Applied and Numerical Harmonic Analysis

Applied and Numerical Harmonic Analysis

J.M. Cooper: *Introduction to Partial Differential Equations with MATLAB* (ISBN 0-8176-3967-5)

C.E. D'Attellis and E.M. Fernández-Berdaguer: *Wavelet Theory and Harmonic Analysis in Applied Sciences* (ISBN 0-8176-3953-5)

H.G. Feichtinger and T. Strohmer: *Gabor Analysis and Algorithms* (ISBN 0-8176-3959-4)

T.M. Peters, J.H.T. Bates, G.B. Pike, P. Munger, and J.C. Williams: *Fourier Transforms and Biomedical Engineering* (ISBN 0-8176-3941-1)

A.I. Saichev and W.A. Woyczyński: *Distributions in the Physical and Engineering Sciences* (ISBN 0-8176-3924-1)

R. Tolimieri and M. An: *Time-Frequency Representations* (ISBN 0-8176-3918-7)

G.T. Herman: *Geometry of Digital Spaces* (ISBN 0-8176-3897-0)

A. Procházka, J. Uhlíř, P.J. W. Rayner, and N.G. Kingsbury: *Signal Analysis and Prediction* (ISBN 0-8176-4042-8)

J. Ramanathan: *Methods of Applied Fourier Analysis* (ISBN 0-8176-3963-2)

A. Teolis: *Computational Signal Processing with Wavelets* (ISBN 0-8176-3909-8)

W.O. Bray and Č.V. Stanojević: *Analysis of Divergence* (ISBN 0-8176-4058-4)

G.T. Herman and A. Kuba: *Discrete Tomography* (ISBN 0-8176-4101-7)

J.J. Benedetto and P.J.S.G. Ferreira: *Modern Sampling Theory* (ISBN 0-8176-4023-1)

A. Abbate, C.M. DeCusatis, and P.K. Das: *Wavelets and Subbands* (ISBN 0-8176-4136-X)

L. Debnath: *Wavelet Transforms and Time-Frequency Signal Analysis* (ISBN 0-8176-4104-1)

K. Gröchenig: *Foundations of Time-Frequency Analysis* (ISBN 0-8176-4022-3)

D.F. Walnut: *An Introduction to Wavelet Analysis* (ISBN 0-8176-3962-4)

O. Brattelli and P. Jorgensen: *Wavelets through a Looking Glass* (ISBN 0-8176-4280-3)

H.G. Feichtinger and T. Strohmer: *Advances in Gabor Analysis* (ISBN 0-8176-4239-0)

O. Christensen: *An Introduction to Frames and Riesz Bases* (ISBN 0-8176-4295-1)

L. Debnath: *Wavelets and Signal Processing* (ISBN 0-8176-4235-8)

J. Davis: *Methods of Applied Mathematics with a MATLAB Overview* (ISBN 0-8176-4331-1)

G. Bi and Y. Zeng: *Transforms and Fast Algorithms for Signal Analysis and Representations* (ISBN 0-8176-4279-X)

J.J. Benedetto and A. Zayed: *Sampling, Wavelets, and Tomography* (ISBN 0-8176-4304-4)

(Continued after index)

Jeffrey A. Hogan
Joseph D. Lakey

Time–Frequency and Time–Scale Methods

*Adaptive Decompositions,
Uncertainty Principles,
and Sampling*

Birkhäuser
Boston • Basel • Berlin

Jeffrey A. Hogan
University of Arkansas
Department of Mathematical Sciences
Fayetteville, AR 72701
USA

Joseph D. Lakey
New Mexico State University
Department of Mathematical Sciences
Las Cruces, NM 88003-8001
USA

AMS Subject Classifications: 42B05, 42B20, 42B25, 42C10, 42C15, 42C20, 65T50, 65T60, 81Q05, 81Q10, 81Q15, 81Q20, 94Axx

Library of Congress Cataloging-in-Publication Data
Hogan, Jeffrey A., 1963-
Time–frequency and time–scale methods : adaptive decompositions, uncertainty principles, and sampling / Jeffrey A. Hogan, Joseph D. Lakey.
p. cm. – (Applied and numerical harmonic analysis)
Includes bibliographical references and index.
ISBN 0-8176-4276-5 (alk. paper)
1. Time–series analysis. I. Lakey, Joseph D., 1963- II. Title. III. Series.

TK5102.9.H66 2003
519.5'5–dc22

2003063458

ISBN 0-8176-4276-5 Printed on acid-free paper.

Birkhäuser ®

Printed in the United States of America. (SB)

9 8 7 6 5 4 3 2 1 SPIN 10850075

www.birkhauser.com

To Mysie and to Ellen

Contents

Preface

By time–frequency (TF) analysis we mean, loosely, techniques and principles used in signal analysis, PDE and harmonic analysis that combine consideration of spatial/temporal content, on the one hand, and spectral content on the other, in ways that yield more powerful results than from considering the two domains separately. Time–scale analysis encompasses those aspects involving wavelets and other multiscale methods. Wavelets provide particularly useful TF-decompositions, and the first two chapters of this book focus on certain aspects of wavelets. The most appealing aspects of TF-analysis, in our view, are those that bring out connections between deep theoretical principles on the one hand and intended uses of the methods on the other. These aspects include adaptive decompositions that come from viewing wavelets and Fourier bases as special instances of a larger class of building blocks for function and signal decompositions, the sampling methods that can be attached to such building blocks, and the uncertainty principles (UPs) that different decompositions necessarily entail.

This is not a textbook. There are numerous resources, both on wavelets and on time–frequency analysis in general, that are more foundational. David Walnut's *An Introduction to Wavelet Analysis* [356] and Karlheinz Gröchenig's *Foundations of Time–Frequency Analysis* [168] are two excellent texts in this book series. This monograph is really geared toward readers who have some knowledge of the foundations of wavelets and time–frequency analysis.

Although applications are mentioned, the word "methods" in the title refers to "mathematical methods" that readers working in applications will find useful when thinking of how to fine-tune their own analysis tools. The first part of this book (Chapters 1–4) builds up material that, while technical, is more readily accessible to a more diverse audience. The second part of the book (Chapters 5–7) builds on the first, but also makes use of important elements of harmonic analysis.

Conceptually, three main threads intertwine throughout. These are: (i) adaptive decompositions, (ii) methods for passing from continuous to discrete information and vice versa, and (iii) uncertainty principles—theoretical limitations on TF-localization. The following three pictures serve to illustrate in broad—and perhaps familiar—terms the main ideas of the book.

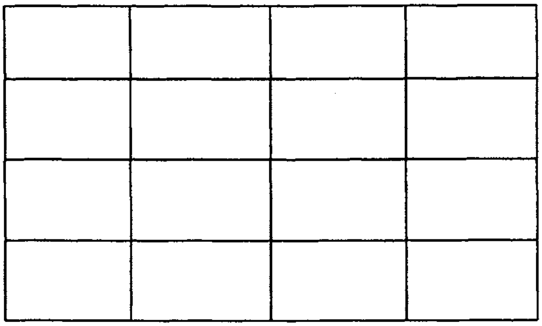

Fig. 0.1. A rectangular or "Gabor" grid

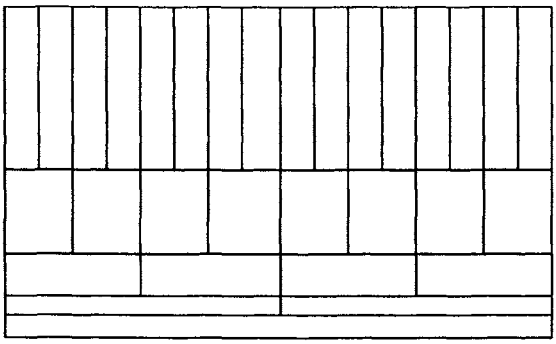

Fig. 0.2. A hyperbolic or "wavelet" grid

Figure 0.1 represents "classical" time–frequency analysis while Figure 0.2 represents wavelet analysis. The dyadic tiling in Figure 0.3 represents one of a vast family of pictures, of which the first two are especially important examples. When interpreted as having unit area, the cells in each picture are called *Heisenberg tiles.* Consider associating to such a picture a family of building blocks—one per cell—from which all signals or functions can be built. The building blocks are thought to be localized on the corresponding cell or tile. Time–frequency analysis addresses a whole raft of issues surrounding the signal or function representations that can arise in this way.

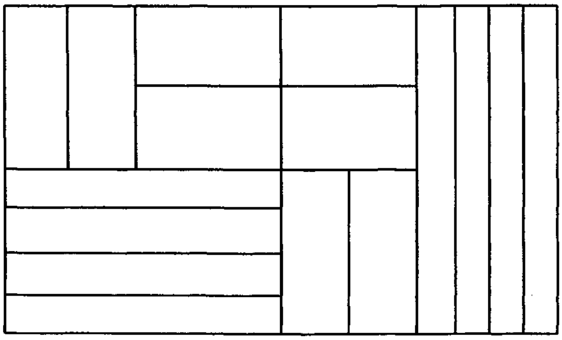

Fig. 0.3. A dyadic tiling

One aspect of the uncertainty principle concerns the possibility of localizing functions on a single rectangle $R = I_R \times \omega_R$ in which I_R and ω_R denote the respective time and frequency intervals of R. If f is well localized about I then its Fourier transform $\hat{f}$ cannot be well localized about ω (and vice versa). Working at Bell Labs in the 1960s, Landau, Slepian and Pollak showed that, when appropriately phrased, the functions that are optimally localized on R are *prolate spheroidal wave functions.* Moreover, when the area $|R|$ of R is large, the dimension of the space of functions "well localized" on R is essentially the area $|R|$ of R.

Functions "localized on R" are said to be essentially time- and bandlimited. Functions f that are truly bandlimited live on a single horizontal strip—one row of rectangles—in Figure 0.1. Shannon's sampling theorem says that, suitably normalized, such an f can be expressed as a sum of its sample values via $f(t) = \sum_k f(k)\ \mathrm{sinc}(t-k)$ where $\mathrm{sinc}(t) = \sin(\pi t)/(\pi t)$. The functions $\mathrm{sinc}(t-k)$ can be thought of as being localized about rectangles $[k, k+1) \times [-1/2, 1/2)$. However, the localization is poor in time. This is the price one must pay for the other nice properties of $\mathrm{sinc}(t)$. Specifically, the functions $\mathrm{sinc}(t-k)$ are orthogonal. Hence, for bandlimited f, $f(k) = \langle f, \mathrm{sinc}(\cdot - k)\rangle$. The orthogonality is best seen by *dualizing* Figure 0.1, i.e., interchanging the roles of f and $\hat{f}$ (rotating by $\pi/2$). Then one is looking at a vertical strip; $\hat{f}$ is localized on this strip, on whose rectangles live the modulated Haar functions $e^{2\pi i n t}\chi_{[0,1)}(t)$, which form an orthonormal basis for $L^2([0,1))$.

By considering all integer translates and modulates $s_{nk}(t) = e^{2\pi i n t}\mathrm{sinc}(t-k)$ of the sinc function or all modulates and shifts $h_{nk}(t) = e^{2\pi i n t}\chi_{[k,k+1)}$ of the Haar function, one obtains orthonormal bases for all of $L^2(\mathbb{R})$. However, the basis elements are either poorly localized in time (s_{nk}) or in frequency (h_{nk}). Can a (Riesz) basis—not necessarily orthonormal—for $L^2(\mathbb{R})$ be formed of functions g_{nk} that are "well localized" on the unit rectangles $[k, k+1) \times [n, n+1)$? The possibility of doing so with unit TF-shifts $e^{2\pi i n x}G(x-k)$ of the

Gaussian $G(x) = e^{-\pi x^2}$ was suggested by Denis Gabor in 1946. This is why Figure 0.1 is often called the *Gabor picture*. However, the TF-shifted Gaussians turn out not to form a Riesz basis. The now famous Balian–Low theorem says that, in fact, no well-localized basis of TF-shifts can exist. Ideas due to Wilson actually furnish bases whose elements are "exponentially localized" on symmetric pairs of time–frequency tiles, but this was not realized until after the wavelet revolution.

Thus, it came as a surprise that by modifying the geometry of the rectangles—replacing the squares in the Gabor picture by the "hyperbolic squares" of Figure 0.2—the so-called *wavelet picture*—one could generate orthonormal bases of $L^2(\mathbb{R})$ whose elements $\psi_{jk}(t) = 2^{j/2}\psi(2^j t - k)$ are well localized about corresponding unit TF-rectangles.

It is constructive to think of the tiling in Figure 0.3 as being obtained from those of the first two figures by a sequence of *recombinations* of adjacent pairs of rectangles, trading *time sibling* tiles sharing the same frequency interval for *frequency siblings* or vice versa. Two specific methods for doing so conform to respective *wavelet packet* and *local trigonometric* bases. Recombination criteria lead to adaptive signal decomposition techniques including best basis algorithms.

The preceding paragraphs give only a broad view of time–frequency analysis. The beauty of its methods comes out best in particular contexts and intricate technical constructions. Once immersed in details it may be difficult to follow the main threads and their connections, several of which are summarized in Figure 0.4, and to discover surprisingly diverse uses of common tools and methods, including the Zak transform (Sections 1.2, 3.5, 4.1, 5.1, 5.7), rearrangements (Sections 5.1, 5.4, 5.6, 6.2, 6.6, 7.3), and combinatorial arguments based on dyadic trees (Sections 3.5, 4.3, 4.5, 6.4, 6.5, 7.2, 7.3, 7.5).

What follows is an outline of the first four chapters, plus some comments pertaining to the remainder of the book. More detailed descriptions can be found in the introductions of the individual chapters.

Though it is not an introduction, Chapter 1 addresses several basic properties of wavelets that crop up later. The technical heart of the chapter lies in Sections 1.1.4 and 1.2. Section 1.1.4 addresses the problem of pointwise regularity of wavelets, elaborating on the discussion in Daubechies' book [99]. Regularity properties of wavelets play a fundamental role in their approximation properties discussed in Chapter 6 and their role in establishing uncertainty relations in Chapter 7. Wavelets having any regularity are associated with multiresolution analyses (MRAs). An MRA is *generated* by a *scaling function* φ whose shifts $\varphi(x-k)$ form a basis for a basic space $V(\varphi)$. When the shifts $\varphi_k = \varphi(x-k)$ are orthogonal, the orthogonal projection of $f \in L^2(\mathbb{R})$ onto $V(\varphi)$ is defined by $f \mapsto \sum_k \langle f, \varphi_k\rangle \varphi_k$. The *Shannon MRA* is generated by the function $\varphi(x) = \mathrm{sinc}(x)$. In this case the space $V(\mathrm{sinc})$ is the space of functions bandlimited to $[-1/2, 1/2)$. As noted above, if $f \in V(\mathrm{sinc})$ then $\langle f, \mathrm{sinc}(\cdot - k)\rangle = f(k)$. The problem of relating expansion coefficients of elements $f \in V(\varphi)$ and samples of f has important ramifications, some of which

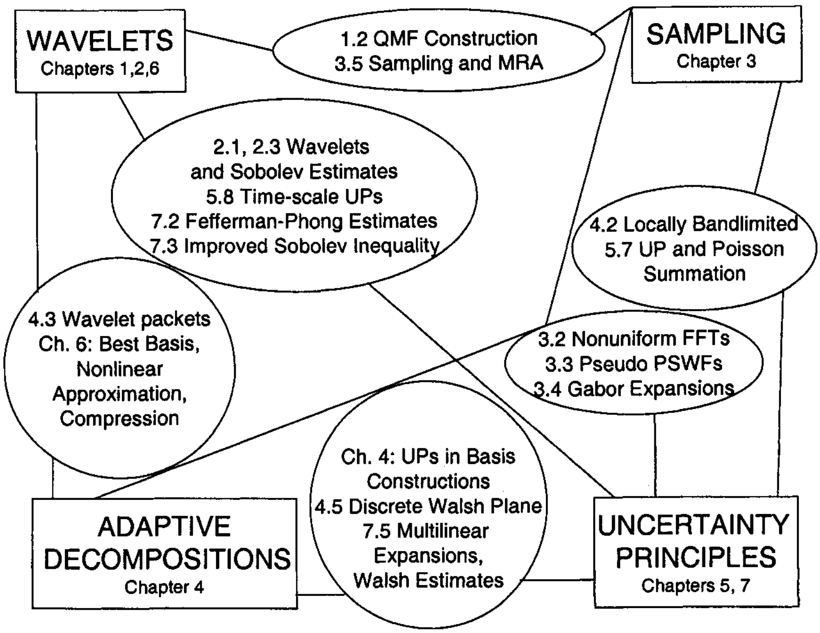

Fig. 0.4. Some relationships among main topics

are considered in Chapter 3. With these issues in mind, we present a construction of MRAs in Section 1.2 displaying some new aspects that boil down to determining the MRA from the integer values of the basic scaling function φ.

In Chapter 2 we investigate further regularity properties of wavelets, with a focus on Sobolev spaces. Although we do not address any specific applications to PDEs, the properties that we consider have their origins in questions of how to apply wavelets in numerical analysis. We start out reviewing the fact that, for suitably regular wavelets, L^2-Sobolev space norms are characterized in terms of magnitudes of wavelet coefficients. Half of this characterization is one form of Sobolev inequality. The role of wavelets in proving sharper Sobolev inequalities and, consequently, their role in establishing sharper forms of the uncertainty principle, is documented in detail later in Chapter 7.

The goal of constructing new wavelets should not simply be to have *more* wavelets—but rather to have ones that can do what could not be done before. One of the major concerns with first-generation wavelets was their inability to line up at boundaries of intervals, thus posing serious questions about their possible application to boundary value problems. Boundary adapted wavelets and biorthogonal wavelets were constructed to address this issue, but they each posed their own tradeoffs. Starting in Section 2.3 we address these concerns by means of the DGHM multiwavelets. These multiwavelets combine symmetry and minimal support properties in crucial ways to produce bases for Sobolev spaces on half-lines and intervals. Implicit to this approach is the question of building, from a dual pair of MRAs, a new pair of MRAs related to the first by differentiating (roughening) one of the dual MRAs and integrating (smoothing) the other.

In Chapter 3 we return to issues of sampling. The focus of the first part of the chapter is on finite sample data using uniform and nonuniform Fourier techniques. Several Fourier-based algorithms for processing sampled data are outlined. Frame expansions play a useful role here in iterative reconstruction algorithms. The second part of Chapter 3 addresses connections between phase space density and the existence of time–frequency localized building blocks for signal approximation and reconstruction. This includes a review of the work of Landau–Slepian–Pollak, identifying prolate spheroidal wave functions (PSWFs) as bases for spaces of approximately time- and bandlimited functions. Approximation properties of Gabor expansions are also considered. The third and last part of Chapter 3 addresses the problem of sampling in multiresolution and, more generally, shift-invariant, spaces. This includes iterative reconstruction as well as interpolation schemes. In the latter case, the full structure of an MRA is shown to play an important and natural role in validating reconstruction from samples. Because of their nestedness, MRA spaces also furnish a natural context for discussing aliasing errors.

Chapter 4 develops specific time–frequency methods that underlie the general tilings represented by Figure 0.3. As the discussion to this point is disproportionately skewed toward the wavelet picture, we first review a construction of the Wilson bases conforming to the Gabor picture. We then review the local trigonometric bases (LTBs) and wavelet packets, which bear parallel relationships to the respective Gabor and wavelet pictures.

Though it is constructive to think of Figure 0.3 as being derived from the respective Gabor and wavelet pictures via recombination, the utility of the corresponding decomposition tools thus derived is obscured by uncertainty issues. The *discrete Walsh plane* provides a model for the time–frequency plane. Orthogonal Walsh functions are indexed by *sequency*—a discrete but imperfect substitute for frequency. Unlike exponentials that must be cut-off in order to achieve time-localization, Walsh functions are truly supported in $[0,1)$. After the substitution of sequency for frequency is accepted, shifted and dilated *Walsh packets* W_P associated with dyadic tiles P can then be interpreted concretely as wavelet packets on the one hand and as a version of LTB functions on the other. The advantage is that there is no uncertainty. Packets associated with disjoint tiles are orthogonal. Thus, one is able to develop precise combinatorial statements, associating time–sequency projections to regions of the phase space. Specifically, if a region $\mathcal{R}$ in the plane can be written as a finite pairwise disjoint union of tiles $P \in \mathcal{P}$, then there is a well-defined orthogonal projection associated to $\mathcal{R}$—defined as the span of those W_P, $P \in \mathcal{P}$ comprising a pairwise disjoint cover of $\mathcal{R}$. Any other pairwise disjoint covering of $\mathcal{R}$ defines the same projection. The Walsh model thus leads to a clean geometrical interpretation of best basis algorithms normally associated with LTBs or wavelet packets.

Important theoretical uses of the ideas surrounding Figure 0.3 are discussed in Chapters 6 and 7. For example, associating LTBs to specific tilings leads to a simple proof of boundedness properties of so-called *exotic* operators discussed in Section 6.8—ones that are actually rather fundamental as regards certain problems in PDE. To some extent, the technique of associating basis functions with tilings of regions of the time–frequency plane has its origins in the problem of *stability of matter*, which boils down to eigenvalue estimates corresponding to bound states of Schrödinger operators. Associating function decompositions to tilings leads to even more sophisticated techniques for estimating multilinear operators that arise naturally in nonlinear PDE. We refer here to Lacey and Thiele's solution of the famous Calderón conjecture [239] regarding boundedness of the bilinear Hilbert transform. The Walsh model for this problem, as well as connections of these ideas to perturbation theory of Schrödinger operators, are also discussed in Chapter 7.

Chapter 5 is about the uncertainty principle as a limitation on joint localization of a function and its Fourier transform. The chapter is intended to serve as a resource on methods for proving uncertainty principles and for establishing relationships among different forms of Fourier uncertainty. A major theme is how to transform statements about joint localization of a function f and its Fourier transform into, sometimes sharper, statements about phase space decay of some *time–frequency distribution* of f.

Chapter 6 addresses consequences of regularity and geometric organization of wavelets in terms of approximation—and thereby compression—of functions and operators. The accompanying techniques culminate in an effective coding scheme for wavelet approximations of signals due to Cohen, Dahmen, Daubechies and DeVore [81]. The same set of techniques lead, surprisingly, to a wavelet proof of a sharp Sobolev inequality (Section 7.3) due to Cohen, DeVore, Petrushev and Yu [85]. On the surface, this inequality bears no relation to signal compression. Other parallels between compression and operator bounds in the wavelet and Gabor pictures are also noted in Chapter 6.

Notational issues and some background results are mentioned in the Appendix. One point is worth mentioning here: in assigning notation, we have tried to be consistent with standard usage in the literature. Because of the breadth of subject matter in this monograph, this approach inevitably leads us to assign multiple meanings to particular symbols such as "$*$" and "δ." The intended usage should always be clear from the context.

Insofar as this is a book about mathematical methods, we have chosen to include proofs only of selected results that illuminate the power of the methods. Many other results are reported in order to sketch the mathematical landscape of which they serve as landmarks. We have omitted or merely outlined proofs of many important results due either to length, technical complication or redundancy. Finally, we are guilty of sins of omission on many counts. We apologize for not including the multitude of brilliant and relevant ideas that we have failed even to mention.

Many persons have contributed to this book in one form or another. Lauren Schultz helped us to get off to an enthusiastic start.

It is a pleasure to thank Luca Capogna, Mark Craddock, Chris Heil, Bill Heller, Loredana Lanzani, Chris Meaney, Sofian Obeidat, Cristina Pereyra, David Walnut and Ying Wang not only for sharing their brilliance but, most importantly, for their friendship and encouragement.

On technical matters we are indebted to Gregory Beylkin, Pete Casazza, Sam Efromovich, Hans Feichtinger, Charly Gröchenig, Palle Jorgensen, Zioma Rseszotnik, Xiaoping Shen, Roy Slaven, Mark Tygert, Chris Weaver and Hong Xiao for discussion and feedback. We also owe special thanks to several anonymous reviewers who set us straight on some important points.

We are grateful to David Donoho et al. for creating WaveLab and to Vasily Strela for creating MWMP which allowed us to produce many of the graphics without a hitch.

We would also like to acknowledge the contributions of several past and present colleagues and mentors, including Lolina Alvarez, Dick Bagby, Charles Chui, Michael Cowling, Garth Gaudry, Doug Kurtz, Alan M^{c}Intosh, John Price, Adam Sikora, Caroline Sweezy and Guido Weiss for their moral support and help in understanding different aspects of the material.

Hogan is grateful for the support of a Macquarie University Research Grant which helped fuel this long-term collaboration with his good friend, and an NSF conference grant which brought the very best of modern applied harmonic analysis to the foothills of the Ozarks.

Lakey is indebted to the Mathematics Department at the University of Texas at Austin, and particularly to Bruce Palka for arranging support for his sabbatical in 2002-2003. He would also like to acknowledge support from the Army Research Office.

We would like to thank the editorial staff at Birkhäuser Boston, particularly, Regina Gorenshteyn, Elizabeth Loew, and Tom Grasso for the amazing job they did in overseeing the development of this project.

Final words of appreciation go to John Benedetto and to John Gilbert, who have profoundly influenced our development as mathematicians. And, of course, to our parents: thank you for making us do our homework.

Fayetteville, Arkansas, *Jeff Hogan*
Las Cruces, New Mexico *Joe Lakey*
October, 2004

Time–Frequency
and
Time–Scale Methods

1

Wavelets: Basic properties, parameterizations and sampling

Wavelets play a fundamental role in decomposing the function spaces—and operators that act on them—that are considered throughout this book. Morlet and colleagues (e.g., [278,279]) coined the term *ondelettes* to describe families of shifted and modulated pulses generated from a single function ψ. Their discovery of the benefits of wavelets in geoexploration was quickly seen as a germination of similar ideas incubating collectively in the mathematics (e.g., Calderón–Zygmund theory), mathematical physics (e.g., coherent states) and electrical engineering (vis-à-vis subband coding) communities. The key features of wavelets that enable their exploitation in discrete signal analysis have long since been distilled into the conceptual framework of a multiresolution analysis (MRA).

Some basic questions regarding wavelets that will be important throughout this book are: what regularity properties can they have, and to what extent does regularity complement or obstruct other desirable properties? Statements regarding the inability of wavelets to possess simultaneous smoothness and localization properties are a manifestation of the uncertainty principle. Nevertheless, wavelets can have some degree of smoothness together with compact support as Daubechies [97] first showed. Another basic question that will be addressed in more detail in Chapter 3 is: how can multiresolution methods be exploited when analyzing sampled data?

This chapter is not intended as an introduction to wavelets. We assume that the reader has some familiarity with them already. The purpose, rather, is to develop properties of wavelets that can be used to address the questions just asked. Nevertheless, some basic discussion is needed both to set notation and to help establish the perspective needed to answer these questions. This discussion, including a brief review of orthogonal and biorthogonal multiresolution analyses, subband coding schemes and computation of scaling functions via the *cascade algorithm*, will comprise the first few parts of Section 1.1.

Daubechies' construction of continuous, compactly supported orthogonal wavelets [97] was a highlight among theoretical developments. Before then, the very existence of such wavelets was suspect. The dependence of regularity—

as measured by local Hölder exponents—on scaling filters was first addressed in contraction estimates due to Daubechies and Lagarias [107, 108] showing, essentially, how the rate of convergence of the cascade algorithm depends on the spectral structure of the scaling filter. Section 1.1.4 contains details of these estimates.

In Section 1.2 we present a "new" way of constructing quadrature mirror filters predicated on a given sequence of integer sample values of the corresponding scaling function. This construction is motivated by the desire to have scaling functions amenable to extrapolation in the sense of Papoulis and Gerchberg [96, 154, 159, 293] and to sampling (see [197, 198, 200, 201] and Chapter 3). Such properties are not easily derived from standard constructions. Technically, this new construction hinges on properties of the Zak transform. The techniques also provide an alternative method for computing the values of the scaling function, as is discussed in Section 1.3.

The notes of this chapter focus largely on two other ways of parameterizing wavelets by building them from basic components. *Pollen's product* provides a means of building orthogonal, compactly supported wavelets based on a certain factorization of unitary matrix-valued Laurent polynomials. A different factorization, based on the Euclidean algorithm, provides a means of building biorthogonal filters from lower-degree factors. This decomposition is the basis for Sweldens' *lifting* construction.

1.1 Scaling and multiresolution analysis

This section contains a more or less standard approach to the construction of a wavelet basis from a multiresolution analysis (MRA) of $L^2(\mathbb{R})$. A two-scale MRA consists of a sequence of closed subspaces $\{V_j\}_{j\in\mathbb{Z}}$ of $L^2(\mathbb{R})$ that are *nested* in the sense that $V_j \subset V_{j+1}$ and, additionally, $f(x) \in V_j$ if and only if $f(2x) \in V_{j+1}$. The spaces also must satisfy: $\cap_j V_j = \{0\}$ while $\overline{\cup_j V_j} = L^2(\mathbb{R})$. The base space V_0 should be *shift-invariant*, in the sense that $f \in V_0$ if and only if $f(\cdot - k) \in V_0$ $(k \in \mathbb{Z})$. Moreover, we insist that there exists a function $\phi \in V_0$ whose integer shifts $\{\phi(\cdot - k)\}_{k\in\mathbb{Z}}$ form a *Riesz basis* for V_0, i.e., there exist constants $A, B > 0$ such that for any sequence $\{a_k\} \in \ell^2(\mathbb{Z})$,

$$A \sum_k |a_k|^2 \le \left\| \sum_k a_k\, \phi(\cdot - k) \right\|_2^2 \le B \sum_k |a_k|^2.$$

We write $V_0 = V(\phi)$ when we wish to emphasize that V_0 is the principal shift-invariant space generated by ϕ (see Chapter 3). The nestedness and Riesz basis properties imply the existence of $\{h_k\} \in \ell^2(\mathbb{Z})$ such that

$$\frac{1}{2}\phi\left(\frac{x}{2}\right) = \sum_k h_k\, \phi(x - k). \tag{1.1}$$

This is known as the *two-scale relation* or *scaling equation* or *dilation equation.* In the Fourier domain this equation becomes

$$\widehat{\phi}(2\xi) = H(\xi)\,\widehat{\phi}(\xi) \tag{1.2}$$

in which the Fourier series $H(\xi) = \sum_k h_k \mathrm{e}^{-2\pi i k\xi}$ is called the *symbol, scaling filter*, refinement mask, etc., of ϕ, which itself is called a *scaling function.* Here we normalize the Fourier transform so that $\widehat{f}(\xi) = \int f(x)\mathrm{e}^{-2\pi i x\xi}\,dx$ when $f \in L^1(\mathbb{R})$. The simplest example of a scaling function ϕ in $L^2(\mathbb{R})$ is $\phi(x) = \chi_{[0,1]}(x)$, the Haar scaling function, with $H(\xi) = \mathrm{e}^{-\pi i\xi}\cos\pi\xi$.

Since the integer shifts of ϕ generate a Riesz basis for $L^2(\mathbb{R})$, a version of Gram–Schmidt can be used to find an element φ of $V(\phi)$ whose shifts form an orthonormal basis for V_0. The idea goes back to Schweinler and Wigner [315] and is discussed by Daubechies (in Section 5.3.1 of [99]). Consider the Gram matrix with entries

$$G(k,l) = \langle\, \phi(\cdot - k),\, \phi(\cdot - l)\,\rangle.$$

By a change of variables, $G(k,l) = G(k-l,0)$, while

$$\begin{aligned} G(k,0) &= \int \widehat{\phi}(\xi)\,\overline{\widehat{\phi}}(\xi)\,\mathrm{e}^{-2\pi i k\xi}\,d\xi \\ &= \sum_l \int_l^{l+1} |\widehat{\phi}(\xi)|^2\,\mathrm{e}^{-2\pi i k\xi}\,d\xi = \int_0^1 \sum_l |\widehat{\phi}(\xi+l)|^2\,\mathrm{e}^{-2\pi i k\xi}\,d\xi. \end{aligned}$$

Thus, $G(k) = G(k,0)$ is the kth Fourier coefficient of the *overlap function* $\Phi(\xi)^2 = \sum_l |\widehat{\phi}(\xi+l)|^2$. That the shifts of ϕ form a Riesz basis for $V(\phi)$ is equivalent to Φ being essentially bounded above and below since, by Plancherel's theorem on $\mathbb{R}$,

$$\Big\|\sum_k a_k\phi(\cdot - k)\Big\|_2^2 = \Big\|\sum_k a_k \mathrm{e}^{-2\pi i k\xi}\widehat{\phi}(\xi)\Big\|_2^2 = \int_0^1 \Big|\sum_k a_k \mathrm{e}^{-2\pi i k\xi}\Big|^2\,\Phi(\xi)^2\,d\xi.$$

The claim then follows from Plancherel's theorem on $\mathbb{T}$.

In fact, the integer shifts of ϕ are orthonormal precisely when $\Phi(\xi) \equiv 1$. Since $1/\Phi$ is bounded and periodic, it can be expressed as the Fourier series of some ℓ^2-sequence $\{v_k\}$. Define $\varphi(x) = \sum_k v_k\phi(x-k)$. Then $\varphi \in V(\phi)$ and the integer shifts of φ form an orthonormal basis for $V(\phi) = V(\varphi)$. We say φ is an *orthogonal generator* of $V_0 = V(\varphi)$.

Although an orthogonal generator always exists, certain properties of an MRA are often easier to deduce by referring to some other generator. For example, in the space $V = \{f \in L^2(\mathbb{R}) \cap C(\mathbb{R}) : f|_{[k,k+1]} \text{ is linear}\}$, the function $\phi(x) = (1-|x|)_+$ serves as a nonorthogonal generator. It is cardinal in the sense that $\phi(k) = \delta_k$ so that each $f \in V$ admits the sampling formula $f(x) = \sum_k f(k)\phi(x-k)$. While there are orthogonal generators of V, none are simultaneously compactly supported and cardinal. In the remainder of this section, though, we consider properties of MRAs predicated on the assumption that φ is an orthogonal generator with scaling filter H.

Suppose now that φ is an orthogonal scaling function. Since $\sum_l |\widehat{\varphi}(\xi + l)|^2 \equiv 1$, (1.2) applied to φ yields

$$\begin{aligned}
1 &= \sum_l |\widehat{\varphi}(2\xi + l)|^2 = \sum_l \left|H\left(\xi + \frac{l}{2}\right)\right|^2 \left|\varphi\left(\xi + \frac{l}{2}\right)\right|^2 \\
&= \sum_l \left(|H(\xi + l)|^2 |\varphi(\xi + l)|^2 + \left|H\left(\xi + l + \frac{1}{2}\right)\right|^2 \left|\varphi\left(\xi + l + \frac{1}{2}\right)\right|^2\right) \\
&= \sum_l \left(|H(\xi)|^2 |\varphi(\xi + l)|^2 + \left|H\left(\xi + \frac{1}{2}\right)\right|^2 \left|\varphi\left(\xi + l + \frac{1}{2}\right)\right|^2\right) \\
&= \left|H(\xi)\right|^2 + \left|H\left(\xi + \frac{1}{2}\right)\right|^2.
\end{aligned}$$

That is, *orthogonality plus scaling implies*

$$\left|H(\xi)\right|^2 + \left|H\left(\xi + \frac{1}{2}\right)\right|^2 \equiv 1. \tag{1.3}$$

For purposes of wavelet construction, one associates to H a filter G such that H and G satisfy $|H(\xi)|^2 + |G(\xi)|^2 \equiv 1$. Then (H, G) is known as a *quadrature mirror filter pair* (QMF). One of the possible choices for G is $G(\xi) = -\mathrm{e}^{-2\pi i \xi} \bar{H}(\xi + 1/2)$. One often refers to H itself as a quadrature mirror filter.

Now consider the converse problem: when does a QMF give rise to an orthogonal scaling function? Iterating (1.2) yields, at least formally,

$$\widehat{\varphi}(\xi) = \prod_{j=1}^{\infty} H\left(\frac{\xi}{2^j}\right). \tag{1.4}$$

For convergence of (1.4) it is necessary that $H(0) = 1$, i.e., $\sum_k h_k = 1$.

The product (1.4) will converge locally uniformly if H is sufficiently well behaved at zero. Suppose, for example, that H is Hölder continuous at zero of some positive order α. That is, there is a $C > 0$ such that

$$\frac{|H(\xi) - H(0)|}{|\xi - 0|^\alpha} = \frac{|H(\xi) - 1|}{|\xi|^\alpha} \leq C.$$

Then, since

$$|\log|H(\xi)|| \leq \log(1 + C|\xi|^\alpha) \leq C'|\xi|^\alpha, \qquad (|\xi| \ll 1)$$

it follows that

$$\sum_{j=N}^{\infty} \left|\log\left|H\left(\frac{\xi}{2^j}\right)\right|\right| \leq C'|\xi|^\alpha \sum_{j=N}^{\infty} 2^{-j\alpha}$$

converges absolutely which, in turn, implies absolute convergence of (1.4). Thus, if H is a trigonometric polynomial or if its coefficients $\{h_k\}$ decay at some rate, then $\widehat{\varphi}$ is well defined as a pointwise product.

Convergence in L^2 of the infinite product in (1.4) then follows from the QMF condition (1.3). Consider the sequence of truncated partial products defined by

$$\widehat{\varphi}_J(\xi) = \prod_{j=1}^{J} H\Big(\frac{\xi}{2^j}\Big)\,\chi_{[-2^{J-1},2^{J-1}]}(\xi). \tag{1.5}$$

Since $\prod_{j=1}^{J} H(\xi/2^j)$ is periodic with period 2^J, it follows that

$$\begin{aligned}
\|\varphi_J\|_2^2 &= \int_{-2^{J-1}}^{2^{J-1}} \prod_{j=1}^{J} \Big|H\Big(\frac{\xi}{2^j}\Big)\Big|^2 d\xi = \int_{-2^{J-1}}^{2^{J-1}} \Big|H\Big(\frac{\xi}{2^J}\Big)\Big|^2 \prod_{j=1}^{J-1} \Big|H\Big(\frac{\xi}{2^j}\Big)\Big|^2 d\xi \\
&= \int_{0}^{2^{J-1}} \left(\Big|H\Big(\frac{\xi}{2^J}\Big)\Big|^2 + \Big|H\Big(\frac{\xi}{2^J}-\frac{1}{2}\Big)\Big|^2\right) \prod_{j=1}^{J-1} \Big|H\Big(\frac{\xi}{2^j}\Big)\Big|^2 d\xi \\
&= \int_{0}^{2^{J-1}} \prod_{j=1}^{J-1} \Big|H\Big(\frac{\xi}{2^j}\Big)\Big|^2 d\xi = \int_{-2^{J-2}}^{2^{J-2}} \prod_{j=1}^{J-1} \Big|H\Big(\frac{\xi}{2^j}\Big)\Big|^2 d\xi = \|\varphi_{J-1}\|_2^2
\end{aligned}$$

because of the QMF condition (1.3). Therefore, by induction from the base case $\|\varphi_1\|_2 = 1$ one has $\|\varphi_J\|_2 = 1$. It follows from the weak-star compactness of the unit ball in $L^2(\mathbb{R})$ that φ_J has a weak-star limit in L^2 which must agree with the pointwise limit defined by the infinite product (1.4).

Thus, any QMF whose coefficients h_k are well behaved gives rise to a scaling function $\varphi \in L^2(\mathbb{R})$. The only remaining question is whether φ defined through (1.3) and (1.4) must be orthogonal to its shifts. In fact, *this is not always the case.* The filter $H(\xi) = (1 + \mathrm{e}^{-6\pi i\xi})/2$ satisfies (1.3) but gives rise through (1.4) to the stretched Haar function $3^{-1/2}\chi_{[0,3]}(x)$ which is not orthogonal to its integer shifts. More sophisticated examples can be found in Daubechies [99] where the pathology is described in more detail. For a trigonometric polynomial H (cf. Proposition 1.4.1 for the more general case), A. Cohen characterized the pathology in the following simple way (cf. Daubechies, [99], p. 188). Let $\tau : [0,1) \to [0,1)$ be given by $\tau(\xi) = 2\xi$ mod 1. A τ-cycle is a collection $\{\xi_1, \dots, \xi_N\} \in [0,1)$ such that $\xi_{j+1} = \tau(\xi_j)$ $1 \le j \le N-1$ and $\xi_1 = \xi_N$. The trivial τ-cycle consists of the single point $\{0\}$. Cohen proved the following.

Theorem 1.1.1. *Suppose that $H(\xi)$ is a trigonometric polynomial satisfying (1.3) with $H(0) = 1$ and φ is the scaling function defined via (1.4). Then φ is orthogonal to its integer shifts if and only if there is no nontrivial τ-cycle $\{\xi_1, \dots, \xi_N\}$ such that $|H(\xi_j)| = 1$ for $1 \le j \le N-1$.*

The criterion of the theorem will be called the *τ-cycle condition.* A separate characterization of this nondegeneracy of φ was found by Lawton [252] (cf. [99], p. 190 and [251]).

Theorem 1.1.2. *Suppose that $H(\xi) = \sum_k h_k e^{-2\pi i k\xi}$ is a trigonometric polynomial satisfying (1.3) with $H(0) = 1$ and φ is the scaling function defined via (1.4). Then φ is orthogonal to its integer shifts when the eigenvalue $1/2$ of the matrix $A_{kl} = \sum_m h_m h_{l-2k+m}$ possesses a one-dimensional eigenspace.*

Failure of the Cohen and Lawton conditions is easy to check for $H(\xi) = (1 + e^{-6\pi i\xi})/2$. Cohen's condition fails because $\{1/3, 2/3\}$ forms a nontrivial τ-cycle. For Lawton's criterion, one has $h_0 = h_3 = 1/2$ and $h_k = 0$ otherwise. Thus $A_{kl} = \delta_{k-2l}/2 + \delta_{k-2l\pm 3}/4$ (where $-2 \le k, l \le 2$). Lawton's condition then fails because $[0,0,1,0,0]^T$ and $[1,1,0,1,1]^T$ are both eigenvectors of A with eigenvalue $1/2$.

1.1.1 Orthonormal wavelet bases for $L^2(\mathbb{R})$

Not every wavelet comes from an MRA, but every MRA gives rise to an orthonormal wavelet. As before, assume that φ is an orthogonal scaling function with QMF $H(\xi) = \sum_k h_k e^{-2\pi i k\xi}$. As the spaces $V_0(\varphi) \subset V_1(\varphi)$ themselves are Hilbert spaces, one can define the relative orthogonal complement $W_0 = (V_0^{\perp}|V_1)$ where the notation denotes the orthogonal complement of V_0 as a subspace of V_1. The space W_0 is *shift-invariant* since $g \in W_0$ implies that $\langle g(\cdot - l), \varphi(\cdot - k)\rangle = \langle g, \varphi(\cdot - (k - l))\rangle = 0$ for all $k, l \in \mathbb{Z}$.

Suppose now that $g(x) = \sum_k c_k \varphi(2x - k) \in W_0$. Setting $C(\xi) = \sum_k c_k e^{-2\pi i k\xi}$ one has, for each $l \in \mathbb{Z}$:

$$\begin{aligned}
\langle g, \varphi(\cdot - l)\rangle &= \int \widehat{g}(\xi)\,\overline{\widehat{\varphi}}(\xi)\, e^{2\pi i l\xi}\, d\xi \\
&= \frac{1}{2}\int C\Big(\frac{\xi}{2}\Big)\, \bar{H}\Big(\frac{\xi}{2}\Big)\, \Big|\widehat{\varphi}\Big(\frac{\xi}{2}\Big)\Big|^2 e^{2\pi i l\xi}\, d\xi \\
&= \int_0^1 C(\xi)\, \bar{H}(\xi)\, e^{4\pi i l\xi} \sum_k |\widehat{\varphi}(\xi + k)|^2\, d\xi \\
&= \int_0^{1/2} \Big(C(\xi)\,\bar{H}(\xi) + C\Big(\xi + \frac{1}{2}\Big)\,\bar{H}\Big(\xi + \frac{1}{2}\Big)\Big) e^{4\pi i l\xi}\, d\xi = 0,
\end{aligned}$$

where we have used the fact that $\sum_k |\widehat{\varphi}(\xi + k)|^2 = 1$ for a.e. ξ. That is, all Fourier coefficients of the $1/2$-periodic function $C(\xi)\bar{H}(\xi) + C(\xi + 1/2)\bar{H}(\xi + 1/2)$ vanish. Thus, $C(\xi)\bar{H}(\xi) = -C(\xi + 1/2)\bar{H}(\xi + 1/2)$ a.e. on $[0, 1/2)$. Since, by (1.3), $H(\xi)$ and $H(\xi + 1/2)$ cannot vanish simultaneously, there is a scalar function M on $[0, 1)$ such that

$$\begin{bmatrix} C(\xi) \\ C(\xi + \frac{1}{2}) \end{bmatrix} = M(\xi) \begin{bmatrix} \bar{H}(\xi + \frac{1}{2}) \\ -\bar{H}(\xi) \end{bmatrix}. \tag{1.6}$$

Moreover, (1.3) implies $|C(\xi)|^2 + |C(\xi + 1/2)|^2 = |M(\xi)|^2$ so $M \in L^2(\mathbb{T})$. Replacing ξ by $\xi + 1/2$ in (1.6), one also has $M(\xi + 1/2) = -M(\xi)$ a.e. so that

$M(\xi) = e^{2\pi i\xi}N(2\xi)$ for some $N \in L^2(\mathbb{T})$. Thus, any $g \in W_0$ can be expressed by $\widehat{g}(\xi) = C(\xi/2)\widehat{\varphi}(\xi/2) = e^{\pi i\xi}\bar{H}(\xi/2 + 1/2)N(\xi)\widehat{\varphi}(\xi/2)$.

Setting $\widehat{\psi}(\xi) = -e^{-\pi i\xi}\bar{H}(\xi/2 + 1/2)\widehat{\varphi}(\xi/2)$, any $g \in W_0$ then satisfies $\widehat{g}(\xi) = N(\xi)\widehat{\psi}(\xi)$ for some $N \in L^2(\mathbb{T})$. In particular, the functions $\{\psi(x - k)\}_{k\in\mathbb{Z}}$ form an orthonormal basis for W_0. Actually, if ψ is defined by $\widehat{\psi}(\xi) = \mu(\xi)\bar{H}(\xi/2 + 1/2)\widehat{\varphi}(\xi/2)$ in which $|\mu(\xi)| \equiv 1$ and $\mu(\xi + 1/2) = -\mu(\xi)$, then the collection $\{\psi(x - k)\}_{k\in\mathbb{Z}}$ will also form an orthonormal basis for W_0. If φ is supported on $[0, M]$ (M odd), then taking $\mu(\xi) = -e^{-\pi i\xi}$ puts the support of ψ in $[(1 - M)/2, (1 + M)/2]$.

From now on we set $G(\xi) = \sum_l (-1)^l \bar{h}_{1-l} e^{-2\pi i l\xi}$ so that $\widehat{\psi}(2\xi) = G(\xi)\widehat{\varphi}(\xi)$. For any $j \in \mathbb{Z}$, let $W_j = (V_j^{\perp}|V_{j+1})$. Arguing just as before, the functions $\psi_{jk}(x) = 2^{j/2}\psi(2^j x - k)$ form an orthonormal basis for W_j. Moreover, if $j \neq j'$ then the spaces W_j and $W_{j'}$ are automatically orthogonal to each other: If, say $j' > j$, then W_j is a subspace of V_{j+1} which is, in turn, a subspace of $V_{j'}$ that is orthogonal to $W_{j'}$. It follows then from the limit properties of the spaces V_j that the functions ψ_{jk} form an orthonormal basis for $L^2(\mathbb{R})$.

1.1.2 Subband coding and FWT

Well before Mallat's discovery of MRA, quadrature mirror filters were utilized by Esteban and Galand [132] to address a concrete application in speech processing. The idea is simple. Start with a sequence $s(k)$ thought of as integer sample values of a continuous-time signal and form the Fourier series $S(\xi) = \sum_k s(k)e^{-2\pi i k\xi}$. The idea of subband coding is essentially to replace S by the subband elements SH and SG which can be processed separately. Reconstruction is achieved by conjugate filtering $SH \mapsto SH\bar{H}$, $SG \mapsto SG\bar{G}$ and adding the components. The QMF condition (1.3) shows that this process recovers S, i.e., $S = SH\bar{H} + SG\bar{G}$. Strang and Nguyen's book [332] is one of numerous excellent sources containing detailed descriptions of subband coding and related algorithms and applications.

Discrete wavelet transform. The spaces V_j ($j \in \mathbb{Z}$) may be thought of as providing a scale of details or resolutions of images and the orthogonal projection P_j onto V_j as an operator that removes details above some level specified by the index j. Elements of V_0 are, roughly speaking, as detailed as φ itself. The space V_1 contains signals that are twice as detailed. The differences (details) between signals in V_1 and their orthogonal projections onto V_0 are contained in the *wavelet space* $W_0 = V_1 \ominus V_0$. This heuristic leads to an efficient, hierarchical decomposition of sequences.

The collection $\{\varphi_{jk}\}_{k\in\mathbb{Z}}$ where $\varphi_{jk}(x) = 2^{j/2}\varphi(2^j x - k)$ forms an orthonormal basis for V_j. Similarly, $\{\psi_{jk}\}_{k\in\mathbb{Z}}$ forms an orthonormal basis for W_j. Hence, orthogonal projections P_j and Q_j of $f \in L^2(\mathbb{R})$ onto these respective spaces are achieved by

$$P_j f = \sum_k \langle f, \varphi_{jk}\rangle \varphi_{jk}; \quad Q_j f = \sum_k \langle f, \psi_{jk}\rangle \psi_{jk}.$$

Given $f \in V_j$, let $c_k^j = c_k^j(f) = \langle f, \varphi_{jk}\rangle$ and $d_k^j = \langle f, \psi_{jk}\rangle$. Since φ satisfies the dilation equation (1.1), it follows that

$$\varphi_{jk}(x) = \sqrt{2}\sum_l h_{l-2k}\,\varphi_{j+1,l}(x).$$

Define filters $\mathcal{H},\ \mathcal{G} : \ell^2(\mathbb{Z}) \to \ell^2(\mathbb{Z})$ by

$$(\mathcal{H}a)_k = \sqrt{2}\sum_l \bar{h}_{l-2k}\,a_l; \quad (\mathcal{G}a)_k = \sqrt{2}\sum_l \bar{g}_{l-2k}\,a_l. \tag{1.7}$$

Notice that $\mathcal{H}$ acts by convolution against the sequence $h_l^* = \bar{h}_{-l}$ followed by decimation/downsampling, and similarly for $\mathcal{G}$. Then

$$c_k^j = \langle f, \varphi_{jk}\rangle = \sqrt{2}\sum_l \bar{h}_{l-2k}\,c_l^{j+1},$$

i.e., $c^j = \mathcal{H}c^{j+1}$. Similarly, $d^j = \mathcal{G}c^{j+1}$. This leads to a hierarchical scheme for the computation of wavelet coefficients, symbolically represented in the following diagram:

$$\begin{tikzcd}
c^j \arrow[r, "\mathcal{H}"] \arrow[dr, "\mathcal{G}"'] & c^{j-1} \arrow[r, "\mathcal{H}"] \arrow[dr, "\mathcal{G}"'] & c^{j-2} \arrow[r, "\mathcal{H}"] \arrow[dr, "\mathcal{G}"'] & c^{j-3} \arrow[r, dashed] \arrow[dr, dashed] & c^L \\
 & d^{j-1} & d^{j-2} & d^{j-3} & d^L
\end{tikzcd}$$

Given the sequence c^j we compute c^{j-1} and d^{j-1} using the decimation and convolution filters $\mathcal{H}$ and $\mathcal{G}$ as in (1.7). Repeating the process on c^{j-1} gives the next layer of coefficients c^{j-2} and d^{j-2}. Continuing gives $c^L = \mathcal{H}^{j-L}c^j$ and $d^L = \mathcal{G}\mathcal{H}^{j-L-1}c^j$. This process is known as the *discrete wavelet transform.*

Inverse discrete wavelet transform. Reconstruction of c^j from c^L, d^L, $d^{L+1}, \ldots, d^{j-1}$ may be achieved with the aid of the filters $\mathcal{H}^*,\ \mathcal{G}^* : \ell^2(\mathbb{Z}) \to \ell^2(\mathbb{Z})$ given by

$$(\mathcal{H}^*a)_k = \sqrt{2}\sum_l h_{k-2l}\,a_l; \quad (\mathcal{G}^*a)_k = \sqrt{2}\sum_l g_{k-2l}\,a_l. \tag{1.8}$$

The filters $\mathcal{H}^*$ and $\mathcal{G}^*$ are ℓ^2-adjoints of $\mathcal{H}$ and $\mathcal{G}$ and act via upsampling followed by convolution, i.e., $(\mathcal{H}^*a)_k = \sqrt{2}(h * \tilde{a})_k$ where

$$\tilde{a}_l = \begin{cases} 0, & \text{if } l \text{ is odd,} \\ a_{l/2}, & \text{if } l \text{ is even,} \end{cases}$$

and similarly for $\mathcal{G}^*$. Now c^{j-1}, d^{j-1} determine c^j via

$$\begin{aligned}
c_k^j = \langle f, \varphi_{jk}\rangle &= \langle P_{j-1}f + Q_{j-1}f, \varphi_{jk}\rangle \\
&= \sum_l c_l^{j-1} \langle \varphi_{j-1,l}, \varphi_{jk}\rangle + \sum_l d_l^{j-1} \langle \psi_{j-1,l}, \varphi_{jk}\rangle \\
&= \sqrt{2} \sum_l h_{k-2l}\, c_l^{j-1} + \sqrt{2} \sum_l g_{k-2l}\, d_l^{j-1} \\
&= (\mathcal{H}^* c^{j-1})_k + (\mathcal{G}^* d^{j-1})_k,
\end{aligned}$$

i.e., $c^j = \mathcal{H}^* c^{j-1} + \mathcal{G}^* d^{j-1}$. Note that $\mathcal{H}\mathcal{H}^* = \mathcal{G}\mathcal{G}^* = I$ on $\ell^2(\mathbb{Z})$ and $\mathcal{H}^*\mathcal{H} + \mathcal{G}^*\mathcal{G} = I$ on $\ell^2(\mathbb{Z})$, an equation equivalent to (1.3). Repeating this process on the sequences c^{j-2} and d^{j-2} gives $c^{j-1} = \mathcal{H}^* c^{j-2} + \mathcal{G}^* d^{j-2}$ and continuing we find

$$c^j = (\mathcal{H}^*)^{j-L}\, c^L + \sum_{m=0}^{j-L-1} (\mathcal{H}^*)^m\, \mathcal{G}^*\, d^{j-m-1}. \tag{1.9}$$

This formula, known as the *discrete inverse wavelet transform*, is represented in the following diagram.

$$\begin{tikzcd}
c^L \arrow[r, "\mathcal{H}^*"] & c^{L+1} \arrow[r, "\mathcal{H}^*"] & c^{L+2} \arrow[r, "\mathcal{H}^*"] & c^{L+3} \arrow[r, dashed] & c^{j-1} \arrow[r, "\mathcal{H}^*"] & c^j \\
d^L \arrow[ur, "\mathcal{G}^*"'] & d^{L+1} \arrow[ur, "\mathcal{G}^*"'] & d^{L+2} \arrow[ur, "\mathcal{G}^*"'] & d^{j-2} \arrow[ur, dashed, "\mathcal{G}^*"'] & d^{j-1} \arrow[ur, "\mathcal{G}^*"'] &
\end{tikzcd}$$

Fast wavelet transform. In order to implement these decompositions in a practical way, one must preprocess signals in a suitable fashion. As in the case of the fast Fourier transform, one can work with periodic signals and periodized basis functions to build a fast algorithm for the wavelet transform and its inverse. Other preprocessing, including truncation or zero padding, will be addressed implicitly in Chapter 2.

Suppose (V_j, φ) is an MRA of $L^2(\mathbb{R})$ with orthogonal generator $\varphi \in L^1 \cap L^\infty$ and ψ is the associated wavelet. We define the periodizations $\varphi_{jk}^{\text{per}}$ and ψ_{jk}^{per} of φ_{jk} and ψ_{jk}, respectively, by

$$\varphi_{jk}^{\text{per}}(x) = \sum_l \varphi_{jk}(x+l); \quad \psi_{jk}^{\text{per}}(x) = \sum_l \psi_{jk}(x+l)$$

and the periodizations V_j^{per} and W_j^{per} of V_j and W_j, respectively, as the closed subspaces of $L^2(\mathbb{T})$ given by

$$V_j^{\text{per}} = \overline{\operatorname{span}}\,\{\varphi_{jk}^{\text{per}}\}; \quad W_j^{\text{per}} = \overline{\operatorname{span}}\,\{\psi_{jk}^{\text{per}}\}.$$

Since $H(1/2) = 0$ and $H(0) = 1$, we have $\sum_k h_{2k} = \sum_k h_{2k+1} = 1/2$. Hence, if $F(x) = \sum_l \varphi(x+l) = \varphi_{0l}^{\text{per}}(x)$, an application of the dilation equation (1.1) gives

$$F(x) = 2\sum_l \sum_k h_k\,\varphi(2x+2l-k) = 2\sum_k \sum_l h_{k+2l}\,\varphi(2x-k) = F(2x). \tag{1.10}$$

However $F \in L^1(\mathbb{T})$, and (1.10) implies that its Fourier coefficients satisfy $\widehat{F}(m) = \widehat{F}(2^j m)$ for non-negative integers j, thus contradicting the Riemann–Lebesgue lemma unless F is constant. Since $1 = \widehat{\varphi}(0) = \int_{-\infty}^{\infty} \varphi(x)\,dx = \int_0^1 F(x)\,dx$, this constant must be 1, i.e., $\sum_l \varphi(x+l) \equiv 1$. As a consequence, for $j \leq 0$, the spaces V_j^{per} are one-dimensional spaces containing only the constant functions. Similarly,

$$\sum_l \psi(x+l/2) = 2\sum_l \sum_k (-1)^k\,\bar{h}_{1-k}\,\varphi(2x+l-k) = 2\sum_k (-1)^k\,\bar{h}_{1-k} = 0$$

from which we see that $W_j^{\mathrm{per}} = \{0\}$ for $j \leq -1$. We need then only concern ourselves with the spaces V_j^{per} and W_j^{per} for $j \geq 0$.

The nestedness of the multiresolution spaces V_j is inherited by their periodizations V_j^{per} as is the orthogonal decomposition $V_j^{\mathrm{per}} = V_{j-1}^{\mathrm{per}} \oplus W_{j-1}^{\mathrm{per}}$. Further, each V_j^{per} has an orthonormal basis $\{\varphi_{jk}^{\mathrm{per}}\}_{k=0}^{2^j-1}$ and each W_j^{per} has an orthonormal basis $\{\psi_{jk}^{\mathrm{per}}\}_{k=0}^{2^j-1}$.

We denote 2^j-periodized versions $h_k^{(j)}$ and $g_k^{(j)}$ of the filter sequences h_k and g_k by $h_k^{(j)} = \sum_l h_{k+2^j l}$ and similarly for $g_k^{(j)}$. Then $g_k^{(j)} = (-1)^k \bar{h}_{1-k}^{(j)}$ where the subscripts are now taken modulo 2^j. Periodized filters $\mathcal{H}^{(j)}$ and $\mathcal{G}^{(j)}$ acting on 2^j-periodic sequences are then defined by

$$(\mathcal{H}^{(j)}a)_k = \sqrt{2}\sum_{l=0}^{2^j-1} \bar{h}_{l-2k}^{(j)}\,a_l; \quad (\mathcal{G}^{(j)}a)_k = \sqrt{2}\sum_{l=0}^{2^j-1} \bar{g}_{l-2k}^{(j)}\,a_l,$$

and their adjoints $(\mathcal{H}^{(j)})^*$ and $(\mathcal{G}^{(j)})^*$ by

$$((\mathcal{H}^{(j)})^*a)_k = \sqrt{2}\sum_{l=0}^{2^j-1} h_{k-2l}^{(j)}\,a_l; \quad ((\mathcal{G}^{(j)})^*a)_k = \sqrt{2}\sum_{l=0}^{2^j-1} g_{k-2l}^{(j)}\,a_l.$$

Given a discrete signal c^j of length 2^j, we compute the signals c^{j-1} and d^{j-1} by $c^{j-1} = \mathcal{H}^{(j)}c^j$, $d^{j-1} = \mathcal{G}^{(j)}c^j$. Both c^{j-1} and d^{j-1} are signals of length 2^{j-1} (or, more precisely, are signals with period 2^{j-1}). Continuing in this way we construct sequences $d^{j-1}, d^{j-2}, \ldots, d^0$ of lengths $2^{j-1}, 2^{j-2}, \ldots, 2, 1$, respectively, and a constant c^0. The sum of the lengths of these sequences is $2^{j-1}+2^{j-2}+\cdots+2+1+1 = 2^j$, the length of the original sequence c^j. When the fast Fourier transform is used to compute the convolutions that appear in the action of the operator $\mathcal{W}_h : \mathbb{C}^{2^j} \to \mathbb{C}^{2^j}$ which decomposes a signal $c^j \in \mathbb{C}^{2^j}$ to the sequence $(d^{j-1}, d^{j-2}, \ldots, d^0, c^0)$, it is easily shown that the algorithm has complexity $O(N \log N)$ where $N = 2^j$ is the length of the data sequence c^j. If the low-pass filter $\{h_k\}_k$ has finite impulse response (FIR) in the sense that $h_k = 0$ if $k < 0$ or $k > M$ for some positive integer M, then the algorithm has complexity $O(MN)$. Under either of these circumstances, $\mathcal{W}_h$ is known as the *fast wavelet transform* (FWT).

In analogy with (1.9), c^j may be recovered from $(d^{j-1}, d^{j-2}, \ldots, d^0, c^0)$ with the aid of the adjoint operators $(\mathcal{H}^{(m)})^*$, $(\mathcal{G}^{(m)})^*$ $(1 \le m \le j-1)$. The operator $\mathcal{W}_h^{-1}$ that implements this mapping is known as the *fast inverse wavelet transform* (FIWT).

The cascade algorithm. There are several schemes for computing the values of the scaling function φ given the QMF H. The first is the spectral method as outlined in (1.4) and subsequent discussion. Another method will be given in Section 1.3. For now we concentrate on the cascade algorithm that arises directly from (1.1). For reasonable H, iterating $\mathcal{H}^*$ starting from the delta sequence provides convergence to the values of φ at dyadic rationals.

Given a scaling function φ with associated QMF $H(\xi) = \sum_k h_k \mathrm{e}^{-2\pi i k \xi}$, define a bounded operator T on $L^2(\mathbb{R})$ by

$$Tf(x) = 2 \sum_k h_k\, f(2x - k).$$

By (1.1), φ is a fixed point of T and the condition $\sum_k h_k = 1$ ensures that T preserves the first moment, i.e., $\int Tf = \int f$. The iterates $\varphi_n = T^n \varphi_0$ $(n \ge 0)$ of $\varphi_0 \in L^2(\mathbb{R})$ having integral one will converge to φ under reasonable hypotheses on H, φ_0.

Alternatively, suppose $\int \varphi = 1$, φ is Hölder continuous of order $\alpha > 0$ and $\int |t|^\alpha |\varphi(t)|\, dt < \infty$. Then if $k/2^J$ is a dyadic rational and j is large enough,

$$\begin{aligned}\left|\varphi\left(\frac{k}{2^J}\right) - 2^{j/2}\langle \varphi, \varphi_{j,k2^{j-J}}\rangle\right| &\le \int \left|\varphi\left(\frac{k}{2^j}\right) - \varphi\left(y + \frac{k}{2^j}\right)\right| 2^j |\varphi(2^j y)|\, dy \\ &\le C2^{-j\alpha}.\end{aligned}$$

Recall that $\mathcal{H}^* : \ell^2(\mathbb{Z}) \to \ell^2(\mathbb{Z})$ acts via $(\mathcal{H}^* a)_k = \sqrt{2} \sum_l h_{k-2l} a_l$. The scaling function φ is unique in $L^2(\mathbb{R})$ with the properties

$$\langle \varphi, \varphi_{0k}\rangle = \delta_k, \qquad \langle \varphi, \psi_{jk}\rangle = 0 \quad (j \ge 0,\ k \in \mathbb{Z}).$$

We now run the inverse discrete wavelet transform on these sequences. Let c^0 be the delta sequence $c_k^0 = \delta_k$ and define sequences c^j by $c^j = (\mathcal{H}^*)^j c^0$ and $d^j = 0$ $(j \ge 0)$. Then

$$\langle \varphi, \varphi_{1k}\rangle = c_k^1 = (\mathcal{H}^* c^0)_k + (\mathcal{G}^* d^0)_k = (\mathcal{H}^* c^0)_k.$$

More generally, $\langle \varphi, \varphi_{jk}\rangle = c_k^j = ((\mathcal{H}^*)^j c^0)_k$ and as a consequence,

$$\left|\varphi\left(\frac{k}{2^J}\right) - 2^{j/2}((\mathcal{H}^*)^j \delta_0)_{k2^{j-J}}\right| = \left|\varphi\left(\frac{k}{2^J}\right) - 2^{j/2} c^j_{k2^{j-J}}\right| \le C2^{-j\alpha}$$