UNITARY MATRIX FUNCTIONS, WAVELET ALGORITHMS, AND STRUCTURAL PROPERTIES OF WAVELETS

Palle E. T. Jorgensen

Department of Mathematics
The University of Iowa
Iowa City, Iowa 52242, U.S.A.
E-mail: jorgen@math.uiowa.edu
URL: http://www.math.uiowa.edu/~jorgen/

Program: "Functional and harmonic analyses of wavelets and frames,"
4–7 August 2004.
Organizers: Judith Packer, Qiyu Sun, Wai Shing Tang
Contribution by Palle E. T. Jorgensen to the Tutorial Sessions:

Abstract: Some connections between operator theory and wavelet analysis: Since the mid eighties, it has become clear that key tools in wavelet analysis rely crucially on operator theory. While isolated variations of wavelets, and wavelet constructions had previously been known, since Haar in 1910, it was the advent of multiresolutions, and sub-band filtering techniques which provided the tools for our ability to now easily create efficient algorithms, ready for a rich variety of applications to practical tasks. Part of the underpinning for this development in wavelet analysis is operator theory. This will be presented in the lectures, and we will also point to a number of developments in operator theory which in turn derive from wavelet problems, but which are of independent interest in mathematics. Some of the material will build on chapters in a new wavelet book, co-authored by the speaker and Ola Bratteli, see http://www.math.uiowa.edu/~jorgen/.

Contents

One cannot expect any serious understanding of what wavelet analysis means without a deep knowledge of the corresponding operator theory.

—Yves Meyer*

1. Introduction

While this series of four lectures will be on the subject of wavelets, the emphasis will be on some interconnections between topics in the mathematics of wavelets and other areas, both within mathematics and outside. Connections to operator theory, to quantum theory, and especially to signal processing will be studied. Concepts such as high-pass and low-pass filters have become synonymous with wavelet tools, but they have also had a significance from the very start of signal processing, for example early telephone signals over transatlantic cables. This was long before the much more recent advances in wavelets which started in the mid-1980's (as a resumption, in fact, of ideas going back to Alfred Haar [Haa10] much earlier).

2000 *Mathematics Subject Classification:* Primary 42C40, 46L60, 47L30, 42A16, 43A65; Secondary 46L45, 42A65, 41A15 .
Key words and phrases: signal processing, matrix functions, infinite products, pyramid algorithm, subdivision algorithm, multiresolution, generalized multiresolution, wavelet packets, library of bases, wavelet filters, high-pass, low-pass filters, filter bank, Gabor frames, fractal measures, wavelet sets, transfer operator, Ruelle operator, Perron–Frobenius, dimension function, homotopy, winding number, index theorem, spectral representation, translation invariance, Hilbert space, biorthogonal wavelet, Cuntz algebra, completely positive map, Fock space, creation operators.
Work supported in part by the U.S. National Science Foundation under grants DMS-9987777 and DMS-0139473(FRG); financial support from the National University of Singapore.
*[Mey00]; see also the web page http://www.math.uiowa.edu/~jorgen/quotes.html.

1.1. *Index of terminology in math and in engineering*

Since the mid-1980's wavelet mathematics has served to some extent as a clearing house for ideas from diverse areas from mathematics, from engineering, as well as from other areas of science, such as quantum theory and optics. This makes the interdisciplinary communication difficult, as the lingo differs from field to field; even to the degree that the same term might have a different name to some wavelet practitioners from what is has to others. In recognition of this fact, Chapter 1 in the recent wavelet book [BrJo02b] samples a little dictionary of relevant terms. Parts of it are reproduced here:

Terminology

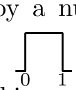
- **multiresolution:** —*real world:* a set of band-pass-filtered component images, assembled into a mosaic of resolution bands, each resolution tied to a finer one and a coarser one.
  —*mathematics:* used in wavelet analysis and fractal analysis, multiresolutions are systems of closed subspaces in a Hilbert space, such as $L^2(\mathbb{R})$, with the subspaces nested, each subspace representing a resolution, and the relative complement subspaces representing the detail which is added in getting to the next finer resolution subspace.
- **matrix function:** a function from the circle, or the one-torus, taking values in a group of N-by-N complex matrices.
- **wavelet:** a function ψ, or a finite system of functions $\{\psi_i\}$, such that for some scale number N and a lattice of translation points on $\mathbb{R}$, say $\mathbb{Z}$, a basis for $L^2(\mathbb{R})$ can be built consisting of the functions $N^{\frac{j}{2}}\psi_i\left(N^j x - k\right)$, $j, k \in \mathbb{Z}$.

  Then dulcet music swelled
  Concordant with the life-strings of the soul;
  It throbbed in sweet and languid beatings there,
  Catching new life from transitory death;
  Like the vague sighings of a wind at even
  That wakes the wavelets of the slumbering sea...
  —Shelley, *Queen Mab*

- **subband filter:** —*engineering:* signals are viewed as functions of time and frequency, the frequency function resulting from a transform of the time function; the frequency variable is broken up into bands, and up-sampling and down-sampling are combined with a filtering of the frequencies in making the connection from one band to the next.

—*wavelets:* scaling is used in passing from one resolution V to the next; if a scale N is used from V to the next finer resolution, then scaling by $\frac{1}{N}$ takes V to a coarser resolution V_1 represented by a subspace of V, but there is a set of functions which serve as multipliers when relating V to V_1, and they are called subband filters.

- **cascades:** —*real world:* a system of successive refinements which pass from a scale to a finer one, and so on; used for example in graphics algorithms: starting with control points, a refinement matrix and masking coefficients are used in a cascade algorithm yielding a cascade of masking points and a cascade approximation to a picture.

  —*wavelets:* in one dimension the scaling is by a number and a fixed simple function, for example of the form  is chosen as the initial step for the cascades; when the masking coefficients are chosen the cascade approximation leads to a scaling function.
- **scaling function:** a function, or a distribution, φ, defined on the real line $\mathbb{R}$ which has the property that, for some integer $N > 1$, the coarser version $\varphi\left(\frac{x}{N}\right)$ is in the closure (relative to some metric) of the linear span of the set of translated functions $\ldots, \varphi(x+1)$, $\varphi(x)$, $\varphi(x-1)$, $\varphi(x-2), \ldots$.
- **logic gates:** —*in computation* the classical logic gates are realized as computers, for example as electronic switching circuits with two-level voltages, say high and low. Several gates have two input voltages and one output, each one allowing switching between high and low: The output of the AND gate is high if and only if both inputs are high. The XOR gate has high output if and only if one of the inputs, but not more than one, is high.
- **qubits:** —*in physics and in computation:* qubits are the quantum analogue of the classical bits 0 and 1 which are the letters of classical computers, the qubits are formed of two-level quantum systems, electrons in a magnetic field or polarized photons, and they are represented in Dirac's formalism $|0\rangle$ and $|1\rangle$; quantum theory allows superpositions, so states $|\psi\rangle = a\,|0\rangle + b\,|0\rangle$, $a, b \in \mathbb{C}$, $|a|^2 + |b|^2 = 1$, are also admitted, and computation in the quantum realm allows a continuum of states, as opposed to just the two classical bits.

—*mathematics:* a chosen and distinguished basis for the two-dimensional Hilbert space $\mathbb{C}^2$ consisting of orthogonal unit vectors, denoted $|0\rangle$, $|1\rangle$.

- **universality:** —*classical computing:* the property of a set of logic gates that they suffice for the implementation of every program; or of a single gate that, taken together with the NOT gate, it suffices for the implementation of every program.

  —*quantum computing:* the property of a set S of basic quantum gates that every (invertible) gate can be written as a sequence of steps using only gates from S. Usually S may be chosen to consist of one-qubit gates and a distinguished tensor gate t. An example of a choice for t is CNOT. An alternative universal one is the Toffoli gate.

  —*mathematics:* the property of a set S of basic unitary matrices that for every n and every $u \in \mathrm{U}_{2^n}(\mathbb{C})$, there is a factorization $u = s_1 s_2 \cdots s_k$, $s_i \in S$, with the understanding that the factors s_i are inserted in a chosen tensor configuration of the quantum register $\underbrace{\mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2}_{n \text{ times}}$. Note that the factors s_i, the number k, and the configuration of the s_i's all depend on n and the gate $u \in \mathrm{U}_{2^n}(\mathbb{C})$ to be studied. The quantum wavelet algorithm (2.2.6) is an example of such a matrix u.
- **chaos:** a small variation or disturbance in the initial states or input of some system giving rise to a disproportionate, or exponentially growing, deviation in the resulting output trajectory, or output data. The term is used more generally, denoting rather drastic forms of instability; and it is measured by the use of statistical devices, or averaging methods.
- $\mathrm{GL}_N(\mathbb{C})$**:** the *general linear group* of all complex $N \times N$ invertible matrices.
- $\mathrm{U}_N(\mathbb{C})$**:** $= \{ A \in \mathrm{GL}_N(\mathbb{C}) \mid AA^* = 1_{\mathbb{C}^N} \}$ where A^* denotes the adjoint matrix, i.e., $(A^*)_{i,j} = \bar{A}_{j,i}$.
- **transfer operator (transition operator):** —*in probability:* An operator which transforms signals s from input s_{in} to output s_{out}. The signals are represented as functions on some set E. In the simplest case, the operator is linear and given in terms of conditional probabilities $p(x,y)$. The number $p(x,y)$ may represent the probability of a transition from y to x where x and y are points in the

set E. Then

$$s_{\text{out}}(x) = \sum_{y \in E} p(x, y)\, s_{\text{in}}(y).$$

—in computation: Let X and Y be functions on a set E, both taking values in $\{0, 1\}$. Let Y be the initial state of the bit, and X the final state of the bit. If the process is governed by a probability distribution P, then the transition probabilities $p(x, y) := P(\{ X = x \mid Y = y \})$ are conditional probabilities: i.e., $p(x, y)$ is the probability of a final bit value x given an initial value y, and we have

$$P(\{X = x\}) = \sum_{y \in E} p(x, y)\, P(\{Y = y\}).$$

—in wavelet theory: Let $N \in \mathbb{Z}_+$, and let W be a positive function on $\mathbb{T} = \{ z \in \mathbb{C} \mid |z| = 1 \}$, for example $W = |m_0|^2$ where m_0 is some low-pass wavelet filter with N bands. (Positivity is only in the sense $W \geq 0$, nonnegative, and the function W may vanish on a subset of $\mathbb{T}$.) Then define a function p on $\mathbb{T} \times \mathbb{T}$ as follows:

$$p(z, w) = \begin{cases} \left(\frac{1}{N}\right) W(w) & \text{if } w^N = z, \\ 0 & \text{for all other values of } w. \end{cases}$$

We arrive at the transfer operator R_W, i.e., the operator transforming functions on $\mathbb{T}$ as follows:

$$s_{\text{out}}(z) = (R_W s_{\text{in}})(z) = \frac{1}{N} \sum_{w^N = z} W(w)\, s_{\text{in}}(w).$$

- **coherence:** *—in mathematics and physics:* The vectors ψ_i that make up a tight frame, one which is not an orthonormal basis, are said to be subjected to *coherence*. So coherent vector systems in Hilbert space are viewed as bases which generalize the more standard concept of orthonormal bases from harmonic analysis. A striking feature of the wavelets with compact support, which are based on scaling, is that the varieties of the two kinds of bases can be well understood geometrically. For example, the collapse of the wavelet orthogonality relations, degenerating into coherent vectors, happens on a subvariety of a lower dimension.
  More generally, coherent vectors in mathematical physics often arise with a continuous index, even if the Hilbert space is separable, i.e., has a countable orthonormal basis. This is illustrated by

a vector system $\{\psi_{r,s}\}$, which should be thought of as a continuous analogue, i.e., a version where a sum gets replaced with an integral

$$C_\psi^{-1} \iint_{\mathbb{R}^2} \frac{dr\,ds}{r^2} \left|\langle\, \psi_{r,s} \mid f \,\rangle\right|^2 = \|f\|^2 .$$

For more details, see also Section 3.3 of [Dau92] and Chapter 3 of [Kai94].

In quantum mechanics, one talks, for example, about coherent states in connection with wavefunctions of the harmonic oscillator. Combinations of stationary wavefunctions from different energy eigenvalues vary periodically in time, and the question is which of the continuously varying wavefunctions one may use to expand an unknown function in without encountering overcompleteness of the basis. The methods of "coherent states" are methods for using these kinds of functions (which fit some problems elegantly) while avoiding the difficulties of overcompleteness. The term "coherent" applies when you succeed in avoiding those difficulties by some means or other. Of course, for students who have just learned about the classic complete orthonormal basis of stationary eigenfunctions, "coherent state" methods at first may seem like a daring relaxation of the rules of orthogonality, so that the term seems to stand for total freedom!

1.1.1. *Some background on Hilbert space*

Wavelet theory is the art of finding a special kind of basis in Hilbert space. Let $\mathcal{H}$ be a Hilbert space over $\mathbb{C}$ and denote the inner product $\langle \cdot \mid \cdot \rangle$. For us, it is assumed linear in the second variable. If $\mathcal{H} = L^2(\mathbb{R})$, then

$$\langle\, f \mid g \,\rangle := \int_{\mathbb{R}} \overline{f(x)}\, g(x)\; dx. \tag{1.1.1}$$

If $\mathcal{H} = \ell^2(\mathbb{Z})$, then

$$\langle\, \xi \mid \eta \,\rangle := \sum_{n\in\mathbb{Z}} \bar{\xi}_n \eta_n. \tag{1.1.2}$$

Let $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$. If $\mathcal{H} = L^2(\mathbb{T})$, then

$$\langle\, f \mid g \,\rangle := \frac{1}{2\pi} \int_{-\pi}^{\pi} \overline{f(\theta)}\, g(\theta)\; d\theta. \tag{1.1.3}$$

Functions $f \in L^2(\mathbb{T})$ have Fourier series: Setting $e_n(\theta) = e^{in\theta}$,

$$\hat{f}(n) := \langle e_n \mid f \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-in\theta} f(\theta)\, d\theta, \tag{1.1.4}$$

and

$$\|f\|_{L^2(\mathbb{T})}^2 = \sum_{n \in \mathbb{Z}} \left| \hat{f}(n) \right|^2. \tag{1.1.5}$$

Similarly if $f \in L^2(\mathbb{R})$, then

$$\hat{f}(t) := \int_{\mathbb{R}} e^{-ixt} f(x)\, dx, \tag{1.1.6}$$

and

$$\|f\|_{L^2(\mathbb{R})}^2 = \frac{1}{2\pi} \int_{\mathbb{R}} \left| \hat{f}(t) \right|^2 dt. \tag{1.1.7}$$

Let J be an index set. We shall only need to consider the case when J is countable. Let $\{\psi_\alpha\}_{\alpha \in J}$ be a family of nonzero vectors in a Hilbert space $\mathcal{H}$. We say it is an *orthonormal basis* (ONB) if

$$\langle \psi_\alpha \mid \psi_\beta \rangle = \delta_{\alpha,\beta} \qquad \text{(Kronecker delta)} \tag{1.1.8}$$

and if

$$\sum_{\alpha \in J} |\langle \psi_\alpha \mid f \rangle|^2 = \|f\|^2 \qquad \text{holds for all } f \in \mathcal{H}. \tag{1.1.9}$$

If only (1.1.9) is assumed, but not (1.1.8), we say that $\{\psi_\alpha\}_{\alpha \in J}$ is a (normalized) *tight frame*. We say that it is a *frame* with *frame constants* $0 < A \le B < \infty$ if

$$A \|f\|^2 \le \sum_{\alpha \in J} |\langle \psi_\alpha \mid f \rangle|^2 \le B \|f\|^2 \qquad \text{holds for all } f \in \mathcal{H}.$$

Introducing the rank-one operators $Q_\alpha := |\psi_\alpha\rangle \langle \psi_\alpha|$ of Dirac's terminology, see [BrJo02b], we see that $\{\psi_\alpha\}_{\alpha \in J}$ is an ONB if and only if the Q_α's are projections and

$$\sum_{\alpha \in J} Q_\alpha = I \qquad (= \text{the identity operator in } \mathcal{H}). \tag{1.1.10}$$

It is a (normalized) tight frame if and only if (1.1.10) holds but with no further restriction on the rank-one operators Q_α. It is a frame with frame constants A and B if the operator

$$S := \sum_{\alpha \in J} Q_\alpha \tag{1.1.11}$$

satisfies

$$AI \leq S \leq BI$$

in the order of hermitian operators. (We say that operators $H_i = H_i^*$, $i = 1, 2$, satisfy $H_1 \leq H_2$ if $\langle f \mid H_1 f \rangle \leq \langle f \mid H_2 f \rangle$ holds for all $f \in \mathcal{H}$).

Wavelets in $L^2(\mathbb{R})$ are generated by simple operations on one or more functions ψ in $L^2(\mathbb{R})$, the operations come in pairs, say scaling and translation, or phase-modulation and translations. If $N \in \{2, 3, \dots\}$ we set

$$\psi_{j,k}(x) := N^{j/2}\psi\left(N^j x - k\right) \qquad \text{for } j, k \in \mathbb{Z}. \tag{1.1.12}$$

1.1.2. *Connections to group theory*

We stress the discrete wavelet transform. But the first line in the two tables below is the continuous one. It is the only treatment we give to the continuous wavelet transform, and the corresponding *coherent vector decompositions*. But, as is stressed in [Dau92], [Kai94], and [KaLe95], the continuous version came first.

Summary of and variations on the resolution of the identity operator 1 in L^2 or in ℓ^2, for ψ and $\tilde{\psi}$ where $\psi_{r,s}(x) = r^{-\frac{1}{2}}\psi\left(\frac{x-s}{r}\right)$, $C_\psi = \int_{\mathbb{R}} \frac{d\omega}{|\omega|} \left|\hat{\psi}(\omega)\right|^2 < \infty$, similarly for $\tilde{\psi}$ and $C_{\psi,\tilde{\psi}} = \int_{\mathbb{R}} \frac{d\omega}{|\omega|} \overline{\hat{\psi}(\omega)}\, \hat{\tilde{\psi}}(\omega)$:

$N = 2$	Overcomplete Basis	Dual Bases
continuous resolution	$C_\psi^{-1} \iint_{\mathbb{R}^2} \frac{dr\, ds}{r^2} \vert\psi_{r,s}\rangle\langle\psi_{r,s}\vert = 1_{L^2}$	$C_{\psi,\tilde{\psi}}^{-1} \iint_{\mathbb{R}^2} \frac{dr\, ds}{r^2} \vert\psi_{r,s}\rangle\langle\tilde{\psi}_{r,s}\vert = 1_{L^2}$
discrete resolution	$\sum_{j\in\mathbb{Z}} \sum_{k\in\mathbb{Z}} \vert\psi_{j,k}\rangle\, \langle\psi_{j,k}\vert = 1_{L^2}$, $\psi_{j,k}$ corresponding to $r = 2^{-j}$, $s = k2^{-j}$	$\sum_{j\in\mathbb{Z}} \sum_{k\in\mathbb{Z}} \vert\psi_{j,k}\rangle\, \langle\tilde{\psi}_{j,k}\vert = 1_{L^2}$
$N \geq 2$	Isometries in ℓ^2	Dual Operator System in ℓ^2
sequence spaces	$\sum_{i=0}^{N-1} S_i S_i^* = 1_{\ell^2}$, where $S_0, \dots, S_{N-1}$ are adjoints to the quadrature mirror filter operators F_i, i.e., $S_i = F_i^*$	$\sum_{i=0}^{N-1} S_i \tilde{S}_i^* = 1_{\ell^2}$, for a dual operator system $S_0, \dots, S_{N-1}$, $\tilde{S}_0, \dots, \tilde{S}_{N-1}$

Consult Chapter 3 of [Kai94] for the continuous resolution, and Section 2.2 of [BrJo02b] for the discrete resolution. If h, k are vectors in a Hilbert space $\mathcal{H}$, then the operator $A = |h\rangle \langle k|$ is defined by the identity $\langle\, u \mid Av \,\rangle = \langle\, u \mid h \,\rangle \langle\, k \mid v \,\rangle$ for all $u, v \in \mathcal{H}$. Then the assertions in the first table amount to:

$C_\psi^{-1} \iint\limits_{\mathbb{R}^2} \dfrac{dr\, ds}{r^2} \left|\langle\, \psi_{r,s} \mid f \,\rangle\right|^2 = \left\|f\right\|_{L^2}^2 \quad \forall\, f \in L^2(\mathbb{R})$	$C_{\psi,\tilde\psi}^{-1} \iint\limits_{\mathbb{R}^2} \dfrac{dr\, ds}{r^2} \langle\, f \mid \psi_{r,s} \,\rangle \langle\, \tilde\psi_{r,s} \mid g \,\rangle = \langle\, f \mid g \,\rangle \quad \forall\, f, g \in L^2(\mathbb{R})$
$\sum\limits_{j\in\mathbb{Z}} \sum\limits_{k\in\mathbb{Z}} \left|\langle\, \psi_{j,k} \mid f \,\rangle\right|^2 = \left\|f\right\|_{L^2}^2 \quad \forall\, f \in L^2(\mathbb{R})$	$\sum\limits_{j\in\mathbb{Z}} \sum\limits_{k\in\mathbb{Z}} \langle\, f \mid \psi_{j,k} \,\rangle \langle\, \tilde\psi_{j,k} \mid g \,\rangle = \langle\, f \mid g \,\rangle \quad \forall\, f, g \in L^2(\mathbb{R})$
$\sum\limits_{i=0}^{N-1} \left\|S_i^* c\right\|^2 = \left\|c\right\|^2 \quad \forall\, c \in \ell^2$	$\sum\limits_{i=0}^{N-1} \langle\, S_i^* c \mid \tilde S_i^* d \,\rangle = \langle\, c \mid d \,\rangle \quad \forall\, c, d \in \ell^2$

A function ψ satisfying the resolution identity is called a *coherent vector* in mathematical physics. The representation theory for the $(ax+b)$-group, i.e., the matrix group $G = \left\{ \left(\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}\right) \mid a \in \mathbb{R}_+,\ b \in \mathbb{R} \right\}$, serves as its underpinning. Then the tables above illustrate how the $\{\psi_{j,k}\}$ wavelet system arises from a discretization of the following unitary representation of G:

$$\left(U_{\left(\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}\right)} f\right)(x) = a^{-\frac{1}{2}} f\left(\frac{x-b}{a}\right) \tag{1.1.13}$$

acting on $L^2(\mathbb{R})$. This unitary representation also explains the discretization step in passing from the first line to the second in the tables above. The functions $\{\, \psi_{j,k} \mid j, k \in \mathbb{Z} \,\}$ which make up a wavelet system result from the choice of a suitable coherent vector $\psi \in L^2(\mathbb{R})$, and then setting

$$\psi_{j,k}(x) = \left(U_{\left(\begin{smallmatrix} 2^{-j} & k\cdot 2^{-j} \\ 0 & 1 \end{smallmatrix}\right)} \psi\right)(x) = 2^{\frac{j}{2}} \psi\left(2^j x - k\right). \tag{1.1.14}$$

Even though this representation lies at the historical origin of the subject of wavelets (see [DGM86]), the $(ax+b)$-group seems to be now largely forgotten in the next generation of the wavelet community. But Chapters 1–3 of [Dau92] still serve as a beautiful presentation of this (now much ignored) side of the subject. It also serves as a link to mathematical physics and to classical analysis.

Since the representation U in (1.1.13) on $L^2(\mathbb{R})$, when a unitary U is defined from (1.1.13) setting $a = 2$, $b = 0$, $(Uf)(x) := 2^{-\frac{1}{2}} f\left(\frac{x}{2}\right)$, leaves invariant the Hardy space

$$\mathcal{H}_+ = \left\{ f \in L^2(\mathbb{R}) \mid \operatorname{supp}(\hat{f}) \subset [0,\infty\rangle \right\}, \tag{1.1.15}$$

formula (1.1.14) suggests that it would be simpler to look for wavelets in $\mathcal{H}_+$. After all, it is a smaller space, and it is natural to try to use the causality features of $\mathcal{H}_+$ implied by the support condition in (1.1.15). Moreover, in the world of the Fourier transform, the two operations of the formulas (1.1.13) and (1.1.14) take the simpler forms

$$\hat{f} \longmapsto a^{\frac{1}{2}} e^{-ibt} \hat{f}(at) \quad \text{and} \quad \hat{\psi} \longmapsto 2^{\frac{j}{2}} e^{-i2^j kt} \hat{\psi}\left(2^j t\right). \tag{1.1.16}$$

So in the early nineties, this was an open problem in the theory, i.e., whether or not there are wavelets in the Hardy space; but it received a beautiful answer in [Aus95]. Auscher showed that there are no wavelet functions ψ in $\mathcal{H}_+$ which satisfy the following mild regularity properties:

- (R_0) $\hat{\psi}$ is continuous;
- (R_ε) for some $\varepsilon \in \mathbb{R}_+$, $\hat{\psi}(t) = \mathcal{O}\left(|t|^\varepsilon\right)$ and $\hat{\psi}(t) = \mathcal{O}\left(\left(1+|t|\right)^{-\varepsilon-\frac{1}{2}}\right)$, $t \in \mathbb{R}$.

Comparison of formulas (1.1.13) and (1.1.14) shows that The traditional discrete wavelet transform may be viewed as the restriction to a subgroup H of a classical unitary representation of G. The unitary representations of G are completely understood: the set of irreducible unitary representations consists of two infinite-dimensional inequivalent subrepresentations of the representation (1.1.13) on $L^2(\mathbb{R})$, together with the one-dimensional representations $\left(\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}\right) \to a^{ik}$ parameterized by $k \in \mathbb{R}$. (The two subrepresentations of (1.1.13) are obtained by restricting to $f \in L^2(\mathbb{R})$ with $\operatorname{supp}\hat{f} \subseteq \langle -\infty, 0]$ and $\operatorname{supp}\hat{f} \subseteq [0,\infty\rangle$, respectively.) However, the subgroup H of G has a rich variety of inequivalent infinite-dimensional representations that do not arise as restrictions of (1.1.13), or of any representation of G. The group H considered in (1.1.14) is a semidirect product (as is G): it is of the form

$$H_N = \left\{ \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} \;\middle|\; a = N^j,\ b = \sum_{i\in\mathbb{Z}} n_i N^i,\ j \in \mathbb{Z},\ n_i \in \mathbb{Z}, \text{ where the } \sum_i \text{ summation is finite} \right\}. \tag{1.1.17}$$

(In the jargon of pure algebra, the nonabelian group H_N is the semidirect product of the two abelian groups $\mathbb{Z}$ and $\mathbb{Z}\left[\frac{1}{N}\right]$, with a naturally defined action of $\mathbb{Z}$ on $\mathbb{Z}\left[\frac{1}{N}\right]$.)

The papers [DaLa98], [Jor01a], [BaMe99], [HLPS99], [LPT01], and [BreJo91] show that it is possible to use these nonclassical representations of H for the construction of unexpected classes of wavelets, the wavelet sets being the most notable ones. Recall that a subset $E \subset \mathbb{R}$ of finite measure is a *wavelet set* if $\hat{\psi} = \chi_E$ is such that, for some $N \in \mathbb{Z}_+$, $N \geq 2$, the functions $\left\{ N^{\frac{j}{2}} \psi\left(N^j x - k\right) \bigm| j, k \in \mathbb{Z} \right\}$ form an orthonormal basis for $L^2(\mathbb{R})$. Until the work of Larson and others, see [DaLa98] and [HLPS99], it was not even clear that wavelet sets E could exist in the case $N > 2$. The paper [LPT01] develops and extends the representation theory for the subgroups H_N independently of the ambient group G and shows that each H_N has continuous series of representations which account for the wavelet sets. The role of the representations of the groups H_N and their generalizations for the study of wavelets was first stressed in [BreJo91].

There is a different transform which is analogous to the wavelet transform of (1.1.13)–(1.1.14), but yet different in a number of respects. It is the Gabor transform, and it has a history of its own. Both are special cases of the following construction: Let G be a nonabelian matrix group with center C, and let U be a unitary irreducible representation of G on the Hilbert space $L^2(\mathbb{R})$. When $\psi \in L^2(\mathbb{R})$ is given, we may define a transform

$$(T_\psi f)(\xi) := \langle U(\xi)\psi \mid f \rangle, \quad \text{for } f \in L^2(\mathbb{R}) \text{ and } \xi \in G/C. \tag{1.1.18}$$

It turns out that there are classes of matrix groups, such as the $ax + b$ group, or the 3-dimensional group of upper triangular matrices, which have transforms T_ψ admitting effective discretizations. This means that it is possible to find a vector $\psi \in L^2(\mathbb{R})$, and a discrete subgroup $\Lambda \subset G/C$, such that the restriction to Λ of the transform T_ψ in (1.1.18) is injective from $L^2(\mathbb{R})$ into functions on Λ.

There are many books on transform theory, and here we are only making the connection to wavelet theory. The book [Per86] contains much more detail on the group-theoretic approach to these continuous and discrete coherent vector transforms.

1.1.3. *Some background on matrix functions in mathematics and in engineering*

One of our coordinates for the landscape of multiresolution wavelets takes the form of a geometric index. In fact, it involves a traditional operator-theoretic index with values in $\mathbb{Z}$. When it is identified with a winding number or a counting of homotopy classes, it serves also as a Fredholm index of an associated Toeplitz operator. An orthogonal dyadic wavelet basis has its wavelet function ψ satisfying the normalization $\|\psi\|_{L^2(\mathbb{R})} = 1$, i.e., ψ is a vector of norm one in the Hilbert space $L^2(\mathbb{R})$. In the lingo of quantum theory, ψ is therefore a pure state, and the x-coordinate is an observable called the position. The integral $E_\psi(x) = \int_{\mathbb{R}} x\,|\psi(x)|^2\,dx$ is the expected value of the position. If ψ_H denotes the standard Haar function in (1.2.15), then clearly $E_{\psi_H}(x) = \frac{1}{2}$. Also note the translation formula $E_{\psi(\,\cdot\,-k)}(x) = E_\psi(x) + k$. We showed in Corollary 2.4.11 of [BrJo02b], completely generally, that the other orthonormal wavelets ψ have expected values in the set $\frac{1}{2} + \mathbb{Z}$. Hence, after ψ is translated by an integer, you cannot distinguish it from the Haar wavelet ψ_H in (1.2.15) by looking only at the expected value of its position coordinate. The translation integer k turns out to be a winding number. Our result holds more generally when the definition of $E_\psi(x)$ is adapted to a wider wavelet context, as we showed in Chapter 6 of [BrJo02b]; but in all cases, there is a winding number which produces the above-mentioned integer translate k.

The issue of connectedness for various classes of wavelets is a general question which has been addressed previously in the wavelet literature; see, e.g., [HLPS99], [HeWe96], [StZh01], and [ReWe98]. Here we bring homotopy to bear on the question, and we identify the connected components when the compact support is fixed and given. We show among other things that for a fixed K_1-class a homotopy may take place within a variety of wavelets which is specified by a slightly bigger support than the initially given one.

An important point of our present discussion, beyond the mere fact of compact support, is the size of the support of the wavelets in question. Consider two wavelets A and B of a certain support size. Then our first results in this section also specify the paths $C(t)$, if any, which connect A and B, and in particular the size of the support of the wavelets corresponding to $C(t)$. In [BrJo02b], we treat connectivity in the wider context of noncompactly supported wavelets, following at the outset [Gar99], which considers scale number $N = 2$, and wavelets ψ satisfying

$$\left\{2^{\frac{j}{2}}\psi\left(2^j x - k\right)\right\}_{j,k\in\mathbb{Z}} \text{ is an orthonormal basis (ONB) for } L^2(\mathbb{R})\,. \tag{1.1.19}$$

Garrigós considers, for $\frac{1}{2} < \alpha \leq \infty$, the class $\mathcal{W}_\alpha$ of wavelets ψ such that

$$\int_{\mathbb{R}} |\psi(x)|^2 \left(1 + |x|^2\right)^\alpha dx < \infty, \tag{1.1.20}$$

and there is an $\varepsilon = \varepsilon(\psi)$ such that

$$\int_{\mathbb{R}} \left|\hat{\psi}(t)\right|^2 \left(1 + |t|^2\right)^{\varepsilon} dt < \infty, \tag{1.1.21}$$

i.e., the wavelet is supposed to have some degree of smoothness in the sense of Sobolev.

We now turn to the group of functions $U\colon \mathbb{T} \to \mathrm{U}(N)$, where $\mathrm{U}(N)$ denotes the group of all complex N-by-N matrices. The functions will not be assumed continuous in general. The continuous functions will be designated $C(\mathbb{T}, \mathrm{U}(N))$. Each function in $C(\mathbb{T}, \mathrm{U}(N))$ has a K_1-class, also called a winding number; see [BrJo02b]. The functions in $C(\mathbb{T}, \mathrm{U}(N))$ with finite Fourier expansion will be called *Fourier polynomials*, also if they are functions which take values in $\mathrm{U}(N)$.

Proposition 1.1.3.1: *Let $U \in C(\mathbb{T}, \mathrm{U}(N))$ be a Fourier polynomial, and assume that $K_1(U) = d \in \mathbb{Z}$. Then U is homotopic in $C(\mathbb{T}, \mathrm{U}(N))$ to*

$$V(z) = z^d p \oplus (1_N - p) \tag{1.1.22}$$

where p is the one-dimensional projection onto the first coordinate slot in $\mathbb{C}^N$, and if U has the form

$$U(z) = \sum_{k=-D}^{D} z^k a_k, \tag{1.1.23}$$

then U may be homotopically deformed to V in $C(\mathbb{T}, \mathrm{U}(N))$ through Fourier polynomials of degree at most $|d| + ND$.

This proposition remains true if the word "Fourier polynomial" is replaced by "polynomial" and $a_k = 0$ for $k = -D, -D+1, \dots, -1$. In that case $d \in \mathbb{Z}_+$ and U may be homotopically deformed to V in the loop semigroup of polynomial unitaries in $C(\mathbb{T}, \mathrm{U}(N))$ through polynomials of degree at most d.

Proof: Multiplying U by z^D, we obtain a polynomial $z^D U(z)$ of degree $2D$ mapping $\mathbb{T}$ into $\mathrm{U}(N)$. Then $K_1\left(z^D U\right) = d + ND$. By Proposition 3.3 of [BrJo02a], there exist $d+ND$ one-dimensional projections $p_1, p_2, \dots, p_{d+ND}$ in $M_N(\mathbb{C})$ and a unitary $V_0 \in M_N(\mathbb{C})$ such that

$$z^D U(z) = V_0 \prod_{k=1}^{d+ND} (1 - p_i + z p_i). \tag{1.1.24}$$

(See § 2.2.4 for a related, but different, decomposition.) Now, deforming each of the p_i's continuously through one-dimensional projections to the projection p_0 onto the first coordinate direction, and deforming V_0 in $\mathrm{U}(N)$ into 1_N, we see that $z^D U(z)$ can be deformed into

$$\prod_{k=1}^{d+ND}(1-p_0+zp_0)=1-p_0+z^{d+ND}p_0. \tag{1.1.25}$$

Thus $U(z)$ itself is deformed into

$$z^{-D}(1-p_0)+z^{d+(N-1)D}p_0. \tag{1.1.26}$$

But writing $(1-p_0)$ as a sum of $N-1$ one-dimensional projections $q_1,\dots,q_{N-1}$, we have that the unitary that $U(z)$ is deformed into is

$$\prod_{k=1}^{N-1}\left((1-q_k)+z^{-D}q_k\right)\cdot\left(1+z^{d+(N-1)D}p_0\right), \tag{1.1.27}$$

and next deforming each of the q_k in this decomposition into p_0, we see that $U(z)$ is deformed into

$$\prod_{k=1}^{N-1}\left((1-p_0)+z^{-D}p_0\right)\cdot\left(1+z^{d+(N-1)D}p_0\right)=(1-p_0)+z^{d}p_0. \tag{1.1.28}$$

The crude estimate $|d|+ND$ on the degree of the Fourier polynomials occurring during the deformation is straightforward.

To prove the last statement in the proposition one does not need to multiply U by z^D, and the proof simplifies. Note in particular that $D\le d$ (assuming $a_D\ne 0$). $\square$

Remark 1.1.3.1: We do not know if Proposition 1.1.3.1 is true if $C(\mathbb{T},\mathrm{U}(N))$ is replaced by $C(\mathbb{T},\mathrm{GL}(N))$. It is known from Lemma 11.2.12 of [RLL00] that if $A\in C(\mathbb{T},\mathrm{GL}(N))$ is a polynomial of degree 1 in z, then A can be homotopically deformed through first-order polynomials in $C(\mathbb{T},\mathrm{GL}(N))$ to a unitary of the form $z\to zp+(1_N-p)$ for some projection p, and hence Proposition 1.1.3.1 for $C(\mathbb{T},\mathrm{GL}(N))$ would follow if any polynomial $A\in C(\mathbb{T},\mathrm{GL}(N))$ could be factored into first-order polynomials. It is also clear, since any element $A\in C(\mathbb{T},\mathrm{GL}(N))$ can be homotopically deformed into $z^d p\oplus(1_N-p)$ in $C(\mathbb{T},\mathrm{GL}(N))$, that if A is a Fourier polynomial, then A can be homotopically deformed into $z^d p\oplus(1_N-p)$ through Fourier polynomials. This follows by compactness and the Stone–Weierstraß theorem (Lemma 11.2.3 of [RLL00]). For our purposes in wavelet theory, though, we would need a computable upper bound for the degree of the Fourier polynomials.

For ease of reference we will now list the correspondences between the various objects that interest us in this case. These objects are:

(i) matrix functions, $A\colon \mathbb{T} \to \mathrm{U}_N(\mathbb{C})$, satisfying the normalization

$$A(1) = H, \qquad H_{k,l} = \frac{1}{\sqrt{N}} e^{i2\pi kl/N},\ k,l = 0,\dots,N-1, \tag{1.1.29}$$

(ii) high- and low-pass wavelet filters m_i, $i = 0,1,\dots,N-1$, satisfying

$$\sum_{w^N=z} \overline{m_i(w)}\, m_j(w) = N\delta_{ij}, \qquad i,j = 0,\dots,N-1, \tag{1.1.30}$$

and

$$m_0(1) = \sqrt{N}, \tag{1.1.31}$$

(iii) scaling functions φ together with wavelet generators ψ_i.

We did not specify the continuity and regularity requirements of the functions A, m_i, φ, ψ_i above. This will be done differently in different contexts and the classes clearly depend on these added requirements. We will now restrict to the case that the functions φ and ψ_i have compact support in $[0,\infty\rangle$, i.e., that A and m_i are polynomials in z. Thus $z \to A(z)$ is a polynomial function with

$$(A(z))^* A(z) = 1, \qquad z \in \mathbb{T}. \tag{1.1.32}$$

Scaling functions/wavelet generators to wavelet filters $(\varphi,\psi) \mapsto m$

One defines a_n by

$$\varphi(x) = \sqrt{N} \sum_{n\in\mathbb{Z}} a_n \varphi(Nx - n), \tag{1.1.33}$$

(cf. (2.3.7)) and then m_0 by

$$m_0(z) = \sum_n a_n z^n, \tag{1.1.34}$$

or one uses

$$\sqrt{N}\hat{\varphi}(Nt) = m_0(t)\,\hat{\varphi}(t) \tag{1.1.35}$$

directly. Then the high-pass filters m_i, $i = 1,\dots,N-1$, can be derived from (2.3.10) below. If we are in the generic case (2.3.6), we may also recover the Fourier coefficients $a_n^{(i)}$ of m_i by

$$\begin{aligned} a_n^{(i)} &= \left(1/\sqrt{N}\right) \langle\, \varphi(\,\cdot\, - n) \mid \psi_i(\,\cdot\,/N) \rangle \\ &= \langle\, \varphi(\,\cdot\, - n) \mid U\psi_i \,\rangle \qquad (\text{with } \psi_0 = \varphi), \end{aligned}$$

where $U\psi_i(x) := N^{-1/2}\psi_i(x/N)$. In particular it follows in this generic case that if the scaling and wavelet functions have compact support and the filters are Lipschitz, then the filters are Fourier polynomials. Is this true also in the nongeneric tight frame case?

Now, if $D \in \mathbb{N}$, define:

- $\mathrm{MF}(D) =$ the set of polynomial functions in $z \in \mathbb{T}$ in $C(\mathbb{T}, \mathrm{U}_N(\mathbb{C}))$ of degree at most D satisfying (1.1.29); (1.1.36)
- $\mathrm{WF}(D) =$ the set of N-tuples of wavelet filters $(m_0, \dots, m_{N-1})$ such that all m_i are polynomials in $z \in \mathbb{T}$ of degree at most D satisfying (1.1.30) and (1.1.31); (1.1.37)
- $\mathrm{SF}(D) =$ the set of N-tuples $(\varphi, \psi_1, \dots, \psi_{N-1})$ of scaling functions/ wavelet functions with support in $[0, D]$. (1.1.38)

The spaces $\mathrm{MF}(D)$, $\mathrm{WF}(D)$, and $\mathrm{SF}(D)$ may be equipped with the obvious topologies, coming in the first two cases from, for example, the L^∞-norm over z, and in the last case either from the $L^2(\mathbb{R})$-norm or, as will be more relevant, the tempered-distribution topology. By virtue of Proposition 3.2 in [BrJo02a], $\mathrm{MF}(D)$ has the structure of a compact algebraic variety, and so by (2.3.4) below, $\mathrm{WF}(D)$ is a compact algebraic variety. It is clear from (2.3.4) that the map $A \to m$ maps $\mathrm{MF}(D)$ into $\mathrm{WF}((D+1)N-1)$, and that $m \to A$ maps $\mathrm{WF}((D+1)N-1)$ into $\mathrm{MF}(D)$. Furthermore, it is clear from (1.1.33) and (2.3.10) that $m \to (\varphi, \psi)$ maps $\mathrm{WF}((N-1)D)$ into $\mathrm{SF}(D)$, and conversely $(\varphi, \psi) \to m$ maps $\mathrm{SF}(D)$ into $\mathrm{WF}((N-1)D)$.

Now, let a subindex 0 denote the subsets of these various spaces such that the condition

$$\mathrm{Spec}(R_0) \cap \mathbb{T} = \{1\} \quad \text{and} \quad \dim\left\{g \in \mathcal{K}_{\left\lfloor \frac{D}{N-1} \right\rfloor},\ R(g) = g\right\} = 1 \tag{1.1.39}$$

holds. It is known that the set of points such that (1.1.39) does not hold is a lower-dimensional subvariety of the various varieties, see Section 6 of [Jor01b], and hence $\mathrm{MF}_0(D)$, $\mathrm{WF}_0(D)$, and $\mathrm{SF}_0(D)$ contain the generic points in $\mathrm{MF}(D)$, $\mathrm{WF}(D)$, and $\mathrm{SF}(D)$.

We now summarize the local connectivity results by stating the following theorem. The proof may be found in [BrJo02b], where this is Theorem 2.1.3.

Theorem 1.1.3.1: Let $k \in \mathbb{N}$. Equip the space $\mathrm{SF}(kN+1)$ of scaling functions/wavelet functions with support in $[0, kN+1]$ with the tempered-distribution topology. Then $\mathrm{SF}(kN+1)$ is homeomorphic to a compact algebraic variety. Furthermore, for two elements $(\varphi_0, \psi_0), (\varphi_1, \psi_1) \in \mathrm{SF}(kN+1)$, the following conditions are equivalent:

(a) The elements (φ_0, ψ_0) and (φ_1, ψ_1) can be connected to each other by a continuous path in $\mathrm{SF}(NkN+1)$;
(b) $K_1(\varphi_0, \psi_0) = K_1(\varphi_1, \psi_1)$;
(c) The elements (φ_0, ψ_0) and (φ_1, ψ_1) can be connected to each other by a continuous path in some $\mathrm{SF}(K)$.

Thus, $\mathrm{SF}(kN+1)$ is divided into $Nk(N-1)+1$ components which are connected over $\mathrm{SF}(NkN+1)$.

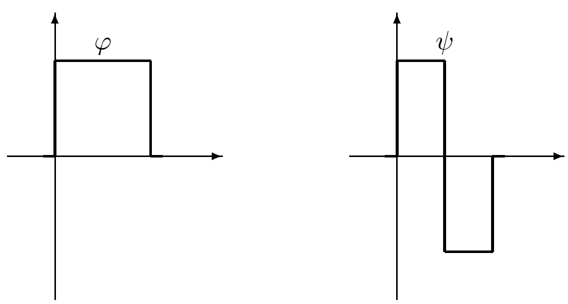
1.2. *Motivation*

In addition to the general background material in the present section, the reader may find a more detailed treatment of some of the current research trends in wavelet analysis in the following papers: [Jor03a] (a book review), [Jor03b] (a survey), and the research papers [DuJo03], [DuJo04a], [DuJo04b], [DuJo04c], [Jor04a], and [Jor04b].

As a mathematical subject, the theory of wavelets draws on tools from mathematics itself, such as harmonic analysis and numerical analysis. But in addition there are exciting links to areas outside mathematics. The connections to electrical and computer engineering, and to image compression and signal processing in particular, are especially fascinating. These interconnections of research disciplines may be illustrated with the two subjects (1) wavelets and (2) subband filtering [from signal processing]. While they are quite different, and have distinct and independent lives, and even have different aims, and different histories, they have in recent years found common ground. It is a truly amazing success story. Advances in one area have helped the other: subband filters are absolutely essential in wavelet algorithms, and in numerical recipes used in subdivision schemes, for example, and especially in JPEG 2000—an important and extraordinarily successful image-compression code. JPEG uses nonlinear approximations and harmonic analysis in spaces of signals of bounded variation. Similarly, new wavelet approximation techniques have given rise to the kind of data-compression which is now used by the FBI [via a patent held by two mathematicians] in digitizing fingerprints in the U.S. It is the happy marriage of the two disciplines, signal processing and wavelets, that enriches the union of the subjects, and the applications, to an extraordinary degree. While the use of high-pass and low-pass filters has a long history in signal processing, dating back more than fifty years, it is only relatively recently, say the mid-1980's, that the connections to wavelets have been made. Multiresolutions from optics are the bread and butter of wavelet algorithms, and they in turn thrive on methods from signal processing, in the quadrature mirror filter construction, for example. The effectiveness of multiresolutions in data compression is related to the fact that multiresolutions are modelled on the familiar positional number system: the digital, or dyadic, representation of numbers. Wavelets are created from scales of closed subspaces of the Hilbert space $L^2(\mathbb{R})$ with a scale of subspaces corresponding to the progression of bits in a number representation. While oversimplified here, this is the key to the use of wavelet algorithms in digital representation of signals and images. The digits in the classical number representation in fact are quite analogous to the frequency subbands that are used *both* in signal processing and in wavelets.

The two functions

$$\varphi(x) = \begin{cases} 1 & 0 \le x < 1 \\ 0 & \text{elsewhere} \end{cases} \text{ and } \psi(x) = \begin{cases} 1 & 0 \le x < \frac{1}{2} \\ -1 & \frac{1}{2} \le x < 1 \\ 0 & \text{elsewhere} \end{cases}$$



Father function (a) Mother function (b) (1.2.1)

capture in a glance the refinement identities

$$\varphi(x) = \varphi(2x) + \varphi(2x-1) \qquad \text{and} \qquad \psi(x) = \varphi(2x) - \varphi(2x-1).$$

The two functions are clearly orthogonal in the inner product of $L^2(\mathbb{R})$, and the two closed subspaces $\mathcal{V}_0$ and $\mathcal{W}_0$ generated by the respective integral translates

$$\{\varphi(\cdot - k) : k \in \mathbb{Z}\} \qquad \text{and} \qquad \{\psi(\cdot - k) : k \in \mathbb{Z}\} \tag{1.2.2}$$

satisfy

$$U\mathcal{V}_0 \subset \mathcal{V}_0 \qquad \text{and} \qquad U\mathcal{W}_0 \subset \mathcal{V}_0 \tag{1.2.3}$$

where U is the dyadic scaling operator $Uf(x) = 2^{-1/2} f(x/2)$. The factor $2^{-1/2}$ is put in to make U a unitary operator in the Hilbert space $L^2(\mathbb{R})$. This version of Haar's system naturally invites the question of what other pairs of functions φ and ψ with corresponding orthogonal subspaces $\mathcal{V}_0$ and $\mathcal{W}_0$ there are such that the same invariance conditions (1.2.3) hold. The invariance conditions hold if there are coefficients a_k and b_k such that the scaling identity

$$\varphi(x) = \sum_{k\in\mathbb{Z}} a_k \varphi(2x-k) \tag{1.2.4}$$

is solved by the father function, called φ, and the mother function ψ is given by

$$\psi(x) = \sum_{k\in\mathbb{Z}} b_k \varphi(2x-k). \tag{1.2.5}$$

A fundamental question is the converse one: Give simple conditions on two sequences (a_k) and (b_k) which guarantee the existence of $L^2(\mathbb{R})$-solutions φ and ψ which satisfy the orthogonality relations for the translates (1.2.2). How do we then get an orthogonal basis from this? The identities for Haar's functions φ and ψ of (1.2.1)(a) and (1.2.1)(b) above make it clear that the answer lies in a similar tiling and matching game which is implicit in the more general identities (1.2.4) and (1.2.5). Clearly we might ask the same question for other scaling numbers, for example $x \to 3x$ or $x \to 4x$ in place of $x \to 2x$. Actually a direct analogue of the visual interpretation from (1.2.1) makes it clear that there are no nonzero locally integrable solutions to the simple variants of (1.2.4),

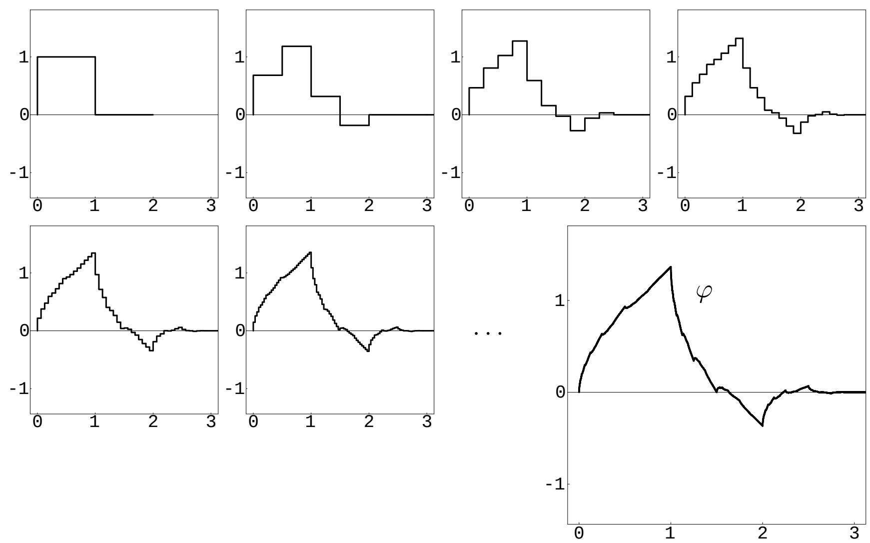
$$\varphi(x) = \frac{3}{2}\left(\varphi(3x) + \varphi(3x-2)\right) \tag{1.2.6}$$

or

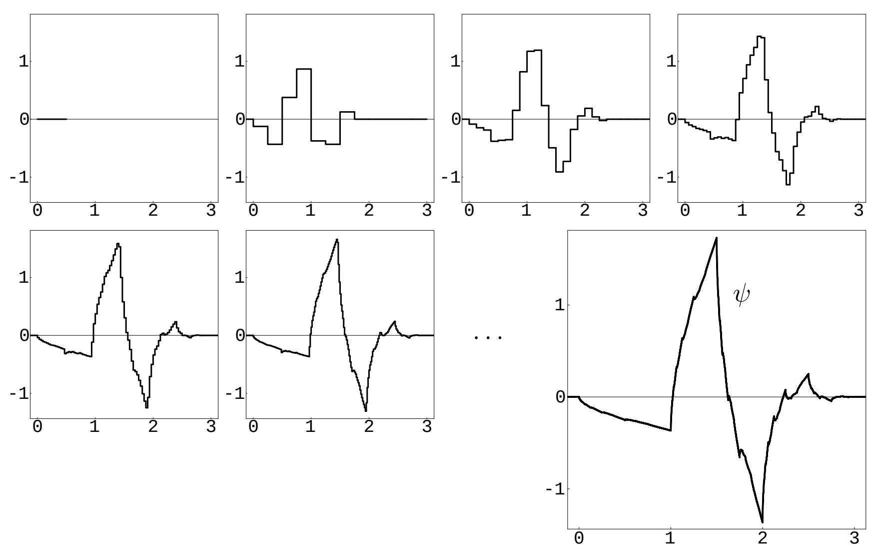
$$\varphi(x) = 2\left(\varphi(4x) + \varphi(4x-2)\right). \tag{1.2.7}$$

There *are* nontrivial solutions to (1.2.6) and (1.2.7), to be sure, but they are versions of the Cantor Devil's Staircase functions, which are prototypes of functions which are not locally integrable.

Since the Haar example is based on the fitting of copies of a fixed "box" inside an expanded one, it would almost seem unlikely that the system (1.2.4)–(1.2.5) admits finite sequences (a_k) and (b_k) such that the corresponding solutions φ and ψ are continuous or differentiable functions of compact support. The discovery in the mid-1980's of compactly supported



(a): Father function



(b): Mother function

Fig. 1. Daubechies wavelet functions and series of cascade approximants

differentiable solutions, see [Dau92], was paralleled by applications in seismology, acoustics [EsGa77], and optics [Mar82], as discussed in [Mey93], and once the solutions were found, other applications followed at a rapid pace: see, for example, the ten books in Benedetto's review [Ben00]. It is the solution ψ in (1.2.5) that the fuss is about, the mother function; the other one, φ, the father function, is only there before the birth of the wavelet. The most famous of them are named after Daubechies, and look like the graphs in Figure 1. With the multiresolution idea, we arrive at the closed subspaces

$$\mathcal{V}_j := U^{-j}\mathcal{V}_0, \qquad j \in \mathbb{Z}, \tag{1.2.8}$$

as noted in (1.2.2)–(1.2.3), where U is some scaling operator. There are extremely effective iterative algorithms for solving the scaling identity (1.2.4): see, for example, Example 2.5.3, pp. 124–125, of [BrJo02b]*, [Dau92], and [StNg96], and Figure 1. A key step in the algorithms involves a clever choice of the kind of resolution pictured in (1.2.13), but digitally encoded. The orthogonality relations can be encoded in the numbers (a_k) and (b_k) of (1.2.4)–(1.2.5), and we arrive at the doubly indexed functions

$$\psi_{j,k}(x) := 2^{j/2}\psi\left(2^j x - k\right), \qquad j,k \in \mathbb{Z}. \tag{1.2.9}$$

It is then not difficult to establish the combined orthogonality relations

$$\int_{\mathbb{R}} \overline{\psi_{j,k}(x)}\,\psi_{j',k'}(x)\;dx = \left\langle \psi_{j,k} \mid \psi_{j',k'} \right\rangle = \delta_{j,j'}\delta_{k,k'} \tag{1.2.10}$$

plus the fact that the functions in (1.2.9) form an orthogonal basis for $L^2(\mathbb{R})$. This provides a painless representation of $L^2(\mathbb{R})$-functions

$$f = \sum_{j\in\mathbb{Z}}\sum_{k\in\mathbb{Z}} c_{j,k}\psi_{j,k} \tag{1.2.11}$$

where the coefficients $c_{j,k}$ are

$$c_{j,k} = \int_{\mathbb{R}} \overline{\psi_{j,k}(x)}\,f(x)\;dx = \left\langle \psi_{j,k} \mid f \right\rangle. \tag{1.2.12}$$

What is more significant is that the resolution structure of closed subspaces of $L^2(\mathbb{R})$

$$\cdots \subset \mathcal{V}_{-2} \subset \mathcal{V}_{-1} \subset \mathcal{V}_0 \subset \mathcal{V}_1 \subset \mathcal{V}_2 \subset \cdots \tag{1.2.13}$$

facilitates powerful algorithms for the representation of the numbers $c_{j,k}$ in (1.2.12). Amazingly, the two sets of numbers (a_k) and (b_k) which were used

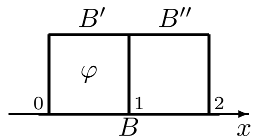
*See an implementation of the "cascade" algorithm using Mathematica, and a "cartoon" of wavelets computed with it, at
http://www.math.uiowa.edu/~jorgen/wavelet_motions.pdf .

in (1.2.4)–(1.2.5), and which produced the magic basis (1.2.9), the wavelets, are the same magic numbers which encode the quadrature mirror filters of signal processing of communications engineering. On the face of it, those signals from communication engineering really seem to be quite unrelated to the issues from wavelets—the signals are just sequences, time is discrete, while wavelets concern $L^2(\mathbb{R})$ and problems in mathematical analysis that are highly non-discrete. Dual filters, or more generally, subband filters, were invented in engineering well before the wavelet craze in mathematics of recent decades. These dual filters in engineering have long been used in technology, even more generally than merely for the context of quadrature mirror filters (QMF's), and it turns out that other popular dual wavelet bases for $L^2(\mathbb{R})$ can be constructed from the more general filter systems; but the best of the wavelet bases are the ones that yield the strongest form of orthogonality, which is (1.2.10), and they are the ones that come from the QMF's. The QMF's in turn are the ones that yield perfect reconstruction of signals that are passed through filters of the analysis-synthesis algorithms of signal processing. They are also the algorithms whose iteration corresponds to the resolution sytems (1.2.13) from wavelet theory.

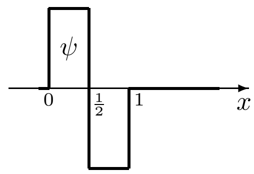
While Fourier invented his transform for the purpose of solving the heat equation, i.e., the partial differential equation for heat conduction, the wavelet transform (1.2.11)–(1.2.12) does not diagonalize the differential operators in the same way. Its effectiveness is more at the level of computation; it turns integral operators into sparse matrices, i.e., matrices which have "many" zeros in the off-diagonal entry slots. Again, the resolution (1.2.13) is key to how this matrix encoding is done in practice.

1.2.1. *Some points of history*

The first wavelet was discovered by Alfred Haar long ago, but its use was limited since it was based on step-functions, and the step-functions jump from one step to the next. The implementation of Haar's wavelet in the approximation problem for continuous functions was therefore rather bad, and for differentiable functions it is atrocious, and so Haar's method was forgotten for many years. And yet it had in it the one idea which proved so powerful in the recent rebirth (since the 1980's) of wavelet analysis: the idea of a *multiresolution*. You see it in its simplest form by noticing that a box function B of (1.2.14) may be scaled down by a half such that two copies B' and B'' of the smaller box then fit precisely inside B. See (1.2.14).



(1.2.14)



(1.2.15)

This process may be continued if you scale by powers of 2 in both directions, i.e., by 2^k for integral k, $-\infty < k < \infty$. So for every $k \in \mathbb{Z}$, there is a finer resolution, and if you take an up- and a shifted mirror image down-version of the dyadic scaling as in (1.2.15), and allow all linear combinations, you will notice that arbitrary functions f on the line $-\infty < x < \infty$, with reasonable integrability properties, admit a representation

$$f(x) = \sum_{k,n} c_{k,n} \psi\left(2^k x - n\right), \tag{1.2.16}$$

where the summation is over all pairs of integers $k, n \in \mathbb{Z}$, with k representing scaling and n translation. The very simple idea of turning this construction into a multiresolution ("multi" for the variety of scales in (1.2.16)) leads not only to an algorithm for the analysis/synthesis problem,

$$f(x) \longleftrightarrow c_{k,n}, \tag{1.2.17}$$

in (1.2.16), but also to a construction of the single functions ψ which solve the problem in (1.2.16), and which can be chosen differentiable, and yet with support contained in a fixed finite interval. These two features, the algorithm and the finite support (called *compact* support), are crucial for computations: Computers do algorithms, but they do not do infinite intervals well. Computers do summations and algebra well, but they do not do integrals and differential equations, unless the calculus problems are discretized and turned into algorithms.

In the discussion to follow, the multiresolution analysis viewpoint is dominant, which increases the role of algorithms; for example, the so-called pyramid algorithm for analyzing signals, or shapes, using wavelets, is an outgrowth of multiresolutions.

Returning to (1.2.14) and (1.2.15), we see that the scaling function φ itself may be expanded in the wavelet basis which is defined from ψ, and we arrive at the infinite series

$$\varphi(x) = \sum_{k=1}^{\infty} 2^{-k} \psi\left(2^{-k} x\right) \tag{1.2.18}$$

which is pointwise convergent for $x \in \mathbb{R}$. (It is a special case of the expansion (1.2.16) when $f = \varphi$.) In view of the picture ( ) below, (1.2.18) gives an alternative meaning to the traditional concept of a *telescoping* infinite sum. If, for example, $0 < x < 1$, then the representation (1.2.18) yields $\varphi(x) = 1 = \frac{1}{2} + \left(\frac{1}{2}\right)^2 + \cdots$, while for $1 < x < 2$, $\varphi(x) = 0 = -\frac{1}{2} + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^3 + \cdots$. More generally, if $n \in \mathbb{N}$, and $2^{n-1} < x < 2^n$, then

$$\varphi(x) = 0 = -\left(\frac{1}{2}\right)^n + \sum_{k>n} \left(\frac{1}{2}\right)^k .$$

So the function φ is itself in the space $\mathcal{V}_0 \subset L^2(\mathbb{R})$, and φ represents the *initial resolution*. The tail terms in (1.2.18) corresponding to

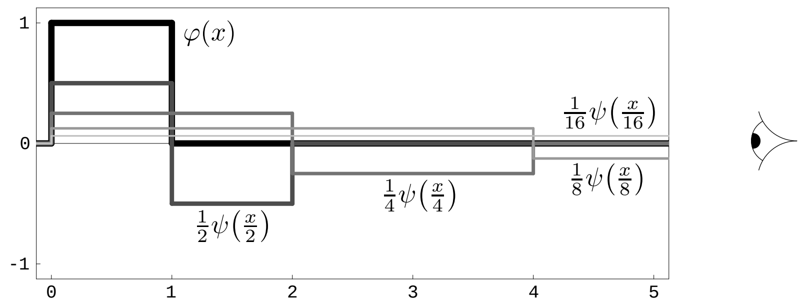
$$\sum_{k>n} 2^{-k} \psi\left(2^{-k} x\right) = \frac{1}{2^n} \varphi\left(\frac{x}{2^n}\right) \tag{1.2.19}$$

represent the *coarser resolution*. The finite sum

$$\sum_{k=1}^{n} 2^{-k} \psi\left(2^{-k} x\right)$$

represents the *missing detail* of φ as a "bump signal". While the sum on the left-hand side in (1.2.19) is *infinite*, i.e., the summation index k is in the range $n < k < \infty$, the expression $2^{-n} \varphi\left(2^{-n} x\right)$ on the right-hand side is merely a coarser scaled version of the original function φ from the subspace $\mathcal{V} \subset L^2(\mathbb{R})$ which specifies the initial resolution. Infinite sums are *analysis problems* while a scale operation is a single simple *algorithmic step*. And so we have encountered a first (easy) instance of the magic of a resolution algorithm; i.e., an instance of a transcendental step (the analysis problem) which is converted into a programmable operation, here the operation of scaling. (Other more powerful uses of the scaling operation may be found in the recent book [Mey98] by Yves Meyer, especially Ch. 5, and [HwMa94].)

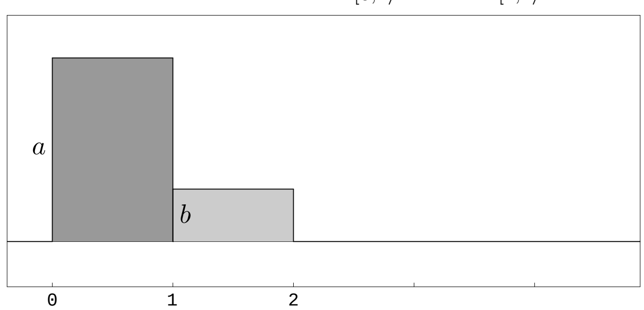
The sketch below allows you to visualize more clearly this resolution versus detail concept which is so central to the wavelet algorithms, also for general wavelets which otherwise may be computationally more difficult than the Haar wavelet.



The wavelet decomposition of Haar's bump function φ in (1.2.14) and (1.2.18)

Using the sketch we see for example that the simple step function

$$f\left(x\right) = a\varphi\left(x\right) + b\varphi\left(x-1\right) = a\chi_{[0,1\rangle}\left(x\right) + b\chi_{[1,2\rangle}\left(x\right) \tag{1.2.20}$$



has the wavelet decomposition into a sum of a *coarser resolution* and an *intermediate detail* as follows:

$$f\left(x\right) = \underbrace{\frac{a-b}{2}\psi\left(\frac{x}{2}\right)}_{\text{intermediate detail}} + \underbrace{\frac{a+b}{2}\varphi\left(\frac{x}{2}\right)}_{\text{coarser version}}, \qquad x \in \mathbb{R}. \tag{1.2.21}$$

Thus the details are measured as differences. This is a general feature that is valid for other functions and other wavelet resolutions. See, for instance, § 2.2 below.

1.2.2. *Some early applications*

While the Haar wavelet is built from flat pieces, and the orthogonality properties amount to a visual tiling of the graphs of the two functions φ and ψ, this is not so for the Daubechies wavelet nor the other compactly supported smooth wavelets. By the Balian–Low theorem [Dau92], a time-frequency wavelet cannot be simultaneously localized in the two dual variables: if ψ is a time-frequency Gabor wavelet, then the two quantities $\int_{\mathbb{R}} |x\psi(x)|^2 \, dx$ and $\int_{\mathbb{R}} \left|t\hat{\psi}(t)\right|^2 dt$ cannot both be finite. Since $\left(\frac{d\psi}{dx}\right)^{\widehat{}}(t) = it\hat{\psi}(t)$, this amounts to poor differentiability properties of well-localized Gabor wavelets, i.e., wavelets built using the two operations translation and frequency modulation over a lattice.

But with the multiresolution viewpoint, we can understand the first of Daubechies's scaling functions as a one-sided differentiable solution φ to

$$\varphi(x) = h_0\varphi(2x) + h_1\varphi(2x-1) + h_2\varphi(2x-2) + h_3\varphi(2x-3), \quad (1.2.22)$$

where the four real coefficients satisfy

$$\left.\begin{aligned} h_0 + h_1 + h_2 + h_3 &= 2, \\ h_3 - h_2 + h_1 - h_0 &= 0, \\ h_3 - 2h_2 + 3h_1 - 4h_0 &= 0, \\ h_1h_3 + h_0h_2 &= 0. \end{aligned}\right\} \quad (1.2.23)$$

The system (1.2.23) is easily solved:

$$\left.\begin{aligned} 4h_0 &= 1+\sqrt{3}, \quad 4h_2 = 3-\sqrt{3}, \\ 4h_1 &= 3+\sqrt{3}, \quad 4h_3 = 1-\sqrt{3}, \end{aligned}\right\} \quad (1.2.24)$$

and Daubechies showed that (1.2.22) has a solution φ which is supported in the interval $[0,3]$, is one-sided differentiable, and satisfies the conditions

$$\int_{\mathbb{R}} \varphi(x)\, dx = 1, \quad \int_{\mathbb{R}} \psi(x)\, dx = 0, \quad \text{and} \quad \int_{\mathbb{R}} x\psi(x)\, dx = 0. \quad (1.2.25)$$

The first applications served as motivating ideas as well: optics, seismic measurements, dynamics, turbulence, data compression; see the book [KaLe95] Actually, it is two books: the first one (primarily by Kahane) is classical Fourier analysis, and the second one (primarily by P.-G. Lemarié-Rieusset) is the wavelet book. It will help you, among other things, to get a better feel for the French connection, the Belgian connection, and the diverse and early impulses from applications in the subject. Enjoy!

For a list of more recent applications we recommend [Mey00].

2. Signal processing

If we idealize and view time as discrete, a copy of $\mathbb{Z}$, then a signal is a sequence $(\xi_n)_{n\in\mathbb{Z}}$ of numbers. A filter is an operator which calculates weighted averages

$$(\xi_n) \longmapsto \sum_{k\in\mathbb{Z}} a_k \xi_{n-k}. \quad (2.1)$$

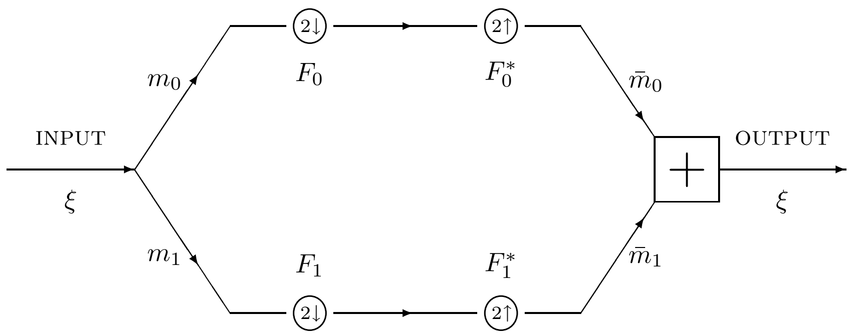


Fig. 2. Perfect reconstruction of signals

But working instead with functions of $z \in \mathbb{T}$, this is multiplication, $f(z) \mapsto m(z) f(z)$, where $m(z) = \sum_{k\in\mathbb{Z}} a_k z^k$ and $f(z) = \sum_{k\in\mathbb{Z}} \xi_k z^k$ are the usual Fourier representation of the corresponding generating functions. Similarly, down-sampling $N\downarrow$ and up-sampling $N\uparrow$ as operators on sequences take the form

$$f \longmapsto \frac{1}{N} \sum_{w\in\mathbb{T},\, w^N = z} f(w) \tag{2.2}$$

and

$$f \longmapsto f\left(z^N\right). \tag{2.3}$$

Since the operators $N\downarrow$ and $N\uparrow$ are clearly dual to one another on the Hilbert space $\ell^2(\mathbb{Z})$ of sequences (i.e., time-signals), we get the corresponding duality for $L^2(\mathbb{T})$, i.e.,

$$\int_{\mathbb{T}} f\left(z^N\right) g(z)\, d\mu(z) = \int_{\mathbb{T}} f(z) \frac{1}{N} \sum_{w^N = z} g(w)\, d\mu(z), \tag{2.4}$$

where μ denotes the normalized Haar measure on $\mathbb{T}$, or equivalently the following identity for 2π-periodic functions:

$$\int_0^{2\pi} f(N\theta) g(\theta)\, d\theta = \int_0^{2\pi} f(\theta) \frac{1}{N} \sum_{k=0}^{N-1} g\left(\frac{\theta + k\cdot 2\pi}{N}\right) d\theta. \tag{2.5}$$

Quadrature mirror filters with N frequency subbands $m_0, m_1, \dots, m_{N-1}$ give perfect reconstruction when signals are analyzed into subbands and then reconstructed via the up-sampling and corresponding dual filters. In engineering formalism this is expressed in the diagram in Fig. 2, for $N = 2$, and m_0, resp. m_1, are called low-pass, resp. high-pass, filters. In operator language, this takes the form

$$F_0^* F_0 + F_1^* F_1 = I,$$

where F_0 and F_1 are the operators in Fig. 2, with dual operators F_0^* and F_1^*. The quadrature conditions may be expressed as

$$F_0 F_0^* = F_1 F_1^* = I \tag{2.6}$$

and

$$F_0 F_1^* = F_1 F_0^* = 0. \tag{2.7}$$

In operator theory there is tradition for working instead with the operators $S_j := F_j^*$. When viewed as operators on $L^2(\mathbb{T})$ they are therefore isometries with orthogonal ranges, and they satisfy

$$\sum_{j=0}^{N-1} S_j S_j^* = I \tag{2.8}$$

with I now representing the identity operator acting on $L^2(\mathbb{T})$. The relations on the S_j-operators are known as the Cuntz relations because of their use in C^*-algebra theory; see [Cun77]. In the present application they take the form

$$(S_j f)(z) = m_j(z) f\left(z^N\right), \qquad f \in L^2(\mathbb{T}), \tag{2.9}$$

and

$$\left(S_j^* f\right)(z) = \frac{1}{N} \sum_{w^N = z} \overline{m_j(w)} f(w), \tag{2.10}$$

and the Cuntz relations are equivalent to the conditions

$$\sum_{w^N = z} |m_j(w)|^2 = N \tag{2.11}$$

and

$$\sum_{w^N = z} \overline{m_j(w)} m_k(w) = 0 \qquad \text{for all } z \in \mathbb{T} \text{ and } j \neq k. \tag{2.12}$$

The last conditions are known in engineering as the quadrature conditions for the subband filters $m_0, m_1, \ldots, m_{N-1}$, with m_0 denoting the low-pass filter. The low-pass and band-pass conditions on the functions m_j are perhaps more familiar in the additive notation given by the substitution $z := e^{-i\theta}$. Then the functions m_j are viewed as 2π-periodic, and

$$m_j\left(j \cdot \frac{2\pi}{N}\right) = \sqrt{N},$$

while

$$m_j\left(k \cdot \frac{2\pi}{N}\right) = 0 \qquad \text{for } j \neq k,$$

with both of the indices j, k ranging over $0, 1, \dots, N-1$.

2.1. *Filters in communications engineering*

The coefficients of the functions $m_j(\,\cdot\,)$ are called *impulse response coefficients* in communications engineering, and when used in wavelets and in subdivision algorithms, they are called *masking coefficients*. In the finite case, the $m_j(\,\cdot\,)$'s are also called FIR for finite impulse response. The model illustrated in Fig. 2 is used in filter design in either hardware or software:

⟦1⟧ Try filters m_0, m_1 in Fig. 2, and approximate the output to the input;
⟦2⟧ Choose a specific structure in which the filter will be realized and then quantize the coefficients, length and numerical values;
⟦3⟧ Verify by simulation that the resulting design meets given performance specifications.

Once filters are constructed, we saw that they are also providing us with wavelet algorithms. When the steps of Fig. 2 are iterated, we arrive at wavelet subdivision algorithms. Relative to a given resolution (pictured as a closed subspace $\mathcal{V}_1$, say, in $L^2(\mathbb{R})$), signals, i.e., functions in $L^2(\mathbb{R})$, decompose into coarser ones and intermediate details. Relative to the subspaces $\mathcal{W}_0$ and $\mathcal{V}_1$, this amounts to

$$\begin{array}{ccccc} \mathcal{V}_1 & = & \mathcal{V}_0 & + & \mathcal{W}_0. \\ \uparrow & & \uparrow & & \uparrow \\ \text{given} & & \text{coarser} & & \text{intermediate} \\ \text{resolution} & & \text{resolution} & & \text{detail} \end{array} \tag{2.1.1}$$

Ideally, we wish the decomposition in (2.1.1) to be orthogonal in the sense that

$$\langle\, f \mid g \,\rangle = 0 \qquad \text{for all } f \in \mathcal{V}_0 \text{ and all } g \in \mathcal{W}_0. \tag{2.1.2}$$

Since the subdivisions involve translations by discrete steps, we specialize the resolution such that both of the spaces $\mathcal{V}_0$ and $\mathcal{W}_0$ are invariant under translations by points in $\mathbb{Z}$, i.e., such that

$$T\colon f \longmapsto f\left(\,\cdot\, - 1\right) \tag{2.1.3}$$

leaves both of the subspaces $\mathcal{V}_0$ and $\mathcal{W}_0$ invariant. The *multiresolution analysis* case (MRA) corresponds to the setup when $\mathcal{V}_0$ is singly generated, i.e., there is a function $\varphi \in \mathcal{V}_0$ such that the closed linear span of

$$T_n \varphi\left(\,\cdot\,\right) = \varphi\left(\,\cdot\, - n\right), \qquad n \in \mathbb{Z}, \tag{2.1.4}$$

is all of $\mathcal{V}_0$. If $N = 2$, then there is then also a $\psi \in \mathcal{W}_0$ such that the closed linear span of $\{\, \psi\left(\,\cdot\, - n\right) : n \in \mathbb{Z} \,\}$ is all of $\mathcal{W}_0$. If $N > 2$, we may need functions $\psi_1, \dots, \psi_{N-1}$ in $\mathcal{W}_0$ such that $\{\, \psi_i\left(\,\cdot\, - n\right) : i = 1, \dots, N-1,\ n \in \mathbb{Z} \,\}$ has a closed span equal to $\mathcal{W}_0$.

2.2. *Algorithms for signals and for wavelets*

The pyramid algorithm and the Cuntz relations. Since the two Hilbert spaces $L^2\left(\mathbb{T}\right)$ and $\ell^2\left(\mathbb{Z}\right)$ are isomorphic via the Fourier series representation, it follows that the system $\{S_i\}_{i=0}^{1}$ is equivalent to a system $\{\hat{S}_i\}_{i=0}^{1}$ acting on $\ell^2\left(\mathbb{Z}\right)$. Specifically, $\left(S_i f\right)^{\widehat{}} = \hat{S}_i \hat{f}$, $i = 0, 1$, where $\hat{f}\left(n\right) := \int_{\mathbb{T}} z^{-n} f\left(z\right)\, d\mu\left(z\right)$. For $c := \left(c_n\right)_{n\in\mathbb{Z}}$ in $\ell^2\left(\mathbb{Z}\right)$, and functions f on $\mathbb{R}$, set

$$\begin{aligned} f_{-1}\left(x\right) &:= \left(Uf\right)\left(x\right) = 2^{-\frac{1}{2}} f\left(\frac{x}{2}\right), \text{ and} \\ \left(c * f\right)\left(x\right) &:= \sum_{n\in\mathbb{Z}} c_n f\left(x - n\right). \end{aligned}$$

For the present, let $\{m_i\}_{i=0}^{1}$ be the low-pass and high-pass wavelet filters, and let φ, ψ be the corresponding scaling function, resp., wavelet function, also called father function, resp., mother function. Now introduce the corresponding operators S_i and their cousins $\hat{S}_i$. The adjoints $\hat{S}_i^*$ are also called *filters*.

Then

$$c * \varphi = \underbrace{\left(\left(\hat{S}_0^* c\right) * \varphi\right)_{-1}}_{\text{coarser resolution}} + \underbrace{\left(\left(\hat{S}_1^* c\right) * \psi\right)_{-1}}_{\text{detail}} \qquad \text{for all } c \in \ell^2\left(\mathbb{Z}\right). \tag{2.2.1}$$

Define $W\colon \ell^2 \to \ell^2$ by

$$W\left(c\right)\left(x\right) = \left(c * \varphi\right)\left(x\right) = \sum_{n\in\mathbb{Z}} c_n \varphi\left(x - n\right). \tag{2.2.2}$$

Then W maps ℓ^2 isometrically onto $\mathcal{V}_0$ in the orthogonal case and

$$W\hat{S}_0 = UW.$$

Further

$$W\hat{S}_0\hat{S}_0^* c = \left(\hat{S}_0^* c * \varphi\right)_{-1}.$$

Embedding ℓ^2 into $\ell^2 \oplus \ell^2$ as $\ell^2 \oplus 0$, extend W to $\ell^2 \oplus \ell^2$ by putting

$$W\left(c \oplus d\right) = c * \varphi + d * \psi.$$

Then the extended W maps $\ell^2 \oplus \ell^2$ isometrically onto $U^{-1}\mathcal{V}_0$ and

$$W\left(\hat{S}_0 c + \hat{S}_1 d\right) = UW\left(c \oplus d\right)$$

for all $c, d \in \ell^2$, where the left W is the one from (2.2.2) and the right is the extension of W to $\ell^2 \oplus \ell^2$.

At this point you can use $1_{\ell^2} = \hat{S}_0\hat{S}_0^* + \hat{S}_1\hat{S}_1^*$ to show (2.2.1). Note that if $c_0 = a$ and $c_1 = b$ and $c_i = 0$ for other i, the formula (2.2.1) reduces to (1.2.21).

The subdivision relations (2.2.1) are equivalent to the system

$$\sqrt{2}\varphi\left(2x\right) = \sum_{k\in\mathbb{Z}} \bar{a}_{2k}\varphi\left(x+k\right) + \sum_{k\in\mathbb{Z}} \bar{b}_{2k}\psi\left(x+k\right), \tag{2.2.3}$$

$$\sqrt{2}\varphi\left(2x-1\right) = \sum_{k\in\mathbb{Z}} \bar{a}_{2k+1}\varphi\left(x+k\right) + \sum_{k\in\mathbb{Z}} \bar{b}_{2k+1}\psi\left(x+k\right), \tag{2.2.4}$$

where the coefficients a_n, b_n are those of the quantum wavelet algorithm, i.e., the coefficients in the "large" unitary matrix (2.2.5). Thus the quantum algorithm does the wavelet decomposition within a fixed resolution subspace.

The scaling function φ defines a resolution subspace $\mathcal{V}_0 \subset L^2\left(\mathbb{R}\right)$. Then (2.2.1), or equivalently (2.2.3)–(2.2.4), represents the orthogonal decomposition of functions in $\mathcal{V}_0$ into an orthogonal sum of a function with coarser resolution and a function in the intermediate detail subspace.

Let m_0, m_1 be a dyadic wavelet filter, and let $\mathbb{T} \ni z \mapsto A\left(z\right) \in \mathrm{U}_2\left(\mathbb{C}\right)$ be the corresponding matrix function, $A_{i,j}\left(z\right) = \frac{1}{2}\sum_{w^2=z} w^{-j} m_i\left(w\right)$. If the low-pass filter $m_0\left(z\right) = a_0 + a_1 z + \cdots + a_{2n+1}z^{2n+1}$, then a choice for $m_1\left(z\right) = \sum_{k=0}^{2n+1} b_k z^k$ is $b_k = \left(-1\right)^k \bar{a}_{2n+1-k}$. We then have $A\left(z\right) = \sum_{k=0}^{n} A_k z^k$ where $A_k = \begin{pmatrix} a_{2k} & a_{2k+1} \\ b_{2k} & b_{2k+1} \end{pmatrix}$, and the following $2^{n+2} \times 2^{n+2}$ scalar matrix can be checked to be unitary:

$$\left(\begin{array}{c|c|c|c|c|c|c|c|c|c|c|c}
\begin{matrix}a_1\\b_1\end{matrix} & A_1 & A_2 & \cdots & A_{n-1} & A_n & 0 & & \cdots & & 0 & \begin{matrix}a_0\\b_0\end{matrix}\\ \hline
\begin{matrix}0\\0\end{matrix} & A_0 & A_1 & \cdots & A_{n-2} & A_{n-1} & A_n & 0 & & \cdots & 0 & \begin{matrix}0\\0\end{matrix}\\ \hline
\begin{matrix}0\\0\end{matrix} & 0 & A_0 & \cdots & A_{n-3} & A_{n-2} & A_{n-1} & A_n & 0 & \cdots & 0 & \begin{matrix}0\\0\end{matrix}\\ \hline
\begin{matrix}0\\0\end{matrix} & & & & & & & & & & & \begin{matrix}0\\0\end{matrix}\\ \hline
\vdots & & & & \ddots & & & & & \ddots & & \vdots\\ \hline
\begin{matrix}0\\0\end{matrix} & & & & & & & & & & & \begin{matrix}0\\0\end{matrix}\\ \hline
\begin{matrix}a_{2n+1}\\b_{2n+1}\end{matrix} & 0 & & \cdots & & 0 & A_0 & A_1 & A_2 & \cdots & A_{n-1} & \begin{matrix}a_{2n}\\b_{2n}\end{matrix}\\ \hline
\begin{matrix}a_{2n-1}\\b_{2n-1}\end{matrix} & A_n & 0 & & \cdots & & 0 & A_0 & A_1 & \cdots & A_{n-2} & \begin{matrix}a_{2n-2}\\b_{2n-2}\end{matrix}\\ \hline
\begin{matrix}a_{2n-3}\\b_{2n-3}\end{matrix} & A_{n-1} & A_n & 0 & & \cdots & & 0 & A_0 & \cdots & A_{n-3} & \begin{matrix}a_{2n-4}\\b_{2n-4}\end{matrix}\\ \hline
\vdots & & & \ddots & & & & & & \ddots & & \vdots\\ \hline
\begin{matrix}a_3\\b_3\end{matrix} & A_2 & A_3 & \cdots & A_n & 0 & & \cdots & & 0 & A_0 & \begin{matrix}a_2\\b_2\end{matrix}
\end{array}\right) \tag{2.2.5}$$

Except for the scalar entries in the two extreme left and right columns, all the other entries of the big combined matrix U_A are taken from the cyclic arrangements of the 2×2 matrices of coefficients $A_0, A_1, \ldots, A_n$ in the expansion of $A(z)$. For the case of $n = 1$ this amounts to the simple 8×8 wavelet matrix

$$A_0 \qquad \left(\begin{array}{c|c|c|c|c}
\begin{matrix}a_1\\b_1\end{matrix} & A_1 & 0 & 0 & \begin{matrix}a_0\\b_0\end{matrix}\\ \hline
\begin{matrix}0\\0\end{matrix} & A_0 & A_1 & 0 & \begin{matrix}0\\0\end{matrix}\\ \hline
\begin{matrix}0\\0\end{matrix} & 0 & A_0 & A_1 & \begin{matrix}0\\0\end{matrix}\\ \hline
\begin{matrix}a_3\\b_3\end{matrix} & 0 & 0 & A_0 & \begin{matrix}a_2\\b_2\end{matrix}
\end{array}\right) \qquad A_1, \tag{2.2.6}$$

which is the one that produces the sequence of quantum gates. The quantum algorithm of a wavelet filter is thus represented by a $2^{n+2} \times 2^{n+2}$ unitary matrix U_A acting on the quantum qubit register $\underbrace{\mathbb{C} \otimes \cdots \otimes \mathbb{C}}_{n+2 \text{ times}} = \mathbb{C}^{2(n+2)}$, i.e., it acts on a configuration of $n+2$ qubits. The realization of a wavelet algorithm in the quantum realm thus amounts to spelling out the steps in factoring U_A into a product of qubit gates. By Shor's theorem, we know that this can be done, and U_A may be built out of one-qubit gates and CNOT gates following the ideas sketched above. The reader may find more discussion of the matrix U_A in Section 3 of [Fre02].

The generalization of classical and quantum wavelet resolution algorithms from $N = 2$ to $N > 2$ is immediate: Then $m_i(z) = \sum_{k\in\mathbb{Z}} a_k^{(i)} z^k$,

$$(S_i f)(z) = m_i(z) f\left(z^N\right), \qquad i = 0, \dots, N-1, \tag{2.2.7}$$

and the transformation rules

$$\xi_{Nk+i} = \sum_{l\in\mathbb{Z}} a_{l-Nk}^{(i)} \varepsilon_l, \qquad i = 0, 1, \dots, N-1, \tag{2.2.8}$$

permute the set of ONB's in $\ell^2(\mathbb{Z})$ and define a unitary commuting with the N-shift. Hence, the standard formulas from [Wic93], [Kla99], and [FiWi99] for the quantum computing algorithm naturally generalize to the case $N > 2$ via (2.2.8). Instead of k-registers $\underbrace{\mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2}_{k \text{ times}} = \mathbb{C}^{2^k}$ over $\mathbb{C}^2$, we will now have to work rather with $\underbrace{\mathbb{C}^N \otimes \cdots \otimes \mathbb{C}^N}_{k \text{ times}} = \mathbb{C}^{N^k}$.

The use of the algorithmic relations in engineering and operator algebra theory predates their more recent use in wavelet theory and wavepacket analysis.

2.2.1. *Pyramid algorithms*

For $N > 2$, the algorithm of the previous section takes the following form.

The pyramid algorithm and the Cuntz relations revisited. By Fourier equivalence of $L^2(\mathbb{T})$ and $\ell^2(\mathbb{Z})$ via the Fourier series, it follows that the system $\{S_i\}_{i=0}^{N-1}$ is equivalent to a system $\{\hat{S}_i\}_{i=0}^{N-1}$ acting on $\ell^2(\mathbb{Z})$. Specifically, $(S_i f)\widehat{\ } = \hat{S}_i \hat{f}$, $i = 0, \dots, N-1$, where $\hat{f}(n) := \int_{\mathbb{T}} z^{-n} f(z)\, d\mu(z)$. For $c := (c_n)_{n\in\mathbb{Z}}$ in $\ell^2(\mathbb{Z})$, and functions f on $\mathbb{R}$, set

$$f_{-1}(x) := N^{-\frac{1}{2}} f\left(\frac{x}{N}\right),$$

and

$$(c * f)(x) := \sum_{n \in \mathbb{Z}} c_n f(x - n).$$

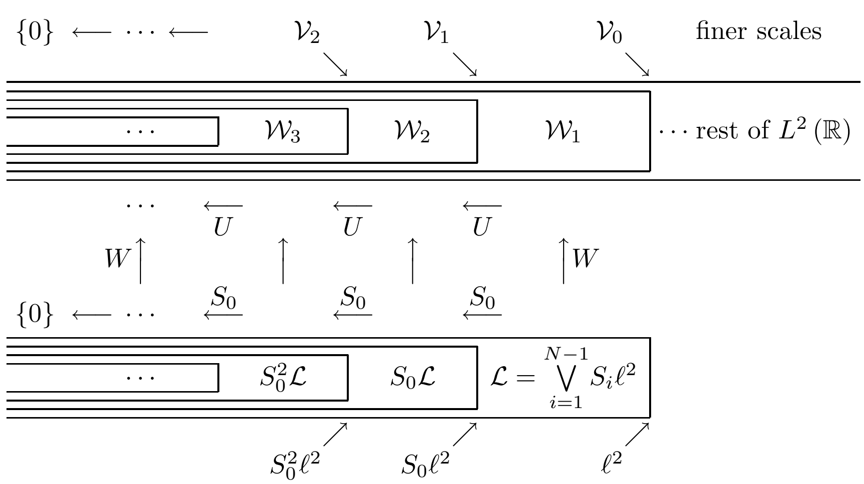
Let $\{m_i\}_{i=0}^{N-1}$ be low-pass and high-pass wavelet filters, and let φ, $\psi_1, \ldots, \psi_{N-1}$ be the corresponding scaling function, resp., wavelet functions. Now introduce the corresponding operators S_i, and their cousins $\hat{S}_i$. The adjoints $\hat{S}_i^*$ are also called *filters*.

Then

$$c*\varphi = \underbrace{\left(\left(\hat{S}_0^* c\right) * \varphi\right)_{-1}}_{\text{coarser resolution}} + \underbrace{\sum_{i=1}^{N-1} \left(\left(\hat{S}_i^* c\right) * \psi_i\right)_{-1}}_{\text{detail}} \qquad \text{for all } c \in \ell^2(\mathbb{Z}). \quad (2.2.9)$$

The scaling function φ defines a resolution subspace $\mathcal{V}_0 \subset L^2(\mathbb{R})$. For the case $N > 2$:

Discrete vs. continuous wavelets, i.e., ℓ^2 vs. $L^2(\mathbb{R})$:



More refined pyramid algorithms yield wavelet packets as follows.

The Haar wavelet is supported in $[0, 1]$, and if $j \in \mathbb{Z}_+$ and $k \in \mathbb{Z}$, then the modified function $x \mapsto \psi\left(2^j x - k\right)$ is supported in the smaller interval $\frac{k}{2^j} \leq x \leq \frac{k+1}{2^j}$. When j is fixed, these intervals are contained in $[0, 1]$ for $k \in \left\{0, 1, \ldots, 2^j - 1\right\}$. This is not the case for the other wavelet functions. For one thing, the non-Haar wavelets ψ have support intervals of length more than one, and this forces periodicity considerations; see [CDV93]. For this reason, Coifman and Wickerhauser [CoWi93] invented the concept of wavelet packets. They are built from functions with prescribed smoothness, and yet they have localization properties that rival those of the (discontinuous) Haar wavelet.

There are powerful but nontrivial theorems on restriction algorithms for wavelets $\psi_{j,k}(x) = 2^{\frac{j}{2}}\psi\left(2^j x - k\right)$ from $L^2(\mathbb{R})$ to $L^2(0,1)$. We refer the reader to [CDV93] and [MiXu94] for the details of this construction. The underlying idea of Alfred Haar has found a recent renaissance in the work of Wickerhauser [Wic93] on *wavelet packets*. The idea there, which is also motivated by the Walsh function algorithm, is to replace the refinement equation (1.1.33) by a related recursive system as follows: Let $m_0(z) = \sum_k a_k z^k$, $m_1(z) = \sum_k b_k z^k$, for example $b_k = (-1)^k \bar{a}_{1-k}$, $k \in \mathbb{Z}$, be a given low-pass/high-pass system, $N = 2$. Then consider the following *refinement system* on $\mathbb{R}$:

$$W_{2n}(x) = \sqrt{2}\sum_{k\in\mathbb{Z}} a_k W_n(2x-k), \tag{2.2.10}$$

$$W_{2n+1}(x) = \sqrt{2}\sum_{k\in\mathbb{Z}} b_k W_n(2x-k). \tag{2.2.11}$$

Clearly the function W_0 can be identified with the traditional scaling function φ of (2.3.7). A theorem of Coifman and Wickerhauser (Theorem 8.1, [CoWi93]) states that if $\mathcal{P}$ is a partition of $\{0,1,2,\dots\}$ into subsets of the form

$$I_{k,n} = \left\{2^k n, 2^k n + 1, \dots, 2^k(n+1) - 1\right\},$$

then the function system

$$\left\{2^{\frac{k}{2}} W_n\left(2^k x - l\right) \;\middle|\; I_{k,n} \in \mathcal{P},\ l \in \mathbb{Z}\right\}$$

is an orthonormal basis for $L^2(\mathbb{R})$. Although it is not spelled out in [CoWi93], this construction of bases in $L^2(\mathbb{R})$ divides itself into the two cases, the true orthonormal basis (ONB), and the weaker property of forming a function system which is only a tight frame. As in the wavelet case, to get the $\mathcal{P}$-system to really be an ONB for $L^2(\mathbb{R})$, we must assume the transfer operator $R_{|m_0|^2}$ to have *Perron–Frobenius spectrum* on $C(\mathbb{T})$. This means that the intersection of the point spectrum of $R_{|m_0|^2}$ with $\mathbb{T}$ is the singleton $\lambda = 1$, and that $\dim\ker((1 - R_{|m_0|^2})|_{C(\mathbb{T})}) = 1$.

2.2.2. *Subdivision algorithms*

The algorithms for wavelets and wavelet packets involve the pyramid idea as well as subdivision. Each subdivision produces a multiplication of subdivision points. If the scaling is by N, then j subdivisions multiply the number of subdivision points by N^j. If the scaling is by a $d \times d$ integral matrix $\mathbf{N}$, then the multiplicative factor is $|\det \mathbf{N}|^j$ in the number of subdivision points placed in $\mathbb{R}^d$.

In the discussion below, we restrict attention to $d = 1$, but the conclusions hold with only minor modification in the general case of $d > 1$ and matrix scaling.

If W is a continuous function on $\mathbb{T}$, the *transfer operator* or *kneading operator* R_W

$$R_W \xi (z) = \frac{1}{N} \sum_{w^N = z} W (w) \xi (w) = S_0^* W \xi (z) , \tag{2.2.12}$$

with the alias

$$(R_W f)_n = \sum_k c_{Nn-k} f_k \tag{2.2.13}$$

in the Fourier transformed space, has an adjoint which is the *subdivision operator* or chopping operator

$$(R_W^* \xi)(z) = \overline{W (z)} \xi \left(z^N\right) \tag{2.2.14}$$

on functions ξ on $\mathbb{T}$, with the alias

$$(R_W^* f)_n = \sum_k \overline{c_{Nk-n}} f_k \tag{2.2.15}$$

on sequences.

We will analyze the duality between R_W and R_W^* and their spectra. Specializing to $W = |m_0|^2$, we note that R_W is then the transfer operator of orthogonal type wavelets. In the following, W is assumed only to satisfy $W \in \mathrm{Lip}_1(\mathbb{T})$ and $W \geq 0$. Other conditions are discussed in [BrJo02b].

In the engineering terminology of § 2.2, the operation (2.2.13) is composed of a local filter with the numbers c_k as coefficients, followed by the down-sampling $N\downarrow$, while (2.2.15) is composed of up-sampling $N\uparrow$, followed by an application of a dual filter. In signal processing, $N\downarrow$ is referred to as "decimation" even if N is not 10.

The operator S $(= R_W^*)$ is called the subdivision operator, or the *woodcutter operator*, because of its use in computer graphics. Iterations of S will generate a shape which (in the case of one real dimension) takes the form of the graph of a function f on $\mathbb{R}$. If $\xi \in \ell^\infty(\mathbb{Z})$ is given, and if the differences

$$D_n (i) = f\left(\frac{i}{2^n}\right) - (S^n \xi)(i) , \qquad i \in \mathbb{Z}, \tag{2.2.16}$$

are small, for example if

$$\lim_{n\to\infty}\sup_{i\in\mathbb{Z}}\left|D_n\left(i\right)\right| = 0, \tag{2.2.17}$$

then we say that ξ represents *control points*, or a control polygon, and the function f is the limit of the *subdivision scheme*.

It follows that the subdivision operator S on the sequence spaces, especially on $\ell^\infty\left(\mathbb{Z}\right)$, governs *pointwise approximation* to refinable limit functions. The dual version of S, i.e., $R = S^*$ (= the transfer operator) governs the corresponding *mean approximation* problem, i.e., approximation relative to the $L^2\left(\mathbb{R}\right)$-norm.

In Scholium 4.1.2 of [BrJo02b], we consider the eigenvalue problem

$$S\xi = \lambda\xi, \qquad \lambda \in \mathbb{C}, \tag{2.2.18}$$

and $\xi \neq 0$ in some suitably defined space of sequences. The formula (2.2.16) for the limit of a given subdivision scheme S makes it clear that the case (2.2.18) must be excluded. For if (2.2.18) holds, for some $\lambda \in \mathbb{C}$, and some sequence ξ of control points, then there is not a corresponding regular function f on $\mathbb{R}$ with its values given on the finer grids $2^{-n}\mathbb{Z}$, $n = 1, 2, \dots,$ by

$$f_\xi\left(i2^{-n}\right) \approx \left(S^n\xi\right)\left(i\right) = \lambda^n\xi\left(i\right), \qquad i \in \mathbb{Z}. \tag{2.2.19}$$

We show in Example 4.1.3 of [BrJo02b] that there are no such control points ξ in $\ell^2\left(\mathbb{Z}\right)\setminus\{0\}$. Hence the stability of the algorithm!

2.2.3. *Wavelet packet algorithms*

The main difference between the algorithms of wavelets and those of wavelet packets is that for the wavelets the path in the pyramid is to one side only: a given resolution is split into a coarser one and the intermediate detail. The intermediate detail may further be broken down into frequency bands. With the operators $S_jf\left(z\right) = m_j\left(z\right)f\left(z^N\right)$ acting on $L^2\left(\mathbb{T}\right)$, the coarser subspace after j steps is modelled on $S_0^jL^2\left(\mathbb{T}\right)$, and the projection onto this subspace is $S_0^jS_0^{*\,j}$ where S_0 is the isometry of $L^2\left(\mathbb{T}\right) \cong \mathcal{V}_0$ defined by the low-pass filter m_0. But in the construction of the wavelet packet, the subspace resulting by running the algorithm j times is $S_{i_1}S_{i_2}\cdots S_{i_j}L^2\left(\mathbb{T}\right)$, and the projection onto this subspace is

$$S_{i_1}S_{i_2}\cdots S_{i_j}S_{i_j}^*\cdots S_{i_2}^*S_{i_1}^*.$$

If $n \in \mathbb{Z}_+$, the wavelet function W_n is computed from the iteration $i_1, \ldots, i_j$ corresponding to the representation

$$n = i_1 + i_2 N + i_3 N^2 + \cdots + i_j N^{j-1},$$

where $i_1, \ldots, i_j \in \{0, 1, \ldots, N-1\}$ are unique from the Euclidean algorithm.

2.2.4. *Lifting algorithms: Sweldens and more*

The discussion centers around the matrix functions $A \colon \mathbb{T} \to \mathrm{GL}_2(\mathbb{C})$.

The case $\det A \equiv 1$**.** Recall that we call a finite sum $\sum_{k=-n_0}^{n_1} A_k z^k$, $n_0, n_1 \geq 0$, a Fourier polynomial both if the coefficients A_k are numbers, and if they are matrices. The matrix-valued Fourier polynomials $\mathbb{T} \ni z \mapsto A(z) \in M_2(\mathbb{C})$ such that $\det A(z) \equiv 1$ form a subgroup of $C(\mathbb{T}, \mathrm{GL}_2(\mathbb{C}))$ which we denote $\mathcal{SL}_2$.

For every $A(z)$ in $\mathcal{SL}_2$ there are $m \in \mathbb{Z}_+$, $K \in \mathbb{C} \setminus \{0\}$, and scalar-valued Fourier polynomials $u_1(z), \ldots, u_m(z), l_1(z), \ldots, l_m(z)$ such that

$$\begin{aligned} A(z) = \begin{pmatrix} K & 0 \\ 0 & K^{-1} \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ l_1(z) & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & u_1(z) \\ 0 & 1 \end{pmatrix} \cdot \\ \cdot \begin{pmatrix} 1 & 0 \\ l_2(z) & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & u_2(z) \\ 0 & 1 \end{pmatrix} \cdots \begin{pmatrix} 1 & 0 \\ l_m(z) & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & u_m(z) \\ 0 & 1 \end{pmatrix}. \end{aligned} \tag{2.2.20}$$

See [DaSw98]. This is the first step in the Daubechies–Sweldens lifting algorithm for the discrete wavelet transform. Thus the case $\det(A(z)) = 1$ gives a constructive lifting algorithm for wavelets, and such an algorithm has not been established in the $C(\mathbb{T}, \mathrm{GL}_2(\mathbb{C}))$ case. The decomposition could also be compared with Proposition 3.3 of [BrJo02a], which was mentioned in connection with the proof of (1.1.24).

Recall the correspondence between matrix functions and wavelet filters: If $A \colon \mathbb{T} \to \mathrm{GL}_2(\mathbb{C})$ is a matrix function, then the corresponding dyadic wavelet filters are

$$m_i^{(A)}(z) = \sum_{j=0}^{1} A_{i,j}\left(z^2\right) z^j, \qquad i = 0, 1.$$

It follows that the two matrix functions A and B satisfy

$$A = \begin{pmatrix} 1 & 0 \\ l & 1 \end{pmatrix} B$$

for some l in the ring $\mathcal{F}$ of Fourier polynomials if and only if $m_0^{(A)} = m_0^{(B)}$ and $m_1^{(A)}(z) = m_1^{(B)}(z) + l\left(z^2\right) m_0^{(A)}(z)$.

Similarly note that the two matrix functions A and B satisfy

$$A = \begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix} B$$

for some $u \in \mathcal{F}$ if and only if $m_1^{(A)} = m_1^{(B)}$ and $m_0^{(A)}(z) = m_0^{(B)}(z) + u(z^2)\, m_1^{(A)}(z)$.

Remark. The conclusion is that the wavelet algorithm for a general wavelet filter corresponding to a matrix function, say A, may be broken down in a sequence of zig-zag steps acting alternately on the high-pass and the low-pass signal components.

2.3. *Factorization theorems for matrix functions*

We mentioned that for matrix functions corresponding to finite impulse response (FIR) filters which are unitary, we need only the constant matrix (which is chosen such as to achieve the high-pass and low-pass conditions) and factors of the form

$$U_P(z) = zP + P^{\perp} \cong \left(\begin{array}{c|c} z & 0 \\ \hline 0 & 1 \end{array}\right)$$

where P is a rank-one projection in $\mathbb{C}^N$ and N is the scaling number of the subdivision.

Unfortunately, no such factorization theorem is available for the non-unitary FIR filters. But the matrix functions take values in the non-singular complex $N \times N$ matrices. The Sweldens–Daubechies factorization and the lifting algorithm serve as a substitute. There are still the general non-unimodular FIR-matrix functions where factorizations are so far a bit of a mystery. The matrix functions are called *polyphase matrices* in the engineering literature. The following summary serves as a classification theorem for the orthogonal wavelets of compact support: the wavelets correspond to FIR polyphase matrices which are unitary.

In summary, an algorithm to construct all the wavelet functions ψ of scale 2 with support in $[0, 2k+1]$ can be established as follows:

⟦1⟧ Pick k one-dimensional orthogonal projections $Q_1, \dots, Q_k$ in $M_2(\mathbb{C})$ and define the unitary-valued matrix function $A(z)$ on $\mathbb{T}$ by

$$A(z) = V(1 - Q_1 + zQ_1)(1 - Q_2 + zQ_2) \cdots (1 - Q_k + zQ_k), \quad (2.3.1)$$

where

$$V = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \quad (2.3.2)$$

Then each Q_j has the form

$$Q_j = \begin{pmatrix} \lambda_j & \sqrt{\lambda_j (1-\lambda_j)} e^{i\theta_j} \\ \sqrt{\lambda_j (1-\lambda_j)} e^{-i\theta_j} & 1-\lambda_j \end{pmatrix}, \tag{2.3.3}$$

where $\lambda_j \in [0,1]$ and $\theta_j \in [0, 2\pi)$. (See Proposition 3.3 of [BrJo02a].)

⟦2⟧ Define the filters $m_0(z)$ and $m_1(z)$ by

$$m_i(z) = \sum_{j=0}^{N-1} z^j A_{ij}\left(z^N\right), \qquad i, j = 0, \ldots, N-1, \tag{2.3.4}$$

with $N = 2$.

⟦3⟧ Define $\hat{\varphi}$ by

$$\hat{\varphi}(t) = \prod_{k=1}^{\infty} \left(\frac{m_0\left(tN^{-k}\right)}{\sqrt{N}} \right). \tag{2.3.5}$$

If the condition

$$\operatorname{PER}\left(|\hat{\varphi}|^2\right)(t) := \sum_{n\in\mathbb{Z}} \left|\hat{\varphi}(t + 2\pi n)\right|^2 = 1 \tag{2.3.6}$$

fails, then the algorithm stops.

⟦4⟧ If the condition (2.3.6) holds, one may alternatively define φ by the cascade algorithm

$$\varphi(x) = \sqrt{N} \sum_{n\in\mathbb{Z}} a_n \varphi(Nx - n), \tag{2.3.7}$$

$$\chi(x) = \begin{cases} 1, \ 0 \leq x < 1, \\ 0, \ x \in \mathbb{R} \setminus [0,1\rangle, \end{cases} \tag{2.3.8}$$

$$M_a \colon \psi \longmapsto \sqrt{N} \sum_n a_n \psi(Nx - n). \tag{2.3.9}$$

⟦5⟧ The wavelet function ψ is then defined by

$$\psi_i(x) = \sqrt{N} \sum_{n\in\mathbb{Z}} a_n^{(i)} \varphi(Nx - n), \tag{2.3.10}$$

where $a_n^{(i)}$ are the Fourier coefficients of m_i,

$$m_i(z) = \sum_n a_n^{(i)} z^n, \tag{2.3.11}$$

and $z = e^{-it}$; this is the most general wavelet function with support in $[0, 2k+1]$.

⟦6⟧ All other wavelet functions with compact support can be obtained from the ones in ⟦5⟧ by integer translation.

2.3.1. *The case of polynomial functions [the polyphase matrix, joint work with Ola Bratteli]*

One problem occurring in the biorthogonal context which does not have an analogue in the orthogonal setting stems from the fact that the duality relations

$$\sum_{w^N=z} \overline{m_i(w)}\,\tilde{m}_j(w) = N\delta_{i,j} \qquad \text{for } i,j = 0,\dots,N-1 \tag{2.3.12}$$

do not give any absolute restrictions on the size of m_i and $\tilde{m}_j$, e.g., a bound on the inner product of two vectors in $\mathbb{C}^N$ does not give a bound on the size of the vectors if they are not equal. This is reflected in the bi-Cuntz relations defined by m_i, $\tilde{m}_i$. Let us now define

$$(S_i f)(z) = m_i(z)\, f\left(z^N\right), \qquad (\tilde{S}_i f)(z) = \tilde{m}_i(z)\, f\left(z^N\right) \tag{2.3.13}$$

for $z \in \mathbb{T}$, $f \in L^2(\mathbb{T})$. Instead of the usual Cuntz relations, the S_i, $\tilde{S}_i$ now satisfy

$$S_i^* \tilde{S}_j = \delta_{i,j} 1, \tag{2.3.14}$$

$$\sum_i S_i \tilde{S}_i^* = 1. \tag{2.3.15}$$

If $A, \tilde{A} \in C(\mathbb{T}, \mathrm{GL}_N(\mathbb{C}))$ are the matrix-valued functions associated to m_i and $\tilde{m}_i$ by

$$m(z) = A\left(z^N\right) v(z), \qquad \tilde{m}(z) = \tilde{A}\left(z^N\right) v(z), \tag{2.3.16}$$

we compute

$$S_i^* S_j = (AA^*)_{j,k} \tag{2.3.17}$$

in the sense that $S_i^* S_j$ is contained in the commutative algebra of multiplication operators on $L^2(\mathbb{T})$ defined by $C(\mathbb{T})$, and $(AA^*)_{j,i} \in C(\mathbb{T})$. Correspondingly,

$$\tilde{S}_i^* \tilde{S}_j = (\tilde{A}\tilde{A}^*)_{j,i} \tag{2.3.18}$$

so all the operators $S_i^* S_j$, $\tilde{S}_i^* \tilde{S}_j$ are contained in the abelian algebra $C(\mathbb{T})$. We may introduce operators S, $\tilde{S}$ from

$$L^2(\mathbb{T})^N = \underset{0}{L^2(\mathbb{T})} \oplus \cdots \oplus \underset{N-1}{L^2(\mathbb{T})} \tag{2.3.19}$$

into $L^2(\mathbb{T})$ by

$$S = (S_0, S_1, \dots, S_{N-1}), \qquad \tilde{S} = (\tilde{S}_0, \dots, \tilde{S}_{N-1}) \tag{2.3.20}$$

and then S^* maps $L^2(\mathbb{T})$ into (2.3.19), etc., and the relations (2.3.14)–(2.3.18) take the form

$$\begin{cases} S^*\tilde{S} = 1, \text{ where } 1 \text{ is the identity in } M_N(\mathbb{C}) \otimes C(\mathbb{T}), \\ S\tilde{S}^* = 1, \text{ where } 1 \text{ is the identity in } C(\mathbb{T}), \end{cases} \tag{2.3.21}$$

$$\begin{cases} S^*S = AA^*, \\ \tilde{S}^*\tilde{S} = \tilde{A}\tilde{A}^*. \end{cases} \tag{2.3.22}$$

These relations say that all combinations of products of S and S^* with $\tilde{S}$ and $\tilde{S}^*$ lie in the algebra $M_N(\mathbb{C}) \otimes C(\mathbb{T})$. But in addition A and $\tilde{A}$ are matrix-valued functions on $\mathbb{T}$, so

$$AA^*\tilde{A}\tilde{A}^* = A\tilde{A}^* = 1 = \tilde{A}\tilde{A}^*AA^* \tag{2.3.23}$$

and hence

$$S^*S = \left(\tilde{S}^*\tilde{S}\right)^{-1} \tag{2.3.24}$$

and all the matrix-valued functions commute.

This discussion can be summarized by saying that the bi-Cuntz relations are much less rigid than the original Cuntz relations, i.e.:

Scholium 2.3.1.1: Given *any* bijective operator S from $L^2(\mathbb{T})^N$ into $L^2(\mathbb{T})$ one may define $\tilde{S} = (S^*)^{-1}$ and the bi-Cuntz relations (2.3.21) are satisfied. If, more specifically, S is given by (2.3.20) and (2.3.13), then operators $\tilde{S}_0, \dots, \tilde{S}_{N-1}$ exist such that the bi-Cuntz relations (2.3.14)–(2.3.15) are satisfied if and only if the operator $A \in M_N(\mathbb{C}) \otimes C(\mathbb{T})$ defined by (2.3.16) is invertible, in which case one must use $\tilde{A} = (A^*)^{-1}$, (2.3.16), and (2.3.13) to define $\tilde{S}_0, \dots, \tilde{S}_{N-1}$.

Let us now connect the filters to the wavelets. We have already defined the scaling functions φ, $\tilde{\varphi}$ and wavelet functions ψ_i, ψ_i, $i = 1, \dots, N$. The expansions for φ and $\tilde{\varphi}$ converge uniformly on compacts, thus $\hat{\varphi}$ and $\hat{\tilde{\varphi}}$ are continuous functions on $\mathbb{R}$. To decide that these functions are in $L^2(\mathbb{R})$ one again forms

$$f_\varphi(t) = \mathrm{PER}\left(|\hat{\varphi}|^2\right)(t) = \sum_{n\in\mathbb{Z}} |\hat{\varphi}(t - 2\pi n)|^2 \tag{2.3.25}$$

and $f_{\tilde{\varphi}}$ similarly, and one deduces again from the nonlinear intertwining relation

$$R^k(p(\psi_1, \psi_2)) = p\left(M^k_{m_0}\psi_1, M^k_{m_0}\psi_2\right), \qquad k \in \mathbb{N} \tag{2.3.26}$$

that

$$R_{m_0}(f_\varphi) = f_\varphi, \qquad R_{\tilde{m}_0}(f_{\tilde{\varphi}}) = f_{\tilde{\varphi}}. \tag{2.3.27}$$

2.3.2. *General results in mathematics on matrix functions*

In the standard case of the good old orthogonal wavelets in $L^2(\mathbb{R})$ of N subbands, we will look for functions $\psi_1, \ldots, \psi_{N-1}$ in $L^2(\mathbb{R})$ such that, if k and n run independently over all the integers $\mathbb{Z}$, i.e., $-\infty < k, n < \infty$, then the countably infinite system of functions

$$\left\{ N^{k/2} \psi_i \left(N^k x - n \right) \bigm| i = 1, \ldots, N-1,\ k, n \in \mathbb{Z} \right\} \tag{2.3.28}$$

is an *orthonormal basis* in the Hilbert space $L^2(\mathbb{R})$. The second half of the word "orthonormal" refers to the restricting requirement that all the functions $\psi_1, \ldots, \psi_{N-1}$ satisfy

$$\int_{\mathbb{R}} \left| \psi_i(x) \right|^2 \, dx = 1, \tag{2.3.29}$$

or stated more briefly,

$$\left\| \psi_i \right\|_{L^2(\mathbb{R})} = 1; \tag{2.3.30}$$

or yet more briefly,

$$\left\| \psi_i \right\| = 1. \tag{2.3.31}$$

From familiar properties of the Lebesgue measure on $\mathbb{R}$, it then follows that all the functions

$$\psi_{i,k,n}(x) := N^{k/2} \psi_i \left(N^k x - n \right), \qquad 1 \le i < N,\ k, n \in \mathbb{Z}, \tag{2.3.32}$$

satisfy the normalization, i.e., that

$$\left\| \psi_{i,k,n} \right\| = 1 \qquad \text{for all } i, k, n. \tag{2.3.33}$$

The functions (2.3.32) are said to be *orthogonal* if

$$\int_{\mathbb{R}} \overline{\psi_{i,k,n}(x)} \, \psi_{i',k',n'}(x) \, dx = 0 \tag{2.3.34}$$

whenever $(i, k, n) \neq (i', k', n')$. We say that the two triple indices are different if $i \neq i'$ or $k \neq k'$ or $n \neq n'$. If, for example, $i = i'$ and $k = k'$, then when the same function is translated by different amounts n and n', the two resulting functions are required to be orthogonal. It is an elementary geometric fact from the theory of Hilbert space that if the functions in (2.3.32) form an orthonormal basis, then for every function $f \in L^2(\mathbb{R})$, i.e., every measurable function f on $\mathbb{R}$ such that

$$\left\| f \right\|^2 = \int_{\mathbb{R}} \left| f(x) \right|^2 \, dx < \infty, \tag{2.3.35}$$

we have the identity

$$\|f\|^2 = \sum_{i,k,n} \left| \int_{\mathbb{R}} \overline{\psi_{i,k,n}(x)}\, f(x)\; dx \right|^2, \tag{2.3.36}$$

where the triple summation in (2.3.36) is over all configurations $1 \leq i < N$, $k, n \in \mathbb{Z}$. It is convenient to rewrite (2.3.36) in the following more compact form:

$$\|f\|^2 = \sum_{i,k,n} |\langle\, \psi_{i,k,n} \mid f \,\rangle|^2 . \tag{2.3.37}$$

Surprisingly, it turns out that (2.3.37) may hold even if the functions $\psi_{i,k,n}$ of (2.3.32) do not form an orthonormal basis. It may happen that one of the initial functions $\psi_1, \ldots,$ or ψ_{N-1} satisfies $\|\psi_i\| < 1$, and yet that (2.3.37) holds for all $f \in L^2(\mathbb{R})$. These more general systems are still called wavelets, but since they are special, they are referred to as *tight frames*, as opposed to orthonormal bases. In either case, we will talk about a *wavelet expansion* of the form

$$f(x) = \sum_{i,k,n} \langle\, \psi_{i,k,n} \mid f \,\rangle\, \psi_{i,k,n}(x) . \tag{2.3.38}$$

It follows that the sum on the right-hand side in (2.3.38) converges in the norm of $L^2(\mathbb{R})$ for all functions f in $L^2(\mathbb{R})$ if (2.3.37) holds.

But there is a yet more general form of wavelets, called *biorthogonal*. The conditions on the functions $\psi_1, \ldots, \psi_{N-1}$ are then much less restrictive than the orthogonality axioms. Hence these wavelets are more flexible and adapt better to a variety of applications, for example, to data compression, or to computer graphics. But the biorthogonality conditions are also a little more technical to state. We say that some given functions ψ_i, $i = 1, \ldots, N-1$, in $L^2(\mathbb{R})$ are part of a biorthogonal wavelet system if there is a second system of functions $\tilde{\psi}_i$, $i = 1, \ldots, N-1$, in $L^2(\mathbb{R})$, such that every $f \in L^2(\mathbb{R})$ admits a representation

$$f(x) = \sum_{i,k,n} \langle\, \psi_{i,k,n} \mid f \,\rangle\, \tilde{\psi}_{i,k,n}(x) = \sum_{i,k,n} \langle\, \tilde{\psi}_{i,k,n} \mid f \,\rangle\, \psi_{i,k,n}(x) , \tag{2.3.39}$$

and

$$\tilde{\psi}_{i,k,n}(x) = N^{k/2} \tilde{\psi}_i\left(N^k x - n\right) . \tag{2.3.40}$$

In the standard normalized case where $\langle\, \psi_i \mid \tilde{\psi}_i \,\rangle = 1$, then you will notice that condition (2.3.37) turns into

$$\|f\|^2 = \sum_{i,k,n} \overline{\langle\, \psi_{i,k,n} \mid f \,\rangle}\, \langle\, \tilde{\psi}_{i,k,n} \mid f \,\rangle \tag{2.3.41}$$

for all $f \in L^2(\mathbb{R})$.

The orthogonal wavelets correspond to matrix functions $\mathbb{T} \to \mathrm{U}_N(\mathbb{C})$, while the wider class of biorthogonal wavelets corresponds to the much bigger group of matrix functions $\mathbb{T} \to \mathrm{GL}_N(\mathbb{C})$, via the associated wavelet filters. You may ask, why bother with the more technical-looking biorthogonal systems? It turns out that they are forced on us by the engineers. They tell us that the real world is not nearly as orthogonal as the mathematicians would like to make it out to be. There is a paucity of symmetric orthogonal wavelets, and symmetry ("linear phase") is prized by engineers and workers in image processing, where the more general wavelet families and their duality play a crucial role. Now what if we could change the biorthogonal wavelets into the orthogonal ones, and still keep the essential spectral properties intact? Then everyone will be happy. This last chapter shows that it is possible, and even in a fairly algorithmic fashion, one that is amenable to computations.

Wavelet filters may be understood as matrix functions, i.e., functions from the one-torus $\mathbb{T} \subset \mathbb{C}$ into some group of invertible matrices. If the scale number is N, then there are three such matrix groups which are especially relevant for wavelet analysis:

$\mathrm{U}_N(\mathbb{C})$: all unitary $N \times N$ complex matrices	$\subset$	$\mathrm{GL}_N(\mathbb{C})$: all invertible $N \times N$ complex matrices	$\supset$	$\mathrm{SL}_N(\mathbb{C})$: all $N \times N$ complex matrices A with $\det A = 1$.

It is possible to reduce some questions in the GL_N case to better understood results for $\mathrm{U}_N(\mathbb{C})$; see Chapter 6 of [BrJo02b]. The SL_2 case is especially interesting in view of Daubechies–Sweldens lifting for dyadic wavelets; see § 2.2.4.

2.3.3. *Connection between matrix functions and wavelets*

Definitions: A function, or a distribution, φ satisfying (2.3.7) is said to be *refinable*, the equation (2.3.7) is called the *refinement equation*, or also, as noted above, the "scaling identity", and φ is called the scaling function. The coefficients a_n of (2.3.7) are called the *masking coefficients*.

We will mainly concentrate on the case when the set $\{a_n\}$ is finite. But in general, a function $\varphi \in L^2(\mathbb{R})$ is said to be refinable with scale number N if $\varphi(x/N)$ is in the L^2-closed linear span of the translates $\{\varphi(x-k)\}_{k\in\mathbb{Z}} \subset$

$L^2(\mathbb{R})$; see, e.g., [HSS96, SSZ99, StZh98, StZh01].

Since there are refinement operations which are more general than scaling (see for example [DLLP01]), there are variations of (2.3.7) which are correspondingly more general, with regard to both the refinement steps that are used and the dimension of the spaces. The term "scaling identity" is usually, but not always, reserved for (2.3.7), while more general refinements lead to "refinement equations". However, (2.3.7) often goes under both names. The vector versions of the identities get the prefix "multi-", for example *multiscaling* and *multiwavelet*.

If m_0 satisfies a condition for obtaining orthogonal wavelets,

$$\sum_{w^N=z} |m_0(w)|^2 = N, \tag{2.3.42}$$

together with the normalization

$$m_0(1) = \sqrt{N}, \tag{2.3.43}$$

then (2.3.7) has a solution φ in $L^2(\mathbb{R})$ which can be obtained by taking the inverse Fourier transform of the product expansion

$$\hat{\varphi}(t) = \prod_{k=1}^{\infty} \left(\frac{m_0\left(tN^{-k}\right)}{\sqrt{N}} \right). \tag{2.3.44}$$

(Here and later we use the convention that if $m(z)$ is a function of $z \in \mathbb{T}$, then $m(t) = m\left(e^{-it}\right)$.) That (2.3.44) gives a solution φ of (2.3.7) follows from the relation

$$\hat{\varphi}(t) = \frac{1}{\sqrt{N}} m_0\left(\frac{t}{N}\right) \hat{\varphi}\left(\frac{t}{N}\right). \tag{2.3.45}$$

2.3.3.1 *Multiresolution wavelets*

We mentioned that there is a direct connection between $m_0 = \sum a_n z^n$ and the scaling function φ on $\mathbb{R}$ given in (1.1.34), (2.3.7), and (2.3.44). There is a similar correspondence between the high-pass filters m_i and the wavelet generators $\psi_i \in L^2(\mathbb{R})$. In the *biorthogonal* case, there is a second system $\tilde{m}_i \leftrightarrow \tilde{\psi}_i$ and the two systems

$$\left\{ N^{\frac{j}{2}} \psi_i\left(N^j x - k\right) \right\} \quad \text{and} \quad \left\{ N^{\frac{j'}{2}} \tilde{\psi}_{i'}\left(N^{j'} x - k'\right) \right\},$$
$$i, i' \in \{1, 2, \dots, N-1\}, \; j, j', k, k' \in \mathbb{Z}, \tag{2.3.46}$$

then form a dual wavelet basis, or dual wavelet frame for $L^2(\mathbb{R})$ in the sense of [Dau92], Chapter 5. We considered this biorthogonal case in more detail in § 2.3.1 above. Much more detail can be found in Chapter 6 of [BrJo02b].

The idea of constructing maximally smooth wavelets when some side conditions are specified has been central to much of the activity in wavelet analysis and its applications since the mid-1980's. As a supplement to [Dau92], the survey article [Stra93] is enjoyable reading. The paper [LaHe96] treats the issue in a more specialized setting and is focussed on the moment method. Some of the early applications to data compression and image coding are done very nicely in [HSS + 95], [SHS + 99], and [HSW95]. An interesting, related but different, algebraic and geometric approach to the problem is offered in [PeWi99].

We now turn to an interesting variation of this setup, which includes higher dimensions, i.e., when the Hilbert space is $L^2\left(\mathbb{R}^d\right)$, $d = 2, 3, \dots$. Staying for the moment with $d = 1$, and N fixed, we will take the viewpoint of what is called *resolutions*, but here understood in a broad sense of closed subspaces: A closed linear subspace $\mathcal{V} \subset L^2(\mathbb{R})$ is said to be an N-resolution if it is invariant under the unitary operator

$$U = U_N : f \longmapsto N^{-\frac{1}{2}} f\left(\frac{x}{N}\right), \tag{2.3.47}$$

i.e., if U maps $\mathcal{V}$ into a proper subspace of itself. The subspace $\mathcal{V}$ is said to be *translation invariant* if

$$f \in \mathcal{V} \iff f(\,\cdot - k) \in \mathcal{V} \qquad \text{for all } k \in \mathbb{Z}. \tag{2.3.48}$$

If there is a function φ such that $\mathcal{V} = \mathcal{V}_\varphi$ is the closed linear span of

$$\{\varphi(\,\cdot - k) \mid k \in \mathbb{Z}\}, \tag{2.3.49}$$

then clearly $\mathcal{V}$ is translation invariant. The translation-invariant resolution subspaces $\mathcal{V}$ are actively studied and reasonably well understood. If $\mathcal{V}$ is of the form $\mathcal{V}_\varphi$ in (2.3.49), then we say that it is *singly generated*, and that φ is a scaling function of scale N.

2.3.3.2 *Generalized multiresolutions [joint work with L. Baggett, K. Merrill, and J. Packer]*

The case when the resolution subspace $\mathcal{V}$ is not singly generated is also interesting, and these resolution subspaces are frequently called *generalized multiresolution subspaces* (GMRA). There is much current and very active research on them; see, for example, [BaLa99], [LPT01], [BaMe99], [HLPS99], [HSS01], [SSZ99], and [Jor01a]. The case when $\mathcal{V}$ is not singly generated as a resolution subspace of scale $N > 2$, i.e., when $\mathcal{V}$ is not of the form (2.3.49), occurs in the study of *wavelet sets*. A wavelet set in $\mathbb{R}^d$ is defined relative to an expansive $d \times d$ matrix $\mathbf{N}$ over $\mathbb{Z}$. A subset $E \subset \mathbb{R}^d$ is said to be an $\mathbf{N}$-wavelet set if there is a single wavelet function $\psi \in L^2\left(\mathbb{R}^d\right)$ such that $\hat{\psi} = \chi_E$. Specifically, the condition states that the family

$$\left\{ |\det \mathbf{N}|^{j/2}\, \psi\left(\mathbf{N}^j x - \mathbf{k}\right) : j \in \mathbb{Z},\ \mathbf{k} \in \mathbb{Z}^d \right\} \tag{2.3.50}$$

is an orthonormal basis for $L^2\left(\mathbb{R}^d\right)$. This can be checked to be equivalent to the combined set of two tiling properties for E as a subset of $\mathbb{R}^d$:

(a) the family of subsets $\left\{ \mathbf{N}^j E : j \in \mathbb{Z} \right\}$ tiles $\mathbb{R}^d$;
(b) the translates $\left\{ E + 2\pi\mathbf{k} : \mathbf{k} \in \mathbb{Z}^d \right\}$ tile $\mathbb{R}^d$.

We define tiling by the requirement that the sets in the family have overlap at most of measure zero relative to Lebesgue measure on $\mathbb{R}^d$. Similarly, the union

$$\mathbb{R}^d = \bigcup_{j\in\mathbb{Z}} \mathbf{N}^j E = \bigcup_{\mathbf{k}\in\mathbb{Z}^d} E + 2\pi\mathbf{k} \tag{2.3.51}$$

is understood to be only up to measure zero.

It is easy to see that compactly supported wavelets in $L^2\left(\mathbb{R}^d\right)$ are MRA wavelets, while most wavelets $\psi = \left(\chi_E\right)^\vee$ from wavelet sets E are not. Thess wavelets are typically (but not always) frequency localized.

The main difference between the GMRA (stands for generalized multiresolution analysis) wavelets and the more traditional MRA ones may be understood in terms of multiplicity. Both come from a fixed resolution subspace $\mathcal{V}_0 \subset L^2\left(\mathbb{R}^d\right)$ which is invariant under the translations $\left\{ T_n : n \in \mathbb{Z}^d \right\}$ where

$$\left(T_n f\right)(x) := f\,(x - n) \qquad \text{for } x \in \mathbb{R}^d \text{ and } n \in \mathbb{Z}^d. \tag{2.3.52}$$

Hence $\{T_n|_{\mathcal{V}_0}\}_{n\in\mathbb{Z}^d}$ is a unitary representation of $\mathbb{Z}^d$ on the Hilbert space $\mathcal{V}_0$. As a result of Stone's theorem, we find that there are subsets

$$E_1 \supset E_2 \supset \cdots \supset E_j \supset \cdots$$

of $\mathbb{T}^d$ such that the spectral measure of the (restricted) representation has multiplicity $\geq j$ on the subset E_j, $j = 1, 2, \ldots$. It can be checked that the projection-valued spectral measure is absolutely continuous. Moreover, there is an intertwining unitary operator

$$J\colon \mathcal{V}_0 \longrightarrow \sum_{j\geq 1}^{\oplus} L^2\left(E_j\right) \tag{2.3.53}$$

such that

$$P_{L^2(E_j)} J T_n f\,(z) = z^n\,(Jf)\,(z) \tag{2.3.54}$$

holds for all $f \in \mathcal{V}_0$ and $z \in E_j$. We may then consider the functions $\varphi_j \in \mathcal{V}_0$ $(\subset L^2(\mathbb{R}^d))$ defined by

$$\varphi_j := J^{-1}(0, \dots, 0, \underbrace{\chi_{E_j}}_{j\text{'th place}}, 0, 0, \dots). \tag{2.3.55}$$

It was proved by Baggett and Merrill [BaMe99] that $\{ \varphi_j : j \geq 1 \}$ generates a normalized tight frame for $\mathcal{V}_0$: specifically, that

$$\sum_{j \geq 1} \sum_{n \in \mathbb{Z}^d} |\langle T_n \varphi_j \mid f \rangle|^2 = \|f\|^2_{L^2(\mathbb{R}^d)} \tag{2.3.56}$$

holds for all $f \in \mathcal{V}_0$.

Treating $(\varphi_1, \varphi_2, \dots)$ as a vector-valued function, denoted simply by φ, we see that there is a matrix function

$$H : \mathbb{T}^d \longrightarrow (\text{complex square matrices})$$

such that

$$\hat{\varphi}\left(\mathbf{N}^{\overset{\text{for transpose}}{\nearrow}\text{tr}} t\right) = H\left(e^{it}\right) \hat{\varphi}(t), \tag{2.3.57}$$

where $t = (t_1, \dots, t_d) \in \mathbb{R}^d$, and $e^{it} := \left(e^{it_1}, e^{it_2}, \dots, e^{it_d}\right)$.

But this method takes the Hilbert space $L^2(\mathbb{R}^d)$ as its starting point, and then proceeds to the construction of wavelet filters in the form (2.3.57). Our current joint work with Baggett, Merrill, and Packer reverses this. It begins with a matrix function H defined on $\mathbb{T}^d$, and then offers subband conditions on the matrix function which allow the construction of a GMRA for $L^2(\mathbb{R}^d)$ with generator $\varphi = (\varphi_1, \varphi_2, \dots)$ given by (2.3.57). So the Hilbert space $L^2(\mathbb{R}^d)$ shows up only at the end of the construction, in the conclusions of the theorems.

2.3.4. *Matrix completion*

In using the polyphase matrices, one may only have the first few rows, and be faced with the problem of completing to get the entire function A from a torus into the matrices of the desired size. The case when only the first row is given, say corresponding to a specified low-pass filter, is treated in [BrJo02b] and [BrJo02a], and we refer the reader to the references given there, especially [JiSh94], [RiSh91], [RiSh92], and [Vai93].

The wavelet transfer operator is used in a variety of wavelet applications not covered here, or only touched upon tangentially: stability of refinable functions, regularity, approximation order, unitary matrix extension principles, to mention only a few. The reader is referred to the following papers for more details on these subjects: [DHRS03], [RoSh03], [RST01], [JJS01], [RoSh00], [RiSh00], [JiSh99], [She98], [RoSh98], [LLS98], [RoSh97], [BJMP03], and [BJMP04].

The unitary extension principle (UEP) of [DHRS03] involves the interplay between a finite set of filters (functions on $\mathbb{R}/\mathbb{Z}$), and a corresponding tight frame (alias Parseval frame) in $L^2\left(\mathbb{R}^d\right)$.

For the sake of illustration, let us take $d = 1$, and scaling number $N = 2$, i.e., the case of dyadic framelets. Naturally, the notion of tight frame is weaker than that of an orthonormal basis (ONB), and it is shown in [DHRS03] that when a system of wavelet filters m_i, $i = 0, 1, \ldots, r$ is given (m_0 must be low-pass), then the orthogonality condition on the m_i's always gets us a framelet in $L^2(\mathbb{R})$, i.e., the functions ψ_i corresponding to the high-pass filters m_i, $i = 1, \ldots, r$, generate a tight frame for $L^2(\mathbb{R})$, also called a framelet. The correspondence m_i to ψ_i is called the UEP in [DHRS03].

The orthogonality condition for m_i, $i = 0, 1, \ldots, r$, referred to in the UEP is simply this: Form an $(r+1)$-by-2 matrix-valued function $F(x)$ by using $m_i(x)$, $i = 0, 1, \ldots, r$ in the first column, and the translates of the m_i's by a half period, i.e., $m_i(x + 1/2)$, $i = 0, 1, \ldots, r$ in the second. The condition on this matrix function $F(x)$ is that the two columns are orthogonal and have unit norm in ℓ^2 for all x. Note that we still get the unitary matrix functions acting on these systems, in the way we outlined above. But there is redundancy as the unitary matrices are $(r+1)$-by-$(r+1)$. The reader is refered to [DHRS03] for further details.

We emphasize that several of these, and other related topics, invite the kind of probabilistic tools that we have stressed here. But a more systematic discussion is outside the scope of this brief set of notes. We only hope to offer a modest introduction to a variety of more specialized topics.

Remark 2.3.4.1: The orthogonality condition for m_i, $i = 0, 1, \ldots, r$, may be stated in terms of the operators S_i from equation (2.9), $N = 2$. For each $i = 0, 1, \ldots, r$, define an operator on $L^2(\mathbb{R}/\mathbb{Z})$ as in (2.9). Then the arguments from Section 2 show that the orthogonality condition for m_i, $i = 0, 1, \ldots, r$, i.e., the UEP condition, is equivalent to the operator identity (2.8) where the summation now runs from 0 to r. Operator systems S_i satisfying (2.8) are called row-isometries.

Remark 2.3.4.2: There are two properties of the low-pass filter m_0 which we have glossed over. First, m_0 must be such that the corresponding scaling function φ is in $L^2(\mathbb{R})$. Without an added condition on m_0, φ might only be a distribution. Secondly, when the dyadic scaling in $L^2(\mathbb{R})$ is restricted to the resolution subspace $V_0(\varphi)$, the corresponding unitary part must be zero. These two issues are addressed in [BJMP03], [BJMP04], and [DHRS03].

2.3.5. *Connections between matrix functions and signal processing*

Since our joint work with Baggett, Merrill, and Packer on the GMRA wavelets is still in progress, we restrict the discussion of matrix functions here to the MRA case.

The two groups of matrix functions $C(\mathbb{T}, \mathrm{U}_N(\mathbb{C}))$ and $C(\mathbb{T}, \mathrm{GL}_N(\mathbb{C}))$, i.e., the continuous functions from the torus into the respective groups, enter wavelet analysis via the associated wavelet filters $(m_i)_{i=0}^{N-1}$.

In [BrJo02b] (see also § 1.1.3 above), we give the details of the multiple correspondence between:

(i) matrix functions, $A\colon \mathbb{T} \to \mathrm{GL}_N(\mathbb{C})$,
(ii) high- and low-pass wavelet filters m_i, $\tilde{m}_{i'}$, $i, i' = 0, 1, \dots, N-1$, and
(iii) wavelet generators ψ_i, $\tilde{\psi}_{i'}$, $i, i' = 1, \dots, N-1$, together with scaling functions φ, $\tilde{\varphi}$.

In particular,

$$A_{i,j}(z) = \frac{1}{N} \sum_{w^N = z} m_i(w)\, w^{-j}, \qquad z \in \mathbb{T}, \tag{2.3.58}$$

$$\left(A^{-1}\right)_{i,j} = \frac{1}{N} \sum_{w^N = z} \overline{\tilde{m}_j(w)}\, w^{i}, \qquad z \in \mathbb{T}. \tag{2.3.59}$$

The dependence of the $L^2(\mathbb{R})$-functions in (iii) on the group elements A from (i) gives rise to homotopy properties. The standard orthogonal wavelets represent the special case when $m_i = \tilde{m}_i$, or equivalently, $A(z) = \left((A(z))^*\right)^{-1}$, $z \in \mathbb{T}$. Hence, the matrix functions are unitary in this case.

The scaling/wavelet functions $\varphi, \psi_1, \dots, \psi_{N-1}$ with support on a fixed compact interval, say $[0, kN+1]$, $k = 0, 1, \dots$, can be parameterized with a finite number of parameters since the unitary-valued function $z \to A(z)$ in (2.3.58) then is a polynomial in z of degree at most $k(N-1)$. It is well-known folklore from computer-generated pictures that the shape of the scaling/wavelet functions depends continuously on these parameters; see Figures 1.1–1.7 in [BrJo02b] and [Tre01].

The scaling function $\varphi \in L^2(\mathbb{R})$ of (2.3.7) is illustrated there, in the case $N = 2$, and for orthogonal $\mathbb{Z}$-translates, i.e., the case (2.3.42). These pictures illustrate the dependence of φ on the masking coefficients (a_n) in the case of [Tre01]:

$$
\begin{aligned}
a_0 &= (\eta_0 - \eta_1 - \eta_2 + \eta_3 + \eta_4)/4, \\
a_1 &= (\eta_0 + \eta_1 - \eta_2 + \eta_3 - \eta_4)/4, \\
a_2 &= (\eta_0 - \eta_3 - \eta_4)/2, \\
a_3 &= (\eta_0 - \eta_3 + \eta_4)/2, \\
a_4 &= (\eta_0 + \eta_1 + \eta_2 + \eta_3 + \eta_4)/4, \\
a_5 &= (\eta_0 - \eta_1 + \eta_2 + \eta_3 - \eta_4)/4,
\end{aligned}
\tag{2.3.60}
$$

where

$$
\begin{aligned}
&\eta_0 = 1/\sqrt{2}, \\
&\eta_1 = (\cos 2\theta + \cos 2\rho)/\sqrt{2}, \qquad \eta_2 = (\sin 2\theta + \sin 2\rho)/\sqrt{2}, \\
&\eta_3 = \cos(2\theta - 2\rho)/\sqrt{2}, \qquad \eta_4 = \sin(2\theta - 2\rho)/\sqrt{2}.
\end{aligned}
\tag{2.3.61}
$$

These formulas arise from an independent pair of rotations by angles θ and ρ of two "spin vectors", i.e., by taking the matrix function A in (2.3.58) unitary, $\mathbb{T} \ni z \to A_{\theta,\rho}(z) \in \mathrm{U}_2(\mathbb{C})$, and setting

$$A(z) = V(Q_\theta^\perp + zQ_\theta)(Q_\rho^\perp + zQ_\rho) = VU_\theta(z)U_\rho(z) \tag{2.3.62}$$

with

$$V = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \tag{2.3.63}$$

$$
\begin{aligned}
Q_\theta &= \begin{pmatrix} \cos^2\theta & \cos\theta\sin\theta \\ \cos\theta\sin\theta & \sin^2\theta \end{pmatrix} \\
&= \frac{1}{2}\left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix} \right),
\end{aligned}
\tag{2.3.64}
$$

and the orthogonal complement to the one-dimensional projection Q_θ,

$$Q_\theta^\perp = Q_{\theta+(\pi/2)}. \tag{2.3.65}$$

With the coefficients $a_0, a_1, a_2, a_3, a_4, a_5$ given by (2.3.60), the algorithmic approach to graphing the solution φ to the scaling identity (2.3.7) is as follows (see [Jor01b], [Tre01] for details): the relation (2.3.7) for $N = 2$ is interpreted as giving the values of the left-hand φ by an operation performed on those of the φ on the right, and a binary digit inversion transforms this into the form

$$\mathbf{f}'_{k+1}\left(x + \frac{1}{2^{k+1}}\right) = \mathbf{A}\mathbf{f}_k(x), \tag{2.3.66}$$

where $\mathbf{A}$ is the 2×3 matrix $\mathbf{A}_{i,j} = \sqrt{2}a_{4+i-2j}$ constructed from the coefficients in (2.3.7), and $\mathbf{f}_j$ and $\mathbf{f}'_j$ are the vector functions

$$\mathbf{f}_j(x) = \begin{pmatrix} \varphi\left(x - \frac{2}{2^j}\right) \\ \varphi\left(x - \frac{1}{2^j}\right) \\ \varphi(x) \end{pmatrix}, \qquad \mathbf{f}'_j(x) = \begin{pmatrix} \varphi\left(x - \frac{1}{2^j}\right) \\ \varphi(x) \end{pmatrix}. \tag{2.3.67}$$

The signal processing aspect can be understood from the description of subband filters in the analusis and synthesis of time signals, or more general signals for images. In either case, we have two subband systems $m = (m_0, m_1, \dots)$ and $\tilde{m} = (\tilde{m}_0, \tilde{m}_1, \dots)$ where the functions

$$m_j(z) = \sum_n a_n^{(j)} z^n \quad \text{and} \quad \tilde{m}_j(z) = \sum_n \tilde{a}_n^{(j)} z^n$$

are the generating functions defined from the filter coefficients $\left(a_n^{(j)}\right)$ and $\left(\tilde{a}_n^{(j)}\right)$, $n \in \mathbb{Z}^d$.

Appendix A: Topics for further research

Originally we had anticipated adding two more chapters to these tutorials, but time and space prevented this. Instead we include the table of contents for this additional material. The details for the remaining chapters will be published elsewhere. But as the items in the list of contents suggest, there are still many exciting open problems in the subject that the reader may wish to pursue on his/her own. We feel that the following list of topics offers at least an outline of several directions that the reader, could take in his/her own study and research on wavelet-related mathematics.

Appendix B: Duality principles in analysis

Several versions of spectral duality are presented. On the two sides we present (1) a basis condition, with the basis functions indexed by a frequency variable, and giving an orthonormal basis; and (2) a geometric notion which takes the form of a tiling, or a Iterated Function System (IFS). Our initial motivation derives from the Fuglede conjecture, see [Fug74, Jor82, JoPe92]: For a subset D of $\mathbb{R}^n$ of finite positive measure, the Hilbert space $L^2(D)$ admits an orthonormal basis of complex exponentials, i.e., D admits a Fourier basis with some frequencies L from $\mathbb{R}^n$, if and only if D tiles $\mathbb{R}^n$ (in the measurable category) where the tiling uses only a set T of vectors in $\mathbb{R}^n$. If some D has a Fourier basis indexed by a set L, we say that (D, L) is a spectral pair. We recall from [JoPe99] that if D is an n-cube, then the sets L in (1) are precisely the sets T in (2). This begins with work of Jorgensen and Steen Pedersen [JoPe99] where the admissible sets $L = T$ are characterized. Later it was shown, [IoPe98] and [LRW00] that the identity $T = L$ holds for all n. The proofs are based on general Fourier duality, but they do not reveal the nature of this common set $L = T$. A complete list is known only for $n = 1$, 2, and 3, see [JoPe99].

We then turn to the scaling IFS's built from the n-cube with a given expansive integral matrix A. Each A gives rise to a fractal in the small, and a dual discrete iteration in the large. In a different paper [JoPe98], Jorgensen and Pedersen characterize those IFS fractal limits which admit Fourier duality. The surprise is that there is a rich class of fractals that do have Fourier duality, but the middle third Cantor set does not. We say that an affine IFS, built on affine maps in $\mathbb{R}^n$ defined by a given expansive integral matrix A and a finite set of translation vectors, admits Fourier duality if the set of points L, arising from the iteration of the A-affine maps in the large, forms an orthonormal Fourier basis (ONB) for the corresponding fractal μ in the small, i.e., for the iteration limit built using the inverse contractive maps, i.e., iterations of the dual affine system on the inverse matrix A^{-1}. By "fractal in the small", we mean the Hutchinson measure μ and its compact support, see [Hut81]. (The best known example of this is the middle-third Cantor set, and the measure μ whose distribution function is corresponding Devil's staircase.)

In other words, the condition is that the complex exponentials indexed by L form an ONB for $L^2(\mu)$. Such duality systems are indexed by complex Hadamard matrices H, see [JoPe99] and [JoPe98]; and the duality issue is connected to the spectral theory of an associated Ruelle transfer operator, see [BrJo02b]. These matrices H are the same Hadamard matrices which index a certain family of quasiperiodic spectral pairs (D, L) studied in [Jor82] and [JoPe92]. They also are used in a recent construction of Terence Tao [Tao04] of a Euclidean spectral pair (D, L) in $\mathbb{R}^5$ for which D does not a tile $\mathbb{R}^5$ with any set of translation vectors T in $\mathbb{R}^5$; see also [IKT03].

We finally report on joint research with Dorin Dutkay [DuJo03], [DuJo04a], [DuJo04b], [DuJo04c] where we show that all the affine IFS's, and more general limit systems from dynamics and probability theory, admit wavelet constructions, i.e., admit orthonormal bases of wavelet functions in Hilbert spaces which are constructed directly from the geometric data. A substantial part of the picture involves the construction of limit sets and limit measures, a part of geometric measure theory.

Acknowledgements: We are happy to thank the organizing committee at the National University of Singapore for all their dedicated work in planning and organizing a successful conference, of which this tutorial is a part. We especially thank Professors Wai Shing Tang, Judith Packer, Zuowei Shen, and the head of the Department of Mathematics of the NUS for all their work in making my visit to Singapore possible. We thank the Institute for Mathematical Sciences at the NUS, and the US National Science Foundation, for partial financial support in the preparation of these lecture notes. We discussed various parts of the mathematics with our colleagues, Professors Larry Baggett, David Larson, Ola Bratteli, Kathy Merrill, Judy Packer, and we thank them for their encouragements and suggestions. The typesetting and graphics were expertly done at the University of Iowa by Brian Treadway. We also thank Brian Treadway for a number of corrections, and for very helpful suggestions, leading to improvements of the presentation.

References

Aus95. P. Auscher, Solution of two problems on wavelets, J. Geom. Anal. 5 (1995), 181–236.

BJMP03. L.W. Baggett, P.E.T. Jorgensen, K.D. Merrill, and J.A. Packer, Construction of Parseval wavelets from redundant filter systems, preprint, 2003, submitted to J. Amer. Math. Soc., arXiv:math.CA/0405301.

BJMP04. L.W. Baggett, P.E.T. Jorgensen, K.D. Merrill, and J.A. Packer, An analogue of Bratteli-Jorgensen loop group actions for GMRA's, in Wavelets, Frames, and Operator Theory (College Park, MD, 2003), ed. C. Heil, P.E.T. Jorgensen, and D.R. Larson, Contemp. Math., vol. 345, American Mathematical Society, Providence, 2004, pp. 11–25.

BaLa99. L.W. Baggett and D.R. Larson (eds.), The Functional and Harmonic Analysis of Wavelets and Frames: Proceedings of the AMS Special Session on the Functional and Harmonic Analysis of Wavelets Held in San Antonio, TX, January 13–14, 1999, Contemp. Math., vol. 247, American Mathematical Society, Providence, 1999.

BaMe99. L.W. Baggett and K.D. Merrill, Abstract harmonic analysis and wavelets in $\mathbf{R}^n$, in The Functional and Harmonic Analysis of Wavelets and Frames (San Antonio, 1999), ed. L.W. Baggett and D.R. Larson, Contemp. Math., vol. 247, American Mathematical Society, Providence, 1999, pp. 17–27.

Ben00. J.J. Benedetto, Ten books on wavelets, SIAM Rev. 42 (2000), 127–138.

BrJo99. O. Bratteli and P.E.T. Jorgensen, Convergence of the cascade algorithm at irregular scaling functions, in The Functional and Harmonic Analysis of Wavelets and Frames (San Antonio, 1999), ed. L.W. Baggett and D.R. Larson, Contemp. Math., vol. 247, American Mathematical Society, Providence, 1999, pp. 93–130.

BrJo02a. ———, Wavelet filters and infinite-dimensional unitary groups, in

Wavelet Analysis and Applications (Guangzhou, China, 1999), ed. D. Deng, D. Huang, R.-Q. Jia, W. Lin, and J. Wang, AMS/IP Studies in Advanced Mathematics, vol. 25, American Mathematical Society, Providence, International Press, Boston, 2002, pp. 35–65.

BrJo02b. ———, Wavelets through a Looking Glass: The World of the Spectrum, Applied and Numerical Harmonic Analysis, Birkhäuser, Boston, 2002.

BreJo91. B. Brenken and P.E.T. Jorgensen, A family of dilation crossed product algebras, J. Operator Theory 25 (1991), 299–308.

CDV93. A. Cohen, I. Daubechies, and P. Vial, Wavelets on the interval and fast wavelet transforms, Appl. Comput. Harmon. Anal. 1 (1993), 54–81.

CoWi93. R.R. Coifman and M.V. Wickerhauser, Wavelets and adapted waveform analysis: A toolkit for signal processing and numerical analysis, in Different Perspectives on Wavelets (San Antonio, TX, 1993), ed. I. Daubechies, Proc. Sympos. Appl. Math., vol. 47, American Mathematical Society, Providence, 1993, pp. 119–153.

Cun77. J. Cuntz, Simple C^*-algebras generated by isometries, Comm. Math. Phys. 57 (1977), 173–185.

DaLa98. X. Dai and D.R. Larson, Wandering vectors for unitary systems and orthogonal wavelets, Mem. Amer. Math. Soc. 134 (1998), no. 640.

Dau92. I. Daubechies, Ten Lectures on Wavelets, CBMS-NSF Regional Conf. Ser. in Appl. Math., vol. 61, SIAM, Philadelphia, 1992.

DGM86. I. Daubechies, A. Grossmann, and Y. Meyer, Painless nonorthogonal expansions, J. Math. Phys. 27 (1986), 1271–1283.

DHRS03. I. Daubechies, B. Han, A. Ron, and Z. Shen, Framelets: MRA-based constructions of wavelet frames, Appl. Comput. Harmon. Anal. 14 (2003), 1–46.

DaSw98. I. Daubechies and W. Sweldens, Factoring wavelet transforms into lifting steps, J. Fourier Anal. Appl. 4 (1998), 247–269.

DuJo03. D.E. Dutkay and P.E.T. Jorgensen, Wavelets on fractals, preprint, University of Iowa, 2003, to appear in Rev. Mat. Iberoamericana, arXiv:math.CA/0305443.

DuJo04a. ———, Martingales, endomorphisms, and covariant systems of operators in Hilbert space, preprint, University of Iowa, 2004, submitted to Amer. J. Math., arXiv:math.CA/0407330.

DuJo04b. ———, Operators, martingales, and measures on projective limit spaces, preprint, University of Iowa, 2004, submitted to Trans. Amer. Math. Soc., arXiv:math.CA/0407517.

DuJo04c. ———, Disintegration of projective measures, preprint, University of Iowa, 2004, submitted to Proc. Amer. Math. Soc., arXiv:math.CA/0408151.

DLLP01. N. Dyn, D. Leviatan, D. Levin, and A. Pinkus (eds.), Multivariate Approximation and Applications, Cambridge University Press, Cambridge, England, 2001.

EsGa77. D. Esteban and C. Galand, Application of quadrature mirror filters to split band voice coding systems, in IEEE International Conference on Acoustics, Speech, and Signal Processing (Washington, DC, May 1977), Institute of Electrical and Electronics Engineers, Piscataway, NJ, 1977, pp. 191–195.

FiWi99. A. Fijany and C.P. Williams, Quantum wavelet transforms: Fast algorithms and complete circuits, in Quantum Computing and Quantum Communications (Palm Springs, CA, 1998), ed. C.P. Williams, Lecture Notes in Computer Science, vol. 1509, Springer, Berlin, 1999, pp. 10–33.

Fre02. M.H. Freedman, Poly-locality in quantum computing, Found. Comput. Math. 2 (2002), 145–154.

Fug74. B. Fuglede, Commuting self-adjoint partial differential operators and a group-theoretic problem, J. Funct. Anal. 16 (1974), 101–121.

Gar99. G. Garrigós, Connectivity, homotopy degree, and other properties of α-localized wavelets on $\mathbb{R}$, Publ. Mat. 43 (1999), 303–340.

Haa10. A. Haar, Zur Theorie der orthogonalen Funktionensysteme, Math. Ann. 69 (1910), 331–371.

HLPS99. D. Han, D. R. Larson, M. Papadakis, and Th. Stavropoulos, Multiresolution analyses of abstract Hilbert spaces and wandering subspaces, in The Functional and Harmonic Analysis of Wavelets and Frames (San Antonio, 1999), ed. L.W. Baggett and D.R. Larson, Contemp. Math., vol. 247, American Mathematical Society, Providence, 1999, pp. 259–284.

HSS01. X.M. He, L.X. Shen, and Z. Shen, A data-adaptive knot selection scheme for fitting splines, IEEE Signal Processing Letters 8 (2001), 137–139.

HSS96. C. Heil, G. Strang, and V. Strela, Approximation by translates of refinable functions, Numer. Math. 73 (1996), 75–94.

HSW95. P.N. Heller, J.M. Shapiro, and R.O. Wells, Optimally smooth symmetric quadrature mirror filters for image coding, in Wavelet Applications II (Orlando, 1995), ed. H.H. Szu, Proceedings of SPIE, vol. 2491, Society of Photo-optical Instrumentation Engineers, Bellingham, WA, 1995, pp. 119–130.

HSS+95. P.N. Heller, V. Strela, G. Strang, P. Topiwala, C. Heil, and L.S. Hills, Multiwavelet filter banks for data compression, in IEEE International Symposium on Circuits and Systems, 1995 (ISCAS '95), Institute of Electrical and Electronics Engineers, New York, 1995, vol. 3, pp. 1796–1799.

HeWe96. E. Hernandez and G. Weiss, A First Course on Wavelets, Studies in Advanced Mathematics, CRC Press, Boca Raton, Florida, 1996.

Hub98. B.B. Hubbard, The World According to Wavelets: The Story of a Mathematical Technique in the Making, second ed., A.K. Peters, Wellesley, MA, 1998.

Hut81. J.E. Hutchinson, Fractals and self similarity, Indiana Univ. Math. J. 30 (1981), 713–747.

HwMa94. W.-L. Hwang and S. Mallat, Characterization of self-similar multifractals with wavelet maxima, Appl. Comput. Harmon. Anal. 1 (1994), 316–328.

IKT03. A. Iosevich, N.Katz, and T. Tao, The Fuglede spectral conjecture holds for convex planar domains, Math. Res. Lett. 10 (2003), 559–569.

IoPe98. A. Iosevich and S. Pedersen, Spectral and tiling properties of the unit cube, Internat. Math. Res. Notices 1998 (1998), 819–828.

JJS01. R.-Q. Jia, Q. Jiang, and Z. Shen, Convergence of cascade algorithms associated with nonhomogeneous refinement equations, Proc. Amer. Math. Soc. 129 (2001), 415–427.

JiSh94. R.Q. Jia and Z. Shen, Multiresolution and wavelets, Proc. Edinburgh Math. Soc. (2) 37 (1994), 271–300.

JiSh99. Q. Jiang and Z. Shen, On existence and weak stability of matrix refinable functions, Constr. Approx. 15 (1999), 337–353.

Jor82. P.E.T. Jorgensen, Spectral theory of finite-volume domains in $\mathbb{R}^n$, Adv. in Math. 44 (1982), 105–120.

Jor01a. ———, Ruelle operators: Functions which are harmonic with respect to a transfer operator, Mem. Amer. Math. Soc. 152 (2001), no. 720.

Jor01b. ———, Minimality of the data in wavelet filters, Adv. Math. 159 (2001), 143–228.

Jor03a. ———, Invited featured book review of *An Introduction to Wavelet Analysis* by David F. Walnut, Applied and Numerical Harmonic Analysis, Birkhäuser, 2002, Bull. Amer. Math. Soc. (N.S.) 40 (2003), 421–427.

Jor03b. ———, Matrix factorizations, algorithms, wavelets, Notices Amer. Math. Soc. 50 (2003), 880–894.

Jor04a. ———, Measures in wavelet decompositions, preprint, University of Iowa, 2004, submitted to Adv. in Appl. Math., arXiv:math.CA/0402024.

Jor04b. ———, Iterated function systems, representations, and Hilbert space, preprint, University of Iowa, 2004, to appear in Internat. J. Math., arXiv:math.CA/0402175.

JoPe92. P.E.T. Jorgensen and S. Pedersen, Spectral theory for Borel sets in $\mathbb{R}^n$ of finite measure, J. Funct Anal. 107 (1992), 72–104.

JoPe98. ———, Dense analytic subspaces in fractal L^2-spaces, J. Analyse Math. 75 (1998), 185–228.

JoPe99. ———, Spectral pairs in Cartesian coordinates, J. Fourier Anal. Appl. 5 (1999), 285–302.

KaLe95. J.-P. Kahane and P.-G. Lemarié-Rieusset, Fourier Series and Wavelets, Studies in the Development of Modern Mathematics, vol. 3, Gordon and Breach, Luxembourg, 1995.

Kai94. G. Kaiser, A Friendly Guide to Wavelets, Birkhäuser, Boston, 1994.

Kla99. A. Klappenecker, Wavelets and wavelet packets on quantum computers, in Wavelet Applications in Signal and Image Processing VII (Denver, 1999), ed. M.A. Unser, A. Aldroubi, and A.F. Laine, Proceedings of SPIE, vol. 3813, Society of Photo-optical Instrumentation Engineers, Bellingham, WA, 1999, pp. 703–713.

LRW00. J.C. Lagarias, J.A. Reeds, and Y. Wang, Orthonormal bases of exponentials for the n-cube, Duke Math. J. 103 (2000), 25–37.

LaHe96. M. Lang and P.N. Heller, The design of maximally smooth wavelets, in IEEE International Conference on Acoustics, Speech, and Signal Processing, 1996 (ICASSP-96), Institute of Electrical and Electronics Engineers, New York, 1996, vol. 3, pp. 1463–1466.

LLS98. W. Lawton, S.L. Lee, and Z. Shen, Convergence of multidimensional cascade algorithm, Numer. Math. 78 (1998), 427–438.

LPT01. L.-H. Lim, J.A. Packer, and K.F. Taylor, A direct integral decomposition of the wavelet representation, Proc. Amer. Math. Soc. 129 (2001), 3057–3067.

Mar82. D. Marr, Vision: A Computational Investigation into the Human Representation and Processing of Visual Information, W.H. Freeman, San Francisco, 1982.

Mey93. Y. Meyer, Wavelets: Algorithms & Applications, SIAM, Philadelphia, 1993, translated from the French and with a foreword by Robert D. Ryan.

Mey98. ———, Wavelets, Vibrations and Scalings, CRM Monograph Series, vol. 9, American Mathematical Society, Providence, 1998.

Mey00. ———, Wavelets and functions with bounded variation from image processing to pure mathematics, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl. 11 (2000), 77–105, special issue: *Mathematics Towards the Third Millennium* (Papers from the International Conference held in Rome, May 27–29, 1999).

MiXu94. C.A. Micchelli and Y. Xu, Using the matrix refinement equation for the construction of wavelets on invariant sets, Appl. Comput. Harmon. Anal. 1 (1994), 391–401.

Per86. A.M. Perelomov, Generalized Coherent States and Their Applications, Texts and Monographs in Physics, Springer-Verlag, Berlin, 1986.

PeWi99. V. Perrier and M.V. Wickerhauser, Multiplication of short wavelet series using connection coefficients, in Advances in Wavelets (Hong Kong, 1997), ed. K.-S. Lau, Springer-Verlag, Singapore, 1999, pp. 77–101.

ReWe98. H.L. Resnikoff and R.O. Wells, Wavelet Analysis: The Scalable Structure of Information, Springer-Verlag, New York, 1998.

RiSh91. S.D. Riemenschneider and Z. Shen, Box splines, cardinal series, and wavelets, in Approximation Theory and Functional Analysis (College Station, Texas, 1990), ed. C.K. Chui, Academic Press, Boston, 1991, pp. 133–149.

RiSh92. ———, Wavelets and pre-wavelets in low dimensions, J. Approx. Theory 71 (1992), 18–38.

RiSh00. ———, Interpolatory wavelet packets, Appl. Comput. Harmon. Anal. 8 (2000), 320–324.

RoSh97. A. Ron and Z. Shen, Affine systems in $L_2(\mathbf{R}^d)$: the analysis of the analysis operator, J. Funct. Anal. 148 (1997), 408–447.

RoSh98. ———, Compactly supported tight affine spline frames in $L_2(\mathbf{R}^d)$, Math. Comp. 67 (1998), 191–207.

RoSh00. ______, The Sobolev regularity of refinable functions, J. Approx. Theory 106 (2000), 185–225.

RoSh03. ______, The wavelet dimension function is the trace function of a shift-invariant system, Proc. Amer. Math. Soc. 131 (2003), 1385–1398.

RST01. A. Ron, Z. Shen, and K.-C. Toh, Computing the Sobolev regularity of refinable functions by the Arnoldi method, SIAM J. Matrix Anal. Appl. 23 (2001), 57–76.

RLL00. M. Rørdam, F. Larsen, and N.J. Laustsen, An Introduction to K-theory for C^*-algebras, Cambridge University Press, Cambridge, 2000.

She98. Z. Shen, Refinable function vectors, SIAM J. Math. Anal. 29 (1998), 235–250.

Stra93. G. Strang, Wavelet transforms versus Fourier transforms, Bull. Amer. Math. Soc. (N.S.) 28 (1993), 288–305.

StNg96. G. Strang and T. Nguyen, Wavelets and Filter Banks, Wellesley-Cambridge Press, Wellesley, Massachusetts, 1996.

SSZ99. G. Strang, V. Strela, and D.-X. Zhou, Compactly supported refinable functions with infinite masks, in The Functional and Harmonic Analysis of Wavelets and Frames (San Antonio, 1999), ed. L.W. Baggett and D.R. Larson, Contemp. Math., vol. 247, American Mathematical Society, Providence, 1999, pp. 285–296.

StZh98. G. Strang and D.-X. Zhou, Inhomogeneous refinement equations, J. Fourier Anal. Appl. 4 (1998), 733–747.

StZh01. ______, The limits of refinable functions, Trans. Amer. Math. Soc. 353 (2001), 1971–1984.

SHS+99. V. Strela, P.N. Heller, G. Strang, P. Topiwala, and C. Heil, The application of multiwavelet filterbanks to image processing, IEEE Transactions on Image Processing 8 (1999), 548–563.

Tao04. T. Tao, Fuglede's conjecture is false in 5 and higher dimensions, Math. Res. Lett. 11 (2004), 251–258.

Tre01. B.F. Treadway, Appendix to [Jor01b].

Vai93. P.P. Vaidyanathan, Multirate Systems and Filter Banks, Prentice Hall, Englewood Cliffs, NJ, 1993.

Wic93. M.V. Wickerhauser, Best-adapted wavelet packet bases, in Different Perspectives on Wavelets (San Antonio, TX, 1993), ed. I. Daubechies, Proc. Sympos. Appl. Math., vol. 47, American Mathematical Society, Providence, 1993, pp. 155–171.