}(m_2) \neq \cdots \neq \nu_{n_q,N}(m_q)$, in general. For such indices n,m for which the sampling points $\theta_m(n,N)$ are equal, we define $\theta_k(N) := \theta_{m_1}(n_1,N) = \cdots = \theta_{m_q}(n_q,N)$ in which the index $k = 1,2,\cdots,K(N)$ numbers all distinguished sampling points $\theta_k(N)$, such that now $\theta_{k_1}(N) \neq \theta_{k_2}(N)$ whenever $k_1 \neq k_2$. Moreover, we define the new *kernel functions*

$$\kappa_{k,N}(z) := \nu_{n_1,N}(m_1)\, \varphi_{n_1}(z) + \nu_{n_2,N}(m_2)\, \varphi_{n_2}(z) + \cdots + \nu_{n_q,N}(m_q)\, \varphi_{n_q}(z)$$

for $k = 1,2,\cdots,K(N)$. With these kernel functions, the approximation operator (8.15) can be rewritten with only one summation over all distinguished sampling points as

$$(\mathfrak{A}_N f)(z) = \textstyle\sum_{k=1}^{K(N)} f(\mathrm{e}^{\mathrm{i}\theta_k(N)})\, \kappa_{k,N}(z) \ . \tag{8.16}$$

Note that the kernel functions $\kappa_{k,N}$ are elements of the disk algebra $A(\mathbb{D})$ since they are linear combinations of the basis functions $\varphi_n \in A(\mathbb{D})$. Moreover, from (8.16) follows at once that $\mathfrak{A}_N f_1 = \mathfrak{A}_N f_2$ for all functions f_1 and f_2 which coincide on all sampling points $\mathrm{e}^{\mathrm{i}\theta_k(N)}$, $k = 1, 2, \cdots, K(N)$. The representation (8.16) of $\mathfrak{A}_N$, which is the discrete equivalent of the integral representation (8.7), will be used subsequently.

Example 8.8. Discrete operators of the form (8.16) arise naturally in practical applications from the continuous case (8.7), due to the numerical integration of (8.7) on digital computers. For example, assume that a certain approximation method of the form (8.7) with a corresponding kernel (8.8) is given (e.g. the method in Example 8.4). Then, the integral in (8.7) may be approximated by its Riemann sum at the equidistant sampling points $\theta_k(N) := \frac{2\pi k}{2N+1}$ with $k = 0, 1, 2, \cdots, 2N$. Therewith, an approximation

$$(\widetilde{\mathfrak{A}}_N f)(z) = \frac{1}{2N+1} \sum_{k=0}^{2N} f(\mathrm{e}^{\mathrm{i}\frac{2\pi k}{2N+1}})\, K_N(\mathrm{e}^{\mathrm{i}\frac{2\pi k}{2N+1}}, z)$$

of (8.7) is obtained. With $\kappa_{k,N}(z) := \frac{1}{2N+1} K_N(\mathrm{e}^{\mathrm{i}\frac{2\pi k}{2N+1}}, z)$, this operator has exactly the form (8.16). It is clear that by increasing the number of sampling points, the error due to the numerical integration can be made as small as desired, and one would expect that if the continuous operator (8.7) converges to f, than also the discrete operator $\widetilde{\mathfrak{A}}_N f$ converges to f as $N \to \infty$ (at most with finite approximation error $\lim_{N\to\infty} \|f - \widetilde{\mathfrak{A}}_N f\|_\infty$).

As in the case of the absolute continuous measures in the coefficient functionals, we present at the beginning a consequence of the required M–property of the approximation method $\mathfrak{A}_N$.

Proposition 8.9. *Let $\mathfrak{A}_N$ be a discrete approximation operator of the form* (8.16). *If $\mathfrak{A}_N$ satisfies the property M then it holds for the Fourier coefficients $\hat{\kappa}_{k,N}(l)$ of the kernel functions in* (8.16) *that*

$$\lim_{N\to\infty} \sum_{k=1}^{K(N)} \hat{\kappa}_{k,N}(l)\, \mathrm{e}^{\mathrm{i}\theta_k(N)\cdot m} = \begin{cases} 1 & \text{for } l = m \\ 0 & \text{for } l \neq m \end{cases} .$$

Proof. Since the kernel functions $\kappa_{k,N}$ are elements of $A(\mathbb{D})$, they can be written as a Taylor series $\kappa_{k,N}(z) = \sum_{l=0}^{\infty} \hat{\kappa}_{k,N}(l)\, z^l$. Now, for a fixed integer $m \geq 0$, we apply (8.16) to the polynomial $p_m = z^m$, and replace the kernel functions $\kappa_{k,N}$ by the corresponding Taylor series. This yields

$$(\mathfrak{A}_N p_m)(z) = \sum_{l=0}^{\infty} \Big(\sum_{k=1}^{K(N)} \hat{\kappa}_{k,N}(l)\, \mathrm{e}^{\mathrm{i}\theta_k(N)\cdot m} \Big) z^l \tag{8.17}$$

Since $\mathfrak{A}_N$ satisfies the M-property, it holds that $\lim_{N\to\infty} \mathfrak{A}_N p_m(z) = z^m$. Therewith the statement of the proposition follows from (8.17). □

The following theorem is the discrete equivalent to Theorem 8.7. It will show that every discrete approximation operator (8.16) with property (B) and (C) will not be robust, i.e. it will not have property (A).

Theorem 8.10. *To every discrete approximation operator* $\mathfrak{A}_N : \mathcal{C}(\mathbb{T}) \to A(\mathbb{D})$ *of the form* (8.16) *there exist functions* $f \in \mathcal{C}(\mathbb{T})$ *such that* $\sup_{N\in\mathbb{N}} \|\mathfrak{A}_N f\|_\infty = \infty$ *and such that* $\limsup_{N\to\infty} \|f - \mathfrak{A}_N f\|_\infty = \infty$.

Proof. The operator norm of $\mathfrak{A}_N : \mathcal{C}(\mathbb{T}) \to A(\mathbb{D})$ is given as

$$\|\mathfrak{A}_N\|_{\mathcal{C}(\mathbb{T})\to A(\mathbb{D})} = \sup_{\substack{f\in\mathcal{C}(\mathbb{T})\\ \|f\|_\infty\leq 1}} \|\mathfrak{A}_N f\|_\infty = \sup_{\substack{f\in\mathcal{C}(\mathbb{T})\\ \|f\|_\infty\leq 1}} \sup_{|z|<1} \left| \sum_{k=1}^{K(N)} f(\mathrm{e}^{\mathrm{i}\theta_k(N)})\, \kappa_{k,N}(z) \right| .$$

Write the kernel functions as $\kappa_{k,N}(z) = |\kappa_{k,N}(z)|\, \mathrm{e}^{\mathrm{i}\arg[\kappa_{k,N}(z)]}$ and choose for an arbitrary $z \in \mathbb{D}$ the function $f \in \mathcal{C}(\mathbb{T})$ for which holds that $f(\mathrm{e}^{\mathrm{i}\theta_k(N)}) = \exp(-\mathrm{i}\arg[\kappa_{k,N}(z)])$ and which is continuous between the the sampling points $\mathrm{e}^{\mathrm{i}\theta_k(N)}$. For this function holds obviously that $f \in \mathcal{C}(\mathbb{T})$ and that $\|f\|_\infty \leq 1$. Using this function in the above relation for the operator norm, one sees that

$$\|\mathfrak{A}_N\|_{\mathcal{C}(\mathbb{T})\to A(\mathbb{D})} \geq \sup_{|z|<1} \underbrace{\left(\sum_{k=1}^{K(N)} |\kappa_{k,N}(z)| \right)}_{=:L_N(z)} = \sup_{|z|<1} L_N(z) .$$

Instead of the infinity norm, we consider first the L^1-norm of $L_N(\rho^{\mathrm{i}\tau})$:

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |L_N(\rho \mathrm{e}^{\mathrm{i}\tau})|\, \mathrm{d}\tau = \sum_{k=1}^{K(N)} \frac{1}{2\pi} \int_{-\pi}^{\pi} |\kappa_{k,N}(\rho \mathrm{e}^{\mathrm{i}\tau})|\, \mathrm{d}\tau .$$

Since all $\kappa_{k,N}$ belong to the disk algebra, we can apply Hardy's inequality to obtain a lower bound on the L^1-norm of $L_N(\rho \mathrm{e}^{\mathrm{i}\tau})$ for any fixed $0 < \rho < 1$. This gives

$$\begin{aligned} \frac{1}{2\pi} \int_{-\pi}^{\pi} |L_N(\rho \mathrm{e}^{\mathrm{i}\tau})|\, \mathrm{d}\tau &\geq \sum_{k=1}^{K(N)} \frac{1}{\pi} \sum_{l=0}^{\infty} \frac{\rho^l}{l+1} |\hat{\kappa}_{k,N}(l)| \\ &\geq \frac{1}{\pi} \sum_{l=0}^{\infty} \frac{\rho^l}{l+1} \left| \sum_{k=1}^{K(N)} \hat{\kappa}_{k,N}(l)\, \mathrm{e}^{\mathrm{i}\theta_k(l)} \right| . \end{aligned}$$

To obtain the second line, we first interchanged the order of the summations, than we insert the term $\mathrm{e}^{\mathrm{i}\theta_k(l)}$ inside the modulus operation, which does not change the value of the expression. Finally, the triangle inequality was applied in reverse direction. Now, we proceed exactly as in the proof of Theorem 8.7. Applying Proposition 8.9 and using the relation (8.13) one finally obtains

$$\liminf_{N\to\infty} \|\mathfrak{A}_N\|_{\mathcal{C}(\mathbb{T})\to A(\mathbb{D})} \geq \liminf_{N\to\infty} \sup_{|z|<1} \left(\sum_{k=1}^{K(N)} |\kappa_{k,N}(z)| \right) = \infty \, .$$

Together with the uniform boundedness principle, this proves the theorem. □

Thus, the discrete operators $\mathfrak{A}_N : \mathcal{C}(\mathbb{T}) \to A(\mathbb{D})$ of the form (8.16) show the same behavior as the operators (8.7) which correspond to an absolute continuous measure. In particular, to every discrete operator there exist functions $f \in \mathcal{C}(\mathbb{T})$ (in fact a whole dense subset of functions in $\mathcal{C}(\mathbb{T})$) such that the approximation error $\|f - \mathfrak{A}_N f\|_\infty$ tends to infinity as the approximation degree is increased. However, the discrete approximation operators show an even worse behavior than the operators of the form (8.7). This is because in the latter case, there exist approximation methods of the form (8.7) with property (B) and (C) which are stable as operators from $A(\mathbb{D}) \to A(\mathbb{D})$. This means that there exist linear and causal approximation methods $\mathfrak{A}_N$ of the form (8.7) such that $\lim_{N\to\infty} \|f - \mathfrak{A}_N f\|_\infty = 0$ for all $f \in A(\mathbb{D}) \subset \mathcal{C}(\mathbb{T})$. In the discrete case however, all approximation operators $\mathfrak{A}_N$ of the form (8.16) with property (B) and (C) are unstable, even if they are restricted to the disk algebra $A(\mathbb{D})$. This is a straight forward consequence of the previous result and of the fact that the approximation operator is defined on a set (the sampling set) of measure zero.

Corollary 8.11 (Somorjai). *For every discrete approximation operator* $\mathfrak{A}_N : A(\mathbb{D}) \to A(\mathbb{D})$ *of the form* (8.16) *there exist functions* $f \in A(\mathbb{D})$ *such that* $\sup_{N\in\mathbb{N}} \|\mathfrak{A}_N f\|_\infty = \infty$ *and such that* $\limsup_{N\to\infty} \|f - \mathfrak{A}_N f\|_\infty = \infty$.

Proof (Corollary 8.11). The operators $\mathfrak{A}_N$ are defined only on the sampling sets $\mathcal{T}_N = \bigcup_{n=1}^N \mathcal{T}_{n,N}$. These sets are of (Lebesgue) measure zero and closed. Moreover, since every function which is defined on a discrete set is continuous there, we can apply a theorem of Rudin [71], which states that for every continuous function defined on a compact subset $\mathcal{T}_N \subset \mathbb{T}$ of measure zero, there exists a function $F \in A(\mathbb{D})$ with $F(z) = f(z)$ for all $z \in \mathcal{T}_N$ and with $\|F\|_\infty = \sup_{z\in\mathcal{T}_N} |f(z)|$. Since f and F coincide on the sampling sets $\mathcal{T}_N$, it holds that $\mathfrak{A}_N F = \mathfrak{A}_N f$, and consequently $\|\mathfrak{A}_N F\|_\infty = \|\mathfrak{A}_N f\|_\infty \leq \|\mathfrak{A}_N\|_{A(\mathbb{D})\to A(\mathbb{D})} \|F\|_\infty = \|\mathfrak{A}_N\|_{A(\mathbb{D})\to A(\mathbb{D})} \|f\|_\infty$. Therewith, one obtains on the one hand that

$$\begin{aligned}\|\mathfrak{A}_N\|_{\mathcal{C}(\mathbb{T})\to A(\mathbb{D})} &= \sup_{\substack{f\in\mathcal{C}(\mathbb{T})\\ \|f\|\leq 1}} \|\mathfrak{A}_N f\|_\infty \\ &\leq \sup_{\substack{f\in\mathcal{C}(\mathbb{T})\\ \|f\|\leq 1}} \|\mathfrak{A}_N\|_{A(\mathbb{D})\to A(\mathbb{D})} \|f\|_\infty = \|\mathfrak{A}_N\|_{A(\mathbb{D})\to A(\mathbb{D})} \, .\end{aligned}$$

On the other hand, because $A(\mathbb{D}) \subset \mathcal{C}(\mathbb{T})$, it is clear that

$$\|\mathfrak{A}_N\|_{\mathcal{C}(\mathbb{T})\to A(\mathbb{D})} = \sup_{\substack{f\in\mathcal{C}(\mathbb{T})\\ \|f\|\le 1}} \|\mathfrak{A}_N f\|_\infty \ge \sup_{\substack{f\in A(\mathbb{D})\\ \|f\|\le 1}} \|\mathfrak{A}_N f\|_\infty = \|\mathfrak{A}_N\|_{A(\mathbb{D})\to A(\mathbb{D})} \, .$$

Altogether $\|\mathfrak{A}_N\|_{A(\mathbb{D})\to A(\mathbb{D})} = \|\mathfrak{A}_N\|_{\mathcal{C}(\mathbb{T})\to A(\mathbb{D})}$ and the corollary follows from Theorem 8.10. □

If the previous proof (together with the proofs of Theorem 8.7 and 8.10) is carefully analyzed, one sees that we even have a lower bound on the maximum amplitude of the approximation $\mathfrak{A}_N f$ inside the unit disk

$$\liminf_{N\to\infty} \sup_{\substack{f\in A(\mathbb{D})\\ \|f\|_\infty\le 1}} \sup_{|z|\le\rho} |(\mathfrak{A}_N f)(z)| \ge \frac{1}{\pi\,\rho} \log \frac{1}{1-\rho}$$

for all $0 \le \rho < 1$.

Theorem 8.7 showed that the approximation operators $\|\mathfrak{A}_N\|_{\mathcal{C}(\mathbb{T})\to A(\mathbb{D})}$ of the form (8.7) are always unbounded as $N \to \infty$. However, this does not in general mean that the operator norms $\|\mathfrak{A}_N\|_{A(\mathbb{D})\to A(\mathbb{D})}$ are also always unbounded. Quite the contrary, Example 8.4 for instance gave an example of an approximation operator for which $\sup_{N\in\mathbb{N}} \|\mathfrak{A}_N\|_{A(\mathbb{D})\to A(\mathbb{D})} < \infty$.

Corollary 8.11 shows now that in the discrete case also the norms of the approximation operators $\mathfrak{A}_N : A(\mathbb{D}) \to A(\mathbb{D})$ are always unbounded. This means that there exists no discrete linear, causal, and stable approximation method for all $f \in A(\mathbb{D}) \subset \mathcal{C}(\mathbb{T})$.

Note that especially the discrete approximation operators are of practical importance because even if there exists an approximation method of the form (8.7) for which $\sup_{N\in\mathbb{N}} \|\mathfrak{A}_N\|_{A(\mathbb{D})\to A(\mathbb{D})}$ is bounded, a numerical integration method has to be applied to calculate (8.7). If this numerical integration is done on a digital computer, one ends up with a discrete approximation operator (cf. Example 8.8) for which $\sup_{N\in\mathbb{N}} \|\mathfrak{A}_N\|_{A(\mathbb{D})\to A(\mathbb{D})}$ is always unbounded, according to the above corollary.

Finally (and without any proof), we remark that if one requires additionally that for all polynomials $p \in \mathcal{P}_{M(N)}$ with a degree of at most $M(N)$ the approximation method is perfect, i.e. if one requires that $\mathfrak{A}_N p = p$ for all $p \in \mathcal{P}_{M(N)}$, then one can even derive a lower bound on the growth behavior of the operator norms [22], namely that

$$\|\mathfrak{A}_N\|_{A(\mathbb{D})\to A(\mathbb{D})} \ge \tfrac{1}{\pi} \log M(N) \, .$$

Theorems 8.7 and 8.10 show that one cannot have all three desired properties (robustness, causality, and linearity) for the approximation methods (8.1). However, the discussions in Section 8.1 show that there always exist robust and causal, but non linear approximation methods, whereas Section 8.2 showed that one can always have linear, and causal, but non robust methods. We summarize this observation by the following corollary.

Corollary 8.12. *There exists no linear, causal, and robust approximation method of the form* (8.1) *and with coefficients of the form* (8.2)*. If one of the three constraints is relaxed, an approximation method with the remaining two properties always exits.*

8.4 Causal Approximations for Smooth Functions

It was shown in the last section that for all linear and causal approximation methods $\mathfrak{A}_N : \mathcal{C}(\mathbb{T}) \to A(\mathbb{D})$ of the form (8.1) there exist continuous functions $g \in \mathcal{C}(\mathbb{T})$ such that $\sup_{N\in\mathbb{N}} \|\mathfrak{A}_N g\|_\infty = \infty$. This implies that the approximation operator is not robust against errors in the given data. To see this, assume that $f \in A(\mathbb{D})$ is a causal function which should be approximated as $\mathfrak{A}_N f$. However, due to disturbances only $\widetilde{f} = f + g$ with a disturbance $g \in \mathcal{C}(\mathbb{T})$ is known. Therefore, the approximation error $\|\mathfrak{A}_N \widetilde{f} - f\|_\infty$ is lower bounded by (7.12) which shows that the approximation error may become infinite due to the first term on the right hand side of (7.12). This implies in particular that the approximation error cannot be controlled by the size of the disturbance $\|g\|_\infty$. Even for very small values of $\|g\|_\infty$, the term $\|\mathfrak{A}_N g\|_\infty$ may get arbitrarily large.

To ensure that the error term $\|\mathfrak{A}_N g\|_\infty$ remains bounded, the disturbance g has to be from a subset of $\mathcal{C}(\mathbb{T})$. The following result shows that under the assumption $g \in \mathcal{C}_\omega(\mathbb{T})$ with a regular majorant ω, the error $\|\mathfrak{A}_N g\|_\infty$ remains always bounded for every linear and causal approximation operator.

Theorem 8.13. *Let ω be a regular majorant. Then there exist approximation methods $\mathfrak{A}_N : \mathcal{C}_\omega(\mathbb{T}) \to A_\omega(\mathbb{D})$ of the form* (8.2) *with property (B) and (C) such that*

$$\|\mathfrak{A}_N\, g\|_\infty \leq \|\mathfrak{A}_N\, g\|_{\mathcal{C}_\omega(\mathbb{T})} < \infty \qquad \textit{for all } g \in \mathcal{C}_\omega(\mathbb{T})\ .$$

Proof. Let $g \in \mathcal{C}_\omega(\mathbb{T})$ and consider the non-causal approximation g_N in terms of the Fejér means (cf. Section 2.1.2) given by

$$g_N(\mathrm{e}^{\mathrm{i}\theta}) = (\mathfrak{F}_N g)(\mathrm{e}^{\mathrm{i}\theta}) = \frac{1}{2\pi}\int_{-\pi}^{\pi} g(\mathrm{e}^{\mathrm{i}(\theta-\tau)})\, \mathcal{F}_N(\tau)\, \mathrm{d}\tau$$

where $\mathcal{F}_N$ is the Fejér kernel defined in (2.16). Using that $\mathcal{F}_N$ is an approximate identity (Prop. 2.3) it follows for arbitrary $\theta_1, \theta_2 \in [-\pi, \pi)$ that

$$\begin{aligned} |g_N(\mathrm{e}^{\mathrm{i}\theta_1}) - g_N(\mathrm{e}^{\mathrm{i}\theta_2})| &\leq \frac{1}{2\pi}\int_{-\pi}^{\pi} |g(\mathrm{e}^{\mathrm{i}(\theta_1-\tau)}) - g(\mathrm{e}^{\mathrm{i}(\theta_2-\tau)})|\, \mathcal{F}_N(\tau)\, \mathrm{d}\tau \\ &\leq \|g\|_{\mathcal{C}_\omega(\mathbb{T})}\, \omega(|\theta_1 - \theta_2|)\ . \end{aligned}$$

This shows that $\|\mathfrak{F}_N g\|_{\mathcal{C}_\omega(\mathbb{T})} \leq \|g\|_{\mathcal{C}_\omega(\mathbb{T})}$ for every $g \in \mathcal{C}_\omega(\mathbb{T})$. Next, we consider the causal operator defined by

$$(\mathfrak{A}_N g)(e^{i\theta}) = (\mathfrak{P}_+ \mathfrak{F}_N\, g)(e^{i\theta}) , \qquad \theta \in [-\pi, \pi)$$

with the Riesz projection $\mathfrak{P}_+$. Using Theorem 6.16 on the boundedness of the Riesz projection on $\mathcal{C}_\omega(\mathbb{T})$, one obtains

$$\|\mathfrak{A}_N g\|_{\mathcal{C}_\omega(\mathbb{T})} \leq \|\mathfrak{P}_+\|_{\mathcal{C}_\omega(\mathbb{T})\to\mathcal{C}_\omega(\mathbb{T})} \, \|\mathfrak{F}_N g\|_{\mathcal{C}_\omega(\mathbb{T})} \leq C(\omega)\|g\|_{\mathcal{C}_\omega(\mathbb{T})}$$

with a certain constant $C(\omega)$, dependent only on the majorant ω. Since $\mathcal{C}_\omega(\mathbb{T})$ is continuously embedded in $\mathcal{C}(\mathbb{T})$, one has therefore $\|\mathfrak{A}_N g\|_\infty < \infty$ for all $g \in \mathcal{C}_\omega(\mathbb{T})$. By the construction of $\mathfrak{A}_N$, it has obviously properties (B) and (C) as well as the M–property. □

Note that by the properties of the Fejér means, the approximation operator $\mathfrak{A}_N$ used in the previous proof satisfies also (8.3). Thus there exist linear and causal approximation methods which are robust with respect to smooth functions in $\mathcal{C}_\omega(\mathbb{T})$ with a regular majorant ω.

Notes

Some of the results of this section can also be found in in the book of Partington [64]. Corollary 8.11 is due to Somorjai [82].

9

On Algorithms for Calculating the Hilbert Transform

Corollary 8.11 of the previous chapter shows that there exists no robust linear approximation method for causal and stable transfer functions $f \in A(\mathbb{D})$ which are defined only on a finite set of discrete sampling points. It was discussed that these convergence problems of the approximation methods are a consequence of the fact that the approximation operator is only defined on a finite, discrete sampling set. However, the sampling of the given data is essential in practical applications, since nowadays numerical calculations are (almost) exclusively carried out on digital computers and such digital computers can process only a finite number of input data. For these reasons, the present chapter will discuss the consequences of the sampling of the given data for the behavior of certain numerical algorithms, a little bit more. Thereby, we will mainly focus on the calculation of the Hilbert transform from sampled data. However, these results carry over directly to algorithms for the calculation of the spectral factorization, Wiener filter or any other algorithm which involves explicitly or implicitly the determination of the algebraic conjugate of a given function.

Thus, we investigate algorithms which determine the Hilbert transform $\widetilde{f} = \mathfrak{H}f$ of a function f given on the unit circle[1]. Since both f and $\widetilde{f}$ are defined only on $\mathbb{T}$, we will write $f(\theta)$ instead of $f(\mathrm{e}^{\mathrm{i}\theta})$ with $\theta \in [-\pi, \pi]$, throughout this chapter, to simplify the notations. Moreover, $\mathbb{T}$ will now stand for the interval $[-\pi, \pi)$ of the real axis $\mathbb{R}$ and $\mathcal{C}(\mathbb{T})$ denotes the set of all continuous functions f on $\mathbb{T}$ with $f(-\pi) = f(\pi)$ or equivalently for all continuous, 2π-periodic functions on $\mathbb{R}$.

We consider linear operators $\mathfrak{T}$ which determine an approximation of the Hilbert transform $\widetilde{f} = \mathfrak{H}f$ from the values $f(\tau_k)$ of the given function f on a finite set $S = \{\tau_k \in \mathbb{T} : k = 1, 2, \cdots, N\}$ of sampling points τ_k, only. Denote by $\mathfrak{T}_N$ such a linear operator which approximates the Hilbert transform based on a sampling set $\mathcal{S}$ of cardinality N. Then we say that a sequence $\{\mathfrak{T}_N\}_{N\in\mathbb{N}}$ of such operators approximates the conjugate function $\widetilde{f}$ of $f \in \mathcal{C}(\mathbb{T})$ arbitrarily

[1] See Section 5.3 for the definition of the Hilbert transform.

V. Pohl, H. Boche, *Advanced Topics in System and Signal Theory*, Foundations in Signal Processing, Communications and Networking 4,
DOI 10.1007/978-3-642-03639-2_9,

well (in the norm of $\mathcal{C}(\mathbb{T})$), if

$$\lim_{N\to\infty} \left\|\widetilde{f} - \mathfrak{T}_N f\right\|_\infty = 0\,.$$

In the following it is assumed that the operators $\mathfrak{T}$ have the following two natural properties.

Definition 9.1 (Property I). *We say that an operator $\mathfrak{T}$ has the* property I *if it is linear, i.e.*

$$\mathfrak{T}(f_1 + f_2) = \mathfrak{T}f_1 + \mathfrak{T}f_2 \quad \textit{and} \quad \mathfrak{T}(\lambda f) = \lambda\mathfrak{T}(f)$$

for all $\lambda \in \mathbb{C}$, and if $\mathfrak{T}$ is concentrated on the set S of sampling points, i.e. if two functions f_1 and f_2 coincide on the sampling point $f_1(\tau_k) = f_2(\tau_k)$ for all $\tau_k \in S$ then $(\mathfrak{T}f_1)(t) = (\mathfrak{T}f_2)(t)$ for all $t \in \mathbb{T}$.

This property requires the linearity of the operator $\mathfrak{T}$ and the concentration of the operator on the finite set S, i.e. if two functions f_1 and f_2 coincide on the sampling set $\mathcal{S}$, the operator $\mathfrak{T}$ will give the same result for both functions. It is immediately clear that all practical algorithms for the calculation of the conjugate function $\widetilde{f}$ from f which can be implemented on a digital computer, have to satisfy this property since generally only a finite number of values can be taken into account during the calculation on such a computer. Therefore, this property is practically no limitation on the linear operators under consideration.

Example 9.2. The *conjugate Shannon sampling series*, given by

$$(\mathfrak{S}_N f)(t) = \frac{1}{2N+1} \sum_{k=0}^{2N} f(\tfrac{2\pi k}{2N+1}) \frac{\cos\frac{2N+1}{2}(t - \frac{2\pi k}{2N+1}) - \cos\frac{1}{2}(t - \frac{2\pi k}{2N+1})}{\sin\frac{1}{2}(t - \frac{2\pi k}{2N+1})}$$

is one example of such an operator. It calculates an approximation of the conjugate function $\widetilde{f}$ based on the function f given only at the points of the sampling set $S = \{\tau_k = 2\pi k/(2N+1) : k = 0, 1, \dots, 2N\}$.

Example 5.9 shows that there exist continuous functions $f \in \mathcal{C}(\mathbb{T})$ such that the conjugate $\widetilde{f} = \mathfrak{H}f$ is not continuous on $\mathbb{T}$. For this reason, it cannot be expected that a linear operator $\mathfrak{T}$ approximates the Hilbert transform of such functions arbitrary well, in general. Therefore the operators $\mathfrak{T}$ are considered only on the set $\mathcal{B}$ of all continuous functions which have a continuous Hilbert transform:

$$\mathcal{B} := \left\{f \in \mathcal{C}(\mathbb{T}) : \widetilde{f} = \mathfrak{H}f \in \mathcal{C}(\mathbb{T})\right\}$$

The norm in $\mathcal{B}$ is defined by $\|f\|_{\mathcal{B}} := \max(\|f\|_\infty, \|\widetilde{f}\|_\infty)$. Thus, the Hilbert transform $\mathfrak{H}$ maps every function $f \in \mathcal{B}$ onto the continuous function $\widetilde{f} = \mathfrak{H}f$ which shows that $\mathfrak{H} : \mathcal{B} \to \mathcal{C}(\mathbb{T})$ is a continuous mapping with $\|\mathfrak{H}\|_{\mathcal{B}\to\mathcal{C}(\mathbb{T})} \le 1$. Note, that if a function $f \in \mathcal{B}$ is approximated by a trigonometric polynomial

$P_N f$ of degree N, then the Hilbert transform $\widetilde{P_N f}$ of this trigonometric polynomial is an approximation of the conjugate function $\widetilde{f}$. This property of the set $\mathcal{B}$ is used later.

Thus we consider operators $\mathfrak{T} : \mathcal{B} \to \mathcal{C}(\mathbb{T})$ with property I and which map every function $f \in \mathcal{B}$ to a continuous function $\mathfrak{T}f \in \mathcal{C}(\mathbb{T})$. The corresponding operator norm is defined, as usual, by

$$\|\mathfrak{T}\|_{\mathcal{B}\to\mathcal{C}(\mathbb{T})} := \sup_{f\in\mathcal{B}} \frac{\|\mathfrak{T}f\|_\infty}{\|f\|_{\mathcal{B}}} \ .$$

In particular sequences $\{\mathfrak{T}_N\}_{N=1}^\infty$ of such linear operators are considered. For these sequences of operators, we will require the following property II.

Definition 9.3 (Property II). *We say that a sequence $\{\mathfrak{T}_N\}_{N=1}^\infty$ of linear operators has the* property II *if there exists a dense subsets $\mathcal{M} \subset \mathcal{B}$ such that*

$$\lim_{N\to\infty} \left\|\widetilde{f} - \mathfrak{T}_N f\right\|_\infty = 0 \ , \quad \textit{for all } f \in \mathcal{M} \ .$$

This property requires the desired approximation behavior of the sequence of operators for a dense subset of $\mathcal{B}$. Clearly, this is no restriction on the operators since this behavior should hold for the whole set $\mathcal{B}$ and therefore also at least for a dense subset.

Example 9.4. Let $f \in \mathcal{C}(\mathbb{T})$ and let

$$(\mathfrak{V}_N f)(t) = \frac{2}{2N+1} \sum_{k=0}^{2N} f\left(\tfrac{2\pi k}{2N+1}\right) V_N\left(t - \tfrac{2\pi k}{2N+1}\right)$$

with the de la Vallée-Poussin kernel V_N given in (2.25), be the *de la Vallée-Poussin series* of f. Then $\widetilde{\mathfrak{V}}_N f := \mathfrak{H}(\mathfrak{V}_N f) = \widetilde{\mathfrak{V}_N f}$ defines an operator sequence $\{\widetilde{\mathfrak{V}}_N\}_{N=1}^\infty$, in which every single operator $\widetilde{\mathfrak{V}}_N$ has property I. Moreover, if $\mathcal{P}_N$ denotes the set of all trigonometric polynomials with a degree of at most N, then it holds $\widetilde{\mathfrak{V}}_N f = \widetilde{f}$ for all $f \in \mathcal{P}_{N/2}$. This shows that this sequence has also property II since it approximates the conjugate function $\widetilde{f}$ arbitrarily well for every polynomial in $\mathcal{B}$, and since this set of polynomials is a dense subset of $\mathcal{B}$.

Example 9.5. Let $\varphi_N(\tau) := \max(1 - \frac{N\,|\tau|}{\pi}\,, 0)$ and define by

$$(\mathfrak{L}_N f)(t) := \sum_{k=0}^{2N} f(\tfrac{2\pi k}{2N+1})\, \varphi_N(t - \tfrac{2\pi k}{2N+1}) \tag{9.1}$$

an operator which linearly interpolates the function f between the sampling points $\{(2\pi k)/(2N+1)\}_{k=0}^{2N}$. It is clear that the Hilbert transform of $\mathfrak{L}_N f$ always exists. Therefore,

$$\widetilde{\mathfrak{L}}_N f := \mathfrak{H}(\mathfrak{L}_N f) = \widetilde{\mathfrak{L}_N f}$$

defines a sequence of linear operators with property I. To show that this operator sequence has also property II is more complicated to show. This will be done later.

Continuous functions

First we investigate whether it is always possible to calculate the conjugate function $\widetilde{f}$ arbitrarily well for all $f \in \mathcal{B}$ from a finite set of sampled values $\{f(\tau_k)\}_{k=1}^N$ by a linear method. That this is actually not possible is a direct consequence of the following theorem.

Theorem 9.6. *Let $\{\mathfrak{T}_N\}_{N\in\mathbb{N}}$ be a sequence of operators with property I and II, then the set of all $f \in \mathcal{B}$ for which*

$$\limsup_{N\to\infty} \|\mathfrak{T}_N f\|_\infty = \infty$$

is of second category (a non-meager set) and dense in $\mathcal{B}$.

Thus even though the sequence $\{\mathfrak{T}_N\}_{N\in\mathbb{N}}$ approximates the conjugate function for a dense subset of $\mathcal{B}$ (as assumed by property II), it fails to converge for all functions f in another dense subset $\mathcal{B}$. Two direct consequences of this result with respect to the operator norms of $\{\mathfrak{T}_N\}$ and the approximation behavior of the conjugate function are formulated in the next two corollaries.

Corollary 9.7. *Let $\{\mathfrak{T}_N\}_{N\in\mathbb{N}}$ be a sequence of operators with property I and II. Then $\lim_{N\to\infty} \|\mathfrak{T}_N\|_{\mathcal{B}\to\mathcal{C}(\mathbb{T})} = \infty$.*

Corollary 9.8. *Let $\{\mathfrak{T}_N\}_{N\in\mathbb{N}}$ be a sequence of operators with property I. Then there exists a dense subset $\mathcal{E}$ of $\mathcal{B}$ such that*

$$\limsup_{N\to\infty} \left\|\widetilde{f} - \mathfrak{T}_N f\right\|_\infty > 0$$

for all $f \in \mathcal{E}$.

Now we prove the above results. To this end, the following auxiliary lemma is needed. It shows that to every continuous function f there exists a function $\varphi \in \mathcal{B}$ such that f and φ coincide at the finite sampling set $\mathcal{S}$ and such that the norm of φ is at most twice the norm of f.

Lemma 9.9. *Let $\{\tau_k\}_{k=1}^N$ be an arbitrary finite set of sampling points in $\mathbb{T}$. Then to every function $f \in \mathcal{C}(\mathbb{T})$ there exists a function $\varphi \in \mathcal{B}$ such that $\varphi(\tau_k) = f(\tau_k)$ for all $k = 1, 2, \dots, N$ and with $\|\varphi\|_{\mathcal{B}} \leq 2\,\|f\|_\infty$.*

The main problem in finding such a function φ is that not only the behavior of φ but also the behavior of the conjugate $\widetilde{\varphi}$ has to be controlled so that $\|\varphi\|_{\mathcal{B}}$ does not become larger than $2\|f\|_\infty$.

Proof. For a given function $f \in \mathcal{C}(\mathbb{T})$, we construct an $f_{\eta_0} \in \mathcal{B}$ which satisfies the lemma. For an arbitrary $0 < \eta \leq 1$ consider the function

$$g_\eta(\tau) = \max\left(1 - \tfrac{|\tau|}{\eta\,\pi}\;,\;0\right)\;, \qquad \tau \in [-\pi, \pi)\;.$$

This function can also be written as a Fourier series

$$g_\eta(\tau) = \tfrac{a_0(\eta)}{2} + \textstyle\sum_{k=1}^{\infty} a_k(\eta)\cos(k\tau)$$

with $a_0 = \eta$ and

$$a_k(\eta) = \tfrac{2}{\eta\pi^2 k^2}\left[1 - \cos(\eta k \pi)\right]\;, \qquad k = 1, 2, \dots\;.$$

Note that all Fourier coefficients $a_k(\eta)$ are non-negative. Moreover, it holds that $\|g_\eta\|_{\ell^1} = 1$ because $\left\|g_\eta\right\|_{\ell^1} = a_0/2 + \sum_{k>1}|a_k(\eta)| = g_\eta(0) = 1$, using that all a_k are non-negative. The conjugate function $\widetilde{g}_\eta$ is then

$$\widetilde{g}_\eta(\tau) = \textstyle\sum_{k=1}^{\infty} a_k(\eta)\sin(k\tau)\;.$$

It follows that

$$|\widetilde{g}_\eta(\tau)| \leq \textstyle\sum_{k=1}^{\infty}|a_k(\eta)| = 1 - \eta \tag{9.2}$$

for all $\tau \in [-\pi, \pi)$, where the last inequality follows from $\|g_\eta\|_{\ell^1} = 1$.

Consider now the function $f_\eta(\tau) = \sum_{k=1}^{N} f(\tau_k)\, g_\eta(\tau - \tau_k)$ with the corresponding conjugate function given by

$$\widetilde{f}_\eta(\tau) = \textstyle\sum_{k=1}^{N} f(\tau_k)\,\widetilde{g}_\eta(\tau - \tau_k)\;. \tag{9.3}$$

It is clear that for sufficiently small η ($\eta\,\pi \leq \min_k|\tau_{k+1} - \tau_k|$) it can be achieved that $\|f_\eta\|_\infty \leq \|f\|_\infty$. Since $g_\eta(\tau)$ is zero for $|\tau| \geq \eta\pi$, the conjugate functions $\widetilde{g}_\eta$ can be written as

$$\widetilde{g}_\eta(\tau) = \frac{1}{2\pi}\int_{-\eta\pi}^{\eta\pi} g_\eta(s)\;\tan^{-1}\left(\tfrac{s-\tau}{2}\right)\,\mathrm{d}s$$

provided that $|\tau| \geq \eta\,\pi$. A straight forward calculation gives

$$|\widetilde{g}_\eta(\tau)| \leq \frac{1}{2\pi}\int_{-\eta\pi}^{\eta\pi}\frac{\mathrm{d}s}{\tan\left(\frac{|\tau|-s}{2}\right)} = \frac{1}{\pi}\log\frac{\sin\left(\frac{|\tau|+\eta\pi}{2}\right)}{\sin\left(\frac{|\tau|-\eta\pi}{2}\right)}$$

for all τ with $|\tau| \geq \eta\,\pi$, where the first inequality follows from $\|\widetilde{g}_\eta\|_\infty \leq 1$ by (9.2). From this, it becomes clear that for a fixed $\delta > \eta\,\pi$ and for an arbitrary $\epsilon > 0$ there exists an $\eta_0 = \eta(\epsilon, \delta)$ such that

$$|\widetilde{g}_{\eta_0}(\tau)| < \epsilon/N \tag{9.4}$$

for all $\delta \leq |\tau| \leq \pi$ and all $0 < \eta \leq \eta_0$. Now choose δ_0 such that for all $k \neq l$

$$[\tau_k - \delta_0, \tau_k + \delta_0] \cap [\tau_l - \delta_0, \tau_l + \delta_0] = \emptyset. \tag{9.5}$$

For this δ_0 and an arbitrary $\epsilon > 0$ choose $\eta_0 = \eta(\epsilon, \delta_0)$ such that (9.4) holds.

Finally, we consider $\left|f_{\eta_0}(\tau)\right|$ and $\left|\widetilde{f}_{\eta_0}(\tau)\right|$ at an arbitrary $\tau \in [-\pi, \pi)$. Because of (9.5), there exists at most one τ_k such that $\tau \in [\tau_k - \delta_0, \tau_k + \delta_0]$, and from (9.3) follows

$$\left|\widetilde{f}_{\eta_0}(\tau)\right| \leq \|f\|_\infty \textstyle\sum_{k=1}^{N} |\widetilde{g}_{\eta_0}(\tau - \tau_k)| \tag{9.6}$$

and a similar inequality holds also for $|f_{\eta_0}(\tau)|$. From the inequality for $|f_{\eta_0}(\tau)|$ follows at once that $\|f_{\eta_0}\|_\infty \leq \|f\|_\infty$, but for $\left|\widetilde{f}_{\eta_0}(\tau)\right|$ we have to distinguish two cases: First $\tau \in [\tau_l - \delta_0, \tau_l + \delta_0]$ for a certain l. Then (9.6) becomes

$$\begin{aligned}\left|\widetilde{f}_{\eta_0}(\tau)\right| &\leq \|f\|_\infty \Big(\sum_{k=1, k\neq l}^{N} |\widetilde{g}_{\eta_0}(\tau - \tau_k)| + |\widetilde{g}_{\eta_0}(\tau - \tau_l)| \Big) \\ &\leq \|f\|_\infty \left[\tfrac{\epsilon}{N}(N-1) + 1 - \eta_0 \right] \leq 2\, \|f\|_\infty\end{aligned}$$

using (9.2) and (9.4). In the second case, $\tau \notin [\tau_l - \delta_0, \tau_l + \delta_0]$ for every $l = 1, 2, \ldots, N$. Then (9.6) together with (9.4) show that $\left|\widetilde{f}_{\eta_0}(\tau)\right| \leq \epsilon\, \|f\|_\infty$. This proves that $\|\widetilde{f}_{\eta_0}\|_\infty \leq 2\, \|f\|_\infty$ and gives the statement of the lemma. □

With these preparations we prove Theorem 9.6.

Proof (Theorem 9.6). The theorem is proved by contradiction. Assume that there exists a set $\mathcal{B}_0 \subset \mathcal{B}$ of second category which is dense in $\mathcal{B}$ and a constant $C_0 < \infty$ such that

$$\limsup_{N\to\infty} \|\mathfrak{T}_N f\|_\infty \leq C_0 \tag{9.7}$$

for all $f \in \mathcal{B}_0$. The norm $\|\mathfrak{T}_N f\|_\infty$ is finite for finite N. Therefore, it follows from (9.7) that there exists a constant $C_1 < \infty$ such that $\sup_{N\in\mathbb{N}} \|\mathfrak{T}_N f\|_\infty \leq C_1$ for all $f \in \mathcal{B}_0$, and from the uniform boundedness principle (theorem of Banach-Steinhaus, see e.g. [54]) follows that the operator norms are uniformly bounded, i.e. there exists a constant $C_2 < \infty$ such that

$$\|\mathfrak{T}_N\|_{\mathcal{B}\to\mathcal{C}(\mathbb{T})} \leq C_2 \tag{9.8}$$

for all $N \in \mathbb{N}$. Next it is shown that from the uniform boundedness of the operators $\mathfrak{T}_N$ and from property II follows that the conjugate function $\widetilde{f}$ of every $f \in \mathcal{B}$ can be approximated arbitrarily well by $\mathfrak{T}_N f$. To this end let $f \in \mathcal{B}$ and $g \in \mathcal{M}$ arbitrary, where $\mathcal{M}$ is a dense subset of $\mathcal{B}$. Then it holds

$$\begin{aligned}\|\widetilde{f} - \mathfrak{T}_N f\|_\infty &\leq \|\widetilde{f} - \widetilde{g}\|_\infty + \|\widetilde{g} - \mathfrak{T}_N g\|_\infty + \|\mathfrak{T}_N(g - f)\|_\infty \\ &\leq (1 + C_2)\, \|f - g\|_{\mathcal{B}} + \|\widetilde{g} - \mathfrak{T}_N g\|_\infty\end{aligned} \tag{9.9}$$

using (9.8) and the obvious relation $\|\widetilde{f}-\widetilde{g}\|_\infty \le \|f-g\|_{\mathcal{B}}$. Let $\epsilon > 0$ arbitrary and choose $g \in \mathcal{M}$ such that $\|f-g\|_{\mathcal{B}} < \epsilon$. Then it follows from (9.9) that $\limsup_{N\to\infty} \|\widetilde{f}-\mathfrak{T}_N f\|_\infty \le (1+C_2)\cdot\epsilon$ using that $\{\mathfrak{T}_N\}$ has property II. Since the left hand side is independent of ϵ, and ϵ was chosen arbitrary, we have

$$\lim_{N\to\infty} \|\widetilde{f}-\mathfrak{T}_N f\|_\infty = 0 \quad \text{for all } f \in \mathcal{B}\,. \tag{9.10}$$

Lemma 9.9 shows that to every $f \in \mathcal{C}(\mathbb{T})$ there exists a $\varphi \in \mathcal{B}$ with $\|\varphi\|_{\mathcal{B}} \le 2\,\|f\|_\infty$ such that $\varphi(\tau_k) = f(\tau_k)$ for all $k = 1, 2, \cdots, N$. Since the sequence $\{\mathfrak{T}_N\}$ has property I it follows that $\mathfrak{T}_N\varphi = \mathfrak{T}_N f$. Therewith, we get

$$\|\mathfrak{T}_N\|_{\mathcal{C}(\mathbb{T})\to\mathcal{C}(\mathbb{T})} = \sup_{\substack{f\in\mathcal{C}(\mathbb{T})\\ \|f\|_\infty\le 1}} \|\mathfrak{T}_N f\|_\infty \le \sup_{\substack{\varphi\in\mathcal{B}\\ \|\varphi\|_{\mathcal{B}}\le 2}} \|\mathfrak{T}_N \varphi\|_\infty \le 2\,\|\mathfrak{T}_N\|_{\mathcal{B}\to\mathcal{C}(\mathbb{T})}\,.$$

Since $\mathcal{B} \subset \mathcal{C}(\mathbb{T})$ and using (9.8), we thus have $\|\mathfrak{T}_N\|_{\mathcal{B}\to\mathcal{C}(\mathbb{T})} \le \|\mathfrak{T}_N\|_{\mathcal{C}(\mathbb{T})\to\mathcal{C}(\mathbb{T})} \le 2C_2$. Since $\varphi \in \mathcal{B}$ also $\widetilde{\varphi} \in \mathcal{B}$ and it holds

$$\begin{aligned}\|\widetilde{\varphi}\|_\infty &\le \|\widetilde{\varphi}-\mathfrak{T}_N\varphi\|_\infty + \|\mathfrak{T}_N\|_{\mathcal{B}\to\mathcal{C}(\mathbb{T})}\|\varphi\|_\infty\\ &\le \|\widetilde{\varphi}-\mathfrak{T}_N\varphi\|_\infty + 2\,C_2\|\varphi\|_\infty\,.\end{aligned}$$

Because of (9.10) we get for $N \to \infty$ that

$$\|\widetilde{\varphi}\|_\infty \le 2\,C_2\|\varphi\|_\infty\,. \tag{9.11}$$

Finally, we consider for an arbitrary $N \in \mathbb{N}$ a special function $\varphi_N \in \mathcal{B}$ and its conjugate function $\widetilde{\varphi}_N$ given by

$$\varphi_N(\tau) = \tfrac{1}{\pi}\sum_{k=1}^{N}\frac{\sin(k\tau)}{k} \quad \text{and} \quad \widetilde{\varphi}_N(\tau) = -\tfrac{1}{\pi}\sum_{k=1}^{N}\frac{\cos(k\tau)}{k}$$

respectively. It is well known [92, Chapter II.9] that $\|\varphi_N\|_\infty \le 1$ and for the modulus of $\widetilde{\varphi}_N$ holds at $\tau = 0$ the inequality $|\widetilde{\varphi}_N(0)| = \frac{1}{\pi}\sum_{k=1}^{N}\frac{1}{k} \ge \frac{1}{\pi}\log N$. This gives together with (9.11) the inequality

$$\tfrac{1}{\pi}\log N \le \|\widetilde{\varphi}_N\|_\infty \le 2\,C_2\,\|\varphi_N\|_\infty \le 2\,C_2\,.$$

But this is a contradiction for sufficiently large $N > \exp(2\pi C_2)$. □

Proof (Corollary 9.7). It was already shown in the proof of Theorem 9.6 that $\|\mathfrak{T}_N\|_1 \le 2\,\|\mathfrak{T}_N\|_{\mathcal{B}\to\mathcal{C}(\mathbb{T})}$. Assume now (in contradiction to the statement of the corollary) that there exists a constant C_2 such that

$$\lim_{N\to\infty} \|\mathfrak{T}_N\|_{\mathcal{B}\to\mathcal{C}(\mathbb{T})} \le C_2\,. \tag{9.12}$$

Then it follows that $\liminf_{N\to\infty}\|\mathfrak{T}_N\|_1 \le 2\,C_2$. But this means that there exists a sequence N_k such that $\|\mathfrak{T}_{N_k}\|_1 \le 2\,C_2$ for all $k \in \mathbb{N}$. Now, we can proceed with this sequence $\{\mathfrak{T}_{N_k}\}$ as in the proof of Theorem 9.6. Again, this results in a contradiction which shows that assumption (9.12) was wrong. □

Proof (Corollary 9.8). This corollary was proved implicitly in the proof of Theorem 9.6. Because if we assume that the statement of the corollary is wrong, equation (9.10) holds. Following the above proof, starting with (9.10), shows that this assumption yields a contradiction for all $\{\mathfrak{T}_N\}$ with property I. This shows that the assumption was wrong and proves the statement of the corollary. □

Smooth functions

Assume that ω is a regular majorant and let $f \in \mathcal{C}_\omega(\mathbb{T})$. Since the Hilbert transform is a continuous operator on $\mathcal{C}_\omega(\mathbb{T})$ (cf. Section 6) it is clear that $\mathcal{C}_\omega(\mathbb{T})$ as well as $\mathcal{C}_{\omega,0}(\mathbb{T})$ are subsets of $\mathcal{B}$. For the later, the following positive result is obtained.

Theorem 9.10. *Let ω be a regular majorant and let $\{\widetilde{\mathfrak{L}}_N\}$ be the sequence of operators defined in Example 9.5. Then there exists a constant C such that*

$$\left\|\widetilde{\mathfrak{L}}_N f\right\|_{\mathcal{C}_\omega(\mathbb{T})} \leq C \, \|f\|_{\mathcal{C}_\omega(\mathbb{T})} \tag{9.13}$$

and for all $f \in \mathcal{C}_{\omega,0}(\mathbb{T})$, and it holds

$$\lim_{N\to\infty} \left\|\widetilde{f} - \widetilde{\mathfrak{L}}_N f\right\|_{\mathcal{C}_\omega(\mathbb{T})} = 0 \tag{9.14}$$

for all $f \in \mathcal{C}_{\omega,0}(\mathbb{T})$.

This result shows that for functions f which are sufficiently smooth, the corresponding conjugate function $\widetilde{f}$ can always be approximated arbitrarily well from the values $f(\tau_k)$ at the N sampling points $\{\tau_k\}$ by means of the linear operator $\widetilde{\mathfrak{L}}_N$ which was given in Example 9.5. The theorem is proved using an auxiliary lemma which characterizes the behavior of the linear interpolation operator $\mathfrak{L}_N$ defined in Example 9.5, Equation (9.1).

Lemma 9.11. *Let ω be an arbitrary majorant and let $\mathfrak{L}_N$ be the interpolation operator defined by* (9.1). *Then there exists a constant C such that*

$$\|\mathfrak{L}_N f\|_{\mathcal{C}_\omega(\mathbb{T})} \leq C \, \|f\|_{\mathcal{C}_\omega(\mathbb{T})} \tag{9.15}$$

for all $f \in \mathcal{C}_\omega(\mathbb{T})$. Moreover, for all $f \in \mathcal{C}_{\omega,0}(\mathbb{T})$ holds

$$\lim_{N\to\infty} \|f - \mathfrak{L}_N f\|_{\mathcal{C}_\omega(\mathbb{T})} = 0 \; . \tag{9.16}$$

Proof (Lemma 9.11). Let $\delta > 0$ arbitrary and consider two points $t_1, t_2 \in \mathbb{T}$ with $t_1 < t_2$ and with $|t_1 - t_2| = \delta$. The function $(\mathfrak{L}_N f)(t)$ is piecewise linear on certain disjoint intervals $T_k \subset \mathbb{T}$. Two cases have to be distinguished: 1) t_1 and t_2 belong to the same interval T_k or 2) t_1 and t_2 belong to different intervals.

1) Assume that $t_1, t_2 \in T = [t_0^l, t_0^u]$ belong to the same interval T. Since $\mathfrak{L}_N f$ is linear over T, it follows at once that

$$\begin{aligned} |(\mathfrak{L}_N f)(t_1) - (\mathfrak{L}_N f)(t_2)| &= \frac{\left|f(t_0^l) - f(t_0^u)\right|}{\left|t_0^l - t_0^u\right|} \, |t_1 - t_2| \leq \left|f(t_0^l) - f(t_0^u)\right| \\ &\leq \|f\|_{\mathcal{C}_\omega(\mathbb{T})} \; \omega\left(\left|t_0^l - t_0^u\right|\right) \end{aligned} \tag{9.17}$$

where for the second line it was used that $f \in \mathcal{C}_\omega(\mathbb{T})$. Because $t_1, t_2 \in T$ there exists a constant $K \geq 1$ such that $\left|t_0^l - t_0^u\right| = K\,|t_1 - t_2|$, and since ω is a majorant, the function $\omega(x)/x$ is not increasing, from which follows that $\omega(K\,x) \leq K\,\omega(x)$ provided that $K \geq 1$. Allowing for these relations, it follows from (9.17)

$$|(\mathfrak{L}_N f)(t_1) - (\mathfrak{L}_N f)(t_2)| \leq K \, \|f\|_{\mathcal{C}_\omega(\mathbb{T})} \; \omega\left(|t_1 - t_2|\right) \leq K \, \|f\|_{\mathcal{C}_\omega(\mathbb{T})} \; \omega\left(\delta\right) \, . \tag{9.18}$$

2) Next it is assumed that $t_1 \in T_1 = [t_1^l, t_1^u]$ and $t_2 \in T_2 = [t_2^l, t_2^u]$ belong to disjoint intervals. Then it follows

$$\begin{aligned} |(\mathfrak{L}_N f)(t_1) - (\mathfrak{L}_N f)(t_2)| &\leq |(\mathfrak{L}_N f)(t_1) - (\mathfrak{L}_N f)(t_1^u)| + \\ &\left|(\mathfrak{L}_N f)(t_1^u) - (\mathfrak{L}_N f)(t_2^l)\right| + \left|(\mathfrak{L}_N f)(t_2^l) - (\mathfrak{L}_N f)(t_2)\right| \, . \end{aligned} \tag{9.19}$$

Next, we consider the three terms on the right hand side. For the first and third term, we can proceed as under point 1) since $t_1, t_1^u \in T_1$ and $t_2, t_2^l \in T_2$. This will give an inequality as in (9.18) for both terms but with different constants K_1 and K_3, respectively. For the second term on the right hand side of (9.19), it is used that the operator $\mathfrak{L}_n$ is concentrated on the boundary points of the intervals T_k. It follows

$$\begin{aligned} \left|(\mathfrak{L}_N f)(t_1^u) - (\mathfrak{L}_N f)(t_2^l)\right| &= \left|f(t_1^u) - f(t_2^l)\right| \\ &\leq \|f\|_{\mathcal{C}_\omega(\mathbb{T})} \, \omega\left(\left|t_1^u - t_2^l\right|\right) \leq \|f\|_{\mathcal{C}_\omega(\mathbb{T})} \, \omega\left(\delta\right) \, . \end{aligned}$$

Altogether, (9.19) becomes $|(\mathfrak{L}_N f)(t_1) - (\mathfrak{L}_N f)(t_2)| \leq (K_1 + 1 + K_3) \, \|f\|_{\mathcal{C}_\omega(\mathbb{T})} \, \omega\left(\delta\right)$. This together with (9.18) shows that there exists a constant C such that (9.15) holds for all $f \in \mathcal{C}_\omega(\mathbb{T})$.

It remains to show (9.16). To this end, let $f \in \mathcal{C}_{\omega,0}(\mathbb{T})$ fixed and $\epsilon > 0$ arbitrary. Then there exists a polynomial p such that $\|f - p\|_{\mathcal{C}_{\omega,0}(\mathbb{T})} \leq \epsilon$. The triangular inequality gives now

$$\|f - \mathfrak{L}_N f\|_{\mathcal{C}_\omega(\mathbb{T})} \leq \|f - p\|_{\mathcal{C}_{\omega,0}(\mathbb{T})} + \|p - \mathfrak{L}_N p\|_{\mathcal{C}_{\omega,0}(\mathbb{T})} + \|\mathfrak{L}_N (p - f)\|_{\mathcal{C}_{\omega,0}(\mathbb{T})} \, . \tag{9.20}$$

By the choice of p, the first term on the right hand side is smaller than or equal to ϵ, and the third term is smaller than or equal to $C \, \|f\|_{\mathcal{C}_\omega(\mathbb{T})} \, \epsilon$, using (9.15). Furthermore, there exists an $N_0 \in \mathbb{N}$ such that the second term on the right hand side of (9.20) becomes smaller than ϵ for all $N > N_0$. Altogether, (9.20) becomes $\|f - L_N f\|_{\mathcal{C}_\omega(\mathbb{T})} \leq (2 + C)\epsilon$ for all $N > N_0$ which proves (9.16). □

Proof (Theorem 9.10). Theorem 5.11 shows that there exists a constant c_1 such that $\|\widetilde{f}\|_{\mathcal{C}_\omega(\mathbb{T})} \leq c_1 \|f\|_{\mathcal{C}_\omega(\mathbb{T})}$ for all $f \in \mathcal{C}_\omega(\mathbb{T})$ provided that ω is a regular majorant. Therewith and together with Lemma 9.11 follows

$$\left\|\widetilde{\mathfrak{L}}_N f\right\|_{\mathcal{C}_\omega(\mathbb{T})} = \left\|\widetilde{\mathfrak{L}_N f}\right\|_{\mathcal{C}_\omega(\mathbb{T})} \leq c_1 \left\|\mathfrak{L}_N f\right\|_{\mathcal{C}_\omega(\mathbb{T})} \leq c_1\, c_2 \left\|f\right\|_{\mathcal{C}_\omega(\mathbb{T})}$$

which proves (9.13). Similarly, the second statement of the theorem follows from

$$\left\|\widetilde{f} - \widetilde{\mathfrak{L}}_N f\right\|_{\mathcal{C}_\omega(\mathbb{T})} = \left\|\widetilde{f} - \widetilde{\mathfrak{L}_N f}\right\|_{\mathcal{C}_\omega(\mathbb{T})} \leq c_1 \left\|f - \mathfrak{L}_N f\right\|_{\mathcal{C}_\omega(\mathbb{T})} .$$

The right hand side of this inequality goes to zero as $N \to \infty$ by Lemma 9.11 which proves (9.14). □

Notes

The results in this section was discussed in [12] by the authors.

10

Spectral Factorization

Spectral factorization is an important tool in the theory of stochastic processes, in information theory, signal processing, control theory and many other fields. This operation factorizes a given function into a causal and an anticausal part. In signal processing it is necessary for example, for the determination of the causal Wiener filter. Despite a clear and simple derivation of the spectral factorization operator, it shows a quite complicated analytic behavior. The main reason behind this complicated behavior is the non-linearity of the spectral factorization operator. It implies in particular that the boundedness of the spectral factorization operator does not imply its continuity and vice versa. In this chapter, we will investigate the relation between boundedness and continuity of the spectral factorization mapping in detail. It turns out that continuity and boundedness are alternative properties, i.e. the spectral factorization mapping is either bounded or continuous, but never both, at least on the function spaces considered in this chapter.

10.1 Regularity of Stochastic Sequences

As the name indicates, the spectral factorization is an operation usually applied to spectral densities of stochastic processes. This subsection shortly reviews some results from the theory of stationary stochastic sequences, as far as they will be needed. Thereby, we are especially interested in the relation between the so called regularity of the stochastic sequence and the smoothness of its spectral density which will get important in Subsection 10.5 and 10.6.

Let $(\Omega, \mathcal{F}, \nu)$ be a probability space, i.e. Ω is the sample space of elementary events, $\mathcal{F}$ is a σ-algebra of subsets of Ω, and ν is a probability measure on $\mathcal{F}$. A *random variable* is an $\mathcal{F}$-measurable (real or complex) function on Ω. The *expectation* (or the *mean*) and the *variance* of a random variable x are defined by

$$\mathcal{E}[x] := \int_\Omega x(\omega)\,\mathrm{d}\nu(\omega) \qquad \text{and} \qquad \mathcal{V}[x] := \mathcal{E}\left[\,(x - \mathcal{E}[x])^2\right]$$

V. Pohl, H. Boche, *Advanced Topics in System and Signal Theory*, Foundations in Signal Processing, Communications and Networking 4,
DOI 10.1007/978-3-642-03639-2_10,

respectively, and the *covariance* of a pair x, y of random variables is

$$\operatorname{cov}[x,y] := \mathcal{E}\left[\,(x-\mathcal{E}[x])\,(y-\mathcal{E}[y])\right] \, .$$

Two random variables x, y are said to be *uncorrelated* if $\operatorname{cov}[x,y] = 0$. If the variance of both random variables is nonzero, one may normalize the covariance of x and y to obtain the so called *correlation coefficient*

$$\operatorname{cor}[x,y] := \frac{\operatorname{cov}[x,y]}{\sqrt{\mathcal{V}[x]}\,\sqrt{\mathcal{V}[y]}}$$

of x and y.

If not mentioned otherwise, we will always consider complex random variables x with zero mean $\mathcal{E}[x] = 0$ and with finite second moments, i.e. with

$$\mathcal{E}[|x|^2] = \int_\Omega |x(\omega)|^2 \,\mathrm{d}\nu(\omega) < \infty \, .$$

The set of all such random variables is denoted by $\mathcal{L}^2 = \mathcal{L}^2(\Omega, \mathcal{F}, \nu)$. If one defines for arbitrary elements $x, y \in \mathcal{L}^2$ an inner product on $\mathcal{L}^2$ by

$$\langle x, y\rangle_{\mathcal{L}^2} := \int_\Omega x(\omega)\,\overline{y(\omega)}\,\mathrm{d}\nu(\omega) = \mathcal{E}[x\,\overline{y}]$$

then $\mathcal{L}^2$ becomes a Hilbert space. Note that since the mean of every random variable in $\mathcal{L}^2$ is zero, one has that $\mathcal{V}[x] = \|x\|^2_{\mathcal{L}^2} = \langle x, x\rangle_{\mathcal{L}^2}$ and one obtains particular simple realtions for the covariance and for the correlation coefficient of two random variables $x, y \in \mathcal{L}^2$:

$$\operatorname{cov}[x,y] = \langle x, y\rangle_{\mathcal{L}^2} \quad \text{and} \quad \operatorname{cor}[x,y] = \frac{\langle x, y\rangle_{\mathcal{L}^2}}{\|x\|_{\mathcal{L}^2}\,\|y\|_{\mathcal{L}^2}} \, .$$

In particular, two random variables $x, y \in \mathcal{L}^2$ are uncorrelated if and only if they are orthogonal.

A set $X = \{x(k)\}_{k\in\mathbb{Z}}$ of random variables in $\mathcal{L}^2$ is called a *stochastic process with discrete time* or a *stochastic sequence*. Such a stochastic sequence X is called *stationary* (in the wide sense) if the covariance $\operatorname{cov}[x(m), x(n)]$ of two of its elements depends only on the difference $m - n$ but not on the absolute position n in the sequence, i.e. if $\operatorname{cov}[x(n+k), x(n)] = \operatorname{cov}[x(k), x(0)]$ for all $k \in \mathbb{Z}$ and every arbitrary $n \in \mathbb{Z}$. Thus, the *covariance function* of a stationary stochastic sequence X is given by $r_X(m-n) = \mathcal{E}[x(m)\,\overline{x(n)}]$ for all $m, n \in \mathbb{Z}$. The following classical result due to Herglotz gives the so called *spectral representation* of the covariance function of stationary random sequences.

Theorem 10.1 (Herglotz). *Let X be a stationary (wide sense) random sequence with zero mean, and let r_X be its covariance function. Then there exists a unique positive probability measure μ_X on $\mathbb{T}$ such that for every $k \in \mathbb{Z}$*

$$r_X(k) = \int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}\theta k}\,\mathrm{d}\mu_X(\mathrm{e}^{\mathrm{i}\theta}) \, . \tag{10.1}$$

The measure μ_X corresponding to the stochastic sequence X is known as the *spectral measure* of X. It is assumed subsequently that the spectral measure is absolute continuous (with respect to Lebesgue measure)[1]. Then by the Radon-Nikodym theorem there exists a function $\phi_X \in L^1(\mathbb{T})$ such that (10.1) can be written as

$$r_X(k) = \int_{-\pi}^{\pi} \phi_X(\mathrm{e}^{\mathrm{i}\theta})\, \mathrm{e}^{-\mathrm{i}\theta k}\, \mathrm{d}\theta\ . \tag{10.2}$$

The function ϕ_X in this representation of the covariance function is the *Radon-Nikodym derivative* of μ_X with respect to the Lebesgue measure on $\mathbb{T}$, and it is called the *spectral density* of the random sequence X.

Let $X = \{x(k)\}_{k\in\mathbb{Z}}$ be a stochastic sequence in $\mathcal{L}^2$, i.e. every element $x(k)$ is an element of the Hilbert space $\mathcal{L}^2$. Then X spans the subspace

$$\mathcal{L}^2(X) := \mathrm{span}\{x(k) : k \in \mathbb{Z}\}$$

of $\mathcal{L}^2$ wherein span denotes the closed linear span. Moreover, every collection $\{x(k)\}_{k\in I}$ from a certain index set I spans a subspace of $\mathcal{L}^2(X)$. A large part of the analysis of stationary stochastic sequences is based on the fact that $\mathcal{L}^2(X)$ can be identified with the Hilbert space $L^2(\mu_X)$ of complex functions on the unit circle with the inner product

$$\langle f, g\rangle_{L^2(\mu_X)} := \int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\theta})\, \overline{g(\mathrm{e}^{\mathrm{i}\theta})}\, \mathrm{d}\mu_X(\mathrm{e}^{\mathrm{i}\theta})\ . \tag{10.3}$$

Lemma 10.2. *Let $X = \{x(k)\}_{k\in\mathbb{Z}}$ be a stationary (in the wide sense) stochastic sequence with spectral measure μ_X. Then $\mathcal{L}^2(X)$ is isometricly isomorphic to $L^2(\mu_X)$.*

Proof. We construct a Hilbert space isomorphism $\Lambda : \mathcal{L}^2(X) \to L^2(\mu_X)$ by setting

$$\Lambda\, x(k) = e_k\ , \qquad k \in \mathbb{Z}$$

where $e_k(\mathrm{e}^{\mathrm{i}\theta}) = \mathrm{e}^{\mathrm{i}k\theta}$ with $\theta \in [-\pi, \pi)$. The mapping Λ has to be linear. Therefore for finite linear combinations of the basis vectors, one has

$$\Lambda\,[\textstyle\sum_k \alpha_k\, x(k)\,] = \sum_k \alpha_k\, e_k\ .$$

It is clear that Λ is one-to-one in the sense that $\sum_k \alpha_k\, e_k = 0$ almost everywhere with respect to μ_X if and only if $\sum_k \alpha_k\, x(k) = 0$ almost surely. Moreover, Λ preserves the inner product, since by (10.3) and (10.1) one has

$$\begin{aligned}\langle\, \Lambda\, x(k), \Lambda\, x(m)\,\rangle_{L^2(\mu_X)} = \langle e_k, e_m\rangle_{L^2(\mu_X)} &= \int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}(k-m)\theta}\, \mathrm{d}\mu_X(\mathrm{e}^{\mathrm{i}\theta}) \\ &= r_X(k-m) = \langle\, x(k), x(m)\,\rangle_{\mathcal{L}^2}\end{aligned}$$

[1] The singular part of the spectral measure μ_X can be associated with the so called *deterministic* part of the random sequence X (see eg. [81, §VI.5]). This part is of no importance in the following considerations and therefore omitted.

and by the linearity of the inner product, this holds also for finite linear combinations of the basis vectors.

Up to now we defined Λ only for finite linear combinations of the basis vectors $\{x(k)\}_{k\in\mathbb{Z}}$. However, since the closure of $\{x(k)\}_{k\in\mathbb{Z}}$ in $\mathcal{L}^2$ and the closure of $\{e_k\}_{k\in\mathbb{Z}}$ in $L^2(\mu_X)$ are dense in $\mathcal{L}^2(X)$ and $L^2(\mu_X)$, respectively, and because Λ is continuous, it can be extended to a Hilbert space isomorphism on the whole $\mathcal{L}^2(X)$. In fact, let $\eta \in \mathcal{L}^2(X)$ be arbitrary. Then there exists a sequence $\{\eta_N\}_{N\in\mathbb{N}}$ in $\mathcal{L}^2(X)$ of the form $\eta_N = \sum_{k=-N}^{N} \alpha_k\, x(k)$ such that $\|\eta_N - \eta\|_{\mathcal{L}^2} \to 0$ as $N \to \infty$. In particular $\{\eta_N\}_{N\in\mathbb{N}}$ is a Cauchy sequence. Set $f_N := \Lambda\,\eta_N$ for all $N \in \mathbb{N}$. Then $\|f_N - f_M\|_{L^2(\mu_X)} = \|\Lambda(\eta_N - \eta_M)\|_{L^2(\mu_X)} = \|\eta_N - \eta_M\|_{\mathcal{L}^2}$ since Λ is an isometry. Thus $\{f_N\}_{N\in\mathbb{N}}$ is a Cauchy sequence in $L^2(\mu_X)$ which converges to a unique (up to equivalences almost everywhere) $f \in L^2(\mu_X)$. Setting $f := \Lambda\,\eta$ gives the desired extension of Λ to $\mathcal{L}^2(X)$. □

By the previous Lemma, it is clear that not only $\mathcal{L}^2(X)$, spanned by the whole stochastic sequence X, can be identified with $L^2(\mu_X)$ but also every subspace $\operatorname{span}\{x(k) : k \in I\}$ of $\mathcal{L}^2(X)$ spanned by a certain collection of elements from the sequence X can be identified with the subspace

$$\{f \in L^2(\mu_X) : \hat{f}(k) = 0 \text{ for all } k \notin I\}$$

of $L^2(\mu_X)$, wherein $\hat{f}(k) = \int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\theta})\,\mathrm{e}^{-\mathrm{i}k\theta}\,\mathrm{d}\mu_X(\mathrm{e}^{\mathrm{i}\theta})$ denotes the k-th Fourier coefficients of f in $L^2(\mu_X)$. In particular, the future of the sequence X, i.e. the subspace $\operatorname{span}\{x(k) : k \geq 0\}$, can be identified with the Hardy space $H^2(\mu_X)$. These identifications of the different subspaces allows one to use powerful methods from complex analysis to study several properties of stochastic sequences. One example where this was successfully applied is the following characterization of completely regular stochastic sequences.

Let X be a stationary stochastic sequence, let $m \in \mathbb{N}$ be an arbitrary positive integer, and consider the two subspaces of $\mathcal{L}^2(X)$ defined by

$$\mathcal{P}_0 := \operatorname{span}\{x(k) : k < 0\} \qquad \text{and} \qquad \mathcal{F}_m := \operatorname{span}\{x(k) : k \geq m\}\,. \tag{10.4}$$

The first subspace $\mathcal{P}_0$ is usually called the "*past*" of the sequence X whereas $\mathcal{F}_m$ is said to be the "*future*" of X. How does the future of the stochastic sequence X depends on its past? This question is of considerable interest in several applications. For example, in detection and estimation or in financial mathematics it is desired to predict future values of X from already observed values of X. If there is a strong correlation between the past and the future, the error of such a prediction can be made smaller (by an appropriated prediction method) than if their correlation is low. If, on the other hand, the future is almost independent of the past, it will not be possible to make any useful prediction of future values of X. So it seems natural to measure the dependency between the two subspaces $\mathcal{P}_0$ and $\mathcal{F}_m$ by the maximal correlation between pairs of vectors from each subspace, i.e. by the expression

$$\rho_m(X) := \sup_{\xi \in \mathcal{P}_0\,,\eta \in \mathcal{F}_m} \operatorname{cor}[\xi, \eta] = \sup_{\xi \in \mathcal{P}_0\,,\eta \in \mathcal{F}_m} \frac{|\langle \xi, \eta \rangle_{\mathcal{L}^2}|}{\|\xi\|_{\mathcal{L}^2}\,\|\eta\|_{\mathcal{L}^2}} \,. \tag{10.5}$$

This maximal correlation coefficient between the two subspaces $\mathcal{P}_0$ and $\mathcal{F}_m$ is also known as *regularity coefficient* of X or as the *angle between* $\mathcal{P}_0$ *and* $\mathcal{F}_m$. The last name is obviously motivated by the right hand side expression of the last equation[2]. Now let $\mathfrak{P}_0$ and $\mathfrak{F}_m$ be the orthogonal projection from $\mathcal{L}^2(X)$ onto $\mathcal{P}_0$ and $\mathcal{F}_m$, respectively. Therewith the angle $\rho_m(X)$ between the past and future of the stochastic sequence X can be expressed as

$$\begin{aligned} \rho_m(X) &= \sup_{\substack{x,y \in \mathcal{L}^2(X) \\ \|x\| \le 1, \|y\| \le 1}} \left| \langle \mathfrak{P}_0\, x, \mathfrak{F}_m\, y \rangle_{\mathcal{L}^2(X)} \right| = \sup_{x \in \mathcal{L}^2(X)\,, \|x\| \le 1} \left| \langle x, \mathfrak{P}_0\, \mathfrak{F}_m\, x \rangle_{\mathcal{L}^2(X)} \right| \\ &= \|\mathfrak{P}_0\, \mathfrak{F}_m\|_{\mathcal{L}^2(X) \to \mathcal{L}^2(X)} \end{aligned}$$

where it was only used that every projection is self adjoint. Thus the angle between past and future of the stochastic sequence X is given by the operator norm of $\mathfrak{P}_0\, \mathfrak{F}_m$. It is clear from the definition that $\rho_m(X) = 0$ if both subspaces are disjoint and that for a fixed $\mathcal{P}_0$ the angle $\rho_m(X)$ is maximized if $\mathcal{F}_m = \mathcal{P}_0$. Based on the angle between past and future, one classifies stochastic sequences as follows.

Definition 10.3 (Completely Regular Stochastic Sequences). *Let X be a stationary (in the wide sense) stochastic sequence and let $\rho_m(X)$ be its regularity coefficient* (10.5)*. Then X is called* completely regular *(or* strong mixing*) if* $\lim_{m\to\infty} \rho_m(X) = 0$*, and it is called* completely regular of order α (fast mixing) *if there exists a constant C such that*

$$\rho_m(X) \le C\, m^{-\alpha} \;, \qquad \alpha > 0 \,.$$

Thus the order α of a completely regular sequence characterizes how fast the correlation between past and future $\rho_m(\mu_X)$ approaches zero as $m \to \infty$.

It is natural to ask whether it is possible to characterize the completely regular sequences in terms of their spectral densities. Such a characterization was obtained by Ibragimov [49] for fast mixing sequences and by Helson and Sarason [46, 77] for strong mixing sequences. The results are given (without proof) in the following theorem.

Theorem 10.4. *Let X be a stationary stochastic sequence with density ϕ_X.*

(a) X is completely regular if and only ϕ_X admits a representation of the form

$$\phi_X = |p|^2\, \mathrm{e}^f \tag{10.6}$$

with a polynomial p which has zeros only on $\mathbb{T}$ and with $f \in VMO$.

[2] Of course, strictly speaking it is the cosine of the angle between $\mathcal{P}_0$ and $\mathcal{F}_m$.

(*b*) *X is completely regular of order α if and only if ϕ_X admits a representation of the form* (10.6) *with a polynomial p with zeros on $\mathbb{T}$ and with a real function $f \in \Lambda_\alpha$.*

Therein, *VMO* is the space of *vanishing mean oscillation.* Since part (*a*) of the above theorem will not be needed subsequently, we only refer to corresponding textbooks (e.g. [41]) for a definition of this space. The Hölder-Zygmund class Λ_α, appearing in part (*b*) of the theorem was defined at the end of Section 1.3.

In some cases, the spectral density ϕ_X is assumed to have no zeros on the unit circle. For example, assuming that ϕ_X has no zeros on $\mathbb{T}$ is sufficient (not necessary) to guaranty the existence of the spectral factorization of ϕ_X (cf. also the discussion in Section 10.2 below). Then the polynomial p in (10.6) is a constant function and part (*b*) (and likewise part (*a*)) of Theorem 10.4 states that X is completely regular of order α if and only if $\phi_X \in \Lambda_\alpha$.

10.2 Definition and Basic Properties

In this subsection, we collect the main definitions and properties of the spectral factorization mapping. Let $\mathcal{A} \subset L^1$ be a Banach algebra of integrable functions on the unit circle $\mathbb{T}$ with unity e. Then for every $\phi \in \mathcal{A}$ the Fourier coefficients $\{\hat{\phi}(k)\}_{k=-\infty}^{\infty}$ exist and we define the causal and the anti-causal subspaces $\mathcal{A}_+$ and $\mathcal{A}_-$ of $\mathcal{A}$, respectively, by

$$\begin{aligned} \mathcal{A}_+ &:= \{\phi \in \mathcal{A} : \hat{\phi}(k) = 0 \text{ for all } k < 0\} \\ \mathcal{A}_- &:= \{\phi \in \mathcal{A} : \hat{\phi}(k) = 0 \text{ for all } k > 0\} \,. \end{aligned}$$

Note[3] that the intersection $\mathcal{A}_+ \cap \mathcal{A}_-$ contains all constant functions of $\mathcal{A}$. In the particular case of $\mathcal{A} = L^p$, $1 \le p \le \infty$ the causal subspace L^p_+ coincides with the *Hardy space* H^p and if $\mathcal{A} = \mathcal{C}(\mathbb{T})$ then $\mathcal{C}(\mathbb{T})_+ = A(\mathbb{D})$ is the disk algebra. Note that every function $\phi_+ \in \mathcal{A}_+$ and $\phi_- \in \mathcal{A}_-$ can be identified with the function

$$\phi_+(z) = \textstyle\sum_{k=0}^{\infty} \hat{\phi}(k)\, z^k \qquad \text{and} \qquad \phi_-(z) = \textstyle\sum_{k=0}^{\infty} \hat{\phi}(-k)\, z^{-k}$$

which is analytic for all $z \in \mathbb{D}$ and all $z \in \mathbb{C}\backslash\overline{\mathbb{D}}$, respectively. Moreover, the natural projection $\mathcal{A} \to \mathcal{A}_+$ is the Riesz projection $\mathfrak{P}_+$ which was studied in detail in Section 6.

Definition 10.5 (Spectral factorization). *Let $\mathcal{A} \subset L^1$ be a Banach algebra of functions on $\mathbb{T}$, and let $\phi \in \mathcal{A}$ be a real valued function. Then ϕ is said to possess a* spectral factorization *if theres exists a $\phi_+ \in \mathcal{A}_+$ with $\phi_+(z) \neq 0$ for all $z \in \mathbb{D}$ and a $\phi_- \in \mathcal{A}_-$ with $\phi_-(z) \neq 0$ for all $z \in \mathbb{C}\backslash\overline{\mathbb{D}}$ such that*

[3] Note also that this definition of the subspaces $\mathcal{A}_+$ and $\mathcal{A}_-$ differs slightly from the definition given at the beginning of Section 6.

$$\phi(\zeta) = \phi_+(\zeta)\,\phi_-(\zeta) \qquad \textit{for all} \qquad \zeta \in \mathbb{T}\,.$$

Every $\phi \in \mathcal{A}$ which possesses a spectral factorization is said to be a spectral density *and the functions ϕ_+ and ϕ_- are called the* spectral factors *of ϕ. The mapping $\mathfrak{S} : \phi \mapsto \phi_+$ is the* spectral factorization mapping.

Remark 10.6. It is clear that the above definition specifies the spectral factors ϕ_+ and ϕ_- (if they exist) only up to a unitary constant because ϕ_+ and ϕ_- may be multiplied by certain factors $\mathrm{e}^{\mathrm{i}\theta_0}$ and $\mathrm{e}^{-\mathrm{i}\theta_0}$, respectively. To make the factorization unique, one often considers the *canonical spectral factorization* which requires additionally that $\phi_+(0)$ is real and positive.

Which elements of a Banach algebra $\mathcal{A}$ posses a spectral factorization and how can we obtain the spectral factors for a given spectral density? To get a first idea, we consider the special case that the given function ϕ is a trigonometric polynomial of a certain degree N. Then the spectral factorization is equivalent to the following celebrated result of L. Fejér and M. Riesz.

Theorem 10.7 (Theorem of Fejér-Riesz). *Let $\phi \in \mathcal{P}_{pos}(N)$ be a nonnegative trigonometric polynomial. Then there exists a unique analytic polynomial $\phi_+ \in \mathcal{P}_+(N)$ such that*

$$\phi(\zeta) = |\phi_+(\zeta)|^2 \qquad \textit{for all} \qquad \zeta \in \mathbb{T}$$

and such that $\phi_+(z) \neq 0$ for all $z \in \mathbb{D}$.

Thus, every non-negative trigonometric polynomial $\phi \in \mathcal{P}_{\mathrm{pos}}(N)$ possesses a spectral factorization. One spectral factor is the polynomial

$$\phi_+(z) = \textstyle\sum_{k=0}^{N} a_k\, z^k\,, \qquad z \in \mathbb{D}$$

with certain complex coefficients $\{a_k\}_{k=0}^{N}$. The second spectral factor ϕ_- is obtained as the parahermitian conjugate $\phi_+^{\star}$ of ϕ_+:

$$\phi_-(z) = \phi_+^{\star}(z) := \overline{\phi_+(1/\overline{z})} = \textstyle\sum_{k=0}^{N} \overline{a_k}\, z^{-k}\,. \tag{10.7}$$

Also for a non-polynomial function ϕ, the spectral factorization can easily be derived formally: Let $\phi \in \mathcal{A}$ and assume that $\log\phi$ is again an element of the algebra $\mathcal{A}$. Then $\log\phi$ can be written as a Fourier series $\log\phi(\mathrm{e}^{\mathrm{i}\theta}) = \sum_{k=-\infty}^{\infty} \alpha(k)\,\mathrm{e}^{\mathrm{i}k\theta}$ with $\theta \in [-\pi,\pi)$ and with the Fourier coefficients

$$\alpha(k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \log\phi(\mathrm{e}^{\mathrm{i}\tau})\,\mathrm{e}^{-\mathrm{i}k\tau}\,\mathrm{d}\tau\,, \qquad k \in \mathbb{Z}\,.$$

Since $\log\phi$ is a real valued function, the Fourier coefficients satisfy the relation $\alpha(-k) = \overline{\alpha(k)}$ for all $k = 1,2,\dots$. Next we define the functions

$$g_+(z) = \tfrac{\alpha(0)}{2} + \textstyle\sum_{k=1}^{\infty} \alpha(k)\, z^k \quad \text{and} \quad g_-(z) = \tfrac{\alpha(0)}{2} + \textstyle\sum_{k=1}^{\infty} \alpha(-k)\, z^{-k}\,.$$

Assuming that g_+ and g_- are again elements of $\mathcal{A}$, it is clear that $g_+ \in \mathcal{A}_+$ and $g_- = g_+^\star \in \mathcal{A}_-$ and it holds $\log \phi = g_+ + g_-$. Taking the exponential function of both sides, gives finally $\phi = \exp(g_+) \exp(g_-)$, which shows that $\phi_+ = \exp(g_+)$ and $\phi_- = \exp(g_-)$ both of which are candidates for the spectral factors of ϕ. Inserting the Fourier coefficients α_k into the series of g_+ gives the following closed form formula for the spectral factorization mapping

$$\phi_+(z) = (\mathfrak{S}\,\phi)(z) = \exp\left(\frac{1}{4\pi}\int_{-\pi}^{\pi} \log\phi(\mathrm{e}^{\mathrm{i}\tau})\,\frac{\mathrm{e}^{\mathrm{i}\tau}+z}{\mathrm{e}^{\mathrm{i}\tau}-z}\,\mathrm{d}\tau\right)\,, \quad z \in \mathbb{D} \tag{10.8}$$

and ϕ_- is obtained from ϕ_+ by (10.7). In order that (10.8) is well defined, ϕ has to satisfy the so called *Paley-Wiener (or Szegö) condition*

$$\textstyle\int_{-\pi}^{\pi} \log\phi(\mathrm{e}^{\mathrm{i}\tau})\,\mathrm{d}\tau > -\infty\,. \tag{10.9}$$

However, this condition guarantees only that $(\mathfrak{S}\phi)(z)$ exists for every $z \in \mathbb{D}$. For an arbitrary $\phi \in \mathcal{A}$ which satisfies the Szegö condition, it is not necessarily true that $\mathfrak{S}\phi$ belongs again to the algebra $\mathcal{A}$. Nevertheless, the above formal derivation of the spectral factorization mapping shows immediately sufficient conditions for the existence of the spectral factorization mapping in arbitrary Banach algebras $\mathcal{A} \subset L^1$.

Lemma 10.8. *Let $\mathcal{A} \subset L^1$ be a Banach algebra on which the Riesz projection $\mathfrak{P}_+ : \mathcal{A} \to \mathcal{A}_+$ is bounded. Then every real valued function $\phi \in \exp(\mathcal{A})$ possesses a spectral factorization in $\mathcal{A}$, and one spectral factor is given by* (10.8)*.*

Proof. Assume $\phi \in \exp(\mathcal{A})$ is an arbitrary real valued function. Then there exists a $g \in \mathcal{A}$ such that $\phi = \exp(g)$ and the spectral factorization mapping (10.8) becomes

$$(\mathfrak{S}\,\phi)(z) = \exp\left(\frac{1}{4\pi}\int_{-\pi}^{\pi} g(\mathrm{e}^{\mathrm{i}\tau})\,\frac{\mathrm{e}^{\mathrm{i}\tau}+z}{\mathrm{e}^{\mathrm{i}\tau}-z}\,\mathrm{d}\tau\right) = \exp\left(\frac{1}{2}(\mathfrak{R}g)(z)\right)$$

with the the Herglotz-Riesz transform $\mathfrak{R}$, defined in (5.2), of g. Since $g \in \mathcal{A} \subset L^1$ it is well defined. Using relation (6.8) between the Herglotz-Riesz transform and the Riesz projection $\mathfrak{P}_+$, the spectral factorization mapping of ϕ can be written as

$$(\mathfrak{S}\,\phi)(z) = \exp\left(-\tfrac{1}{2}\hat{g}(0)\right)\,\exp\left[(\mathfrak{P}_+ g)(z)\right]\,, \quad z \in \mathbb{D}\,.$$

From which follows that $\|\mathfrak{S}\,\phi\|_{\mathcal{A}} \leq \exp\left(-\frac{1}{2}\|g\|_1\right)\,\exp\|\mathfrak{P}_+\,g\|_{\mathcal{A}} < \infty$ since $\mathfrak{P}_+$ is assumed to be bounded and using (3.14). Consequently $\mathfrak{S}\,\phi \in \mathcal{A}$ for every $\phi \in \exp(\mathcal{A})$. □

Definition 10.9 (Decomposing Banach algebra). *A Banach algebra $\mathcal{A} \subset L^1$ on which the Riesz projection $\mathfrak{P}_+ : \mathcal{A} \to \mathcal{A}_+$ is bounded is said to be* decomposing*.*

The notation of a decomposing Banach algebra is motivated by the statement of Theorem 6.7 which shows that $\mathcal{A}_+$ is a complemented subspace in every decomposing Banach algebra $\mathcal{A}$ and $\mathcal{A} = \mathcal{A}_+ \oplus \mathcal{A}_-$ where $\mathcal{A}_-$ is the complement of $\mathcal{A}_+$ in $\mathcal{A}$. In the context of spectral factorization, the decomposing property of an algebra $\mathcal{A}$ guarantees the existence of the spectral factor (10.8) for every $f \in \exp(\mathcal{A})$ by Lemma 10.8.

We saw in Section 6 that there exist important Banach algebras on which the Riesz projection is unbounded. For these algebras, the above lemma cannot be applied. However, for the important case of $\mathcal{A} = L^\infty$ (on which, according to Theorem 6.13, $\mathfrak{P}_+$ is unbounded) we have the following result

Lemma 10.10. *Every real valued $\phi \in L^\infty$ which satisfies the Paley-Wiener condition* (10.9) *posseses a spectral factorization in L^∞ and one spectral factor is given by* (10.8)*.*

Proof. Defining the spectral factor ϕ_+ by (10.8), it is clear that ϕ_+ is an outer function (cf. Definition 2.22) which is well defined since $\log \phi \in L^1$. Consequently ϕ_+ is analytic in $\mathbb{D}$, and by Theorem 2.25, $\phi_+ \in H^\infty$ and $|\phi_+(\mathrm{e}^{\mathrm{i}\theta})|^2 = \phi(\mathrm{e}^{\mathrm{i}\theta})$ for all $\theta \in [-\pi, \pi)$. Thus, ϕ_+ is indeed the spectral factor of ϕ and $\phi_-(z) = \phi_+^\star(z)$. □

10.3 Factorization on Algebras of Continuous Functions

In this section, we investigate the boundedness and continuity of the spectral factorization mapping. To obtain a quite general framework, we consider the spectral factorization on a class of Banach algebras $\mathcal{B} \subset L^1$ which are defined by the following four axioms.

Definition 10.11 ($\mathcal{S}$–algebra). *A commutative Banach algebra $\mathcal{B} \subset L^1$ is called an $\mathcal{S}$*–algebra *if*

(B1) $\mathcal{B}$ is a Banach algebra with respect to pointwise multiplication.
(B2) If $f \in \mathcal{B}$, then $\overline{f} \in \mathcal{B}$.
(B3) The set of all trigonometric polynomials is dense in $\mathcal{B}$.
(B4) Every multiplicative functional on $\mathcal{B}$ coincides with a functional $f \mapsto f(\zeta)$ defined as the value of f at some point $\zeta \in \mathbb{T}$.

Remark 10.12. The above axioms are similar to the set of axioms introduced by Peller and Khrushchev [66] in the context of best approximation of analytic functions[4]. Algebras satisfying the axioms of Peller and Khrushchev are sometimes called *decomposing Banach algebras* [27, 52] (which should not be confused with Def. 10.9), and Jacob and Partington [52] studied the spectral factorization mapping on such algebras, in detail. The $\mathcal{S}$-algebras as defined above, contain the class of decomposing Banach algebras, but an $\mathcal{S}$–algebras needs not be decomposing in the sense of Def. 10.9.

[4] There one requires additionally to $(B1)$–$(B4)$ that $f \in \mathcal{B}$ implies $\mathfrak{P}_+ f \in \mathcal{B}$.

Throughout this section, the symbol $\mathcal{B}$ will always stand for an $\mathcal{S}$–algebra. The four axioms $(B1)$–$(B4)$ of an $\mathcal{S}$–algebra imply certain properties on its elements. Some of these properties are given in the following propositions.

Proposition 10.13. *Let $\mathcal{B}$ be an $\mathcal{S}$–algebra. Every $\phi \in \mathcal{B}$ with $\phi(\zeta) \neq 0$ for all $\zeta \in \mathbb{T}$ belongs to $\mathcal{G}(\mathcal{B})$, i.e. is invertible in $\mathcal{B}$.*

Proof. Let h be an arbitrary multiplicative functional on $\mathcal{B}$. By axiom $(B4)$ there exists a $\zeta \in \mathbb{T}$ such that $h(\phi) = \phi(\zeta)$ for all $\phi \in \mathcal{B}$, and by $(B1)$

$$h(\phi\,\psi) = (\phi\,\psi)(\zeta) = \phi(\zeta)\,\psi(\zeta) = h(\phi)\,h(\psi) \qquad \text{for all } \phi, \psi \in \mathcal{B}\;.$$

This shows that h is even a homomorphism on $\mathcal{B}$ and $(B4)$ shows that all homomorphisms on $\mathcal{B}$ are obtained in this way. Therefore the proposition follows from Theorem 3.21. □

Proposition 10.14. *Every $\mathcal{S}$–algebra $\mathcal{B}$ is continuously embedded in $\mathcal{C}(\mathbb{T})$ with $\|\phi\|_\infty \leq \|\phi\|_{\mathcal{B}}$ for all $\phi \in \mathcal{B}$.*

Proof. Let ϕ be an arbitrary element of $\mathcal{B}$. By $(B1)$ and $(B4)$, to every $\zeta \in \mathbb{T}$ there exists an homomorphism $h \in \Gamma(\mathcal{B})$ such that $\phi(\zeta) = h(\phi)$. Since the norm of each complex homomorphism on $\mathcal{B}$ is upper bounded by 1 (see Part c of Theorem 3.15), one obtains

$$|\phi(\zeta)| = |h(\phi)| \leq \|h\|\,\|\phi\|_{\mathcal{B}} = \|\phi\|_{\mathcal{B}}\;.$$

This relation holds for arbitrary $\zeta \in \mathbb{T}$. Therefore it implies $\|\phi\|_\infty \leq \|\phi\|_{\mathcal{B}}$ which shows that the embedding is continuous. It remains to show that $\phi \in \mathcal{C}(\mathbb{T})$. By $(B3)$, to every $\varepsilon > 0$ there exists a trigonometric polynomial ψ_N such that

$$\|\phi - \psi_N\|_\infty \leq \|\phi - \psi_N\|_{\mathcal{B}} < \varepsilon$$

which shows that ϕ is continuous on $\mathbb{T}$. □

The next proposition shows that the operation of taking the parahermitian conjugate is a continuous operation on every $\mathcal{S}$–algebra.

Proposition 10.15. *Let $\mathcal{B}$ be an $\mathcal{S}$–algebra, then there exists a constant $m_1 > 0$ such that*

$$\|\phi^\star\|_{\mathcal{B}} \leq m_1\,\|\phi\|_{\mathcal{B}} \qquad \textit{for all } \phi \in \mathcal{B}\;.$$

Proof. The elements of an $\mathcal{S}$-algebra are functions defined on $\mathbb{T}$. Therefore, the parahermitian conjugate $f^\star$ corresponds to a pointwise conjugate complex of f: $f^\star(\zeta) = \overline{f(\zeta)}$ for every $\zeta \in \mathbb{T}$. It is easily verified that the operation $\phi \mapsto \overline{\phi}$ is an involution on $\mathcal{B}$ (cf. Def. 3.23). Moreover, by axiom $(B4)$, every multiplicative functional $h_\zeta \in \Gamma(\mathcal{B})$ is defined by a point $\zeta \in \mathbb{T}$ on the unit circle, according to $h_\zeta(\phi) := \phi(\zeta)$. The null space of such a homomorphisms is obviously given by $\ker h_\zeta = \{\phi \in \mathcal{B} : \phi(\zeta) = 0\}$. Therefore, the intersection of all null spaces contains only the zero function:

$$\bigcap_{h_\zeta \in \Gamma(\mathcal{B})} \ker h_\zeta = \bigcap_{\zeta \in \mathbb{T}} \{\phi \in \mathcal{B} : \phi(\zeta) = 0\} = \{\phi \equiv 0\} .$$

Thus, $\mathcal{B}$ is a semisimple commutative algebra and the continuity of the involution $\phi \mapsto \overline{\phi}$ follows form Theorem 3.26. □

Next we give some examples of $\mathcal{S}$–algebras. More examples may be found in [52] or [66].

Example 10.16 (Continuous functions on $\mathbb{T}$). The set of all continuous functions on $\mathbb{T}$ is obviously an $\mathcal{S}$–algebra. However, it is not decomposing since the Riesz projection is unbounded on $\mathcal{C}(\mathbb{T})$ (cf. Theorem 6.14).

Example 10.17 (Wiener algebra). The Wiener algebra (cf. Example 3.5) is an $\mathcal{S}$–algebra which is also decomposing.

Example 10.18 (Hölder continuous functions). For a certain $0 < \alpha < 1$ denote by Λ_α the set of all Hölder continuous functions (cf. Section 1.3). This is a non-separable space and therefore the axiom $(B3)$ of $\mathcal{S}$–algebras is not satisfied. However, if λ_α denotes the closure of all trigonometric polynomials under the norm $\|\cdot\|_\alpha$ of Λ_α then this space satisfies axiom $(B3)$ and one can verify that also the other axioms of an $\mathcal{S}$–algebra are satisfied. Moreover, it follows from Theorem 6.16 that the Riesz projection is bounded on Λ_α, such that λ_α is a decomposing $\mathcal{S}$ algebra.

10.3.1 Continuity of the spectral factorization mapping

In this section, we want to characterize $\mathcal{S}$–algebras on which the spectral factorization mapping $\mathfrak{S}$ is continuous. Since the trigonometric polynomials are dense in every $\mathcal{S}$–algebra $\mathcal{B}$, we know from the theorem of Fejér-Riesz that the spectral factorization exists at least for all non-negative trigonometric polynomials in $\mathcal{B}$ and that one spectral factor is given by (10.8). However, since we do not know whether $\mathfrak{S}$ is continuous, it is not clear at the outset whether the spectral factorization exists in $\mathcal{B}$ for non-polynomials. For this reason, we will assume in the following that the spectral factorization exists at least in a small neighborhood of the unity $e(\zeta) \equiv 1$, for which the spectral factorization is known. Then, it is shown that if the spectral factorization mapping is continuous at e, it will exist for all real valued functions in $\exp(\mathcal{B})$ and will be continuous there.

Since $e \in \exp(\mathcal{B})$ and since $\exp(\mathcal{B})$ is open, there exists an $\epsilon > 0$ such that the neighborhood $U_\epsilon(e) := \{\phi = \mathrm{e}^g : g \in \mathcal{B}, \|g\|_\mathcal{B} < \epsilon\}$ of e is an open subset of $\exp(\mathcal{B})$.

Definition 10.19. *Let $\mathcal{B}$ be an $\mathcal{S}$–algebra. The spectral factorization mapping $\mathfrak{S}$ is said to be continuous at e if there exists an $\epsilon > 0$ such that for every sequence $\{\phi_n\}_{n\in\mathbb{N}} \subset U_\epsilon(e)$ with $\phi_n \to e$ in $\mathcal{B}$ the sequence $\{\mathfrak{S}\,\phi_n\}_{n\in\mathbb{N}}$ of spectral factors exists in $\mathcal{B}$ and converges to $\mathfrak{S}e = e$ in $\mathcal{B}$.*

Theorem 10.20. *Let $\mathcal{B}$ be an $\mathcal{S}$–algebra on which the spectral factorization mapping $\mathfrak{S}$ is continuous at e. Then the Riesz projection $\mathfrak{P}_+$ is a bounded mapping $\mathcal{B} \to \mathcal{B}_+$.*

Proof. Since $\mathfrak{S}$ is assumed to be continuous at e, there exists an $\epsilon > 0$ such that the spectral factorization exists for each $\phi \in U_\epsilon(e)$, i.e. for each $\phi = \mathrm{e}^g$ with $g \in \mathcal{B}$ and $\|g\|_{\mathcal{B}} < \epsilon$. Let $\{g_n\}_{n\in\mathbb{N}} \in \mathcal{B}$ be an arbitrary sequence of real valued functions with $\|g_n\|_{\mathcal{B}} < \epsilon$ and with $\lim_{n\to\infty} \|g_n\|_{\mathcal{B}} = 0$. Then the continuity of the exponential function (cf. Lemma 3.12) implies that $\lim_{n\to\infty} \| \exp(g_n) - e\|_{\mathcal{B}} = 0$ and the continuity of $\mathfrak{S}$ at e implies

$$\lim_{n\to\infty} \|(\exp g_n)_+ - e\|_{\mathcal{B}} = 0 \tag{10.10}$$

for every such sequence $\{g_n\}_{n\in\mathbb{N}}$ with $\|g_n\|_{\mathcal{B}} < \epsilon$ and $\|g_n\|_{\mathcal{B}} \to 0$.

We choose an arbitrary $g \in \mathcal{B}$ and define $g_\mu := \mu\, g$ with a positive real number $\mu \in \mathbb{R}_+$ with $\mu < \epsilon$. Obviously it holds that $\|g_\mu\|_{\mathcal{B}} \to 0$ as $\mu \to 0$. By (10.10) to every $\delta > 0$ there exists a $\mu_0 > 0$ such that

$$\|(\exp g_\mu)_+ - e\|_{\mathcal{B}} < \delta \tag{10.11}$$

for all $\mu \le \mu_0$. Set $\phi_\mu := \exp g_\mu$ and consider

$$h_N := \textstyle\sum_{k=1}^{N} \frac{(-1)^k}{k} \left[(\phi_\mu)_+ - e\right]^k \ .$$

Then for all $M > N$ and all $\mu \le \mu_0$, it holds

$$\begin{aligned} \|h_M - h_N\|_{\mathcal{B}} &\le \sum_{k=N+1}^{M} \tfrac{1}{k} \, \|(\phi_\mu)_+ - e\|_{\mathcal{B}}^k \\ &\le \frac{1}{N+1} \sum_{k=N+1}^{M} \delta^k \le \frac{1}{N+1} \frac{1}{1-\delta} \end{aligned}$$

which shows that $\{h_N\}_{N\in\mathbb{N}}$ is a Cauchy sequence in $\mathcal{B}$ which converges to the function $\log[(\phi_\mu)_+]$ in $\mathcal{B}$.
On the other hand, the spectral factor $(\phi_\mu)_+$ exists and it is given by (10.8). Therewith, one has

$$\left[\log(\phi_\mu)_+\right](\zeta) = \frac{\mu}{2} \frac{1}{2\pi} \int_{-\pi}^{\pi} g(\mathrm{e}^{\mathrm{i}\tau}) \frac{\mathrm{e}^{\mathrm{i}\tau} + \zeta}{\mathrm{e}^{\mathrm{i}\tau} - \zeta} \, \mathrm{d}\tau = \frac{\mu}{2} \, (\mathfrak{R} g)(\zeta) \ , \qquad \zeta \in \mathbb{T}$$

with the Herglotz-Riesz transform $\mathfrak{R}g$ of g. Since we already saw that $\log[(\phi_\mu)_+] \in \mathcal{B}$, this shows that $\mathfrak{R}g \in \mathcal{B}$ for every $g \in \mathcal{B}$. The Herglotz-Riesz transform can also be written as (5.7) in terms of the Poisson and the conjugate Poisson integral. On the unit circle $\mathbb{T}$, this relation becomes (see Section 5).

$$(\mathfrak{R}g)(\zeta) = g(\zeta) + \mathrm{i}\,(\mathfrak{H}g)(\zeta) \ , \qquad \zeta \in \mathbb{T} \tag{10.12}$$

with the Hilbert transform $\mathfrak{H}g$ of g which is given by

$$(\mathfrak{H}g)(\zeta) = \sum_{k=-\infty}^{\infty} -\mathrm{i}\ \mathrm{sgn}(k)\, \hat{g}(k)\, \zeta^k\ , \qquad \zeta \in \mathbb{T}$$

or which may be expressed as the principal value integral (5.19). Since $\mathfrak{R}g \in \mathcal{B}$ for every $g \in \mathcal{B}$, (10.12) shows that $\mathfrak{H}g \in \mathcal{B}$ whenever $g \in \mathcal{B}$. Moreover, since $\mathfrak{H}\mathfrak{H}g = -g$ for all $g \in \mathcal{B}_0 := \{g \in \mathcal{B} : \hat{g}(0) = 0\}$, it is clear that to every $f \in \mathcal{B}_0$ there exists a $g \in \mathcal{B}_0$ such that $f = \mathfrak{H}g$. Thus $\mathfrak{H} : \mathcal{B}_0 \to \mathcal{B}_0$ is one-to-one and onto which implies that the graph of $\mathfrak{H}$ is closed, and the closed graph theorem shows that $\mathfrak{H} : \mathcal{B}_0 \to \mathcal{B}_0$ is continuous, which in turn implies the continuity of $\mathfrak{R}$ by (10.12). The Herglotz-Riesz transform $\mathfrak{R}$ can also be expressed in terms of the Riesz projection (see (6.8)) $(\mathfrak{R}g)(\zeta) = 2\,(\mathfrak{P}_+ g)(\zeta) - \hat{g}(0)$ with the zeroth Fourier coefficient $\hat{g}(0)$ of g. Since $|\hat{g}(0)| \leq \|g\|_\infty \leq \|g\|_{\mathcal{B}}$ and since $\mathfrak{R}$ is bounded on $\mathcal{B}_0$, the Riesz projection $\mathfrak{P}_+$ is bounded on $\mathcal{B}$. □

We already saw in Lemma 10.8 that if the Riesz projection $\mathfrak{P}_+$ is bounded on an Banach algebras $\mathcal{A} \subset L^1$ then the spectral factorization exists for every $\phi \in \exp(\mathcal{A})$. Therewith, the assumed continuity of $\mathfrak{S}$ at e in the previous theorem can be extended to the continuity of $\mathfrak{S}$ at every point $\phi \in \exp(\mathcal{B})$.

Theorem 10.21. *Let $\mathcal{B}$ be an $\mathcal{S}$-algebra. If the spectral factorization mapping $\mathfrak{S}$ is continuous at e in $\mathcal{B}$, then $\phi_+ = \mathfrak{S}\phi$ exists for every real valued $\phi \in \exp(\mathcal{B})$. Moreover, $\mathfrak{S}$ is locally continuous on $\mathcal{B}$, i.e. to every $\phi \in \exp(\mathcal{B})$ there exist constants $C(\phi)$ and $r(\phi)$ such that*

$$\|\mathfrak{S}\phi - \mathfrak{S}\psi\|_{\mathcal{B}} \leq C(\phi)\, \|\phi - \psi\|_{\mathcal{B}}$$

for all $\psi \in \exp(\mathcal{B})$ with $\|\phi - \psi\|_{\mathcal{B}} < r(\phi)$.

In particular, the theorem is satisfied by the constants

$$C(\phi) = 2\,(1 + \tfrac{1}{2}\,\|e\|_{\mathcal{B}})\, \|\mathfrak{P}_+\|\, \|\phi^{-1}\|_{\mathcal{B}}\, \|\mathfrak{S}\phi\|_{\mathcal{B}} \quad \text{and} \quad r(\phi) = \frac{\|\mathfrak{S}\phi\|_{\mathcal{B}}}{C(\phi)}\ . \tag{10.13}$$

Proof. By Theorem 10.20, the Riesz projection is bounded on $\mathcal{B}$. Therefore Lemma 10.8 implies that every real valued $\phi \in \exp(\mathcal{B})$ possesses a spectral factorization $\phi_+ = \mathfrak{S}\phi$ in $\mathcal{B}$ which is given by

$$(\mathfrak{S}\phi)(z) = \exp[(\mathfrak{P}_+ \log\phi)(z) - \tfrac{1}{2}\,(\mathfrak{P}_+ \log\phi)(0)]\ . \tag{10.14}$$

Choose $\phi_1, \phi_2 \in \exp(\mathcal{B})$ arbitrary, and denote the corresponding argument of the exponential function in the representation (10.14) of the spectral factor with q_1 and q_2, respectively. Then $\mathfrak{S}\phi_2 - \mathfrak{S}\phi_1 = \exp(q_1)[\exp(q_2 - q_1) - e]$ and the continuity of the exponential function (cf. Lemma 3.12) implies

$$\|\mathfrak{S}\phi_2 - \mathfrak{S}\phi_1\|_{\mathcal{B}} \leq \|\exp q_1\|_{\mathcal{B}}\, \|\exp(q_2 - q_1) - e\|_{\mathcal{B}} \tag{10.15}$$

$$\leq 2\, \|\mathfrak{S}\phi_1\|_{\mathcal{B}}\, \|q_2 - q_1\|_{\mathcal{B}} \tag{10.16}$$

provided that $\|q_2 - q_1\|_{\mathcal{B}} < 1$. Next, we investigate the term $\|q_2 - q_1\|_{\mathcal{B}}$. For it holds obviously

$$\|q_2 - q_1\|_{\mathcal{B}} \leq \|\mathfrak{P}_+[\log\phi_2 - \log\phi_1]\|_{\mathcal{B}} - \tfrac{1}{2}\,|\mathfrak{P}_+[\log\phi_2 - \log\phi_1](0)|\,\|e\|_{\mathcal{B}}$$

using only the linearity of $\mathfrak{P}_+$. To get an upper bound for the second term on the right hand side, we note that $|(\mathfrak{P}_+ f)(0)| \leq \|\mathfrak{P}_+ f\|_\infty \leq \|\mathfrak{P}_+ f\|_{\mathcal{B}}$ for every $f \in \mathcal{B}$ using the maximum modulus principle for analytic functions and that $\mathcal{B}$ is continuously embedded in $\mathcal{C}(\mathbb{T})$. Therefore and together with the boundedness of the Riesz projection one obtains

$$\|q_2 - q_1\|_{\mathcal{B}} \leq \left(1 + \tfrac{1}{2}\,\|e\|_{\mathcal{B}}\right)\,\|\mathfrak{P}_+\|\,\|\log\phi_2 - \log\phi_1\|_{\mathcal{B}} \tag{10.17}$$

where $\|\mathfrak{P}_+\|$ is the common operator norm of $\mathfrak{P}_+ : \mathcal{B} \to \mathcal{B}_+$. Next, set $f = \phi_1$ and $f - h = \phi_2$ in Lemma 3.13 on the continuity of the logarithm. This shows

$$\|\log\phi_2 - \log\phi_1\|_{\mathcal{B}} \leq \|\phi_1^{-1}\|_{\mathcal{B}}\,\|\phi_2 - \phi_1\|_{\mathcal{B}} \tag{10.18}$$

for all $\phi_1, \phi_2 \in \exp\mathcal{B}$ with $\|\phi_2 - \phi_1\|_{\mathcal{B}} < \|\phi_1^{-1}\|_{\mathcal{B}}^{-1}$. Combining (10.15), (10.17), and (10.18) one obtains the statement of the theorem with the constants (10.13). □

The previous proof implies in particular that the spectral factorization mapping is continuous on every decomposing $\mathcal{S}$–algebra.

Corollary 10.22. *The spectral factorization mapping* $\mathfrak{S}$ *is continuous on an* $\mathcal{S}$*–algebra* $\mathcal{B}$ *if and only if* $\mathcal{B}$ *is decomposing, i.e. if and only if* $\mathfrak{P}_+ : \mathcal{B} \to \mathcal{B}_+$ *is bounded.*

10.3.2 Boundedness of the Spectral Factorization mapping

Since $\mathfrak{S}$ is a non-linear operator, the continuity of $\mathfrak{S}$ does not imply the boundedness of $\mathfrak{S}$. Therefore, the boundedness of the spectral factorization mapping has to be investigated separately. This will be done in the present section. In particular, we want to characterize $\mathcal{S}$–algebras on which the spectral factorization mapping $\mathfrak{S}$ is bounded.

Let $\mathcal{B}$ be an arbitrary $\mathcal{S}$–algebra. Since the trigonometric polynomials are dense in $\mathcal{B}$, the spectral factorization exists in $\mathcal{B}$ at least for all non-negative trigonometric polynomials, and we define the boundedness of $\mathfrak{S}$, at the moment, only on the set $\mathcal{P}(N)$ of trigonometric polynomials of degree N. To this end, we define for every $N \in \mathbb{N}$ the *boundedness constant* of $\mathfrak{S}$ on $\mathcal{P}(N)$ by

$$C(N, \mathcal{B}) := \sup_{\phi \in \mathcal{P}_{\text{pos}}(N)\,,\,\|\phi\|_{\mathcal{B}} \leq 1} \|\mathfrak{S}\phi\|_{\mathcal{B}}\,. \tag{10.19}$$

By this definition, it is clear that $C(N+1, \mathcal{B}) \geq C(N, \mathcal{B})$ for all N. Based on these constants, we define the boundedness of the spectral factorization mapping on $\mathcal{B}$.

Definition 10.23. *Let* $\mathcal{B}$ *be an* $\mathcal{S}$*–algebra, then the spectral factorization mapping* $\mathfrak{S}$ *is said to be* p-bounded *on* $\mathcal{B}$ *if*

$$C(\mathcal{B}) := \sup_{N \in \mathbb{N}} C(N, \mathcal{B}) < \infty\,. \tag{10.20}$$

Of course, the p-boundedness of $\mathfrak{S}$ is a necessary requirement for the boundedness of $\mathfrak{S}$ on $\mathcal{B}$. Obviously $\lim_{N\to\infty} C(N,\mathcal{B}) = C(\mathcal{B})$, and by the definition of $C(N,\mathcal{B})$, it is clear that

$$\|\phi_+\|_{\mathcal{B}} = \|\mathfrak{S}\phi\|_{\mathcal{B}} \leq C(N,\mathcal{B}) \qquad \text{for all } \phi \in \mathcal{P}_{\mathrm{pos}}(N) \text{ with } \|\phi\|_{\mathcal{B}} \leq 1\ . \quad (10.21)$$

Let $\phi \in \mathcal{P}_{\mathrm{pos}}(N)$ be arbitrary, set $\psi := \phi / \|\phi\|_{\mathcal{B}}$, and apply (10.21) to ψ. This shows that

$$\|\phi_+\|_{\mathcal{B}}^2 \leq C(N,\mathcal{B})^2\, \|\phi_+\, \phi_+^\star\|_{\mathcal{B}} \quad (10.22)$$

for all spectral factors $\phi_+ \in \mathcal{F}[\mathcal{P}(N)]$ in $\mathcal{P}_+(N)$. We define for every $N \in \mathbb{N}$ the constants

$$D_+(N,\mathcal{B}) := \inf_{\substack{\phi_+ \in \mathcal{F}[\mathcal{P}(N)] \\ \|\phi_+\|_{\mathcal{B}}=1}} \|\phi_+\, \phi_+^\star\|_{\mathcal{B}} \qquad \text{and} \qquad D_+(\mathcal{B}) := \lim_{N\to\infty} D_+(N,\mathcal{B})\ .$$

By this definition, it is clear that the sequence $D_+(N,\mathcal{B})$ is monotone decreasing and in view of (10.22) one sees that

$$C(N,\mathcal{B}) = \frac{1}{\sqrt{D_+(N,\mathcal{B})}} \qquad \text{and} \qquad C(\mathcal{B}) = \frac{1}{\sqrt{D_+(\mathcal{B})}}\ .$$

Therewith, the boundedness condition (10.20) can be stated also in terms of $D_+(\mathcal{B})$.

Lemma 10.24. *The spectral factorization mapping $\mathfrak{S}$ is p-bounded on $\mathcal{B}$ if and only if $D_+(\mathcal{B}) > 0$.*

The difficulty in the definition of $D_+(N,\mathcal{B})$ is that the infinium is taken over the set $\mathcal{F}[\mathcal{P}(N)]$ of all functions $\phi_+ \in P_+(N)$ which are obtained by a spectral factorization from a polynomial spectra $\phi \in \mathcal{P}(N)$, but this set is unknown in general. However, we will show next that the p-boundedness condition remains unchanged even if one takes the infimum over all analytic polynomials $\phi \in \mathcal{P}_+(N)$ instead of $\mathcal{F}[\mathcal{P}(N)]$. To this end, we define the constants

$$D(N,\mathcal{B}) := \inf_{\substack{\phi \in \mathcal{P}_+(N) \\ \|\phi\|_{\mathcal{B}}=1}} \|\phi\, \phi^\star\|_{\mathcal{B}} \quad \text{and} \quad D(\mathcal{B}) := \inf_{\substack{\phi \in \mathcal{B}_+ \\ \|\phi\|_{\mathcal{B}}=1}} \|\phi\, \phi^\star\|_{\mathcal{B}}\ . \quad (10.23)$$

It is immediately clear that $D(N,\mathcal{B}) \leq D_+(N,\mathcal{B})$ and $D(\mathcal{B}) \leq D_+(\mathcal{B})$. Moreover $D(\mathcal{B}) = \lim_{N\to\infty} D(N,\mathcal{B})$.

Therewith, we are able to give necessary and sufficient conditions for the p-boundedness of the spectral factorization mapping $\mathfrak{S}$ on $\mathcal{S}$–algebras in the following proposition.

Proposition 10.25. *Let $\mathcal{B}$ be an $\mathcal{S}$–algebra. Then the spectral factorization mapping $\mathfrak{S}$ is p-bounded on $\mathcal{B}$ if and only if*

$$D(\mathcal{B}) = \inf_{\phi \in \mathcal{B}_+\,,\, \|\phi\|_{\mathcal{B}}=1} \|\phi\, \phi^\star\|_{\mathcal{B}} > 0\ . \quad (10.24)$$

Proof. If (10.24) is satisfied, then $D_+(\mathcal{B}) \geq D(\mathcal{B}) > 0$ and Lemma 10.24 implies the p-boundedness of $\mathfrak{S}$.

The necessity of (10.24) for the p-boundedness of $\mathfrak{S}$ is shown by contradiction. We assume that $\mathfrak{S}$ is p-bounded but that $D(\mathcal{B}) = 0$. Let $N \in \mathbb{N}$ be a fixed degree and let $D(N, \mathcal{B})$ be the constant defined by (10.23). Then to every $\delta > 0$ there exists a $\phi \in \mathcal{P}_+(N)$ with $\|\phi\|_{\mathcal{B}} = 1$ and

$$\|\phi\, \phi^\star\|_{\mathcal{B}} \leq D(N, \mathcal{B}) + \delta\,. \tag{10.25}$$

Moreover, since $|\phi(\zeta)\, \overline{\phi(\zeta)}| = |\phi(\zeta)|^2$ for all $\zeta \in \mathbb{T}$, Proposition 10.14 implies that $\|\phi\|_\infty^2 = \|\phi\, \phi^\star\|_\infty \leq \|\phi\, \phi^\star\|_{\mathcal{B}}$. Together with (10.25) this gives

$$\|\phi\|_\infty \leq \sqrt{D(N, \mathcal{B}) + \delta}\,. \tag{10.26}$$

Let $\mu > 0$ be arbitrary and define the function

$$\phi_\mu := \frac{\phi + \|\phi\|_\infty + \mu}{\|\phi + \|\phi\|_\infty + \mu\|_{\mathcal{B}}}\,.$$

By this definition, it is clear that $\phi_\mu \in \mathcal{P}_+(N)$, that $\|\phi_\mu\|_{\mathcal{B}} = 1$, and that $|\phi_\mu(z)| > 0$ for all $z \in \mathbb{D}$. Therefore $\phi_\mu \in \mathcal{F}[\mathcal{P}(N)]$ is a spectral factor. For ϕ_μ it holds

$$\|\phi_\mu\, \phi_\mu^\star\|_{\mathcal{B}} = \frac{\|\phi\, \phi^\star + (\|\phi\|_\infty + \mu)(\phi + \phi^\star) + (\|\phi\|_\infty + \mu)^2\|_{\mathcal{B}}}{\|\phi + \|\phi\|_\infty + \mu\|_{\mathcal{B}}^2}$$

and since $D_+(N, \mathcal{B}) \leq \|\phi_\mu\, \phi_\mu^\star\|_{\mathcal{B}}$, one gets

$$D_+(N, \mathcal{B}) \leq \frac{\|\phi\, \phi^\star\|_{\mathcal{B}} + (\|\phi\|_\infty + \mu)(1 + m_1)\|\phi\|_{\mathcal{B}} + (\|\phi\|_\infty + \mu)^2\|e\|_{\mathcal{B}}}{\|\phi + (\|\phi\|_\infty + \mu)\|_{\mathcal{B}}^2} \tag{10.27}$$

for every $\mu > 0$. Moreover, since

$$1 = \|\phi\|_{\mathcal{B}} = \|\,\phi + \|\phi\|_\infty - \|\phi\|_\infty\,\|_{\mathcal{B}} \leq \|\,\phi + \|\phi\|_\infty\,\|_{\mathcal{B}} + \|\phi\|_\infty\,\|e\|_{\mathcal{B}}$$

one gets together with (10.26) that

$$\|\,\phi + \|\phi\|_\infty\,\|_{\mathcal{B}} \geq 1 - \|\phi\|_\infty\,\|e\|_{\mathcal{B}} \geq 1 - \|e\|_{\mathcal{B}}\,\sqrt{D(N, \mathcal{B}) + \delta}\,. \tag{10.28}$$

Letting $\mu \to 0$ in (10.27), using that $\|\phi\|_{\mathcal{B}} = 1$, and applying the bounds (10.25), (10.26), and (10.28) one obtains from (10.27) that

$$D_+(N, \mathcal{B}) \leq \frac{(1 + \|e\|_{\mathcal{B}})[D(N, \mathcal{B}) + \delta] + (1 + m_1)\,\sqrt{D(N, \mathcal{B}) + \delta}}{1 - \|e\|_{\mathcal{B}}\sqrt{D(N, \mathcal{B}) + \delta}} \tag{10.29}$$

for arbitrary $\delta > 0$. Since we assumed that $D(\mathcal{B}) = 0$ and because $\lim_{N\to\infty} D(N, \mathcal{B}) = D(\mathcal{B})$, the right hand side of (10.29) converges to zero, which shows that

$$D_+(\mathcal{B}) = \lim_{N\to\infty} D_+(N,\mathcal{B}) = 0 \,. \tag{10.30}$$

However, this contradicts the assumption that $\mathfrak{S}$ is bounded since by Lemma 10.24 the boundedness implies that $D_+(\mathcal{B}) > 0$. This proves that the assumption $D(\mathcal{B}) = 0$ was wrong and that $D(\mathcal{B}) > 0$ whenever $\mathfrak{S}$ is bounded. □

Assume that $\mathcal{B}$ is an $\mathcal{S}$–algebra on which the spectral factorization mapping is p-bounded, let $f \in \mathcal{B}_+$ be arbitrary, and define $g := f/\|f\|_{\mathcal{B}}$. For this function obviously holds that $g \in \mathcal{B}_+$ and that $\|g\|_{\mathcal{B}} = 1$. Since $\mathfrak{S}$ is assumed to be p-bounded, Proposition 10.25 implies $\|g\, g^\star\|_{\mathcal{B}} = \|f\|_{\mathcal{B}}^{-2}\, \|f\, f^\star\|_{\mathcal{B}} \geq D(\mathcal{B}) > 0$, which shows that

$$D(\mathcal{B})\, \|f\|_{\mathcal{B}}^2 \leq \|f\, f^\star\|_{\mathcal{B}} \qquad \text{for all } f \in \mathcal{B}_+ \,. \tag{10.31}$$

Recall that an *inner function* is a $\varphi \in L_+^\infty$ such that $|\varphi(\zeta)| = 1$ for almost all $\zeta \in \mathbb{T}$, which implies that $|\varphi(\zeta)\varphi^\star(\zeta)| = |\varphi(\zeta)\overline{\varphi}(\zeta)| = 1$ for almost all $\zeta \in \mathbb{T}$. Assume that $\varphi \in \mathcal{B}$ is an arbitrary inner function which belongs to $\mathcal{B}$. Then it is clear that $\varphi \in \mathcal{B}_+$ and that $\varphi\, \varphi^\star = e$. Since $\mathfrak{S}$ is assumed to be p-bounded on $\mathcal{B}$, (10.31) implies that there exists a universal upper bound on the norm for every inner function in $\mathcal{B}$

$$\|\varphi\|_{\mathcal{B}} \leq \sqrt{\tfrac{\|e\|_{\mathcal{B}}}{D(\mathcal{B})}} =: C_3$$

which depends only on the algebra $\mathcal{B}$. Define for $n = 0, 1, 2, \cdots$ the functions $s_n(z) := z^n$ in the complex plane. It is clear that $s_n \in \mathcal{B}_+$ for all $n \in \mathbb{N}$ and each s_n is an inner function. Moreover, the functions $s_{-n}(z) := s_n^\star(z) = z^{-n}$, $n \in \mathbb{N}$ belong to $\mathcal{B}$ and because of Proposition 10.15 one obtains a uniform upper bound

$$\|s_n\|_{\mathcal{B}} \leq m_1\, C_3 \qquad \text{for all } n \in \mathbb{Z} \tag{10.32}$$

for the norms of these particular inner functions on every $\mathcal{S}$–algebra on which $\mathfrak{S}$ is bounded.

Lemma 10.26. *Let $\mathcal{B}$ be an $\mathcal{S}$–algebra on which $\mathfrak{S}$ is p-bounded. Then there exists a constant $m_6 > 0$ such that*

$$m_6\, \|f\|_{\mathcal{B}}^2 \leq \|f\, f^\star\|_{\mathcal{B}} \qquad \textit{for all}\, f \in \mathcal{B} \tag{10.33}$$

or equivalently that

$$\inf_{f\in\mathcal{B}\,,\, \|f\|_{\mathcal{B}}=1} \|f\, f^\star\|_{\mathcal{B}} \geq m_6 > 0 \,.$$

Proof. Let $f \in \mathcal{B}$ arbitrary. Since the trigonometric polynomials are dense in $\mathcal{B}$, there exists a sequence $\{p_n\}_{n=1}^\infty$ of polynomials $p_n \in \mathcal{P}(n)$ such that $\lim_{n\to\infty} \|f - p_n\|_{\mathcal{B}}$. Consequently, to every $\epsilon > 0$ there exists an N_0 such that

$$\|f\|_{\mathcal{B}} - \epsilon \leq \|p_n\|_{\mathcal{B}} \leq \|f\|_{\mathcal{B}} + \epsilon \qquad \text{for all } n \geq N_0 \; . \tag{10.34}$$

Let $s_n(z) = z^n$ and define the function $g_n := s_n\, p_n$. Then $g_n \in \mathcal{P}_+(2n) \subset \mathcal{B}_+$ and

$$\|p_n\|_{\mathcal{B}} = \|s_n^\star\, s_n\, p_n\|_{\mathcal{B}} \leq \|s_n^\star\|_{\mathcal{B}}\, \|g_n\|_{\mathcal{B}} \leq m_1\, C_3\, \|g_n\|_{\mathcal{B}} \; . \tag{10.35}$$

Since $g_n \in \mathcal{B}_+$ and $\mathfrak{S}$ is p-bounded, Proposition 10.25 implies

$$D(\mathcal{B})\, \|g_n\|_{\mathcal{B}}^2 \leq \|g_n\, g_n^\star\|_{\mathcal{B}} = \|s_n^\star\, g_n \; \cdot s_n\, g_n^\star\|_{\mathcal{B}} = \|p_n\, p_n^\star\|_{\mathcal{B}} \; . \tag{10.36}$$

Next we consider the expression $f\,\overline{f} - p_n\, p_n^\star$. With the triangle inequality and together with Proposition 10.15 and relation (10.34), one obtains

$$\begin{aligned} \|f\, f^\star - p_n\, p_n^\star\|_{\mathcal{B}} &\leq \|f^\star\|_{\mathcal{B}}\, \|f - p_n\|_{\mathcal{B}} + \|p_n\|_{\mathcal{B}}\, \|(f - p_n)^\star\|_{\mathcal{B}} \\ &\leq m_1\, (\|f\|_{\mathcal{B}} + \|p_n\|_{\mathcal{B}})\, \|f - p_n\|_{\mathcal{B}} \\ &\leq m_1\, (2\, \|f\|_{\mathcal{B}} + \epsilon)\, \|f - p_n\|_{\mathcal{B}} \; . \end{aligned}$$

Since the right hand side of the last inequality converges to 0 as $n \to \infty$, this shows that for $\epsilon > 0$ there exists an $N_1 \geq N_0$ such that

$$\|p_n\, p_n^\star\|_{\mathcal{B}} \leq \|f\, f^\star\|_{\mathcal{B}} + \epsilon \qquad \text{for all } n \geq N_1 \; . \tag{10.37}$$

Putting together all the previous steps, one obtains for an arbitrary $n \geq N_1$

$$\begin{aligned} (\|f\|_{\mathcal{B}} - \epsilon)^2 &\overset{(10.34)}{\leq} \|p_n\|_{\mathcal{B}}^2 \overset{(10.35)}{\leq} m_1^2\, C_3^2\, \|g_n\|_{\mathcal{B}}^2 \overset{(10.36)}{\leq} \frac{m_1^2\, C_3^2}{D(\mathcal{B})}\, \|p_n\, p_n^\star\|_{\mathcal{B}} \\ &\overset{(10.37)}{\leq} \frac{m_1^2\, C_3^2}{D(\mathcal{B})}\, (\|f\, f^\star\|_{\mathcal{B}} + \epsilon) \end{aligned}$$

Since ϵ was chosen arbitrary this shows that $\frac{D(\mathcal{B})}{m_1^2\, C_3^2}\, \|f\|_{\mathcal{B}}^2 \leq \|f\, f^\star\|_{\mathcal{B}}$ for all $f \in \mathcal{B}$, which is equivalent to the statement of the lemma with the constant $m_6 = D(\mathcal{B})/(m_1\, C_3)^2 > 0$. $\square$

The main result of this section is the following theorem, which shows that $\mathcal{C}(\mathbb{T})$ is essentially the only $\mathcal{S}$–algebra on which the spectral factorization mapping $\mathfrak{S}$ is bounded.

Theorem 10.27. *Let $\mathcal{B}$ be an $\mathcal{S}$–algebra. If the spectral factorization mapping $\mathfrak{S}$ is p-bounded on $\mathcal{B}$ then there exists a constant m_2 such that*

$$\|f\|_\infty \;\leq\; \|f\|_{\mathcal{B}} \;\leq\; m_2\, \|f\|_\infty \qquad \textit{for all } f \in \mathcal{B} \, . \tag{10.38}$$

That is, $\mathcal{B}$ is isomorphic to $\mathcal{C}(\mathbb{T})$.

Proof. The lower bound in (10.38) is equivalent to Proposition 10.14. At the beginning let $h \in \mathcal{B}$ be a real valued function. Then by Lemma 10.26 it holds that $m_6\, \|h\|_{\mathcal{B}}^2 \leq \|h\, h^\star\|_{\mathcal{B}} = \|h^2\|_{\mathcal{B}}$. Moreover, since $h\, h \in \mathcal{B}$, Lemma 10.26

can be applied to h^2 which gives $m_6\,\|h^2\|_{\mathcal{B}}^2 \leq \|h^4\|_{\mathcal{B}}$. Together with the previous inequality, one gets $m_6\,m_6^2\,\|h\|_{\mathcal{B}}^4 \leq \|h^4\|_{\mathcal{B}}$. Applying this upper bound repeatedly, one obtains

$$\|h\|_{\mathcal{B}} \leq C_5\,\|h^{2^n}\|_{\mathcal{B}}^{1/2^n} \qquad \text{with} \quad C_5 = \left(\prod_{k=0}^{n-1} (m_6)^{2^k}\right)^{-1/2^n} = (m_6)^{-\frac{2^n-1}{2^n}} \;.$$

By the spectral radius formula (Theorem 3.10), the spectral radius of h is given by $r_\sigma(h) = \lim_{n\to\infty} \|h^n\|_{\mathcal{B}}^{1/n}$, and property $(B4)$ of an $\mathcal{S}$–algebra implies that $\|h\|_\infty = r_\sigma(h)$. Thus, for $n \to \infty$ one obtains

$$\|h\|_{\mathcal{B}} \leq \tfrac{1}{m_6}\,\|h\|_\infty \tag{10.39}$$

for every real valued $h \in \mathcal{B}$. Let now $f = f_1 + \mathrm{i}\,f_2$ be a complex function in $\mathcal{B}$ with real functions $f_1, f_2 \in \mathcal{B}$. Then it follows from (10.39) that $\|f\|_{\mathcal{B}} \leq \frac{1}{m_6}\,(\|f_1\|_\infty + \|f_2\|_\infty) \leq \frac{2}{m_6}\,\|f\|_\infty$ which is the upper bound in (10.38). □

We know that the the Riesz projection is unbounded on $\mathcal{C}(\mathbb{T})$ (cf. Theorem 6.14). This means that there exists a sequence $\{f_n\}_{n=1}^\infty \subset \mathcal{C}(\mathbb{T})$ with $\|f_n\|_\infty \leq 1$ for all $n \in \mathbb{N}$ but such that $\|\mathfrak{P}_+ f_n\|_\infty \to \infty$ as $n \to \infty$. Then Theorem 10.27 implies $\frac{1}{m_2}\|f_n\|_{\mathcal{B}} \leq \|f_n\|_\infty$ which shows that $\{f_n\}_{n\in\mathbb{N}}$ is a uniformly bounded sequence in every $\mathcal{S}$–algebra $\mathcal{B}$ on which $\mathfrak{S}$ is bounded. Moreover, by the continuous embedding of $\mathcal{B}$ into $\mathcal{C}(\mathbb{T})$ one has that $\|\mathfrak{P}_+ f_n\|_\infty \leq \|\mathfrak{P}_+ f_n\|_{\mathcal{B}}$ which shows that $\|\mathfrak{P}_+ f_n\|_{\mathcal{B}} \to \infty$ as $n \to \infty$. Consequently, the Riesz projection $\mathfrak{P}_+$ is also unbounded on $\mathcal{B}$ and we have the following corollaries.

Corollary 10.28. *Let $\mathcal{B}$ be an $\mathcal{S}$–algebra. If the spectral factorization mapping $\mathfrak{S}$ is bounded on $\mathcal{B}$ then the Riesz projection $\mathfrak{P}_+$ is unbounded on $\mathcal{B}$.*

Corollary 10.29. *The spectral factorization mapping $\mathfrak{S}$ is unbounded on every decomposing $\mathcal{S}$–algebra.*

Proof. Let $\mathcal{B}$ be an $\mathcal{S}$–algebra and assume that $\mathfrak{S}$ is bounded on $\mathcal{B}$. Then $\mathfrak{S}$ is p-bounded and Corollary 10.28 implies that $\mathfrak{P}_+$ is unbounded on $\mathcal{B}$, which contradicts the decomposing assumption of $\mathcal{B}$. Therefore, $\mathfrak{S}$ cannot be bounded on $\mathcal{B}$. □

To recapitulate and summarize the continuity and boundedness behavior of the spectral factorization mapping $\mathfrak{S}$ on $\mathcal{S}$–algebras $\mathcal{B}$, we see that $\mathfrak{S}$ is either bounded or continuous on every $\mathcal{S}$–algebra but never both. Because a necessary condition for the boundedness of $\mathfrak{S}$ is the unboundedness of the Riesz projection $\mathfrak{P}_+$ on $\mathcal{B}$. Conversely, the boundedness of $\mathfrak{P}_+$ is necessary and sufficient for the continuity of the spectral factorization mapping. This conclusion may be summarized as follows

Corollary 10.30. *Let $\mathcal{B}$ be an $\mathcal{S}$–algebra and let $\mathfrak{S}$ be the spectral factorization mapping on $\mathcal{B}$. Then*

- *if $\mathfrak{S}$ is continuous on $\mathcal{B}$ then $\mathfrak{S}$ is unbounded.*
- *if $\mathfrak{S}$ is bounded on $\mathcal{B}$ then $\mathfrak{S}$ is discontinuous.*

The following two subsections consider the spectral factorization on two particular examples of an $\mathcal{S}$–algebra in more detail. On the first algebra $\mathfrak{S}$ is bounded on the second algebra $\mathfrak{S}$ is continuous.

10.3.3 Example A – Factorization on $\mathcal{C}(\mathbb{T})$

If the spectral factorization mapping should be bounded on an $\mathcal{S}$–algebra $\mathcal{B}$, then $\mathcal{B}$ has to be isomorphic to $\mathcal{C}(\mathbb{T})$ (cf. Theorem 10.27). Therefore, $\mathcal{C}(\mathbb{T})$ is in a sense the only $\mathcal{S}$–algebra on which the spectral factorization mapping is bounded. On the other hand, since the Riesz projection $\mathfrak{P}_+$ is unbounded on $\mathcal{C}(\mathbb{T})$, the spectral factorization mapping is discontinuous on $\mathcal{C}(\mathbb{T})$. In this paragraph, it is shown that the continuity behavior of $\mathfrak{S}$ is even worse on $\mathcal{C}(\mathbb{T})$, in the sense that every non-negative function in $\mathcal{C}(\mathbb{T})$ is a discontinuity point of the spectral factorization mapping.

Definition 10.31 (Continuity Point). *Let $\mathcal{B}$ be an $\mathcal{S}$–algebra and let $\phi \in \mathcal{B}_{pos}$. Then ϕ is called a* continuity point *of the spectral factorization mapping $\mathfrak{S} : \mathcal{C}(\mathbb{T}) \to A(\mathbb{D})$ if for all $\varepsilon > 0$ there exists a constant $C = C(\phi, \varepsilon)$ with $\lim_{\varepsilon\to 0} C(\phi, \varepsilon) = 0$ such that*

$$\|\mathfrak{S}\phi - \mathfrak{S}\psi\|_{\mathcal{B}} < C(\phi, \varepsilon) \ . \tag{10.40}$$

for all $\psi \in \mathcal{B}_{pos}$ and with $\|\phi - \psi\|_{\mathcal{B}} < \varepsilon$.

Conversely, ϕ is called a discontinuity point *of the spectral factorization mapping if there exists a constant $C > 0$ such that for all $0 < \varepsilon \leq 1$ there exists a $\psi \in \mathcal{B}_{pos}$ with $\|\phi - \psi\|_{\mathcal{B}} < \varepsilon$ such that $\|\mathfrak{S}\phi - \mathfrak{S}\psi\|_{\mathcal{B}} > C$.*

Of course, this notion of continuity is equivalent to the definition of continuity in terms of sequences, as it was used in Section 10.3.1 because if $\phi \in \mathcal{B}_{\text{pos}}$ is a continuity point of the spectral factorization mapping and $\{\phi_n\}_{n\in\mathbb{N}}$ is a sequence of spectra which converges to ϕ in $\mathcal{B}$, then the corresponding sequence $\{\mathfrak{S}\phi_n\}_{n\in\mathbb{N}}$ of the spectral factors converges in $\mathcal{B}$ to the spectral factor $\phi_+ = \mathfrak{S}\phi$ of ϕ. Note that in contrast to linear operators, the continuity constant $C(\phi, \varepsilon)$ depends on the actual spectrum ϕ, in general. For a linear operator, the continuity constant would be independent of ϕ and it would linear in ε, i.e. it would have the form $C(\varepsilon) = C_0\,\varepsilon$, with a universal constant C_0 equal to the operator norm of $\mathfrak{S}$.

Theorem 10.32. *Let $\phi \in \mathcal{C}_{pos}(\mathbb{T})$ be a real valued continuous function on $\mathbb{T}$ with $\phi(\zeta) > 0$ for all $\zeta \in \mathbb{T}$. Then ϕ is a discontinuity point of the spectral factorization, i.e.*

$$\liminf_{\varepsilon\to 0} \sup_{\substack{\psi\in\mathcal{C}_{pos}(\mathbb{T}) \\ \|\phi-\psi\|_\infty<\varepsilon}} \|\phi_+ - \psi_+\|_\infty > 0 \ . \tag{10.41}$$

This theorem shows that *every* continuous and strictly positive function ϕ on $\mathbb{T}$ is a discontinuity point of the spectral factorization mapping on $\mathcal{C}(\mathbb{T})$. This means that to every $\phi \in \mathcal{C}_{\text{pos}}(\mathbb{T})$ there exists a sequence $\{\psi_n\}_{n\in\mathbb{N}}$ of positive spectra in $\mathcal{C}_{\text{pos}}(\mathbb{T})$ which converge to ϕ

$$\lim_{n\to\infty} \|\phi - \psi_n\|_\infty = 0$$

but for which the sequence $\{(\psi_n)_+\}_{n\in\mathbb{N}}$ of spectral factors does not converge to $\phi_+ = \mathfrak{S}\phi$. Thus, there exists a constant $\varepsilon > 0$ such that

$$\lim_{n\to\infty} \|\phi_+ - (\psi_n)_+\|_\infty > \varepsilon \ .$$

The following proof of the theorem is constructive in the sense that it constructs a sequence $\{\psi_n\}_{n\in\mathbb{N}}$ with those properties.

Proof. Let $\phi \in \mathcal{C}_{\text{pos}}(\mathbb{T})$ be an arbitrary function with $\phi(\zeta) > 0$ for all $\zeta \in \mathbb{T}$, and denote by $\phi_{max} := \max_{\zeta\in\mathbb{T}} \phi(\zeta)$ the maximum value of ϕ. Without loss of generality, we assume that $\phi(1) = \phi_{max}$.

1) In the first part, we assume that $\phi_+(r\mathrm{e}^{\mathrm{i}\theta})$ converges as $r \to 1$ for all $\theta \in [-\pi,\pi)$. Let $\varepsilon > 0$ be an arbitrary number. We define a function $g \in \mathcal{C}_{\text{pos}}(\mathbb{T})$ and an interval $I \subset [-\pi,\pi)$ with $0 \in I$ such that the function g satisfies the following three conditions

$$\begin{aligned}
(i)&\quad \phi(\zeta) + g(\zeta) = \phi_{max} \qquad \text{for all } \zeta \in \mathbb{T}\\
(ii)&\quad \min_{\zeta\in\mathbb{T}} [\phi(\zeta) + g(\zeta)] > 0\\
(iii)&\quad \max_{\theta\in I} \left|g(\mathrm{e}^{\mathrm{i}\theta})\right| < \tfrac{\varepsilon}{2} \ .
\end{aligned}$$

Since ϕ is assumed to be continuous and strictly positive, it is clear that such a g can always be found by choosing the size of the interval I appropriately. Furthermore, let $q \in \mathcal{C}(\mathbb{T})$ be a real valued function with $\left|q(\mathrm{e}^{\mathrm{i}\theta})\right| < \frac{\varepsilon}{2}$ for all $\theta \in [-\pi,\pi)$, with $q(1) = 0$, and such that

$$\lim_{r\to 1} \frac{1}{2\pi} \int_{-\pi}^{\pi} \log\left[\phi_{max} + q(\mathrm{e}^{\mathrm{i}\theta})\right] \mathcal{Q}_r(\theta)\,\mathrm{d}\theta = \infty \ . \tag{10.42}$$

in which $\mathcal{Q}_r$ denotes the conjugate Poisson kernel as given in (5.4). That such functions q exist was shown in [8].

Consider now the function $\phi_\varepsilon(\zeta) := \phi(\zeta) + g(\zeta) + q(\zeta)$ for $\zeta \in \mathbb{T}$. By choosing ε sufficiently small, it can always be achieved that $\phi_\varepsilon(\zeta) > 0$ for all $\zeta \in \mathbb{T}$. Moreover, for this function it holds that $\|\phi - \phi_\varepsilon\|_\infty < \varepsilon$ and $\phi_\varepsilon(1) = \phi_{max}$. Now, we analyze $\log\phi_\varepsilon$ and determine the conjugate Poisson integral of $\log\phi_\varepsilon$ at the point $z = r\mathrm{e}^{\mathrm{i}0}$

$$\mathfrak{Q}(\log\phi_\varepsilon)(r) = \frac{1}{2\pi}\int_I \log\phi_\varepsilon(\mathrm{e}^{\mathrm{i}\theta})\,\mathcal{Q}_r(-\theta)\,\mathrm{d}\theta + \frac{1}{2\pi}\int_{[-\pi,\pi)\setminus I} \log\phi_\varepsilon(\mathrm{e}^{\mathrm{i}\theta})\,\mathcal{Q}_r(-\theta)\,\mathrm{d}\theta \ .$$

Next, we let $r \to 1$. Because of the property (10.42) of q, it follows that the first integral over the interval I diverges for $r \to 1$. The second integral is bounded, because $\theta = 0$ does not belong to the integration region and because $\phi_\varepsilon(e^{i\theta}) > 0$ in the whole integration region. It follows for ϕ_ε that

$$\lim_{r \to 1} (\mathfrak{Q} \log \phi_\varepsilon)(r) = \lim_{r \to 1} \frac{1}{2\pi} \int_{-\pi}^{\pi} \log \phi_\varepsilon(e^{i\theta})\, \mathcal{Q}_r(-\theta)\, d\theta = -\infty\,. \tag{10.43}$$

Since the spectral factorization mapping $\mathfrak{S}$ can be written in terms of the Herglotz-Riesz transform (5.2), the spectral factor of ϕ_ϵ becomes $(\phi_\varepsilon)_+ = \exp(\frac{1}{2}\mathfrak{R}[\log \phi_\varepsilon])$. Using the decomposition (5.7) of $\mathfrak{R}$ into the Poisson and conjugate Poisson integral and considering the point $z = re^{i0}$, one gets

$$(\phi_\varepsilon)_+(r) = e^{\frac{1}{2}(\mathfrak{P} \log \phi_\varepsilon)(r)} \left\{ \cos\left[\tfrac{1}{2}(\mathfrak{Q} \log \phi_\varepsilon)(r)\right] + i \sin\left[\tfrac{1}{2}(\mathfrak{Q} \log \phi_\varepsilon)(r)\right] \right\}\,. \tag{10.44}$$

Since ϕ_ε is continuous and strictly positive, the Poisson integral of $\log \phi_\varepsilon$ converges to $\log \phi_\varepsilon$ as $r \to 1$. In particular, it follows that $\lim_{r \to 1} \exp\left[\frac{1}{2}(\mathfrak{P} \log \phi_\varepsilon)(r)\right] = \sqrt{\phi_{max}}$. However, as (10.43) shows, the conjugate Poisson integral of $\log \phi_\varepsilon$ diverges as $r \to 1$. Therewith, (10.44) gives for the real part of $(\phi_\epsilon)_+$ that there exists a constant $C > 0$ such that

$$\limsup_{r \to 1} \Re\{(\phi_\varepsilon)_+(r)\} = C\sqrt{\phi_{max}} \quad \text{and} \quad \liminf_{r \to 1} \Re\{(\phi_\varepsilon)_+(r)\} = -C\sqrt{\phi_{max}}$$

and a similar result is obtained for the imaginary part of $(\phi_\varepsilon)_+$

$$\limsup_{r \to 1} \Im\{(\phi_\varepsilon)_+(r)\} = C\sqrt{\phi_{max}} \quad \text{and} \quad \liminf_{r \to 1} \Im\{(\phi_\varepsilon)_+(r)\} = -C\sqrt{\phi_{max}}\,.$$

It follows that

$$\|(\phi_\varepsilon)_+ - \phi_+\|_\infty = \sup_{|z|<1} |\phi_+(z) - (\phi_\varepsilon)_+(z)| \geq C\sqrt{\phi_{max}}\,.$$

The right hand side is independent of ε, which shows that (10.41) holds.

2) Assume now that $\phi_+(re^{i\theta})$ does not converge as $r \to 1$. Then it holds

$$\sup_{\substack{g \in \mathcal{C}(\mathbb{T}),\ g>0 \\ \|\phi - g\|_\infty < \varepsilon}} \|\phi_+ - g_+\|_\infty \geq \|\phi_+ - f_+\|_\infty$$

in which f is such that $\|\phi - f\|_\infty < \varepsilon$, $f(\zeta) > 0$ for all $\zeta \in \mathbb{T}$ and such that $f_+ \in A(\mathbb{D})$. □

10.3.4 Example B – Factorization on the Wiener algebra

As a second example, we consider the spectral factorization on the Wiener algebra $\mathcal{W}$ in some detail. Recall from Example 3.5 that $\mathcal{W}$ is the set of all functions of the form $\phi(e^{i\theta}) = \sum_{k=-\infty}^{\infty} \hat{\phi}(k)\, e^{ik\theta}$, $\theta \in [-\pi, \pi)$ with $\|\phi\|_{\mathcal{W}} = \sum_{k=-\infty}^{\infty} |\hat{\phi}(k)| < \infty$. Moreover, $\mathcal{W}_+$ is the subset of all $\phi \in \mathcal{W}$ for which $\hat{\phi}(k) =$

0 for all $k < 0$. As before, a functions in $\phi \in \mathcal{W}_+$ will be identified with the function $\phi(z) = \sum_{k=0}^{\infty} \hat{\phi}(k)\, z^k$ where $z \in \overline{\mathbb{D}}$.

Since $\mathcal{W}$ is a decomposing $\mathcal{S}$–algebra, the spectral factorization exists at least for each real valued $\phi \in \exp(\mathcal{W})$ by Lemma 10.8, and Corollary 10.22 shows that the spectral factorization mapping $\mathfrak{S}$ is continuous on $\mathcal{W}$. On the other hand, Corollary 10.29 states that the spectral factorization mapping $\mathfrak{S}$ is unbounded on $\mathcal{W}$. This unboundedness of the spectral factorization mapping $\mathfrak{S}$ implies that the norm of the spectral factor ϕ_+ may become arbitrarily large even though the norm of the given spectrum is bounded by $\|\phi\|_{\mathcal{W}} \leq 1$. Subsequently, we will characterize subsets of $\mathcal{W}$ on which $\mathfrak{S}$ is uniformly bounded. These subsets of $\mathcal{W}$ are specified by the minimum value which is attained by the spectral densities in these subsets.

We introduce the following three subsets of the Wiener algebra:

$$\begin{aligned} \mathcal{M}_W(c_0) &:= \{\phi \in \mathcal{W} \ :\ \|\phi\|_{\mathcal{W}} \leq 1, \quad |\hat{\phi}(0)| \geq c_0\} \\ \mathcal{M}_+(c_0) &:= \{\phi \in \mathcal{W}_+ \ :\ \|\phi\|_{\mathcal{W}} \leq 1;\ |\phi(\zeta)| \geq c_0,\ \forall \zeta \in \mathbb{T}\} \\ \mathcal{M}_{\mathbb{R}}(c_0) &:= \{\phi \in \mathcal{W},\ \text{real valued}\ :\ \|\phi\|_{\mathcal{W}} \leq 1,\ \phi(\zeta) \geq c_0, \forall \zeta \in \mathbb{T}\} \end{aligned}$$

in which $\hat{\phi}(0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \phi(\mathrm{e}^{\mathrm{i}\theta})\, \mathrm{d}\theta$ is the zeroth Fourier coefficient of ϕ and $0 < c_0 \leq 1$ is an arbitrary positive constant. The two sets $\mathcal{M}_+(c_0)$ and $\mathcal{M}_{\mathbb{R}}(c_0)$ are obviously strict subsets of $\mathcal{M}_W(c_0)$. However, they are not just the restriction of $\mathcal{M}_W(c_0)$ to causal and real valued functions, respectively. Indeed there exists functions $\phi \in \mathcal{M}_W(c_0) \cap \mathcal{W}_+$ but with $\phi \notin \mathcal{M}_+(c_0)$, e.g. the function $\phi(z) = c_0 - (1 - c_0)\, z$, and likewise there are real valued functions $\phi \in \mathcal{M}_W(c_0)$ which belong not to $\mathcal{M}_{\mathbb{R}}(c_0)$, e.g. the function $\phi(\mathrm{e}^{\mathrm{i}\theta}) = c_0 - (1 - c_0)\, \cos(\theta)$.

Preliminaries

To investigate the norm of the spectral factors on the above sets, it is necessary to determine upper bounds on the logarithm of the functions from these sets. Nevertheless, we will give here a much more general result on the Wiener norm of a large class of functions on these sets. The logarithm will only be one special function from this set. For that purpose the set $\mathcal{G}$ of functions is defined as follows.

Definition 10.33. *The symbol $\mathcal{G}$ denotes the set of all functions G such that for all $\lambda \in (0,1)$ the function $Q_\lambda(z) := G(\lambda[1 - z])$ is analytic for all $|z| < 1$ and such that the power series representation*

$$Q_\lambda(z) = G(\lambda[1 - z]) = \textstyle\sum_{k=0}^{\infty} q_k(\lambda)\, z^k$$

has only non-negative coefficients. That is $q_k(\lambda) \geq 0$ for all k.

Set $w := \lambda[1 - z]$, then $G(w)$ is analytic for all w with $|\lambda - w| < \lambda$ because $Q_\lambda(z) = G(\lambda[1 - z])$ is an analytic function for all $|z| < 1$. Thus, every $G \in \mathcal{G}$ is a function which is analytic inside a circle with center λ and radius λ. From the above definition, it is clear that $Q_\lambda(x) = G(\lambda[1-x]) \geq 0$ for all $x \in [-1, 1)$ whenever $G \in \mathcal{G}$. Moreover, the function $Q_\lambda(x)$ is monotone increasing in x because $Q'_\lambda(x) \geq 0$. Consequently, the function $G(x)$ is monotone decreasing in x, since $\mathrm{d}Q_\lambda/\mathrm{d}x = -\lambda\,\mathrm{d}G/\mathrm{d}x$.

Example 10.34. The function $G(z) = 1/z$ belongs to $\mathcal{G}$, since

$$Q_\lambda(z) = \tfrac{1}{\lambda}\,\tfrac{1}{1-z} = \tfrac{1}{\lambda}\textstyle\sum_{k=0}^{\infty} z^k \ .$$

Also the function $G(z) = \log(1/z)$ belongs to $\mathcal{G}$, since

$$Q_\lambda(z) = \log \tfrac{1}{\lambda(1-z)} = \log \tfrac{1}{\lambda} + \textstyle\sum_{k=1}^{\infty} \tfrac{1}{k}\, z^k \ .$$

Since every function $G \in \mathcal{G}$ has a power series representation, it is clear what is meant by the expression $G(\phi)$ for a certain $\phi \in \mathcal{W}$ (cf. Section 3.1.2). Therewith, can can formulate the following lemma.

Lemma 10.35. *Let $G \in \mathcal{G}$ and let c_0 be a constant with $1/2 < c_0 < 1$. Then*

$$\|G(\phi)\|_{\mathcal{W}} \leq G(2\,c_0 - 1) \tag{10.45}$$

for all $\phi \in \mathcal{M}_W(c_0)$. Moreover for the function $\phi_0(z) := c_0 - (1-c_0)\,z$, which belongs to $\mathcal{M}_W(c_0)$, even equality holds, i.e.

$$\|G(\phi_0)\|_{\mathcal{W}} = G(\phi_0(1)) = G(2\,c_0 - 1) \ .$$

Proof. Let $\hat{\phi}(0)$ be the zeroth Fourier coefficient of the function $\phi \in \mathcal{M}_W(c_0)$. Without loss of generality, it can be assumed that $\hat{\phi}(0)$ is real and positive. Otherwise, if $\hat{\phi}(0) = |\hat{\phi}(0)|\,\mathrm{e}^{\mathrm{i}\alpha_0}$, we would consider the function $\phi\,\mathrm{e}^{-\mathrm{i}\alpha_0}$ without any further change. Now $G(\phi)$ is written as

$$G(\phi(z)) = G\left(\hat{\phi}(0)\left[1 - \tfrac{\hat{\phi}(0)-\phi(z)}{\hat{\phi}(0)}\right]\right) = \sum_{k=0}^{\infty} g_k[\hat{\phi}(0)] \left(\tfrac{\hat{\phi}(0)-\phi(z)}{\hat{\phi}(0)}\right)^k$$

wherein all coefficients $g_k[\hat{\phi}(0)]$ are positive since $G \in \mathcal{G}$. Using the triangle inequality and the submultiplicative condition of the Banach algebra, the Wiener norm of $G(\phi)$ is upper bounded by

$$\|G(\phi)\|_{\mathcal{W}} \leq \sum_{k=0}^{\infty} g_k[\hat{\phi}(0)] \left\| \left(\tfrac{\hat{\phi}(0)-\phi}{\hat{\phi}(0)}\right)^k \right\|_{\mathcal{W}} \leq \sum_{k=0}^{\infty} g_k[\hat{\phi}(0)] \left(\tfrac{\|\hat{\phi}(0)-\phi\|_{\mathcal{W}}}{\hat{\phi}(0)}\right)^k \ .$$

Since $\|\phi\|_{\mathcal{W}} \leq 1$, we hae $\|\hat{\phi}(0) - \phi(\cdot)\|_{\mathcal{W}} = \sum_{k\neq 0} |\phi_k| \leq 1 - |\hat{\phi}(0)|$. Therefore, the upper bound becomes

$$\|G(\phi)\|_{\mathcal{W}} \leq \sum_{k=0}^{\infty} g_k[\hat{\phi}(0)] \left(\tfrac{1-\hat{\phi}(0)}{\hat{\phi}(0)}\right)^k = G\left(\hat{\phi}(0)\left[1 - \tfrac{1-\hat{\phi}(0)}{\hat{\phi}(0)}\right]\right) = G\left(2\hat{\phi}(0) - 1\right) \ .$$

Moreover, since it was assumed that $\hat{\phi}(0) > 0$ and that $\phi \in \mathcal{M}_W(c_0)$, we have $\hat{\phi}(0) \geq c_0$ and therefore $2\,\hat{\phi}(0) - 1 \geq 2\,c_0 - 1$. Therewith and using that $G(x)$ is a monotone decreasing function, the upper bound (10.45) follows from the last inequality.

Next, the function $\phi_0(z)$ is considered.

$$G(\phi_0(z)) = G\big(c_0\big[1 - \big(\tfrac{1-c_0}{c_0}\big)z\big]\big) = \sum_{k=0}^{\infty} g_k(c_0)\big(\tfrac{1-c_0}{c_0}\big)^k z^k \ .$$

This shows that $G(\phi_0) \in \mathcal{W}_+$ and therefore its norm becomes

$$\|G(\phi_0)\|_{\mathcal{W}} = \sum_{k=0}^{\infty} g_k[c_0] \left(\tfrac{1-c_0}{c_0}\right)^k = G(\phi_0(1)) = G(2\,c_0 - 1) \ ,$$

which is indeed the equality in (10.45). □

Since $\mathcal{M}_+(c_0), \mathcal{M}_{\mathbb{R}}(c_0) \subset \mathcal{M}_W(c_0)$, the upper bound (10.45) holds also for all functions in the subsets $\mathcal{M}_+(c_0)$ and $\mathcal{M}_{\mathbb{R}}(c_0)$. However, the function ϕ_0, defined in Lemma 10.35, is not an element of $\mathcal{M}_+(c_0)$ or $\mathcal{M}_{\mathbb{R}}(c_0)$ since $\phi_0(1) = 2\,c_0 - 1 < c_0$. Therefore, it is not clear at the outset whether the upper bound (10.45) is also sharp for $\mathcal{M}_+(c_0)$ and $\mathcal{M}_{\mathbb{R}}(c_0)$. The next lemma will prove that the bound (10.45) is also sharp for the set $\mathcal{M}_+(c_0)$. The question whether the bound (10.45) is sharp for $\mathcal{M}_{\mathbb{R}}(c_0)$ is still open.

Lemma 10.36. *Let $G \in \mathcal{G}$ and let c_0 be a constant with $1/2 < c_0 < 1$. Then*

$$\sup_{\phi \in \mathcal{M}_+(c_0)} \|G(\phi)\|_{\mathcal{W}} = G(2\,c_0 - 1) \ .$$

There exists in general no function in $\mathcal{M}_+(c_0)$ for which the supremum is attained. However, the upper bound can be achieved arbitrarily close by functions from $\mathcal{M}_+(c_0)$. To prove this lemma an auxiliary result is needed.

Lemma 10.37. *Let $\mu > 0$ and $0 < \delta < 1$ arbitrary, and let $b_1, b_2, \cdots, b_N \geq 0$ be arbitrary positive numbers. Then there exists a function $g \in \mathcal{W}_+$ with $\|g\|_{\mathcal{W}} = 1$ and with $\|g\|_\infty \leq \mu$ such that*

$$\Big\| \sum_{k=0}^N b_k\, g^k \Big\|_{\mathcal{W}} \geq (1 - \delta) \sum_{k=0}^N b_k \ . \tag{10.46}$$

The proof of this lemma is given in the appendix. With this lemma, we are able to prove Lemma 10.36.

Proof (Lemma 10.36). Define $\lambda_0 := \frac{1}{c_0} - 1$, which implies $0 < \lambda_0 < 1$. Furthermore let $\mu > 0$, $\delta \in (0,1)$, and $N \in \mathbb{N}$ be given. At the beginning let $g \in \mathcal{W}_+$ with $\|g\|_{\mathcal{W}} = 1$ and $\|g\|_\infty \leq \mu$ arbitrary (that such functions exist follows from Lemma 10.37 or Lemma 10.44 below). Choose $\lambda \leq \lambda_0$ and consider the function

$$\phi_\lambda(z) = \tfrac{1}{1+\lambda}\left[1 - \lambda\, g(z)\right] \ . \tag{10.47}$$

Clearly, for this function holds that $\|\phi_\lambda\|_{\mathcal{W}} \leq 1$ and that

$$|\phi_\lambda(z)| \geq \frac{1 - \lambda\,|g(z)|}{1+\lambda} \geq \frac{1-\lambda\,\mu}{1+\lambda} \geq \frac{1}{1+\lambda_0} = c_0 \ , \quad \forall z \in \overline{\mathbb{D}} \ .$$

where the last inequality holds for sufficiently small μ since $\lambda < \lambda_0$. This shows that $\phi_\lambda \in \mathcal{M}_+(c_0)$.

Now $G(\phi_\lambda)$ is considered. Since $G \in \mathcal{G}$, it has the following series representation:

$$G(\phi_\lambda) = G\left(\tfrac{1}{1+\lambda}\left[1 - \lambda g\right]\right) = \textstyle\sum_{k=0}^{\infty} q_k(\lambda)\, \lambda^k g^k$$

with positive coefficients $q_k(\lambda)$, $k \in \mathbb{N}$. Note that the numbers $b_k := q_k(\lambda)\lambda^k$ depend only on λ and the given function G but not on g. Consequently we can apply Lemma 10.37, which shows that there exists a particular $g \in \mathcal{W}_+$ with the specified properties ($\|g\|_{\mathcal{W}} = 1$ and $\|g\|_\infty \leq \mu$) such that

$$\left\| \sum_{k=0}^{N} q_k(\lambda)\, \lambda^k\, g^k \right\|_{\mathcal{W}} \geq (1-\delta) \sum_{k=0}^{N} q_k(\lambda)\, \lambda^k \ . \tag{10.48}$$

From now on, g is assumed to be that function for which (10.48) holds. By the triangle inequality, we have

$$\|G(\phi_\lambda)\|_{\mathcal{W}} \geq \left|\ \left\|\textstyle\sum_{k=0}^{N} q_k(\lambda)\, \lambda^k\, g^k\right\|_{\mathcal{W}} - \left\|\textstyle\sum_{k=N+1}^{\infty} q_k(\lambda)\, \lambda^k\, g^k\right\|_{\mathcal{W}}\ \right|$$

Using the lower bound (10.48) for the first term on the right hand side and the upper bound

$$\left\|\textstyle\sum_{k=N+1}^{\infty} q_k(\lambda)\, \lambda^k g^k\right\|_{\mathcal{W}} \leq \textstyle\sum_{k=N+1}^{\infty} q_k(\lambda)\, \lambda^k\, \|g\|_{\mathcal{W}}^k \leq \textstyle\sum_{k=N+1}^{\infty} q_k(\lambda)\, \lambda^k$$

for the second term, one finally obtains

$$\|G(\phi_\lambda)\|_{\mathcal{W}} \geq (1-\delta) \textstyle\sum_{k=0}^{N} q_k(\lambda)\, \lambda^k - \textstyle\sum_{k=N+1}^{\infty} q_k(\lambda)\, \lambda^k \ .$$

Such a lower bound exists for arbitrary $\delta > 0$ and $N \in \mathbb{N}$. Therefore it follows for $N \to \infty$ that

$$\begin{aligned}\liminf_{N\to\infty} \sup_{g \in \Gamma(\mu,\delta,N)} \|G(\phi_\lambda)\|_{\mathcal{W}} &\geq \sum_{k=0}^{\infty} q_k(\lambda)\, \lambda^k = G\left(\tfrac{1-\lambda}{1+\lambda}\right) \\ &\geq G\left(\tfrac{1-\lambda_0}{1+\lambda_0}\right) = G(2\, c_0 - 1)\end{aligned}$$

using that $\lambda \leq \lambda_0$ and that $G(x)$ is monotone increasing in x. This last inequality together with (10.45) gives the statement of the theorem. □

It should be noted that the previous proof shows how the "worst case" function ϕ, for which $\|G(\phi)\|_{\mathcal{W}}$ becomes arbitrarily close to the upper bound $G(2\,c_0-1)$ given by Lemma 10.35, can be constructed. One has to choose λ close to λ_0 and define the function ϕ_λ (10.47). Then $G(\phi_\lambda)$ becomes arbitrarily close to $G(2\,c_0-1)$.

The following corollary is an immediate consequence of Lemma 10.35 for the special functions $G(z)=1/z$ and $G(z)=\log(1/z)$.

Corollary 10.38. *Let* $\phi\in\mathcal{M}_W(c_0)$ *with* $1/2<c_0\leq 1$*. Then* $\log\phi\in\mathcal{W}$ *and* $\phi^{-1}\in\mathcal{W}$ *with*

$$\|\log\phi\|_{\mathcal{W}}=\|\log(1/\phi)\|_{\mathcal{W}}\leq\log\left(\tfrac{1}{2\,c_0-1}\right)\qquad\textit{and}\qquad\left\|\phi^{-1}\right\|_{\mathcal{W}}\leq\tfrac{1}{2\,c_0-1}\;.$$

Since $\mathcal{M}_+(c_0)\subset\mathcal{M}_W(c_0)$ this corollary holds also for all $\phi\in\mathcal{M}_+(c_0)$ with $c_0>1/2$. Moreover, Lemma 10.36 shows that the upper bounds given in this corollary are also sharp for the smaller set $\mathcal{M}_+(c_0)$, i.e. there exist functions $\phi\in\mathcal{M}_+(c_0)$ for which $\|\log\phi\|_{\mathcal{W}}$ and $\|\phi^{-1}\|_{\mathcal{W}}$ are arbitrary close to the bounds in Corollary 10.38, respectively.

Boundedness behavior

Based on the previous results, we will now characterize subsets of $\mathcal{W}$ on which the spectral factorization mapping $\mathfrak{S}$ is uniformly bounded. Since $\mathcal{W}$ is a decomposing $\mathcal{S}$–algebra, the spectral factorization exists in $\mathcal{W}$ for all $\phi\in\exp(\mathcal{W})$, which includes all strictly positive functions on $\mathbb{T}$, i.e. the set $\mathcal{M}_{\mathbb{R}}(c_0)$.

The following theorem gives a lower and upper bound for the Wiener norm of the spectral factors $\phi_\pm$ as a function of the minimum c_0.

Theorem 10.39. *Let* $\phi\in\mathcal{M}_{\mathbb{R}}(c_0)$ *with* $1/2<c_0<1$*. Then*

$$\sqrt{\|\phi\|_{\mathcal{W}}}\;\leq\;\|\phi_+\|_{\mathcal{W}}=\|\phi_-\|_{\mathcal{W}}\;\leq\;\frac{1}{\sqrt{2c_0-1}}\;. \tag{10.49}$$

Proof. The lower bound follows from the relations $\|\phi\|_{\mathcal{W}}\leq\|\phi_+\|_{\mathcal{W}}\,\|\phi_-\|_{\mathcal{W}}$ and $\|\phi_+\|_{\mathcal{W}}=\|\phi_-\|_{\mathcal{W}}$.

Since $\phi(\zeta)>1/2$ for all $\zeta\in\mathbb{T}$ it follows that $\hat{\phi}(0)>1/2$ and Corollary 10.38 shows that $\log\phi\in\mathcal{W}$. Let

$$(\log\phi)(\zeta)=\textstyle\sum_{k=-\infty}^{\infty}a(k)\,\zeta^k\;,\qquad\zeta\in\mathbb{T}$$

be the Fourier series of $\log\phi$ with the Fourier coefficients $a(k)$. Since $\log\phi$ is real valued, it holds $a(-k)=\overline{a(k)}$ for all k. Therewith the two functions

$$g_+(\zeta):=\tfrac{a(0)}{2}+\textstyle\sum_{k=1}^{\infty}a(k)\,\zeta^k\quad\text{and}\quad g_-(\zeta):=\tfrac{a(0)}{2}+\textstyle\sum_{k=1}^{\infty}a(-k)\,\zeta^k$$

are defined such that $\log\phi=g_++g_-$. Obviously $\|g_+\|_{\mathcal{W}}=\|g_-\|_{\mathcal{W}}$ and $\|\log\phi\|_{\mathcal{W}}=\|g_+\|_{\mathcal{W}}+\|g_-\|_{\mathcal{W}}$. Consequently, one obtains that

$$\|g_+\|_{\mathcal{W}} = \tfrac{1}{2}\,\|\log\phi\|_{\mathcal{W}} \ . \tag{10.50}$$

With the two functions g_+ and g_-, the spectral factors become $\phi_\pm = \exp(g_\pm)$, and by the boundedness of the exponential function (3.14) follows $\|\phi_\pm\|_{\mathcal{W}} \leq \exp\left(\|g_\pm\|_{\mathcal{W}}\right)$. With (10.50) and Corollary 10.38 the upper bound (10.49) finally follows. □

It is important to note that this theorem gives an upper bound only in the case that ϕ never becomes smaller than $c_0 > 1/2$ on $\mathbb{T}$. For $c_0 \to 1/2$ the upper bound in (10.49) goes to infinity and it is not clear whether there exists a corresponding upper bound on the norm of the spectral factors if $c_0 \leq 1/2$.

Continuity behavior

Since $\mathcal{W}$ is decomposing, Corollary 10.22 implies that the spectral factorization mapping $\mathfrak{S}$ is continuous on $\mathcal{W}$. Moreover, Theorem 10.21 gives an explicit expression (10.13) for the continuity constant $C(\phi)$ of the spectral factorization mapping $\mathfrak{S}$. This continuity constant $C(\phi)$ depends on the actual spectrum ϕ, which means that $\mathfrak{S}$ is not uniformly continuous on $\mathcal{W}$, in general. In particular, $C(\phi)$ depends on $\|\phi^{-1}\|_{\mathcal{W}}$ and $\|\mathfrak{S}\phi\|_{\mathcal{W}}$. However, Corollary 10.38 and Theorem 10.39 gave uniform upper bounds on $\|\phi^{-1}\|_{\mathcal{W}}$ and $\|\mathfrak{S}\phi\|_{\mathcal{W}}$, respectively, provided that the spectrum ϕ belongs to $\mathcal{M}_{\mathbb{R}}(c_0)$ with $c_0 > 1/2$. Therewith, we obtain the following corollary of Theorem 10.21.

Corollary 10.40. *Let $\phi, \psi \in \exp(\mathcal{W})$ be real valued functions with the spectral factors $\phi_+ = \mathfrak{S}\phi$ and $\psi_+ = \mathfrak{S}\psi$. Then*

$$\|\phi_+ - \psi_+\|_{\mathcal{W}} \leq 3\,\|\phi^{-1}\|_{\mathcal{W}}\,\|\phi_+\|_{\mathcal{W}}\,\|\phi - \psi\|_{\mathcal{W}} \ .$$

for all $\phi, \psi \in \exp(\mathcal{W})$ with $\|\phi - \psi\|_{\mathcal{W}} < \frac{1}{3}\,\|\phi^{-1}\|_{\mathcal{W}}^{-1}$.

If even $\phi, \psi \in \mathcal{M}_{\mathbb{R}}(c_0)$ with $1/2 < c_0 < 1$, then

$$\|\phi_+ - \psi_+\|_{\mathcal{W}} \leq \frac{3}{(2\,c_0 - 1)^{3/2}}\,\|\phi - \psi\|_{\mathcal{W}}$$

for all $\phi, \psi \in \mathcal{M}_{\mathbb{R}}(c_0)$ with $\|\phi - \psi\|_{\mathcal{W}} < \frac{1}{3}(2\,c_0 - 1)$.

Proof. The first part is just the specialization of Theorem 10.21 to the Wiener algebra, using that $\|\mathfrak{P}_+\|_{\mathcal{W}\to\mathcal{W}_+} = 1$ and $\|e\|_{\mathcal{W}} = 1$. The second part follows by applying Corollary 10.38 and Theorem 10.39 to the first part. □

Altogether, on the subset $\mathcal{M}_{\mathbb{R}}(c_0)$ of the Wiener algebra with $c_0 > 1/2$, the spectral factorization mapping is uniformly bounded (Theorem 10.39) and uniformly continuous (Corollary 10.40). Both, the upper bound on the norm and the continuity constant depend only on the minimum c_0 of the spectra.

10.4 Error Bounds for Polynomial Data

The previous section showed that the spectral factorization is either unbounded or discontinuous on $\mathcal{S}$–algebras. In many practical applications however, it is assumed that the given spectra are polynomials of a certain finite degree N. Since the set $\mathcal{P}(N)$ of all polynomials with a degree not larger than N is finite dimensional, it is clear that the spectral factorization is always bounded and continuous for all spectral densities in $\mathcal{P}(N)$. However, it is also clear that then either the continuity constant or the boundedness constant will depend strongly on the degree N of the polynomials and will go to infinity as the degree N goes to infinity. This dependency of the continuity and boundedness constant will be investigated in the present section.

In the first part, we consider the boundedness behavior of $\mathfrak{S}$ on the Wiener algebra $\mathcal{W}$. Since the concrete dependency on the degree N is influenced by the norm in the algebra, we do not investigate the boundedness behavior for decomposing Banach algebras, in general, but only for the most important example of such an algebra, the Wiener algebra. In the second part, the continuity behavior of $\mathfrak{S}$ with respect to the supremum norm for all polynomials is investigated in detail.

10.4.1 Factorization in the Wiener norm

We consider the spectral factorization mapping $\mathfrak{S}$ for polynomial spectral densities $\phi \in \mathcal{P}_{\text{pos}}(N)$ in the Wiener algebra. By the theorem of Fejér-Riesz, the spectral factorization exists for non-negative polynomials and the spectral factor $\mathfrak{S}\phi \in \mathcal{P}_{+}(N)$ is a causal polynomial of degree N. It follows immediately that $\mathfrak{S}$ is bounded on $\mathcal{P}_{\text{pos}}(N)$, i.e. there exists a constant $C(N) < \infty$ such that

$$\|\mathfrak{S}\phi\|_{\mathcal{W}}^2 \leq C(N)\, \|\phi\|_{\mathcal{W}} \qquad \text{for all} \quad \phi \in \mathcal{P}_{\text{pos}}(N)\,. \tag{10.51}$$

Definition 10.41. *Let $\mathfrak{S} : \mathcal{W}_{pos} \to \mathcal{W}_{+}$ be the spectral factorization mapping given by* (10.8)*. Then the constant*

$$C(N) := \sup_{\substack{\phi \in \mathcal{P}_{pos}(N) \\ \|\phi\|_{\mathcal{W}}=1}} \|\mathfrak{S}\phi\|_{\mathcal{W}}^2 \tag{10.52}$$

is called the boundedness constant *of the spectral factorization mapping for polynomials of degree N on $\mathcal{W}$.*

Since $\mathfrak{S}$ is unbounded on $\mathcal{W}$, the boundedness constant $C(N)$ will depend strongly on the degree N of the given spectra and it will go to infinity as $N \to \infty$. We want to investigate how $C(N)$ depends on N. In particular, we are going to show (Theorem 10.49) that the boundedness constant $C(N)$ is lower and upper bounded by

$$C_1 \sqrt{N+1} \;\leq\; C(N) \;\leq\; N+1$$

with a constant C_1 independent of N. Therein the upper bound is easily derived, whereas the proof of the lower bound is somewhat intricate. Therefore, we give a short outline of the main ideas and steps of this proof. Since $\mathcal{W}$ is a decomposing $\mathcal{S}$–algebra, the spectral factorization mapping is unbounded on $\mathcal{W}$. In view of relation (10.51) and since the trigonometric polynomials are dense in $\mathcal{W}$, this unboundedness of $\mathfrak{S}$ implies that there exists a sequence $\{\phi_N\}_{N\in\mathbb{N}}$ of polynomial spectral densities with $\lim_{N\to\infty} \|\phi_N\|_{\mathcal{W}} = 0$ but for which the norm of spectral factors $(\phi_N)_+$ are uniformly lower bounded by a positive constant $c_0 > 0$, i.e. for which

$$\|\mathfrak{S}\phi_N\|_{\mathcal{W}} = \|(\phi_N)_+\|_{\mathcal{W}} \geq c_0 > 0 \qquad \text{for all } N \in \mathbb{N}\,.$$

If such a sequence is known, a lower bound for the boundedness constant is obtained from (10.51) by

$$C(N) \geq \frac{\|(\phi_N)_+\|^2_{\mathcal{W}}}{\|\phi_N\|_{\mathcal{W}}} = \frac{\|(\phi_N)_+\|^2_{\mathcal{W}}}{\|(\phi_N)_+\,(\phi_N)_+^\star\|_{\mathcal{W}}}\,.$$

To obtain as tight a bound as possible, one needs such a sequence for which $\|\phi_N\|_{\mathcal{W}}$ converges to zero as fast as possible as $N \to \infty$.

For these reasons, we will start our investigations by constructing a sequence of polynomials $\{g_N\}_{N\in\mathbb{N}}$ with $g_N \in \mathcal{P}_+(N)$, with $\|g_N\|_{\mathcal{W}} = 1$, and for which $\|g_N\, g_N^\star\|_{\mathcal{W}}$ converges to zero as $N \to \infty$ with almost the fastest possible convergence rate. However, it turns out that these polynomials might not be spectral factors of $g_N\, g_N^\star$, in general. Therefore, these polynomials g_N will be modified such that they belong to the set $\mathcal{G}(\mathcal{W}_+)$ of invertible elements of $\mathcal{W}_+$. This will guarantee that they are spectral factors and still have the desired properties. With these modified polynomials, the lower bound of the boundedness constant $C(N)$ is proved in Theorem 10.49.

Functions with small peak value

Consider the following problem. We look for trigonometric polynomials $\phi_N \in \mathcal{P}(N)$ of degree N with constant Wiener norm $\|\phi_N\|_{\mathcal{W}} = 1$ (independent of the degree N) but with the property that the peak value $\|\phi_N\|_\infty$ becomes as small as possible. The first lemma derives a lower bound, on the minimal achievable peak value of polynomials $\phi_N \in \mathcal{P}(N)$.

Lemma 10.42. *Let $\phi \in \mathcal{P}(N)$ be a trigonometric polynomial of degree N. Then its peak value is bounded by*

$$\frac{1}{\sqrt{2N+1}}\,\|\phi\|_{\mathcal{W}} \;\leq\; \|\phi\|_\infty \;\leq\; \|\phi\|_{\mathcal{W}}\,. \tag{10.53}$$

Proof. Let $\phi \in \mathcal{P}(N)$. Then by the Cauchy-Schwarz inequality and by Parseval's identity it holds that

$$\begin{aligned}\|\phi\|_{\mathcal{W}} = \sum_{k=-N}^{N} |\hat{\phi}(k)| &\leq \sqrt{\sum_{k=-N}^{N} 1}\sqrt{\sum_{k=-N}^{N} |\hat{\phi}(k)|^2} \\ &= \sqrt{2N+1}\,\|\phi_+\|_2 \leq \sqrt{2N+1}\,\|\phi_+\|_\infty \ ,\end{aligned}$$

which gives the lower bound in (10.53). The upper bound is just a consequence of the continuous embedding of $\mathcal{W}$ in $\mathcal{C}(\mathbb{T})$ (cf. Proposition 10.14). □

Remark 10.43. Of course, if one considers analytic polynomials $\phi(\mathrm{e}^{\mathrm{i}\omega}) = \sum_{k=0}^{N} \hat{\phi}(k)\,\mathrm{e}^{\mathrm{i}k\omega}$ in $\mathcal{P}_+(N)$, the lower bound in (10.53) becomes $\|\phi\|_\infty \geq \frac{\|\phi\|_{\mathcal{W}}}{\sqrt{N+1}}$.

Thus for a fixed degree N, the peak value of a polynomial $\phi_N \in \mathcal{P}(N)$ with $\|\phi_N\|_{\mathcal{W}} = 1$ cannot be made arbitrary small, since the above lemma shows that there exists no such polynomial with a peak value smaller than $1/\sqrt{2N+1}$. However, as the degree N of the polynomial ϕ_N increases, the lower bound (10.53) decreases and gets arbitrary small as $N \to \infty$. The following lemma gives concrete polynomials g_N with a constant norm $\|g_N\|_{\mathcal{W}}$ and with small peak value $\|g_N\|_\infty$. In particular, it is shown that these polynomials achieve almost the lower bound (10.53) for their peak value.

Lemma 10.44. *Let $g_N \in \mathcal{P}_+(N)$ be defined by*

$$g_N(\mathrm{e}^{\mathrm{i}\omega}) = \frac{1}{N+1}\sum_{k=0}^{N} \exp\left(\mathrm{i}\frac{k^2\pi}{N+1}\right)\mathrm{e}^{\mathrm{i}k\omega} \ , \qquad \omega \in [-\pi, \pi) \ . \tag{10.54}$$

Then $\|g_N\|_{\mathcal{W}} = 1$ and there exists a constant C_2 such that

$$\|g_N\|_\infty \leq \frac{C_2}{\sqrt{N+1}} \tag{10.55}$$

for all $N \in \mathbb{N}$ with $N \geq N_0$.

The polynomials (10.54) are also known as *chirp sequences*. Originally they were obtained in search for signals with a flat power spectrum and with a low peak value [80]. Moreover, according to (10.55) these polynomials achieve almost the fastest possible convergence rate (10.53) as $N \to \infty$. Only the constant C_2 is not optimal, i.e. $C_2 > 1$.

To verify (10.55), one has to find an upper bound on the modulus of an exponential sum of the form

$$S[q_\omega; a, b] := \sum_{k=a}^{b} \exp(\mathrm{i}\,2\pi\, q_\omega(k)) \ . \tag{10.56}$$

In our case, the particular function $q_\omega(k)$ is given by

$$q_\omega(k) = \frac{k^2}{2(N+1)} + \frac{k\omega}{2\pi}\,, \qquad \omega \in [-\pi,\pi]\,, \quad k = 0,1,\ldots,N \quad (10.57)$$

and the bounds of the summation index are $a = 0$ and $b = N$. Here ω is looked upon a fixed parameter and one has to find an upper bound on $|S[q_\omega; a,b]|$ for all parameters $\omega \in [-\pi,\pi)$. To this end, we apply an extension of a technique which finds an upper bound on the corresponding integral

$$I[q_\omega; a,b] := \int_a^b \exp(\mathrm{i}\,2\pi\, q_\omega(\tau))\,\mathrm{d}\tau\,. \qquad (10.58)$$

By a result of van der Corput, it is possible to control the difference $D[q_\omega; a,b] := I[q_\omega; a,b] - S[q_\omega; a,b]$ between the sum (10.56) and the integral (10.58), under some conditions on the function q_ω.

Lemma 10.45 (van der Corput). *If $q'_\omega(k)$ is monotone and if $|q'_\omega(k)| \leq 1-\varepsilon$ for some $\varepsilon > 0$ and for all $k \in [a,b]$, then*

$$|D[q_\omega; a,b]| \leq C_3 \tfrac{1}{\varepsilon} + C_4 \qquad (10.59)$$

with $C_3 = 4/\pi$ and $C_4 = 1 + 4/\pi$.

This lemma due to van der Corput [86] is taken from [92, Chapter V, Lemma 4.4] where also a proof can be found. As a further preparation, we derive an upper bound on the modulus of the integral $I[q_\theta; a,b]$. This is done in the next lemma.

Lemma 10.46. *Let $I[q_\omega; a,b]$ be given by (10.58) with the function q_ω defined by (10.57). Then there exists a constant C_5 such that*

$$|I[q_\omega; a,b]| \leq C_5\sqrt{N+1}$$

for all ω and for arbitrary integration bounds a and b.

Proof. To see this, one only has to write the integral in another form. First, the function q_ω in the exponent of the integral is written as

$$q_\omega(\tau) = \left(\frac{\tau}{\sqrt{2(N+1)}} + \sqrt{\frac{N+1}{2}}\,\frac{\omega}{2\pi}\right)^2 - \frac{N+1}{2}\,\frac{\omega^2}{(2\pi)^2}\,.$$

Since the last term is independent of the integration variable τ, the modulus of the integral becomes

$$\begin{aligned} |I(q_\omega; a,b)| &= \left|\int_a^b \exp\left(\mathrm{i}\,2\pi\left[\tfrac{\tau}{\sqrt{2(N+1)}} + \sqrt{\tfrac{N+1}{2}}\,\tfrac{\omega}{2\pi}\right]^2\right)\mathrm{d}\tau\right| \\ &= \sqrt{\tfrac{N+1}{2\pi^2}}\left|\int_{D(a,\omega)}^{D(b,\omega)} \exp(\mathrm{i}\tfrac{\pi}{2}x^2)\,\mathrm{d}x\right| \qquad (10.60)\end{aligned}$$

in which the integration limits are given by the function

$$D(\gamma,\omega) = \sqrt{\tfrac{2}{N+1}}\,\gamma + \sqrt{\tfrac{N+1}{2}}\,\tfrac{\omega}{\pi}\ .$$

It remains to investigate the integral in (10.60). For simplicity, it is denoted by $Q(\alpha,\beta) := \int_\alpha^\beta \exp(\mathrm{i}\frac{\pi}{2}x^2)\,\mathrm{d}x$ with the two real numbers $\alpha = D(a,\omega)$ and $\beta = D(b,\omega)$. First, we note that $Q(0,\xi) = F_C(\xi) + \mathrm{i}\,F_S(\xi)$ in which

$$F_C(\xi) = \int_0^\xi \cos(\tfrac{\pi}{2}\,x^2)\,\mathrm{d}x \qquad \text{and} \qquad F_S(\xi) = \int_0^\xi \sin(\tfrac{\pi}{2}\,x^2)\,\mathrm{d}x$$

are the so called *Fresnel integrals*. And for all real number $\xi \in \mathbb{R}$ holds that

$$|F_C(\xi)| \le |F_C(1)| \approx 0.779 \qquad \text{and} \qquad |F_S(\xi)| \le |F_S(\sqrt{2})| \approx 0.714\ .$$

Therefore, one has that

$$|Q(0,\xi)| \le \sqrt{|F_C(\xi)|^2 + |F_S(\xi)|^2} \le \sqrt{|F_C(1)|^2 + |F_S(\sqrt{2})|^2} =: C_{11}$$

for all $\xi \in \mathbb{R}$. Consequently, since $Q(\alpha,\beta) = Q(0,\beta) - Q(0,\alpha)$, one obtains that $|Q(\alpha,\beta)| \le 2\,C_{11}$ for arbitrary $\alpha,\beta \in \mathbb{R}$. Together with (10.60) this gives the statement of Lemma 10.46 with $C_5 = \sqrt{2}\,C_{11} \approx 1.5$. □

Note that the upper bound on $|I[q_\omega;a,b]|$ given by Lemma 10.46 is independent of the integration bounds a and b. After these preparations, we are able to prove Lemma 10.44.

Proof (Lemma 10.44). We consider the functions q_ω, $S[q_\omega;a,b]$, and $I[q_\omega;a,b]$ as defined above. Therewith the function g_N, defined by (10.54), can be written as

$$g_N(\mathrm{e}^{\mathrm{i}\omega}) = \frac{1}{N+1}\,S[q_\omega;0,N]\ . \tag{10.61}$$

Assume for the moment that q_ω satisfies the conditions of Lemma 10.45. Then by applying Lemma 10.45, one obtains an upper bound for the exponential sum

$$|S[q_\omega;0,N]| \le |I[q_\omega;0,N]| + C_3\,\tfrac{1}{\varepsilon} + C_4 \tag{10.62}$$

and Lemma 10.46 shows that there exists a constant C_5 such that $|I[q_\omega;a,b]| \le C_5\,\sqrt{N+1}$ for arbitrary integration limits a and b.

It remains to verify whether the function q_ω satisfies the conditions of Lemma 10.45 and to determine the corresponding ε. Since $q'_\omega(k) = \frac{k}{N+1} + \frac{\omega}{2\pi}$, it is not hard to see that $|q'_\omega(k)|$ becomes larger than 1 for certain parameters $\omega \in [-\pi,\pi)$ and some $k \in [0,N]$. Nevertheless, since the polynomials $g_N(\mathrm{e}^{\mathrm{i}\omega})$ are 2π-periodic, one can consider the problem not only for $\omega \in [-\pi,\pi)$ but equivalently for $\omega + 2\pi\,n$ with an arbitrary integer n. This property is used, and the integration $I[q_\omega; 0, N]$ over the interval $[0, N]$ in (10.62) is split into integrations over the intervals $I_1 = [0, N/2]$ and $I_2 = [N/2, N]$. Then it can be shown that for each $i \in \{1, 2\}$ there exists an $n \in \mathbb{Z}$ such that for $\omega_i = \omega + 2\pi\, n$ always

$$|q'_{\omega_i}(k)| = \left| \frac{k}{N+1} + \frac{\omega}{2\pi} + n \right| \leq 1 - \tfrac{1}{8} \qquad \text{for each } k \in I_i\,. \tag{10.63}$$

Indeed, one easily verifies that for $i = 1$ and if

$$-\frac{7}{8} - \frac{\omega}{2\,\pi} \leq n \leq \frac{3}{8} - \frac{\omega}{2\,\pi}$$

then (10.63) is always satisfied. Similarly in the case $i = 2$, (10.63) is fulfilled if n satisfies

$$-\frac{9}{8} - \frac{\omega}{2\,\pi} \leq n \leq \frac{1}{8} - \frac{\omega}{2\,\pi}\,.$$

This proves for both cases ($i = 1, 2$) that there exists an integer $n \in \mathbb{Z}$ such that (10.63) holds. This means that Lemma 10.45 is satisfied for each interval I_1 and I_2 with $\epsilon = 1/8$, separately. Therefore, we can combine (10.61) and (10.62) and apply Lemma 10.46 to obtain

$$\begin{aligned} |g_N(\mathrm{e}^{\mathrm{i}\omega})| &\leq \frac{|I[q_{\omega_1}; 0, N/2]| + |I[q_{\omega_2}; N/2, N]| + 8\,C_3 + C_4}{N+1} \\ &\leq \frac{2\,C_5}{\sqrt{N+1}} + \frac{8\,C_3 + C_4}{N+1}\,. \end{aligned}$$

Since this bound is independent of ω, it proves (10.55). □

The polynomials g_N, defined in Lemma 10.44, have the property that the Wiener norm $\|g_N\|_{\mathcal{W}} = 1$ is independent of the degree N, whereas the peak value of these polynomials is upper bounded by $C_2/\sqrt{N+1}$ and decreases with increasing degree N. Fig. 10.1 illustrates this behavior for three different degrees N. Therewith, we have solved the problem of finding a function $g \in \mathcal{W}$ with $\|g\|_{\mathcal{W}} = 1$ and with arbitrarily small peak value. By Lemma 10.44, for every $\varepsilon > 0$ there exists a degree $N \in \mathbb{N}$ such that the polynomial (10.54) satisfies $\|g_N\|_\infty < \varepsilon$ but $\|g_N\|_{\mathcal{W}} = 1$.

Subsequently, it will be important that the Wiener norm of the polynomial $g_N\, g_N^\star$ tends to zero as $N \to \infty$. As the proof of the following lemma will show, this is a direct consequence of the decreasing peak value of g_N as $N \to \infty$.

Lemma 10.47. *Let g_N be the polynomial defined by* (10.54). *Then*

$$\|g_N\|_{\mathcal{W}} = 1 \qquad \textit{and} \qquad \|g_N\, g_N^\star\|_{\mathcal{W}} \leq \frac{3 + 2\log([N+1]/2)}{N+1}\,.$$

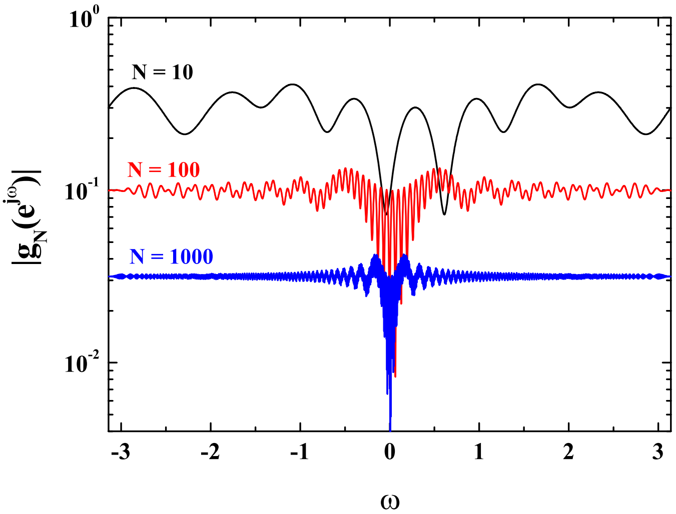

Fig. 10.1. The modulus of the polynomials g_N defined in (10.54) for different degrees N.

Proof. It remains to prove the second statement. On the unit circle $\mathbb{T}$, it holds that $g_N^\star(\mathrm{e}^{\mathrm{i}\omega}) = \overline{g_N(\mathrm{e}^{\mathrm{i}\omega})}$. We define $h_N \in \mathcal{P}(N)$ by

$$h_N(\mathrm{e}^{\mathrm{i}\omega}) := g_N(\mathrm{e}^{\mathrm{i}\omega})\,\overline{g_N(\mathrm{e}^{\mathrm{i}\omega})} = \textstyle\sum_{k=-N}^{N} \hat{h}_N(k)\,\mathrm{e}^{\mathrm{i}k\omega} \,. \tag{10.64}$$

Since $\hat{g}_N(k) = \frac{1}{N+1}\exp(\mathrm{i}\frac{k^2\pi}{N+1})$, $k = 0, 1, \cdots, N$ are the Fourier coefficients of g_N, a straight forward calculation shows that the Fourier coefficients of h_N are given by

$$\begin{aligned} \hat{h}_N(k) &= \sum_{l=0}^{N-k} \overline{\hat{g}_N(l)}\,\hat{g}_N(l+k) \\ &= \frac{\exp(\mathrm{i}\frac{\pi k^2}{N+1})}{(N+1)^2}\,\frac{1-\exp(-\mathrm{i}\frac{2\pi k^2}{N+1})\exp(\mathrm{i}\,2\pi k)}{1-\exp(\mathrm{i}\,\frac{2\pi k}{N+1})} \end{aligned}$$

and by $\hat{h}_N(-k) = \overline{\hat{h}_N(k)}$ for all $k = 0, 1, 2, \cdots, N$. For the modulus of these Fourier coefficients one obtains

$$|\hat{h}_N(k)| = \frac{1}{(N+1)^2}\left|\frac{\sin\left(\pi\,\frac{k^2-k[N+1]}{N+1}\right)}{-\sin\left(\frac{\pi k}{N+1}\right)}\right| \leq \frac{1}{(N+1)^2}\,\frac{1}{\left|\sin\left(\frac{\pi k}{N+1}\right)\right|}$$

for all $|k| = 1, 2, \cdots, N$ and $|\hat{h}_N(0)| = 1/(N+1)$. Therewith, the norm of h_N is given by $\|h_N\|_{\mathcal{W}} = \frac{1}{N+1} + 2\sum_{k=1}^{N} |\hat{h}_N(k)|$. Let L_N be the largest integer for which $L_N < [N+1]/2$ and note that $\sin(\pi x) \geq 2x$ for all $x \in [0, 1/2]$. Therewith, one obtains the upper bound

$$\begin{aligned}\|h_N\|_{\mathcal{W}} &\leq \frac{1}{N+1} + \frac{4}{(N+1)^2} \sum_{k=1}^{L_N} \frac{N+1}{2k} \\ &\leq \frac{1}{N+1} \left(3 + 2 \sum_{k=2}^{L_N} \frac{1}{k} \right) \leq \frac{3 + 2\log L_N}{N+1}\end{aligned}$$

for the norm of h_N, which is equivalent to the statement of the lemma, since $L_N \leq (N+1)/2$. □

By Lemma 10.47, $\|g_N\, g_N^{\star}\|_{\mathcal{W}}$ tends to zero as $N \to \infty$ with a convergence rate which is almost as fast as possible. Because for an arbitrary $\phi_N \in \mathcal{W}_+$, it always holds

$$\|\phi_N\, \phi_N^{\star}\|_{\mathcal{W}} \geq \|\phi_N\|_{\infty}^2 \geq \frac{\|\phi_N\|_{\mathcal{W}}^2}{N+1}$$

where the first inequality is a consequence of the continuous embedding of $\mathcal{W}$ in $\mathcal{C}(\mathbb{T})$ (cf. Proposition 10.14) and the second inequality follows from Lemma 10.42.

Bounds of the boundedness constant

The polynomials g_N defined in (10.54) are not spectral factors, because g_N may have zeros inside the unit disk $\mathbb{D}$. For this reason the sequence $\{g_N\}_{N\in\mathbb{N}}$ cannot be used directly for our intention, and we have to choose a slightly different approach. To this end, we define the functions

$$\varphi_N := g_N + \|g_N\|_{\infty} \qquad \text{and} \qquad \psi_N := \varphi_N\, \varphi_N^{\star}\ . \tag{10.65}$$

By this definition, it is clear that $\varphi_N \in \mathcal{W}_+$ and the maximum modulus principle implies that $\varphi_N(z) \neq 0$ for all $z \in \mathbb{D}$. Therefore $\varphi_N = (\psi_N)_+$ is a spectral factor of the trigonometric polynomial $\psi_N \in \mathcal{P}_{\text{pos}}(N)$. The next lemma gives upper bounds on the $\|\cdot\|_{\mathcal{W}}$-norm of ψ_N and φ_N.

Lemma 10.48. *Let $\varphi_N \in \mathcal{P}_+(N)$ and $\psi_N \in \mathcal{P}_{pos}(N)$ be the polynomials defined in* (10.65)*. Then there exist two constants C_8 and C_9 such that*

$$1 - \frac{C_8}{\sqrt{N+1}} \leq \|\varphi_N\|_{\mathcal{W}} \leq 1 + \frac{C_8}{\sqrt{N+1}}$$

$$\|\psi_N\|_{\mathcal{W}} \leq \frac{C_9}{\sqrt{N+1}}\ .$$

Proof. The first statement is a consequence of Lemma 10.44. By the definition of φ_N in (10.65) one has

$$\|\varphi_N\|_{\mathcal{W}} \le \|g_N\|_{\mathcal{W}} + \|g_N\|_\infty \le 1 + \tfrac{C_2}{\sqrt{N+1}}$$
$$\text{and} \quad \|\varphi_N\|_{\mathcal{W}} \ge |\ \|g_N\|_{\mathcal{W}} - \|g_N\|_\infty\ | \ge 1 - \tfrac{C_2}{\sqrt{N+1}}\ .$$

To prove the second statement, note that by the definition of ψ_N one has $\psi_N = g_N\, g_N^\star + (g_N + g_N^\star)\, \|g_N\|_\infty + \|g_N\|_\infty^2$. Applying Lemma 10.47 and 10.44 one obtains the upper bound

$$\|\psi_N\|_{\mathcal{W}} \le \frac{3 + 2\log([N+1]/2)}{N+1} + \frac{2\, C_2}{\sqrt{N+1}} + \frac{C_2^2}{N+1}\ ,$$

which proves the lemma with an adequate constant C_9. □

Lemma 10.48 shows that $\|\varphi_N\|_{\mathcal{W}}$ converges to 1 as $N \to \infty$ and that $\|\psi_N\|_{\mathcal{W}}$ converges to 0 as $N \to \infty$. Thus the sequences $\{\varphi_N\}_{N\in\mathbb{N}}$ and $\{\psi_N\}_{N\in\mathbb{N}}$ have all the desired properties which are needed to prove our main result on the behavior of the boundedness constant $C(N)$ of the spectral factorization mapping.

Theorem 10.49. *Let $\mathfrak{S} : \mathcal{W}_{pos} \to \mathcal{W}_+$ be the spectral factorization mapping on the Wiener algebra, and let $C(N)$ be the boundedness constant of $\mathfrak{S}$ defined in* (10.52)*. Then there exists a constant C_1, independent of N, such that*

$$C_1 \sqrt{N+1}\ \le\ C(N)\ \le\ N+1\,.$$

Proof. Upper bound: Let $\phi \in \mathcal{P}_{\mathrm{pos}}(N)$ be a spectral density and let $\phi_+ = \mathfrak{S}\phi$ its spectral factor. Lemma 10.42 implies that $\|\phi_+\|_{\mathcal{W}} \le \sqrt{N+1}\, \|\phi_+\|_\infty$. Since $\phi(\mathrm{e}^{\mathrm{i}\omega}) = |\phi_+(\mathrm{e}^{\mathrm{i}\omega})|^2$ for all $\omega \in [-\pi, \pi)$ it follows that $\|\phi_+\|_\infty = \|\phi\|_\infty^{1/2}$, and since $\mathcal{W}$ is continuously embedded in $\mathcal{C}(\mathbb{T})$ one obtains $\|\phi_+\|_{\mathcal{W}}^2 \le (N + 1)\, \|\phi\|_{\mathcal{W}}$, which shows that $C(N) \le N+1$.

Lower bound: We consider the polynomial $\psi_N \in \mathcal{P}(N)$ and its spectral factor $\varphi_N = \mathfrak{S}\, \psi_N \in \mathcal{P}_+(N)$, both defined in (10.65). By (10.51) and by the boundedness of $\mathfrak{S}$ on $\mathcal{P}(N)$ there exists a constant $C(N)$ such that $\|\varphi_N\|_{\mathcal{W}}^2 \le C(N)\, \|\psi_N\|_{\mathcal{W}}$. If one applies the lower and upper bound of φ_N and ψ_N, given by Lemma 10.48, one obtains the lower bound

$$C(N) \ge \frac{\sqrt{N+1}}{C_9} - 2\,\frac{C_8}{C_9} + \frac{C_8^2}{C_9}\frac{1}{\sqrt{N+1}}$$

for the boundedness constant. This is equivalent to the statement of the lemma with an appropriate constant C_1. □

Therewith, we have found the desired lower and upper bound on the boundedness constant (10.52) of the spectral factorization mapping in the Wiener algebra. At the end, we want to briefly review the above approach to prove Theorem 10.49 and give some explanatory remarks. The most important step was the construction of the sequence of polynomials $\{\varphi_N\}_{N\in\mathbb{N}}$ with $\|\varphi_N\|_{\mathcal{W}} \geq c_0 > 0$ for all $N \in \mathbb{N}$ and such that $\|\varphi_N\, \varphi_N^\star\|_{\mathcal{W}} \to 0$ as $N \to \infty$ as fast as possible. The construction of these functions was based on the polynomials g_N defined in (10.54). For the proof of Lemma 10.48, it was important that these polynomials have the property that $\|g_N\|_\infty$ as well as $\|g_N\, g_N^\star\|_{\mathcal{W}}$ converges fast to zero as $N \to \infty$. We saw that $\|g_N\, g_N^\star\|_{\mathcal{W}}$ has almost the best possible convergence behavior as $N \to \infty$. However, the proof of Lemma 10.48 shows that the upper bound of $\|\psi_N\|_{\mathcal{W}} = \|\varphi_N\, \varphi_N^\star\|_{\mathcal{W}}$ is mainly determined by the convergence behavior of $\|g_N\|_\infty$. For this convergence behavior a very simple upper bound can already be obtained in terms of Lemma 10.47 and by the continuous embedding of $\mathcal{W}$ in $\mathcal{C}(\mathbb{T})$ as follows

$$\frac{3 + 2\log([N+1]/2)}{N+1} \geq \|g_N\, g_N^\star\|_{\mathcal{W}} \geq \|g_N\|_\infty^2 \,.$$

Therewith, one gets immediately that

$$\|g_N\|_\infty \leq \sqrt{\frac{3 + 2\log([N+1]/2)}{N+1}} \,.$$

If this bound were used for the proof of Lemma 10.48, the lower bound in Theorem 10.49 would only be

$$\frac{\sqrt{N+1}}{\sqrt{3 + 2\log([N+1]/2)}} \leq C(N) \,.$$

For this reason, a better upper bound for $\|g_N\|_\infty$ was derived in Lemma 10.44.

Apart from that, the functions φ_N were obtained by adding $\|g_N\|_\infty$ to the function g_N. This was necessary because g_N itself was not an outer function. But due to this definition, $\|g_N\|_\infty$ determines the convergence behavior of $\|\varphi_N\, \overline{\varphi_N}\|_{\mathcal{W}}$, and since $\|g_N\|_\infty$ converges to zero slower than $\|g_N\, \overline{g_N}\|_{\mathcal{W}}$ (compare Lemma 10.44 and Lemma 10.47), this may indicate that the lower bound given in Theorem 10.49, is not yet the best possible lower bound. However, for the technique which we use for the proof (adding $\|g_N\|_\infty$ to g_N), the bound of Theorem 10.49 is the best possible bound.

Given a $\phi \in \mathcal{P}(N)$ with $\|\phi\|_{\mathcal{W}} = 1$, the upper bound of Theorem 10.49 shows that the Wiener norm of the spectral factor ϕ_+ never becomes larger than $\|\mathfrak{S}\phi\|_{\mathcal{W}} \leq \sqrt{N+1}$. On the other hand, the lower bound of Theorem 10.49 shows that for every degree N, there exist trigonometric polynomials $\phi \in \mathcal{P}(N)$ for which the norm of the spectral factor becomes larger than $\sqrt{C_1}\,(N+1)^{1/4}$. Both, upper and lower bound tend to infinity as $N \to \infty$, which shows in particular that the spectral factorization is unbounded on $\mathcal{W}$ since the trigonometric polynomials are dense in $\mathcal{W}$.

Corollary 10.50. *The spectral factorization is unbounded on the Wiener algebra $\mathcal{W}$.*

Of course, this result is a special case of Corollary 10.29. However, since Corollary 10.29 holds for every decomposing $\mathcal{S}$–algebra, the proof was rather abstract. The derivation here for the Wiener algebra, as a corollary of Theorem 10.49, gives concrete functions for which the factorization is unbounded and shows explicitly the growth behavior of the norm of the spectral factor $\mathfrak{S}\phi$ as the degree N of the spectral data increases.

10.4.2 Factorization in the infinity norm

This section considers the spectral factorization mapping $\mathfrak{S}$ on the space $\mathcal{P}(N)$ of all trigonometric polynomials with degree of at most N, in the infinity norm $\|\cdot\|_\infty$. It follows from Theorem 10.27 that the spectral factorization mapping is bounded in $\mathcal{C}(\mathbb{T})$. Therefore, it will be bounded a fortiori for all polynomials $\mathcal{P}(N)$ with respect to the infinity norm. Theorem 10.32, on the other hand showed that $\mathfrak{S}$ is discontinuous on $\mathcal{C}(\mathbb{T})$. Subsequently, we want to investigate how the continuity behavior of $\mathfrak{S}$ depends on the degree N of the polynomials. It turns out, that the continuity constants will also depend on the minimum and maximum of the polynomials under consideration. Therefore we will use the following notation: $\mathcal{P}_{\text{pos}}(N; c_1, c_2)$ denotes the set of all trigonometric polynomials $\phi \in \mathcal{P}(N)$ with $c_1 \le \phi(\mathrm{e}^{\mathrm{i}\theta}) \le c_2$ for all $\theta \in [-\pi, \pi)$. Throughout, and without loss of generality, it is always assumed that $c_1 < c_2$. For $c_1 = c_2$, the set $\mathcal{P}_{\text{pos}}(N; c_1, c_2)$ contains only the constant function $f(\mathrm{e}^{\mathrm{i}\theta}) = c_1$.

Assume that ϕ and ψ are two positive trigonometric polynomials of a certain degree N such that the difference between them is smaller than a certain value ε, i.e. $\|\phi - \psi\|_\infty < \varepsilon$. What is the difference $\|\phi_+ - \psi_+\|_\infty$ in the corresponding spectral factors? To answer this question, we are going to determine two constants C_{S1} and C_{S2} such that

$$C_{S1}\,\|\phi - \psi\|_\infty \;\le\; \|\phi_+ - \psi_+\|_\infty \;\le\; C_{S2}\,\|\phi - \psi\|_\infty \tag{10.66}$$

for all polynomials $\phi, \psi \in \mathcal{P}_{\text{pos}}(N; c_1, c_2)$. In general, both constants depend on the degree N and on the minimal and maximal values of the polynomials under consideration. The constant C_{S1} can easily be determined. Since $\phi = \phi_+\phi_+^\star$, the difference in the given spectra can be written as $\|\phi - \psi\|_\infty = \|\phi_+\phi_+^\star - \psi_+\psi_+^\star\|_\infty$ and some straight forward algebraic manipulations show that

$$\begin{aligned} \|\phi - \psi\|_\infty &\le \left\|\{\phi_+ - \psi_+\}\phi_+^\star + \psi_+\{\phi_+^\star - \psi_+^\star\}\right\|_\infty \\ &\le \|\phi_+ - \psi_+\|_\infty \,\left(\|\phi_+\|_\infty + \|\psi_+\|_\infty\right) \,. \end{aligned}$$

Since we assumed that $\phi, \psi \in \mathcal{P}_{\text{pos}}(N; c_1, c_2)$, it follows that both spectral factors are upper bounded by $\sqrt{c_2}$ and therefore, the last inequality becomes $\frac{1}{2\sqrt{c_2}}\,\|\phi - \psi\|_\infty \le \|\phi_+ - \psi_+\|_\infty$, which shows that the constant C_{S1} in (10.66) is given by $C_{S1} = \frac{1}{2\sqrt{c_2}}$. Thus, the lower bound in (10.66) is independent of the degree N of the polynomials. In fact, this lower bound holds for all spectral densities ϕ (not just polynomials) which possess a spectral factorization and which are upper bounded by c_2.

Of much more interest is the upper bound on $\|\phi_+ - \psi_+\|_\infty$. Since we consider trigonometric polynomials $\mathcal{P}(N)$ of degree N and since this space has a finite dimension, it is clear that for every fixed N there exists a constant $C_{S2}(N) < \infty$ such that (10.66) holds.

However, does this constant depend on the degree N? If so, how does $C_{S2}(N)$ depend on N? The following theorem gives an upper bound on the error $\|\phi_+ - \psi_+\|_\infty$ in the spectral factor as a function of the degree N of the spectra and as a function of the error $\|\phi - \psi\|_\infty$ in the given spectra. The same theorem shows that there exist spectra for which this upper bound is almost achieved. Based on this theorem, lower and upper bounds on the continuity constant $C_{S2}(N)$ are derived in a subsequent corollary.

Theorem 10.51. *Let $\mathfrak{S} : \phi \mapsto \phi_+$ be the spectral factorization mapping on the space $\mathcal{P}_{pos}(N; c_1, c_2)$. Then for all $\phi, \psi \in \mathcal{P}_{pos}(N; c_1, c_2)$ holds*

$$\|\phi_+ - \psi_+\|_\infty \leq \left(K_1 + K_2 \log \frac{1}{\sin \frac{\pi}{2N}} \right) \|\phi - \psi\|_\infty \tag{10.67}$$

with the constants

$$K_1 = \frac{1}{\sqrt{c_1}} + \frac{2}{\sqrt{c_1}} \left(\frac{c_2}{c_1} \right)^{3/2} \qquad \text{and} \qquad K_2 = \frac{1}{\pi \sqrt{c_1}} \left(\frac{c_2}{c_1} \right)^{1/2} .$$

Moreover, to every $\delta > 0$ there exist polynomials $\Phi, \Psi \in \mathcal{P}_{pos}(N; c_1, c_2)$ such that

$$\|\Phi_+ - \Psi_+\|_\infty \ \geq \ \left(\tfrac{1}{2\pi} \log(N+1) - \delta \right) \|\Phi - \Psi\|_\infty \, . \tag{10.68}$$

The first part of this theorem shows that the error $\|\phi_+ - \psi_+\|_\infty$ in the spectral factor growth proportional with the error $\|\phi - \psi\|_\infty$ in the given data, and that the proportionality constant C_{S2} in (10.66) depends on the degree N of the spectral data. For large degrees N, one has $C_{S2} \sim \log N$. The second statement of Theorem 10.51 shows that this statement is tight with respect to the growth behavior proportional to $\log N$, because it shows that there exist polynomials $\Phi, \Psi \in P_N$ for which the error $\|\phi_+ - \psi_+\|_\infty$ also grows at least proportional to $\log N$. This observation is summarized in

Corollary 10.52. *Let $\mathfrak{S}$ be the spectral factorization mapping on the space $\mathcal{P}_{pos}(N; c_1, c_2)$. Then for the continuity constant $C_{S2} = C_{S2}(N)$ in (10.66) holds*

$$\tfrac{1}{2\pi} \log(N+1) \ \leq \ C_{S2}(N) \ \leq \ K_1 + K_2 \log N \tag{10.69}$$

with the same constants K_1 and K_2 as in Theorem 10.51.

The proof of Theorem 10.51 will show that the polynomials Φ, Ψ for which (10.68) holds are very simple polynomials (the constant polynomial with a small error). It should be noted that the lower bound in (10.68) holds only for a sufficiently small error $\|\Phi - \Psi\|_\infty$ since the error in the spectral factor is always upper bounded by $\|\Phi_+ - \Psi_+\|_\infty \leq \|\Phi_+\|_\infty + \|\Psi_+\|_\infty \leq 2\sqrt{c_2}$ for all Φ, Ψ which are upper bounded by c_2. Therefore (10.68) holds only if Φ and Ψ satisfy

$$\|\Phi - \Psi\|_\infty \leq \frac{2\sqrt{c_2}}{\frac{1}{2\pi}\log(N+1) - \delta} \,.$$

It is clear that Theorem 10.51 also implies the following, already known result (cf. Theorem 10.32).

Corollary 10.53. *The spectral factorization mapping is discontinuous on* $\mathcal{C}(\mathbb{T})$.

Proof (Theorem 10.51). Without loss of generality, we assume that $0 < c_1 < 1 < c_2$ throughout this proof.
a) *Lower bound*: First, the second statement of the theorem is proved. Let N and $0 < \varepsilon < 1$ be fixed and consider the two trigonometric polynomials $\Phi, \Psi \in \mathcal{P}_{\mathrm{pos}}(N)$ given by

$$\Phi(\mathrm{e}^{\mathrm{i}\theta}) = 1 + \tfrac{\varepsilon}{2}\, g_N(\mathrm{e}^{\mathrm{i}\theta}) \qquad \text{and} \qquad \Psi(\mathrm{e}^{\mathrm{i}\theta}) = 1 - \tfrac{\varepsilon}{2}\, g_N(\mathrm{e}^{\mathrm{i}\theta}) \tag{10.70}$$

with the trigonometric polynomial $g_N \in \mathcal{P}(N)$ given by

$$g_N(\mathrm{e}^{\mathrm{i}\theta}) = \tfrac{1}{\pi}\textstyle\sum_{k=1}^{N} \frac{\sin(k\theta)}{k} \,.$$

The polynomial g_N has the following three properties [92, Chapter II.9 and Chapter V.1], which are needed subsequently:

i)	$\lvert g_N(\mathrm{e}^{\mathrm{i}\theta})\rvert \leq 1$	for all $\theta \in [-\pi, \pi)$
ii)	$g_N(\mathrm{e}^{\mathrm{i}\theta}) \geq 0$	for all $\theta \in [0, \pi]$
iii)	$g_N(\mathrm{e}^{-\mathrm{i}\theta}) = -g_N(\mathrm{e}^{\mathrm{i}\theta})$	for all $\theta \in [0, \pi]$

From the first of these properties follows that

$$\|\Phi - \Psi\|_\infty \leq \varepsilon \qquad \text{and} \qquad 1 - \tfrac{\varepsilon}{2} \leq |\Phi(\mathrm{e}^{\mathrm{i}\theta})|, |\Psi(\mathrm{e}^{\mathrm{i}\theta})| \leq 1 + \tfrac{\varepsilon}{2} \,.$$

Thus the two spectra Φ and Ψ are separated at most by ε from each other in the infinity norm, and we have to investigate the separation of the corresponding spectral factors. To this end, we consider first the difference $|\Phi_+(\zeta) - \Psi_+(\zeta)|$ at points $\zeta \in \mathbb{T}$ on the unit circle

$$|\Phi_+(\zeta) - \Psi_+(\zeta)| = |\Psi_+(\zeta)| \left| \frac{\Phi_+(\zeta)}{\Psi_+(\zeta)} - 1 \right| \tag{10.71}$$

Next, we write the ratio of the spectral factors as

$$\frac{\Phi_+(\zeta)}{\Psi_+(\zeta)} = A(\zeta)\, \mathrm{e}^{-\mathrm{i}\frac{1}{2}\,\beta(\zeta)} \tag{10.72}$$

in which $A(\zeta) = \sqrt{\Phi(\zeta)/\Psi(\zeta)}$ and $\beta(\zeta) = 2\,[\arg \Psi_+(\zeta) - \arg \Phi_+(\zeta)]$ are the modulus and the phase of Φ_+/Ψ_+, respectively. Now we consider (10.71) at the point $\zeta = 1$ and use the representation (10.72) for the ratio of the spectral factors. By the definition of Φ and Ψ, it holds that $\Phi(1) = \Psi(1) = 1$ and therefore $A(1) = 1$. Therewith, the difference (10.71) at $\zeta = 1$ becomes

$$|\Phi_+(1) - \Psi_+(1)| = \left|\mathrm{e}^{-\mathrm{i}\frac{1}{2}\beta(1)} - 1\right| = 2\left|\sin\left(\tfrac{\beta(1)}{4}\right)\right| . \tag{10.73}$$

The function β has to be analyzed next. By the definition (10.8) of the spectral factor, it holds that

$$\frac{\Phi_+(z)}{\Psi_+(z)} = \exp\left(\frac{1}{4\pi}\int_{-\pi}^{\pi} \left[\log \Phi(\mathrm{e}^{\mathrm{i}\tau}) - \log \Psi(\mathrm{e}^{\mathrm{i}\tau})\right] \frac{\mathrm{e}^{\mathrm{i}\tau} + z}{\mathrm{e}^{\mathrm{i}\tau} - z}\,\mathrm{d}\tau\right) .$$

Comparing this expression with (10.72), it is clear that the function β is equal to the conjugate Poisson integral (5.6) of $\log \Psi - \log \Phi$:

$$\beta(r\mathrm{e}^{\mathrm{i}\theta}) = -\frac{1}{2\pi}\int_{-\pi}^{\pi} \left[\log \Phi(\mathrm{e}^{\mathrm{i}\tau}) - \log \Psi(\mathrm{e}^{\mathrm{i}\tau})\right] \mathcal{Q}_r(\theta - \tau)\,\mathrm{d}\tau . \tag{10.74}$$

Using the series expansion $\log(1-x) = -\sum_{k=1}^{\infty} \frac{x^k}{k}$ and the definition of F and G, the argument of the above conjugate Poisson integral becomes

$$\begin{aligned} \log \Phi(\mathrm{e}^{\mathrm{i}\tau}) - \log \Psi(\mathrm{e}^{\mathrm{i}\tau}) &= -\sum_{k=1}^{\infty} \tfrac{1}{k} \left[\tfrac{\varepsilon}{2}\, g_N(\mathrm{e}^{\mathrm{i}\tau})\right]^k \left([-1]^k - 1\right) \\ &= 2\sum_{k=0}^{\infty} \tfrac{1}{2k+1} \left[\tfrac{\varepsilon}{2}\, g_N(\mathrm{e}^{\mathrm{i}\tau})\right]^{2k+1} . \end{aligned} \tag{10.75}$$

Note that because $|\frac{\varepsilon}{2}\, g_N(\mathrm{e}^{\mathrm{i}\tau})| \leq \frac{\varepsilon}{2} < 1$, the above series converges uniformly on $[-\pi, \pi)$. Therefore, (10.75) can be used in (10.74) and the order of integration and summation can be exchanged. This gives for β the expression

$$\begin{aligned} \beta(r\mathrm{e}^{\mathrm{i}\theta}) &= -2\sum_{k=0}^{\infty} \frac{1}{2k+1} \left(\frac{\varepsilon}{2}\right)^{2k+1} (\mathfrak{Q}\, g_N^{2k+1})(r\mathrm{e}^{\mathrm{i}\theta}) \\ &= -\varepsilon\, (\mathfrak{Q} g_N)(r\mathrm{e}^{\mathrm{i}\theta}) + \varepsilon^3 R(\varepsilon, r\mathrm{e}^{\mathrm{i}\theta}) \end{aligned}$$

where for the second line, the sum was split up into its first term ($k = 0$) and all the remaining terms ($k = 1, 2, \dots$) which are collected in the function

$$R(\varepsilon, z) := -\frac{1}{4}\sum_{k=1}^{\infty} \frac{1}{2k+1} \left(\frac{\varepsilon}{2}\right)^{2k-2} (\mathfrak{Q}\, g_N^{2k+1})(z) . \tag{10.76}$$

The conjugate Poisson integral $(\mathfrak{Q}g_N)(z)$ of our special function g_N can easily be determined. It is given by $(\mathfrak{Q}g_N)(re^{i\theta}) = -\frac{1}{\pi}\sum_{k=1}^{N} r^k \frac{\cos(k\theta)}{k}$. Therewith, the function β becomes

$$\beta(re^{i\theta}) = \frac{\varepsilon}{\pi}\sum_{k=1}^{N} r^k \frac{\cos(k\theta)}{k} + \varepsilon^3 R(\varepsilon, re^{i\theta}) \ . \tag{10.77}$$

Next we analyze the term $R(\varepsilon, z)$ given in (10.76) at the real axis, i.e. for $z = r$. In particular, it will be shown that $R(\varepsilon, r)$ is positive and bounded for all $0 \le \varepsilon < 1$ and all $0 \le r \le 1$. Recall first that g_N is an odd function and that $0 \le g_N(\zeta) \le 1$ for all $\zeta \in \mathbb{T}$. From this follows that g_N^{2k+1} is an odd function and that $g_N^{2k+1}(\zeta) \le g_N(\zeta)$ for all $k \ge 0$ and for all $\zeta \in \mathbb{T}$. Furthermore, note that also the conjugate Poisson kernel $\mathcal{Q}(\tau)$, given in (5.4), is an odd function and that $\mathcal{Q}(\tau)$ is non-negative for all $\tau \in [0, \pi]$. Therewith, it follows that

$$-(\mathfrak{Q}\, g_N^{2k+1})(r) \le \frac{1}{\pi}\int_0^{\pi} g_N(e^{i\tau})\, \frac{2r\sin\tau}{1 - 2r\cos\tau + r^2}\, d\tau$$

for all $k \ge 0$. Therein, the right hand side is equal to $-(\mathfrak{Q}\, g_N)(r) = \sum_{k=1}^{N} \frac{r^k}{k}$. Using this in expression (10.76) for the remainder $R(\varepsilon, z)$, one obtains

$$0 \le R(\varepsilon, r) \le \underbrace{\frac{1}{4}\sum_{k=1}^{\infty} \frac{1}{2k+1}\left(\frac{\varepsilon}{2}\right)^{2k-2}}_{\le D_1} \sum_{k=1}^{N} \frac{r^k}{k}$$

where the first sum converges uniformly for all $\varepsilon/2 < 1$ and it is clear that there exists a universal upper bound D_1 (independent of ε) for this sum. Consequently, we have that

$$0 \le R(\varepsilon, r) \le D_1\,(1 + \log N) \ .$$

The last inequality shows that $R(\varepsilon, r)$ is bounded for any fixed N and for all $0 \le \varepsilon \le 1$ and all $0 \le r \le 1$.

We come back to (10.73). Together with (10.77) one obtains that

$$|\Phi_+(1) - \Psi_+(1)| = 2\left|\sin\left[\tfrac{\varepsilon}{4\pi}\textstyle\sum_{k=1}^{N}\tfrac{1}{k} + \tfrac{\varepsilon^3}{4}R(\varepsilon, 1)\right]\right| \ . \tag{10.78}$$

Next, we use that $\sin x \ge x - \frac{1}{6}x^3$ for all $x > 0$. Therewith, the previous equality becomes

$$\begin{aligned} |\Phi_+(1) - \Psi_+(1)| &\ge \frac{\varepsilon}{2\pi}\sum_{k=1}^{N}\frac{1}{k} - O(\varepsilon^3) \\ &\ge \left[\tfrac{1}{2\pi}\log(N+1) - O(\varepsilon^2)\right]\, \|\Phi - \Psi\|_\infty \ . \end{aligned}$$

Since $\|\Phi_+ - \Psi_+\|_\infty \geq |\Phi_+(1) - \Psi_+(1)|$, this shows that

$$\liminf_{\|\Phi-\Psi\|_\infty \to 0} \frac{\|\Phi_+ - \Psi_+\|_\infty}{\|\Phi - \Psi\|_\infty} \geq \frac{1}{2\pi} \log(N+1) \ .$$

Thus, to every $\delta > 0$ there exist trigonometric polynomials $\Phi, \Psi \in \mathcal{P}(N; c_1, c_2)$ such that (10.68) holds.

b) *Upper bound*: We consider two arbitrary spectra $\phi, \psi \in P_N(c_1, c_2)$ with $\|\phi - \psi\|_\infty \leq \varepsilon$. Because of the maximum principle for analytic functions, the maximum of the modulus of $\phi_+ - \psi_+$ is attained on the unit circle $\mathbb{T}$. Therefore, we consider $|\phi_+(\zeta) - \psi_+(\zeta)|$ for $\zeta \in \mathbb{T}$. A straight forward calculation shows that

$$|\phi_+(\zeta) - \psi_+(\zeta)| \leq \tfrac{1}{2\sqrt{c_1}} \, |\phi(\zeta) - \psi(\zeta)| + \tfrac{\sqrt{c_2}}{2} \, |\beta(\zeta)| \qquad (10.79)$$

wherein the function $\beta(z) = 2\left[\arg \phi_+(z) - \arg \psi_+(z)\right]$ is defined as in part a) of this proof. Next, we derive an upper bound for the modulus of $\beta(\zeta)$. By the definition (10.8) of the spectral factor, β is equal to the Hilbert transform (5.19) of the function $h := \log \phi - \log \psi$.

$$\beta(\mathrm{e}^{\mathrm{i}\theta}) = \lim_{\delta \to 0} \frac{1}{2\pi} \int_{\delta < |\theta - \tau| \leq \pi} h(\mathrm{e}^{\mathrm{i}\tau}) \, \frac{1}{\tan \frac{\theta - \tau}{2}} \, \mathrm{d}\tau \ .$$

Note that because ϕ and ψ are trigonometric polynomials in $\mathcal{P}_{\mathrm{pos}}(N; c_1, c_2)$, the functions $\log \phi$, $\log \psi$, and h are 2π-periodic and infinitely differentiable. Therefore, the Hilbert transform of h exists for all $\theta \in [-\pi, \pi)$, i.e. the above integral converges for $\delta \to 0$. We are going to find an upper bound on the Hilbert transform of h. To this end, the integral in the Hilbert transform is split up into an integration over all τ with $|\theta - \tau| < \pi/N$ and an integration over all τ with $|\theta - \tau| \geq \pi/N$. Since the kernel $\cot \frac{\theta - \tau}{2}$ is an odd function with respect to $\tau = \theta$, we can subtract the constant $h(\mathrm{e}^{\mathrm{i}\theta})$ from the argument of the first integral (i.e. the integration over $|\theta - \tau| < \frac{\pi}{N}$) without changing its value. All this, together with the triangle inequality, gives

$$\begin{aligned} \left|\beta(\mathrm{e}^{\mathrm{i}\theta})\right| \leq \underbrace{\frac{1}{2\pi} \int_{|\theta - \tau| \leq \frac{\pi}{N}} \frac{\left|h(\mathrm{e}^{\mathrm{i}\tau}) - h(\mathrm{e}^{\mathrm{i}\theta})\right|}{\left|\tan \frac{\theta - \tau}{2}\right|} \, \mathrm{d}\tau}_{=:T_1} \\ + \underbrace{\frac{1}{2\pi} \int_{\frac{\pi}{N} \leq |\theta - \tau| \leq \pi} \frac{\left|h(\mathrm{e}^{\mathrm{i}\tau})\right|}{\left|\tan \frac{\theta - \tau}{2}\right|} \, \mathrm{d}\tau}_{=:T_2} \qquad (10.80) \end{aligned}$$

for the modulus of β. First the second term T_2 is analyzed. To this end, we note that

$$\left|h(\mathrm{e}^{\mathrm{i}\theta})\right| = \left|\log \frac{\phi(\mathrm{e}^{\mathrm{i}\theta})}{\psi(\mathrm{e}^{\mathrm{i}\theta})}\right| \leq \frac{1}{c_1} \|\phi - \psi\|_\infty \qquad \text{for all } \theta \in [-\pi, \pi)$$

using the relation $|\log(x/y)| \leq \frac{1}{c}\,|x-y|$ with $c = \min(x,y)$, which may easily be verified[5]. With this upper bound on $|h(\mathrm{e}^{\mathrm{i}\theta})|$ one obtains for T_2

$$T_2 \;\leq\; \frac{1}{c_1}\,\|\phi-\psi\|_\infty \frac{1}{\pi}\int_{\frac{\pi}{N}}^{\pi} \frac{\mathrm{d}\tau}{\tan(\tau/2)} \;=\; \frac{2}{\pi c_1}\|\phi-\psi\|_\infty \,\log\frac{1}{\sin\frac{\pi}{2N}} \;. \quad (10.81)$$

Next T_1 is analyzed. The mean value theorem states, that there exists a $\xi \in [\min(\tau,\theta), \max(\tau,\theta)]$ such that $\left|h(\mathrm{e}^{\mathrm{i}\tau}) - h(\mathrm{e}^{\mathrm{i}\theta})\right| = \left|h'(\mathrm{e}^{\mathrm{i}\xi})\right|\,|\tau-\theta|$. This shows that

$$h'(\mathrm{e}^{\mathrm{i}\xi}) = \frac{\phi'(\mathrm{e}^{\mathrm{i}\xi})\left[\psi(\mathrm{e}^{\mathrm{i}\xi}) - \phi(\mathrm{e}^{\mathrm{i}\xi})\right] + \phi(\mathrm{e}^{\mathrm{i}\xi})\left[\phi'(\mathrm{e}^{\mathrm{i}\xi}) - \psi'(\mathrm{e}^{\mathrm{i}\xi})\right]}{\phi(\mathrm{e}^{\mathrm{i}\xi})\,\psi(\mathrm{e}^{\mathrm{i}\xi})} \;.$$

Now we apply *Bernstein's inequality* for trigonometric polynomials [92, Chapter X, §3]. It states that for a trigonometric polynomial ϕ of order N and with $|\phi(\mathrm{e}^{\mathrm{i}\xi})| \leq c_2$ for all ξ, the modulus of the first derivative with respect to ξ is upper bounded by $|\phi'(\mathrm{e}^{\mathrm{i}\xi})| \leq N\,c_2$. Therefore

$$|\phi'(\mathrm{e}^{\mathrm{i}\xi}) - \psi'(\mathrm{e}^{\mathrm{i}\xi})| \leq N|\phi(\mathrm{e}^{\mathrm{i}\xi}) - \psi(\mathrm{e}^{\mathrm{i}\xi})| \leq N\,\|\phi-\psi\|_\infty \leq N\,\varepsilon$$

and one obtains finally $\left|h'(\mathrm{e}^{\mathrm{i}\xi})\right| \;\leq\; 2\,\frac{c_2}{c_1^2}\,N\,\|\phi-\psi\|_\infty$. Therewith the upper bound for T_1 becomes

$$T_1 \;\leq\; \frac{c_2}{c_1^2}\,N\,\|\phi-\psi\|_\infty \frac{2}{\pi}\int_0^{\pi/N} \frac{\tau}{\tan\tau/2}\,\mathrm{d}\tau \;\leq\; 4\frac{c_2}{c_1^2}\,\|\phi-\psi\|_\infty \quad (10.82)$$

using that the function $\tau/\tan(\tau/2)$ is positive and bounded on $[0,\pi/N]$. If the results (10.82) and (10.81) for T_1 and T_2 are plugged into (10.80), one obtains

$$|\beta(\mathrm{e}^{\mathrm{i}\theta})| = 2\,|\arg\psi_+ - \arg\phi_+| \leq \frac{2}{c_1}\left(\frac{2\,c_2}{c_1} + \frac{1}{\pi}\log\frac{1}{\sin\frac{\pi}{2N}}\right)\,\|\phi-\psi\|_\infty$$

for all $\phi, \psi \in \mathcal{P}(N)$. If this bound is used in (10.79) one obtains (10.67). □

Proof (Corollary 10.52). a) *Lower bound*: The upper bound in (10.66) should hold for all $\phi, \psi \in \mathcal{P}_{\mathrm{pos}}(N; c_1, c_2)$. This implies that

$$\sup_{\substack{\phi,\psi\in\mathcal{P}_{\mathrm{pos}}(N;c_1,c_2)\\ \phi\neq\psi}} \frac{\|\phi_+ - \psi_+\|_\infty}{\|\phi-\psi\|_\infty} \leq C_{S2}(N)\;. \quad (10.83)$$

In Theorem 10.51, it was shown that to every $\delta > 0$ there exist polynomials $\Phi, \Psi \in \mathcal{P}_{\mathrm{pos}}(N; c_1, c_2)$ such that

$$\tfrac{1}{2\pi}\,\log(N+1) - \delta \leq \frac{\|\Phi_+ - \Psi_+\|_\infty}{\|\Phi-\Psi\|_\infty}\;. \quad (10.84)$$

[5] Without loss of generality, assume that $x \geq y$ and set $x = \alpha\,y$ for some $\alpha \geq 1$. Then $\log(x/y) = \log\alpha \leq \alpha - 1 = x/y - 1 = (x-y)/y \leq (x-y)/c$.

It is clear that the right hand side of (10.84) is always lower or equal than the left hand side of (10.83). Therefore, combining (10.83) and (10.84) gives $\frac{1}{2\pi}\log(N+1) - \delta \leq C_{S2}(N)$, and since δ was arbitrary, one obtains the lower bound in (10.69) for $\delta \to 0$.

b) *Upper bound*: The upper bound for C_{S2} follows directly from the first statement (10.67) of Theorem 10.51. Using that $\sin x \geq \frac{2}{\pi}x$ for all $x \leq \pi/2$ gives the upper bound in (10.69). □

10.5 Approximation of Spectral Densities

If the given spectral density ϕ is a trigonometric polynomial, there exists a variety of different efficient methods for the determination of the spectral factorization of ϕ (see e.g. the overview in [78]). Therefore, it seems to be very natural to proceed as follows in the case of non-polynomial spectra: First one approximates the given spectral density ϕ by a trigonometric polynomial ϕ_N of a certain degree N. Secondly, one determines the spectral factorization of ϕ_N which will (hopefully) give an approximation $(\phi_N)_+$ of the true spectral factor ϕ_+. Following this approach, the first question which arises is how do we approximate ϕ by the polynomial ϕ_N? Such approximation methods

$$\mathfrak{A}_N : \phi \mapsto \phi_N$$

of the spectral density ϕ are investigated in the present section in some detail. Thereby, we consider only the approximation of spectral densities of completely regular stochastic sequences of a certain order α (cf. Section 10.1), and we consider the approximation only in the infinity norm $\|\cdot\|_\infty$.

In order to obtain a simple representation of the approximation operator $\mathfrak{A}_N$, it is desirable that $\mathfrak{A}_N$ be linear. Another requirement on the approximation method is certainly that the approximation error $\|\phi - \phi_N\|_\infty$ converges to zero as fast as possible as the degree N of the approximation polynomial is increased. This would allow the use of approximation polynomials ϕ_N with small degrees, which are usually easier to factorize. Assume that ϕ is the spectral density of a completely regular stochastic sequence of order α and assume that ϕ has no zeros on $\mathbb{T}$. Then it follows from Theorem 10.4 that ϕ is a real valued function in the Hölder-Zygmund class $\Lambda_\alpha(\mathbb{T})$. For such densities, it was shown in Corollary 2.9 that there exists a constant C_α, which dependents only on the order α, such that

$$B_N[\phi] = \min_{\phi_N \in \mathcal{P}_N} \|\phi - \phi_N\|_\infty \leq C_\alpha\, N^{-\alpha}\,. \tag{10.85}$$

Thus, the maximal rate of convergence of $\|\phi - \phi_N\|_\infty$ is determined by the order α. The higher the order α of the stochastic sequence X, the faster the best possible approximation error decreases as the degree of the approximation polynomial is increased. However, the mapping $\phi \mapsto \phi_N$ that gives the optimal approximation polynomial ϕ_N is non-linear and unknown, in general. Nevertheless, it was shown in Theorem 2.7 that the approximation of ϕ by its delayed arithmetic mean $\sigma_{N,N}(\phi)$ shows almost (up to a factor of at most 4) the optimal convergence behavior according to (10.85).

Besides the desired linearity and fast convergence of the approximation operation, there is also a more stringent property which the approximation operator must satisfy. Since the approximated spectrum ϕ_N should be factorized, it has to be non-negative. Consequently, the approximation operator $\mathfrak{A}_N : \phi \mapsto \phi_N$ needs to have the property that it always maps non-negative spectral densities onto non-negative polynomials ϕ_N. But this property of $\mathfrak{A}_N$ may to be at odds with the desired linearity and fast convergence of $\mathfrak{A}_N$. This is what we want to investigate in the present section.

To make the statement precise, we require that the approximation operator $\mathfrak{A}_N : \mathcal{C}_{\text{pos}}(\mathbb{T}) \to \mathcal{P}(N)$ satisfies the following four properties.

(A) *Linearity*: We require that the approximation operator $\mathfrak{A}_N$ is linear, i.e.

$$\mathfrak{A}_N(a\,\phi + b\,\psi) = a(\mathfrak{A}_N\phi) + b(\mathfrak{A}_N\psi)$$

for all spectral densities ϕ, ψ and for arbitrary complex numbers a and b. The linearity of $\mathfrak{A}_N$ ensures a sufficiently simple calculation of the approximate spectrum ϕ_N.

(B) *Translation invariance*: If we write $\phi^\tau(\mathrm{e}^{\mathrm{i}\theta}) := \phi(\mathrm{e}^{\mathrm{i}(\theta+\tau)})$ for the spectrum obtained from ϕ by a translation[6] by τ, then $\mathfrak{A}_N$ should satisfy the relation

$$(\mathfrak{A}_N\phi^\tau)(\mathrm{e}^{\mathrm{i}\theta}) = (\mathfrak{A}_N\phi)(\mathrm{e}^{\mathrm{i}(\theta+\tau)})\ .$$

(C) *Positivity*: To ensure that the approximated spectrum $\phi_N = \mathfrak{A}_N\,\phi$ possesses a spectral factorization, we have to require that $\mathfrak{A}_N$ maps every non-negative spectrum ϕ onto a non-negative trigonometric polynomial. Thus, from $\phi \geq 0$ should always follow that $\mathfrak{A}_N\phi \geq 0$. This condition on $\mathfrak{A}_N$ is obviously a minimal necessary requirement which cannot be abdicated since it arises from the desired application of spectral factorization.

(D) *Optimal rate of convergence*: The optimal rate is determined by the regularity of the stochastic sequence. We therefore require that that for every completely regular sequence of order $\alpha > 0$ with spectral density ϕ the approximation error satisfies

$$\|\phi - \mathfrak{A}_N\phi\|_\infty \leq C\,N^{-\alpha} \tag{10.86}$$

with a certain constant C independent of ϕ.

[6] Since the spectra ϕ is defined on $\mathbb{T}$, and therefore 2π-periodic, the translation is equivalent to a rotation of ϕ on the unit circle.

10.5.1 General form of the approximation method

At the beginning, we derive the general from of the approximation operator $\mathfrak{A}_N$ which is imposed by the required properties (A), (B), and (C)[7]. To this end, we consider for $k = 0, 1, 2, \dots$ the functions

$$c_k(\mathrm{e}^{\mathrm{i}\theta}) := \cos(k\,\theta) \qquad \text{and} \qquad s_k(\mathrm{e}^{\mathrm{i}\theta}) := \sin(k\,\theta)\ , \qquad \theta \in [-\pi, \pi)\ .$$

A simple calculation shows that

$$\begin{aligned} c_k^{\tau}(\mathrm{e}^{\mathrm{i}\theta}) &= \cos(k\tau)\, c_k(\mathrm{e}^{\mathrm{i}\theta}) - \sin(k\tau)\, s_k(\mathrm{e}^{\mathrm{i}\theta}) \\ s_k^{\tau}(\mathrm{e}^{\mathrm{i}\theta}) &= \cos(k\tau)\, s_k(\mathrm{e}^{\mathrm{i}\theta}) + \sin(k\tau)\, c_k(\mathrm{e}^{\mathrm{i}\theta})\ . \end{aligned}$$

Next, we apply the approximation operator $\mathfrak{A}_N$ onto these two functions. By the required linearity and translation invariance of $\mathfrak{A}_N$ follows that

$$\begin{aligned} (\mathfrak{A}_N c_k)(\mathrm{e}^{\mathrm{i}(\theta+\tau)}) &= \cos(k\tau)\, (\mathfrak{A}_N\, c_k)(\mathrm{e}^{\mathrm{i}\theta}) - \sin(k\tau)\, (\mathfrak{A}_N\, s_k)(\mathrm{e}^{\mathrm{i}\theta}) \\ (\mathfrak{A}_N s_k)(\mathrm{e}^{\mathrm{i}(\theta+\tau)}) &= \cos(k\tau)\, (\mathfrak{A}_N\, s_k)(\mathrm{e}^{\mathrm{i}\theta}) + \sin(k\tau)\, (\mathfrak{A}_N\, c_k)(\mathrm{e}^{\mathrm{i}\theta})\ . \end{aligned}$$

If these equations are evaluated at $\theta = 0$, one obtains the two equations

$$\begin{aligned} (\mathfrak{A}_N c_k)(\mathrm{e}^{\mathrm{i}\tau}) &= \gamma_k(N)\ \cos(k\tau) - \delta_k(N)\ \sin(k\tau) \\ (\mathfrak{A}_N s_k)(\mathrm{e}^{\mathrm{i}\tau}) &= \delta_k(N)\ \cos(k\tau) + \gamma_k(N)\ \sin(k\tau) \end{aligned} \tag{10.87}$$

with the constants $\gamma_k(N) := (\mathfrak{A}_N c_k)(1)$ and $\delta_k(N) := (\mathfrak{A}_N s_k)(1)$ which are uniquely determined by the approximation operator $\mathfrak{A}_N$. Thus $\gamma_k(N)$ and $\delta_k(N)$ are just the Fourier coefficients of the trigonometric polynomials $\mathfrak{A}_N c_k$ and $\mathfrak{A}_N s_k$. By the definition of the approximation operation, $\mathfrak{A}_N \phi$ is a trigonometric polynomial with a degree of at most N for every continuous spectrum ϕ. Therefore equations (10.87) imply that

$$\gamma_k(N) = 0 \quad \text{and} \quad \delta_k(N) = 0 \qquad \text{for all } k > N\ . \tag{10.88}$$

Now let now $\phi \in \mathcal{C}(\mathbb{T})$ be an arbitrary spectral density with the Fourier series representation

$$\begin{aligned} \phi(\mathrm{e}^{\mathrm{i}\theta}) &= \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k\ \cos(k\theta) + b_k\ \sin(k\theta) \\ &= \frac{a_0}{2}\, c_0(\theta) + \sum_{k=1}^{\infty} a_k\, c_k(\mathrm{e}^{\mathrm{i}\theta}) + b_k\, s_k(\mathrm{e}^{\mathrm{i}\theta}) \end{aligned} \tag{10.89}$$

Applying the approximation operator onto ϕ and using the properties (10.87) and (10.88) gives the following general form of $\mathfrak{A}_N \phi$

[7] The following deviation is a simple variation of the so called *Zygmund-Martsinkevich-Berman identity*, which may be found in [54, Chapter VII].

$$(\mathfrak{A}_N\phi)(\theta) = \frac{a_0}{2}\,\gamma_0(N) + \sum_{k=1}^{N}[a_k\,\gamma_k(N) + b_k\,\delta_k(N)]\,\cos(k\theta) + [b_k\,\gamma_k(N) - a_k\,\delta_k(N)]\,\sin(k\theta)\ . \tag{10.90}$$

We summarize this result in the following lemma.

Lemma 10.54. *Let $\mathfrak{A}_N : \mathcal{C}(\mathbb{T}) \to \mathcal{P}(N)$ be an arbitrary linear and translation invariant approximation operator and let $\phi \in \mathcal{C}(\mathbb{T})$ be an arbitrary continuous function of the form* (10.89)*. Then there exist uniquely defined constants $\gamma_k(N)$ and $\delta_k(N)$, $k = 1, 2, \ldots, N$ such that $\mathfrak{A}_N\phi$ is equal to* (10.90)*.*

If the Fourier coefficients (2.1) of the function ϕ are plugged into the general representation (10.90) of the operator $\mathfrak{A}_N$, one obtains an integral representation of $\mathfrak{A}_N$

$$(\mathfrak{A}_N\phi)(\mathrm{e}^{\mathrm{i}\theta}) = \frac{1}{2\,\pi}\int_{-\pi}^{\pi}\phi(\mathrm{e}^{\mathrm{i}\tau})\,A_N(\theta-\tau)\,\mathrm{d}\tau\ , \qquad \theta\in[-\pi,\pi) \tag{10.91}$$

with the kernel

$$A_N(\theta) = \gamma_0(N) + 2\sum_{k=1}^{N}\gamma_k(N)\,\cos(k\theta) - \delta_k(N)\,\sin(k\theta)\ . \tag{10.92}$$

Thus, the kernel A_N is a trigonometric polynomial. For simplicity we assume in the following that all $\delta_k(N)$ are equal to zero such that the approximation methods are assumed to have the form

$$(\mathfrak{A}_N\phi)(\mathrm{e}^{\mathrm{i}\theta}) = \frac{a_0}{2}\,\gamma_0(N) + \sum_{k=1}^{N}\gamma_k(N)\,[a_k\,\cos(k\theta) + b_k\,\sin(k\theta)]\ . \tag{10.93}$$

and such that the kernel A_N in the integral representation (10.91) is a pure cosine polynomial of the form

$$A_N(\theta) = \gamma_0(N) + 2\sum_{k=1}^{N}\gamma_k(N)\,\cos(k\theta)\ . \tag{10.94}$$

This restriction to cosine polynomials will result in no limitation on our results. We will see later that the optimal kernel (with respect to the convergence rate) is a purely cosine polynomial. Moreover, among all positive kernels (see below) the cosine polynomials always possess such an optimality behavior among all trigonometric polynomials [37]. Notice that the particular approximation methods considered in Section 2.1 had (of course) the form (10.94) (compare with 2.7) where the sequence $\{\gamma_k\}_{k=0}^N$ was called the *window function* of the particular approximation method.

Up to now, we used only the linearity (A) and the translation invariance (B) to obtain the representation (10.91) of $\mathfrak{A}_N$. Next, it is shown that from requirement (C) on $\mathfrak{A}_N$, namely that $\phi(\theta) \geq 0$ always implies $(\mathfrak{A}_N\phi)(\theta) \geq 0$ for all $\theta \in [-\pi, \pi)$, follows that the kernel A_N of $\mathfrak{A}_N$ is positive.

Lemma 10.55. *Let $\mathfrak{A}_N$ be a linear operator of the form* (10.91) *such that $\phi \geq 0$ always implies that $\mathfrak{A}_N\phi \geq 0$. Then*

$$A_N(\tau) \geq 0 \qquad \text{for all } \tau \in [-\pi, \pi) \,.$$

Proof. This lemma is proved indirectly by contradiction. Assume that there exists a point $\tau_0 \in [-\pi, \pi)$ such that $A_N(\tau_0) < 0$. Since A_N is continuous, there exists a whole interval $I = [\tau_1, \tau_2]$ with $\tau_0 \in I$ and a constant $c_0 > 0$ such that $A_N(\tau) \leq -c_0 < 0$ for all $\tau \in I$. Next, consider for an arbitrary $\mu > 0$ the function ϕ_μ defined by

$$\phi_\mu(\mathrm{e}^{\mathrm{i}\tau}) := \begin{cases} \mu \,, & \tau_1 + \frac{1}{\mu} \leq -\tau \leq \tau_2 - \frac{1}{\mu} \\ \mu^{-1} \,, & -\tau \notin I \\ \text{linear} \,, & \text{elsewhere} \end{cases}$$

and such that $\phi_\mu(\mathrm{e}^{\mathrm{i}\tau})$ is continuous for all $\tau \in [-\pi, \pi)$. This function satisfies the conditions of the lemma, and we consider the approximation $(\mathfrak{A}_N\phi_\mu)(\mathrm{e}^{\mathrm{i}\theta})$ at the point $\theta = 0$. If the integration in (10.91) is split up into an integral over I and an integral over $[-\pi, \pi)\backslash I$ one obtains that

$$\begin{aligned} (\mathfrak{A}_N\phi_\mu)(1) &= \frac{1}{2\pi}\int_I \phi_\mu(\mathrm{e}^{-\mathrm{i}\tau})\, A_N(\tau)\,\mathrm{d}\tau + \frac{1}{2\pi}\int_{[-\pi,\pi)\backslash I} \phi_\mu(\mathrm{e}^{-\mathrm{i}\tau})\, A_N(\tau)\,\mathrm{d}\tau \\ &\leq -c_0\,\mu\,|\tau_1 - \tau_2| \;+\; \frac{1}{\mu\,\pi}\int_{[-\pi,\pi)\backslash I} A_N(\tau)\,\mathrm{d}\tau \,. \end{aligned}$$

Since the integral of A_N in the last line is finite, the last inequality shows that for sufficiently large μ the approximation $(\mathfrak{A}_N\phi_\mu)(\theta)$ becomes negative at $\theta = 0$ which contradicts the assumption that $\phi \geq 0$ always implies that $\mathfrak{A}_N\phi \geq 0$. Therefore, $A_N(\tau) \geq 0$ for all $\tau \in [-\pi, \pi)$. □

10.5.2 No free lunch with positive approximation methods

All approximation methods $\mathfrak{A}_N$ with properties (A), (B), and (C) have the form (10.91) with a positive kernel A_N. The question arises whether there exists an approximation method which satisfies all four requirement (A), (B), (C), and (D). Before we answer this question, the existence of a non-linear method with the properties (B), (C), and (D) is investigated. To this end, we define $\mathfrak{A}_N\phi$ as the polynomial $\phi_N \in \mathcal{P}(N)$ which achieves the best approximation, i.e. which minimizes $\|\phi - \phi_N\|_\infty$. Thus $\mathfrak{A}_N\phi$ is defined by the relation

$$\inf_{p\in\mathcal{P}(N)} \|\phi - p\|_\infty = \|\phi - \mathfrak{A}_N\phi\|_\infty \,. \tag{10.95}$$

Since, according to Proposition 2.6, for every continuous function ϕ there exists a polynomial $\phi_N = \mathfrak{A}_N\phi \in \mathcal{P}(N)$ such that (10.95) holds, $\mathfrak{A}_N\phi$ is well defined. It is clear that $\mathfrak{A}_N$ is translation invariant and that it satisfies property (D). However, $\mathfrak{A}_N$ does not satisfy property (C), i.e. $\mathfrak{A}_N\phi$ needs not be non-negative for every non-negative ϕ. For this reason, we define the operator

$$(\mathfrak{A}_N^+\phi)(\mathrm{e}^{\mathrm{i}\theta}) := (\mathfrak{A}_N\phi)(\mathrm{e}^{\mathrm{i}\theta}) + \|\phi - \mathfrak{A}_N\phi\|_\infty \,, \qquad \theta\in[-\pi,\pi) \,. \tag{10.96}$$

The so defined non-linear approximation operator has all three desired properties. This is shown by the following theorem.

Theorem 10.56. *The operator* $\mathfrak{A}_N^+$ *defined by* (10.96) *possesses property* (B)*,* (C)*, and* (D)*.*

Proof. Since $\mathfrak{A}_N$ is translation invariant, so is $\mathfrak{A}_N^+$ and therefore it satisfies (B). Moreover, for any spectral density ϕ of a regular stochastic process of order α holds that

$$\|\phi - \mathfrak{A}_N^+\phi\|_\infty = \|\phi - \mathfrak{A}_N\phi + \|\phi - \mathfrak{A}_N\phi\|_\infty\,\|_\infty \le 2\,\|\phi - \mathfrak{A}_N\phi\|_\infty \le \frac{2\,C_1}{N^\alpha} \,.$$

This shows that $\mathfrak{A}_N$ satisfies property (D). Furthermore, by the definition of $\mathfrak{A}_N^+$ it obviously holds that

$$\begin{aligned}(\mathfrak{A}_N^+\phi)(\mathrm{e}^{\mathrm{i}\theta}) &= \phi(\mathrm{e}^{\mathrm{i}\theta}) - \phi(\mathrm{e}^{\mathrm{i}\theta}) + (\mathfrak{A}_N\phi)(\mathrm{e}^{\mathrm{i}\theta}) + \|\phi - \mathfrak{A}_N\phi\|_\infty \\ &= \underbrace{\phi(\mathrm{e}^{\mathrm{i}\theta})}_{>0} + \underbrace{\|\phi - \mathfrak{A}_N\phi\|_\infty - \big[\,\phi(\mathrm{e}^{\mathrm{i}\theta}) - (\mathfrak{A}_N\phi)(\mathrm{e}^{\mathrm{i}\theta})\big]}_{\ge 0} \;>\; 0\end{aligned}$$

for every $\phi > 0$ and for all $\theta \in [-\pi,\pi)$. Therefore, $\mathfrak{A}_N$ satisfies property (C). □

Thus there exists a non-linear approximation method with properties (B), (C), and (D). Does there also exist a *linear* method with these three properties? The following theorem gives a negative answer, in general.

Theorem 10.57. *There exists no approximation method with properties* (A)*,* (B)*,* (C) *and which satisfies property* (D) *for completely regular spectral densities of order* $\alpha > 2$*. However, there exist approximation methods which satisfies the properties* (A)*,* (B)*,* (C)*, and* (D) *for spectral densities with an order of regularity* $\alpha \le 2$*.*

Comparing Theorem 10.56 and 10.57 one sees that an approximation method with properties (A), (B), (C) will not necessarily achieve the optimal approximation rate according to the regularity of the stochastic process (cf. (10.86)). Only for processes with a regularity $\alpha \le 2$ does such a method always exist. However, if one gives up the linearity (A) of the method, the optimal convergence rate can always be achieved for every completely regular process of order $\alpha > 0$ (cf. Theorem 10.56).

Example 10.58. The method of the de-la-Vallée-Poussion mean (cf. Section 2.1.3) of the Fourier series is a method which satisfies properties (A), (B), and (D), but which is non-positive. Therefore this method may not be applied in the present context of spectral factorization.

Proof (Theorem 10.57). For $0 < \mu < 1$, we consider the spectral density $\phi_\mu(\mathrm{e}^{\mathrm{i}\theta}) := 1 - \mu\cos(\theta)$, which is positive and infinitely often differentiable, and we consider an approximation method (10.91) with a positive kernel of the form (10.94). Without loss of generality we assume that $\gamma_0(N)$ is normalized to 1. From (10.90), one obtains that $(\mathfrak{A}_N\phi_\mu)(\mathrm{e}^{\mathrm{i}\theta}) = 1 - \mu\,\gamma_1(N)\cos(\theta)$. Therewith, the approximation error becomes

$$\|\phi_\mu - \mathfrak{A}_N\phi_\mu\|_\infty = \mu\,[1 - |\gamma_1(N)|]\ . \tag{10.97}$$

It was shown by *L. Fejér* [37] that the first coefficient $\gamma_1(N)$ of every positive trigonometric polynomial of the form (10.94) with $\gamma_0(N) = 1$ satisfies

$$|\gamma_1(N)| \le \cos\frac{\pi}{N+2}\ . \tag{10.98}$$

Moreover, there exists exactly one trigonometric polynomial of the form (10.94) for which equality holds in (10.98), and it was pointed out by *G. Szegö* [84], that there exists no other trigonometric polynomial of the more general form (10.92) for which equality holds in (10.98)[8]. The window function for which equality holds in (10.98) is given by

$$\gamma_k(N) = \frac{\sum_{n=0}^{N-k}\sin\left[\frac{(n+1)\pi}{N+2}\right]\,\sin\left[\frac{(n+k+1)\pi}{N+2}\right]}{\sum_{n=0}^{N}\sin^2\left[\frac{(n+1)\pi}{N+2}\right]}\ ,\qquad k = 0,1,\dots,N \tag{10.99}$$

and the corresponding kernel (10.94) in closed form becomes (cf. [25, Sec. 1.6]) equal to

$$\mathcal{K}_N(\theta) = \frac{2\,\sin^2\left(\frac{\pi}{N+2}\right)}{N+2}\left(\frac{\cos(N+2)\frac{\theta}{2}}{\cos\theta - \cos\frac{\pi}{N+2}}\right)^2\ . \tag{10.100}$$

This kernel is usually called *Fejér-Korovkin*. Using (10.98) in (10.97) shows that

$$\|\phi_\mu - \mathfrak{A}_N\phi_\mu\|_\infty \ge \mu\left(1 - \cos\frac{\pi}{N+2}\right) \ge \frac{\mu}{2}\,\frac{\pi^2}{(N+1)^2} \tag{10.101}$$

which proves that even for the very smooth (infinitely often differentiable) trigonometric polynomial ϕ_μ only a convergence behavior can be achieved by which the approximation error decreases proportional to N^{-2}.

If on the other hand, the stochastic process has a regularity of $0 \le \alpha \le 2$, the approximation method (10.91) with the kernel (10.100) achieves the optimal approximation behavior proportional to $N^{-\alpha}$ [25]. □

[8] This justifies by hindsight the restriction to cosine kernels of the form (10.94).

Remark 10.59. In the above proof, it was shown that for the special density $\phi_\mu = 1 - \mu \cos\theta$ the approximation behavior is never better than (10.101) and in particular never better than N^{-2}, even though the density ϕ_μ is infinitely often continuously differentiable. If the densities are only twice continuously differentiable it holds [25, Sec. 1.6] that

$$\lim_{N\to\infty} N^2 \left[\phi(\theta) - (\mathfrak{A}_N^{\text{opt}}\phi)(\theta)\right] = \tfrac{\pi^2}{2}\, \phi''(\theta)$$

in which $\mathfrak{A}_N^{\text{opt}}$ denotes the optimal approximation method with the Fejér-Korovkin kernel (10.100). Clearly, also in this case, the optimal approximation behavior is proportional to N^{-2}. Only for the trivial case that $\phi(\theta)$ is constant for all θ does the approximation error become zero, independent of the approximation degree N. However, this behavior holds also for the more general case that ϕ is differentiable and that ϕ' satisfies a generalized Lipschitz condition (a *Zygmund condition*) [25], i.e. for spectra for which there exists a constant C_2 such that

$$|\phi'(\theta+\tau) + \phi'(\theta-\tau) - \phi'(\theta)| \le C_2\,|\tau| \qquad \text{for all } \theta \in [-\pi,\pi)\ .$$

It might be interesting to compare the optimal positive method given by the window (10.99) and by the kernel (10.100) with approximation methods studied in Section 2.1. In Fig. 10.2 and Fig. 10.3 we show again the window function $\{\gamma_k(N)\}_{k=0}^\infty$ and the corresponding kernel A_N, given by (10.94), for the de-la-Vallée-Poussin and Fejér approximation method, together with the optimal Fejér-Korovkin kernel used in the proof of Theorem 10.57. This Figure shows that the window function of the de-la-Vallée-Poussin method has a *passband property*, i.e. the $\gamma_k(N) = 1$ for all $k \le K(N)$ for a certain natural number $K(N) \le N$ (In the case of the de-la-Vallée-Poussin window is $K = K(N) = N/2$ for even N.). Thus, the first $K(N)$ Fourier coefficients are unchanged by the approximation operator. Methods with such a passband property will give the exact approximation for all polynomials with a degree less or equal to $K(N)$, i.e. $\mathfrak{A}_N p = p$ for all $p \in P(K)$. As a further consequence, methods with a passband property always achieve the optimal convergence rate according to the smoothness of the given density ϕ. To see this, let ϕ be a spectral density of a complete regular stochastic sequence of order α. Moreover, assume for simplicity that $K = K(N) = cN$ with an arbitrary constant $0 < c < 1$ and denote with $p_K \in \mathcal{P}(K)$ that polynomial which achieves the best approximation $B_K[\phi]$ of ϕ. Then it holds that

$$\begin{aligned}\|\phi - \mathfrak{A}_N\phi\|_\infty &\le \|\phi - p_K\|_\infty + \|\mathfrak{A}_N(p_K - \phi)\|_\infty \\ &\le (1 + \|\mathfrak{A}_N\|)\,\|\phi - p_K\|_\infty \le (1 + \|\mathfrak{A}_N\|)\,\frac{C_\alpha}{c^\alpha}\,\frac{1}{N^\alpha}\end{aligned}$$

using that $\mathfrak{A}_N p_K = p_K$ and property (10.85) of the best approximation. This shows that approximation methods with a passband property achieve the optimal convergence rate (according to the regularity of the spectral density) as long as they are uniformly bounded, i.e. as long as there exists a

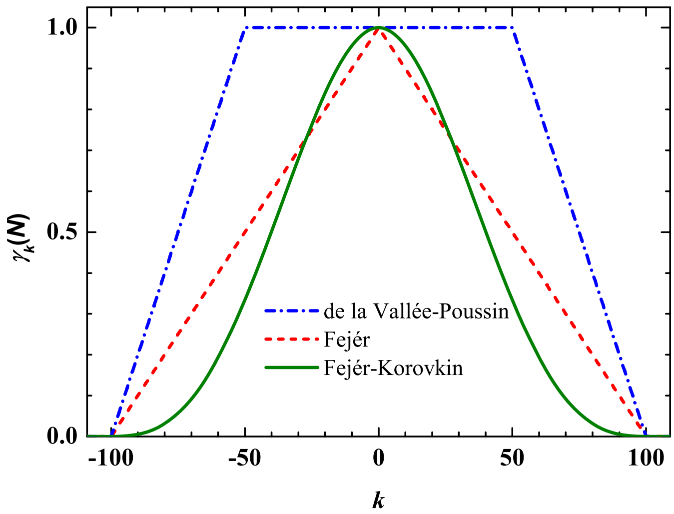

Fig. 10.2. Window functions corresponding to the de-la-Vallée-Poussin, Fejér, and Fejér-Korovkin approximation method of order $N = 100$.

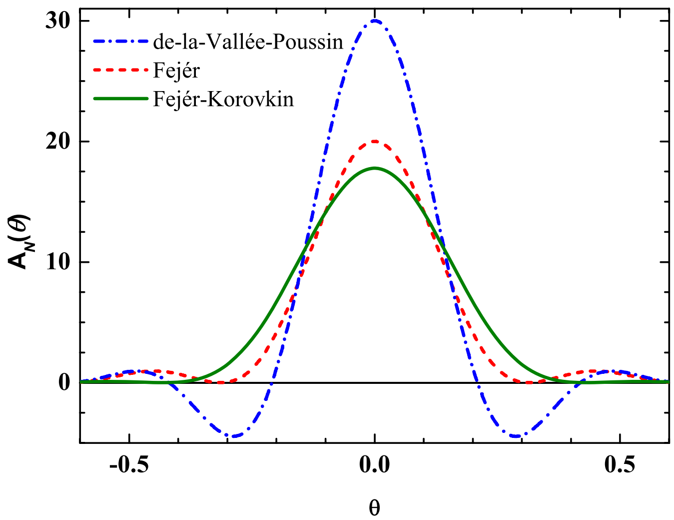

Fig. 10.3. De-la-Vallée-Poussin, Fejér, and Fejér-Korovkin kernel for the order $N = 20$.

constant $C_0 < \infty$ such that $\|\mathfrak{A}_N\| \leq C_0$ for all N. However, the result of Fejér (10.98), which was used in the proof of Theorem 10.57, shows that a positive approximation method can never have a passband property because the first coefficient $\gamma_1(N)$ of the window function is always smaller or equal to $\cos\frac{\pi}{N+2} < 1$ if the zeroth coefficient $\gamma_0(N)$ is normalized to 1. As the above proof shows, this particular decrease of the window function according to (10.98) in the neighborhood of 0 is responsible for the approximation behavior according to (10.101). The Fejér kernel, on the other hand, is a positive kernel (cf. Fig. 10.3). Therefore, it possesses no passband property (cf. Fig. 10.2). Moreover, Fig. 10.2 shows that the window function of the Fejér method decreases faster than the Fejér-Korovkin window in the neighborhood of zero. As a consequence, the convergence rate will be worse compared with the Fejér-Korovkin method. Indeed, it can be shown [92, Chap. III,§ 13.32] that approximation error using the Fejér means decreases at most proportional t0 N^{-1} even for densities with a regularity of $\alpha > 1$.

Assume that the given spectral density ϕ belongs to a certain subset $\mathcal{C}_\omega(\mathbb{T})$ of smooth functions on $\mathbb{T}$, characterized by a majorant ω (cf. Section 1.3). Is it true that also the approximation $\phi_N = \mathfrak{A}_N\phi$ belongs to the same class $\mathcal{C}_\omega(\mathbb{T})$ of smooth functions? The affirmative answer to this question is given by the next lemma, which will be needed in the next section.

Lemma 10.60. *Let $\mathfrak{A}_N$ be an approximation operator of the form* (10.91) *with a positive kernel A_N, let ω be an arbitrary majorant, and let $\phi \in \mathcal{C}_\omega(\mathbb{T})$ be a spectral density. Then the approximation $\phi_N = (\mathfrak{A}_N\phi)$ also belongs to $\mathcal{C}_\omega(\mathbb{T})$.*

Proof. By the definition (10.91) of the approximation operator, one has that

$$\begin{aligned}
\left|\phi_N(\mathrm{e}^{\mathrm{i}\tau}) - \phi_N(\mathrm{e}^{\mathrm{i}\theta})\right| &= \left|\frac{1}{2\pi}\int_{-\pi}^{\pi}\left[\phi(\mathrm{e}^{\mathrm{i}(\tau-x)}) - \phi(\mathrm{e}^{\mathrm{i}(\theta-x)})\right] A_N(x)\,\mathrm{d}x\right| \\
&\leq \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\phi(\mathrm{e}^{\mathrm{i}(\tau-x)}) - \phi(\mathrm{e}^{\mathrm{i}(\theta-x)})\right| A_N(x)\,\mathrm{d}x \\
&\overset{(a)}{\leq} \frac{1}{2\pi}\int_{-\pi}^{\pi}\omega(|\tau-\theta|)\,A_N(x)\,\mathrm{d}x \;=\; \gamma_0(N)\,\omega(|\tau-\theta|)
\end{aligned}$$

where for (a) it was used that $\phi \in \mathcal{C}_\omega(\mathbb{T})$. This inequality proves that $\phi_N \in \mathcal{C}_\omega(\mathbb{T})$. □

10.6 Spectral Factorization of Approximated Spectra

Assume that ϕ is a given spectral density and that $\phi_N = \mathfrak{A}_N\phi$ is an approximation of degree N obtained by a linear, translation invariant, and positive approximation method $\mathfrak{A}_N$. In the previous section it was studied how the approximation error $\|\phi - \phi_N\|_\infty$ depends on the degree N and how the convergence rate is influenced by the regularity of the density ϕ. Now, let $(\phi_N)_+$

be the spectral factor of the approximated spectrum ϕ_N. What can be said about the error $\|\phi_+ - (\phi_N)_+\|_\infty$ between the spectral factor ϕ_+ and its approximation $(\phi_N)_+$. We saw in Section 10.3 that the spectral factorization mapping is discontinuous on the space of all continuous functions with respect to the supremum norm. Thus, it seems not to be obvious at the outset whether the error $\|\phi_+ - (\phi_N)_+\|_\infty$ even converges for $N \to \infty$. Although we consider here the spectral factorization only on the subset of smooth, i.e. regular spectral densities, one might expect at least a certain loss in the convergence rate due to the additional spectral factorization. However, in this section we will derive an upper bound on the error $\|\phi_+ - (\phi_N)_+\|_\infty$ which will show that this loss in convergence rate (if it exists) cannot be to large, i.e. $\|\phi_+ - (\phi_N)_+\|_\infty$ shows a similar dependency on the approximation degree N as $\|\phi - \phi_N\|_\infty$.

As before, $\mathfrak{A}_N : \phi \mapsto \phi_+$ denotes an approximation operator (10.91) with the properties (A), (B), (C) defined in the previous section. We start with the observation that the convergence rate of the error $\|\phi_+ - (\phi_N)_+\|_\infty$ in the spectral factor can never be better than the convergence rate of $\|\phi - \phi_N\|_\infty$, i.e. the error $\|\phi_+ - (\phi_N)_+\|_\infty$ in the spectral factor never decreases faster than the approximation error $\|\phi - \phi_N\|_\infty$ of the spectral density itself. In particular it decreases never faster than proportional to N^{-2} as the approximation degree N is increased, even for spectral densities with a regularity larger than 2. Thus, we obtain a lower bound on the error in the spectral factor.

Theorem 10.61 (Lower bound). *Let ϕ be a spectral density of a completely regular stochastic sequence of order $\alpha > 0$ and let $\phi_N = \mathfrak{A}_N \phi$ be an approximation of ϕ by a method with the properties (A), (B), and (C). Then the approximation error $\|\phi_+ - (\phi_N)_+\|_\infty$ of the spectral factor never decreases faster than proportional to N^{-2}, i.e. there exists a constant C_0 such that*

$$\|\phi_+ - (\phi_N)_+\|_\infty \geq C_0\, N^{-2} .$$

Proof. Using the relation $\phi = \phi_+ \overline{\phi_+}$ for the spectral factors, a simple calculation shows that

$$\begin{aligned} \|\phi - \phi_N\|_\infty &\leq \|\phi_+ \overline{\phi_+} - \overline{\phi_+}\,(\phi_N)_+\|_\infty + \|\overline{\phi_+}\,(\phi_N)_+ - (\phi_N)_+\,\overline{(\phi_N)_+}\|_\infty \\ &= (\|\phi_+\|_\infty + \|(\phi_N)_+\|_\infty)\; \|\phi_+ - (\phi_N)_+\|_\infty \ . \end{aligned}$$

Theorem 10.57 (and its proof) show that there exists a constant C_1 such that every approximation method with property (A), (B), and (C) and for all densities ϕ with a regularity $\alpha \geq 2$ always $\|\phi - \phi_N\|_\infty \geq C_1/N^2$ holds, and for densities with a regularity $\alpha < 2$ the approximation error is lower bounded by C_2/N^α with a certain constant C_2. Altogether, this shows that there exists a constant C_0 such that $C_0/N^2 \leq \|\phi_+ - (\phi_N)_+\|_\infty$ for every approximation method with properties (A), (B), (C). □

Thus, the convergence rate for the error in the spectral factor $\|\phi_+ - (\phi_N)_+\|_\infty$ cannot be better than the rate for the approximation error $\|\phi -$

$\phi_N\|_\infty$ of the spectrum itself. The question is whether the approximation behavior of the spectral factor is even worse due to spectral factorization mapping, applied to the approximation ϕ_N. Note that for continuous spectral densities $\phi \in \mathcal{C}(\mathbb{T})$ and in view of Theorem 10.32, which states that every continuous spectral density is a discontinuous point of the spectral factorization mapping, it is not even clear whether the approximation series $(\phi_N)_+ = \mathfrak{S}\,\phi_N = \mathfrak{S}\,\mathfrak{A}_N\,\phi$ always converges to the desired spectral factor ϕ_+ as N tends to infinity. However, for smooth spectral densities $\phi \in \mathcal{C}_\omega(\mathbb{T})$ with a weak regular majorant ω of type 1, we are able to show that the approximation series of the spectral factor converges to the actual ϕ_+, and the following theorem presents an upper bound on the error $\|\phi_+ - (\phi_N)_+\|_\infty$ induced by the approximation of the original spectrum. This upper bound of the error $\|\phi_+ - (\phi_N)_+\|_\infty$ decreases slightly slower than the approximation error $\|\phi - \phi_N\|_\infty$ as N increases.

Theorem 10.62 (Upper Bound). *Let ω be a weak regular majorant of type 1, let $\phi \in \mathcal{C}_\omega(\mathbb{T})$ with $0 < c_1 \le \phi(\zeta) \le c_2 < \infty$ for all $\zeta \in \mathbb{T}$, and let $\phi_N = \mathfrak{A}_N\phi$ be an approximation of ϕ by a method with properties (A), (B), and (C). Then there exist constants C_1, C_2, and C_3, which depend on c_1, c_2, and ω, such that*

$$\|\phi_+ - (\phi_N)_+\|_\infty \le (C_1 + C_2\,\log N)\ \|\phi - \phi_N\|_\infty + C_3\,\omega(\tfrac{\pi}{N})\ . \quad (10.102)$$

This upper bound on the convergence rate holds for all smooth densities ϕ with a weak regular majorant of type 1. This includes in particular the spectra of all completely regular stochastic processes with an order $0 < \alpha \le 1$. Thus, the error in the spectral factor $\|\phi_+ - (\phi_N)_+\|_\infty$ depends on a term $\omega(\pi/N)$ determined by the majorant ω of the spectra which approaches zero as N tends to infinity, and on a term which depends linearly on the approximation error $\|\phi - \phi_N\|_\infty$ and which is proportional to $\log N$. We know from Theorem 2.8 that the convergence rate of $\|\phi - \phi_N\|_\infty$ is upper bounded by $\omega(1/N)$. Thus, if an optimal approximation method is used the upper bound on the convergence rate of the spectral factor is proportional to $\omega(1/N)\,\log N$, i.e. it is worse by the factor $\log N$ compared with the approximation error $\|\phi - \phi_N\|_\infty$. However, the last section showed that due to the required positivity of the approximation operator, the optimal convergence rate of $\|\phi - \phi_N\|_\infty$ may not be achievable. In this case, the convergence rate of $\|\phi_+ - (\phi_N)_+\|_\infty$ is determined by the term $\log N\,\|\phi - \phi_N\|_\infty$. Thus it is determined by the used approximation method of ϕ but again, it is slightly worse by the factor $\log N$.

If for an actual approximation method $\mathfrak{A}_N$, an upper bound on the approximation error $\|\phi - \phi_N\|_\infty$ is known, one may express (10.102) in terms of the smoothness of the spectra and in terms of the approximation degree N only.

Example 10.63. Assume that the Fejér mean (cf. Sect. 2.1.2) is used as approximation method, and assume that the spectra are Hölder continuous of order $0 < \alpha < 1$, which means that the majorant ω is given by $\omega(\tau) = \tau^\alpha$. For this case, it is known [92, Chap. 3, § 13.32] that there exists a constant C such that $\|\phi - \phi_N\|_\infty \leq K\,\omega(1/N) = K\,N^{-\alpha}$ for all $\phi \in \mathcal{C}_\omega(\mathbb{T})$. Therewith, the upper bound for the error in the spectral factor becomes

$$\|\phi_+ - (\phi_N)_+\|_\infty \leq [K_1 + K_2 \log N]\ N^{-\alpha}$$

with certain constants K_1, K_2. Thus, the upper bound is worse than the lower bound by the factor $\log N$.

Proof (Theorem 10.62). On the unit circle $\mathbb{T}$, the spectral factor can be written as $\phi_+(\mathrm{e}^{\mathrm{i}\theta}) = \sqrt{\phi(\mathrm{e}^{\mathrm{i}\theta})}\,\mathrm{e}^{\mathrm{i}\arg\phi_+(\mathrm{e}^{\mathrm{i}\theta})}$. Therefore, a straight forward calculation shows that

$$\left|\phi_+(\mathrm{e}^{\mathrm{i}\theta}) - (\phi_N)_+(\mathrm{e}^{\mathrm{i}\theta})\right| \leq \left|\sqrt{\phi(\mathrm{e}^{\mathrm{i}\theta})} - \sqrt{\phi_N(\mathrm{e}^{\mathrm{i}\theta})}\right| + \sqrt{\phi_N(\mathrm{e}^{\mathrm{i}\theta})}\left|\mathrm{e}^{\mathrm{i}p(\mathrm{e}^{\mathrm{i}\theta})} - 1\right|$$

wherein $p(\mathrm{e}^{\mathrm{i}\theta}) = \arg\phi_+(\mathrm{e}^{\mathrm{i}\theta}) - \arg(\phi_N)_+(\mathrm{e}^{\mathrm{i}\theta})$. Next, it is used that $|\mathrm{e}^{\mathrm{i}p} - 1| = 2\,|\sin(p/2)|$ and that $|\sqrt{\phi} - \sqrt{\phi_N}| = |\phi - \phi_N|/(\sqrt{\phi} + \sqrt{\phi_N})$. Together with the maximum modulus principle for analytic functions one obtains that

$$\|\phi_+ - (\phi_N)_+\|_\infty \leq \frac{1}{2\sqrt{c_1}}\,\|\phi - \phi_N\|_\infty + \sqrt{c_2}\,\|p\|_\infty\ . \tag{10.103}$$

By definition (10.8) of the spectral factor, the function p is given by the Hilbert transform of the function $h(\mathrm{e}^{\mathrm{i}\theta}) := \frac{1}{2}[\log\phi(\mathrm{e}^{\mathrm{i}\theta}) - \log\phi_N(\mathrm{e}^{\mathrm{i}\theta})]$, i.e

$$p(\mathrm{e}^{\mathrm{i}\theta}) = \lim_{\varepsilon\to 0}\frac{1}{2\pi}\int_{\varepsilon<|\theta-\tau|\leq\pi} h(\mathrm{e}^{\mathrm{i}\tau})\,\frac{1}{\tan\frac{\theta-\tau}{2}}\,\mathrm{d}\tau\ . \tag{10.104}$$

Since we assumed that $\phi \in \mathcal{C}_\omega(\mathbb{T})$ with a weak regular majorant ω, Lemma 10.60 shows that $\phi_N \in \mathcal{C}_\omega(\mathbb{T})$, and since $\phi \geq c_1 > 0$ the approximation ϕ_N is also strictly positive, by the assumption on the approximation operator $\mathfrak{A}_N$. It is easily verified, therefore, that also $\log\phi$ and $\log\phi_N$ and consequently also h belong to $\mathcal{C}_\omega(\mathbb{T})$. Next we apply Lemma 5.10 which states that the Hilbert transform of every function $f \in \mathcal{C}_\omega(\mathbb{T})$ with a weak regular majorant ω of type 1 exists and is continuous. Therefore, the integral in (10.104) exists for $\varepsilon \to 0$ and p is a continuous function on $\mathbb{T}$. Next, we derive an upper bound on $|p|$. To this end, the integral (10.104) is split up into an integration over all τ with $|\theta - \tau| < \pi/N$ and an integration over all τ with $|\theta - \tau| \geq \pi/N$. Since the kernel $1/\tan\frac{\theta-\tau}{2}$ is an odd function with respect to $\tau = \theta$, we can subtract the constant $h(\theta)$ from the argument of the first integral (i.e. the integration over $|\theta - \tau| < \pi/N$) without changing its value. All this, together with the triangle inequality gives

$$\left|p(\mathrm{e}^{\mathrm{i}\theta})\right| \leq \underbrace{\frac{1}{2\pi}\int_{|\theta-\tau|\leq\frac{\pi}{N}}\frac{\left|h(\mathrm{e}^{\mathrm{i}\tau}) - h(\mathrm{e}^{\mathrm{i}\theta})\right|}{\left|\tan\frac{\theta-\tau}{2}\right|}\,\mathrm{d}\tau}_{=:T_1} + \underbrace{\frac{1}{2\pi}\int_{\frac{\pi}{N}\leq|\theta-\tau|\leq\pi}\frac{\left|h(\mathrm{e}^{\mathrm{i}\tau})\right|}{\left|\tan\frac{\theta-\tau}{2}\right|}\,\mathrm{d}\tau}_{=:T_2} \tag{10.105}$$

for the modulus of p. First, we analyze the second term T_2. Using the relation $|\log(x/y)| \leq (1/c)|x-y|$ in which $c = min(x,y)$ and which is easily verified[9], one finds that

$$T_2 \leq \frac{1}{2\,c_1}\,\|\phi - \phi_N\|_\infty\,\frac{1}{2\pi}\int_{\frac{\pi}{N}}^{\pi}\frac{\mathrm{d}\tau}{\tan(\tau/2)}$$
$$\leq \frac{1}{2\,\pi\,c_1}\,\|\phi - \phi_N\|_\infty\ \log(2N/\pi)\ . \tag{10.106}$$

To obtain an upper bound for T_1, we notice that

$$\left|h(\mathrm{e}^{\mathrm{i}\tau}) - h(\mathrm{e}^{\mathrm{i}\theta})\right| \leq \tfrac{1}{2}\left|\log\phi(\mathrm{e}^{\mathrm{i}\tau}) - \log\phi(\mathrm{e}^{\mathrm{i}\theta})\right| + \tfrac{1}{2}\left|\log\phi_N(\mathrm{e}^{\mathrm{i}\tau}) - \log\phi_N(\mathrm{e}^{\mathrm{i}\theta})\right|\ .$$

Therewith, T_1 is upper bounded by

$$T_1 \leq \frac{1}{4\pi}\int_{|\theta-\tau|\leq\frac{\pi}{N}}\frac{\left|\log\phi(\mathrm{e}^{\mathrm{i}\tau}) - \log\phi(\mathrm{e}^{\mathrm{i}\theta})\right|}{\left|\tan\frac{\theta-\tau}{2}\right|}\,\mathrm{d}\tau +$$
$$+\frac{1}{4\pi}\int_{|\theta-\tau|\leq\frac{\pi}{N}}\frac{\left|\log\phi_N(\mathrm{e}^{\mathrm{i}\tau}) - \log\phi_N(\mathrm{e}^{\mathrm{i}\theta})\right|}{\left|\tan\frac{\theta-\tau}{2}\right|}\,\mathrm{d}\tau\ .$$

Since $\phi \in \mathcal{C}_\omega(\mathbb{T})$ with $\phi(\zeta) \geq c_1 > 0$ for all $\zeta \in \mathbb{T}$, $\log\phi$ also belongs to $\mathcal{C}_\omega(\mathbb{T})$. Moreover, Lemma 10.60 shows that ϕ_N as well as $\log\phi_N$ are elements of $\mathcal{C}_\omega(\mathbb{T})$. Therewith, the upper bound for T_1 becomes

$$T_1 \overset{(a)}{\leq} \frac{1}{4\pi}\int_{|\theta-\tau|\leq\frac{\pi}{N}}\frac{2\,\omega(|\tau-\theta|)}{\left|\tan\frac{\theta-\tau}{2}\right|}\,\mathrm{d}\tau \overset{(b)}{\leq} \frac{1}{\pi}\int_0^{\pi/N}\frac{\omega(x)}{\tan\frac{x}{2}}\,\mathrm{d}x$$
$$\overset{(c)}{\leq} \frac{2}{\pi}\int_0^{\pi/N}\frac{\omega(x)}{x}\,\mathrm{d}x \overset{(d)}{\leq} \tfrac{2}{\pi}\,C\,\omega(\tfrac{\pi}{N})\ . \tag{10.107}$$

Therein, (a) follows from $\log\phi, \log\phi_N \in \mathcal{C}_\omega(\mathbb{T})$. For (b) it was used that the integrand is an even function, and (c) follows from the inequality $\tan x \geq x$. Since ω is assumed to be weak regular, one obtains (d) with a certain constant C. Finally, one has to use the upper bounds (10.107) and (10.106) in (10.105). Then (10.103) together with (10.105) gives the desired upper bound (10.102). □

Sampled Data

We still investigate the approximative determination of the spectral factor ϕ_+ of a spectral density ϕ by approximating first ϕ by a trigonometric polynomial ϕ_N and determine afterward the spectral factor $(\phi_N)_+$ of the approximative polynomial ϕ_N. Up to now, we assumed implicitly that the density $\phi(\mathrm{e}^{\mathrm{i}\theta})$

[9] See footnote on page 207.

is given on the whole unit circle, i.e. for all $\theta \in [-\pi.\pi)$. However, if a digital computer is used to calculate the approximation $\phi \mapsto \phi_N$, only a finite number of sampling points $\phi(\mathrm{e}^{\mathrm{i}\theta_k})$, $k = 1, 2, \ldots, M$ can be taken into account for the determination of ϕ_N and $(\phi_N)_+$. Whereas in the case that ϕ is know on the whole unit circle, the error in the spectral factor could be controlled by the approximation degree N (cf. Theorem 10.62), will will show now that this is no longer possible if ϕ is known only at discrete sampling points.

For technical reasons, the spectral factorization mapping is considered only on the space of all spectral $\phi \in \mathcal{C}(\mathbb{T})$ which posses a continuous spectral factor $\phi_+ \in \mathcal{C}(\mathbb{T})$, i.e. on the space

$$\mathcal{B}(\mathbb{T}) = \{\phi \in \mathcal{C}(\mathbb{T}) \ : \ \phi_+ \in \mathcal{C}(\mathbb{T})\} .$$

Clearly, this is no restriction on the generality, because if the spectral factor ϕ_+ is not continuous, a uniform approximation of ϕ_+ would not be possible, anyway.

Since ϕ is given only on a discrete sampling set, we can no longer use the approximation methods of Section 10.5. Instead, we consider in the following sequences $\{\mathfrak{A}_N\}_{N\in\mathbb{N}}$ of *linear* approximation operators, with the following three properties

(a) *Concentration on a discrete sampling set*: Let $N \in \mathbb{N}$ and let $\mathcal{S}_N = \{\theta_1[N], \theta_2[N], \cdots, \theta_{M(N)}[N]\}$ be the set of sampling points $\theta_k[N] \in [-\pi, \pi)$, $k = 1, 2, \cdots, M(N)$. Then for every two densities ϕ_1 and ϕ_2 which coincide on $\mathcal{S}_N$, i.e. for which

$$\phi_1(\mathrm{e}^{\mathrm{i}\theta_k[N]}) = \phi_2(\mathrm{e}^{\mathrm{i}\theta_k[N]}) \qquad \text{for all} \qquad k = 1, 2, \cdots, M(N)$$

the approximation operators $\mathfrak{A}_N$ should give the same result, i.e $\mathfrak{A}_N\phi_1 = \mathfrak{A}_N\phi_2$.

(b) *Positivity*: If the spectrum ϕ is non-negative at all sampling points, i.e. $\phi(\mathrm{e}^{\mathrm{i}\theta_k[N]}) \geq 0$, $k = 1, 2, \cdots, M(N)$, the approximation $\phi_N(\mathrm{e}^{\mathrm{i}\theta}) = (\mathfrak{A}_N\phi)(\mathrm{e}^{\mathrm{i}\theta})$ should be non-negative at all $\theta \in [-\pi, \pi)$.

(c) *Perfect approximation on* $\mathcal{B}(\mathbb{T})$: The approximation method should be perfect for all functions in $\mathcal{B}(\mathbb{T})$, i.e.

$$\lim_{N\to\infty} \|\mathfrak{A}_N\phi - \phi\|_\infty = 0 \quad \text{for all } \phi \in \mathcal{B}(\mathbb{T}) .$$

This is a minimum requirement for any useful approximation method. If the approximation $\mathfrak{A}_N\phi$ already does not converge to ϕ, one cannot expect that $(\mathfrak{A}_N\phi)_+$ to converge to ϕ_+. Note that by the Theorem of Banach-Steinhaus, this assumption on the approximation method implies that the norm of operators $\mathfrak{A}_N$ are uniformly bounded.

Because the approximation operator $\mathfrak{A}_N$ is assumed to be concentrated on the sampling set $\mathcal{S}_N$, one can give a simple canonical form of every $\mathfrak{A}_N$. To every $N > 0$, we consider functions $\eta_{k,N} \in \mathcal{B}(\mathbb{T})$ with the property that

$$\eta_{k,N}(\theta_l[N]) = \begin{cases} 1, & l = k \\ 0, & l \neq k \end{cases}$$

and define the functions $G_{k,N}(\theta) := (\mathfrak{A}_N \eta_{k,N})(\mathrm{e}^{\mathrm{i}\theta})$ with $k = 1, 2, \cdots, M(N)$ and for $\theta \in [-\pi, \pi)$. Therewith, it is clear that the approximation of an arbitrary spectral density $\phi \in \mathcal{B}(\mathbb{T})$ can be written as

$$(\mathfrak{A}_N \phi)(\mathrm{e}^{\mathrm{i}\theta}) = \textstyle\sum_{k=1}^{M(N)} \phi(\mathrm{e}^{\mathrm{i}\theta_k[N]})\, G_{k,N}(\theta)$$

and from the required positivity (b) of $\mathfrak{A}_N$ follows that $G_{k,N}(\theta)$ is non-negative and satisfies the Paley-Wiener condition. Additionally, we require that the constant function $1(\mathrm{e}^{\mathrm{i}\theta}) = 1$ for all $\theta \in [-\pi, \pi)$ is always approximated by the constant function, i.e. we require that $(\mathfrak{A}_N\, 1)(\mathrm{e}^{\mathrm{i}\theta}) = 1$ for all $\theta \in [-\pi, \pi)$ and for all N. This property of $\mathfrak{A}_N$ clearly implies that $\sum_{k=1}^{M(N)} G_{k,N}(\theta) = 1$ for all θ and N, and that

$$(\mathfrak{A}_N \phi)(\mathrm{e}^{\mathrm{i}\theta}) \geq \min_k \phi(\mathrm{e}^{\mathrm{i}\theta_k[N]}) \quad \text{and} \quad (\mathfrak{A}_N \phi)(\mathrm{e}^{\mathrm{i}\theta}) \leq \max_k \phi(\mathrm{e}^{\mathrm{i}\theta_k[N]})$$

for all $\theta \in [-\pi, \pi)$.

With these preparations we can prove the following theorem, which shows that the error $\|\phi_+ - (\phi_N)_+\|_\infty$ in the spectral factor cannot be controlled by the approximation degree N if the given density is known only on a discrete sampling set.

Theorem 10.64. *Let* $\{\mathfrak{A}_N\}_{N\in\mathbb{N}}$ *be a sequence of linear discrete approximation operators with properties* (a), (b), *and* (c). *Then there exists no constant* $C < \infty$ *such that*

$$\|\phi_+ - (\mathfrak{A}_N \phi)_+\|_\infty \leq C\, \|\phi - \mathfrak{A}_N \phi\|_\infty \tag{10.108}$$

for every $\phi \in \mathcal{B}(\mathbb{T})$ *with* $\|\phi\|_\infty \leq 1$ *and with* $\phi(\mathrm{e}^{\mathrm{i}\theta}) \geq c_1 > 0$.

Remark 10.65. Thus the relative error in the spectral factor is unbounded in $\mathcal{B}(\mathbb{T})$, i.e. if one defines the set $\mathcal{B}_{\mathrm{pos}} := \{\phi \in \mathcal{B}(\mathbb{T}) : \phi(\mathrm{e}^{\mathrm{i}\theta}) > 0\ ,\ \|\phi\|_\infty \leq 1\}$, then the theorem shows that

$$\lim_{N\to\infty} \sup_{\phi \in \mathcal{B}_{\mathrm{pos}}} \frac{\|\phi_+ - (\mathfrak{A}_N \phi)_+\|_\infty}{\|\phi - \mathfrak{A}_N \phi\|_\infty} = \infty\ . \tag{10.109}$$

To see this, assume that (10.109) does not hold. Then there would exist a constant K and a subsequence N_k such that

$$\lim_{k\to\infty} \sup_{\phi \in \mathcal{B}_{\mathrm{pos}}} \frac{\|\phi_+ - (\mathfrak{A}_{N_k} \phi)_+\|_\infty}{\|\phi - \mathfrak{A}_{N_k} \phi\|_\infty} \leq K\ .$$

But then we could define the sequence $\mathfrak{B}_k := \mathfrak{A}_{N_k}$, $k = 1, 2, \ldots$ of operators which possess all properties (a), (b), and (c) assumed in the above Theorem 10.64. However, this would imply that $\|\phi_+ - (\mathfrak{B}_k \phi)_+\|_\infty \leq K\, \|\phi - \mathfrak{B}_k \phi\|_\infty$ for all $\phi \in \mathcal{B}_{\mathrm{pos}}$ and for all k, which contradicts the statement of Theorem 10.64.

The theorem shows that even if we have an approximation operator $\mathfrak{A}_N$ which approximates the function ϕ as close as desired from the sampling points, it will in general not be possible to control the error in the spectral factor by the number N of sampling points. Thus, assume ϵ is the maximal allowed error in the spectral factor. Then it is not possible to find a finite number N of sampling points such that $\|\phi_+ - (\mathfrak{A}_N\phi)_+\|_\infty < \epsilon$. There always exists functions $\phi \in \mathcal{B}(\mathbb{T})$ which violate the desired error bound.

Proof (Theorem 10.64). Let $c_1 > 0$ be an arbitrary but fixed constant. In contradiction to the statement of the theorem, we assume that there exists a constant $C < \infty$ such that (10.108) holds for all $\phi \in \mathcal{C}^\infty(\mathbb{T})$ with $\|\phi\|_\infty \le 1$ and $\phi \ge c_1$. Note that $\mathcal{C}^\infty(\mathbb{T})$ is a dense subset of $\mathcal{B}(\mathbb{T})$. Then it holds for all such functions $\phi \in \mathcal{C}^\infty(\mathbb{T})$ that

$$\limsup_{N\to\infty} \frac{\|\phi_+ - (\mathfrak{A}_N\phi)_+\|_\infty}{\|\phi - \mathfrak{A}_N\phi\|_\infty} \le C\,. \tag{10.110}$$

Let $\epsilon > 0$ and let $g \in \mathcal{C}^\infty(\mathbb{T})$ such that the function $\psi := 1 - \epsilon\, g$ satisfies $\psi(\mathrm{e}^{\mathrm{i}\theta}) \ge c_1$ for all $\theta \in [-\pi, \pi)$. Clearly, for sufficiently small ϵ this can always be achieved. Since $\mathfrak{A}_N$ is linear and because we assumed that $\mathfrak{A}_N\, 1 = 1$, it holds that $\mathfrak{A}_N\psi = 1 - \epsilon\,(\mathfrak{A}_N g) = 1 - \epsilon\, g_N$ where $g_N = \mathfrak{A}_N g$. Therewith, it is clear that $\|\psi - \psi_N\|_\infty = \epsilon\,\|g - g_N\|_\infty$ and consequently

$$\frac{\|\psi_+ - (\psi_N)_+\|_\infty}{\|\psi - \psi_N\|_\infty} = \frac{1}{\epsilon}\,\frac{\|\psi_+ - (\psi_N)_+\|_\infty}{\|g - g_N\|_\infty}\,. \tag{10.111}$$

Moreover, for the error in the spectral factor holds

$$\|\psi_+ - (\psi_N)_+\|_\infty \ge \tfrac{2}{\pi}\,\sqrt{c_1}\,\|p\|_\infty - \tfrac{1}{2\sqrt{c_1}}\,\|\psi - \psi_N\|_\infty \tag{10.112}$$

in which $p := \arg(\psi_N)_+ - \arg\psi_+$. This lower bound is obtained in a similar way as the upper bound (10.103) in the proof of Theorem 10.62 but using the lower bound $|\,\|x\| - \|y\|\,| \le \|x + x\|$ of the triangle inequality instead of the upper bound $\|x + y\| \le \|x\| + \|y\|$. Combining (10.112) with (10.111) gives

$$\frac{\|\psi_+ - (\psi_N)_+\|_\infty}{\|\psi - \psi_N\|_\infty} \ge \frac{2\sqrt{c_1}}{\epsilon\,\pi}\,\frac{\|p\|_\infty}{\|g - g_N\|_\infty} - \frac{1}{2\sqrt{c_1}}\,.$$

Together with the assumption (10.110), one obtains the inequality

$$\frac{\pi}{2\sqrt{c_1}}\left(C + \frac{1}{2\sqrt{c_1}}\right)\|g - g_N\|_\infty \ge \frac{1}{\epsilon}\,\|p\|_\infty\,. \tag{10.113}$$

Next we analyze the term $\|p\|_\infty$. By the definition of the spectral factor, the function p is equal to the Hilbert transform of $h := (\log\psi_N - \log\psi)/2$ (cf. also the proof of Theorem 10.62). By the definition of ψ and using the power series expansion of the logarithm, one obtains that

$$h(\mathrm{e}^{\mathrm{i}\tau}) = \frac{1}{2}[\log \psi_N(\mathrm{e}^{\mathrm{i}\tau}) - \log \psi(\mathrm{e}^{\mathrm{i}\tau})] = \frac{1}{2}\sum_{k=1}^{\infty}\frac{1}{k}\,\epsilon^k[g^k(\mathrm{e}^{\mathrm{i}\tau}) - g_N^k(\mathrm{e}^{\mathrm{i}\tau})]$$
$$= \frac{1}{2}\,\epsilon\,[g(\mathrm{e}^{\mathrm{i}\tau}) - g_N(\mathrm{e}^{\mathrm{i}\tau})] + \epsilon^2\,R_N(\epsilon,\tau)$$

in which the rest term R_N is given by

$$R_N(\epsilon,\tau) = \frac{1}{2}\left[\frac{1}{\epsilon}\log\frac{1-\epsilon\, g_N(\mathrm{e}^{\mathrm{i}\tau})}{1-\epsilon\, g(\mathrm{e}^{\mathrm{i}\tau})} - g(\mathrm{e}^{\mathrm{i}\tau}) + g_N(\mathrm{e}^{\mathrm{i}\tau})\right]$$
$$= \frac{1}{2}\sum_{k=2}^{\infty}\frac{1}{k}\,\epsilon^{k-2}[g^k(\mathrm{e}^{\mathrm{i}\tau}) - g_N^k(\mathrm{e}^{\mathrm{i}\tau})]\;.$$

Therewith, the function p can be written as

$$p(\mathrm{e}^{\mathrm{i}\theta}) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{h(\mathrm{e}^{\mathrm{i}\tau})}{\tan\frac{\theta-\tau}{2}}\,\mathrm{d}\tau$$
$$= \frac{\epsilon}{4\pi}\int_{-\pi}^{\pi}\frac{g(\mathrm{e}^{\mathrm{i}\tau}) - g_N(\mathrm{e}^{\mathrm{i}\tau})}{\tan\frac{\theta-\tau}{2}}\,\mathrm{d}\tau + \frac{\epsilon^2}{2\pi}\int_{-\pi}^{\pi}\frac{R_N(\epsilon,\tau)}{\tan\frac{\theta-\tau}{2}}\,\mathrm{d}\tau\;. \quad (10.114)$$

Since $g \in \mathcal{C}^\infty(\mathbb{T})$ the Hilbert transforms $\widetilde{g}$, $\widetilde{g}_N$ exist. Also the Hilbert transform of R_N exists. To see this, we consider the first derivate of R_N

$$R_N'(\epsilon,\tau) = \frac{1}{2\,\epsilon}\left[\frac{g'(\mathrm{e}^{\mathrm{i}\tau})}{1-\epsilon\, g(\mathrm{e}^{\mathrm{i}\tau})} - \frac{g_N'(\mathrm{e}^{\mathrm{i}\tau})}{1-\epsilon\, g_N(\mathrm{e}^{\mathrm{i}\tau})} - g'(\mathrm{e}^{\mathrm{i}\tau}) + g_N'(\mathrm{e}^{\mathrm{i}\tau})\right]\;.$$

Since $g \in \mathcal{C}^\infty(\mathbb{T})$, this shows that there exists a constant C_2 such that $|R_N'(\epsilon,\tau)| \leq C_2$ independent on ϵ and τ. Thus, $R_N(\epsilon,\cdot)$ is Lipschitz continuous, which implies that it is Hölder continuous of an order $\alpha < 1$, i.e. $\|R_N(\epsilon,\cdot)\|_{\Lambda_\alpha} < \infty$. Therefore, it follows from Theorem 5.11 that the mapping $f \mapsto \widetilde{f}$ is continuous on Λ_α. Consequently, there exists a constant C_3 such that

$$\|\widetilde{R}_N(\epsilon,\cdot)\|_\infty \leq \|\widetilde{R}_N(\epsilon,\cdot)\|_{\Lambda_\alpha} \leq C_3\,\|R_N(\epsilon,\cdot)\|_{\Lambda_\alpha}\;.$$

All this show that the integrals on the right hand side of (10.114) converge such that the function p becomes

$$p(\theta) = \tfrac{\epsilon}{2}\,[\widetilde{g}(\theta) - \widetilde{g}_N(\theta)] + \epsilon^2\,\widetilde{R}_N(\epsilon,\theta)\;.$$

Moreover, from (10.114) follows for its norm

$$\tfrac{1}{\epsilon}\,\|p\|_\infty = \tfrac{1}{2}\|[\widetilde{g}(\theta) - \widetilde{g}_N(\theta)] + 2\epsilon R_N(\epsilon,\cdot)\|_\infty\;.$$

Therewith, inequality (10.113) becomes for $\epsilon \to 0$ equal to

$$\frac{\pi}{\sqrt{c_1}}\left(C + \frac{1}{2\sqrt{c_1}}\right)\|g - g_N\|_\infty \geq \|\widetilde{g} - \widetilde{g}_N\|_\infty\;. \quad (10.115)$$

Since $g \in \mathcal{C}^\infty(\mathbb{T}) \subset \mathcal{B}(\mathbb{T})$ the approximation $g_N = \mathfrak{A}_N g$ converges to g as N tends to infinity by the required property (c) of the approximation method $\mathfrak{A}_N$. Therefore, the last inequality gives

$$0 = \limsup_{N\to\infty} \frac{\pi}{\sqrt{c_1}} \left(C + \frac{1}{2\sqrt{c_1}} \right) \|g - g_N\|_\infty \geq \limsup_{N\to\infty} \|\widetilde{g} - \widetilde{g}_N\|_\infty \,,$$

which shows that $\limsup_{N\to\infty} \|\widetilde{g} - \widetilde{g}_N\|_\infty = 0$ for all $g \in \mathcal{C}^\infty(\mathbb{T})$.

Now, define the linear operator

$$(\widetilde{\mathfrak{A}}_N g)(\mathrm{e}^{\mathrm{i}\theta}) := (\mathfrak{H}[\mathfrak{A}_N g])(\mathrm{e}^{\mathrm{i}\theta}) = \widetilde{g}_N(\mathrm{e}^{\mathrm{i}\theta})\,, \qquad \theta \in [-\pi, \pi)$$

which is concentrated on the sampling set $\mathcal{S}_N$. Such operators were studied in Section 9. The above deduction shows that under the assumption (10.110), the series $\widetilde{\mathfrak{A}}_N g$ converges to the Hilbert transform $\widetilde{g}$ for every function $g \in \mathcal{C}^\infty(\mathbb{T})$. However, it was shown in Corollary 9.8 that for every such operator series $\mathfrak{A}_N$, there exists a dense subset $\mathcal{X}$ in $\mathcal{B}(\mathbb{T})$ such that

$$\limsup_{N\to\infty} \|\widetilde{\mathfrak{A}}_N f\|_\infty = \infty \tag{10.116}$$

for all $f \in \mathcal{X}$. Since also $\mathcal{C}^\infty(\mathbb{T})$ is a dense subset of $\mathcal{B}(\mathbb{T})$, it follows from (10.116) that there exists a function $g^* \in \mathcal{C}^\infty(\mathbb{T}) \cap \mathcal{X}$ with $\|g^*\|_\infty \leq 1$ and with $\|\widetilde{g}^*\|_\infty \leq 2$ such that to every $K > 0$ there exists an N_0 such that $\|\widetilde{g}^*_{N_0}\|_\infty = \|\widetilde{\mathfrak{A}}_{N_0} g^*\|_\infty \geq K$, and by the triangle inequality therefore holds that

$$\|\widetilde{g}^* - \widetilde{g}^*_{N_0}\|_\infty \geq \left|\, \|\widetilde{g}^*_{N_0}\|_\infty - \|\widetilde{g}^*\|_\infty \,\right| \geq K - 2\,.$$

Since by property (c) of the approximation operator, the norms $\|\mathfrak{A}_N\|$ are uniformly bounded, it holds also that

$$\|g^* - g^*_{N_0}\|_\infty \;\leq\; \|g^*\|_\infty + \|g^*_{N_0}\|_\infty \;\leq\; 1 + \|\mathfrak{A}_{N_0}\| \;\leq\; C_3\,.$$

Using these inequalities in (10.115), one obtains that

$$\frac{\pi}{\sqrt{c_1}} \left(C + \frac{1}{2\sqrt{c_1}} \right) (1 - \|\mathfrak{A}_{N_0}\|) \;\geq\; K - 2\,.$$

Since K can be chosen arbitrarily, whereas the left hand side is a fixed constant, this gives a contradiction for sufficiently large K. Therefore, the assumption (10.110) cannot hold. This proves that there exists no constant $C < \infty$ such that (10.108) holds for all $\phi \in \mathcal{C}^\infty(\mathbb{T})$ with $\phi \geq c_1$. $\square$

Notes

There exist numerous excellence textbooks for an introduction to stochastic sequences e.g. [29, 33, 69, 81]. The characterization of completely regular stochastic sequences in terms of the regularity coefficients (10.5) goes back to Gelfand and Yaglom [42], following by a series of publications [46, 47, 49, 50, 56, 77] which yields the characterization given in Theorem 10.4.

The spectral factorization goes back to Wiener and Hopf [89]. Later it was applied by Wiener to estimation and prediction problems [88]. Nowadays it is a common tool in applied mathematics and engineering and is covered in many different textbooks, e.g. [29, 38, 53]. A survey of algorithms for spectral factorization may be found in [78]. Theorem 10.7 is due to L. Fejér and M. Riesz [37].

The $\mathcal{S}$–algebras, we used in this chapter, are similar to the so called *decomposing Banach algebras* introduces by Clancey and Gohberg [27]. Peller and Khrushchev gave in [66] a similar systems of axioms as in Def. 10.11. These axioms were also used by Jacob and Partington in [52].

The discontinuity of the spectral factorization mapping on L^∞, $\mathcal{C}(\mathbb{T})$ was shown by Anderson [4] but it follows also from the counterexample given by Treil [85]. A detailed investigation of the continuity and boundedness of the spectral factorization mapping on decomposing Banach algebras was done by Jacob and Partington in [52]. In particular, they showed that the spectral factorization is continuous on every decomposing Banach algbera. The unboundedness of the spectral factorization on decomposing Banach algebras was proved in [17] using ideas from [52]. The results concerning the spectral factorization on $\mathcal{S}$-algebras were presented in [21]. The two particular examples discussed in some detail in Section 10.3, i.e. the factorization on $\mathcal{C}(\mathbb{T})$ and $\mathcal{W}$, were taken from [19] and [10], respectively. Lemma 10.35 is a generalization of a classical result of Wiener [87]. He showed the statement of Lemma 10.35 for the special function $G(z) = 1/z$.

The error bounds for polynomial data in Section 10.4 were derived in [15] for the Wiener algebra and in [14, 16] Lemma 10.46 due to van der Corput appeared in [86]. The polynomials (10.54) used in Section 10.4.1 are also known as *chirp sequences* or *sequences with quadratic phase*. They appear for example in design signals with a flat power spectrum and with a low peak value [80]. The presentation in Section 10.5 and 10.6 is basically taken from [13].

Appendix – Proof of Lemma 10.37

This appendix contain the proof of Lemma 10.37. This proof is based on four preparatory lemmas. These lemmas (especially the first two) are almost classical and can also be found elsewhere, in slightly different form (see e.g. [35]).

Lemma 10.66. *Let $f, g \in \mathcal{W}$ and let $s_n(z) := z^n$, $z \in \mathbb{D}$, then*

$$\lim_{n\to\infty} \|f + s_n\, g\|_{\mathcal{W}} = \|f\|_{\mathcal{W}} + \|g\|_{\mathcal{W}} \ . \tag{10.117}$$

A similar results holds of course for $n \to -\infty$.

Remark 10.67. From the above lemma follows for every sequence $\{f_k\}_{k\in\mathbb{Z}}$ of functions in $\mathcal{W}$ with $\sum_k \|f_k\|_{\mathcal{W}} < \infty$ immediately that

$$\lim_{n\to\infty} \left\|\sum_{k=-\infty}^{\infty} s_{k\cdot n}\, f_k\right\|_{\mathcal{W}} = \sum_{k=-\infty}^{\infty} \|f_k\|_{\mathcal{W}} \ .$$

Proof. Let $f(\zeta) = \sum_{k=-\infty}^{\infty} \hat{f}(k)\,\zeta^k$ and $g(\zeta) = \sum_{k=-\infty}^{\infty} \hat{g}(k)\,\zeta^k$ where $\zeta \in \mathbb{T}$ be two arbitrary functions in $\mathcal{W}$ then it is clear that $(s_n\, g)(\zeta) = \sum_{k=-\infty}^{\infty} g_{k-n}\,\zeta^k$. Using the triangle inequality, one obtains

$$\|f + s_n\, g\|_{\mathcal{W}} \leq \sum_{k=-\infty}^{\infty}(|\hat{f}(k)| + |\hat{g}(k-n)|) = \|f\|_{\mathcal{W}} + \|g\|_{\mathcal{W}}$$

It remains to show that equality is achieved as $n \to \infty$. To this end, let $\epsilon > 0$ then there exists a constant $N_0 = N_0(\epsilon)$ such that $\sum_{k=N_0+1}^{\infty} |\hat{f}(k)| < \epsilon$ and $\sum_{k=-\infty}^{-(N_0+1)} |\hat{f}(k)| < \epsilon$ and $\sum_{k=N_0+1}^{\infty} |\hat{g}(k)| < \epsilon$ and $\sum_{k=-\infty}^{-(N_0+1)} |\hat{g}(k)| < \epsilon$. Moreover, for such a fixed $N_0 = N_0(\epsilon)$ there exists an $n_0 = n_0(\epsilon)$ such that for all $|k| < N_0$

$$|\hat{g}(k-n)| < \frac{\epsilon}{2\,N_0 + 1} \qquad \text{for all } n \geq n_0(\epsilon) \ .$$

Now, we consider for $n \geq n_0(\epsilon)$ the obvious equation

$$\begin{aligned} \|f + s_n\, g\|_{\mathcal{W}} = & \underbrace{\sum_{k=-\infty}^{-(N_0+1)} |\hat{f}(k) - \hat{g}(k-n)|}_{T_1} + \\ & + \underbrace{\sum_{k=-N_0}^{N_0} |\hat{f}(k) - \hat{g}(k-n)|}_{T_2} + \underbrace{\sum_{k=N_0+1}^{\infty} |\hat{f}(k) - \hat{g}(k-n)|}_{T_3} \end{aligned} \tag{10.118}$$

and analyze the three terms on the right hand side separately. From the choice of N_0, one obtains for the first summand that

$$T_1 \leq \sum_{k=-\infty}^{-(N_0+1)} |\hat{f}(k)| + \sum_{k=-\infty}^{-(N_0+1+n)} |\hat{g}(k)| \leq 2\epsilon \ .$$

For T_2, the triangle inequality gives

$$T_2 \geq \left| \sum_{k=-N_0}^{N_0} |\hat{f}(k)| - \sum_{k=-N_0}^{N_0} |\hat{g}(k-n)| \right| \ .$$

Now, the choice of N_0 shows that the first sum on the right hand side is larger or equal than $\|f\|_{\mathcal{W}} - 2\epsilon$ whereas the choice of $n > n_0$ shows that $\sum_{k=-N_0}^{N_0} |\hat{g}(k-n)| \leq (2\,N_0+1)\frac{\epsilon}{2\,N_0+1}$. Therewith one obtains the lower bound

$$T_2 \geq \|f\|_{\mathcal{W}} - 3\epsilon$$

for T_2. With similar arguments, the triangle inequality and the choice of N_0 and n_0 show that

$$T_3 \geq \left| \sum_{k=N_0+1-n}^{\infty} |\hat{g}(k)| - \sum_{k=N_0+1}^{\infty} |\hat{f}(k)| \right| \geq \|g\|_{\mathcal{W}} - 2\epsilon \ .$$

Using the bounds for T_1, T_2, and T_3 in (10.118), one obtains that

$$\|f + s_n\, g\|_{\mathcal{W}} \ \geq \ \|f\|_{\mathcal{W}} + \|g\|_{\mathcal{W}} - 7\epsilon \ .$$

Since ϵ was chosen arbitrary, this proves (10.117). □

Lemma 10.68. *Let $f, g \in \mathcal{W}$ and let $f_n(z) := f(z^n)$. Then*

$$\lim_{n\to\infty} \|f_n\, g\|_{\mathcal{W}} = \|f\|_{\mathcal{W}}\, \|g\|_{\mathcal{W}} \ .$$

Proof. For this proof, we use Lemma 10.66 and the functions $s_n(z) := z^n$. Note that $\|s_n\|_{\mathcal{W}} = 1$ for every $n \in \mathbb{N}$. Let $f(\zeta) = \sum_{k=-\infty}^{\infty} \hat{f}(k)\, \zeta^k$, $\zeta \in \mathbb{T}$ then

$$f_n(\zeta)\, g(\zeta) = \left(\sum_{k=-\infty}^{\infty} \hat{f}(k)\, \zeta^{k\cdot n} \right) g(\zeta) = \sum_{k=-\infty}^{\infty} h_k(\zeta)\, \zeta^{k\, n} = \sum_{k=-\infty}^{\infty} s_{k\cdot n}(\zeta)\, h_k(\zeta)$$

with $h_k(\zeta) := \hat{f}(k)\, g(\zeta) \in \mathcal{W}$, for which $\|h_k\|_{\mathcal{W}} = |\hat{f}(k)|\, \|g\|_{\mathcal{W}}$. Now Remark 10.67 to Lemma 10.66 is applied. This gives

$$\lim_{n\to\infty} \|f_n\, g\|_{\mathcal{W}} = \lim_{n\to\infty} \| \sum_{k=-\infty}^{\infty} s_{k\cdot n} h_k \|_{\mathcal{W}}$$

$$= \sum_{k=-\infty}^{\infty} \|h_k\|_{\mathcal{W}} = \sum_{k=-\infty}^{\infty} |\hat{f}(k)|\, \|g\|_{\mathcal{W}} = \|f\|_{\mathcal{W}}\, \|g\|_{\mathcal{W}} \, .$$

□

Lemma 10.69. *To every $\epsilon > 0$ and $A > 0$ there exists a real valued function $u \in \mathcal{W}$ such that $\|u\|_{\mathcal{W}} = 1$ and such that*

$$\|\mathrm{e}^{\mathrm{i}Au}\|_{\mathcal{W}} > (1 - \epsilon)\, \mathrm{e}^{A} \ . \qquad (10.119)$$

Proof. Let $u_0(\zeta) := \frac{1}{2}(\zeta + \zeta^{-1})$, which simply means that $u_0(\mathrm{e}^{\mathrm{i}\omega}) = \cos\omega$, and define

$$u(\zeta) = \tfrac{1}{N} \sum_{k=1}^{N} u_0(\zeta^{m(k)}) \qquad (\zeta \in \mathbb{T})$$

with certain integers $m(1) < m(2) < \cdots < m(N)$ which properties are characterized subsequently. It is clear that for the so defined function u holds that $\|u\|_{\mathcal{W}} = 1$, and that

$$\exp(\mathrm{i}A\, u(z)) = \exp\left(\mathrm{i}\tfrac{A}{N} \sum_{k=1}^{N} u_0(\zeta^{m(k)})\right) = \prod_{k=1}^{N} \exp\left(\mathrm{i}\tfrac{A}{N} u_0(\zeta^{m(k)})\right) \ .$$

Now, Lemma 10.68 shows that to every $\epsilon > 0$ there exist sufficiently large integers $m(1) < \cdots < m(N)$ such that

$$\|e^{iA\,u}\|_{\mathcal{W}} \geq (1-\epsilon)^{1/2}\big(\|e^{i\frac{A}{N}\,u_0}\|_{\mathcal{W}}\big)^N \ . \tag{10.120}$$

On the other hand, applying the Taylor expansion of the exponential function, one has that

$$\begin{aligned}\|e^{i\frac{A}{N}u_0} - (1 + i\tfrac{A}{N}u_0)\|_{\mathcal{W}} &= \|\textstyle\sum_{k=2}^{\infty} \frac{1}{k!}(i\frac{A}{N}u_0)^k\|_{\mathcal{W}} \\ &= \textstyle\sum_{k=2}^{\infty} \frac{1}{k!}(\frac{A}{N})^k \leq (\frac{A}{N})^2 \sum_{k=0}^{\infty} \frac{1}{k!}(\frac{A}{N})^k \ .\end{aligned}$$

Together with the triangle inequality, the last relation gives

$$\|e^{i\frac{A}{N}u_0}\|_{\mathcal{W}} \geq \|1 + i\tfrac{A}{N}\,u_0\|_{\mathcal{W}} - (\tfrac{A}{N})^2\,e^{A/N} = 1 + \tfrac{A}{N} - (\tfrac{A}{N})^2\,e^{A/N}$$

and consequently

$$\|e^{i\frac{A}{N}u_0}\|_{\mathcal{W}}^N \geq \big(1 + \tfrac{A}{N} - (\tfrac{A}{N})^2\,e^{A/N}\big)^N \geq (1-\epsilon)^{1/2}\,e^A$$

for sufficiently large N. Finally, we have to combine this last inequality with (10.120) to obtain (10.119). □

Lemma 10.70. *Let* $\{f_n\}_{n\in\mathbb{N}}$ *and* $\{g_n\}_{n\in\mathbb{N}}$ *be two sequences of uniformly bounded functions in an arbitrary Banach algebra* $\mathcal{B}$ *for which holds that* $\lim_{n\to\infty}\|f_n - g_n\|_{\mathcal{B}} = 0$. *Then for every* $k \in \mathbb{N}$ *holds that*

$$\lim_{n\to\infty} \|(f_n)^k - (g_n)^k\|_{\mathcal{B}} = 0 \ . \tag{10.121}$$

Proof. The proof is done by induction. For $k = 1$, the statement is trivial. Assume that (10.121) holds for $k-1$, then the triangle inequality and the continuity of the multiplication on $\mathcal{B}$ gives

$$\begin{aligned}\|(f_n)^k - (g_n)^k\|_{\mathcal{B}} &\leq \|f_n(f_n^{k-1} - g_n^{k-1})\|_{\mathcal{B}} + \|g_n^{k-1}(f_n - g_n)\|_{\mathcal{B}} \\ &\leq \|f_n\|_{\mathcal{B}}\|f_n^{k-1} - g_n^{k-1}\|_{\mathcal{B}} + \|g_n\|_{\mathcal{B}}^{k-1}\|f_n - g_n\|_{\mathcal{B}} \ .\end{aligned}$$

The right hand side goes to zero for $n \to \infty$ since (10.121) holds for $k-1$ and for $k = 1$, and because $\|f_n\|_{\mathcal{B}}$ and $\|g_n\|_{\mathcal{B}}$ are uniformly bounded. □

After these preparations, we are able to prove Lemma 10.37.

Proof (Lemma 10.37). Let $\delta > 0$, set $\epsilon = \delta/2$, and choose a $B > \log(2/\mu) > 0$, then Lemma 10.69 shows that there exists a $u \in \mathcal{W}$ with $\|u\|_{\mathcal{W}} = 1$ such that

$$\|e^{i\,B\,N\,u}\|_{\mathcal{W}} \geq (1-\epsilon)\,e^{B\,N} \ .$$

With this u, we define the function $Q(z) := e^{iBu(z)}$. Since B and u are real, it follows that $\|Q\|_\infty = 1$. Then, for all $1 \leq k \leq N$ holds that

$$(1-\epsilon)\, e^{B\,N} \leq \|e^{\mathrm{i}\,B\,k\,u}\|_{\mathcal{W}}\, \|e^{\mathrm{i}\,B\,(N-k)\,u}\|_{\mathcal{W}} \leq \|e^{\mathrm{i}\,B\,k\,u}\|_{\mathcal{W}}\, e^{B(N-k)}$$

For $k=1$ this gives $\|Q\|_{\mathcal{W}} \geq (1-\epsilon)\, e^{B}$, and for all other k, one has

$$\|Q^k\|_{\mathcal{W}} \geq (1-\epsilon)\, e^{B\,k} \geq (1-\epsilon)\, \|Q\|_{\mathcal{W}}^k\ . \tag{10.122}$$

Next, we define for an arbitrary $n > 0$ the function

$$G_n := \mathfrak{P}_+(s_n\, Q) \tag{10.123}$$

in which $s_n(z) = z^n$ and where $\mathfrak{P}_+$ denotes the *Riesz projection* which cuts of the left side Fourier series, i.e. if $Q(\zeta) = \sum_{k=-\infty}^{\infty} \hat{Q}(k)\, \zeta^k$, one has

$$(\mathfrak{P}_+[s_n\, Q])(\zeta) = P_+(\textstyle\sum_{k=-\infty}^{\infty} Q_{k-n}\, \zeta^k) = \sum_{k=0}^{\infty} Q_{k-n}\, \zeta^k\ .$$

By this definition of $G_n \in \mathcal{W}_+$, it is clear that $\lim_{n\to\infty} \|G_n - s_n Q\|_{\mathcal{W}} = 0$, and together with Lemma 10.68 one obtains that

$$\lim_{n\to\infty} \|G_n\|_{\mathcal{W}} = \lim_{n\to\infty} \|s_n Q\|_{\mathcal{W}} = \|Q\|_{\mathcal{W}}\ .$$

This last result together with Lemma 10.70 implies also that

$$\lim_{n\to\infty} \|(G_n)^k\|_{\mathcal{W}} = \|Q^k\|_{\mathcal{W}}\ .$$

for all $1 \leq k \leq N$. Therefore, there exists an $n_0 = n_0(\epsilon, N)$ such that

$$\|(G_n)^k\|_{\mathcal{W}} \geq (1-\epsilon)\, \|Q^k\|_{\mathcal{W}} \tag{10.124}$$

for all $1 \leq k \leq N$ and for all $n > n_0$.

Now, we fix $n > n_0(\epsilon, N)$ and define $G := G_n = s_n Q$. From $\|Q\|_\infty = 1$ follows that $\|G\|_\infty = 1$ and inequality (10.122) shows that

$$\|G\|_{\mathcal{W}} \geq (1-\epsilon)\, e^{B} \geq \tfrac{1}{2}\, e^{B}\ .$$

Let $\mu > 0$ as given by the lemma and define the function $H := G\, \|G\|_{\mathcal{W}}^{-1}$. By this definition, it is clear that $\|H\|_W = 1$, and from the above lower bound of $\|G\|_{\mathcal{W}}$ and due to the choice $B > \log(2/\mu)$, it holds that $\|H\|_\infty \leq \mu$. Moreover, for its L^1-norm holds

$$\|H^k\|_{\mathcal{W}} = \tfrac{1}{\|G\|_{\mathcal{W}}^k}\, \|G^k\|_{\mathcal{W}} \geq (1-\epsilon)\, \|G\|_{\mathcal{W}}^k = (1 - \tfrac{\delta}{2}) \tag{10.125}$$

using (10.124) and (10.122). The so defined function $H \in \mathcal{W}_+$ has already the desired properties, apart from property (10.46). However, the desired g is obtained from this H. To this end, define for an arbitrary $m \geq 0$ the function $H_m(z) := z^m\, H(z)$ which belongs to $\mathcal{W}_+$ with $\|H_m\|_\infty = 1$. Next we consider the expression

$$\textstyle\sum_{k=1}^{N} b_k\, (H_m)^k = \sum_{k=1}^{N} z^{k\cdot m}\, [b_k\, H^k]\ .$$

Applying Lemma 10.66 to this expression, one obtains that

$$\lim_{m\to\infty} \| \sum_{k=1}^{N} b_k\,(H_m)^k \|_{\mathcal{W}} = \sum_{k=1}^{N} b_k \|H^k\|_{\mathcal{W}}\ .$$

This shows that there exists an $m_0 = m_0(\delta, N, b_1, \cdots, b_N)$ such that for all $m > m_0$

$$\begin{aligned} \| \sum_{k=1}^{N} b_k\,(H_m)^k \|_{\mathcal{W}} &\geq \sum_{k=1}^{N} b_k \|H^k\|_{\mathcal{W}} - \tfrac{\delta}{2} \sum_{k=1}^{N} b_k \\ &= (1-\delta) \sum_{k=1}^{N} b_k \end{aligned}$$

using (10.125) to obtain the last line. Thus, choose $m > m_0$ and set $g := H_m$ gives the desired function. $\square$

List of Symbols

The numbers that follow the symbols indicate the page numbers where these symbols where defined.

General Notations

χ_A		indicator function of set A
$\mathrm{i} = \sqrt{-1}$		imaginary unit
$\Re(z)$		real part of z
$\Im(z)$		imaginary part of z
$\overline{z}$		complex conjugate of z
$f^\star$	169	parahermitian conjugate of f: $f^\star(z) = \overline{f(1/\overline{z})}$
$\hat{f}(n)$	15	n-th Fourier coefficients of a function f
$\mathrm{sgn}(k)$	16	signum function
$B_N[f]$	24	best approximation of a function f of degree N
$\oplus$	100	direct sum

Sets

$\mathbb{C}$	complex plane
$\mathbb{C}^N$	N-dimensional complex Euclidean space
$\mathbb{C}^{M\times N}$	set of all complex matrices with M rows and N columns
$\mathbb{N}$	natural numbers
$\mathbb{D} = \{z \in \mathbb{C} : \lvert z\rvert < 1\}$	unit disk in $\mathbb{C}$
$\mathbb{T} = \{z \in \mathbb{C} : \lvert z\rvert = 1\}$	unit circle in $\mathbb{C}$
$\overline{\mathbb{D}} = \mathbb{D} \cup \mathbb{T}$	closed unit disk in $\mathbb{C}$
$\mathbb{D}_r(a) = \{z \in \mathbb{C} : \lvert z - a\rvert < r\}$	open disk in $\mathbb{C}$ with center a and radius r

V. Pohl, H. Boche, *Advanced Topics in System and Signal Theory*, Foundations in Signal Processing, Communications and Networking 4,
DOI 10.1007/978-3-642-03639-2_BM2,

Function Spaces

$A(\mathbb{D})$	52	disk algebra
$\mathcal{B}_+$	99	causal subspace of the Banach space $\mathcal{B}$
$\mathcal{C}(\mathbb{T})$	5	continuous functions on $\mathbb{T}$
$\mathcal{C}_\omega(\Omega)$, $\mathcal{C}_{\omega,0}(\Omega)$	11	smooth functions on Ω
H^p	30	Hardy space of complex functions
$H^p(\mathcal{H})$	46	Hardy space of function with values in the Hilbert space $\mathcal{H}$
ℓ^p, ℓ^p_+	4	p-summable complex sequences
L^p	5	p-integrable functions on $\mathbb{T}$
Λ_α	14	Hölder-Zygmund class
Lip_K	13	Lipschitz continuous functions
$\mathcal{P}(N)$	24	trigonometric polynomials of degree N
$\mathcal{W}$	53	Wiener algebra

Operators

$\mathbf{O_H}$, $\mathbf{O_H^+}$	47	Multiplication operator with symbol $\mathbf{H}$
$\mathcal{R}$	72	Right shift
$\mathfrak{H}$	89	Hilbert transform
$\mathfrak{I}$		Identity operator
$\mathfrak{P}_+$	103	Riesz projection
$\mathfrak{P}$	82	Poisson integral
$\mathfrak{Q}$	82	conjugate Poisson integral
$\mathfrak{R}$	81	Herglotz Riesz transform
$\mathfrak{S}$	170	Spectral factorization mapping
$\mathcal{T}_\varphi$	111	Toeplitz operator with symbol φ

Functional analytic notions

Let $\mathcal{X}, \mathcal{Y}$ arbitrary normed spaces and $\mathfrak{T}$ a map from $\mathcal{X}$ into $\mathcal{Y}$.

$\mathscr{B}(\mathcal{X}, \mathcal{Y})$	8	set of linear bounded operators from $\mathcal{X}$ to $\mathcal{Y}$
$\mathscr{D}(\mathfrak{T})$	7	domain of $\mathfrak{T}$
$\mathscr{R}(\mathfrak{T})$	7	the range of $\mathfrak{T}$
$\mathscr{N}(\mathfrak{T})$	7	null space (the kernel) of $\mathfrak{T}$
$\mathcal{X}^*$	9	the dual space of $\mathcal{X}$ (the set of all bounded linear functionals on $\mathcal{X}$)
$\langle\cdot,\cdot\rangle_{\mathcal{H}}$	4	inner product on the Hilbert space $\mathcal{H}$
$\|\cdot\|_{\mathcal{X}\to\mathcal{Y}}$	8	norm in $\mathscr{B}(\mathcal{X}, \mathcal{Y})$

Banach algebra notions

Let $\mathcal{A}$ be an arbitrary Banach algebra.

$\mathcal{G}(\mathcal{A})$	54	set of all invertible elements of $\mathcal{A}$
$\Gamma(\mathcal{A})$	60	set of all complex homomorphism of $\mathcal{A}$
$\exp(\mathcal{A})$	59	set of all $f \in \mathcal{A}$ which posses a logarithm in $\mathcal{A}$
$\sigma(f)$	55	spectrum of $f \in \mathcal{A}$
$\rho(f)$	55	resolvent set of $f \in \mathcal{A}$
$r_\sigma(f)$	55	spectral radius of $f \in \mathcal{A}$
$\breve{f}$	64	Gelfand transformed of f

References

[1] L. Ahlfors. *Complex Analysis.* McGraw-Hill, New York, 1953.

[2] H. Akcay. On the existence of a disk algebra basis. *Signal Processing*, 80:903–907, 2000.

[3] H. Akcay and B. Ninness. Rational Basis Functions for Robust Identification from Frequency and Time-Domain Measurements. *Automatica*, 34(9):1101–1117, 1998.

[4] B. D. O. Anderson. Continuity of the Spectral Factorization Operation. *Math. Appl. Comput.*, 4(2):139–156, 1985.

[5] S. Banach. *Théorie des Operationes Linéaires.* PWN-Panstwowe Wydawnictwo Naukowe, Warsaw, Poland, 1979.

[6] D. L. Berman. A class of linear operations. *Dokl. Akad. Nauk SSSR*, 85 (1):13–16, 1952.

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

[8] H. Boche. Verhalten der Cauchy-Transformation und der Hilbert-Transformation für auf dem Einheitskreis stetige Funktionen. *Arch. Math.*, 75:349–357, 2000.

[9] H. Boche and V. Pohl. Signal Representation and Approximation – Fundamental Limits. *European Trans. Telecommun. (ETT)*, 18(5):445–456, Aug. 2007. Special Issue on Turbo Coding 2006.

[10] H. Boche and V. Pohl. The Stability and Continuity Behavior of the Spectral Factorization in the Wiener Algebra with Applications in Wiener Filtering. *IEEE Trans. Circuits Syst.*, 55(10):3063–3076, Nov. 2008.

[11] H. Boche and V. Pohl. Structural Properties of the Wiener Filter – Stability, Smoothness Properties, and FIR Approximation Behavior. *IEEE Trans. Inf. Theory*, 52(9):4272–4282, Sept. 2006.

[12] H. Boche and V. Pohl. On the Calculation of the Hilbert Transform from Interpolated Data. *IEEE Trans. Inf. Theory*, 54(5):2358–2366, May 2008.

[13] H. Boche and V. Pohl. Rate of Convergence in Approximating the Spectral Factor of Regular Stochastic Sequences. *IEEE Trans. Inf. Theory*, 2009.

[14] H. Boche and V. Pohl. Spectral Factorization for Polynomials Spectral Densities – Impact of Dimension. *IEEE Trans. Inf. Theory*, 53(11):4236–4241, Nov. 2007.

[15] H. Boche and V. Pohl. Boundedness Behavior for the Spectral Factorization for Polynomial Data in the Wiener Algebra. *IEEE Trans. Signal Process.*, 56(1):274–283, Jan. 2008.

[16] H. Boche and V. Pohl. Robustness of the Inner-Outer Factorization and of the Spectral Factorization for FIR Data. *IEEE Trans. Signal Process.*, 56(1):274–283, Jan. 2008.

[17] H. Boche and V. Pohl. On the Boundedness of the Spectral Factorization Mapping on Decomposing Banach Algebras. *SIAM J. Control Optim.*, 47(3):1415–1427, 2008.

[18] H. Boche and V. Pohl. On the Behavior of Disk Algebra Bases with Applications. *Signal Processing*, 86(12):3915–3922, Dec. 2006.

[19] H. Boche and V. Pohl. Behaviour of the Spectral Factorization for Continuous Spectral Densities. *Signal Processing*, 87(5):1078–1088, May 2007.

[20] H. Boche and V. Pohl. On the Behavior of Causal Projections with Applications. *Signal Processing*, 87(12):3108–3130, Dec. 2007.

[21] H. Boche and V. Pohl. Continuity versus Boundedness of the Spectral Factorization Mapping. *Studia Math.*, 189(2):131–145, 2008.

[22] H. Boche and G. Wunder. Abschätzung der Normen von Operatoren analytischer Funktionen. *Complex Var. Theory Appl.*, 48(3):211–219, Mar. 2003.

[23] S. V. Bockarev. Existence of a Basis in the space of functions analytic in the disk, and some properties of Franklin's system. *Math. of USSR, Sbornik*, 24(1):1–16, 1974.

[24] J. Bourgain. On bases in the disk algebra. *Trans. Amer. Math. Soc.*, 285 (1):133–139, Sept. 1984.

[25] P. L. Butzer and R. J. Nessel. *Fourier Analysis and Approximation.* Birkhäuser Verlag, Basel, Stuttgart, 1971.

[26] M. Ciampa, M. Franciosi, and M. Poletti. A Note on Impulse Response for Continuous, Linear, Time-Invariant, Continuous-Time Systems. *IEEE Trans. Circuits Syst. I*, 53(1):106–113, Jan. 2006.

[27] K. F. Clancey and I. Gohberg. *Factorization of Matrix Functions and Singular Integral Operators.* Birkhäuser Verlag, Basel, 1981.

[28] L. Debnath and P. Mikusinski. *Hilbert Spaces with Applications.* Elsevier, Academic Press, Amsterdam, 3rd edition, 2005.

[29] J. L. Doob. *Stochastic Processes.* John Wiley, New York, 1953.

[30] P. L. Duren. *Theory of H^p Spaces.* Academic Press, New York, 1970.

[31] K. M. Dyakonov. Holomorphic functions and quasiconformal mappings with smooth moduli. *Adv. Math.*, 187:146–172, 2004.

[32] K. M. Dyakonov. Equivalent norms on Lipschitz-type spaces of holomorphic functions. *Acta Math.*, 178:143–167, 1997.

[33] H. Dym and H. P. McKean. *Gaussian Processes, Function Theory, and the Inverse Spectral Problem.* Academic Press, New York, etc., 1976.

[34] E. M. Dyn'kin. The Pseudoanalytic Extension. *J. Anal. Math.*, 60:45–70, 1993.

[35] O. El-Fallah, N. K. Nikolski, and M. Zarrabi. Resolvent Estimates in Beurling-Sobolev Algebras. *St. Petersburg Math. J.*, 10(9):901–964, 1999.

[36] G. Faber. Über die Orthogonalfunktionen des Herrn Haar. *Jahresber. Deutsch. Math.-Verein.*, 19:104–112, 1910.

[37] L. Fejér. Über nichtnegative trigonometrische Polynome. *J. Reine Angew. Math.*, 146:53–82, 1915.

[38] B. A. Francis. *A Course in H_∞ Control Theory.* Springer Verlag, New York, 1987.

[39] P. Franklin. A Set of Continuous Orthogonal Functions. *Math. Ann.*, 100:522–529, 1928.

[40] T. W. Gamelin. *Uniform Algebras.* Prentice-Hall, Englewood Cliffs, NJ, 1969.

[41] J. B. Garnett. *Bounded Analytic Functions.* Academic Press, New York, 1981.

[42] I. M. Gelfand and A. M. Yaglom. Calculation of the amount of information about a random function contained in another such function. *Amer. Math. Soc. Transl. Ser. 2.*, 12:199–246, 1959.

[43] P. R. Halmos. Shifts on Hilbert spaces. *J. für Math.*, 208(1/2):102–112, 1961.

[44] H. Helson. *Lectures on Invariant Subspaces.* Academic Press, New York and London, 1964.

[45] H. Helson. *Harmonic Analysis.* Addison-Wesley, London, 1983.

[46] H. Helson and D. Sarason. Past and Future. *Math. Scand.*, 21:5–16, 1967.

[47] H. Helson and G. Szegö. A Problem in Prediction Theory. *Ann. Mat. Pura Appl.*, 51:107–138, 1960.

[48] K. Hoffman. *Banach Spaces of Analytic Functions.* Prentice-Hall, Englewood Cliffs, NJ, 1962.

[49] I. A. Ibragimov. On the Spectrum of Stationary Gaussian Sequences Satisfying the Strong Mixing Condition; I. Necessary Conditions. *Theory Probab. Appl.*, 10(1):85–106, 1965.

[50] I. A. Ibragimov. Completely Regular Multidimensional Stationary Processes with Discrete Time. *Proc. Steklov Inst. Math.*, 111:269–301, 1970.

[51] D. Jackson. On approximations by trigonometric sums and polynomials. *Trans. Amer. Math. Soc.*, 13:491–515, 1912.

[52] B. Jacob and J. R. Partington. On the Boundedness and Continuity of the Spectral Factorization Mapping. *SIAM J. Control Optim.*, 40(1): 88–106, 2001.

[53] T. Kailath, A. H. Sayed, and B. Hassibi. *Linear Estimation.* Prentice-Hall, Inc., Upper Saddle River, NJ, 2000.

[54] L. V. Kantorovich and G. P. Akilow. *Functional Analysis in Normed Spaces.* Pergamon Press, Oxford, etc., 1964.

[55] A. N. Kolmogorov. Sur les functions harmoniques conjuguées et les séries de Fourier. *Fund. Math.*, 7:24–29, 1925.

[56] A. N. Kolmogorov and Y. A. Rozanov. On the strong mixing for stationary gaussian processes. *Theory Probab. Appl.*, 5(2):204–208, 1960.

[57] P. Lax. Translation Invariant Spaces. *Acta Math.*, 101:163–178, 1959.

[58] J. Lindenstrauss and L. Tzafriri. On the Complemented Subspaces Problem. *Israel J. Math.*, 9:263–269, 1971.

[59] S. M. Lozinskii. Spaces $\widetilde{\mathbf{C}}_\omega$ and $\mathbf{C}^*_\omega$ and the convergence of interpolation processes in them. *Dokl. Akad. Nauk SSSR*, 59(8):1389–1392, 1948.

[60] M. A. Naimark. *Normed Rings.* Wolters-Noordhoff, Groningen, Netherlands, 1970.

[61] I. P. Natanson. *Constructive Function Theory.* Frederik Ungar Publishing, New York, 1964.

[62] N. K. Nikolskii. *Operators, Functions, and Systems: An Easy Reading.* American Mathematical Society, Providence, 2002.

[63] A. V. Oppenheim, A. S. Willsky, and S. H. Nawab. *Signals and Systems.* Prentice-Hall, Upper Saddle River, NJ, 1997.

[64] J. R. Partington. *Interpolation, Identification, and Sampling.* Clarendon Press, Oxford, UK, 1997.

[65] J. R. Partington. *Linear Operators and Linear Systems: An Analytical Approach to Control Theory.* Cambridge University Press, Cambridge, UK, 2004.

[66] V. V. Peller and S. V. Khrushchev. Hankel operators, best approximations, and stationary Gaussian processes. *Russian Math. Surveys*, 37(1): 61–144, 1982.

[67] J. G. Proakis and D. G. Maolakis. *Digital Signal Processing.* McMillan, New York, 1992.

[68] C. E. Rickart. *General Theory of Banach Algebras.* D. Van Nostrand, New York, 1960.

[69] Y. A. Rozanov. *Stationary Random Processes.* Holden-Day, San Francisco, 1967.

[70] W. Rudin. *Real and Complex Analysis.* McGraw-Hill, Boston, etc., 3rd edition, 1987.

[71] W. Rudin. Boundary values of continuous analytic functions. *Proc. Amer. Math. Soc.*, 7:808–811, Oct. 1956.

[72] W. Rudin. *Functional Analysis.* McGraw-Hill, Boston, 2nd edition, 1991.

[73] I. W. Sandberg. A Note on Representation Theorems for Linear Discrete-Space Systems. *Circuits Systems Signal Process*, 17(6):703–708, 1998.

[74] I. W. Sandberg. Bounded inputs and the representation of linear system maps. *Circuits Systems Signal Process*, 24(6):103–115, 2005.

[75] I. W. Sandberg. Causality and the Impulse Response Scandal. *IEEE Trans. Circuits Syst. I*, 50(6):810–811, 2003.

[76] I. W. Sandberg. A Note on Properties of BIBO Stable Linear Discrete-Space Systems. *IEEE Signal Process. Lett.*, 7(10):291–292, 2000.

[77] D. Sarason. An Addendum to "Past and Future". *Math. Scand.*, 30:62–64, 1972.

[78] A. H. Sayed and T. Kailath. A Survey of Spectral Factorization Methods. *Numer. Linear Algebra Appl.*, 8:467–496, 2001.

[79] J. Schauder. Eine Eigenschaft des Haarschen Orthogonalsystems. *Math. Z.*, 28:317–320, 1928.

[80] M. R. Schroeder. Synthesis of Low-Peak-Factors Signals and Binary Sequences With Low Autocorrelations. *IEEE Trans. Inf. Theory*, 16(1): 85–89, Jan. 1970.

[81] A. N. Shiryayev. *Probability.* Springer-Verlag, New York, 2nd edition, 1996.

[82] G. Somorjai. On discrete linear operators in the function Space A. In *Proc. Constructive Function Theory*, pages 489–496, Sofia, Bulgaria, 1980.

[83] D. Sz.-Nagy and C. Foias. *Harmonic Analysis of Operators on Hilbert Spaces.* North Holland, Amsterdam, Netherlands, 1970.

[84] G. Szegö. Koeffizientenabschätzungen bei ebenen und räumlichen harmonischen Entwicklungen. *Math. Ann.*, 96:601–632, 1927.

[85] S. Treil. A Counterexample on Continuous Coprime Factors. *IEEE Trans. Autom. Control*, 39(6):1262–1263, June 1994.

[86] J. G. van der Corput. Zahlentheoretische Abschätzungen. *Math. Ann.*, 84:53–79, 1921.

[87] N. Wiener. Tauberian Theorems. *Ann. of Math.*, 33:1–100, 1932.

[88] N. Wiener. *Extrapolation, Interpolation, and Smoothing of Stationary Time Series with Engineering Applications.* The MIT Press and John Wiley & Sons, Inc., Cambridge, 1949.

[89] N. Wiener and E. Hopf. Über eine Klasse singulärer Integralgleichungen. *Sitzungsberichte der Königlich Preußische Akademie der Wissenschaften zu Berlin*, pages 696–706, 3. Dez. 1931.

[90] P. Wojtaszczyk. *Banach Spaces for Analysts.* Cambridge University Press, Cambridge, 1991.

[91] R. M. Young. *An Introduction to Nonharmonic Fourier Series.* Academic Press, New York, 1980.

[92] A. Zygmund. *Trigonometric Series.* Cambridge University Press, Cambridge, UK, 3rd edition, 2002.