BIRKHÄUSER

APPLIED AND NUMERICAL HARMONIC ANALYSIS

Four Short Courses on Harmonic Analysis

Wavelets, Frames, Time-Frequency Methods, and Applications to Signal and Image Analysis

With Contributions by
Ole Christensen
Karlheinz Gröchenig
Demetrio Labate
Pierre Vandergheynst
Guido Weiss
Yves Wiaux

Brigitte Forster
Peter Massopust
Editors

BIRKHÄUSER

Applied and Numerical Harmonic Analysis

Brigitte Forster • Peter Massopust
Editors

Four Short Courses on Harmonic Analysis

Wavelets, Frames, Time-Frequency Methods, and Applications to Signal and Image Analysis

With Contributions by

Ole Christensen
Karlheinz Gröchenig
Demetrio Labate
Pierre Vandergheynst
Guido Weiss
Yves Wiaux

Birkhäuser
Boston • Basel • Berlin

Editors
Brigitte Forster
Zentrum Mathematik, M6
Technische Universität München
Boltzmannstraße 3
85748 Garching b. München
Germany
forster@ma.tum.de

Peter Massopust
Institute of Biomathematics and Biometry
Helmholtz Zentrum München
Ingolstädter Landstraße 1
85764 Neuherberg
Germany
peter.massopust@helmholtz-muenchen.de

and

Zentrum Mathematik, M6
Technische Universität München
Boltzmannstraße 3
85748 Garching b. München
Germany
massopust@ma.tum.de

ISBN 978-0-8176-4890-9 e-ISBN 978-0-8176-4891-6
DOI 10.1007/978-0-8176-4891-6
Springer New York Dordrecht Heidelberg London

Library of Congress Control Number: 2009939428

Mathematics Subject Classification (2000): 42B35, 42Cxx, 41A15, 43A90, 68U10, 92C55, 94Axx

Printed on acid-free paper

Birkhäuser Boston is part of Springer Science+Business Media (www.birkhauser.com)

ANHA Series Preface

The *Applied and Numerical Harmonic Analysis (ANHA)* book series aims to provide the engineering, mathematical, and scientific communities with significant developments in harmonic analysis, ranging from abstract harmonic analysis to basic applications. The title of the series reflects the importance of applications and numerical implementation, but richness and relevance of applications and implementation depend fundamentally on the structure and depth of theoretical underpinnings. Thus, from our point of view, the interleaving of theory and applications and their creative symbiotic evolution is axiomatic.

Harmonic analysis is a wellspring of ideas and applicability that has flourished, developed, and deepened over time within many disciplines and by means of creative cross-fertilization with diverse areas. The intricate and fundamental relationship between harmonic analysis and fields such as signal processing, partial differential equations (PDEs), and image processing is reflected in our state-of-the-art *ANHA* series.

Our vision of modern harmonic analysis includes mathematical areas such as wavelet theory, Banach algebras, classical Fourier analysis, time-frequency analysis, and fractal geometry, as well as the diverse topics that impinge on them.

For example, wavelet theory can be considered an appropriate tool to deal with some basic problems in digital signal processing, speech and image processing, geophysics, pattern recognition, biomedical engineering, and turbulence. These areas implement the latest technology from sampling methods on surfaces to fast algorithms and computer vision methods. The underlying mathematics of wavelet theory depends not only on classical Fourier analysis, but also on ideas from abstract harmonic analysis, including von Neumann algebras and the affine group. This leads to a study of the Heisenberg group and its relationship to Gabor systems, and of the metaplectic group for a meaningful interaction of signal decomposition methods. The unifying influence of wavelet theory in the aforementioned topics illustrates the justification for providing a means for centralizing and disseminating information from the broader, but still focused, area of harmonic analysis. This will be a key role of *ANHA*. We intend to publish with the scope and interaction that such a host of issues demands.

Along with our commitment to publish mathematically significant works at the frontiers of harmonic analysis, we have a comparably strong commitment to publish major advances in the following applicable topics in which harmonic analysis plays a substantial role:

Antenna theory	*Prediction theory*
Biomedical signal processing	*Radar applications*
Digital signal processing	*Sampling theory*
Fast algorithms	*Spectral estimation*
Gabor theory and applications	*Speech processing*
Image processing	*Time-frequency and*
Numerical partial differential equations	*time-scale analysis*
	Wavelet theory

The above point of view for the *ANHA* book series is inspired by the history of Fourier analysis itself, whose tentacles reach into so many fields.

In the last two centuries Fourier analysis has had a major impact on the development of mathematics, on the understanding of many engineering and scientific phenomena, and on the solution of some of the most important problems in mathematics and the sciences. Historically, Fourier series were developed in the analysis of some of the classical PDEs of mathematical physics; these series were used to solve such equations. In order to understand Fourier series and the kinds of solutions they could represent, some of the most basic notions of analysis were defined, e.g., the concept of "function." Since the coefficients of Fourier series are integrals, it is no surprise that Riemann integrals were conceived to deal with uniqueness properties of trigonometric series. Cantor's set theory was also developed because of such uniqueness questions.

A basic problem in Fourier analysis is to show how complicated phenomena, such as sound waves, can be described in terms of elementary harmonics. There are two aspects of this problem: first, to find, or even define properly, the harmonics or spectrum of a given phenomenon, e.g., the spectroscopy problem in optics; second, to determine which phenomena can be constructed from given classes of harmonics, as done, for example, by the mechanical synthesizers in tidal analysis.

Fourier analysis is also the natural setting for many other problems in engineering, mathematics, and the sciences. For example, Wiener's Tauberian theorem in Fourier analysis not only characterizes the behavior of the prime numbers, but also provides the proper notion of spectrum for phenomena such as white light; this latter process leads to the Fourier analysis associated with correlation functions in filtering and prediction problems, and these problems, in turn, deal naturally with Hardy spaces in the theory of complex variables.

Nowadays, some of the theory of PDEs has given way to the study of Fourier integral operators. Problems in antenna theory are studied in terms of unimodular trigonometric polynomials. Applications of Fourier analysis abound in signal processing, whether with the fast Fourier transform (FFT), or filter design, or the adaptive modeling inherent in time-frequency-scale methods such as wavelet theory. The coherent states of mathematical physics are translated and modulated Fourier transforms, and these are used, in conjunction with the uncertainty principle, for dealing with signal reconstruction in communications theory. We are back to the raison d'être of the *ANHA* series!

John J. Benedetto
Series Editor
University of Maryland
College Park

Preface

Classical harmonic analysis studies problems related to series expansions of signals or functions using trigonometric polynomials. The theory of Fourier series and Fourier integrals forms the core of harmonic analysis and extends from there to other mathematical areas such as the theory of singular integrals, approximation theory, and sampling theory, just to mention a few. Harmonic analysis is also used in numerous applications where it can be thought of as the mathematical backbone for a large number of modern methods in signal analyis and signal processing as well as image analysis and image processing. Its internal growth has seen generalizations to nontrigonometric expansions and noncommutative group settings, but its basic role in other areas of mathematics (differential equations, number theory, probability theory, and statistics), physics and chemistry (wave phenomena, crystallography, and optics), financial analysis (time series), medicine (tomography, brain and heart wave analyses), and biological signal processing has made harmonic analysis the main fundamental contributor to all of 20th century's human-based technologies. These include telephone, radio, television, radar and sonar, satellite and wireless communications, medical imaging, the Internet, and multimedia.

The applications of harmonic analysis to medical image processing have been undergoing a rapid change primarily driven by better hardware and software. Part of this development is an attempt by researchers to base medical engineering principles on solid and rigorous mathematical foundations, and to develop mathematical methods that allow the creation of effective software programs that reduce or replace invasive medical procedures.

Approximation theory and harmonic analysis benefit from each other. The latter provides the means that the former uses to approximate complicated functions or signals and surfaces or images, and to estimate the errors of this approximation. On the other hand, harmonic analysis problems often require methods or input from approximation theory. Like harmonic analysis, approximation theory has seen decades of rapid development and growth, again, primarily driven by applications, such as computer-aided geometric design (CAGD) and its various ramifications.

Recently, a great deal of emphasis has been put into the digitization, transmission, and processing of three-dimensional data sets. One-dimensional methods developed in harmonic analysis and approximation theory in the past do not easily carry over to this higher-dimensional setting. Instead, new ideas and methods need to be found to take into account the nonisotropy and nonhomogeneities inherent in such data sets. In order for these generalizations to take place, new ideas from lower-dimensional problems need to be reconsidered. As an example, we take the effective design of wave forms that is essential to the simultaneous transmission of clear messages on the same frequency band. Constructive approximations of unimodular sequences whose autocorrelations vanish on prescribed sets are introduced, and their analysis depends signifantly on Wiener's generalized harmonic analysis (see [19]).

Signal analysis and image analysis have greatly benefited from the theory of wavelets and their generalizations to frames. These multiscale methods use representations based on two specific groups that are used to transfer information between the scales and within each scale. It has become clear that for multidimensional data, more general groups and multiscale methods need to be employed. The geometry involved in such a high-dimensional setting is more complicated and challenging than in the one-dimensional case, as spatial and, in the video setting, even temporal features need to be taken into account. A first step toward such an improvement in representation is undertaken in [130, 233].

This advanced textbook is intended for graduate students, pure and applied mathematicians, mathematical physicists, and engineers working in image/signal processing and communication theory. The book may be used in an advanced topics course or in a seminar on harmonics analysis and its applications to image and signal analysis. The prerequisites are a solid background in linear algebra and real analysis and knowledge of the fundamentals of functional analysis and metric topology.

Chapters 2, 3, 4, and 5 in this book are based on lectures given by their authors at the summer school on New Trends and Directions in Harmonic Analysis, Approximation Theory, and Image Analysis, which took place in Inzell, Germany, from September 17–21, 2007. One of the goals of this summer school was to bring together a distinguished group of highly established international researchers to present their latest cutting-edge research, and, in conjunction with a small group of scientists including young researchers, to establish new and exciting directions for future investigation into the topics described above.

A short introduction to the mathematical aspects of time-frequency analysis paves the way for the above-mentioned chapters. The reader is exposed to the main themes presented in this book and provided with a summary of those mathematical notions and concepts needed to fully appreciate the contents of Chapters 2 to 5. In addition, the material in these chapters is put into perspective in this introductory chapter.

Chapters 2 to 5 were written by internationally renowned mathematicians and have an expository and interdisciplinary character, allowing the reader to understand the theory behind modern image and signal processing methodologies. In detail, the chapters cover the following.

Ole Christensen considers B-spline generated frames. He exploits the flexibility of frames and combines them with the elegant representations for B-splines. In the first part of his chapter, he introduces the terminology of Bessel sequences, Riesz bases, and frames and exhibits their central properties. In the second part, he considers concrete constructions for Gabor systems and other tight frames, before he finally deduces the wavelet frames generated by B-splines via the so-called unitary extension principle.

Demetrio Labate and Guido Weiss consider the theory and applications of composite wavelets. They first describe the unified theory of reproducing systems, a simple and flexible mathematical framework to characterize and analyze wavelets, Gabor systems, and other reproducing systems in a unified manner. These systems can be rewritten as a countable family of translations applied to a countable collection of functions. The authors then define wavelets with composite dilations, a novel class of reproducing systems that provide truly multidimensional generalizations of traditional wavelets, and discuss so-called shearlets as a special case of optimally sparse representations for 2D. Applications in edge detection and considerations on the continuous analogues of composite wavelets are also considered.

Pierre Vandergheynst and Yves Wiaux introduce wavelets on the sphere and therefore leave the classical Cartesian space. For many applications such as astrophysics, geophysics, neuroscience, computer vision, and computer graphics, data are given as functions on the sphere. In all these situations, one is compelled to design data analysis tools that are adapted to spherical geometry, for one cannot simply project the data into Euclidean geometry without having to deal with severe distortions. The authors provide a generalization of the wavelet transform to signals on the sphere. This generalization is not trivial, as the dilation operator is not well defined on the sphere. In addition, any algorithm faces the problem of how to sample data on the sphere. This chapter discusses some recently developed methods for the analysis and reconstruction of signals on the sphere with wavelets, on the basis of theory, implementation, and applications.

Karlheinz Gröchenig gives various new and interesting aspects of Wiener's Lemma. This result is one of the main theorems of Banach algebra theory. In the first part of his chapter, he discusses Wiener's Lemma in detail and investigates equivalent formulations for convolution operators. In the second part, he considers various variations, especially in noncommutative settings. He also shows the importance of the lemma for time-varying systems and pseudodifferential operators and concludes with applications in mobile communications.

One of the main features of this book is its emphasis on the interdependence of these four modern research directions. Each chapter ends with exercises that allow for a more in-depth understanding of the material and are intended to stimulate the reader to further research.

We would like to thank the VolkswagenStiftung for generously providing the funds and support for the summer school on New Trends and Directions in Harmonic Analysis, Approximation Theory, and Image Analysis in Inzell, Germany.

Our thanks also go to the Institute for Biomathematics and Biometry at the Helmholtz Zentrum München and the Centre of Mathematics, Research Unit M6 – Mathematical Modelling, at the Technische Universität München.

We also would like to acknowledge that this work was partially supported by the grant MEXT-CT-2004-013477, Acronym MAMEBIA, of the European Commission.

We wish to express our gratitude to Birkhäuser and its professional staff, in particular, Tom Grasso, Regina Gorenshteyn, and Patrick Keene, for their support and help during the preparation of this book.

And last but not least, we heartily thank John Benedetto, who had the initial idea for this book and invested the time and energy to launch it.

Munich, Germany *Brigitte Forster*
August 2009 *Peter Massopust*

Contents

List of Contributors

Ole Christensen
Technical University of Denmark
Department of Mathematics
Building 303
2800 Lyngby
Denmark
e-mail: ole.christensen@mat.dtu.dk

Brigitte Forster
Zentrum Mathematik, M6
Technische Universität München
Boltzmannstraße 3
85748 Garching b. München
Germany
e-mail: forster@ma.tum.de

Karlheinz Gröchenig
Faculty of Mathematics
University of Vienna
Nordbergstrasse 15
1090 Vienna
Austria
e-mail: karlheinz.groechenig@univie.ac.at

Demetrio Labate
Department of Mathematics
245 Harrelson Hall
2610 Stinson Drive
Box 8205
North Carolina State University
Raleigh, NC 27695-8205
U.S.A.
e-mail: dlabate@math.ncsu.edu

Peter Massopust
Institute of Biomathematics and Biometry
Helmholtz Zentrum München
Ingolstädter Landstraße 1
85764 Neuherberg
Germany
e-mail: peter.massopust@helmholtz-muenchen.de

and

Zentrum Mathematik, M6
Technische Universität München
Boltzmannstraße 3
85748 Garching b. München
Germany
e-mail: massopust@ma.tum.de

Pierre Vandergheynst
Signal Processing Laboratory (LTS2)
Ecole Polytechnique Fédérale de Lausanne (EPFL)
Station 11
1015 Lausanne
Switzerland
e-mail: pierre.vandergheynst@epfl.ch

Guido L. Weiss
Department of Mathematics
Washington University
Campus Box 1146
One Brookings Drive
St. Louis, MO 63130-4899
U.S.A.
e-mail: guido@math.wustl.edu

Yves Wiaux
Signal Processing Laboratory (LTS2)
Ecole Polytechnique Fédérale de Lausanne (EPFL)
Station 11
1015 Lausanne
Switzerland
e-mail: yves.wiaux@epfl.ch

Chapter 1
Introduction: Mathematical Aspects of Time-Frequency Analysis

Peter Massopust and Brigitte Forster

Abstract Time-frequency analysis of signals or images deals with mathematical transforms of continuous or discrete data with the aim of having more information accessible after than before the transform. The possible choice of the respective transform strongly depends on the mathematical model of the signal or image source. The modeling scheme affects which analyzing transforms can be applied in a mathematically sensible way, e.g., the mapping should be continuous, and also which transforms give access to new and, in particular, well-formulated interpretations of the data.

In this chapter, we present the ideas behind the optimal choice of an appropriate modeling scheme, the standard signal and image models, and the most important mathematical analysis transforms. This covers aspects of Fourier series and integrals, sampling or discretization problems, and various windowed transforms, such as the short-time Fourier transform, the Gabor transform, and wavelets.

The chapter gives an introduction to the main mathematical terms used in the subsequent four chapters of the book. It shows their relation and interplay and gives entrance points to the lectures presented in subsequent chapters. We provide citations to references for further and more in-depth reading and conclude with a list of exercises.

1.1 Aims of Time-Frequency Analysis

Time-frequency analysis deals with the characterization and manipulation of signals whose frequency components vary in time. By a signal, we understand a complex-valued functional $f : X \to \mathbb{C}$, where X is a Banach or Hilbert space and $\mathbb{C}$ denotes

Peter Massopust
Institute of Biomathematics and Biometry, Helmholtz Zentrum München, Ingolstädter Landstraße 1, 85764 Neuherberg, Germany, e-mail: massopust@ma.tum.de

Brigitte Forster
Zentrum Mathematik, M6, Technische Universität München, Boltzmannstraße 3, 85748 Garching, b. München Germany, e-mail: forster@ma.tum.de

B. Forster and P. Massopust (eds.), *Four Short Courses on Harmonic Analysis*,
Applied and Numerical Harmonic Analysis, DOI 10.1007/978-0-8176-4891-6_1,

the field of complex numbers. The choice of the time domain X determines different types of signals. For instance,

- $X := \mathbb{R}$ describes a time-continuous signal;
- $X := \mathbb{Z}$ or $X := \mathbb{N}$ a discrete signal in time;
- $X := [a,b]$, $-\infty < a < b < +\infty$, a signal that is time-limited.

Among these signals, one also distinguishes the following classes:

- T-periodic signals: $f(t) = f(t+kT)$, $T \in \mathbb{R}^+$, and $k \in \mathbb{Z}$;
- finite energy signals: $f \in L^2(\mathbb{R})$ or $f \in l^2(\mathbb{Z})$;
- bounded signals: $f \in L^\infty(\mathbb{R})$ or $f \in l^\infty(\mathbb{Z})$;
- integrable or summable signals: $f \in L^1(\mathbb{R})$, resp., $f \in l^1(\mathbb{Z})$.

In many applications, signals are measured in order to use them to monitor or regulate a time-varying process, or to ensure and manage its quality. The measured signals must be brought into a form that allows for efficient and quick evaluation and interpretation. In order to achieve this, a signal f is transformed so that its image $\widetilde{f}$ under the transform is more easily interpretable and analyzable. In particular, $\widetilde{f}$ should be such that any unwanted noise can be filtered out or removed, and the characteristic system parameters can be estimated. In addition, more information should be extractable from the transformed image $\widetilde{f}$ than from f itself.

1.1.1 Signal and Model

There are many families of functions that can be used to analyze a signal by means of a series expansion or an integral transform. Which family of functions is to be chosen to analyze a signal depends on the underlying model for the system that generates the signal. For instance, for time-limited signals a different function family is used than for time-unlimited signals, and for time-continuous signals a different one than for time-discrete signals. However, the procedure is the same in each case. Figure 1.1 shows a schematic sketch of this procedure.

One of the first analyzing systems was developed by J. B. J. Fourier and is known as the Fourier series. Originally, Fourier was interested in finding an elegant way to solve the heat equation. He expanded the initial value function or distribution in a series of complex exponentials and then could easily derive the solution from the series' coefficients [162, 200, 223]. In fact, the Fourier series of a function $f \in L^p[-\pi,\pi]$, $1 \le p < \infty$, or $f \in C[-\pi,\pi]$ is given by

$$f \sim \sum_{k\in\mathbb{Z}} \widehat{f}(k) e^{ik\bullet}, \tag{1.1}$$

where

$$\widehat{f}(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-ikt}\, dt$$

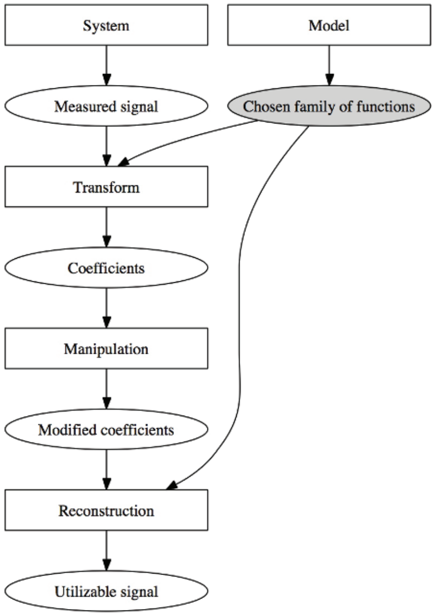


Fig. 1.1: The different methods for signal analysis employ the same scheme. Coefficients are associated with a signal, and these coefficients are transformed. From these new coefficients, a transformed signal is reconstructed.

are the Fourier coefficients. They allow the following interpretation: $|\widehat{f}(k)|$ is the amplitude and $\arg(\widehat{f}(k))$ is the phase corresponding to the frequency $k \in \mathbb{Z}$.

The series (1.1) converges in norm for the L^p-spaces with $1 < p < \infty$, such that the Fourier transform $\widehat{\ }: L^p[-\pi,\pi] \to c_0(\mathbb{Z}), f \to \{\widehat{f}(k)\}_{k\in\mathbb{Z}}$, is a linear, continuous, and on its image continuously invertible operator. [Here, $c_0(\mathbb{Z})$ denotes the space of all sequences with domain $\mathbb{Z}$ converging to zero.] For $L^1[-\pi,\pi]$ and $C[-\pi,\pi]$, convergence of (1.1) is attained with so-called approximate identities [151, 162, 248].

A drawback of the Fourier series for signal analysis is that a local change of the function f results in a global change of the coefficients. Therefore, mathematical interest began to focus more on localized transforms, such as, e.g., the Haar system (see Section 1.1.2.3) or the Rademacher system (see Section 1.1.2.4), and more recently wavelet multiresolution analyses (see Section 1.5.3).

The same drawback as for the Fourier series applies to integral transforms, such as, e.g., the Fourier transform for $L^2(\mathbb{R}^n)$ functions (see Section 1.3). To cope with this problem, local transforms such as the short-time Fourier transform (Section 1.4.1) and, as a special case, the Gabor transform (Section 1.4.2) were developed.

1.1.2 Transforms

For the better interpretability of a signal, a whole variety of possible transforms may be considered. But the question is, which ones are appropriate? One major point is that the transform should not hide or cut off any information. This means that if, in the diagram in Fig. 1.1, we choose the identity operator as the "manipulation," i.e., we leave the coefficients as they are, then the measured signal and the output, the utilizable signal, should coincide. This fact gives rise to several mathematical requirements:

- The transform should be continuous: Quantitatively small changes in the signal should cause only quantitatively small effects in the transform's image.
- The transform should be continuously invertible.
- There should exist an invertible discrete version of the transform.
- There should exist a stable numerical algorithm.

As examples, we review three common transforms based on Fourier series, Haar wavelets, and Rademacher functions.

1.1.2.1 Fourier Series

Let $\mathbb{T} := \{z \in \mathbb{C} \,|\, |z| = 1\} = \{e^{it} \,|\, t \in [0, 2\pi)\}$ be the torus. Then $\mathbb{T} = \mathbb{S}^1$ is a compact subset of $\mathbb{R}^2$ and a commutative group with respect to multiplication. Since the multiplication $\cdot : \mathbb{T} \times \mathbb{T} \to \mathbb{T}$ is continuous with respect to the metric topology on $\mathbb{R}^n$, $\mathbb{T}$ is a topological group.

Every function $f : \mathbb{T} \to \mathbb{C}$ can be uniquely identified with a 2π-periodic function $\widetilde{f} : \mathbb{R} \to \mathbb{C}$ via $\widetilde{f}(t + 2\pi n) = \widetilde{f}(t) = f(e^{it})$, where $t \in [0, 2\pi)$, $n \in \mathbb{Z}$. The Lebesgue measure on $[0, 2\pi)$ is mapped onto $\mathbb{T}$ via

$$\int_{\mathbb{T}} f(z) dz = \int_0^{2\pi} f(e^{it})\, dt = \int_0^{2\pi} \widetilde{f}(t)\, dt,$$

provided the Lebesgue integral exists. We denote by $L^2(\mathbb{T})$ the space of all complex-valued functions on $\mathbb{T}$ that are square-integrable with respect to the Lebesgue measure. Endowed with the inner product

$$\langle f, g \rangle := \frac{1}{2\pi} \int_{\mathbb{T}} f(z)\, \overline{g(z)}\, dz, \quad \text{for } f, g \in L^2(\mathbb{T}),$$

$L^2(\mathbb{T})$ becomes a Hilbert space.

Remark 1.1. Let $1 \le p < \infty$. Then $L^p(\mathbb{T}) = \mathscr{L}^p(\mathbb{T})/\mathscr{N}$, where

$$\mathscr{L}^p(\mathbb{T}) = \left\{ f : \mathbb{T} \to \mathbb{C} \,|\, f \text{ Lebesgue-measurable and } \int_{\mathbb{T}} |f(t)|^p dt < \infty \right\}$$

and $\mathscr{N} = \{f : \mathbb{T} \to \mathbb{C} \,|\, f = 0 \text{ a.e.}\}$.

Convention: In the following, we identify $\widetilde{f}$ and f.

The complex trigonometric system $\{e^{in\bullet}\}_{n\in\mathbb{Z}}$ is an orthonormal basis for the Hilbert space $L^2(\mathbb{T})$. (This can be easily verified by a direct computation.) Orthogonality then implies that the family is minimal and complete. It can be shown (see Section 1.2.1) that this implies that each $f \in L^2(\mathbb{T})$ has a unique Fourier series representation of the form

$$f = \sum_{n\in\mathbb{Z}} \langle f, e^{in\bullet}\rangle e^{in\bullet} \quad \text{in } L^2(\mathbb{T}),$$

where

$$\langle f, e^{in\bullet}\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\, e^{-int}\, dt, \quad n \in \mathbb{Z}.$$

The Parseval equality, which is given in (1.10) in a general form, implies that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(t)|^2\, dt = \sum_{n\in\mathbb{Z}} |\langle f, e^{in\bullet}\rangle|^2, \quad \forall f \in L^2(\mathbb{T}),$$

and that the mapping

$$T : L^2(\mathbb{T}) \to l^2(\mathbb{Z}),\ f \mapsto \{\langle f, e^{in\bullet}\rangle\}_{n\in\mathbb{Z}}$$

is a Hilbert space isomorphism.

Therefore, we have the following result.

Theorem 1.2. *The trigonometric system $\{e^{in\bullet}\}_{n\in\mathbb{Z}}$ is complete in $L^2(\mathbb{T})$.*

Proof. We employ the Weierstraß theorem. Suppose that $f \in L^2(\mathbb{T})$ is an integrable function with the property that $f \notin \overline{\operatorname{span}\{e^{in\bullet}\}_{n\in\mathbb{Z}}}$. Then

$$\langle f, e^{in\cdot}\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\, e^{-int}\, dt = 0, \quad \forall n \in \mathbb{Z}.$$

We show that $f = 0$ almost everywhere. To this end, let

$$g(t) = \int_{-\pi}^{t} f(u)\, du, \quad \text{for } t \in [-\pi, \pi].$$

Note that g is continuous. Suppose that $c \in \mathbb{C}$ is a constant. Integration by parts yields

$$\int_{-\pi}^{\pi} (g(t) - c)\, e^{-int}\, dt = 0, \quad \text{for all } n \in \mathbb{Z} \setminus \{0\}. \tag{1.2}$$

Now choose c such that (1.2) also holds for $n = 0$ and set $F(t) := g(t) - c$. Then F is continuous on $[-\pi, \pi]$ and $F(\pi) = F(-\pi) = -c$. The Weierstraß theorem now implies that for all $\varepsilon > 0$ there exists a trigonometric sum

$$T(t) = \sum_{k=-N}^{N} c_k e^{ikt},$$

so that

$$|F(t) - T(t)| < \varepsilon, \quad \text{for } |t| \le \pi.$$

Thus,

$$\|F\|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |F(t)|^2\,dt = \frac{1}{2\pi}\int_{-\pi}^{\pi} \overline{F(t)}\,(F(t)-T(t))\,dt \le \frac{\varepsilon}{2\pi}\int_{-\pi}^{\pi} |F(t)|\,dt \le \varepsilon\,\|F\|.$$

Hence, $\|F\| \le \varepsilon$ and since F was arbitrary, we have that $F = 0$. Thus, $g = c$ and therefore $f = 0$ almost everywhere. □

Remark 1.3. Analogously, one establishes the completeness of $\{e^{in\bullet}\}_{n\in\mathbb{Z}}$ in $L^p(\mathbb{T})$ for all $1 \le p < \infty$.

The results obtained in this section can be summarized in the following theorem.

Theorem 1.4. *The trigonometric system $\{e^{in\bullet}\}_{n\in\mathbb{Z}}$ is an orthonormal basis for the Hilbert space $L^2(\mathbb{T})$.*

1.1.2.2 Convolution

Denote by $L^1(\mathbb{T})$ the Banach space of all complex-valued Lebesgue-measurable functions on the torus $\mathbb{T}$ endowed with the L^1-norm.

Theorem 1.5. *Suppose that $f, g \in L^1(\mathbb{T})$. Then the function*

$$t \mapsto f(s-t)g(t)$$

is absolutely integrable for almost all $s \in [-\pi,\pi]$. Setting

$$h(s) := \frac{1}{2\pi}\int_{-\pi}^{\pi} f(s-t)g(t)\,dt,$$

one has that $h \in L^1(\mathbb{T})$ and $\|h\|_1 \le \|f\|_1\|g\|_1$. The Fourier coefficients satisfy

$$\widehat{h}(n) = \widehat{f}(n)\cdot\widehat{g}(n) \tag{1.3}$$

for all $n \in \mathbb{Z}$.

Proof. Exercise! □

Definition 1.6. Assume that $f, g \in L^1(\mathbb{T})$. The a.e. defined function $h : \mathbb{T} \to \mathbb{C}$ in Theorem 1.5 is denoted by $f * g$ and is called the *convolution* of f and g.

*Example 1.7 (***Filtering)***.* Consider a 2π-periodic signal $f \in L^1(\mathbb{T})$, i.e., a function $f : \mathbb{T} \to \mathbb{C}$, that contains the frequencies $n \in \mathbb{Z}$, i.e., $\widehat{f}(n) \ne 0$.

A filter is a function $g \in L^1(\mathbb{T})$ that removes certain frequencies $I \subset \mathbb{Z}$ and preserves all others: $\widehat{g}(n) = 0$, for $n \in I$, and $\widehat{g}(n) = 1$, for $n \in \mathbb{Z}\setminus I$.

The filtered signal $h = f * g$ contains on $\mathbb{Z}\setminus I$ exactly the frequencies of f and on I no frequencies.

1.1.2.3 Haar System

In 1910, Haar [131] introduced a Schauder basis for $L^p[0,1)$, $1 \le p < \infty$, which is an unconditional basis if $p > 1$. (Details on Schauder bases and unconditional bases are explained in Section 1.2.1.) Let

$$\psi_H(t) := \begin{cases} 1, & 0 \le t < \frac{1}{2}; \\ -1, & \frac{1}{2} \le t < 1; \\ 0, & \text{otherwise.} \end{cases}$$

The Haar system is defined as the family of functions

$$\{2^{n/2}\,\psi_H(2^n t - k) \,|\, k = 0, 1, \ldots, 2^n;\; n \in \mathbb{N}_0\}.$$

It is easy to verify that the Haar system satisfies the L^2-orthogonality conditions

$$\int_{\mathbb{R}} 2^{m/2}\psi_H(2^m t - k)\, 2^{n/2}\psi_H(2^n t - l) = \delta_{mn}\delta_{kl}, \quad k, l \in \mathbb{Z};\; m, n \in \mathbb{N}_0.$$

If $\phi := \chi_{[0,1)}$ denotes the characteristic function on $[0,1)$, then one has the relation

$$\psi_H = \phi(2\bullet) - \phi(2\bullet - 1).$$

The disadvantage of the Haar system is that it is not continuous.

1.1.2.4 Rademacher System

The Rademacher system [199] is given by the family of functions $R := \{r_n \,|\, n \in \mathbb{N}_0\}$, where

$$r_n(t) := \operatorname{sgn} \sin 2^n \pi t, \quad t \in [0,1).$$

Here $\operatorname{sgn} : \mathbb{R} \to \mathbb{R}$ denotes the signum function

$$\operatorname{sgn}(t) := \begin{cases} \frac{t}{|t|}, & t \ne 0; \\ 0, & t = 0. \end{cases}$$

It can be shown that the family R constitutes an orthonormal system for $L^2[0,1)$ with respect to the L^2-inner product, but not an orthonormal basis. An obvious disadvantage of the Rademacher functions is their discontinuity.

Both Haar and Rademacher functions have support contained in $[0,1]$ and are orthogonal with respect to the L^2-inner product, and both families are generated by a *single* function. The difference between the two systems lies in the way the family of functions is generated: For the Haar systems, one needs the dyadic dilates and integer-translates of ψ_H, whereas for the Rademacher system, only the dyadic dilates of r_0 are needed.

1.1.3 Signal Manipulations—Filters

In order to use and process information obained from a measured or received signal, the sequence of real or complex values describing the signal must be manipulated to yield a utilizable representation. This manipulation of a signal is called *filtering*. Mathematically, it corresponds to transforming a given signal f into a new signal $\widetilde{f}$ via a transform (operator) $\mathscr{T}$. The transform $\mathscr{T}$ can be linear, as in the case of a bandpass filter (see Example 1.7), or nonlinear, as in the case of denoising, where coefficients whose values are in magnitude less than a given threshold are set equal to zero. As a simple example of a transform or *filter* $\mathscr{T}$, we mention the Fourier series introduced in Section 1.1.2.1. Here the signal is represented by its frequency content.

Let σ_τ denote the time shift (by $\tau > 0$), i.e., $\sigma_\tau(f) := f(\bullet + \tau)$. A transform $\mathscr{T}$ is called *time-invariant* if

$$\mathscr{T}(\sigma_\tau f) = \sigma_\tau(\mathscr{T} f).$$

Any linear time-invariant transform $\mathscr{T} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ acting on a signal $f \in L^2(\mathbb{R})$ can be represented as a convolution:

$$\mathscr{T}(f) = F * f, \tag{1.4}$$

where the function $F \in L^2(\mathbb{R})$ is usually called the *impulse response* of the signal f (see [175, Chapter II]). The fact that F is in $L^2(\mathbb{R})$ is a consequence of the Riesz representation theorem. (Show this!)

In case the signal f is representable by a sequence, i.e., $f \in l^2(\mathbb{Z})$, and $\mathscr{T} : l^2(\mathbb{Z}) \to l^2(\mathbb{Z})$ is again a linear time-invariant transform, the discrete analogue of (1.4) then reads

$$\mathscr{T}(f)(n) = (F * f)(n) := \sum_{k\in\mathbb{Z}} F(n-k) f(k), \quad n \in \mathbb{Z}.$$

1.1.4 Why Discretizing? Techniques, Challenges, Pitfalls

One approach for obtaining a numerical solution of a mathematical problem is to discretize it. The discrete problem can then, in many cases, be solved computationally efficiently.

The general procedure of discretization is as follows. Let B_1, B_2, B_1^n, and B_2^n, $n \in \mathbb{N}$, be Banach spaces. Let $\mathscr{T}$ be a linear, not necessarily continuous operator on B_1. Suppose that for all $n \in \mathbb{N}$, the operator $\mathscr{T}^n : B_1^n \to B_2^n$ is linear and that $(\mathscr{T}^n)^{-1}$ exists and is linear. Moreover, assume that the discretization operators $\mathscr{D}_i^n : B_i^n \to B_i^n$, $i = 1, 2$, are linear.

$$\begin{array}{ccc} B_1 & \xrightarrow{\mathscr{T}} & B_2 \\ {\scriptstyle \mathscr{D}_1^n}\downarrow & & \downarrow{\scriptstyle \mathscr{D}_2^n} \\ B_1^n & \xrightarrow{\mathscr{T}^n} & B_2^n \end{array} \tag{1.5}$$

The spaces B_i^n, $i = 1,2$, are usually finite-dimensional, thus reducing the problem to a linear algebraic setting.

From the above scheme one requires that the operators $\mathscr{T}$ and $\mathscr{T}^n$ are *consistent*, i.e., $\|\mathscr{D}_2^n \mathscr{T}(f) - \mathscr{T}^n \mathscr{D}_1^n(f)\| \to 0$, as $n \to \infty$. Note that consistency implies that diagram (1.5) commutes in the limit. If, in addition, the sequence of discretized operators $\{(\mathscr{T}^n)^{-1}\}_{n\in\mathbb{N}}$ is uniformly bounded, then the discretization procedure is called *stable*.

The next theorem gives conditions under which a discretization procedure yields utilizable results.

Theorem 1.8. *Let $f_i \in B_i$ and $f_i^n \in B_i^n$ be such that $\mathscr{T} f_1 = f_2$ and $\mathscr{T}^n f_1^n = f_2^n$, for all $n \in \mathbb{N}$. Suppose that $\|\mathscr{D}_2^n f_2 - f_2^n\| \to 0$ as $n \to \infty$. Furthermore, suppose that $\|\mathscr{D}_2^n \mathscr{T}(f) - \mathscr{T}^n \mathscr{D}_1^n(f)\| \to 0$ as $n \to \infty$ and that $\{(\mathscr{T}^n)^{-1}\}_{n\in\mathbb{N}}$ is uniformly bounded. Then*

$$\|\mathscr{D}_1^n f_1 - f_1^n\| \to 0 \quad \text{as } n \to \infty.$$

Proof. Exercise! □

Remark 1.9. Theorem 1.8 can also be reformulated as: consistency plus stability implies convergence.

In order to represent a continuous signal in a unique manner, the discretized continuous transform $\mathscr{T}^n$ must yield a basis for the spaces B^n. This is exemplified by the Fourier series and, in particular, by Theorem 1.4. Below, several other bases are described that come from discretized operators: the translates of the sinc-function in the sampling theorem (Section 1.3.4), the Gabor system (Section 1.4), and the wavelet system (Section 1.5).

Time- or space-based measurements of signals yield only finitely many discrete values although the mathematical models are usually continuous. Applying the correct discretization procedure is therefore very important, as is the analysis of the different types of measurement error. There are four basic types of measurement error:

1. **truncation error:** arises if only a finite number of samples is taken into account;
2. **amplitude error:** arises since in general the exact ordinate value of the signal is not known but is contaminated by noise, or falsified due to round-off or quantization;
3. **time-/space-jitter error:** arises if the sample points are not met correctly;
4. **aliasing error:** arises if the signal is not exactly band-limited or the bandwidth is larger than assumed.

For the precise analysis of these four types of measurement error, we refer the reader to [35–37] and the references given therein.

1.2 Basic Methods of Time-Frequency Analysis: Orthonormal Bases and Generalized Fourier Series

In this section, we introduce Fourier series using general bases. In the following sections, concrete Banach and Hilbert spaces are considered. For this purpose, we first need to consider the concept of a function basis in Banach spaces.

1.2.1 Schauder Bases in Banach Spaces

In the following, X always denotes a separable Banach space and H a separable Hilbert space. The topological dual to X is denoted by X' and consists of all linear functionals $\phi : X \to \mathbb{C}$. Endowed with the operator norm

$$\|\phi\|_{op} = \sup_{\substack{0\neq x\in X\\ \|x\|\leq 1}} \frac{|\phi(x)|}{\|x\|}, \quad y \in X',$$

X' becomes a Banach space.

One of the most important concepts of a basis in analysis is that of a Schauder basis.

Definition 1.10 (Schauder 1927 [135, 246]). A sequence of elements $\{x_n\}_{n\in\mathbb{N}}$ in an infinite-dimensional Banach space X is called a *Schauder basis* for X if, for every $x \in X$, there exists a unique sequence of scalars $\{c_n\}_{n\in\mathbb{N}}$ so that

$$\left\| x - \sum_{i=1}^{n} c_i x_i \right\| \to 0 \quad \text{as } n \to \infty. \tag{1.6}$$

A Schauder basis is called *bounded* if $0 < \inf_{n\in\mathbb{N}} \|x_n\| \leq \sup_{n\in\mathbb{N}} \|x_n\| < \infty$. In case $\{x_n\}_{n\in\mathbb{N}}$ is a Schauder basis, the linear functionals

$$f_k : X \to \mathbb{C}, \quad x = \sum_{n\in\mathbb{N}} c_n x_n \mapsto c_k, \quad k \in \mathbb{N}, \tag{1.7}$$

are called *coefficient functionals*.

Example 1.11. 1. The sequence spaces $\ell^p(\mathbb{N})$, $1 \leq p < \infty$, have as a Schauder basis the canonical basis $\{e_n\}_{n\in\mathbb{N}}$, where the sequence

$$e_n = \{0, \ldots, 0, 1, 0, \ldots\}$$

has a "1" in the nth position.

2. Banach spaces that possess a Schauder basis are separable. Since the sequence space $\ell^\infty(\mathbb{N})$ is not separable, it cannot have a Schauder basis.
3. Orthonormal bases $\{e_n\}_{n\in\mathbb{N}}$ in a Hilbert space H are Schauder bases having the additional property that $\langle e_n, e_m\rangle = \delta_{nm}$ and $\|e_n\| = 1$.

For Schauder bases and their coefficient functionals, the following theorem holds.

Theorem 1.12. *Let X be a Banach space and $\{x_n\}_{n\in\mathbb{N}}$ a Schauder basis for X. Then the following hold:*

1. *$\{x_n\}_{n\in\mathbb{N}}$ is complete and minimal; i.e., the Schauder basis spans the space and the coefficients in a presentation $x = \sum_{n\in\mathbb{N}} a_n x_n$ are unique for all $x \in X$ [167].*
2. *The coefficient functionals $\{f_n\}_{n\in\mathbb{N}}$ are continuous, hence elements of the dual space X'. Moreover, one has the estimate*

$$1 \le \|x_n\| \cdot \|f_n\| \le K \tag{1.8}$$

for a positive constant K and all $n \in \mathbb{N}$.

Proof. 1. The first statement follows immediately from the definition.

2. To prove the second statement [246], let Y be the vector space

$$Y := \left\{ \{c_n\}_{n\in\mathbb{N}} \subset \mathbb{C} \;\middle|\; \sum_{n=1}^{\infty} c_n x_n \text{ converges in } X \right\} \tag{1.9}$$

endowed with the norm

$$\|\{c_n\}_{n\in\mathbb{N}}\| := \sup_{n\in\mathbb{N}} \left\| \sum_{i=1}^{n} c_i x_i \right\|.$$

Then Y is a Banach space that is isomorphic to X: The mapping

$$T : Y \to X, \quad \{c_n\}_{n\in\mathbb{N}} \mapsto \sum_{n=1}^{\infty} c_n x_n$$

is linear and since $\{x_n\}_{n\in\mathbb{N}}$ is a Schauder basis for X, it follows from (1.9) that T is bijective. Moreover,

$$\|T\{c_n\}_{n\in\mathbb{N}}\| = \left\| \sum_{n=1}^{\infty} c_n x_n \right\| \le \sup_{n\in\mathbb{N}} \left\| \sum_{i=1}^{n} c_i x_i \right\| = \|\{c_n\}_{n\in\mathbb{N}}\|,$$

implying the continuity of T. Thus, by the open mapping theorem, T is a Banach space isomorphism.

Let $x = \sum_{n=1}^{\infty} c_n x_n \in X$. Then we have for every $n \in \mathbb{N}$,

$$\begin{aligned} |f_n(x)| = |c_n| = \frac{\|c_n x_n\|}{\|x_n\|} &\le \frac{\|\sum_{i=1}^{n} c_i x_i\| + \|\sum_{i=1}^{n-1} c_i x_i\|}{\|x_n\|} \\ &\le \frac{2\sup_k \|\sum_{i=1}^{k} c_i x_i\|}{\|x_n\|} = \frac{2\|T^{-1}x\|}{\|x_n\|} \le \frac{2\|T\|^{-1}\|x\|}{\|x_n\|}. \end{aligned}$$

Hence, $\|f_n\| \leq 2\|T\|^{-1}/\|x_n\|$. This implies that f_n is continuous and we have proven the right-hand side of the inequality. The left-hand side holds because of

$$1 = f_n(x_n) = |f_n(x_n)| \leq \|f_n\| \cdot \|x_n\|. \quad \square$$

Remark 1.13. Not every bounded, complete, and minimal sequence in a Banach space is a Schauder basis. Below, we will encounter an example of this fact.

In order to compare Schauder bases in a Banach space, the concept of the equivalence of Schauder bases is needed.

Definition 1.14. Two Schauder bases $\{x_n\}_{n\in\mathbb{N}}$ and $\{y_n\}_{n\in\mathbb{N}}$ in a Banach space are called *equivalent* if there exists a bounded, invertible operator $T : X \to X$ such that $Tx_n = y_n$ for all $n \in \mathbb{N}$.

Every basis in a finite-dimensional vector space can be mapped onto the canonical basis via an invertible operator. In infinite-dimensional vector spaces, this is no longer true. In such spaces, the convergence with respect to a basis depends in general on the order of summation. This and the question on how to find the coefficients with respect to a Schauder basis will be considered next.

1.2.1.1 Biorthogonality

For the computation of the coefficients c_n in an expansion of the form $f \sim \sum_{n\in\mathbb{Z}} c_n x_n$, certain linear functionals y_n with $c_n = y_n(f)$ are used. These functionals are to satisfy the following condition.

Definition 1.15. Two sequences $\{x_n\}_{n\in\mathbb{Z}} \subset X$ and $\{y_n\}_{n\in\mathbb{Z}} \subset X'$ are called *biorthogonal* provided that

$$y_m(x_n) = \delta_{mn}, \qquad \text{for all } m,n \in \mathbb{Z}.$$

Using the Hahn–Banach theorems, it can be shown that in a Banach space a sequence $\{x_n\}_{n\in\mathbb{Z}}$ has a biorthogonal sequence if it is minimal and that this biorthogonal sequence is unique if $\{x_n\}_{n\in\mathbb{Z}}$ is complete.

The coefficient functionals satisfy additional properties.

Theorem 1.16 ([167, 246]).

1. *If $\{x_n\}_{n\in\mathbb{N}}$ is a Schauder basis in a Banach space X, then the associated coefficient functionals $\{f_n\}_{n\in\mathbb{N}}$ are biorthogonal.*
2. *Suppose that $\{x_n\}_{n\in N}$ is a Schauder basis in a reflexive Banach space X. Then the coefficient functionals $\{f_n\}_{n\in\mathbb{N}}$ form a Schauder basis of the dual space X'.*

Proof. Exercise! $\square$

The connection between Schauder bases and their coefficient functionals gives rise to the next definition.

Definition 1.17. Let X be a reflexive Banach space and $\{x_n\}_{n\in\mathbb{N}} \subset X$ a Schauder basis of X. The Schauder basis of X' consisting of the coefficient functionals $\{f_n\}_{n\in\mathbb{N}}$ is called the *dual Schauder basis* to $\{x_n\}_{n\in\mathbb{N}} \subset X$.

Example 1.18. Complete minimal sequences are not necessarily Schauder bases.

To verify this statement, consider a Hilbert space H and an orthonormal sequence $\{e_n\}_{n\in\mathbb{N}}$. This orthonormal sequence is minimal and complete. The family $\{x_n\}_{n\in\mathbb{N}} \subset H$ given by

$$x_n = \sum_{k=1}^{n} \frac{1}{k} e_k, \quad n \in \mathbb{N},$$

is also bounded, complete, and minimal, but not a Schauder basis. For otherwise, the dual sequence $\{y_n\}_{n\in\mathbb{N}}$ with

$$y_n = n e_n - (n+1) e_{n+1}, \quad n \in \mathbb{N},$$

would also be a Schauder basis, hence complete. This, however, is not possible since $f := \sum_{n\in\mathbb{N}} \frac{1}{n} e_n$ does not lie in the span of $\{y_n\}_{n\in\mathbb{N}}$.

1.2.1.2 Unconditional Convergence

In our consideration of Schauder bases, we have so far assumed conditional convergence: The Schauder basis consists of an indexed sequence and the convergence of a series with respect to this basis depends on the order of the basis elements. In the case of unconditional bases, the order of the basis elements is irrelevant.

Definition 1.19. A series $\sum_{n\in\mathbb{Z}} a_n$ in a Banach space X is called *unconditionally convergent* if every permutation $\sigma : \mathbb{N} \to \mathbb{N}$ of the series $\sum_{n\in\mathbb{Z}} a_{\sigma(n)}$ converges to the same element in X.

A Schauder basis $\{x_n\}_{n\in\mathbb{Z}}$ for X is called an *unconditional basis* for X if every convergent series of the form $\sum_{n\in\mathbb{Z}} c_n x_n$ converges unconditionally.

Remark 1.20. In signal analysis it is often customary to sum up the N largest manipulated coefficients and to consider the limit $N \to \infty$ to obtain an approximation of the signal. If conditional bases are employed, then it is no longer guaranteed that the result of the summation, i.e., the synthesized signal, is interpretable. The associated series is summed up using an unpredictable order and may not converge.

Example 1.21. Recall that $c_0(\mathbb{N}) := \{x \in \mathbb{R}^{\mathbb{N}} \,|\, \lim_{n\to\infty} x_n = 0\}$. The sequence $\{d_n\}_{n\in\mathbb{N}}$, where

$$d_1 = (1,0,0,\ldots), \quad d_2 = (1,1,0,\ldots), \quad d_3 = (1,1,1,0,\ldots), \quad \ldots,$$

is a conditional basis for the sequence space $c_0(\mathbb{N})$.

For suppose that $\{a_n\}_{\in\mathbb{N}}$ is a sequence in $\mathbb{R}$ so that $\sum_{n\in\mathbb{N}} a_n$ converges, but does not converge absolutely. Then the series $\sum_{n\in\mathbb{N}} a_n d_n$ is not unconditionally convergent. (Show this!)

In Hilbert spaces H, there are generalizations of unconditional bases in several aspects. The most important ones are the notions of a Riesz basis and of a frame. A frame in H is a redundant family of functions $\{x_n\}_{n\in\mathbb{N}}$ that spans H and fulfills the following stability condition: There exist constants $A, B > 0$ such that

$$A\|x\|^2 \leq \sum_{n\in\mathbb{N}} |\langle x, x_n\rangle|^2 \leq B\|x\|^2, \quad \forall x \in H.$$

It can be shown that frame representations of functions are unconditionally convergent. A Riesz basis is a frame, which is a Schauder basis, i.e., a minimal frame, and allows for unconditionally convergent series representations. Moreover, it is equivalent to an orthonormal basis. A detailed description on frames and Riesz bases is given in Ole Christensen's Chapter 2. Absolutely and therefore unconditionally convergent Fourier series are explored in Karlheinz Gröchenig's Chapter 5.

1.2.2 Generalized Fourier Series

The most important property of orthonormal bases as compared to other bases is the simplicity of the terms in a basis representation. If $\{e_n\}_{n\in\mathbb{N}}$ is an orthonormal basis in a Hilbert space H, then every $f \in H$ has a Fourier series representation of the form

$$f = \sum_{n=1}^{\infty} \langle f, e_n\rangle e_n,$$

and this series converges in the induced norm on H. The inner product $\langle f, e_n\rangle$ is called the nth Fourier coefficient of f. The theorem of Pythagoras implies Parseval's equality:

$$\|f\|^2 = \sum_{n=1}^{\infty} |\langle f, e_n\rangle|^2. \tag{1.10}$$

In particular, Parseval's equality means that the linear mapping

$$S:\ H \to l^2(\mathbb{N}),\ f \mapsto \{\langle f, e_n\rangle\}_{n\in\mathbb{N}}$$

is an isometric Hilbert space isomorphism. Therefore, S also preserves scalar products and the weak Parseval equality holds:

$$\langle f, g\rangle = \sum_{n=1}^{\infty} \langle f, e_n\rangle \overline{\langle g, e_n\rangle}, \quad \forall f, g \in H.$$

Example 1.22. In $l^2(\mathbb{N})$ is the canonical basis in Example 1.11, item 1, an orthonormal basis.

Theorem 1.23. *For every finite orthonormal system, we have that*

$$\sum_{n=1}^{N} |\langle f, e_n\rangle|^2 \leq \|f\|^2.$$

This result has the immediate corollary

Lemma 1.24 (Lemma of Riemann–Lebesgue in Hilbert spaces).

$$\lim_{|n|\to\infty} \langle f, e_n \rangle = 0, \quad \forall f \in H.$$

1.3 The Fourier Integral Transform

In this section, we introduce the Fourier integral transform on $L^1(\mathbb{R})$ and discuss some of its properties. In addition, we define the convolution between two integrable functions and the concept of a summation kernel. Extending the Fourier transform to a Hilbert space setting gives rise to the Plancherel transform, which is presented next. The theorem of Paley–Wiener and the Poisson summation formula conclude this section.

1.3.1 Definition and Properties

We consider time-continuous, Lebesgue-integrable functions $f \in L^1(\mathbb{R})$.

Definition 1.25. Assume that $f \in L^1(\mathbb{R})$. The *Fourier transform* $\mathscr{F}(f)$ of f is defined by

$$\mathscr{F}(f)(\omega) = \int_{\mathbb{R}} f(x)e^{-i\omega x}\,dx, \quad \text{for all } \omega \in \mathbb{R}.$$

Theorem 1.26. *Assume that $f, g \in L^1(\mathbb{R})$, $\omega \in \mathbb{R}$, and $\lambda \in \mathbb{C}$.*

1. *$\mathscr{F}$ is linear: $\mathscr{F}(\lambda f + g)(\omega) = \lambda \mathscr{F}(f)(\omega) + \mathscr{F}(g)(\omega)$.*
2. *Let $\overline{f}(t) := \overline{f(t)}$. Then $\mathscr{F}(\overline{f})(\omega) = \overline{\mathscr{F}(-\omega)}$.*
3. *Let $L_y f(t) := f(t-s)$, for $s \in \mathbb{R}$. Then $\mathscr{F}(L_s f)(a) = e^{-is\omega}\mathscr{F}(f)(\omega)$.*
4. *$|\mathscr{F}(f)(\omega)| \leq \|f\|_1$.*
5. *Let $f_\lambda(t) := \lambda f(\lambda t)$, for $\lambda \in \mathbb{R}\setminus\{0\}$. Then $\mathscr{F}(f_\lambda)(\omega) = \mathscr{F}(f)(\omega/\lambda)$.*

Proof. Exercise! □

Remark 1.27. In higher dimensions and $f \in L^1(\mathbb{R}^n)$, $n \geq 1$, one defines for $\omega \in \mathbb{R}^n$ the Fourier transform of f by

$$\mathscr{F}(f)(\omega) := \int_{\mathbb{R}^n} f(t)e^{-i\langle t,\omega\rangle}\,dt.$$

Theorem 1.28. *Suppose that $f \in L^1(\mathbb{R})$. Then the function $\mathscr{F}(f) : \mathbb{R} \to \mathbb{C}$ is bounded and uniformly continuous.*

Proof. Exercise! □

Remark 1.29. Note that in case of the Fourier transform in $L^1(\mathbb{T})$ we dealt with a discrete-frequency spectrum, whereas here in $L^1(\mathbb{R})$ we consider a continuous-frequency spectrum $\omega \in \mathbb{R}$.

Completely analogous to the case $L^1(\mathbb{T})$, the Banach space $L^1(\mathbb{R})$ can be endowed with a convolution structure and thus becomes a Banach algebra.

Theorem 1.30. *Suppose that $f,g \in L^1(\mathbb{R})$. For almost all $t \in \mathbb{R}$ is the mapping $s \mapsto f(t-s)\cdot g(s)$ absolutely integrable. Let*

$$h(t) := \int_{\mathbb{R}} f(t-s)g(s)\,ds.$$

Then $h \in L^1(\mathbb{R})$ and

$$\|h\|_1 \le \|f\|_1 \cdot \|g\|_1 \quad \text{and} \quad \mathscr{F}(h)(\omega) = \mathscr{F}(f)(\omega)\cdot\mathscr{F}(g)(\omega), \quad \forall \omega \in \mathbb{R}.$$

Proof. Analogous to Theorem 1.5. Exercise! □

Definition 1.31. Assume that $f,g \in L^1(\mathbb{R})$. The function h in Theorem 1.30 is denoted by $f * g$ and called the *convolution* of f and g.

Remark 1.32. The Banach space $L^1(\mathbb{T})$ together with the operation of convolution becomes a commutative Banach algebra without unity. The Fourier transform is a Banach algebra homomorphism

$$\mathscr{F} : L^1(\mathbb{T}) \to l^\infty(\mathbb{Z}),$$

with respect to the Banach algebra $l^\infty(\mathbb{Z})$ with elementwise multiplication.

Remark 1.33. The Fourier transform is an algebra homomorphism in $C_u^b(\mathbb{R})$, i.e., the Banach algebra of pointwise multiplication of all uniformly continuous (u) and bounded (b) functions on $\mathbb{R}$.

Theorem 1.34. *Let $f,k \in L^1(\mathbb{R})$ and suppose that*

$$k(t) = \int_{\mathbb{R}} K(\omega)e^{i\omega t}\,da, \quad \text{where } K \in L^1(\mathbb{R}).$$

Then

$$k * f(t) = \int_{\mathbb{R}} K(\omega)\mathscr{F}(f)(\omega)e^{i\omega t}\,d\omega.$$

Proof. Exercise! □

Theorem 1.35. *Let $f \in L^1(\mathbb{R})$.*

1. Let

$$F(t) = \int_{-\infty}^{t} f(s)\,ds = \int_{\mathbb{R}} f(s)\chi_{(-\infty,t]}(s)\,ds, \quad \text{for } t \in \mathbb{R}.$$

If $F \in L^1(\mathbb{R})$, then

$$\mathscr{F}(F)(\omega) = \frac{1}{i\omega}\mathscr{F}(f)(\omega),$$

for all $\omega \in \mathbb{R}\setminus\{0\}$.

2. *Suppose that f is a differentiable function on $\mathbb{R}$ whose derivative $f' \in L^1(\mathbb{R})$. Then*
$$\mathscr{F}(f)(\omega) = \frac{1}{i\omega}\mathscr{F}(f')(\omega),$$
for all $\omega \in \mathbb{R} \setminus \{0\}$.

Proof. It suffices to show 1. We know that $F'(t) = f(t)$ holds for almost all $t \in \mathbb{R}$. Integration by parts and the dominated convergence theorem yield

$$\mathscr{F}(F)(\omega) = \lim_{A\to\infty} F(t)\frac{1}{-i\omega}e^{-i\omega t}\Big|_{t=-A}^{A} + \int_{\mathbb{R}} f(t)\frac{1}{i\omega}e^{-i\omega t}\,dt.$$

Obviously, $\lim_{A\to-\infty} F(A) = 0$. Since f is integrable, the limit $\lim_{A\to\infty} F(A)$ exists and is finite, for
$$\lim_{A\to\infty} F(A) = \int_{\mathbb{R}} f(t)\,dt < \infty$$
exists.

Assume that $\lim_{A\to\infty} F(A) = \alpha \neq 0$. Then there exists $A_0 > 0$ so that $|F(A)| \geq |\alpha|/2 > 0$ for all $A > A_0$. This, however, is a contradiction to $F \in L^1(\mathbb{R})$. □

Theorem 1.36. *Suppose that $f \in L^1(\mathbb{R})$. Set $g(t) := tf(t)$, and assume that $g \in L^1(\mathbb{R})$. Then, for all $\omega \in \mathbb{R}$, we have that $\mathscr{F}(f)$ is differentiable and $(\mathscr{F}(f))'(\omega) = \mathscr{F}(-ig)(\omega)$.*

Proof. Consider

$$\frac{\mathscr{F}(f)(\omega+h) - \mathscr{F}(f)(\omega)}{h} = \int_{\mathbb{R}} f(t)\,e^{-i\omega t}\,\frac{e^{-iht}-1}{h}\,dt.$$

Now,
$$\left|\frac{e^{-iht}-1}{h}\right| \leq |t|,$$
and
$$\frac{e^{-iht}-1}{h} \to -it \quad \text{for } h \to 0.$$
Since $g(t) = tf(t) \in L^1(\mathbb{R})$ is absolutely integrable, the dominated convergence theorem yields

$$(\mathscr{F}(f))'(\omega) = -i\int_{\mathbb{R}} f(t)e^{-i\omega t}t\,dt = -i\mathscr{F}(g)(\omega). \quad \square$$

Induction implies a corollary to the above theorem.

Corollary 1.37. *Suppose $f \in L^1(\mathbb{R})$ is such that $t \mapsto t^n f(t) =: g(t)$ is absolutely integrable for an $n \in \mathbb{N}$. Then $\mathscr{F}(f)$ is n-times differentiable and*

$$(\mathscr{F}(f))^{(n)}(\omega) = (i)^n\mathscr{F}(g)(\omega), \quad \textit{for all } \omega \in \mathbb{R}.$$

In particular, for $\omega = 0$, we obtain

$$(\mathscr{F}(f))^{(n)}(0) = (-i)^n m_n,$$

where

$$m_n := \int_{\mathbb{R}} t^n f(t)\,dt$$

is the nth moment of f.

Corollary 1.38. *Suppose that $f : \mathbb{R} \to \mathbb{C}$ has compact support and is twice continuously differentiable. Then $\mathscr{F}(f) \in L^1(\mathbb{R})$.*

Proof. Since f has compact support and is an element of C^2, the derivatives f' and f'' are also compactly supported. Theorem 1.35 implies that

$$|\mathscr{F}(f)(\omega)| = \left|\frac{\mathscr{F}(f'')(\omega)}{\omega^2}\right| \leq \frac{1}{\omega^2}\|f''\|_1.$$

Divide $\mathbb{R}$ into the sets $[-1,1]$ and $\mathbb{R}\setminus[-1,1]$. On the compact set $[-1,1]$ the continuous function $\mathscr{F}(f)$ is bounded and thus integrable. On $\mathbb{R} \setminus [-1,1]$ the function $\mathscr{F}(f)(\omega)$ is also integrable since it decays like $1/\omega^2$. Hence, $\mathscr{F}(f) \in L^1(\mathbb{R})$. □

Our next goal is to find approximations for the missing unity in the Banach algebra $(L^1(\mathbb{R}), *)$. For this purpose, we need a definition.

Definition 1.39. A *summation kernel* on $\mathbb{R}$ is a family $\{k_\lambda\}_{\lambda\in(0,\infty)}$ of continuous functions with the following properties.

(S1) $\int_{\mathbb{R}} k_\lambda(t)\,dt = 1$, for all $\lambda \in (0,\infty)$.

(S2) $\int_{\mathbb{R}} |k_\lambda(t)|\,dt \leq M$, for all $\lambda \in (0,\infty)$ and a constant $M > 0$.

(S3) For all $\delta > 0$, we have that

$$\lim_{\lambda\to\infty}\int_{|t|>\delta} |k_\lambda(t)|\,dt = 0.$$

Example 1.40. 1. **All summation kernels are found via the following procedure.** Choose an $f \in L^1(\mathbb{R}) \cap C(\mathbb{R})$ with $\int_{\mathbb{R}} f(t)\,dt = 1$ and set $k_\lambda(t) := \lambda f(\lambda t)$. Then

- $\int_{\mathbb{R}} K_\lambda(t)\,dt = \int_{\mathbb{R}} \lambda f(\lambda t)\,dt = \int_{\mathbb{R}} f(s)\,ds = 1$. Hence, (S1) is satisfied.
- $\int_{\mathbb{R}} |K_\lambda(t)|\,dt = \int_{\mathbb{R}} |\lambda f(\lambda t)|\,dt = \int_{\mathbb{R}} |f(s)|\,ds = \|f\|_1$. Thus, (S2) holds.
- For $\delta > 0$, we have

$$\int_{|t|>\delta} |K_\lambda(t)|\,dt = \int_{|t|>\delta} |\lambda f(\lambda t)|\,dt = \int_{|s|>\delta\lambda} |f(s)|\,ds \to 0, \quad \text{as } \lambda \to \infty.$$

Therefore, (S3) is also valid.

2. **The Fejér kernel:**

$$F(t) := \frac{1}{2\pi}\left(\frac{\sin\frac{t}{2}}{t/2}\right)^2, \quad F_\lambda(t) = \lambda F(\lambda t), \quad \lambda \in (0,\infty).$$

$(F_\lambda)_{\lambda\in(0,\infty)}$ is called the Fejér kernel on $\mathbb{R}$ and it is a summation kernel. To verify this last claim, it suffices to show that $\int_{\mathbb{R}} F_\lambda(t)\,dt = 1$.
To this end, we use the Fejér kernel

$$\frac{1}{n+1}\left(\frac{\sin\frac{n+1}{2}t}{\sin\frac{t}{2}}\right)^2 = \sum_{k=-n}^{n}\left(1-\frac{|k|}{n+1}\right)e^{ikt}$$

defined on $\mathbb{T}$. It is easy to verify that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{n+1}\left(\frac{\sin\frac{n+1}{2}t}{\sin\frac{t}{2}}\right)^2 dt = 1$$

and

$$\lim_{n\to\infty}\frac{1}{2\pi}\int_{\delta}^{2\pi-\delta}\frac{1}{n+1}\left(\frac{\sin\frac{n+1}{2}t}{\sin\frac{t}{2}}\right)^2 dt = 0.$$

Thus,

$$\lim_{n\to\infty}\frac{1}{2\pi}\int_{-\delta}^{\delta}\frac{1}{n+1}\left(\frac{\sin\frac{n+1}{2}t}{\sin\frac{t}{2}}\right)^2 dt = 1. \tag{1.11}$$

For every $0<\varepsilon<1$ exists a $\delta>0$, so that for all $|t|<\delta$, the following inequalities hold:

$$\left(\sin\frac{t}{2}\right)^2 \le \left(\frac{t}{2}\right)^2 \le (1+\varepsilon)\left(\sin\frac{t}{2}\right)^2.$$

For such a $\delta>0$, we thus have that

$$\begin{aligned}
&\frac{1}{1+\varepsilon}\frac{1}{2\pi}\int_{-\delta}^{\delta}\frac{1}{n+1}\left(\frac{\sin\frac{n+1}{2}t}{\sin\frac{t}{2}}\right)^2 dt\\
&\qquad\le\frac{1}{2\pi}\int_{-\delta}^{\delta}\frac{1}{n+1}\left(\frac{\sin\frac{n+1}{2}t}{\frac{t}{2}}\right)^2 dt\\
&\qquad\le\frac{1}{2\pi}\int_{-\delta}^{\delta}\frac{1}{n+1}\left(\frac{\sin\frac{n+1}{2}t}{\sin\frac{t}{2}}\right)^2 dt.
\end{aligned} \tag{1.12}$$

Setting $\lambda := n+1$ yields

$$\int_{\mathbb{R}} F(t)\,dt = \int_{\mathbb{R}} F_{n+1}(t)\,dt = \int_{-\delta}^{\delta} F_{n+1}(t)\,dt + \int_{\mathbb{R}\setminus[-\delta,\delta]} F_{n+1}(t)\,dt. \tag{1.13}$$

The second integral in (1.13) vanishes:

$$\int_{\mathbb{R}\setminus[-\delta,\delta]} F_{n+1}(t)\,dt = \frac{1}{2\pi}\int_{-\delta}^{\delta}\frac{1}{n+1}\left(\frac{\sin(\frac{n+1}{2}t)}{\frac{t}{2}}\right)^2 dt \to 0 \quad \text{for } n\to\infty.$$

Hence,

$$\int_{\mathbb{R}} F(t)\,dt = \lim_{n\to\infty}\int_{-\delta}^{\delta} F_{n+1}(t)\,dt = \lim_{n\to\infty}\frac{1}{2\pi}\int_{-\delta}^{\delta}\frac{1}{n+1}\left(\frac{\sin\frac{n+1}{2}t}{\frac{t}{2}}\right)^2 dt.$$

Equations (1.11) and (1.12) now imply that

$$\frac{1}{1+\varepsilon} \le \lim_{n\to\infty}\frac{1}{2\pi}\int_{-\delta}^{\delta}\frac{1}{n+1}\left(\frac{\sin\frac{n+1}{2}t}{\frac{t}{2}}\right)^2 dt \le 1.$$

Since $0<\varepsilon<1$ was arbitrary, the claim that $\int_{\mathbb{R}} F(t)\,dt = 1$ follows.

Definition 1.41. A family $(k_\lambda)_{\lambda\in(0,\infty)}$ is called an *approximate identity* for $L^1(\mathbb{R})$ if

$$\lim_{\lambda\to\infty}\|f-k_\lambda * f\|_1 = 0, \quad \forall f\in L^1(\mathbb{R}).$$

It can be shown that every summation kernel on $\mathbb{R}$ is an approximate identity for $L^1(\mathbb{R})$.

Theorem 1.42. *Assume that $\{k_\lambda\}_{\lambda\in(0,\infty)}$, is a summation kernel. Then*

$$\lim_{\lambda\to\infty}\|f-k_\lambda * f\|_1 = 0, \quad \forall f\in L^1(\mathbb{R}).$$

Proof. See, e.g., [151, 162]. □

The special case $k_\lambda := F_\lambda$ yields a corollary.

Corollary 1.43. *Let $f\in L^1(\mathbb{R})$. For a given $\lambda>0$, set*

$$\sigma_\lambda(f)(t) := F_\lambda * f(t) = \frac{1}{2\pi}\int_{-\lambda}^{\lambda}\left(1-\frac{|\omega|}{\lambda}\right)\mathscr{F}(f)(\omega)e^{i\omega t}\,d\omega.$$

Then

$$\lim_{\lambda\to\infty}\|f-F_\lambda * f\|_1 = 0.$$

Proof. Only the validity of the second equality needs to be shown. For this purpose, define

$$\varDelta(t) := \begin{cases} \frac{1}{2\pi}(1-|t|), & \text{for } |t|\le 1,\\ 0, & \text{else.}\end{cases}$$

Then, for $\omega \neq 0$, one has that

$$\int_0^1 (1-t)e^{-i\omega t}\,dt = -\frac{1}{i\omega} - \frac{1}{\omega^2}(e^{i\omega}-1)$$

and

$$\int_{-1}^0 (1+t)e^{-i\omega t}\,dt = \frac{1}{i\omega} - \frac{1}{\omega^2}(e^{-i\omega}-1).$$

Hence, for all $\omega \neq 0$,

$$\begin{aligned}\mathscr{F}(\Delta)(\omega) &= \frac{1}{2\pi}\int_{-1}^1 (1-|t|)e^{-it\omega}\,dt \\ &= \frac{1}{2\pi}\int_{-1}^1 \frac{2-2\cos\omega}{\omega^2} \quad = \quad \frac{1}{2\pi}\left(\frac{\sin\frac{\omega}{2}}{\frac{\omega}{2}}\right)^2 \quad = \quad F(\omega).\end{aligned}$$

In addition, $\mathscr{F}(\Delta)(0) = 1/2\pi = F(0)$. Therefore, $F = \mathscr{F}(\Delta)$ and

$$F_\lambda(\omega) = \lambda F(\lambda\omega) = \lambda\mathscr{F}(\Delta)(\lambda\omega) = \mathscr{F}(\Delta^\lambda)(\omega),$$

where

$$\Delta^\lambda(t) = \begin{cases} \frac{1}{2\pi}\left(1-\frac{|t|}{\lambda}\right), & \text{for } |t| < \lambda, \\ 0, & \text{else.}\end{cases}$$

Thus,

$$F_\lambda(t) = \frac{1}{2\pi}\int_{-\lambda}^{\lambda}\left(1-\frac{|\omega|}{\lambda}\right)e^{-i\omega t}\,d\omega = \frac{1}{2\pi}\int_{-\lambda}^{\lambda}\left(1-\frac{|\omega|}{\lambda}\right)e^{i\omega t}\,d\omega,$$

since Δ is an even function.

Theorem 1.34 implies

$$F_\lambda * f(t) = \frac{1}{2\pi}\int_{-\lambda}^{\lambda}\left(1-\frac{|\omega|}{\lambda}\right)\mathscr{F}(f)(\omega)e^{i\omega t}\,d\omega.$$

Now, F_λ is a summation kernel and according to Theorem 1.42 an approximate identity for $L^1(\mathbb{R})$, which implies the statement. □

Corollary 1.44 (Uniqueness).
Let $f \in L^1(\mathbb{R})$ satisfy $\mathscr{F}(f)(\omega) = 0$ for all $\omega \in \mathbb{R}$. Then $f = 0$ almost everywhere.

Corollary 1.45 (Inversion formula).
Let $f \in L^1(\mathbb{R})$ satisfy $\mathscr{F}(f) \in L^1(\mathbb{R})$. Then

$$f(t) = \frac{1}{2\pi}\int_{\mathbb{R}} \mathscr{F}(f)(\omega)e^{i\omega t}\,d\omega, \quad \textit{for almost all } t \in \mathbb{R},$$

with equality at the points of continuity of f.

Proof. For fixed $t \in \mathbb{R}$, the function

$$\mathbb{R} \to \mathbb{C}: \quad \omega \mapsto \chi_{[-\lambda,\lambda]}(\omega) \left(1 - \frac{|\omega|}{\lambda}\right) \mathscr{F}(f)(\omega) e^{i\omega t}$$

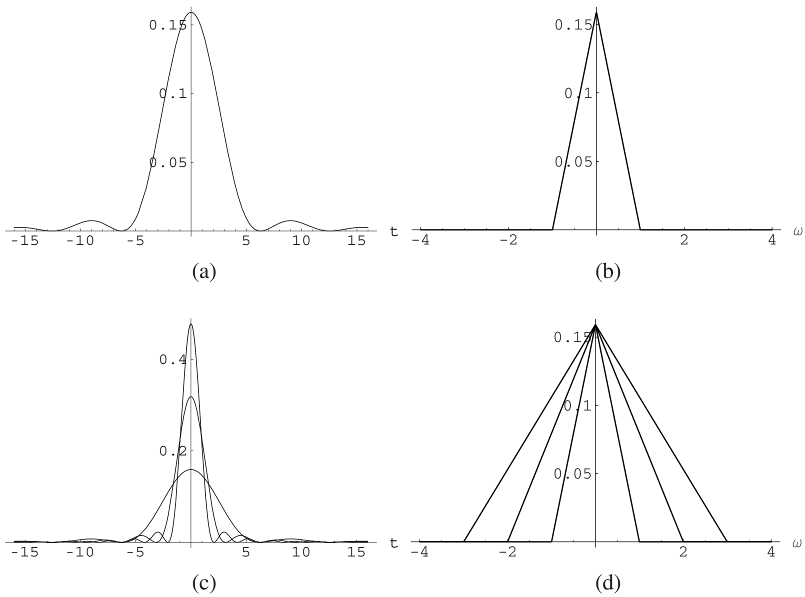


Fig. 1.2: (a) The Fejér kernel $F(t)$ and (b) its Fourier integral transform $\Delta(\omega)$ in $L^1(\mathbb{R})$. (c) For scales $\lambda \cdot F(\lambda t)$ with increasing $\lambda = 1, 2, 3$, the Fejér kernel becomes narrower and higher and (d) the Fourier integral transform wider.

has as its limit the function

$$\omega \mapsto \mathscr{F}(f)(\omega) e^{i\omega t},$$

as $\lambda \to \infty$. The modulus of this function can be bounded above by $|\mathscr{F}(f)|$. As by assumption $\mathscr{F}(f) \in L^1(\mathbb{R})$, the dominated convergence theorem in the L^1-norm implies

$$\begin{aligned}\lim_{\lambda\to\infty} F_\lambda * f(t) &= \lim_{\lambda\to\infty} \frac{1}{2\pi} \int_{\mathbb{R}} \chi_{[-\lambda,\lambda]}(\omega) \left(1 - \frac{|\omega|}{\lambda}\right) \mathscr{F}(f)(\omega) e^{i\omega t}\, d\omega \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \mathscr{F}(f)(\omega) e^{i\omega t}\, d\omega.\end{aligned}$$

Note that if a sequence $g_n \to g$ in $L^1(\mathbb{R})$ and also $\|g_n - g\| \to 0$ as $n \to \infty$, then there exists a subsequence g_{n_k} such that

$$\lim_{k\to\infty} g_{n_k}(t) = g(t) \quad \text{a.e.}$$

Applying this to the setting at hand, we deduce the existence of a subsequence $\{F_{\lambda_k}\}_{k\in\mathbb{N}}$ with the property that $F_{\lambda_k} * f \to f$ almost everywhere as $k \to \infty$.

The statement regarding the points of continuity of f follows from a theorem ahead. □

Definition 1.46. For a $g \in L^1(\mathbb{R})$, we denote the *inverse Fourier transform* of g by

$$\mathscr{F}^{-1}(g)(t) := \frac{1}{2\pi}\int_{\mathbb{R}} g(\omega)e^{i\omega t}\,d\omega.$$

Remark 1.47. For functions $f \in L^1(\mathbb{R})$ such that $\mathscr{F}(f) \in L^1(\mathbb{R})$, one has that

$$\mathscr{F}^{-1}(\mathscr{F}(f))(t) = f(t) \quad \text{a.e.}$$

Example 1.48 (The Fejér kernel revisited). We already know that $\mathscr{F}(\Delta) = F$. Since $\mathscr{F}(\Delta)$ has compact support, $\mathscr{F}(\Delta) \in L^1(\mathbb{R})$ and the inversion formula

$$\Delta(\omega) = \frac{1}{2\pi}\mathscr{F}(F)(\omega)$$

holds.

Corollary 1.49 (Continuous analogue of the theorem of Weierstraß). *Let $C_c(\mathbb{R})$ denote the space of all continuous functions $\mathbb{R} \to \mathbb{C}$ with compact support. The set of functions $f \in L^1(\mathbb{R})$ with $\mathscr{F}(f) \in C_c(\mathbb{R})$ forms a dense (with respect to the norm $\|\bullet\|_1$) subspace of $L^1(\mathbb{R})$.*

Proof.

$$\mathscr{F}(F_\lambda * f)(\omega) = \mathscr{F}(F_\lambda)(\omega)\mathscr{F}(f)(\omega) = \chi_{[-\lambda,\lambda]}(\omega)\left(1 - \frac{|\omega|}{\lambda}\right)\mathscr{F}(f)(\omega),$$

i.e., $\mathscr{F}(F_\lambda * f) \in C_c(\mathbb{R})$. The claim follows now from the fact that the Fejér kernel is a summation kernel. □

Corollary 1.50 (Riemann–Lebesgue lemma).
Let $f \in L^1(\mathbb{R})$. Then $\lim_{|\omega|\to\infty}\mathscr{F}(f)(\omega) = 0$.

Proof. Use Corollary 1.49. The details are left to the reader. □

Theorem 1.51. *Let $f \in L^1(\mathbb{R})$ and define*

$$\sigma_\lambda(f)(t) := \frac{1}{2\pi}\int_{-\lambda}^{\lambda}\left(1 - \frac{|\omega|}{\lambda}\right)\mathscr{F}(f)(\omega)e^{it\omega}\,da.$$

Then $\sigma_\lambda(f)(t) \to f(t)$ as $\lambda \to \infty$, for almost all $t \in \mathbb{R}$.

If t is a point of continuity of f, then $\sigma_\lambda(f)(t)$ converges to $f(t)$.

Proof. See, for instance, [162]. □

Example 1.52 (Additional summation kernels on $\mathbb{R}$).

(i) **Cauchy–Poisson kernel**

$$P(t) := \frac{1}{\pi}\left(\frac{1}{1+t^2}\right), \quad P_\lambda(t) := \lambda P(\lambda t), \quad \lambda \in (0,\infty),$$

is a summation kernel (see Fig. 1.3).

$$\frac{1}{\pi}\int_{\mathbb{R}} \frac{1}{1+t^2}\,dt = \frac{1}{\pi}\arctan(t)\Big|_{t=-\infty}^{\infty} = 1.$$

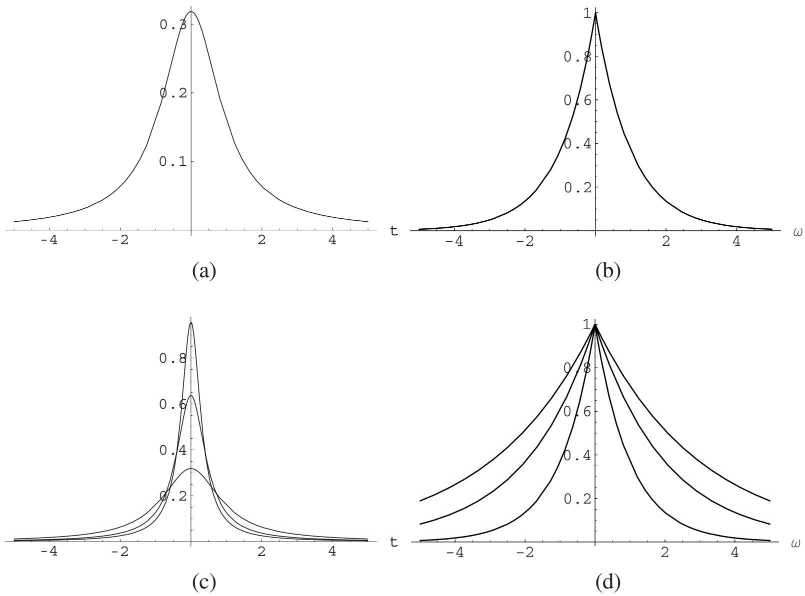


Fig. 1.3: (a) The Cauchy–Poisson kernel P and (b) its Fourier transform in $L^1(\mathbb{R})$. (c) For scales $\lambda \cdot P(\lambda t)$ with increasing $\lambda = 1,2,3$, the Cauchy–Poisson kernel becomes narrower and higher, and (d) the Fourier transform wider.

Hence, $(P_\lambda)_{\lambda\in(0,\infty)}$ is a summation kernel. The Fourier transform is given by

$$\mathscr{F}(P_\lambda)(\omega) = \exp\left(-\left|\frac{\omega}{\lambda}\right|\right),$$

since

$$f(t) = \chi_{[0,\infty)}(t)e^{-t}, \quad \mathscr{F}(f)(\omega) = \frac{1}{1+i\omega},$$
$$g(t) = \chi_{(-\infty,0]}(t)e^{t}, \quad \mathscr{F}(g)(\omega) = \frac{1}{1-i\omega}.$$

Set $h(t) = f(t) + g(t) = e^{-|t|}$. Then $\mathscr{F}(h)(\omega) = 2/(1+\omega^2)$. However, h, $\mathscr{F}(h) \in L^1(\mathbb{R})$ and application of the inversion theorem yields

$$\begin{aligned}\mathscr{F}(P_\lambda)(\omega) &= \int_{\mathbb{R}} P_\lambda(t) e^{-i\omega t}\, dt \\ &= \frac{\lambda}{2\pi} \int_{\mathbb{R}} \mathscr{F}(h)(\lambda t) e^{-i\omega t}\, dt \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \mathscr{F}(h)(s) e^{-i\frac{\omega}{\lambda}s}\, ds \\ &= \mathscr{F}^{-1}\mathscr{F} h\left(-\frac{\omega}{\lambda}\right) \quad = \quad e^{-|\frac{\omega}{\lambda}|}.\end{aligned}$$

(ii) **Gauß kernel**

$$G(t) := \frac{1}{\sqrt{\pi}} e^{-t^2}, \quad G_\lambda(t) := \lambda G(\lambda t), \quad \lambda \in (0,\infty)$$

(see Fig. 1.4). Note that $\int_{\mathbb{R}} G(t)\, dt = 1$. Thus, $\{G_\lambda\}_{\lambda \in (0,\infty)}$ is a summation kernel. Moreover,

$$\mathscr{F}(G)(\omega) = \exp\left(-\left(\frac{\omega}{2}\right)^2\right).$$

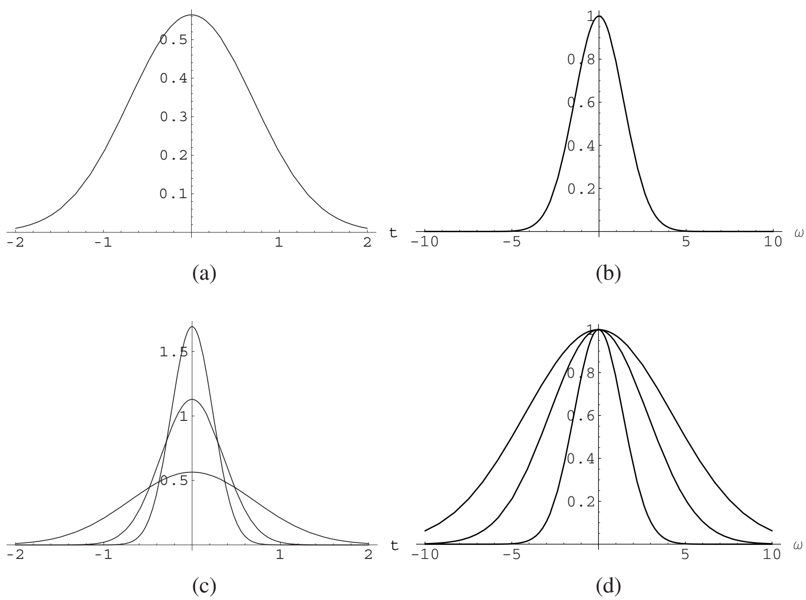


Fig. 1.4: (a) The Gauß kernel G and (b) its Fourier transform in $L^1(\mathbb{R})$. (c) For scales $\lambda \cdot G(\lambda t)$ with increasing $\lambda = 1, 2, 3$, the Gauß kernel becomes narrower and higher, and (d) the Fourier transform wider.

1.3.2 The Plancherel Transform

For many applications, a Hilbert space model is more appropriate since in this case an inner product is available. In this section, we extend the Fourier transform to the Hilbert space $L^2(\mathbb{R})$.

To this end, we use the fact that $C_c(\mathbb{R})$, the space of all continuous functions with compact support, is a subspace of $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ and dense in $L^1(\mathbb{R})$ and $L^2(\mathbb{R})$ in the respective norm topology.

Lemma 1.53. *If $f \in C_c(\mathbb{R})$, then $\mathscr{F}(f) \in L^2(\mathbb{R})$, and*

$$\frac{1}{2\pi} \int_{\mathbb{R}} |\mathscr{F}(f)(\omega)|^2 \, d\omega = \int_{\mathbb{R}} |f(t)|^2 \, dt.$$

Proof. See, for instance, [209]. □

Definition 1.54. Suppose that $f \in L^2(\mathbb{R})$ and $\{f_n\}_{n\in\mathbb{N}} \subset C_c(\mathbb{R})$ is an arbitrary sequence that converges in $L^2(\mathbb{R})$ to f. The limit

$$\mathscr{P}(f) = \lim_{n\to\infty} \mathscr{F}(f_n) \in L^2(\mathbb{R})$$

in the $L^2(\mathbb{R})$-norm is called the *Plancherel transform* of f.

Remark 1.55. The Plancherel transform and the Fourier transform are both defined on $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ and agree there. Indeed, let $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ and choose a sequence $\{f_n\}_{n\in\mathbb{N}} \subset C_c(\mathbb{R})$ that converges in $L^1(\mathbb{R})$ and $L^2(\mathbb{R})$ to f. Then

$$\|\mathscr{F}(f_n) - \mathscr{P}(f)\|_2 \to 0 \quad \text{and} \quad \mathscr{F}(f_n)(\omega) \to \mathscr{F}(f)(\omega), \quad \text{for all } \omega \in \mathbb{R} \text{ and } n \to \infty.$$

Since $\mathscr{F}(f_n) \to \mathscr{P}(f)$ in $L^2(\mathbb{R})$, there exists a subsequence $\{n_k\}_{k\in\mathbb{N}} \subset \mathbb{N}$ so that $\mathscr{F}(f_n)(\omega) \to \mathscr{P}(f)(\omega)$ for almost all $\omega \in \mathbb{R}$. Hence, $\mathscr{F}(f) = \mathscr{P}(f)$ almost everywhere.

Lemma 1.56. *Let $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. Then $\mathscr{F}(f) = \mathscr{P}(f)$ almost everywhere. Moreover,*

$$\frac{1}{2\pi} \int_{\mathbb{R}} |\mathscr{F}(f)(\omega)|^2 \, d\omega = \int_{\mathbb{R}} |f(t)|^2 \, dt.$$

Theorem 1.57 (An equivalent description of the Plancherel transform). *Let $f \in L^2(\mathbb{R})$. For $\lambda > 0$, let $f_\lambda := \chi_{[-\lambda,\lambda]} f$. Then $f_\lambda \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ and $\mathscr{F}(f_\lambda) \in L^2(\mathbb{R})$. In particular,*

$$\lim_{\lambda\to\infty} \|\mathscr{P}(f) - \mathscr{F}(f_\lambda)\|_2 = 0.$$

Proof. Exercise! (Hint: Use Lemma 1.56.) □

So far we know that the Plancherel transform is a bounded linear operator $\mathscr{P} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$. In addition, we know that

$$\frac{1}{\sqrt{2\pi}} \|\mathscr{P}(f)\|_2 = \|f\|_2, \quad \forall f \in L^2(\mathbb{R}).$$

Hence, $\mathscr{P}$ is injective. To show that $\mathscr{P}$ is also surjective, we prove an inversion formula.

Lemma 1.58 (Parseval). *Let* $f,g \in L^2(\mathbb{R})$. *Then*

$$\langle \mathscr{P}(f), \overline{g} \rangle = \langle \mathscr{P}(g), \overline{f} \rangle.$$

Proof. Let $f,g \in C_c(\mathbb{R})$. Fubini's theorem yields

$$\begin{aligned} \langle \mathscr{F}(f), \overline{g} \rangle = \int_{\mathbb{R}} \mathscr{F}(f)(a) g(\omega)\, d\omega \quad &= \quad \int_{\mathbb{R}} \int_{\mathbb{R}} f(t) e^{-i\omega t}\, dt\, g(\omega)\, d\omega \\ = \int_{\mathbb{R}} f(t) \mathscr{F}(g)(t)\, dt \quad &= \quad \langle \mathscr{F}(g), \overline{f} \rangle. \end{aligned}$$

For $f,g \in L^2(\mathbb{R})$, choose convergent sequences $\{f_n\}_{n\in\mathbb{N}}, \{g_n\}_{n\in\mathbb{N}} \subset C_c(\mathbb{R})$ so that $f_n \to f$ and $g_n \to g$ as $n \to \infty$ in the $L^2(\mathbb{R})$-norm. From the continuity of the inner product, it follows that

$$\langle \mathscr{P}(f), \overline{g} \rangle = \lim_{n,m\to\infty} \langle \mathscr{F}(f_n), \overline{g_m} \rangle = \lim_{n,m\to\infty} \langle \mathscr{F}(g_m), \overline{f_n} \rangle = \langle \mathscr{P}(g), \overline{f} \rangle. \quad \square$$

Theorem 1.59 (Inversion formula for the Plancherel transform). *Let* $f \in L^2(\mathbb{R})$. *Then*

$$f = \frac{1}{2\pi} \overline{\mathscr{P}(\overline{\mathscr{P}(f)})} \quad \textit{almost everywhere.}$$

Proof. Set $g := \overline{\mathscr{P}(f)}$. The linearity of the inner product yields

$$\left\| f - \frac{1}{2\pi} \overline{\mathscr{P}(g)} \right\|_2^2 = \|f\|_2^2 - \left\langle f, \frac{1}{2\pi} \overline{\mathscr{P}(g)} \right\rangle - \left\langle \frac{1}{2\pi} \overline{\mathscr{P}(g)}, f \right\rangle + \left\| \frac{1}{2\pi} \overline{\mathscr{P}(g)} \right\|_2^2.$$

By Parseval's equality, we have that

$$\left\langle f, \frac{1}{2\pi} \overline{\mathscr{P}(g)} \right\rangle = \frac{1}{2\pi} \langle \mathscr{P}(g), \overline{f} \rangle = \frac{1}{2\pi} \langle \mathscr{P}(f), \overline{g} \rangle = \frac{1}{2\pi} \|\mathscr{P}(f)\|_2^2 = \|f\|_2^2$$

and

$$\left\langle \frac{1}{2\pi} \overline{\mathscr{P}(g)}, f \right\rangle = \frac{1}{2\pi} \overline{\langle \mathscr{P}(g), \overline{f} \rangle} = \|f\|_2^2,$$

as well as

$$\left\| \frac{1}{2\pi} \overline{\mathscr{P}(g)} \right\|_2^2 = \frac{1}{4\pi^2} \|\mathscr{P}(g)\|_2^2 = \frac{1}{2\pi} \|g\|_2^2 = \|f\|_2^2.$$

Thus,

$$\left\| f - \frac{1}{2\pi} \overline{\mathscr{P}(g)} \right\|_2^2 = 0,$$

which implies the statement in the theorem. $\square$

Hence, $\mathscr{P} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ is a topological automorphism.

In summary, we arrive at the following theorem.

Theorem 1.60 (Plancherel). *The linear operator $\mathscr{P} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ is an isomorphism from $L^2(\mathbb{R})$ onto $L^2(\mathbb{R})$. It has the following properties.*

1. *For all $f, g \in L^2(\mathbb{R})$, one has that*
$$\frac{1}{2\pi}\langle \mathscr{P}(f), \mathscr{P}(g)\rangle = \langle f, g\rangle.$$
In particular,
$$\frac{1}{\sqrt{2\pi}}\|\mathscr{P}(f)\|_2 = \|f\|_2.$$
2. *$\mathscr{P}(f) = \mathscr{F}(f)$ for $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$.*
3. *$\mathscr{P}(f)(\omega) = \lim_{\lambda\to\infty} \int_{-\lambda}^{\lambda} f(t)e^{-i\omega t}\,dt$ in the L^2-norm.*
4. *$f(t) = \frac{1}{2\pi}\lim_{\lambda\to\infty} \int_{-\lambda}^{\lambda} \mathscr{P}(f)(\omega)e^{i\omega t}\,d\omega$ in the L^2-norm.*

Proof. Only item 1 needs to be shown. However, Plancherel's formula and the inversion theorem imply that
$$\frac{1}{2\pi}\langle \mathscr{P}(f), \mathscr{P}(g)\rangle = \frac{1}{2\pi}\left\langle \mathscr{P}(\overline{\mathscr{P}(g)}), \overline{f}\right\rangle = \langle \overline{g}, \overline{f}\rangle = \langle f, g\rangle. \quad \square$$

We introduce the convolution $* : L^1(\mathbb{R}) \times L^p(\mathbb{R}) \to L^p(\mathbb{R})$, $1 < p < \infty$. For $f \in L^p(\mathbb{R})$ and $g \in L^1(\mathbb{R})$, define
$$g * f(t) := \int_{\mathbb{R}} f(t-s)g(s)\,ds.$$
One can show that $g * f$ exists for almost all $t \in \mathbb{R}$ and that $g * f \in L^p(\mathbb{R})$, satisfying the estimate
$$\|g * f\|_p \le \|g\|_1 \|f\|_p.$$
In the special case $p := 2$, we obtain the next result.

Theorem 1.61. *Let $g \in L^1(\mathbb{R})$ and $f \in L^2(\mathbb{R})$. Then*
$$g * f(t) = \int_{\mathbb{R}} f(t-s)g(s)\,ds$$
is an element of $L^2(\mathbb{R})$ and the estimate
$$\|g * f\|_2 \le \|g\|_1 \|f\|_2$$
holds. Moreover, the Plancherel transform satisfies
$$\mathscr{P}(g * f) = \mathscr{F}(g)\mathscr{P}(f).$$

Proof. Only the last statement needs to be proven. This, however, is left as an exercise to the reader. $\square$

Note that if $f \in L^2(\mathbb{R})$, then $F_\lambda * f \in L^2(\mathbb{R})$. In particular,

$$F_\lambda * f(t) = \frac{1}{2\pi}\int_{-\lambda}^{\lambda}\left(1-\frac{|\omega|}{\lambda}\right)\mathscr{P}(f)(\omega)e^{i\omega t}\,d\omega.$$

This last equality can be shown as follows. Set $\varphi(\omega) := \Delta^\lambda(\omega)e^{i\omega t}$. Then

$$\mathscr{F}(\varphi)(s) = \int_{\mathbb{R}} \Delta^\lambda(\omega)e^{-i\omega(s-t)}\,d\omega = F_\lambda(s-t).$$

Since φ is continuous and has compact support, $\mathscr{F}(\varphi) = \mathscr{P}(\varphi)$ holds. Plancherel's theorem then implies

$$\begin{aligned} F_\lambda * f(t) &= \int_{\mathbb{R}} F_\lambda(t-s)f(s)\,ds = \int_{\mathbb{R}} F_\lambda(s-t)f(s)\,ds \\ &= \int_{\mathbb{R}} \mathscr{P}(\varphi)(s)f(s)\,ds = \int_{\mathbb{R}} \mathscr{P}(f)(\omega)\varphi(\omega)\,d\omega \\ &= \frac{1}{2\pi}\int_{-\lambda}^{\lambda}\left(1-\frac{|\omega|}{\lambda}\right)\mathscr{P}(f)(\omega)e^{i\omega t}\,d\omega. \end{aligned}$$

Using arguments analogous to those in the proof of Theorem 1.51, one can establish the next result.

Theorem 1.62. *Let $f \in L^2(\mathbb{R})$ and let*

$$\sigma_\lambda(f)(t) := \frac{1}{2\pi}\int_{-\lambda}^{\lambda}\left(1-\frac{|\omega|}{\lambda}\right)\mathscr{P}(f)(\omega)e^{i\omega t}\,d\omega.$$

Then $\sigma_\lambda(f)(t) \to f(t)$ as $\lambda \to \infty$, for almost all $t \in \mathbb{R}$.

1.3.3 The Theorem of Paley–Wiener

In this section, we exhibit the connection between Fourier series and the sampling theorem. The link is the theorem of Paley and Wiener.

Theorem 1.63 (*Paley–Wiener*). *Suppose that $f \in L^2(\mathbb{R})$. Then the following are equivalent.*

1. *$\mathscr{P}(f)|_{\mathbb{R}\setminus[-\delta,\delta]} = 0$, a.e., for some $\delta > 0$.*
2. *f can be extended to an entire function $f : \mathbb{C} \to \mathbb{C}$ of exponential type δ, i.e.,*

$$|F(z)| \leq M\exp(\delta|z|), \quad \forall z \in \mathbb{C},$$

and some constant $M > 0$.

Proof. See [166, 209]. □

Definition 1.64. An $L^2(\mathbb{R})$-function satisfying $\mathscr{F}(f)|_{\mathbb{R}\setminus[-\delta,\delta]} = 0$, a.e., is called *band-limited*. The number 2δ is referred to as the *bandwidth* of f. The space of all band-limited functions is called a *Paley–Wiener space* and is denoted by PW^2_δ.

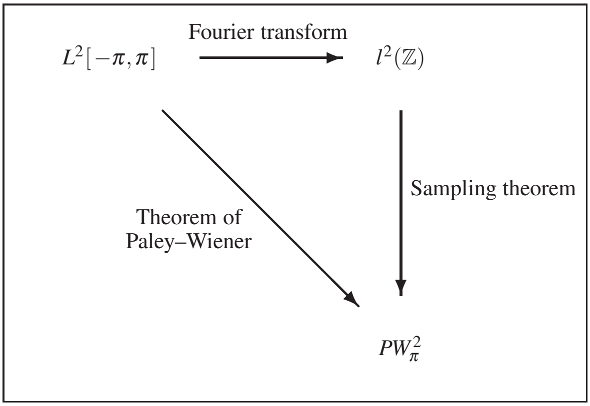


Fig. 1.5: The isometry of the Fourier transform on the torus $\mathbb{T}$ in both the theorem of Paley–Wiener and in the sampling theorem generates a commutative diagram.

1.3.4 Discretization: The Poisson Summation Formula and the Sampling Theorem

Let $T > 0$. Suppose that $f \in L^1(\mathbb{R})$ and that the following two conditions are satisfied:

(P1) The series $\sum_{n=-\infty}^{\infty} f(t+2nT)$ converges everywhere to a continuous function.
(P2) The Fourier series $\sum_{n=-\infty}^{\infty} \mathscr{F}(f)(n/2T)\, e^{int}$ converges everywhere.

Under the above conditions, a relationship between Fourier coefficients and the Fourier transform can be established. The result is the *Poisson summation formula*, which can be stated in the form

$$\sum_{n=-\infty}^{\infty} f(t+2nT) = \frac{1}{2T} \sum_{n=-\infty}^{\infty} \mathscr{F}(f)\left(\frac{n}{2T}\right) e^{int}.$$

A proof of the Poisson summation formula can be found in, e.g., [151].

The next theorem gives conditions on f and its Fourier transform $\mathscr{F}(f)$ under which (P1) and (P2) automatically hold.

Theorem 1.65. *Suppose that f is a Lebesgue-measurable function satisfying*

$$f(t) = \mathscr{O}\left(\frac{1}{1+|t|^\alpha}\right) \quad \textit{and} \quad \mathscr{F}(f)(\omega) = \mathscr{O}\left(\frac{1}{1+|\omega|^\alpha}\right),$$

for some $\alpha > 1$ as $|t| \to \infty$ and $|\omega| \to \infty$. Then conditions (P1) and (P2) hold.

Here, $\mathscr{O}$ denotes the Landau symbol.

Theorem 1.66. *Suppose that $f \in L^2(\mathbb{R}) \cap C(\mathbb{R})$ and that $\mathscr{P}(f) \in L^1(\mathbb{R})$. Then the following estimate holds for every $\delta > 0$:*

$$\sup_{t\in\mathbb{R}} \left| f(t) - \sum_{n=-\infty}^{\infty} f\left(\frac{n}{\delta}\right) \operatorname{sinc}(\delta t - n) \right| \leq \frac{1}{\pi} \int_{|\omega|>\pi\delta} |\mathscr{P}(f)(\omega)| d\omega.$$

Proof. See [162]. □

Considering band-limited functions in Theorem 1.66 immediately yields the Shannon–Whittaker–Kotel'nikov sampling theorem.

Corollary 1.67 (Shannon–Whittaker–Kotel'nikov sampling theorem). *Let $f \in L^2(\mathbb{R}) \cap C(\mathbb{R})$ be a band-limited function with bandwidth $2\pi\delta$, where $\delta > 0$. Then, for every $t \in \mathbb{R}$,*

$$f(t) = \sum_{n=-\infty}^{\infty} f\left(\frac{n}{\delta}\right) \operatorname{sinc}(\delta t - n).$$

Moreover, the series on the right-hand side converges uniformly and absolutely on $\mathbb{R}$.

Defining B to be the smallest positive real number for which $\mathscr{P}(f)|_{\mathbb{R}\setminus[-\pi B,\pi B]} = 0$, a.e., Corollary 1.67 expresses the fact that B is the largest sampling rate allowing the exact reconstruction of f in terms of the samples $f(n/B)$, $n \in \mathbb{Z}$. This value of B is called the *Nyquist rate.*

1.4 Windowed Fourier Transforms

The Fourier, resp. the Plancherel, transform is not stable with respect to local changes in the time or frequency domain. As an example, consider Fig. 1.6, where a small local change of the signal spreads over the whole frequency spectrum. This is due to the fact the analyzing function family consisting of sine and cosine functions is not local, but global.

In order to obtain a better location of a signal in both the time and frequency domains, the ordinary Fourier transform is modified by multiplying the signal f by a window function φ. The present section defines such windowed Fourier transforms and discusses the dependence of the filtered signal on the parameters defining the window function. Particular emphasis is placed on a specific windowed Fourier transform, namely the Gabor transform.

1.4.1 The Short-Time Fourier Transform (STFT)

Definition 1.68. Suppose that $\varphi \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ and $f \in L^2(\mathbb{R})$. For $\omega \in \mathbb{R}$ and $b \in \mathbb{R}$, define

$$\Phi_b(f)(\omega) := \int_{\mathbb{R}} f(t)\varphi(t-b)e^{-i\omega t}\, dt = \langle f, W_{b,\omega}\rangle, \tag{1.14}$$

where $W_{b,\omega}(t) := e^{i\omega t}\overline{\varphi(t-b)}$.

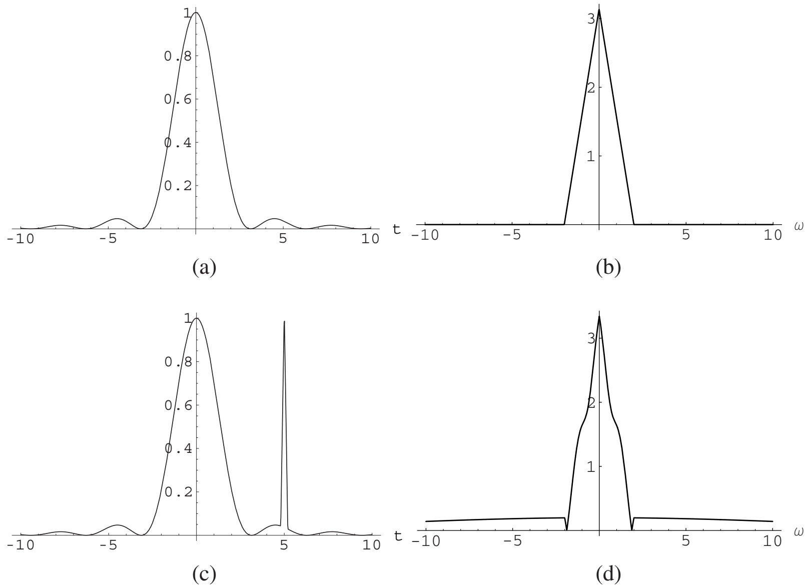


Fig. 1.6: (a) is a function f and (b) the modulus of its frequency spectrum with respect to the Fourier transform in $L^1(\mathbb{R})$. (c): A perturbation was added to the function f, causing a global change in the modulus of its frequency spectrum (d).

Then $\Phi_b(f)(\omega)$ is called the *short-time fourier transform* (STFT) of f.

Hölder's inequality and the Cauchy–Schwarz inequality imply that $f \cdot \varphi(\bullet - b) \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. Thus, the STFT has properties analogous to those of the Fourier and Plancherel transforms.

1.4.2 The Gabor Transform

Consider the Gaussian

$$g_s(t) := \frac{1}{2\sqrt{\pi s}} e^{\frac{-t^2}{4s}} \tag{1.15}$$

with the parameter $s > 0$. Recall that $g_s(t) = G_\lambda(t)$, where $\lambda = 1/2\sqrt{s}$. With this window in the STFT, we get the following special case.

Definition 1.69. Let $f \in L^2(\mathbb{R})$, and let $s > 0$ and $b \in \mathbb{R}$. Define

$$\mathscr{G}_b^s(f)(\omega) := \int_{\mathbb{R}} f(t) g_s(t-b) e^{-i\omega t}\, dt. \tag{1.16}$$

Then $\mathscr{G}_b^s(f)(\omega)$ is called the *Gabor transform* of f with parameters s and b.

The Gabor transform is a so-called time-frequency representation, where the parameter b models the time variable, i.e., the position of the window g_s, and ω the frequency part. The parameter s describes the "width" of the time window.

1.4.2.1 The Gabor Transform as a Fourier and Plancherel Transform—From the View point of the Variable ω

The variable ω appears in (1.16) in the exponential term, similar to the usual Fourier or Plancherel transform. Since the Gabor transform is localized with the integral kernel g_s, the variable ω can be interpreted as a local frequency and $|\mathscr{G}_b^s(f)(\omega)|$ as a local amplitude for a signal in a neighborhood of b. In fact, the variable ω allows for similar properties as the Fourier and the Plancherel transform, as the following proposition shows.

Proposition 1.70. *Let $f \in L^2(\mathbb{R})$. Then*

1. $\mathscr{G}_b^s(f) \in C_u^b(\mathbb{R})$.
2. $\lim_{|\omega|\to\infty} \mathscr{G}_b^s(f)(\omega) = 0$ *(variation of the Riemann–Lebesgue lemma).*
3. $\mathscr{G}_b^s(f) \in L^2(\mathbb{R})$.

Proof. The Hölder and Cauchy–Schwarz inequalities imply $f \cdot g_s(\bullet - b) \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. Thus,

$$\mathscr{G}_b^s(f) = \mathscr{F}(f \cdot g_s(\bullet - b))$$

and statements 1 and 2 follow from the properties of the Fourier transform on $L^1(\mathbb{R})$.

Moreover,

$$\mathscr{G}_b^s(f) = \mathscr{P}(f \cdot g_s(\bullet - b)),$$

together with the fact that the Plancherel transform is an isometry on $L^2(\mathbb{R})$, yields statement 3. □

1.4.2.2 The Gabor Transform from the View point of the Variable b : Window Translation

The family of functions $\{\mathscr{G}_b^s(f)(\omega) : b \in \mathbb{R}\}$ partitions $\mathscr{F}(f)(\omega)$ into a set of Gabor transforms. To see this, we first define the modulation operator

$$\begin{aligned} M_\omega : L^p(\mathbb{R}) &\to L^p(\mathbb{R}), \\ f &\mapsto M_\omega f := e^{i\omega\bullet} f, \end{aligned}$$

for all $\omega \in \mathbb{R}$, and the translation operator

$$\begin{aligned} T_b : L^p(\mathbb{R}) &\to L^p(\mathbb{R}), \\ f &\mapsto T_b f := f(\bullet - b), \end{aligned}$$

for all $b \in \mathbb{R}$, where in both cases $1 \le p \le \infty$.

Theorem 1.71. *Let $f \in L^2(\mathbb{R})$. Then the following hold:*

1. *The mapping $\mathscr{G}_\bullet^s(f)(\omega) : \mathbb{R} \to \mathbb{C}$, $b \mapsto \mathscr{G}_b^s(f)(\omega)$ is continuous.*
2. $\mathscr{G}_\bullet^s(f)(\omega) \in L^2(\mathbb{R})$ *and* $\|\mathscr{G}_\bullet^s(f)(\omega)\|_2 \le \|g_s\|_1 \cdot \|f\|_2$, $\forall s > 0, \forall \omega \in \mathbb{R}$.
3. $\mathscr{P}(\mathscr{G}_\bullet^s(f)(\omega)) = \mathscr{F}(g_s)\mathscr{P}(M_\omega f)$ *is in* $L^2(\mathbb{R})$, *where* $\mathscr{F}(g_s)(\omega) = e^{-s\omega^2}$.

4. If $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, *then* $\mathscr{G}^s_\bullet(f)(\omega) \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ *and*

$$\int_{\mathbb{R}} \mathscr{G}^s_b(f)(\omega)\, db = \mathscr{F}(f)(\omega), \quad \forall \omega \in \mathbb{R}, s > 0.$$

Proof. 1. Follows from the continuity of the function $\mathbb{R} \to L^2(\mathbb{R}), b \mapsto T_b g_s = g_s(\bullet - b)$ and the dominated convergence theorem.
2. $\mathscr{G}^s_b(f)(\omega) = \int_{\mathbb{R}} e^{-i\omega t} f(t) g_s(t-b)\, dt = g_s * (M_{-\omega} f)(b)$, with the convolution $* : L^1(\mathbb{R}) \times L^2(\mathbb{R}) \to L^2(\mathbb{R})$. Thus,

$$\|\mathscr{G}^s_\bullet(f)(\omega)\|_2 \le \|g_s\|_1 \|M_{-\omega} f\|_2 = \|g_s\|_1 \|e^{-i\omega\bullet} f\|_2 = \|g_s\|_1 \|f\|_2.$$

3. $\mathscr{P}(\mathscr{G}^s_\bullet(f)(\omega)) = \mathscr{P}(g_s)\mathscr{P}(M_{-\omega} f) = \mathscr{F}(g_s)\mathscr{P}(M_{-\omega} f)$.
4. $\mathscr{G}^s_\bullet(f)(\omega) = g_s * M_{-\omega} f \in L^1(\mathbb{R})$, since g_s and f are in $L^1(\mathbb{R})$. Hence,

$$\mathscr{F}(\mathscr{G}^s_\bullet(f)(\omega))(\rho) = \mathscr{F}(g_s)(\rho)\mathscr{F}(M_{-\omega} f)(\rho), \quad \forall \rho \in \mathbb{R}.$$

Setting $\rho = 0$ yields the claim. □

1.4.2.3 The Window Parameter *s* and Time-Frequency Localization

The Gabor transform is a special case of the localized STFT. A measure for the width of the window is given by the standard deviation of the "density" g_s^2:

$$\Delta_{g_s} := \frac{1}{\|g_s\|_2} \left(\int_{\mathbb{R}} t^2 (g_s(t))^2\, dt \right)^{\frac{1}{2}}.$$

The quantity Δ_{g_s} is called the *radius* of g_s. It is easy to verify that $\Delta_{g_s} = \sqrt{s}$. (Show this!)

Theorem 1.72. *Let* $f \in L^2(\mathbb{R})$ *and let* $b, \omega \in \mathbb{R}$. *Then*

$$\mathscr{G}^s_b(f)(\omega) = \frac{1}{2\sqrt{\pi s}} e^{-ib\omega} \mathscr{G}^{\frac{1}{4s}}_\omega(\mathscr{P}(f))(-b).$$

Proof. Exercise! □

Remark 1.73. $\mathscr{G}^s_\omega(\mathscr{P}(f))(-b)$ localizes in the frequency domain with radius $1/\sqrt{4s}$, whereas $\mathscr{G}^s_b(f)(\omega)$ localizes in the time domain with radius $\sqrt{s}$.

A measure for the simultaneous localization in the time and frequency domains is given by $\sqrt{s} \cdot 1/\sqrt{4s} = 1/2$.

The time-frequency window is thus $[b - \sqrt{s}, b + \sqrt{s}] \times [\omega - (1/2\sqrt{s}), \omega + (1/2\sqrt{s})]$.

Theorem 1.74. *Let* $f, g \in L^2(\mathbb{R})$. *Then*

$$\int_{\mathbb{R}} \int_{\mathbb{R}} \mathscr{G}^s_b(f)(\omega) \overline{\mathscr{G}^s_b(g)(\omega)}\, db\, d\omega = \frac{\sqrt{\pi}}{\sqrt{2s}} \langle f, g \rangle.$$

Proof. Exercise! □

Theorem 1.75 (Inversion theorem). *Assume that* $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. *At all points* $t \in \mathbb{R}$ *at which* f *is continuous, the following inversion formula holds:*

$$f(t) = \sqrt{\frac{2s}{\pi}} \int_{\mathbb{R}} \int_{\mathbb{R}} \mathscr{G}_b^s(f)(\omega) g_s(t-b) e^{i\omega t}\, db\, d\omega.$$

Proof. Let $\{F_\lambda\}_{\lambda \in (0,\infty)}$ be the Fejér kernel on $\mathbb{R}$. Then $F_\lambda \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. Approximating f with the Fejér kernel yields

$$\langle f, T_a F_\lambda \rangle = \int_{\mathbb{R}} f(t) F_\lambda(t-a)\, dt = \int_{\mathbb{R}} f(t) F_\lambda(a-t)\, dt = f * F_\lambda(a).$$

This equation holds almost everywhere and pointwise at all points of continuity of f. At these points, we also have pointwise convergence:

$$\lim_{\lambda \to \infty} \langle f, T_a F_\lambda \rangle = f(a). \tag{1.17}$$

Moreover,

$$\begin{aligned}
\mathscr{G}_b^s(T_a F_\lambda)(\omega) &= \int_{\mathbb{R}} F_\lambda(t-a)\, g_s(t-b)\, e^{-i\omega t}\, dt \\
&= e^{-i\omega b} \int_{\mathbb{R}} F_\lambda(t-a)\, g_s(t-b)\, e^{-i\omega(t-b)}\, dt \\
&= e^{-i\omega b} \int_{\mathbb{R}} F_\lambda(y+b-a)\, g_s(y)\, e^{-i\omega y}\, dy \\
&= e^{-i\omega b} \int_{\mathbb{R}} F_\lambda(a-b-y)\, g_s(y)\, e^{-i\omega y}\, dy \\
&= e^{-i\omega b} (F_\lambda * M_{-\omega} g_s)(a-b) \\
&\to e^{-i\omega b} M_{-\omega} g_s(a-b) = g_s(a-b)\, e^{-i\omega a} \quad \text{as } \lambda \to \infty.
\end{aligned} \tag{1.18}$$

Now apply Theorem 1.72 to f and $F_\lambda(\bullet - a)$ and use the dominated convergence theorem to obtain

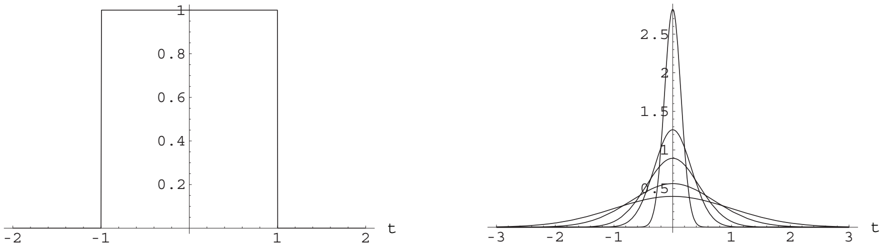
$$\lim_{\lambda \to 0} \int_{\mathbb{R}} \int_{\mathbb{R}} \mathscr{G}_b^s(f)(\omega) \overline{\mathscr{G}_b^s(T_a F_\lambda)(\omega)}\, db\, d\omega = \int_{\mathbb{R}} \int_{\mathbb{R}} \mathscr{G}_b^s(f)(\omega)\, g_s(a-b)\, e^{i\omega a}\, db\, d\omega.$$

On the other hand, by Theorem 1.74 and Eq. (1.17),

$$\begin{aligned}
&\int_{\mathbb{R}} \int_{\mathbb{R}} \mathscr{G}_b^s(f)(\omega) \overline{\mathscr{G}_b^s(T_a F_\lambda)(\omega)}\, db\, d\omega \\
&= \sqrt{\frac{\pi}{2s}} \langle f, T_a F_\lambda \rangle \to \sqrt{\frac{\pi}{2s}} f(a), \quad \text{for } \lambda \to \infty,
\end{aligned}$$

at all points a, where f is continuous. This gives the result. □

For an illustration of the dependence of the Gabor transform on the parameter s, see Figs. 1.7 and 1.8.



(a) (b)

Fig. 1.7: (a) The function f and (b) the Gauß kernels for different values of the window parameter s. Figure 1.8 depicts the corresponding Gabor transforms.

1.4.3 The Heisenberg Uncertainty Principle

In the context of localizating the Fourier transform, immediately the following question arises: Is it possible to choose a window function φ for the STFT whose energy is well localized in both time and frequency? Unfortunately, it is impossible to tell which frequencies are present at a specific point in time. This is the content of the so-called uncertainty principle, which states that *a function f and its Plancherel transform $\mathscr{P}(f)$ cannot both have arbitrarily small support.*

Definition 1.76. Assume that $\varphi \in L^2(\mathbb{R})$ and that $\sqrt{|\bullet|} \cdot |\varphi| \in L^2(\mathbb{R})$.

1. $a^* := \frac{1}{\|\varphi\|_2^2} \int_{\mathbb{R}} t \cdot |\varphi(t)|^2 \, dt$ is called the *center* of φ.
2. $\Delta_\varphi := \frac{1}{\|\varphi\|_2} \left(\int_{\mathbb{R}} (t-a^*)^2 |\varphi(t)|^2 \, dt \right)^{1/2}$ is called the *radius* of φ, occasionally also the *mean bandwidth* or the *mean running time*.

Theorem 1.77 (The Heisenberg uncertainty relation). *Let $f \in L^2(\mathbb{R})$. Then*

$$\frac{1}{2\pi} \int_{\mathbb{R}} t^2 \, |f(t)|^2 \, dt \cdot \int_{\mathbb{R}} \omega^2 \, |\mathscr{P}(f)(\omega)|^2 \, d\omega \geq \frac{1}{4} \|f\|_2^4. \tag{1.19}$$

(We allow the left-hand side to assume the value "∞.") The left-hand side equals the right-hand side iff $f(t) = c \cdot e^{-kt^2}$ for $k > 0$ and $c \in \mathbb{C}$.

Proof. We refer to [46] for the proof. □

The Heisenberg uncertainty relation was first proved in this form but under stronger assumptions by H. Weyl.

For the proof of Theorem 1.77, one requires the concept of Schwartz space. The Schwartz space $\mathscr{S}$ consists of all functions $f \in C^\infty(\mathbb{R})$ that satisfy the condition

$$\sup_{|\alpha| \leq N} \sup_{t \in \mathbb{R}} (1+t^2)^N |D^\alpha(f)(t)| < \infty \tag{1.20}$$

for every $N \in \mathbb{N}_0$.

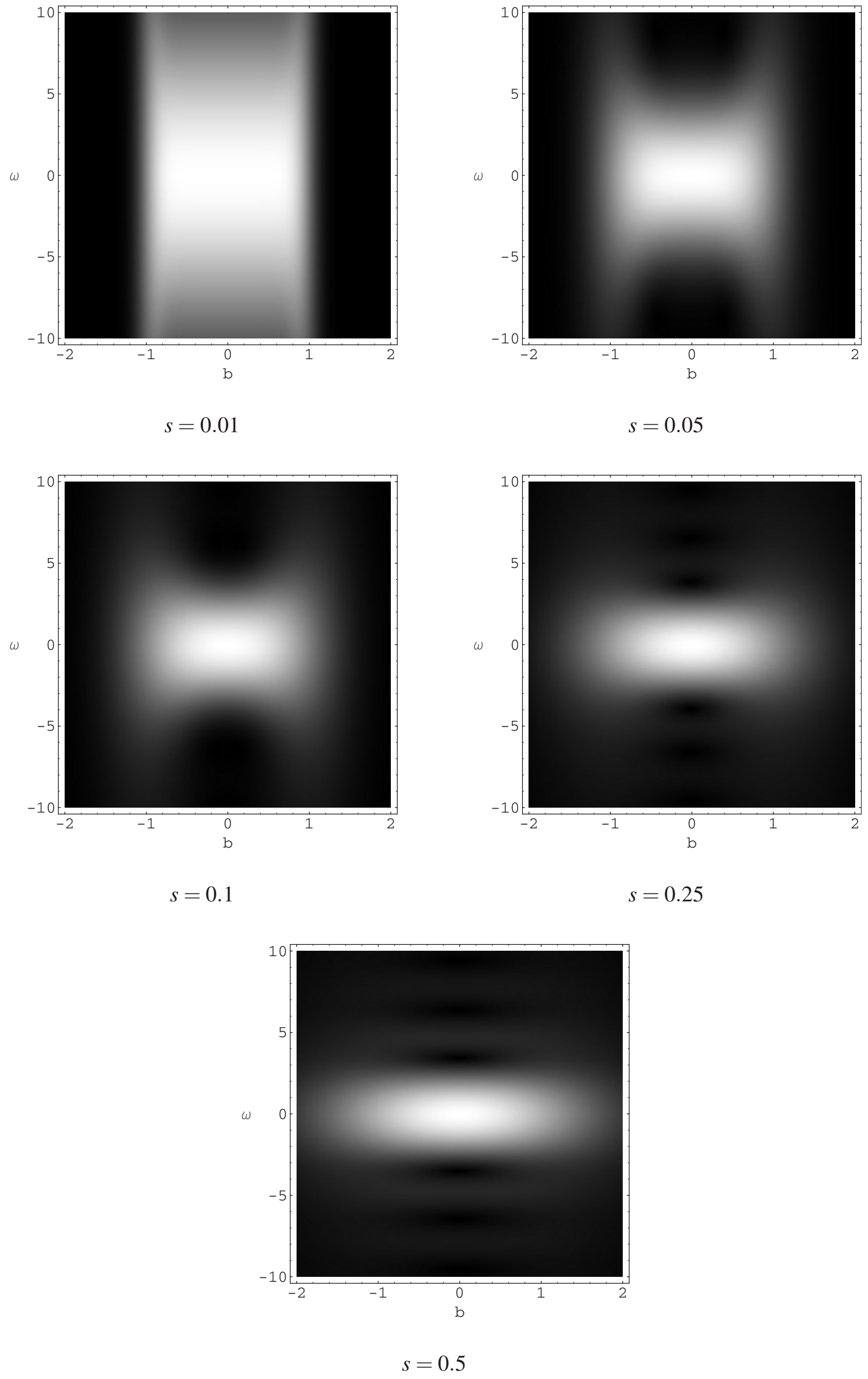


Fig. 1.8: Gabor transform for the functions in Fig. 1.7 for different values of the window parameter s. For small s, e.g., $s = 0.01$, the window is well localized in the time domain, and the edges of the function f appear nicely (horizontal axis), whereas the Fourier spectrum (vertical axis) is blurred. For large s, e.g., $s = 0.5$, on the contrary, the sinc character of $\mathscr{F}(f)$ is well recovered, but the time-domain representation of f is blurred.

In other words, $p \cdot D^\alpha f$ is a bounded function on $\mathbb{R}$ for every polynomial p and every differential operator D^α of order $\alpha \in N_0$. Since this is also true for $(1+t^2)^N p(t)$ instead of $p(t)$, it follows that every $p \cdot D^\alpha f \in L^1(\mathbb{R})$. For,

$$|p(t) \cdot D^\alpha f(t)| \leq \text{const.} \frac{1}{(1+t^2)^N}.$$

Therefore, functions in $\mathscr{S}$ decay faster than any power $1/|t|^m$ as $|t| \to \infty$.

Functions with this property form a vector space $\mathscr{S}$ for which the countably many norms (1.20) generate a locally convex topology.

Theorem 1.78. *1. $\mathscr{S}$ is a Fréchet space, i.e., a complete, locally convex, metrizable space.*

2. Suppose that p is a polynomial, $g \in \mathscr{S}$, and $\alpha \in \mathbb{N}$. Then each of the following three maps is continuous and linear from $\mathscr{S} \to \mathscr{S}$:

$$f \mapsto p \cdot f, \quad f \mapsto g \cdot f, \quad f \mapsto D^\alpha f.$$

3. The Fourier transform is a continuous linear mapping from $\mathscr{S} \to \mathscr{S}$. In particular, $\mathscr{F}(f) \in \mathscr{S}$ for $f \in \mathscr{S}$.

4. $\mathscr{S}$ is dense in $L^2(\mathbb{R})$.

Proof. See, for instance, [209]. □

1.4.4 Discretization: Gabor Frames

Up to now, the Gabor transform has been an integral transform operator with continuous parameters and variables s, b, and ω. This family of parameters yields a highly redundant system; therefore, the question arises if there exists a discretized representation in the form of a series instead of an integral. Various aspects of such a discretization are considered in Chapters 2, 3, and 5. We refer to these chapters for the ideas and details. It is, however, worth noticing that a discretization to a Gabor Riesz basis with fast decaying basis elements in the time domain *as well as* in the frequency domain is not possible. This is the famous Balian–Low theorem, which is stated in Chapter 2 (Theorem 2.27).

1.4.5 Shortcomings of the Windowed Fourier Transform

It follows from the definition (1.14) of the windowed Fourier transform that at each point in the time-frequency domain, a window is translated to the time location and frequency location under consideration. The duration and the bandwidth of the window and thus the resolution do not change. The resolution in the time and frequency domains depends only on the form of the window and, by the Heisenberg uncertainty relation, it is not possible to achieve an optimal resolution simultaneously in the time and frequency domains. For a signal that contains both low- and high-frequency components, however, it would be desirable to obtain a good resolution in the low-frequency spectrum since there small changes are relevant, whereas in the high-frequency spectrum, a good time resolution is more important since a complete oscillation requires less time, causing a faster change of the instantaneous frequency. Thus, at low frequencies one needs a good frequency resolution, taking into account a bad time resolution, whereas at high frequencies a good time resolution but a bad frequency resolution is desired. The STFT does not provide a means of achieving this.

A generalization of the STFT is provided by the wavelet transform. Instead of comparing a signal with a window function that is translated and modulated, the wavelet transform compares a signal with a window function that is translated and scaled. The scaling induces, analogous to modulation, a frequency shift. However, a frequency increase causes a simultaneous reduction of the time duration. At high frequencies this results in a better time resolution and at low frequencies in a better frequency resolution but a worse time resolution.

1.5 The Wavelet Transform

The idea of the wavelet transform consists of comparing the analyzed signal or image with one single pattern, the wavelet. The pattern is dilated and translated, such that its shape works like a looking glass, which is moved (translated) over a signal or an image in various distances (dilations) from the signal or image. The wavelet is a function that is well localized in time or space as well as in the frequency domain. Therefore, in contrast to the Fourier transform, it allows for a local analysis.

1.5.1 Definition and Properties

Definition 1.79. A function $\psi \in L^2(\mathbb{R})$ is called a *wavelet* if it fulfills the *admissibility condition*:

$$0 < c_\psi = \int_{\mathbb{R}} \frac{|\mathscr{P}(\psi)(\omega)|^2}{|\omega|} \, d\omega < \infty.$$

The *wavelet transform* of a function $f \in L^2(\mathbb{R})$ with respect to the wavelet ψ is defined as

$$\mathscr{W}_\psi f(a,b) = \frac{1}{\sqrt{c_\psi}} \frac{1}{\sqrt{|a|}} \int_{\mathbb{R}} f(t)\, \psi\left(\frac{t-b}{a}\right) dt,$$

for all $a \in \mathbb{R} \setminus \{0\}$ and all $b \in \mathbb{R}$.

Let $\psi \in L^1(\mathbb{R})$ be a wavelet. Then $\mathscr{P}(\psi) = \mathscr{F}(\psi)$, and $\mathscr{F}(\psi)$ is continuous. The admissibility condition therefore implies $\mathscr{F}(\psi)(0) = 0$, which is equivalent to ψ having zero mean:

$$0 = \mathscr{F}(\psi)(0) = \int_{\mathbb{R}} \psi(t)\, dt.$$

This is a necessary condition for wavelets in $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$.

Example 1.80. 1. The generator of the Haar system from Section 1.1.2.3 is a wavelet, the so-called Haar wavelet:

$$\psi(t) = \begin{cases} 1, & \text{for } 0 \le t < \frac{1}{2}, \\ -1, & \text{for } \frac{1}{2} \le t \le 1, \\ 0, & \text{else.} \end{cases}$$

For the Fourier transform, $\mathscr{F}(\psi)(\omega) = ie^{-i\frac{\omega}{2}} \sin(\omega/4)(\sin\frac{\omega}{4}/\frac{\omega}{4})$. The admissibility constant is $c_\psi = 2\ln 2$.

2. A C^∞-example for a wavelet is the *Mexican hat wavelet*:

$$\psi(x) = -\frac{d^2}{dx^2} e^{-x^2/2} = (1-x^2)e^{-x^2/2}.$$

The Fourier transform has the form

$$\mathscr{F}(\psi)(\omega) = \omega^2 e^{-\omega^2/2};$$

the admissibility constant $c_\psi = 1$.

Lemma 1.81. *The set of wavelets* $\{\psi \in L^2(\mathbb{R}) \mid \psi$ *is admissible*$\}$ *is dense in* $L^2(\mathbb{R})$.

Proof. Exercise! □

The wavelet transform operates as a linear time-invariant filter. To see this, we consider the Plancherel transform with respect to the second variable:

$$\begin{aligned} \mathscr{P}(\mathscr{W}_\psi f(a,\bullet))(\omega) &= \sqrt{|a|}\frac{1}{\sqrt{c_\psi}} \mathscr{P}(\psi)(-a\omega)\mathscr{P}(f)(\omega) \\ &= \sqrt{|a|}\frac{1}{\sqrt{c_\psi}} \overline{\mathscr{P}(\overline{\psi})(a\omega)}\mathscr{P}(f)(\omega) \quad \text{in } L^2(\mathbb{R}). \end{aligned}$$

From this we can deduce that the wavelet transform is an isometry:

Theorem 1.82. *The wavelet transform corresponding to the wavelet* ψ *is an isometry:*

$$\mathscr{W}_\psi : L^2(\mathbb{R}) \to L^2\left(\mathbb{R}^2, \frac{da\,db}{a^2}\right).$$

Proof. By definition of the wavelet transform, for $f \in L^2(\mathbb{R})$,

$$\begin{aligned} \mathscr{W}_\psi f(a,b) &= \frac{1}{\sqrt{c_\psi}} \frac{1}{\sqrt{|a|}} \int_{\mathbb{R}} f(t)\psi\left(\frac{t-b}{a}\right) dt \\ &= \frac{1}{\sqrt{c_\psi}} \frac{1}{\sqrt{|a|}} \left\langle f, \overline{\psi\left(\frac{\bullet - b}{a}\right)} \right\rangle. \end{aligned}$$

$\mathscr{W}_\psi f(a,b)$ is well defined, because $\psi \in L^2(\mathbb{R})$ and therefore also $\psi(\bullet - b/a) \in L^2(\mathbb{R})$. By the Parseval inequality,

$$\|\mathscr{W}_\psi f\|^2_{L^2\left(\mathbb{R}^2, \frac{da\,db}{a^2}\right)} = \int_{\mathbb{R}} \int_{\mathbb{R}} |\mathscr{W}_\psi f(a,b)|^2 \frac{da\,db}{a^2}$$

$$
\begin{aligned}
&= \frac{1}{2\pi}\int_{\mathbb{R}}\int_{\mathbb{R}} |\mathscr{P}(\mathscr{W}_\psi f(a,\bullet))(\omega)|^2 \frac{da\,d\omega}{a^2}\\
&= \frac{1}{2\pi}\int_{\mathbb{R}}\int_{\mathbb{R}} \left|\sqrt{|a|}\frac{1}{\sqrt{c_\psi}}\mathscr{P}(\psi)(-a\omega)\mathscr{P}(f)(\omega)\right|^2 \frac{da\,d\omega}{a^2}\\
&= \frac{1}{2\pi}\frac{1}{c_\psi}\int_{\mathbb{R}}\int_{\mathbb{R}} |a||\mathscr{P}(\psi)(-a\omega)|^2|\mathscr{P}(f)(\omega)|^2 \frac{da\,d\omega}{a^2}\\
&= \frac{1}{2\pi}\frac{1}{c_\psi}\int_{\mathbb{R}}\int_{\mathbb{R}} \frac{|\mathscr{P}(\psi)(s)|^2}{|s|}|\mathscr{P}(f)(\omega)|^2\,ds\,d\omega\\
&= \frac{1}{2\pi}\|\mathscr{P}(f)\|^2_{L^2(\mathbb{R})} = \|f\|^2_{L^2(\mathbb{R})}.
\end{aligned}
$$

Theorem 1.83 (Inverse wavelet transform). *The adjoint operator*

$$
\begin{aligned}
\mathscr{W}_\psi^* : L^2\left(\mathbb{R}^2, \frac{da\,db}{a^2}\right) &\to L^2(\mathbb{R}),\\
g &\mapsto \frac{1}{\sqrt{c_\psi}}\int_{\mathbb{R}}\int_{\mathbb{R}} \frac{1}{\sqrt{|a|}}\psi\left(\frac{t-b}{a}\right) g(a,b)\frac{da\,db}{a^2} \qquad (1.21)
\end{aligned}
$$

is the inverse operator of the wavelet transform on the image $\mathscr{W}_\psi(L^2(\mathbb{R}))$.

In fact, for all $f \in L^2(\mathbb{R})$ and all $g \in L^2\left(\mathbb{R}^2, da\,db/a^2\right)$,

$$
\begin{aligned}
\langle \mathscr{W}_\psi f, g\rangle_{L^2\left(\mathbb{R}^2,\frac{da\,db}{a^2}\right)} &= \int_{\mathbb{R}}\int_{\mathbb{R}} \mathscr{W}_\psi f(a,b) g(a,b)\frac{da\,db}{a^2}\\
&= \int_{\mathbb{R}}\int_{\mathbb{R}} \frac{1}{\sqrt{c_\psi}}\int_{\mathbb{R}} f(t)\frac{1}{\sqrt{|a|}}\psi\left(\frac{t-b}{a}\right) g(a,b)\,dt\frac{da\,db}{a^2}\\
&= \int_{\mathbb{R}} f(t)\frac{1}{\frac{1}{\sqrt{c_\psi}}}\int_{\mathbb{R}}\int_{\mathbb{R}} \frac{1}{\sqrt{|a|}}\psi\left(\frac{t-b}{a}\right) g(a,b)\,dt\frac{da\,db}{a^2}\,dt\\
&= \langle f, \mathscr{W}_\psi^* g\rangle.
\end{aligned}
$$

The parameter b shifts the pattern, the wavelet ψ, over the signal. The parameter a scales the wavelet and therefore adapts its shape. If the scaled wavelet at a certain place b has a similar local shape as the analyzed signal, then the wavelet coefficient has a large absolute value. Small $|a|$ describe small details, while larger $|a|$ generate approximations of the analyzed function in a neighborhood of the point b. In fact, the wavelet transform operates as a filter:

$$
\mathscr{W}_\psi f(a,b) = \frac{1}{\sqrt{c_\psi}}\frac{1}{\sqrt{|a|}}\int_{\mathbb{R}} f(t)\psi\left(\frac{t-b}{a}\right)dt = \frac{1}{\sqrt{c_\psi}}(D_{-a}\psi * f)(b).
$$

For fixed a, this corresponds to filtering with $\psi(\bullet/a)$ at the point b. $\psi \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ is a bandpass filter, because due to the admissibility condition, $\mathscr{F}(\psi)(0) = 0$ and by the Riemann–Lebesgue lemma, $\lim_{|\omega|\to\infty}\mathscr{F}(\psi)(\omega) = 0$.

1.5.2 Scale Discretization—The Dyadic Wavelet Transform

To discretize the wavelet transform $\mathscr{W}_\psi$, we will have to discretize both parameters a and b. To this end, we consider the affine operators D_a and T_b for dilations and translations, respectively. Obviously, $\mathscr{W}_\psi f(a,b) = (1/\sqrt{c_\psi})\langle f, \overline{T_b D_a \psi}\rangle_{L^2(\mathbb{R})}$. Therefore, the question arises, if there exists a discrete set of wavelets $\{T_b D_a \psi \mid (a,b)$ in a discrete set $I\}$ such that the inversion formula still holds; i.e., no information about the analyzed function f is lost by considering the discrete set of coefficients $\{\mathscr{W}_\psi f(a,b) \mid (a,b) \in I\}$.

The wavelet transform is translation-invariant. Let $\tau \in \mathbb{R}$. Then

$$\mathscr{W}_\psi(T_\tau f)(a,b) = \mathscr{W}_\psi(f(\bullet - \tau))(a,b) = \mathscr{W}_\psi(f)(a,b-\tau).$$

The idea of this section is to discretize the scale parameter a while keeping the translation invariance of the wavelet transform.

Definition 1.84. Let $f \in L^2(\mathbb{R})$. Its dyadic wavelet transform is defined as

$$\mathscr{W}_\psi f(2^j,b) = \frac{1}{\sqrt{c_\psi}} \frac{1}{\sqrt{2^j}} \int_{\mathbb{R}} f(t)\psi\left(\frac{t-b}{2^j}\right) dt, \quad j \in \mathbb{Z}.$$

Theorem 1.85. *Suppose there exist two positive constants A and B such that*

$$A \le \sum_{j\in\mathbb{Z}} |\mathscr{F}(\psi)(2^j\omega)|^2 \le B, \quad \forall \omega \in \mathbb{R}.$$

Then

$$A\|f\|^2 \le \sum_{j\in\mathbb{Z}} \frac{c_\psi}{2^j} \|\mathscr{W}_\psi f(2^j,\bullet)\|^2 \le B\|f\|^2$$

with respect to the $L^2(\mathbb{R})$*-norm.*

Proof. Consider the Plancherel transform

$$\mathscr{P}(\mathscr{W}_\psi f(2^j,\bullet))(\omega) = \sqrt{2^j}\frac{1}{\sqrt{c_\psi}} \overline{\mathscr{P}(\overline{\psi})(2^j\omega)}\mathscr{P}(f)(\omega).$$

Summation together with the assumption gives

$$A|\mathscr{P}(f)(\omega)|^2 \le \sum_{j\in\mathbb{Z}} \frac{c_\psi}{2^j} |\mathscr{P}(\mathscr{W}_\psi f(2^j,\bullet))(\omega)|^2 \le B|\mathscr{P}(f)(\omega)|^2,$$

for all $\omega \in \mathbb{R}$. Integration, the theorem of dominated convergence, and the Parseval equation yield the claim. □

The theorem shows that the normalized dyadic wavelet transform

$$\frac{\sqrt{c_\psi}}{\sqrt{2^j}} \mathscr{W}_\psi f(2^j,b) = \left\langle f, \overline{\frac{1}{2^j}\psi\left(\frac{\bullet - b}{2^j}\right)}\right\rangle$$

has the same properties as a frame.

1.5.3 Multiresolution Analyses

We have seen early above that the discretization of the scale parameter a yields the dyadic wavelet transform that behaves similarly to a frame. In this section, we discretize the translation parameter b. Our aim is to obtain a wavelet basis of the form

$$\left\{ \psi_{j,n}(t) = \frac{\sqrt{c_\psi}}{\sqrt{2^j}} \psi\left(\frac{t-2^j n}{2^j}\right) \right\}_{(j,n)\in\mathbb{Z}^2}.$$

A classical approach to this end is multiresolution analysis.

Definition 1.86. A sequence of nonempty closed subspaces $\{V_j\}_{j\in\mathbb{Z}}$ of $L^2(\mathbb{R})$ is called a *multiresolution analysis* if all of the following properties are satisfied:

1. For all $j \in \mathbb{Z}$, $V_{j+1} \subset V_j$; i.e., the spaces are nested.
2. The spaces are translation-invariant: For all $(j,k) \in \mathbb{Z}^2$,
$$f \in V_j \iff f(\bullet - 2^j k) \in V_j.$$
3. Scaling or refinement relation: For all $j \in \mathbb{Z}$,
$$f \in V_j \iff f(\bullet/2) \in V_{j+1}.$$
4. The subspaces span $L^2(\mathbb{R})$ and separate the space:
$$\lim_{j\to-\infty} V_j = \mathrm{cl}\left(\bigcup_{j\in\mathbb{Z}} V_j\right) = L^2(\mathbb{R});$$
$$\lim_{j\to\infty} V_j = \bigcap_{j\in\mathbb{Z}} V_j = \{0\}.$$
5. There is a so-called scaling function $\varphi \in L^2(\mathbb{R})$, such that the family
$$\{\varphi(\bullet - n)\}_{n\in\mathbb{Z}}$$
forms a Riesz basis of V_0.

Because of the scaling relation, multiresolution analyses of this form with dilation 2 are sometimes called *dyadic multiresolution analyses*. Since the spaces $\{V_j\}_{j\in\mathbb{Z}}$ are nested, the approximation at scale 2^{-j} contains all information of the coarser scale 2^{-j-1}.

There is practical criterion to check, whether a function fulfills the Riesz property 5.

Proposition 1.87. *The following are equivalent:*

1. *The family $\{\varphi(\bullet - n)\}_{n\in\mathbb{Z}}$ is a Riesz basis of $V_0 = cl\{span\{\varphi(\bullet - n)\}_{n\in\mathbb{Z}}\}$.*
2. *There are constants $A, B > 0$ such that for all $\omega \in [-\pi, \pi]$,*
$$\frac{1}{B} \le \sum_{k\in\mathbb{Z}} |\mathscr{F}(\varphi)(\omega - 2k\pi)|^2 \le \frac{1}{A}.$$

Proof. E.g., [244, Prop. 2.8].

Example 1.88. 1. The piecewise-constant functions generate multiresolution analyses via the nested spaces

$$V_j = \{g \in L^2(\mathbb{R}) \mid g|_{[2^j n, 2^j(n+1))} = \text{const. a.e.}, n \in \mathbb{Z}\}.$$

The scaling function is $\varphi = \chi_{[0,1)}$.

2. In an analogous manner, the piecewise polynomial functions generate multiresolution analyses [244].

Now the questions arises of how the wavelets come into play. In fact, they form a Riesz basis for the complementary spaces W_{j+1} of V_{j+1} in V_j:

$$V_{j+1} \oplus W_{j+1} = V_j, \quad j \in \mathbb{Z}.$$

Properties 1 and 4 of Definition 1.86 give the decomposition

$$L^2(\mathbb{R}) = \bigoplus_{j \in \mathbb{Z}} W_j.$$

A wavelet can be generated from a scaling function in various ways. A common approach is to consider

$$\psi(t) = \sum_{n \in \mathbb{Z}} (-1)^n \overline{a_n} \varphi(2t + n + 1),$$

where $a_n = \int_{\mathbb{R}} \varphi(\frac{t}{2}) \overline{\varphi(t-n)}\, dt$. Thus, the wavelet can be calculated directly from the scaling function.

The ideas behind this construction, as well as many other possibilities, are discussed, e.g., in [60, 170, 175, 244] and many other books on wavelets.

There are many families of scaling function/wavelet pairs, e.g., B-splines are among them. The Haar wavelet together with its scaling function $\varphi = \chi_{[0,1)}$ belongs to this class. A good introduction on spline wavelets and others and their approximation properties can be found in [226, 244].

Remark 1.89. There exist wavelet bases of $L^2(\mathbb{R})$ that are not associated with any multiresolution analysis. An example are so-called unimodular wavelets, whose Fourier transforms are characteristic functions of certain sets. Such sets are called *wavelet sets*. More about unimodular wavelets and wavelet sets can be found in the list of references in [244, Sect. 3.4]. For interesting connections between composite wavelets as discussed in Chapter 3, wavelet sets, and reflection groups, we refer the reader to [157–159]

Mallat made the important discovery that there is a fast algorithm for the wavelet transform. In his article [174], his proposed pyramidal algorithm based on convolutions brought the breakthrough for the applicability of the wavelet transform theory.

1.6 Other Multiscale Transforms

In this section, we briefly consider the problem of extending wavelets to higher dimensions. Extensions via the tensor product are always possible, but from a modeling viewpoint are not very desirable; images and data would have to be assumed to be homogeneous and to possess a lattice structure. To overcome these impediments, several new types of wavelets have been introduced, some with very specific applications in mind.

After introducing the tensor product approach for the construction of wavelet bases in $\mathbb{R}^2$, a simple case that nevertheless reflects the main ideas, we list some examples of multiscale transforms and give references to the literature.

1.6.1 Tensor Product Wavelets in 2D

To employ wavelets for image analyses in 2D and 3D, one can consider the tensor product of 1D wavelets. To this end, suppose that ϕ is a given (one-dimensional) scaling function for an MRA on $L^2(\mathbb{R})$ and ψ the associated wavelet. Define

$$\boldsymbol{\Phi}_{j_1k_1;j_2k_2}(x,y) = (\phi_{j_1k_1} \otimes \phi_{j_2k_2})(x,y) = \phi_{j_1k_1}(x)\phi_{j_2k_2}(y),$$

where $\phi_{jk} := 2^{j/2}\phi(2^j \bullet -k),\ j,k \in \mathbb{Z}$. It can be shown that the scaling functions $\{\boldsymbol{\Phi}_{jk_1;jk_2} \,|\, j \in \mathbb{Z};\, (k_1,k_2) \in \mathbb{Z}\times\mathbb{Z}\}$ are basis functions for approximation spaces $\mathfrak{V}_j \subset L^2(\mathbb{R}^2)$ by setting

$$\mathfrak{V}_j = V_j \otimes V_j, \qquad j \in \mathbb{Z},$$

where $\otimes$ denotes the tensor product of the vector spaces V_j. It should be clear that if the spaces V_j form an MRA of $L^2(\mathbb{R})$, then the spaces $\mathfrak{V}_j$ form an MRA of $L^2(\mathbb{R}^2)$. Now

$$\begin{aligned}\mathfrak{V}_{j+1} &= V_{j+1} \otimes V_{j+1} = (V_j \oplus W_j) \otimes (V_j \oplus w_j)\\ &= V_j \otimes V_j \oplus (V_j \otimes W_j \oplus W_j \otimes V_j \oplus W_j \otimes W_j)\\ &= \mathfrak{V}_j \oplus \mathfrak{W}_j,\end{aligned}$$

where we set

$$\mathfrak{W}_j = \underbrace{(V_j \otimes W_j)}_{\text{horizontal}} \oplus \underbrace{(W_j \otimes V_j)}_{\text{vertical}} \oplus \underbrace{(W_j \otimes W_j)}_{\text{diagonal}}.$$

Hence, in the two-dimensional setting, there are *three* wavelets:

$$\begin{aligned}\boldsymbol{\Psi}^h(x,y) &= (\phi \otimes \psi)(x,y) = \phi(x)\psi(y), &&\text{(horizontal)}\\ \boldsymbol{\Psi}^v(x,y) &= (\psi \otimes \phi)(x,y) = \psi(x)\phi(y), &&\text{(vertical)}\\ \boldsymbol{\Psi}^d(x,y) &= (\psi \otimes \psi)(x,y) = \psi(x)\psi(y), &&\text{(diagonal)}\end{aligned}$$

if x is taken along the horizontal and y the vertical direction. Moreover,

$$\left\{\Psi^{\lambda}_{jk_1;jk_2}\,\middle|\,(k_1,k_2)\in\mathbb{Z}^2;\ \lambda=h,v,\text{ or } d\right\}$$

is a Riesz basis for $\mathfrak{W}_j$, and function $f\in L^2(\mathbb{R}^2)$ can be written as

$$f(x,y)=\sum_{j\in\mathbb{Z}}\sum_{k_1,k_2\in\mathbb{Z}}\langle c^j_{k_1k_2},\Psi_{jk_1;jk_2}\rangle,$$

where $\Psi=(\Psi^h_{jk_1;jk_2},\Psi^v_{jk_1;jk_2},\Psi^d_{jk_1;jk_2})^T$ and the $c^j_{k_1k_2}=c^j_{k_1k_2}(f)$ are vector coefficients.

A similar construction in 3D yields eight wavelets oriented along the face and space diagonals of the unit 3D cube.

Remark 1.90. To reduce the number of wavelets, dilation matrices A other than the dyadic $2I$ can also be used. The number of wavelets depends on the number of cosets, which is specified by the determinant of A. For example, if in 2D $A=2I$, then $|\det A|=|\det 2I|=4$ and there are $|\det A|-1=3$ wavelets, as we saw above. The matrix $\begin{pmatrix}1&1\\-1&1\end{pmatrix}$ generates the so-called quincunx grid in 2D, and every multiresolution analysis based on this dilation matrix needs only one wavelet to generate the orthogonal complements W_j. For more interesting facts on the quincunx lattice, its dilation matrices in 2D, and extensions to higher dimensions, we refer to [153, 170, 227] and the references therein.

1.6.2 Some Wavelet-Type Transforms

Although 1D wavelets have good resolution in both the time and frequency domains they lack the ability to resolve signals along arbitrary directions in 2D and 3D. In addition, a large number of wavelet coefficients are required to account for edges, i.e., for singularities along lines or curves. In order to retain the multiscale structure of a signal decomposition, the idea of wavelet transform had to be extended to incorporate the resolution of singularities along lines or curves.

Among these extensions are the following transforms, which are briefly described for image decompositions and analysis.

- **Ridgelet transform:** The idea is to choose basis functions that are constant along lines, i.e., ridges, and that transverse to these ridges are wavelets in the regular sense. Details of this construction can be found in, for instance, [39, 67].
- **Curvelet transform:** Here an image is analyzed using different block sizes but employing only a single transform. The image is first decomposed into a set of wavelet bands and then each band is analyzed using a ridgelet transform. The block size can be changed at each scale level. As references, we mention [40–42]

- **Beamlets:** Unlike wavelets which offer a localized scale and position representation near fixed regions in an image with a specified scale and location, beamlets offer a localized scale, position, and orientation based on dyadically organized line segments. For more details, the reader is referred to [71].
- **Wedgelets:** Wedgelet approximations were introduced in [70] as a means to efficiently approximate piecewise-constant images. Generally speaking, wedgelet approximations are obtained by adaptively partitioning the image domain into disjoint sets and by computing an approximation of the image on each of these sets. Optimal approximations are defined using a certain functional that weighs the approximation error against the complexity of the decomposition. As a reference for an application, see [99].
- **Platelets:** The image partition is based on recursive, dyadic squares allowing wedge-shaped final nodes (instead of squares). Like wedgelets, platelets approximate with piecewise-constant functions. They are suited for the approximation of images consisting of smooth regions separated by smooth contours. For more details and an application, see [243].
- **Framelets:** Here, the idea of redundent representations and frames is employed to construct redundant wavelet systems. The interested reader is referred to [61] for a construction and more details.
- **Shearlets:** Shearlets are an affine system with a single generator parameterized by scaling, shear, and translation parameters. The shear parameter captures the direction of singularities, and the shearlet transform can be regarded as matrix coefficients of a unitary representation of a special affine group. In addition, there exists a natural MRA structure associated with the systems. For the construction and a discussion of shearlets, we refer the reader to Chapter 3.

1.6.3 Moving to Other Manifolds—Wavelets on the Sphere

For applications that deal with one- or two-dimensional signals of finite duration or with data that are distributed on spherical surfaces and that require a multiscale approach, the notion of a wavelet has to be extended to encompass the underlying geometry of these applications.

One way of considering wavelets on compact intervals is via periodization, which corresponds to constructing wavelets on the circle $\mathbb{S}^1$. (See, for instance, [60].) There exist several methods of constructing wavelets on compact intervals. One such method adds boundary functions to the collection of wavelets whose support lies in the interior of the interval in order to preserve the orthogonality conditions [53]. Another approach is via multiwavelets based on fractal functions, where such boundary functions are not necessary. This latter approach was also extended to higher-dimensional settings. The interested reader may consult [72, 73, 103, 132, 181, 182] as references.

Extending wavelets to the sphere $\mathbb{S}^2$ is not trivial. One reason is the nonexistence of a homogeneous dilation operator on $\mathbb{S}^2$. In addition, notions such as the

Fourier transform that were introduced earlier in this chapter need to be transferred to $\mathbb{S}^2$. The mathematical details of harmonic analysis on the sphere and also on other manifolds can be found in [93] and [136].

In Chapter 4, a construction of wavelets on $\mathbb{S}^2$ is presented and applications to astrophysics and neuroscience are considered.

Acknowledgments This work was partially supported by the Marie Curie Excellence Team Grant MEXT-CT-2004-013477, Acronym MAMEBIA, funded by the European Commission.

Exercises

1. Prove Theorem 1.8.
2. Prove Theorem 1.16.
3. Verify the claim made in Example 1.21.
4. Prove Theorem 1.26.
5. Prove Theorem 1.28.
6. Show that $(L^1(\mathbb{T}), *)$ is a commutative Banach algebra without unity and that the Fourier transform is a Banach algebra homomorphism from $L^1(\mathbb{T}) \to l^\infty(\mathbb{Z})$.
7. Verify the claim made in Remark 1.33.
8. Prove Theorem 1.34.
9. Complete the proof of Corollary 1.50.
10. Prove Theorem 1.57.
11. Assume that $f \in L^2(\mathbb{R})$ and $g \in L^1(\mathbb{R})$. Show that $\mathscr{P}(g * f) = \mathscr{F}(g)\mathscr{P}(f)$.
12. Prove Theorem 1.62.
13. Let g_s be given as in (1.15). Show that $\Delta_{g_s} = \sqrt{s}$.
14. Prove Theorem 1.72.
15. Prove Theorem 1.74.
16. Prove Lemma 1.81. Hint: For $f \in L^2(\mathbb{R})$ and $\varepsilon > 0$, consider $\mathscr{F}(f) \cdot \chi_{\mathbb{R}\setminus(-\varepsilon,\varepsilon)}$.
17. Show that the integral in (1.21) defines an element of $L^2(\mathbb{R})$.
18. Verify the following equations for the dilation and translation operators introduced in Section 1.5.2, D_a and T_b, respectively.

    a. For the adjoint operators, $(D_a)^* = D_{1/a}$ and $(T_b)^* = T_{-b}$.
    b. $W_\psi(T_B D_A f)(a,b) = W_{T_b D_a \psi} f(1/A, -B/A) = W_\psi f(a/A, (b-B)/A)$.

19. Verify Example 1.88.
20. Show that the nested spaces

$$V_j = \{g \in L^2(\mathbb{R}) \mid \operatorname{supp} \mathscr{F}(g) \subset [-2^{-j}\pi, 2^{-j}\pi]\}, \quad j \in \mathbb{Z},$$

    form a multiresolution analysis of $L^2(\mathbb{R})$.
    Hint: Use $\varphi(t) = (\sin \pi t)/\pi t$ as the scaling function and apply the Shannon sampling theorem.

Chapter 2
B-Spline Generated Frames

Ole Christensen

Abstract B-splines are some of the most versatile functions in applied mathematics. The purpose of this chapter is to present the theory of frames in Hilbert spaces with a direct focus on B-spline generators.

2.1 Introduction

Frames provide a natural way of expanding functions in separable Hilbert spaces: They are more general than orthonormal bases and yield more flexibility. In this chapter we give a short presentation of general frame theory, as well as an introduction to frames in $L^2(\mathbb{R})$ having Gabor structure or wavelet structure. The main body of the chapter concerns explicit frame constructions based on B-splines.

The content can naturally be split into two parts: an introduction to frames in general Hilbert spaces, and concrete constructions in $L^2(\mathbb{R})$. The two parts are tied together by Section 2.7, where the B-splines are introduced.

We begin in Section 2.2 by considering the so-called Bessel condition: It is a technical condition implying that all the series expansions considered in this chapter converge unconditionally. Section 2.3 reminds the reader about bases, in particular, orthonormal bases, in Hilbert spaces; the important case of a Riesz basis is discussed in Section 2.4. Section 2.5 introduces frames and their central properties. Section 2.6 relates frames and Riesz bases; in particular, it turns out that all Riesz bases are frames.

Section 2.7 marks the beginning of the second part of the chapters, where we focus on concrete constructions. Most of these constructions are based on B-splines, so Section 2.7 gives a short presentation on their key properties. Section 2.8

Ole Christensen
Technical University of Denmark, Department of Mathematics, Building 303, 2800 Lyngby, Denmark, e-mail: ole.christensen@mat.dtu.dk

B. Forster and P. Massopust (eds.), *Four Short Courses on Harmonic Analysis*,
Applied and Numerical Harmonic Analysis, DOI 10.1007/978-0-8176-4891-6_2,

deals with the basic properties of systems of functions formed by translates of a single function. Section 2.9 introduces Gabor systems and their frame properties, and Section 2.10 focuses on the case of tight frames. Section 2.11 states the main results from the theory for dual frames associated with a given Gabor frame; in Section 2.12 these results are used to construct explicitly given dual pairs of Gabor frames. Finally, Section 2.13 deals with wavelet frames generated by B-splines, in particular, the constructions obtained via the unitary extension principle due to Ron and Shen.

2.2 Bessel Sequences in Hilbert Spaces

The ultimate goal of the present chapter is to obtain series expansions in infinite-dimensional vector spaces. The purpose of the current section is to introduce a condition that ensures that the relevant infinite series actually converge.

Let $\mathscr{H}$ be a separable Hilbert space, with the inner product $\langle \cdot, \cdot \rangle$ chosen to be linear in the first entry. When speaking about a *sequence* $\{f_k\}_{k=1}^{\infty}$ in $\mathscr{H}$, we mean an *ordered* set, i.e.,

$$\{f_k\}_{k=1}^{\infty} = \{f_1, f_2, \ldots\}.$$

That we have chosen to index the sequence by the natural numbers is just for convenience: Soon, we will see that all results in this section (and all subsequent results based on the Bessel condition) hold with arbitrary countable index sets and the elements f_k ordered in an arbitrary way.

We begin with a technical lemma.

Lemma 2.1. *Let $\{f_k\}_{k=1}^{\infty}$ be a sequence in $\mathscr{H}$, and suppose that $\sum_{k=1}^{\infty} c_k f_k$ is convergent for all $\{c_k\}_{k=1}^{\infty} \in \ell^2(\mathbb{N})$. Then*

$$T : \ell^2(\mathbb{N}) \to \mathscr{H}, \qquad T\{c_k\}_{k=1}^{\infty} := \sum_{k=1}^{\infty} c_k f_k \tag{2.1}$$

defines a bounded linear operator. The adjoint operator is given by

$$T^* : \mathscr{H} \to \ell^2(\mathbb{N}), \qquad T^* f = \{\langle f, f_k \rangle\}_{k=1}^{\infty}. \tag{2.2}$$

Furthermore,

$$\sum_{k=1}^{\infty} |\langle f, f_k \rangle|^2 \leq \|T\|^2 \, \|f\|^2, \qquad \forall f \in \mathscr{H}. \tag{2.3}$$

Proof. Consider the sequence of bounded linear operators

$$T_n : \ell^2(\mathbb{N}) \to \mathscr{H}, \qquad T_n\{c_k\}_{k=1}^{\infty} := \sum_{k=1}^{n} c_k f_k.$$

Clearly, $T_n \to T$ pointwise as $n \to \infty$, so by the principle of uniform boundedness, the map T defines a bounded linear operator. In order to find the expression for T^*, let $f \in \mathscr{H}$ and $\{c_k\}_{k=1}^{\infty} \in \ell^2(\mathbb{N})$. Then

$$\langle f, T\{c_k\}_{k=1}^{\infty}\rangle_{\mathscr{H}} = \left\langle f, \sum_{k=1}^{\infty} c_k f_k \right\rangle_{\mathscr{H}} = \sum_{k=1}^{\infty} \langle f, f_k\rangle \overline{c_k}. \tag{2.4}$$

When $T : \ell^2(\mathbb{N}) \to \mathscr{H}$ is bounded, we know that T^* is a bounded operator from $\mathscr{H}$ to $\ell^2(\mathbb{N})$. Therefore, the kth-coordinate function is bounded from $\mathscr{H}$ to $\mathbb{C}$; by Rieszs' representation theorem, T^* therefore has the form

$$T^* f = \{\langle f, g_k\rangle\}_{k=1}^{\infty}$$

for some $\{g_k\}_{k=1}^{\infty}$ in $\mathscr{H}$. By the definition of T^*, (2.4) now shows that

$$\sum_{k=1}^{\infty} \langle f, g_k\rangle \overline{c_k} = \sum_{k=1}^{\infty} \langle f, f_k\rangle \overline{c_k}, \quad \forall \{c_k\}_{k=1}^{\infty} \in \ell^2(\mathbb{N}),\ f \in \mathscr{H}.$$

It follows from here that $g_k = f_k$.

The adjoint of a bounded operator T is itself bounded, and $\|T\| = \|T^*\|$. Under the assumption in Lemma 2.1, we therefore have

$$\|T^* f\|^2 \leq \|T\|^2 \ \|f\|^2, \quad \forall f \in \mathscr{H},$$

which leads to (2.3). □

Sequences $\{f_k\}_{k=1}^{\infty}$ for which an inequality of the type (2.3) holds will play a crucial role in the sequel.

Definition 2.2. A sequence $\{f_k\}_{k=1}^{\infty}$ in $\mathscr{H}$ is called a *Bessel sequence* if there exists a constant $B > 0$ such that

$$\sum_{k=1}^{\infty} |\langle f, f_k\rangle|^2 \leq B \ \|f\|^2, \quad \forall f \in \mathscr{H}. \tag{2.5}$$

Any number B satisfying (2.5) is called a *Bessel bound* for $\{f_k\}_{k=1}^{\infty}$. The *optimal bound* for a given Bessel sequence $\{f_k\}_{k=1}^{\infty}$ is the smallest possible value of $B > 0$ satisfying (2.5). Except for the case $f_k = 0$, $\forall k \in \mathbb{N}$, the optimal bound always exists. We will now present a useful characterization of Bessel sequences.

Theorem 2.3. *Let $\{f_k\}_{k=1}^{\infty}$ be a sequence in $\mathscr{H}$ and $B > 0$ be given. Then $\{f_k\}_{k=1}^{\infty}$ is a Bessel sequence with Bessel bound B if and only if*

$$T : \{c_k\}_{k=1}^{\infty} \mapsto \sum_{k=1}^{\infty} c_k f_k$$

defines a bounded operator from $\ell^2(\mathbb{N})$ into $\mathscr{H}$ and $\|T\| \leq \sqrt{B}$.

Proof. First, assume that $\{f_k\}_{k=1}^{\infty}$ is a Bessel sequence with Bessel bound B. Let $\{c_k\}_{k=1}^{\infty} \in \ell^2(\mathbb{N})$. First, we want to show that $T\{c_k\}_{k=1}^{\infty}$ is well defined, i.e., that

$\sum_{k=1}^{\infty} c_k f_k$ is convergent. Consider $n, m \in \mathbb{N}, n > m$. Then

$$\left\| \sum_{k=1}^{n} c_k f_k - \sum_{k=1}^{m} c_k f_k \right\| = \left\| \sum_{k=m+1}^{n} c_k f_k \right\|.$$

It follows that

$$\begin{aligned}
\left\| \sum_{k=1}^{n} c_k f_k - \sum_{k=1}^{m} c_k f_k \right\| &= \sup_{\|g\|=1} \left| \left\langle \sum_{k=m+1}^{n} c_k f_k, g \right\rangle \right| \\
&\leq \sup_{\|g\|=1} \sum_{k=m+1}^{n} |c_k \langle f_k, g\rangle| \\
&\leq \left(\sum_{k=m+1}^{n} |c_k|^2 \right)^{1/2} \sup_{\|g\|=1} \left(\sum_{k=m+1}^{n} |\langle f_k, g\rangle|^2 \right)^{1/2} \\
&\leq \sqrt{B} \left(\sum_{k=m+1}^{n} |c_k|^2 \right)^{1/2}.
\end{aligned}$$

Since $\{c_k\}_{k=1}^{\infty} \in \ell^2(\mathbb{N})$, we know that $\left\{\sum_{k=1}^{n} |c_k|^2\right\}_{n=1}^{\infty}$ is a Cauchy sequence in $\mathbb{C}$. The above calculation now shows that $\{\sum_{k=1}^{n} c_k f_k\}_{n=1}^{\infty}$ is a Cauchy sequence in $\mathscr{H}$ and therefore convergent. Thus, $T\{c_k\}_{k=1}^{\infty}$ is well defined. Clearly, T is linear; since $\|T\{c_k\}_{k=1}^{\infty}\| = \sup_{\|g\|=1} |\langle T\{c_k\}_{k=1}^{\infty}, g\rangle|$, a calculation as above shows that T is bounded and that $\|T\| \leq \sqrt{B}$.

For the opposite implication, suppose that T defines a bounded operator with $\|T\| \leq \sqrt{B}$. Then Lemma 2.1 shows that $\{f_k\}_{k=1}^{\infty}$ is a Bessel sequence with Bessel bound B. □

The Bessel condition (2.5) remains the same regardless of how the elements $\{f_k\}_{k=1}^{\infty}$ are numbered. This leads to a very important consequence of Theorem 2.3:

Corollary 2.4. *If $\{f_k\}_{k=1}^{\infty}$ is a Bessel sequence in $\mathscr{H}$, then $\sum_{k=1}^{\infty} c_k f_k$ converges unconditionally for all $\{c_k\}_{k=1}^{\infty} \in \ell^2(\mathbb{N})$.*

Thus, a reordering of the elements in $\{f_k\}_{k=1}^{\infty}$ will not affect the series $\sum_{k=1}^{\infty} c_k f_k$ when $\{c_k\}_{k=1}^{\infty}$ is reordered the same way: The series will converge toward the same element as before. For this reason we can choose an arbitrary indexing of the elements in the Bessel sequence; in particular, it is not a restriction that we present all results with the natural numbers as the index set. As we will see in the sequel, all orthonormal bases, Riesz bases, and frames are Bessel sequences.

2.3 General Bases and Orthonormal Bases

Before we introduce the frame concept in Section 2.5, we shortly remind the reader about bases in Hilbert spaces. In particular, we will discuss orthonormal bases.

Orthonormal bases are widely used in mathematics as well as in physics, signal processing, and many other areas where one needs to represent functions in terms of "elementary building blocks."

Definition 2.5. Consider a sequence $\{e_k\}_{k=1}^{\infty}$ of vectors in a Hilbert space $\mathscr{H}$.

1. The sequence $\{e_k\}_{k=1}^{\infty}$ is a (Schauder) basis for $\mathscr{H}$ if for each $f \in \mathscr{H}$ there exist unique scalar coefficients $\{c_k(f)\}_{k=1}^{\infty}$ such that
$$f = \sum_{k=1}^{\infty} c_k(f) e_k. \tag{2.6}$$
2. A basis $\{e_k\}_{k=1}^{\infty}$ is an unconditional basis if the series (2.6) converges unconditionally for each $f \in \mathscr{H}$.
3. A basis $\{e_k\}_{k=1}^{\infty}$ is an orthonormal basis if $\{e_k\}_{k=1}^{\infty}$ is an orthonormal system, i.e., if
$$\langle e_k, e_j \rangle = \delta_{k,j} = \begin{cases} 1 \text{ if } & k = j, \\ 0 \text{ if } & k \neq j. \end{cases}$$

The next well-known theorem gives equivalent conditions for an orthonormal system $\{e_k\}_{k=1}^{\infty}$ to be an orthonormal basis.

Theorem 2.6. *For an orthonormal system $\{e_k\}_{k=1}^{\infty}$, the following are equivalent:*

1. $\{e_k\}_{k=1}^{\infty}$ *is an orthonormal basis.*
2. $f = \sum_{k=1}^{\infty} \langle f, e_k \rangle e_k, \ \forall f \in \mathscr{H}$.
3. $\langle f, g \rangle = \sum_{k=1}^{\infty} \langle f, e_k \rangle \langle e_k, g \rangle, \ \forall f, g \in \mathscr{H}$.
4. $\sum_{k=1}^{\infty} |\langle f, e_k \rangle|^2 = \|f\|^2, \ \forall f \in \mathscr{H}$.
5. $\overline{\text{span}}\{e_k\}_{k=1}^{\infty} = \mathscr{H}$.
6. *If* $\langle f, e_k \rangle = 0, \ \forall k \in \mathbb{N}$, *then* $f = 0$.

The equality in item 4 is called *Parseval's equation*; in particular, it shows that an orthonormal system $\{e_k\}_{k=1}^{\infty}$ is a Bessel sequence. Via Corollary 2.4, we obtain the following important consequence of Theorem 2.6:

Corollary 2.7. *If $\{e_k\}_{k=1}^{\infty}$ is an orthonormal basis, then each $f \in \mathscr{H}$ has an unconditionally convergent expansion*
$$f = \sum_{k=1}^{\infty} \langle f, e_k \rangle e_k. \tag{2.7}$$

The expansion property (2.7) is the main reason for considering orthonormal bases. In practice, orthonormal bases are certainly the most convenient bases to use: We will later see that, for other types of bases, the representation (2.7) has to be replaced by a more complicated expression. Unfortunately, the conditions for $\{e_k\}_{k=1}^{\infty}$ being an orthonormal basis are strong, and often it is impossible to construct orthonormal bases satisfying extra conditions.

The following theorem characterizes all orthonormal bases for $\mathscr{H}$ in terms of an operator acting on an arbitrary orthonormal basis.

Theorem 2.8. *Let $\{e_k\}_{k=1}^{\infty}$ be an orthonormal basis for $\mathscr{H}$. Then the orthonormal bases for $\mathscr{H}$ are precisely the sets $\{Ue_k\}_{k=1}^{\infty}$, where $U : \mathscr{H} \to \mathscr{H}$ is a unitary operator.*

Proof. Let $\{f_k\}_{k=1}^{\infty}$ be an orthonormal basis for $\mathscr{H}$. Define the operator

$$U : \mathscr{H} \to \mathscr{H},\ U\left(\sum c_k e_k\right) = \sum c_k f_k, \qquad \{c_k\}_{k=1}^{\infty} \in \ell^2(\mathbb{N}).$$

Then U maps $\mathscr{H}$ boundedly and bijectively onto $\mathscr{H}$, and $f_k = Ue_k$. For $f, g \in \mathscr{H}$, write $f = \sum \langle f, e_k \rangle e_k$ and $g = \sum \langle g, e_k \rangle e_k$; then, via the definition of U and Theorem 2.6,

$$\begin{aligned} \langle U^*Uf, g\rangle &= \langle Uf, Ug\rangle \\ &= \left\langle \sum \langle f, e_k\rangle f_k, \sum \langle g, e_k\rangle f_k \right\rangle \\ &= \sum \langle f, e_k\rangle \overline{\langle g, e_k\rangle} = \langle f, g\rangle. \end{aligned}$$

This implies that $U^*U = I$. Since U is surjective, it follows that U is unitary. On the other hand, if U is a given unitary operator, then

$$\langle Ue_k, Ue_j\rangle = \langle U^*Ue_k, e_j\rangle = \langle e_k, e_j\rangle = \delta_{k,j};$$

i.e., $\{Ue_k\}_{k=1}^{\infty}$ is an orthonormal system. That it is a basis follows from Theorem 2.6 and the fact that U is surjective. □

2.4 Riesz Bases

In Theorem 2.8 we characterized all orthonormal bases in terms of unitary operators acting on a single orthonormal basis. Formally, the definition of a *Riesz basis* appears by weakening the condition on the operator:

Definition 2.9. A *Riesz basis* for $\mathscr{H}$ is a family of the form $\{Ue_k\}_{k=1}^{\infty}$, where $\{e_k\}_{k=1}^{\infty}$ is an orthonormal basis for $\mathscr{H}$ and $U : \mathscr{H} \to \mathscr{H}$ is a bounded bijective operator.

A Riesz basis $\{f_k\}_{k=1}^{\infty}$ is actually a basis; this follows from the proof of Theorem 2.10, which we state now. Note that the expansion (2.8) of elements $f \in \mathscr{H}$ in terms of a Riesz basis is more involved than the expression (2.7) we obtained via orthonormal bases:

Theorem 2.10. *If $\{f_k\}_{k=1}^{\infty}$ is a Riesz basis for $\mathscr{H}$, then $\{f_k\}_{k=1}^{\infty}$ is a Bessel sequence. Furthermore, there exists a unique sequence $\{g_k\}_{k=1}^{\infty}$ in $\mathscr{H}$ such that*

$$f = \sum_{k=1}^{\infty} \langle f, g_k\rangle f_k, \quad \forall f \in \mathscr{H}. \tag{2.8}$$

The sequence $\{g_k\}_{k=1}^{\infty}$ *is also a Riesz basis, and the series* (2.8) *converges unconditionally for all* $f \in \mathscr{H}$.

Proof. According to the definition, we can write $\{f_k\}_{k=1}^{\infty} = \{Ue_k\}_{k=1}^{\infty}$, where U is a bounded bijective operator and $\{e_k\}_{k=1}^{\infty}$ is an orthonormal basis. Now let $f \in \mathscr{H}$. By expanding $U^{-1}f$ in the orthonormal basis $\{e_k\}_{k=1}^{\infty}$, we have

$$U^{-1}f = \sum_{k=1}^{\infty} \langle U^{-1}f, e_k\rangle e_k = \sum_{k=1}^{\infty} \langle f, (U^{-1})^* e_k\rangle e_k.$$

Therefore, with $g_k := (U^{-1})^* e_k$,

$$\begin{aligned} f = UU^{-1}f &= \sum_{k=1}^{\infty} \langle f, (U^{-1})^* e_k\rangle Ue_k \\ &= \sum_{k=1}^{\infty} \langle f, g_k\rangle f_k. \end{aligned}$$

Since the operator $(U^{-1})^*$ is bounded and bijective, $\{g_k\}_{k=1}^{\infty}$ is a Riesz basis by definition. Furthermore, for $f \in \mathscr{H}$,

$$\begin{aligned} \sum_{k=1}^{\infty} |\langle f, f_k\rangle|^2 = \sum_{k=1}^{\infty} |\langle f, Ue_k\rangle|^2 &= \|U^* f\|^2 & (2.9) \\ &\leq \|U^*\|^2 \|f\|^2 \\ &= \|U\|^2 \|f\|^2. & (2.10) \end{aligned}$$

This proves that a Riesz basis is a Bessel sequence. Thus, the series (2.8) converges unconditionally by Corollary 2.4. We complete the proof by showing that the sequence $\{g_k\}_{k=1}^{\infty}$ constructed in the proof is the only one that satisfies (2.8). For that purpose, we first note that if

$$f = \sum_{k=1}^{\infty} c_k(f) f_k = \sum_{k=1}^{\infty} d_k(f) f_k \tag{2.11}$$

for some coefficients $c_k(f)$ and $d_k(f)$, then necessarily $c_k(f) = d_k(f)$ for all $k \in \mathbb{N}$; this follows by applying the operator U^{-1} on both sides of the equality and using that $\{e_k\}_{k=1}^{\infty}$ is known to be a basis. This argument shows that a Riesz basis actually is a basis. Now we only have to show that if $\{g_k\}_{k=1}^{\infty}$ and $\{h_k\}_{k=1}^{\infty}$ are sequences in $\mathscr{H}$ such that

$$f = \sum_{k=1}^{\infty} \langle f, g_k\rangle f_k = \sum_{k=1}^{\infty} \langle f, h_k\rangle f_k, \quad \forall f \in \mathscr{H}, \tag{2.12}$$

then $g_k = h_k$ for all $k \in \mathbb{N}$. However, due to the argument above, (2.12) implies that for all $k \in \mathbb{N}$,

$$\langle f, g_k\rangle = \langle f, h_k\rangle, \quad \forall f \in \mathscr{H};$$

the desired result now follows. $\square$

The unique sequence $\{g_k\}_{k=1}^\infty$ satisfying (2.8) is called the *dual Riesz basis* of $\{f_k\}_{k=1}^\infty$. Let us find the dual of $\{g_k\}_{k=1}^\infty$. In the notation used in the proof of Theorem 2.10, we have that the dual of $\{f_k\}_{k=1}^\infty = \{Ue_k\}_{k=1}^\infty$ is given by

$$\{g_k\}_{k=1}^\infty = \{(U^{-1})^* e_k\}_{k=1}^\infty;$$

thus, the dual of $\{g_k\}_{k=1}^\infty$ is

$$\left\{ \left(\left((U^{-1})^* \right)^{-1} \right)^* e_k \right\}_{k=1}^\infty = \{Ue_k\}_{k=1}^\infty = \{f_k\}_{k=1}^\infty.$$

That is, $\{f_k\}_{k=1}^\infty$ and $\{g_k\}_{k=1}^\infty$ are duals of each other. For this reason, we frequently call $\{f_k\}_{k=1}^\infty$ and $\{g_k\}_{k=1}^\infty$ a *pair of dual Riesz bases.*

Already in the proof of Theorem 2.8, we saw that a Riesz basis is a Bessel sequence. For later use, we now state that it also satisfies some kind of "opposite inequality"; the proof is left to the reader (Exercise 5).

Proposition 2.11. *If* $\{f_k\}_{k=1}^\infty = \{Ue_k\}_{k=1}^\infty$ *is a Riesz basis for* $\mathscr{H}$*, there exist constants* $A, B > 0$ *such that*

$$A\,\|f\|^2 \le \sum_{k=1}^\infty |\langle f, f_k\rangle|^2 \le B\,\|f\|^2, \quad \forall f \in \mathscr{H}. \tag{2.13}$$

The largest possible value for the constant A is $1/\|U^{-1}\|^2$*, and the smallest possible value for B is* $\|U\|^2$*.*

For completeness we finally state an equivalent characterization of Riesz bases. Several authors use condition (ii) as the definition of a Riesz basis; the proof of the equivalence with the definition used here can be found, e.g., in [246] or [49].

Theorem 2.12. *For a sequence* $\{f_k\}_{k=1}^\infty$ *in* $\mathscr{H}$*, the following conditions are equivalent:*

1. $\{f_k\}_{k=1}^\infty$ *is a Riesz basis for* $\mathscr{H}$*.*
2. $\{f_k\}_{k=1}^\infty$ *is complete in* $\mathscr{H}$*, and there exist constants* $A, B > 0$ *such that for every finite scalar sequence* $\{c_k\}$*, one has*

$$A \sum |c_k|^2 \le \left\| \sum c_k f_k \right\|^2 \le B \sum |c_k|^2. \tag{2.14}$$

A sequence $\{f_k\}_{k=1}^\infty$ satisfying (2.14) for all finite sequences $\{c_k\}_{k=1}^\infty$ is called a *Riesz sequence.*

If (2.14) holds for all finite scalar sequences $\{c_k\}$, then it automatically holds for all $\{c_k\}_{k=1}^\infty \in \ell^2(\mathbb{N})$; see Exercise 6. If $\{f_k\}_{k=1}^\infty$ is a Riesz basis, numbers $A, B > 0$ that satisfy (2.14) are called *lower Riesz bounds*, respectively, *upper Riesz bounds*. They are clearly not unique, and we define the *optimal Riesz bounds* as the largest possible value for A and the smallest possible value for B.

2.5 Frames and Their Properties

We are now ready to introduce frames. Frames were invented in 1952 by Duffin and Schaeffer [77], but it took several years before the potential was realized by the scientific community. By now, frame theory is well established; we will only give a glimpse of the general theory, and focus on the parts of the theory that are important for our later constructions based on B-splines.

Definition 2.13. A sequence $\{f_k\}_{k=1}^{\infty}$ of elements in $\mathscr{H}$ is a frame for $\mathscr{H}$ if there exist constants $A, B > 0$ such that

$$A\,\|f\|^2 \le \sum_{k=1}^{\infty} |\langle f, f_k\rangle|^2 \le B\,\|f\|^2, \quad \forall f \in \mathscr{H}. \tag{2.15}$$

The numbers A and B are called *frame bounds*. They are not unique. The *optimal upper frame bound* is the infimum over all upper frame bounds, and the *optimal lower frame bound* is the supremum over all lower frame bounds. Note that the optimal bounds actually are frame bounds.

The following lemma shows that it is enough to check the frame condition on a dense set. The proof is left to the reader as Exercise 7.

Lemma 2.14. *Suppose that $\{f_k\}_{k=1}^{\infty}$ is a sequence of elements in $\mathscr{H}$ and that there exist constants $A, B > 0$ such that*

$$A\,\|f\|^2 \le \sum_{k=1}^{\infty} |\langle f, f_k\rangle|^2 \le B\,\|f\|^2 \tag{2.16}$$

for all f in a dense subset V of $\mathscr{H}$. Then $\{f_k\}_{k=1}^{\infty}$ is a frame for $\mathscr{H}$ with bounds A, B.

A special role is played by frames for which the optimal frame bounds coincide:

Definition 2.15. A sequence $\{f_k\}_{k=1}^{\infty}$ in $\mathscr{H}$ is a tight frame if there exists a number $A > 0$ such that

$$\sum_{k=1}^{\infty} |\langle f, f_k\rangle|^2 = A\,\|f\|^2, \quad \forall f \in \mathscr{H}.$$

The (exact) number A is called the frame bound.

Since a frame $\{f_k\}_{k=1}^{\infty}$ is a Bessel sequence, the operator

$$T : \ell^2(\mathbb{N}) \to \mathscr{H}, \qquad T\{c_k\}_{k=1}^{\infty} = \sum_{k=1}^{\infty} c_k f_k \tag{2.17}$$

is bounded by Theorem 2.3; T is called the *preframe operator* or the *synthesis operator.* By Lemma 2.1, the adjoint operator is given by

$$T^* : \mathscr{H} \to \ell^2(\mathbb{N}), \qquad T^* f = \{\langle f, f_k\rangle\}_{k=1}^{\infty}. \tag{2.18}$$

T^* is called the *analysis operator.* Composing T and T^*, we obtain the *frame operator*

$$S:\mathscr{H}\to\mathscr{H},\qquad Sf=TT^*f=\sum_{k=1}^{\infty}\langle f,f_k\rangle f_k. \tag{2.19}$$

Note that because $\{f_k\}_{k=1}^{\infty}$ is a Bessel sequence, the series defining S converges unconditionally for all $f\in\mathscr{H}$ by Corollary 2.4. We state some of the important properties of S; proofs can be found, e.g., in [135] or [49].

Lemma 2.16. *Let $\{f_k\}_{k=1}^{\infty}$ be a frame with frame operator S and frame bounds A,B. Then the following hold:*

1. *S is bounded, invertible, self-adjoint, and positive.*
2. *$\{S^{-1}f_k\}_{k=1}^{\infty}$ is a frame with frame operator S^{-1} and frame bounds B^{-1}, A^{-1}.*

The frame $\{S^{-1}f_k\}_{k=1}^{\infty}$ is called the *canonical dual frame* of $\{f_k\}_{k=1}^{\infty}$. The reason for the name will soon become clear; in fact, Theorem 2.17 will show that $\{S^{-1}f_k\}_{k=1}^{\infty}$ plays the same role in frame theory as the dual basis in the theory of bases.

The *frame decomposition*, stated in (2.20) below, is one of the most important frame results. It shows that if $\{f_k\}_{k=1}^{\infty}$ is a frame for $\mathscr{H}$, then every element in $\mathscr{H}$ has a representation as an infinite linear combination of the frame elements. Thus, it is natural to view a frame as some kind of "generalized basis."

Theorem 2.17. *Let $\{f_k\}_{k=1}^{\infty}$ be a frame with frame operator S. Then*

$$f=\sum_{k=1}^{\infty}\langle f,S^{-1}f_k\rangle f_k,\quad \forall f\in\mathscr{H}, \tag{2.20}$$

and

$$f=\sum_{k=1}^{\infty}\langle f,f_k\rangle S^{-1}f_k,\quad \forall f\in\mathscr{H}. \tag{2.21}$$

Both series converge unconditionally for all $f\in\mathscr{H}$.

Proof. Let $f\in\mathscr{H}$. Using the properties of the frame operator in Lemma 2.16,

$$f=SS^{-1}f=\sum_{k=1}^{\infty}\langle S^{-1}f,f_k\rangle f_k=\sum_{k=1}^{\infty}\langle f,S^{-1}f_k\rangle f_k.$$

Because $\{f_k\}_{k=1}^{\infty}$ is a Bessel sequence and $\{\langle f,S^{-1}f_k\rangle\}_{k=1}^{\infty}\in\ell^2(\mathbb{N})$, the fact that the series converges unconditionally follows from Corollary 2.4. The expansion (2.21) is proved similarly, using that $f=S^{-1}Sf$. □

Theorem 2.17 shows that all information about a given vector $f\in\mathscr{H}$ is contained in the sequence $\{\langle f,S^{-1}f_k\rangle\}_{k=1}^{\infty}$. The numbers $\langle f,S^{-1}f_k\rangle$ are called *frame coefficients.*

Theorem 2.17 also immediately reveals one of the main difficulties in frame theory. In fact, in order for the expansions (2.20) and (2.21) to be applicable in practice, we need to be able to find the operator S^{-1}, or at least to calculate its action on all f_k, $k \in \mathbb{N}$. In general, this is a major problem. One way of circumventing the problem is to consider only tight frames:

Corollary 2.18. *If $\{f_k\}_{k=1}^{\infty}$ is a tight frame with frame bound A, then the canonical dual frame is $\{A^{-1}f_k\}_{k=1}^{\infty}$, and*

$$f = \frac{1}{A}\sum_{k=1}^{\infty}\langle f, f_k\rangle f_k, \quad \forall f \in \mathscr{H}. \tag{2.22}$$

Proof. If $\{f_k\}_{k=1}^{\infty}$ is a tight frame with frame bound A and frame operator S, the definition shows that

$$\langle Sf, f\rangle = \sum_{k=1}^{\infty}|\langle f, f_k\rangle|^2 = A\,\|f\|^2 = \langle Af, f\rangle, \quad \forall f \in \mathscr{H}.$$

Since S is self-adjoint, this implies that $S = AI$; thus, S^{-1} acts by multiplication by A^{-1}, and the result follows from (2.20). □

By a suitable scaling of the vectors $\{f_k\}_{k=1}^{\infty}$ in a tight frame, we can always obtain that $A = 1$; in that case, (2.22) has exactly the same form as the representation via an orthonormal basis; see (2.7). Thus, such frames can be used without any additional computational effort compared with the use of orthonormal bases.

Tight frames have other advantages. For the design of frames with prescribed properties, it is essential to control the behavior of the canonical dual frame, but the complicated structure of the frame operator and its inverse makes this difficult. If, for example, we consider a frame $\{f_k\}_{k=1}^{\infty}$ for $L^2(\mathbb{R})$ consisting of functions with exponential decay, nothing guarantees that the functions in the canonical dual frame $\{S^{-1}f_k\}_{k=1}^{\infty}$ have exponential decay. However, for tight frames, questions of this type trivially have satisfying answers. Also, for a tight frame, the canonical dual frame automatically has the same structure as the frame itself: If the frame has a wavelet structure or a Gabor structure (see Sections 2.9–2.13), the same is the case for the canonical dual frame. In contrast, the canonical dual frame of a nontight wavelet frame might not have the wavelet structure.

Later we will discuss another way to avoid the problem of inverting the frame operator S. In fact, for frames $\{f_k\}_{k=1}^{\infty}$ that are *not* bases, we prove in Theorem 2.21 that one can find other frames $\{g_k\}_{k=1}^{\infty}$ than $\{S^{-1}f_k\}_{k=1}^{\infty}$, for which

$$f = \sum_{k=1}^{\infty}\langle f, g_k\rangle f_k, \quad \forall f \in \mathscr{H}. \tag{2.23}$$

Such a frame $\{g_k\}_{k=1}^{\infty}$ is called a *dual frame* of $\{f_k\}_{k=1}^{\infty}$. Now, there is a chance that even if the canonical dual frame is difficult to find, there exist other duals that are easy to find; or, that it is possible to find duals having more pleasant properties than the canonical dual. In Section 2.12 we discuss such cases.

A note on terminology is in order. In Exercise 9 we ask the reader to prove that if $\{g_k\}_{k=1}^{\infty}$ is a dual frame of $\{f_k\}_{k=1}^{\infty}$, then $\{f_k\}_{k=1}^{\infty}$ is also a dual of $\{g_k\}_{k=1}^{\infty}$. For this reason, we will usually call $\{f_k\}_{k=1}^{\infty}$ and $\{g_k\}_{k=1}^{\infty}$ a *pair of dual frames,* or a *dual frame pair,* when (2.23) holds.

2.6 Frames and Riesz Bases

As we have seen, a frame $\{f_k\}_{k=1}^{\infty}$ in a Hilbert space $\mathscr{H}$ has one of the main properties of a basis: Given $f \in \mathscr{H}$, there exist coefficients $\{c_k\}_{k=1}^{\infty} \in \ell^2(\mathbb{N})$ such that $f = \sum_{k=1}^{\infty} c_k f_k$. This makes it natural to study the relationship between frames and bases. In this section we notice that all Riesz bases are frames and characterize the frames that are actually Riesz bases.

Theorem 2.19. *A Riesz basis $\{f_k\}_{k=1}^{\infty}$ for $\mathscr{H}$ is a frame for $\mathscr{H}$. The dual Riesz basis equals the canonical dual frame $\{S^{-1}f_k\}_{k=1}^{\infty}$.*

Proof. By Proposition 2.11, a Riesz basis $\{f_k\}_{k=1}^{\infty}$ for $\mathscr{H}$ is also a frame for $\mathscr{H}$. The rest follows from the frame decomposition combined with the uniqueness part of Theorem 2.10. □

A frame that is *not* a Riesz basis is said to be *overcomplete;* in the literature, the term "redundant frame" is also used. Theorem 2.20 will explain why the word "overcomplete" is used: In fact, if $\{f_k\}_{k=1}^{\infty}$ is a frame that is not a Riesz basis, there exist coefficients $\{c_k\}_{k=1}^{\infty} \in \ell^2(\mathbb{N}) \setminus \{0\}$ for which

$$\sum_{k=1}^{\infty} c_k f_k = 0. \tag{2.24}$$

That is, for such frames there is some dependency between the frame elements.

Theorem 2.20. *Let $\{f_k\}_{k=1}^{\infty}$ be a frame for $\mathscr{H}$. Then the following are equivalent:*

1. $\{f_k\}_{k=1}^{\infty}$ is a Riesz basis for $\mathscr{H}$.
2. If $\sum_{k=1}^{\infty} c_k f_k = 0$ for some $\{c_k\}_{k=1}^{\infty} \in \ell^2(\mathbb{N})$, then $c_k = 0$, $\forall k \in \mathbb{N}$.

Proof. *1* $\Rightarrow$ *2*: Assume that $\{f_k\}_{k=1}^{\infty}$ is a Riesz basis and that $\sum_{k=1}^{\infty} c_k f_k = 0$ for a sequence $\{c_k\}_{k=1}^{\infty} \in \ell^2(\mathbb{N})$. Writing $\{f_k\}_{k=1}^{\infty} = \{Ue_k\}_{k=1}^{\infty}$ for a certain orthonormal basis for $\mathscr{H}$ and an appropriate bounded bijective operator $\mathscr{H}$, it follows that

$$U \sum_{k=1}^{\infty} c_k e_k = 0.$$

Because U is injective, this implies that $\sum_{k=1}^{\infty} c_k e_k = 0$, and thus $c_k = 0$ for all k.

2 $\Rightarrow$ *1*: Let $\{\delta_k\}_{k=1}^{\infty}$ be the canonical orthonormal basis for $\ell^2(\mathbb{N})$. Assumption (ii) assures that the preframe operator T associated with $\{f_k\}_{k=1}^{\infty}$ is injective, and T is also surjective because $\{f_k\}_{k=1}^{\infty}$ is a frame. Since $T\delta_k = f_k$, $\forall k$, the result follows from the definition of a Riesz basis. □

Much more can be said about the relationship between frames and Riesz bases; the following result and the proof are borrowed from [135].

Theorem 2.21. *Assume that $\{f_k\}_{k=1}^{\infty}$ is an overcomplete frame. Then there exist frames $\{g_k\}_{k=1}^{\infty} \neq \{S^{-1}f_k\}_{k=1}^{\infty}$ for which*

$$f = \sum_{k=1}^{\infty} \langle f, g_k \rangle f_k, \quad \forall f \in \mathscr{H}. \tag{2.25}$$

Proof. We split the proof in two cases and assume first that $f_\ell = 0$ for some $\ell \in \mathbb{N}$; in this case $S^{-1} f_\ell = 0$. Letting $g_k := S^{-1} f_k$ for $k \neq \ell$ and choosing g_ℓ to be any arbitrary nonzero vector, the frame decomposition shows that (2.25) holds; and, clearly, $\{g_k\}_{k=1}^{\infty} \neq \{S^{-1} f_k\}_{k=1}^{\infty}$.

Now we consider the case where $f_k \neq 0$ for all $k \in \mathbb{N}$. By Theorem 2.20, there exists a sequence $\{c_k\}_{k=1}^{\infty} \in \ell^2(\mathbb{N}) \setminus \{0\}$ such that

$$0 = \sum_{k=1}^{\infty} c_k f_k.$$

For a certain $\ell \in \mathbb{N}$, we have $c_\ell \neq 0$, and we can write

$$f_\ell = \frac{-1}{c_\ell} \sum_{k \neq \ell} c_k f_k.$$

We now show that $\{f_k\}_{k \neq \ell}$ is a frame for $\mathscr{H}$; we only have to prove that $\{f_k\}_{k \neq \ell}$ satisfies the lower frame condition. In order to do so, observe that for any $f \in \mathscr{H}$, the Cauchy–Schwarz inequality shows that

$$\begin{aligned} |\langle f, f_\ell \rangle|^2 &= \left| \frac{-1}{c_\ell} \sum_{k \neq \ell} c_k \langle f, f_k \rangle \right|^2 \\ &\leq \frac{1}{|c_\ell|^2} \sum_{k \neq \ell} |c_k|^2 \sum_{k \neq \ell} |\langle f, f_k \rangle|^2 \\ &= C \sum_{k \neq \ell} |\langle f, f_k \rangle|^2, \end{aligned}$$

where $C := (1/|c_\ell|^2) \sum_{k \neq \ell} |c_k|^2$. Letting A denote a lower frame bound for the frame $\{f_k\}_{k=1}^{\infty}$, this implies that

$$\begin{aligned} A\,\|f\|^2 &\leq \sum_{k=1}^{\infty} |\langle f, f_k \rangle|^2 \\ &= \sum_{k \neq \ell} |\langle f, f_k \rangle|^2 + |\langle f, f_\ell \rangle|^2 \\ &\leq (1 + C) \sum_{k \neq \ell} |\langle f, f_k \rangle|^2. \end{aligned}$$

This shows that $\{f_k\}_{k \neq \ell}$ indeed satisfies the lower frame condition.

Denoting the canonical dual frame of $\{f_k\}_{k \neq \ell}$ by $\{g_k\}_{k \neq \ell}$ and defining $g_\ell = 0$, we have found a frame $\{g_k\}_{k=1}^{\infty}$ for which (2.25) holds; it is different from the canonical dual of $\{f_k\}_{k=1}^{\infty}$ because $S^{-1} f_\ell \neq 0$. $\square$

2.7 B-Splines

So far, we have considered frames in general Hilbert spaces. Our purpose is now to consider concrete frames in $L^2(\mathbb{R})$. We will focus on frames generated by B-splines, so in this section we recall the basic properties for these functions. For more information on B-splines (and more general splines), we refer to the books by de Boor [28] and Chui [51].

The B-splines are defined inductively: The first is simply

$$N_1(x) = \chi_{[0,1]}(x), \tag{2.26}$$

and, assuming that we have defined N_n for some $n \in \mathbb{N}$, the next is defined by a convolution:

$$\begin{aligned} N_{n+1}(x) = N_n * N_1(x) &= \int_{-\infty}^{\infty} N_n(x-t) N_1(t)\, dt \\ &= \int_0^1 N_n(x-t)\, dt. \end{aligned} \tag{2.27}$$

The functions N_n defined by (2.26) and (2.27) are called *B-splines*, and n is the *order*. See Fig. 2.1 for graphs of the first few B-splines. We collect some of their fundamental properties; all of the results can be proved by induction (Exercise 13).

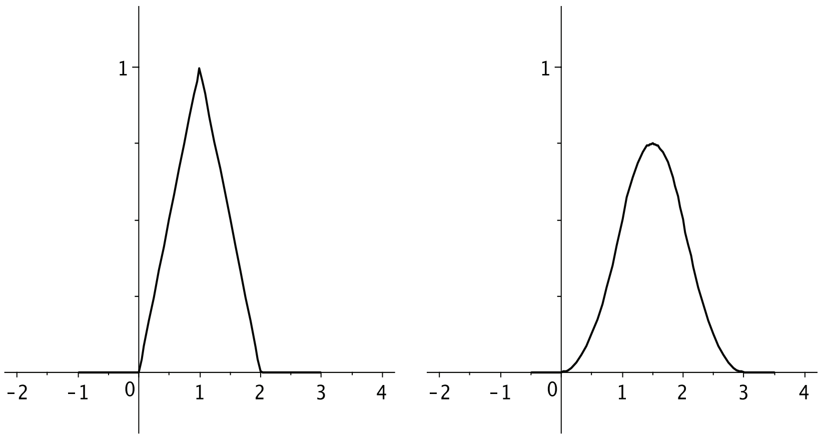


Fig. 2.1: The B-splines N_2 and N_3, respectively.

Theorem 2.22. *Given $n \in \mathbb{N}$, the B-spline N_n has the following properties:*

1. supp $N_n = [0,n]$ *and* $N_n > 0$ *on* $(0,n)$.
2. $\int_{-\infty}^{\infty} N_n(x)\, dx = 1.$

3. *For* $n \geq 2$,

$$\sum_{k\in\mathbb{Z}} N_n(x-k) = 1 \quad \text{for all } x \in \mathbb{R}; \tag{2.28}$$

for $n = 1$, *formula* (2.28) *holds for a.e.* $x \in \mathbb{R}$.

4. *For any* $n \in \mathbb{N}$,

$$\widehat{N_n}(\gamma) = \left(\frac{1-e^{-2\pi i\gamma}}{2\pi i\gamma}\right)^n. \tag{2.29}$$

We will now consider a centered version of the discussed B-splines. For $n \in \mathbb{N}$, let

$$B_n(x) := T_{-\frac{n}{2}} N_n(x) = N_n\left(x + \frac{n}{2}\right). \tag{2.30}$$

We will also call the functions B_n for B-splines. Alternatively, one can define these functions by

$$B_1 := \chi_{[-\frac{1}{2},\frac{1}{2}]}, \quad B_{n+1} := B_n * B_1, \quad n \in \mathbb{N}. \tag{2.31}$$

Thus, for any $n \in \mathbb{N}$, we have that

$$B_{n+1}(x) = \int_{-\frac{1}{2}}^{\frac{1}{2}} B_n(x-t)\,dt.$$

It is clear that B_n has support on the interval $[-n/2,\ n/2]$. We state the following direct consequences of Theorem 2.22:

Corollary 2.23. *For* $n \in \mathbb{N}$, *the B-spline* B_n *has the following properties:*

1. *For* $n \geq 2$,

$$\sum_{k\in\mathbb{Z}} B_n(x-k) = 1 \quad \text{for all } x \in \mathbb{R}.$$

For $n = 1$, *the formula holds for a.e.* $x \in \mathbb{R}$.

2. $\widehat{B_n}(\gamma) = \left(\frac{e^{\pi i\gamma}-e^{-\pi i\gamma}}{2\pi i\gamma}\right)^n = \left(\frac{\sin(\pi\gamma)}{\pi\gamma}\right)^n.$

2.8 Frames of Translates

We will now start the approach to the explicit construction of frames in $L^2(\mathbb{R})$. Our focus will be on frames having Gabor structure or wavelet structure, to be discussed in Sections 2.9–2.13; however, both of these types of systems involve translation of a fixed function, so we first give a short presentation of such systems.

For $b \in \mathbb{R}$, define the translation operator $T_b : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ by

$$(T_b f)(x) = f(x-b), \quad x \in \mathbb{R}.$$

We will consider systems of functions in $L^2(\mathbb{R})$ of the form $\{T_k\phi\}_{k\in\mathbb{Z}}$, where ϕ is a fixed function. Our goal is to state a characterization of Riesz sequences and frames of this form.

Associated with a given function $\phi \in L^2(\mathbb{R})$, we will consider the function

$$\Phi(\gamma) = \sum_{k\in\mathbb{Z}} \left|\hat{\phi}(\gamma+k)\right|^2, \quad \gamma \in \mathbb{R}. \tag{2.32}$$

We will state the announced characterization of the frame properties for $\{T_k\phi\}_{k\in\mathbb{Z}}$ in terms of the function Φ associated with ϕ. It was first proved by Benedetto and Li in [18]. We state the result without proof:

Theorem 2.24. *Let $\phi \in L^2(\mathbb{R})$. For any $A, B > 0$, the following characterizations hold:*

1. $\{T_k\phi\}_{k\in\mathbb{Z}}$ is an orthonormal sequence if and only if

$$\Phi(\gamma) = 1, \quad a.e.\ \gamma \in [0,1].$$

2. $\{T_k\phi\}_{k\in\mathbb{Z}}$ is a Riesz sequence with bounds A, B if and only if

$$A \le \Phi(\gamma) \le B, \quad a.e.\ \gamma \in [0,1].$$

3. $\{T_k\phi\}_{k\in\mathbb{Z}}$ is a frame sequence with bounds A, B if and only if

$$A \le \Phi(\gamma) \le B, \quad a.e.\ \gamma \in [0,1] \setminus N,$$

where $N = \{\gamma \in [0,1] \ : \Phi(\gamma) = 0\}$.

As a very important consequence of Theorem 2.24, we now prove that the integer-translates of any B-spline form a Riesz sequence. We formulate the result for the symmetric B-splines B_n defined in (2.31), but the same result holds for the B-splines N_n in (2.27).

Theorem 2.25. *For each $n \in \mathbb{N}$, the sequence $\{T_kB_n\}_{k\in\mathbb{Z}}$ is a Riesz sequence.*

Proof. For $n = 1$, $\{T_kB_1\}_{k\in\mathbb{Z}}$ is an orthonormal system, and therefore a Riesz sequence. In order to prove the result for $n > 1$, we apply Theorem 2.24 to B_1; this shows that

$$\sum_{k\in\mathbb{Z}} \left|\widehat{B_1}(\gamma+k)\right|^2 = 1, \quad \text{a.e. } \gamma \in \mathbb{R}.$$

Since $|\widehat{B_1}(\gamma)| \le 1$ for all $\gamma \in \mathbb{R}$ and $\widehat{B_n}(\gamma) = (\widehat{B_1}(\gamma))^n$ by Corollary 2.23, it immediately follows that

$$\sum_{k\in\mathbb{Z}} \left|\widehat{B_n}(\gamma+k)\right|^2 \le \sum_{k\in\mathbb{Z}} \left|\widehat{B_1}(\gamma+k)\right|^2 = 1, \quad \text{a.e. } \gamma \in \mathbb{R}.$$

Thus, $\{T_kB_n\}_{k\in\mathbb{Z}}$ is a Bessel sequence. In order to prove that $\{T_kB_n\}_{k\in\mathbb{Z}}$ satisfies the lower Riesz basis condition, we again use Corollary 2.23: It shows that, for a.e. $\gamma \in \mathbb{R}$,

$$\sum_{k\in\mathbb{Z}} \left|\widehat{B_n}(\gamma+k)\right|^2 \geq \inf_{\gamma\in[-\frac{1}{2},\frac{1}{2}]} \left|\widehat{B_n}(\gamma)\right|^2 = \left(\frac{\sin(\pi/2)}{\pi/2}\right)^{2n} = \left(\frac{2}{\pi}\right)^{2n}. \qquad (2.33)$$

The result now follows from Theorem 2.24. □

2.9 Basic Gabor Frame Theory

We are now ready to approach the analysis of Gabor systems. Consider for $b \in \mathbb{R}$ the *modulation operator*

$$E_b : L^2(\mathbb{R}) \to L^2(\mathbb{R}), \qquad (E_b f)(x) = e^{2\pi ibx} f(x).$$

A collection of functions on the form $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$ is called a *Gabor system.* Explicitly, these functions have the form

$$E_{mb}T_{na}g(x) = e^{2\pi imbx} g(x-na).$$

Systems of the Gabor type play a role in time-frequency analysis. In this chapter we will focus on the frame properties for Gabor systems, in particular, for the case where g is a B-spline. For a broader view on Gabor systems, we refer to the book by Gröchenig [108], as well as the collections of research papers in the books [86] and [87] edited by Feichtinger and Strohmer.

Our purpose is to consider frames for $L^2(\mathbb{R})$ having the Gabor structure:

Definition 2.26. A *Gabor frame* is a frame for $L^2(\mathbb{R})$ of the form $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$, where $a,b > 0$ are given and $g \in L^2(\mathbb{R})$ is a fixed function.

Frames of this type are also called *Weyl–Heisenberg frames*. The function g is called the *window function* or the *generator.*

Gabor systems play a role in the context of *time-frequency analysis.* Although bases of the Gabor type exist (take, e.g., $a = b = 1$ and $g = \chi_{[0,1]}$), they are not well suited for the purpose of the time-frequency analysis of functions. One reason is that the generator g cannot be particularly nice; for example, it is known that no continuous and compactly supported function g can generate a Gabor basis, regardless of the choice of the parameters a and b. Another reason is that it is impossible to have a Gabor basis generated by a function g with fast decay in the time domain *and* the frequency domain. This is the famous *Balian–Low theorem:*

Theorem 2.27. *Let* $g \in L^2(\mathbb{R})$. *If* $\{E_mT_ng\}_{m,n\in\mathbb{Z}}$ *is a Riesz basis for* $L^2(\mathbb{R})$, *then*

$$\left(\int_{-\infty}^{\infty} |xg(x)|^2\,dx\right)\left(\int_{-\infty}^{\infty} |\gamma\hat{g}(\gamma)|^2\,d\gamma\right) = \infty. \qquad (2.34)$$

The Balian–Low theorem implies that if $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$ is a Gabor–Riesz basis, then it is not possible that g and $\hat{g}$ satisfy decay conditions like

$$|g(x)| \leq \frac{C}{1+x^2}, \quad x \in \mathbb{R}, \quad |\hat{g}(\gamma)| \leq \frac{C}{1+\gamma^2}, \quad \gamma \in \mathbb{R},$$

simultaneously. The reader can find a proof of the Balian–Low theorem in [59].

This discussion motivates the analysis and construction of Gabor frames: As we have seen, frames lead to expansions that are somewhat similar to the expansions obtained via bases, at least if we restrict our attention to tight frames or to frames for which convenient duals can be found. We will present such constructions in the next sections.

We now state a proposition that gives a necessary condition for $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$ to be a frame for $L^2(\mathbb{R})$. It depends on the interplay between the function g and the translation parameter a and is expressed in terms of the function

$$G(x) := \sum_{n\in\mathbb{Z}} |g(x-na)|^2, \quad x \in \mathbb{R}. \tag{2.35}$$

The proof can be found, e.g., in [135] or [49].

Proposition 2.28. *Let $g \in L^2(\mathbb{R})$ and $a,b > 0$ be given, and assume that the collection of functions $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$ is a frame with bounds A,B. Then*

$$bA \leq \sum_{n\in\mathbb{Z}} |g(x-na)|^2 \leq bB, \quad a.e.\ x \in \mathbb{R}. \tag{2.36}$$

More precisely: If the upper bound in (2.36) *is violated, then $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$ is not a Bessel sequence; if the lower bound is violated, then $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$ does not satisfy the lower frame condition.*

It follows from Proposition 2.28 that a function g generating a Gabor frame $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$ necessarily is bounded. Note also that Proposition 2.28 gives a relationship between the frame bounds and the lower and upper bounds for the function G in (2.35). In Corollary 2.32 we will see that under certain circumstances, the necessary condition (2.36) is also sufficient for $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$ to be a frame for $L^2(\mathbb{R})$.

If we want to check that a Gabor system $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$ forms a frame by hand, we need to be able to estimate the expression $\sum_{m,n\in\mathbb{Z}} |\langle f, E_{mb}T_{na}g\rangle|^2$ for all functions f belonging to $L^2(\mathbb{R})$ (or at least a dense subset thereof). Under certain conditions on the functions f and g, we can find an explicit expression for this infinite sum. The next statement is taken from [44], but we notice that similar results already appear in [59].

Lemma 2.29. *Suppose that f is a bounded, measurable function with compact support and that the function G defined by* (2.35) *is bounded. Then*

$$\begin{aligned}\sum_{m,n\in\mathbb{Z}} |\langle f, E_{mb}T_{na}g\rangle|^2 &= \frac{1}{b}\int_{-\infty}^{\infty} |f(x)|^2 \sum_{n\in\mathbb{Z}} |g(x-na)|^2\,dx \\ &\quad + \frac{1}{b}\sum_{k\neq 0}\int_{-\infty}^{\infty} \overline{f(x)} f(x-k/b) \sum_{n\in\mathbb{Z}} g(x-na)\overline{g(x-na-k/b)}\,dx.\end{aligned}$$

Lemma 2.29 has several important consequences. For example, as shown in [44], it leads to a sufficient condition for $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$ to form a frame (the original proof can be found in [207]):

Theorem 2.30. *Let* $g \in L^2(\mathbb{R}),\ a,b>0$, *and suppose that*

$$B := \frac{1}{b}\sup_{x\in[0,a]}\left\{\sum_{k\in\mathbb{Z}}\left|\sum_{n\in\mathbb{Z}} g(x-na)\overline{g(x-na-k/b)}\right|\right\} < \infty. \tag{2.37}$$

Then $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$ *is a Bessel sequence with bound B. If also*

$$A := \frac{1}{b}\inf_{x\in[0,a]}\left\{\sum_{n\in\mathbb{Z}}|g(x-na)|^2 - \sum_{k\neq 0}\left|\sum_{n\in\mathbb{Z}} g(x-na)\overline{g(x-na-k/b)}\right|\right\} \tag{2.38}$$
$$> 0,$$

then $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$ *is a frame for* $L^2(\mathbb{R})$ *with bounds* A,B.

Condition (2.37) leads to an easy, sufficient condition for $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$ to be a Bessel sequence (Exercise 18):

Corollary 2.31. *Let* $g \in L^2(\mathbb{R})$ *be bounded and compactly supported. Then the collection of functions* $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$ *is a Bessel sequence for any choice of* $a,b>0$.

The condition that the function g is bounded and compactly supported is not sufficient for $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$ to be a frame: In fact, as shown in Proposition 2.28, the associated function G in (2.35) also needs to be bounded below and above. On the other hand, for a function g with compact support, the condition that the function G is bounded below and above for some $a>0$ is enough for $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$ to be a frame for sufficiently small values of b. We also obtain expressions for the frame operator and its inverse in this case:

Corollary 2.32. *Let* $a,b>0$ *be given. Suppose that* $g\in L^2(\mathbb{R})$ *has support in an interval of length* $1/b$ *and that the function G satisfies* (2.36) *for some* $A,B>0$. *Then* $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$ *is a frame for* $L^2(\mathbb{R})$ *with bounds* A,B. *The frame operator S and its inverse* S^{-1} *are given by*

$$Sf = \frac{G}{b}f, \quad S^{-1}f = \frac{b}{G}f, \quad f\in L^2(\mathbb{R}).$$

Proof. That $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$ is a frame follows directly from Lemma 2.29 or Theorem 2.30 because

$$\sum_{n\in\mathbb{Z}} g(x-na)\overline{g(x-na-k/b)} = 0 \quad \text{for all } k\neq 0.$$

Given a continuous function f with compact support, Lemma 2.29 implies that

$$\langle Sf,f\rangle = \sum_{m,n\in\mathbb{Z}} |\langle f,E_{mb}T_{na}g\rangle|^2 = \frac{1}{b}\int_{-\infty}^{\infty} |f(x)|^2 G(x)\,dx = \left\langle \frac{G}{b}f,f \right\rangle.$$

By the continuity of S, this expression even holds for all $f \in L^2(\mathbb{R})$. It follows that S acts by multiplication with the function G/b. □

For a continuous function g, we can be even more explicit. We leave the proof of the following result to the reader (Exercise 22).

Corollary 2.33. *Suppose that $g \in L^2(\mathbb{R})$ is a continuous function with support on an interval I with length $|I|$ and that $g(x) > 0$ on the interior of I. Then $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$ is a frame for all $(a,b) \in (0,|I|) \times (0,1/|I|]$.*

In particular, this result applies to the B-splines. In order to avoid a conflict with our notation for a Gabor system, we will denote the splines by B_ℓ and N_ℓ, $\ell \in \mathbb{N}$, instead of B_n and N_n.

Corollary 2.34. *For $\ell \in \mathbb{N}$, the B-splines B_ℓ and N_ℓ generate Gabor frames for all $(a,b) \in (0,\ell) \times (0,1/\ell]$.*

One might wonder whether the Gabor system $\{E_{mb}T_{na}B_\ell\}_{m,n\in\mathbb{Z}}$ is a frame for $(a,b) \notin (0,\ell) \times (0,1/\ell]$. Surprisingly, only a few partial results are known. For the B-spline B_2, it is proved in [117] that if $b \in \mathbb{N} \setminus \{1\}$, the Gabor system cannot form a frame for any $a > 0$; see also Exercise 19.

The results discussed so far concentrate on the interplay between the function g and the parameters a,b. For completeness, we mention a central result in Gabor analysis, although it does not have a direct influence on the result presented here: It shows that, regardless of the choice of generator $g \in L^2(\mathbb{R})$, the choice of the parameters a and b puts certain restrictions on the possible frame properties for $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$.

Theorem 2.35. *Let $g \in L^2(\mathbb{R})$ and $a,b > 0$ be given. Then the following hold:*

1. *If $ab > 1$, then $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$ cannot be a frame for $L^2(\mathbb{R})$.*
2. *If $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$ is a frame, then*

$$ab = 1 \quad \Leftrightarrow \quad \{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}} \text{ is a Riesz basis.} \tag{2.39}$$

2.10 Tight Gabor Frames

In applications of frames, it is inconvenient that the frame decomposition, stated in Theorem 2.17, requires inversion of the frame operator. As we have seen in the discussion of general frame theory, one way of avoiding the problem is to consider tight frames. We will now characterize tight Gabor frames; the first result is taken from [45].

Theorem 2.36. *Let $g \in L^2(\mathbb{R})$ and $a,b > 0$ be given. Then the following are equivalent:*

1. *$\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$ is a tight frame for $L^2(\mathbb{R})$ with frame bound $A = 1$.*
2. *For a.e. $x \in \mathbb{R}$, the following conditions hold:*

   a. $G(x) := \sum_{n\in\mathbb{Z}} |g(x-na)|^2 = b$;

   b. $G_k(x) := \sum_{n\in\mathbb{Z}} g(x-na)\overline{g(x-na-k/b)} = 0$ for all $k \neq 0$.

Proof. $1 \Rightarrow 2$: Assume that $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$ is a tight frame for $L^2(\mathbb{R})$ with frame bound $A = 1$. Then Proposition 2.28 shows that $G(x) = b$ for a.e. $x \in \mathbb{R}$. Therefore,

$$\sum_{m,n\in\mathbb{Z}} |\langle f, E_{mb}T_{na}g\rangle|^2 = \frac{1}{b}\int_{-\infty}^{\infty} |f(x)|^2 G(x)\, dx$$

for all functions $f \in L^2(\mathbb{R})$. Using Lemma 2.29, we conclude that for all bounded, compactly supported $f \in L^2(\mathbb{R})$,

$$\frac{1}{b}\sum_{k\neq 0}\int_{-\infty}^{\infty} \overline{f(x)} f(x-k/b) \sum_{n\in\mathbb{Z}} g(x-na)\overline{g(x-na-k/b)}dx = 0.$$

A change of variable shows that the contribution in the above sum arising from any value of $k \in \mathbb{Z}$ is the complex conjugate of the contribution from the value $-k$. Therefore,

$$\sum_{k=1}^{\infty} \mathrm{Re}\left(\int_{-\infty}^{\infty} \overline{f(x)} f(x-k/b) \sum_{n\in\mathbb{Z}} g(x-na)\overline{g(x-na-k/b)}dx\right) = 0. \tag{2.40}$$

Now fix $k_0 \geq 1$ and let I be any interval in $\mathbb{R}$ of length at most $1/b$. Define a function $f \in L^2(\mathbb{R})$ by

$$f(x) = \begin{cases} e^{-i\arg(G_{k_0}(x))} & \text{for } x \in I, \\ 1 & \text{for } x \in I + k_0/b, \\ 0 & \text{otherwise.} \end{cases}$$

Then, by (2.40),

$$\begin{aligned} 0 &= \sum_{k=1}^{\infty} \mathrm{Re}\left(\int_{-\infty}^{\infty} \overline{f(x)} f(x-k/b) \sum_{n\in\mathbb{Z}} g(x-na)\overline{g(x-na-k/b)}\, dx\right) \\ &= \mathrm{Re}\left(\int_{-\infty}^{\infty} \overline{f(x)} f(x-k_0/b) G_{k_0}(x)\, dx\right) = \int_I |G_{k_0}(x)|\, dx. \end{aligned}$$

It follows that $G_{k_0}(x) = 0$ for a.e. $x \in I$. Since I was an arbitrary interval of length at most $1/b$, we conclude that $G_{k_0} = 0$. In order to deal with G_k for $k < 0$, a direct computation shows that

$$G_{-k_0}(x) = \overline{G_{k_0}(x + k_0/b)} = 0;$$

this shows that statement $2b$ indeed holds for all $k \neq 0$.

$2 \Rightarrow 1$: The assumptions in 2 imply, again by Lemma 2.29, that for all bounded, compactly supported functions $f \in L^2(\mathbb{R})$,

$$\sum_{m,n\in\mathbb{Z}} |\langle f, E_{mb}T_{na}g\rangle|^2 = \frac{1}{b}\int_{-\infty}^{\infty} |f(x)|^2 \sum_{n\in\mathbb{Z}} |g(x-na)|^2 \, dx = \|f\|^2.$$

Since the bounded, compactly supported functions are dense in $L^2(\mathbb{R})$, Lemma 2.14 implies that $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$ is a tight frame with frame bound $A = 1$, as desired. □

In general, it is not easy to construct functions g such that the conditions in Theorem 2.36 (2) are satisfied for some given $a, b > 0$. A simplification occurs if we assume that g has compact support: In that case, condition $2b$ is automatically satisfied for sufficiently small values of the parameter b. In particular, we obtain the following very useful sufficient condition for $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$ being a tight Gabor frame. We ask the reader to provide the proof in Exercise 23.

Corollary 2.37. *Let $a, b > 0$ be given. Assume that $\varphi \in L^2(\mathbb{R})$ is a real-valued, nonnegative function with support in an interval of length $1/b$, and that*

$$\sum_{n\in\mathbb{Z}} \varphi(x+na) = 1, \quad a.e.\ x \in \mathbb{R}. \tag{2.41}$$

Then the function

$$g(x) := \sqrt{b\varphi(x)}$$

generates a tight Gabor frame $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$ with frame bound $A = 1$.

If (2.41) is satisfied, we say that the functions $\{T_{na}\varphi\}_{n\in\mathbb{Z}}$ form a *partition of unity.* In particular, we can apply the result to B-splines:

Example 2.38. For any $\ell \in \mathbb{N}$, the B-spline $\varphi = N_\ell$ defined in (2.26) satisfies the requirements in Corollary 2.37 with $a = 1$ and any $b \in (0, 1/\ell]$. Thus, for any $b \in (0, 1/\ell]$, the function

$$g(x) = \sqrt{bN_\ell(x)}$$

generates a tight Gabor frame $\{E_{mb}T_n g\}_{m,n\in\mathbb{Z}}$ with frame bound $A = 1$.

We note that the frame generators in Example 2.38 are very suitable for time-frequency analysis: They are given by an explicit formula, have compact support, and can be chosen with polynomial decay of any desired order in the frequency domain, simply by taking the parameter ℓ sufficiently large.

2.11 The Duals of a Gabor Frame

For a Gabor frame $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$ with associated frame operator S, the frame decomposition (see Theorem 2.17) shows that

$$f = \sum_{m,n\in\mathbb{Z}} \langle f, S^{-1}E_{mb}T_{na}g\rangle E_{mb}T_{na}g, \quad \forall f \in L^2(\mathbb{R}). \tag{2.42}$$

In order to use the frame decomposition, we need to be able to calculate the canonical dual frame $\{S^{-1}E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$. This is usually difficult. Via the following lemma, we will be able to obtain a simplification; for the proof, we refer, to [49].

Lemma 2.39. *Let $g \in L^2(\mathbb{R})$ and $a,b > 0$ be given, and assume that $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$ is a Bessel sequence with frame operator S. Then the following hold:*

1. *$SE_{mb}T_{na} = E_{mb}T_{na}S$ for all $m,n \in \mathbb{Z}$.*
2. *If $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$ is a frame, then*

$$S^{-1}E_{mb}T_{na} = E_{mb}T_{na}S^{-1}, \quad \forall m,n \in \mathbb{Z}.$$

Lemma 2.39 has important consequences for the structure of the canonical dual frame of a Gabor frame:

Theorem 2.40. *Let $g \in L^2(\mathbb{R})$ and $a,b > 0$ be given, and assume that the collection of functions $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$ is a Gabor frame. Then the canonical dual frame also has Gabor structure and is given by $\{E_{mb}T_{na}S^{-1}g\}_{m,n\in\mathbb{Z}}$.*

Via Theorem 2.40, the frame decomposition (2.42) associated with a Gabor frame $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$ takes the form

$$f = \sum_{m,n\in\mathbb{Z}} \langle f, E_{mb}T_{na}S^{-1}g\rangle E_{mb}T_{na}g, \quad \forall f \in L^2(\mathbb{R}). \tag{2.43}$$

In practice, this version of the frame decomposition is much more convenient than (2.42): Instead of calculating the *double infinite* family $\{S^{-1}E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$, it is enough to find $S^{-1}g$ and then apply the modulation and translation operators. The function $S^{-1}g$ is called the *dual window function* or the *dual generator.*

We will now leave the discussion of the canonical dual frame and examine the question of how general dual frames of a given Gabor frame $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$ can be found. Our analysis is based on a fundamental result due to Ron and Shen [207], respectively, Janssen [146]; the technical proof can be found in the original papers or in [49].

Theorem 2.41. *Let $g,h \in L^2(\mathbb{R})$ and $a,b > 0$ be given. Two Bessel sequences $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$ and $\{E_{mb}T_{na}h\}_{m,n\in\mathbb{Z}}$ form dual frames if and only if*

$$\sum_{k\in\mathbb{Z}} \overline{g(x-n/b-ka)}h(x-ka) = b\delta_{n,0}, \quad a.e.\ x \in [0,a]. \tag{2.44}$$

2.12 Explicit Construction of Dual Gabor Frame Pairs

So far, we have only seen a few examples of Gabor frames and their dual frames. After the preparation in Section 2.11, we are now ready to provide explicit constructions of certain Gabor frames and some particularly convenient duals. The assumptions are tailored to the properties of the B-splines. The results presented here first appeared in [48]. For convenience, we first consider the case where the translation parameter is $a = 1$:

Theorem 2.42. *Let $N \in \mathbb{N}$. Let $g \in L^2(\mathbb{R})$ be a real-valued, bounded function with* $\operatorname{supp} g \subseteq [0,N]$, *for which*

$$\sum_{k\in\mathbb{Z}} g(x-k) = 1, \quad x \in \mathbb{R}. \tag{2.45}$$

Let $b \in (0, 1/(2N-1)]$. Then the function g and the function h defined by

$$h(x) = bg(x) + 2b\sum_{k=1}^{N-1} g(x+k) \tag{2.46}$$

generate dual frames $\{E_{mb}T_n g\}_{m,n\in\mathbb{Z}}$ and $\{E_{mb}T_n h\}_{m,n\in\mathbb{Z}}$ for $L^2(\mathbb{R})$.

Proof. By assumption, the function g has compact support and is bounded; by definition (2.46), the function h shares these properties. It now follows from Corollary 2.31 that $\{E_{mb}T_n g\}_{m,n\in\mathbb{Z}}$ and $\{E_{mb}T_n h\}_{m,n\in\mathbb{Z}}$ are Bessel sequences. In order to verify that these sequences form dual frames, we use Theorem 2.41: According to (2.44), we need to check that for $x \in [0,1]$,

$$\sum_{k\in\mathbb{Z}} g(x-n/b-k)h(x-k) = b\delta_{n,0}. \tag{2.47}$$

The function g has support in $[0,N]$, so by construction h has support in $[-N+1,N]$; thus, (2.47) is satisfied for $n \neq 0$ whenever $1/b \geq 2N-1$, i.e., if $b \in (0, 1/(2N-1)]$. For $n = 0$, condition (2.47) means that

$$\sum_{k\in\mathbb{Z}} g(x-k)h(x-k) = b, \quad x \in [0,1];$$

because of the compact support of g, this is equivalent to

$$\sum_{k=0}^{N-1} g(x+k)h(x+k) = b, \quad x \in [0,1]. \tag{2.48}$$

Condition (2.48) is indeed satisfied in our setting. To see this, we use that for $x \in [0,1]$,

$$1 = \sum_{k=0}^{N-1} g(x+k).$$

This implies that, again for $x \in [0,1]$,

$$\begin{aligned} 1 &= \left(\sum_{k=0}^{N-1} g(x+k)\right)^2 \\ &= (g(x)+g(x+1)+\cdots+g(x+N-1)) \\ &\quad \times (g(x)+g(x+1)+\cdots+g(x+N-1)) \end{aligned}$$

$$
\begin{aligned}
&= g(x)\,[g(x)+2g(x+1)+2g(x+2)+\cdots+2g(x+N-1)]\\
&\quad + g(x+1)\,[g(x+1)+2g(x+2)+2g(x+3)+\cdots+2g(x+N-1)]\\
&\quad + g(x+2)\,[g(x+2)+2g(x+3)+2g(x+4)+\cdots+2g(x+N-1)]\\
&\quad + \cdots\\
&\quad + \cdots\\
&\quad + g(x+N-2)\,[g(x+N-2)+2g(x+N-1)]\\
&\quad + g(x+N-1)\,[g(x+N-1)]\\
&= \frac{1}{b}\sum_{k=0}^{N-1} g(x+k)h(x+k).
\end{aligned}
$$

Thus, condition (2.48) is satisfied. □

The assumptions in Theorem 2.42 are tailored to the properties of the B-splines N_ℓ defined in (2.26):

Corollary 2.43. *For any $\ell \in \mathbb{N}$ and $b \in (0, 1/(2\ell-1)]$, the functions N_ℓ and*

$$
h_\ell(x) := bN_\ell(x) + 2b\sum_{k=1}^{\ell-1} N_\ell(x+k) \tag{2.49}
$$

generate dual frames $\{E_{mb}T_nN_\ell\}_{m,n\in\mathbb{Z}}$ and $\{E_{mb}T_nh_\ell\}_{m,n\in\mathbb{Z}}$ for $L^2(\mathbb{R})$.

Some of the important features of the dual pair of frame generators (N_ℓ, h_ℓ) in Corollary 2.43 are as follows:

1. The functions N_ℓ and h_ℓ are splines for all choices of $\ell \in \mathbb{N}$;
2. N_ℓ and h_ℓ are explicitly given functions with compact support, i.e., they have perfect time-localization;
3. By choosing $\ell \in \mathbb{N}$ sufficiently large, polynomial decay of $\widehat{N_\ell}$ and $\widehat{h_\ell}$ of any desired order can be obtained.

Example 2.44. For the B-spline

$$
N_2(x) = \begin{cases} x, & x \in [0,1),\\ 2-x, & x \in [1,2),\\ 0, & x \notin [0,2), \end{cases}
$$

we can use Corollary 2.43 for $b \in (0, 1/3]$. For $b = 1/3$, we obtain the dual generator

$$
h_2(x) = \frac{1}{3}N_2(x) + \frac{2}{3}N_2(x+1) = \begin{cases} \frac{2}{3}(x+1), & x \in [-1,0),\\ \frac{1}{3}(2-x), & x \in [0,2),\\ 0, & x \notin [-1,2). \end{cases} \tag{2.50}
$$

See Fig. 2.2, which also shows a similar construction based on N_3.

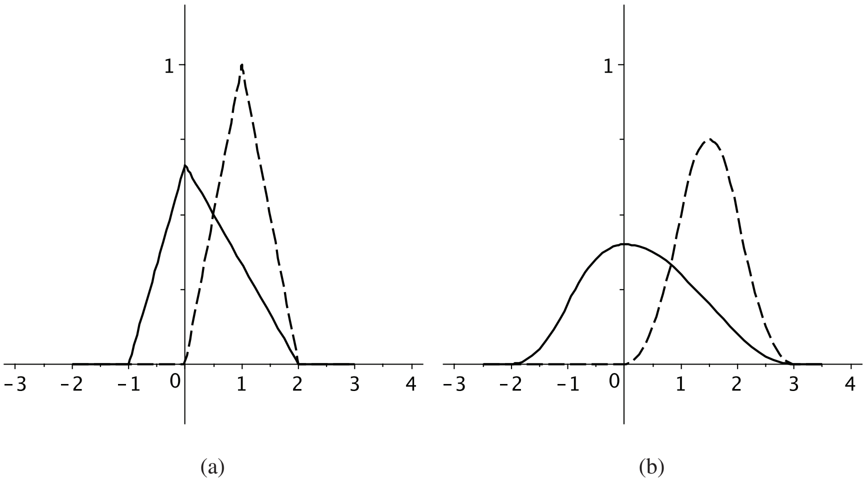


Fig. 2.2: (a) The B-spline N_2 and the dual generator h_2 in (2.50); (b) the B-spline N_3 and the dual generator h_3 in (2.46) with $b = 1/5$.

Via a scaling, one can obtain a version of Theorem 2.42 that is valid for any translation parameter $a > 0$; see [48] for details.

The reader will notice that even though the B-splines N_ℓ are symmetric, the constructed dual generators are not. We state without proof a recent result, due to Christensen and Kim [50], that gives freedom to choose various duals:

Theorem 2.45. *Let $N \in \mathbb{N}$. Let $g \in L^2(\mathbb{R})$ be a real-valued, bounded function with* supp $g \subset [0,N]$, *for which*

$$\sum_{n\in\mathbb{Z}} g(x-n) = 1.$$

Let $b \in (0, 1/(2N-1)]$. Consider any scalar sequence $\{a_n\}_{n=-N+1}^{N-1}$ for which

$$a_0 = b \ \text{ and } \ a_n + a_{-n} = 2b, \ n = 1,2,\ldots,N-1, \tag{2.51}$$

and define $h \in L^2(\mathbb{R})$ by

$$h(x) = \sum_{n=-N+1}^{N-1} a_n g(x+n). \tag{2.52}$$

Then g and h generate dual frames $\{E_{mb}T_n g\}_{m,n\in\mathbb{Z}}$ and $\{E_{mb}T_n h\}_{m,n\in\mathbb{Z}}$ for $L^2(\mathbb{R})$.

In particular, if the generator g is symmetric, it is possible to construct a symmetric dual generator:

Corollary 2.46. *Under the assumptions in Theorem 2.45, the function*

$$h(x) = b \sum_{n=-N+1}^{N-1} g(x+n) \tag{2.53}$$

generates a dual frame of $\{E_{mb}T_n g\}_{m,n\in\mathbb{Z}}$. *The function* h *satisfies that* $h = b$ *on the support of* g. *Furthermore, if* g *is symmetric, then* h *is symmetric.*

See Fig. 2.3 for an illustration of the dual generators, based on the B-splines N_2 and N_3.

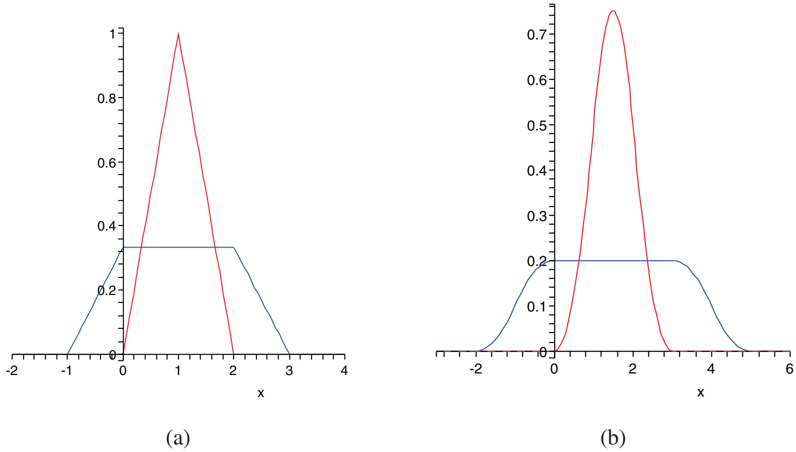


Fig. 2.3: (a) The B-spline N_2 and the dual generator h_2 in (2.53) for $b = 1/3$. (b) The B-spline N_3 and the dual generator h_3 in (2.53) for $b = 1/5$.

2.13 Wavelets and the Unitary Extension Principle

Classical wavelet theory deals with the construction of orthonormal bases for $L^2(\mathbb{R})$ of the form $\{2^{j/2}\psi(2^j x - k)\}_{j,k\in\mathbb{Z}}$ for a suitable function $\psi \in L^2(\mathbb{R})$. Introducing the dilation operator

$$D : L^2(\mathbb{R}) \to L^2(\mathbb{R}), \qquad (Df)(x) = 2^{1/2} f(2x),$$

the *wavelet system* takes the form

$$\{2^{j/2}\psi(2^j x - k)\}_{j,k\in\mathbb{Z}} = \{D^j T_k \psi\}_{j,k\in\mathbb{Z}}.$$

Most wavelet constructions are based on the so-called multiresolution analysis, invented by Mallat in 1989; soon after that, Daubechies [59] presented her famous construction of compactly supported wavelets. We will not go into a discussion of all the aspects of classical multiresolution analysis, but refer to Daubechies' book [60]. Our focus will be on a more recent construction of tight wavelet frames, based on B-splines.

It the context of B-splines, Battle and Lemarié have proven that for any (centered) B-spline B_m, it is possible to construct an orthonormal basis $\{D^j T_k \psi\}_{j,k\in\mathbb{Z}}$ for $L^2(\mathbb{R})$

for a function ψ of the form

$$\psi(x) = \sum_{k\in\mathbb{Z}} c_k B_m(2x-k). \tag{2.54}$$

However, except for the case of the first B-spline B_1, the sequence $\{c_k\}_{k\in\mathbb{Z}}$ is infinite; and one can prove that no orthonormal basis construction for a function ψ of the form (2.54) with a finite sequence $\{c_k\}_{k\in\mathbb{Z}}$ is possible. This motivates the result presented in the current section: In fact, we will construct tight frames of wavelet structure, but based on *two* (or more) generators of the form (2.54).

Our aim is to state the unitary extension principle of Ron and Shen [206], which enables us to construct tight frames for $L^2(\mathbb{R})$ of the form $\{D^j T_k \psi_\ell\}_{j,k\in\mathbb{Z},\ell=1,\ldots,n}$; after doing so, we will show how to construct frames based on B-splines. We follow the approach by Benedetto and Treiber [20].

The following proofs are based on standard Fourier analysis for 1-periodic functions. It will be convenient to write the integrals appearing, e.g., in the expression for the Fourier coefficients and in Parseval's equation, as integrals over the interval $(-1/2,1/2)$ rather than $(0,1)$. The interval $(-1/2,1/2)$ is identified with the torus $\mathbb{T}$, and the class of 1-periodic functions on $\mathbb{R}$ whose restriction to $(-1/2,1/2)$ belongs to $L^p(-1/2,1/2)$, $p=1,2$, is denoted by $L^p(\mathbb{T})$. Similarly, $L^\infty(\mathbb{T})$ consists of the bounded, measurable 1-periodic functions on $\mathbb{R}$. With this notation, $L^\infty(\mathbb{T}) \subset L^2(\mathbb{T})$. We note that the spaces $L^p(\mathbb{T})$ actually consist of equivalence classes of functions that are identical almost everywhere, so when we speak about pointwise relationships between functions, it is understood that they can only be expected to hold almost everywhere.

We now list the standing assumptions and conventions for this section.

General setup: Let $\psi_0 \in L^2(\mathbb{R})$ and assume that

1. There exists a function $H_0 \in L^\infty(\mathbb{T})$ such that

$$\widehat{\psi_0}(2\gamma) = H_0(\gamma)\widehat{\psi_0}(\gamma). \tag{2.55}$$

2. $\lim_{\gamma\to 0} \widehat{\psi_0}(\gamma) = 1$.

Further, let $H_1,\ldots,H_n \in L^\infty(\mathbb{T})$, and define $\psi_1,\ldots,\psi_n \in L^2(\mathbb{R})$ by

$$\widehat{\psi_\ell}(2\gamma) = H_\ell(\gamma)\widehat{\psi_0}(\gamma), \quad \ell = 1,\ldots,n. \tag{2.56}$$

Finally, let H denote the $(n+1)\times 2$ matrix-valued function defined by

$$H(\gamma) = \begin{pmatrix} H_0(\gamma) & T_{1/2}H_0(\gamma) \\ H_1(\gamma) & T_{1/2}H_1(\gamma) \\ \cdot & \cdot \\ \cdot & \cdot \\ H_n(\gamma) & T_{1/2}H_n(\gamma) \end{pmatrix}, \quad \gamma \in \mathbb{R}. \tag{2.57}$$

With this setup, our purpose is to find conditions on the functions $H_1,\ldots,H_n$ such that $\psi_1,\ldots,\psi_n$ defined by (2.56) generate a multiwavelet frame for $L^2(\mathbb{R})$. It turns out to be convenient to formulate the results in terms of the matrices $H(\gamma),\ \gamma \in \mathbb{R}$. Note that if we know the functions H_ℓ, then we can find an explicit expression for the functions ψ_ℓ: In fact, expanding H_ℓ in a Fourier series, $H_\ell(\gamma) = \sum_{k\in\mathbb{Z}} c_{k,\ell} e^{2\pi i k\gamma}$, elementary manipulations with the Fourier transform show that

$$\psi_\ell(x) = \sqrt{2}\sum_{k\in\mathbb{Z}} c_{k,\ell} D T_{-k}\psi_0(x) = 2\sum_{k\in\mathbb{Z}} c_{k,\ell}\psi_0(2x+k). \tag{2.58}$$

Recall that we prefer the functions H_ℓ to be trigonometric polynomials: This implies that the sums in (2.58) are finite and therefore that the functions ψ_ℓ have compact support if ψ_0 has compact support.

We are now ready to formulate the *unitary extension principle;* the quite complicated proof can be found in the original paper [206] by Ron and Shen, in the paper [20] by Benedetto and Treiber, or in [49].

Theorem 2.47. *Let* $\{\psi_\ell, H_\ell\}_{\ell=0}^n$ *be as in the general setup above, and assume that* $H(\gamma)^* H(\gamma) = I$ *for a.e.* $\gamma \in \mathbb{T}$. *Then the multiwavelet system* $\{D^j T_k \psi_\ell\}_{j,k\in\mathbb{Z},\ell=1,\dots,n}$ *constitutes a tight frame for* $L^2(\mathbb{R})$ *with frame bound equal to* 1, *and*

$$f = \sum_{\ell=1}^{n}\sum_{j\in\mathbb{Z}}\sum_{k\in\mathbb{Z}} \langle f, D^j T_k \psi_\ell\rangle D^j T_k \psi_\ell, \quad \forall f \in L^2(\mathbb{R}). \tag{2.59}$$

The matrix $H(\gamma)^* H(\gamma)$ has four entries, so at first glance it seems that we have to solve four scalar equations in order to apply Theorem 2.47. However, it turns out that it is enough to verify two sets of equations (Exercise 24):

Corollary 2.48. *Let* $\{\psi_\ell, H_\ell\}_{\ell=0}^n$ *be as in the general setup given above, and assume that*

$$\begin{cases} \displaystyle\sum_{\ell=0}^{n} |H_\ell(\gamma)|^2 = 1, \\ \displaystyle\sum_{\ell=0}^{n} \overline{H_\ell(\gamma)} T_{1/2} H_\ell(\gamma) = 0, \end{cases} \tag{2.60}$$

for a.e. $\gamma \in \mathbb{T}$. *Then the multiwavelet system* $\{D^j T_k \psi_\ell\}_{j,k\in\mathbb{Z},\ell=1,\dots,n}$ *constitutes a tight frame for* $L^2(\mathbb{R})$ *with frame bound equal to* 1.

As an application of Corollary 2.48, we show how one can construct compactly supported, tight multiwavelet frames based on B-splines. In contrast with the Battle–Lemarié wavelets, the generators will be *finite* linear combinations of the type (2.54), and thus have compact support.

Example 2.49. For any $m = 1, 2, \dots$, we consider the B-spline

$$\psi_0 := B_{2m}$$

of order $2m$ as defined in (2.31). By Corollary 2.23,

$$\widehat{\psi_0}(\gamma) = \left(\frac{\sin(\pi\gamma)}{\pi\gamma}\right)^{2m}.$$

It is clear that $\lim_{\gamma\to 0}\widehat{\psi_0}(\gamma)=1$. Furthermore, the result in Exercise 17 shows that

$$\widehat{\psi_0}(2\gamma)=\cos^{2m}(\pi\gamma)\widehat{\psi_0}(\gamma).$$

Thus, ψ_0 satisfies a refinement equation with the two-scale symbol

$$H_0(\gamma)=\cos^{2m}(\pi\gamma). \tag{2.61}$$

Note that Exercise 17 also explains why we are restricting ourselves to the case of even-order B-splines. Now, consider the binomial coefficient

$$\binom{2m}{\ell}:=\frac{(2m)!}{(2m-\ell)!\ell!},$$

and define the functions $H_1,\dots,H_{2m}\in L^\infty(\mathbb{T})$ by

$$H_\ell(\gamma)=\sqrt{\binom{2m}{\ell}}\sin^\ell(\pi\gamma)\cos^{2m-\ell}(\pi\gamma),\quad \ell=1,\dots,2m. \tag{2.62}$$

Using that $\cos(\pi(\gamma-1/2))=\sin(\pi\gamma)$ and $\sin(\pi(\gamma-1/2))=-\cos(\pi\gamma)$, it follows that

$$T_{1/2}H_\ell(\gamma)=\sqrt{\binom{2m}{\ell}}(-1)^\ell\cos^\ell(\pi\gamma)\sin^{2m-\ell}(\pi\gamma),\quad \ell=1,\dots,2m. \tag{2.63}$$

Thus, the matrix H in (2.57) is given by

$$H(\gamma)=\begin{pmatrix} H_0(\gamma) & T_{1/2}H_0(\gamma)\\ H_1(\gamma) & T_{1/2}H_1(\gamma)\\ \cdot & \cdot\\ \cdot & \cdot\\ H_{2m}(\gamma) & T_{1/2}H_{2m}(\gamma)\end{pmatrix}$$

$$=\begin{pmatrix} \cos^{2m}(\pi\gamma) & \sin^{2m}(\pi\gamma)\\ \sqrt{\binom{2m}{1}}\sin(\pi\gamma)\cos^{2m-1}(\pi\gamma) & -\sqrt{\binom{2m}{1}}\cos(\pi\gamma)\sin^{2m-1}(\pi\gamma)\\ \sqrt{\binom{2m}{2}}\sin^2(\pi\gamma)\cos^{2m-2}(\pi\gamma) & \sqrt{\binom{2m}{2}}\cos^2(\pi\gamma)\sin^{2m-2}(\pi\gamma)\\ \cdot & \cdot\\ \cdot & \cdot\\ \sqrt{\binom{2m}{2m}}\sin^{2m}(\pi\gamma) & \sqrt{\binom{2m}{2m}}\cos^{2m}(\pi\gamma)\end{pmatrix}$$

We now verify the conditions in Corollary 2.48. Using the binomial formula

$$(x+y)^{2m} = \sum_{\ell=0}^{2m} \binom{2m}{\ell} x^{\ell} y^{2m-\ell}, \tag{2.64}$$

we see via (2.62) that

$$\begin{aligned}\sum_{\ell=0}^{2m} |H_\ell(\gamma)|^2 &= \sum_{\ell=0}^{2m} \binom{2m}{\ell} \sin^{2\ell}(\pi\gamma)\cos^{2(2m-\ell)}(\pi\gamma) \\ &= \left(\sin^2(\pi\gamma)+\cos^2(\pi\gamma)\right)^{2m} \\ &= 1, \quad \gamma \in \mathbb{T}.\end{aligned}$$

Using the binomial formula with $x = -1, y = 1$, the expressions in (2.62) and (2.63) yield

$$\begin{aligned}\sum_{\ell=0}^{2m} \overline{H_\ell(\gamma)} T_{1/2} H_\ell(\gamma) &= \sin^{2m}(\pi\gamma)\cos^{2m}(\pi\gamma) \sum_{\ell=0}^{2m} (-1)^{\ell} \binom{2m}{\ell} \\ &= \sin^{2m}(\pi\gamma)\cos^{2m}(\pi\gamma)(1-1)^{2m} \\ &= 0.\end{aligned}$$

Now Corollary 2.48 implies that the $2m$ functions $\psi_1, \ldots, \psi_{2m}$ defined by

$$\begin{aligned}\widehat{\psi_\ell}(\gamma) &= H_\ell(\gamma/2)\widehat{\psi_0}(\gamma/2) \\ &= \sqrt{\binom{2m}{\ell}} \frac{\sin^{2m+\ell}(\pi\gamma/2)\cos^{2m-\ell}(\pi\gamma/2)}{(\pi\gamma/2)^{2m}}\end{aligned}$$

generate a tight multiwavelet frame $\{D^j T_k \psi_\ell\}_{j,k\in\mathbb{Z},\ell=1,\ldots,2m}$ for $L^2(\mathbb{R})$.

We want to study the properties of the frame constructed in Example 2.49, but we first change the definition slightly by multiplying each of the functions H_ℓ in (2.62) with a complex number of absolute value 1. This modification will not change the frame properties for the generated wavelet system.

Example 2.50. We continue Example 2.49, but now we define

$$H_\ell(\gamma) = i^{\ell} \sqrt{\binom{2m}{\ell}} \sin^{\ell}(\pi\gamma)\cos^{2m-\ell}(\pi\gamma), \quad \ell = 1, \ldots, 2m. \tag{2.65}$$

H_ℓ only differs from the choice in (2.62) by a constant of absolute value 1, so the functions $\psi_1, \ldots, \psi_{2m}$ given by

$$\widehat{\psi_\ell}(2\gamma) = H_\ell(\gamma)\widehat{\psi_0}(\gamma), \quad \ell = 1, \ldots, 2m, \tag{2.66}$$

also generate a tight multiwavelet frame. Instead of inserting the expression for $\widehat{\psi_0}$ in (2.66), we now rewrite $H_\ell(\gamma)$ using Euler's formula:

$$H_\ell(\gamma) = i^\ell \sqrt{\binom{2m}{\ell}} \left(\frac{e^{\pi i\gamma} - e^{-\pi i\gamma}}{2i}\right)^\ell \left(\frac{e^{\pi i\gamma} + e^{-\pi i\gamma}}{2}\right)^{2m-\ell}$$
$$= 2^{-2m} \sqrt{\binom{2m}{\ell}} \left(e^{\pi i\gamma} - e^{-\pi i\gamma}\right)^\ell \left(e^{\pi i\gamma} + e^{-\pi i\gamma}\right)^{2m-\ell}. \tag{2.67}$$

Via the binomial formula we see that $H_\ell(\gamma)$ is a finite linear combination of terms:

$$e^{-2\pi i m\gamma}, e^{-2\pi i(m-1)\gamma}, \ldots, e^{2\pi i(m-1)\gamma}, e^{2\pi i m\gamma}.$$

All coefficients in the linear combination are real. Writing

$$H_\ell(\gamma) = \sum_{k=-m}^{m} c_{k,\ell} e^{2\pi i k\gamma},$$

it follows that

$$\psi_\ell = \sqrt{2} \sum_{k=-m}^{m} c_{k,\ell} D T_{-k} \psi_0. \tag{2.68}$$

That is, ψ_ℓ is a real-valued spline. Since $DT_m\psi_0$ has support in $[0,m]$ and $DT_{-m}\psi_0$ has support in $[-m,0]$, the spline ψ_ℓ has support in $[-m,m]$. Our arguments also show that the splines ψ_ℓ inherit other properties from ψ_0: They have degree $2m-1$, belong to $C^{2m-2}(\mathbb{R})$, and have knots at $\mathbb{Z}/2$.

Let us find an explicit expression for the generators in Example 2.50 in the case $m=1$:

Example 2.51. In the case $m=1$, the construction in Example 2.50 leads to two generators, ψ_1 and ψ_2. Via the expression (2.67) for H_1,

$$H_1(\gamma) = \frac{1}{4}\sqrt{\binom{2m}{\ell}} (e^{\pi i\gamma} - e^{-\pi i\gamma})(e^{\pi i\gamma} + e^{-\pi i\gamma})$$
$$= \frac{1}{2\sqrt{2}} (e^{2\pi i\gamma} - e^{-2\pi i\gamma}).$$

Via elementary manipulations with the Fourier transform, we conclude that

$$\psi_1(x) = \frac{1}{\sqrt{2}} (B_2(2x+1) - B_2(2x-1)). \tag{2.69}$$

See Fig.2.4 Similarly, one proves (Exercise 25) that

$$\psi_2(x) = \frac{1}{2} \left(B_2(2x+1) - 2B_2(2x) + B_2(2x-1)\right), \tag{2.70}$$

which is shown in Fig. 2.4.

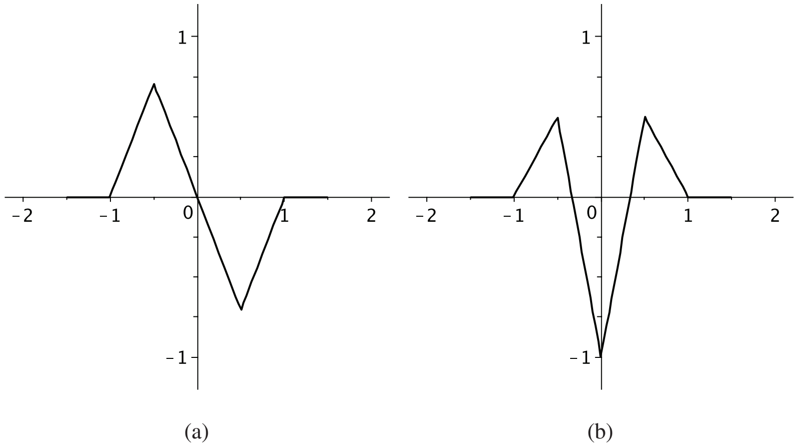


Fig. 2.4: (a) The function ψ_1 given by (2.69); (b) the function ψ_2 given by (2.70).

We note that the computational effort in Example 2.50 increases with the order of the B-spline B_{2m} we start with: The number of generators $\psi_1, \ldots, \psi_{2m}$ increases with the order of the spline B_{2m}, and (2.68) shows that computation of ψ_ℓ involves the calculation of a large number of coefficients for high-order B-splines.

Wavelet orthonormal bases have traditionally been used for approximation-theoretic purposes. In this context it is known that the number of vanishing moments plays a key role. Unfortunately, it has been shown that among the B-spline frame generators constructed using the unitary extension principle, at least one of the generators can have at most one vanishing moment. Using a modification of the unitary extension principle, it has been shown by two groups of researchers [52] and [61] how one can obtain similar constructions with a high number of vanishing moments. Also, one can prove that it is possible to construct multiwavelet frames with two generators based on any B-spline B_{2m}, i.e., with any prescribed regularity; see [52] and [61].

Exercises

1. Assume that $\{e_k\}_{k=1}^{\infty}$ is a sequence of normalized vectors in a Hilbert space $\mathscr{H}$ and that
$$\sum_{k=1}^{\infty} |\langle f, e_k\rangle|^2 = \|f\|^2, \quad \forall f \in \mathscr{H}.$$
Show that $\{e_k\}_{k=1}^{\infty}$ is an orthonormal basis for $\mathscr{H}$.
2. Assume that $\{f_k\}_{k=1}^{\infty}$ is a Bessel sequence with bound B. Prove that
   a. $\|f_k\|^2 \le B$ for all $k \in \mathbb{N}$.
   b. If $\|f_k\|^2 = B$ for some $k \in \mathbb{N}$, then $f_k \perp f_j$ for all $j \in \mathbb{N} \setminus \{k\}$.
3. Let $\{f_k\}_{k=1}^{\infty}$ be a sequence in a Hilbert space $\mathscr{H}$. Prove that
   a. If there exists $B > 0$ such that
   $$\left\|\sum c_k f_k\right\|^2 \le B \sum |c_k|^2$$
   for all finite sequences $\{c_k\}$, then $\sum_{k=1}^{\infty} c_k f_k$ converges for all $\{c_k\}_{k=1}^{\infty} \in \ell^2(\mathbb{N})$ and $\{f_k\}_{k=1}^{\infty}$ is a Bessel sequence with bound B.
   b. If (2.14) holds for all finite scalar sequences $\{c_k\}$, then it holds for all $\{c_k\}_{k=1}^{\infty} \in \ell^2(\mathbb{N})$.
   c. If $\{f_k\}_{k=1}^{\infty}$ is a Riesz basis, then
   $$\sum_{k=1}^{\infty} c_k f_k \text{ is convergent} \quad \Leftrightarrow \quad \{c_k\}_{k=1}^{\infty} \in \ell^2(\mathbb{N}).$$
4. Two sequences $\{f_k\}_{k=1}^{\infty}$ and $\{g_k\}_{k=1}^{\infty}$ in a Hilbert space are biorthogonal if
$$\langle f_k, g_j\rangle = \delta_{k,j}.$$
Show that for a pair of dual Riesz bases $\{f_k\}_{k=1}^{\infty}$ and $\{g_k\}_{k=1}^{\infty}$, the following hold:
   a. $\{f_k\}_{k=1}^{\infty}$ and $\{g_k\}_{k=1}^{\infty}$ are biorthogonal.
   b. For all $f \in \mathscr{H}$,
   $$f = \sum_{k=1}^{\infty} \langle f, g_k\rangle f_k = \sum_{k=1}^{\infty} \langle f, f_k\rangle g_k. \tag{2.71}$$
5. Prove Proposition 2.11.
6. Show that if (2.14) holds for all finite sequences $\{c_k\}_{k=1}^{\infty}$, it automatically holds for all $\{c_k\}_{k=1}^{\infty} \in \ell^2(\mathbb{N})$.
7. Prove Lemma 2.14.
8. Find an example of a sequence in a Hilbert space that is a basis but not a frame.
9. Show that if $\{f_k\}_{k=1}^{\infty}$ is a dual frame of a frame $\{g_k\}_{k=1}^{\infty}$, then $\{g_k\}_{k=1}^{\infty}$ is also a dual frame of $\{f_k\}_{k=1}^{\infty}$.

10. Prove that the upper and lower frame conditions are unrelated: In an arbitrary Hilbert space $\mathscr{H}$, there exists a sequence $\{f_k\}_{k=1}^{\infty}$ satisfying the upper condition for all $f \in \mathscr{H}$, but not the lower condition; and vice versa.
11. Let $\{e_k\}_{k=1}^{\infty}$ be an orthonormal basis and consider the family

$$\{f_k\}_{k=1}^{\infty} := \left\{ e_1 + \frac{1}{k} e_k, e_k \right\}_{k=2}^{\infty}.$$

 Prove that $\{f_k\}_{k=1}^{\infty}$ is not a Bessel sequence.
12. Let $\{f_k\}_{k=1}^{\infty}$ and $\{g_k\}_{k=1}^{\infty}$ be dual frames for a Hilbert space $\mathscr{H}$, and $U : \mathscr{H} \to \mathscr{H}$ a unitary operator. Show that $\{Uf_k\}_{k=1}^{\infty}$ and $\{Ug_k\}_{k=1}^{\infty}$ also form a pair of dual frames for $\mathscr{H}$.
13. Prove Theorem 2.22.
14. We consider the B-splines N_2 and N_3.

 a. Show via the definition that the B-spline N_2 is given by

$$N_2(x) = \begin{cases} x, & \text{if } x \in [0,1], \\ 2 - x, & \text{if } x \in [1,2], \\ 0, & \text{otherwise.} \end{cases}$$

 b. Use (a) to show that

$$N_3(x) = \begin{cases} \frac{1}{2}x^2, & \text{if } x \in [0,1], \\ -x^2 + 3x - \frac{3}{2}, & \text{if } x \in [1,2], \\ \frac{1}{2}x^2 - 3x + \frac{9}{2}, & \text{if } x \in [2,3], \\ 0, & \text{otherwise.} \end{cases}$$

15. Consider the B-spline $N_n,\ n \in \mathbb{N}$.

 a. Show that

$$\widehat{N_n}(\gamma) = e^{-\pi i n \gamma} \left(\frac{\sin \pi\gamma}{\pi\gamma} \right)^n.$$

 b. Show that

$$\widehat{N_n}(2\gamma) = e^{-\pi i n \gamma} (\cos \pi\gamma)^n \widehat{N_n}(\gamma).$$

 c. Show that the function

$$H_0(\gamma) := e^{-\pi i n \gamma} (\cos \pi\gamma)^n$$

 is 1-periodic.

16. Show that the definitions of the centered B-splines B_m in (2.30) and (2.31) coincide.

17. Consider the centered B-spline $B_n,\ n \in \mathbb{N}$.

    a. Show that
    $$\widehat{B_n}(2\gamma) = (\cos \pi\gamma)^n \widehat{B_n}(\gamma).$$

    b. Show that the function
    $$H_0(\gamma) := (\cos \pi\gamma)^n$$
    is 1-periodic if and only if n is even.

18. Prove Corollary 2.31.
19. Show that for the B-spline B_2, the system $\{E_{mb}T_{2n}B_2\}_{m,n\in\mathbb{Z}}$ cannot be a frame for any $b > 0$.
20. Show by an example (maybe with $a = b = 1$) that the necessary condition in Proposition 2.28 does not suffice for $\{E_{mb}T_{na}g\}_{m,n\in\mathbb{Z}}$ being a Gabor frame.
21. Prove that $\{E_m T_{na} \chi_{[0,1]}\}_{m,n\in\mathbb{Z}}$ is a frame for $L^2(\mathbb{R})$ for all $a \in (0,1]$.
22. Prove Corollary 2.33.
23. Prove Corollary 2.37.
24. Verify that (2.60) is equivalent to the condition $H(\gamma) * H(\gamma) = I$, a.e. γ.
25. Derive expression (2.70) for the function ψ_2.
26. Let $\{D^j T_k \psi\}_{j,k\in\mathbb{Z}}$ be a frame with frame operator S. Prove that S commutes with the dilation operator D, and thereby that
    $$\{S^{-1}D^j T_k \psi\}_{j,k\in\mathbb{Z}} = \{D^j S^{-1} T_k \psi\}_{j,k\in\mathbb{Z}}.$$

Chapter 3
Continuous and Discrete Reproducing Systems That Arise from Translations. Theory and Applications of Composite Wavelets

Demetrio Labate and Guido Weiss

Abstract Reproducing systems of functions such as the wavelet and Gabor systems have been particularly successful in a variety of applications from both mathematics and engineering. In this chapter, we review a number of recent results in the study of such systems and their generalizations developed by the authors and their collaborators. We first describe the unified theory of reproducing systems. This is a simple and flexible mathematical framework to characterize and analyze wavelets, Gabor systems, and other reproducing systems in a unified manner. The systems of interest to us are obtained by applying families of translations, modulations, and dilations to a countable set of functions. As the reader will see, we can rewrite such systems as a countable family of translations applied to a countable collection of functions. Building in part on this approach, we define the wavelets with composite dilations, a novel class of reproducing systems that provide truly multidimensional generalizations of traditional wavelets. For example, in dimension 2, the elements of such systems are defined not only at various scales and locations, as traditional wavelet systems, but also at various orientations. The shearlet system is a special case of a composite wavelet system that provides an optimally sparse representation for a large class of bivariate functions. This is useful for a number of applications in image processing, such as image denoising and edge detection. Finally, we discuss some related issues about the continuous wavelet transform and the continuous analogues of composite wavelets.

3.1 Introduction

These lectures present an overview of a research program developed by the authors and their collaborators at Washington University in St. Louis during the past 10 years, which is devoted to the study of reproducing systems of functions.

Demetrio Labate
North Carolina State University, Raleigh, NC 27695, USA, e-mail: dlabate@math.ncsu.edu

Guido Weiss
Washington University in St. Louis, St. Louis, MO 63130, USA,
e-mail: guido@math.wustl.edu

B. Forster and P. Massopust (eds.), *Four Short Courses on Harmonic Analysis*,
Applied and Numerical Harmonic Analysis, DOI 10.1007/978-0-8176-4891-6_3,

By *reproducing systems of functions*, we refer to those families of functions $\{\psi_i : i \in \mathscr{I}\}$ in $L^2(\mathbb{R}^n)$ that are obtained by applying a countable collection of operators to a countable set of "generating" functions and have the property that any function $f \in L^2(\mathbb{R}^n)$ can be recovered from the reproducing formula

$$f = \sum_{i\in\mathscr{I}} \langle f, \psi_i\rangle\, \psi_i,$$

with convergence in the L^2-norm. The *wavelet systems*, for example, have received a great deal of attention in the last 20 years, since their applications in mathematics and engineering have been especially successful. In dimension $n = 1$, they are defined as those collections of the form

$$\Psi = \{\psi_{j,k} = 2^{j/2}\, \psi(2^j \cdot -k) : j,k \in \mathbb{Z}\}, \tag{3.1}$$

where ψ is a fixed function in $L^2(\mathbb{R})$. As the above expression shows, Ψ is obtained by applying dyadic dilations and integer translations to the generating function ψ. For particular choices of the generator ψ, the wavelet system Ψ is an orthonormal basis or a Parseval frame for $L^2(\mathbb{R})$, in which case any $f \in L^2(\mathbb{R}^n)$ can be recovered as

$$f = \sum_{j,k\in\mathbb{Z}} \langle f, \psi_{j,k}\rangle\, \psi_{j,k}, \tag{3.2}$$

with convergence in the L^2-norm. Other important classes of reproducing systems are the Gabor systems, which are obtained by applying translations and modulations to a fixed generator, and the wave packet systems, which involve translations, dilations, and modulations.

One main theme developed in these lectures is that there is a general framework that allows us to describe and analyze wavelet systems, Gabor systems, and many other reproducing systems by using a unified approach. Indeed, for a large class of reproducing systems of the form

$$\{g_p(\cdot - C_p k) : k \in \mathbb{Z}^n, p \in \mathscr{P}\}, \tag{3.3}$$

where $\mathscr{P}$ is countable and $\{C_p\}$ is a set of invertible matrices, there is a relatively simple set of equations that characterizes those generating functions $\{g_p\}_{p\in\mathscr{P}}$ such that the corresponding system (3.3) is an orthonormal basis or, more generally, a Parseval frame for $L^2(\mathbb{R}^n)$. For example, it was discovered by Gripenberg [107] and Wang [232] independently, in 1995, that a function $\psi \in L^2(\mathbb{R})$ is the generator of an orthonormal wavelet system if and only if $\|\psi\|_2 = 1$,

$$\sum_{j\in\mathbb{Z}} |\hat{\psi}(2^j\xi)|^2 = 1 \quad \text{for a. e. } \xi \in \mathbb{R}, \tag{3.4}$$

and

$$t_q(\xi) = \sum_{j\geq 0} \hat{\psi}(2^j\xi)\overline{\hat{\psi}(2^j(\xi+q))} = 0 \quad \text{for a.e. } \xi \in \mathbb{R}, \tag{3.5}$$

whenever q is an odd integer. It is remarkable that a similar set of characterization equations holds not only for wavelet systems in higher dimensions, but also for many other reproducing systems. This topic, and the corresponding *unified theory of reproducing systems,* will be presented in Section 3.2.

Parallel to the unified theory mentioned above, there is another "unifying" perspective to the study of reproducing systems that is provided by representation theory and, more specifically, by the study of the continuous wavelet transform and its generalization. In Section 3.3, we introduce the continuous analogues of the wavelet systems (3.1), which are obtained by applying dilations (with respect to a *dilation* group) and continuous translations to a function $\psi \in L^2(\mathbb{R}^n)$. For example, in dimension $n = 1$, the continuous wavelet system is a system of the form

$$\{\psi_{at} = a^{-\frac{1}{2}}\psi(a^{-1}(\cdot - t)) : a > 0,\, t \in \mathbb{R}\},$$

and the (one-dimensional) *continuous wavelet transform* is the mapping

$$f \mapsto \left\{ \langle f, \psi_{at}\rangle = a^{-\frac{1}{2}} \int_0^\infty f(y)\overline{\psi(a^{-1}(y-t))}\,dy : (a,t) \in \mathbb{R}^+ \times \mathbb{R} \right\}.$$

Then, provided that ψ satisfies a certain admissibility condition, any $f \in L^2(\mathbb{R})$ can be expressed using the *Calderón reproducing formula*:

$$f = \int_{\mathbb{R}} \int_0^\infty \langle f, \psi_{at}\rangle\, \psi_{at}\, \frac{da}{a}\, dt. \tag{3.6}$$

The close relationship between the discrete and continuous frameworks is apparent by comparing the last expression with formula (3.2). A number of observations concerning this relationship, as well as several multidimensional extensions of the continuous wavelet transform, are discussed in Section 3.3.

Traditional multidimensional wavelet systems are obtained by taking tensor products of one-dimensional ones, as a result, they have a very limited capability to deal effectively with those directional features that typically occur in images and other multidimensional data. To overcome such limitations, several extensions and generalizations have been proposed in applied harmonic analysis during the last 10 years. One such approach is the *theory of wavelets with composite dilations*, which was originally introduced by the authors and their collaborators and provides a very flexible and powerful framework to construct "truly" multidimensional extensions of the wavelet approach.

An example of a composite wavelet system, in dimension $n = 2$, is the collection

$$\{\psi_{ijk} = |\det A|^{i/2}\, \psi(B^j A^i \cdot - k) : i, j \in \mathbb{Z},\, k \in \mathbb{Z}^2\}, \tag{3.7}$$

where $A = \begin{pmatrix} 2 & 0 \\ 0 & \sqrt{2} \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. The elements of such systems are defined not only at various scales and locations, as traditional wavelet systems, but also at various orientations, associated with the powers of the *shearing matrix* B. In addition, for appropriate choices of ψ, the elements ψ_{ijk} have the ability to provide very efficient representations for data containing directional and anisotropic features (see Section 3.5). There is a variety of systems of the form (3.7) forming Parseval frames or even orthonormal bases, for many choices of matrices A and B. Indeed, the theory of wavelets with composite dilations encompasses the theory of wavelets, and there is a generalized multiresolution analysis (MRA) associated with this theory. As in the case of the classical MRA, this framework allows one to obtain a variety of constructions with many different geometric and analytic properties. An outline of this theory is presented in Section 3.4.

In Section 3.5, we examine a generalization of the wavelet transform associated with the affine group

$$G = \{(M,t) : M \in \mathscr{D}_\alpha,\ t \in \mathbb{R}^2\},$$

where, for each $0 < \alpha < 1$, $\mathscr{D}_\alpha \subset GL_2(\mathbb{R})$ is the set of matrices

$$\mathscr{D}_\alpha = \left\{ M = M_{as} = \begin{pmatrix} a & -a^\alpha s \\ 0 & a^\alpha \end{pmatrix}, \quad a > 0,\ s \in \mathbb{R} \right\}.$$

Associated with this is the *continuous shearlet transform* $\mathscr{S}_\psi^\alpha$, defined by

$$f \to \{\mathscr{S}_\psi^\alpha f(a,s,t) = \langle f, \psi_{ast} \rangle : a > 0,\ s \in \mathbb{R},\ t \in \mathbb{R}^2\},$$

which is mapping $f \in L^2(\mathbb{R}^2)$ into a transform domain dependent on the scale a, the shearing parameter s, and the location t. The analyzing elements ψ_{ast}, forming a *continuous shearlet system*, are the functions

$$\psi_{ast}(x) = |\det M_{as}|^{-\frac{1}{2}}\, \psi(M_{as}^{-1}(x-t)), \tag{3.8}$$

with $M_{as} \in \mathscr{D}_\alpha$. One remarkable property is that the continuous shearlet transform of a function f has the ability to completely characterize both the location and the geometry of the set of singularities of f.

A *discrete shearlet system* is obtained by appropriately discretizing the functions (3.8). Indeed, such a discrete system can be designed so that it forms a Parseval frame and it provides us with a special case of wavelets with composite dilations (3.7). In addition, the generator ψ can be chosen to be a well-localized function; that is, ψ has fast decay in both the space and frequency domains (see [124, 127]). As a result, the elements of the discrete shearlet systsm form a collection of well-localized waveforms at various scales, locations, and orientations and provide optimally sparse representations for a large class of bivariate functions with distributed discontinuities. Only the curvelets introduced by Candès and Donoho have been proved, to have similar properties; however, the curvelets do not share the simple affine-like structure of wavelets with composite dilations. To illustrate the advantages of the shearlet framework with respect to wavelets and other traditional representations, we describe a number of useful applications of shearlets to the analysis and processing of images, including some representative applications of feature extraction and edge detection.

3.2 Unified Theory of Reproducing Systems

In order to describe the types of reproducing systems that we will consider in this study, it will be useful to introduce the following definitions. We adopt the convention that $x \in \mathbb{R}^n$ is a column vector, i.e., $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, and that $\xi \in \widehat{\mathbb{R}}^n$ is a row vector, i.e., $\xi = (\xi_1, \dots, \xi_n)$. A vector x multiplying a matrix $M \in GL_n(\mathbb{R})$ on the right is understood to be a column vector, while a vector ξ multiplying M on the left is a row vector. Thus, $Mx \in \mathbb{R}^n$ and $\xi M \in \widehat{\mathbb{R}}^n$.

Let $f \in L^2(\mathbb{R}^n)$. For $y \in \mathbb{R}^n$, the *translation operator* T_y is defined by $T_y f(x) = f(x-y)$; for $M \in GL_n(\mathbb{R})$, the *dilation operator* D_M is defined by $D_M f(x) = |\det M|^{-1/2} f(M^{-1}x)$; for $\nu \in \mathbb{R}^n$, the *modulation operator* M_ν is defined by $(M_\nu f)(x) = e^{2\pi i \nu x} f(x)$.

We will use the Fourier transform in the form

$$\hat{f}(\xi) = \int_{\mathbb{R}^n} f(x)\, e^{-2\pi i \xi x}\, dx,$$

for $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$. Thus, the inverse Fourier transform is given by

$$\check{f}(x) = \int_{\widehat{\mathbb{R}}^n} f(\xi)\, e^{2\pi i \xi x}\, d\xi.$$

We remark that $(T_y f)^\wedge(\xi) = (M_y \hat{f})(\xi)$ and $(D_M f)^\wedge(\xi) = (\widehat{D}_M \hat{f})(\xi)$, where $(D_M f)^\wedge(\xi) = (\hat{D}_M \hat{f})(\xi) = |\det M|^{1/2} \hat{f}(\xi M)$.

Virtually all systems of functions used in harmonic analysis to generate subspaces of $L^2(R^n)$ are obtained by applying a certain combination of translations, dilations, and modulations to a finite family of functions in $L^2(\mathbb{R}^n)$. Let us start by recalling the definitions of the systems commonly used in many harmonic analysis applications.

Gabor Systems. Let $\Psi = \{\psi^1, \dots, \psi^L\} \subset L^2(\mathbb{R}^n)$, and $B, C \in GL_n(\mathbb{R})$. The *Gabor systems* are the collections

$$\mathscr{G} = \mathscr{G}_{B,C}(\Psi) = \{M_{Bm} T_{Ck}\, \psi^\ell : m,k \in \mathbb{Z}^n,\ \ell = 1, \dots, L\}$$

or

$$\tilde{\mathscr{G}} = \tilde{\mathscr{G}}_{B,C}(\Psi) = \{T_{Ck} M_{Bm}\, \psi^\ell : m,k \in \mathbb{Z}^n,\ \ell = 1, \dots, L\}.$$

Notice that $\tilde{\mathscr{G}}$ is obtained from $\mathscr{G}$ by interchanging the order of the translation and modulation operators. Also, it is easy to see that

$$M_{Bm} T_{Ck}\, \psi^\ell = e^{-2\pi i BmCk}\, T_{Ck} M_{Bm}\, \psi^\ell.$$

Affine Systems. Given $\Psi = \{\psi^1, \dots, \psi^L\} \subset L^2(\mathbb{R}^n)$, $A \subset GL_n(\mathbb{R})$, and $\Gamma \subset \mathbb{R}^n$, the *affine systems* are the collections

$$\mathscr{F} = \mathscr{F}_{A,\Gamma}(\Psi) = \{D_a T_\gamma \psi^\ell : a \in A,\ \gamma \in \Gamma,\ \ell = 1, \dots, L\}.$$

Very often we use the notation $\mathscr{D} = \{M^j : j \in \mathbb{Z}\}$, where $M \in GL_n(\mathbb{R})$ is expanding (i.e., each proper value λ of M satisfies $|\lambda| > 1$), and Γ is the lattice $C\mathbb{Z}^n$, where $C \in GL_n(\mathbb{R})$.

Wave Packet Systems. These include the above two systems. For $\Psi = \{\psi^1, \dots, \psi^L\}$, they consist of those functions

$$\mathscr{W}\mathscr{P}_{\Gamma,A,S}(\Psi) = \{T_\gamma D_a M_y \psi^\ell : \gamma \in \Gamma,\ a \in \mathscr{D},\ y \in S,\ \ell = 1, \dots, L\},$$

where Γ, S are countable (or finite) subsets of $\mathbb{R}^n$, $A \subset GL_n(\mathbb{R})$. As will be discussed later, the order of the three operators T_γ, D_a, M_y can be permuted.

It is easy to see that each of the above systems can be expressed in the following form.

Let $\mathscr{P}$ be a countable indexing set, $\{g_p : p \in \mathscr{P}\}$ a family of functions in $L^2(R^n)$, and $\{C_p : p \in \mathscr{P}\}$ a corresponding collection of matrices in $GL_n(\mathbb{R})$. Then each of the systems we just described has the form

$$\left\{T_{C_p k}\, g_p : k \in \mathbb{Z}^n,\ p \in \mathscr{P}\right\}. \tag{3.9}$$

Indeed, in order to write down the general wave packet system into the form (3.9), one needs just to use the "commutativity relations" $D_M T_k = T_{Mk} D_M$ and $M_y T_k = e^{2\pi i y k} T_k M_y$ (notice that $e^{2\pi i y k}$ is a constant of absolute value 1).

3.2.1 Unified Theorem for Reproducing Systems

In the theory of wavelets and, more generally, in harmonic analysis, it is of paramount importance to construct such systems that form a reproducing set for the space $L^2(\mathbb{R}^n)$ (or more general function spaces). For example, it is of particular interest to know when a system $\{\phi_j : j \in \mathbb{Z}\}$ of functions in $L^2(\mathbb{R}^n)$ is an orthonormal basis or, more generally, a frame. Many characterizations of systems that are Parseval frames have been given in the literature; most often these results concern affine systems [107, 106, 137, 158, 206, 232].

We shall now give necessary and sufficient conditions for the system (3.3) to be a Parseval frame for $L^2(\mathbb{R}^n)$. For simplicity, we are letting the lattice Γ be $\mathbb{Z}^n$; our arguments below can be easily extended to a more general Γ.

Recall that a countable collection $\{\phi_i\}_{i \in I}$ in a (separable) Hilbert space $\mathscr{H}$ is a *Parseval frame* (sometimes called a *tight frame* with constant 1) for $\mathscr{H}$ if

$$\sum_{i \in I} |\langle f, \phi_i \rangle|^2 = \|f\|^2 \quad \text{for all } f \in \mathscr{H}.$$

This is equivalent to the reproducing formula $f = \sum_i \langle f, \phi_i \rangle \phi_i$, for all $f \in \mathscr{H}$, where the series converges unconditionally in the norm of $\mathscr{H}$. This shows that a Parseval frame provides a basis-like representation even though a Parseval frame need not be a basis in general. We refer the reader to [43, 47] for more details about frames.

We refer to the following result as the "unifying theorem for reproducing systems" [137]:

Theorem 3.1. *Let $\mathscr{P}$ be a countable indexing set, $\{g_p\}_{p\in\mathscr{P}}$ a collection of functions in $L^2(\mathbb{R}^n)$, and $\{C_p\}_{p\in\mathscr{P}} \subset GL_n(\mathbb{R})$. Let*

$$\mathscr{E} = \left\{ f \in L^2(\mathbb{R}^n) : \hat{f} \in L^\infty(\mathbb{R}^n) \text{ and } \operatorname{supp} \hat{f} \text{ is compact} \right\},$$

and suppose that

$$\mathscr{L}(f) = \sum_{p\in\mathscr{P}} \sum_{m\in\mathbb{Z}^n} \int_{\operatorname{supp}\hat{f}} |\hat{f}(\xi + mC_p^{-1})|^2 \frac{1}{|\det C_p|} |\hat{g}_p(\xi)|^2 \, d\xi < \infty \tag{3.10}$$

for all $f \in \mathscr{E}$. Then the system (3.3) is a Parseval frame for $L^2(\mathbb{R}^n)$ if and only if

$$\sum_{p\in\mathscr{P}_\alpha} \frac{1}{|\det C_p|} \overline{\hat{g}_p(\xi)} \, \hat{g}_p(\xi + \alpha) = \delta_{\alpha,0} \quad \text{for a.e. } \xi \in \mathbb{R}^n, \tag{3.11}$$

for each $\alpha \in \Lambda = \bigcup_{p\in\mathscr{P}} \mathbb{Z}^n C_p^{-1}$, where $\mathscr{P}_\alpha = \{p \in \mathscr{P} : \alpha C_p \in \mathbb{Z}^n\}$ and δ is the Kronecker delta for $\mathbb{R}^n$.

Before discussing the proof of this theorem, it will be useful to make a few comments about this result, in order to elucidate its context and impact.

Remark 3.2. It is relatively well known that if $\psi \in L^2(\mathbb{R})$, then $\{\psi_{jk} = D_{2^j} T_k \psi : j,k \in \mathbb{Z}\}$ is an orthonormal basis for $L^2(\mathbb{R})$ (i.e., ψ is an *orthonormal wavelet*) if and only if Eqs. (3.4) and (3.5) hold. As we mentioned above, this result was obtained independently by Gripenberg [107] and Wang [232]. As will be discussed, these equations are a simple consequence of Theorem 3.1 (see Exercise 1).

Remark 3.3. Assumption (3.10) is referred to as the *local integrability condition* (LIC). At first sight, it might appear as a rather formidable technical hypothesis. In some cases, however, it can be shown that it is a simple consequence of the system being considered. For example, let us consider the Gabor system $\tilde{\mathscr{G}}_{B,C}(G)$, where $G = \{g^1, \ldots, g^L\}$, and let us write it in the form (3.3). Namely, let $\mathscr{P} = \mathbb{Z}^n \times \{1,2,\ldots,L\}$, $g_p = g_{j,\ell} = M_{Bj}\, g^\ell$, and $C_p = C$, so that

$$T_{C_p k}\, g_p = T_{Ck} M_{Bj}\, g^\ell.$$

Without loss of generality, we can assume that $L = 1$. Thus, the expression of (3.10) is

$$\begin{aligned}
\mathscr{L}(f) &= \sum_{p\in\mathscr{P}} \sum_{m\in\mathbb{Z}^n} \int_K |\hat{f}(\xi + m(C_p)^{-1}|^2 |\hat{g}_p(\xi)|^2 \frac{d\xi}{|\det C_p|} \\
&= \frac{1}{|\det C|} \sum_{p\in\mathbb{Z}^n} \sum_{m\in\mathbb{Z}^n} \int_K |\hat{f}(\xi + m(C)^{-1}|^2 |\hat{g}(\xi - Bp)|^2 \, d\xi,
\end{aligned}$$

for $f \in \mathscr{E}$ and $K = \operatorname{supp} \hat{f}$ is compact. Since $\xi \in K$, only a finite number of terms in the sum $\sum_{m\in\mathbb{Z}^n}$ are nonzero. Moreover, if $\mathbb{T}^n$ is the n-torus, for each $j \in \mathbb{Z}^n$, the set $\{B(\mathbb{T}^n + j - p) : p \in \mathbb{Z}^n\}$ is a partition of $\mathbb{R}^n$. Thus,

$$\|g\|_2^2 = \int_{\bigcup_{p\in\mathbb{Z}^n} B(\mathbb{T}^n+j-p)} |\hat{g}(\eta)|^2 \, d\eta = \sum_{p\in\mathbb{Z}^n} \int_{B(\mathbb{T}^n+j)} |\hat{g}(\xi - Bp)|^2 \, d\xi .$$

Now observe that a finite union of the sets $\{B(\mathbb{T}^n + j) : j \in \mathbb{Z}^n\}$ covers K. Using this fact and the fact that $\|\hat{f}\|_\infty \leq \infty$ (since $f \in \mathscr{E}$), it is not difficult to show that

$$\mathscr{L}(f) \leq A \, \|g\|_2^2,$$

where A is a positive constant. As a result, the characterization theorem for the Gabor systems can be stated explicitly as follows:

Theorem 3.4. *The system $\mathscr{G}_{B,C}(G)$ [or the system $\widetilde{\mathscr{G}}_{B,C}(G)$] is a Parseval frame for $L^2(\mathbb{R}^n)$ if and only if*

$$\sum_{\ell=1}^{L} \sum_{k\in\mathbb{Z}^n} \frac{1}{|\det C|} \, \hat{g}^\ell(\xi - Bk) \, \overline{\hat{g}^\ell(\xi - Bk + mC^{-1})} = \delta_{m,0}$$

for a.e. $\xi \in \mathbb{R}^n$, all $m \in \mathbb{Z}^n$.

This result is well known and can be found, for example, in [145, 207, 58, 158].

The situation for the "usual" affine systems is somewhat more subtle. Here, by the word "usual," we mean the case where $A = \{a^j : j \in \mathbb{Z}\}$, where $a \in GL_n(\mathbb{R})$ is expanding, and $\Gamma = \mathbb{Z}^n$. In this case, one can show that if conditions (3.11) are true, then the LIC is valid and, conversely, if the system (3.3) is a Parseval frame, then the LIC also holds. Thus, in the characterization of Parseval frames given by Theorem 3.1, it is not needed to assume the LIC. The characterization theorem for these systems can be written down explicitly as

Theorem 3.5. *Let $\Psi = \{\psi^1, \ldots, \psi^L\} \subset L^2(\mathbb{R}^n)$ and $a \in GL_n(\mathbb{R})$ be expanding. Then the system $\mathscr{F}_{A,\Gamma}(\Psi) = \{D_{a^j} T_k \psi^\ell : j \in \mathbb{Z}, k \in \mathbb{Z}^n, \ \ell = 1, \ldots, L\}$ is a Parseval frame for $L^2(\mathbb{R}^n)$ if and only if*

$$\sum_{\ell=1}^{L} \sum_{j\in\mathscr{P}_\alpha} \hat{\psi}^\ell(\xi a^{-j}) \, \overline{\hat{\psi}^\ell((\xi+\alpha) a^{-j})} = \delta_{\alpha,0}, \quad \text{for a.e. } \xi \in \mathbb{R}^n, \tag{3.12}$$

for all $\alpha \in \Lambda = \bigcup_{j\in\mathbb{Z}} \mathbb{Z}^n a^j$, where $\mathscr{P}_\alpha = \{j \in \mathbb{Z} : \alpha a^{-j} \in \mathbb{Z}^n\}$.

Apart from the argument needed to establish the validity of the LIC, which we mentioned above, this last theorem is a simple consequence of Theorem 3.1 once the system $\mathscr{F}_{A,\Gamma}(\Psi)$ is expressed in the form (3.3). Notice that there is a redundancy in condition (3.12). Indeed, an elementary argument shows that (3.12) can be simplified to

$$\sum_{\ell=1}^{L} \sum_{j\in\mathscr{P}_m} \hat{\psi}^\ell(\xi a^{-j}) \, \overline{\hat{\psi}^\ell((\xi+m) a^{-j})} = \delta_{m,0}, \quad \text{for a.e. } \xi \in \mathbb{R}^n, \tag{3.13}$$

for all $m \in \mathbb{Z}^n$, where $\mathscr{P}_m = \{j \in \mathbb{Z} : ma^{-j} \in \mathbb{Z}^n\}$. It follows easily from this form of Theorem 3.5 that the result of Gripenberg and Wang (given in Remark 3.2) holds for $n = 1$ and $a = 2$.

In order to present the ideas involved in the proof of Theorem 3.1, it is useful to introduce the *C-bracket product* of $f, g \in L^2(\mathbb{R}^n)$, which, for $C \in GL_n(\mathbb{R})$, is defined by

$$[f,g](x;C) = \sum_{k\in\mathbb{Z}^n} f(x - Ck)\,\overline{g(x - Ck)}.$$

It is clear that $[f,g]$ is $C\mathbb{Z}^n$- periodic; that is, $[f,g](x + Cm;C) = [f,g](x;C)$ for each $m \in \mathbb{Z}^n$.

That the system (3.3) is a Parseval frame for $L^2(\mathbb{R}^n)$ is equivalent to

$$N^2(f) = \sum_{p\in\mathscr{P}} \sum_{k\in\mathbb{Z}^n} |\langle f, T_{C_p k}\, g_p\rangle|^2 = \|f\|_2^2, \tag{3.14}$$

for all $f \in \mathscr{E}$ [recall that $\mathscr{E}$ is dense in $L^2(\mathbb{R}^n)$].

Using the fact that $\mathbb{R}^n = \bigcup_{l\in\mathbb{Z}^n}\{(\mathbb{T}^n - l)C^{-1}\}$ is a disjoint union, it follows easily that

$$\begin{aligned}
\langle f, T_{Ck}\, g\rangle &= \int_{\mathbb{R}^n} \hat{f}(\xi)\,\overline{\hat{g}(\xi)}\, e^{2\pi i Ck\cdot\xi}\, d\xi \\
&= \sum_{l\in\mathbb{Z}^n} \int_{C^I(\mathbb{T}^n)} \hat{f}(\xi - C^I l)\,\overline{\hat{g}(\xi - C^I l)}\, e^{2\pi i Ck\cdot\xi}\, d\xi \\
&= \int_{C^I(\mathbb{T}^n)} [\hat{f},\hat{g}](\xi;C^I)\, e^{2\pi i Ck\cdot\xi}\, d\xi\,.
\end{aligned}$$

Under all these assumptions, let us consider the function

$$H(x) = \sum_{k\in\mathbb{Z}^n} |\langle T_x f, T_{Ck}\, g\rangle|^2,$$

where $C \in GL_n(\mathbb{R})$. Indeed, it is clear that the function H is $C\mathbb{Z}^n$- periodic. Using the fact that $\hat{f}$ has compact support, one can show that

Lemma 3.6. *The function $H(x)$ is the trigonometric polynomial where*

$$H(x) = \sum_{m\in\mathbb{Z}^n} \hat{H}(m)\, e^{2\pi i (C^I m)\cdot x},$$

where

$$\hat{H}(m) = \frac{1}{|\det C|} \int_{\mathbb{R}^n} \hat{f}(\xi)\,\overline{\hat{f}(\xi + C^I m)}\,\overline{\hat{g}(\xi)}\,\hat{g}(\xi + C^I m)\, d\xi,$$

and only a finite number of these expressions are nonzero.

The fact that $\hat{H}(m) \neq 0$ for finitely many m at most follows from the fact that $\hat{f}$ has compact support.

To show that Eq. (3.14) holds for all $f \in \mathscr{E}$, consider now the function

$$w(x) = N^2(T_x f) = \sum_{p \in \mathscr{P}} H_p(x),$$

where $H_p(x) = \sum_{k \in Z^n} |\langle T_x f, T_{C_p k}\, g_p \rangle|^2$. By Lemma 3.6, for each $p \in \mathscr{P}$,

$$H_p(x) = \sum_{m \in \mathbb{Z}^n} \hat{H}_p(m)\, e^{2\pi i (C_p^I m)\cdot x},$$

where

$$\hat{H}_p(m) = \frac{1}{|\det C_p|} \int_{\mathbb{R}^n} \hat{f}(\xi)\, \overline{\hat{f}(\xi + C_p^I m)}\, \overline{\hat{g}_p(\xi)}\, \hat{g}_p(\xi + C_p^I m)\, d\xi.$$

Thus, using the assumptions of Theorem 3.1, from the observations we made above, we have the expression

$$w(x) = N^2(T_x f) = \sum_{\alpha \in \Lambda} \hat{w}(\alpha)\, e^{2\pi i \alpha \cdot x}, \tag{3.15}$$

where

$$\hat{w}(\alpha) = \int_{\mathbb{R}^n} \hat{f}(\xi)\, \overline{\hat{f}(\xi + \alpha)} \sum_{p \in \mathscr{P}} \frac{1}{|\det C_p|} \overline{\hat{g}_p(\xi)}\, \hat{g}_p(\xi + \alpha)\, d\xi. \tag{3.16}$$

This integral is absolutely convergent, and the series defining $w(x)$ is absolutely and uniformly convergent. Notice that the LIC plays an important role in establishing these convergence properties and the various uses of Fubini's theorem needed for the formulas developed here.

To complete the proof of Theorem 3.1 we argue as follows. Let us assume (3.11). Then, by Eq. (3.16),

$$\hat{w}(\alpha) = \delta_{\alpha,0} \int_{\mathbb{R}^n} \hat{f}(\xi)\, \overline{\hat{f}(\xi + \alpha)}\, d\xi.$$

By Eq. (3.15), this implies

$$w(x) = N^2(T_x f) = \sum_{\alpha \in \Lambda} \hat{w}(\alpha)\, e^{2\pi i \alpha \cdot x} = \hat{w}(0) = \|f\|_2^2.$$

Hence, the system (3.3) is a Parseval frame for $L^2(\mathbb{R}^n)$.

Conversely, let us now assume that the system (3.3) is a Parseval frame for $L^2(\mathbb{R}^n)$. Hence, by our assumptions, we know that

$$N^2(T_x f) = w(x) = \sum_{\alpha \in \Lambda} \hat{w}(\alpha)\, e^{2\pi i \alpha \cdot x} = \|T_x f\|_2^2 = \|f\|_2^2,$$

for all $f \in \mathscr{E}$.

Since Λ is countable and the "Fourier coefficients" $\hat{w}(\alpha)$ of this generalized Fourier series are unique, we must have $\hat{w}(\alpha) = 0$ if $\alpha \neq 0$ and $\hat{w}(0) = 1$. We can then use (3.16) and appropriate choices of $f \in \mathscr{E}$ to show that the Eq. (3.11) must hold. For example, by letting f be such that $\hat{f}(\xi) = \hat{f}_\varepsilon(\xi) = (1/\sqrt{|B(\varepsilon)|})\chi_{B(\varepsilon)}(\xi - \xi_0)$, where $B(\varepsilon)$ is a ball of radius ε about the origin, $\varepsilon > 0$, and ξ_0 is a point of differentiability of the integral of $h(\xi) = \sum_{p\in\mathscr{P}}(1/|\det C_p|)\,|\hat{g_p}(\xi)|^2$, one obtains easily from (3.16) that $h(\xi_0) = 1$. This gives (3.11) when $\alpha = 0$.

This is, to conclude, the basic idea of the proof of Theorem 3.1. The role played by these generalized Fourier series is arrived at naturally; it arises from the importance of the notion of shift invariance, which is essentially related to the structure of these families of reproducing systems.

Theorem 3.1 has many applications, several of which are described in [137, 138]. As mentioned above, they include Gabor, affine, and wave packet systems. Theorem 3.1 applies also to the *quasi-affine systems*. In dimension $n = 1$, these are the systems $\{\tilde{\psi}_{jk} : j,k \in \mathbb{Z}\}$ obtained from $\psi \in L^2(\mathbb{R})$ by setting

$$\widetilde{\psi}_{j,k} = \begin{cases} 2^{j/2}\, T_k D_{2^{-j}}\, \psi^\ell, & j > 0, \\ D_{2^{-j}}\, T_k\, \psi^\ell, & j \leq 0. \end{cases}$$

These systems (as well as their higher-dimensional versions) were introduced by Ron and Shen in [205]. They pointed out that, unlike the affine systems, these systems are shift-invariant. Furthermore, the quasi-affine system $\{\widetilde{\psi}_{j,k}\}$ is a Parseval frame if and only if the corresponding affine system $\{\psi_{j,k}\}$ is a Parseval frame.

Recall that, in higher dimensions, affine and quasi-affine systems are typically defined using dilations of the form D_{M^j}, where M is an expanding matrix: that is, each proper value λ of M satisfies $|\lambda| > 1$. Notice that this condition is equivalent to the existence of constants k and γ, satisfying $0 < k \leq 1 < \gamma < \infty$, such that

$$|M^j x| \geq k\gamma^j\, |x| \tag{3.17}$$

when $x \in \mathbb{R}^n$, $j \in \mathbb{Z}$, $j \geq 0$, and

$$|M^j x| \leq \frac{1}{k}\gamma^j\, |x| \tag{3.18}$$

when $x \in \mathbb{R}^n$, $j \in \mathbb{Z}$, $j \leq 0$. One remarkable property of Theorem 3.1 is that it applies not only to the case of expanding-dilation matrices, but also to a more general class of dilations that are *expanding on a subspace* [137] and are defined as follows.

Definition 3.7. Given $M \in GL_n(\mathbb{R})$ and a nonzero linear subspace F of $\mathbb{R}^n$, we say that M *is expanding on* F if there exists a complementary (not necessarily orthogonal) linear subspace E of $\mathbb{R}^n$ with the following properties[1]

[1] This is the revised definition from [123]. It turned out that the definition initially proposed in [137], with a different condition (iv), was not sufficient to guarantee that the LIC was satisfied.

1. $\mathbb{R}^n = F + E$ and $F \cap E = \{0\}$; that is, for any $x \in \mathbb{R}^n$, there exist unique $x_F \in F$ and $x_E \in E$ such that $x = x_F + x_E$;
2. $M(F) = F$ and $M(E) = E$; that is, F and E are invariant under M;
3. conditions (3.17) and (3.18) hold for all $x \in F$;
4. for any $j \geq 0$, there exists $k_1 = k_1(M) > 0$ such that $|x_E| \leq k_1\,|M^j\,x_E|$.

It is clear that if a matrix M is expanding, then it is also expanding on a subspace. However, there are several examples of matrices that satisfy Definition 3.7 and are not expanding. For example, the following matrices are all expanding on a subspace:

- $M = \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix}$, where $a \in \mathbb{R}$, $|a| > 1$;
- $M = \begin{pmatrix} a & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$, where $a \in \mathbb{R}$, $|a| > 1$.

It is shown in [123, 137] that for affine systems where the dilation matrix M is expanding on a subspace, according to the definition above, then the LIC is "automatically" satisfied. Hence, Theorem 3.5 applies to this class of affine systems as well.

The examples seem to suggest that Theorem 3.5 applies whenever the dilation matrix M has all eigenvalues $|\lambda_k| \geq 1$ and at least one eigenvalue $|\lambda_1| > 1$. However, this is not the case. In [123] there is an example of a 3×3 dilation matrix having eigenvalues $\lambda_1 = a > 1$ and $\lambda_2 = \lambda_3 = 1$, for which the LIC fails. Indeed, it turns out that the information about the eigenvalues of M alone is not sufficient to determine the LIC or even the existence of corresponding affine systems. We refer to [144, 215] for additional results and observations about this topic.

3.3 Continuous Wavelet Transform

The *full affine group of motions on* $\mathbb{R}^n$, denoted by $\mathbf{A}_n$, consists of all pairs $(M,t) \in GL_n(\mathbb{R}) \times \mathbb{R}^n$ (endowed with the product topology) together with the group operation

$$(M,t)\cdot(M',t') = (MM', t' + (M')^{-1}t).$$

This operation is associated with the action $x \to M(x+t)$ on $\mathbb{R}^n$. The subgroup $\mathscr{N} = \{(M,t) \in \mathbf{A}_n : M = I,\ t \in \mathbb{R}^n\}$ is clearly a normal subgroup of $\mathbf{A}_n$.

We consider a class of subgroups $\{G\}$ of $\mathbf{A}_n$ of the form

$$G = \{(M,t) \in \mathbf{A}_n : M \in \mathscr{D},\ t \in \mathbb{R}^n\},$$

where $\mathscr{D}$ is a closed subgroup of $GL_n(\mathbb{R})$. We can identify $\mathscr{D}$ with the subgroup $\{(M,t) \in G : M \in \mathscr{D},\ t = 0\}$. Hence, we refer to $\mathscr{D}$ as the *dilation subgroup* and to $\mathscr{N}$ as the *translation subgroup* of G. If μ is the left Haar measure for $\mathscr{D}$, then $d\lambda(M,t) = d\mu(M)\,dt$ is the element of the left Haar measure for G.

Let U be the unitary representation of G acting on $L^2(\mathbb{R}^n)$ defined by

$$\left(U_{(M,t)}\,\psi\right)(x) = |\det M|^{-\frac{1}{2}}\,\psi(M^{-1}x - t) := \psi_{M,t}(x), \tag{3.19}$$

for $(M,t) \in G$ and $\psi \in L^2(\mathbb{R}^n)$. The elements $\{\psi_{M,t} : (M,t) \in G\}$ are the *continuous affine systems* with respect to G. The corresponding expression in the frequency domain is

$$\left(U_{(M,t)}\,\psi\right)^{\wedge}(\xi) = |\det M|^{\frac{1}{2}}\,\hat{\psi}(\xi M)\,e^{-2\pi i M t}.$$

For a fixed $\psi \in L^2(\mathbb{R}^n)$, the *wavelet transform* associated with G is the mapping

$$f \to (\mathscr{W}_\psi f)(M,t) = \langle f, \psi_{M,t}\rangle = |\det M|^{-\frac{1}{2}} \int_{\mathbb{R}^n} f(y)\,\overline{\psi(M^{-1}y - t)}\,dy,$$

where $f \in L^2(\mathbb{R}^n)$ and $(M,t) \in G$. If there exists a function $\psi \in L^2(\mathbb{R}^n)$ such that, for all $f \in L^2(\mathbb{R}^n)$, the reproducing formula

$$f = \int_G \langle f, \psi_{M,t}\rangle\,\psi_{M,t}\,d\lambda(M,t) \tag{3.20}$$

holds, then ψ is a *continuous wavelet* with respect to G. Expression (3.20) is a generalized version of the Calderón reproducing formula (3.6) presented in section 3.1. Notice that Eq. (3.20) is understood in the weak sense (see the proof of Theorem 3.8 below); the pointwise result is much more subtle.

The following theorem establishes an admissibility condition for ψ that guarantees that (3.20) is satisfied:

Theorem 3.8. *Equation (3.20) is valid for all $f \in L^2(\mathbb{R}^n)$ if and only if, for a.e. $\xi \in \mathbb{R}^n \setminus \{0\}$,*

$$\Delta_\psi(\xi) = \int_{\mathscr{D}} |\hat{\psi}(\xi M)|^2\,d\mu(M) = 1. \tag{3.21}$$

Proof. Suppose that (3.21) is satisfied. Then, by direct computation, we have that

$$\begin{aligned}
\|\mathscr{W}_\psi f\|^2_{L^2(G,\lambda)} &= \int_{\mathscr{D}} \int_{\mathbb{R}^n} |\langle f, \psi_{M,t}\rangle|^2\,dt\,d\mu(M)\\
&= \int_{\mathscr{D}} \int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} \hat{f}(\xi)\,\hat{\psi}(\xi M)\,e^{2\pi i \xi M t}\,d\xi \right|^2 |\det M|\,dt\,d\mu(M)\\
&= \int_{\mathscr{D}} \left(\int_{\mathbb{R}^n} \left| \left(\hat{f}(\xi)\overline{\hat{\psi}(\cdot M)}\right)^{\vee}(Mt) \right|^2 |\det M|\,dt \right) d\mu(M)\\
&= \int_{\mathscr{D}} \int_{\mathbb{R}^n} |\hat{f}(\xi)|^2 |\hat{\psi}(\xi M)|^2\,d\xi\,d\mu(M)\\
&= \int_{\mathbb{R}^n} |\hat{f}(\xi)|^2 \Delta_\psi(\xi)\,d\xi\\
&= \|f\|^2_{L^2(\mathbb{R}^n)}.
\end{aligned}$$

This shows that the mapping $\mathscr{W}_\psi : L^2(\mathbb{R}^n) \to L^2(G,\lambda)$ is an isometry. By polarization, we then obtain

$$\langle \mathscr{W}_\psi f, \mathscr{W}_\psi g\rangle_{L^2(G)} = \langle f, g\rangle_{L^2(\mathbb{R}^n)}, \tag{3.22}$$

for all $f,g \in L^2(\mathbb{R}^n)$.

Conversely, suppose that Eq. (3.20) holds in the weak sense [i.e., (3.22) holds]. Consider the expression

$$\int_{\mathbb{R}^n} |\hat{f}(\xi)|^2 \Delta_\psi(\xi)\, d\xi,$$

with f satisfying $|\hat{f}(\xi)|^2 = |\beta(r,\xi_0)|^{-1}\chi_{\beta(r,\xi_0)}(\xi)$, where $\beta(r,\xi_0)$ is a ball of radius r and center ξ_0, and ξ_0 is a point of differentiability of Δ_ψ. Then, by reversing the chain of equalities above, we obtain that

$$|\beta(r,\xi_0)|^{-1} \int_{\beta(r,\xi_0)} \Delta_\psi(\xi)\, d\xi = 1,$$

for all $r > 0$. By taking $\lim_{r\to 0^+}$, we conclude that $\Delta_\psi(\xi_0) = 1$. Thus, $\Delta_\psi(\xi) = 1$ for a.e. $\xi \in \mathbb{R}^n$. □

Theorem 3.8 can easily be extended to the case where G is not a subgroup of $GL_n(\mathbb{R})$, but simply a subset of $GL_n(\mathbb{R})$. Furthermore, Theorem 3.8 extends to functions on subspaces of $L^2(\mathbb{R}^n)$ of the form

$$L^2(V)^\vee = \{f \in L^2(\mathbb{R}^n) : \operatorname{supp} \hat{f} \subset V\}.$$

The proof of this fact is left as an exercise.

In the special case of Theorem 3.8 where $n = 1$ and $\mathscr{D} = \{2^j : j \in \mathbb{Z}\}$, Eq. (3.21) is $\sum_{j\in\mathbb{Z}} |\hat{\psi}(2^j\xi)|^2 = 1$ for a.e. $\xi \in \mathbb{R}$ (this is the classical Calderón equation), and Eq. (3.20) is

$$f = \sum_{j\in\mathbb{Z}} \int_{\mathbb{R}} \langle f, \psi_{j,t}\rangle\, \psi_{j,t}\, dt, \tag{3.23}$$

where $\psi_{j,t}(x) := 2^{-j/2}\psi(2^{-j}x - t)$, $j \in \mathbb{Z}$, $t \in \mathbb{R}$. Thus, the classical orthonormal wavelet expansion

$$f = \sum_{j\in\mathbb{Z}} \sum_{k\in\mathbb{Z}} \langle f, \psi_{jk}\rangle\, \psi_{jk}$$

is a "discretization" of (3.23). This shows, by Eq. (3.4), that an orthonormal wavelet (in this classical case) is always a continuous wavelet satisfying property (3.23) for all $f \in L^2(\mathbb{R})$. This raises the question of how to "discretize" continuous wavelets associated with general dilation groups $\mathscr{D}$. We refer to [234] for more observations about this topic.

A variant of the affine group $\mathbf{A}_n$ [and the corresponding affine systems (3.19)] is obtained by considering the group G^* consisting of all pairs $(M,t) \in GL_n(\mathbb{R}) \times \mathbb{R}^n$ (endowed with the product topology) together with the group operation

$$(M,t)\cdot(M',t') = (MM', t + M't').$$

This operation is associated with the action $x \to Mx + t$ on $\mathbb{R}^n$. The *co-affine systems* associated with G^* are then defined as the elements

$$\left(V_{(M,t)}\,\psi\right)(x) = |\det M|^{-\frac{1}{2}}\,\psi(M^{-1}(x-t)) := \psi^*_{M,t}(x),$$

for $(M,t) \in G^*$ and $\psi \in L^2(\mathbb{R}^n)$. The corresponding expression in the frequency domain is

$$\left(U_{(M,t)}\,\psi\right)^\wedge(\xi) = |\det M|^{\frac{1}{2}}\,\hat{\psi}(\xi M)\,e^{-2\pi i t}.$$

The left Haar measure, λ^*, for G^* is easily seen to satisfy

$$d\lambda^*(M,t) = |\det M|^{-1}\,d\mu(M)\,dt,$$

where μ is the left Haar measure for $\mathscr{D}$. Then the "co-affine" reproducing formula is

$$f = \int_G \langle f, \psi^*_{M,t}\rangle\,\psi^*_{M,t}\,d\lambda^*(M,t). \tag{3.24}$$

A straightforward calculation shows that (3.24) holds if and only if ψ satisfies condition (3.21). Thus, ψ *is a continuous affine wavelet if and only if it is a continuous co-affine wavelet.*

Notice that the situation observed above is different from the discrete case. In fact, consider the systems $\Psi = \{\psi_{j,k} = 2^{-j/2}\psi(2^{-j}\cdot - k) : j,k \in \mathbb{Z}\}$ and $\Psi^* = \{\psi^*_{j,k} = 2^{-j/2}\psi(2^{-j}(\cdot - k)) : j,k \in \mathbb{Z}\}$. A simple calculation shows that

$$\langle \psi^*_{j,k}, \psi_{-1,-1}\rangle = \langle \psi_{j,0}, \psi_{-1,2k-1}\rangle.$$

This shows that the co-affine systems cannot generate the space $L^2(\mathbb{R})$ if the corresponding affine system Ψ is an orthonormal basis for $L^2(\mathbb{R})$. In fact, since $2k-1$ is never 0, the affine system Ψ is an orthonormal basis for $L^2(\mathbb{R})$ (in which case the right-hand side of the above expression is zero) if and only if the co-affine system Ψ^* has a nonempty orthogonal complement.

3.3.1 Admissible Groups

It is not difficult to show that there are dilation groups $\mathscr{D}$ for which one can find no functions ψ satisfying Eq. (3.21). In particular, if $\mathscr{D}$ is compact, there are no associated functions ψ that satisfy this condition. For example, let $\mathscr{D} = SO(2)$ and suppose that there is a function $\psi \in L^2(\mathbb{R}^2)$ satisfying (3.21). Notice that, in this case, using polar coordinates Eq. (3.21) can be expressed as

$$\int_0^{2\pi} |\hat{\psi}(re^{i\phi}e^{i\theta})|^2\,\frac{d\theta}{2\pi} = 1,$$

for a.e. $\xi = re^{i\phi}$. Multiplying both sides of the equality by $r > 0$ and integrating with respect to $r \in [0, \infty)$, we obtain

$$\begin{aligned}\infty &= \int_0^\infty r\,dr \\ &= \int_0^\infty r \int_0^{2\pi} |\hat{\psi}(re^{i\phi}e^{i\theta})|^2 \frac{d\theta}{2\pi}\,dr \\ &= \int_0^\infty r \int_0^{2\pi} |\hat{\psi}(re^{i\theta})|^2 \frac{d\theta}{2\pi}\,dr \\ &= \|\psi\|^2 < \infty.\end{aligned}$$

This is clearly a contradiction and, thus, there is no ψ satisfying (3.21). In this situation, we say that the group $SO(2)$ is not *admissible*. That a general compact $\mathscr{D} \subset GL_n(\mathbb{R})$ is not admissible is not much harder to prove (see last paragraph in this section).

The observation above leads to the question: What are the groups $\mathscr{D}$ that are admissible? Our result on admissibility involves the notion of ε-*stabilizer* of $x \in \mathbb{R}^n$, which is defined as the set

$$\mathscr{D}_x^\varepsilon = \{M \in \mathscr{D} : |xM - x| \le \varepsilon\},$$

for each $\varepsilon > 0$. The set $\mathscr{D}_x := D_x^0 = \{M \in \mathscr{D} : xM = x\}$ is called the *stabilizer* of x. The *modular function* Δ, on $\mathscr{D}$, defined by the property

$$\mu(EM) = \Delta(M)\,\mu(E)$$

for all μ-measurable $E \subset \mathscr{D}$ and $M \in \mathscr{D}$, also plays an important role in the following basic result about admissible dilation groups.

Theorem 3.9. *(a) If $\mathscr{D}$ is admissible, then $\Delta \neq |\det M|$ and the stabilizer of x is compact for a.e. $x \in \mathbb{R}^n$.*

(b) If $\Delta \neq |\det M|$ and for a.e. $x \in \mathbb{R}^n$ there exists an $\varepsilon > 0$ such that the ε-stabilizer of x is compact, then $\mathscr{D}$ is admissible.

The proof of Theorem 3.9 is rather involved and can be found in [163]. Even though Theorem 3.9 "just fails" to be a characterization of admissibility, still it is quite useful for determining the admissibility or nonadmissibility of particular groups $\mathscr{D}$. For example, if $\mathscr{D}$ is compact, then $\Delta = |\det M| = 1$ and, thus, it cannot be admissible. Another example where Theorem 3.9 can be used effectively is the case where $\mathscr{D}$ is a one-parameter group. Namely, let $\mathscr{D} = \{M_t = e^{tL} : t \in \mathbb{R}\}$, where L is a real $n \times n$ matrix. Then $\mathscr{D}$ is admissible if and only if $\text{trace}(L) \neq 0$. Indeed, since $\det M_t = e^{t\,\text{trace}(L)}$ and $\mathscr{D}$ is Abelian, it follows that the modular function, Δ, is identically 1. Thus, when $\text{trace}(L) \neq 0$, we have that $\det M_t \neq 1 = \Delta$ and $\mathscr{D}$ is admissible.

3.3.2 Wave Packet Systems

In [57], Córdoba and Fefferman introduced "wave packets" as those families of functions obtained by applying certain collections of dilations, modulations, and translations to the Gaussian function. More generally, we will describe as "wave packet systems" any collections of functions that are obtained by applying a combination of dilations, modulations, and translations to a finite family of functions in $L^2(\mathbb{R}^n)$. For $\Psi = \{\psi^\ell : 1 \le \ell \le L\} \subset L^2(\mathbb{R}^n)$, where $L \in \mathbb{N}$, and $S \subset GL_n(\mathbb{R}) \times \mathbb{R}^n$, the *continuous wave packet system* with respect to S that is generated by Ψ is the collection

$$\mathscr{W}\mathscr{P}_S(\Psi) = \{D_A M_\nu T_y \psi^\ell : (A,\nu) \in S,\ y \in \mathbb{R}^n,\ 1 \le \ell \le L\}, \tag{3.25}$$

where M_ν is the modulation operator defined at the beginning of Section 3.2.

Let

$$G = \{U = c D_A M_\nu T_y : c \in \mathbb{C},\ |c| = 1,\ (A,\nu,y) \in GL_n(\mathbb{R}) \times \mathbb{R}^n \times \mathbb{R}^n\}.$$

G is a subgroup of the unitary operators on $L^2(\mathbb{R}^n)$ that is preserved by the action of the mapping $U \to \widehat{U}$, where $\widehat{U} f = (Uf)^\wedge$.

In the definition (3.25), we considered the map $(A,\nu,y) \to U^{(0)}_{(A,\nu,y)} = D_A M_\nu T_y$, which is a one-to-one mapping from $S \times \mathbb{R}^n$ into the group G. By changing the order of the operators, we can also define the following one-to-one mappings from $S \times \mathbb{R}^n$ into G:

$$\begin{aligned}
U^{(1)}_{(A,\nu,y)} &= D_A T_y M_\nu,\\
U^{(2)}_{(A,\nu,y)} &= T_y D_A M_\nu,\\
U^{(3)}_{(A,\nu,y)} &= M_\nu D_A T_y,\\
U^{(4)}_{(A,\nu,y)} &= T_y M_\nu D_A,\\
U^{(5)}_{(A,\nu,y)} &= M_\nu T_y D_A.
\end{aligned}$$

Hence, we can generate alternate continuous wave packet systems, $\mathscr{W}\mathscr{P}^{(i)}_S(\Psi)$, by replacing $U^{(0)}_{(A,\nu,y)}$ with $U^{(i)}_{(A,\nu,y)}$, for $1 \le i \le 5$. The systems $\mathscr{W}\mathscr{P}^{(0)}_S(\Psi)$ and $\mathscr{W}\mathscr{P}^{(1)}_S(\Psi)$ are equivalent in the sense that one is a Parseval frame if and only if the other one is a Parseval frame (in fact, by the commutativity relations of translations and modulations, they only differ by a unimodular scalar factor). The same is true for $\mathscr{W}\mathscr{P}^{(4)}_S(\Psi)$ and $\mathscr{W}\mathscr{P}^{(5)}_S(\Psi)$. The other systems, on the other hand, have substantial differences.

Each subgroup $U^{(i)}_{(A,\nu,y)}$, , $i = 0,\dots,5$, is associated with a continuous wave packet system generated by $\Psi \subset L^2(\mathbb{R}^n)$. We can characterize those Ψ for which we have Parseval frames:

$$\sum_{\ell=1}^{L} \int_{S\times\mathbb{R}^n} \left|\langle f, U^{(i)}_{(A,\nu,y)\psi^\ell}\rangle\right|^2 d\lambda(A,\nu)\,dy = \|f\|_2^2$$

for all $f \in L^2(\mathbb{R}^n)$, where λ is a measure on S. Such a characterization is an extension of Theorem 3.8 and is given by an analogue of Eq. (3.21). Explicitly, we have the result:

Theorem 3.10. *Let* $\Psi = \{\psi^\ell : 1 \le \ell \le L\} \subset L^2(\mathbb{R}^n)$*. The systems* $\mathscr{WP}_S^{(i)}(\Psi)$*,* $i = 0,\dots,5$*, are continuous Parseval frame wave packet systems with respect to* (S,λ)*, for* $L^2(\mathbb{R}^n)$*, if and only if*

$$\Delta_\Psi^{(i)}(\xi) = 1 \quad \textit{for a.e. } \xi \in \mathbb{R}^n,$$

where

$$\Delta_\Psi^{(1)}(\xi) = \Delta_\Psi^{(0)}(\xi) = \sum_{\ell=1}^L \int_S |\hat\psi^\ell(\xi A^{-1} - \nu)|^2 \, d\lambda(A,\nu);$$

$$\Delta_\Psi^{(2)}(\xi) = \sum_{\ell=1}^L \int_S |\hat\psi^\ell(\xi A^{-1} - \nu)|^2 \, |\det A|^{-1} \, d\lambda(A,\nu);$$

$$\Delta_\Psi^{(3)}(\xi) = \sum_{\ell=1}^L \int_S |\hat\psi^\ell((\xi - \nu)A^{-1})|^2 \, d\lambda(A,\nu);$$

$$\Delta_\Psi^{(4)}(\xi) = \Delta_\Psi^{(5)}(\xi) = \sum_{\ell=1}^L \int_S |\hat\psi^\ell((\xi - \nu)A^{-1})|^2 \, |\det A|^{-1} \, d\lambda(A,\nu).$$

3.4 Affine Systems with Composite Dilations

To describe the class of systems that will be considered in this section, it will be useful to begin with one example in $L^2(\mathbb{R}^2)$.

Let $A = \begin{pmatrix} 2 & 0 \\ 0 & \varepsilon \end{pmatrix}$, where $\varepsilon \ne 0$, $B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $G = \{(B^j,k) : j \in \mathbb{Z},\ k \in \mathbb{Z}^2\}$. Then G is a group with group multiplication:

$$(B^\ell,m)\,(B^j,k) = (B^{\ell+j}, k + B^{-j}m). \tag{3.26}$$

In particular, we have $(B^j,k)^{-1} = (B^{-j}, -B^j k)$. The multiplication (3.26) is consistent with the operation that maps $x \to B^j(x+k)$ of $\mathbb{R}^2$ into $\mathbb{R}^2$. Let π be the unitary representation of G, acting on $L^2(\mathbb{R}^2)$, which is defined by

$$\Big(\pi(B^j,k)f\Big)(x) = f((B^j,k)^{-1}x) = f(B^{-j}x - k) = \Big(D_B^j T_k f\Big)(x), \tag{3.27}$$

for $f \in L^2(\mathbb{R}^2)$. Notice that $\det B^j = 1$. The observation that

$$(D_B^\ell T_m)\,(D_B^j T_k) = (D_B^{\ell+j} T_{k+B^{-j}m}),$$

where $\ell, j \in \mathbb{Z}$, $k,m \in \mathbb{Z}^2$, shows how the group operation (3.26) is associated with the unitary representation (3.27).

Let $S_0 = \{\xi = (\xi_1,\xi_2) \in \widehat{\mathbb{R}}^2 : |\xi_1| \le 1\}$ and define

$$V_0 = L^2(S_0)^\vee = \{f \in L^2(\mathbb{R}^n) : \operatorname{supp}\hat f \subset S_0\}.$$

Since, for all $j \in \mathbb{Z}$ and $k \in \mathbb{Z}^2$, we have[2]

$$\left(\pi(B^j,k)f\right)^{\wedge}(\xi) = \left(D_B^j T_k f\right)^{\wedge}(\xi) = e^{-2\pi i \xi B^j k}\hat{f}(\xi B^j),$$

and $\xi B^j = (\xi_1, \xi_2)B^j = (\xi_1, \xi_2 + j\xi_1)$, then the action of B^j maps the vertical strip domain S_0 into itself and, thus, the space V_0 is invariant under the action of $\pi(B^j,k)$. The same invariance property holds for the vertical strips

$$S_i = S_0 A^i = \{\xi = (\xi_1, \xi_2) \in \widehat{\mathbb{R}}^2 : |\xi_1| \le 2^i\},$$

$i \in \mathbb{Z}$, and, as a consequence, the spaces $V_i = L^2(S_i)^\vee$ are also invariant under the action of the operators $\pi(B^j,k)$.

The spaces $\{V_i\}_{i\in\mathbb{Z}}$ defined above satisfy the basic MRA properties:

1. $V_i \subset V_{i+1}$, $i \in \mathbb{Z}$;
2. $D_A^{-i} V_0 = V_i$;
3. $\bigcap_{i\in\mathbb{Z}} V_i = \{0\}$;
4. $\overline{\bigcup_{i\in\mathbb{Z}} V_i} = L^2(\mathbb{R}^n)$.

The complete definition of an MRA includes the assumption that V_0 is generated by the integer-translates of a $\phi \in V_0$, called the *scaling function*, and that these translates $\{T_k \phi : k \in \mathbb{Z}^2\}$ are an orthonormal basis of V_0. In some cases, there is more than one scaling function.

The situation here is a bit different, and the *scaling* property is replaced by an analogous property. Namely, consider $\widehat{V_0} = L^2(S_0)$ and let $\hat{\phi} = \chi_J$, where $J = J^+ \cup J^-$, J^+ is the triangle with vertices $(0,0)$, $(1,0)$, $(1,1)$ and J^- is the triangle with vertices $(0,0)$, $(-1,0)$, $(-1,-1)$. The sets JB^j, $j \in \mathbb{Z}$, form a partition of S_0; that is, $S_0 = \bigcup_{j\in\mathbb{Z}} JB^j$, except for the set of points $\{(0,\xi_2) : \xi_2 \neq 0\}$, which is, however, a set of measure 0. The set J has measure 1 and the collection $\{e^{-2\pi ik\xi}\chi_J : k \in \mathbb{Z}^2\}$ is easily seen to be an ON basis of $L^2(J)$. Since

$$\left(e^{-2\pi ik\cdot}\chi_J(\cdot) :\right)^{\vee}(x) = (T_k\phi)(x) = \phi(x-k),$$

these last functions form an ON basis of $L^2(J)^\vee$. It follows that $\{D_{B^j} T_k \phi : k \in \mathbb{Z}^2\}$ is an ON basis of $L^2(JB^j)^\vee$, for each $j \in \mathbb{Z}^2$. Hence, the set

$$\left\{D_{B^j} T_k \phi : j \in \mathbb{Z},\ k \in \mathbb{Z}^2\right\} = \left\{T_k D_{B^j} \phi : j \in \mathbb{Z},\ k \in \mathbb{Z}^2\right\}$$

is an ON basis of V_0. The sets J^+, J^-, as well as the other sets used in this construction, are illustrated in Fig. 3.1.

Thus, the “complete” definition of the MRA, introduced above, adds to properties 1–4 the property

---

[2] Recall that, according to the notation introduced in Section 3.2, in the frequency domain, the matrices B^j multiply row vectors on the right.

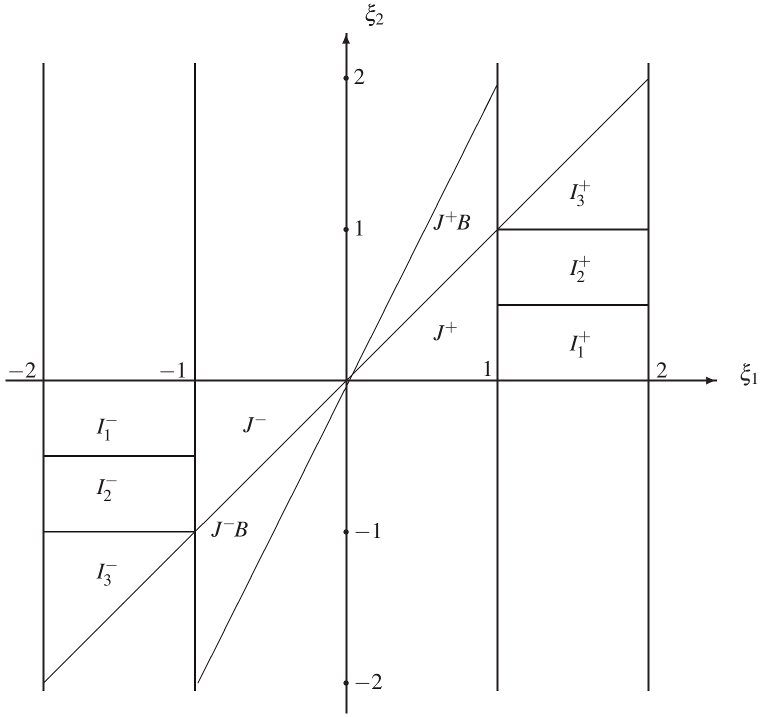


Fig. 3.1: Example of ON AB–MRA. The sets $\{J^+ B^j, J^- B^j : j \in \mathbb{Z}\}$ form a disjoint partition of S_0.

5. V_0 is generated by a "scaling function" ϕ, in the sense that $\{T_k D_{B^j} \phi : j \in \mathbb{Z}, k \in \mathbb{Z}^2\}$ is an ON basis of V_0.

Let G_B be the group $\{B^j : j \in \mathbb{Z}\}$; this is equivalent to the dilation group $\{D_{B^j} : j \in \mathbb{Z}\}$. Then $G = \{(B^j, k) : j \in \mathbb{Z},\ k \in \mathbb{Z}^2\}$ is the semidirect product of G_B and $\mathbb{Z}^2$, denoted by $G_B \ltimes \mathbb{Z}^2$. This shows that the *shift invariance* of the traditional MRA is replaced by a notion of $G_B \ltimes \mathbb{Z}^n$ *invariance*, that is, the space V_0 is invariant with respect to both integer translations and G_B dilations.

We shall now show how the MRA we just introduced can be used to construct a wavelet-like basis of $L^2(\mathbb{R}^2)$. We begin by constructing an ON basis of W_0, defined to be the orthogonal complement of V_0 in V_1, that is, $V_1 = V_0 \oplus W_0$. It will be convenient to work in the frequency domain. We have that $\widehat{V_1} = \widehat{V_0} \oplus \widehat{W_0}$ and, consequently, $\widehat{W_0} = L^2(R_0)$, where $R_0 = S_1 \setminus S_0 = \{\xi = (\xi_1, \xi_2) \in \widehat{\mathbb{R}}^2 : 1 < |\xi_1| \le 2\}$. We define the following subsets of $R_0 = S_1 \setminus S_0$:

$$I_1 = I_1^+ \cup I_1^-, I_2 = I_2^+ \cup I_2^-, I_3 = I_3^+ \cup I_3^-,$$

where

$$I_1^+ = \left\{\xi = (\xi_1, \xi_2) \in \widehat{\mathbb{R}}^2 : 1 < \xi_1 \le 2,\ 0 \le \xi_2 < \frac{1}{2}\right\},$$

$$I_2^+ = \left\{ \xi = (\xi_1, \xi_2) \in \widehat{\mathbb{R}}^2 : 1 < \xi_1 \le 2, \ \frac{1}{2} \le \xi_2 < 1 \right\},$$

$$I_3^+ = \left\{ \xi = (\xi_1, \xi_2) \in \widehat{\mathbb{R}}^2 : 1 < \xi_1 \le 2, \ 1 \le \xi_2 < \xi_1 \right\},$$

and $I_\ell^- = \{\xi \in \widehat{\mathbb{R}}^2 : -\xi \in I_\ell^+\}$, $\ell = 1,2,3$. These sets are illustrated in Fig. 3.1 and 3.2. Observe that each set I_ℓ is a *fundamental domain* for $\mathbb{Z}^2$: The functions $\{e^{2\pi i \xi k} : k \in \mathbb{Z}^2\}$, restricted to I_ℓ, form an ON basis for $L^2(I_\ell)$, $\ell = 1,2,3$. We then define ψ^ℓ, $\ell = 1,2,3$, by setting $\hat{\psi}^\ell = \chi_{I_\ell}$, $\ell = 1,2,3$. It follows from the observations about the sets $\{I_\ell\}$ that the collection

$$\{e^{2\pi i \xi k}\, \hat{\psi}^\ell(\xi) : k \in \mathbb{Z}^2\}$$

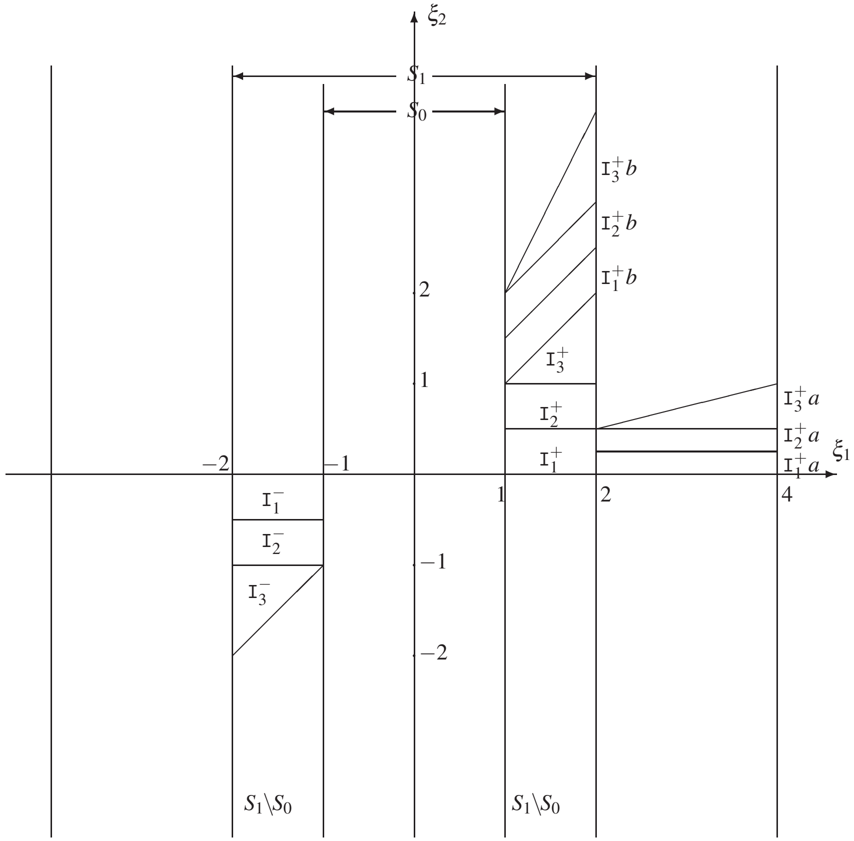


Fig. 3.2: Example of an orthonormal *AB* wavelet.

is an orthonormal basis of $L^2(I_\ell)$, $\ell = 1,2,3$. A simple direct calculation shows that the sets $\{I_\ell b^j : j \in \mathbb{Z}, \ell = 1,2,3\}$ are a partition of R_0, that is,

$$\bigcup_{\ell=1}^{3} \bigcup_{j\in\mathbb{Z}} I_\ell B^j = R_0,$$

where the union is disjoint. As a consequence, the collection

$$\{e^{2\pi i \xi k}\, \hat{\psi}^\ell(\xi B^j) : k \in \mathbb{Z}^2,\ j \in \mathbb{Z},\ \ell = 1,2,3\} \tag{3.28}$$

is an orthonormal basis of $L^2(R_0)$ and, thus, by taking the inverse Fourier transform of (3.28), we have that

$$\{\pi(B^j,k)\,\psi^\ell : k \in \mathbb{Z}^2,\ j \in \mathbb{Z},\ \ell = 1,2,3\} \tag{3.29}$$

is an orthonormal basis of $W_0 = L^2(R_0)^\vee$. Notice that, since, for each $j \in \mathbb{Z}$ fixed, B^j maps $\mathbb{Z}^2$ into itself, the collection $\{e^{2\pi i \xi B^j k} : k \in \mathbb{Z}^2\}$ is equal to the collection $\{e^{2\pi i \xi k} : k \in \mathbb{Z}^2\}$.

It is clear that, by applying the dilations D_{A^i}, $i \in \mathbb{Z}$, to the system (3.29), we obtain an ON basis of $L^2(R_i)^\vee$, where

$$R_i = R_0 A^i = \{\xi = (\xi_1,\xi_2) \in \widehat{\mathbb{R}}^2 : 2^i < |\xi_1| \le 2^{1+i}\}.$$

Furthermore, we have that $\bigcup_{i\in\mathbb{Z}} R_i = \widehat{\mathbb{R}}^2$, where the union is disjoint, and hence we can write $L^2(\mathbb{R}^2) = \bigoplus_{i\in\mathbb{Z}} W_i$. Hence, by combining the observations above, it follows that the collection

$$\{D_{A^i} D_{B^j} T_k\, \psi^\ell : k \in \mathbb{Z}^2,\ i,j \in \mathbb{Z},\ \ell = 1,2,3\} \tag{3.30}$$

is an ON basis of $L^2(\mathbb{R}^2)$.

3.4.1 Affine System with Composite Dilations

The construction given above is a particular example of a general class of affine-like systems called *affine systems with composite dilations*, which have the form

$$\mathscr{A}_{AB}(\Psi) = \{D_A D_B T_k\, \psi^\ell : A \in G_A, B \in G_B,\ k \in \mathbb{Z}^n,\ \ell = 1,\dots,L\}, \tag{3.31}$$

where $\Psi \subset \{\psi^1,\dots,\psi^\ell\} \in L^2(\mathbb{R}^n)$, $G_A \subset GL_n(\mathbb{R})$ (usually, $G_A = \{A^i : i \in \mathbb{Z}\}$, with A expanding or having some "expanding" property), and $G_B \subset GL_n(\mathbb{R})$ with $|\det B| = 1$. Later on, we will show that there are several examples of such systems that form ON bases of $L^2(\mathbb{R}^n)$ or, more generally, Parseval frames of $L^2(\mathbb{R}^n)$.

The roles played by the two families of dilations, G_A and G_B, in definition (3.31), are very different. The elements $A \in G_A$ dilate (at least in some direction), while the elements of G_B affect the geometry of the reproducing system $\mathscr{A}_{AB}(\Psi)$. In the example we worked out, $G_B = \{\left(\begin{smallmatrix}1 & 1\\ 0 & 1\end{smallmatrix}\right)^j : j \in \mathbb{Z}\}$ is the *shear group* and exhibits a "shear geometry," in which objects in the plane are stretched vertically without increasing their size (like the trapezoids in Fig. 3.2). In Section 3.5, we will use this group and a construction similar to the one above to obtain the *shearlets*, whose geometrical properties are similar to the example above and are, in addition, well-localized functions (i.e., they have rapid decay in both the space and frequency domains). They have similarities to the *curvelets* introduced by Candès and Donoho [38] and to the *contourlets* of Do and Vetterli [68]. However, their mathematical construction is simpler, since it derived from the structure of affine systems, and, as a result, their development and applications are "more systematic" [129, 130].

As indicated by the example above, there is a special multiresolution analysis associated with the affine systems with composite dilations that is useful for constructing "composite wavelets." Let us give a proper definition of this new framework. Let G_B be a countable subset of $\widetilde{SL}_n(\mathbb{Z}) = \{B \in GL_n(\mathbb{R}) : |\det B| = 1\}$ and $G_A = \{A^i : i \in \mathbb{Z}\}$, where $A \in GL_n(\mathbb{Z})$ (notice that A is an *integral* matrix). Also, assume that A *normalizes* G_B, that is, $ABA^{-1} \in G_B$ for every $B \in G_B$, and that the quotient space $B/(ABA^{-1})$ is finite. Then the sequence $\{V_i\}_{i\in\mathbb{Z}}$ of closed subspaces of $L^2(\mathbb{R}^n)$ is an *AB-multiresolution analysis (AB-MRA)* if the following hold:

1. $D_B T_k V_0 = V_0$, *for any* $B \in G_B$, $k \in \mathbb{Z}^n$
2. *For each* $i \in \mathbb{Z}$, $V_i \subset V_{i+1}$, *where* $V_i = D_A^{-i} V_0$
3. $\bigcap V_i 1, = \{0\}$ *and* $\overline{\bigcup V_i} = L^2(\mathbb{R}^n)$
4. *There exists* $\phi \in L^2(\mathbb{R}^n)$ *such that* $\Phi_B = \{D_B T_k \phi : B \in G_B, k \in \mathbb{Z}^n\}$ *is a semi-orthogonal Parseval frame for* V_0, *that is,* Φ_B *is a Parseval frame for* V_0 *and, in addition,* $D_B T_k \phi \perp D_{B'} T_{k'} \phi$ *for any* $B \neq B'$, $B, B' \in G_B$, $k, k' \in \mathbb{Z}^n$.

The space V_0 is called an *AB scaling space* and the function ϕ is an *AB scaling function* for V_0. In addition, if Φ_B is an orthonormal basis for V_0, then ϕ is an *orthonormal AB scaling function*.

The number of generators L of an orthonormal MRA *AB*-wavelet is completely determined by the group $G = \{(B^j, k) : j \in \mathbb{Z}, k \in \mathbb{Z}^n\}$. Indeed, we have the following simple fact:

Proposition 3.11. *Let G be a countable group and $u \to T_u$ be a unitary representation of G acting on a (separable) Hilbert space $\mathscr{H}$. Suppose $\Phi = \{\phi^1, \ldots, \phi^N\}$, $\Psi = \{\psi^1, \ldots, \psi^M\} \subset \mathscr{H}$, where $N, M \in \mathbb{N} \bigcup \{\infty\}$. If $\{T_u \phi^k : u \in G,\ 1 \le k \le N\}$ and $\{T_u \psi^i : u \in G,\ 1 \le i \le M\}$ are each orthonormal bases for $\mathscr{H}$, then $N = M$.*

Proof. It follows from the assumptions that, for each $1 \le k \le N$

$$\|\phi^k\|^2 = \sum_{u\in G} \sum_{i=1}^{M} |\langle \phi^k, T_u \psi^i \rangle|^2.$$

Thus, by the properties of T_u, we have

$$\begin{aligned} N = \sum_{k=1}^{N} \|\phi^k\|^2 &= \sum_{k=1}^{N} \sum_{u\in G} \sum_{i=1}^{M} |\langle \phi^k, T_u \psi^i \rangle|^2 \\ &= \sum_{i=1}^{M} \sum_{u\in G} \sum_{k=1}^{N} |\langle T_{u^{-1}} \phi^k, \psi^i \rangle|^2 \\ &= \sum_{i=1}^{M} \|\psi^i\|^2 = M. \quad \square \end{aligned}$$

Using Proposition 3.11, one obtains the following result, which establishes the number of generators needed to obtain an orthonormal MRA AB-wavelet.

Theorem 3.12. *Let $\Psi = \{\psi^1, \ldots, \psi^L\}$ be an orthonormal MRA AB-multiwavelet for $L^2(\mathbb{R}^n)$, and let $N = |B/ABA^{-1}|$ (= the order of the quotient group B/ABA^{-1}). Assume that $|\det A| \in \mathbf{N}$. Then $L = N|\det A| - 1$.*

The composite wavelet system $\mathscr{A}_{AB}(\Psi)$ has associated continuous multiwavelets. The simplest case is the one in which the translations are $\{T_y : y \in \mathbb{R}^n\}$. In this case, we have the reproducing formula corresponding to (3.20):

$$f = \sum_{\ell=1}^{L} \sum_{i,j\in\mathbb{Z}} \int_{\mathbb{R}^n} \langle f, D_{A^i} D_{B^j} T_y \psi^\ell \rangle D_{A^i} D_{B^j} T_y \psi^\ell \, dy, \tag{3.32}$$

for $f \in L^2(\mathbb{R}^n)$. As in Section 3.3, one can show that $\Psi = \{\psi^1, \ldots, \psi^\ell\}$ satisfies (3.32) if and only if it satisfies the Calderón equation

$$\sum_{\ell=1}^{L} \sum_{i,j\in\mathbb{Z}} |\hat{\psi}^\ell(\xi A^i B^j)| = 1 \quad \text{for a.e. } \xi \in \mathbb{R}^n.$$

Some more general examples of continuous composite wavelet systems will be examined in Section 3.5.

3.4.2 Other Examples

There are several other examples of affine systems with composite dilations $\mathscr{A}_{AB}(\Psi)$ that form ON bases or Parseval frames.

In particular, the construction presented above in dimension $n = 2$ extends to the general n-dimensional setting. In this case, the shear group is given by $G_B = \{B^i : i \in \mathbb{Z}\}$, where $B \in GL_n(\mathbb{R})$ is characterized by the equality $(B - I_n)^2 = 0$, and I_n is the $n \times n$ identity matrix. We refer to [130] for more details about these systems.

A different type of affine system with composite dilations arises when G_B is a finite group. For example, let $G_B = \{\pm B_0, \pm B_1, \pm B_2, \pm B_3\}$ be the eight-element group consisting of the isometries of the square $[-1,1]^2$. Specifically: $B_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $B_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $B_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, $B_3 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$. Let U be the parallelogram with vertices $(0,0), (1,0), (2,1)$, and $(1,1)$ and $S_0 = \bigcup_{b\in B} U b$ (see the snowflake region in Fig. 3.3). It is easy to verify that S_0 is B-invariant.

Let A be the quincunx matrix $\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$, and $S_i = S_0 A^i$, $i \in \mathbb{Z}$. Observe that A is expanding, $ABA^{-1} = B$, and $S_0 \subseteq S_0 A = S_1$. In particular, the region $S_1 \setminus S_0$ is the disjoint union $\bigcup_{b\in B} RB$, where the region R is the parallelogram illustrated in Fig. 3.3. Thus, as in the case of the shear composite wavelet that we described

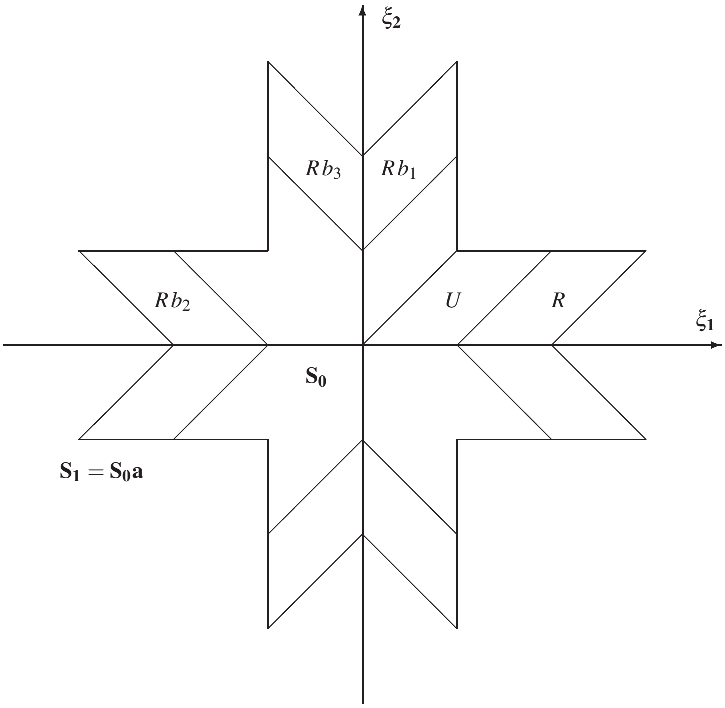


Fig. 3.3: Example of composite wavelet with finite group. $G_A = \{A^i : i \in \mathbb{Z}\}$, where A is the quincunx matrix, and G_B is the group of isometries of the square $[-1,1]^2$.

above, it follows that the system

$$\{D_A^i D_B T_k \psi : i \in \mathbb{Z},\ B \in G_B,\ k \in \mathbb{Z}^2\}, \tag{3.33}$$

where $\hat{\psi} = \chi_R$, is an orthonormal basis for $L^2(\mathbb{R}^2)$.

If the quincunx matrix A is replaced by the matrix $\tilde{A} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$, we obtain a different ON basis. Let B, U, and S_i, $i \in \mathbb{Z}$, be defined as above. Also, in this case, $\tilde{A}$ is expanding, $\tilde{A} B \tilde{A}^{-1} = B$, and $S_1 = S_0 a \supset S_0$. A direct computation shows that the region $S_1 \setminus S_0$ is the disjoint union $\bigcup_{B \in G_B} Rb$, where $R = R_1 \bigcup R_2 \bigcup R_3$ and the regions R_1, R_2, R_3 are illustrated in Fig. 3.4. Observe that each of the regions R_1, R_2, R_3 is a fundamental domain. Thus, the system

$$\{D_{\tilde{A}}^i D_B T_k \psi^\ell : i \in \mathbb{Z},\ B \in G_B,\ k \in \mathbb{Z}^2,\ \ell = 1, \ldots, 3\}, \tag{3.34}$$

where $\hat{\psi}^\ell = \chi_{R_\ell}$, $\ell = 1,2,3$, is an orthonormal basis for $L^2(\mathbb{R}^2)$.

Note that the system in the first example [Eq. (3.33)] was generated by a single function, while the second system [Eq. (3.34)] is generated by three functions ψ^1, ψ^2, ψ^3. This is consistent with Theorem 3.12. In fact, if B is a finite group,

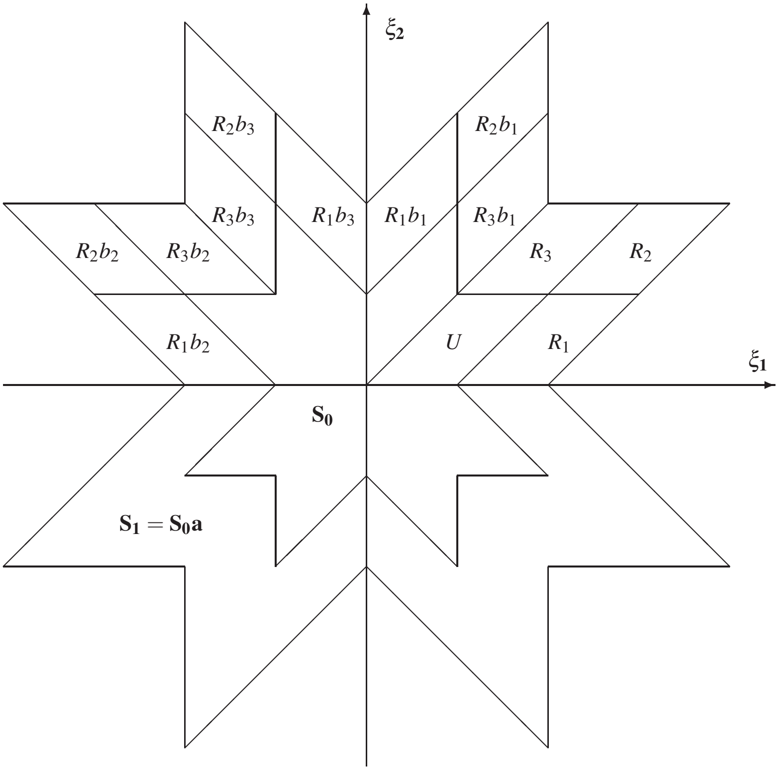


Fig. 3.4: Example of composite wavelet with finite group. $G_A = \{A^i : i \in \mathbb{Z}\}$, where $A = 2I$, and G_B is the group of isometries of the square $[-1,1]^2$.

then $N = |B/ABA^{-1}| = 1$, and so, in this situation, the number of generators is $L = |\det A| - 1$. Thus, by Theorem 3.12, in the first example, we obtain that the number of generators is $L = 1$ since A is the quincunx matrix and $\det A = 2$. In the second example, the number of generators is $L = 3$ since $\tilde{A} = 2I$ and $\det\tilde{A} = 4$. Finally, in the example at beginning of this section, where G_B is the two-dimensional group of shear matrices and $G_A = \{A^i : i \in \mathbb{Z}\}$, with $A = \begin{pmatrix} 2 & A_{1,2} \\ 0 & A_{2,2} \end{pmatrix} \in GL_2(\mathbb{Z})$, a calculation shows that $|B/ABA^{-1}| = 2|A_{2,2}|^{-1}$ and, thus, the number of generators is $L = 2|A_{2,2}|^{-1}\, 2|A_{2,2}| - 1 = 3$.

In higher dimensions, the type of constructions we have just described extends by using the Coxeter groups. These are finite groups (hence, their elements have determinant 1 in magnitude) generated by reflections through hyperplanes.

Other examples of composite wavelets, in dimension $n = 2$, are obtained, for each $\lambda > 1$ fixed, by considering the group

$$G_B = \left\{ B_j = \begin{pmatrix} \lambda^j & 0 \\ 0 & \lambda^{-j} \end{pmatrix} : j \in \mathbb{Z} \right\}$$

and choosing G_A to be a group of expanding matrices; for example, $G_A = \{A^i : i \in \mathbb{Z}\}$, where A is diagonal and $|\det A| > 1$. We refer to [130] for more details about this construction.

All examples of composite wavelets presented so far are "direct" constructions in the frequency domain. Let us now discuss a different class of composite wavelets in the "time domain."

Perhaps the simplest dyadic-dilation wavelet in dimension $n = 1$ is the Haar wavelet. It is produced by the scaling function $\phi = \chi_{[0,1)}$ and is generated by the Haar function $\psi = \chi_{[0,\frac{1}{2})} - \chi_{[\frac{1}{2},1)}$. The *Haar ON basis* of $L^2(\mathbb{R})$ is the affine system

$$\{\psi_{i,k} = D_{2^i} T_k \psi : i,k \in \mathbb{Z}\}.$$

It is a natural question to ask what the extensions are of this compactly supported wavelet ψ in higher dimensions. For example, in dimension $n = 2$, consider the quincunx matrix $A_q = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ and the associated affine system

$$\{\psi_{i,k} = D_{A_q^i} T_k \psi : i \in \mathbb{Z},\ k \in \mathbb{Z}^2\}. \tag{3.35}$$

Then, similarly to the one-dimensional Haar wavelet, one can find an MRA wavelet ψ produced by a scaling function ϕ that is the characteristic function of a compact set $Q \subset \mathbb{R}^2$ of area 1. However, the functions ϕ and ψ are not that simple. In fact, the scaling function ϕ is the characteristic function of a rather complicated fractal set known at the "twin dragon" and ψ is the difference of two similar characteristic functions (see Fig. 3.5).

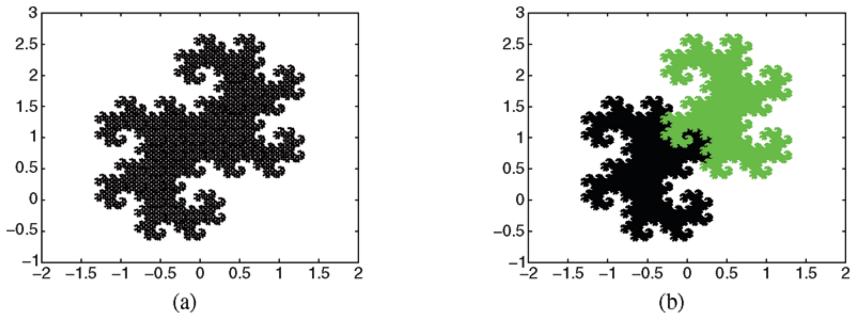


Fig. 3.5: (a) The fractal set known as "twin dragon." (b) Support of the two-dimensional Haar wavelet ψ; $\psi = 1$ on the darker set, $\psi = -1$ on the lighter set.

We can construct an affine system with composite dilations having the same expanding dilation group $G_A = \{A_q^i : i \in \mathbb{Z}\}$ and the same translations that does, however, generate a very simple Haar-type wavelet. For the group G_B, let us again choose the group of symmetries of the unit square given at the beginning of this section. Let R_0 be the triangle with vertices $(0,0)$, $(1/2,0)$, $(1/2,1/2)$ and

$R_\ell = B_\ell R_0$, $\ell = 1,\dots,7$ (see Fig. 3.6). Then, for $\phi = 2\sqrt{2}\chi_{R_0}$, it follows that the system

$$\{D_{B_\ell} T_k \phi : \ell = 0,\dots,7,\ k \in \mathbb{Z}^2\}$$

is an ON basis for the space V_0, which is the closed linear span of the subspace of $L^2(\mathbb{R}^2)$ consisting of the functions that are constant on each $\mathbb{Z}^2$-translate of the triangles R_ℓ, $\ell = 0,1,\dots,7$. Let us now consider the spaces $V_i = D_{A_q^{-i}} V_0$, $i \in \mathbb{Z}$. Then one can verify that each space V_i is the closed linear span of the subspace of $L^2(\mathbb{R}^2)$ consisting of the functions that are constant on each $A_q^{-i}\mathbb{Z}^2$-translate of the triangles $A_q^{-i} R_\ell$, $\ell = 0,1,\dots,7$. Thus, $V_i \subset V_{i+1}$ for each $i \in \mathbb{Z}$, and the spaces $\{V_i\}$ form an AB-MRA, with ϕ as an AB-scaling function. We can now construct a simple Haar-like wavelet obtained from this AB-MRA. Specifically, let

$$R_0 = A_q^{-1} R_1 \cup \left[A_q^{-1} R_6 + \frac{1}{2}\begin{pmatrix}1\\1\end{pmatrix} \right] = A_q^{-1} R_1 \cup A_q^{-1}\left[R_6 + \begin{pmatrix}0\\1\end{pmatrix} \right].$$

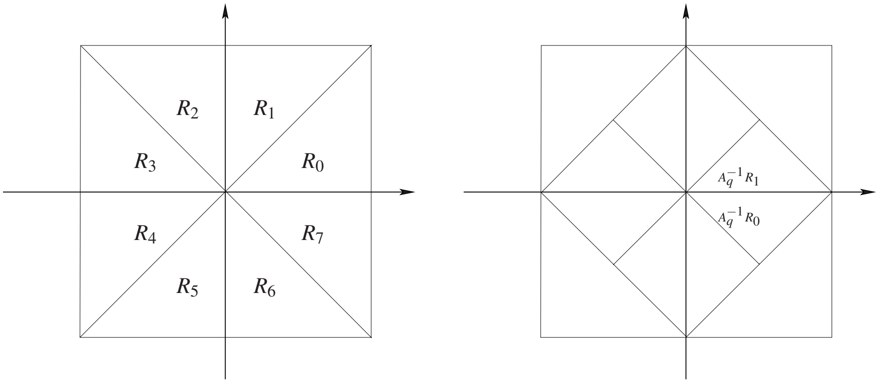


Fig. 3.6: Example of a composite wavelet with finite support.

Thus, $\chi_{R_0} = \chi_{A_q^{-1}R_1} + \chi_{A_q^{-1}R_6 + \frac{1}{2}\binom{1}{1}}$, or, equivalently,

$$\phi^{(0)}(x) = \phi^{(1)}(A_q x) + \phi^{(6)}(A_q x - \tbinom{0}{1}), \tag{3.36}$$

where $\phi^{(\ell)} = D_{B_\ell}\phi$, for $\ell = 0,1,\dots,7$. It is now easy to see that $\psi = \phi^{(1)}(A_q x) - \phi^{(6)}(A_q x - \binom{0}{1})$ is the desired Haar-like AB-wavelet. The space V_0 is generated by applying the translations T_k, $k \in \mathbb{Z}^2$, to the scaling functions $\phi^{(\ell)} = D_{B_\ell}\phi$, $\ell = 0,1,\dots,7$. We see that this is the case by applying D_{B_ℓ} in Eq. (3.36); we obtain

$$\phi^{(0)} = \phi^{(1)}(A_q x) + \phi^{(6)}(A_q x - \tbinom{0}{1}),$$

$$\phi^{(1)} = \phi^{(2)}(A_q x) + \phi^{(5)}(A_q x - \binom{0}{1}),$$
$$\phi^{(2)} = \phi^{(3)}(A_q x) + \phi^{(0)}(A_q x - \binom{0}{1}),$$
$$\phi^{(3)} = \phi^{(4)}(A_q x) + \phi^{(7)}(A_q x - \binom{0}{1}),$$
$$\phi^{(4)} = \phi^{(5)}(A_q x) + \phi^{(2)}(A_q x - \binom{0}{1}),$$
$$\phi^{(5)} = \phi^{(6)}(A_q x) + \phi^{(1)}(A_q x - \binom{0}{1}),$$
$$\phi^{(6)} = \phi^{(7)}(A_q x) + \phi^{(4)}(A_q x - \binom{0}{1}),$$
$$\phi^{(7)} = \phi^{(0)}(A_q x) + \phi^{(3)}(A_q x - \binom{0}{1}).$$

It follows that

$$\{D_{A_q^i} D_{B_\ell} T_k \psi : i \in \mathbb{Z},\ \ell = 0,1,\ldots,7,\ k \in \mathbb{Z}^2\}$$

is an ON basis for $L^2(\mathbb{R}^2)$. This Haar-type *AB*-wavelet is clearly simpler that the twin dragon wavelet obtained earlier. We refer to [23, 154] for more information about this type of construction.

Other complicated fractal wavelets appear in many situations. For example, if the dilation matrix A_q in the affine system (3.35) is replaced by $A_{q1} = \begin{pmatrix} 1 & \sqrt{3} \\ -\sqrt{3} & 1 \end{pmatrix}$ or $A_{q2} = \begin{pmatrix} 3/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 3/2 \end{pmatrix}$, then also in this case there is a compactly supported MRA wavelet generated by a (compactly supported) scaling function ϕ that is the characteristic function of a fractal set (see Fig. 3.7).

The construction given above suggests that also in these cases one should be able to find an *AB*-MRA such that the associated compactly supported *AB*-wavelet has a simpler "nonfractal" support. This is done in [154].

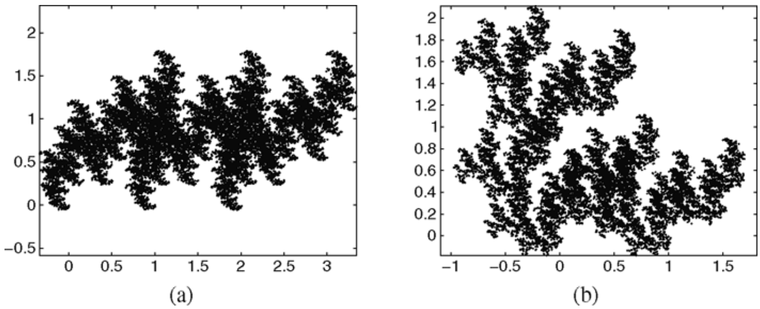


Fig. 3.7: The fractal sets associated with the MRA generated by the dilation matrices (a) A_{q1} and (b) A_{q2}.

3.5 Continuous Shearlet Transform

An important class of subgroups of the affine group $\mathbf{A}_2$ (which was described in Section 3.3) is obtained by considering

$$G = \{(M,t) : M \in \mathscr{D}_\alpha,\ t \in \mathbb{R}^2\}, \tag{3.37}$$

where, for each $0 < \alpha < 1$, $\mathscr{D}_\alpha \subset GL_2(\mathbb{R})$ is the set of matrices

$$\mathscr{D}_\alpha = \left\{ M = M_{as} = \begin{pmatrix} a & -a^\alpha s \\ 0 & a^\alpha \end{pmatrix}, \quad a > 0,\ s \in \mathbb{R} \right\}.$$

The matrices M_{as} can be factorized as $M_{as} = B_s A_a$, where

$$B_s = \begin{pmatrix} 1 & -s \\ 0 & 1 \end{pmatrix}, \quad A_a = \begin{pmatrix} a & 0 \\ 0 & a^\alpha \end{pmatrix}. \tag{3.38}$$

The matrix B_s is called a *shear matrix* and, for each $s \in \mathbb{R}$, is a nonexpanding matrix ($\det B_s = 1$ for each s). The matrix A_a is an anisotropic dilation matrix, that is, the dilation rate is different in the x and y directions. In particular, if $\alpha = 1/2$, the matrix A_a produces *parabolic scaling* since $f(A_a x) = f\left(A_a \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right)$ leaves invariant the parabola $x_1 = x_2^2$. Thus, the action associated with the dilation group $\mathscr{D}_\alpha$ can be interpreted as the superposition of anisotropic dilation and shear transformations.

Using Theorem 3.8 from Section 3.3, we can establish simple conditions on the function ψ so that it will satisfy the Calderón reproducing formula (3.20) with respect to G. This is done in the following proposition.

Proposition 3.13. *Let G be given by (3.37) and, for $\xi = (\xi_1, \xi_2) \in \mathbb{R}^2$, $\xi_2 \neq 0$, let ψ be given by*

$$\hat\psi(\xi) = \hat\psi(\xi_1,\xi_2) = \hat\psi_1(\xi_1)\,\hat\psi_2(\tfrac{\xi_2}{\xi_1}).$$

Suppose that

1. $\psi_1 \in L^2(\mathbb{R})$ *satisfies*

$$\int_0^\infty |\hat\psi_1(a\xi)|^2 \frac{da}{a^{2\alpha}} = 1 \quad \textit{for a.e. } \xi \in \mathbb{R};$$

2. $\|\psi_2\|_{L^2} = 1$.

Then ψ satisfies (3.20) and, hence, is a continuous wavelet with respect to G.

Proof. A direct computation shows that $(\xi_1,\xi_2)M = (a\xi_1, a^\alpha(\xi_2 - s\xi_1))$. Also, notice the element of the left Haar measure for $\mathscr{D}$ is $d\mu(M_{as}) = (da/|\det M_{as}|)\,ds$. Hence, the admissibility condition (3.21) for ψ is

$$\Delta(\psi)(\xi) = \int_{\mathbb{R}} \int_{\mathbb{R}^+} |\hat\psi_1(a\xi_1)|^2 \left| \hat\psi_2\left(a^{\alpha-1}\left(\tfrac{\xi_2}{\xi_1} - s\right)\right)\right|^2 \frac{da}{a^{1+\alpha}}\, ds = 1 \tag{3.39}$$

for a.e. $(\xi_1,\xi_2)\in\mathbb{R}^2$. Thus, by Theorem 3.8, to show that ψ is a continuous wavelet with respect to G, it is sufficient to show that (3.39) is satisfied. Using the assumption on ψ_1 and ψ_2, we have

$$\begin{aligned}\Delta(\psi)(\xi) &= \int_{\mathbb{R}}\int_{\mathbb{R}^+} |\hat{\psi}_1(a\,\xi_1)|^2 \left|\hat{\psi}_2\left(a^{\alpha-1}\left(\tfrac{\xi_2}{\xi_1}-s\right)\right)\right|^2 \frac{da}{a^{1+\alpha}}\,ds \\ &= \int_{\mathbb{R}^+} |\hat{\psi}_1(a\,\xi_1)|^2 \left(\int_{\mathbb{R}} \left|\hat{\psi}_2\left(a^{\alpha-1}\tfrac{\xi_2}{\xi_1}-s\right)\right|^2 ds\right)\frac{da}{a^{2\alpha}} \\ &= \int_{\mathbb{R}^+} |\hat{\psi}_1\,(a\,\xi_1)|^2\,\frac{da}{a^{2\alpha}} = 1\end{aligned}$$

for a.e. $\xi=(\xi_1,\xi_2)\in\mathbb{R}^2$. This shows that Eq. (3.39) is satisfied. □

In the following, to distinguish a continuous wavelet ψ associated with this particular group G from other continuous wavelets, we will refer to such a function as a *continuous shearlet*. Hence, for each $0<\alpha<1$, the *continuous shearlet transform* is the mapping

$$f \mapsto \{\mathscr{S}^{\alpha}_{\psi} f(a,s,t) = \langle f,\psi_{ast}\rangle : a>0,\ s\in\mathbb{R},\ t\in\mathbb{R}^2\},$$

where the analyzing elements,

$$\{\psi_{ast}(x) = |\det M_{as}|^{-\frac{1}{2}}\,\psi(M_{as}^{-1}(x-t)) : a>0,\ s\in\mathbb{R},\ t\in\mathbb{R}^2\},$$

with $M_{as}\in\mathscr{D}_\alpha$, form a *continuous shearlet system*. Notice that, according to the terminology introduced in Section 3.3, the elements $\{\psi_{ast}\}$ are co-affine functions.

A useful variant of the continuous shearlet transform is obtained by restricting the range of the shear variable s associated with the shearing matrices B_s to a finite interval. Namely, for $0<\alpha<1$, let us redefine

$$\mathscr{D}^{(h)}_{\alpha} = \left\{ M_{as} = \begin{pmatrix} a & -a^{\alpha} s \\ 0 & a^{\alpha}\end{pmatrix},\quad 0<a\le\frac{1}{4},\ -\frac{3}{2}\le s\le\frac{3}{2},\right\},$$

and

$$G^{(h)} = \{(M,t) : M\in\mathscr{D}^{(h)}_{\alpha},\ t\in\mathbb{R}^2\}.$$

Also, consider the subspace of $L^2(\mathbb{R}^2)$ given by $L^2(C_h)^{\vee} = \{f\in L^2(\mathbb{R}^2) : \operatorname{supp}\hat{f}\subset C_h\}$, where C_h is the "horizontal cone" in the frequency plane:

$$C_h = \left\{(\xi_1,\xi_2)\in\mathbb{R}^2 : |\xi_1|\ge 1 \text{ and } |\tfrac{\xi_2}{\xi_1}|\le 1\right\}.$$

Hence, we can show that by slightly modifying the assumptions of Proposition 3.13, the function ψ is a continuous shearlet for the subspace $L^2(C_h)^{\vee}$.

Proposition 3.14. *For $\xi=(\xi_1,\xi_2)\in\mathbb{R}^2$, $\xi_2\neq 0$, let ψ be given by*

$$\hat{\psi}(\xi) = \hat{\psi}(\xi_1,\xi_2) = \hat{\psi}_1(\xi_1)\,\hat{\psi}_2\left(\tfrac{\xi_2}{\xi_1}\right),$$

where

1. $\psi_1 \in L^2(\mathbb{R})$ *satisfies*

$$\int_0^\infty |\hat{\psi}_1(a\xi)|^2 \frac{da}{a^{2\alpha}} = 1 \quad \textit{for a.e. } \xi \in \mathbb{R},$$

   and $\operatorname{supp}\hat{\psi}_1 \subset [-2,-1/2] \cup [1/2,2]$*;*

2. $\|\psi_2\|_{L^2} = 1$ *and* $\operatorname{supp}\hat{\psi}_2 \subset [-1,1]$.

Then ψ *satisfies (3.24). That is, for all* $f \in L^2(C_h)^\vee$,

$$f(x) = \int_{\mathbb{R}^2} \int_{-\frac{3}{2}}^{\frac{3}{2}} \int_0^{\frac{1}{4}} \langle f, \psi_{ast} \rangle \, \psi_{ast}(x) \, \frac{da}{a^{2+2\alpha}} \, ds\, dt,$$

with convergence in the L^2 *sense.*

There are several examples of functions ψ_1 and ψ_2 satisfying the assumptions of Propositions 3.13 and 3.14. In particular, we can choose ψ_1, ψ_2 such that $\hat{\psi}_1$, $\hat{\psi}_2 \in C_0^\infty$, and we will make this assumption in the following. We refer to [122, 130] for the construction of these functions.

If the assumptions of Proposition 3.14 are satisfied, we say that the set

$$\Psi^{(h)} = \left\{ \psi_{ast} : 0 < a \le \tfrac{1}{4}, \, -\tfrac{3}{2} \le s \le \tfrac{3}{2}, \, t \in \mathbb{R}^2 \right\}$$

is a *continuous shearlet system* for $L^2(C_h)^\vee$ and that the corresponding mapping from $f \in L^2(C_h)^\vee$ into $\mathscr{S}_\psi^{(h),\alpha} f(a,s,t) = \langle f, \psi_{ast} \rangle$ is the *continuous shearlet transform* on $L^2(C_h)^\vee$.

In the frequency domain, an element of the shearlet system ψ_{ast} has the form

$$\hat{\psi}_{ast}(\xi_1,\xi_2) = a^{\frac{1+\alpha}{2}} \, \hat{\psi}_1(a\,\xi_1) \, \hat{\psi}_2\left(a^{\alpha-1} \left(\tfrac{\xi_2}{\xi_1} - s \right) \right) e^{-2\pi i \xi t}.$$

As a result, each function $\hat{\psi}_{ast}$ has support:

$$\operatorname{supp}\hat{\psi}_{ast} \subset \left\{ (\xi_1,\xi_2) : \xi_1 \in [-\tfrac{2}{a}, -\tfrac{1}{2a}] \cup [\tfrac{1}{2a}, \tfrac{2}{a}], \, |\tfrac{\xi_2}{\xi_1} - s| \le a^{1-\alpha} \right\}.$$

As illustrated in Fig. 3.8, the frequency support is a pair of trapezoids, symmetric with respect to the origin, oriented along a line of slope s. The support becomes increasingly elongated as $a \to 0$.

As shown by Proposition 3.14, the continuous shearlet transform $\mathscr{S}_\psi^{(h),\alpha}$ provides a reproducing formula only for functions in a proper subspace of $L^2(\mathbb{R}^2)$. To extend the transform to all $f \in L^2(\mathbb{R}^2)$, we introduce a similar transform to deal with the functions supported on the "vertical cone":

$$C^{(v)} = \left\{ (\xi_1,\xi_2) \in \mathbb{R}^2 : |\xi_2| \ge 1 \text{ and } \left| \tfrac{\xi_2}{\xi_1} \right| > 1 \right\}.$$

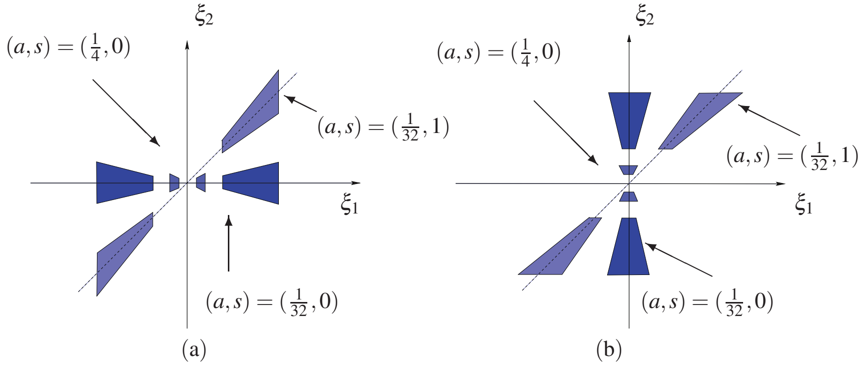


Fig. 3.8: Frequency support of the (a) horizontal shearlets and (b) vertical shearlets for different values of a and s.

Specifically, let

$$\hat{\psi}^{(v)}(\xi) = \hat{\psi}^{(v)}(\xi_1, \xi_2) = \hat{\psi}_1(\xi_2)\, \hat{\psi}_2\left(\frac{\xi_1}{\xi_2}\right),$$

where $\hat{\psi}_1$, $\hat{\psi}_2$ satisfy the same assumptions as in Proposition 3.14, and consider the dilation group

$$\mathscr{D}_\alpha^{(v)} = \left\{ N_{as} = \begin{pmatrix} a^\alpha & 0 \\ -a^\alpha s & a \end{pmatrix} : 0 < a \leq 1,\ -\tfrac{3}{2} \leq s \leq \tfrac{3}{2},\ t \in \mathbb{R}^2 \right\}.$$

Then it is easy to verify that the set

$$\Psi^{(v)} = \left\{ \psi_{ast}^{(v)} : 0 < a \leq \tfrac{1}{4},\ -\tfrac{3}{2} \leq s \leq \tfrac{3}{2},\ t \in \mathbb{R}^2 \right\},$$

where $\psi_{ast}^{(v)} = |\det N_{as}|^{-1/2} \psi^{(v)}(N_{as}^{-1}(x-t))$, is a continuous shearlet system for $L^2(C^{(v)})^\vee$. The corresponding transform $\mathscr{S}_\psi^{(v),\alpha} f(a,s,t) = \langle f, \psi_{ast}^{(v)} \rangle$ is the continuous shearlet transform on $L^2(C^{(v)})^\vee$. Finally, by introducing an appropriate window function W, we can represent the functions with frequency support on the set $[-2,2]^2$ as

$$f = \int_{\mathbb{R}^2} \langle f, W_t \rangle W_t \, dt,$$

where $W_t(x) = W(x-t)$. As a result, any function $f \in L^2(\mathbb{R}^2)$ can be reproduced with respect of the full shearlet system, which consists of the horizontal shearlet system $\Psi^{(h)}$, the vertical shearlet system $\Psi^{(v)}$, and the collection of coarse-scale isotropic functions $\{W_t : t \in \mathbb{R}^2\}$. We refer to [156] for more details about this representation. For our purposes, it is only the behavior of the fine-scale shearlets that matters. Indeed, in the following, we will apply the continuous shearlet transforms $\mathscr{S}_\psi^{(h),\alpha}$ and $\mathscr{S}_\psi^{(v),\alpha}$, at fine scales ($a \to 0$), to resolve and precisely describe the boundaries of certain planar regions. Hence, it will be convenient to redefine the shearlet transform, at "fine scales," as follows. For $0 < a \le 1/4$, $s \in \mathbb{R}$, $t \in \mathbb{R}^2$, the *(fine-scale) continuous shearlet transform* is the mapping from $f \in L^2(\mathbb{R}^2 \setminus [-2,2]^2)^\vee$ into $\mathscr{S}_\psi f$, which is defined by

$$\mathscr{S}_\psi^\alpha f(a,s,t) = \begin{cases} \mathscr{S}_\psi^{(h),\alpha}(a,s,t), & \text{if } |s| \le 1, \\ \mathscr{S}_\psi^{(v),\alpha}(a,\frac{1}{s},t), & \text{if } |s| > 1. \end{cases}$$

3.5.1 Edge Analysis Using the Shearlet Transform

One remarkable property of the continuous shearlet transform is its ability to provide a very precise characterization of the set of singularities of functions and distributions. Indeed, let f be a function on $\mathbb{R}^2$ consisting of several smooth regions Ω_n, $n = 1,\dots,N$, separated by piecewise smooth boundaries $\gamma_n = \partial\Omega_n$:

$$f(x) = \sum_{n=1}^{N} f_n(x)\chi_{\Omega_n}(x),$$

where each function f_n is smooth. Then the continuous shearlet transform $\mathscr{S}_\psi^\alpha f(a,s,t)$ will signal both the location and orientation of the boundaries through its asymptotic decay at fine scales. In fact, $\mathscr{S}_\psi^\alpha f(a,s,t)$ will exhibit fast asymptotic decay $a \to 0$ for all (s,t), except for the values of t on the boundary curves γ_n and for the values of s associated with the normal orientation to the γ_n at t.

The study of these objects is motivated by image applications, where f is used to model an image, and the curves γ_n are the edges of the image f. We will show that the shearlet framework provides a very effective method for the detection and analysis of edges. This is a fundamental problem in many applications from computer vision to image processing.

To illustrate how the shearlet transform can be employed to characterize the geometry of edges, let us consider the case where f is simply the characteristic function of a bounded subset of $\mathbb{R}^2$. Also, to simplify the presentation, we will only present the situation where $\alpha = 1/2$ and use the simplified notation $\mathscr{S}_\psi = \mathscr{S}_\psi^{1/2}$. In more general case where $\alpha \in (0,1)$, the continuous shearlet transform $\mathscr{S}_\psi^\alpha$ is similar and details can be found in [127].

We then have the following result from [127].

Theorem 3.15. *Let $D \subset \mathbb{R}^2$ be a bounded region in $\mathbb{R}^2$, and suppose that the boundary curve $\gamma = \partial D$ is a simple C^3 regular curve. Denote $B = \chi_D$. If $t = t_0 \in \gamma$, and $s_0 = \tan\theta_0$, where θ_0 is the angle corresponding to the normal orientation to γ at t_0, then*

$$\lim_{a \to 0^+} a^{-\frac{3}{4}}\, \mathscr{S}_\psi B(a,s_0,t_0) \ne 0.$$

If $t = t_0 \in \gamma$ and $s \ne \tan\theta_0$, or if $t \notin \gamma$, then

$$\lim_{a \to 0^+} a^{-\beta}\, \mathscr{S}_\psi B(a,s,t) = 0 \quad \textit{for all } \beta > 0.$$

This shows that the continuous shearlet transform $\mathscr{S}_\psi B(a,s,t)$ has "slow" decay only for $t = t_0$ on γ when the value of the shear variable s corresponds exactly to the normal orientation to γ at t_0. For all other values of t and s, the decay is fast. This behavior is illustrated in Fig. 3.9.

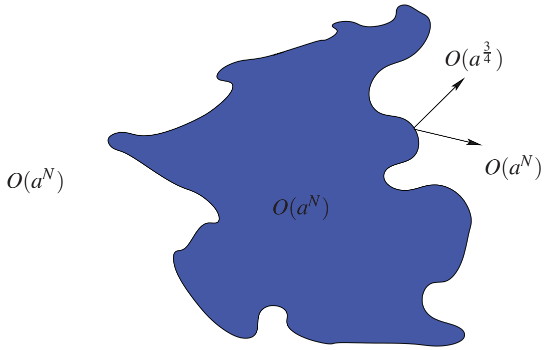


Fig. 3.9: Asymptotic decay of the continuous shearlet transform of the $B(x) = \chi_D(x)$. On the boundary ∂D, for normal orientation, the shearlet transform decays as $O(a^{3/4})$. For all other values of (t,s), the decay is as fast as $O(a^N)$, for any $N \in \mathbb{N}$.

Theorem 3.15 can be generalized to the situation where the boundary curve γ is piecewise smooth and contains finitely many corner points. Also, in this case, the continuous shearlet transform provides a precise description of the geometry of the boundary curve through its asymptotic decay at fine scales. In particular, at the corner points, the asymptotic decay at fine scales is the slowest for values of s corresponding to the normal directions (notice that there are two of them). We refer the interested reader to [128] for a detailed discussion of the shearlet analysis of regions with piecewise smooth boundaries. We also refer to [127, 156] for other related results, including the situation where f is not simply the union of characteristic functions of sets.

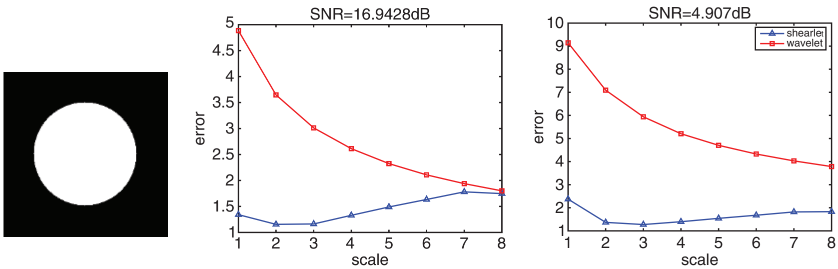
Finally, we recall that the shearlet transform shares some of the features described above with the *continuous curvelet transform*, another directional multiscale transform introduced by Candès and Donoho in [41]. Even if a result like Theorem 3.15 is not known for the curvelet transform, other results in [41] indicate that the curvelet transform is also able to capture the geometry of singularities in $\mathbb{R}^2$ through its asymptotic decay at fine scales. Notice that, unlike the shearlet transform, the curvelet transform is not directly associated with an affine group.

3.5.2 A Shearlet Approach to Edge Analysis and Detection

Taking advantage of the properties of the continuous shearlet transform described above, an efficient numerical algorithm for edge detection was designed by one of the authors and his collaborators [82, 83]. The shearlet approach adapts several ideas from the well-known *wavelet modulus maxima* method of Hwang, Mallat, and Zhong [176, 177], where the edge points of an image f are identified as the locations corresponding to the local maxima of the magnitude of the continuous wavelet transform of f. Recall that, at a single scale, this wavelet-based method is indeed equivalent to the *canny edge detector*, which is a standard edge detection algorithm [62].

As shown above, one main feature of the continuous shearlet transform is its superior directional selectivity with respect to wavelets and other traditional methods. This property plays a very important role in the design of the edge detection algorithm. In fact, one major task in edge detection is to accurately identify the edges of an image in the presence of noise, to perform this task, both the location and the orientation of edge points have to be estimated from a noisy image.

In the usual wavelet modulus maxima approach, the edge orientation of an image f, at the location t, is estimated by looking at the ratio of the vertical over the horizontal components of $W_\psi f(a,t)$, the wavelet transform of f. However, this approach is not very accurate when dealing with discrete data. The advantage of the continuous shearlet transform is that, by representing the image as a function of scale, location, and orientation, the directional information is directly available. A number of tests conducted in [82, 83] show indeed that a shearlet-based approach provides a very accurate estimate of the edge orientation of a noisy image; this method significantly outperforms the wavelet-based approach. A typical numerical experiment is illustrated in Fig. 3.10, where the test image is the characteristic function of a disc. This figure displays the average angular error in the estimate of the edge's orientation, as a function of the scale a. The average angle error is defined by where E is the set of edge points, θ is the exact angle, and $\hat{\theta}$ the estimated angle. The average angle error is indicated for both shearlet- and wavelet-based methods, in the presence of additive Gaussian noise. As the figure shows, the shearlet approach significantly outperforms the wavelet method, especially at finer scales, and is extremely robust to noise.



(a) (b) (c)

Fig. 3.10: (a) Test image. (b–c) Comparison of the average error in angle estimation using the wavelet method versus the shearlet method, as a function of the scale, with different noise levels; (b) PSNR = 16.9 dB, (c) PSNR = 4.9 dB. (Courtesy of Sheng Yi.)

Using these properties, a very competitive algorithm for edge detection was developed in [83] and a representative numerical test is illustrated in Fig. 3.11. We refer to [82, 83] for details about these algorithms and for additional numerical demonstrations.

$$\frac{1}{|E|} \cdot \sum_{t \in E} |\hat{\theta}(t) - \theta(t)|,$$

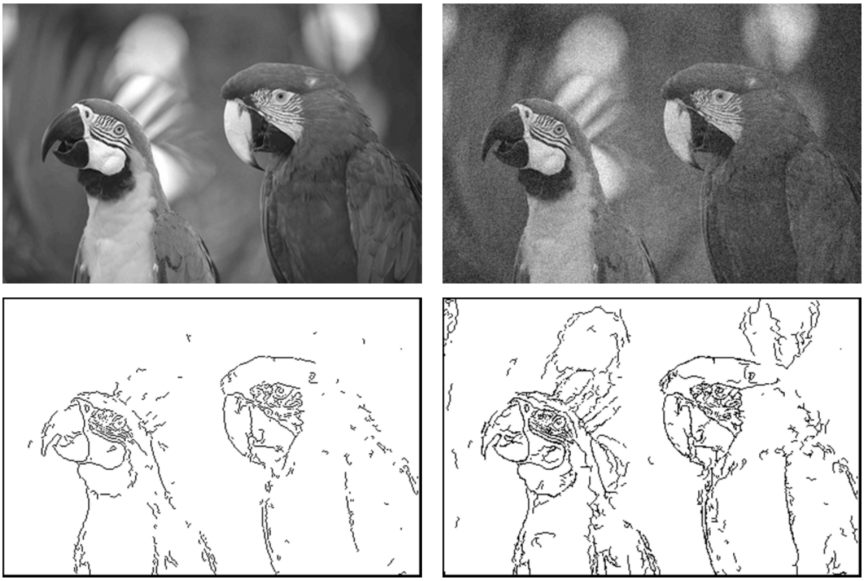


Fig. 3.11: Comparison of edge detection using a shearlet-based method versus a wavelet-based method. From top left, clockwise: original image, noisy image (PSNR = 24.59 dB), shearlet result, and wavelet result. (Courtesy of Glenn Easley.)

3.5.3 Discrete Shearlet System

By sampling the continuous shearlet transform

$$f \mapsto \mathscr{S}_\psi f(a,s,t) = \langle f, \psi_{a,s,t} \rangle$$

on an appropriate discrete set of the scaling, shear, and translation parameters $(a,s,t) \in \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^2$, it is possible to obtain a frame or even a Parseval frame for $L^2(\mathbb{R}^2)$. Notice that, as above, we will only consider the case $\mathscr{S}_\psi = \mathscr{S}_\psi^{1/2}$.

To construct the discrete shearlet system (see [157] for more details), we start by choosing a discrete set of scales $\{a_j\}_{j \in \mathbb{Z}} \subset \mathbb{R}^+$; next, for each fixed j, we choose the shear parameters $\{s_{j,\ell}\}_{\ell\in\mathbb{Z}} \subset \mathbb{R}$ so that the directionality of the representation is allowed to change with the scale. Finally, to provide a "uniform covering" of $\mathbb{R}^2$, we allow the location parameter to describe a different grid depending on j on ℓ; hence, we let $t_{j,\ell,k} = B_{s_{j,\ell}} A_{a_j} k$, $k \in \mathbb{Z}^2$, where the matrices B_s, for $s \in \mathbb{R}$, and A_a, for $a > 0$, are given by (3.38). Observing that

$$T_{\{B_{s_{j,\ell}} A_{a_j} k\}} D_{B_{s_{j,\ell}} A_{a_j}} = D_{B_{s_{j,\ell}} A_{a_j}} T_k,$$

we obtain the discrete system

$$\left\{ \psi_{j,\ell,k} = D_{B_{s_{j,\ell}} A_{a_j}} T_k \psi : \ j,\ell \in \mathbb{Z}, \ k \in \mathbb{Z}^2 \right\}.$$

In particular, we will set $a_j = 2^{2j}$, $s_{j,\ell} = \ell \sqrt{a_j} = \ell 2^j$. Thus, observing that $B_{\ell 2^j} A_{2^{2j}} = A_{2^{2j}} B_\ell$, we finally obtain the *discrete shearlet system*

$$\{ \psi_{j,\ell,k} = D_{A_{4^j}} D_{B_\ell} T_k \psi : \ j,\ell \in \mathbb{Z}, \ k \in \mathbb{Z}^2 \}. \tag{3.40}$$

Notice that (3.40) is an example of the affine systems with composite dilations (3.31), described in Section 3.4. More specifically, the discrete shearlet system obtained above is similar to the "shearlet-like" system (3.30). Unlike the system (3.30), however, whose elements are characteristic functions of sets in the frequency domain, we will show that in this case we obtain a system of well-localized functions.

To do that, we will adapt some ideas from the continuous case. Namely, for any $\xi = (\xi_1, \xi_2) \in \widehat{\mathbb{R}}^2$, $\xi_1 \neq 0$, let

$$\hat{\psi}^{(h)}(\xi) = \hat{\psi}^{(h)}(\xi_1, \xi_2) = \hat{\psi}_1(\xi_1)\, \hat{\psi}_2\left(\frac{\xi_2}{\xi_1}\right),$$

where $\hat{\psi}_1, \hat{\psi}_2 \in C^\infty(\widehat{\mathbb{R}})$, $\operatorname{supp}\hat{\psi}_1 \subset [-1/2, -1/16] \cup [1/16, 1/2]$ and $\operatorname{supp}\hat{\psi}_2 \subset [-2, 2]$. This implies that $\hat{\psi}^{(h)}$ is a compactly supported C^∞ function with support contained in $[-1/2, 1/2]^2$. In addition, we assume that

$$\sum_{j\geq 0} |\hat{\psi}_1(2^{-2j}\omega)|^2 = 1 \quad \text{for } |\omega| \geq \frac{1}{8}, \tag{3.41}$$

and, for each $j \geq 0$,

$$\sum_{\ell=-2^j}^{2^j-1} |\hat{\psi}_2(2^j \omega - \ell)|^2 = 1 \quad \text{for } |\omega| \leq 1. \tag{3.42}$$

From the conditions on the support of $\hat{\psi}_1$ and $\hat{\psi}_2$, one can easily deduce that the functions $\psi_{j,\ell,k}$ have frequency support contained in the set

$$\left\{ (\xi_1, \xi_2) : \xi_1 \in [-2^{2j-1}, -2^{2j-4}] \cup [2^{2j-4}, 2^{2j-1}], \ \left| \tfrac{\xi_2}{\xi_1} + \ell 2^{-j} \right| \leq 2^{-j} \right\}.$$

Thus, each element $\hat{\psi}_{j,\ell,k}$ is supported on a pair of trapezoids of approximate size $2^{2j} \times 2^j$, oriented along lines of slope $\ell 2^{-j}$ [see Fig. 3.12(b)].

From Eqs. (3.41) and (3.42), it follows that the functions $\{\hat{\psi}_{j,\ell,k}\}$ form a tiling of the set

$$\mathscr{D}_h = \left\{ (\xi_1, \xi_2) \in \widehat{\mathbb{R}}^2 : |\xi_1| \geq \tfrac{1}{8}, \left| \frac{\xi_2}{\xi_1} \right| \leq 1 \right\}.$$

Indeed, for $(\xi_1, \xi_2) \in \mathscr{D}_h$,

$$\sum_{j \geq 0} \sum_{\ell=-2^j}^{2^j-1} |\hat{\psi}^{(h)}(\xi A_4^{-j} B_1^{-\ell})|^2 = \sum_{j \geq 0} \sum_{\ell=-2^j}^{2^j-1} \left| \hat{\psi}_1(2^{-2j}\xi_1)|^2 \, |\hat{\psi}_2 \left(2^j \frac{\xi_2}{\xi_1} - \ell \right) \right|^2 = 1. \tag{3.43}$$

An illustration of this frequency tiling is shown in Fig. 3.12(a).

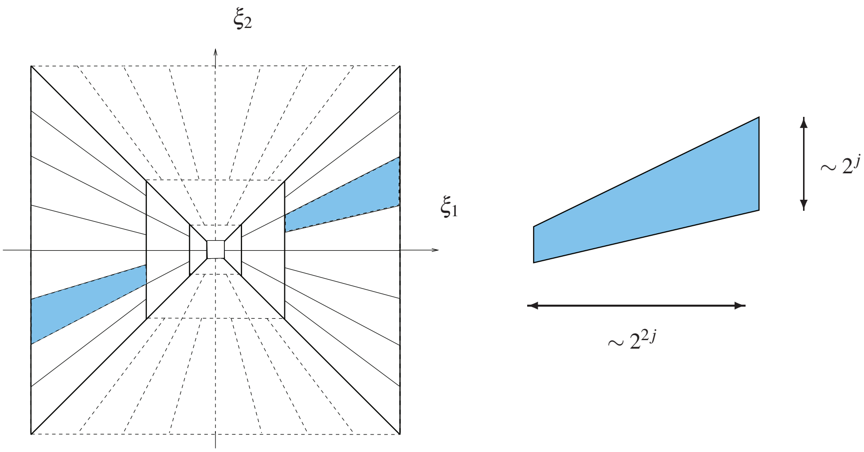


Fig. 3.12: (a) The tiling of the frequency plane $\widehat{\mathbb{R}}^2$ induced by the shearlets. The tiling of $\mathscr{D}_h$ is illustrated by solid lines, while the tiling of $\mathscr{D}_v$ appears is in dashed lines. (b) The frequency support of a shearlet $\psi_{j,\ell,k}$ satisfies parabolic scaling. The figure shows only the support for $\xi_1 > 0$; the other half of the support, for $\xi_1 < 0$, is symmetrical.

Letting $L^2(\mathscr{D}_h)^\vee = \{f \in L^2(\mathbb{R}^2) : \operatorname{supp} \hat{f} \subset \mathscr{D}_h\}$, property (3.43) and the fact that $\hat{\psi}^{(h)}$ is supported inside $[-1/2, 1/2]^2$ imply that the discrete shearlet system

$$\Psi_d^{(h)} = \{\psi_{j,\ell,k} : j \geq 0, -2^j \leq \ell \leq 2^j - 1, \, k \in \mathbb{Z}^2\}$$

is a Parseval frame for $L^2(\mathscr{D}_h)^\vee$. Similarly, we can construct a Parseval frame for $L^2(\mathscr{D}_v)^\vee$, where $\mathscr{D}_v$ is the vertical cone $\mathscr{D}_v = \{(\xi_1, \xi_2) \in \widehat{\mathbb{R}}^2 : |\xi_2| \geq 1/8, |\xi_1/\xi_2| \leq 1\}$.

Specifically, let

$$\tilde{A} = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}, \quad \tilde{B} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix},$$

and let $\psi^{(v)}$ be given by

$$\hat{\psi}^{(v)}(\xi) = \hat{\psi}^{(v)}(\xi_1, \xi_2) = \hat{\psi}_1(\xi_2)\, \hat{\psi}_2\left(\frac{\xi_1}{\xi_2}\right).$$

Then the collection

$$\Psi_d^{(v)} = \{\psi_{j,\ell,k}^{(v)} : j \geq 0,\ -2^j \leq \ell \leq 2^j - 1,\ k \in \mathbb{Z}^2\},$$

where $\psi_{j,\ell,k}^{(v)} = D_{\tilde{A}}^{j} D_{\tilde{B}}^{\ell} T_k \psi^{(v)}$, is a Parseval frame for $L^2(\mathscr{D}_v)^{\vee}$.

Finally, let $\hat{\varphi} \in C_0^{\infty}(\mathbb{R}^2)$ be chosen to satisfy

$$|\hat{\varphi}(\xi)|^2 + \sum_{j \geq 0} \sum_{\ell=-2^j}^{2^j-1} |\hat{\psi}^{(h)}(\xi A_4^{-j} B_1^{-\ell})|^2 \chi_{\mathscr{D}_h}(\xi)$$
$$+ \sum_{j \geq 0} \sum_{\ell=-2^j}^{2^j-1} |\hat{\psi}^{(v)}(\xi \tilde{A}^{-j} \tilde{B}^{-\ell})|^2 \chi_{\mathscr{D}_v}(\xi) = 1, \quad \text{for } \xi \in \widehat{\mathbb{R}}^2,$$

where $\chi_{\mathscr{D}}$ is the indicator function of the set $\mathscr{D}$. This implies that $\operatorname{supp} \hat{\varphi} \subset [-1/8, 1/8]^2$, $|\hat{\varphi}(\xi)| = 1$ for $\xi \in [-1/16, 1/16]^2$, and the collection $\{\varphi_k : k \in \mathbb{Z}^2\}$ defined by $\varphi_k(x) = \varphi(x - k)$ is a Parseval frame for $L^2([-1/16, 1/16]^2)^{\vee}$.

Thus, letting $\widehat{\tilde{\psi}}_{j,\ell,k}^{(\omega)}(\xi) = \widehat{\psi}_{j,\ell,k}^{(\omega)}(\xi)\, \chi_{\mathscr{D}_\omega}(\xi)$, for $\omega = h$ or $\omega = v$, we have the following result.

Theorem 3.16. *The discrete shearlet system*

$$\{\varphi_k : k \in \mathbb{Z}^2\} \bigcup \{\tilde{\psi}_{j,\ell,k}^{(\omega)}(x) : j \geq 0,\ \ell = -2^j, 2^j - 1,\ k \in \mathbb{Z}^2,\ \omega = h, v\}$$
$$\bigcup \{\psi_{j,\ell,k}^{(\omega)}(x) : j \geq 0, -2^j + 1 \leq \ell \leq 2^j - 2,\ k \in \mathbb{Z}^2,\ \omega = h, v\}$$

is a Parseval frame for $L^2(\mathbb{R}^2)$.

The "corner" elements $\tilde{\psi}_{j,\ell,k}^{(\omega)}(x)$, $\ell = -2^j, 2^j - 1$, are simply obtained by truncation on the cones $\chi_{\mathscr{D}_\omega}$ in the frequency domain. Notice that the corner elements in the horizontal cone $\mathscr{D}_v$ match nicely with those in the vertical cone $\mathscr{D}_h$. We refer to [80, 130] for additional details on this construction.

3.5.4 Optimal Representations Using Shearlets

One major feature of shearlet systems is that if f is a compactly supported function that is C^2 away from a C^2 curve, then the sequence of discrete shearlet coefficients $\{\langle f, \psi_{j,\ell,k} \rangle\}$ has (essentially) optimally fast decay. To make this more precise, let

f_N^S be the N-term approximation of f obtained from the N largest coefficients of its shearlet expansion, namely,

$$f_N^S = \sum_{\mu \in I_N} \langle f, \psi_\mu \rangle \, \psi_\mu,$$

where $I_N \subset M$ is the set of indices corresponding to the N largest entries of the sequence $\{|\langle f, \psi_\mu \rangle|^2 : \mu \in M\}$. Also, we follow [38] and introduce $STAR^2(A)$, a class of indicator functions of sets B with C^2 boundaries ∂B. In polar coordinates, let $\rho(\theta) : [0, 2\pi) \to [0,1]^2$ be a radius function and define B by $x \in B$ if and only if $|x| \le \rho(\theta)$. In particular, the boundary ∂B is given by the curve in $\mathbb{R}^2$:

$$\beta(\theta) = \begin{pmatrix} \rho(\theta)\cos(\theta) \\ \rho(\theta)\sin(\theta) \end{pmatrix}. \tag{3.44}$$

The class of boundaries of interest to us is defined by

$$\sup |\rho''(\theta)| \le A, \quad \rho \le \rho_0 < 1. \tag{3.45}$$

We say that a set $B \in STAR^2(A)$ if $B \subset [0,1]^2$ and B is a translate of a set obeying (3.44) and (3.45). Finally, we define the set $\mathscr{E}^2(A)$ of *functions that are C^2 away from a C^2 edge* as the collection of functions of the form

$$f = f_0 + f_1 \, \chi_B,$$

where $f_0, f_1 \in C_0^2([0,1]^2)$, $B \in STAR^2(A)$, and $\|f\|_{C^2} = \sum_{|\alpha| \le 2} \|D^\alpha f\|_\infty \le 1$. We can now state the following result from [124].

Theorem 3.17. *Let $f \in \mathscr{E}^2(A)$ and f_N^S be the approximation to f defined above. Then*

$$\|f - f_N^S\|_2^2 \le C N^{-2} \, (\log N)^3.$$

Notice that the approximation error of shearlet systems significantly outperforms wavelets, in which case the approximation error $\|f - f_N^W\|_2^2$ decays at most as fast as $O(N^{-1})$ [175], where f_N^W is the N-term approximation of f obtained from the N largest coefficients in the wavelet expansion. Indeed, the shearlet representation is essentially optimal for the kind of functions considered here, since the optimal theoretical approximation rate (cf. [65]) satisfies

$$\|f - f_N\|_2^2 \asymp N^{-2}, \qquad N \to \infty.$$

Only the curvelet system of Candès and Donoho is known to satisfy similar approximation properties [38]. However, the curvelet construction has a number of important differences, including the fact that the curvelet system is not associated with a fixed translation lattice and, unlike the shearlet system, is not an affine-like system, since it is not generated from the action of a family of operators on a single or finite family of functions.

The optimal sparsity of the shearlet system plays a fundamental role in a number of applications. For example, the shearlet system can be applied to provide a sparse representation of Fourier integral operators, a very important class of operators that appear in problems from partial differential equations [125, 126]. Another class of applications comes from image processing, where the sparsity of the shearlet representation is closely related to the ability to efficiently separate the relevant features of an image from noise. A number of results in this direction are described in [79–81].

Exercises

1. Show that Eq. (3.12) in Theorem 3.5 can be simplified to obtain (3.13). Next, show that, for $n = 1$, when the dilation matrix a is replaced by the dyadic factor 2, Eq. (3.13) yields the "classical" Gripenberg–Wang equations (3.4) and (3.5).
2. Show that the matrices

$$M = \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad M = \begin{pmatrix} a & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix},$$

   where $a > 1$, are expanding on a subspace (that is, they satisfy Definition 3.7).
3. Show that Theorem 3.8 is valid for functions on subspaces of $L^2(\mathbb{R}^n)$ of the form

$$L^2(V)^\vee = \{f \in L^2(\mathbb{R}^n) : \operatorname{supp} \hat{f} \subset V\}.$$

4. Let $\psi_1 \in L^2(\mathbb{R})$ be a dyadic wavelet with $\operatorname{supp} \hat{\psi}_1 \subset [-1/2, 1/2]$ and $\psi_2 \in L^2(\mathbb{R})$ be such that $\operatorname{supp} \hat{\psi}_1 \subset [-1, 1]$ and

$$\sum_{k \in \mathbb{Z}} |\hat{\psi}_2(\omega + k)|^2 = 1 \quad \text{for a.e. } \omega \in \mathbb{R}.$$

   For $\xi = (\xi_1, \xi_2) \in \mathbb{R}^2$, let ψ be defined by $\hat{\psi}(\xi) = \hat{\psi}_1(\xi_1)\,\hat{\psi}_2(\xi_2/\xi_1)$. Show that the affine system $\{D_A^i D_B^j T_k \psi : i, j \in \mathbb{Z}, k \in \mathbb{Z}^2\}$, where $A = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, is a Parseval frame for $L^2(\mathbb{R}^2)$.
5. Prove Proposition 3.14 by modifying the argument of Proposition 3.14.
6. Let ψ be a Schwarz class function and $\mathscr{S}_\psi$ be the fine-scale continuous shearlet transform (for $\alpha = 1/2$), as defined in this chapter. Show that, for any $s \in \mathbb{R}$, the continuous shearlet transform of the Dirac delta distribution satisfies

$$\mathscr{S}_\psi \delta(a, s, (0,0)) \sim a^{-\frac{3}{4}},$$

   asymptotically as $a \to 0$. Show that if $t \neq (0,0)$, then, for any $N \in \mathbb{N}$, there is a constant $C_N > 0$ such that

$$\mathscr{S}_\psi \delta(a, s, (0,0)) \leq C_N\, a^N,$$

   asymptotically as $a \to 0$.
7. Let ψ and $\mathscr{S}_\psi$ be as in Exercise 6. For $p \in \mathbb{R}$, consider the distribution $\nu_p(x_1, x_2)$ defined by

$$\int_{\mathbb{R}^2} \nu_p(x_1, x_2)\, f(x_1, x_2)\, dx_1\, dx_2 = \int_{\mathbb{R}} f(p x_2, x_2)\, dx_2.$$

Show that, for $s = -p$ and $t_1 = pt_2$, we have

$$\mathscr{S}_\psi \delta(a,s,(t_1,t_2)) \sim a^{-\frac{1}{4}},$$

asymptotically as $a \to 0$. Show that for all other values of $t = (t_1,t_2)$ or s, then, for any $N \in \mathbb{N}$, there is a constant $C_N > 0$ such that

$$\mathscr{S}_\psi \delta(a,s,(0,0)) \leq C_N\, a^N,$$

asymptotically as $a \to 0$.

Chapter 4
Wavelets on the Sphere

Pierre Vandergheynst and Yves Wiaux

Abstract In many application fields, ranging from astrophysics and geophysics to neuroscience, computer vision, and computer graphics, data to be analyzed are defined as functions on the sphere. In all these situations, there are compelling reasons to design dedicated data analysis tools that are adapted to spherical geometry, for one cannot simply project the data in Euclidean geometry without having to deal with severe distortions. The wavelet transform has become a ubiquitous tool in signal processing mostly for its ability to exploit the multiscale nature of many data sets, and it is thus quite natural to generalize it to signals on the sphere. This generalization is not trivial, for the main ingredient of wavelet theory, dilation, is not well defined on the sphere. Moreover, when turning to algorithms, one faces the problem that sampling data on the sphere is not an easy task either. In this chapter, we discuss recently developed results for the analysis and reconstruction of signals on the sphere with wavelets, on the basis of theory, implementation, and applications.

4.1 Introduction

There are many application scenarios where data to be analyzed are defined as a scalar function on the sphere. Some of the most common examples include processing geodesic signals, climate indicators (atmospheric or ocean temperature, for example), or astronomical data defined on the celestial sphere. Recently, with the advent of advanced imaging modalities and devices, data sets defined in spherical geometry have started to appear in many other areas. In computer vision, catadioptric cameras allow one to record omnidirectional images using a regular sensor overlooking a curved mirror. The captured images are most naturally

Pierre Vandergheynst
Yves Wiaux
Signal Processing Laboratory (LTS2), Ecole Polytechnique Fédérale de Lausanne (EPFL), Station 11, 1015 Lausanne, Switzerland, e-mail: pierre.vandergheynst@epfl.ch
e-mail: yves.wiaux@epfl.ch

B. Forster and P. Massopust (eds.), *Four Short Courses on Harmonic Analysis*, Applied and Numerical Harmonic Analysis, DOI 10.1007/978-0-8176-4891-6_4,

expressed in spherical coordinates around the focal point of the system. In computer graphics, complicated but closed genus-zero surfaces are expressed as elevation maps in spherical coordinates. One then seeks to process these surfaces so as to reveal or simplify their shape attributes.

In all these situations, there are compelling reasons to design dedicated data analysis tools that are adapted to spherical geometry, for one cannot simply project the data in Euclidean geometry without having to deal with severe distortions. The wavelet transform has become a ubiquitous tool in signal processing mostly for its ability to exploit the multiscale nature of many data sets, and it is thus quite natural to generalize it to signals on the sphere. However, this generalization is not trivial, for the main ingredient of the wavelet theory, dilation, is not well defined on the sphere. Moreover, when turning to algorithms, one faces the problem that sampling data on the sphere is not an easy task either: There is no preferred sampling grid similar to the $\mathbb{Z}^2$ lattice in the plane.

There have been many attempts at generalizing the wavelet transform to the sphere, and it is well beyond the scope of the present chapter to review all approaches. Instead, we will focus on recently developed results that provide a sensitive mathematical framework and that can be efficiently implemented by provably stable and fast algorithms. We will proceed by first defining and studying several possible dilation operations on the sphere. A continuous wavelet formalism will then be simply defined by generalizing the operation of correlating a signal with suitably dilated waveforms. We will then define a scale-discretized wavelet formalism in order to allow the practical reconstruction of a signal from its wavelet coefficients. We will also discuss fast algorithms allowing efficient analysis and reconstruction of digital data. Finally, we will conclude with applications in astrophysics and neuroscience illustrating the usefulness of the proposed tools.

4.2 Scale-Space Premises

In this section, we discuss the notion of directional correlation on the sphere, we concisely recall the harmonic analysis on the sphere and on the rotation group, and we discuss affine transformations, in particular dilations. These are the essential tools for the definition of a wavelet formalism.

4.2.1 Directional Correlations

We consider a three-dimensional Cartesian coordinate system $(o, o\hat{x}, o\hat{y}, o\hat{z})$ centered on the unit sphere $\mathbb{S}^2$, and where the direction $o\hat{z}$ identifies the north pole. Any point ω on the sphere is identified by its corresponding spherical coordinates (θ, φ), where $\theta \in [0, \pi]$ stands for the colatitude, or polar angle, and $\varphi \in [0, 2\pi)$ for the longitude, or azimuthal angle. We consider signals F and analysis functions Ψ on the sphere as described by elements of the Hilbert space of square-integrable functions $L^2(\mathbb{S}^2, d\Omega)$, with the invariant measure $d\Omega = d\cos\theta d\varphi$. In this space, the scalar product between two functions F_1 and F_2 reads as $\langle F_1|F_2\rangle = \int_{\mathbb{S}^2} d\Omega\, F_1^*(\omega)\, F_2(\omega)$.

In order to perform a scale-space analysis of signals, continuous affine transformations such as translations, rotations, and dilations on the sphere must be applied to the analysis function. These affine transformations are mathematically defined and described in detail below. In a few words, the continuous translations by $\omega_0 = (\theta_0, \varphi_0) \in \mathbb{S}^2$ and rotations by $\chi \in [0, 2\pi)$ are defined by the three Euler angles defining an element $\rho = (\varphi_0, \theta_0, \chi)$ of the group of rotations in three dimensions SO(3). The continuous dilations affect by definition the continuous scale of the function and may be parametrized in terms of some dilation factor $a \in \mathbb{R}_+^*$. The precise definition of spherical dilations is the main challenge for building spherical wavelets, and clean mathematical arguments will be given in Section 4.2.3. Let us simply assume that the operation is formally defined so that we can fix notations.

The analysis of the signal F with an analysis function Ψ defines wavelet coefficients through the scalar products of F with the translated, rotated, and dilated functions $\Psi_{\rho,a}$ as

$$W_\Psi^F(\rho, a) = \langle \Psi_{\rho,a}|F\rangle. \tag{4.1}$$

This relation also defines the so-called directional correlation of F with the dilated functions Ψ_a. At each scale a, the function $W_\Psi^F(\cdot, a)$ of ρ identifying the wavelet coefficients is an element of the Hilbert space of square-integrable functions $L^2(\mathrm{SO}(3), d\rho)$ on the rotation group SO(3), with the invariant measure $d\rho = d\varphi d\cos\theta d\chi$. They characterize the signal around each point ω_0, and in each orientation χ. This defines the scale-space nature of the wavelet decomposition on the sphere. In this context, some basic knowledge of harmonic analysis on both $\mathbb{S}^2$ and SO(3) is absolutely essential.

4.2.2 Harmonic Analysis

4.2.2.1 On the Sphere $\mathbb{S}^2$

As discussed, any point ω on the sphere $\mathbb{S}^2$ may be identified as $\omega = (\theta, \varphi)$, with $\theta \in [0, \pi]$ and $\varphi \in [0, 2\pi)$, and we consider signals in $L^2(\mathbb{S}^2, d\Omega)$. The harmonic analysis in this space may be summarized as follows. The spherical harmonics $Y_{lm}(\omega)$ form an orthonormal basis of $L^2(\mathbb{S}^2, d\Omega)$, with $l \in \mathbb{N}$, $m \in \mathbb{Z}$, and $|m| \leq l$. They are explicitly given in a factorized form in terms of the associated Legendre polynomials $P_l^m(\cos\theta)$ and the complex exponentials $e^{im\varphi}$ as

$$Y_{lm}(\theta, \varphi) = \left[\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}\right]^{1/2} P_l^m(\cos\theta)\, e^{im\varphi}. \tag{4.2}$$

The index l represents an overall frequency on the sphere. The absolute value $|m|$ represents the frequency associated with the azimuthal variable φ. The definition

(4.2) corresponds to the choice of Condon–Shortley phase $(-1)^m$ for the spherical harmonics, ensuring the relation

$$(-1)^m Y_{lm}^*(\omega) = Y_{l(-m)}(\omega). \tag{4.3}$$

This phase is included in the definition of the associated Legendre polynomials [1, 229]. Another convention [31] explicitly transfers it to the spherical harmonics. The orthonormality and completeness relations for the spherical harmonics respectively read as

$$\int_{\mathbb{S}^2} d\Omega\, Y_{lm}^*(\omega)\, Y_{l'm'}(\omega) = \delta_{ll'}\,\delta_{mm'} \tag{4.4}$$

and

$$\sum_{l\in\mathbb{N}} \sum_{|m|\leq l} Y_{lm}^*\left(\omega'\right) Y_{lm}(\omega) = \delta^2\left(\omega' - \omega\right), \tag{4.5}$$

with the notation $\delta^2(\omega' - \omega) \equiv \delta(\cos\theta' - \cos\theta)\delta(\varphi' - \varphi)$.

Any function $G \in \mathrm{L}^2(\mathbb{S}^2, d\Omega)$ is thus uniquely given as a linear combination of spherical harmonics:

$$G(\omega) = \sum_{l\in\mathbb{N}} \sum_{|m|\leq l} \widehat{G}_{lm} Y_{lm}(\omega). \tag{4.6}$$

This combination defines the inverse spherical harmonic transform on $\mathbb{S}^2$. The corresponding spherical harmonic coefficients are given by the scalar products in $\mathrm{L}^2(\mathbb{S}^2, d\Omega)$:

$$\widehat{G}_{lm} = \int_{\mathbb{S}^2} d\Omega\, Y_{lm}^*(\omega)\, G(\omega), \tag{4.7}$$

with $l \in \mathbb{N}$, $m \in \mathbb{Z}$, and $|m| \leq l$.

By definition, any function $G \in \mathrm{L}^2(\mathbb{S}^2, d\Omega)$ explicitly depending on the azimuthal angle φ is said to be directional. It exhibits generic spherical harmonic coefficients $\widehat{G}_{lm}$ for $l \in \mathbb{N}$, $m \in \mathbb{Z}$, and $|m| \leq l$. For a real function G, these spherical harmonic coefficients also satisfy the reality constraint $(-1)^m \widehat{G}_{lm}^* = \widehat{G}_{l(-m)}$ following from the symmetry (4.3). Any function $G \in \mathrm{L}^2(\mathbb{S}^2, d\Omega)$ independent of the azimuthal angle φ is said to be zonal, or axisymmetric: $G = G(\theta)$. It only exhibits nonzero spherical harmonic coefficients for $m = 0$: $\widehat{G}_{lm} = \widehat{G}_{l0}\delta_{m0}$. For a real function G, these spherical harmonic coefficients also satisfy the reality constraint $\widehat{G}_{l0}^* = \widehat{G}_{l0}$, still following from the symmetry (4.3). Such an axisymmetric function is obviously invariant under rotation around itself by any angle $\chi \in [0, 2\pi)$.

Notice that the orthonormality of scalar spherical harmonics implies the following Plancherel relation for $F_1, F_2 \in \mathrm{L}^2(\mathbb{S}^2, d\Omega)$:

$$\langle F_2 | F_1 \rangle = \sum_{l\in\mathbb{N}} \sum_{|m|\leq l} \widehat{(F_2)}_{lm}^* \widehat{(F_1)}_{lm}. \tag{4.8}$$

4.2.2.2 On the Rotation Group SO(3)

Any rotation ρ in the group of rotations in three dimensions SO(3) is given in terms of the three Euler angles $\rho = (\varphi, \theta, \chi)$, with $\theta \in [0, \pi]$, and $\varphi, \chi \in [0, 2\pi)$. We consider signals H in the Hilbert space of square-integrable functions $\mathrm{L}^2(\mathrm{SO(3)}, d\rho)$, with the invariant measure $d\rho = d\varphi d\cos\theta d\chi$. The harmonic analysis in this space may be summarized as follows. The Wigner D-functions are the matrix elements of the irreducible unitary representations of weight l of the group in $\mathrm{L}^2(\mathrm{SO(3)}, d\rho)$. By the Peter–Weyl theorem on compact groups, the matrix elements D^{l*}_{mn} also form an orthogonal basis in $\mathrm{L}^2(\mathrm{SO(3)}, d\rho)$, with $l \in \mathbb{N}$, $m, n \in \mathbb{Z}$, and $|m|, |n| \leq l$. They are explicitly given in a factorized form in terms of the real Wigner d-functions $d^l_{mn}(\theta)$ and the complex exponentials, $e^{-im\varphi}$ and $e^{-in\chi}$, as

$$D^l_{mn}(\varphi, \theta, \chi) = e^{-im\varphi} d^l_{mn}(\theta) e^{-in\chi}. \tag{4.9}$$

Again, l represents an overall frequency on SO(3), and $|m|$ and $|n|$ the frequencies associated with the variables φ and χ, respectively [229, 31]. The Wigner d-functions read as

$$d^l_{mn}(\theta) = \sum_{t=C_1}^{C_2} \frac{(-1)^t \left[(l+m)!\,(l-m)!\,(l+n)!\,(l-n)!\right]^{1/2}}{(l+m-t)!\,(l-n-t)!t!\,(t+n-m)!} \left(\frac{\cos\theta}{2}\right)^{2l+m-n-2t} \times \left(\frac{\sin\theta}{2}\right)^{2t+n-m},$$

with the summation bounds $C_1 = \max(0, m-n)$ and $C_2 = \min(l+m, l-n)$ defined to consider only factorials of positive integers. They satisfy various symmetry properties on their indices. The orthogonality and completeness relations of the Wigner D-functions respectively read as

$$\int_{\mathrm{SO(3)}} d\rho\, D^l_{mn}(\rho) D^{l'*}_{m'n'}(\rho) = \frac{8\pi^2}{2l+1} \delta_{ll'} \delta_{mm'} \delta_{nn'} \tag{4.10}$$

and

$$\sum_{l \in \mathbb{N}} \frac{2l+1}{8\pi^2} \sum_{|m|,|n| \leq l} D^l_{mn}(\rho') D^{l*}_{mn}(\rho) = \delta^3(\rho' - \rho), \tag{4.11}$$

with $\delta^3(\rho' - \rho) \equiv \delta(\varphi' - \varphi)\delta(\cos\theta' - \cos\theta)\delta(\chi' - \chi)$. Notice that for $n = 0$, the Wigner D-functions are independent of χ and simply identify with the spherical harmonics:

$$D^l_{m0}(\omega) = \left[\frac{4\pi}{2l+1}\right]^{1/2} Y^*_{lm}(\omega). \tag{4.12}$$

Any function $H \in \mathrm{L}^2(\mathrm{SO(3)}, d\rho)$ is thus uniquely given as a linear combination of Wigner D-functions:

$$H(\rho) = \sum_{l \in \mathbb{N}} \frac{2l+1}{8\pi^2} \sum_{|m|,|n| \leq l} \widehat{H}^l_{mn} D^{l*}_{mn}(\rho). \tag{4.13}$$

This combination defines the inverse Wigner D-function transform on SO(3). The corresponding Wigner D-function coefficients are given by the scalar products in $\mathrm{L}^2(\mathrm{SO(3)}, d\rho)$:

$$\widehat{H}^l_{mn} = \int_{\mathrm{SO(3)}} d\rho \, D^l_{mn}(\rho) H(\rho), \tag{4.14}$$

with $l \in \mathbb{N}$, $m, n \in \mathbb{Z}$, and $|m|, |n| \leq l$.

4.2.3 *Affine Transformations*

4.2.3.1 Translations and Rotations

Continuous translations and rotations of square-integrable functions on the sphere are described by the three Euler angles defining an element $\rho = (\varphi_0, \theta_0, \chi)$ in SO(3). The operator $R(\omega_0)$ in $\mathrm{L}^2(\mathbb{S}^2, d\Omega)$ for the translation of amplitude $\omega_0 = (\theta_0, \varphi_0)$ of a function G reads as

$$G_{\omega_0}(\omega) = [R(\omega_0) G](\omega) = G\left(R^{-1}_{\omega_0} \omega\right), \tag{4.15}$$

where $R_{\omega_0}(\theta, \varphi) = [R^{\hat{z}}_{\varphi_0} R^{\hat{y}}_{\theta_0}](\theta, \varphi)$ is defined by the three-dimensional rotation matrices $R^{\hat{y}}_{\theta_0}$ and $R^{\hat{z}}_{\varphi_0}$, acting on the Cartesian coordinates (x, y, z) associated with $\omega = (\theta, \varphi)$. The rotation operator $R^{\hat{z}}(\chi)$ in $\mathrm{L}^2(\mathbb{S}^2, d\Omega)$ for the rotation of the function G around itself, by an angle $\chi \in [0, 2\pi)$, is given as

$$G_{\chi}(\omega) = \left[R^{\hat{z}}(\chi) G\right](\omega) = G\left({R^{\hat{z}}_{\chi}}^{-1} \omega\right), \tag{4.16}$$

where $R^{\hat{z}}_{\chi}(\theta, \varphi) = (\theta, \varphi + \chi)$ also follows from the action of the three-dimensional rotation matrix $R^{\hat{z}}_{\chi}$ on the Cartesian coordinates (x, y, z) associated with $\omega = (\theta, \varphi)$. The operator incorporating both the translations and rotations simply reads as $R(\rho) = R(\omega_0) R^{\hat{z}}(\chi)$ and $G_{\rho}(\omega) = [R(\rho) G](\omega) = G(R^{-1}_{\rho} \omega)$, with $R_{\rho} = R_{\omega_0} R^{\hat{z}}_{\chi}$. Notice that the action of the operator $R(\rho)$ on $G \in \mathrm{L}^2(\mathbb{S}^2, d\Omega)$ reads in terms of its spherical harmonic coefficients as

$$\widehat{(G_{\rho})}_{lm} = \sum_{|n| \leq l} D^l_{mn}(\rho) \widehat{G}_{ln}. \tag{4.17}$$

4.2.3.2 Stereographic Dilation

As already stated, there is no natural dilation operator for functions defined on S^2. One of the main compelling reasons behind this difficulty is the compactness of the sphere: It is not possible to linearly scale the geodesic distance between points as one would in $\mathbb{R}^2$. However, intuitively at least, the size of compact features on the sphere may be associated with a given scale. By analogy with the Euclidean case, dilation may a priori be defined both in real or harmonic space on $\mathbb{S}^2$. But as we shall see, and contrary to the Euclidean setting, these definitions do not necessarily coincide. This will lead us to study several proposed dilations, each of which is well defined mathematically and has particular advantages. We will thus formulate the wavelet formalism so that we can incorporate those various definitions into a single, unifying framework.

The stereographic dilation of functions is a natural candidate if one wants to define dilations explicitly in real space on $\mathbb{S}^2$. This naturally appears in the wavelet formalism on the sphere originally proposed by (3–6), further developed by [235, 236], and reviewed in [7, 237]. The stereographic dilation operator $D(a)$ on $G \in \mathrm{L}^2(\mathbb{S}^2, d\Omega)$, for a continuous dilation factor $a \in \mathbb{R}_+^*$, is defined in terms of the inverse of the corresponding stereographic dilation D_a on points in $\mathbb{S}^2$. It reads as

$$\begin{aligned} G_a(\omega) &= [D(a)G](\omega) \\ &= \lambda^{1/2}(a,\theta)\, G\left(D_a^{-1}\omega\right), \end{aligned} \tag{4.18}$$

with $\lambda^{1/2}(a,\theta) = a^{-1}[1+\tan^2(\theta/2)]/[1+a^{-2}\tan^2(\theta/2)]$. The dilated point is given by $D_a(\theta,\varphi) = (\theta_a(\theta),\varphi)$ with the linear relation $\tan(\theta_a(\theta)/2) = a\tan(\theta/2)$. The dilation operator therefore maps the sphere without its south pole on itself: $\theta_a(\theta) : \theta \in [0,\pi) \to \theta_a \in [0,\pi)$. This dilation operator is uniquely defined by the requirement of the following natural properties. The dilation of points on $\mathbb{S}^2$ must be a radial (i.e., only affecting the radial variable θ independently of φ, and leaving φ invariant) and conformal (i.e., preserving the measure of angles in the tangent plane at each point) diffeomorphism (i.e., a continuously differentiable bijection). The normalization by $\lambda^{1/2}(a,\theta)$ in (4.18) is uniquely determined by the requirement that the dilation of functions in $\mathrm{L}^2(\mathbb{S}^2, d\Omega)$ be a unitary operator [i.e., preserving the scalar product in $\mathrm{L}^2(\mathbb{S}^2, d\Omega)$, and specifically the norm of functions]. Notice that the stereographic dilation operation is supported by a group structure for the composition law of the corresponding operator $D(a)$. A group homomorphism also holds with the operation of multiplication by a on $\mathbb{R}_+^*$.

Finally, in the Euclidean limit where a function is localized on a small portion of the sphere, this portion is assimilated to the tangent plane, and the stereographic dilation identifies with the standard dilation in the plane [6, 235], which is the expected geometric behavior in this asymptotic regime.

4.2.3.3 Harmonic Dilation

Another possible definition of the dilation of functions may be considered, that which is explicitly defined in harmonic space on $\mathbb{S}^2$. It was proposed in previous developments relative to the definition of a wavelet formalism of axisymmetric [97, 96] and also directional [140, 186] wavelets on the sphere. The harmonic dilation is defined directly on $G \in \mathrm{L}^2(\mathbb{S}^2, d\Omega)$ through a sequence of prescriptions rather than in terms of the application of a simple operator. First, an arbitrary prescription must be chosen to define a set of generating functions $\tilde{G}_m(k)$ of a continuous variable $k \in \mathbb{R}_+$ for each $m \in \mathbb{Z}$. These functions are identified to the spherical harmonic coefficients of G through $\tilde{G}_m(l) = \widehat{G}_{lm}$ for $l \in \mathbb{N}$, and $|m| \leq l$. Second, the variable k is dilated linearly, $k = l \rightarrow k = al$, just as would be the norm of the Fourier frequency on the plane. For a continuous dilation factor $a \in \mathbb{R}_+^*$, the spherical harmonic coefficients of the dilated function G_a are defined by

$$\widehat{(G_a)}_{lm} = \tilde{G}_m(al). \tag{4.19}$$

Again, in the Euclidean limit where a function is localized on a small portion of the sphere, this portion is assimilated to the tangent plane, and the harmonic dilation identifies with the standard dilation in the plane [140].

Notice that in the framework of scale-space signal processing through the linear heat flow on the sphere [34, 32, 33], the harmonic dilation applied to axisymmetric filters appears to be an extremely natural procedure. Considering the heat diffusion equation on the sphere, one may understand the signal F to be analyzed as an initial temperature distribution, at time $t = 0$. The analysis of the signal is performed through the analysis of the temperature distribution at any instant $t \in \mathbb{R}_+$ in the course of the diffusion process, which reveals larger and larger scales in the signal. The heat kernel is an axisymmetric function Φ_{heat} defined as a function of time t by the following spherical harmonic coefficients:

$$[\widehat{(\Phi_{\text{heat}})_t}]_{lm} = \sqrt{\frac{2l+1}{4\pi}}\, e^{-l(l+1)t}\, \delta_{m0}. \tag{4.20}$$

The solution of the heat equation for an initial condition F simply results from its scalar products with the heat kernel at any $\omega_0 \in \mathbb{S}^2$ and $t \in \mathbb{R}_+$:

$$W^F_{\Phi_{\text{heat}}}(\omega_0, t) = \langle (\Phi_{\text{heat}})_{\omega_0,t} | F \rangle. \tag{4.21}$$

At $t = 0$, the heat kernel is given by the optimally localized Dirac delta distribution and the initial temperature distribution identifies with the signal itself. In the limit $t \rightarrow \infty$, the kernel is a constant function on the sphere with the unique coefficient $[\widehat{(\Phi_{\text{heat}})_{t\rightarrow\infty}}]_{00} = (4\pi)^{-1/2}$ and the well-known constant asymptotic temperature distribution identifies with the mean of the signal. The dilation process in this context is very similar to the harmonic dilation applied to axisymmetric functions. A generating function for Φ_{heat} can be defined as $\tilde{\Phi}_{\text{heat}}(k) = e^{-k}$, and the continuous variable $k \in \mathbb{R}_+$ is dilated linearly. However, this variable is not identified with the spherical harmonic index l itself, but with $l(l+1)$. Notice that the dilation process is additive in the time variable $t \in \mathbb{R}_+$, rather than multiplicative in the corresponding dilation factor $a \in \mathbb{R}_+^*$ in the case of the harmonic dilation. Directional filters were also considered in this context [32].

On the one hand, the very simple action of the harmonic dilation in harmonic space also exhibits several advantages relative to the stereographic dilation. Notably, the harmonic dilation ensures that the band limit of a wavelet and of the corresponding wavelet coefficients is reduced by a factor a. Such a multiresolution property is essential in reducing the memory and computation time requirements for the wavelet analysis of signals. On the other hand, the harmonic dilation lacks some of the important properties that hold under stereographic dilation. Notably, as the harmonic dilation does not act on points, the question of the corresponding properties of a radial and conformal diffeomorphism make no sense. The harmonic dilation of functions is also not a unitary procedure. Moreover, as the harmonic dilation is explicitly defined in harmonic space, the evolution in real space of the localization and directionality properties of functions on the sphere through harmonic dilation is not known analytically. In order to circumvent this last drawback, the definition of harmonic dilation may be slightly amended to obtain the kernel dilation defined next.

4.2.3.4 Kernel Dilation

A function $G \in \mathrm{L}^2(\mathbb{S}^2, d\Omega)$ can be defined to be a factorized function in harmonic space if it can be written in the form

$$\widehat{G}_{lm} = \tilde{K}_G(l)\, S^G_{lm}, \tag{4.22}$$

for $l \in \mathbb{N}$ and $|m| \leq l$. The positive real kernel $\tilde{K}_G(k) \in \mathbb{R}_+$ is a generating function of a continuous variable $k \in \mathbb{R}_+$, initially evaluated on integer values $k = l$. The directionality coefficients S^G_{lm}, for $l \in \mathbb{N}$ and $|m| \leq l$, define the directional split of the function. In particular, for a real function G, they bear the same symmetry relation as the spherical harmonic coefficients $\widehat{G}_{lm}$ themselves: $S^{G*}_{lm} = (-1)^m S^G_{l(-m)}$. Without loss of generality, one can impose

$$\sum_{|m|\leq l} |S^G_{lm}|^2 = 1, \tag{4.23}$$

for the values of l for which S^G_{lm} is nonzero for at least one value of m. Hence, localization properties of a function G, such as a measure of dispersion of angular distances around its central position as weighted by the function values, are governed by the kernel and to a lesser extent by the directional split. Indeed, the power contained in the function G at each allowed value of l is fixed by the kernel only. The norm of $G \in \mathrm{L}^2(\mathbb{S}^2, d\Omega)$ reads as $||G||^2 = \sum_{l\in\mathbb{N}} \tilde{K}^2_G(l)$, where the sum runs over the values of l for which S^G_{lm} is nonzero for at least one value of m. However, the directional split is essential in defining the directionality properties measuring the behavior of the function with the azimuthal variable φ, because it bears the entire dependence of the spherical harmonic coefficients of the function in the index m.

The kernel dilation applied to a factorized function (4.22) is simply defined by application of the harmonic dilation (4.19) to the kernel only. The directionality of the dilated function is defined through the same directional split as the original function. For a continuous dilation factor $a \in \mathbb{R}_+^*$, the dilated function therefore reads as

$$\widehat{(G_a)}_{lm} = \tilde{K}_G(al)\, S_{lm}^G. \tag{4.24}$$

Let us emphasize that the directionality coefficients S_{lm}^G are not affected by dilations, contrary to what the complete action of the harmonic dilation (4.19) would imply. The specific directionality properties of the function may, however, be modified through kernel dilation due to the modification of the values of l identifying the dilated kernel. Also, notice that the kernel and harmonic dilations strictly identify with one another when applied to factorized axisymmetric functions A, for which the directional split takes the trivial values $S_{lm}^A = \delta_{m0}$ for $l \in \mathbb{N}$.

Any function $G \in \mathrm{L}^2(\mathbb{S}^2, d\Omega)$ can be said to have a compact harmonic support in the interval $l \in (\lfloor \alpha^{-1}B \rfloor, B)$, for any $B \in \mathbb{N}^0$ and any real value $\alpha > 1$, if

$$\widehat{G}_{lm} = 0 \quad \text{for all} \quad l,m \quad \text{with} \quad l \notin \left(\lfloor \alpha^{-1}B \rfloor, B\right), \tag{4.25}$$

where $\lfloor x \rfloor$ denotes the largest integer value below $x \in \mathbb{R}$. Notice that the compactness of the harmonic support of G can be defined as the ratio of the band limit to the width of its support interval. For a factorized function G of the form (4.22), the compact harmonic support in the interval $l \in (\lfloor \alpha^{-1}B \rfloor, B)$ is ensured by the choice of a kernel with compact support in the interval $k \in (\alpha^{-1}B, B)$:

$$\tilde{K}_G(k) = 0 \quad \text{for} \quad k \notin (\alpha^{-1}B, B). \tag{4.26}$$

The compactness of the harmonic support of G can simply be estimated from the compact support of the kernel as $c(\alpha) = \alpha/(\alpha - 1) \in [1, \infty)$. One has $c(\alpha) \to \infty$ when $\alpha \to 1$, and $c(\alpha) \to 1$ when $\alpha \to \infty$. Typical values would be $\alpha = 2$, corresponding to a compactness $c(2) = 2$, or $\alpha = 1.1$, leading to a higher compactness $c(1.1) = 11$. By a kernel dilation with a dilation factor $a \in \mathbb{R}_+^*$ in (4.24), the compact support of the dilated kernel $\tilde{K}_G(ak) \in \mathbb{R}_+$ is defined in the interval $k \in (a^{-1}\alpha^{-1}B, a^{-1}B)$. The compact harmonic support of the dilated function G_a itself is thus defined in the corresponding interval $l \in (\lfloor a^{-1}\alpha^{-1}B \rfloor, \lceil a^{-1}B \rceil)$, where $\lceil x \rceil$ denotes the smallest integer value above $x \in \mathbb{R}$. In particular, the compactness of the harmonic support of a function remains invariant through a kernel dilation.

The factorization and compact harmonic support, together with the notion of steerability introduced in Section 4.3.1, have been shown to be important properties that ensure good control of the evolution of localization and directionality properties of functions through kernel dilation [238]. In particular considering functions with directionality coefficients S_{lm}^G that become independent of l in the limit $l \to \infty$, it can be shown that the kernel dilation identifies with the standard dilation in the plane in the Euclidean limit [140].

Several definitions of dilations in real or harmonic space may hold for the development of a wavelet formalism on the sphere, each of which exhibits its specific advantages. In the following, we review various wavelet formalisms with different definitions for a dilation operation. We review both continuous and discrete formalisms, respectively identified by continuous and discrete position, orientation, and scale parameters.

4.3 Continuous Formalism

In this section, we begin with a definition of a continuous wavelet formalism relying on a generic dilation operation. We notably introduce the notion of steerability of the analysis function, which will reveal to be essential in many respects in the context of a wavelet formalism. We then consider the cases of the stereographic and kernel dilations. We also comment on the necessary discretization of the translation, rotation, and dilation parameters in the perspective of practical implementations of the wavelet formalism.

4.3.1 Generic Wavelets

4.3.1.1 Directional Case

We consider the general case of analysis of a real signal F with a real and directional analysis function Ψ. In a continuous wavelet formalism, the wavelet coefficients $W_\Psi^F(\rho,a)$ of F with Ψ are defined at each continuous scale $a \in \mathbb{R}_+^*$, around each continuous point $\omega_0 \in \mathbb{S}^2$, and in each continuous orientation $\chi \in [0,2\pi)$, through the directional correlations (4.1). These coefficients living on SO(3) follow a very simple expression in harmonic space. At each scale a, the directional correlation reads as an inverse Wigner D-function transform:

$$W_\Psi^F(\rho,a) = \sum_{l\in\mathbb{N}} \frac{2l+1}{8\pi^2} \sum_{|m|,|n|\leq l} \widehat{\left(W_\Psi^F\right)}^{l}_{mn}(a)\, D^{l*}_{mn}(\rho)\,. \tag{4.27}$$

The Wigner D-function coefficients in this relation follow from relations (4.8) and (4.17) as the pointwise product of the spherical harmonic coefficients of the signal and the wavelet:

$$\widehat{\left(W_\Psi^F\right)}^{l}_{mn}(a) = \frac{8\pi^2}{2l+1} \widehat{(\Psi_a)}^{*}_{ln} \widehat{F}_{lm}. \tag{4.28}$$

The operation of decomposition of the signal F in its wavelet coefficients W_Ψ^F with the analysis function Ψ may somewhat abusively be called analysis. An essential feature of a wavelet formalism, by opposition with generic filtering, resides in the formal possibility of reconstruction of the signal from its wavelet coefficients. This actually raises the analysis function to the rank of a wavelet. A reconstruction formula

$$F(\omega) = \int_{\mathbb{R}_+^*} d\mu(a) \int_{\mathrm{SO(3)}} d\rho\, W_\Psi^F(\rho,a)\left[R(\rho)\, L_\Psi \Psi_a\right](\omega) \tag{4.29}$$

directly follows from (4.28) for a generic scale integration measure $d\mu(a)$. This measure is fixed by each formalism relying on a specific definition of the dilation operation. The operator L_Ψ in $\mathrm{L}^2(\mathbb{S}^2, d\Omega)$ is defined by its action on the spherical harmonic coefficients of a function G: $\widehat{L_\Psi G}_{lm} = \widehat{G}_{lm}/C_\Psi^l$. The reconstruction formula holds if and only if the analysis function satisfies the following admissibility condition for all $l \in \mathbb{N}$:

$$0 < C_\Psi^l = \frac{8\pi^2}{2l+1} \sum_{|m| \le l} \int_{\mathbb{R}_+^*} d\mu(a)\, |\widehat{(\Psi_a)}_{lm}|^2 < \infty. \tag{4.30}$$

This intuitively requires that the whole wavelet family $\{\Psi_a(\omega)\}$, for $a \in \mathbb{R}_+^*$, covers each frequency index l with a finite and nonzero amplitude, hence preserving the signal information at each frequency.

4.3.1.2 Steerability

The steerability of a function $G \in \mathrm{L}^2(\mathbb{S}^2, d\Omega)$ represents a notion of controlled directionality. By definition, a function $G \in \mathrm{L}^2(\mathbb{S}^2, d\Omega)$ is steerable if any rotation of the function around itself may be expressed as a linear combination of a finite number M of basis functions G_p:

$$G_\chi(\omega) = \sum_{p=0}^{M-1} k_p(\chi)\, G_p(\omega). \tag{4.31}$$

The square-integrable functions $k_p(\chi)$ on the unit circle $\mathrm{S}^1 \equiv [0, 2\pi)$, with $0 \le p \le M-1$, and $M \in \mathbb{N}^0$, are called *interpolation weights*.

The generic continuous wavelet formalism described is obviously directly applicable to steerable analysis functions. If the analysis function Ψ is steerable with M basis functions Ψ_p and weights $k_p(\chi)$, the linearity of the directional correlation (4.1) automatically implies that the steerability relation holds identically on the wavelet coefficients:

$$W_\Psi^F(\rho,a) = \sum_{p=0}^{M-1} k_p(\chi)\, W_{\Psi_p}^F(\omega_0,a), \tag{4.32}$$

for $\rho = (\varphi_0, \theta_0, \chi)$ and $\omega_0 = (\varphi_0, \theta_0)$. In this relation, the wavelet coefficients $W_{\Psi_p}^F(\omega_0,a)$ simply follow from the standard correlations with the basis functions Ψ_p:

$$W_{\Psi_p}^F(\omega_0,a) = \langle (\Psi_p)_{\omega_0,a} | F \rangle. \tag{4.33}$$

Consequently, at a scale a and a point ω_0, the exact value of the wavelet coefficient in any continuous orientation χ is known on the basis of the computation of the finite number M of the coefficients $W^F_{\Psi_p}(\omega_0, a)$.

This property was actually first introduced on the plane [98, 212], and more recently defined on the sphere [235, 238]. It is of great interest in the context of wavelet analysis in a computational perspective both for analysis of local signal orientations (see Section 4.4.2.3) as well as for exact signal reconstruction (see Section 4.6.1). In this perspective, this notion of steerability is further discussed in the following paragraphs and an equivalent definition in harmonic space is established.

Intuitively, steerable functions have a nonzero angular width in the azimuthal angle φ, which renders them sensitive to a range of directions and enables them to satisfy the steerability relation. This nonzero angular width naturally corresponds to an azimuthal band limit $N \in \mathbb{N}^0$ in the frequency index m associated with the azimuthal variable φ:

$$\widehat{G}_{lm} = 0 \quad \text{for all} \quad l, m \quad \text{with} \quad |m| \geq N. \tag{4.34}$$

It can actually be shown that the property of steerability (4.31) is equivalent to the existence of an azimuthal band limit N (4.34). First, if a function G is steerable with M basis functions, then the number T of values of m for which $\widehat{G}_{lm}$ has a nonzero value for at least one value of l is less than or equal to M: $M \geq T$. This was first established for functions on the plane [98], and the proof is absolutely identical on the sphere. As a consequence, the function has some azimuthal band limit N, with $T \leq 2N-1$. Second, if a function G has an azimuthal band limit N, then it is steerable, and the number of basis functions can be reduced at least to $M = 2N-1$. This second part of the equivalence can be proved by explicitly deriving a steerability relation for band-limited functions with an azimuthal band limit N. Any band-limited function G can in particular be steered using M rotated versions $G_{\chi_p} = R^{\hat{z}}(\chi_p)G$ as basis functions, and interpolation weights given by simple translations by χ_p of a unique square-integrable function $k(\chi)$ on the circle $\mathbb{S}^1$:

$$G_\chi(\omega) = \sum_{p=0}^{M-1} k(\chi - \chi_p)\, G_{\chi_p}(\omega), \tag{4.35}$$

for specific rotation angles χ_p with $0 \leq p \leq M-1$. One may choose $M = 2N-1$ equally spaced rotation angles $\chi_p \in [0, 2\pi)$ as $\chi_p = 2\pi p/(2N-1)$, with $0 \leq p \leq 2N-2$. The function $k(\chi)$ is then defined by the Fourier coefficients $\widehat{k}_m = 1/(2N-1)$ for $|m| \leq N-1$, and $\widehat{k}_m = 0$ otherwise. Notice that the angles χ_p and the structure of the function $k(\chi)$ are independent of the explicit nonzero values $\widehat{G}_{lm}$.

Typically, if $\widehat{G}_{lm}$ has a nonzero value for at least one value of l for all m with $|m| \leq N-1$, then $T = 2N-1$ and the function is optimally steered by these $M = T$ angles and the function $k(\chi)$ described. On the contrary, when values of m, with $|m| \leq N-1$, exist for which $\widehat{G}_{lm} = 0$ for all values of l, then $T < 2N-1$ and one might want to reduce the number $M = 2N-1$ of basis functions. Depending on the distribution of the T values of m for which $\widehat{G}_{lm}$ has a nonzero value for at least one value of l, the number of basis functions required to steer the band-limited function may indeed be optimized to its smallest possible value $M = T$. This optimization is notably reachable for functions with specific distributions of the T values of m, corresponding to particular symmetries in real space. For example, a function G is even or odd through rotation around itself by $\chi = \pi$ if and only if $\widehat{G}_{lm}$ has nonzero values only for, respectively, even or odd values of m. This property notably implies that the central position of the function G identifies with the north pole, in the sense that its modulus $|G|$ is always even through rotation around itself by $\chi = \pi$. The combination of an azimuthal band limit N with that symmetry reads as

$$\widehat{G}_{lm} = 0 \quad \text{for all} \quad l,m \quad \text{with} \quad m \notin T_N, \tag{4.36}$$

with

$$T_N = \{-(N-1), -(N-3), \ldots, (N-3), (N-1)\}. \tag{4.37}$$

In this particular case, $T = N$ and one may choose $M = N$ equally spaced rotation angles $\chi_p \in [0, \pi)$ as $\chi_p = \pi p/N$, with $0 \leq p \leq N-1$, and steer the function through relation (4.35). The function $k(\chi)$ is defined by the Fourier coefficients $\widehat{k}_m = 1/N$ for $m \in T_N$, and $\widehat{k}_m = 0$ otherwise.

The lower the azimuthal band limit N of the filter, the smaller the number of basis functions M required for its steerability. In particular, the axisymmetry of a function may be understood as an extreme case of steerability, for an azimuthal band limit $N = 1$, and a number of basis functions $M = 1$. By opposition, one may understand a function as optimally directional if it is only sensitive to the specific direction χ in which it is rotated. Such a function would have an azimuthal dependence $\sim \delta(\varphi)$ thus containing nonzero coefficients for an infinite number of values of m, i.e., $N \to \infty$. An infinite number of weights would thus be required to steer such a function. Optimal directionality and steerability are thus competing concepts.

4.3.1.3 Axisymmetric Case

We consider the particular case of analysis of a real signal F with a real and axisymmetric analysis function Θ, in a continuous wavelet formalism. The directional correlation of F with Θ is obviously independent of the rotation angle χ. As such, it reduces to a so-called standard correlation [236]:

$$W_{\Theta}^{F}(\omega_0, a) = \langle \Theta_{\omega_0,a} | F \rangle. \tag{4.38}$$

At each scale a, the wavelet coefficients identify a square-integrable function on $\mathbb{S}^2$ rather than on SO(3), which reads as an inverse spherical harmonic transform:

$$W_{\Theta}^{F}(\omega_0, a) = \sum_{l \in \mathbb{N}} \sum_{|m| \leq l} \widehat{\left(W_{\Theta}^{F}\right)}_{lm}(a)\, Y_{lm}(\omega). \tag{4.39}$$

The spherical harmonic coefficients in this relation follow from relation (4.28) as the pointwise product of the spherical harmonic coefficients of the signal and the wavelet:

$$\widehat{\left(W_{\Theta}^{F}\right)}_{lm}(a) = \sqrt{\frac{4\pi}{2l+1}}\,\widehat{(\Theta_a)}_{l0}^{*}\widehat{F}_{lm}. \tag{4.40}$$

The reconstruction of F from its wavelet coefficients reads as

$$F(\omega) = \int_{\mathbb{R}_+^*} d\mu(a) \int_{\mathbb{S}^2} d\omega_0 \, W_{\Theta}^{F}(\omega_0, a) \left[R(\omega_0) L_{\Theta} \Theta_a\right](\omega), \tag{4.41}$$

for any scale integration measure $d\mu(a)$, and with the operator L_{Θ} in $\mathrm{L}^2(\mathbb{S}^2, d\Omega)$ defined by $\widehat{L_{\Theta}G}_{l0} = \widehat{G}_{l0}/C_{\Theta}^{l}$. The reconstruction formula holds if and only if the analysis function satisfies the following admissibility condition for all $l \in \mathbb{N}$:

$$0 < C_{\Theta}^{l} = \frac{4\pi}{2l+1} \int_{\mathbb{R}_+^*} d\mu(a) \, |\widehat{(\Theta_a)}_{l0}|^2 < \infty. \tag{4.42}$$

4.3.2 Stereographic Wavelets

4.3.2.1 Correspondence Principle

When the stereographic dilation is considered, the effect of the dilation on the spherical harmonic coefficients of a function is not easily tractable analytically. Consequently, the admissibility condition (4.30) is difficult to check in practice. It can be shown that the nearly zero-mean condition

$$\frac{1}{4\pi} \int_{\mathbb{S}^2} d\Omega \, \frac{\Psi(\omega)}{1+\cos\theta} = 0 \tag{4.43}$$

is a necessary condition for wavelet admissibility. It is, however, formally not sufficient. On the contrary, wavelets on the plane are well known, and may be easily constructed, as the corresponding admissibility condition reduces to a zero-mean condition for a function that is both integrable and square-integrable. In that context, a correspondence principle was proved [235], stating that the inverse stereographic projection of a wavelet on the plane leads to a wavelet on the sphere.

The stereographic projection is the unique radial conformal diffeomorphism mapping the sphere $\mathbb{S}^2$ onto the plane $\mathbb{R}^2$. The unitary stereographic projection operators between functions $G \in \mathrm{L}^2(\mathbb{S}^2, d\Omega)$ and $g \in \mathrm{L}^2(\mathbb{R}^2, d^2\mathbf{x})$, and its inverse, respectively read as

$$[\Pi G](\mathbf{x}) = \left(1 + \left(\frac{r}{2}\right)^2\right)^{-1} G\left(\pi^{-1}\mathbf{x}\right),$$

$$\left[\Pi^{-1} g\right](\omega) = \left(1 + \tan^2\frac{\theta}{2}\right) g(\pi\omega), \tag{4.44}$$

where $\omega = (\theta, \varphi)$ still identify spherical coordinates on the sphere, and $\mathbf{x} = (r, \varphi)$ identify polar coordinates in the plane tangent to the sphere at the north pole. The azimuthal coordinates on the plane and on the sphere are identified to one another: φ. The radial conformal diffeomorphism between points is given as $\pi(\theta, \varphi) = (r(\theta), \varphi)$ for $r(\theta) = 2\tan(\theta/2)$, and its inverse reads $\pi^{-1}(r, \varphi) = (\theta(r), \varphi)$ for $\theta(r) = 2\arctan(r/2)$. The diffeomorphism $r(\theta)$ and its inverse $\theta(r)$ explicitly define the stereographic projection and its inverse. This stereographic projection maps the sphere, without its south pole, on the entire plane: $r(\theta) : \theta \in [0, \pi[\rightarrow [0, \infty[$. Geometrically, it projects a point $\omega = (\theta, \varphi)$ on the sphere onto a point $\mathbf{x} = (r, \varphi)$ on the tangent plane at the north pole, colinear with ω and the south pole (see Fig. 4.1). The prefactors in (4.44) are required to ensure the unitarity of the projection operators Π and Π^{-1}.

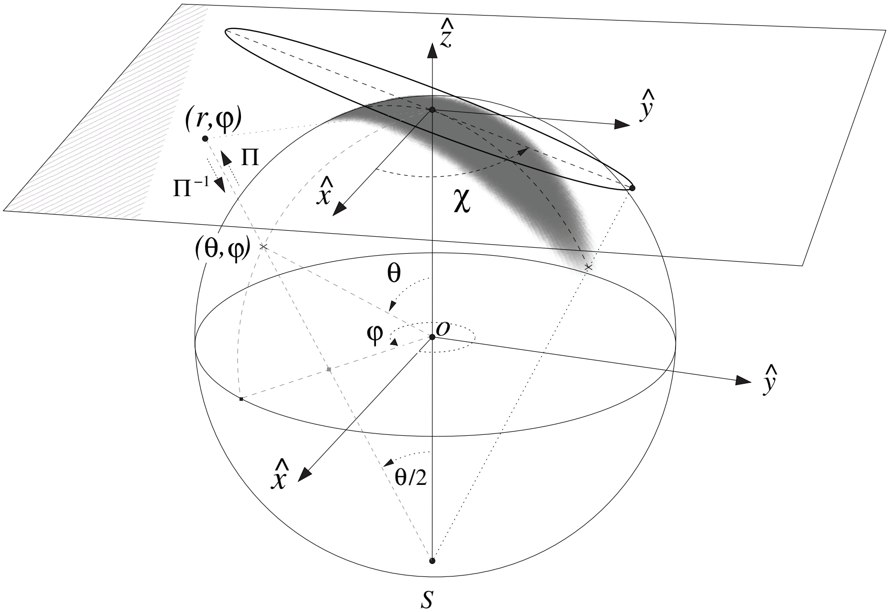


Fig. 4.1: Stereographic projection π and its inverse π^{-1}, relating points (θ, φ) on the sphere and (r, φ) on its tangent plane at the north pole. The same relation holds through Π and Π^{-1} between functions living on each of the two manifolds, as illustrated by the shadow on the sphere and the localized region on the plane. (Figure borrowed from [235].)

In this framework, the correspondence principle established states that if the function $\psi \in \mathrm{L}^2(\mathbb{R}^2, d^2\mathbf{x})$ satisfies the wavelet admissibility condition on the plane, then the function

$$\Psi(\theta, \varphi) = \left[\Pi^{-1}\psi\right](\theta, \varphi), \tag{4.45}$$

in $\mathrm{L}^2(\mathbb{S}^2, d\Omega)$, satisfies the wavelet admissibility condition (4.53) on the sphere. Notice that this correspondence principle requires the definition of a scale integration measure identical to the measure used on the plane: $d\mu(a) = a^{-3}da$.

This enables the construction of wavelets on the sphere by projection of wavelets on the plane. It also transfers wavelet properties from the plane onto the sphere. In particular, as the steerability of functions only depends on their behavior relative to the azimuthal variable φ, it is obviously preserved through stereographic projection, which only affects the function through its dependence in the polar angle θ on the sphere or in the radial variable r on the plane.

4.3.2.2 Example Filters

For the sake of illustration, here we present axisymmetric, directional, and steerable example wavelets on the sphere in the context of the wavelet formalism with stereographic dilation. These wavelets are thus built as inverse stereographic projections of wavelets on the plane.

The axisymmetric Mexican hat wavelet on the plane is defined as the normalized (negative) Laplacian of a Gaussian $e^{-(x^2+y^2)/2}$. Its inverse stereographic projection defines the axisymmetric Mexican hat wavelet on the sphere (see Fig. 4.2). The elliptical Mexican hat wavelet is a directional modification of the axisymmetric Mexican hat, obtained by considering different widths σ_x and σ_y, respectively, in the $\hat{x}$ and $\hat{y}$ directions on the plane for the original Gaussian [187]. The wavelet obtained as the inverse stereographic projection of the (negative) Laplacian of this Gaussian is proportional to (see Fig. 4.2)

$$\Psi(\theta,\varphi) \propto \left(1+\tan^2\frac{\theta}{2}\right)\left[1-\frac{4\tan^2\theta/2}{\sigma_x^2+\sigma_y^2}\left(\frac{\sigma_y^2}{\sigma_x^2}\cos^2\varphi+\frac{\sigma_x^2}{\sigma_y^2}\sin^2\varphi\right)\right] \times e^{-2\tan^2\frac{\theta}{2}\left(\cos^2\varphi/\sigma_x^2+\sin^2\varphi/\sigma_y^2\right)}. \tag{4.46}$$

One can identify the wavelet parameters through the eccentricity of the ellipse defined by the points where the wavelet has zero value (zero-crossing), $\varepsilon = (1-(\sigma_x/\sigma_y)^4)^{1/2}$ (for $\sigma_y \geq \sigma_x$), and the sum $s = \sigma_x^2+\sigma_y^2$. It is alternatively described by the ratio of the semimajor and semiminor axes of the Gaussian $r = \sigma_x/\sigma_y$, and the sum $s = \sigma_x^2+\sigma_y^2$. The axisymmetric Mexican hat is recovered for $\sigma_x = \sigma_y = 1$, in which case $r = 1$ ($\varepsilon = 0$), and $s = 2$.

The real Morlet wavelet on the plane is another typical example of a directional wavelet. Its inverse stereographic projection (see also [64, 187] for similar projections) on the sphere is proportional to (see Figure 4.3)

$$\Psi(\theta,\varphi) \propto \left(1+\tan^2\frac{\theta}{2}\right)\left[\cos\left(\frac{\mathbf{k}\cdot(\pi^{-1}\mathbf{x})}{\sqrt{2}}\right)-e^{-\mathbf{k}^2/4}\right]e^{-2\tan^2(\theta/2)}, \tag{4.47}$$

with $\pi^{-1}\mathbf{x} = (2\tan(\theta/2)\cos\varphi, 2\tan(\theta/2)\sin\varphi)$ in Cartesian coordinates. The arbitrary wave-vector $\mathbf{k} = (k_x,k_y)$ controls the direction and frequency of oscillation of the wavelet ($\mathbf{k}^2 = k_x^2+k_y^2$). Notice that for $|\mathbf{k}| = 2$, the real Morlet wavelet closely approximates at large scales the second Gaussian derivative described in the following.

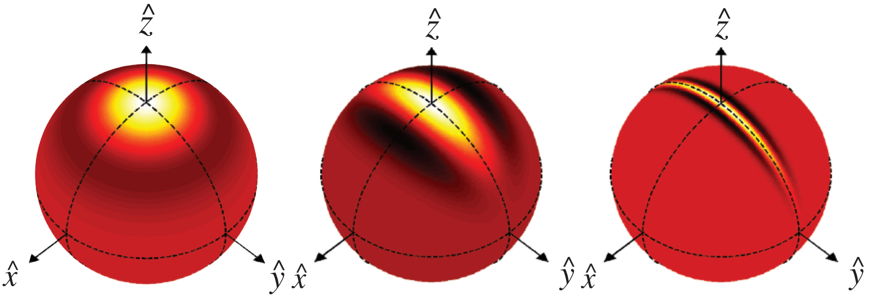


Fig. 4.2: Mexican hat wavelet on the sphere for a dilation factor $a = 0.4$ and different eccentricities. On the left, the axisymmetric Mexican hat: $r = 1$ ($\varepsilon = 0$) and $s = 2$ (left). At the center and on the right, respectively, the elliptical Mexican hat for $r = 0.5$ ($\varepsilon \simeq 0.96825$) and $s = 2$, and $r = 0.1$ ($\varepsilon = 0.99995$) and $s = 2$. Dark and light regions respectively identify negative and positive values. (Figure borrowed from [237].)

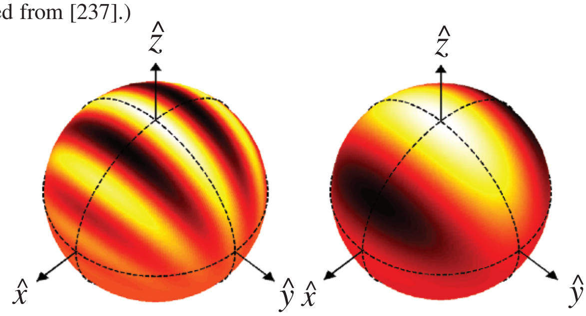


Fig. 4.3: Real Morlet wavelet on the sphere for a dilation factor $a = 0.4$ and a wave-vector $\mathbf{k} = (6, 0)$ on the left, and for a dilation factor $a = 0.4$ and a wave-vector $\mathbf{k} = (2, 0)$ on the right. Dark and light regions respectively identify negative and positive values. (Figure borrowed from [237].)

Derivatives of order $N - 1$ in direction $\hat{x}$ of radial functions on the plane are steerable wavelets. Their inverse stereographic projection thus defines steerable wavelets on the sphere. They have an azimuthal band limit equal to N and may be rotated in terms of $M = N$ basis filters. We give explicit examples of the normalized first and second Gaussian derivatives. A first derivative has a band limit $N = 2$ and only contains the frequencies $m = \{\pm 1\}$. It may be rotated in terms of two specific rotations at $\chi = 0$ and $\chi = \pi/2$, corresponding to the inverse projection of the first derivatives in directions $\hat{x}$ and $\hat{y}$, $\Psi^{\partial_{\hat{x}}}$ and $\Psi^{\partial_{\hat{y}}}$, respectively:

$$\left[R^{\hat{z}}(\chi)\Psi^{\partial_{\hat{x}}}\right](\omega) = \Psi^{\partial_{\hat{x}}}(\omega)\cos\chi + \Psi^{\partial_{\hat{y}}}(\omega)\sin\chi. \tag{4.48}$$

The normalized first derivatives of a Gaussian (see Fig. 4.4) in directions $\hat{x}$ and $\hat{y}$ read

$$\Psi^{\partial_{\hat{x}}}(\theta,\varphi) = \sqrt{\frac{8}{\pi}}\left(1+\tan^2\frac{\theta}{2}\right)\left(\tan\frac{\theta}{2}\cos\varphi\right)e^{-2\tan^2(\theta/2)}$$

$$\Psi^{\partial_{\hat{y}}}(\theta,\varphi) = \sqrt{\frac{8}{\pi}}\left(1+\tan^2\frac{\theta}{2}\right)\left(\tan\frac{\theta}{2}\sin\varphi\right)e^{-2\tan^2(\theta/2)}. \tag{4.49}$$

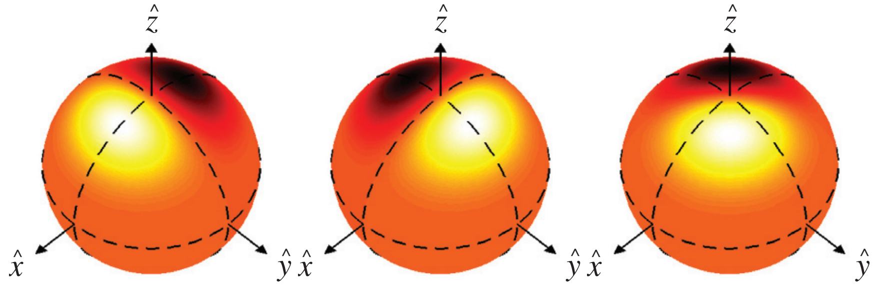


Fig. 4.4: First Gaussian derivative wavelet on the sphere for a dilation factor $a = 0.4$: from left to right, $\Psi^{\partial_{\hat{x}}}$, $\Psi^{\partial_{\hat{y}}}$, and rotation by $\chi = \pi/4$ of $\Psi^{\partial_{\hat{x}}}$. Dark and light regions respectively identify negative and positive values. (Figure borrowed from [235].)

A second derivative has a band limit $N = 3$ and contains the frequencies $m = \{0, \pm 2\}$. It may be rotated in terms of three basis filters. It indeed reads in terms of the inverse projection of the second derivatives in directions $\hat{x}$ and $\hat{y}$, $\Psi^{\partial^2_{\hat{x}}}$ and $\Psi^{\partial^2_{\hat{y}}}$, respectively, and the cross derivative $\Psi^{\partial_{\hat{x}}\partial_{\hat{y}}}$ as

$$\left[R^{\hat{z}}(\chi)\Psi^{\partial^2_{\hat{x}}}\right](\omega) = \Psi^{\partial^2_{\hat{x}}}(\omega)\cos^2\chi + \Psi^{\partial^2_{\hat{y}}}(\omega)\sin^2\chi + \Psi^{\partial_{\hat{x}}\partial_{\hat{y}}}(\omega)\sin 2\chi. \quad (4.50)$$

The correctly normalized second derivatives of a Gaussian (see Fig. 4.5) in directions $\hat{x}$ and $\hat{y}$ read

$$\begin{aligned}
\Psi^{\partial^2_{\hat{x}}}(\theta,\varphi) &= \sqrt{\frac{4}{3\pi}}\left(1+\tan^2\frac{\theta}{2}\right)\left(1-4\tan^2\frac{\theta}{2}\cos^2\varphi\right)e^{-2\tan^2(\theta/2)}\\
\Psi^{\partial^2_{\hat{y}}}(\theta,\varphi) &= \sqrt{\frac{4}{3\pi}}\left(1+\tan^2\frac{\theta}{2}\right)\left(1-4\tan^2\frac{\theta}{2}\sin^2\varphi\right)e^{-2\tan^2(\theta/2)},\\
\Psi^{\partial_{\hat{x}}\partial_{\hat{y}}}(\theta,\varphi) &= -\frac{4}{\sqrt{3\pi}}\left(1+\tan^2\frac{\theta}{2}\right)\left(\tan^2\frac{\theta}{2}\sin 2\varphi\right)e^{-2\tan^2(\theta/2)}.
\end{aligned} \quad (4.51)$$

4.3.3 *Kernel Wavelets*

4.3.3.1 Harmonic Dilation Case

When the harmonic dilation is considered, the analysis function Ψ must satisfy the following form of the admissibility condition (4.30). As a first constraint, one has $\widehat{\Psi}_{00} = \tilde{\Psi}_0(0) = 0$, which corresponds to the requirement that Ψ has a zero mean on the sphere:

$$\frac{1}{4\pi}\int_{\mathbb{S}^2} d\Omega\,\Psi(\omega) = 0. \quad (4.52)$$

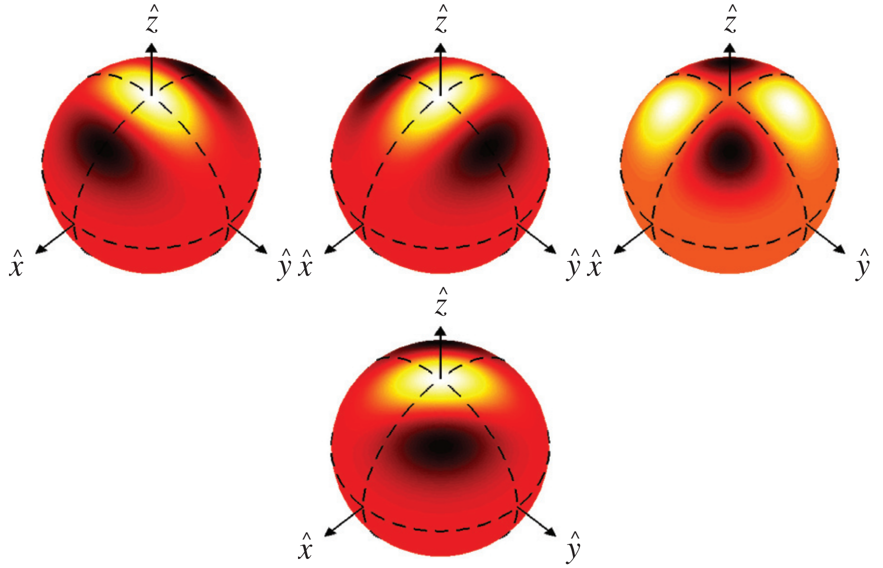


Fig. 4.5: Second Gaussian derivative wavelet on the sphere for a dilation factor $a = 0.4$: from left to right, $\Psi^{\partial_{\hat{x}}^2}$, $\Psi^{\partial_{\hat{y}}^2}$, $\Psi^{\partial_{\hat{x}}\partial_{\hat{y}}}$, and below, rotation by $\chi = \pi/4$ of $\Psi^{\partial_{\hat{x}}^2}$. Dark and light regions respectively identify negative and positive values. (Figure borrowed from [235].)

This zero mean is of course preserved through harmonic dilation. As the zero frequency is not supported by the wavelets, only signals with zero mean can be analyzed in this formalism [see relation (4.28)]. Notice that the scale integration measure can arbitrarily be chosen as $d\mu(a) = a^{-1}da$. This leads to a simple expression of the remaining constraints for $l \in \mathbb{N}^0$ as

$$0 < C_{\Psi}^{l} = \frac{8\pi^2}{2l+1} \sum_{|m| \leq l} \int_{\mathbb{R}_+} \frac{dk'}{k'} \left|\tilde{\Psi}_m\left(k'\right)\right|^2 < \infty. \tag{4.53}$$

The left-hand-side inequality implies $0 < \int_{\mathbb{R}_+} dk'/k' \, |\tilde{\Psi}_{m_0}(k')|^2$ for at least one of the first two generating functions: $m_0 \in \{0,1\}$. In other words, either $\tilde{\Psi}_0$ or $\tilde{\Psi}_1$ must be nonzero on a set of nonzero measure on $\mathbb{R}_+$. The right-hand-side inequality implies $\int_{\mathbb{R}_+} dk'/k' \, |\tilde{\Psi}_m(k')|^2 < \infty$ for all generating functions: $m \in \mathbb{Z}$. Hence, the generating functions must satisfy $\tilde{\Psi}_m(0) = 0$ [this condition encompasses the zero-mean condition (4.52) in the form $\tilde{\Psi}_0(0) = 0$] and tend to zero when $k' \to \infty$. With this choice of scale integration measure, the constraints summarize to the requirement that each generating function satisfies a condition very similar to the wavelet admissibility condition [235, 4] for an axisymmetric wavelet on the plane defined by a Fourier transform identical to $\tilde{\Psi}_m(k)$. Consequently, the wavelet admissibility condition (4.53) can be checked in practice and wavelets associated with the harmonic dilation can be designed easily.

For continuous axisymmetric wavelets, a unique generating function $\tilde{\Theta}_0(k)$ of a continuous variable $k \in \mathbb{R}_+$ is required. The admissibility condition (4.42) reduces to the following expression. The analysis function Θ must have a zero mean and only allows the analysis of signals with zero mean. A unique additional condition holds independently of l:

$$0 < C_\Theta = \int_{\mathbb{R}_+} \frac{dk'}{k'} \, |\tilde{\Theta}_0\left(k'\right)|^2 < \infty. \tag{4.54}$$

This condition actually encompasses the zero-mean condition in the form $\tilde{\Theta}_0(0) = 0$ and also requires that the generating function must tend to zero when $k' \to \infty$. The coefficients entering the reconstruction formula (4.41) read as $C^l_\Theta = 4\pi C_\Theta/(2l+1)$, for $l \in \mathbb{N}^0$.

However, as discussed in Section 4.2.3, the evolution of the localization and directionality properties of functions in real space through harmonic dilation are not explicitly controlled. These requirements are met in the context of the kernel dilation. Consequently, in the following we describe the continuous wavelet formalism based on this kernel dilation. Moreover, this dilation renders the transition between the continuous and scale-discretized formalism much simpler and more transparent than what the harmonic dilation can provide.

4.3.3.2 Kernel Dilation Case

When the kernel dilation is considered, factorized steerable functions $\Psi \in L^2(\mathbb{S}^2, d\Omega)$ with compact harmonic support must be used:

$$\widehat{\Psi}_{lm} = \tilde{K}_\Psi\left(l\right) S^\Psi_{lm}, \tag{4.55}$$

for a continuous kernel defined by a positive real function $\tilde{K}_\Psi(k) \in \mathbb{R}_+$ and a directional split defined by the directionality coefficients S^Ψ_{lm}. The compact harmonic support of the wavelet in the interval $l \in (\lfloor \alpha^{-1}B \rfloor, B)$ is ensured by a kernel $\tilde{K}_\Psi(k)$ with compact support in the interval $k \in (\alpha^{-1}B, B)$, with a compactness $c(\alpha) = \alpha/(\alpha-1) \in [1, \infty)$:

$$\tilde{K}_\Psi\left(k\right) = 0 \quad \text{for} \quad k \notin (\alpha^{-1}B, B). \tag{4.56}$$

The steerability of a wavelet with an azimuthal band limit N is ensured by the directional split:

$$S^\Psi_{lm} = 0 \quad \text{for all} \quad l, m \quad \text{with} \quad |m| \geq N, \tag{4.57}$$

with

$$\sum_{|m| \leq \min(N-1, l)} |S^\Psi_{lm}|^2 = 1, \tag{4.58}$$

for all $l \in \mathbb{N}^0$. Continuous axisymmetric wavelets $\Theta(\theta)$ with compact harmonic support are simply obtained by the trivial directional split with $S^\Theta_{lm} = \delta_{m0}$ for all $l \in \mathbb{N}^0$.

The analysis of a signal $F \in \mathrm{L}^2(\mathbb{S}^2, d\Omega)$ with the analysis function Ψ gives the wavelet coefficients $W_{\Psi}^{F}(\rho, a)$ at each continuous scale a, around each point ω_0, and in each orientation χ, through the directional correlation (4.1). The reconstruction of F from its wavelet coefficients results from relation (4.29). The zero-mean condition (4.52) for the admissibility of Ψ implies $\tilde{K}_{\Psi}(0) = 0$. One can also arbitrarily set $S_{00}^{\Psi} = 0$. The admissibility condition (4.53) summarizes to

$$0 < C_{\Psi} = \int_{(\alpha^{-1}B,B)} \frac{dk'}{k'} \tilde{K}_{\Psi}^{2}\left(k'\right) < \infty, \tag{4.59}$$

which actually also encompasses the zero-mean condition. The coefficients entering the reconstruction formula are $C_{\Psi}^{l} = 8\pi^2 C_{\Psi}/(2l+1)$ for $l \in \mathbb{N}^0$. In other words, the kernel must formally be identified with the Fourier transform of an axisymmetric wavelet on the plane.

Notice that for a factorized wavelet Ψ, the directional correlation defining the analysis of a signal may also be understood as a double correlation, by the kernel and the directional split successively. The standard correlation (4.38) of the signal F and the axisymmetric wavelets defined by the kernel of Ψ provides intermediate wavelet coefficients $W_{\tilde{K}_{\Psi}}^{F}(\omega_0, a)$ on $\mathbb{S}^2$ at each scale $a \in \mathbb{R}_{+}^{*}$. The spherical harmonic transform of these coefficients reads as

$$\widehat{\left(W_{\tilde{K}_{\Psi}}^{F}\right)}_{lm}(a) = \sqrt{\frac{4\pi}{2l+1}} \tilde{K}_{\Psi}(al) \widehat{F}_{lm}. \tag{4.60}$$

At each scale a, the directional correlation of the intermediate signal $W_{\tilde{K}_{\Psi}}^{F}(\omega_0, a)$ and a directional wavelet defined by the directional split of Ψ provides the final wavelet coefficients on SO(3):

$$\widehat{\left(W_{\Psi}^{F}\right)}_{mn}^{l}(a) = \frac{8\pi^2}{2l+1} \left(\sqrt{\frac{2l+1}{4\pi}} S_{ln}^{\Psi}\right)^{*} \widehat{\left(W_{\tilde{K}_{\Psi}}^{F}\right)}_{lm}(a). \tag{4.61}$$

This reasoning obviously holds independently of the steerability or compact harmonic support properties of Ψ.

Let us finally emphasize that even though the steerability of the wavelet is initially set by the directional split, this steerability may be affected through kernel dilation due to the modified compact harmonic support associated with the dilated kernel. However, the computational advantage in relation (4.32) introduced by the existence of an azimuthal band limit in the definition of the directional split is preserved through kernel dilation.

4.3.4 Discretization of Variables

The directional correlation defining wavelet coefficients at each position, orientation, and scale in (4.1) requires integration of functions on a continuous variable on $\mathbb{S}^2$. The reconstruction of the signal at each point on $\mathbb{S}^2$ from its wavelet coefficients in (4.29) also requires integration on a continuous variable on SO(3), as well as on the dilation parameter on $\mathbb{R}_+^*$. Practical implementations must obviously be based on a choice of discretization for each of these variables.

On the one hand, adequate pixelizations of the parameter spaces of the sphere and the rotation group may be designed for approximate or even exact integration on $\omega \in \mathbb{S}^2$ or $\rho \in \mathrm{SO}(3)$ by finite weighted summations, generically called *quadratures*. Quadrature rules obviously rely on the fact that the spaces of integration are compact. On the other hand, continuous scale integration on $a \in \mathbb{R}_+^*$ may not be suitably approximated by quadrature rules. As a consequence, in the framework of a wavelet formalism relying on a continuous dilation parameter, signals may be analyzed at specifically chosen scales by computing of the corresponding wavelet coefficients, but reconstruction is not accessible in practice. In the next section, we describe fast and potentially exact algorithms for the analysis of a signal with a wavelet in that context. The definition of a wavelet formalism in which the dilation operation relies on a discrete dilation parameter is a necessary condition in order to reach in practice signal reconstruction from wavelet coefficients. Applications such as denoising or deconvolution with wavelets obviously require a discrete formalism where the signal under scrutiny may in practice be reconstructed after modification of its wavelet coefficients. Such a discrete wavelet formalism is discussed later in this chapter, along with a corresponding fast and exact algorithm for both analysis and reconstruction.

4.4 Analysis Algorithms

In this section, we discuss choices of pixelizations on the sphere and on the rotation group, and describe two fast algorithms for the analysis of signals in the context of the continuous wavelet formalism developed.

4.4.1 Pixelization

4.4.1.1 Sampling Theorems

We generally consider band-limited signals. Any function $G \in \mathrm{L}^2(\mathbb{S}^2, d\Omega)$ is said to be band-limited with band limit B, for any $B \in \mathbb{N}^0$, if $\widehat{G}_{lm} = 0$ for all l, m with $l \geq B$. Any function $H \in \mathrm{L}^2(\mathrm{SO}(3), d\rho)$ is said to be band-limited with band limit B, for any $B \in \mathbb{N}^0$, if $\widehat{H}^l_{mn} = 0$ for all l, m, n with $l \geq B$. From relation (4.28), if the signal F or the wavelet Ψ is band-limited on $\mathbb{S}^2$, then the wavelet coefficients W^F_Ψ are automatically band-limited on SO(3), with the same band limit B. A continuous band-limited signal $F \in \mathrm{L}^2(\mathbb{S}^2, d\Omega)$ and a continuous wavelet $\Psi \in \mathrm{L}^2(\mathbb{S}^2, d\Omega)$ are respectively identified by the $\mathscr{O}(B^2)$ spherical harmonic coefficients $\widehat{F}_{lm}$ and

$\widehat{\Psi}_{lm}$ with $l \in \mathbb{N}$, $m \in \mathbb{Z}$, and $|m| \leq l$. A simple extrapolation of the Nyquist–Shannon theorem on the line would suggest that the same number $\mathcal{O}(B^2)$ of sampled values $F(\omega_i)$ on points $\omega_i \in \mathbb{S}^2$ is required in order to describe the signal completely. The index i simply identifies the points of the sampling. Typically, $\mathcal{O}(B)$ values are required for both samplings in θ and φ. The directional correlation associated with the wavelet coefficients at each scale $W_\Psi^F(\cdot,a) \in \mathrm{L}^2(\mathrm{SO}(3), d\rho)$ is identified by $\mathcal{O}(B^3)$ Wigner D-function coefficients $\widehat{(W_\Psi^F)}^{l}_{mn}(a)$ with $l \in \mathbb{N}$, $m,n \in \mathbb{Z}$, and $|m|,|n| \leq l$. A similar Nyquist–Shannon extrapolation would also suggest that a number $\mathcal{O}(B^3)$ of sampled values $W_\Psi^F(\rho_i,a)$ on points $\rho_i \in \mathrm{SO}(3)$ are required for the exact description of each directional correlation. Again, $\mathcal{O}(B)$ values are required for samplings in θ_0, φ_0, and χ.

These considerations raise the question of the choice of pixelization of $\mathbb{S}^2$ on which the original signal should be sampled in order to provide precise computation of the required directional correlation at each scale. The same question is raised for the choice of pixelization of SO(3) on which the wavelet coefficients should be computed so that the continuous counterpart of the sampled values is known.

On the sphere $\mathbb{S}^2$, $2B \times 2B$ equiangular pixelizations are defined on points $\omega_{ij} = (\theta_i, \varphi_j)$ for $0 \leq i,j \leq 2B-1$, with a uniform discretization of the coordinates: $\Delta\theta = \theta_{i+1} - \theta_i = \pi/2B$ and $\Delta\varphi = \varphi_{j+1} - \varphi_j = 2\pi/2B$. The specific choice $\theta_0 = \pi/4B$ and $\varphi_0 = 0$ can be made for convenience. It gives $\theta_i = (2i+1)\pi/4B$ and $\varphi_j = 2j\pi/2B$, and excludes the poles of the sampling, which can be convenient for numerical reasons. The pixel centers are identified with the sampling points ω_{ij} defined above. The pixel's edges are identified by meridians shifted by $\Delta\theta/2 = \pi/4B$, and parallel shifted by $\Delta\varphi/2 = 2\pi/4B$ relative to ω_{ij}. The poles therefore appear as pixel corners. Equiangular pixelizations enjoy the so-called iso-latitude property; i.e., the sampling in θ is independent of φ, of interest for computational purposes discussed below. Also, notice that the pixel area varies drastically with θ as $\sin\theta d\theta d\varphi$.

A sampling theorem ensures that exact quadrature rules for integration of signals with band limit B on $\mathbb{S}^2$ exist on $2B \times 2B$ equiangular grids [76]. This sampling theorem represents a generalization of the Nyquist–Shannon theorem on the line. One way of stating it is to say that the spherical harmonic coefficients of a band-limited function on the sphere may be computed exactly up to a band limit B, through an equi-angular sampling, as a finite weighted sum, i.e., a quadrature, of the sampled values of that function [76]. The weights are defined from the structure of the Legendre polynomials $P_l(\cos\theta)$ on the interval $[0,\pi]$.

Other pixelization schemes widely used in astrophysics and cosmology may be considered. The HEALPix pixelization[1] (Hierarchical Equal Area iso-Latitude Pixelization) [105], and the GLESP pixelization[2] (Gauss–Legendre Sky Pixelization) [75, 74] are two major examples. GLESP pixelizations are defined by a sampling of θ on the roots of the Legendre polynomials of some order related to B, and by an equiangular sampling on φ for each value of θ. This scheme provides pixels of

[1] `http://healpix.jpl.nasa.gov/`

[2] `http://www.glesp.nbi.dk/`

nearly equal areas. A sampling theorem also ensures exact quadrature rules on such pixelizations. HEALPix pixelizations are defined through a hierarchical pixelization scheme, containing $12N_{\rm side}^2$ pixels of exactly equal areas at resolution $N_{\rm side} = 2^k$ with $k \in \mathbb{N}$. Only approximate quadrature rules exist, which nevertheless can be made very precise thanks to an iteration process.

On the rotation group SO(3), pixelizations may, for instance, be defined by combining pixelizations on $\mathbb{S}^2$ with an equiangular sampling of χ. A sampling theorem exists on the pixelizations based on equiangular and Gauss–Legendre pixelizations on $\mathbb{S}^2$, hence providing exact quadrature rules for the integration of band-limited signals on SO(3). Again, quadrature rules for pixelizations based on HEALPix pixelizations are approximate. This extension basically relies on the separation of the integration variables [180, 179, 152] from relation (4.9). The sampling theorem notably ensures that the Wigner D-function coefficients of a band-limited function on SO(3) may be computed exactly by quadrature up to a band limit B. A corresponding choice of pixelization for computation of the wavelet coefficients $W_{\Psi}^{F}(\rho_i, a)$ at each scale a will thus provide exact knowledge of the corresponding function in $\mathrm{L}^2(\mathrm{SO}(3), d\rho)$.

4.4.1.2 A Priori Computational Complexity

The computational complexity of an algorithm, or of a part of it, may be defined as the number of basic summation or multiplication operations required to obtain the result from initial data. Each two-dimensional scalar product on $\mathbb{S}^2$ required in the directional correlation relation (4.1) may a priori be computed by quadrature with computational complexity $\mathcal{O}(B^2)$ on $\mathcal{O}(B^2)$ positions and $\mathcal{O}(B)$ orientations, at each analysis scale a. The computation would be exact on those pixelizations where a sampling theorem holds. However, such a $\mathcal{O}(B^5)$ complexity appears to be absolutely unaffordable for fine samplings on the sphere with $B \geq 10^3$. The definition of fast analysis algorithms is consequently essential.

4.4.2 Fast Algorithms

4.4.2.1 Separation of Variables

Let us consider a function $G \in \mathrm{L}^2(\mathbb{S}^2, d\Omega)$ with band limit B and given in terms of its sampled values $G(\omega_i)$ on the $\mathcal{O}(B^2)$ discrete points ω_i of the chosen pixelization of $\mathbb{S}^2$. The factorized form (4.2) of the spherical harmonics naturally enables one to compute a direct spherical harmonic transform by quadrature through separation of the integrations on the variables θ and φ. Conversely, an inverse spherical harmonic transform may be computed as successive summations on the indices l and m, up to the band limit B. Correctly ordering the corresponding operations provides a calculation of direct and inverse spherical harmonic transforms in $\mathcal{O}(B^3)$

operations [76]. This separation of variables in the spherical harmonic transforms simply requires iso-latitude pixelizations on the sphere. The computation of the spherical harmonic coefficients of a band-limited function is theoretically exact only on those pixelization where a sampling theorem holds. The inverse spherical harmonic transform is of course exact independently of the pixelization as it is a simple finite sum truncated by the band limit.

Let us consider a function $H \in \mathrm{L}^2(\mathrm{SO}(3), d\rho)$ with band limit B and given in terms of its sampled values $H(\rho_i)$ on the $\mathscr{O}(B^3)$ discrete points ρ_i of the chosen pixelization of SO(3). The factorized form (4.9) of the Wigner D-functions enables the computation of a direct Wigner D-function transform by quadrature through separation of the integrations on the variables θ_0, φ_0, and χ. Conversely, an inverse Wigner D-function transform may be computed as successive summations on the indices l, m, and n, up to the band limit B. Correctly ordering the corresponding operations provides a calculation of direct and inverse Wigner D-function transform in $\mathscr{O}(B^4)$ operations [180, 179]. Again, iso-latitude pixelizations are required for θ_0 and φ_0 on the sphere, while the separation of variable may be performs for any structure of the sampling in the third Euler angle χ, potentially depending on θ_0 and φ_0. The computation of the Wigner D-function coefficients of a band-limited function is theoretically exact only on those pixelizations where a sampling theorem holds. The inverse Wigner D-function transform is of course exact independently of the pixelization as it is a simple finite sum truncated by the band limit.

In this context, relation (4.28) provides a simple expression for the Wigner D-function coefficients of the directional correlation defining the wavelet coefficients $W_{\Psi}^{F}(\rho, a)$ on SO(3) for a signal F with a wavelet Ψ, at each analysis scale a. This relation is essential to avoid the large $\mathscr{O}(B^5)$ computational complexity for the wavelet coefficients through simple quadrature in real space. A corresponding harmonic space algorithm can be designed. The band-limited signal F is given in terms of its sampled values $F(\omega_i)$ on the $\mathscr{O}(B^2)$ discrete points ω_i of the chosen pixelization of $\mathbb{S}^2$. The algorithm first computes required direct spherical harmonic transforms (4.7) for $\widehat{F}_{lm}$ and $\widehat{(\Psi_a)}_{ln}$, second performs the pointwise product (4.28), and finally computes the inverse Wigner D-function transform (4.27) to obtain the wavelet coefficients. The resulting band-limited coefficients $W_{\Psi}^{F}(\rho, a)$ at each scale a are given in terms of sampled values $W_{\Psi}^{F}(\rho_i, a)$ on the $\mathscr{O}(B^3)$ discrete points ρ_i of the chosen pixelization of SO(3). Using the separation of variables, the computational complexity of the spherical harmonic coefficients is of order $\mathscr{O}(B^3)$. The computational complexity of the pointwise product (4.28) is also of order $\mathscr{O}(B^3)$. Again using the separation of variables, the computation complexity of the inverse Wigner D-function transform is of order $\mathscr{O}(B^4)$, which consequently sets the overall computational complexity of the algorithm.

Consequently, a harmonic space algorithm relying on the separation of variables on iso-latitude pixelizations on the sphere allows computation of a directional correlation in $\mathscr{O}(B^4)$ instead of $\mathscr{O}(B^5)$ operations. The exactness of the computation relies only on the exactness of computation of the spherical harmonic coefficients of the signal and the wavelet, which depends on the existence of a sampling theorem on the pixelization chosen.

4.4.2.2 Factorization of Rotations

A second fast algorithm for the directional correlation of a band-limited signal F with a band limit B and a wavelet Ψ may be designed through the technique of factorization of three-dimensional rotations detailed below.

The rotation operator $R(\rho)$ on functions in $L^2(\mathbb{S}^2, d\Omega)$ may be factorized as [204, 231, 188]

$$R(\varphi_0, \theta_0, \chi) = R\left(\varphi_0 - \frac{\pi}{2}, -\frac{\pi}{2}, \theta_0\right) R\left(0, \frac{\pi}{2}, \chi + \frac{\pi}{2}\right). \tag{4.62}$$

Let us recall that the Wigner D-functions are the matrix elements of the operator $R(\rho)$. The factorized form (4.9) thus provides an alternative expression for the wavelet coefficients at scale a relative to the explicit inverse Wigner D-function transform (4.27) as

$$W_\Psi^F(\rho, a) = \sum_{|m|,|m'|,|n|<B} \widehat{\langle R\Psi_a|F\rangle}_{mm'n} e^{i(m\varphi_0 + m'\theta_0 + n\chi)}. \tag{4.63}$$

The coefficients in this relation read as

$$\widehat{\langle R\Psi_a|F\rangle}_{mm'n} = e^{i(n-m)\pi/2} \sum_{l\geq C} d^l_{m'm}\left(\frac{\pi}{2}\right) d^l_{m'n}\left(\frac{\pi}{2}\right) \widehat{(\Psi_a)}^*_{ln} \widehat{F}_{lm}, \tag{4.64}$$

where $C = \max(|m|, |m'|, |n|)$, and where the symmetry relation $d^l_{m'm}(\theta) = d^l_{mm'}(-\theta)$ was used.

In this context, a harmonic space algorithm must firstly compute required spherical harmonic transforms for $\widehat{F}_{lm}$ and $\widehat{(\Psi_a)}_{ln}$, secondly perform the summation (4.64), and finally compute the inverse transform (4.63) to obtain the wavelet coefficients. The band-limited signal F is given in terms of its sampled values $F(\omega_i)$ on the $\mathscr{O}(B^2)$ discrete points ω_i of the chosen pixelization of $\mathbb{S}^2$. Considering an iso-latitude pixelization for θ and φ, the spherical harmonic transforms may be computed by quadrature in $\mathscr{O}(B^3)$ operations through separation of variables in the spherical harmonics. The computational complexity of the summation (4.64) for all required values of m, m', and n is of order $\mathscr{O}(B^4)$. Again considering an iso-latitude pixelization for θ_0 and φ_0, the computation of the inverse transform (4.63) may be performed by quadrature in $\mathscr{O}(B^4)$ operations through separation of variables in the imaginary exponentials, instead of an explicit separation of variables in the Wigner D-functions themselves. Again, the resulting band-limited coefficients $W_\Psi^F(\rho, a)$ at each scale a are given in terms of sampled values $W_\Psi^F(\rho_i, a)$ on the $\mathscr{O}(B^3)$ discrete points ρ_i of the chosen pixelization of SO(3).

Consequently, a harmonic space algorithm relying on the factorization of rotations on iso-latitude pixelizations on the sphere also allows computation of a directional correlation in $\mathscr{O}(B^4)$ operations, now driven by the intermediate summation (4.64) and the inverse transform (4.63), instead of $\mathscr{O}(B^5)$ operations. The exactness of the computation again relies only on the exactness of the computation of the spherical harmonic coefficients of the signal and the wavelet, which still depends on the existence of a sampling theorem on the pixelization chosen.

Notice that, while the Euler angles φ_0 and χ are in the range $\varphi_0, \chi \in [0, 2\pi)$, the original range for θ_0 is $\theta_0 \in [0, \pi]$, in order to cover the parameter space of SO(3). A formal extension of this interval $\theta_0 \in [0, 2\pi)$ provides the parameter space of the three-torus $\mathbb{T}^3$, covering twice the parameter space of SO(3). In this context, the relation (4.63) represents a three-dimensional inverse Fourier transform, which can be calculated in $\mathscr{O}(B^3 \log_2 B)$ operations on a $2B \times 2B \times 2B$ equiangular grid on $\mathbb{T}^3$ by the use of the standard Cooley–Tukey fast Fourier transform (FFT) algorithm. This optimization, however, does not reduce the overall computational complexity for the directional correlation, still driven by the summation (4.64).

4.4.2.3 Steerable Optimization

Steerable wavelets may be used for further optimization of the algorithmic complexity. Wavelets Ψ are considered with a small number M of basis functions Ψ_p, for $0 \le p \le M-1$ in (4.31), which may actually represent rotations of a unique filter in M basis orientations in (4.35). In other words, the azimuthal band limit N of the wavelets is small relative to the overall band limit B: $M, N \ll B$. The directional correlation with a steerable wavelet reduces to a linear combination of M standard correlations (4.33) with the basis functions. The computational complexity of a directional correlation reduces to that of M standard correlations, with total a priori computational complexity of order $\mathscr{O}(MB^4)$, to which is simply added the $\mathscr{O}(MB^3)$ linear combination that arises from (4.32). Either the technique of separation of variables or the factorization of rotations can be applied to the standard correlation on iso-latitude pixelizations on the sphere, by setting $\chi = 0$ in the relation (4.27) or (4.63), respectively. The computational complexity count for the index n, with $|n| < N$, is also reduced from B down to M. It readily appears that the corresponding computational complexity for the two algorithms hence reduces to $\mathscr{O}(M^2 B^3)$ rather than $\mathscr{O}(MB^4)$ for the directional correlation. When the basis functions actually represent rotations of a unique filter in M basis orientations in (4.35), coefficients in all basis orientations may be computed at once, hence reducing the overall computational complexity for the directional correlation to $\mathscr{O}(MB^3)$. The steerable optimization thus renders the computation more easily affordable, even when multiple signals and multiple scales are considered.

Details on the algorithmic structure, computation times, memory requirements, and numerical stability of the corresponding implementations on HEALPix and equiangular grids on the sphere may be found in [188] for the factorization of rotations, and in [236] for the technique of separation of variables and the optimization with steerable wavelets. Further possible optimization of the algorithm to a computational complexity of order $\mathscr{O}(M^2 B^2 \log_2^2 B)$ can formally be reached through separation of variables on equiangular pixelizations on the sphere. It notably relies on the Driscoll and Healy algorithm for fast spherical harmonic transform [76, 134, 133, 236].

Considering an axisymmetric wavelet Θ, the azimuthal band limit is reduced to $N = 1$: $\widehat{\Theta}_{ln} = 0$ for $|n| \ge 1$. The proper rotation by χ has no effect on the filter, and the directional correlation reduces to a unique standard correlation. At each scale, the wavelet coefficients of a signal with an axisymmetric filter live on the sphere $\mathbb{S}^2$. On iso-latitude pixelizations, the direct spherical harmonic transforms for $\widehat{F}_{lm}$ and $\widehat{(\Theta_a)}_{l0}$, the pointwise product (4.40), and the inverse spherical harmonic transform (4.39) can simply be computed by separation of variables in the spherical harmonics (4.2). This provides an algorithmic structure with $\mathcal{O}(B^3)$ asymptotic complexity, which again can formally be reduced to $\mathcal{O}(B^2 \log^2 B)$ on equiangular pixelizations.

4.5 Discrete Formalism

In this section, we define a discrete wavelet formalism following from a discretization of the scales in the continuous formalism based on the kernel dilation. We also comment on other possible constructions of a discrete wavelet formalism.

4.5.1 Discrete Wavelets

4.5.1.1 Scale Discretization

In the context of the wavelet formalism relying on the kernel dilation, scale-discretized wavelets Γ can simply be obtained from continuous factorized steerable wavelets with compact harmonic support, through an integration by slices of the dilation factor $a \in \mathbb{R}_+^*$. Through this transition procedure, scale-discretized wavelets remain factorized steerable functions with compact harmonic support and are dilated through the same kernel dilation.

We consider the analysis of a signal $F \in \mathrm{L}^2(\mathbb{S}^2, d\Omega)$ with band limit B. The original continuous wavelet $\Psi \in \mathrm{L}^2(\mathbb{S}^2, d\Omega)$ with a compact support is defined in the interval $k \in (\alpha^{-1}B, B)$. The value $\alpha > 1$ regulates the compactness $c(\alpha)$ of Ψ. It is also taken as a basis dilation factor. The discrete dilation factors for the scale-discretized wavelet will correspond to integer powers α^j, for analysis depths $j \in \mathbb{N}$.

The scale-discretized wavelet $\Gamma \in \mathrm{L}^2(\mathbb{S}^2, d\Omega)$ is thus defined in factorized form:

$$\widehat{\Gamma}_{lm} = \tilde{K}_\Gamma(l)\, S^\Gamma_{lm}, \tag{4.65}$$

for a scale-discretized kernel defined by a positive real function $\tilde{K}_\Gamma(k) \in \mathbb{R}_+$ and a directional split defined by the directionality coefficients S^Γ_{lm}. The directional split of Γ is identified with the split of Ψ:

$$S^\Gamma_{lm} = S^\Psi_{lm}, \tag{4.66}$$

also giving

$$S^\Gamma_{lm} = 0 \quad \text{for all} \quad l,m \quad \text{with} \quad |m| \geq N, \tag{4.67}$$

and

$$\sum_{|m|\leq\min(N-1,l)} |S^{\Gamma}_{lm}|^2 = 1, \tag{4.68}$$

for $l \in \mathbb{N}^0$, while $S^{\Gamma}_{00} = 0$. In this relation, N stands for the azimuthal band limit of the steerable wavelets. The identification of the directional splits ensures that the same steerability properties are shared by the continuous wavelet and the scale-discretized wavelet. The scale-discretized kernel $\tilde{K}_{\Gamma}(k)$ is obtained from the continuous kernel $\tilde{K}_{\Psi}(k)$ through an integration by slices of the dilation factor $a \in \mathbb{R}^*_+$ of the continuous wavelet formalism.

As a first step, a positive real scaling function $\tilde{\Phi}_{\Gamma}(k) \in \mathbb{R}_+$ of a continuous variable $k \in \mathbb{R}_+$ is defined that gathers the largest dilation factors $a \in (1,\infty)$, or correspondingly the lowest values of k. This generating function reads for $k \in \mathbb{R}^*_+$ as

$$\tilde{\Phi}^2_{\Gamma}(k) = \frac{1}{C_{\Psi}} \int_1^{\infty} \frac{da}{a} \tilde{K}^2_{\Psi}(ak) = \frac{1}{C_{\Psi}} \int_{(\alpha^{-1}B,B)\cap(k,\infty)} \frac{dk'}{k'} \tilde{K}^2_{\Psi}\left(k'\right) \tag{4.69}$$

and is continuously extended at $k = 0$ by $\tilde{\Phi}^2_{\Gamma}(k) = 1$. The scaling function $\tilde{\Phi}_{\Gamma}(k)$ therefore decreases continuously from unity down to zero in the interval $k \in (\alpha^{-1}B, B)$:

$$\begin{aligned} \tilde{\Phi}_{\Gamma}(k) &= 1 && \text{for} \quad 0 \leq k \leq \alpha^{-1}B, \\ \tilde{\Phi}_{\Gamma}(k) &\in (0,1) && \text{for} \quad \alpha^{-1}B < k < B, \\ \tilde{\Phi}_{\Gamma}(k) &= 0 && \text{for} \quad k \geq B. \end{aligned} \tag{4.70}$$

Notice that similar procedures of scale integration by slices were already proposed in the development of corresponding formalisms for directional wavelets on the plane [78, 190, 228], and in the particular case of axisymmetric wavelets on the sphere [97].

As a second step, a simple Littlewood–Paley decomposition [95] is used to define the scale-discretized kernel $\tilde{K}_{\Gamma}(k)$ by subtracting the scaling function $\tilde{\Phi}_{\Gamma}(k)$ to its contracted version $\tilde{\Phi}_{\Gamma}(\alpha^{-1}k)$. This implicitly sets the value α as the basis dilation factor. The scale-discretized kernel also reads as an integration of the continuous kernel over a slice $a \in (\alpha^{-1}, 1)$ for the dilation factor, or equivalently over a slice $k \in (\alpha^{-1}B, B) \cap (\alpha^{-1}k, k)$ of the compact support interval:

$$\begin{aligned} \tilde{K}^2_{\Gamma}(k) &= \tilde{\Phi}^2_{\Gamma}\left(\alpha^{-1}k\right) - \tilde{\Phi}^2_{\Gamma}(k) \\ &= \frac{1}{C_{\Psi}} \int_{\alpha^{-1}}^{1} \frac{da}{a} \tilde{K}^2_{\Psi}(ak) = \frac{1}{C_{\Psi}} \int_{(\alpha^{-1}B,B)\cap(\alpha^{-1}k,k)} \frac{dk'}{k'} \tilde{K}^2_{\Psi}\left(k'\right). \end{aligned} \tag{4.71}$$

The scale-discretized kernel therefore has a compact support in the interval $k \in (\alpha^{-1}B, \alpha B)$:

$$\tilde{K}_{\Gamma}(k) = 0 \quad \text{for} \quad k \notin \left(\alpha^{-1}B, \alpha B\right). \tag{4.72}$$

This support is wider than that for the original continuous kernel and the scaling function. The corresponding compactness reads as $c(\alpha^2) = \alpha^2/(\alpha^2 - 1) \in [1,\infty)$. The compact harmonic support of the scale-discretized wavelet Γ itself is thus defined in the interval $l \in (\lfloor \alpha^{-1}B \rfloor, \lceil \alpha B \rceil)$. The kernel also satisfies $\tilde{K}_{\Gamma}^2(0) = 0$, leading to a scale-discretized wavelet Γ with a zero mean on the sphere:

$$\frac{1}{4\pi} \int_{\mathbb{S}^2} d\Omega \, \Gamma(\omega) = 0. \tag{4.73}$$

The dilations by α^j of the scale-discretized wavelet are defined by the kernels $\tilde{K}_{\Gamma}(\alpha^j k)$ for any analysis depth $j \in \mathbb{N}$. Each kernel has compact support in the interval $k \in (\alpha^{-(1+j)}B, \alpha^{(1-j)}B)$ and exhibits a maximum at $k = \alpha^{-j}B$, with $\tilde{K}_{\Gamma_{\alpha^j}}(\alpha^{-j}B) = 1$. The scale-discretized wavelet Γ_{α^j} at each analysis depth j thus has a compact harmonic support in the interval $l \in \left(\left\lfloor \alpha^{-(1+j)}B \right\rfloor, \left\lceil \alpha^{(1-j)}B \right\rceil\right)$. The property $\tilde{K}_{\Gamma}(0) = 0$ still ensures that each scale-discretized wavelet has a zero mean on the sphere. Notice that for $j \geq 1$, one gets a dilation factor strictly greater than unity $\alpha^j > 1$, and the scale-discretized wavelet has a band limit less than or equal to the assumed band limit B for the signal F to be analyzed. At $j = 0$, only the values of the kernel in the interval $l \in (\lfloor \alpha^{-1}B \rfloor, B)$ are of interest, as higher frequencies l are truncated by the signal F itself through the directional correlation. One can equivalently consider that the compact support of the kernel is restricted to $k \in (\alpha^{-1}B, B)$ in the definition of the scale-discretized wavelet at this first analysis depth $j = 0$. For $j \leq -1$, the lower bound of the compact harmonic support of the scale-discretized wavelet is larger than the band limit B. The scale-discretized wavelets with negative analysis depths can therefore be discarded, as the result of their directional correlation with the signal F would be identically zero.

4.5.1.2 Invertible Filter Bank

The admissibility condition (4.59) for continuous wavelets simply turns into a resolution of the identity below the band limit by a set of dilated wavelets at various analysis depths j, with $0 \leq j \leq J$, and a dilated scaling function at some total analysis depth $J \in \mathbb{N}$. This defines what one can call a filter bank on the sphere. One gets in particular for $0 \leq k = l < B$:

$$\tilde{\Phi}_{\Gamma}^2 \left(\alpha^J l\right) + \sum_{j=0}^{J} \tilde{K}_{\Gamma}^2 \left(\alpha^j l\right) = 1. \tag{4.74}$$

The scaling function values $\tilde{\Phi}_{\Gamma}(\alpha^J l)$ are equal to unity in the interval $l \in \left[0, \left\lfloor \alpha^{-(1+J)}B \right\rfloor\right]$, then decrease in the interval $l \in \left(\left\lfloor \alpha^{-(1+J)}B \right\rfloor, \left\lceil \alpha^{-J}B \right\rceil\right)$, and are equal to zero for $l \geq \left\lceil \alpha^{-J}B \right\rceil$. The kernel values $\tilde{K}_{\Gamma}(\alpha^j l)$ are nonzero only in the compact harmonic support interval $l \in \left(\left\lfloor \alpha^{-(1+j)}B \right\rfloor, \left\lceil \alpha^{(1-j)}B \right\rceil\right)$. The scaling function typically retains the low-frequency part of the signal, which will not be analyzed. All signal information at frequencies $l \leq \left\lfloor \alpha^{-(1+J)}B \right\rfloor$ is kept only in the scaling function, equal to unity. The wavelets are equal to zero at these frequencies.

All signal information at frequencies $l \geq \lceil \alpha^{-J} B \rceil$ is fully analyzed by the wavelets, while the scaling function is equal to zero. Intermediate frequencies are also analyzed by the wavelets, but the scaling function is required for the reconstruction of the corresponding signal information.

Let us define the maximum analysis depth $J_B(\alpha)$ as the lowest integer value such that $\alpha^{-J_B(\alpha)} B \leq 1$:

$$J_B(\alpha) = \lceil \log_\alpha B \rceil . \tag{4.75}$$

In a case where the total analysis depth would be chosen strictly above $J_B(\alpha)$, all wavelets at analysis depths j with $J \geq j \geq J_B(\alpha) + 1$ would be identically null, as their kernels have a compact support strictly included in the interval $k \in (0,1)$. The total analysis depth is consequently naturally limited by $J \leq J_B(\alpha)$. In the case $J = J_B(\alpha)$, the dilated scaling function evaluated at $\alpha^{J_B(\alpha)} l$ has a nonzero value only at $l = 0$, $\tilde{\Phi}_\Gamma^2(\alpha^{J_B(\alpha)} l) = \delta_{l0}$, while all wavelets are equal to zero at $l = 0$ as they have a zero mean. Hence, the identity can be resolved with $J_B(\alpha) + 1$ dilated wavelets and a trivial scaling function that simply retains the spherical harmonic coefficient $\widehat{F}_{00}$ out of the analysis, or equivalently the mean of the signal over the sphere. One gets in particular for $0 \leq k = l < B$:

$$\delta_{l0} + \sum_{j=0}^{J_B(\alpha)} \tilde{K}_\Gamma^2 \left(\alpha^j l\right) = 1. \tag{4.76}$$

4.5.1.3 Analysis

Following the scale discretization defining the wavelets $\Gamma \in \mathrm{L}^2(\mathbb{S}^2, d\Omega)$, a new scale-discretized wavelet formalism is provided for the analysis and exact reconstruction of band-limited signals.

The analysis of a band-limited signal $F \in \mathrm{L}^2(\mathbb{S}^2, d\Omega)$ with band limit B, with a scale-discretized wavelet Γ, is performed by directional correlations just as in the continuous wavelet formalism. At each analysis depth j with $0 \leq j \leq J \leq J_B(\alpha)$, the wavelet coefficients $W_\Gamma^F(\rho, \alpha^j)$ characterizing the signal around each continuous point $\omega_0 = (\theta_0, \varphi_0) \in \mathbb{S}^2$, and in each continuous orientation $\chi \in [0, 2\pi)$, are still defined by the directional correlation of F with the analysis functions Γ_{α^j} dilated through the kernel dilation by the dilation factor α^j:

$$W_\Gamma^F \left(\rho, \alpha^j\right) = \langle \Gamma_{\rho,\alpha^j} | F \rangle, \tag{4.77}$$

with $\rho = (\varphi_0, \theta_0, \chi)$. Once more, at each analysis depth j, the wavelet coefficients read as an inverse Wigner D-function transform:

$$W_\Gamma^F \left(\rho, \alpha^j\right) = \sum_{l \in \mathbb{N}} \frac{2l+1}{8\pi^2} \sum_{|m|,|n| \leq l} \widehat{\left(W_\Psi^F\right)}^l_{mn} \left(\alpha^j\right) D_{mn}^{l*}(\rho), \tag{4.78}$$

with the Wigner D-function coefficients

$$\widehat{\left(W_{\Gamma}^{F}\right)}_{mn}^{l}\left(\alpha^{j}\right)=\frac{8\pi^{2}}{2l+1}\widehat{\left(\Gamma_{\alpha^{j}}\right)}_{ln}^{*}\widehat{F}_{lm}. \tag{4.79}$$

Again, the factorization relation (4.65) allows one to understand the directional correlation (4.79) as a double correlation, by the kernel and the directional split successively.

4.5.1.4 Exact Reconstruction

The reconstruction of the band-limited signal F from its wavelet coefficients reads in terms of a summation on a finite number $J+1$ of discrete dilation factors:

$$F(\omega)=\left[\Phi_{\alpha^{J}}F\right](\omega)+\sum_{j=0}^{J}\int_{SO(3)}d\rho\, W_{\Gamma}^{F}\left(\rho,\alpha^{j}\right)\left[R(\rho)L^{d}\Gamma_{\alpha^{j}}\right](\omega). \tag{4.80}$$

The approximation $[\Phi_{\alpha^{J}}F](\omega)$ accounts for the part of the signal retained in the scaling function $\tilde{\Phi}_{\Gamma}(\alpha^{J}l)$. In a very similar way to the part of the signal analyzed by the wavelets, it can be written as

$$\left[\Phi_{\alpha^{J}}F\right](\omega)=2\pi\int_{\mathbb{S}^{2}}d\Omega_{0}\, W_{\Phi}^{F}\left(\omega_{0},\alpha^{J}\right)\left[R(\omega_{0})L^{d}\Phi_{\alpha^{J}}\right](\omega), \tag{4.81}$$

with $W_{\Phi}^{F}(\omega_{0},\alpha^{J})=\langle\Phi_{\omega_{0},\alpha^{J}}|F\rangle$, and for an axisymmetric function $\Phi\in L^{2}(\mathbb{S}^{2},d\Omega)$ defined by $\widehat{(\Phi_{\Gamma})}_{lm}=\tilde{\Phi}_{\Gamma}(l)\delta_{m0}$. In the particular case where $J=J_{B}(\alpha)$, one gets $\widehat{(\Phi_{\Gamma})}_{lm}=\delta_{l0}\delta_{m0}$ and the approximation simply reduces to the mean of the signal over the sphere: $[\Phi_{\alpha^{J_{B}(\alpha)}}F]=(4\pi)^{-1}\int_{\mathbb{S}^{2}}d\Omega\, F(\omega)$. The zero-mean signal is completely analyzed by the scale-discretized wavelets. The operator L^{d} in $L^{2}(\mathbb{S}^{2},d\Omega)$ in the present scale-discretized wavelet formalism is defined by the following action on the spherical harmonic coefficients of functions: $\widehat{L^{d}G}_{lm}=(2l+1)\widehat{G}_{lm}/8\pi^{2}$. This operator defining the scale-discretized wavelets $L^{d}\Gamma_{\alpha^{j}}$ used for reconstruction is independent of Γ, contrary to the operator L_{Ψ} for continuous wavelets. This simply comes from the fact that the scale-discretized wavelets are, through their definition (4.71), normalized by C_{Ψ}.

Just as in the continuous wavelet formalism where the admissibility condition (4.59) is required, the present reconstruction formula holds if and only if the scale-discretized wavelet satisfies the constraints (4.68), and (4.74) or (4.76). These constraints are automatically satisfied by construction of the scale-discretized wavelets through the integration by slices. Again, this corresponds to the requirement that the wavelet family as a whole, including the scaling function, preserves the signal information at each frequency $l\in\mathbb{N}$.

Equivalently, the decomposition of the reconstruction relation (4.80) into spherical harmonic coefficients reads as a finite summation:

$$\widehat{\bar{F}}_{lm} = \widehat{[\Phi_{\alpha^J} F]}_{lm} + \frac{2l+1}{8\pi^2} \sum_{j=0}^{J} \sum_{|n|\leq \min(N-1,l)} \widehat{(\Gamma_{\alpha^j})}_{ln} \widehat{\left(\bar{W}_{\Gamma}^{F}\right)}^{l}_{mn} \left(\alpha^j\right), \quad (4.82)$$

with

$$\widehat{[\Phi_{\alpha^J} F]}_{lm} = \tilde{\Phi}_{\Gamma}^{2} \left(\alpha^J l\right) \widehat{F}_{lm}. \quad (4.83)$$

Let us emphasize the fact that a finite number of discrete dilation factors are required for the analysis and reconstruction of a band-limited signal. Contrary to the case of the continuous dilation factors, this allows the exact reconstruction of band-limited signals from relation (4.80), on pixelizations of the sphere where a sampling theorem holds.

4.5.1.5 Example Filter

We describe a real scale-discretized factorized steerable wavelet Γ with compact harmonic support, designed through relations (4.65)–(4.71) from a real continuous wavelet Ψ, for a band limit $B = 1024$, a basis dilation factor $\alpha = 2$, and an azimuthal band limit $N = 3$ of steerability. The function is imposed to be real and to be even both under rotation around itself by π and under a change of sign on φ. This corresponds to the constraints that only the $T = 3$ values $m \in \{-2,0,2\}$ are allowed and S_{lm}^{Γ} is real. We also impose that the directionality coefficients are independent of l for $l \geq 2$. This ensures that the directionality and steerability properties are preserved for the analysis depths for which the lower bound of the compact harmonic support is larger than 2. The continuous kernel is defined with a compact support in the interval $k \in (512,1024)$. The scaling function $\tilde{\Phi}_{\Gamma}(k)$ for each analysis depth j is obtained by numerical integration, with nonzero values in the intervals $k \in (512/2^j, 1024/2^j)$. The scale-discretized kernel $\tilde{K}_{\Gamma}(k)$ then follows with nonzero values in the intervals $k \in (512/2^j, 2048/2^j)$. For $j = 0$, the corresponding compact support interval is cut at the band limit: $k \in (512,1024)$. For $1 \leq j \leq 9$, the intervals progressively move to lower frequencies and shrink. At the maximum analysis depth $j = J_B(\alpha) = 10$, the compact support is shrunk to $k \in (0.5,2)$ and the scale-discretized kernel only contains the frequency $l = 1$. A specific choice of directional split and continuous kernel is made under all these constraints, as in [238]. Corresponding graphs are reported in Fig. 4.6.

Plots of the scale-discretized wavelet are also reported in Fig. 4.7. The wavelets are represented at the four largest analysis depths, $7 \leq j \leq 10$, identifying the four largest scales. At $j = 7$, $j = 8$, and $j = 9$, the compact supports of the scale-discretized kernels respectively contain the frequencies $l = 5$ to $l = 15$ with a kernel maximum at $l = 8$, $l = 3$ to $l = 7$ with a kernel maximum at $l = 4$, and $l = 2$ to $l = 3$ with a kernel maximum at $l = 2$. At $j = 10$, the scale-discretized kernel only contains the frequency $l = 1$. For the depths j with $7 \leq j \leq 9$, the lowest frequencies

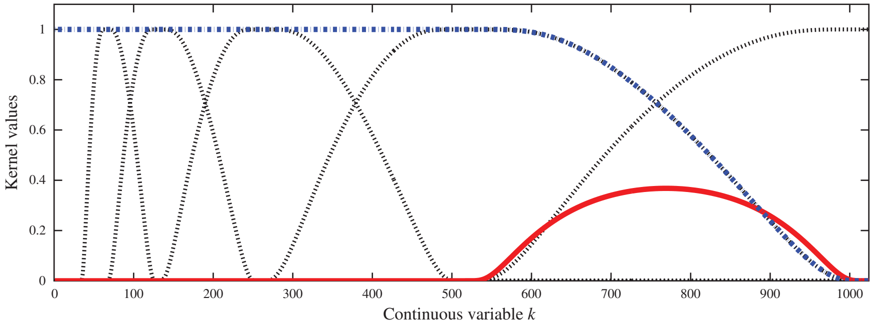


Fig. 4.6: Graphs of the continuous kernel (continuous red line), scaling function (dot-dashed blue line), and scale-discretized kernels (dotted black lines) for the example scale-discretized wavelet designed at a band limit $B = 1024$ and with a basis dilation factor $\alpha = 2$. The scale-discretized kernels are plotted for the first five analysis depths j, with $0 \leq j \leq 4$. (Figure borrowed from [238].)

l are greater than or equal to 2 and the azimuthal frequency indices contained in the directional split are $m \in \{-2, 0, 2\}$. These wavelets all have the same directionality and steerability properties. For the depth $j = 10$, the scale-discretized wavelet is a pure dipole ($l = 1$). The azimuthal frequency index is restricted to $m = 0$ due to the constraint $|m| \leq l$, ensuring that the harmonic structure on the sphere is respected, and the wavelet is simply axisymmetric.

4.5.2 Other Constructions

First, a scale-discretized wavelet formalism with relations (4.74) and (4.76) for factorized steerable wavelets with compact harmonic support can be developed by simply relying on a Littlewood–Paley decomposition, without any contact with the continuous wavelet formalism. One simply needs to choose any arbitrary scaling function satisfying relation (4.70) and define the corresponding scale-discretized kernels by differences of scaling functions at successive scales.

Such invertible filter banks based on the harmonic dilation were already developed in the case of axisymmetric wavelets [97, 216]. Our definition of factorized steerable wavelets with compact harmonic support allows a straightforward generalization to directional wavelets with the kernel dilation. Also, notice that the constraints of steerability and compact harmonic support for the scale-discretized wavelets can technically be relaxed without affecting the Littlewood–Paley decomposition. However, both properties are essential for the control of localization and directionality properties through kernel dilation. Moreover, in the absence of compact harmonic support, relation (4.74) turns into a resolution of the contracted scaling function $\tilde{\Phi}_{\Gamma}(\alpha^{-1}l)$, which differs from unity below the band limit. In other words, the filter bank developed in such a case analyzes the part of the signal

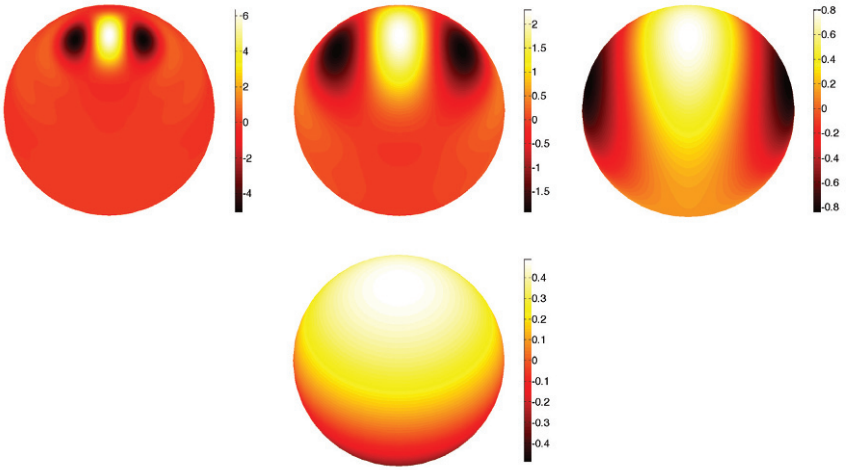


Fig. 4.7: Plots of the example scale-discretized wavelet designed with $B = 1024$, $\alpha = 2$, and $N = 3$. The wavelets are represented at the four largest analysis depths: from left to right $j = 7$, $j = 8$, and $j = 9$, and below $j = 10$. Light and dark regions respectively correspond to positive and negative values of the functions (see value bars). The wavelets are neither translated, i.e., they have their central position at the north pole, nor rotated, i.e., they are in their original orientation $\chi = 0$ (the meridian $\varphi = 0$ corresponds to a vertical line passing by the north pole). (Figure borrowed from [238].)

corresponding to its standard correlation with the contracted scaling function, rather than the signal itself. In the absence of compact harmonic support and steerability, essential multiresolution properties are also lost (see Section 4.6.1). The memory and computation time requirements of the algorithm for the analysis and reconstruction of signals therefore increase significantly and may rapidly become overwhelming.

Second, also notice that scale-discretized axisymmetric wavelets with compact harmonic support and dilated through kernel dilation were introduced under the name of *needlets* [193, 11, 178]. It is possible to show that the needlet coefficients of a wide class of random signals on the sphere are uncorrelated in the asymptotic limit of small scales, at any fixed angular distance on $\mathbb{S}^2$. The scale-discretized steerable wavelets with compact harmonic support, thanks to their factorized form and to the choice of the kernel dilation, are also good candidates for a directional extension of needlets.

Third, invertible filter banks based on the stereographic dilation have also recently been proposed [245], but they do not share these essential multiresolution properties.

Finally, frames of stereographic wavelets have been constructed in [25] by direct discretization of the translation, rotation, and dilation parameters. Notice though that the obtained frames are not tight, which means that numerical reconstruction can only be achieved by resorting to rather heavy algorithms that seek to approximate the pseudo-inverse of the forward analysis wavelet operator. By contrast, the formalism discussed here leads to an efficient and simple numerical scheme.

4.6 Reconstruction Algorithm

In this section, we identify the multiresolution properties of the scale-discretized wavelet formalism developed. We also describe a corresponding fast algorithm for the analysis and reconstruction of signals.

4.6.1 Multiresolution

We consider the analysis and reconstruction of a signal $F \in \mathrm{L}^2(\mathbb{S}^2, d\Omega)$ with a scale-discretized wavelet $\Gamma \in \mathrm{L}^2(\mathbb{S}^2, d\Omega)$, which is a factorized steerable function with compact harmonic support. The band limit and basis dilation factor of interest are denoted B and $\alpha > 1$, respectively.

The compact harmonic support of the scale-discretized wavelet Γ_{α^j} is reduced in the intervals $l \in\leq \left(\left\lfloor \alpha^{-(1+j)}B \right\rfloor, \left\lceil \alpha^{(1-j)}B \right\rceil\right)$ through the kernel dilation at each analysis depth j. As a function on SO(3), the wavelet coefficients at depth j exhibit the same compact harmonic support as the scale-discretized wavelet Γ_{α^j}. From relation (4.79), the Wigner D-transform $\widehat{(W_\Gamma^F)}^{l}_{mn}(\alpha^j)$ of the wavelet coefficients is indeed nonzero only in the same interval as the wavelet. In particular, the band limit of the wavelet coefficients is decreased to $\left\lceil \alpha^{(1-j)}B \right\rceil$ at depth j. Consequently, the number of sampled values required for the wavelet coefficients is reduced at each increase of the analysis depths j to $\alpha^{2(1-j)}\,\mathscr{O}(B^2)$ discrete points of the form $(\omega_0)_{i(j)}$ on $\mathbb{S}^2$, where $i(j)$ simply indexes these points. The number of operations required for their computation is reduced correspondingly. Hence, the kernel dilation applied to scale-discretized wavelets with compact harmonic support provides a first strong multiresolution property for the formalism. The steerability of the wavelet is also important in the algorithmic structure of the analysis, beyond its interest in preserving directionality properties through kernel dilation. At each point $(\omega_0)_{i(j)}$ and at each analysis depth j, the wavelet coefficients of a signal F with the scale-discretized wavelet Γ_{α^j} are known for all continuous rotation angles $\chi \in [0, 2\pi)$ as a linear combination of the wavelet coefficients of F with M basis wavelets, which can be taken as specific rotations Γ_{χ_p,α^j} of the wavelet Γ_{α^j}, with $0 \leq p \leq M-1$. From this perspective, steerability provides a second strong multiresolution property for the formalism.

In summary, when multiresolution properties of the formalism are fully accounted for, a reduced number of discrete points of the form $\rho_{I(j)} = ((\omega_0)_{i(j)}, \chi_p)$

on SO(3) are required for the sampled values $W_{\Gamma}^{F}(\rho_{I(j)},\alpha^{j})$ of the wavelet coefficients, where $I(j)=\{i(j),p\}$ simply indexes these points at each analysis depth j.

4.6.2 Fast Algorithm

We describe here an algorithm for both analysis and reconstruction. As for the analysis algorithms described earlier in this chapter, it is designed in harmonic space in order to take advantage of the directional correlation relation (4.79).

Some precalculations are required in order to build the wavelets Γ_{α^j} in real space, from the spherical harmonic coefficients $\widehat{\Psi}_{lm}$ of a continuous wavelet in relation (4.69). As it clearly appears in the following, the cost of these operations is negligible relative to the cost of the analysis and reconstruction themselves. Moreover, it must be performed only once for all signals to be analyzed.

The analysis may be performed by application of either the separation-of-variables algorithm or the factorization-of-rotations algorithm defined in Section 4.4, at all analysis depths j separately. The band-limited signal F is given in terms of its sampled values $F(\omega_i)$ on the $\mathcal{O}(B^2)$ discrete points ω_i of the chosen pixelization of $\mathbb{S}^2$. The resulting band-limited coefficients $W_{\Gamma}^{F}(\rho,\alpha^{j})$ at each depth j are given in terms of sampled values $W_{\Gamma}^{F}(\rho_{I(j)},\alpha^{j})$ on the $\alpha^{2(1-j)}\,\mathcal{O}(MB^2)$ discrete points $\rho_{I(j)}$ of the chosen pixelization of SO(3). The exactness of the computation relies only on the exactness of the computation of the spherical harmonic coefficients of the signal, which depends on the existence of a sampling theorem on the initial pixelization on which the original signal was sampled. The a priori computational complexity for the directional correlation (4.77) by quadrature is of order $\alpha^{4(1-j)}\times\mathcal{O}(MB^4)$ at each analysis depth j. This cost is reduced to $\alpha^{3(1-j)}\mathcal{O}(MB^3)$ with the steerable optimization of the fast analysis algorithms.

The reconstruction part of the algorithm proceeds through the exact same operations as the analysis, in reverse order. The Wigner D-function coefficients $\widehat{(W_{\Gamma}^{F})}_{mn}^{l}(\alpha^{j})$ of the wavelet coefficients are computed by quadrature through a direct Wigner D-function transform at each analysis depth j. The spherical harmonic coefficients of the reconstructed signal $\widehat{F}_{lm}$ are then obtained as a finite summation following from relations (4.82) and (4.83). The samples $F(\omega_i)$ of the signal are finally recovered by a simple inverse spherical harmonic transform. The reconstruction is symmetric to the analysis and therefore requires the same number of operations.

The total computational complexity of the algorithm is obtained by summing over all analysis depths j with $0\leq j\leq J$. In the most exacting case where $J=J_B(\alpha)$, it simply reads as

$$C(\alpha,B,M)=\left[1+c\left(\alpha^{3}\right)\left(1-\alpha^{-3J_B(\alpha)}\right)\right]\mathcal{O}(MB^{3}). \tag{4.84}$$

In this expression, the impact of the compact harmonic support of the scale-discretized wavelet is concentrated in $c(\alpha^3)=\alpha^3/(\alpha^3-1)\in[1,\infty)$. For example, a complete decomposition at a high band limit with a dyadic decomposition of the scales $\alpha = 2$ gives $C \simeq 15/7 \times \mathscr{O}(MB^3)$. Obviously, the more compact the support of the scale-discretized wavelet, the larger the computational complexity.

Let us finally comment on the exactness of the analysis and reconstruction algorithm described. If pixelizations of $\mathbb{S}^2$ and SO(3) are chosen where a sampling theorem holds, exact quadrature rules hold both for the direct spherical harmonic transform of a band-limited signal in the analysis part, and for the direct Wigner D-function transform of its wavelet coefficients in the reconstruction part. Such exact quadrature is accessible on equiangular or Gauss–Legendre pixelizations on $\mathbb{S}^2$ for sampling of the original signal on points ω_i, and for sampling the wavelet coefficients at each analysis depth j and for each value χ_p on points $(\omega_0)_{i(j)}$. In this context, if the computed wavelet coefficients are not altered before reconstruction, the exact same samples are obtained after reconstruction as for the original signal F. Again, HEALPix pixelizations provide nonexact but very precise quadrature rules.

Details on the algorithmic structure, computation times, memory requirements, and numerical stability of the corresponding implementation may be found in [238].[3]

4.7 Applications

In this section, we illustrate the usefulness of the wavelet formalisms developed for analysis and reconstruction of signals in the context of applications in astrophysics and neuroscience.

4.7.1 Cosmic Microwave Background Analysis

The aim of cosmology is the study of the structure and evolution of the universe. The last decades have led us to the dawn of a new era of precision cosmology, characterized by access to more and more precise observations of our universe. One of our best laboratories is the cosmic microwave background radiation (CMB).

The CMB is a relic black-body radiation, a unique realization of a random process that occurred in the early universe. The radiation is observed in all directions of the celestial sphere. The corresponding data crystallize various forms of complexities. The data are distributed on the surface of the celestial sphere. Current and forthcoming experiments give access to high angular resolutions on the celestial sphere, and therefore large volumes of data. The radiation is described not only by a scalar temperature field but also by tensor polarization parameters. The observed CMB signal is inevitably contaminated by galactic and extragalactic foreground emissions that allow only partial sky coverage. Moreover, data are inevitably

[3] Code for Steerable and Scale-Discretized Wavelets on the sphere (S2DW) is available for download at the following URL: `http://www.mrao.cam.ac.uk/~jdm57/software.html`

contaminated by instrumental noise, blurred by an experimental beam, and affected by other sources of systematic errors.

In this context, the analysis of CMB data is of major importance to obtain a precise picture of the universe through the definition of a cosmological model that best fits the data. Wavelets on the sphere appeared in recent years as an essential tool for CMB data analysis. Corresponding analyses of the Wilkinson Microwave Anisotropy Probe satellite mission (WMAP) data have allowed the probe of the fundamental pillars of our models. Gaussianity and statistical isotropy of the random process from which the CMB arises have been studied and corresponding anomalies were found in the data, thereby confirming and synthesizing other analyses. The presence of a very poorly understood form of energy in the universe, called *dark energy*, was also assessed and confirmed independently of other probes through wavelet analyses. These analyses were based on the decomposition of the CMB temperature signal on the sphere with continuous axisymmetric, directional, and steerable wavelets. Statistical analyses of the corresponding wavelet coefficients led to the physical conclusions. The first of these analyses is reviewed in [185]. More recent analyses are specifically based on steerable wavelets [240, 230, 189, 239].

To quote only one of these applications, steerable wavelets have allowed the identification of anomalously preferred directions in the CMB data, through the following process. The steerability gives access to morphological measures of local features of the signal at a given analysis scale a and around each point ω_0. The orientation $\chi^*(a,\omega_0)$ of a local feature may notably be defined as the direction in which the signal response to the wavelet has maximum absolute value. This orientation may simply be obtained from relation (4.32). The corresponding absolute value of the signal identifies the intensity of the feature. The alignment of local CMB features toward specific directions on the celestial sphere can then be probed by combining the information on the orientation and on the intensity of local features. First, the great circle is defined, which passes by the point ω_0 and admits as a tangent the local direction defined by $\chi^*(a,\omega_0)$. All directions on that great circle are considered to be seen by the local feature at ω_0 with a weight naturally given as the intensity at that point. At scale a, the degree of preference of each direction ω is defined as the sum of the weights originating from all points in the original signal for which the defined great circle crosses the direction considered. Fig. 4.8 represents the result of this alignment analysis with a second Gaussian derivative wavelet at a typical scale of 10° of angular opening, as performed on WMAP data in comparison with simulations based on an assumption of isotropy of the universe. Further analysis of this result allowed the identification of a mean preferred plane with a normal direction close to the CMB dipole axis, and a mean preferred direction in this plane, very close to the ecliptic poles' axis [240, 230].

4.7.2 Human Cortex Image Denoising

Subtle changes in human cortical thickness are thought to be associated with neurological or clinical deficiencies. It is therefore important to detect these changes

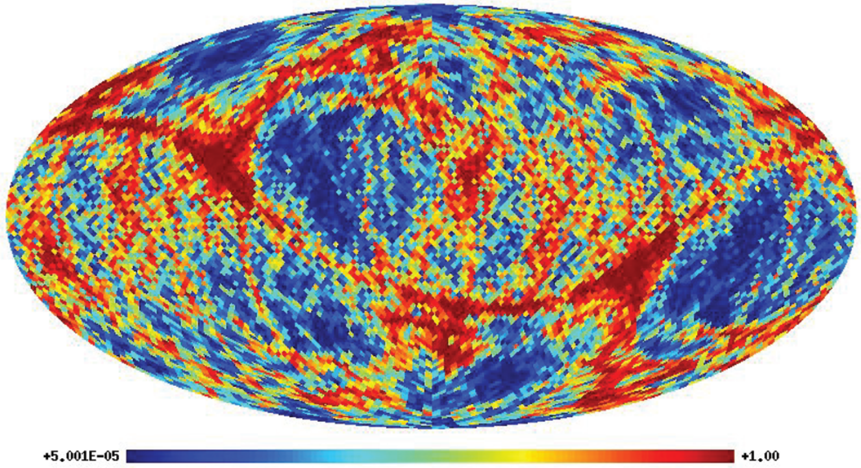


Fig. 4.8: Cumulative probabilities map in Mollweide projection of a HEALPix pixelization ($N_{\text{side}} = 32$), for the degree of preference of each direction on the celestial sphere as obtained by an alignment analysis of the CMB signal using continuous steerable wavelets. The observed pattern presents several great circles of anomalously high (red) and low (blue) preference with a value well higher, or lower, than the median value of the simulations. (Figure borrowed from [230].)

using bioimaging modalities. Typically, cortical thickness maps are inferred over the brain surface through magnetic resonance imaging (MRI) examinations. Cortical thickness features are extracted from these maps and studied using statistical tests. The acquisition process is, however, unavoidably affected by noise. As discussed in [21], the outcome of the statistical tests is known to be greatly improved by spatial smoothing, specifically with low-pass Gaussian filters, which enhances the signal-to-noise ratio. Since the cortex is a very convoluted surface, there is no easy way directly to apply simple low-pass filtering. One successful method, illustrated in Fig. 4.9, is to first map the cortex to a sphere where scalar data can be analyzed in a simple way.

In this context, scale-discretized wavelets on the sphere offer a very flexible and computationally efficient way to denoise data while preserving the most salient features and most important spatial variations. Denoising with wavelets obviously require a discrete formalism where the signal may be reconstructed after modification of its wavelet coefficients. In [21], the spherical cortical thickness map is processed using scale-discretized axisymmetric wavelets on the sphere such as those discussed in Section 4.5. The wavelet coefficients are then thresholded by soft thresholding [69] in order to remove the noise, which is assumed to be uniformly distributed over the coefficients, while the most important spatial variations of the original data are encoded in the strongest wavelet coefficients only. This study shows

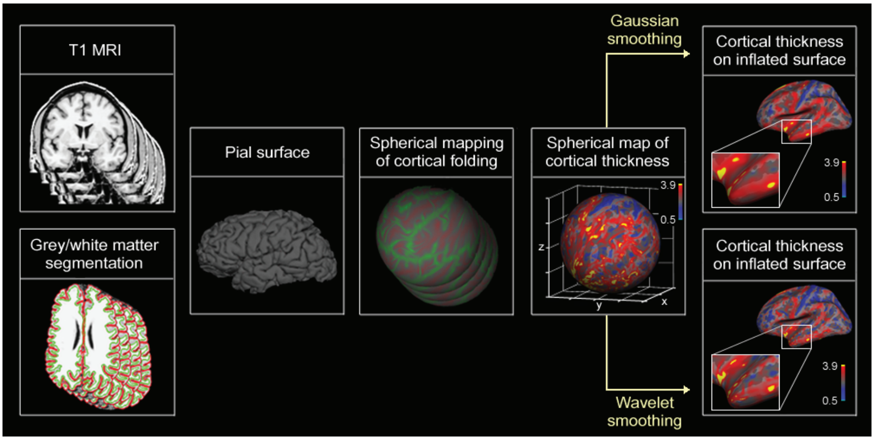


Fig. 4.9: A cortical thickness map is measured from T1 MRI sequences. The scalar map is then inflated to a spherical surface before being processed. (Figure reproduced from [21], with the kind permission of the authors.)

that wavelet denoising yields a significant improvement over spatial smoothing by Gaussian filtering, as illustrated in Fig. 4.10.

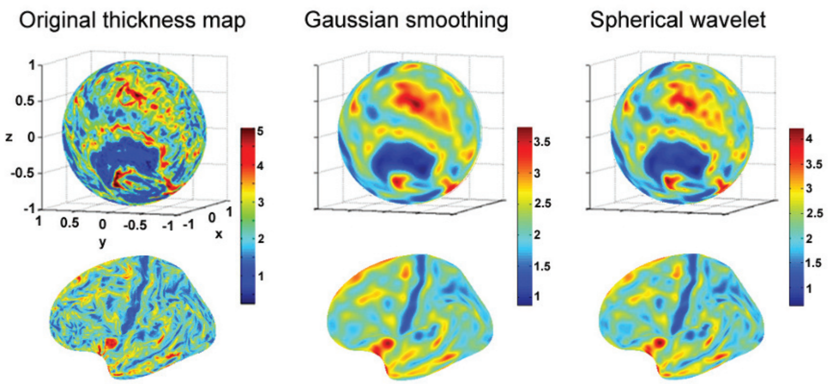


Fig. 4.10: Comparison of processing of a cortical thickness map by spatial smoothing using a Gaussian filter or by denoising using soft thresholding of wavelet coefficients. Wavelets allow for better reconstruction of morphological features. (Figure reproduced from [21], with the kind permission of the authors.)

4.8 Conclusion

The wavelet transform has grown into a mature tool for data analysis and has enjoyed large success in an amazingly broad spectrum of applications. In this chapter, we have reviewed several constructions that aim at extending wavelets to data defined in spherical geometry, which is a desirable generalization for many practical problems.

Wavelets on the sphere implement multiresolution through specific dilation mechanisms, but the concept of directional correlation provides a flexible framework that unifies the various families of wavelets discussed here. We have seen that the continuous wavelet formalism on the sphere is a powerful analysis tool, while the scale-discretized wavelet formalism offers a full-fledged framework for digital data processing on the sphere, allowing for reconstruction from wavelet coefficients. Finally, this theoretical formalism is also backed up by computationally efficient algorithms. Virtually all applications involving wavelets can now be generalized to data on the sphere. In particular, solving inverse problems in spherical geometry, notably denoising and deconvolution problems, by imposing some sort of sparsity of wavelet coefficients is of significant practical interest in application fields ranging from astrophysics to neuroscience.

Exercises

1. Prove the Plancherel relation (4.8) from the orthonormality of the spherical harmonics.
2. Prove that the stereographic dilation D_a on points in (4.18) is the unique radial conformal diffeomorphism on $\mathbb{S}^2$.
3. Prove that the expression $\lambda^{1/2}(a,\theta) = a^{-1}[1+\tan^2(\theta/2)]/[1+a^{-2}\tan^2(\theta/2)]$ is required to ensure unitarity of the dilation $D(a)$ of functions in $\mathrm{L}^2(\mathbb{S}^2, d\Omega)$.
4. Prove the pointwise product form (4.28) for the Wigner D-function transform of the directional correlation defining wavelet coefficients of a signal on $\mathbb{S}^2$.
5. Prove that the admissibility condition (4.30) is a necessary and sufficient condition for the reconstruction formula (4.29).
6. Prove that the stereographic projection π in (4.44) is the unique radial conformal diffeomorphism mapping the sphere $\mathbb{S}^2$ onto the plane $\mathbb{R}^2$.
7. Prove that the prefactors in (4.44) are required to ensure the unitarity of the projection operator Π between $\mathrm{L}^2(\mathbb{R}^2, d^2\mathbf{x})$ and $\mathrm{L}^2(\mathbb{S}^2, d\Omega)$.
8. Prove the relation (4.62) for the factorization of rotations.
9. Prove that the admissibility condition (4.59) for continuous wavelets simply turns into the resolution of the identity (4.74) after scale discretization.
10. Prove the expression (4.75) for the maximum analysis depth $J_B(\alpha)$.

Chapter 5
Wiener's Lemma: Theme and Variations. An Introduction to Spectral Invariance and Its Applications

Karlheinz Gröchenig

Abstract Wiener's Lemma is a classical statement about absolutely convergent Fourier series and remains one of the driving forces in the development of Banach algebra theory. In the first part of the chapter—the theme—we discuss Wiener's Lemma in detail. We prove Wiener's Lemma and discuss equivalent formulations about convolution operators. We then extract the underlying abstract concepts from Banach algebras. In the second part of the chapter—the variations—we discuss several, mostly noncommutative reincarnations of Wiener's Lemma. We will develop some of the theoretical background and explain why Wiener's Lemma is still useful and inspiring. The topics cover weighted versions of Wiener's Lemma, infinite matrix algebras, noncommutative tori and time-frequency analysis, convolution operators on noncommutative groups, and time-varying systems and pseudodifferential operators.

5.1 Introduction

Wiener's Lemma is a classical and seemingly innocent result about absolutely convergent Fourier series. In its original version it asserts that the pointwise inverse of an absolutely convergent Fourier series without zeros is again an absolutely convergent Fourier series. This result is contained in every text about Fourier series and in every treatment of commutative Banach algebras. Norbert Wiener needed this lemma for his proof of a "Tauberian theorem" [241, 242].

But Wiener's Lemma is much more: Wiener's Lemma is central in the development of the abstract theory of Banach algebras and has inspired Gel'fand's theory of commutative Banach algebras [100, 101]. The generalizations of Wiener's Lemma are now legion, and this chapter bears witness to some recent developments.

Karlheinz Gröchenig
Faculty of Mathematics, University of Vienna, Nordbergstrasse 15, A-1090 Vienna, Austria,
e-mail: karlheinz.groechenig@univie.ac.at

B. Forster and P. Massopust (eds.), *Four Short Courses on Harmonic Analysis*, Applied and Numerical Harmonic Analysis, DOI 10.1007/978-0-8176-4891-6_5,

Why is Wiener's Lemma so important?

Wiener's Lemma is a deep result about the invertibility and spectrum of certain operators.

For the direct verification that the function $1/f$ possesses an absolutely convergent Fourier series, we would have to calculate or estimate the Fourier coefficients of $1/f$ and check that they are summable. Wiener's Lemma offers a much easier test. We only need to make sure that f does not have any zeros; this means that $1/f$ exists as a continuous function. This is the heart of Wiener's Lemma: A difficult condition for the invertibility can be replaced by an easier, more evident, and more convenient condition.

Clearly, the understanding of the invertibility is of tremendous importance for solving systems of linear equations or operator equations. Indeed, parallel to our journey through the manifold aspects of Wiener's Lemma, we will discuss its relevance in signal analysis. In engineering Wiener's Lemma plays a vital role in the analysis of signal transmission, time-invariant and time-varying channels, and for the signal recovery in wireless communications.

The realm of Wiener's Lemma is a vector space of functions that can be multiplied by each other. Adding a norm and completeness, the mathematical structure in the background is that of a Banach algebra. Thus, from an abstract point of view Wiener's Lemma is about invertibility in a Banach algebra and provides an easy criterion. A function f is invertible as an absolutely convergent Fourier series if and only if it is invertible as a continuous function.

The invertibility of a function or an operator is closely related to its spectrum, and we will see how Wiener's Lemma can be formulated as a result about spectral invariance. The spectrum of an absolutely convergent Fourier series f is independent of the underlying Banach algebra. The concept of spectral invariance is fundamental in many fields. One of our objectives is a unified treatment of spectral invariance in several areas of mathematics and to trace back several fundamental results on spectral invariance to Wiener's Lemma.

Following the title, this chapter is divided into two parts, the theme and its variations.

Section 5.2 discusses the classical version of Wiener's Lemma under different angles. We introduce absolutely convergent Fourier series and then elaborate an elementary proof of Wiener's Lemma. This proof is void of abstract theory and requires only elementary estimates from analysis. Only then do we provide the structural background and interpret Wiener's Lemma in the context of Banach algebra theory. In Section 5.2.5 we discuss the main concepts of spectral invariance and the immediate consequences. Finally, we convert Wiener's Lemma to a statement about the spectrum of convolution operators.

The material is fundamental for an appreciation of the variations, because the sequence of definitions and results sets the pattern for the treatment of the variations, which are usually considered parts of different fields of mathematics.

Section 5.2 is elementary and requires only basic analysis. We do not even need the Gel'fand theory of commutative Banach algebras, although some knowledge of Banach algebras may be helpful to appreciate the context and to open more perspectives on Wiener's Lemma.

Section 5.3 is devoted to variations of Wiener's Lemma in several areas of mathematics, namely in Fourier analysis, infinite matrix algebras and operator theory, in the noncommutative geometry of tori and time-frequency analysis, in the harmonic analysis of locally compact groups, and in the theory of pseudodifferential operators.

In each of the variations we will set up the basic concepts, and then draw the analogy between Fourier series and the new object, series of time-frequency shifts, say. The analogy to Fourier series usually suggests some natural questions; in particular, we will always wonder whether a version of Wiener's Lemma holds. We then will proceed to formulate answers that are similar to the classical Wiener's Lemma. The proofs are usually much more advanced, and some proofs are outside the scope of this introduction. Wiener's Lemma is often hidden in the proofs, and we will try to make the connections visible. Section 5.3 can no longer be self-contained, because the mathematical concepts are drawn from disjoint mathematical worlds. Our goal is to reveal the common structure and convince the reader that the topics treated are indeed variations of a classical theme, albeit the variations may be highly nontrivial.

In Section 5.3.1 we introduce weights and investigate weighted versions of Wiener's Lemma for Fourier series. We discuss the dichotomy between subexponential and exponential weights and define the Gel'fand–Raikov–Shilov condition. This is a new concept that arises invariably in spectral problems with weights.

Section 5.3.2 deals with the spectral invariance of matrices with some form of off-diagonal decay and is a first version of a noncommutative Wiener's Lemma. In Section 5.3.3 we turn to time-frequency analysis and investigate series of time-frequency shifts instead of Fourier series. The next section, 5.3.4, treats convolution operators on general locally compact groups. Here we have to content ourselves with explaining the concepts of harmonic analysis, stating the results, and discussing their meaning. Section 5.3.5 is devoted to pseudodifferential operators and their invertibility. We discuss a class of symbols (the so-called Sjöstrand class) that resembles absolutely convergent Fourier series and formulate the correct generalization of Wiener's Lemma for pseudodifferential operators.

In the last section, 5.3.6, we explain how and why the results on pseudodifferential operators can be used for the analysis of time-varying systems and in wireless communications.

5.2 Wiener's Lemma—Classical

Let us motivate Wiener's Lemma with a familiar statement from calculus. Let $C^k(\mathbb{T})$ be the space of k-times differentiable functions with period 1. We identify the interval $[0,1)$ of a period with the torus $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$ when necessary.

The product rule $(fg)' = f'g + fg'$ implies that $C^k(\mathbb{T})$ is an *algebra*. The quotient rule for differentiation implies the following property of $C^k(\mathbb{T})$.

Lemma 5.1. *If $f \in C^k(\mathbb{T})$ and $f(t) \neq 0$ for all $t \in [0,1)$, then $1/f \in C^k(\mathbb{T})$.*

Proof. Since $(1/f)' = -f'/f^2$, $1/f$ is continuously differentiable whenever $f(t) \neq 0$. Now proceed by induction. Assume that we already know that $1/f \in C^\ell(\mathbb{T})$ for $\ell < k$. Then $(1/f)' = -f'/f^2$ is continuous and in $C^\ell(\mathbb{T})$ by the induction hypothesis. Therefore $1/f \in C^{\ell+1}(\mathbb{T})$. □

5.2.1 Definitions from Banach Algebras

The quotient rule and Lemma 5.1 are about the invertibility of differentiable functions. The abstract discussion of invertibility is best carried out in the context of Banach algebras. To begin with, let us recall the standard definitions.

Definition 5.2. A Banach space $\mathscr{A}$ is called a *Banach algebra* if it possesses a multiplication $\mathscr{A} \times \mathscr{A} \to \mathscr{A}$ that satisfies the following properties for all $a,b,c \in \mathscr{A}$ and $\lambda \in \mathbb{C}$:

1. $(a+b)c = ac + bc$ and $a(b+c) = ab + ac$;
2. $(ab)c = a(bc)$;
3. $(\lambda a)b = a(\lambda b) = \lambda(ab)$;
4. $\|ab\| \leq \|a\| \, \|b\|$.

We always assume that $\mathscr{A}$ possesses a unit element e that satisfies $ae = ea = a$ for all $a \in \mathscr{A}$. In a unital algebra an element a is called *invertible* if there exists an element $b \in \mathscr{A}$ such that $ab = ba = e$. If such a b exists, it is unique and called the *inverse* of a and denoted by a^{-1}.

If we endow $C^k(\mathbb{T})$ with the norm $\|f\|_{C^k} = \sum_{j=0}^k \frac{1}{j!}\|f^{(j)}\|_\infty$, then $C^k(\mathbb{T})$ becomes a Banach algebra with respect to pointwise multiplication (see Exercises). Lemma 5.1 provides a simple and in this case rather obvious condition for the invertibility of an element in the algebra $C^k(\mathbb{T})$.

5.2.2 Absolutely Convergent Fourier Series

Let us now introduce $\mathscr{A}(\mathbb{T})$, the main object of Wiener's Lemma.

Definition 5.3. A periodic function f possesses an absolutely convergent Fourier series if it can be written as $f(t) = \sum_{k\in\mathbb{Z}} a_k e^{2\pi i k t}$ with coefficients in $\mathbf{a} \in \ell^1(\mathbb{Z})$. In this case we write $f \in \mathscr{A}(\mathbb{T})$ and endow $\mathscr{A}(\mathbb{T})$ with the norm

$$\|f\|_{\mathscr{A}} = \|\mathbf{a}\|_1 = \sum_{k\in\mathbb{Z}} |a_k| \,.$$

It is not completely obvious that $\|\cdot\|_{\mathscr{A}}$ is a norm. This fact follows from the uniqueness of the Fourier coefficients.

Lemma 5.4. *The space $\mathscr{A}(\mathbb{T})$ is a Banach algebra under pointwise multiplication.*

Proof. Let $f(t) = \sum_{k\in\mathbb{Z}} a_k e^{2\pi ikt}$ and $g(t) = \sum_{k\in\mathbb{Z}} b_k e^{2\pi ikt}$ with norms $\|f\|_{\mathscr{A}} = \|\mathbf{a}\|_1$ and $\|g\|_{\mathscr{A}} = \|\mathbf{b}\|_1$. Then

$$\begin{aligned} f(t)g(t) &= \left(\sum_{k\in\mathbb{Z}} a_k e^{2\pi ikt}\right)\left(\sum_{l\in\mathbb{Z}} b_l e^{2\pi ilt}\right) \\ &= \sum_{k,l\in\mathbb{Z}} a_k b_l e^{2\pi i(k+l)t} \\ &= \sum_{n\in\mathbb{Z}} \Bigg(\underbrace{\sum_{k\in\mathbb{Z}} a_k b_{n-k}}_{(\mathbf{a}*\mathbf{b})(n)}\Bigg) e^{2\pi int}. \end{aligned} \tag{5.1}$$

The interchange of the summation is justified because both series converge absolutely. Thus, the coefficients of the pointwise product fg are given by the convolution of the sequences $\mathbf{a}$ and $\mathbf{b}$. Now

$$\begin{aligned} \|\mathbf{a}*\mathbf{b}\|_1 &= \sum_{n\in\mathbb{Z}} \Big|\sum_{k\in\mathbb{Z}} a_k b_{n-k}\Big| \\ &\le \sum_{k\in\mathbb{Z}}\sum_{n\in\mathbb{Z}} |a_k|\,|b_{n-k}| = \|\mathbf{a}\|_1\,\|\mathbf{b}\|_1, \end{aligned}$$

and consequently,

$$\|fg\|_{\mathscr{A}} = \|\mathbf{a}*\mathbf{b}\|_1 \le \|\mathbf{a}\|_1\,\|\mathbf{b}\|_1 = \|f\|_{\mathscr{A}}\,\|g\|_{\mathscr{A}}.$$

The other properties of Definition 5.2 are obvious, and thus $\mathscr{A}(\mathbb{T})$ is a Banach algebra. □

In Section 5.3 we will encounter several variations of this proof in rather different contexts.

5.2.3 Wiener's Lemma

As with the algebra $C^k(\mathbb{T})$, we may now investigate the invertibility for absolutely convergent Fourier series. The formulation of Lemma 5.1 suggests the following question: If $f \in \mathscr{A}(\mathbb{T})$ and $f(t) \neq 0$ for all $t \in [0,1)$, is f then invertible in $\mathscr{A}(\mathbb{T})$? In the absence of a quotient rule the answer is by no means obvious. It is given by the following theorem, which, with historical understatement, is now called Wiener's Lemma.

Theorem 5.5 (Classical formulation). *If $f \in \mathscr{A}(\mathbb{T})$ and $f(t) \neq 0$ for all $t \in \mathbb{T}$, then also $1/f \in \mathscr{A}(\mathbb{T})$, i.e., $1/f(t) = \sum_{k\in\mathbb{Z}} b_k e^{2\pi ikt}$ with $\mathbf{b} \in \ell^1(\mathbb{Z})$.*

Wiener's original proof [241, 242] from 1932 uses a localization property and a partition-of-unit argument. Today's standard proof is an abstract proof via Gel'fand theory [56, 150, 151, 209].

Wiener's Lemma has intrigued many analysts and is the seed for the abstract theory of Banach algebras by Gel'fand. In fact, Gel'fand [100, 101] developed his theory of commutative Banach algebras for the specific purpose of finding a conceptual proof of Wiener's Lemma. Variations of Wiener's lemma and its proof were found by Levy [165] and Zygmund [248]. An interesting proof of Wiener's Lemma without the use of Fourier series was given by Hulanicki [142]. A very short, elementary proof was found by Newman [194].

Quantitative aspects of Wiener's Lemma, namely norm estimates for the inverse f^{-1}, were investigated by Nikolski [195] and Tao [222].

5.2.4 Proof of Wiener's Lemma

Here we give an elementary proof of Wiener's Lemma following Newman [194] and Hulanicki [142]. It is of interest in its own right because it does not use Fourier series and avoids the abstract notions of Gel'fand theory.

Step 1. Reduction to special case. If $f \in \mathscr{A}(\mathbb{T})$, then also $\overline{f} \in \mathscr{A}(\mathbb{T})$ and $|f|^2 = f \cdot \overline{f} \in \mathscr{A}(\mathbb{T})$. Since $1/f = \overline{f}/|f|^2$, it suffices to show that $1/|f|^2 \in \mathscr{A}(\mathbb{T})$. By replacing f by $|f|^2$, we may assume without loss of generality that f is nonnegative. By normalizing, we may further assume that $0 \le f(t) \le 1$ for $t \in \mathbb{T}$.

Now note that since f is continuous, the assumption $f(t) \neq 0$ for all t implies that

$$\inf_{t \in \mathbb{T}} |f(t)| = \delta > 0\,. \tag{5.2}$$

Step 2. Analyze the invertibility of f in $C(\mathbb{T})$ by a geometric series. Let $h = 1 - f$; then

$$0 \le h(t) = 1 - f(t) \le 1 - \delta\,.$$

Hence, the geometric series $\sum_{n=0}^{\infty} h(t)^n$ converges in $C(\mathbb{T})$ and possesses the limit

$$\sum_{n=0}^{\infty} h(t)^n = \frac{1}{1-h(t)} = \frac{1}{f(t)} \in C(\mathbb{T})\,.$$

Our goal is to show that $\sum h^n$ also converges in $\mathscr{A}(\mathbb{T})$.

Step 3. Approximate h by a trigonometric polynomial. Given $\varepsilon > 0$, choose a trigonometric polynomial $p(t)$ such that

$$\|h - p\|_{\mathscr{A}} < \varepsilon\,.$$

If fact, if $f(t) = \sum_{k \in \mathbb{Z}} a_k e^{2\pi i k t}$, we may choose $p(t) = 1 - \sum_{|k| \le N} a_k e^{2\pi i k t}$ for sufficiently large N and obtain $\|h - p\|_{\mathscr{A}} = \sum_{|k|>N} |a_k| < \varepsilon$.

Set $r = h - p$; then $h = p + r$ and $\|r\|_{\mathscr{A}} < \varepsilon$.

As to the choice of ε, we will see that we must have $1-\delta+2\varepsilon<1$, where δ is given by (5.2).

Step 4. Some elementary estimates. First, if $q(t)=\sum_{|k|\le N} b_k e^{2\pi i k t}$ is a trigonometric polynomial of degree N, then

$$\begin{aligned}\|q\|_{\mathscr{A}} &= \sum_{|k|\le N} |b_k| \le \|b\|_2\,(2N+1)^{1/2} \\ &= \|q\|_2\,(2N+1)^{1/2} \le \|q\|_\infty\,(2N+1)^{1/2}\,.\end{aligned}$$

Second, if q is trigonometric polynomial of degree N, then q^k is a trigonometric polynomial of degree kN.

Step 5. Estimate the $\mathscr{A}$-norm of h^n. By the binomial theorem we have

$$h^n=\sum_{k=0}^{n}\binom{n}{k}p^k r^{n-k},$$

so we may estimate

$$\begin{aligned}\|h^n\|_{\mathscr{A}} &\le \sum_{k=0}^{n}\binom{n}{k}\|p^k r^{n-k}\|_{\mathscr{A}} \\ &\le \sum_{k=0}^{n}\binom{n}{k}\|p^k\|_{\mathscr{A}}\|r^{n-k}\|_{\mathscr{A}}.\end{aligned}$$

Now by our choice of p in Step 3, we have $\|r^{n-k}\|_{\mathscr{A}}\le \|r\|_{\mathscr{A}}^{n-k}<\varepsilon^{n-k}$. By Step 4 applied to the trigonometric polynomials p^k, we have

$$\|p^k\|_{\mathscr{A}}\le \|p^k\|_\infty\,(2Nk+1)^{1/2}\le (2Nn+1)^{1/2}\|p\|_\infty^k.$$

Step 6. Complete the estimate for the $\mathscr{A}$-norm of h^n.

$$\begin{aligned}\|h^n\|_{\mathscr{A}} &\le (2Nn+1)^{1/2}\sum_{k=0}^{n}\binom{n}{k}\varepsilon^{n-k}\|p\|_\infty^k \\ &= (2Nn+1)^{1/2}\,(\|p\|_\infty+\varepsilon)^n \le (2Nn+1)^{1/2}\,(\|h-r\|_\infty+\varepsilon)^n \\ &\le (2Nn+1)^{1/2}\,(\|h\|_\infty+2\varepsilon)^n \le (2Nn+1)^{1/2}(\underbrace{1-\delta+2\varepsilon}_{<1})^n.\end{aligned} \tag{5.3}$$

Step 7. Convergence of geometric series in $\mathscr{A}$-norm. Using the estimate from the previous step, we finally obtain that

$$\sum_{n=0}^{\infty}\|h^n\|_{\mathscr{A}}\le \sum_{n=0}^{\infty}(2Nn+1)^{1/2}\,(1-\delta+2\varepsilon)^n<\infty. \tag{5.4}$$

Thus, the geometric series $\sum_{n=0}^{\infty} h^n$ converges in $\mathscr{A}(\mathbb{T})$, and we have proved that $1/f=\sum_{n=0}^{\infty} h^n\in\mathscr{A}(\mathbb{T})$. $\square$

Remark 5.6. Let us take nth roots in (5.3). Then we obtain

$$\lim_{n\to\infty} \|h^n\|_{\mathscr{A}}^{1/n} \le \|h\|_\infty + 2\varepsilon .$$

Since $\varepsilon > 0$ was arbitrary, this implies that

$$\lim_{n\to\infty} \|h^n\|_{\mathscr{A}}^{1/n} \le \|h\|_\infty = \lim_{n\to\infty} \|h^n\|_\infty^{1/n} .$$

Since $\|h^n\|_\infty \le \|h^n\|_{\mathscr{A}}$ always holds, the estimate in (5.3) implies that

$$\lim_{n\to\infty} \|h^n\|_{\mathscr{A}}^{1/n} = \lim_{n\to\infty} \|h^n\|_\infty^{1/n} . \tag{5.5}$$

This identity can be interpreted as a statement about spectral radii. It is fundamental for the generalizations and abstract versions of Wiener's Lemma. See Proposition 5.11.

Remark 5.7. The proof above is certainly not the simplest proof and lacks the elegance of Gel'fand theory. However, in view of the many variations and generalizations of Wiener's Lemma, it is important to understand which properties of $\mathscr{A}(\mathbb{T})$ come into play. First, we have studied the problem on the dense subspace of trigonometric polynomials; second, we have compared several norms, namely the L^2-norm, the L^∞-norm, and the $\mathscr{A}$-norm. The $\mathscr{A}$-norm is rather tricky, because it is defined indirectly via the Fourier coefficients. The comparison of $\|\cdot\|_{\mathscr{A}}$ with more accessible norms is therefore natural. Last but not least, we used that $\mathscr{A}(\mathbb{T})$ is commutative, when we applied the binomial theorem in Step 5.

Remark 5.8. The final estimate (5.4) of the proof leads to an estimate for the norm of $1/f$ in $\mathscr{A}(\mathbb{T})$. The norm $\|1/f\|_{\mathscr{A}}$ depends on δ, on ε, and on N, which in turn is a function of ε. It can be shown that there is no control of $\|1/f\|_{\mathscr{A}}$ in terms of δ alone [195]. The problem of norm-controlled inversion in Banach algebras is rather difficult, and in general little can be said. We refer to the beautiful work of Nikolski [195] and an essay by Tao [222].

5.2.5 Abstract Concepts—Inverse-Closedness

Following Naimark, let us now take a very abstract look at Wiener's Lemma. Naimark [191, 192] turned Wiener's Lemma into a definition. This is not a cheap trick (to avoid a proof), but Naimark's procedure conveys an important insight into Wiener's Lemma. Naimark understood that Wiener's Lemma is a result about the relationship between **two** Banach algebras, namely, the algebras $\mathscr{A}(\mathbb{T})$ and $C(\mathbb{T})$. In particular, the condition "$f(t) \neq 0, \forall t \in \mathbb{T}$" occurring in Theorem 5.5 simply means that f is invertible in $C(\mathbb{T})$. This observation justifies the following definition.

Definition 5.9. Let $\mathscr{A} \subseteq \mathscr{B}$ be two Banach algebras with a common identity. Then $\mathscr{A}$ is called *inverse-closed* in $\mathscr{B}$ if

$$a \in \mathscr{A} \text{ and } a^{-1} \in \mathscr{B} \quad \Longrightarrow \quad a^{-1} \in \mathscr{A} .$$

In other words, $\mathscr{A}$ is inverse-closed in $\mathscr{B}$ if an element a in the smaller algebra that is invertible in the bigger algebra is automatically invertible in the smaller algebra. Or to put it differently, an element $a \in \mathscr{A}$ is invertible in $\mathscr{A}$ if and only if a is invertible in $\mathscr{B}$.

The inverse-closedness is often extremely useful for the study of invertibility. The large algebra $\mathscr{B}$ contains more invertible elements; it may therefore be easier to check the invertibility of an element. If we start with $a \in \mathscr{A}$, then a^{-1} is automatically in $\mathscr{A}$. The direct verification that a is invertible in $\mathscr{A}$ may be much more difficult.

In the light of Definition 5.9, Wiener's Lemma states that the algebra of absolutely convergent Fourier series $\mathscr{A}(\mathbb{T})$ is inverse-closed in the algebra $C(\mathbb{T})$. For most mathematicians it is easier and more natural to see that f does not have any zeros on the interval $[0,1)$. The direct verification of $1/f \in \mathscr{A}(\mathbb{T})$ would require finding the Fourier coefficients of $1/f$ and checking whether they are absolutely summable.

Inverse-closedness is an important concept in many area of mathematics where Banach algebra arguments and spectra of operators play a role. Each area uses its own terms, and so there is a Babylonian confusion in the terminology. We follow Barnes [13]. Naimark uses the term *Wiener pair* $(\mathscr{A},\mathscr{B})$ when $\mathscr{A}$ is inverse-closed in $\mathscr{B}$. Palmer [196] says that $\mathscr{A}$ is a *spectral subalgebra* of $\mathscr{B}$. In K-theory one says that $\mathscr{A}$ is a *local subalgebra* of $\mathscr{B}$ [22]; in the Russian literature (or rather its translations into English) $\mathscr{A}$ is a *full subalgebra* of $\mathscr{B}$. Connes [55] says that $\mathscr{A}$ is *invariant* under holomorphic calculus in $\mathscr{B}$, and Arveson [8] calls $\mathscr{A}$ a *spectrally invariant subalgebra* of $\mathscr{B}$ and uses the term *spectral permanence*. Some of the terminology will become clearer when we discuss the properties of inverse-closedness in more detail.

5.2.5.1 Spectral Invariance

Recall that the *spectrum* of an element a in Banach algebra $\mathscr{A}$ (with unit e) is defined to be the set

$$\sigma_{\mathscr{A}}(a) = \{\lambda \in \mathbb{C} : a - \lambda e \text{ is not invertible in } \mathscr{A}\}.$$

The *spectral radius* of a is

$$r_{\mathscr{A}}(a) = \max\{|\lambda| : \lambda \in \sigma_{\mathscr{A}}(a)\} = \lim_{n\to\infty} \|a^n\|_{\mathscr{A}}^{1/n}.$$

The last identity is the fundamental spectral radius formula. See [27, 56, 150, 209].

Lemma 5.10. *Let $\mathscr{A} \subseteq \mathscr{B}$ be two Banach algebras with a common unit e. Then the following statements are equivalent:*

1. *$\mathscr{A}$ is inverse-closed in $\mathscr{B}$.*
2. *$\sigma_{\mathscr{A}}(a) = \sigma_{\mathscr{B}}(a)$ for all $a \in \mathscr{A}$.*

Proof. (1) $\Rightarrow$ (2): If $\lambda \notin \sigma_{\mathscr{A}}(a)$, then $(a-\lambda e)^{-1} \in \mathscr{A} \subseteq \mathscr{B}$, so $\lambda \notin \sigma_{\mathscr{B}}(a)$. This means that

$$\sigma_{\mathscr{B}}(a) \subseteq \sigma_{\mathscr{A}}(a).$$

This argument shows that the inclusion $\mathscr{A} \subseteq \mathscr{B}$ always implies that $\sigma_{\mathscr{B}}(a) \subseteq \sigma_{\mathscr{A}}(a)$.

Conversely, if $a \in \mathscr{A}$ and $\lambda \notin \sigma_{\mathscr{B}}(a)$, then $(a - \lambda e)^{-1} \in \mathscr{B}$. Since $\mathscr{A}$ is inverse-closed in $\mathscr{B}$, $(a-\lambda e)^{-1}$ must also be in $\mathscr{A}$, and so $\lambda \notin \sigma_{\mathscr{A}}(a)$ and thus

$$\sigma_{\mathscr{A}}(a) \subseteq \sigma_{\mathscr{B}}(a).$$

(2) $\Rightarrow$ (1): $a \in \mathscr{A}, a^{-1} \in \mathscr{B}$ means $0 \notin \sigma_{\mathscr{B}}(a)$, so $0 \notin \sigma_{\mathscr{A}}(a)$ and $a^{-1} \in \mathscr{A}$. Thus, $\mathscr{A}$ is inverse-closed in $\mathscr{B}$. □

Lemma 5.10 explains why the term *spectral invariance* is often used in connection with an inverse-closed subalgebra.

In the light of Lemma 5.10, Wiener's Lemma states that

$$\sigma_{\mathscr{A}(\mathbb{T})}(f) = \sigma_{C(\mathbb{T})}(f) = f(\mathbb{T}). \tag{5.6}$$

In general, it is very difficult to verify when an algebra $\mathscr{A}$ is inverse-closed in $\mathscr{B}$. Hulanicki's lemma [143] yields an important criterion for, and offers a strategy to prove, inverse-closedness. In this regard Hulanicki's result lies somewhat deeper and requires some additional property of the larger algebra $\mathscr{B}$.

Recall that an involution $a \to a^*$ of an algebra $\mathscr{A}$ is a mapping that satisfies the following properties: (a) $(\lambda a + \mu b)^* = \bar{\lambda} a^* + \bar{\mu} b^*$ for all $a, b \in \mathscr{A}$ and $\lambda, \mu \in \mathbb{C}$; (b) $(a^*)^* = a$ for all $a \in \mathscr{A}$; and (c) $(ab)^* = b^* a^*$ for all $a, b \in \mathscr{A}$. A Banach algebra with a continuous involution is called a *Banach $*$-algebra*. A Banach $*$-algebra $\mathscr{A}$ is called *symmetric* if $\sigma_{\mathscr{A}}(a^*a) \subseteq [0,\infty)$ for all $a \in \mathscr{A}$; i.e., the spectrum of positive elements is positive.

Proposition 5.11 (Hulanicki's lemma). *Assume that $\mathscr{A} \subseteq \mathscr{B}$ are two Banach $*$-algebras with a common unit element and common involution. Assume that $\mathscr{B}$ is symmetric. Then the following are equivalent:*

1. *$\mathscr{A}$ is inverse-closed in $\mathscr{B}$.*
2. *$r_{\mathscr{A}}(a) = r_{\mathscr{B}}(a)$ for all $a = a^* \in \mathscr{A}$.*
3. *$r_{\mathscr{A}}(a) \leq r_{\mathscr{B}}(a)$ for all $a = a^* \in \mathscr{A}$.*

If one of these conditions is satisfied, then $\mathscr{A}$ is also symmetric.

Thus, instead of verifying the spectral identity of Lemma 5.10, it suffices to verify the equality of two spectral radii. The spectral radius is an analytic concept (whereas the spectrum is an algebraic notion), and the equality of spectral radii in condition (2) can be attacked with analytic methods. In this way Hulanicki's lemma offers a strategy to verify inverse-closedness.

Proof. The implication (1) $\Rightarrow$ (2) follows from Lemma 5.10, and the implication (2) $\Rightarrow$ (3) is obvious. The implication (3) $\Rightarrow$ (2) follows from the inclusion $\sigma_{\mathscr{B}}(a) \subseteq \sigma_{\mathscr{A}}(a)$ and the ensuing inequality $r_{\mathscr{B}}(a) \leq r_{\mathscr{A}}(a)$, which always hold when $\mathscr{A} \subseteq \mathscr{B}$ (see the proof of Lemma 5.10).

The heart of Hulanicki's lemma is the nontrivial implication (2) $\Rightarrow$ (1). The idea is that the identity of spectral radii $r_{\mathscr{A}}(a) = r_{\mathscr{B}}(a)$ implies that power series have same radius of convergence in $\mathscr{A}$ and in $\mathscr{B}$.

We first treat a special case: For $b = b^* \in \mathscr{A}$ consider the geometric series $c = \sum_{k=0}^{\infty}(e-b)^k$. If $r_{\mathscr{B}}(e-b) < 1$, then this series converges in $\mathscr{B}$. In this case the sum is $c = b^{-1}$. By the spectral identity $r_{\mathscr{A}}(e-b) = r_{\mathscr{B}}(e-b) < 1$; hence, this series also converges in $\mathscr{A}$. Consequently, its sum $c = b^{-1}$ belongs to the smaller algebra $\mathscr{A}$.

To treat the general case, assume that $a \in \mathscr{A}$ is invertible and $a^{-1} \in \mathscr{B}$. Consider the element $b = (2\|a^*a\|_{\mathscr{B}})^{-1}a^*a \in \mathscr{A} \subseteq \mathscr{B}$. Then b is invertible in $\mathscr{B}$, and $\|b\|_{\mathscr{B}} = 1/2$. Since $\mathscr{B}$ is assumed to be symmetric, the spectrum of b is contained in $[0,\infty)$. The invertibility of b implies that 0 is not in the spectrum, and the norm bound implies that the spectrum is contained in a disc of radius $1/2$. Since the spectrum is a compact set, there exists a $\delta > 0$ such that

$$\sigma_{\mathscr{B}}(b) \subseteq [\delta, \tfrac{1}{2}] \,.$$

Consequently, $\sigma_{\mathscr{B}}(e-b) \subseteq [1/2, 1-\delta]$ and

$$\rho_{\mathscr{B}}(e-b) \leq 1-\delta < 1 \,.$$

This is the situation of the special case above, and we may conclude that $b^{-1} = \sum_{n=0}^{\infty}(e-b)^n$ converges simultaneously in $\mathscr{B}$ and in $\mathscr{A}$, whence $b^{-1} \in \mathscr{A}$.

Now $e = b^{-1}b = ((2\|a^*a\|_{\mathscr{B}})^{-1}b^{-1}a^*)a$ and thus a possesses the left inverse $c = (2\|a^*a\|_{\mathscr{B}})^{-1}b^{-1}a^*$. By applying the same argument to $\tilde{b} = (2\|aa^*\|_{\mathscr{B}})^{-1}aa^* \in \mathscr{A} \subseteq \mathscr{B}$, we obtain a right inverse of a of the form $(2\|aa^*\|_{\mathscr{B}})^{-1}a^*\tilde{b}^{-1} \in \mathscr{A}$. Thus, a is invertible in $\mathscr{A}$.

Finally, since $\mathscr{B}$ is symmetric, we know that $\sigma_{\mathscr{B}}(a^*a) \subseteq [0,\infty)$ for all $a \in \mathscr{A}$. Since $\mathscr{A}$ is inverse-closed in $\mathscr{B}$, $\sigma_{\mathscr{A}}(a^*a) = \sigma_{\mathscr{B}}(a^*a) \subseteq [0,\infty)$ for all $a \in \mathscr{A}$, and so $\mathscr{A}$ must be symmetric. $\square$

Note that the structure of the proof of Hulanicki's lemma is identical to that of the proof of Wiener's Lemma (Theorem 5.5). In both cases we first restricted to positive elements for which we used geometric series to investigate their invertibility.

Since the symmetry of a Banach algebra with involution is usually difficult to verify (and still a topic of many unsolved problems), Hulanicki's lemma is usually applied in the following form.

Lemma 5.12. *Let $\mathscr{A}$ be an involutive Banach algebra with identity e. Suppose that there exists a one-to-one $*$-homomorphism π from $\mathscr{A}$ into $\mathscr{B}(\mathscr{H})$, the C^*-algebra of bounded operators on a Hilbert space $\mathscr{H}$, such that $\pi(e) = \mathrm{Id}_{\mathscr{H}}$. If*

$$r_{\mathscr{A}}(a) = \|\pi(a)\|_{op} \qquad \textit{for all } a = a^* \in \mathscr{A}\,, \tag{5.7}$$

then

$$\sigma_{\mathscr{A}}(a) = \sigma_{\mathscr{B}(\mathscr{H})}(\pi(a)) \qquad \textit{for all } a \in \mathscr{A}\,. \tag{5.8}$$

In particular, $\mathscr{A}$ is symmetric.

Proof. Since π is assumed to be faithful (i.e., one-to-one), we may identify $\mathscr{A}$ with a *-subalgebra of $\mathscr{B} = \mathscr{B}(\mathscr{H})$. With this identification and Proposition 5.11, (5.7) implies the spectral identity (5.8). Furthermore, $\mathscr{B}(\mathscr{H})$ is symmetric, because we know from functional analysis that the spectrum of operators of the form T^*T is positive. Consequently, $\sigma_{\mathscr{A}}(a^*a) = \sigma_{\mathscr{B}(\mathscr{H})}(\pi(a)^*\pi(a)) \subseteq [0,\infty)$, and $\mathscr{A}$ is symmetric.

Remark 5.13. In the jargon π is called a *faithful* $*$-representation of $\mathscr{A}$ by bounded operators on the Hilbert space $\mathscr{H}$. To prove the symmetry of a Banach $*$-algebra, one often constructs a faithful representation π on a Hilbert space and then tries to establish the identity of spectral radii (5.7). Lemma 5.12 thus gives us a glimpse of the important relationship between symmetry and representation theory.

At first glance, symmetry is a property of a single algebra, whereas inverse-closedness is a relationship between two algebras. Nevertheless the two concepts are closely related. Let us digress for a moment and describe their connection. In the abstract theory of Banach $*$-algebras one can assign a C*-algebra to every Banach $*$-algebra. Consider the set $\mathscr{S}$ of all *-seminorms on $\mathscr{A}$; i.e., $\mathscr{S}$ contains all seminorms p on $\mathscr{A}$ satisfying the C*-condition $p(a^*a) = p(a)^2$ for all $a \in \mathscr{A}$. Note that $\mathscr{A}$ is usually not complete with respect to such a seminorm. Now define the maximal C*-seminorm on $\mathscr{A}$, the so-called Gel'fand seminorm, by

$$\gamma_{\mathscr{A}}(a) = \sup\{p(a) : p \in \mathscr{S}\}.$$

The completion of the quotient $\mathscr{A}/\{a \in \mathscr{A} : \gamma_{\mathscr{A}}(a) = 0\}$ with respect to the maximal C*-seminorm $\gamma_{\mathscr{A}}$ is a C*-algebra and is called the *enveloping C*-algebra* of $\mathscr{A}$, denoted by $C^*(\mathscr{A})$. Now we can formulate the relationship between symmetry and inverse-closedness.

Proposition 5.14. *Assume that $\gamma_{\mathscr{A}}$ is a norm on $\mathscr{A}$. Then $\mathscr{A}$ is symmetric if and only if $\mathscr{A}$ is inverse-closed in the enveloping C*-algebra $C^*(\mathscr{A})$.*

For a proof see [196, 11.4]. One implication is easy. If $\mathscr{A}$ is inverse-closed in $C^*(\mathscr{A})$, then Lemma 5.10 implies that $\sigma_{\mathscr{A}}(a^*a) = \sigma_{C^*(\mathscr{A})}(a^*a)$ for all $a \in \mathscr{A}$. Since every C*-algebra is symmetric [27, 196], the spectrum of positive elements a^*a is contained in $[0,\infty)$. This means that $\mathscr{A}$ is symmetric.

Functional Calculus. Recall how the Riesz functional calculus (or holomorphic functional calculus) works. Fix an element $a \in \mathscr{B}$ with spectrum $\sigma_{\mathscr{B}}(a)$. Let f be an analytic functions on an open neighborhood O of $\sigma_{\mathscr{B}}(a)$ and let $\gamma \subseteq O$ be a contour of $\sigma_{\mathscr{B}}(a)$; i.e., γ is a rectifiable curve and points in $\sigma_{\mathscr{B}}(a)$ have winding number 1, and points in the complement of O have winding number 0. Define the Banach-algebra-valued path integral

$$f(a) = \frac{1}{2\pi i}\int_\gamma f(z)(ze-a)^{-1}\,dz. \tag{5.9}$$

Here the integral can be understood as a Riemann integral (limit of Riemann sums); in particular, it is also defined weakly. If $a^* \in \mathscr{A}^*$ is in the dual space of $\mathscr{A}$, then

$$\langle a^*, f(a)\rangle = \frac{1}{2\pi i}\int_\gamma f(z)\langle a^*, (ze-a)^{-1}\rangle\,dz.$$

The mapping $f \to f(a)$ is an algebra homomorphism from the commutative algebra of functions analytic on some neighborhood of $\sigma_{\mathscr{B}}(a)$ into a commutative subalgebra of $\mathscr{A}$. For a detailed exposition of the functional calculus, see [56, 209].

Usually the functional calculus depends sensitively on the underlying algebra. We therefore note an immediate, but important, consequence of the spectral invariance.

Corollary 5.15. *Assume that $\mathscr{A}$ is continuously embedded in $\mathscr{B}$ and inverse-closed in $\mathscr{B}$. Then the Riesz functional calculi for $\mathscr{A}$ and $\mathscr{B}$ coincide.*

Proof. Since $\sigma_{\mathscr{A}}(a) = \sigma_{\mathscr{B}}(a)$ by Lemma 5.10, the set of analytic functions f for which $f(a)$ is well defined by (5.9) is the same for $\mathscr{A}$ as for $\mathscr{B}$. Thus, the path integral in (5.9) defines an element in both $\mathscr{A}$ and $\mathscr{B}$. Since $\mathscr{A}$ is continuously embedded $\mathscr{B}$, the limit of Riemann sums is the same in $\mathscr{A}$ as in $\mathscr{B}$. Thus, $f(a)$ is defined unambiguously and does not depend on the algebra. □

Corollary 5.15 is extremely useful in situations when the existence of $f(a)$ is known by other means. The most common situation is when $\mathscr{B}$ is $\mathscr{B}(\mathscr{H})$, the algebra of bounded operators on a Hilbert space. In this case we have the continuous functional calculus at our disposal and know how to establish the existence of (square) roots, absolute values, and other functions of positive operators. Assume that $\mathscr{A}$ is inverse-closed in $\mathscr{B}(\mathscr{H})$ and that a is a positive invertible element in $\mathscr{A}$, i.e., $a = b^*b$ for some $b \in \mathscr{A}$. Then $f(z) = z^\sigma$ is analytic on a neighborhood of $\sigma_{\mathscr{A}}(a) \subseteq [\alpha, \beta]$, $\alpha > 0$, and a^σ makes sense in $\mathscr{B}$ for arbitrary $\sigma \in \mathbb{R}$. By Corollary 5.15, $a^\sigma \in \mathscr{A}$ as well.

As another consequence we state an early result from the theory of absolutely convergent series, which is known as the *theorem of Wiener–Levy* [165, 248].

Theorem 5.16. *Assume that $h \in \mathscr{A}(\mathbb{T})$ and that f is holomorphic on an open set containing the image $h(\mathbb{T})$. Then $f \circ h \in \mathscr{A}(\mathbb{T})$.*

Proof. We use the Riesz functional calculus to compute $f(h)$. Choose a contour γ of $\sigma_{\mathscr{A}(\mathbb{T})}(h) = h(\mathbb{T})$ [by (5.6)] and let $\delta_t \in \mathscr{A}(\mathbb{T})^*$ be the point evaluation $\delta_t(h) = h(t)$. Then with the weak definition of the Banach-algebra-valued integral we obtain

$$
\begin{aligned}
f(h)(t) &= \langle \delta_t, f(h) \rangle \\
&= \frac{1}{2\pi i} \int_\gamma f(z) \langle \delta_t, (ze - h)^{-1} \rangle \, dz \\
&= \frac{1}{2\pi i} \int_\gamma f(z)(z - h(t))^{-1} \, dz \; = f\big(h(t)\big),
\end{aligned}
$$

where in the last step we used Cauchy's integral formula. By the properties of the functional calculus, $f(h) \in \mathscr{A}(\mathbb{T})$, and by the above computation, $f(h) = f \circ h$. Hence, $f \circ h$ possesses an absolutely convergent Fourier series. □

Note that if we take $f(z) = 1/z$, we recover the classical formulation of Wiener's Lemma.

5.2.6 Convolution Operators

Wiener's Lemma can be reformulated as a statement about the spectrum of convolution operators. This formulation is useful for concrete problems in signal analysis.

As a motivation let us introduce some engineering terminology. In engineering a *system* is a black box that transforms an input signal $\mathbf{a}$ into an output signal $\mathbf{b}$. Usually this transformation is assumed to be linear, so for a mathematician a system is just a linear operator A. In other applications $\mathbf{a}$ is a signal that is transmitted by a sender and $\mathbf{b}$ is the received signal. In this case, the operator A describes the distortion of $\mathbf{a}$. In this context one speaks of a *channel* rather than of a system.

The goal is to understand the properties of the "channel" and the input–output relationship. A specific goal in signal processing is to calculate or estimate the input $\mathbf{a}$ from a measured output $\mathbf{b}$. This process is called *equalization* and amounts to solving the equation $A\mathbf{a} = \mathbf{b}$ for $\mathbf{a}$ or to inverting A.

The simplest systems are discrete time-invariant systems corresponding to a black box with constant characteristics or to a stationary transmission environment. As before we denote sequences with boldface letters $\mathbf{a}, \mathbf{b}$, etc., and their entries with $a(k), b(k)$ or $a_k, b_k, k \in \mathbb{Z}$. Let T_r denote the translation operator; it acts on a sequence $\mathbf{a}$ by $(T_r\mathbf{a})(k) = a(k-r)$ for $k, r \in \mathbb{Z}$. Time invariance means that if the input $\mathbf{a}$ results in the output $\mathbf{b}$, then the shifted input $T_r\mathbf{a}$ results in the output $T_r\mathbf{b}$. For a linear system we then have

$$AT_r\mathbf{a} = T_rA\mathbf{a}, \qquad \forall r \in \mathbb{Z}, \tag{5.10}$$

for "all" sequences $\mathbf{a}$. Mathematically, a time-invariant system is therefore an operator that commutes with translations.

Let $\delta_k, k \in \mathbb{Z}$, be the standard basis of $\ell^2(\mathbb{Z})$ defined by $\delta_k(l) = 1$, if $l = k$, and $\delta_k(l) = 0$, if $l \neq k$. Then $\delta_k = T_k\delta_0$ and every sequence on $\mathbb{Z}$ can formally be written as $\mathbf{a} = \sum_{k\in\mathbb{Z}} a(k)\delta_k = \sum_{k\in\mathbb{Z}} a(k)T_k\delta_0$.

Then by (5.10) we find that

$$\begin{aligned}
(A\mathbf{a})(l) &= A\Big(\sum_k a(k)T_k\delta_0\Big)(l) \\
&= \sum_k a(k)T_k(A\delta_0)(l) \\
&= \sum_k a(k)(A\delta_0)(l-k) \\
&= \Big(\mathbf{a} * (A\delta_0)\Big)(l).
\end{aligned}$$

Thus, the time-invariant system A is the convolution with the sequence $\mathbf{h} := A\delta_0$. This sequence is the response of the system to the "pulse" δ_0 and therefore is called the *impulse response*.

We denote the convolution operator $C_\mathbf{h}\mathbf{a} = \mathbf{a} * \mathbf{h}$ and call $\mathbf{h}$ the *symbol* of $C_\mathbf{h}$. Our informal argument shows that every time-invariant linear system is uniquely defined by its impulse response $\mathbf{h}$, and conversely that every sequence $\mathbf{h}$ defines a unique time-invariant system $A = C_\mathbf{h}$.

For physical reasons we may assume that signals of finite energy (i.e., finite ℓ^2-norm) are mapped to signals of finite energy; in other words, A is bounded on ℓ^2. Usually the symbol $\mathbf{h}$ is a finite sequence (a system with "finite impulse response"), so it is safe to assume that $\mathbf{h}$ is in $\ell^1(\mathbb{Z})$. Under this assumption our deduction is rigorous, because all sums converge absolutely.

In particular, we note the following boundedness property.

Lemma 5.17 (Young's inequality).

1. *If* $\mathbf{h} \in \ell^1(\mathbb{Z})$ *and* $\mathbf{a} \in \ell^p(\mathbb{Z})$*, then* $\mathbf{h} * \mathbf{a} \in \ell^p(\mathbb{Z})$ *and* $\|\mathbf{h} * \mathbf{a}\|_p \le \|\mathbf{h}\|_1 \|\mathbf{a}\|_p$.
2. *Thus, if* $\mathbf{h} \in \ell^1(\mathbb{Z})$*, then* $C_{\mathbf{h}}$ *is bounded on* ℓ^p *for* $1 \le p \le \infty$*. The operator norm is bounded uniformly by*
$$\|C_{\mathbf{h}}\|_{\ell^p \to \ell^p} \le \|\mathbf{h}\|_1 .$$

By Young's inequality a convolution operator with ℓ^1-symbol is bounded simultaneously on all ℓ^p-spaces. So for $\mathbf{h} \in \ell^1$, $C_{\mathbf{h}}$ is an element of the Banach algebra $\mathscr{B}(\ell^p)$, the Banach algebra of bounded operators on $\ell^p(\mathbb{Z})$. Let

$$\sigma_{\mathscr{B}(\ell^p)}(C_{\mathbf{h}}) = \{\lambda \in \mathbb{C} : C_{\mathbf{h}} - \lambda \, \mathrm{Id}_{\ell^p} \text{ is not invertible on } \ell^p(\mathbb{Z})\}$$

be the spectrum of $C_{\mathbf{h}}$ as an operator acting on $\ell^p(\mathbb{Z})$.

An immediate question is how the spectrum of $C_{\mathbf{h}}$ depends on the domain space $\ell^p(\mathbb{Z})$. This question is important for signal analysis and interesting in its own right. In signal analysis one would like to deduce properties of the input $\mathbf{a}$ from the output $\mathbf{b} = C_{\mathbf{h}}\mathbf{a}$. In particular, if $\mathbf{b} \in \ell^p(\mathbb{Z})$, can we assert that $\mathbf{a}$ is in the same $\ell^p(\mathbb{Z})$?

The mathematical analysis of this problem brings good news for the engineer: The spectrum is independent of $\ell^p(\mathbb{Z})$, and in particular the invertibility of a time-invariant system is independent of the domain space $\ell^p(\mathbb{Z})$.

We now come to the main point: Wiener's Lemma is the main tool to understand and compute the spectrum of convolution operators. We first give a formulation of Wiener's Lemma for convolution operators that is equivalent to the classical Wiener's Lemma.

Theorem 5.18 (Wiener's Lemma for convolution operators). *If* $\mathbf{h} \in \ell^1(\mathbb{Z})$ *and* $C_{\mathbf{h}}$ *is invertible on* $\ell^2(\mathbb{Z})$*, then the inverse operator is again a convolution operator* $C_{\mathbf{h}}^{-1} = C_{\mathbf{g}}$ *with a symbol* $\mathbf{g} \in \ell^1(\mathbb{Z})$*. Consequently,* $C_{\mathbf{h}}$ *is invertible simultaneously on all* $\ell^p(\mathbb{Z})$*,* $1 \le p \le \infty$*.*

Proof. As a preparation, consider the Fourier series $\widehat{\mathbf{h}}(t) = \sum_{k \in \mathbb{Z}} h_k e^{2\pi i k t}$ of the sequence $\mathbf{h}$. Since $\{e^{2\pi i k t} : k \in \mathbb{Z}\}$ is an orthonormal basis for $L^2(\mathbb{T})$, the mapping $\mathbf{h} \to \widehat{\mathbf{h}}$ is a unitary operator from $\ell^2(\mathbb{Z})$ onto $L^2(\mathbb{T})$. Consequently, if $\mathbf{h} \in \ell^2(\mathbb{Z})$, then the Fourier series $\widehat{\mathbf{h}}$ converges in $L^2(\mathbb{T})$ and $\widehat{\mathbf{h}}$ is defined almost everywhere. If, in addition, $\mathbf{h} \in \ell^1(\mathbb{Z})$, then the Fourier series $\widehat{\mathbf{h}}$ converges absolutely and $\widehat{\mathbf{h}} \in \mathscr{A}(\mathbb{T})$.

Next, by reading (5.1) backwards, we know that the Fourier series of $\mathbf{h} * \mathbf{g}$, $\mathbf{h}, \mathbf{g} \in \ell^1(\mathbb{Z})$, is just the pointwise product

$$\widehat{\mathbf{h} * \mathbf{g}} = \widehat{\mathbf{h}} \, \widehat{\mathbf{g}}. \tag{5.11}$$

If $\mathbf{h} \in \ell^1(\mathbb{Z})$ and $\mathbf{g} \in \ell^2(\mathbb{Z})$, then this formula holds almost everywhere and is an identity of two L^2-functions.

Consequently, the Fourier series of $C_{\mathbf{h}}\mathbf{a}$ is

$$\widehat{C_{\mathbf{h}}\mathbf{a}} = \widehat{\mathbf{h} * \mathbf{a}} = \widehat{\mathbf{h}}\widehat{\mathbf{a}}. \tag{5.12}$$

Thus, on the Fourier side, $C_{\mathbf{h}}$ becomes the multiplication operator $M_{\widehat{\mathbf{h}}}\widehat{\mathbf{a}} = \widehat{\mathbf{h}}\,\widehat{\mathbf{a}}$. More precisely, we may write the identity (5.12) by means of a commutative diagram:

$$\begin{array}{ccc} \ell^2(\mathbb{Z}) & \xrightarrow{C_{\mathbf{h}}} & \ell^2(\mathbb{Z}) \\ \downarrow\widehat{\ } & & \downarrow\widehat{\ } \\ L^2(\mathbb{T}) & \xrightarrow{M_{\widehat{\mathbf{h}}}} & L^2(\mathbb{T}). \end{array} \tag{5.13}$$

Since $\widehat{\ }$ is unitary, $C_{\mathbf{h}}$ has the same spectrum as $M_{\widehat{\mathbf{h}}}$:

$$\sigma_{\mathscr{B}(L^2)}(M_{\widehat{\mathbf{h}}}) = \sigma_{\mathscr{B}(\ell^2)}(C_{\mathbf{h}}) = \operatorname{ran}\widehat{\mathbf{h}}. \tag{5.14}$$

In particular, $C_{\mathbf{h}}$ is invertible on $\ell^2(\mathbb{Z})$ if and only if $M_{\widehat{\mathbf{h}}}$ is invertible on $L^2(\mathbb{T})$. This is the case if and only if $|\widehat{\mathbf{h}}(t)| \geq \delta > 0$ for almost all $t \in \mathbb{T}$.

Here Wiener's Lemma makes its decisive appearance: Since $C_{\mathbf{h}}$ is assumed to be invertible and $\mathbf{h} \in \ell^1(\mathbb{Z})$, we have $\widehat{\mathbf{h}}(t) \neq 0$ for all $t \in \mathbb{T}$. Theorem 5.5 asserts that $1/\widehat{\mathbf{h}}$ possesses an absolutely convergent Fourier series. This means that there exists a $\mathbf{g} \in \ell^1(\mathbb{Z})$ such that $\widehat{\mathbf{g}} = 1/\widehat{\mathbf{h}}$. By (5.11), the equation $\widehat{\mathbf{g}}\widehat{\mathbf{h}} \equiv 1$ implies that

$$\mathbf{g} * \mathbf{h} = \mathbf{h} * \mathbf{g} = \delta_0.$$

Consequently, for $\mathbf{a} \in \ell^2(\mathbb{Z})$ we find, using (5.11) repeatedly,

$$\begin{aligned} \mathbf{a} &= \delta_0 * \mathbf{a} = (\mathbf{h} * \mathbf{g}) * \mathbf{a} \\ &= \mathbf{h} * (\mathbf{g} * \mathbf{a}) = C_{\mathbf{h}}C_{\mathbf{g}}\mathbf{a}. \end{aligned}$$

Thus, $C_{\mathbf{h}}C_{\mathbf{g}} = Id_{\ell^2}$ and likewise $C_{\mathbf{g}}C_{\mathbf{h}} = \mathrm{Id}_{\ell^2}$. So $C_{\mathbf{h}}^{-1} = C_{\mathbf{g}}$ with $\mathbf{g} \in \ell^1(\mathbb{Z})$ as claimed. Since both $C_{\mathbf{h}}$ and $C_{\mathbf{g}}$ are bounded on every $\ell^p(\mathbb{Z})$, $1 \leq p \leq \infty$, the convolution operator $C_{\mathbf{h}}$ is also invertible on $\ell^p(\mathbb{Z})$ with the inverse $C_{\mathbf{g}}$. □

Spectral Invariance of Convolution Operators. Finally, let us draw some consequences of Wiener's Lemma for convolution operators.

Theorem 5.19. *Assume that* $\mathbf{h} \in \ell^1(\mathbb{Z})$. *Then the following statements are equivalent.*

1. $C_{\mathbf{h}}$ *is invertible on* $\ell^2(\mathbb{Z})$.
2. $C_{\mathbf{h}}$ *is invertible on* $\ell^p(\mathbb{Z})$ *for* all $p \in [1,\infty]$.
3. $C_{\mathbf{h}}$ *is invertible on* $\ell^p(\mathbb{Z})$ *for* some $p \in [1,\infty]$.

Proof. The implication (1) $\Rightarrow$ (2) is Wiener's Lemma for convolution operators (Theorem 5.18) and the implication (2) $\Rightarrow$ (3) is trivial. To verify (3) $\Rightarrow$ (1), let us assume that $C_{\mathbf{h}}$ is invertible on $\ell^p(\mathbb{Z})$ for some $p \in [1,\infty]$.

For the implication (3) $\Rightarrow$ (1) we use a duality argument and interpolation. First let us prepare the argument. Let $\mathbf{a} \to \mathbf{a}^*$ be the usual involution $\mathbf{a}^*(k) = \overline{a(-k)}$ extended to $\ell^p(\mathbb{Z})$. This involution is a bijective conjugate-linear isometry on $\ell^p(\mathbb{Z})$ and satisfies the relation $(\mathbf{a} * \mathbf{b})^* = \mathbf{b}^* * \mathbf{a}^* = \mathbf{a}^* * \mathbf{b}^*$ whenever the convolution is defined, in particular, when $\mathbf{a} \in \ell^1(\mathbb{Z})$ and $\mathbf{b} \in \ell^p(\mathbb{Z})$. Next, using the inner product $\langle \mathbf{a}, \mathbf{b} \rangle = \sum_{k\in\mathbb{Z}} a_k \overline{b_k}$ for the duality between $\ell^p(\mathbb{Z})$ and $\ell^{p'}(\mathbb{Z})$, $1/p + 1/p' = 1$, it is easily checked that

$$\langle C_{\mathbf{h}}\mathbf{a}, \mathbf{b} \rangle = \langle \mathbf{a}, C_{\mathbf{h}^*}\mathbf{b} \rangle \qquad \text{for } \mathbf{h} \in \ell^1, \mathbf{a} \in \ell^p, \mathbf{b} \in \ell^{p'}.$$

Thus, the adjoint operator of a convolution operator with respect to $\langle \cdot, \cdot \rangle$ is $(C_{\mathbf{h}})^* = C_{\mathbf{h}^*}$.

Claim: For $\mathbf{h} \in \ell^1(\mathbb{Z})$ and fixed $p \in [1,\infty]$, the convolution operator $C_{\mathbf{h}}$ is invertible on $\ell^p(\mathbb{Z})$ if and only if $C_{\mathbf{h}^*}$ is invertible on $\ell^p(\mathbb{Z})$.

Since $C_{\mathbf{h}^*}\mathbf{a} = \mathbf{h}^* * \mathbf{a} = (\mathbf{h} * \mathbf{a}^*)^* = (C_{\mathbf{h}}\mathbf{a}^*)^*$ and * is a bijection on $\ell^p(\mathbb{Z})$, $C_{\mathbf{h}^*}$ is one-to-one if and only if $C_{\mathbf{h}}$ is one-to-one, and likewise $C_{\mathbf{h}^*}$ is onto if and only if $C_{\mathbf{h}}$ is onto.

Now assume that $C_{\mathbf{h}}$ is invertible on $\ell^p(\mathbb{Z})$. By the claim, $C_{\mathbf{h}^*}$ is also invertible on $\ell^p(\mathbb{Z})$ and thus by a general principle [56, Prop. VI.1.4] its adjoint $(C_{\mathbf{h}^*})^* = C_{\mathbf{h}}$ is invertible on the dual space $\ell^{p'}(\mathbb{Z})$. Let M be the inverse of $C_{\mathbf{h}}$ on $\ell^{\max(p,p')}(\mathbb{Z})$; then clearly M is also the inverse of $C_{\mathbf{h}}$ on $\ell^{\min(p,p')}(\mathbb{Z})$. Since M is bounded on both $\ell^p(\mathbb{Z})$ and $\ell^{p'}(\mathbb{Z})$, the Riesz–Thorin interpolation theorem [151, 248] implies that M is bounded on $\ell^2(\mathbb{Z})$. The factorization $MC_{\mathbf{h}} = C_{\mathbf{h}}M = \mathrm{Id}_{\ell^2}$ holds on the dense subspace $\ell^{\min(p,p')}(\mathbb{Z})$, and thus $C_{\mathbf{h}}$ is invertible on $\ell^2(\mathbb{Z})$ with inverse M, as was to be shown. □

Corollary 5.20. *If* $\mathbf{h} \in \ell^1(\mathbb{Z})$*, then*

$$\sigma_{\mathscr{B}(\ell^p)}(C_{\mathbf{h}}) = \sigma_{\mathscr{B}(\ell^2)}(C_{\mathbf{h}}) = \sigma_{\ell^1(\mathbb{Z})}(\mathbf{h}) = \widehat{\mathbf{h}}(\mathbb{T}), \qquad \forall p \in [1,\infty].$$

Proof. Theorem 5.19 says that the algebra $\{C_{\mathbf{h}} : \mathbf{h} \in \ell^1(\mathbb{Z})\}$ is inverse-closed in the Banach algebra $\mathscr{B}(\ell^p(\mathbb{Z}))$. The spectral identity $\sigma_{\mathscr{B}(\ell^p)}(C_{\mathbf{h}}) = \sigma_{\mathscr{B}(\ell^1)}(C_{\mathbf{h}})$ now follows from Lemma 5.10. Finally, $\sigma_{\mathscr{B}(\ell^2)}(C_{\mathbf{h}}) = \widehat{\mathbf{h}}(\mathbb{T})$ follows from (5.14). □

Summarizing, we may say that convolution operators obey a strong form of spectral invariance: Namely, the spectrum of a convolution operator is independent of the domain space $\ell^p(\mathbb{Z})$. For a version on noncommutative groups, see Section 5.3.4.

Symbolic Calculus. Wiener's Lemma may be seen as a primitive form of a symbol calculus.

Usually, by a symbol calculus we understand a mapping from functions to operators. To each function a is associated an operator $\mathrm{Op}(a)$. Then a is called the *symbol* of the operator. The mapping $a \to \mathrm{Op}(a)$ is assumed to be linear. A nice symbolic calculus satisfies some additional desirable properties. Whereas the pointwise product of functions is commutative, the composition of operators is noncommutative in general; thus, the mapping $a \to \mathrm{Op}(a)$ usually fails to be an algebra homomorphism. However, for a "good" symbolic calculus or for a suitable class of "nice" symbols one can often show that the mapping is close to an algebra homomorphism by showing that $\mathrm{Op}(ab) - \mathrm{Op}(a)\mathrm{Op}(b)$ is small in some sense. In particular, if the operator $\mathrm{Op}(a)$ is invertible, then $\mathrm{Op}(a^{-1})$ is an approximate inverse. This idea is fundamental in the symbolic calculus for pseudodifferential operators.

Wiener's Lemma is the prototype of a symbolic calculus. In this case, we map a sequence $\mathbf{h}$ to the convolution operator $C_{\mathbf{h}}$. The distinguished class of symbols is $\ell^1(\mathbb{Z})$. This symbolic calculus is particularly simple, because it is an algebra homomorphism from $\ell^1(\mathbb{Z})$ (with respect to convolution) to bounded operators on $\ell^2(\mathbb{Z})$.

The inverse of a convolution operator $C_{\mathbf{h}}$ is again a convolution operator $C_{\mathbf{h}}^{-1} = C_{\mathbf{g}}$. If $\mathbf{h} \in \ell^1$, then also $\mathbf{g} \in \ell^1$. Thus, Wiener's Lemma shows that the inverse of a convolution operator possesses the same form; i.e., it is again a convolution operator. If the symbol of the operator is "nice" (in ℓ^1), then the symbol of the inverse is also "nice."

Exercises for Section 5.2

1. Let $\mathscr{A}$ be an algebra with a unit element e and $a \in \mathscr{A}$. If there exist $b, c \in \mathscr{A}$ such that $ac = e$ and $ba = e$, then a is invertible and $b = c = a^{-1}$.
2. Consider the algebra $C(\mathbb{T})$ of continuous functions of period 1. Show that the spectrum of a function $f \in C(\mathbb{T})$ coincides with its range: $\sigma_{C(\mathbb{T})}(f) = f(\mathbb{T}) = \{f(t) : t \in \mathbb{T}\}$.
3. (a) Show that both $f(t) = 2 + \cos 2\pi t$ and $g(t) = 1 - |2t - 1|$ (for $t \in [0,1]$ and extended with period 1) possess an absolutely convergent Fourier series.
   (b) Show that f^{α} possesses an absolutely convergent Fourier series for every $\alpha \in \mathbb{R}$.
   (c) Show that $g^{1/2}$ does not have an absolutely convergent Fourier series. (Hint: Use integration by parts to estimate the Fourier coefficients of $g^{1/2}$.)
4. Let $0 < s \leq 1$. We say that a function f on $\mathbb{T}$ is Hölder continuous with exponent s if
$$|f(x) - f(y)| \leq C|x - y|^s \qquad \text{for all } x, y \in \mathbb{T}.$$
   Let $C^s(\mathbb{T})$ be the space of all Hölder continuous functions with exponent s and the norm
$$\|f\|_{C^s} := \|f\|_\infty + \sup_{x,y \in \mathbb{T}, x \neq y} \frac{|f(x) - f(y)|}{|x - y|}.$$
   (a) Show that C^s is a Banach algebra contained in $C(\mathbb{T})$.
   (b) Show that C^s is inverse-closed in $C(\mathbb{T})$.
   (c) Find a bound for the norm of $1/f$ in $C^s(\mathbb{T})$ in terms of $\|f\|_{C^s}$.
5. Let $\mathscr{A}^p(\mathbb{T})$ be the space of all absolutely convergent Fourier series with coefficients in $\ell^p(\mathbb{Z})$ for $0 < p < 1$; i.e., a Fourier series $f(t) = \sum_{k \in \mathbb{Z}} a_k e^{2\pi i k t}$ belongs to $\mathscr{A}^p(\mathbb{T})$ whenever $\sum_{k \in \mathbb{Z}} |a_k|^p < \infty$. Endow $\mathscr{A}^p(\mathbb{T})$ with the norm
$$\|f\|_{\mathscr{A}^p} = \sum_{k \in \mathbb{Z}} |a_k|^p = \|\mathbf{a}\|_p^p.$$
   (a) Show that $\mathscr{A}^p$ satisfies all properties of a Banach algebra, except that the homogeneity of the norm is replaced by the property $\|cf\|_{\mathscr{A}^p} = |c|^p \|f\|_{\mathscr{A}^p}$ for $c \in \mathbb{C}$ and $f \in \mathscr{A}^p(\mathbb{T})$. [Such an algebra is called a p-normed algebra.]
   (b) Show Wiener's Lemma for $\mathscr{A}^p(\mathbb{T})$ and $0 < p < 1$: If $f \in \mathscr{A}^p(\mathbb{T})$ and $f(t) \neq 0$ for all $t \in \mathbb{T}$, then $1/f \in \mathscr{A}^p(\mathbb{T})$.
   Hint: Verify and use the inequality $|a+b|^p \leq |a|^p + |b|^p$ for $a, b \in \mathbb{C}$ and $0 < p \leq 1$. Follow the proof of Wiener's Lemma in Section 5.2.4.
   For the original statement and result, see [247].
6. Brandenburg's trick [30]: Let $\mathscr{A} \subseteq \mathscr{B}$ be two Banach algebras with a common unit element. Assume for every $a \in \mathscr{A}$ there exists a sequence $c_n = c_n(a) > 0$, such that $\lim_{n \to \infty} c_n^{1/n} = 1$ and
$$\|a^{2n}\|_{\mathscr{A}} \leq c_n \|a^n\|_{\mathscr{A}} \|a^n\|_{\mathscr{B}}.$$
   Show that $\mathscr{A}$ is inverse-closed in $\mathscr{B}$. Hint: Apply Hulanicki's Lemma 5.11.

7. Let $\mathscr{A}_s^1(\mathbb{T})$ be the space of all absolutely convergent Fourier series whose coefficients decay like $|a_k| \le C|k|^{-s}$ and norm

$$\|f\|_{\mathscr{A}_s^1} = \sum_{k\in\mathbb{Z}} |a_k| + \sup_{k\in\mathbb{Z}} |a_k||k|^s .$$

(a) Show that $\mathscr{A}_s^1(\mathbb{T})$ is a Banach algebra for any $s \ge 0$.
(b) Show that $\mathscr{A}_s^1(\mathbb{T})$ is inverse-closed in $C(\mathbb{T})$.
Hint: Use that $|k+l|^s \le C_s(|k|^s + |l|^s)$ for all $k, l \in \mathbb{Z}$ and prove that

$$\sup_{k\in\mathbb{Z}} |(\mathbf{a} * \mathbf{b})(k)|\,|k|^s \le C\left(\|\mathbf{a}\|_1 \sup_{k\in\mathbb{Z}} |b_k|\,|k|^s + \|\mathbf{b}\|_1 \sup_{k\in\mathbb{Z}} |a_k|\,|k|^s\right).$$

Now apply Brandenburg's trick from Exercise 6.

5.3 Variations

In this chapter we discuss a number of variations of Wiener's Lemma. We will cover the following variations:

- weighted versions of Wiener's Lemma for Fourier series;
- noncommutative versions for matrix algebras;
- a version of Wiener's Lemma for twisted convolution and its relation to noncommutative tori;
- the analysis of the spectrum of convolution operators in certain non-Abelian locally compact groups; and
- a symbolic calculus for pseudodifferential operators and their spectrum.

We will develop each subject in parallel to the exposition of Section 5.2.

- First, we will define the basic concepts and unravel a relevant Banach algebra.
- Then we will formulate a version of Wiener's Lemma and turn it into a statement of inverse-closedness between two Banach algebras. Although we will not be able to give the complete proofs in each case, we will explain the main ideas and establish the connection to the classical version of Wiener's Lemma.
- Finally, we will make explicit the consequences for spectral invariance.

Since each subsection draws material from a different field of mathematics, the exposition is not always self-contained. Our goal is to provide a synthetic and unifying view of related topics in apparently unrelated fields.

5.3.1 Weighted Versions of Wiener's Lemma

When considering Fourier series, the ℓ^1-condition on the coefficients guarantees that the series converges absolutely. In particular, the partial sums converge in the supremum norm. To obtain faster convergence of the partial sums, it is natural to impose decay conditions on the coefficients. This is done with weight functions.

In general, a weight is simply a nonnegative function. To consider weighted absolutely convergent Fourier series, we use the following definition. From now on, we work with multivariate Fourier series $f(t) = \sum_{k\in\mathbb{Z}^d} a_k e^{2\pi i k\cdot t}$ for $t = (t_1,\dots,t_d) \in \mathbb{T}^d$ and replace the index set $\mathbb{Z}$ by $\mathbb{Z}^d$.

A weight v on $\mathbb{Z}^d$ is called *submultiplicative* if

$$v(k+l) \le v(k)v(l) \qquad \text{for } k,l \in \mathbb{Z}^d\,. \tag{5.15}$$

For simplicity we consider only symmetric weights satisfying $v(-k) = v(k)$. Associated to each weight function on $\mathbb{Z}^d$ is the weighted ℓ^1-space ℓ^1_v defined by the norm

$$\|\mathbf{a}\|_{\ell^1_v} = \|\mathbf{a}v\|_1 = \sum_{k\in\mathbb{Z}^d} |a_k| v(k)\,. \tag{5.16}$$

In analogy to $\mathscr{A}(\mathbb{T}^d)$, we define the weighted absolutely convergent Fourier series as follows: We say that $f \in \mathscr{A}_v(\mathbb{T}^d)$ if $f(t) = \sum_{k\in\mathbb{Z}^d} a_k\, e^{2\pi i k\cdot t}$ with norm

$$\|f\|_{\mathscr{A}_v} = \|\mathbf{a}\|_{\ell^1_v}\,.$$

Example 5.21. The typical submultiplicative weights on $\mathbb{Z}^d$ are of the form

$$v(k) = e^{a|k|^b}\,(1+|k|)^s, \qquad k \in \mathbb{Z}^d,$$

for $a,s \geq 0$ and $0 \leq b \leq 1$.

Weights are useful for the study of convergence properties of Fourier series. They can be seen as a parameter for the rate of convergence of the partial sums. For simplicity take an absolutely convergent Fourier series of one variable $f(t) = \sum_{k\in\mathbb{Z}} a_k e^{2\pi i k t} \in \mathscr{A}_v(\mathbb{T})$ and let $S_N f(t) = \sum_{|k|\leq N} a_k e^{2\pi i k t}$ be the Nth partial sum of f. Then

$$\begin{aligned}
\|f - S_N f\|_\infty &= \Big\|\sum_{|k|>N} a_k e^{2\pi i k t}\Big\|_\infty \\
&\leq \sum_{|k|>N} |a_k| v(k) v(k)^{-1} \\
&\leq \Big(\sup_{|k|>N} v(k)^{-1}\Big) \sum_{|k|>N} |a_k| v(k) \\
&\leq \Big(\sup_{|k|>N} v(k)^{-1}\Big) \|f\|_{\mathscr{A}_v}\,.
\end{aligned}$$

For increasing weight v, such as the standard weights in Example 5.21, we have $\sup_{|k|>N} v(k)^{-1} \leq v(N)^{-1}$, and thus the partial sums $S_N f$ converge to f at the rate $v(N)^{-1}$. The precise connection between the approximability by trigonometric polynomials and the decay of the Fourier coefficients is treated in approximation theory. See, for instance, [66].

The following simple lemma explains why we need the above conditions on the weight function v.

Lemma 5.22. *1. If v is submultiplicative, then $\mathscr{A}_v(\mathbb{T}^d)$ is a Banach algebra with respect to pointwise multiplication.*

2. If, in addition, v is symmetric, then complex conjugation $f \to \bar{f}$ is an isometry on $\mathscr{A}_v$.

Proof. The proof is almost identical to the proof of Lemma 5.4. Fix $\mathbf{a},\mathbf{b} \in \ell^1_v(\mathbb{Z}^d)$. Since $v(n) \leq v(k)v(n-k)$ by the submultiplicativity of v, we have

$$\begin{aligned}
\|\mathbf{a} * \mathbf{b}\|_{\ell^1_v} &= \sum_{n\in\mathbb{Z}} \Big|\sum_{k\in\mathbb{Z}} a_k b_{n-k}\Big| v(n) \\
&\leq \sum_{k\in\mathbb{Z}} \sum_{n\in\mathbb{Z}} |a_k|\,|b_{n-k}| v(k) v(n-k) \\
&= \|\mathbf{a}\|_{\ell^1_v} \|\mathbf{b}\|_{\ell^1_v}\,.
\end{aligned}$$

Now let $f, g \in \mathscr{A}_v$ with coefficient sequences $\mathbf{a}$ and $\mathbf{b} \in \ell^1_v(\mathbb{Z}^d)$. Then

$$\|fg\|_{\mathscr{A}_v} = \|\mathbf{a} * \mathbf{b}\|_{\ell^1_v} \leq \|\mathbf{a}\|_{\ell^1_v} \|\mathbf{b}\|_{\ell^1_v} = \|f\|_{\mathscr{A}_v} \|g\|_{\mathscr{A}_v}$$

and $\mathscr{A}_v(\mathbb{T}^d)$ is a Banach algebra. Item (2) is obvious. □

At this point arises the natural question of whether Wiener's Lemma holds for the weighted absolutely convergent series $\mathscr{A}_v(\mathbb{T}^d)$. In other words, suppose we know that $f \in \mathscr{A}_v(\mathbb{T}^d)$ and $f(t) \neq 0$ for all $t \in \mathbb{T}^d$. What can we say about $1/f$?

Though this seems a harmless variation that is typical for the mathematical mind, the answer to this question has led to a new idea in Banach algebra theory. A striking result of Gel'fand, Raikov, and Shilov characterizes all weights for which Wiener's Lemma remains true [102].

Before we state their fundamental result, we need a new concept about weights.

Definition 5.23. A submultiplicative weight v is said to satisfy the *GRS condition* (Gel'fand-Raikov-Shilov condition) if

$$\lim_{n \to \infty} v(n\mathbf{k})^{1/n} = 1, \quad \forall \mathbf{k} \in \mathbb{Z}^d \,. \tag{5.17}$$

(When dealing with the GRS condition and only in this context, we write integer vectors in boldface as $\mathbf{k} \in \mathbb{Z}^d$ to distinguish them from positive integers.)

The limit in (5.17) exists always because v is submultiplicative. If v is symmetric, then $v(\mathbf{k}) \geq 1$ for all $\mathbf{k} \in \mathbb{Z}^d$ and so always $\lim_{n \to \infty} v(n\mathbf{k})^{1/n} \geq 1$.

Considering the standard weight functions $v(\mathbf{k}) = e^{a|\mathbf{k}|^b}(1+|\mathbf{k}|)^s$, we see immediately that v satisfies the GRS condition if and only if $0 \leq b < 1$. On the other hand, if $b = 1$ and $v(\mathbf{k}) = e^{a|\mathbf{k}|}$ for $a > 0$, then obviously $v(n\mathbf{k})^{1/n} = e^{a|\mathbf{k}|} > 1$ for all n and $\mathbf{k} \neq 0$, and thus the exponential weight violates the GRS condition.

In fact, exponential growth in some direction is the only reason why the GRS condition may fail. Assume that the weight v violates the GRS condition. Then there exist a $\mathbf{k} \in \mathbb{Z}^d$ and $a > 0$ such that

$$v(n\mathbf{k})^{1/n} \geq e^a > 1 \quad \text{ for } n \geq N_0 \,.$$

Thus, $v(n\mathbf{k}) \geq e^{an}$ and the weight v grows exponentially along the subgroup $\mathbf{k}\mathbb{Z}$.

To summarize, the GRS condition is a precise technical condition that excludes the exponential growth of a weight.

We can now formulate the weighted version of Wiener's lemma.

Theorem 5.24. *Let v be a submultiplicative weight on $\mathbb{Z}^d$. If v satisfies the GRS condition, then Wiener's Lemma holds for $\mathscr{A}_v$: If $f \in \mathscr{A}_v(\mathbb{T}^d)$ and $f(t) \neq 0$ for all $t \in \mathbb{T}^d$, then $1/f \in \mathscr{A}_v(\mathbb{T}^d)$.*

Proof. We follow the proof of Wiener's Lemma in Section 5.2.4 and make the necessary modifications when the weight occurs.

As in Steps 1 and 2, we may assume that $h \in \mathscr{A}_v(\mathbb{T})$ and $0 \leq h(t) \leq 1 - \delta$ for some $\delta > 0$. We then have to show that the geometric series $\sum_{n=0}^{\infty} h^n$ converges in $\mathscr{A}_v(\mathbb{T})$, and not just in $C(\mathbb{T})$.

Now choose a trigonometric polynomial p (of degree N in each variable) such that $\|f-p\|_{\mathscr{A}_v} < \varepsilon$. Set $r = h - p$, and then $\|r\|_\infty \le \|r\|_{\mathscr{A}_v} < \varepsilon$.

The main modification is the following estimate for the comparison of various norms: If q is a trigonometric polynomial of degree N in each of the d variables, then—writing $|k|_\infty = \max_{j=1,\dots,d} |k_j|$ for the maximum norm on $\mathbb{R}^d$—we obtain

$$\begin{aligned} \|q\|_{\mathscr{A}_v} &= \sum_{|k|_\infty \le N} |b_k| v(k) \\ &\le \|b\|_2 \left(2N+1\right)^{d/2} \max_{|k|_\infty \le N} v(k) \\ &= \|q\|_2 \left(2N+1\right)^{d/2} \max_{|k|_\infty \le N} v(k) \\ &\le \|q\|_\infty \left(2N+1\right)^{d/2} \max_{|k|_\infty \le N} v(k)\,. \end{aligned} \tag{5.18}$$

For further use, we set $\tilde{v}(n) = \max_{|k|_\infty \le n} v(k)$ and formulate the properties of $\tilde{v}$ as a sublemma [89].

Lemma 5.25. *The weight $\tilde{v}$ is submultiplicative and increasing on $\mathbb{N}$. If v satisfies the GRS condition, then $\tilde{v}$ also satisfies the GRS condition.*

The proof is elementary, but not instructive. For completeness we will give it at the end of this section.

To estimate the $\mathscr{A}_v$-norm of h^n, we use the binomial theorem and obtain as in Step 5 of Section 5.2.4

$$\|h^n\|_{\mathscr{A}_v} \le \sum_{l=0}^{n} \binom{n}{l} \|p^l\|_{\mathscr{A}_v} \|r^{n-l}\|_{\mathscr{A}_v} \le \sum_{l=0}^{n} \binom{n}{l} \|p^l\|_{\mathscr{A}_v}\, \varepsilon^{n-l}\,.$$

By (5.18) we have

$$\|p^l\|_{\mathscr{A}_v} \le \|p^l\|_\infty \,(2Nl+1)^{d/2}\, \tilde{v}(Nl)\,.$$

Therefore, the complete estimate for the $\mathscr{A}_v$-norm of h^n becomes

$$\begin{aligned} \|h^n\|_{\mathscr{A}_v} &\le (2Nn+1)^{d/2} \sum_{l=0}^{n} \binom{n}{l} \varepsilon^{n-l} \|p\|_\infty^l \,\tilde{v}(lN) \\ &= (2Nn+1)^{d/2}\, \tilde{v}(nN)\, \left(\|p\|_\infty + \varepsilon\right)^n \\ &\le (2Nn+1)^{d/2}\, \tilde{v}(nN)\, \left(\|h-r\|_\infty + \varepsilon\right)^n \\ &\le (2Nn+1)^{d/2} \tilde{v}(nN)(1-\delta+2\varepsilon)^n\,. \end{aligned} \tag{5.19}$$

Finally, the $\mathscr{A}_v$-norm of the geometric series $\sum h^n$ is majorized by

$$\sum_{n=0}^{\infty} \|h^n\|_{\mathscr{A}_v} \le \sum_{n=0}^{\infty} (2Nn+1)^{d/2}\, \tilde{v}(Nn)\, \left(1-\delta+2\varepsilon\right)^n\,. \tag{5.20}$$

This series converges provided that

$$1 > \limsup_{n\to\infty} \Big((2Nn+1)^{d/2} \tilde{v}(Nn)\, (1-\delta+2\varepsilon)^n \Big)^{1/n} = \limsup_{n\to\infty} \tilde{v}(Nn)^{1/n} (1-\delta+2\varepsilon)\,.$$

This is the case, because $\tilde{v}$ satisfies the GRS condition by Lemma 5.25. □

Remark 5.26. The expression in (5.20) offers some insight into the nature of the GRS condition. The radius of convergence of the power series $\sum_{n=0}^{\infty} \tilde{v}(nN) z^n$ is exactly $\big(\limsup_{n\to\infty} \tilde{v}(nN)^{1/n}\big)^{-1}$. For this series to converge for all $z, |z| < 1$, we need $\big(\limsup_{n\to\infty} \tilde{v}(nN)^{1/n}\big)^{-1} \geq 1$. This is exactly the GRS condition for $\tilde{v}$.

Corollary 5.27. *The algebra $\mathscr{A}_v(\mathbb{T}^d)$ is inverse-closed in $C(\mathbb{T}^d)$ if and only if the weight v satisfies the GRS condition.*

Proof. The sufficiency of the GRS condition is the content of Theorem 5.24.

To show the necessity of the GRS condition, we assume that v violates this condition and give a counterexample to Wiener's Lemma. The following example illustrates the nature of the GRS condition and will return in several further variations.

If v violates the GRS condition, then there are $\mathbf{k} \in \mathbb{Z}^d$ and $a > 0$, and $n_0 \in \mathbb{N}$ such that

$$v(n\mathbf{k}) \geq e^{an} \qquad \text{for } n \geq n_0\,.$$

Now fix $\delta \in (0, a]$ and set

$$f(t) = 1 - e^{-\delta} e^{2\pi i \mathbf{k}\cdot t} \in \mathscr{A}_v(\mathbb{T}^d).$$

Then $|f(t)| \geq 1 - e^{-\delta} > 0$ and thus $f(t) \neq 0$ for all $t \in \mathbb{T}^d$. Furthermore, the inverse of f is given by the trigonometric series

$$\frac{1}{f(t)} = (1 - e^{-\delta} e^{2\pi i \mathbf{k}\cdot t})^{-1} = \sum_{n=0}^{\infty} e^{-\delta n} e^{2\pi i n \mathbf{k} t}\,. \tag{5.21}$$

Calculating the $\mathscr{A}_v$-norm of $1/f$, we find that

$$\Big\| \frac{1}{f} \Big\|_{\mathscr{A}_v} = \sum_{n=0}^{\infty} e^{-\delta n} v(n\mathbf{k}) \geq \sum_{n=n_0}^{\infty} e^{-\delta n} e^{an} = \infty\,.$$

Thus, $1/f$ does not belong to $\mathscr{A}_v$, and Wiener's Lemma does not hold in $\mathscr{A}_v$. □

The GRS condition is ubiquitous in the investigation of inverse-closed subalgebras with weights. This condition draws the fine line between exponential and subexponential growth. Exponential growth is special and usually implies the existence of some analytic structure.

What happens in the case of exponential weights? If $v_a(k) = e^{a|k|}$ is an exponential weight with growth constant a, then $\mathscr{A}_{v_a}(\mathbb{T}^d)$ is still a Banach algebra with respect to pointwise multiplication (Lemma 5.22). One can show the following weak version of Wiener's Lemma for exponential weights. If $f \in \mathscr{A}_{v_a}(\mathbb{T})$ and $f(t) \neq 0$ for all $t \in \mathbb{T}^d$, then there exists $\delta > 0$, $\delta \leq a$, such that $1/f \in \mathscr{A}_{v_\delta}(\mathbb{T}^d)$. Thus, the Fourier coefficients of the inverse still have exponential decay, but the growth constant δ may be arbitrarily small and depends on f. This is also shown in our counterexample, especially (5.21).

Proof (of Sublemma 5.25). Set $\tilde{v}(n) = \max_{|k|_\infty \leq n} v(k)$. Then the weight $\tilde{v}$ is submultiplicative and increasing on $\mathbb{N}$. If v satisfies the GRS condition, then $\tilde{v}$ also satisfies the GRS condition.

The submultiplicativity and monotonicity are clear; we only show the GRS condition for $\tilde{v}$. Let $\mathbf{e}_j, j = 1, \ldots, d$, be the standard basis for $\mathbb{R}^d$ and let $v_j(l) = v(l\mathbf{e}_j)$ be the restriction of v to the subgroup $\{0\} \times \cdots \times \{0\} \times \mathbb{Z} \times \{0\} \times \cdots \times \{0\}$. Then

$$v(\mathbf{k}) = v\left(\sum_{j=1}^{d} k_j \mathbf{e}_j\right) \leq \prod_{j=1}^{d} v_j(k_j)\,.$$

Now assume that the GRS condition is not satisfied for $\tilde{v}$; then for some $N \in \mathbb{N}$ and $a > 0$ we have $\tilde{v}(nN) \geq e^{an}$ for n large enough. Consequently, there exists a sequence $\mathbf{k}_n \in \mathbb{Z}^d$ such that $|\mathbf{k}_n|_\infty \leq nN$ and

$$v(\mathbf{k}_n) = \tilde{v}(nN) \geq e^{an}\,.$$

Since $e^{an} \leq v(\mathbf{k}_n) \leq \prod_{j=1}^{d} v_j(\mathbf{k}_{n,j})$, there exist a coordinate j_0 and a subsequence n_r of $\mathbb{N}$ such that, with $\ell_r = |\mathbf{k}_{n_r,j_0}|$,

$$e^{an_r/d} \leq v_{j_0}(\ell_r) \quad \text{and} \quad |\ell_r| \leq n_r N\,.$$

For this subsequence we obtain

$$v(\ell_r \mathbf{e}_{j_0}) = v_{j_0}(\ell_r) \geq e^{an_r/d} \geq e^{a\ell_r/(dN)}$$

and

$$\lim_{r\to\infty} v(\ell_r \mathbf{e}_{j_0})^{1/\ell_r} \geq e^{a/(dN)} > 1\,,$$

in contradiction to the assumed GRS condition of v. □

5.3.2 Matrix Algebras

In Section 5.2.6 we argued that a discrete time-invariant channel is modeled by a convolution operator and discussed the relevance of Wiener's Lemma for the analysis of the input–output relationship. In this section we study time-varying channels and the corresponding mathematical model, namely matrix algebras. We present a version of Wiener's Lemma in this context.

5.3.2.1 Discrete Time-Varying Channels

Recall that a time-invariant system corresponds to a convolution operator

$$(C_{\mathbf{h}}\mathbf{a})(k) = \sum_{k\in\mathbb{Z}} h(k-l)a(l).$$

The infinite matrix M corresponding to this linear operator has the entries $m_{kl} = h(k-l)$. The system matrix is thus constant along diagonals; such a matrix is usually called a *Toeplitz matrix.*

When we deal with time-varying systems, the corresponding matrix will no longer be constant along diagonals, but for a slowly time-varying system, it will have small variations along the diagonals and will still be close to a convolution operator.

If M is a matrix over the index set $\mathbb{Z}^d$ with entries $m_{kl}, k,l \in \mathbb{Z}^d$, then the lth diagonal has the entries $m_{k,k-l}$. We may write the matrix-vector multiplication in a way that resembles a convolution, namely,

$$(M\mathbf{a})(k) = \sum_{l\in\mathbb{Z}^d} m_{kl}a_l = \sum_{l\in\mathbb{Z}^d} m_{k,k-l}a_{k-l}\,.$$

If M is a Toeplitz matrix, then this is a convolution. If M is "almost constant" along diagonals, the action of M resembles a convolution.

This observation motivated Gohberg, Kaashoek, and Woerdeman [104] to introduce a *nonstationary Wiener algebra*. It is defined as the class of all matrices for which the norm

$$\|M\|_{\mathscr{C}} = \sum_{l\in\mathbb{Z}^d} \sup_{k\in\mathbb{Z}^d} |m_{k,k-l}|$$

is finite. This class of matrices was studied simultaneously and independently by Baskakov [14] and Kurbatov [155], and was later rediscovered by Sjöstrand [214]. It is often named after the inventors as the Baskakov–Gohberg–Sjöstrand matrix algebra. This class of matrices has recently appeared in several applications in frame theory [10, 94, 109, 119] and in the analysis of the finite section method in numerical analysis [121, 198].

Let us look in more detail at the class $\mathscr{C}$. The number

$$d(l) = \sup_{k\in\mathbb{Z}^d} |m_{k,k-l}|$$

is the supremum of the lth diagonal of M and so

$$|m_{kl}| \le d(k-l), \qquad k,l \in \mathbb{Z}^d\,. \tag{5.22}$$

Hence, the action of M on a vector $\mathbf{c}$ can be estimated as follows:

$$|(M\mathbf{c})(k)| = \left|\sum_{l\in\mathbb{Z}^d} m_{kl}c_l\right| \le \sum_{l\in\mathbb{Z}^d} d(k-l)|c_l| = (\mathbf{d}*|\mathbf{c}|)(k)\,. \tag{5.23}$$

This inequality says that the action of M is dominated (pointwise) by the convolution with the sequence $\mathbf{d}\in\ell^1(\mathbb{Z}^d)$. This explains our notation; $\mathscr{C}$ is the class of *convolution-dominated matrices*. Young's inequality (Lemma 5.17) implies that a matrix $M\in\mathscr{C}$ is bounded on every $\ell^p(\mathbb{Z}^d)$, $1\le p\le\infty$.

Next let us consider a weighted version of convolution-dominated matrices. Let v be a submultiplicative weight on $\mathbb{Z}^d$. Then $\mathscr{C}_v$ contains all matrices for which the norm

$$\|M\|_{\mathscr{C}_v} = \sum_{l\in\mathbb{Z}^d}\sup_{k\in\mathbb{Z}^d}|m_{k,k-l}|\,v(l)$$

is finite. As above, this means that M is dominated by a convolution with an ℓ^1_v-sequence. In particular, the matrix of a convolution operator $C_{\mathbf{h}}$ with symbol $\mathbf{h}\in\ell^1_v(\mathbb{Z}^d)$ belongs to $\mathscr{C}_v$. Moreover, identifying the operator with its matrix, we have

$$\|C_{\mathbf{h}}\|_{\mathscr{C}_v} = \|\mathbf{h}\|_{\ell^1_v}\,.$$

Lemma 5.28. *If v is submultiplicative and symmetric, then $\mathscr{C}_v$ is a Banach $*$-algebra with respect to matrix multiplication and taking the adjoint matrix as the involution.*

Proof. Again, the proof is similar to the one of Lemmas 5.4 and 5.22 and in fact makes direct use of Lemma 5.22. Let $M,N\in\mathscr{C}_v$ and set $d(l)=\sup_{k\in\mathbb{Z}^d}|m_{k,k-l}|$ and $e(l)=\sup_{k\in\mathbb{Z}^d}|n_{k,k-l}|$. Then by (5.22)

$$\begin{aligned}|(MN)_{k,k-l}| &= \Big|\sum_{r\in\mathbb{Z}^d} m_{kr}n_{r,k-l}\Big| \\ &\le \sum_{r\in\mathbb{Z}^d} d(k-r)\,e(r-k+l) \;=\; (\mathbf{d}*\mathbf{e})(l)\end{aligned}$$

and

$$\begin{aligned}\|MN\|_{\mathscr{C}_v} &= \sum_{l\in\mathbb{Z}^d}\sup_{k\in\mathbb{Z}^d}|(MN)_{k,k-l}|\,v(l) \\ &\le \sum_{l\in\mathbb{Z}^d}(\mathbf{d}*\mathbf{e})(l)\,v(l) = \|\mathbf{d}*\mathbf{e}\|_{\ell^1_v} \\ &\le \|\mathbf{d}\|_{\ell^1_v}\,\|\mathbf{e}\|_{\ell^1_v} = \|M\|_{\mathscr{C}_v}\,\|N\|_{\mathscr{C}_v}\,.\end{aligned}$$

For the involution $M\to M^*$ we find

$$\begin{aligned}\|M^*\|_{\mathscr{C}_v} &= \sum_{l\in\mathbb{Z}^d}\sup_{k\in\mathbb{Z}^d}|(M^*)_{k,k-l}|\,v(l) \;=\; \sum_{l\in\mathbb{Z}^d}\sup_{k\in\mathbb{Z}^d}|m_{k-l,k}|\,v(l) \\ &= \sum_{l\in\mathbb{Z}^d}\sup_{k\in\mathbb{Z}^d}|m_{k,k+l}|\,v(-l) = \|M\|_{\mathscr{C}_v}\,.\end{aligned}$$

So the involution is an isometry. □

5.3.2.2 A Nonstationary Version of Wiener's Lemma

For convolution-dominated matrices we may now pose the same questions as for convolution operators in Section 5.2.6. The question is now whether the inverse of a convolution-dominated matrix is again convolution-dominated. In the context of time-varying systems we are interested in a meaningful input–output relationship. So if M is convolution-dominated and the output $\mathbf{y} = M\mathbf{c}$ is in ℓ^p, is it true that the input $\mathbf{c}$ was also in ℓ^p? As for convolution operators in Section 5.2.6, a satisfactory answer requires the independence of the spectrum of M of the domain space $\ell^p(\mathbb{Z}^d)$.

After the discussion of several versions of Wiener's Lemma, it is perhaps no longer surprising that the answers to these questions will be the same as for time-invariant systems (convolution operators). The techniques and proofs, however, are significantly more involved, because the matrix algebra $\mathscr{C}$ is highly noncommutative.

Theorem 5.29. *If $M \in \mathscr{C}$ and M is invertible on $\ell^2(\mathbb{Z}^d)$, then $M^{-1} \in \mathscr{C}$.*

This result was obtained independently and almost simultaneously by Gohberg, Kaashoek, and Woerdemann [104], by Baskakov [14] and Kurbatov [155], and a little later by Sjöstrand [214] with a completely different proof.

As with the classical Wiener's Lemma, we next consider a variation of Theorem 5.29 with weights. The weighted version is due to Baskakov, who studied convolution-dominated operators on Banach spaces with unconditional (block) bases. The following theorem is often referred to as Baskakov's theorem [14].

Theorem 5.30 (Baskakov [14]). *Assume that v satisfies the GRS condition and $1 \leq p \leq \infty$. If $M \in \mathscr{C}_v$ and M is invertible on $\ell^p(\mathbb{Z}^d)$, then $M^{-1} \in \mathscr{C}_v$. In other words, $\mathscr{C}_v$ is inverse-closed in $\mathscr{B}(\ell^p)$ for all $p \in [1, \infty]$.*

We cannot give the complete proof of this theorem, but will sketch the proof idea at the end of this section. In particular, we will elaborate the relationship of Theorems 5.29 and 5.30 with the classical Wiener's Lemma.

As with convolution operators, the GRS condition characterizes those weights for which Wiener's Lemma holds.

Corollary 5.31. *The algebra $\mathscr{C}_v$ is inverse-closed in $\mathscr{B}(\ell^2)$ if and only if v satisfies the GRS condition.*

This is just a reformulation of Theorem 5.30. The hard part is to show that the GRS condition is sufficient for the inverse-closedness of $\mathscr{C}_v$.

To verify the necessity of the GRS condition, we return to the example in the proof of Corollary 5.27. If v violates the GRS conditions, then there are $\mathbf{k} \in \mathbb{Z}^d$ and $a > 0$ such that $v(n\mathbf{k}) \geq e^{an}$ for $n \geq n_0$.

We construct an invertible convolution operator $C_{\mathbf{h}}$ with $\mathbf{h} \in \ell^1_v$ with $C_{\mathbf{h}}^{-1} \notin \mathscr{C}_v$. Set $h(0) = 1$ and $h(\mathbf{k}) = e^{-\delta}$ and $h(l) = 0$ for $l \neq 0$ and $l \neq k$ and consider the convolution operator $C_{\mathbf{h}}$. Since $\widehat{\mathbf{h}}(t) = 1 - e^{-\delta} e^{2\pi i \mathbf{k} \cdot t} \neq 0$ for all $t \in \mathbb{T}^d$, the operator $C_{\mathbf{h}}$ is invertible on $\ell^2(\mathbb{Z}^d)$ by Corollary 5.20. Its inverse is the operator $C_{\mathbf{g}} = C_{\mathbf{h}}^{-1}$

with $\widehat{\mathbf{g}}(t) = \widehat{\mathbf{h}}(t)^{-1} = \sum_{n=0}^{\infty} e^{-\delta n} e^{2\pi i n \mathbf{k}\cdot t}$. Since $\mathbf{g} \notin \ell_v^1(\mathbb{Z}^d)$, the matrix of $C_{\mathbf{g}}$ is not in $\mathscr{C}_v$. So $\mathscr{C}_v$ is not inverse-closed in $\mathscr{B}(\ell^2)$.

Further variations on convolution-dominated matrices, including a complete characterization of inverse-closed Banach algebras of convolution dominated matrices, can be found in [120]. For a recent extension to noncommutative groups as index sets, see [90].

5.3.2.3 Spectral Invariance

We next formulate Baskakov's theorem as a statement about the spectral invariance of matrices.

We first note that, like convolution operators, a matrix in $\mathscr{C}_v$ is not only bounded on $\ell^2(\mathbb{Z}^d)$, but acts on a whole class of weighted ℓ^p-spaces (and other spaces as well). For this we introduce another class of weight functions. We say that a nonnegative function m on $\mathbb{Z}^d$ is v-moderate if

$$m(k+l) \le Cv(k)m(l), \qquad \text{for all } k,l \in \mathbb{Z}^d\,. \tag{5.24}$$

A weight is called *moderate* if it is v-moderate with respect to some submultiplicative weight v.

Let the weighted ℓ^p-space $\ell^p_m(\mathbb{Z}^d)$ be defined by the norm $\|\mathbf{c}\|_{\ell^p_m} = \|\mathbf{c}m\|_p$.

The relevance of moderate weight functions is explained by the next lemma.

Lemma 5.32. *Let v be a submultiplicative weight on $\mathbb{Z}^d$.*

1. *Then ℓ^p_m is invariant under all translations $T_k, k \in \mathbb{Z}^d$, if and only if m is moderate.*
2. *If m is v-moderate, then $\ell^1_v * \ell^p_m \subseteq \ell^p_m$; i.e., if $\mathbf{a} \in \ell^1_v$ and $\mathbf{b} \in \ell^p_m$, then $\mathbf{a} * \mathbf{b} \in \ell^p_m$ with the convolution estimate*

$$\|\mathbf{a} * \mathbf{b}\|_{\ell^p_m} \le C\|\mathbf{a}\|_{\ell^1_v}\,\|\mathbf{b}\|_{\ell^p_m}\,. \tag{5.25}$$

3. *If $M \in \mathscr{C}_v$ and m is v-moderate, then M is bounded on every ℓ^p_m for $1 \le p \le \infty$, and*

$$\|M\|_{\ell^p_m \to \ell^p_m} \le C\|M\|_{\mathscr{C}_v}\,,$$

where C is the constant in (5.24)*.*

Proof. Items (1) and (2) are elementary and left to the reader (see Exercises). In fact, (2) is just a modified version of Young's inequality (Lemma 5.17).

Since $|(M\mathbf{c})(k)| \le (\mathbf{d} * |\mathbf{c}|)(k)$ by (5.23), the weighted Young inequality (5.25) yields

$$\|M\mathbf{c}\|_{\ell^p_m} \le C\,\|\mathbf{d}\|_{\ell^1_v}\,\|\mathbf{c}\|_{\ell^p_m} = C\,\|M\|_{\mathscr{C}_v}\,\|\mathbf{c}\|_{\ell^p_m}\,. \quad \square$$

We may reformulate (3) by saying that the matrix algebra $\mathscr{C}_v$ is continuously embedded in the algebra $\mathscr{B}(\ell^p_m)$ of bounded operators on ℓ^p_m. Given a matrix M that is bounded on ℓ^p_m, we denote its spectrum as an operator on ℓ^p_m by $\sigma_{\mathscr{B}(\ell^p_m)}(M)$.

We can now formulate the spectral invariance for convolution-dominated matrices.

Corollary 5.33. *Assume that v satisfies the GRS condition and $M \in \mathscr{C}_v$. Then the spectral invariance*

$$\sigma_{\mathscr{B}(\ell^p_m)}(M) = \sigma_{\mathscr{B}(\ell^2)}(M) = \sigma_{\mathscr{C}_v}(M)$$

holds for every $p \in [1,\infty]$ and every v-moderate weight m.

Proof. The statement follows from a weighted version of Baskakov's theorem (Theorem 5.30). In fact, Baskakov's general result implies that $\mathscr{C}_v$ is inverse-closed in $\mathscr{B}(\ell^p_m)$ for $1 \le p < \infty$ and every v-moderate weight m. Now Lemma 5.10 implies that $\sigma_{\mathscr{B}(\ell^p_m)}(M) = \sigma_{\mathscr{C}_v}(M)$ for all $M \in \mathscr{C}_v$. For $p = \infty$ we use duality and find that $\sigma_{\mathscr{B}(\ell^\infty_m)}(M) = \sigma_{\mathscr{B}(\ell^1_{1/m})}(M^*) = \sigma_{\mathscr{C}_v}(M^*) = \sigma_{\mathscr{C}_v}(M)$. Thus, the spectrum is independent of the domain space ℓ^p_m. □

Corollary 5.33 says that the spectrum of a matrix M is independent of the domain space provided that the matrix has sufficient off-diagonal decay. We may also state the corollary in the style of Theorem 5.19 as follows: *A matrix $M \in \mathscr{C}_v$ is invertible on $\ell^2(\mathbb{Z}^d)$ if and only if M is invertible on $\ell^p_m(\mathbb{Z}^d)$ for some $p \in [1,\infty]$, and some v-moderate weight m if and only if M is invertible on all $\ell^p_m(\mathbb{Z}^d)$ for all $p \in [1,\infty]$ and all v-moderate weights m.*

This result is analogous to Corollary 5.20 for convolution operators. However, whereas convolution operators on $\mathbb{Z}^d$ form a commutative algebra of operators, $\mathscr{C}_v$ is a highly noncommutative matrix algebra.

For some recent results in the line of Baskakov's theorem, see [2, 211].

5.3.2.4 The Idea of the Proof of Baskakov's Theorem

Although we cannot give a complete proof of Baskakov's theorem, we will indicate the main ideas. Our goal is to show how the classical version of Wiener's Lemma for absolutely convergent Fourier series enters the field of matrix algebras.

It is tempting to imitate the proof of the classical Wiener's Lemma (Theorem 5.5) in the noncommutative setting. Though many of the steps carry over, the proof cannot be saved because the binomial theorem is not applicable in noncommutative algebras.

Following deLeeuw [63], we first associate to every matrix A a matrix-valued function. Define *modulation* M_t, $t \in \mathbb{R}^d$, acting on a sequence $\mathbf{c}$ by

$$(M_t\mathbf{c})(k) = e^{2\pi i k\cdot t} c(k) \quad \text{for } k \in \mathbb{Z}^d .$$

Given a matrix A, we next consider the matrix-valued function

$$\mathbf{f}(t) = M_t A M_{-t} .$$

Clearly, M_tAM_{-t} is periodic with period 1 in each coordinate of $t = (t_1,\dots,t_d) \in \mathbb{R}^d$ and is a continuous matrix-valued function on $\mathbb{T}^d$. Instinctively, the first question to ask is, what are the Fourier coefficients of $\mathbf{f}$?

Lemma 5.34 (Fourier coefficients of f). *The nth Fourier coefficient is the nth side diagonal of A. Precisely, let D_n be the matrix with entries $(D_n)_{k,k-n} = a_{k,k-n}$ and $(D_n)_{kl} = 0$ for $l \neq k-n$, where $k,l,n \in \mathbb{Z}^d$. Then*

$$\int_{[0,1]^d} \mathbf{f}(t) e^{-2\pi i n\cdot t}\, dt = D_n\,. \tag{5.26}$$

Proof. The integral in (5.26) is a matrix-valued integral; we interpret it entrywise. First note that

$$\begin{aligned}\big(M_tAM_{-t}\mathbf{c}\big)(k) &= e^{2\pi i k\cdot t} \sum_{l\in\mathbb{Z}^d} a_{kl} e^{-2\pi i l\cdot t} c_l \\ &= \sum_{l\in\mathbb{Z}^d} a_{kl} e^{2\pi i (k-l)\cdot t} c_l\,.\end{aligned}$$

The matrix $\mathbf{f}(t) = M_tAM_{-t}$ has the entries $a_{kl}e^{2\pi i(k-l)\cdot t}$. Therefore, the nth Fourier coefficient of the (k,l)th entry is

$$\begin{aligned}\widehat{\mathbf{f}}(n)_{kl} &= \int_{[0,1]^d} \mathbf{f}(t)_{kl} e^{-2\pi i n\cdot t} dt \\ &= \int_{[0,1]^d} a_{kl} e^{2\pi i (k-l)\cdot t} e^{-2\pi i n\cdot t} dt \\ &= a_{kl}\,\delta_{k-l-n} = a_{k,k-n}\,\delta_{k-l-n}\,,\end{aligned}$$

and so $\widehat{\mathbf{f}}(n) = D_n$. □

Therefore, the matrix-valued function $\mathbf{f}(t)$ possesses the formal Fourier series $\mathbf{f}(t) = \sum_{n\in\mathbb{Z}^d} D_n e^{2\pi i k\cdot t}$. In particular, for $t = 0$ we recover $A = \sum_{n\in\mathbb{Z}^d} D_n$ as a sum of its diagonals. As always with Fourier series, we must be cautious in which sense the Fourier series represents the given function.

This question is easy to answer for convolution-dominated matrices. Recall that

$$\begin{aligned}\|A\|_{\mathscr{C}} &= \sum_{n\in\mathbb{Z}} \sup_{k\in\mathbb{Z}} |m_{k,k-n}| \\ &= \sum_{n\in\mathbb{Z}} \|D_n\|_{\ell^p\to\ell^p}\,.\end{aligned}$$

Thus, the Fourier series of $\mathbf{f}$ converges in the operator norm, and $\mathbf{f}$ possesses an absolutely convergent Fourier series. The difference is that the coefficients of $\mathbf{f}$ are infinite matrices (or operators).

Let us go a step further and introduce the space of absolutely convergent matrix-valued Fourier series $\mathscr{A}_v(\mathbb{T}^d, \mathscr{B}(\ell^p))$. A matrix-valued function $\mathbf{f}$ belongs to

$\mathscr{A}_v(\mathbb{T}^d, \mathscr{B}(\ell^p))$ if it has a Fourier series $\mathbf{f}(t) = \sum_{k\in\mathbb{Z}^d} A_k e^{2\pi i k\cdot t}$ with coefficients $A_k \in \mathscr{B}(\ell^p)$ satisfying

$$\|\mathbf{f}\|_{\mathscr{A}_v(\mathbb{T}^d,\mathscr{B}(\ell^p))} := \sum_{k\in\mathbb{Z}^d} \|A_k\|_{\ell^p\to\ell^p}\, v(k) < \infty .$$

Using this notation, we quickly obtain the following.

Lemma 5.35. *1. A matrix A belongs to $\mathscr{C}_v$ if and only if the matrix-valued function $\mathbf{f}$ belongs to $\mathscr{A}_v(\mathbb{T}^d, \mathscr{B}(\ell^p))$. In this case,*

$$\|A\|_{\mathscr{C}_v} = \|\mathbf{f}\|_{\mathscr{A}_v(\mathbb{T}^d,\mathscr{B}(\ell^p))} \,.$$

2. Assume that Wiener's Lemma holds for $\mathscr{A}_v(\mathbb{T}^d, \mathscr{B}(\ell^p))$. [This means that if $\mathbf{f} \in \mathscr{A}_v(\mathbb{T}^d, \mathscr{B}(\ell^p))$ and $\mathbf{f}(t)$ is invertible on $\ell^p(\mathbb{Z}^d)$ for all $t \in \mathbb{T}^d$, then $\mathbf{f}^{-1}$ is in $\mathscr{A}_v(\mathbb{T}^d, \mathscr{B}(\ell^p))$.] Then $\mathscr{C}_v$ is inverse-closed in $\mathscr{B}(\ell^p)$.

Proof. (1) follows directly from the definitions.

(2): If $A \in \mathscr{C}_v \subseteq \mathscr{B}(\ell^p)$ is invertible, then the operator-valued function associated to its inverse is just $M_t A^{-1} M_{-t} = \mathbf{f}(t)^{-1}$, and $\mathbf{f}(t)$ is invertible on $\ell^p(\mathbb{Z}^d)$ for all $t \in \mathbb{T}^d$. Wiener's Lemma for $\mathscr{A}_v(\mathbb{T}^d, \mathscr{B}(\ell^p))$ implies that $\mathbf{f}^{-1} \in \mathscr{A}_v(\mathbb{T}^d, \mathscr{B}(\ell^p))$. Consequently, by (1), $A^{-1} \in \mathscr{C}_v$. □

Lemma 5.35 establishes a surprisingly direct connection between the topic of matrix algebras and the classical formulation of Wiener's Lemma. To understand the inverse of a convolution-dominated matrix $A \in \mathscr{C}_v$, we need a version of Wiener's Lemma for operator-valued Fourier series. Such a generalization of Wiener's Lemma exists indeed and was obtained by Bochner and Phillips [24] already in 1946. Though this generalization is "only" from scalar-valued functions to matrix-valued functions, it is highly nontrivial, because the algebra of operator-valued absolutely convergent Fourier series is noncommutative. Therefore, neither Gel'fand theory nor the elegant arguments of Newman and Hulanicki used in Section 5.2.4 can be applied. The proof of the operator-valued version of Wiener's Lemma requires several results from noncommutative Banach algebras and their representations. For the details we refer to the original sources [14, 24] or the appendix of [120].

Our main point was to reveal the connection of the nonstationary version of Wiener's Lemma to its classical version.

5.3.2.5 Off-Diagonal Decay of Matrices

So far we have discussed convolution-dominated matrices in analogy to absolutely convergent Fourier series. In applications, many other conditions are used to measure off-diagonal decay. We mention just a few of them.

(a) Strict off-diagonal decay is measured by the norm

$$\|M\|_{\mathscr{A}_v^\infty} = \sup_{k,l\in\mathbb{Z}^d} |m_{kl}|\, v(k-l) \,.$$

The weight function v measures the rate of off-diagonal decay. The typical weights are the polynomial weights $v(k) = (1+|k|)^s$ or a combination of polynomial and subexponential weights $v(k) = (1+|k|)^s e^{a|k|^b}$, $s > d$, $a \geq 0$, and $0 \leq b < 1$.

(b) Schur-type conditions: In imitation of Schur's test, which involves the column sums and row sums of a matrix, we may also study the norm

$$\|M\|_{\mathscr{A}_v^1} = \max\Big\{ \sup_{l\in\mathbb{Z}^d} \sum_{k\in\mathbb{Z}^d} |m_{kl}|v(k-l), \sup_{k\in\mathbb{Z}^d} \sum_{l\in\mathbb{Z}^d} |m_{kl}|v(k-l) \Big\}. \tag{5.27}$$

If v is submultiplicative on $\mathbb{Z}^d$, then $\mathscr{A}_v^1$ is a Banach algebra with respect to matrix multiplication (see Exercises).

We note that the matrix algebras defined by off-diagonal conditions obey the following inclusion relations:

$$\mathscr{C}_v \subset \mathscr{A}_v^1 \subset \mathscr{A}_v^\infty . \tag{5.28}$$

As in the case of convolution-dominated matrices, the off-diagonal decay is preserved under suitable conditions on the weight involved. We quote a typical theorem from [119] and refer also to the work of Baskakov [14–16] and Sun [221].

Theorem 5.36. *1. Assume that $v^{-1} \in \ell^1(\mathbb{Z})$, $v^{-1} * v^{-1} \leq Cv^{-1}$, and v satisfies the GRS condition. Then $\mathscr{A}_v^\infty$ is inverse-closed in $\mathscr{B}(\ell^2)$.*
2. Assume that v is submultiplicative, $v(k) \geq (1+|k|)^\delta$ for some $\delta > 0$, and v satisfies the GRS condition. Then $\mathscr{A}_v^1$ is inverse-closed in $\mathscr{B}(\ell^2)$.

Again, the GRS condition is necessary and sufficient for the validity of Theorem 5.36. The necessity follows from the counterexample constructed after Corollary 5.31.

5.3.3 Absolutely Convergent Series of Time-Frequency Shifts

We now turn to time-frequency analysis and noncommutative geometry.

5.3.3.1 The Basic Definitions

Recall the definition of time-frequency shifts. Given $x, \xi \in \mathbb{R}^d$, the translation operator or time shift T_x and the modulation or frequency shift M_ξ act on a function f on $\mathbb{R}^d$ by

$$T_x f(t) = f(t-x) \quad \text{and} \quad M_\xi f(t) = e^{2\pi i \xi \cdot t} f(t), \quad x, \xi, t \in \mathbb{R}^d .$$

Combining the parameter x and ξ into a single point $z = (x, \xi) \in \mathbb{R}^{2d}$ in the time-frequency "plane," their composition is the time-frequency shift

$$\pi(z) f(t) = M_\xi T_x f(t) = e^{2\pi i \xi \cdot t} f(t-x), \quad t \in \mathbb{R}^d . \tag{5.29}$$

The time-frequency shifts $\pi(z)$, $z \in \mathbb{R}^{2d}$, are unitary operators on $L^2(\mathbb{R}^d)$ and isometries on $L^p(\mathbb{R}^d)$ for $p \neq 2$. The translations and modulations satisfy the commutation relations

$$T_x M_\xi = e^{-2\pi i x\cdot\xi} M_\xi T_x\,.$$

As a consequence the composition of two time-frequency shifts $\pi(z)$ and $\pi(w)$ with $w = (w_1, w_2) \in \mathbb{R}^{2d}$ and $z = (z_1, z_2) \in \mathbb{R}^{2d}$ is

$$\begin{aligned}\pi(w)\pi(z) &= (M_{w_2} T_{w_1})\,(M_{z_2} T_{z_1}) \\ &= e^{-2\pi i w_1\cdot z_2} M_{w_2+z_2} T_{w_1+z_1} = e^{-2\pi i w_1\cdot z_2}\pi(w+z)\,.\end{aligned} \tag{5.30}$$

Thus, the composition of time-frequency shifts is a multiple of a time-frequency shift. However, since the $\pi(z), z \in \mathbb{R}^{2d}$, do not commute, the mathematics of time-frequency shifts always leads to noncommutative structures.

Next we fix a lattice Λ in $\mathbb{R}^{2d}$. A lattice is a discrete subgroup $\Lambda \subseteq \mathbb{R}^{2d}$ with compact quotient $\mathbb{R}^{2d}/\Lambda$. Choosing a basis $\mathbf{a}_j, j = 1,\ldots,2d$, we can write every $\lambda \in \Lambda$ as $\lambda = \sum_{j=1}^{2d} k_j \mathbf{a}_j$ with integer coefficients $k_j \in \mathbb{Z}$. Consequently, every lattice can be represented as $\Lambda = A\mathbb{Z}^{2d}$ for some invertible $2d \times 2d$-matrix A, the columns of which are just the basis vectors $\mathbf{a}_j$.

5.3.3.2 The Rotation Algebra

In time-frequency analysis and in noncommutative geometry one considers formal sums of time-frequency shifts on a lattice Λ, i.e., operators of the form

$$A = \sum_{\lambda\in\Lambda} c_\lambda \pi(\lambda)\,.$$

We should think of sums of time-frequency shifts as a noncommutative analogue of Fourier series. The complex exponentials $e^{2\pi i k\cdot t}$ are replaced by the unitary operators $\pi(\lambda)$. Whereas a Fourier series is a function on the torus $\mathbb{T}^d$, a sum of time-frequency shifts on a lattice Λ yields an operator. The structure and properties of function spaces (and algebras) on the torus $\mathbb{T}^d$ completely describe the (topological) properties of $\mathbb{T}^d$ (this is the content of the theorem of Gel'fand–Naimark [27, 56, 150, 209]). By analogy the sums of time-frequency shifts are interpreted as "functions" on some "exotic" structure. Since by (5.30) the time-frequency shifts $\pi(\lambda)$ do not commute for a general lattice, this structure is taken to be a *noncommutative torus*.

Keeping the analogy between Fourier series and time-frequency shifts in mind, it is now time to make some precise definitions.

Definition 5.37. Let $\mathscr{A}_0(\Lambda)$ be the vector space of all finite linear combinations of time-frequency shifts $\pi(\lambda)$. The *rotation algebra* or *noncommutative torus* $C^*(\Lambda)$ is the closure of $\mathscr{A}_0(\Lambda)$ in the operator norm on $L^2(\mathbb{R}^d)$.

By definition $C^*(\Lambda)$ is a closed subspace of $\mathscr{B}(L^2(\mathbb{R}^d))$, the algebra of all bounded operators on $L^2(\mathbb{R}^d)$. An operator $A \in \mathscr{B}(L^2)$ belongs to $C^*(\Lambda)$ if and only if for every $\varepsilon > 0$ there exists a finite linear combination $P = \sum_{\lambda\in\Lambda} c_\lambda \pi(\lambda) \in \mathscr{A}_0(\Lambda)$ with $\operatorname{supp}\mathbf{c}$ finite, such that $\|A-P\|_{L^2\to L^2} < \varepsilon$.

Let us collect some basic properties of series of time-frequency shifts to ensure that such series are well defined.

Lemma 5.38. *1. If the sequence $\mathbf{c} = (c_\lambda)_{\lambda\in\Lambda}$ grows polynomially, $|c_\lambda| = \mathscr{O}(1+|\lambda|^N)$ for some $N \geq 0$, then the sum $\sum_{\lambda\in\Lambda} c_\lambda \pi(\lambda)$ is a well-defined continuous operator from the Schwartz class $\mathscr{S}(\mathbb{R}^d)$ to $\mathscr{S}'(\mathbb{R}^d)$.*

2. Strong linear independence: If $\sum_{\lambda\in\Lambda} c_\lambda \pi(\lambda) = 0$ for a sequence $\mathbf{c}$ of polynomial growth, then $\mathbf{c} = 0$.

Proof. 1. We use the following property of time-frequency shifts ([108, Thm. 11.2.5] or [92]): If $f,g \in \mathscr{S}(\mathbb{R}^d)$, then the function $z \in \mathbb{R}^{2d} \to \langle \pi(z)f, g\rangle$ belongs to $\mathscr{S}(\mathbb{R}^{2d})$ and depends continuously on f and g. In particular, for all $M \geq 0$,

$$|\langle \pi(\lambda)f, g\rangle| = \mathscr{O}\left((1+|\lambda|)^{-M}\right).$$

Consequently, $|\langle \sum_{\lambda\in\Lambda} c_\lambda \pi(\lambda)f, g\rangle| \leq \sum_{\lambda\in\Lambda} |c_\lambda| |\langle \pi(\lambda)f,g\rangle|$ is well defined and $\sum_{\lambda\in\Lambda} c_\lambda \pi(\lambda)$ makes sense as an operator from $\mathscr{S}(\mathbb{R}^d)$ to $\mathscr{S}'(\mathbb{R}^d)$.

2. The proof of the linear independence is taken from [113]. By assumption we have, for all $g,h \in \mathscr{S}(\mathbb{R}^d)$ and $z \in \mathbb{R}^{2d}$,

$$\sum_{\lambda\in\Lambda} c_\lambda \langle \pi(\lambda)\pi(z)g, \pi(z)h\rangle = 0.$$

Now $\pi(z)^{-1}\pi(\lambda)\pi(z) = e^{2\pi i[z,\lambda]}\pi(\lambda)$, where $[z,\lambda] = z_1\cdot\lambda_2 - z_2\cdot\lambda_1$ is the symplectic form on $\mathbb{R}^{2d}$. This implies that

$$\sum_{\lambda\in\Lambda} c_\lambda \langle \pi(\lambda)g, h\rangle\, e^{2\pi i(\lambda_2\cdot z_1 - \lambda_1\cdot z_2)} = 0 \tag{5.31}$$

for all $z \in \mathbb{R}^{2d}$ and all $g,h \in \mathscr{S}(\mathbb{R}^d)$.

Equation (5.31) is an absolutely convergent Fourier series on $\mathbb{R}^{2d}/\Lambda$. Since it vanishes everywhere, we must have

$$c_\lambda \langle \pi(\lambda)g, h\rangle = 0 \qquad \text{for all } \lambda \in \Lambda,$$

from which we deduce that $c_\lambda = 0$ for all λ. □

Lemma 5.38 guarantees that formal series of time-frequency shifts with polynomially growing coefficients are always well defined in a distributional sense. In particular, every $A \in C^*(\Lambda)$ possesses a unique expansion $A = \sum_{\lambda\in\Lambda} a_\lambda \pi(\lambda)$. It can be shown that the coefficient sequence must be in $\ell^2(\Lambda)$. However, not every $\mathbf{c} \in \ell^2(\Lambda)$ defines an operator in $\mathscr{B}(L^2(\mathbb{R}^d)) \supseteq C^*(\Lambda)$. See [147] for details.

To pursue the analogy between Fourier series and series of time-frequency shifts further, we next study the composition of sums of time-frequency shifts. Let $A = \sum_{\lambda\in\Lambda} a_\lambda \pi(\lambda) \in \mathscr{A}_0(\Lambda)$ and $B = \sum_{\mu\in\Lambda} b_\mu \pi(\mu) \in \mathscr{A}_0(\Lambda)$ be two finite sums of time-frequency shifts. We now mimic the calculation in Lemma 5.4 and see what we get. Using (5.30), we have

$$AB = \left(\sum_{\lambda\in\Lambda} a_\lambda \pi(\lambda)\right)\left(\sum_{\mu\in\Lambda} b_\mu \pi(\mu)\right)$$
$$= \sum_{\lambda,\mu\in\Lambda} a_\lambda b_\mu e^{-2\pi i\lambda_1\cdot\mu_2}\pi(\lambda+\mu) \tag{5.32}$$
$$= \sum_{\nu\in\Lambda}\left(\sum_{\lambda\in\Lambda} a_\lambda b_{\nu-\lambda}\, e^{-2\pi i\lambda_1\cdot(\nu_2-\lambda_2)}\right)\pi(\nu)\,.$$

Except for the phase factor $e^{-2\pi i\lambda_1\cdot(\nu_2-\lambda_2)}$, the coefficient sequence of AB looks like the convolution of the sequence **a** and **b**. Thus, we make the following definition.

Definition 5.39. The *twisted convolution* $\natural_\Lambda$ of two (finite) sequences **a** and **b** over Λ is defined by

$$(\mathbf{a}\,\natural_\Lambda\,\mathbf{b})(\nu) = \sum_{\lambda\in\Lambda} a_\lambda b_{\nu-\lambda} e^{-2\pi i\lambda_1\cdot(\nu_2-\lambda_2)}\,. \tag{5.33}$$

Remark 5.40. 1. If $\Lambda = \mathbb{Z}^{2d}$, then the phase factor disappears and $\mathbf{a}\,\natural_{\mathbb{Z}^{2d}}\,\mathbf{c} = \mathbf{a}*\mathbf{c}$ is just the ordinary convolution and thus commutative.
2. However, in general, the twisted convolution $\natural_\Lambda$ is not commutative.
3. By pulling in absolute values, we have

$$|(\mathbf{a}\natural_\Lambda\mathbf{b})(\nu)| \le \sum_{\lambda\in\Lambda}|a_\lambda|\,|b_{\nu-\lambda}| = (|\mathbf{a}|*|\mathbf{b}|)(\nu)$$

for all $\nu\in\Lambda$. We may therefore apply Young's inequality (Lemma 5.17) and obtain

$$\|\mathbf{a}\natural_\Lambda\mathbf{b}\|_p \le \|\mathbf{a}\|_1\,\|\mathbf{b}\|_p \tag{5.34}$$

whenever $\mathbf{a}\in\ell^1(\Lambda)$ and $\mathbf{b}\in\ell^p(\Lambda)$, $1\le p\le\infty$. Thus, convolution inequalities for the ordinary convolution imply immediately analogous inequalities for the twisted convolution, and $\natural_\Lambda$ is well defined on many sequence spaces.

Proposition 5.41. *The subspace $C^*(\Lambda)$ is a C^*-subalgebra of $\mathscr{B}(L^2(\mathbb{R}^d))$.*

Proof. If $A,B\in\mathscr{A}_0(\Lambda)$, then by (5.32) AB is again a finite linear combination of time-frequency shifts and thus $AB\in\mathscr{A}_0(\Lambda)$.

Next note that $\pi(z)^* = (M_\xi T_x)^* = T_{-x}M_{-\xi} = e^{-2\pi i x\cdot\xi}M_{-\xi}T_{-x} = e^{-2\pi i x\cdot\xi}\pi(-z)$. If $A=\sum_\lambda a_\lambda\pi(\lambda)\in\mathscr{A}_0(\Lambda)$, then the adjoint operator A^* is

$$A^* = \sum_{\lambda\in\Lambda}\overline{a_\lambda}e^{-2\pi i\lambda_1\cdot\lambda_2}\pi(-\lambda)\in\mathscr{A}_0(\Lambda)\,.$$

As a consequence, $\mathscr{A}_0(\Lambda)$ is a $*$-subalgebra of $\mathscr{B}(L^2(\mathbb{R}^d))$. Hence, its closure in the operator norm $C^*(\Lambda)$ is a C^*-subalgebra of $\mathscr{B}(L^2(\mathbb{R}^d))$. □

We can now go back to Section 5.2 and imitate all the constructions about Fourier series and repeat our previous questions in the context of the rotation algebras $C^*(\Lambda)$.

Let us first introduce some interesting subalgebras of $C^*(\Lambda)$ that are easier to work with than the full C^*-algebra.

To avoid convergence questions, one often resorts to the algebra of absolutely convergent series of time-frequency shifts:

$$\mathscr{A}(\Lambda) = \left\{ A \in \mathscr{B}(L^2(\mathbb{R}^d)) : A = \sum_{\lambda \in \Lambda} a_\lambda \pi(\lambda),\, \mathbf{a} \in \ell^1(\Lambda) \right\}.$$

If $A \in \mathscr{A}(\Lambda)$, the sum of time-frequency shifts converges absolutely in the operator norm on $L^2(\mathbb{R}^d)$. Note that this in complete analogy with the procedure of Fourier series.

Since the coefficients of an operator in $C^*(\Lambda)$ are unique, we may endow $\mathscr{A}(\Lambda)$ with the norm

$$\|A\|_{\mathscr{A}} = \|\mathbf{a}\|_1 .$$

By Lemma 5.38(2), this is indeed a norm on $\mathscr{A}(\Lambda)$.

By (5.33) and (5.34) $\mathscr{A}(\Lambda)$ is a Banach algebra embedded in $C^*(\Lambda)$. Since there are no convergence issues, $\mathscr{A}(\Lambda)$ might be called the "lazy man's rotation algebra."

As a variation we may also consider algebras of weighted absolutely convergent series of time-frequency shifts. If v is a submultiplicative weight on Λ, we introduce

$$\mathscr{A}_v(\Lambda) = \left\{ A \in \mathscr{B}(L^2(\mathbb{R}^d)) : A = \sum_{\lambda \in \Lambda} a_\lambda \pi(\lambda),\, \mathbf{a} \in \ell^1_v(\Lambda) \right\}$$

with norm

$$\|A\|_{\mathscr{A}_v} = \|\mathbf{a}\|_{\ell^1_v} .$$

In noncommutative geometry one considers the smooth noncommutative torus

$$\mathscr{A}_\infty(\Lambda) = \left\{ A \in \mathscr{B}(L^2(\mathbb{R}^d)) : A = \sum_{\lambda \in \Lambda} a_\lambda \pi(\lambda),\; |a_\lambda| = \mathscr{O}((1+|\lambda|)^{-N}),\; \forall N \geq 0 \right\}.$$

If we write v_s for the polynomial weights $v_s(z) = (1+|z|)^s$, then clearly the smooth noncommutative torus is the intersection

$$\mathscr{A}_\infty(\Lambda) = \bigcap_{s \geq 0} \mathscr{A}_{v_s}(\Lambda) . \tag{5.35}$$

The analogy between Fourier series and series of time-frequency shifts leads to the natural questions: Is there a version of Wiener's Lemma for the subalgebras of the rotation algebras? Is there a corresponding version of Wiener's Lemma for twisted convolution? These questions will be answered affirmatively in the following.

In contrast to Section 5.2 (where we treated absolutely convergent Fourier series before convolution operators), we consider convolution operators first and then turn to absolutely convergent series of time-frequency shifts.

5.3.3.3 Wiener's Lemma for Twisted Convolution

Given $\mathbf{h} \in \ell^1(\Lambda)$ or $\mathbf{h} \in \ell^1_v(\Lambda)$, define the *twisted convolution operator* $C^\natural_{\mathbf{h}}$ acting on $\mathbf{a} \in \ell^2(\Lambda)$ by

$$C^\natural_{\mathbf{h}}\mathbf{a} = \mathbf{h} \natural_\Lambda \mathbf{a}\,.$$

By Young's inequality, $C^\natural_{\mathbf{h}}$ maps $\ell^p(\Lambda)$ into $\ell^p(\Lambda)$. Likewise one can use convolution from the right and define a twisted convolution operator $\mathbf{a} \to \mathbf{a} \natural_\Lambda \mathbf{h}$. The results are the same.

We now have the following noncommutative counterpart of Theorem 5.18 [118].

Theorem 5.42. *Fix a weight v on Λ that satisfies the GRS condition. Assume that $\mathbf{h} \in \ell^1_v(\Lambda)$ and that $C^\natural_{\mathbf{h}}$ is invertible on $\ell^2(\Lambda)$; then $\mathbf{h}$ is invertible in the algebra $(\ell^1_v(\Lambda), \natural_\Lambda)$ and there exists a $\mathbf{g} \in \ell^1_v(\Lambda)$ such that $(C^\natural_{\mathbf{h}})^{-1} = C^\natural_{\mathbf{g}}$. As a consequence, $C^\natural_{\mathbf{h}}$ is invertible on all $\ell^p(\Lambda)$ simultaneously.*

Proof. Although the formulation is identical to that of Theorem 5.18 (we have only replaced $*$ by $\natural_\Lambda$), the proof is radically different, because we have lost commutativity and we can no longer use Fourier series. Instead we will use the results on matrix algebras from the previous section. The following proof is taken from [119].

Let us interpret $C^\natural_{\mathbf{h}}$ as a matrix acting on $\ell^2(\Lambda)$ and find its entries:

$$\begin{aligned} C^\natural_{\mathbf{h}}\mathbf{a} &= \sum_{\lambda \in \Lambda} h_\lambda a_{\nu-\lambda} e^{-2\pi i \lambda_1 \cdot (\nu_2 - \lambda_2)} \\ &= \sum_{\mu \in \Lambda} h_{\nu-\mu} a_\mu e^{-2\pi i (\nu_1 - \mu_1)\cdot \mu_2}\,. \end{aligned}$$

Thus, the matrix M of $C^\natural_{\mathbf{h}}$ has the entries

$$M_{\nu\mu} = h_{\nu-\mu} e^{-2\pi i (\nu_1 - \mu_1)\cdot \mu_2}\,.$$

Clearly, M is convolution-dominated. Moreover, since

$$\sup_{\nu \in \Lambda} |M_{\nu,\nu-\rho}| = |h_\rho|\,,$$

we find that $M \in \mathscr{C}_v$ and that

$$\|M\|_{\mathscr{C}_v} = \|\mathbf{h}\|_{\ell^1_v}\,. \tag{5.36}$$

Since $\mathscr{C}_v$ is inverse-closed in $\mathscr{B}(\ell^2(\Lambda))$ and $M = C^\natural_{\mathbf{h}}$ is invertible on $\ell^2(\Lambda)$, Theorem 5.30 guarantees that $(C^\natural_{\mathbf{h}})^{-1} = M^{-1} \in \mathscr{C}_v$.

It is left to show that M^{-1} is again a twisted convolution operator. Let $\mathbf{g} \in \ell^2(\Lambda)$ be the unique element such that $C^\natural_{\mathbf{h}}\mathbf{g} = \delta_0$. As in the proof of Theorem 5.18, we argue that $C^\natural_{\mathbf{g}}$ is the inverse of $C^\natural_{\mathbf{h}}$.

The (twisted) convolution operator $C^\natural_{\mathbf{g}}$ is certainly defined on the dense subspace $\ell^0(\Lambda) = \{\mathbf{a} : \operatorname{supp}\mathbf{a} \text{ is finite }\}$, and maps $\ell^0(\Lambda)$ into $\ell^2(\Lambda)$. Then for all $\mathbf{a} \in \ell^0(\Lambda)$

$$C^\natural_{\mathbf{h}}\big(C^\natural_{\mathbf{g}} - M^{-1}\big)\mathbf{a} = \mathbf{h} \natural_\Lambda (\mathbf{g} \natural_\Lambda \mathbf{a}) - C^\natural_{\mathbf{h}} M^{-1}\mathbf{a} = \mathbf{a} - \mathbf{a} = 0\,.$$

Since we have $C^\natural_{\mathbf{g}} = M^{-1}$ on the dense subspace $\ell^0(\Lambda)$, the matrix of $C^\natural_{\mathbf{g}}$ coincides with M, and so (5.36) implies that $\mathbf{g} \in \ell^1_v(\Lambda)$. □

Note that the deduction of Theorem 5.42 from Baskakov's Theorem 5.30 is rigorous. So far the only results that we have not proved completely are the results of Baskakov (or, equivalently, the operator-valued version of Wiener's Lemma by Bochner and Phillips [24]).

Since the matrix of $C^\natural_{\mathbf{h}}$ is convolution-dominated and belongs to $\mathscr{C}_v$, Corollary 5.33 can be rephrased as the spectral invariance of twisted convolution operators.

Corollary 5.43. *Fix* $\mathbf{h} \in \ell^1_v(\Lambda)$. *Then the following are equivalent:*

1. $C^\natural_{\mathbf{h}}$ *is invertible on* $\ell^2(\Lambda)$.
2. $C^\natural_{\mathbf{h}}$ *is invertible on* $\ell^p_m(\Lambda)$ *for* some $p \in [1,\infty]$ *and some v-moderate weight m.*
3. $C^\natural_{\mathbf{h}}$ *is invertible simultaneously on* $\ell^p_m(\Lambda)$ *for* all $p \in [1,\infty]$ *and all v-moderate weights m.*

5.3.3.4 Wieners Lemma for the Rotation Algebra

The following result was proved in [118].

Theorem 5.44. *Assume that v is submultiplicative and satisfies the GRS condition. If* $A \in \mathscr{A}_v(\Lambda)$ *and A is invertible on* $L^2(\mathbb{R}^d)$, *then* $A^{-1} \in \mathscr{A}_v(\Lambda)$.

Let us make plausible why the result follows from Wiener's Lemma for twisted convolution (Theorem 5.42). Consider the mapping $\pi : \ell^1_v(\Lambda) \to \mathscr{A}_v(\Lambda)$ defined by

$$\pi(\mathbf{a}) = \sum_{\lambda\in\Lambda} a_\lambda \pi(\lambda)\,.$$

By (5.32) and Lemma 5.38(2), π is an isometric $*$-isomorphism between $\ell^1_v(\Lambda)$ (with respect to $\natural_\Lambda$) and $\mathscr{A}_v(\Lambda)$ (with composition of operators). Thus, $\big(\ell^1_v(\Lambda), \natural_\Lambda\big)$ and $\mathscr{A}_v(\Lambda)$ are just different realizations of the same abstract Banach $*$-algebra. Clearly, Wiener's Lemma for one realization should imply Wiener's Lemma for the other realization.

The subtle part is in the hypotheses: In Theorem 5.42 we assume that $C_{\mathbf{h}}^{\natural}$ is invertible on $\ell^2(\Lambda)$, whereas in Theorem 5.44 we assume the invertibility of $A = \pi(\mathbf{h})$ on $L^2(\mathbb{R}^d)$.

The technical part of the proof requires the spectral identity

$$\sigma_{\mathscr{B}(L^2)}(\pi(\mathbf{h})) = \sigma_{\mathscr{B}(\ell^2(\Lambda))}(C_{\mathbf{h}}^{\natural}) \tag{5.37}$$

for all $\mathbf{h} \in \ell_v^1(\Lambda)$. This is a question of representation theory and requires some effort; see [118]. Once (5.37) has been proved, Theorem 5.44 follows from

$$\sigma_{\mathscr{A}_v(\Lambda)}(\pi(\mathbf{h})) = \sigma_{\ell_v^1}(\mathbf{h}) = \sigma_{\mathscr{B}(\ell^2(\Lambda))}(C_{\mathbf{h}}^{\natural}) = \sigma_{\mathscr{B}(L^2)}(\pi(\mathbf{h}))\,.$$

The interaction between time-frequency analysis and noncommutative geometry goes much further. The Banach algebras $\mathscr{A}_v(\Lambda)$ and their spectral invariance in the noncommutative torus $C^*(\Lambda)$ play a central role in the theory and are the starting point for many developments. We refer to [55, 170, 171, 199–201, 206] for noncommutative geometry. A generalization of Theorem 5.44 to time-frequency shifts not supported on a lattice is given in [9].

Once more the GRS condition characterizes those weights for which Wiener's Lemma for the rotation algebra holds.

Corollary 5.45. *$\mathscr{A}_v(\Lambda)$ is inverse-closed in $C^*(\Lambda)$ and in $\mathscr{B}(L^2(\mathbb{R}^d))$ if and only if v satisfies the GRS condition.*

Again the necessity of the GRS condition follows from a counterexample. If v violates the GRS condition, then there are $\lambda \in \Lambda$ and $a > 0$ such that $v(n\lambda) \geq e^{an}$ for $n \geq n_0$. Let $A = \mathrm{Id}_{\ell^2} - e^{-\delta}\pi(\lambda) \in \mathscr{A}_v(\Lambda)$. Then A is invertible in $\mathscr{B}(L^2(\mathbb{R}^d))$ with inverse

$$A^{-1} = \sum_{n=0}^{\infty} e^{-n\delta}\pi(\lambda)^n = \sum_{n=0}^{\infty} e^{-n\delta}\gamma_n \pi(n\lambda)\,,$$

where the γ_n are phase factors, $|\gamma_n| = 1$, resulting from the commutation rule (5.30). Then

$$\|A^{-1}\|_{\mathscr{A}_v(\Lambda)} = \sum_{n=0}^{\infty} e^{-n\delta} v(n\lambda)$$

and thus $A^{-1} \notin \mathscr{A}_v(\Lambda)$ whenever $\delta \leq a$.

This counterexample shows that the noncommutative torus $\mathscr{A}_v(\Lambda)$ with exponential weight v is not inverse-closed in $C^*(\Lambda)$. This observation was already made in [210] by means of a rather subtle argument.

In the motivating section 5.2 we discussed the quotient rule and proved that $C^k(\mathbb{T})$ is inverse-closed in $C(\mathbb{T})$ (Lemma 5.1). Consequently, if $f \in C^\infty(\mathbb{T})$ and $f(t) \neq 0$ for all $t \in \mathbb{T}$, then $1/f \in C^\infty(\mathbb{T})$. The statement for the smooth noncommutative torus $\mathscr{A}_\infty(\Lambda)$ is completely analogous and is a celebrated result of Connes [54].

Corollary 5.46. *If $A \in \mathscr{A}_\infty(\Lambda)$ and is invertible on $L^2(\mathbb{R}^d)$, then $A^{-1} \in \mathscr{A}_\infty(\Lambda)$, i.e., $A^{-1} = \sum_{\lambda \in \Lambda} b_\lambda \pi(\lambda)$ with rapidly decaying* **b**.

Proof. Recall from (5.35) that $\mathscr{A}_\infty(\Lambda) = \bigcap_{s\geq 0} \mathscr{A}_{v_s}(\Lambda)$ for the polynomial weight $v_s(\lambda) = (1+|\lambda|)^s$. Since $A \in \mathscr{A}_\infty(\Lambda) \subseteq \mathscr{A}_{v_s}(\Lambda)$ for all $s \geq 0$ and A is invertible on $L^2(\mathbb{R}^d)$, Theorem 5.44 asserts that $A^{-1} \in \mathscr{A}_{v_s}(\Lambda)$. This is true for all $s \geq 0$, so $A^{-1} \in \mathscr{A}_\infty(\Lambda)$. □

In view of (5.35) the above statement is a simple corollary of Theorem 5.44. For further discussion and background we refer to [172] and [202]; an independent time-frequency proof was given by Janssen [147].

5.3.4 Convolution Operators on Groups

Another variation concerns the form of Wiener's Lemma for convolution operators. For this variation we replace the Abelian group $\mathbb{Z}$ by a general locally compact group $\mathscr{G}$.

5.3.4.1 The Basics

Let us first mention the basic facts about locally compact groups [93, 139]. We write $x, y, \dots$ for the elements of $\mathscr{G}$. The group multiplication is $(x,y) \to xy$ and is continuous by definition. Every locally compact group possesses a left-invariant measure, the *Haar measure* dx, which satisfies

$$\int_{\mathscr{G}} f(ax)\,dx = \int_{\mathscr{G}} f(x)\,dx, \qquad \text{for all } a \in \mathscr{G},$$

for all continuous functions with compact support in $\mathscr{G}$. As usual, the L^p-norm is defined as $\|f\|_p = \left(\int_{\mathscr{G}} |f(x)|^p\,dx \right)^{1/p}$ and $L^p(\mathscr{G})$ is the completion of the continuous functions with compact support with respect to the p-norm. We write $|U| = \int_{\mathscr{G}} \chi_U(x)\,dx$ for the Haar measure of the set $U \subseteq \mathscr{G}$. The convolution is

$$(f * g)(x) = \int_{\mathscr{G}} f(y) g(y^{-1}x)\,dy. \tag{5.38}$$

As in the case of $\mathbb{Z}$, we may now study the convolution operator C_f with symbol f acting on a function h: $C_f h = f * h$. Young's inequality holds for arbitrary locally compact groups; therefore, we have

$$\|f * h\|_p \leq \|f\|_1 \, \|h\|_p$$

for all $f \in L^1(\mathscr{G})$ and $h \in L^p(\mathscr{G})$. Consequently, whenever $f \in L^1(\mathscr{G})$, the convolution operator C_f is bounded on $L^p(\mathscr{G})$ for all $p \in [1,\infty]$. In particular, $L^1(\mathscr{G})$ is a Banach algebra with respect to convolution. It is commutative if and only if $\mathscr{G}$ is a locally compact Abelian group (see Exercises).

The algebra $L^1(\mathscr{G})$ can be equipped with an involution. Let $\Delta : \mathscr{G} \to \mathbb{R}^+$ be the Haar modulus of $\mathscr{G}$, which is defined by the right translation invariance $\int_{\mathscr{G}} f(xa)dx = \Delta(a^{-1}) \int_{\mathscr{G}} f(x)\,dx$. Then $f^*(x) = \overline{f(x^{-1})}\Delta(x^{-1})$ is an involution on $L^1(\mathscr{G})$, and $L^1(\mathscr{G})$ becomes a Banach $*$-algebra.

Again we may study the spectrum of C_f as an operator acting on $L^p(\mathscr{G})$. To make the possible dependence of the spectrum on p explicit, we denote

$$\sigma_{\mathscr{B}(L_p)}(C_f) = \left\{\lambda \in \mathbb{C} : C_f - \lambda \mathrm{I} \text{ not invertible on } L^p(\mathscr{G})\right\} .$$

Now the fundamental question is whether the spectrum is independent of p. As we have seen in Section 5.2, for the group $\mathscr{G} = \mathbb{Z}$ this form of spectral invariance is equivalent to Wiener's Lemma.

What happens when we replace $\mathbb{Z}$ by more general groups; in particular, what happens for non-Abelian groups $\mathscr{G}$? The answer to this question is a far-reaching generalization of Wiener's Lemma and has led to very deep mathematics.

We start with an abstract answer [12].

Lemma 5.47. *The spectral invariance $\sigma_{\mathscr{B}(L^p)}(C_f) = \sigma_{\mathscr{B}(L^2)}(C_f)$ for $1 \le p \le \infty$ holds for all $f \in L^1(\mathscr{G})$ if and only if G is amenable and symmetric.*

Thus, the appropriate formulation of Wiener's Lemma on locally compact groups holds only for groups with certain properties. As a first insight we note that Wiener's Lemma does not generalize to arbitrary locally compact groups. Its validity depends subtly on the group structure.

But let us first explain the terms in Lemma 5.47.

A locally compact group is called *amenable* if there exists a continuous linear functional m on $L^\infty(\mathscr{G})$ such that $m(1) = 1$, m is positive ($f \ge 0 \Rightarrow m(f) \ge 0$), and $m(T_x f) = m(f)$ holds for all $f \in L^\infty(\mathscr{G})$ and $x \in \mathscr{G}$. One says that $\mathscr{G}$ possesses a translation-invariant mean on $L^\infty(\mathscr{G})$. Here T_x is the translation operator on $\mathscr{G}$ defined by $T_x f(y) = f(x^{-1}y), x, y \in \mathscr{G}$.

The group $\mathscr{G}$ is called *symmetric* if the Banach $*$-algebra $L^1(\mathscr{G})$ is symmetric. Recall from Section 5.2.5 that this means that the spectrum of positive elements is positive:

$$\sigma_{\mathscr{B}(L^1)}(C_{f^* * f}) \subseteq [0,\infty), \quad \forall f \in L^1(\mathscr{G}).$$

Both properties have been studied extensively and constitute independent directions of harmonic analysis. For many classes of locally compact groups it is known whether or not they possess these properties. Every compact group and every locally compact Abelian group is both symmetric and amenable. Roughly speaking, amenability and symmetry are properties that indicate the distance of a given group $\mathscr{G}$ to the class of commutative groups or to compact groups. Currently there is no example of a locally compact group that is symmetric, but not amenable. It is conjectured that every symmetric group is amenable.

To get a feeling for these two concepts, let us verify that every compact group is amenable. Indeed, the Haar measure is an invariant mean on $L^\infty(\mathscr{G})$. Since $L^\infty(\mathscr{G}) \subset L^1(\mathscr{G})$ for compact $\mathscr{G}$, the invariance properties of the Haar measure imply that

$$m(T_x f) = \int_{\mathscr{G}} T_x f(y)dy = \int_{\mathscr{G}} f(x^{-1}y)dy = \int_{\mathscr{G}} f(y)dy = m(f), \;\; f \in L^\infty(\mathscr{G}), \; x \in \mathscr{G} .$$

To construct an invariant mean on $\mathbb{R}^d$, we start with the functionals

$$m_n(f) = \mathrm{vol}(B_n)^{-1} \int_{B_n} f(x)\,dx,$$

where $B_n = \{x \in \mathbb{R}^d : |x| \le n\}$. The functionals m_n are in the unit ball of $L^\infty(\mathbb{R}^d)^*$. Since the unit ball of a dual Banach space is weak*-compact, the sequence m_n possesses at least one point of accumulation m. Since $m_n(1) = 1$, we also have $m(1) = 1$ and so $m \neq 0$. This limit point is translation-invariant and positive, and is thus an invariant mean on $L^\infty(\mathscr{G})$. See Exercises.

Next, every compact group is symmetric and every locally compact Abelian group is symmetric. However, symmetry is more subtle. The symmetry of compact groups requires some representation theory, and the symmetry of locally compact Abelian groups requires several properties of the Fourier transform.

Many classes of groups are known to be amenable and symmetric. "Extremely noncommutative" groups, in particular all semisimple Lie groups including $\mathrm{SL}(2,\mathbb{R})$ or $\mathrm{SL}(2,\mathbb{C})$, are neither amenable nor symmetric. Not surprisingly, Wiener's Lemma fails for these groups, and the spectrum of a convolution operator depends crucially on the domain space $L^p(\mathscr{G})$.

5.3.4.2 Convolution Operators on Groups of Polynomial Growth

A complete structural characterization of all groups that are symmetric and amenable seems to be completely out of reach. Therefore, one restricts the investigation to certain natural classes of locally compact groups and studies convolution operators on these groups. This is what we will do in the following.

A natural condition to impose on a group is how the size of a sequence of neighborhoods grows. We will consider groups of polynomial growth.

We say that $\mathscr{G}$ is compactly generated if there exists a neighborhood $U \subseteq G$ of the identity element such that $\mathscr{G} = \bigcup_{n=1}^\infty U^n$, where $U^n = \{u = u_1 u_2 \dots u_n : u_j \in U\}$. Such a neighborhood is called a *generating neighborhood*.

A group $\mathscr{G}$ is said to have *polynomial growth* if for some generating and relatively compact neighborhood U of the identity $\mathscr{G}$ there exist positive constants C, d such that

$$|U^n| \le C n^d \qquad \text{for all } n \in \mathbb{N}\,.$$

Every compact group has polynomial growth, because the Haar measure is finite and we may take $U = \mathscr{G}$ as a generating neighborhood. Also, every compactly generated, locally compact Abelian group possesses polynomial growth. This is easy to see for the elementary groups $\mathscr{G} = \mathbb{R}^d$ and $\mathscr{G} = \mathbb{Z}^d$. Let $U = [-1,1]^d \subseteq \mathbb{R}^d$, then $U^n = [-n,n]^d$, and $|U^n| = (2n)^d$, and thus $\mathbb{R}^d$ possesses polynomial growth. Likewise $\mathbb{Z}^d$ and every finitely generated Abelian group possess polynomial growth. The proof for general locally compact Abelian groups requires the structure theorem for such groups [139], but the proof can be carried out similarly.

Every compactly generated group of polynomial growth is amenable. The construction of an invariant mean is similar to the one on $\mathbb{R}^d$. Let

$m_n(f) = |U^n|^{-1} \int_{U^n} f(x)\, dx$; then $m_n(1) = 1$ and m_n is positive. It is not easy to show that every weak-* limit point of this sequence is an invariant mean on $L^\infty(\mathscr{G})$. See [148, 149].

In view of Lemma 5.47 the next question is whether every compactly generated group of polynomial growth is symmetric. If this is true, then the version of Wiener's Lemma for convolution operators will also hold for every group of polynomial growth.

The study of the symmetry of locally compact groups has a long history. It was known early on that highly non-Abelian groups, such as the matrix groups $GL(d,\mathbb{R}), SL(d,\mathbb{R})$, and more generally the semisimple Lie groups, cannot be symmetric. On the other hand, groups that are "almost Abelian," such as nilpotent groups, are symmetric [164, 171].

So far the final touch in the quest for symmetric groups was obtained by Losert. He derived a deep structure theorem for groups of polynomial growth and as a consequence showed that such groups are symmetric [168, 169]. The following theorem is a milestone in noncommutative harmonic analysis and is one of the deepest generalizations of the original Wiener's Lemma.

Theorem 5.48 (Losert [169]). *Every compactly generated group of polynomial growth is symmetric.*

Combining all properties of groups of polynomial growth and applying Lemma 5.47, we obtain the spectral invariance of convolution operators.

Corollary 5.49. *Assume that $\mathscr{G}$ is a compactly generated group of polynomial growth and $f \in L^1(\mathscr{G})$. Then for $1 \le p \le \infty$,*

$$\sigma_{\mathscr{B}(L^p)}(C_h) = \sigma_{\mathscr{B}(L^2)}(C_h)\,.$$

In other words, the spectrum of the convolution operator is independent of the L^p-space, and the version of Wiener's Lemma for convolution operators holds in all groups of polynomial growth.

Next we consider weighted versions of Wiener's Lemma on groups of polynomial growth. Recall that a locally bounded weight function v on a locally compact group is called submultiplicative if $v(xy) \le v(x)v(y)$ for all $x, y \in \mathscr{G}$, and symmetric if $v(x^{-1}) = v(x)$ for all $x \in \mathscr{G}$. The weighted L^1-space $L^1_v(\mathscr{G})$ is defined by the norm

$$\|f\|_{L^1_v} = \int_G |f(x)|v(x)\, dx.$$

If v is submultiplicative and symmetric, then L^1_v is a Banach $*$-algebra with respect to convolution and the involution $f^*(x) = \overline{f(x^{-1})}\Delta(x^{-1})$ and $L^1_v(\mathscr{G})$ is embedded in $L^1(\mathscr{G})$. See Exercises.

We have already seen in Section 5.3.1 that the GRS condition characterizes those weights for which a weighted version of Wiener's Lemma holds. This pattern carries over to the much more difficult situation of groups of polynomial growth.

Theorem 5.50 ([91, 89]). *Assume that $\mathscr{G}$ is a compactly generated group of polynomial growth. Then the following conditions are equivalent:*

1. *The Banach $*$-algebra $L^1_v(\mathscr{G})$ is symmetric.*
2. *Spectral invariance holds:*

$$\sigma_{L^1_v}(C_h) = \sigma_{L^2}(C_h) \quad \textit{for all } h \in L^1_v(\mathscr{G}).$$

3. *The weight v satisfies the GRS condition*

$$\lim_{n\to\infty} v(x^n)^{1/n} = 1 \quad \textit{for all } x \in \mathscr{G}.$$

This theorem is much deeper than the original version of Wiener's Lemma for weighted absolutely convergent Fourier series. The proof of the implication $(2) \Rightarrow (1)$ requires the full structure theorem of Losert [169] and a detailed analysis and corresponding modifications of Ludwig's proof that nilpotent groups are symmetric [171].

The necessity of the GRS condition for nondiscrete groups is similar in spirit to the counterexamples we have seen before, but it is tricky and requires Gaussian estimates for the heat kernel on $\mathscr{G}$. For discrete groups, the following argument proves the implication $(2) \Rightarrow (3)$.

The spectral invariance of (2) implies that $r_{\mathscr{B}(\ell^1_v)}(C_h) = r_{\mathscr{B}(\ell^2)}(C_h)$. Choose $h = \delta_x$ for $x \in \mathscr{G}$ and consider the convolution operator $C_{\delta_x} f = \delta_x * f$. Since $(\delta_x * f)(y) = f(x^{-1}y)$ is the translation by x, C_{δ_x} is unitary on $\ell^2(\mathscr{G})$ and thus $r_{\mathscr{B}(\ell^2)}(C_{\delta_x}) = 1$ for all $x \in \mathscr{G}$.

To treat $\ell^1_v(\mathscr{G})$, we first note that $r_{\mathscr{B}(\ell^1_v)}(C_h) = r_{\ell^1_v}(h)$, because

$$\|h\|_{\ell^1_v} = \|h * \delta_e\|_{\ell^1_v} \le \|C_h\|_{\ell^1_v \to \ell^1_v} \le \|h\|_{\ell^1_v}.$$

So let us compute the spectral radius of δ_x in $\ell^1_v(\mathscr{G})$:

$$r_{\ell^1_v}(\delta_x) = \lim_{n\to\infty} \|\delta_x * \cdots * \delta_x\|_{\ell^1_v}^{1/n} = \lim_{n\to\infty} \|\delta_{x^n}\|_{\ell^1_v}^{1/n} = \lim_{n\to\infty} v(x^n)^{1/n}.$$

Combining these observations, we find that

$$\lim_{n\to\infty} v(x^n)^{1/n} = r_{\ell^1_v}(\delta_x) = r_{\mathscr{B}(\ell^1_v)}(C_{\delta_x}) = r_{\mathscr{B}(\ell^2)}(C_{\delta_x}) = 1,$$

which is the GRS condition.

5.3.5 Pseudodifferential Operators

We now turn from harmonic analysis to a topic in classical analysis and discuss pseudodifferential operators. These arise in partial differential equations, in quantum mechanics, or in wireless communications. So far we have dealt with sequences and matrices on the index set $\mathbb{Z}^d$; now we deal with functions on $\mathbb{R}^d$ and operators acting on functions. Instead of Fourier series we use the Fourier transform

$$\hat{f}(\xi) = \int_{\mathbb{R}^d} f(x)\, e^{-2\pi i x\cdot\xi}\, dx.$$

In this section we will not prove any result. The goal is to tell the story of spectral invariance for pseudodifferential operators and put it in the context of Wiener's Lemma. The sequence of statements is the same as in the previous sections.

5.3.5.1 The Basics

A *pseudodifferential operator* in the Kohn–Nirenberg calculus is formally defined by the integral

$$K_\sigma f(x) = \int_{\mathbb{R}^d} \sigma(x,\xi)\hat{f}(\xi)e^{2\pi i \xi\cdot x}\, d\xi\,. \tag{5.39}$$

The function σ is called the (Kohn–Nirenberg) *symbol* of the operator.

This integral is certainly well defined whenever $\sigma \in L^\infty(\mathbb{R}^{2d})$ and $\hat{f}$ belongs to the Schwartz class $\mathscr{S}(\mathbb{R}^d)$. For more general symbols one may resort to a distributional interpretation. If $f,g \in \mathscr{S}(\mathbb{R}^d)$, then the function $R(g,f)(x,\xi) = e^{-2\pi i x\cdot\xi} g(x)\overline{\hat{f}(\xi)}$ belongs to $\mathscr{S}(\mathbb{R}^{2d})$. Then we may take a tempered distribution $\sigma \in \mathscr{S}'(\mathbb{R}^{2d})$ and define K_σ weakly by the formula

$$\langle K_\sigma f, g\rangle = \langle \sigma, R(g,f)\rangle \quad \text{for } f,g \in \mathscr{S}(\mathbb{R}^d)\,. \tag{5.40}$$

This weak interpretation defines a continuous operator from $\mathscr{S}(\mathbb{R}^d)$ to $\mathscr{S}'(\mathbb{R}^d)$. Clearly, (5.40) extends the definition (5.39) to general symbols. (As in Section 5.2.6, we take the duality conjugate-linear in the second term.)

If the symbol σ depends only on the first variable, $\sigma(x,\xi) = m(x)$, then $K_\sigma f(x) = \int_{\mathbb{R}^d} m(x)\hat{f}(\xi)\, e^{2\pi i x\cdot\xi}\, d\xi = m(x)f(x)$ is a just a multiplication operator. If σ depends only on the second variable, $\sigma(x,\xi) = \hat{\mu}(\xi)$, where $\hat{\mu}$ is the Fourier transform of a measure or distribution on $\mathbb{R}^d$, then $K_\sigma f = \mu * f$ is a convolution operator or a so-called Fourier multiplier.

Writing

$$\begin{aligned} K_\sigma f(x) &= \int_{\mathbb{R}^d} \sigma(x,\xi)\hat{f}(\xi)e^{2\pi i\xi\cdot x}\, d\xi \\ &= \int_{\mathbb{R}^d} \Big(\int_{\mathbb{R}^d} \sigma(x,\xi)e^{2\pi i\xi\cdot(x-y)}\, d\xi\Big) f(y)\, dy \\ &= \int_{\mathbb{R}^d} k(x,x-y) f(y)\, dy\,, \end{aligned}$$

we may interpret $K_\sigma f$ as a "time-dependent" convolution with kernel $k(x,y)$. This is the reason why pseudodifferential operators are used to model time-varying continuous systems in signal processing. See Section 5.3.6 for a detailed discussion.

By using the time-frequency shifts $M_\xi T_x f(t) = e^{2\pi i\xi\cdot t} f(t-x)$, we may write every pseudodifferential operator formally as a superposition of time-frequency shifts as follows:

$$K_\sigma f = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \widehat{\sigma}(\eta, u) M_\eta T_{-u} f \, du d\eta \,. \tag{5.41}$$

To understand (5.41), assume that $\sigma, \widehat{\sigma} \in L^1(\mathbb{R}^{2d})$ and use the Fourier inversion formula $\int_{\mathbb{R}^d} \sigma(x,\xi) e^{-2\pi i u \cdot \xi} \, d\xi = \int_{\mathbb{R}^d} \widehat{\sigma}(\eta, u) \, e^{2\pi i \eta \cdot x} \, d\eta$. Now the following computation is rigorous:

$$\begin{aligned} K_\sigma f(x) &= \int_{\mathbb{R}^d} \sigma(x,\xi) \hat{f}(\xi) e^{2\pi i x \cdot \xi} d\xi \\ &= \iint_{\mathbb{R}^{2d}} \sigma(x,\xi) \, e^{2\pi i (x-y)\cdot \xi} \, f(y) \, dy d\xi \\ &= \iint_{\mathbb{R}^{2d}} \hat{\sigma}(\eta, y-x) \, e^{2\pi i \eta \cdot x} f(y) \, dy d\eta \\ &= \iint_{\mathbb{R}^{2d}} \hat{\sigma}(\eta, u) \, e^{2\pi i \eta \cdot x} f(u+x) \, du d\eta \\ &= \iint_{\mathbb{R}^{2d}} \hat{\sigma}(\eta, u) M_\eta T_{-u} f(x) \, du d\eta \,. \end{aligned} \tag{5.42}$$

For general symbols $\sigma \in \mathscr{S}'(\mathbb{R}^{2d})$, (5.42) can be proved with a distributional argument. The form (5.41) is sometimes called the *spreading representation* of K_σ. For a survey of the time-frequency approach to pseudodifferential operators, we refer to [111].

The theory of pseudodifferential operators is usually treated as a subject of classical "hard" analysis, as is exemplified in the treatises of Hörmander [141] and Stein [217]. The spreading representation (5.41) suggests an alternative approach to pseudodifferential operators with time-frequency methods. The time-frequency approach has been particularly successful in the study of time-varying systems (Section 5.3.6) and is highly relevant for our discussion of Wiener's Lemma.

5.3.5.2 The Sjöstrand Class $M^{\infty,1}(\mathbb{R}^{2d})$

The mapping $\sigma \mapsto K_\sigma$ is an example of a symbolic calculus (as discussed at the end of Section 5.2). Our first task is the identification of "nice" symbols. Here classical analysis and time-frequency analysis offer rather different answers. Whereas the classical Hörmander classes are defined by differentiability properties, the time-frequency approach defines symbol classes via properties of the short-time Fourier transform.

Fix a nonzero window function $\Phi \in \mathscr{S}(\mathbb{R}^{2d})$ of $2d$-variables, e.g., the Gaussian. The short-time Fourier transform (STFT) of a symbol σ is

$$V_\Phi \sigma(z,\zeta) = (\sigma \cdot \overline{T_z \Phi})\widehat{\ }(\zeta) = \langle \sigma, M_\zeta T_z \Phi \rangle, \quad z, \zeta \in \mathbb{R}^{2d} \,. \tag{5.43}$$

The STFT of a symbol is a function on $\mathbb{R}^{4d}$. We say that a symbol σ belongs to the *Sjöstrand class* $M^{\infty,1}(\mathbb{R}^{2d})$ if

$$\|\sigma\|_{M^{\infty,1}} = \int_{\mathbb{R}^{2d}} \sup_{z\in\mathbb{R}^{2d}} |(\sigma \cdot T_z\Phi)\hat{}(\zeta)|\,d\zeta = \int_{\mathbb{R}^{2d}} \sup_{z\in\mathbb{R}^{2d}} |V_\Phi\sigma(z,\zeta)|\,d\zeta < \infty. \quad (5.44)$$

To obtain better control of the smoothness of the symbol, we will also consider the weighted Sjöstrand class $M_v^{\infty,1}(\mathbb{R}^{2d})$. We assume that v is a submultiplicative function on $\mathbb{R}^{2d}$. Then the *weighted Sjöstrand class* $M_v^{\infty,1}(\mathbb{R}^{2d})$ is defined by the norm

$$\|\sigma\|_{M_v^{\infty,1}} = \int_{\mathbb{R}^{2d}} \sup_{z\in\mathbb{R}^{2d}} |(\sigma \cdot T_z\Phi)\hat{}(\zeta)|\,v(\zeta)\,d\zeta\,.$$

It can be shown that the definition of $M_v^{\infty,1}(\mathbb{R}^{2d})$ is independent of the particular window function Φ as long as Φ belongs to a suitable space of test functions. Then different windows yield equivalent norms on $M_v^{\infty,1}(\mathbb{R}^{2d})$ [108, Thm. 11.3.7].

Note that if $\sigma \in M^{\infty,1}(\mathbb{R}^{2d})$ and $z \in \mathbb{R}^{2d}$ is fixed, then $(\sigma \cdot T_z\Phi)\hat{}(\zeta) = V_\Phi\sigma(z,\zeta) \in L^1(\mathbb{R}^{2d})$. This means that σ *coincides locally with the Fourier transform of an L^1-function*. At this stage one may already sense an analogy between Fourier series with ℓ^1-coefficients and the symbol class $M^{\infty,1}$. This analogy should alert us for yet another version of Wiener's Lemma.

Before approaching Wiener's Lemma, we first have to find two Banach algebras [a small one corresponding to the absolutely convergent Fourier series and a big one corresponding to $C(\mathbb{T})$]. This preliminary work was easy for convolution operators [Lemma 5.17) and for matrices [Lemma 5.28 and (5.23)]. For pseudodifferential operators the Banach algebra property is nontrivial and interesting in its own right.

We first state the algebra property of the Sjöstrand class [110, 213, 224]. The composition of two pseudodifferential operators K_σ and K_τ defines a product on the level of symbols via $K_\sigma K_\tau = K_{\sigma\circ\tau}$. Likewise, taking the adjoint operator yields an involution on the level of symbols by $(K_\sigma)^* = K_{\sigma^*}$. Explicit formulas are available [141], but are not necessary for the time-frequency approach.

Theorem 5.51. *If v is submultiplicative and $\sigma,\tau \in M_v^{\infty,1}$, then $K_\sigma K_\tau = K_{\sigma\circ\tau}$ with $\sigma\circ\tau \in M_v^{\infty,1}$. In fact, $M_v^{\infty,1}$ is a Banach $*$-algebra with respect to $\circ$.*

The following boundedness result has been proved many times [29, 108, 112, 116, 213].

Theorem 5.52. *If $\sigma \in M^{\infty,1}(\mathbb{R}^{2d})$, then K_σ is bounded on $L^2(\mathbb{R}^d)$ and $\|K_\sigma\|_{L^2\to L^2} \le C\|\sigma\|_{M^{\infty,1}}$.*

When dealing with convolution operators, we identified the sequence $\mathbf{h}$ with the corresponding convolution operator $C_{\mathbf{h}}$ and thus obtained an embedding of $\ell^1(\mathbb{Z})$ into $\mathscr{B}(\ell^2)$. In this spirit let us define $\mathrm{Op}(M_v^{\infty,1})$ as the set of all operators T from $\mathscr{S}(\mathbb{R}^d)$ to $\mathscr{S}'(\mathbb{R}^d)$ that can be written as a pseudodifferential operator $T = K_\sigma$ with a symbol $\sigma \in M_v^{\infty,1}$. Then the two previous results (Theorems 5.51 and 5.52) can be rephrased by saying that

$$\mathrm{Op}(M_v^{\infty,1}) \text{ is a Banach } *\text{-subalgebra of } \mathscr{B}(L^2(\mathbb{R}^d))$$
$$\text{with norm } \|K_\sigma\|_{\mathrm{Op}(M_v^{\infty,1})} = \|\sigma\|_{M_v^{\infty,1}}.$$

5.3.5.3 Wiener's Lemma for Pseudodifferential Operators

We have reached exactly the same point as in our discussions of convolution operators and matrix algebras. We study an interesting, noncommutative Banach algebra of operators with two striking properties: It is embedded into the C^*-algebra of bounded operators on L^2 and it possesses a hidden ℓ^1-like structure. So the logical next question is whether this algebra $\mathrm{Op}(M^{\infty,1})$ is again inverse-closed in $\mathscr{B}(L^2(\mathbb{R}^d))$. This fundamental discovery is due to Sjöstrand [214].

Theorem 5.53. *If $\sigma \in M^{\infty,1}(\mathbb{R}^{2d})$ and K_σ is invertible on $L^2(\mathbb{R}^d)$, then $K_\sigma^{-1} = K_\tau$ for some $\tau \in M^{\infty,1}$.*

Sjöstrand's proof works with a decomposition of the pseudodifferential operator K_σ into small localized pieces and with Wiener's Lemma for the matrix algebra $\mathscr{C}$ of Gohberg, Baskakov, and Kurbatov (Section 5.3.2). Our recent proof [112] applies time-frequency methods and structural Banach algebra arguments. It extends and sharpens Sjöstrand's result to the weighted case and clarifies the precise spectral invariance properties.

Theorem 5.54 ([112]). *Assume that v is a submultiplicative weight on $\mathbb{R}^{2d}$ satisfying the GRS condition $\lim_{n\to\infty} v(nz)^{1/n} = 1$ for all $z \in \mathbb{R}^{2d}$.*

If $\sigma \in M_v^{\infty,1}(\mathbb{R}^{2d})$ and K_σ is invertible on $L^2(\mathbb{R}^d)$, then $K_\sigma^{-1} = K_\tau$ for some $\tau \in M_v^{\infty,1}$.

Theorem 5.55 ([112, 114]). *Assume that v is submultiplicative on $\mathbb{R}^{2d}$. Then $\mathrm{Op}(M_v^{\infty,1})$ is inverse-closed in $\mathscr{B}(L^2)$ if and only if v satisfies the GRS condition*

$$\lim_{n\to\infty} v(nz)^{1/n} = 1$$

for all $z \in \mathbb{R}^{2d}$.

Once again the counterexample follows the pattern established in Section 5.3.1. If for some $z_0 = (u_0, \eta_0) \in \mathbb{R}^{2d}$ we have $\lim_{n\to\infty} v(nz_0)^{1/n} = e^a > 1$, then we consider the operator $\mathrm{Id}_{L^2} - e^{-\delta}\pi(z_0)$ for $\delta \le a$. Its symbol $\sigma(x,\xi) = 1 - e^{-\delta} e^{2\pi i(\eta_0\cdot x - u_0\cdot\xi)}$ is in $M_v^{\infty,1}(\mathbb{R}^{2d})$, but the symbol of the inverse operator $\sum_{n=0}^\infty e^{-\delta n}\pi(z_0)^n$ is not in $M_v^{\infty,1}(\mathbb{R}^{2d})$. See Exercises.

Further generalizations were obtained in [120]. Let us also mention that the theory of the rotation algebra discussed in Section 5.3.3 and the time-frequency analysis of pseudodifferential operators are related: Theorem 5.44 can be derived from Theorem 5.54 [115].

5.3.5.4 Spectral Invariance

As with convolution operators, pseudodifferential operators with "nice" symbols are bounded on a much larger class of function spaces, the so-called modulation spaces. In time-frequency analysis these are defined by properties of the short-time

Fourier transform. Let $\varphi(t) = e^{-\pi t \cdot t}$ be the Gaussian on $\mathbb{R}^d$, $1 \le p, q < \infty$, and m be a v-moderate weight. Then the modulation space $M_m^{p,q}$ is defined as the completion of the space of finite linear combinations $\left\{ f : f = \sum_{j=1}^r c_j \pi(z_j)\varphi, c_j \in \mathbb{C}, z_j \in \mathbb{R}^{2d} \right\}$ with respect to the norm

$$\|f\|_{M_m^{p,q}} = \left(\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |V_\varphi f(x,\xi)|^p \, m(x,\xi)^p \, dx \right)^{q/p} d\xi \right)^{1/q} .$$

For $p = \infty$ or $q = \infty$ a small modification of the definition is necessary. For a detailed exposition of modulation spaces, see [108, Chap. 11–13], for a historical account with an extensive list of reference, see [84].

The following general boundedness result for pseudodifferential operators on modulation spaces is the analogue of Young's inequality for convolution (Lemma 5.17). See [108, Chap. 14] or [112, 120].

Theorem 5.56. *If $\sigma \in M_v^{\infty,1}(\mathbb{R}^{2d})$, then K_σ is bounded on all modulation spaces $M_m^{p,q}$ with $1 \le p, q \le \infty$ and v-moderate m.*

As a consequence of Theorem 5.53, we obtain the complete spectral invariance for pseudodifferential operators.

Corollary 5.57. *Assume that v satisfies the GRS condition and $\sigma \in M_v^{\infty,1}$. Then the spectral invariance*

$$\sigma_{\mathscr{B}(M_m^{p,q})}(K_\sigma) = \sigma_{\mathscr{B}(L^2)}(K_\sigma)$$

holds for every $p, q \in [1, \infty]$ and every v-moderate weight m.

The argument is similar to the proof of Theorem 5.19 and does not require new ideas. One shows that K_σ is invertible on $M_m^{p,q}$ if and only if $(K_\sigma)^*$ is invertible on $M_m^{p,q}$ and then uses duality and interpolation of modulation spaces.

To sum up, the spectrum of a "nice" pseudodifferential operator does not depend on the space on which it acts.

A Connection to Classical Pseudodifferential Operators. Recall that a symbol σ belongs to the *Hörmander class* $S_{0,0}^0$ if and only if $\partial^\alpha \sigma \in L^\infty(\mathbb{R}^{2d})$ for all multi-indices $\alpha \ge 0$. Toft [225] observed that the Hörmander class can be written as an intersection of modulation spaces. If $v_s(\zeta) = (1 + |\zeta|)^s$, then

$$S_{0,0}^0 = \bigcap_{s \ge 0} M_{v_s}^{\infty,1} .$$

Since $\mathrm{Op}(M_{v_s}^{\infty,1})$ is inverse-closed in $\mathscr{B}(L^2)$, we find that the intersection $\bigcap_{s\ge 0} M_{v_s}^{\infty,1}$ is also inverse-closed in $\mathscr{B}(L^2)$. Formulated explicitly, this is a famous result of Beals [17] and probably the earliest result on spectral invariance in the theory of pseudodifferential operators.

Corollary 5.58. *If $\sigma \in S_{0,0}^0$ and K_σ is invertible on $L^2(\mathbb{R}^d)$, then $K_\sigma^{-1} = K_\tau$ for some $\tau \in S_{0,0}^0$.*

5.3.5.5 Almost Diagonalization of Pseudodifferential Operators

The results about pseudodifferential operators may look fairly technical. In this section we would like to make them a bit more plausible by using time-frequency analysis. Since a pseudodifferential operator K_σ is defined as a superposition of time-frequency shifts, it is almost obligatory to ask how K_σ acts on a time-frequency shift $\pi(z)$. This idea culminates in the insight that pseudodifferential operators are almost diagonal with respect to time-frequency "bases."

For the formulation of this result we need an additional concept. Fix a lattice $\Lambda = A\mathbb{Z}^{2d} \subseteq \mathbb{R}^{2d}$ and a "nice" basis function g. Ideally we choose the Gaussian $e^{-\pi t \cdot t}$, but any nonzero function satisfying $\int_{\mathbb{R}^{2d}} |\langle g, \pi(z)g\rangle|\, v(z)\, dz < \infty$ works. Such a g belongs to the modulation space $M_v^{1,1}(\mathbb{R}^d)$. We say that the set $\{\pi(\lambda)g : \lambda \in \Lambda\}$ is a *Gabor frame* if there exist constants $A, B > 0$ such that

$$A\|f\|_2^2 \le \sum_{\lambda\in\Lambda} |\langle f, \pi(\lambda)g\rangle|^2 \le B\|f\|_2^2 \quad \text{for all } f \in L^2(\mathbb{R}^d).$$

General frame theory and the explicit construction of Gabor frames are discussed in detail in Ole Christensen's Chapter 1. The construction of Gabor frames with a basis function in $M_v^{1,1}(\mathbb{R}^d)$ for rapidly increasing weights is more difficult and in fact requires Wiener's Lemma for twisted convolution! See [85, 118] and [108, Chap. 13].

The following characterization of the Sjöstrand class is the key to understanding many properties of pseudodifferential operators and to proving the main results in Section 5.3.5.

Theorem 5.59. *Assume that $g \in M_v^{1,1}(\mathbb{R}^d)$, $g \neq 0$, and that $\{\pi(\lambda)g : \lambda \in \Lambda\}$ is a Gabor frame for $L^2(\mathbb{R}^d)$. Then the following properties are equivalent.*

1. $\sigma \in M_v^{\infty,1}$.
2. *There exists a continuous function $H \in L_v^1(\mathbb{R}^{2d})$ such that*

$$|\langle K_\sigma(\pi(z)g), \pi(w)g\rangle| \le H(w-z) \quad \text{for all } w, z \in \mathbb{R}^{2d}. \tag{5.45}$$

3. *There is a $h \in \ell_v^1(\Lambda)$ such that*

$$|\langle K_\sigma(\pi(\mu)g), \pi(\lambda)g\rangle| \le h(\lambda-\mu) \quad \text{for all } \lambda, \mu \in \Lambda. \tag{5.46}$$

Theorem 5.59 shows that the time-frequency shift $\pi(z)$ is almost an eigenvector of K_σ and that K_σ is almost diagonalized by frames of time-frequency shifts.

Let us rewrite (5.46) and connect it to the topic of matrix algebras. Let $M(\sigma)$ be the matrix indexed by $\mathbb{Z}^{2d}$ with the entries

$$M(\sigma)_{kl} = \langle K_\sigma(\pi(Al)g), \pi(Ak)g\rangle, \quad \lambda = Ak, \mu = Al \in \Lambda. \tag{5.47}$$

Then (5.46) states that $|M(\sigma)_{kl}| \le h(A(k-l))$, and thus the matrix of K_σ is dominated by convolution with $\tilde{h} = h \circ A$. In the light of Section 5.3.2 we may recast Theorem 5.59 in a more compact way.

Theorem 5.60. *Assume that* $\{\pi(\lambda)g : \lambda \in \Lambda\}$ *is a frame for* $L^2(\mathbb{R}^d)$ *and* $g \in M_v^{1,1}(\mathbb{R}^d)$*. Then a symbol* σ *belongs to the weighted Sjöstrand class* $M_v^{\infty,1}$ *if and only if* $M(\sigma)$ *belongs to the matrix algebra* $\mathscr{C}_v$*.*

$$\sigma \in M_v^{\infty,1}(\mathbb{R}^{2d}) \iff M(\sigma) \in \mathscr{C}_v\,.$$

Furthermore, it can be shown that $\|\sigma\|_{M_v^{\infty,1}}$ and $\|M(\sigma)\|_{\mathscr{C}_v}$ are equivalent norms on the symbol class $M_v^{\infty,1}$.

By means of Theorem 5.59, the algebra property of $M_v^{\infty,1}$ and Wiener's Lemma for pseudodifferential operators are now much easier to understand and follow from the corresponding properties of the matrix algebra $\mathscr{C}_v$.

Assume that $\sigma, \tau \in M_v^{\infty,1}$ and thus $M(\sigma), M(\tau) \in \mathscr{C}_v$. The composition of operators corresponds to matrix multiplication. Consequently, $K_{\sigma\circ\tau} = K_\sigma K_\tau$ corresponds to the matrix $M(\sigma)M(\tau) \in \mathscr{C}_v$ and thus $\sigma \circ \tau \in M_v^{\infty,1}$.

Likewise, the inverse of K_σ corresponds to the inverse of $M(\sigma)$. Since $\mathscr{C}_v$ is inverse-closed in $\mathscr{B}(\ell^2)$, $M(\sigma)^{-1} \in \mathscr{C}_v$ and by Theorem 5.59 the symbol τ of $K_\sigma^{-1} = K_\tau$ is in $M_v^{\infty,1}$. The rigorous proof requires much more work, because strictly speaking, $M(\sigma)$ is not invertible. It possesses a nontrivial kernel and is invertible only on a certain subspace of ℓ^2. See [112] for the precise details.

Our main point here is that Wiener's Lemma for matrix algebras enters crucially and directly in the proof of Wiener's Lemma for pseudodifferential operators.

5.3.6 Time-Varying Systems and Wireless Communications

In Section 5.2.6 we used discrete time-invariant systems as a motivation for convolution operators on $\mathbb{Z}$, and in Section 5.3.2 discrete time-varying systems served as a motivation to study matrix algebras. In both cases, we assumed a "digital" world, and a signal was understood to be a sequence of numbers. The "physical" world, however, is continuous, and therefore we now turn to the discussion of "analog" signals and continuous time-varying systems. For the mathematician, signals are functions on $\mathbb{R}$ or $\mathbb{R}^d$, and systems are operators on $L^2(\mathbb{R}^d)$.

The goal of this section is to discuss how the results about pseudodifferential operators of Section 5.3.5 can be applied in an engineering context.

5.3.6.1 Time-Varying Systems

Let us make a simple model of a time-varying system as it is used in mobile communications.

In Fig. 5.1 a signal f is transmitted by an antenna to a cellular phone in a moving tramway car. The signal is an electromagnetic wave and the propagation of f is governed by the wave equation. Thus, we are forced to work with analog signals.

During transmission the signal is distorted and transformed by various effects, of which we model two main effects:

(a) The signal is reflected at various obstacles and the signal arrives at the receiver with a delay caused by different path lengths. Formally, the received signal $\tilde{f}$ is a weighted superposition of time shifts of the transmitted signal f with some weight V:

$$\tilde{f}(t) = \int_{\mathbb{R}^d} V(u) \cdots f(t+u)\, du\,.$$

The weight V depends on the physical characteristics of the transmission, such as the path length or the absorption at reflectors.

(b) If the sender and receiver are in motion with respect to each other, then the Doppler effect will result in a frequency shift proportional to the relative velocity of sender and receiver. Since $\widehat{M_\xi f}(\tau) = \hat{f}(\tau - \xi)$, the received signal $\tilde{f}$ is a superposition of modulations (= frequency shifts) with some weight W:

$$\tilde{f}(t) = \int_{\mathbb{R}^d} W(\eta) \cdots e^{2\pi i \eta t} f(t) d\eta\,.$$

The differing path lengths and the Doppler effect are illustrated in Fig. 5.1.

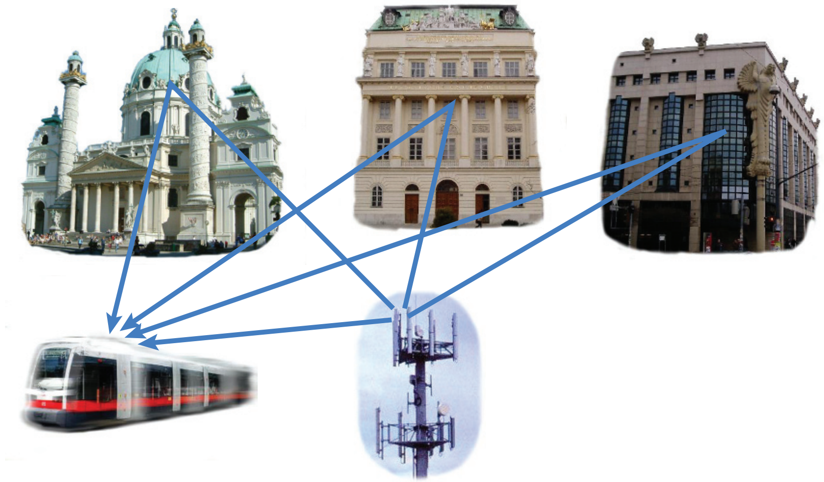


Fig. 5.1: Signal distortion caused by variable path length and by the Doppler effect. (Courtesy of Gerald Matz, Technical University of Vienna.)

Combining the two types of distortion, we find that the received signal $\tilde{f}$ is a superposition of time-frequency shifts, which we write as

$$\tilde{f} = \int_{\mathbb{R}^{2d}} \widehat{\sigma}(\eta, u)\, e^{2\pi i \eta \cdot t} f(t+u)\, du d\eta\,. \tag{5.48}$$

The weighting factor $\widehat{\sigma}$ is the spreading function of the time-varying system. It depends on the physical characteristics of the transmission and is usually estimated with statistical methods.

Comparing with (5.42), we understand the awkward notation for the weighting factor. The distortion $f \to \tilde{f}$ is nothing more than the *pseudodifferential operator* K_σ with the Kohn–Nirenberg symbol σ, and thus

$$\tilde{f} = K_\sigma f = \int_{\mathbb{R}^d} \sigma(x,\xi)\hat{f}(\xi)e^{2\pi i\xi\cdot x}\,d\xi\,.$$

Although mathematicians and engineers study the same object, there is a big difference between pseudodifferential operators in analysis and time-varying systems in engineering. Pseudodifferential operators are used for the construction of approximate inverses (parametrices) of partial differential operators and for the study of regularity properties in partial differential equations. Thus, all conditions and arguments in analysis involve derivatives and smoothness.

In engineering, however, differentiability properties do not play a role. The spreading representation (5.48) rather suggests time-frequency methods as a tool for the investigation of time-varying systems.

Symbol Classes in Mobile Communications. Let us first discuss the question of symbol classes. For physical reasons there are a maximum Doppler shift ν_0 and also a maximum time delay τ_0; consequently, the spreading function $\widehat{\sigma}$ is compactly supported in the rectangle $[-\nu_0,\nu_0]\times[0,\tau_0]$. A time-varying system with compactly supported spreading function is called *underspread* [183]. Such operators play an important role in communication theory.

Concerning the nature of $\hat{\sigma}$, the modeling of engineers is at odds with the mathematician's need for rigor. The standard assumption of engineers is that $\sigma \in L^2$. This assumption is doubtful, because then by Pool's theorem [197] K_σ must be a Hilbert–Schmidt operator. This is definitely a problem, because the class of Hilbert–Schmidt operators excludes the distortion-free channel (the identity operator) and time-invariant channels (convolution operators). Furthermore, a Hilbert–Schmidt operator cannot have a bounded inverse on $L^2(\mathbb{R})$, or in engineering terms, the recovery of f from $\tilde{f}$ is ill-posed, and the equalization will be extremely unstable.

To make a more satisfactory model for the symbols arising in wireless communications and time-varying channels, we follow Strohmer [219]. We keep the assumption that $\operatorname{supp}\widehat{\sigma}$ is compact, but we admit $\widehat{\sigma}$ to be a distribution in $M^\infty(\mathbb{R}^{2d})$. This means that $\widehat{\sigma}$ is a tempered distribution with bounded short-time Fourier transform. For instance, if $\hat{\sigma} = \delta$ (the point measure at 0), then $K_\sigma = \mathrm{Id}_{L^2}$.

Lemma 5.61. *Assume that* $\widehat{\sigma} \in M^\infty(\mathbb{R}^{2d})$ *and that* $\operatorname{supp}\widehat{\sigma}$ *is compact in some ball* $B(0,R) = \{z \in \mathbb{R}^{2d} : |z| \le R\}$. *If* $\widehat{\Phi} \in \mathscr{S}(\mathbb{R}^{2d})$ *and* $\operatorname{supp}\widehat{\Phi} \subseteq B(0,R)$, *then*

$$\int_{\mathbb{R}^{2d}} \sup_{z\in\mathbb{R}^{2d}} |V_\Phi\sigma(z,\zeta)|\,v(\zeta)\,d\zeta < \infty \tag{5.49}$$

for any nonnegative locally bounded function v. Consequently, $\sigma \in M_v^{\infty,1}(\mathbb{R}^{2d})$ *for any weight v.*

Proof. Observe that $(\sigma \cdot T_z \Phi)\widehat{\ }(\zeta) = (\widehat{\sigma} * M_{-z}\widehat{\Phi})(\zeta)$ and hence for fixed z has its support in $B(0,2R)$. We find that

$$V_\Phi \sigma(z,\zeta) = 0 \qquad \text{for } |z| > 2R\,.$$

Since $V_\Phi \sigma$ is bounded, (5.49) follows. □

5.3.6.2 Transmission of Information by OFDM

Next we explain how pseudodifferential operators enter in the process of data transmission.

1. The data. Given is a discrete set of data (digital information), $c_{kl} \in \mathbb{C}$, where $k,l \in \mathbb{Z}$. Usually the numbers c_{kl} are taken from a finite alphabet, either $c_{kl} \in \{-1,1\}$ or $c_{kl} \in \{i^\ell(1+i) : \ell = 0,1,2,3\}$. The parameter k indicates the time when the coefficient c_{kl} is transmitted, and the parameter l labels the frequency band over which the coefficient is sent. For fixed l, we may think of the sequence $\{c_{kl} : k \in \mathbb{Z}\}$ as a "word" that is sent over the lth frequency band. For fixed k, the set $\{c_{kl} : l \in \mathbb{Z}\}$ is the symbol group that is transmitted at time k.

2. Digital–analog conversion. In the first step the digital information c_{kl} is converted to an analog signal. The data c_{kl} serve as coefficients in a series expansion. Fix a suitable pulse g; then the transmitted analog signal is

$$f(t) = \sum_{k\in\mathbb{Z}} \Big(\sum_{l\in\mathbb{Z}} c_{kl} e^{2\pi i\beta l t}\Big) g(t-\alpha k) = \sum_{k,l\in\mathbb{Z}} c_{kl} M_{\beta l} T_{\alpha k} g(t)\,. \tag{5.50}$$

A series of this form is called a *Gabor series*. (Of course, in practice the sum is finite.)

It is easy to understand why Gabor series are a convenient way to transform a discrete set of data into an analog signal. The symbol group transmitted at time αk is $\sum_{l\in\mathbb{Z}} c_{kl} e^{2\pi i\beta l t}\, g(t-\alpha k)$; this is the Fourier series of the coefficients c_{kl} and can be calculated easily with a fast Fourier transform (FFT). For fixed l, the lth word is transmitted as $f_l(t) = M_{\beta l}\big(\sum_{k\in\mathbb{Z}} c_{kl} g(t-\alpha k)\big)$. If $\operatorname{supp} \hat{g} \subseteq [-\beta'/2, \beta'/2]$, then $\operatorname{supp} \widehat{f_l} \subseteq [\beta l - \beta'/2, \beta l + \beta'/2]$. If we choose $\beta' < \beta$, then each word is transmitted on a different frequency band.

This method for the simultaneous transmission of several independent data sets is called *frequency-division multiplexing*. If the time-frequency shifts $M_{\beta l} T_{\alpha k} g$ form an orthogonal set, then one speaks of orthogonal frequency-division multiplexing, OFDM in short.

In industrial applications the pulse g is usually chosen to be a characteristic function $g = \chi_{[0,\alpha']}$ with $\alpha' < \alpha$. Such a pulse achieves a good separation of consecutive symbol groups and works optimally in stationary environments. However, in nonstationary environments the Fourier transform of the characteristic function $\chi_{[0,\alpha']}$ decays slowly, and adjacent frequency bands are not separated adequately.

For nonstationary time-varying environments better pulse shapes are required. The ideal pulse is compactly supported in time and frequency, and its time-frequency shifts $M_{\beta l}T_{\alpha k}g$ are mutually orthogonal. These properties exclude each other because of the uncertainty principle. Thus, it becomes a relevant mathematical problem to construct appropriate pulse shapes that are compatible with the uncertainty principle. See [26, 184, 218, 220] for a small sample of papers by both engineers and mathematicians. Let us mention that Wiener's Lemma both for Fourier series and for matrix algebras is used implicitly in several pulse-shaping constructions.

Figure 5.2 shows a formal representation of the transmitted signal in the time-frequency plane. Each coefficient c_{kl} belongs to a different cell in the time-frequency plane; the goal of pulse design is to ensure that the cells are well separated from each other.

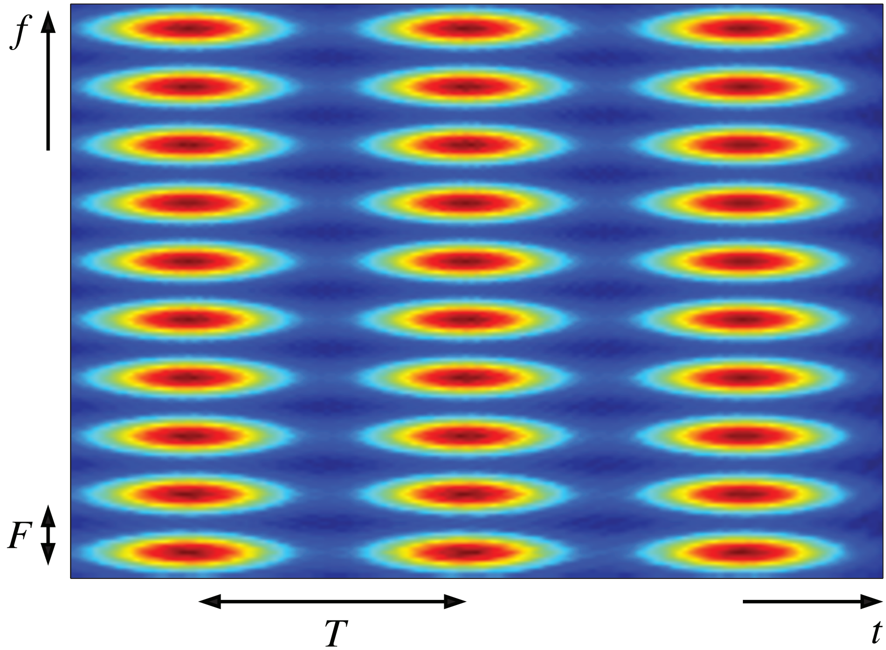


Fig. 5.2: Each coefficient c_{kl} of the transmitted signal occupies a cell in the time-frequency plane.

3. Transmission of *f* and signal distortion. Next the analog signal generated from the digital information is transmitted by a sender and, in the course of its propagation, undergoes various distortions. Let us emphasize that the transmission is a physical process subject to the wave equation. This is not a phenomenon in a discrete space, and any discretization is an approximation that has to be justified.

As explained above, the distortion is described by a pseudodifferential operator. The received signal is of the form

$$\tilde{f} = K_\sigma f = \int_{\mathbb{R}^d} \widehat{\sigma}(u,\eta) M_\eta T_{-u} f \, du d\eta \,,$$

and the symbol may be assumed to be in a weighted Sjöstrand class $M_v^{\infty,1}(\mathbb{R}^{2d})$.

4. Analog–digital conversion. In terms of the coefficients c_{kl}, the received signal is

$$\tilde{f} = K_\sigma f = \sum_{k',l' \in \mathbb{Z}} c_{k',l'} \, K_\sigma (M_{\beta l'} T_{\alpha k'} g) \,.$$

The goal is now to recover the coefficients c_{kl} from $\tilde{f}$. This task is usually approached by taking correlations, i.e., by taking inner products with time-frequency shifts of g (or with other known pulse shapes). After taking correlations, we obtain the sequence

$$\begin{aligned} y_{kl} &= \langle \tilde{f}, M_{\beta l} T_{\alpha k} g \rangle \\ &= \sum_{k',l' \in \mathbb{Z}} c_{k',l'} \langle K_\sigma (M_{\beta l'} T_{\alpha k'} g), M_{\beta l} T_{\alpha k} g \rangle \,. \end{aligned}$$

Let $\mathbf{y}$ be the vector with components y_{kl} and $M(\sigma)$ be the matrix with entries

$$M(\sigma)_{kl,k'l'} = \langle K_\sigma (M_{\beta l'} T_{\alpha k'} g), M_{\beta l} T_{\alpha k} g \rangle \,. \tag{5.51}$$

Then we may write the relation between the original data sequence $\mathbf{c}$ and the output sequence $\mathbf{y}$ as an infinite system of linear equations

$$M(\sigma)\mathbf{c} = \mathbf{y} \,. \tag{5.52}$$

The matrix $M(\sigma)$ is the matrix of K_σ with respect to a set of time-frequency shifts. In the context of data transmission it is called the *channel matrix*. Its entries describe the interference between different cells in the time-frequency plane caused by the distortion operator K_σ.

5. Equalization. Finally, we have to recover the original data $\mathbf{c}$. This amounts to solving the system (5.52) or explicitly

$$\mathbf{c} = M(\sigma)^{-1} \mathbf{y} \,.$$

In order to write the inverse of $M(\sigma)$ in a meaningful way, we have to assume that $M(\sigma)$ is an invertible matrix. The usual assumptions are that K_σ is invertible and that the set of time-frequency shifts $M_{\beta l} T_{\alpha k} g$ forms a Riesz basis (or orthogonal basis) for some subspace of L^2. (Otherwise the input–output relationship $f \to \tilde{f} = K_\sigma f$ would be ill-posed and the reconstruction of the data c_{kl} from the received signal $\tilde{f}$ would be unstable.)

In general, the solution of the system (5.52) or the inversion of the channel matrix poses a challenging computational problem. In practice this is *not* a problem, because it is generally assumed that the channel matrix $M(\sigma)$ is diagonal. Thus, $\langle K_\sigma(M_{\beta l'}T_{\alpha k'}g), M_{\beta l}T_{\alpha k}g\rangle = 0$ for $(k,l) \neq (k',l')$. Consequently, the original data are simply given by

$$c_{kl} = \langle K_\sigma(M_{\beta l}T_{\alpha k}g), M_{\beta l}T_{\alpha k}g\rangle^{-1} y_{kl}\,.$$

This assumption of diagonality seems truly miraculous. There is no reason why the channel matrix should be exactly diagonal. At first glance it is really amazing that the OFDM method should work for time-varying systems.

Nevertheless, the engineering intuition is correct, and the described equalization method works in practice. The mathematical reason is, once again, Wiener's Lemma.

Proposition 5.62. *Assume that $\widehat{\sigma}$ has compact support and $M(\sigma)$ is invertible on $L^2(\mathbb{R}^d)$, that $v(x,\xi) = e^{a(|x|+|\xi|)^b}$ for $a>0, 0<b<1$, and that $g \in M^{1,1}_v(\mathbb{R}^d)$. Then*

$$|M(\sigma)_{kl,k'l'}| \leq C\, e^{-a(|k-k'|+|l-l'|)^b}, \quad \forall k,k',l,l' \in \mathbb{Z},$$

and likewise

$$|\big(M(\sigma)^{-1}\big)_{kl,k'l'}| \leq C'\, e^{-a(|k-k'|+|l-l'|)^b}, \quad \forall k,k',l,l' \in \mathbb{Z}.$$

Proof. By Lemma 5.61 the symbol describing the channel is in every $M^{\infty,1}_v(\mathbb{R}^{2d})$. Consequently, by Theorem 5.59 the matrix of K_σ is almost diagonal and its entries are dominated by $|M(\sigma)_{kl,k'l'}| \leq h(k-k', l-l')$ for some sequence $h \in \ell^1_v$. In particular, $|M(\sigma)_{kl,k'l'}| \leq h(k-k',l-l') \leq Cv(k-k',l-l')^{-1}$. Since the subexponential weight v satisfies the GRS condition, Wiener's Lemma for pseudodifferential operators (Theorem 5.53) implies that the same decay property holds for the inverse matrix. □

Proposition 5.62 says that the channel matrix *and its inverse* are almost diagonal with respect to arbitrary subexponential weights. Consequently, it suffices to use the diagonal (and possibly a few side-diagonals) for matrix computations. In particular, the inverse of $M(\sigma)$ is almost a diagonal matrix. Thus, Proposition 5.62 can be taken as a mathematical justification of the engineering practice.

At the core of this justification is Wiener's Lemma for pseudodifferential operators.

Acknowledgments K. G. was supported by the Marie-Curie Excellence Grant MEXT-CT 2004-517154. He thanks Michael Leinert for reading the manuscript and providing useful feedback.

Exercises for Section 5.3

1. Show that the weight function $v(k) = e^{a|k|^b}$ $a,b > 0$, satisfies the GRS condition if and only if $b < 1$. Show that the function $v(k) = e^{|k|/\log(e+|k|)}$ is submultiplicative and satisfies the GRS condition.
2. Formulate and prove a version of the theorem of Wiener–Levy (Theorem 5.16) for the algebra of weighted absolutely convergent Fourier series $\mathscr{A}_v(\mathbb{T})$.
3. (a) Show that the space of matrices defined by the Schur-type conditions (5.27) is a Banach algebra with respect to matrix multiplication.
   (b) Prove the continuous embeddings $\mathscr{C}_v \subseteq \mathscr{A}_v^1 \subseteq \mathscr{A}_v^\infty$ of (5.28) and give examples to show that the inclusions are proper.
4. Let $\mathscr{J} = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$ (consisting of $d \times d$ blocks). Given a lattice $\Lambda = A\mathbb{Z}^{2d}$ for some invertible matrix A, define the *adjoint lattice* Λ° by
$$\Lambda^\circ = \mathscr{J}(A^T)^{-1}\mathbb{Z}^{2d}, \tag{5.53}$$
   where A^T is the transpose of A. Show that a time-frequency shift $\pi(z)$ commutes with all $\pi(\lambda), \lambda \in \Lambda$, if and only if $z \in \Lambda^\circ$ [88].
5. Given are a lattice $\Lambda \subseteq \mathbb{R}^{2d}$ and the exponential weight $v(\lambda) = e^{|\lambda|}$. Find a sequence $\mathbf{h} \in \ell_v^1(\Lambda)$ such that the twisted convolution operator $C_{\mathbf{h}}^\natural$ is invertible on $\ell^2(\Lambda)$ with inverse $C_{\mathbf{g}}^\natural$, but $\mathbf{g} \notin \ell_v^1(\Lambda)$.
6. Show that the set $\{C_{\mathbf{h}}^\natural : \mathbf{h} \in \ell_v^1(\Lambda)\}$ is a closed $*$-subalgebra of $\mathscr{C}_v$.
7. Prove statements (1) and (2) of Lemma 5.32.
8. For a locally compact group $\mathscr{G}$, show that $L^1(\mathscr{G})$ is commutative if and only if $\mathscr{G}$ is Abelian.
9. Let $m_n(f) = \mathrm{vol}(B_n)^{-1} \int_{B_n} f(x)\,dx$ for $f \in L^\infty(\mathbb{R}^d)$. Show that any limit point m of the sequence m_n is an invariant mean on $L^\infty(\mathbb{R}^d)$, i.e., $m(1) = 1$, $|m(f)| \le \|f\|_\infty$, $f \ge 0 \Rightarrow m(f) \ge 0$, and $m(T_x f) = m(f)$ for all $f \in L^\infty(\mathbb{R}^d)$ and $x \in \mathbb{R}^d$.
10. Let $\mathscr{G}$ be a locally compact group and v be a submultiplicative, even weight on $\mathscr{G}$. Show that $L_v^1(\mathscr{G})$ is a Banach $*$-algebra with respect to convolution and the involution $f^*(x) = \overline{f(x^{-1})}\Delta(x^{-1})$.
11. Let $\sigma(x,\xi) = \sum_{k,l \in \mathbb{Z}^d} a_{kl} e^{2\pi i(k\cdot x + l\cdot\xi)}$.
    a. Describe the associated pseudodifferential operator K_σ.
    b. Show that $\sigma \in M_v^{\infty,1}(\mathbb{R}^{2d})$ if and only if $\mathbf{a} \in \ell_v^1(\mathbb{Z}^{2d})$.

References

1. Abramowitz, M., Stegun, I.: Handbook of Mathematical Functions. Dover Publications Inc., New York (1965)
2. Aldroubi, A., Baskakov, A.G., Krishtal, I.A.: Slanted matrices, Banach frames, and sampling. J. Funct. Anal. **255**, 1667–1691 (2008)
3. Antoine, J.P., Demanet, L., Jacques, L., Vandergheynst, P.: Wavelets on the sphere: Implementation and approximations. Appl. Comput. Harmon. Anal. **13**(3), 177–200 (2002)
4. Antoine, J.P., Murenzi, R., Vandergheynst, P., Ali, S.T.: Two-Dimensional Wavelets and Their Relatives. Cambridge University Press, Cambridge, U.K. (2004)
5. Antoine, J.P., Vandergheynst, P.: Wavelets on the n-sphere and related manifolds. J. Math. Phys. **39**(8), 3987–4008 (1998)
6. Antoine, J.P., Vandergheynst, P.: Wavelets on the 2-sphere: A group-theoretical approach. Appl. Comput. Harmon. Anal. **7**(3), 262–291 (1999)
7. Antoine, J.P., Vandergheynst, P.: Wavelets on the two-sphere and other conic sections. J. Fourier Anal. Appl. **13**(4), 369–386 (2007)
8. Arveson, W.: Discretized CCR algebras. J. Oper. Theory **26**(2), 225–239 (1991)
9. Balan, R.: The noncommutative Wiener lemma, linear independence, and spectral properties of the algebra of time-frequency shift operators. Trans. Amer. Math. Soc. **360**(7), 3921–3941 (2008)
10. Balan, R., Casazza, P.G., Heil, C., Landau, Z.: Density, overcompleteness, and localization of frames. II. Gabor systems. J. Fourier Anal. Appl. **12**(3), 309–344 (2006)
11. Baldi, P., Kerkyacharian, G., Marinucci, D., Picard, D.: Asymptotics for spherical needlets. Ann. Stat. **37**(3), 1150–1171 (2009)
12. Barnes, B.A.: When is the spectrum of a convolution operator on L^p independent of p? Proc. Edinburgh Math. Soc. **33**(2), 327–332 (1990)
13. Barnes, B.A.: Symmetric Banach $*$-algebras: invariance of spectrum. Studia Math. **141**(3), 251–261 (2000)
14. Baskakov, A.G.: Wiener's theorem and asymptotic estimates for elements of inverse matrices. Funktsional. Anal. i Prilozhen. **24**(3), 64–65 (1990)
15. Baskakov, A.G.: Estimates for the elements of inverse matrices, and the spectral analysis of linear operators. Izv. Ross. Akad. Nauk Ser. Mat. **61**(6), 3–26 (1997)
16. Baskakov, A.G.: Theory of representations of Banach algebras, and Abelian groups and semigroups in the spectral analysis of linear operators. Sovrem. Mat. Fundam. Napravl. **9**, 3–151 (2004)
17. Beals, R.: Characterization of pseudodifferential operators and applications. Duke Math. J. **44**(1), 45–57 (1997)
18. Benedetto, J., Li, S.: The theory of multiresolution analysis frames and applications to filter banks. Appl. Comput. Harmon. Anal. **5**, 389–427 (1998)

19. Benedetto, J., Powell, A., Yilmaz, Ö.: Sigma-delta quantization and finite frames. IEEE Inform. Theory **52**, 1990–2005 (2006)
20. Benedetto, J., Treiber, O.: Wavelet frames: Multiresolution analysis and extension principles. In: L. Debnath (ed.) Wavelet Transforms and Time-Frequency Signal Analysis, pp. 1–36. Birkhäuser, Boston (2001)
21. Bernal-Rusiel, J.L., Atienza, M., Cantero, J.L.: Detection of focal changes in human cortical thickness: Spherical wavelets versus Gaussian smoothing. NeuroImage **41**(4), 1278–1292 (2008)
22. Blackadar, B.: *K*-Theory for Operator Algebras, Volume 5. Mathematical Sciences Research Institute Publications, Cambridge University Press, Cambridge, U.K. (1998)
23. Blanchard, J.: Minimally supported frequency composite dilation Parseval frame wavelets. To appear in J. Geom. Anal.
24. Bochner, S., Phillips, R.S.: Absolutely convergent Fourier expansions for non-commutative normed rings. Ann. Math. **43**, 409–418 (1942)
25. Bogdanova, I., Vandergheynst, P., Antoine, J.P., Jacques, L., Morvidone, M.: Stereographic wavelet frames on the sphere. Appl. Comput. Harmon. Anal. **9**(2), 223–252 (2005)
26. Bölcskei, H., Duhamel, P., Hleiss, R.: Design of pulse shaping OFDM/OQAM systems for high data-rate transmission over wireless channels. Proc. IEEE Int. Conf. Communications (ICC) **1**, 559–564 (1999)
27. Bonsall, F.F., Duncan, J.: Complete Normed Algebras. Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 80. Springer-Verlag, New York (1973)
28. de Boor, C.: A Practical Guide to Splines. Springer-Verlag, New York (2001)
29. Boulkhemair, A.: Remarks on a Wiener type pseudodifferential algebra and Fourier integral operators. Math. Res. Lett. **4**(1), 53–67 (1997)
30. Brandenburg, L.H.: On identifying the maximal ideals in Banach algebras. J. Math. Anal. Appl. **50**, 489–510 (1975)
31. Brink, D.M., Satchler, G.R.: Angular Momentum, 3rd ed. Clarendon Press, Oxford (1993)
32. Bülow, T.: Multiscale image processing on the sphere. In: L. Van Gool (ed.) Pattern Recognition, 24th DAGM Symposium, *Lecture Notes in Computer Science*, vol. 2449, pp. 609–617. Springer-Verlag (2002)
33. Bülow, T.: Spherical diffusion for 3D surface smoothing. IEEE Trans. Patt. Anal. Mach. Intell. **26**(12), 1650–1654 (2004)
34. Bülow, T.B.T.: Spherical diffusion. Tech. Rep. MS-CIS-01-38, University of Pennsylvania, Department of Computer and Information Science (2001)
35. Butzer, P., Schmeisser, G., Stens, R.: An introduction to sampling analysis. In: F.A. Marvasti (ed.) Nonuniform Sampling: Theory and Practice. Kluwer Academic/Plenum Publishers, New York (2001)
36. Butzer, P.L., Lei, J.: Approximation of signals using measured sampled values and error analysis. Commun. Appl. Math. **4**, 245–255 (2000)
37. Butzer, P.L., Splettstößer, W., Stens, R.L.: The sampling theorem and linear prediction in signal analysis. Jber. d. Dt. Math.-Verein. **90**, 1–70 (1988)
38. Candès, E., Donoho, D.: New tight frames of curvelets and optimal representations of objects with C^2 singularities. Commun. Pure Appl. Math. **56**, 219–266 (2004)
39. Candès, E.J.: Ridgelets: theory and applications. Ph.D. thesis, Stanford University (1998)
40. Candès, E.J., Donoho, D.L.: Curvelets and curvilinear integrals. J. Approx. Theory **113**, 59–90 (2001)
41. Candès, E.J., Donoho, D.L.: Continuous curvelet transform. I. Resolution of the wavefront set. Appl. Comput. Harmon. Anal. **19**(2), 162–197 (2005)
42. Candès, E.J., Donoho, D.L.: Continuous curvelet transform. II. Discretization and frames. Appl. Comput. Harmon. Anal. **19**(2), 198–222 (2005)
43. Casazza, P.G.: The art of frame theory. Taiwanese J. Math. **4**, 129–201 (2000)
44. Casazza, P.G., Christensen, O.: Weyl–Heisenberg frames for subspaces of $L^2(\mathbb{R})$. Proc. Amer. Math. Soc. **129**, 145–154 (2001)
45. Casazza, P.G., Janssen, A., Christensen, O.: Classifying tight Weyl–Heisenberg frames. Contemp. Math **247**, 131–148 (1999)

46. Chandrasekharan, K.: Classical Fourier Transforms. Springer-Verlag, Berlin (1989)
47. Christensen, O.: An Introduction to Frames and Riesz Bases. Birkhäuser, Boston (2003)
48. Christensen, O.: Pairs of dual Gabor frames with compact support and desired frequency localization. Appl. Comput. Harmon. Anal. **20**, 403–410 (2006)
49. Christensen, O.: Frames and Bases: An Introductory Course (Applied and Numerical Harmonic Analysis). Birkhäuser, Boston (2008)
50. Christensen, O., Kim, R.: On dual Gabor frame pairs generated by polynomials (2008) Preprint, J. Fourier Anal. Appl. To appear 2009
51. Chui, C.: Wavelets—A Tutorial in Theory and Practice. Academic Press, San Diego (1992)
52. Chui, C., He, W., Stöckler, J.: Compactly supported tight and sibling frames with maximum vanishing moments. Appl. Comput. Harmon. Anal. **13**(3), 226–262 (2002)
53. Cohen, A., Daubechies, I., Vial, P.: Wavelets on the interval and fast wavelet transforms. Appl. Comput. Harmon. Anal. **1**, 54–81 (1993)
54. Connes, A.: C^* algèbres et géométrie différentielle. C. R. Acad. Sci. Paris Sér. A-B **290**, A599–A604 (1980)
55. Connes, A.: Noncommutative Geometry. Academic Press, San Diego (1994)
56. Conway, J.B.: A Course in Functional Analysis, 2nd ed. Springer-Verlag, New York (1990)
57. Córdoba, A., Fefferman, C.: Wave packets and Fourier integral operators. Commun. Partial Diff. Eq. **3**, 979–1005 (1978)
58. Czaja, W.: Characterizations of Gabor systems via the Fourier transform. Collect. Math. **51**(2), 205–224 (2000)
59. Daubechies, I.: The wavelet transformation, time-frequency localization and signal analysis. IEEE Trans. Inform. Theory **36**, 961–1005 (1990)
60. Daubechies, I.: Ten Lectures on Wavelets. CBMS-NSF Regional Conference series in applied mathematics. Society for Industrial and Applied Mathematics, Philadelphia (1992)
61. Daubechies, I., Han, B., Ron, A., Shen, Z.: Framelets: MRA-based constructions of wavelet frames. Appl. Comput. Harmon. Anal. **14**(1), 1–46 (2003)
62. Davies, E.: Machine Vision: Theory, Algorithms and Practicalities. Academic Press, San Diego (1990)
63. de Leeuw, K.: An harmonic analysis for operators. I. Formal properties. Illinois J. Math. **19(4)**, 593–606 (1975)
64. Demanet, L., Vandergheynst, P.: Gabor wavelets on the sphere. In: M.A. Unser, A. Aldroubi, A.F. Laine (eds.) Proc. SPIE Conf. Wavelets: Applications in Signal and Image Processing X, vol. 5207, pp. 208–215. SPIE, Bellingham, WA (2003)
65. DeVore, R.A.: Nonlinear approximation. In: A. Iserles (ed.) Acta Numerica, pp. 51–151. Cambridge University Press (1998)
66. DeVore, R.A., Lorentz, G.G.: Constructive Approximation. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 303. Springer-Verlag, Berlin (1993)
67. Do, M.N., Vetterli, M.: The finite ridgelet transform for image representation. IEEE Trans. Image Process. **12**, 16–28 (2003)
68. Do, M.N., Vetterli, M.: The contourlet transform: An efficient directional multiresolution image representation. IEEE Trans. Image Process. **14**, 2091–2106 (2005)
69. Donoho, D.L.: De-noising by soft-thresholding. IEEE Trans. Inform. Theory **4**(3), 613–627 (1995)
70. Donoho, D.L.: Wedgelets: Nearly minimax estimation of edges. Ann. Stat. **27**(3), 859–897 (1999)
71. Donoho, D.L., Huo, X.: Beamlets and multiscale image analysis. In Multiscale and Multiresolution Methods. *Lect. Notes Comput. Sci. Eng.*, 20, Springer-Verlag, Berlin (2002)
72. Donovan, G., Geronimo, J., Hardin, D.: Compactly supported, piecewise affine scaling functions on triangulations. Constr. Approx. **16**, 201–219 (2000)
73. Donovan, G., Geronimo, J.S., Hardin, D.P., Massopust, P.R.: Construction of orthogonal wavelets using fractal functions. SIAM J. Math. Anal. **27**(4), 1158–1192 (1996)
74. Doroshkevich, A.G., Naselsky, P.D., Verkhodanov, O.V., Novikov, D.I., Turchaninov, V.I., Novikov, I.D., Christensen, P.R., Chiang, L.Y.: First release of Gauss–Legendre sky pixelization (GLESP) software package for CMB analysis (2005). ArXiv:astro-ph/0501494v2

75. Doroshkevich, A.G., Naselsky, P.D., Verkhodanov, O.V., Novikov, D.I., Turchaninov, V.I., Novikov, I.D., Christensen, P.R., Chiang, L.Y.: Gauß–Legendre sky pixelation (GLESP) for CMB maps. Int. J. Mod. Phys. D **14**(2), 275–290 (2005)
76. Driscoll, J.R., Healy, D.M.J.: Computing Fourier transforms and convolutions on the 2-sphere. Adv. Appl. Math. **15**(2), 202–250 (1994)
77. Duffin, R., Schaeffer, A.: A class of nonharmonic Fourier series. Trans. Amer. Math. Soc. **72**, 341–366 (1952)
78. Duval-Destin, M., Muschietti, M., Torresani, B.: Continuous wavelet decompositions, multiresolution, and contrast analysis. SIAM J. Math. Anal. **24**(3), 739–755 (1993)
79. Easley, G., Lim, W., Labate, D.: Optimally sparse image representations using shearlets. Proc. 40th Asilomar Conf. on Signals, Systems and Computers, Monterey (2006)
80. Easley, G., Lim, W., Labate, D.: Sparse directional image representations using the discrete shearlet transform. Appl. Comput. Harmon. Anal. **25**, 25–46 (2008)
81. Easley, G. R., Labate, D., and Colonna, F., Shearlet based total variation for denoising. IEEE Trans. Image Process. **18** (2), 260–268 (2009)
82. Easley, G. R., Krim, H., Labate, D. and Yi, S., Edge detection and processing using shearlets. Proceedings of IEEE Int. Conf. on Image Processing (ICIP), San Diego, October 12–15 (2008)
83. Easley, G. R., Krim, H., Labate, D. and Yi, S., A shearlet approach to edge analysis and detection. IEEE Trans. Image Processing **18** (5), 929–941 (2009)
84. Feichtinger, H.: Modulation spaces: Looking back and ahead. Sampl. Theory Signal Image Process. **5**(2), 109–140 (2006)
85. Feichtinger, H., Gröchenig, K.: Gabor frames and time-frequency analysis of distributions. J. Funct. Anal. **146**(2), 464–495 (1997)
86. Feichtinger, H., Strohmer, T. (eds.): Gabor Analysis and Algorithms: Theory and Applications. Birkhäuser, Boston (1998)
87. Feichtinger, H., Strohmer, T. (eds.): Advances in Gabor Analysis. Birkhäuser, Boston (2002)
88. Feichtinger, H.G., Kozek, W.: Quantization of TF lattice-invariant operators on elementary LCA groups. In: Gabor Analysis and Algorithms, pp. 233–266. Birkhäuser, Boston (1998)
89. Fendler, G., Gröchenig, K., Leinert, M.: Symmetry of weighted L^1-algebras and the GRS-condition. Bull. London Math. Soc. **38**(4), 625–635 (2006)
90. Fendler, G., Gröchenig, K., Leinert, M.: Convolution-dominated operators on discrete groups. Integr. Equ. Oper. Theory. **61**, 493–509 (2008)
91. Fendler, G., Gröchenig, K., Leinert, M., Ludwig, J., Molitor-Braun, C.: Weighted group algebras on groups of polynomial growth. Math. Z. **102**(3), 791–821 (2003)
92. Folland, G.B.: Harmonic Analysis in Phase Space. Princeton University Press, Princeton (1989)
93. Folland, G.B.: A Course in Abstract Harmonic Analysis. CRC Press, Boca Raton (1995)
94. Fornasier, M., Gröchenig, K.: Intrinsic localization of frames. Constr. Approx. **22**, 395–415 (2004)
95. Frazier, M., Jawerth, B., Weiss, G.: Littlewood–Paley Theory and the Study of Function Spaces. Regional Conference Series in Mathematics. 79. American Mathematical Society, Providence (1991)
96. Freeden, W., Gervens, T., Schreiner, M.: Constructive Approximation on the Sphere, with Applications to Geomathematics. Numerical Mathematics and Scientific Computation. Clarendon Press, Oxford (1998)
97. Freeden, W., Windheuser, U.: Spherical wavelet transform and its discretization. Adv. Comput. Math. **5**(1), 51–94 (1996)
98. Freeman, W., Adelson, E.: The design and use of steerable filters. IEEE Trans. Patt. Anal. Mach. Intell. **13**(9), 891–906 (1991)
99. Friedrich, F., Demaret, L., Führ, H., Wicker, K.: Efficient moment computation over polygonal domains with an application to rapid wedgelet approximation. SIAM J. Scientific Computing **29**(2), 842–863 (2007)
100. Gel'fand, I.: To the theory of normed rings. II. On absolutely convergent trigonometrical series and integrals. C. R. (Doklady) Acad. Sci. URSS (N.S.) **25**, 570–572 (1939)

101. Gel'fand, I.: Über absolut konvergente trigonometrische Reihen und Integrale. Rec. Math. [Mat. Sbornik] N. S. **9**(51), 51–66 (1941)
102. Gel'fand, I., Raikov, D., Shilov, G.: Commutative Normed Rings. Chelsea Publishing, New York (1964)
103. Geronimo, J.S., Hardin, D.P., Massopust, P.R.: Fractal functions and wavelet expansions based upon several scaling functions. J. Approx. Theory **78**(3), 373–401 (1994)
104. Gohberg, I., Kaashoek, M.A., Woerdeman, H.J.: The band method for positive and strictly contractive extension problems: An alternative version and new applications. Integr. Equ. Oper. Theory **12**(3), 343–382 (1989)
105. Górski, K.M., Hivon, E., Banday, A.J., Wandelt, B.D., Hansen, F.K., Reinecke, M., Bartelmann, M.: HEALPix: A framework for high-resolution discretization and fast analysis of data distributed on the sphere. Astrophys. J. **622**(2), 759–771 (2005)
106. Gressman, P., Labate, D., Weiss, G., Wilson, E.N.: Affine, quasi-affine and co-affine wavelets. In: J. Stoeckler, G. Welland (eds.) Beyond Wavelets, Chap. 8, pp. 215–224. Academic Press, San Diego (2003)
107. Gripenberg, G.: A necessary and sufficient condition for the existence of a father wavelet. Studia Math. **114**, 197–226 (1995)
108. Gröchenig, K.: Foundations of Time-Frequency Analysis. Birkhäuser, Boston (2001)
109. Gröchenig, K.: Localization of frames, Banach frames, and the invertibility of the frame operator. J. Fourier Anal. Appl. **10**(2) (2004)
110. Gröchenig, K.: Composition and spectral invariance of pseudodifferential operators on modulation spaces. J. Anal. Math. **98**, 65–82 (2006)
111. Gröchenig, K.: A pedestrian's approach to pseudodifferential operators. In: Harmonic Analysis and Applications, pp. 139–169. Appl. Numer. Harmon. Anal., Birkhäuser, Boston (2006)
112. Gröchenig, K.: Time-frequency analysis of Sjöstrand's class. Revista Mat. Iberoam. **22**(2), 703–724 (2006)
113. Gröchenig, K.: Gabor frames without inequalities. Int. Math. Res. Not. IMRN **23**, Art. ID rnm111, 21 (2007)
114. Gröchenig, K.: A pedestrian's approach to pseudodifferential operators. In: L. Rodino, M.W. Wong (eds.) Pseudodifferential Operators: Partial Differential Equations and Time-Frequency Analysis, pp. 343–366. Vol. 52, Fields Institute Comm. (2007)
115. Gröchenig, K.: Spectral invariance of non-commutative tori and pseudodifferential operators (2009). In preparation
116. Gröchenig, K., Heil, C.: Modulation spaces and pseudodifferential operators. Integr. Equ. Oper. Theory **34**(4), 439–457 (1999)
117. Gröchenig, K., Janssen, A., Kaiblinger, N., Pfander, G.: Note on B-splines, wavelet scaling functions, and Gabor frames. IEEE Trans. Inform. Theory **49**(12), 3318–3320 (2003)
118. Gröchenig, K., Leinert, M.: Wiener's lemma for twisted convolution and Gabor frames. J. Amer. Math. Soc. **17**(1), 1–18 (2004)
119. Gröchenig, K., Leinert, M.: Symmetry and inverse-closedness of matrix algebras and functional calculus for infinite matrices. Trans. Amer. Math. Soc. **358**(6), 2695–2711 (2006)
120. Gröchenig, K., Rzeszotnik, Z.: Almost diagonalization of pseudodifferential operators. Ann. Inst. Fourier **58**, 2279–2314 (2008)
121. Gröchenig, K., Rzeszotnik, Z., Strohmer, T.: Quantitative analysis of the finite section method (2006). Preprint
122. Guo, K., Kutyniok, G., Labate, D.: Sparse multidimensional representations using anisotropic dilation and shear operators. In: G. Chen, M. Lai (eds.) Wavelets and Splines, pp. 189–201. Nashboro Press, Nashville (2006)
123. Guo, K., Labate, D.: Some remarks on the unified characterization of reproducing systems. Collect. Math. **57**, 279–293 (2006)
124. Guo, K., Labate, D.: Optimally sparse multidimensional representation using shearlets. SIAM J. Math. Anal. **39**, 298–318 (2007)
125. Guo, K., Labate, D.: Sparse shearlet representation of Fourier integral operators. ERA-MS **14**, 7–19 (2007)

126. Guo, K., Labate, D.: Representation of Fourier integral operators using shearlets. J. Fourier Anal. Appl. **14**, 327–371 (2008)
127. Guo, K., Labate, D., and Lim, W., Edge analysis and identification using the continuous shearlet transform. Appl. Comput. Harmonic Anal. **27** (1), 24–46 (2009)
128. Guo, K. and Labate, D. Characterization and analysis of edges using the continuous shearlet transform. To appear in SIAM on Imaging Sciences (2008).
129. Guo, K., Lim, W., Labate, D., Weiss, G., Wilson, E.: Wavelets with composite dilations. Electron. Res. Announc. Amer. Math. Soc. **10**, 78–87 (2004)
130. Guo, K., Lim, W., Labate, D., Weiss, G., Wilson, E.: Wavelets with composite dilations and their MRA properties. Appl. Comput. Harmon. Anal. **20**, 220–236 (2006)
131. Haar, A.: Zur Theorie der orthogonalen Funktionensystme. Math. Ann. **69**, 331–371 (1910)
132. Hardin, D., Hong, D.: On the constructions of wavelets and prewavelets over a triangulation. J. Comput. Appl. Math. **155**(1), 91–109 (2003)
133. Healy, D.J., Kostelec, P., Rockmore, D.: Towards safe and effective high-order Legendre transforms with applications to FFTs for the 2-sphere. Adv. Comput. Math. **21**(1-2), 59–105 (2004)
134. Healy, D.M.J., Rockmore, D.N., Kostelec, P.J., Moore, S.: FFTs for the 2-sphere-improvements and variations. J. Fourier Anal. Appl. **9**(4), 341–385 (2003)
135. Heil, C., Walnut, D.: Continuous and discrete wavelet transforms. SIAM Rev. **31**, 628–666 (1989)
136. Helgason, S.: Groups and Geometric Analysis. Academic Press, San Diego (1984)
137. Hernández, E., Labate, D., Weiss, G.: A unified characterization of reproducing systems generated by a finite family, II. J. Geom. Anal. **12**(4), 615–662 (2002)
138. Hernández, E., Labate, D., Weiss, G., Wilson, E.: Oversampling, quasi affine frames and wave packets. Appl. Comput. Harmon. Anal. **16**, 111–147 (2004)
139. Hewitt, E., Ross, K.A.: Abstract Harmonic Analysis. Vol. I, Structure of Topological Groups, Integration Theory, Group Representations, 2nd ed. Springer-Verlag, Berlin (1979)
140. Holschneider, M.: Continuous wavelet transforms on the sphere. J. Math. Phys. **37**(8), 4156–4165 (1996)
141. Hörmander, L.: The Analysis of Linear Partial Differential Operators. III: Pseudo-differential Operators, Grundlehren der Mathematischen Wissenschaften, vol. 274. Springer-Verlag, Berlin (1994)
142. Hulanicki, A.: On the spectral radius of Hermitian elements in group algebras. Pacific J. Math. **18**, 277–288 (1966)
143. Hulanicki, A.: On the spectrum of convolution operators on groups with polynomial growth. Invent. Math. **17**, 135–142 (1972)
144. Ionascu, E., Wang, Y.: Simultaneous translations and multiplicative tiling and wavelet sets in $\mathbb{R}^n$. Preprint
145. Janssen, A.: Signal analytic proof of two basic results on lattice expansions. Appl. Comput. Harmon. Anal. **1**(4), 350–354 (1994)
146. Janssen, A.: The duality condition for Weyl–Heisenberg frames. In [86] pp. 453–488 (1998)
147. Janssen, A.J.E.M.: Duality and biorthogonality for Weyl–Heisenberg frames. J. Fourier Anal. Appl. **1**(4), 403–436 (1995)
148. Jenkins, J.W.: Folner's condition for exponentially bounded groups. Math. Scand. **35**, 165–174 (1974)
149. Jenkins, J.W.: Representations of exponentially bounded groups. Amer. J. Math. **98**(1), 29–38 (1976)
150. Kaniuth, E.: A Course in Commutative Banach Algebras. Springer, New York (2009)
151. Katznelson, Y.: An Introduction to Harmonic Analysis. John Wiley & Sons, Inc., New York (1968)
152. Kostelec, P.J., Rockmore, D.N.: FFTs on the rotation group. Tech. Rep. 03-11-060, Santa Fe Institute (2003)
153. Kovačević, J., Vetterli, M.: Nonseparable multidimensional perfect reconstruction filter banks and wavelet bases for $\mathscr{R}^n$. IEEE Trans. Inform. Th., sp. iss. Wavelet Transforms and Multiresolution Signal Analysis **38**(2), 533–555 (1992)

154. Kryshtal, I., Robinson, B., Weiss, G., Wilson, E.: Compactly supported wavelets with composite dilations. J. Geom. Anal. **17**, 87–96 (2006)
155. Kurbatov, V.G.: Algebras of difference and integral operators. Funktsional. Anal. i Prilozhen. **24**(2), 87–88 (1990)
156. Kutyniok, G. and Labate, D., Resolution of the wavefront set using the continuous shearlet transform. Trans. AMS, **361**, 2719–2754 (2009)
157. Kutyniok, G., Labate, D.: Construction of regular and irregular shearlet frames. J. Wavelet Theory Appl. **1**, 1–10 (2007)
158. Labate, D.: A unified characterization of reproducing systems generated by a finite family. J. Geom. Anal. **12**(3), 469–491 (2002)
159. Larson, D., Dai, X.: Wandering vectors for unitary systems and orthogonal wavelets. Mem. Amer. Math. Soc. **134**(640) (1998)
160. Larson, D., Massopust, P.: Coxeter groups and wavelet sets. Contemp. Math. **451**, 187–218 (2008)
161. Larson, D., Massopust, P., Ólafsson, G.: Three-way tiling sets in two dimensions. To appear in Acta Appl. Math. (2009)
162. Lasser, R.: Introduction to Fourier Series, vol. 199. Marcel Dekker, New York (1996)
163. Laugesen, R.S., Weaver, N., Weiss, G., Wilson, E.: A characterization of the higher dimensional groups associated with continuous wavelets. J. Geom. Anal. **12**, 89–102 (2001)
164. Leptin, H.: A pedestrian's approach to pseudodifferential operators. In: Operator Algebras and Group Representations, Vol. II, pp. 48–61. Monogr. Stud. Math., Pitman, Boston (1984)
165. Lévy, P.: Sur la convergence absolue des séries de Fourier. Compositio Math. **1**, 1–14 (1934)
166. Lewin, B.J.: Nullstellenverteilung ganzer Funktionen. Akademie-Verlag, Berlin (1962)
167. Lindenstrauss, J., Tzafriri, L.: Classical Banach Spaces I, Sequence Spaces. Springer-Verlag, Berlin (1977)
168. Losert., V.: On the structure of groups with polynomial growth. Math. Z. **195**(1), 109–117 (1987)
169. Losert., V.: On the structure of groups with polynomial growth. II. J. London Math. Soc. (2) **63**(3), 640–654 (2001)
170. Louis, A.K., Maass, P., Rieder, A.: Wavelets: Theory and Applications. John Wiley & Sons, Inc. New York (1997)
171. Ludwig, J.: A class of symmetric and a class of Wiener group algebras. J. Funct. Anal. **31**(2), 187–194 (1979)
172. Luef, F.: On spectral invariance of non-commutative tori. In: Operator Theory, Operator Algebras, and Applications, pp. 131–146. Vol. 414, Contemp. Math., Amer. Math. Soc., Providence (1984)
173. Luef, F.: Projective modules. J. Functional Anal. **257**(6), 1921–1946 (2009)
174. Mallat, S.: A theory for multiresolution signal decomposition: The wavelet representation. IEEE Trans. Patt. Anal. Mach. Intell. **11**, 674–693 (1989)
175. Mallat, S.: A Wavelet Tour of Signal Processing. Academic Press, San Diego (1997)
176. Mallat, S., Hwang, W.L.: Singularity detection and processing with wavelets. IEEE Trans. Inform. Theory **38**, 617–643 (1992)
177. Mallat, S., Zhong, S.: Characterization of signals from multiscale edges. IEEE Trans. Patt. Anal. Mach. Intell. **14**(7), 710–732 (1992)
178. Marinucci, D., Pietrobon, D., Balbi, A., Baldi, P., Cabella, P., Kerkyacharian, G., Natoli, P., Picard, D., Vittorio, N.: Spherical needlets for cosmic microwave background data analysis. MNRAS **383**(2), 539–545 (2008)
179. Maslen, D.K., Rockmore, D.N.: Generalized FFTs — A survey of some recent results. In: L.F. et al. (eds.) Groups and Computation II. Workshop on Groups and Computation, June 1995, New Brunswick, NJ, DIMACS, Ser. Discrete Math. Theor. Comput. Sci., vol. 28, pp. 183–237. Amer. Math. Soc. Providence (1997)
180. Maslen, D.K., Rockmore, D.N.: Separation of variables and the computation of Fourier transforms on finite groups. I. J. Amer. Math. Soc. **10**(1), 169–214 (1997)
181. Massopust, P.: Fractal Functions, Fractal Surfaces, and Wavelets. Academic Press, San Diego (1994)

182. Massopust, P.: Multiwavelets: Some approximation-theoretic properties, sampling on the interval, and translation invariance. In: N.M. Chuong et al. (edso.) Harmonic, Wavelet and p-adic Analysis, pp. 37–58. World Scientific Publishing Co., Singapore (2007)
183. Matz, G., Hlawatsch, F.: Time-frequency transfer function calculus (symbolic calculus) of linear time-varying systems (linear operators) based on a generalized underspread theory. J. Math. Phys. **39**(8), 4014–4070 (1998)
184. Matz, G., Schafhuber, D., Gröchenig, K., Hartmann, M., Hlawatsch, F.: Analysis, optimization, and implementation of low-interference wireless multicarrier systems. IEEE Trans. Wireless Commun. **6**(4), 1–11 (2007)
185. McEwen, J., Vielva, P., Wiaux, Y., Barreiro, R., Cayón, L., Hobson, M., Lasenby, A., Martínez-González, E., Sanz, J.: Cosmological applications of a wavelet analysis on the sphere. J. Fourier Anal. Appl. **13**(4), 495–510 (2007)
186. McEwen, J.D., Hobson, M.P., Lasenby, A.N.: A directional continuous wavelet transform on the sphere (2006). ArXiv:astro-ph/0609159v1
187. McEwen, J.D., Hobson, M.P., Lasenby, A.N., Mortlock, D.J.: A high-significance detection of non-Gaussianity in the Wilkinson Microwave Anisotropy Probe 1-yr data using directional spherical wavelets. MNRAS **359**(4), 1583–1596 (2005)
188. McEwen, J.D., Hobson, M.P., Mortlock, D.J., Lasenby, A.N.: Fast directional continuous spherical wavelet transform algorithms. IEEE Trans. Signal Process. **55**(2), 520–529 (2007)
189. McEwen, J.D., Wiaux, Y., Hobson, M.P., Vandergheynst, P., Lasenby, A.N.: Probing dark energy with steerable wavelets through correlation of WMAP and NVSS local morphological measures. MNRAS **384**(4), 1289–1300 (2008)
190. Muschietti, M., Torrésani, B.: Pyramidal algorithms for Littlewood–Paley decompositions. SIAM J. Math. Anal. **26**(4), 925–943 (1995)
191. Naĭmark, M.A.: Normed Rings. Wolters-Noordhoff Publishing, Groningen, The Netherlands (1970)
192. Naĭmark, M.A.: Normed Algebras. Wolters-Noordhoff Publishing, Groningen, The Netherlands (1972)
193. Narcowich, F.J., Petrushev, P., , Ward, J.D.: Localized tight frames on spheres. SIAM J. Math. Anal. **38**(2), 574–594 (2006)
194. Newman, D.J.: A simple proof of Wiener's $1/f$ theorem. Proc. Amer. Math. Soc. **48**, 264–265 (1975)
195. Nikolski, N.: In search of the invisible spectrum. Ann. Inst. Fourier (Grenoble) **49**(6), 1925–1966 (1999)
196. Palmer, T.W.: Banach Algebras and the General Theory of $*$-Algebras. Vol. 2: $*$-Algebras. Cambridge University Press, Cambridge, U.K. (2001)
197. Pool, J.C.T.: Mathematical aspects of the Weyl correspondence. J. Math. Phys. **7**, 66–76 (1966)
198. Rabinovich, V., Roch, S., Silbermann, B.: Limit Operators and their Applications in Operator Theory, Vol. 150, Operator Theory: Advances and Applications. Birkhäuser-Verlag, Basel (2004)
199. Rademacher, H.: Einige Sätze über Reihen von allgemeinen Orthogonalfunktionen. Math. Ann. **87**, 112–138 (1922)
200. Ramanathan, J.: Methods of Applied Fourier Analysis. Birkhäuser, Boston (1998)
201. Rieffel, M.A.: C^*-algebras associated with irrational rotations. Pacific J. Math. **93**(2), 415–429 (1981)
202. Rieffel, M.A.: Projective modules over higher-dimensional noncommutative tori. Canad. J. Math. **40**(2), 257–338 (1988)
203. Rieffel, M.A.: Noncommutative tori—A case study of noncommutative differentiable manifolds. In: Geometric and Topological Invariants of Elliptic Operators (Brunswick, ME, 1988), pp. 191–211. Vol. 105, Contemp. Math., Amer. Math. Soc., Providence (1990)
204. Risbo, T.: Fourier transform summation of Legendre series and D-functions. J. Geodesy **70**(7), 383–396 (1996)
205. Ron, A., Shen, Z.: Frames and stable bases for shift-invariant subspaces of $L_2(\mathbb{R}^d)$. Canad. J. Math. **47**, 1051–1094 (1995)

206. Ron, A., Shen, Z.: Affine systems in $L_2(\mathbb{R}^d)$: The analysis of the analysis operator. J. Funct. Anal. **148**, 408–447 (1997)
207. Ron, A., Shen, Z.: Weyl–Heisenberg systems and Riesz bases in $L^2(\mathbb{R}^d)$. Duke Math. J. **89**, 237–282 (1997)
208. Rosenberg, J.: Noncommutative variations on Laplace's equation (2009). To appear in Analysis and PDE
209. Rudin, W.: Functional Analysis. McGraw-Hill, New York (1973)
210. Schweitzer, L.B.: A nonspectral dense Banach subalgebra of the irrational rotation algebra. Proc. Amer. Math. Soc. **120**(2), 811–813 (1994)
211. Shin, C.E., Sun, Q.: Stability of localized operators (2009). To appear in J. Funct. Anal.
212. Simoncelli, E., Freeman, W., Adelson, E., Heeger, D.: Shiftable multiscale transforms. IEEE Trans. Inform. Theory. **38**(2), 587–607 (1992)
213. Sjöstrand, J.: An algebra of pseudodifferential operators. Math. Res. Lett. **1**(2), 185–192 (1994)
214. Sjöstrand, J.: Wiener type algebras of pseudodifferential operators. In: Séminaire sur les Équations aux Dérivées Partielles, 1994–1995, pp. Exp. No. IV, 21. École Polytech., Palaiseau, France (1995)
215. Speegle, D.: On the existence of wavelets for non–expansive dilation matrices. Collect. Math. **54**, 163–179 (2003)
216. Starck, J.L., Moudden, Y., Abrial, P., Nguyen, M.: Wavelets, ridgelets and curvelets on the sphere. Astronomy and Astrophysics **446**, 1191–1204 (2006)
217. Stein, E.M.: Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals. Princeton University Press, Princeton (1993)
218. Strohmer, T.: Rates of convergence for the approximation of shift-invariant systems in $\ell^2(\mathbb{Z})$. J. Fourier Anal. Appl. **5**(6), 599–616 (2000)
219. Strohmer, T.: Pseudodifferential operators and Banach algebras in mobile communications. Appl. Comput. Harmon. Anal. **20**(2), 237–249 (2006)
220. Strohmer, T., Beaver, S.: Optimal OFDM design for time-frequency dispersive channels. IEEE Trans. Commun. **51**(7), 1111–1122 (2003)
221. Sun, Q.: Wiener's lemma for infinite matrices. Trans. Amer. Math. Soc. **359**(7), 3099–3123 (2007)
222. Tao, T.: A quantitative proof of Wiener's theorem. Preprint, http://www.math.ucla.edu/∼tao/preprints/harmonic.html
223. The MacTutor History of Mathematics Archive: Jean Baptiste Joseph Fourier. http://www-groups.dcs.st-andrews.ac.uk/∼history/Mathematicians/Fourier.html
224. Toft, J.: Subalgebras to a Wiener type algebra of pseudo-differential operators. Ann. Inst. Fourier (Grenoble) **51**(5), 1347–1383 (2001)
225. Toft, J.: Continuity properties for modulation spaces, with applications to pseudo-differential calculus. II. Ann. Global Anal. Geom. **26**(1), 73–106 (2004)
226. Unser, M., Blu, T.: Wavelet theory demystified. IEEE Trans. Signal Process. **51**(2), 470–483 (2003)
227. Van De Ville, D., Blu, T., Unser, M.: On the multidimensional extension of the quincunx subsampling matrix. IEEE Signal Process. Lett. **12**(2), 112–115 (2005)
228. Vandergheynst, P., Gobbers, J.F.: Directional dyadic wavelet transforms: Design and algorithms. IEEE Trans. Image Process. **11**(4), 363–372 (2002)
229. Varshalovich, D., Moskalev, A., Khersonskij, V.: Quantum theory of angular momentum. World Scientific Publishing, Singapore (1989). Transl. from the Russian.
230. Vielva, P., Wiaux, Y., Martínez-González, E., Vandergheynst, P.: Alignment and signed-intensity anomalies in Wilkinson Microwave Anisotropy Probe data. MNRAS **381**(3), 932 (2007)
231. Wandelt, B.D., Górski, K.M.: Fast convolution on the sphere. Phys. Rev. D **63**(12) (2001). Article 123002
232. Wang, X.: The study of wavelets from the properties of their Fourier transforms. Ph.D. Thesis, Washington University in St. Louis (1995)

233. Weiss, G., Labate, D.: Examples of wavelets with composite dilations. Proceedings of a symposium dedicated to M. de Guzman, Univ. Complutense de Madrid, pp. 231–249 (2005)
234. Weiss, G., Wilson, E.: The Mathematical Theory of Wavelets. Proceedings of the NATO–ASI Meeting. Harmonic Analysis 2000 — A Celebration. Kluwer (2001)
235. Wiaux, Y., Jacques, L., Vandergheynst, P.: Correspondence principle between spherical and Euclidean wavelets. The Astrophysical Journal **632**(1), 15–28 (2005)
236. Wiaux, Y., Jacques, L., Vielva, P., Vandergheynst, P.: Fast directional correlation on the sphere with steerable filters. The Astrophysical Journal **652**(1), 820–832 (2006)
237. Wiaux, Y., McEwen, J., Vielva, P.: Complex data processing: Fast wavelet analysis on the sphere. J. Fourier Anal. Appl. **13**(4), 477–493 (2007)
238. Wiaux, Y., McEwen, J.D., Vandergheynst, P., Blanc, O.: Exact reconstruction with directional wavelets on the sphere. MNRAS **388**(2), 770–788 (2008)
239. Wiaux, Y., Vielva, P., Barreiro, R.B., Martínez-González, E., Vandergheynst, P.: Non-Gaussianity analysis on local morphological measures of WMAP data. MNRAS **385**(2), 939–947 (2008)
240. Wiaux, Y., Vielva, P., Martínez-González, E., Vandergheynst, P.: Global universe anisotropy probed by the alignment of structures in the cosmic microwave background. Phys. Rev. Lett. **96**(15) (2006). Article 151303
241. Wiener, N.: Tauberian theorems. Ann. Math. **33**, 1–100 (1932)
242. Wiener, N.: Generalized Harmonic Analysis and Tauberian Theorems. MIT Press, Cambridge (1966)
243. Willett, R.M., Nowak, R.D.: A multiscale approach for recovering edges and surfaces in photon-limited medical imaging. IEEE Trans. Med. Imaging **22**(3), 332–350 (2003)
244. Wojtaszczyk, P.: A Mathematical Introduction to Wavelets. No. 37 in London Mathematical Society Student Texts. Cambridge University Press, Cambridge, U.K. (1997)
245. Yeo, B., Ou, W., Golland, P.: On the construction of invertible filter banks on the 2-sphere. IEEE Trans. Image Process. **17**(3), 283–300 (2008)
246. Young, R.: An Introduction to Nonharmonic Fourier Series. Academic Press, New York [1980 (revised first edition 2001)]
247. Żelazko, W.: On the analytic functions in *p*-normed algebras. Studia Math. **21**, 345–350 (1961/62)
248. Zygmund, A.: Trigonometric Series, vol. I & II, 2nd ed. Cambridge University Press (1988). First Edition Warsaw 1935

Index

Applied and Numerical Harmonic Analysis

J.M. Cooper: *Introduction to Partial Differential Equations with MATLAB* (ISBN 978-0-8176-3967-9)

C.E. D'Attellis and E.M. Fernández-Berdaguer: *Wavelet Theory and Harmonic Analysis in Applied Sciences* (ISBN 978-0-8176-3953-2)

H.G. Feichtinger and T. Strohmer: *Gabor Analysis and Algorithms* (ISBN 978-0-8176-3959-4)

T.M. Peters, J.H.T. Bates, G.B. Pike, P. Munger, and J.C. Williams: *The Fourier Transform in Biomedical Engineering* (ISBN 978-0-8176-3941-9)

A.I. Saichev and W.A. Woyczyński: *Distributions in the Physical and Engineering Sciences* (ISBN 978-0-8176-3924-2)

R. Tolimieri and M. An: *Time-Frequency Representations* (ISBN 978-0-8176-3918-1)

G.T. Herman: *Geometry of Digital Spaces* (ISBN 978-0-8176-3897-9)

A. Procházka, J. Uhlíř, P.J.W. Rayner, and N.G. Kingsbury: *Signal Analysis and Prediction* (ISBN 978-0-8176-4042-2)

J. Ramanathan: *Methods of Applied Fourier Analysis* (ISBN 978-0-8176-3963-1)

A. Teolis: *Computational Signal Processing with Wavelets* (ISBN 978-0-8176-3909-9)

W.O. Bray and Č.V. Stanojević: *Analysis of Divergence* (ISBN 978-0-8176-4058-3)

G.T Herman and A. Kuba: *Discrete Tomography* (ISBN 978-0-8176-4101-6)

J.J. Benedetto and P.J.S.G. Ferreira: *Modern Sampling Theory* (ISBN 978-0-8176-4023-1)

A. Abbate, C.M. DeCusatis, and P.K. Das: *Wavelets and Subbands* (ISBN 978-0-8176-4136-8)

L. Debnath: *Wavelet Transforms and Time-Frequency Signal Analysis* (ISBN 978-0-8176-4104-7)

K. Gröchenig: *Foundations of Time-Frequency Analysis* (ISBN 978-0-8176-4022-4)

D.F. Walnut: *An Introduction to Wavelet Analysis* (ISBN 978-0-8176-3962-4)

O. Bratteli and P. Jorgensen: *Wavelets through a Looking Glass* (ISBN 978-0-8176-4280-8)

H.G. Feichtinger and T. Strohmer: *Advances in Gabor Analysis* (ISBN 978-0-8176-4239-6)

O. Christensen: *An Introduction to Frames and Riesz Bases* (ISBN 978-0-8176-4295-2)

L. Debnath: *Wavelets and Signal Processing* (ISBN 978-0-8176-4235-8)

J. Davis: *Methods of Applied Mathematics with a MATLAB Overview* (ISBN 978-0-8176-4331-7)

G. Bi and Y. Zeng: *Transforms and Fast Algorithms for Signal Analysis and Representations* (ISBN 978-0-8176-4279-2)

J.J. Benedetto and A. Zayed: *Sampling, Wavelets, and Tomography* (ISBN 978-0-8176-4304-1)

E. Prestini: *The Evolution of Applied Harmonic Analysis* (ISBN 978-0-8176-4125-2)

O. Christensen and K.L. Christensen: *Approximation Theory* (ISBN 978-0-8176-3600-5)

L. Brandolini, L. Colzani, A. Iosevich, and G. Travaglini: *Fourier Analysis and Convexity* (ISBN 978-0-8176-3263-2)

W. Freeden and V. Michel: *Multiscale Potential Theory* (ISBN 978-0-8176-4105-4)

O. Calin and D.-C. Chang: *Geometric Mechanics on Riemannian Manifolds* (ISBN 978-0-8176-4354-6)

Applied and Numerical Harmonic Analysis (Cont'd)

J.A. Hogan and J.D. Lakey: *Time-Frequency and Time-Scale Methods* (ISBN 978-0-8176-4276-1)

C. Heil: *Harmonic Analysis and Applications* (ISBN 978-0-8176-3778-1)

K. Borre, D.M. Akos, N. Bertelsen, P. Rinder, and S.H. Jensen: *A Software-Defined GPS and Galileo Receiver* (ISBN 978-0-8176-4390-4)

T. Qian, V. Mang I, and Y. Xu: *Wavelet Analysis and Applications* (ISBN 978-3-7643-7777-9)

G.T. Herman and A. Kuba: *Advances in Discrete Tomography and Its Applications* (ISBN 978-0-8176-3614-2)

M.C. Fu, R.A. Jarrow, J.-Y. J. Yen, and R.J. Elliott: *Advances in Mathematical Finance* (ISBN 978-0-8176-4544-1)

O. Christensen: *Frames and Bases* (ISBN 978-0-8176-4677-6)

P.E.T. Jorgensen, K.D. Merrill, and J.A. Packer: *Representations, Wavelets, and Frames* (ISBN 978-0-8176-4682-0)

M. An, A.K. Brodzik, and R. Tolimieri: *Ideal Sequence Design in Time-Frequency Space* (ISBN 978-0-8176-4737-7)

S.G. Krantz: *Explorations in Harmonic Analysis* (ISBN 978-0-8176-4668-4)

G.S. Chirikjian: *Stochastic Models, Information Theory, and Lie Groups, Volume 1* (ISBN 978-0-8176-4802-2)

C. Cabrelli and J.L. Torrea: *Recent Developments in Real and Harmonic Analysis* (ISBN 978-0-8176-4531-1)

B. Luong: *Fourier Analysis on Finite Abelian Groups* (ISBN 978-0-8176-4915-9)

M.V. Wickerhauser: *Mathematics for Multimedia* (ISBN 978-0-8176-4879-4)

P. Massopust and B. Forster: *Four Short Courses on Harmonic Analysis* (ISBN 978-0-8176-4890-9)