Wavelets: Basic properties, parameterizations and sampling

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

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

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

Daubechies’ construction of continuous, compactly supported orthogonal wavelets [97] was a highlight among theoretical developments. Before then, the very existence of such wavelets was suspect. The dependence of regularity—
as measured by local Hölder exponents—on scaling filters was first addressed in contraction estimates due to Daubechies and Lagarias [107, 108] showing, essentially, how the rate of convergence of the cascade algorithm depends on the spectral structure of the scaling filter. Section 1.1.4 contains details of these estimates.

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

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

1.1 Scaling and multiresolution analysis

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

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

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

$$\frac{1}{2} \phi \left( \frac{x}{2} \right) = \sum_k h_k \phi(x - k). \quad (1.1)$$
This is known as the two-scale relation or scaling equation or dilation equation. In the Fourier domain this equation becomes

\[ \hat{\phi}(2\xi) = H(\xi) \hat{\phi}(\xi) \]  

(1.2)

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

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

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

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

\[ G(k, 0) = \int \hat{\phi}(\xi) \hat{\phi}(\xi) e^{-2\pi i k\xi} d\xi = \sum_l \int_l^{l+1} |\hat{\phi}(\xi)|^2 e^{-2\pi i k\xi} d\xi = \int_0^1 \sum_l |\hat{\phi}(\xi + l)|^2 e^{-2\pi i k\xi} d\xi. \]

Thus, \( G(k) = G(k, 0) \) is the \( k \)th Fourier coefficient of the overlap function \( \Phi(\xi)^2 = \sum_l |\hat{\phi}(\xi + l)|^2 \). That the shifts of \( \phi \) form a Riesz basis for \( V(\phi) \) is equivalent to \( \Phi \) being essentially bounded above and below since, by Plancherel’s theorem on \( \mathbb{T} \),

\[ \left\| \sum_k a_k \phi(\cdot - k) \right\|_2^2 = \left\| \sum_k a_k e^{-2\pi i k\xi} \hat{\phi}(\xi) \right\|_2^2 = \int_0^1 \left| \sum_k a_k e^{-2\pi i k\xi} \right|^2 |\Phi(\xi)|^2 d\xi. \]

The claim then follows from Plancherel’s theorem on \( \mathbb{T} \).

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

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

section, though, we consider properties of MRAs predicated on the assumption that \( \psi \) is an orthogonal generator with scaling filter \( H \).

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

\[
1 = \sum_l |\hat{\psi}(2\xi + l)|^2 = \sum_l \left( |H(\xi + l)|^2 |\psi(\xi + l)|^2 + |H(\xi + l + \frac{1}{2})|^2 |\psi(\xi + l + \frac{1}{2})|^2 \right)
\]

That is, orthogonality plus scaling implies

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

For purposes of wavelet construction, one associates to \( H \) a filter \( G \) such that \( H \) and \( G \) satisfy
\[
|H| + |G(\xi)|^2 \equiv 1.
\]

One of the possible choices for \( G \) is

\[
G(\xi) = -e^{-2\pi i\xi}H(\xi + 1/2).
\]

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

\[
\hat{\psi}(\xi) = \prod_{j=1}^{\infty} H\left(\frac{\xi}{2^j}\right).
\] (1.4)

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

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

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

Then, since

\[
|\log |H(\xi)|| \leq \log(1 + C|\xi|^\alpha) \leq C'|\xi|^\alpha, \quad (|\xi| \ll 1)
\]

it follows that

\[
\sum_{j=N}^{\infty} \left| \log H\left(\frac{\xi}{2^j}\right) \right| \leq C'|\xi|^\alpha \sum_{j=N}^{\infty} 2^{-j\alpha}
\]
1.1 Scaling and multiresolution analysis

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

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

\[
\varphi_J(\xi) = \prod_{j=1}^{J} H\left(\frac{\xi}{2^j}\right) \chi_{[-2^{J-1},2^{J-1}]}(\xi).
\]

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

\[
\|\varphi_J\|_2^2 = \int_{-2^J}^{2^J} \prod_{j=1}^{J} |H\left(\frac{\xi}{2^j}\right)|^2 d\xi = \int_{-2^J}^{2^J} \left|H\left(\frac{\xi}{2^j}\right)\right|^2 \prod_{j=1}^{J-1} |H\left(\frac{\xi}{2^j}\right)|^2 d\xi
\]

\[
= \int_{0}^{2^J} \left(\left|H\left(\frac{\xi}{2^j}\right)\right|^2 + \left|H\left(\frac{\xi}{2^j} - \frac{1}{2}\right)\right|^2\right) \prod_{j=1}^{J-1} |H\left(\frac{\xi}{2^j}\right)|^2 d\xi
\]

\[
= \int_{0}^{2^J} \prod_{j=1}^{J-1} |H\left(\frac{\xi}{2^j}\right)|^2 d\xi = \int_{-2^J}^{2^J} \prod_{j=1}^{J-1} |H\left(\frac{\xi}{2^j}\right)|^2 d\xi = \|\varphi_{J-1}\|_2^2
\]

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

Thus, any QMF whose coefficients \( h_k \) are well behaved gives rise to a scaling function \( \varphi \in L^2(\mathbb{R}) \). The only remaining question is whether \( \varphi \) defined through (1.3) and (1.4) must be orthogonal to its shifts. In fact, this is not always the case. The filter \( H(\xi) = (1 + e^{-6\pi i \xi})/2 \) satisfies (1.3) but gives rise through (1.4) to the stretched Haar function \( 3^{-1/2} \chi_{[0,3]}(x) \) which is not orthogonal to its integer shifts. More sophisticated examples can be found in Daubechies [99] where the pathology is described in more detail.

For a trigonometric polynomial \( H \) (cf. Proposition 1.4.1 for the more general case), A. Cohen characterized the pathology in the following simple way (cf. Daubechies, [99], p. 188). Let \( \tau : [0,1) \to [0,1) \) be given by \( \tau(\xi) = 2\xi \mod 1 \). A \( \tau \)-cycle is a collection \( \{\xi_1,\ldots,\xi_N\} \in [0,1) \) such that \( \xi_{j+1} = \tau(\xi_j) \) for \( 1 \leq j \leq N - 1 \) and \( \xi_1 = \xi_N \). The trivial \( \tau \)-cycle consists of the single point \{0\}. Cohen proved the following.

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

The criterion of the theorem will be called the \( \tau \)-cycle condition. A separate characterization of this nondegeneracy of \( \varphi \) was found by Lawton [252] (cf. [99], p. 190 and [251]).
Theorem 1.1.2. Suppose that \( H(\xi) = \sum_k h_k e^{-2\pi ik\xi} \) is a trigonometric polynomial satisfying (1.3) with \( H(0) = 1 \) and \( \varphi \) is the scaling function defined via (1.4). Then \( \varphi \) is orthogonal to its integer shifts when the eigenvalue 1/2 of the matrix \( A_{kl} = \sum_m h_m h_{l-2k+m} \) possesses a one-dimensional eigenspace.

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

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

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

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

\[
\langle g, \varphi(-l) \rangle = \int \hat{g}(\xi) \hat{\varphi}(\xi) e^{2\pi il\xi} d\xi
= \frac{1}{2} \int C\left(\frac{\xi}{2}\right) \hat{H}\left(\frac{\xi}{2}\right) |\hat{\varphi}\left(\frac{\xi}{2}\right)|^2 e^{2\pi il\xi} d\xi
= \int_0^1 C(\xi) \hat{H}(\xi) e^{4\pi il\xi} \sum_k |\hat{\varphi}(\xi + k)|^2 d\xi
= \int_0^{1/2} \left( C(\xi) \hat{H}(\xi) + C(\xi + \frac{1}{2}) \hat{H}\left(\xi + \frac{1}{2}\right) \right) e^{4\pi il\xi} d\xi = 0,
\]

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

\[
\begin{bmatrix}
C(\xi) \\
C(\xi + \frac{1}{2})
\end{bmatrix}
= M(\xi)
\begin{bmatrix}
\hat{H}(\xi + \frac{1}{2}) \\
-\hat{H}(\xi)
\end{bmatrix}.
\]

Moreover, (1.3) implies \( |C(\xi)|^2 + |C(\xi + 1/2)|^2 = |M(\xi)|^2 \) so \( M \in L^2(\mathbb{T}) \).
Replacing \( \xi \) by \( \xi + 1/2 \) in (1.6), one also has \( M(\xi + 1/2) = -M(\xi) \) a.e. so that
between signals in itself. The space specified by the index spaces are achieved by orthogonal projections $V$mal basis for efficient, hierarchical decomposition of sequences. Contained in the wavelet space as providing a scale of details or resolutions of images and the orthogonal $V_j$ say Discrete wavelet transform. The spaces are achieved by Esteban and Galand [132] to address a concrete application in speech coding and related algorithms and applications. Well before Mallat’s discovery of MRA, quadrature mirror filters were utilized by Esteban and Galand [132] to address a concrete application in speech processing. The idea is simple. Start with a sequence $s(k)$ thought of as integer sample values of a continuous-time signal and form the Fourier series $S(\xi) = \sum_k s(k)e^{-2\pi ik\xi}$. The idea of subband coding is essentially to replace $S$ by the subband elements $SH$ and $SG$ which can be processed separately. Reconstruction is achieved by conjugate filtering $SH \mapsto SHH$, $SG \mapsto SGG$ and adding the components. The QMF condition (1.3) shows that this process recovers $S$, i.e., $S = SHH + SGG$. Strang and Nguyen’s book [332] is one of numerous excellent sources containing detailed descriptions of subband coding and related algorithms and applications.

1.1 Subband coding and FWT

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

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

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

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

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

1.1.2 Subband coding and FWT

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

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

Given \( f \in V_j \), let \( c^j_k = c^j_k(f) = \langle f, \varphi_{jk} \rangle \) and \( d^j_k = \langle f, \psi_{jk} \rangle \). Since \( \varphi \) satisfies the dilation equation (1.1), it follows that

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

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

\[ (\mathcal{H}a)_k = \sqrt{2} \sum_l h_{1-2k} a_l; \quad (\mathcal{G}a)_k = \sqrt{2} \sum_l g_{2k} a_l. \quad (1.7) \]

Notice that \( \mathcal{H} \) acts by convolution against the sequence \( h^*_l = \hat{h}_{-1} \) followed by decimation/downsampling, and similarly for \( \mathcal{G} \). Then

\[ c^j_k = \langle f, \varphi_{jk} \rangle = \sqrt{2} \sum_l h_{1-2k} c^{j+1}_l, \]

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

\[ \begin{array}{cccccc}
  c^j & \xrightarrow{\mathcal{H}} & c^{j-1} & \xrightarrow{\mathcal{H}} & c^{j-2} & \cdots \\
  d^j & \xrightarrow{\mathcal{G}} & d^{j-1} & \xrightarrow{\mathcal{G}} & d^{j-2} & \cdots \\
  & & & & & d^L
\end{array} \]

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

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

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

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

\[ \tilde{a}_l = \begin{cases} 
  0, & \text{if } l \text{ is odd}, \\
  a_{l/2}, & \text{if } l \text{ is even}, 
\end{cases} \]
and similarly for $G^*$. Now $c^{j-1}, d^{j-1}$ determine $c^j$ via
\[
c_k^j = \langle f, \varphi_{jk} \rangle = \langle P_{j-1} f + Q_{j-1} f, \varphi_{jk} \rangle = \sum_l c_l^{j-1} \langle \varphi_{j-1,l}, \varphi_{jk} \rangle + \sum_l d_l^{j-1} \langle \psi_{j-1,l}, \varphi_{jk} \rangle = \sqrt{2} \sum_l h_{k-2l} c_l^{j-1} + \sqrt{2} \sum_l g_{k-2l} d_l^{j-1} = (\mathcal{H}^* c^{j-1})_k + (G^* d^{j-1})_k,
\]
i.e., $c^j = \mathcal{H}^* c^{j-1} + G^* d^{j-1}$. Note that $\mathcal{H}\mathcal{H}^* = GG^* = I$ on $\ell^2(\mathbb{Z})$ and $\mathcal{H}^*\mathcal{H} = \hat{G}\hat{G} = I$ on $\ell^2(\mathbb{Z})$, an equation equivalent to (1.3). Repeating this process on the sequences $c^{j-2}$ and $d^{j-2}$ gives $c^{j-1} = \mathcal{H}^* c^{j-2} + G^* d^{j-2}$ and continuing we find
\[
c_j = (\mathcal{H}^*)^{j-L} c_L + \sum_{m=0}^{j-L-1} (\mathcal{H}^*)^m G^* d^{j-m-1}.
\]
This formula, known as the discrete inverse wavelet transform, is represented in the following diagram.

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

Suppose $(V_j, \varphi)$ is an MRA of $L^2(\mathbb{R})$ with orthogonal generator $\varphi \in L^1 \cap L^\infty$ and $\psi$ is the associated wavelet. We define the periodizations $\varphi_{jk}^{per}$ and $\psi_{jk}^{per}$ of $\varphi_{jk}$ and $\psi_{jk}$, respectively, by
\[
\varphi_{jk}^{per}(x) = \sum_l \varphi_{jk}(x + l); \quad \psi_{jk}^{per}(x) = \sum_l \psi_{jk}(x + l)
\]
and the periodizations $V_j^{per}$ and $W_j^{per}$ of $V_j$ and $W_j$, respectively, as the closed subspaces of $L^2(\mathbb{T})$ given by
\[
V_j^{per} = \text{span} \{ \varphi_{jk}^{per} \}; \quad W_j^{per} = \text{span} \{ \psi_{jk}^{per} \}.
\]
Since $H(1/2) = 0$ and $H(0) = 1$, we have $\sum_k h_{2k} = \sum_k h_{2k+1} = 1/2$. Hence, if $F(x) = \sum_l \varphi(x + l) = \varphi_{01}^{per}(x)$, an application of the dilation equation (1.1) gives
\[ F(x) = 2 \sum_l \sum_k h_k \varphi(2x + 2l - k) = 2 \sum_k h_{k+2l} \varphi(2x - k) = F(2x). \]

However \( F \in L^1(\mathbb{T}) \), and (1.10) implies that its Fourier coefficients satisfy \( \hat{F}(m) = \hat{F}(2^j m) \) for non-negative integers \( j \), thus contradicting the Riemann–Lebesgue lemma unless \( F \) is constant. Since \( 1 = \hat{\varphi}(0) = \int_\mathbb{R} \varphi(x) \, dx = \int_0^1 F(x) \, dx \), this constant must be 1, i.e., \( \sum_l \varphi(x + l) \equiv 1 \). As a consequence, for \( j \leq 0 \), the spaces \( V_j^{\text{per}} \) are one-dimensional spaces containing only the constant functions. Similarly, \[
\sum_l \psi(x + l/2) = 2 \sum_l \sum_k (-1)^k \tilde{h}_{1-k} \varphi(2x + l - k) = 2 \sum_k (-1)^k \tilde{h}_{1-k} = 0
\]
from which we see that \( W_j^{\text{per}} = \{0\} \) for \( j \leq -1 \). We need then only concern ourselves with the spaces \( V_j^{\text{per}} \) and \( W_j^{\text{per}} \) for \( j \geq 0 \).

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

We denote \( 2^j \)-periodized versions \( h_k^{(j)} \) and \( g_k^{(j)} \) of the filter sequences \( h_k \) and \( g_k \) by \( h_k^{(j)} = \sum_l h_{k+2jl} \) and similarly for \( g_k^{(j)} \). Then \( g_k^{(j)} = (-1)^k h_{1-k}^{(j)} \) where the subscripts are now taken modulo \( 2^j \). Periodized filters \( \mathcal{H}^{(j)} \) and \( \mathcal{G}^{(j)} \) acting on \( 2^j \)-periodic sequences are then defined by
\[
(\mathcal{H}^{(j)} a)_k = \sqrt{2} \sum_{l=0}^{2^j-1} h_{k-2l}^{(j)} a_l; \quad (\mathcal{G}^{(j)} a)_k = \sqrt{2} \sum_{l=0}^{2^j-1} g_{k-2l}^{(j)} a_l,
\]
and their adjoints \( (\mathcal{H}^{(j)})^* \) and \( (\mathcal{G}^{(j)})^* \) by
\[
((\mathcal{H}^{(j)})^* a)_k = \sqrt{2} \sum_{l=0}^{2^j-1} h_{k-2l}^{(j)} a_l; \quad ((\mathcal{G}^{(j)})^* a)_k = \sqrt{2} \sum_{l=0}^{2^j-1} g_{k-2l}^{(j)} a_l.
\]

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

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

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

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

\[
Tf(x) = 2 \sum_k h_k f(2x - k).
\]

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

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

\[
\left| \varphi\left(\frac{k}{2^j}\right) - 2^{j/2} \langle \varphi, \psi_{j,k2^j-j}\rangle \right| \leq \int \left| \varphi\left(\frac{k}{2^j}\right) - \varphi\left(\frac{k}{2^j} + \frac{k}{2^j}\right) \right| 2^j |\varphi(2^j y)| \, dy \leq C 2^{-j\alpha}.
\]

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

\[
\langle \varphi, \varphi_{0k} \rangle = \delta_k, \quad \langle \varphi, \psi_{jk} \rangle = 0 \quad (j \geq 0, k \in \mathbb{Z}).
\]

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

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

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

\[
\left| \varphi\left(\frac{k}{2^j}\right) - 2^{j/2} ((H^*)^j \delta_0)_{k2^j-j} \right| \leq \left| \varphi\left(\frac{k}{2^j}\right) - 2^{j/2} c^j_{k2^j-j} \right| \leq C 2^{-j\alpha}.
\]
which is the desired convergence estimate.

For visualization, one typically plots the piecewise linear interpolant \( \eta' \) of the pairs \((k/2^j, 2^{j/2}c_k^j)\). The following result and its proof appear in [99] (Proposition 6.5.2).

**Proposition 1.1.3.** If \( \varphi \) is Hölder continuous of order \( \alpha > 0 \), then there exist \( C > 0 \) and \( j_0 \in \mathbb{N} \) such that, for \( j \geq j_0 \), \( \| \varphi - \eta' \|_\infty \leq C 2^{-j\alpha} \).

### Initialization and design.

Typical examples of MRAs, scaling functions, wavelets and QMFs illustrate the tradeoffs involved in using FWT algorithms in signal processing. First we consider the so-called Shannon MRA.

The Shannon scaling function \( \varphi_S(x) = \sin \pi x / (\pi x) \) has the Fourier transform \( \hat{\varphi}_S(\xi) = \chi_{[-1/2, 1/2]}(\xi) \). The associated QMF is the one-periodic function \( H_S(\xi) = \chi_{[-1/4, 1/4]}(\xi) \) on \( \mathbb{T} \sim [-1/2, 1/2] \). The space \( V_0(\varphi_S) \) is the space of \( L^2 \)-functions \( f \) bandlimited to \([-2^{j-1}, 2^{j-1}]\), i.e., \( \hat{f}(\xi) = 0 \) for \( |\xi| > 2^{j-1} \). The statement that \( \{ \varphi_S(-k) \}_{k \in \mathbb{Z}} \) is an orthonormal basis for \( V_0 \) is the classical sampling theorem (see Chapter 3), which also states that if \( f \in V_0(\varphi_S) \) then \( f(x) = \sum_k f(k) \varphi_S(x-k) \), a consequence of the cardinality of the Shannon scaling function, i.e., \( \varphi_S(k) = \delta_{k0} \). Xia and Zhang [368] proved that there is no compactly supported continuous orthogonal scaling function with this property (see Theorem 1.2.1 for an alternative proof). The Haar scaling function \( \varphi_H = \chi_{[0,1]} \) is cardinal, compactly supported and has orthogonal shifts, but is, of course, not continuous, while the Shannon scaling function is continuous, orthogonal and cardinal, but not compactly supported.

In implementing the FWT, it is standard practice to regard the values \( s_k \) of an integer-sampled sequence as the input coefficients \( c_k \) of some \( f \in V_0 \). Strang and Nguyen refer to this as a wavelet crime (see [332], p. 232). For a fixed scaling function \( \varphi \) and \( f(x) = \sum_k c_k \varphi(x-k) \in V(\varphi) \), the samples \( s_k = f(k) \) are effectively never the same as the coefficients \( c_k \). Even in the Shannon MRA, real signals are only approximately bandlimited. If one wishes only to compute, store, transmit and reconstruct samples from wavelet coefficients, then there is nothing wrong with this practice. However, as soon as one modifies wavelet coefficients (e.g., for compression or denoising) there is difficulty in interpreting the signal errors thus encumbered. Moreover, such errors can be magnified rapidly when fed back into iterative schemes. Even in noniterative applications, though, such errors can be difficult to interpret, particularly when they are measured by some norm other than \( L^2 \).

Suppose, for example, that one wishes to approximate a sampled signal from some subset of wavelet terms of \( f(x) = \sum_k c_k \varphi(x-k) \) in such a way that the reconstruction captures important fluctuations of the original samples. This is possible if variation can be expressed in terms of the magnitude of wavelet coefficients. It is important to keep in mind that one is measuring the variational error between \( f(x) = \sum_k c_k \varphi(x-k) \) and \( f_{\text{reconst}}(x) = \sum_k c'_k \varphi(x-k) \) where \( c'_k \) are the reconstructed approximate co-
1.1 Scaling and multiresolution analysis

The sample variation of \( f_{\text{reconst}} \) is not the same as the discrete variation of \( c_k' \). This is one reason why prefiltering seems desirable.

The ideal manner of prefiltering would be to find a convolutional inverse \( d_k \) for the sequence of samples \( \varphi(k) \), i.e., \( \sum_l d_l \varphi(k - l) = \delta_k \). Then \( f(x) = \sum_k (\sum_l d_l s_k - l) \varphi(x - k) \) at least agrees with \( f \) at the integers. As we will see in Chapter 3, this problem is not always well posed, and even if it is, the filter coefficients \( d_l \) may have undesirable properties.

Strang and Nguyen [332] suggest the following prefiltering scheme that at least serves to approximate the desired coefficients \( c_k(f) \). In the case of an orthogonal generator one replaces the desired \( c_k(f) = \int f(t) \varphi(t - k) \, dt \) in which \( f(t) \) is a function having samples \( s_k \), by \( \gamma_k = \sum_l s_l \varphi(k - l) \). This discrete approximation of the desired processing coefficients is accurate on polynomials up to degree \( d \) whenever the filter \( H \) of \( \varphi \) has the following property (see [268, 330, 331], cf. Theorem 2.1.1).

**Theorem 1.1.4.** Suppose that \( \varphi \) is an orthogonal scaling function. Then the space \( P_{d-1} \) of polynomials of degree less than or equal to \( d - 1 \) is reproduced by \( \varphi \) in the sense that

\[
p(t) = \sum_k p(k) \varphi(t - k)
\]

for all \( p \in P_{d-1} \), if and only if the QMF \( H(\xi) \) associated with \( \varphi \) has a zero of order \( d \) at \( \xi = 1/2 \).

In wavelet-based signal processing algorithms, one would postprocess as well. As mentioned, in order that the wavelet \( \psi \) have certain properties desirable in signal analysis and processing, constraints must be placed on the scaling function \( \varphi \). For example, symmetry is an important consideration in signal analysis when trying to avoid artifacts caused by the edges of images, and when solving boundary value problems as will be discussed below. Daubechies (see [99], p. 252) proved the following.

**Theorem 1.1.5.** If \( \phi \) generates an MRA for \( L^2(\mathbb{R}) \) such that \( \phi \) and the associated wavelet \( \psi \) are both real and have compact support, and if \( \psi \) has an axis either of symmetry or antisymmetry, then \( \psi \) is the Haar wavelet.

Although orthogonality can have important ramifications for processing transformed data, from the point of view of straight subband coding, the perfect reconstruction property is what matters. This property is expressed naturally in terms of biorthogonal filterbanks.

### 1.1.3 Biorthogonal multiresolution analyses

The conditions for biorthogonal MRA filters can be deduced in much the same way as in the orthogonal case. One starts with a pair of scaling functions \( \phi \) and \( \psi \) satisfying \( \langle \phi, \psi(\cdot - k) \rangle = \delta_k \). In the Fourier domain this becomes
$\sum_l \hat{\phi}(\xi + l) \hat{\tilde{\phi}}(\xi + l) \equiv 1$.

Following the orthogonal case routinely yields the mirror filter condition:

$$H(\xi) \tilde{H}(\xi) + H\left(\xi + \frac{1}{2}\right) \tilde{H}\left(\xi + \frac{1}{2}\right) \equiv 1 \quad (1.11)$$

with $H$ and $\tilde{H}$ the scaling filters associated with $\phi$ and $\tilde{\phi}$, respectively. Often $H$ is called the primal filter and $\tilde{H}$ the dual filter.

B-splines furnish a natural family of examples (e.g., Cohen [79]). The B-splines are symmetric, compactly supported scaling functions whose integer shifts form Riesz bases for their respective spans. For example, the linear B-spline $\phi_1(x) = (1 - |x - 1|)_+$ satisfies the two-scale equation

$$\phi_1(x) = \frac{1}{2} \phi_1(2x) + \phi_1(2x - 1) + \frac{1}{2} \phi_1(2x - 2).$$

The inner products $\|\phi_1\|^2 = 2/3$, $\langle \phi_1, \phi_1(\cdot \pm 1) \rangle = 1/12$ and $\langle \phi_1, \phi_1(\cdot \pm k) \rangle = 0$ ($k > 1$) are determined by explicit integrations. Thus, if $f = \sum_k c_k \phi_1(x - k) \in V(\phi_1)$ with real coefficients $c_k$ then

$$\|f\|^2 = \frac{2}{3} \sum_k |c_k|^2 + \frac{1}{6} \sum_k c_k c_{k-1}$$

and the Riesz basis property of $\{\phi_1(\cdot - k)\}$ follows directly from the Cauchy–Schwarz inequality.

The standard orthogonal generator of $V(\phi_1)$ fails to have compact support. One seeks instead a dual generator $\tilde{\phi}_1$ that is biorthogonal to $\phi_1$ in the sense that $\langle \phi_1, \tilde{\phi}_1(\cdot - k) \rangle = \delta_k$. Such a biorthogonal generator is not necessarily unique, but instead might be chosen among a family of generators that allow for tradeoffs between regularity and support length.

Suppose now that $\tilde{\phi}$ is a scaling function whose shifts are biorthogonal to those of $\phi$. Then (1.11) must be satisfied by the dual scaling pair and this is the condition that one seeks to solve. For the linear spline centered at $x = 1$, the scaling filter has the form

$$H_1(\xi) = \frac{1}{4} + \frac{1}{2} e^{-2\pi i \xi} + \frac{1}{4} e^{-4\pi i \xi} = \left(\frac{1 + e^{-2\pi i \xi}}{2}\right)^2 = e^{-2\pi i \xi} \cos^2 2\pi \xi.$$  

More generally, the B-spline of order $N$ is defined as the $N+1$-fold convolution of $\chi_{[0,1)}$ with itself. It is a scaling function with symbol

$$H_N(\xi) = \left(\frac{1 + e^{-2\pi i \xi}}{2}\right)^{N+1} = e^{-(N+1)\pi i \xi} \cos^{N+1} \pi \xi.$$  

The Bezout equation
1.1 Scaling and multiresolution analysis

\[(1 - y)^L P_L(y) + y^L P_L(1 - y) = 1\]

has the solution [79,99]:

\[P_L(y) = \sum_{j=0}^{L-1} \binom{L - 1 + j}{j} y^j.\]

Upon setting \(y = \cos^2 \pi \xi\) one sees that

\[(\cos^2 \pi \xi)^L P_L(\sin^2 \pi \xi) + (\sin^2 \pi \xi)^L P_L(\cos^2 \pi \xi) \equiv 1,\]

which can be rewritten as

\[\left(\frac{1 + e^{-2\pi i \xi}}{2}\right)^{2L} P_L(\sin^2 \pi \xi) + \left(\frac{1 - e^{-2\pi i \xi}}{2}\right)^{2L} (-1)^L P_L(\cos^2 \pi \xi) \equiv e^{2\pi i L \xi}.\]

Subject to \(2L \geq N + 1\), the dual filters

\[\tilde{H}_{N,L}(\xi) = e^{-2\pi i L \xi} \left(\frac{1 + e^{-2\pi i \xi}}{2}\right)^{2L-N-1} P_L(\sin^2 \pi \xi)\]

provide all solutions to (1.11) for \(H(\xi) = H_N(\xi)\) (see [99], p. 272 for plots of corresponding scaling functions and wavelets). The minimal choice \(2L = N+1\) \((N\text{ odd})\) does not lead to a convergent scaling function when \(N = 1\). In this case \(\tilde{H}_{1,1}(\xi) = e^{-2\pi i \xi}\), the scaling filter for the (shifted) \(\delta\) distribution. However, \(\tilde{H}_{1,2}(\xi) = e^{-6\pi i \xi} \cos^2 \pi \xi(1 + 2 \sin^2 \pi \xi)\) gives rise to a square-integrable scaling function. It is worth pointing out that in the biorthogonal case the FWT follows the same pattern as in the orthogonal case. After the corresponding high-pass filters \(G(\xi) = e^{-2\pi i \xi} H(\xi + 1/2)\) and \(\tilde{G}(\xi) = e^{-2\pi i \xi} \tilde{H}(\xi + 1/2)\) are fixed, one uses the dual filters for the forward transform and the primal filters for the inverse transform (e.g., [79,99]).

1.1.4 Regularity for scaling distributions

Consequences of the fact that wavelets form unconditional bases for a large scale of function spaces will be a basic theme of Chapter 6. This fact also goes a long way towards explaining their utility as tools for solving PDEs with initial conditions in suitable function spaces. To form convergent expansions, the wavelets themselves should belong to the space in question. It becomes important to be able to analyze local and global regularity properties of wavelets. One needs to know, for example, when a scaling function belongs to a Sobolev space—having a given number of derivatives in \(L^2\)—on the one hand, and a Hölder space—a uniform pointwise condition on divided differences—on the other. Here we will review methods for analyzing pointwise (i.e., local) regularity of a scaling distribution that has compact support. Global regularity considerations based on Fourier methods will be addressed in Chapter 2.
Refinement methods are used to establish pointwise and difference estimates on scaling distributions based on the eigenspace structure of certain matrices attached to the scaling coefficients. Such methods are discussed in some detail in Chapter 7 of Daubechies [99], but first appeared in [107, 108] (see also Cabrelli et al. [63]). What follows is, to some extent, a modest elaboration of the discussion in [99].

**Refinement methods.** In this section we consider questions of existence and regularity of solutions of the refinement or scaling equation (cf. (1.1))

\[ \phi(x) = 2 \sum_{k=0}^{M} h_k \phi(2x - k). \]  

(1.12)

There are only \( M + 1 \) terms in the sum on the right-hand side of (1.12), i.e., \( H \) is an FIR filter. A solution of (1.12) is a fixed point of the refinement operator

\[ Tf(x) = 2 \sum_{k=0}^{M} h_k f(2x - k). \]  

(1.13)

associated to the scaling sequence \( \{h_k\}_{k=0}^{M} \).

Notice that if \( f \) is supported on \([0, M]\), then so is \( Tf \). Hence, if a solution \( \phi \) of (1.12) can be expressed as a limit of iterates \( Tf \) of a function supported in \([0, M]\) then \( \text{supp}(\phi) \subset [0, M] \). Daubechies and Lagarias [107,108] construct \( \phi \) as a limit of piecewise linear splines having desired values at dyadic rationals. Very roughly speaking, differences in values in passing from one iteration to the next are closely related to eigenvectors of certain transition matrices associated with \( \{h_k\} \) and sum rules on \( \{h_k\} \) enable one to avoid large eigenvalues that lead to irregularity.

The refinement operator \( T \) in (1.13) induces a refinement operator \( T \) on vector-valued functions as follows. Let \( G = (G_0, \ldots, G_{M-1}) : [0, 1) \to \mathbb{R}^M \). For \( 0 \leq x < 1/2 \) set \( (TG)_j(x) = 2 \sum_k h_k G_{2j-k}(2x) \) while, for \( 1/2 < x \leq 1 \), set \( (TG)_j(x) = 2 \sum_k h_k G_{2j-k+1}(2x - 1) \). Then \( T \) can be expressed in terms of the \( M \times M \) matrices

\[ T_0 = 2 \begin{bmatrix} h_0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ h_1 & h_0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & h_M & h_{M-1} \end{bmatrix}, \quad T_1 = 2 \begin{bmatrix} h_1 & h_0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ h_2 & h_1 & h_0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & h_M \end{bmatrix}, \]

i.e., \( TG(x) = T_0 G(2x) \) \((0 \leq x < 1/2)\); \( TG(x) = T_1 G(2x - 1) \) \((1/2 \leq x < 1)\).

The action of \( T \) on vector-valued functions may be written more economically with the aid of the operator \( \tau : [0, 1) \to [0, 1) \) which acts via \( \tau x = 2x \mod 1 \) (cf. Theorem 1.1.1). We have

\[ TG(x) = T_{\tau_1(x)} G(\tau x) \]
where \( x = \sum_{j=1}^{\infty} \xi_j(x)2^{-j} \) is the binary expansion of \( x \). The ambiguity at \( x = 1/2 \) causes no problems if \( G_k(0) = 0 = G_{k-1}(1) \).

If \( \phi \) is a continuous solution of (1.12) then \( \phi(0) = \phi(M) = 0 \). By evaluating both sides of (1.12) at integers \( k \), the vector \( \mathbf{v} = (\phi(1), \phi(2), \ldots, \phi(M-1))^T \in \mathbb{R}^{M-1} \) becomes an eigenvector of the \((M-1) \times (M-1)\) matrix \( A \) with \((j,k)\)-th entry \( A_{jk} = 2h_{2j-k} \) \((1 \leq j, k \leq M - 1)\) and eigenvalue 1, i.e., \( Av = v \). Note that \( A \) is a submatrix of both \( T_0 \) and \( T_1 \); it may be obtained by removing the first row and column of \( T_0 \) or the last row and column of \( T_1 \). We insist, as before, that the sequence \( \{h_k\}_{k=0}^M \) satisfies the sum rule

\[
\sum_k h_{2k} = \sum_k h_{2k+1} = \frac{1}{2}.
\]  

(1.14)

Thus \( \sum_{j=1}^{M-1} A_{jk} = \sum_{j=1}^{M-1} 2h_{2j-k} = 1 \) for all \( k \). Equivalently, the row vector \( \mathbf{u} = (1, 1, \ldots, 1) \in \mathbb{R}^{M-1} \) is a left-eigenvector of \( A \) with eigenvalue 1. Similarly, \( \mathbf{e}_0 = (1, 1, \ldots, 1) \in \mathbb{R}^M \) is left-fixed by both \( T_0 \) and \( T_1 \).

Suppose for the moment that the eigenvalue 1 of \( A \) is nondegenerate (i.e., has multiplicity one). Later we will give a natural condition that ensures this. Then there is a right eigenvector \( \mathbf{w} \) of \( A \) with eigenvalue 1 that one sees, upon consideration of the Jordan form of \( A \), cannot be orthogonal to \( \mathbf{u} \). Normalize \( \mathbf{w} \) so that \( \mathbf{uw} = \sum_{j=1}^{M-1} u_j w_j = 1 \). With \( \mathbf{w} = (w_1, w_2, \ldots, w_{M-1})^T \) so normalized and extended to \( \mathbb{R}^{M+1} \) by setting \( w_0 = w_M = 0 \), define the piecewise linear interpolant \( f^{(0)}(x) \) on \( \mathbb{R} \) by

\[
f^{(0)}(k) = \begin{cases} 
  w_k, & \text{if } 0 \leq k \leq M + 1, \\
  0, & \text{else.}
\end{cases}
\]

Then \( f^{(0)} \) is supported on \([0, M]\). Define a (column) vector-valued function \( F^{(0)} : [0, 1) \to \mathbb{R}^M \) by \( F^{(0)}_k(x) = f^{(0)}(x + k) \) \((0 \leq x < 1, 0 \leq k \leq M - 1)\), i.e.,

\[
F^{(0)}(x) = (f^{(0)}(x), f^{(0)}(x+1), \ldots, f^{(0)}(x+M-1))^T \quad (0 \leq x < 1).
\]

Then \( F^{(0)}_k(0) = f^{(0)}(k) = f^{(0)}(1 + (k-1)) = F^{(0)}_{k-1}(1) \) and \( F^{(0)}_0(0) = f^{(0)}(0) = 0 = f(M) = F^{(0)}_{M-1}(1) \).

Vector-valued functions \( F^{(j)} \) are now defined recursively by \( F^{(j)}(x) = T_{\varepsilon_1(x)} F^{(j-1)}(\tau x) \) \((j \geq 1)\). Each \( F^{(j)} \) may be unfolded to generate a piecewise linear function \( f^{(j)} \) on the line with \( f^{(j)}(x + k) = F^{(j)}_k(x) \) \((0 \leq x < 1)\). Since \( F^{(0)}_k(x) = xw_{k+1} + (1 - x)w_k \),

\[
\mathbf{e}_0 F^{(0)}_k(x) = \sum_k F^{(0)}_k(x) = x \sum_{k=0}^{M-1} w_{k+1} + (1 - x) \sum_{k=0}^{M-1} w_k = \sum_{k=0}^{M-1} w_k = 1
\]

and, since \( \mathbf{e}_0 T_{\varepsilon} = \mathbf{e}_0 \) for \( \varepsilon = 0, 1 \) and \( \varepsilon_1(\tau^j x) = \varepsilon_{j+1}(x) \),
First, as Daubechies showed in [99], the assumption (1.16) implies that so has a continuous limit. Observe that by (1.15), φ is a solution of (1.12), is supported on \(E_0 \) and is Hölder continuous of order \(\alpha = -\log_2(\lambda)\).

**Proof.** First, as Daubechies showed in [99], the assumption (1.16) implies that the eigenvalue 1 is nondegenerate. Observe that by (1.15), \(F^{(j)}(x) - F^{(k)}(x) \in E_0\) for all \(j,k\) and \(x\). Hence, with an application of (1.16),

\[
\|F^{(j+1)}(x) - F^{(j)}(x)\| = \|T_{\varepsilon_1}(x) \cdots T_{\varepsilon_j}(x) (F^{(1)}(\tau^j x) - F^{(0)}(\tau^j x))\| 
\leq C \lambda^j \|F^{(1)}(\tau^j x) - F^{(0)}(\tau^j x)\|
\]

where \(\|\cdot\|\) denotes the Euclidean norm on \(\mathbb{R}^M\). Therefore,

\[
\|F^{(j)}(x)\| \leq \sum_{k=0}^{j-1} \|F^{(k+1)}(x) - F^{(k)}(x)\| + \|F^{(0)}(x)\| 
\leq \sum_{k=0}^{j-1} C \lambda^k \|F^{(1)}(x) - F^{(0)}(x)\| + \|F^{(0)}(x)\| 
\leq \frac{C}{1 - \lambda} \sup_{0 \leq y < 1} \|F^{(1)}(y) - F^{(0)}(y)\| + \sup_{0 \leq y < 1} \|F^{(0)}(y)\|
\]

so that the norms \(\|F^{(j)}(x)\|\) are bounded independent of \(x, j\). Furthermore,

\[
\|F^{(j+k)}(x) - F^{(k)}(x)\| = \|T_{\varepsilon_1}(x) \cdots T_{\varepsilon_k}(x) (F^{(j)}(\tau^k x) - F^{(0)}(\tau^k x))\| 
\leq C \lambda^k \|F^{(j)}(\tau^k x) - F^{(0)}(\tau^k x)\| 
\leq C \lambda^k.
\]

Since each \(F^{(j)}(x)\) is continuous, \(\{F^{(j)}\}_{j=0}^\infty\) is Cauchy in \(L^\infty([0,1], \mathbb{R}^M)\) and so has a continuous limit \(\Phi(x) = \lim_{j \to \infty} F^{(j)}(x)\). Now unfold \(\Phi(x)\) to obtain a continuous solution \(\phi\) of (1.12) by \(\phi(x + k) = \Phi_k(x)\) (0 ≤ \(\phi < 1\)). Since \(\Phi_k(0) = \lim_{j \to \infty} F^{(j)}_k(0) = \lim_{j \to \infty} F^{(j)}_{k-1}(1) = \Phi_{k-1}(1)\), the definition of \(\phi\)
at the integers is consistent. By (1.15) and the definition of $\phi$, the solution satisfies

$$\int_0^M \phi(x) \, dx = \sum_{j=0}^{M-1} \int_0^1 \phi(x+j) \, dx = \sum_{j=0}^{M-1} \int_0^1 \Phi_j(x) \, dx$$

$$= \int_0^1 \phi(x) \, dx = \lim_{j \to \infty} \int_0^1 \Phi_j(x) \, dx = 1.$$

To obtain an estimate of the Hölder continuity of $\phi$, observe that since $\phi$ is a fixed point of $T$ in (1.13) and $\Phi(x) = F(0) \in E_0^\perp$, for all $x$,

$$\|\Phi(x) - \Phi(y)\| \leq \|\Phi(x) - F(0)\| + \|F(0) - \Phi(y)\| \leq 2C \lambda^j + \|T_{\epsilon_1(x)} \cdots T_{\epsilon_L(x)} (F(0) - \Phi(y))\| \leq 2C \lambda^j + C' \lambda \|F(0) - \Phi(y)\| \leq C' \lambda \|F(0) - \Phi(y)\|,$$

since $F(0) - \Phi(y) \in E_0^\perp$. Hence $\phi$ is Hölder continuous of order $\alpha = -\log_2(\lambda)$. This proves the theorem.

In [97] and [99] Daubechies constructs continuous scaling functions $\varphi^N$ ($N \geq 1$) supported on $[0,2N-1]$. To see how Theorem 1.1.6 works in a particular example, let $\mu = (1 + \sqrt{3})/2$ and $\phi = \varphi$ be the lowest-order Daubechies scaling function supported on $[0,3]$ with scaling coefficients $(h_0, h_1, h_2, h_3) = (\mu, 1 + \mu, 2 - \mu, 1 - \mu)/4$. Then $T_0$, $T_1$ are given by

$$T_0 = \frac{1}{2} \begin{bmatrix} \mu & 0 & 0 \\ 2 - \mu & 1 + \mu & \mu \\ 0 & 1 - \mu & 2 - \mu \end{bmatrix}, \quad T_1 = \frac{1}{2} \begin{bmatrix} 1 + \mu & \mu & 0 \\ 1 - \mu & 2 - \mu & 1 + \mu \\ 0 & 0 & 1 - \mu \end{bmatrix}$$

and have simple eigenvalues $1, 1/2, (1 + \sqrt{3})/4$ and $1, 1/2, (1 - \sqrt{3})/4$, respectively. It is a simple matter to show that $\|T_0|_{E_0^{\perp}} \approx 0.7954 < 1$; however $\|T_1|_{E_0^{\perp}} \approx 1.2646 > 1$. Thus, the simple estimate $\|T_{\epsilon_1} \cdots T_{\epsilon_L}|_{E_0^{\perp}} \leq \prod_{j=1}^{L} \|T_{\epsilon_j}|_{E_0^{\perp}}$ is insufficient to obtain (1.16). Nevertheless, Daubechies [99] gives an estimate that does handle this example.

**Theorem 1.1.7.** Let $\lambda_m = \max_{\epsilon_j=0} \| \prod_{j=1}^{m} T_{\epsilon_j}|_{E_0^{\perp}}^{1/m} (m \geq 0)$. Then a necessary and sufficient condition for (1.16) is that $\lambda_m < 1$ for some positive integer $m$. 

When \( \phi = 2\varphi, \lambda_1 = \|T_0\|_{E_0^r} \approx 1.2646 > 1 \) and \( \lambda_2 = \|T_0T_1\|_{E_0^r} \|1/2 \approx 1.0028 > 1 \), but \( \lambda_3 = \|T_0^2T_1\|_{E_0^r} \|1/3 \approx 0.9104 \). Consequently, \( 2\varphi \) is Hölder continuous of order \( \alpha = \log_2((0.9104)^{-1}) = 0.1354 \). Higher values of \( m \) can produce lower values of \( \lambda_m \) and better estimates of the Hölder exponent of continuity.

For the optimal Hölder exponent, one needs to take account of extra sum rules satisfied by QMFs with higher-order zeros at \( \xi = 1/2 \). We assume in what follows that \( H \) has a zero of order two at \( \xi = 1/2 \), i.e., apart from the sum rule (1.14), the sequence \( \{h_k\}_{k=0}^M \) also satisfies

\[
\sum_k (-1)^k k h_k = 0. \tag{1.18}
\]

Set \( \sigma_2 = \sum_k 2k h_{2k} \). If \( u = (1, 2, \ldots, M) \in \mathbb{R}^M \) then, by sum rules (1.14) and (1.18), \( uT_0 = u/2 + \sigma_2 e_0 \). Similarly, \( uT_1 = u/2 + (\sigma_2 - 1/2)e_0 \). Putting \( e_0^0 = u - 2\sigma_2 e_0 \) gives

\[
e_0^0 T_0 = u T_0 - 2 \sigma_2 e_0 T_0 = \frac{1}{2} u - \sigma_2 e_0 = \frac{1}{2} e_1^0.
\]

Further,

\[
e_1^0 T_1 = u T_1 - 2 \sigma_2 e_0 T_1 = \frac{1}{2} (e_1^0 - e_0),
\]

i.e., \( e_1^0 T_2 = (e_1^0 - \varepsilon e_0)/2 \) with \( \varepsilon = 0, 1 \). Hence, if \( e_1^1 = e_1^0 + e_0 \), then

\[
e_1^1 T_1 = e_1^0 T_1 + e_0 T_1 = \frac{1}{2} (e_1^0 - e_0) + e_0 = \frac{1}{2} (e_1^0 + e_0) = \frac{1}{2} e_1^1.
\]

Thus, we have constructed left eigenvectors \( e_1^j \) for \( T_2 \) (\( \varepsilon = 0, 1 \)) with eigenvalues \( 1/2^k \) (\( k = 0, 1 \)). With \( E_1 = \text{span} \{e_0, e_1^0\} = \text{span} \{e_0, e_1^j\} \), Daubechies [99] showed that if

\[
\max_{\varepsilon_j = 0 \text{ or } 1, j = 1, \ldots, m} \left\| T_{\varepsilon_1}\cdot T_{\varepsilon_2}\cdot \cdots T_{\varepsilon_m} \right\|_{E_1^r} \leq C \lambda^m \tag{1.19}
\]

for some \( C > 0 \) and \( \lambda < 1 \) and the sum rules (1.14), (1.18) hold, then the left and right eigenspaces of \( T_0 \) and \( T_1 \) with eigenvalues 1 and 1/2 are nondegenerate. For \( j, k = 0, 1 \), let \( v_j^k \) be the right eigenvector of \( T_k \) with eigenvalue \( 2^{-k} \). By considering the Jordan forms of \( T_0 \) and \( T_1 \), it is clear that \( e_1^j \) and \( v_j^k \) cannot be orthogonal \( (e_1^k v_j^k \neq 0) \), and one may normalize the right eigenvectors so that

\[
e_1^j v_k^{j'} = \delta_{k, k'} \quad (\varepsilon, k, k' = 0, 1). \tag{1.20}
\]

The remaining inner products \( e_1^0 v_0^j \) may be deduced from (1.20), the eigenproperties of these vectors and the relationship \( e_1^1 = e_1^0 + e_0 \), and may be summarized by

\[
e_1^1 v_0^0 = 1, \quad e_1^0 v_1^k = (-1)^{k+1}, \quad e_0 v_1^1 = 0 \quad (\varepsilon, k = 0, 1). \tag{1.21}
\]
Consider the operators \( P_\varepsilon : \mathbb{R}^M \to \mathbb{R}^M \) given by \( P_\varepsilon u = (e_0 u) v_0^\varepsilon + (e_1^\varepsilon u) v_1^\varepsilon \) \((\varepsilon = 0, 1)\) and let \( P_\varepsilon^* u = (I - P_\varepsilon) u = u - P_\varepsilon u \). By (1.20) and (1.21) one has \( e_0 P_\varepsilon^* u = e_1^\varepsilon P_\varepsilon^* u = 0 \) so that \( \text{Ran} (P_\varepsilon^*) \subset E_1^\perp \).

Returning now to the question of the continuity of solutions of refinement equations, observe that

\[
e_1^0 F^{(0)}(x) = (1 - x) e_1^0 F^{(0)}(0) + x e_1^0 F^{(0)}(1)
= (1 - x) \left( \sum_{l=1}^{M} l w_l - 2 \sigma_2 \sum_{l=1}^{M} w_{l-1} \right)
+ x \left( \sum_{l=1}^{M} l w_{l-1} - 2 \sigma_2 \sum_{l=1}^{M} w_l \right) = x
\]

since \( \sum_{l} l w_l = 2 \sigma_2 \) and \( \sum_{l} w_l = 1 \). Similarly, \( e_1^0 F^{(0)}(x) = x + 1 \), and therefore

\[
P_\varepsilon F^{(0)}(x) = (e_0 F^{(0)}(x)) v_0^\varepsilon + (e_1^\varepsilon F^{(0)}(x)) v_1^\varepsilon = v_0^\varepsilon + (x + \varepsilon) v_1^\varepsilon \quad (\varepsilon = 0, 1).
\]

Hence, for all \( x, y \),

\[
P_\varepsilon (F^{(0)}(x) - F^{(0)}(y)) = (x - y) v_1^\varepsilon.
\] (1.23)

Suppose now that \( 2^{-(j+1)} \leq |x - y| \leq 2^{-j} \). As in the proof of Theorem 1.1.6, one may assume that \( x \) and \( y \) have the same binary expansion up to the first \( j \) terms. Then because of (1.23), assumption (1.19), the uniform bound on \( F^{(0)} \) and the fact that \( P_\varepsilon \) maps \( \mathbb{R}^M \) into \( E_1^\perp \),

\[
\| F^{(j)}(x) - F^{(j)}(y) \| = \| T_{\varepsilon_1(x)} \cdots T_{\varepsilon_j(x)} (F^{(0)}(\tau^j x) - F^{(0)}(\tau^j y)) \|
\leq \| T_{\varepsilon_1(x)} \cdots T_{\varepsilon_j(x)} P_{\varepsilon_j(x)} (F^{(0)}(\tau^j x) - F^{(0)}(\tau^j y)) \|
+ \| T_{\varepsilon_1(x)} \cdots T_{\varepsilon_j(x)} P'_{\varepsilon_j(x)} (F^{(0)}(\tau^j x) - F^{(0)}(\tau^j y)) \|
\leq \| T_{\varepsilon_1(x)} \cdots T_{\varepsilon_j(x)} (\tau^j x - \tau^j y) v_1^\varepsilon(x) \| + C \lambda^j
= \frac{1}{2} |\tau^j x - \tau^j y| \| T_{\varepsilon_1(x)} \cdots T_{\varepsilon_{j-1}(x)} v_1^\varepsilon(x) \| + C \lambda^j
\]

where we have used the fact that \( T_{\varepsilon_j(x)} v_1^\varepsilon(x) = v_1^\varepsilon(x) / 2 \). However, \( v_1^\varepsilon(x) = P_{\varepsilon_{j-1}(x)} v_1^\varepsilon(x) + P'_{\varepsilon_{j-1}(x)} v_1^\varepsilon(x) = v_1^{\varepsilon_{j-1}(x)} + P'_{\varepsilon_{j-1}(x)} v_1^\varepsilon(x) \), so

\[
\| T_{\varepsilon_1(x)} \cdots T_{\varepsilon_{j-1}(x)} v_1^\varepsilon(x) \| \leq \| T_{\varepsilon_1(x)} \cdots T_{\varepsilon_{j-1}(x)} P_{\varepsilon_{j-1}(x)} v_1^{\varepsilon_{j-1}(x)} \|
+ \| T_{\varepsilon_1(x)} \cdots T_{\varepsilon_{j-1}(x)} P'_{\varepsilon_{j-1}(x)} v_1^{\varepsilon_{j-1}(x)} \|
\leq \| T_{\varepsilon_1(x)} \cdots T_{\varepsilon_{j-1}(x)} v_1^{\varepsilon_{j-1}(x)} \| + C \lambda^{j-1}
\leq \frac{1}{2} \| T_{\varepsilon_1(x)} \cdots T_{\varepsilon_{j-1}(x)} v_1^{\varepsilon_{j-1}(x)} \| + C \lambda^{j-1}.
\]

Combining these estimates yields
\[
\|F^{(j)}(x) - F^{(j)}(y)\| \leq \frac{1}{4} |\tau^j x - \tau^j y| \|T_{\varepsilon_1(x)} \cdots T_{\varepsilon_3(x)} \varphi^{(j-1)}(x)\|
\]
\[
+ \frac{C}{2} |\tau^j x - \tau^j y| \lambda^{j-1} + C \lambda^j.
\]
Continuing in this manner, one finds that
\[
\|F^{(j)}(x) - F^{(j)}(y)\| \leq C |\tau^j x - \tau^j y| \sum_{k=1}^{j} \frac{\lambda^{j-k}}{2^k} + C \lambda^j. \tag{1.24}
\]
If \( \lambda < 1/2 \), estimating the sum in (1.24) gives
\[
\|F^{(j)}(x) - F^{(j)}(y)\| \leq C \frac{\lambda^j |x-y|}{2^j} + C \lambda^j \leq C |x-y| + C 2^{-j} \leq C |x-y|
\]
so that each \( F^{(j)} \) is Lipschitz continuous. If \( \lambda = 1/2 \), one finds
\[
\|F^{(j)}(x) - F^{(j)}(y)\| \leq C j |x-y| + C \lambda^j \leq C |x-y| \log(|x-y|^{-1}).
\]
Finally, for \( 1/2 < \lambda < 1 \),
\[
\|F^{(j)}(x) - F^{(j)}(y)\| \leq C |\tau^j x - \tau^j y| \lambda^{j} \sum_{k=1}^{j} (2\lambda)^{-k} + C \lambda^j \leq C \lambda^j \leq C |x-y|^{\alpha}
\]
where \( \alpha = -\log_2(\lambda) \), i.e., each \( F^{(j)} \) is Hölder continuous of order \( \alpha = -\log_2(\lambda) \). As in the proof of Theorem 1.1.6, this implies the continuity of \( \Phi \) and the scaling function \( \phi \). In summary, if \( 0 < |x-y| < 1 \),
\[
|\phi(x) - \phi(y)| \leq C \left\{ \begin{array}{ll}
|x-y|, & \text{if } \lambda < 1/2, \\
|x-y| \log(|x-y|^{-1}), & \text{if } \lambda = 1/2, \\
|x-y|^{-\log_2(\lambda)}, & \text{if } 1/2 < \lambda < 1.
\end{array} \right.
\]
In the case of the Daubechies scaling function \( \phi = \varphi_2 \), the eigenvalues 1, 1/2 and \((1 + \sqrt{3})/4\) of \( T_0 \) and 1, 1/2 and \((1 - \sqrt{3})/4\) are all nondegenerate. The space \( E^+ \) is one-dimensional so that \( T_0 |_{E^+} \) and \( T_1 |_{E^+} \) commute and act as multiplication by \((1 + \sqrt{3})/4\) and \((1 - \sqrt{3})/4\), respectively. Therefore the restriction \( T_{\varepsilon_1} T_{\varepsilon_2} \cdots T_{\varepsilon_j} |_{E^+} \) acts as multiplication by \((1 + \sqrt{3})/4)^{-s_j}((1 - \sqrt{3})/4)^{s_j} \) where \( s_j = \sum_{l=1}^{j} \varepsilon_l \). This observation allows one to compute local Hölder estimates. For the worst case one obtains
\[
\|T_{\varepsilon_1} T_{\varepsilon_2} \cdots T_{\varepsilon_j} |_{E^+}\| \leq \left( \frac{1 + \sqrt{3}}{4} \right)^j
\]
which gives Hölder continuity of order \( -\log_2((1 + \sqrt{3})/4) \approx 0.5500 \). This is the best global estimate and is attained at \( x = 0 \).
Statements and proofs of the general versions of these results (for QMFs with higher-order zeros at \( \xi = 1/2 \)—equivalently with extra sum rules) as well as local estimates of Hölder continuity, differentiability and calculations relating to scaling functions with longer support, can be found in [99].

As has already been hinted, convergence properties of refinement schemes can be recast in terms of joint spectral radius properties (cf. [63]). If \( \varepsilon_j(x) = \varepsilon_j(y) \) for all \( j = 1, \ldots, L \) then

\[
\Phi(x) - \Phi(y) = T_{\varepsilon_1} \cdots T_{\varepsilon_L} (\Phi(x^L x) - \Phi(x^L y)).
\]

In particular, if \( x_L = 0.\varepsilon_1(x) \ldots \varepsilon_L(x) \) in binary notation then \( x - x_L \) tends to zero as \( L \to \infty \) so, assuming that \( \phi \) is continuous and supported in \([0, M] \), it follows that \( \phi(x + k) - \phi(xL + k) \) tends to zero as \( L \to \infty \) for each \( k = 0, \ldots, M - 1 \). That is, \( \Phi(x) - \Phi(x_L) \) tends to the zero vector. On the other hand, the points \( \tau^L x \) may be arbitrarily distributed in \([0, 1] \), so it must be that the product \( T_{\varepsilon_1} \cdots T_{\varepsilon_L} \) applied to any vector of the form \( \Phi(x) - \Phi(y) \) tends to zero. The subspace

\[
W = \text{span} \{ \Phi(x) - \Phi(y) : x, y \in [0, 1] \} \subset \mathbb{C}^M
\]

is invariant under \( T_0 \) and \( T_1 \). Consequently \( T_{\varepsilon_1} \cdots T_{\varepsilon_L} |_W \to 0 \) as \( L \to \infty \). Define the uniform joint spectral radius (JSR) \( \tilde{\rho}(A_0, A_1) \) of matrices \( A_0, A_1 \) to be

\[
\tilde{\rho}(A_0, A_1) = \limsup_{m \to \infty} \max_{\varepsilon_j = 0, 1} \| A_{\varepsilon_1} A_{\varepsilon_2} \ldots A_{\varepsilon_m} \|^{1/m}.
\]

The JSR for a subspace \( W \) is defined similarly, in terms of the norms of the restrictions of \( A_1 \) to \( W \). It is known that arbitrary products of the form \( A_{\varepsilon_1} \cdots A_{\varepsilon_L} \) tend to zero if and only if the JSR of \( A_0, A_1 \) is strictly less than one. Consequently, a necessary condition for the continuity of \( \phi \) is that \( \tilde{\rho}(T_0|_W, T_1|_W) < 1 \).

The space \( W \) turns out to have a characterization as the smallest common invariant subspace of the \( T_\varepsilon \) that also contains the vector \( \Phi(1) - \Phi(0) \) (see [182], Proposition 4.2). Since \( T_\varepsilon \) fixes \( \Phi(\varepsilon) (\varepsilon = 0, 1) \), it is possible to determine these vectors directly from the scaling coefficients \( b_k \). Using this approach one can show that \( \tilde{\rho}(T_0|_W, T_1|_W) < 1 \) is also sufficient for the continuity of \( \phi \) [182].

In general, the JSR can be difficult to estimate precisely, but matters simplify in important cases including the present refinement setting. Specifically, \( W = \{ w \in \mathbb{C}^N : v_1 + v_2 + \cdots + v_N = 0 \} \) precisely when the integer shifts \( \phi(\cdot - k) \) are linearly independent [63].

### 1.2 A construction of quadrature mirror filters

In Section 1.1.4 we saw that, in the case of a finite length scaling filter \( H \), the integer values of its scaling function \( \phi \) arise as an eigenvector of the matrix \( A_{kl} = 2h_{2k-l} \). Here we propose to work in reverse, parameterizing QMF
filters in terms of the possible integer values of the QMF scaling function \( \varphi \).

The present construction offers a different perspective and certainly a different set of techniques from the standard constructions of Daubechies [99] or Mallat [267] for designing QMFs and is motivated by applications to sampling and extrapolation of signals in wavelet subspaces. Sampling in wavelet subspaces will be considered in further detail in Chapter 3 (cf. [201]). As QMF coefficients are not arbitrary, constraints must be imposed on \( \{ \varphi(k) \} \) in order to ensure the existence of an orthogonal scaling function with these integer samples. Understanding these constraints will allow one to design QMFs suitable for signal extrapolation and for understanding when signals in \( V(\varphi) \) can be represented efficiently in terms of their samples. The construction is based on the Zak transform.

### 1.2.1 The Zak transform

Here we define the Zak transform and consider some of its properties of immediate relevance. Other properties will be considered later in this book as needs arise. More details and insights, including some history and applications to signal processing, can be found in Janssen’s tutorial [211]. The Zak transform has a natural place in Gabor analysis and accounts of this role are found in Daubechies’ book [99] and the tutorial by Heil and Walnut [183]. Here we develop its role in the multiresolution context.

Given \( f \) in the Schwartz space \( \mathcal{S}(\mathbb{R}) \), the function

\[
Zf(x, \xi) = \sum_{k} f(x + k) e^{2\pi ik \xi} \quad ((x, \xi) \in \mathbb{R} \times \mathbb{R})
\]

(1.25)

is called the Zak transform of \( f \). For such an \( f \), the Poisson summation formula can be expressed as \( Zf(x, \xi) = e^{-2\pi ik \xi} \hat{Zf}(-\xi, x) \). The Zak transform is quasi-periodic in the sense that

\[
Zf(x + k, \xi + l) = e^{-2\pi ik \xi} Zf(x, \xi) \quad (l, k \in \mathbb{Z}).
\]

Consequently, the values of \( Zf \) on the square \( Q = [0,1) \times [0,1) \) determine the values of \( f \) on the whole time–frequency plane and one thinks of \( Q \) as the domain of \( Zf \). A nontrivial consequence of quasi-periodicity—one that has well-established ramifications in Gabor theory and sampling and will play an important role in Chapter 3 as well—is that if \( Zf(x, \xi) \) is continuous on \( Q \) then \( Zf \) has a zero in \( Q \) (see [183]).

The Zak transform is unitary: if \( f, g \in \mathcal{S}(\mathbb{R}) \), then

\[
\int f(x) \bar{g}(x) dx = \int_{0}^{1} \int_{0}^{1} Zf(x, \xi) \overline{Zg(x, \xi)} dx d\xi.
\]

Hence \( Z \) extends to a unitary mapping from \( L^{2}(\mathbb{R}) \) to the space
and wavelet equations satisfied by 
be a scaling function for an MRA of 
wavelet generated by a QMF. Under suitable a priori conditions on the integer 
the conjugate filter is 

\[ Z\phi \]

and wavelet

\[ Zf \]

notation, write

\[ 1.2.2 \text{ Scaling functions in the Zak domain} \]

is to the unit circle, one recovers the previous definition of the Zak transform,

\[ z \]

support, for fixed

the sequence of samples

the fundamental scaling and wavelet equations can be expressed in matrix form:

\[ F : \mathbb{R}^2 \to \mathbb{C} : F(x + 1, \xi + 1) = e^{-2\pi i \xi} F(x, \xi), \]

and

\[ \int_0^1 \int_0^1 |F(x, \xi)|^2 dx d\xi < \infty \]  \hspace{1cm} (1.26)

The inversion formula is particularly simple: \( \int_0^1 Zf(x, \xi) d\xi = f(x) \) whenever the integral converges. It will certainly do so for \( f \in \mathcal{S}(\mathbb{R}) \); the formula extends to \( f \in L^2(\mathbb{R}) \) by a limiting argument.

It will be important later (in the context of sampling) and convenient now to extend the definition of the Zak transform from \( \mathbb{R} \times \mathbb{R} \) to \( \mathbb{R} \times \mathbb{C} \). To this end we define the complexified Zak transform \( Z_Cf(x, z) \) for \( x \in \mathbb{R}, z \in \mathbb{C} \) as the Laurent series

\[ Z_Cf(x, z) = \sum_k f(x + k) z^k \]

whenever the sum converges. For fixed \( x \), \( Z_Cf(x, z) \) is the \( z \)-transform of the sequence of samples \( \{f(x + k)\}_{k \in \mathbb{Z}} \). Notice also that if \( f \) has compact support, for fixed \( x \) the Zak transform \( Z_Cf(x, z) \) is a Laurent polynomial in \( z \). Whenever it is clear from context, we omit the subscript \( \mathbb{C} \) and, abusing notation, write \( Zf(x, z) \) for \( Z_Cf(x, z) \). As such, when \( z = e^{2\pi i \xi} \) is restricted to the unit circle, one recovers the previous definition of the Zak transform, i.e., \( Zf(x, z) = Z_Cf(x, e^{2\pi i \xi}) = Zf(x, \xi) \).

\[ 1.2.2 \text{ Scaling functions in the Zak domain} \]

Because of the fundamental role that the \( z \)-transforms of the integer values of scaling function \( \varphi \) and wavelet \( \psi \) will play, we denote in this section \( \Phi(z) = Z\varphi(0, z) \) and \( \Psi(z) = Z\psi(0, z) \) where \( \varphi \) and \( \psi \) are the scaling function and wavelet generated by a QMF. Under suitable a priori conditions on the integer samples of \( \varphi \), its QMF can in fact be defined in terms of these samples. Let \( \varphi \) be a scaling function for an MRA of \( L^2(\mathbb{R}) \). In the Zak domain, the dilation and wavelet equations satisfied by \( \varphi \) and its wavelet \( \psi \) may be written

\[ Z_C\varphi(x, z^2) = H(z) Z_C\varphi(2x, z) + H(-z) Z_C\varphi(2x, -z), \]

\[ Z_C\psi(x, z^2) = G(z) Z_C\varphi(2x, z) + G(-z) Z_C\varphi(2x, -z), \]

where \( H \) is the QMF associated to \( \varphi \) and \( G \) is a conjugate filter (see [212]). We make frequent use of the involution \( z^* = 1/z \). Given a function \( F \) defined on a region of the complex plane, we also define \( F^*(z) = \bar{F}(z^*) \). This should not be confused with the conjugate transpose for matrices that will use the same \( ^* \)-notation. When \( F(z) = \sum_k c_k z^k \), \( F^*(z) = \sum_k \bar{c_k} z^{-k} \). A common choice for the conjugate filter is \( G(z) = -z^{2Q+1} H^*(-z) \) for some integer \( Q \) whose role is to center \( \psi \) as will be clear in a few examples below. With this choice, the fundamental scaling and wavelet equations can be expressed in matrix form:

\[ \begin{bmatrix} Z_C\varphi(x, z^2) \\ Z_C\psi(x, z^2) \end{bmatrix} = M(z) \begin{bmatrix} Z_C\varphi(2x, z) \\ Z_C\varphi(2x, -z) \end{bmatrix} \]  \hspace{1cm} (1.27)
with $M(z)$ the $2 \times 2$ matrix

$$M(z) = \begin{bmatrix} H(z) & H(-z) \\ -z^{2Q+1} H^*(z) & z^{2Q+1} H^*(-z) \end{bmatrix}. \tag{1.28}$$

Let $u, v, w : \mathbb{C} \to \mathbb{C}^2$ be defined by

$$u(z) = \begin{bmatrix} H(z) \\ H(-z) \end{bmatrix}, \quad v(z) = \begin{bmatrix} \Phi(z) \\ \Phi(-z) \end{bmatrix}, \quad w(z) = \begin{bmatrix} \Phi(z^2) \\ \Psi(z^2) \end{bmatrix}. \tag{1.29}$$

Putting $x = 0$ in (1.27) then gives

$$w(z) = M(z) v(z). \tag{1.29}$$

$M(z)$ is not necessarily unitary off the unit circle, but it is invertible since the QMF condition on $H$ extends to $\mathbb{C} \setminus \{0\}$ as

$$H(z) H^*(z) + H(-z) H^*(-z) = 1. \tag{1.30}$$

Consequently, det $M(z) = z^{2Q+1}$ and

$$M(z) \overline{M}(z) = I_2 \tag{1.31}$$

where $\overline{M}(z) = (M(z^*))^*$ is the paraconjugate of $M(z)$, i.e., the conjugate transpose of $M(z^*)$. Matrix-valued functions $M(z)$ satisfying (1.31) are said to be paraunitary in that they are invertible extensions of matrices that are unitary on the unit circle. With $\langle \cdot, \cdot \rangle$ now representing the usual complex inner product on $\mathbb{C}^2$, and with $c(z) = \Phi(z^2)\Phi^*(z^2) + \Phi(z^2)\Psi^*(z^2)$, (1.29) gives

$$c(z) = \langle w(z), w(z^*) \rangle = \langle M(z) v(z), M(z^*) v(z^*) \rangle = \langle v(z), M(z^*)^* M(z^*) v(z^*) \rangle = \langle v(z), v(z^*) \rangle = \Phi(z) \Phi^*(z) + \Phi(-z) \Phi^*(-z) \tag{1.32}$$

by the paraunitarity of $M(z)$. Notice also that equation (1.29) may be rewritten in the form

$$w(z) = Q(z) u(z) \tag{1.33}$$

where

$$Q(z) = \begin{bmatrix} \Phi(z) & \Phi(-z) \\ z^{-2Q-1} \Phi^*(-z) & -z^{-2Q-1} \Phi^*(z) \end{bmatrix}$$

so that $Q(z) \overline{Q}(z) = c(z) I_2$. Hence, if $c(z) \neq 0$ then $Q(z)$ is invertible with inverse $Q(z)^{-1} = c(z)^{-1} \overline{Q}(z)$ and we may multiply both sides of (1.33) by $c(z)^{-1} \overline{Q}(z)$ to obtain

$$u(z) = c(z)^{-1} \overline{Q}(z) w(z). \tag{1.33}$$

In particular, equating the top entries in this identity yields
Choose a smooth

Define

Its associated wavelet

(1.34) becomes

imposed to guarantee any regularity of the resulting wavelets.

To generate orthogonal wavelets, one still must verify Cohen’s condition

This is the fundamental equation defining $H$ in terms of the samples of $\varphi$ and its associated wavelet.

To see that $H$, as defined by (1.34) satisfies the QMF condition (1.30), observe that by (1.32), $c(z) = c(-z) = c^*(z) = c^*(-z)$. Also, since $Q(z)\tilde{Q}(z) = c(z)I_2$, (1.33) yields

$$H(z) H^*(z) + H(-z) H^*(-z) = \langle u(z), u(z^*) \rangle$$

$$= \langle \tilde{Q}(z) w(z)/c(z), \tilde{Q}(z^*) w(z^*)/c(z^*) \rangle$$

$$= \langle w(z), \tilde{Q}(z)^* \tilde{Q}(z) w(z^*)/c^2(z) \rangle$$

$$= \langle w(z), \tilde{c}(z) w(z^*)/c^2(z) \rangle = 1$$

since $c(z) = \langle w(z), w(z^*) \rangle$. Furthermore, the operator $T_H$ given by $T_H f(z^2) = H(z)f(z) + H(-z)f(-z)$ fixes $\Phi$, since

$$T_H \Phi(z^2) = \langle u(z), \bar{v}(z) \rangle = \langle c(z)^{-1} \tilde{Q}(z) w(x), \bar{v}(z) \rangle$$

$$= c(z)^{-1} \langle w(z), Q(z^*) \bar{v}(z) \rangle$$

$$= c(z)^{-1} \left\langle w(z), \begin{bmatrix} \bar{c}(z) \\ 0 \end{bmatrix} \right\rangle = \Phi(z^2).$$

1.2.3 QMF construction algorithm

In summary, one has the following algorithm for constructing QMFs:

1. Choose a smooth $\Phi$ on the complex plane with $\Phi(1) = 1$ and such that $c(z) = \Phi(z)\Phi^*(z) + \Phi(-z)\Phi^*(-z)$ is nonvanishing in a neighborhood of $|z| = 1$.
2. Choose $\Psi$ satisfying $\Psi(z^2)\Psi^*(z^2) = c(z) - \Phi(z^2)\Phi^*(z^2)$. If $\Phi(-1) \neq 0$ then one also requires $\Psi(1) = \Phi(-1)$.
3. Define $H$ by (1.34).

These conditions only insure that $H(0) = 1$ and $|H(\xi)|^2 + |H(\xi + 1/2)|^2 = 1$. To generate orthogonal wavelets, one still must verify Cohen’s condition (Proposition 1.4.1) independently. Moreover, additional conditions must be imposed to guarantee any regularity of the resulting wavelets.

Though the algorithm is quite general, design becomes nontrivial when properties such as the compact support of scaling functions are sought. Then

$$H(z) = \frac{1}{c(z)} (\Phi^*(z)\Phi(z^2) + z^{2Q+1} \Phi(-z)\Psi^*(z^2)).$$

(1.34)

Since $c(\xi) = |\Phi(\xi)|^2 + |\Phi(\xi + 1/2)|^2$ on the unit circle ($z = e^{2\pi i \xi}$), there

(1.34) becomes

$$H(-\xi) = \frac{\Phi(2\xi)\Phi(\xi) + e^{2\pi i (2Q+1)\xi} \Phi(2\xi)\Phi(\xi + 1/2)}{|\Phi(\xi)|^2 + |\Phi(\xi + 1/2)|^2}.$$
Φ should also satisfy certain constraints. If one can choose \( c(z) = c \), a constant, and still verify \( Ψ(1) = 1 \), then \( c = c(1) = 1 + |Φ^{-1}|^2 \geq 1 \). Then, by definition of \( c, Ψ \) should satisfy

\[
Ψ(z)Ψ^*(z) = c - Φ(z)Φ^*(z) = Φ(-z)Φ^*(z).
\]

Possible choices for such \( Ψ \) include \( Ψ(z) = Ψ^{(1)}(z) = ωz^NΦ(−z) \) or \( Ψ(z) = Ψ^{(2)}(z) = ωz^NΦ^*(−z) \) for some integer \( N \) and \( ω \in \mathbb{C} \) with \( |ω| = 1 \). If \( Φ(-1) \neq 0 \), then the compatibility condition becomes \( ωΦ(−1) = Φ(−1), \) i.e., \( ω = 1 \) for \( Ψ = Ψ^{(1)} \) and \( ωΦ(−1) = Φ(−1), \) i.e., \( ω = Φ(−1)/Φ(−1) \) for \( Ψ = Ψ^{(2)} \). In either case, \( H \) can be defined by

\[
H(z) = \frac{(Φ(z^2)Φ^*(z) + z^{2Q+1}Ψ^*(z^2)Φ(−z))}{1 + |Φ(−1)|^2}.
\]

### 1.2.4 Constraints on samples imposed by QMFs

For a concrete construction and parameterization, let \( Φ \) be the trigonometric polynomial \( Φ(z) = \sum_{k=1}^{M−1} a_k z^k \). The condition \( Φ(1) = 1 \) translates to

\[
\sum_{k=1}^{M−1} a_k = 1
\]

on the coefficient side. The condition \( c(z) = c = \text{const.} \) can be expressed as

\[
\sum_{k=1}^{M−1} a_k \bar{a}_{k−2m} = \frac{c}{2} \delta_m.
\]

If \( M = 2N + 1 \) is an odd integer, (1.37) imposes \( M - 2 \) nontrivial quadratic constraints and (1.36) imposes an additional linear constraint. Thus, an admissible \( Φ \) can be identified through its coefficients with a vector \( a = (a_1, a_2, \ldots, a_{M−1}) \in \mathbb{C}^{M−1} \) satisfying these constraints for some \( c \geq 1 \). This vector contains the integer values of the scaling function. With such a \( Φ \) given, \( Ψ \) must be chosen so that \( Ψ(z)Ψ^*(z) = Φ(z)Φ^*(z) \) and satisfying the compatibility condition. Then \( H \) is determined by (1.35).

### 1.2.5 Parameterization of four-coefficient systems

When \( M = 3 \), the family (1.37) contains just one nontrivial equation, namely

\[
|a_1|^2 + |a_2|^2 = c/2.
\]

From (1.36) we have \( a_3 = 1 - a_1 \), hence \( 2|a_1|^2 - 2\text{Re}(a_1) + 1 - c/2 = 0 \). For simplicity we assume \( a_1 \) and \( a_2 \) real. Then \( 2a_1^2 - 2a_1 + 1 - c/2 = 0 \) and therefore \( a_1 = (1 + \sqrt{c - 1})/2 \) (\( c \geq 1 \)). Let \( a_1^{(+)} = (1 + \sqrt{c - 1})/2 \), \( a_1^{(-)} = (1 - \sqrt{c - 1})/2 \). Then \( a_2^{(+)} = a_1^{(-)} \) and \( a_2^{(-)} = a_1^{(+)} \) so that \( Φ^{(+)}(z) = a_1^{(+)} z + a_2^{(+)} z^2 \) and \( Φ^{(-)}(z) = a_1^{(-)} z + a_2^{(-)} z^2 \) satisfy \( z^3 Φ^{(−)}(z^{−1}) = Φ^{(+)}(z) \).
and the construction starting from $\Phi(\cdot)$ will lead simply to a time reversal of the scaling function determined by $\Phi(\cdot)$. From now on $\Phi = \Phi(\cdot)$ and $\nu = -1/\sqrt{c} - 1$. Then

$$\Phi(z) = \left(\frac{\nu+1}{2\nu}\right) z + \left(\frac{\nu-1}{2\nu}\right) z^2.$$ 

Let $\Psi(z) = \Psi^{(1)}(z) = z^N \Phi(-z)$. Then

$$H(z) = e^{-1} (\Phi(z^2) \Phi^*(z) + z^{2Q-2N+1} \Phi^*(-z) \Phi(z))$$

$$= e^{-1} ((a_1 z^4 + a_2 z^4)(a_1 z + a_2 z^2)$$

$$+ z^{-2Q-1} (a_2 z^{-4} - a_1 z^{-2})(a_2 z^2 - a_1 z))$$

which is a polynomial for $N = Q - 1$. With this value of $N$, $H$ becomes

$$H(z) = \frac{1}{c} \left( (a_1^2 + a_2^2) z + (a_1^2 + a_2^2) z^2 \right) = \frac{1}{2} (z + z^2),$$

which produces a translation of the Haar scaling function. On the other hand, if $\Psi(\cdot)(z) = z^n \Phi^*(-z)$,

$$H(z) = \frac{1}{c} \left( (a_1 z^4 + a_2 z^4) \left( \frac{a_1}{z} + \frac{a_2}{z^2} \right) + z^{2Q-2N+1} (a_2 z^{-4} - a_1 z^{-2}) (a_2 z^2 - a_1 z) \right)$$

which is a polynomial of degree 3 if $N = Q + 2$. With this value of $N$, $H$ becomes

$$H(z) = \frac{1}{c} \left( a_1 (a_1 + a_2) + a_1 (a_1 - a_2) z + a_2 (a_2 - a_1) z^2 + a_2 (a_1 + a_2) z^3 \right)$$

$$= \frac{1}{2(\nu^2 + 1)} \left( \nu(\nu - 1) + (1 - \nu) z + (\nu + 1) z^2 + \nu(\nu + 1) z^3 \right). \quad (1.38)$$

As $\nu$ ranges over $(-1, 0)$ one recovers the Daubechies 4-tap QMFs (cf. [358]).

1.2.6 Cardinal scaling functions

These ideas furnish a simple proof of the nonexistence of orthogonal, continuous, compactly supported cardinal scaling functions (cf. Xia and Zhang [358]). Recall that a continuous function $f$ on $\mathbb{R}$ is cardinal if $f(k) = \delta_k$ ($k \in \mathbb{Z}$).

**Theorem 1.2.1.** Suppose that $\varphi$ is a compactly supported orthogonal cardinal scaling function. Then $\varphi$ must be the Haar function or its time reversal.

**Proof.** If $\varphi$ is cardinal, then $\Phi(\cdot) \equiv 1$, $c(\cdot) \equiv 2$ and by (1.32), $\Psi$ must satisfy $\Psi(\cdot) \Phi^*(\cdot) \equiv 1$. Since $\Psi$ is a trigonometric polynomial, it must be of the form $\Psi(z) = \lambda z^P$ for some $\lambda \in \mathbb{C}$ with $|\lambda| = 1$ and integer $P$. The compatibility condition $\Psi(1) = \Phi(-1)$ requires $\lambda = 1$ so that $\Psi(z) = z^P$. From (1.35) we have $H(z) = (1 + z^{2R+1}) / 2$ for some integer $R$ and when $z = e^{-2\pi \xi}$,
$H(\xi) = \frac{1}{2}(1 + e^{-2\pi i(2R+1)\xi}) = e^{-\pi i(2R+1)\xi} \cos \pi(2R+1)\xi = H_{\text{Haar}}((2R+1)\xi)$

where $H_{\text{Haar}}(\xi) = e^{-\pi i\xi} \cos \pi \xi$ is the QMF associated to the Haar scaling function $\phi_{\text{Haar}} = \chi_{[0,1]}$. Then

$$\hat{\phi}(\xi) = \prod_{j=1}^{\infty} H_{\text{Haar}} \left( \frac{(2R+1)\xi}{2^j} \right) = \hat{\phi}_{\text{Haar}}((2R+1)\xi).$$

In particular, $\phi(t) = \phi_{\text{Haar}}(t/(2R+1))/|2R+1|$ is not continuous at 0. If $\phi$ is cardinal in the sense of one-sided limits, then $\lim_{t\to 0^+} \phi(t) = 1$ or $\lim_{t\to 0^-} \phi(t) = 1$, and this requires $1/(|2R+1|) = 1$. Hence $R = 0$ or $R = -1$. If $R = 0$ then $\phi = \phi_{\text{Haar}}$ while if $R = -1$ then $\phi(t) = \phi_{\text{Haar}}(-t)$.

### 1.3 Computing the scaling function

The cascade algorithm generates approximate values of the scaling function $\phi$ at the dyadic rationals $D$. Exact values can be computed by the scaling relation (1.1): the values $\phi(k)$ solve $\phi(k) = \sum_i h_i \phi(2k-l)$ and values at dyadic rationals $k/2^j$ are then computed by iterating (1.1).

Exact values of $\phi$ along $D$ can also be obtained using Fourier/Zak transform techniques, as we will see here. In Chapter 3 this approach will be put to use in verifying sampling formulas in $V(\phi)$. The integer samples of the scaling function $\phi$ induced from $\Phi(\xi)$ as in the QMF construction algorithm above are just the Fourier coefficients of $\Phi$, that is, $\Phi(\xi) = \sum \phi(k)e^{2\pi ik\xi} = Z\phi(0, \xi)$. Let $H$ be the QMF associated with $\Phi$ by (1.35). Let $E$ denote the operation of multiplication by $e^{2\pi i\xi}$ and $T$ the operator $Tf(\xi) = H(\xi/2)f(\xi/2) + H(\xi/2 + 1/2)f(\xi/2 + 1/2)$ acting on $L^2([0,1])$. By (1.27), $\phi$ must satisfy $Z\phi(x, \xi) = T(Z\phi(2x, \cdot))(\xi)$. By quasiperiodicity, $Z\phi(l, \xi) = E^{-l}Z\phi(0, \xi) = E^{-l}\Phi(\xi)$ ($l \in \mathbb{Z}$). Combining and iterating these facts and the reconstruction formula $\phi(x) = \int_0^1 Z\phi(x, \xi) \, d\xi$, one must then have

$$\phi \left( \frac{l}{2^j} + k \right) = \int_0^1 E^{-k}T^l E^{-l} \Phi(\xi) \, d\xi.$$

Suppose now that $\phi$ is supported on $[0, M]$. Recall that if $A(\xi) = \sum_{j=1-M}^{M-1} a_j e^{2\pi ij\xi}$, then

$$a_0 = \int_0^1 A(\xi) \, d\xi = \frac{1}{M} \sum_{k=0}^{M-1} A \left( \frac{k}{M} \right).$$

(1.39)

If $0 \leq l \leq 2^j - 1$ then $Z\phi(l/2^j, \xi)$ is a forward trigonometric polynomial of degree at most $M - 1$. It follows that, for each $0 \leq k \leq M - 1$, $e^{-2\pi ik\xi} Z\phi(l/2^j, \xi) = \sum_{j=1-M}^{M-1} b_j e^{2\pi ij\xi}$ for some constants $b_j$ and we may use the quadrature formula (1.39) to obtain
\[
\varphi\left(\frac{l}{2^J} + k\right) = \frac{1}{M} \sum_{j=0}^{M-1} e^{-2\pi i j k/M} Z \varphi\left(\frac{l}{2^J}, \frac{j}{M}\right)
\]
\[
= \frac{1}{M} \sum_{j=0}^{M-1} e^{-2\pi i j k/M} T^J E^{-i} \Phi\left(\frac{j}{M}\right)
\]
\[
= \frac{1}{M} \sum_{j=0}^{M-1} e^{-2\pi i j k/M} \sum_{n=0}^{2^j-1} e^{-2\pi i n/2^j} F(j, n) \quad (1.40)
\]

where \( F \) is the \( M \times 2^J \) matrix with \((j, n)\)-th entry

\[
F(j, n) = \Phi\left(\frac{j + M n}{M 2^J}\right) \prod_{p=1}^{J} H\left(\frac{j + M n}{M 2^p}\right) \quad (0 \leq j \leq M - 1, 0 \leq n \leq 2^J - 1).
\]

Hence, values of \( \varphi \) at dyadic rationals may be computed from (1.40) by

\[
\varphi\left(\frac{l}{2^J} + k\right) = \sqrt{\frac{2^J}{M}} F_m^{(1)} F_n^{(2)} F(k, l) \quad (1.41)
\]

where \( F_n^{(i)} \) denotes the \( N \)-point discrete Fourier transform in the \( i \)th variable.

In summary: if the sequence \( a_l \) satisfies some mild constraints, then one can construct a scaling function \( \varphi \) such that \( \varphi(l) = a_l \). The QMF \( H \) of \( \varphi \) is obtained by passing through the Zak transform.

While defining a scaling function \( \varphi \) in terms of its integer samples has useful applications that we will see in Chapter 3, the previous considerations leave open one important problem, namely the possible regularity of a scaling function \( \varphi \) when it is defined in terms of its integer values as in Section 1.2.3.

### 1.4 Notes

**Frame multiresolution analysis.** Lawton [251] made an initial investigation of QMF wavelets of the type of the stretched Haar filter that give rise to frames for \( L^2(\mathbb{R}) \). The theory was developed further by Benedetto and Li [41] in terms of the overlap function \( \Phi(\xi) = \sum_k |\hat{\varphi}(\xi + k)|^2 \). The shifts of \( \varphi \) form a frame for \( V(\varphi) \) precisely if \( \Phi(\xi) \) is bounded and essentially bounded below on the set on which it does not vanish. In the case of an orthogonal scaling function, of course, one has \( \Phi \equiv 1 \). MRAs associated with such frames are studied in [44].

**Basic properties of wavelets.** In her book [99], Daubechies pointed out several basic problems in characterizing relationships between QMFs and wavelets on the one hand and wavelets and MRAs on the other. The latter
was already settled by 1990 and Lemarié-Rieusset characterized those wavelets that have an associated MRA by the simple condition
\[ \sum_{j=1}^{\infty} \sum_{k} |\hat{\psi}(2^j(\xi + k))|^2 = 1 \text{ a.e.} \]
Another basic question is: which filters \( H \) satisfying \( H(0) = 1 \) and \( |H(\xi)|^2 + |H(\xi + 1/2)|^2 = 1 \text{ a.e.} \) give rise to orthogonal scaling functions?

We saw that, when \( H \) is a trigonometric polynomial, the \( \tau \)-cycle condition must be satisfied. When \( H \) is only assumed to be \( C^1 \), a necessary and sufficient condition was discovered by Cohen (cf. [189], Section 7.4), namely:

**Proposition 1.4.1.** Under the hypotheses above, \( H \) generates an orthogonal scaling function if and only if there a finite union \( K \) of closed and bounded intervals containing 0 its interior such that \( \sum_k \chi_K(\xi + k) = 1 \text{ a.e. and } H(\xi/2^j) \neq 0 \) for all \( \xi \in K \) and \( j = 1, 2, \ldots \).

Still, several nagging issues remained. The WUTAM consortium (e.g., [367]) was a group of waveleteers at Washington University and Texas A&M who set out to settle several basic questions about wavelets once and for all. Some of the very basic properties appear in the book of Hernández and Weiss [189]. Here are a few others. A **wavelet multiplier** is a function \( \nu \) such that \((\nu \hat{\psi})^\vee \) is an orthonormal wavelet whenever \( \psi \) itself is an orthonormal wavelet. WUTAM characterized such \( \nu \) as those unimodular functions for which \( \nu(2\xi)/\nu(\xi) \) is periodic of period one. They also used such multipliers to prove that MRA wavelets are arc-connected in the sense that if \( \psi_0 \) and \( \psi_1 \) are two MRA wavelets then there is a continuous path \( A(t) : [0, 1] \to W \) (with \( W \) the class of MRA wavelets on the line) such that \( A(0) = \psi_0 \) and \( A(1) = \psi_1 \). A characterization of scaling filters was subsequently obtained by part of WUTAM [292].

**M-band wavelets.** In this chapter we considered only two-scale dilation equations. If one replaces 2 by \( m \) in the scaling equation (1.1) one obtains an \( m \)-scale dilation equation. In many cases it is possible to construct, but now in a less canonical way, orthogonal \( m \)-scale wavelet bases having \( m - 1 \) generators. There are some advantages to doing so. For example, Chui and Lian [78] constructed a scaling function with \( m = 3 \) leading to a pair of wavelets, one of which is symmetric and the other antisymmetric. In terms of subband coding, this approach corresponds to using \( m \) subbands rather than two; see [332].

**Convolution structure on scaling distributions.** As discussed in Section 1.1.3, the \( N \)-th order B-splines are all scaling functions and, at the same time, are \( N \)-th order autoconvolutions of the Haar scaling function. They are, however, not orthogonal to their shifts. These observations apply to more general convolutions of scaling functions. Suppose that \( \phi^{(1)} \) and \( \phi^{(2)} \) are scaling distributions, meaning that they are distributional solutions of \( \phi = 2 \sum_k h_k \phi(2-k) \) in the sense that, for every test function \( f \) one has
\[ \langle f, \phi \rangle = 2 \langle f, \sum_k h_k \phi(2-k) \rangle. \]
If, in addition, \( \phi \) is tempered then its Fourier transform is well defined and satisfies \( \hat{\phi}(\cdot) = H(\cdot/2)\hat{\phi}(\cdot/2) \) where \( H \) is the Fourier series of \( \{h_k\} \), provided this product exists. In what follows we shall assume at the very least that \( \hat{\phi} \) is defined as a locally square-integrable function and that \( H \) is locally bounded and continuous. Then the pointwise product \( \hat{\phi}^{(1)} \hat{\phi}^{(2)} \) of two such functions is well defined. Suppose now that we can justify rearranging the terms of the formal product to write

\[
\hat{\phi}^{(1)}(\xi)\hat{\phi}^{(2)}(\xi) = \prod H^{(1)}(\xi/2^n) \prod H^{(2)}(\xi/2^n) = \prod H^{(1)}(\xi/2^n) H^{(2)}(\xi/2^n)
\]
as is the case when the individual products converge locally uniformly.

In short, reasonable subclasses of scaling distributions will be closed under convolution. However, the product of two orthogonal QMFs \( H^{(1)}, H^{(2)} \) will never itself be orthogonal since, typically, \( |H^{(1)}H^{(2)}(\xi)|^2 + |H^{(1)}H^{(2)}(\xi + 1/2)|^2 < 1 \) (cf. (1.3)). Nevertheless, one can still attach to the convolution \( \hat{\phi} = \hat{\phi}^{(1)} * \hat{\phi}^{(2)} \) the orthogonal scaling function \( \varphi^{(1,2)}(\xi) = \hat{\phi}^{(1)}(\xi) \hat{\phi}^{(2)}(\xi)/\Phi(\xi) \) where \( \Phi(\xi) = \sum_k |\hat{\phi}^{(1)}(\xi + k)|^2 \) as was originally suggested by Aldroubi and Unser (e.g., [4,352]).

**Polyphase representation.** Compactly supported orthogonal scaling functions and wavelets are associated with a Laurent series \( H \) satisfying (1.30). We define \( G(z) = cz^{2Q+1}H^*(-z) \) for some integer \( Q \) and constant \( c \) with \( |c| = 1 \). Then \( (H,G) \) form a QMF pair, i.e.,

\[
H(z)H^*(z) + G(z)G^*(z) = 1 \tag{1.42}
\]

wherever the series defining \( H(z) \) and \( G(z) \) are defined (a set containing the unit circle). If the scaling function has compact support then \( H \) and \( G \) are finite polynomials and (1.42) is well defined on \( \mathbb{C} \setminus \{0\} \). As we have seen in Section 1.2.2, the QMF condition (1.42) can be expressed in terms of the paraunitarity of the modulation matrix (1.28), i.e., \( M(z)\bar{M}(z) = I \) with \( \bar{M}(z) \) the conjugate of \( M(z) \). If \( H \) and \( G \) are Laurent polynomials, this is equivalent to the unitarity of \( M(z) \) for all \( z \in \mathbb{T} \).

Any Laurent series \( P(z) = \sum_n p_n z^n \) can be expressed in *polyphase* form:

\[
P(z) = p_c(z^2) + z p_o(z^2)
\]

where \( p_c(z) = \sum_n p_{2n} z^n \) and \( p_o(z) = \sum_n p_{2n+1} z^n \). Given a QMF pair \( L = (H,G) \), one associates the *polyphase matrix*

\[
P_L(z) = \sqrt{2} \begin{bmatrix} h_c(z) & g_c(z) \\ h_o(z) & g_o(z) \end{bmatrix}
\]

For the Haar QMF, \( H(z) = (z + 1)/2 \) and \( G(z) = (z - 1)/2 \), the polyphase matrix simply takes the form \( P_{\text{Haar}}(z) = \left[ \begin{array}{cc} 1 & 1 \\ -1 & 1 \end{array} \right]/\sqrt{2} \).
Pollen’s parameterization. Pollen [299] revealed a group structure on polyphase matrices that leads to a unique factorization of orthogonal QMF pairs into so-called elementary factors. This should be contrasted with lifting—outlined below—which emphasizes factorization of biorthogonal filters. What follows is a brief outline of Pollen’s program.

Let $U(2,\mathbb{Z}_{1/2})$ be the group of paraunitary $2 \times 2$ matrices $A(z)$ whose entries are Laurent polynomials. The modulation matrices $M(z)$ of (1.28) are in $U(2,\mathbb{Z}_{1/2})$. The key to the factorization is the construction of a mapping that assigns the identity matrix to a nontrivial QMF.

Define a mapping $\Delta$ on $2 \times 2$ matrices by $\Delta(A) = P_{\text{Haar}} A^T$. Since $P_{\text{Haar}} P_{\text{Haar}}^T = I_2$, $\Delta$ maps $U(2,\mathbb{Z}_{1/2})$ to itself and $\Delta^{-1}(A) = A^T P_{\text{Haar}}$. Also, $\Delta$ acts on the polyphase matrix $P_L(z)$ of a QMF pair $L = (H, G)$ by

$$
\Delta(P_L) = P_{\text{Haar}} P_L^T = \begin{bmatrix}
    h_c(z) + h_o(z) & g_c(z) + g_o(z) \\
    h_o(z) - h_c(z) & g_o(z) - g_c(z)
\end{bmatrix}.
$$

Pollen defined the (Pollen) product $K_1 \sharp K_2$ of two QMF pairs $K_1 = (H_1, G_1)$ and $K_2 = (H_2, G_2)$ via the polyphase matrix

$$
P_{K_1 \sharp K_2}(z) = \Delta^{-1}(\Delta(P_{K_1}(z)) \Delta(P_{K_2}(z))) = P_{K_2}(z) P_{\text{Haar}}^T P_{K_1}(z).
$$

As $P_{K_1 \sharp K_2} \in U(2,\mathbb{Z}_{1/2})$, the product $\sharp$ defines a group multiplication on $U(2,\mathbb{Z}_{1/2})$. In this group, the Haar QMF acts via $P_{\text{Haar}}$ as group identity. Furthermore, if $H_1(z) = h_c(z) + z h_o(z^2)$, $G_1(z) = g_c(z) + z g_o(z^2)$ ($i = 1, 2$) then $K_3 = K_1 \sharp K_2$ is defined through its polyphase terms

$$
b_i^3 = [h_i(z) - g_i(z)], \quad g_i^3 = [h_i(z) + g_i(z)], \quad a_i^3 = [h_i(z) + g_i(z)]
$$

Pollens inverses have the form $P_{K^{-1}} = P_{\text{Haar}} \tilde{P}_K(z) P_{\text{Haar}}$ with $\tilde{P}_K(z)$ the paraconjugate of $P_K(z)$.

In fact, given any QMF $L = (H, G)$ one could define a Pollen product $\sharp_L$ on $U(2,\mathbb{Z}_{1/2})$ by putting $\Delta_L(A)(z) = P_L(z) A(z)^T$ and $P_{M_L B}(z) = \Delta_L^{-1}(\Delta_L(A) \Delta_L(B)) = B(z) P_L(z)^T A(z)$. Choosing $L$ to be the Haar QMF though is important for building wavelets from a set of minimal factors.

Pollens’s factorization of $SU(2,\mathbb{Z}_{1/2})$. Finite-degree QMFs are special pairs of Laurent polynomials. Pollen discovered a basic set of building blocks, referred to as factors, such that any QMF has a unique factorization under Pollens product. These factors have minimal degree. Specifically, if $p(z) = \sum_{m=-1}^1 a_m z^m$, one defines $d_{\text{max}}(p) = \max(k, l)$. The degree function $d_{\text{max}}$ can be extended to matrices $A(z, 1/z)$ by setting $d_{\text{max}}(A) = \max_{i,j} d_{\text{max}}(a_{ij}(z, 1/z))$.

**Proposition 1.4.2.** Every $P \in U(2,\mathbb{Z}_{1/2})$ can be factored in the form $P = \text{ABS}$ such that $A$ is a scalar matrix in $U(2)$, $B$ has the form $B = \begin{bmatrix} 1 & 0 \\ 0 & z^k \end{bmatrix}$ for some integer $k$, and $S \in SU_1(2,\mathbb{Z}_{1/2})$. 

Here $SU(2, [z, 1/z])$ denotes the subgroup of $U(2, [z, 1/z])$ consisting of parauitary matrices $A(z)$ of determinant 1 and $SU_I(2, [z, 1/z])$ is the subgroup of $SU(2, [z, 1/z])$ satisfying $A(1) = I$.

To factorize $SU_I(2, [z, 1/z])$, Pollen defined a set $F$ of basic factors

$$X = \begin{bmatrix} p(z) & q(z) \\ -q^*(z) & p^*(z) \end{bmatrix},$$

such that $p(z) = a + cz$, $q(z) = b + dz$ subject to the constraint that $X \in SU_I(2, [z, 1/z])$. The set $F$ itself does not contain all elements of $SU_I(2, [z, 1/z])$ having Laurent degree $d_{\text{max}} = 1$. In fact, $F^* \cap F = \emptyset$, but it turns out that the mapping $(X, Y) \mapsto XY^*$ maps $F \times F$ onto the subset of $SU_I(2, [z, 1/z])$ consisting of those matrices of degree at most one. Pollen’s unique factorization theorem is as follows.

**Theorem 1.4.3.** Any $P \in SU_I(2, [z, 1/z])$ of Laurent degree $m$ can be written in the form $P = \prod_{k=1}^{m} A_k^* B_k$ in which $A_k, B_k \in F$. Moreover this factorization is unique when considering $P$ as a product of elements of $F$.

The theorem is proved constructively by a series of lemmas providing conditions for pulling off factors of specific types. Existence and uniqueness are both proved by induction on the Laurent degree of $P$. We refer to [299] for its proof.

**More on parameterizing wavelets.** Scaling sequences form a subset of the unit ball in $\ell^2(\mathbb{Z})$. As such they form a metric space under the induced metric; however, this metric is not necessarily invariant under any transformation such as Pollen multiplication. To find a suitable metric, take the angle in $\ell^2$ between two scaling sequences $h = \{h_k\}$ and $k = \{k_l\}$, that is, $\theta = \arccos(\langle h, k \rangle)$. One then defines the new distance to be $\sqrt{2} |\sin(\theta/2)|$. It can be shown that Pollen multiplication is an isometry on the set of QMF sequences under this metric [361]. This result is of a different nature from the characterizations of wavelets by the WUTAM consortium.

**Wavelet systems are not ideal.** The group structure arising from Pollen’s product is sometimes called the **wavelet group**. It would be better termed the MRA group in the sense of Benedetto and Li’s frame MRAs [41], since the product does not preserve Cohen’s $\tau$-cycle condition (see Theorem 1.1.1). To give a simple example, the stretched Haar filter $K(z) = (1 + z^3)/2$ has Pollen square $K^2 K = (z^{-2} - z + z^2 + 2z^3 + z^5)/4$. The quantities $|K|^2$ and $K^2 K^*|^2$ are plotted in Figure 1.1. Clearly the $\tau$-cycle condition fails for $K$ but is satisfied by $K^2 K$.

**Pollen’s factorization: m × m case.** Pollen’s factorization has been considered in the case of integer dilations by Heller et al. [185], who also proved that the so-called $m$-band orthogonal wavelets form a group under the appropriate analogue of Pollen’s product. Their work was extended further by Kautsky and Turcajova [226] who addressed the problem of factoring $SL(m, [z, 1/z])$. Pollen’s factorization takes certain advantages of properties of $2 \times 2$ matrices.
that do not extend to larger matrices. In particular, when \( m > 2 \) the analog of Pollen’s factorization no longer produces unique factorization. In [226], a quasi-canonical factorization was considered. It was extended to the biorthogonal case by Resnikoff et al. [305]. One calls a pair \((C(z), D(z))\) a factorization of \( I \) provided (i) both \( C(z) \) and \( D(z) \) are matrix polynomials (not Laurent polynomials!), (ii) \( C(1) = 1 = D(1) \) and (iii) \( C(z)(D(\bar{z}))^* = I \). The quasi-canonical factorization states that a pair of polyphase matrices \((L, R)\) forms a biorthogonal matrix pair of rank \( m \) if and only if there exist primitive matrices \( V_1, \ldots, V_d \), a factorization \((C(z), D(z))\) of \( I \) and a \( G \) in the group \( GL(m-1) \) of invertible \((m-1) \times (m-1)\) matrices, such that

\[
L(z) = z^{-k_0}V_1(z) \cdots V_d(z)C(z) \begin{bmatrix} 1 & 0 \\ 0 & G \end{bmatrix} H, \quad \text{and}
\]

\[
R(z) = z^{-k_0}V_1(z) \cdots V_d(z)D(z) \begin{bmatrix} 1 & 0 \\ 0 & (G^{-1})^* \end{bmatrix} H.
\]

Here \( H \) denotes a canonical Haar matrix of size \( m - 1 \). Assuming that \( \det L(z) = cz^{-b} \), one has \( d = b - mk_0 \) while the degree of \( C(z) \) is at most the maximal degree of the polyphase components of \( L \).

**Euclidean algorithm for Laurent polynomials.** Pollen’s parameterization of QMF wavelets relied on a unique factorization of \( SU(2, [z, 1/z]) \). A second type of factorization of Laurent series—which is not unique—leads to a natural and rapidly implementable construction of biorthogonal wavelets.
The Laurent sum degree of \( q(z) = \sum_{k=M}^{N} c_k z^k \in \mathbb{C}[z, 1/z] \) is defined as 
\[ d_{\text{sum}}(q) = N - M. \]
Then \( d_{\text{sum}} \) is a homomorphism from the multiplicative group \( \mathbb{C}[z, 1/z] \) to the integers. Henceforth we will write \( D(p) = d_{\text{sum}}(p) \). Just as in the case of ordinary polynomials, given \( a, b \in \mathbb{C}[z, 1/z] \) with \( D(a) \geq D(b) \), one has the Euclidean algorithm factorization
\[
\begin{align*}
  a(z) &= b(z) q(z) + r(z) \\
  D(b) + D(q) &= D(a) \\
  D(r) &< D(b).
\end{align*}
\]
However, in contrast with \( \mathbb{C}[z] \), such a factorization is not necessarily unique. For example, 
\[
z + z^{-1} = (z + 1)(1 + z^{-1}) - 2 = (2z + 1)(1/2 + z^{-1}) - 5/2.
\]
This nonuniqueness actually provides some freedom in defining lifting steps to pass from lower-degree filters to higher-degree filters.

**Theorem 1.4.4. (Euclidean algorithm)** Let \( a, b \) be Laurent polynomials with \( D(a) \geq D(b) \). Starting with \( a = a_0 \) and \( b = b_0 \), inductively define \( a_{i+1} = b_i \) and \( b_{i+1} \) the remainder obtained from dividing \( b_i \) into \( a_i \). Let \( n \) be the smallest integer for which \( b_n(z) = 0 \). Then \( a_n(z) \) is the gcd of \( a \) and \( b \).

If, eventually, \( a_i \) is a monomial then \( a, b \) are relatively prime. When \( P = \begin{pmatrix} h^e & g^e \\ h^o & g^o \end{pmatrix} \) belongs to \( SL(2, \mathbb{C}[z, 1/z]) \), the group of matrix-valued Laurent polynomials having constant determinant one, the Laurent polynomials \( h^e \) and \( h^o \) are necessarily relatively prime: otherwise \( \det P \) has a nonzero root. Therefore, \( \gcd(h^e, h^o) \) is a monomial and in fact can be normalized to be a constant \( K \) as above. After \( n = D(h^e) \) steps of the Euclidean algorithm one has
\[
\begin{pmatrix} h^e \\ h^o \end{pmatrix} = \prod_{i=1}^{n} \begin{pmatrix} q_i(z) & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} K \\ 0 \end{pmatrix}.
\]
In case \( D(h^e) < D(h^o) \), take \( q_1 = 0 \), while if \( n \) is odd, multiplying \( h^e, h^o \) by \( z \) and dividing \( g^e, g^o \) by \( z \) does not affect the determinant. Thus, \( n \) may be assumed to be even.

With these conventions, starting with a filter \( H \) one can readily define a complementary filter \( G \) in such a way that
\[
P(z) = \begin{pmatrix} h^e & g^e \\ h^o & g^o \end{pmatrix} = \prod_{i=1}^{n} \begin{pmatrix} q_i(z) & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} K & 0 \\ 0 & 1/K \end{pmatrix} \in SL(2, \mathbb{C}[z, 1/z]).
\]
Because
\[
\begin{pmatrix} q(z) & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & q(z) \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ q(z) & 1 \end{pmatrix},
\]
once can take advantage of the fact that \( n \) is assumed even to write
Suppose \( G \) from which it is clear that \( h \) and by reading off the polyphase components of both sides we have \( L \).

\[ \text{Proof.} \]

Complementary filters. We have observed that the Euclidean algorithm provides a means of attaching to \( H \) a complementary filter \( G \) such that \( P(z) = \begin{bmatrix} h^e(z) & g^e(z) \\ h^o(z) & g^o(z) \end{bmatrix} \in SL(2, [z, 1/z]) \). From now on, we say that the filter pair \((H, G)\) is complementary when \( P \in SL(2, [z, 1/z]) \). Then \( G \) is said to complement \( H \), and vice versa. Nonuniqueness of the factorization indicates that there is more than one way to complement \( H \).

**Theorem 1.4.5.** If two Laurent polynomials \( G_1 \) and \( G_2 \) complement \( H \), then \( G_2(z) = G_1(z) + H(z)s(z^2) \) for some Laurent polynomial \( s \). Conversely, if \( G_1 \) complements \( H \) and \( G_2(z) = G_1(z) + H(z)s(z^2) \), then \( G_2 \) complements \( H \).

**Proof.** Suppose \( G_1, G_2 \) complement \( H \) and let

\[ P_1(z) = \begin{bmatrix} h^e(z) & g^e(z) \\ h^o(z) & g^o(z) \end{bmatrix}, \quad P_2(z) = \begin{bmatrix} h^e(z) & g^e(z) \\ h^o(z) & g^o(z) \end{bmatrix} \]

be the polyphase matrices associated with the pairs \( L_1 = (H, G_1) \) and \( L_2 = (H, G_2) \), respectively. By considering the explicit form of \( P^{-1} \) for \( P \in SL(2, [z, 1/z]) \), we see that \( P^{-1}_{1}(z)P_2(z) = \begin{bmatrix} 1 & s(z) \\ 0 & 1 \end{bmatrix} \). Hence \( P_2(z) = P_1(z)S(z) \) and by reading off the polyphase components of both sides we have \( g^e(z) = h^e(z)s(z)g^o(z) + g^e(z)g^o(z) \). Combining these polyphase terms gives \( G_2(z) = g^e(z)z^2 + zg^o(z)z^2 = G_1(z) + H(z)s(z^2) \). On the other hand, if \( G_1 \) complements \( H \) and \( G_2(z) = G_1(z) + H(z)s(z^2) \), then the polyphase components of \( G_2 \) are \( g^e(z) = h^e(z)s(z) + g^e(z) \) and \( g^o(z) = h^o(z)s(z) + g^o(z) \). The polyphase matrix for the pair \((H, G_2)\) is then

\[ P_2(z) = \begin{bmatrix} h^e(z) & g^e(z) \\ h^o(z) & g^o(z) \end{bmatrix} \begin{bmatrix} s(z) \\ 1 \end{bmatrix} \]

from which it is clear that \( P_2(z) \in SL(2, [z, 1/z]) \), i.e., \( G_2 \) complements \( H \).

\[ P(z) = \begin{bmatrix} K & 0 \\ 0 & 1/K \end{bmatrix} \prod_{i=1}^{n/2} \begin{bmatrix} 1 & q_{2i-1}(z) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ q_{2i}(z) & 1 \end{bmatrix}. \]
Theorem 1.4.5 is symmetric in $H$ and $G$. If two Laurent polynomials $H_1$ and $H_2$ complement $G$, then $H_2(z) = H_1(z) + G(z)t(z^2)$ for some Laurent polynomial $t$. Conversely, if $H_1$ complements $G$ and $H_2(z) = H_1(z) + G(z)t(z^2)$, then $H_2$ complements $G$. It is also worth noting that if $K_1 = (H_1, G)$ and $K_2 = (H_2, G)$ then $H_1$ and $H_2$ complement $G$ if and only if the polyphase matrices $Q_1(z)$ and $Q_2(z)$ associated to $K_1$ and $K_2$ respectively satisfy $Q_2(z) = Q_1(z) \left[ \begin{array}{cc} 1 & 0 \\ t(z) & 1 \end{array} \right]$ for some Laurent polynomial $t(z)$.

**Lifting.** While paraunitarity of the polyphase matrix $P$ of $(H, G)$ identifies the pair as a QMF, invertibility of $P$ indicates an ability to complete $H$ to a perfect reconstruction filter bank. In particular, we require a polyphase matrix $Q$ such that $P(z)Q(z) = I$. The Laurent polynomials $\det P$ and $\det Q$ satisfy $\det P(z)\det Q(z) = 1$ so that $\det P$ must be a monomial $cz^k$ since it is invertible in $\mathbb{C}[z, 1/z]$. One can then renormalize $P$ and $Q$ so that they belong to $SL(2, [z, 1/z])$.

The factorization technique just considered is the foundation of Sweldens’ lifting scheme [109]. Here, perfect reconstruction subband filters are decomposed into (alternatively lifted from) readily implemented lifting steps.

**Definition 1.4.6.** A polyphase matrix $P_2 \in SL(2, [z, 1/z])$ is said to be lifted from $P_1 \in SL(2, [z, 1/z])$ via $S(z) = \left[ \begin{array}{cc} 1 & \ast(z) \\ 0 & 1 \end{array} \right]$ if $P_2(z) = P_1(z)S(z)$. Similarly, $P_2 \in SL(2, [z, 1/z])$ is dual-lifted from $P_1$ via $S$ if $P_2 = P_1S^T$.

Theorem 1.4.5 tells us how to write a polyphase matrix $P$ as a product of lifting steps. That the same Laurent polynomial can be used to factor both the even and odd terms takes advantage of nonuniqueness of factorization.

Upon reassembling the modulation matrix (1.28) from the polyphase matrix $P(z)$, the lifting scheme can be summarized as: given complementary finite filters $H, G$, any other finite filter $G^{\text{new}}$ complementary to $H$ is obtained from $G$ by a lifting step:

$$G^{\text{new}}(z) = G(z) + H(z)s(z^2),$$

while any other finite filter $H^{\text{new}}$ complementary to $G$ is obtained from $H$ by a dual lifting step:

$$H^{\text{new}}(z) = H(z) + G(z)t(z^2).$$

Starting with the Lazy wavelet defined as $P = I$, other wavelets may be built via such lifting steps, so named because these steps lift the degree of the Laurent filters in question. The Haar wavelet itself requires a nontrivial lifting step. Examples of biorthogonal wavelets, and special considerations leading to orthogonal ones through lifting, are provided in [109].

Subband coding is improved through lifting: rather than implementing the filters $H, G$ at once, their lifting steps may be implemented in series. The scheme is also adaptable to other multiscale contexts that lack the shift-invariance of the MRA spaces, e.g., [102].