1 Basics of Wavelets

The first theoretical results in wavelets are connected with continuous wavelet decompositions of $\mathbb{L}_2$ functions and go back to the early 1980s. Papers of Morlet *et al.* (1982) and Grossmann and Morlet (1985) were among the first on this subject.

Let $\psi_{a,b}(x),\ a \in \mathbb{R}\backslash\{0\}, b \in \mathbb{R}$ be a family of functions defined as translations and re-scales of a single function $\psi(x) \in \mathbb{L}_2(\mathbb{R})$,

$$\psi_{a,b}(x) = \frac{1}{\sqrt{|a|}}\psi\left(\frac{x-b}{a}\right). \tag{1}$$

Normalization by $\frac{1}{\sqrt{|a|}}$ ensures that $||\psi_{a,b}(x)||$ is independent of a and b. The function ψ (called *the wavelet function* or *the mother wavelet*) is assumed to satisfy the *admissibility condition,*

$$C_\psi = \int_{\mathbb{R}} \frac{|\Psi(\omega)|^2}{|\omega|} d\omega < \infty, \tag{2}$$

where $\Psi(\omega) = \int_R \psi(x)e^{-ix\omega}dx$ is the Fourier transformation of $\psi(x)$. The admissibility condition (2) implies

$$0 = \Psi(0) = \int \psi(x)dx.$$

Also, if $\int \psi(x)dx = 0$ and $\int (1 + |x|^\alpha)|\psi(x)|dx < \infty$ for some $\alpha > 0$, then $C_\psi < \infty$.

Wavelet functions are usually normalized to "have unit energy", i.e., $||\psi_{a,b}(x)|| = 1$.

For any $\mathbb{L}_2$ function $f(x)$, the continuous wavelet transformation is defined as a function of two variables

$$\mathcal{CWT}_f(a,b) = \langle f, \psi_{a,b}\rangle = \int f(x)\overline{\psi_{a,b}(x)}dx.$$

Here the dilation and translation parameters, a and b, respectively, vary continuously over $\mathbb{R}\backslash\{0\} \times \mathbb{R}$.

Resolution of Identity. When the admissibility condition is satisfied, i.e., $C_\psi < \infty$, it is possible to find the inverse continuous transformation via the relation known as *resolution of identity* or *Calderón's reproducing identity,*

$$f(x) = \frac{1}{C_\psi}\int_{\mathbb{R}^2} \mathcal{CWT}_f(a,b)\psi_{a,b}(x)\frac{da\, db}{a^2}.$$

If a is restricted to $\mathbb{R}^+$, which is natural since a can be interpreted as a reciprocal of frequency, (2) becomes

$$C_\psi = \int_0^\infty \frac{|\Psi(\omega)|^2}{\omega} d\omega < \infty, \tag{3}$$

and the *resolution of identity* relation takes the form

$$f(x) = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \int_0^{\infty} \mathcal{CWT}_f(a,b) \psi_{a,b}(x) \frac{1}{a^2} da\, db. \tag{4}$$

Next, we list a few important properties of continuous wavelet transformations.

Shifting Property. If $f(x)$ has a continuous wavelet transformation $\mathcal{CWT}_f(a,b)$, then $g(x) = f(x-\beta)$ has the continuous wavelet transformation $\mathcal{CWT}_g(a,b) = \mathcal{CWT}_f(a, b-\beta)$.

Scaling Property. If $f(x)$ has a continuous wavelet transformation $\mathcal{CWT}_f(a,b)$, then $g(x) = \frac{1}{\sqrt{s}} f\left(\frac{x}{s}\right)$ has the continuous wavelet transformation $\mathcal{CWT}_g(a,b) = \mathcal{CWT}_f\left(\frac{a}{s}, \frac{b}{s}\right)$.

Both the shifting property and the scaling property are simple consequences of changing variables under the integral sign.

Energy Conservation. From (4),

$$\int_{-\infty}^{\infty} |f(x)|^2 dx = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \int_0^{\infty} |\mathcal{CWT}_f(a,b)|^2 \frac{1}{a^2} da\, db.$$

Localization. Let $f(x) = \delta(x - x_0)$ be the Dirac pulse at the point x_0. Then, $\mathcal{CWT}_f(a,b) = \frac{1}{\sqrt{a}} \psi(\frac{x_0 - b}{a})$.

Reproducing Kernel Property. Define $\mathbb{K}(u,v;a,b) = \langle \psi_{u,v}, \psi_{a,b} \rangle$. Then, if $F(u,v)$ is a continuous wavelet transformation of $f(x)$,

$$F(u,v) = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \int_0^{\infty} \mathbb{K}(u,v;a,b) F(a,b) \frac{1}{a^2} da\, db,$$

i.e., $\mathbb{K}$ is a reproducing kernel. The associated reproducing kernel Hilbert space (RKHS) is defined as a $\mathcal{CWT}$ image of $\mathbb{L}_2(\mathbb{R})$ – the space of all complex-valued functions F on $\mathbb{R}^2$ for which $\frac{1}{C_\psi} \int_{-\infty}^{\infty} \int_0^{\infty} |F(a,b)|^2 \frac{da\, db}{a^2}$ is finite.

Characterization of Regularity. Let $\int (1 + |x|)\, |\psi(x)|\, dx < \infty$ and let $\Psi(0) = 0$. If $f \in \mathbb{C}^\alpha$(Hölder space with exponent α), then

$$|\mathcal{CWT}_f(a,b)| \leq C|a|^{\alpha + 1/2}. \tag{5}$$

Conversely, if a continuous and bounded function f satisfies (5), then $f \in \mathbb{C}^{\alpha.}$

Example 1.1 Mexican hat or Marr's wavelet. The function

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

is a wavelet [known as the "Mexican hat" or Marr's wavelet.

By direct calculation one may obtain $C_\psi = 2\pi$.

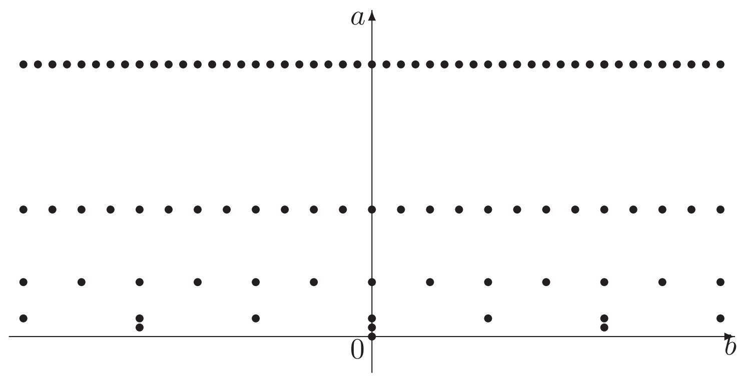

Figure 1: Critical Sampling in $\mathbb{R} \times \mathbb{R}^+$ half-plane ($a = 2^{-j}$ and $b = k\, 2^{-j}$).

Example 1.2 Poisson wavelet. The function $\psi(x) = -(1 + \frac{d}{dx})\frac{1}{\pi}\ \frac{1}{1+x^2}$ is a wavelet [known as the Poisson wavelet. The analysis of functions with respect to this wavelet is related to the boundary value problem of the Laplace operator.

The continuous wavelet transformation of a function of one variable is a function of two variables. Clearly, the transformation is redundant. To "minimize" the transformation one can select discrete values of a and b and still have a transformation that is invertible. However, sampling that preserves all information about the decomposed function cannot be coarser than the *critical sampling.*

The critical sampling (Fig. 1) defined by

$$a = 2^{-j},\ b = k 2^{-j},\ j, k \in \mathbb{Z}, \tag{6}$$

will produce the minimal basis. Any coarser sampling will not give a unique inverse transformation; that is, the original function will not be uniquely recoverable. Moreover under mild conditions on the wavelet function ψ, such sampling produces an orthogonal basis $\{\psi_{jk}(x) = 2^{j/2}\psi(2^j x - k),\ j, k \in \mathbb{Z}\}$.

There are other discretization choices. For example, selecting $a = 2^{-j},\ b = k$ will lead to non-decimated (or stationary) wavelets. For more general sampling, given by

$$a = a_0^{-j},\ b = k\, b_0\, a_0^{-j},\ j, k \in \mathbb{Z},\ a_0 > 1, b_0 > 0, \tag{7}$$

numerically stable reconstructions are possible if the system $\{\psi_{jk},\ j, k \in \mathbb{Z}\}$ constitutes a frame. Here

$$\psi_{jk}(x) = a_0^{j/2}\, \psi\left(\frac{x - k\, b_0\, a_0^{-j}}{a_0^{-j}}\right) = a_0^{j/2}\psi(a_0^j x - k\, b_0),$$

is (1) evaluated at (7).

Next, we consider wavelet transformations (wavelet series expansions) for values of a and b given by (6). An elegant theoretical framework for critically sampled wavelet transformation is *Mallat's Multiresolution Analysis* (Mallat, 87; 89a, 89b, 98).

1.1 Multiresolution Analysis

A multiresolution analysis (MRA) is a sequence of closed subspaces $V_n, n \in \mathbb{Z}$ in $\mathbb{L}_2(\mathbb{R})$ such that they lie in a containment hierarchy

$$\cdots \subset V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \cdots . \tag{8}$$

The nested spaces have an intersection that contains the zero function only and a union that is dense in $\mathbb{L}(\mathbb{R})$,

$$\cap_n V_j = \{\mathbf{0}\}, \quad \overline{\cup_j V_j} = \mathbb{L}_2(\mathbb{R}).$$

[With $\overline{A}$ we denoted the closure of a set A]. The hierarchy (8) is constructed such that (i) V-spaces are self-similar,

$$f(2^j x) \in V_j \text{ iff } f(x) \in V_0. \tag{9}$$

and (ii) there exists *a scaling function* $\phi \in V_0$ whose integer-translates span the space V_0,

$$V_0 = \left\{ f \in \mathbb{L}_2(\mathbb{R}) |\;\; f(x) = \sum_k c_k \phi(x-k) \right\},$$

and for which the set $\{\phi(\bullet - k),\; k \in \mathbb{Z}\}$ is an orthonormal basis.[1]

Mild technical conditions on ϕ are necessary for future developments. It can be assumed that $\int \phi(x)dx \geq 0$. Later, we will prove that this integral is in fact equal to 1. Since $V_0 \subset V_1$, the function $\phi(x) \in V_0$ can be represented as a linear combination of functions from V_1, i.e.,

$$\phi(x) = \sum_{k \in \mathbb{Z}} h_k \sqrt{2} \phi(2x - k), \tag{10}$$

for some coefficients $h_k,\; k \in \mathbb{Z}$. This equation is called the *scaling equation* (or two-scale equation) and it is fundamental in constructing, exploring, and utilizing wavelets.

ADD BASES FOR V_j.

ANY L_2 can be projected on V_j

In the wavelet literature, the reader may encounter an indexing of the multiresolution subspaces, which is the reverse of that in (8),

$$\cdots \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \subset \cdots . \tag{11}$$

FORM OF $\phi_{jk}(x)$.

[1] It is possible to relax the orthogonality requirement. It is sufficient to assume that the system of functions $\{\phi(\bullet - k),\; k \in \mathbb{Z}\}$ constitutes a Riesz basis for V_0.

Theorem 1.1 *For the scaling function it holds*

$$\int_{\mathbb{R}} \phi(x)dx = 1,$$

or, equivalently,

$$\Phi(0) = 1,$$

where $\Phi(\omega)$ *is Fourier transformation of* ϕ*,* $\int_{\mathbb{R}} \phi(x)e^{-i\omega x}dx$*.*

The coefficients h_n in (10) are important in connecting the MRA to the theory of signal processing. The (possibly infinite) vector $\mathbf{h} = \{h_n,\ n \in \mathbb{Z}\}$ will be called a *wavelet filter.* It is a low-pass (averaging) filter as will become clear later by considerations in the Fourier domain.

To further explore properties of multiresolution analysis subspaces and their bases, we will often work in the Fourier domain. Define the function m_0 as follows:

$$m_0(\omega) = \frac{1}{\sqrt{2}} \sum_{k\in\mathbb{Z}} h_k e^{-ik\omega} = \frac{1}{\sqrt{2}} H(\omega). \tag{12}$$

The function in (12) is sometimes called the *transfer function* and it describes the behavior of the associated filter $\mathbf{h}$ in the Fourier domain. Notice that the function m_0 is periodic with the period 2π and that the filter taps $\{h_n,\ n \in \mathbb{Z}\}$ are the Fourier coefficients of the function $H(\omega) = \sqrt{2}\, m_0(\omega)$.

In the Fourier domain, the relation (10) becomes

$$\Phi(\omega) = m_0\left(\frac{\omega}{2}\right) \Phi\left(\frac{\omega}{2}\right), \tag{13}$$

where $\Phi(\omega)$ is the Fourier transformation of $\phi(x)$. Indeed,

$$\begin{aligned}
\Phi(\omega) &= \int_{-\infty}^{\infty} \phi(x)e^{-i\omega x}dx \\
&= \sum_k \sqrt{2}\, h_k \int_{-\infty}^{\infty} \phi(2x-k)e^{-i\omega x}dx \\
&= \sum_k \frac{h_k}{\sqrt{2}} e^{-ik\omega/2} \int_{-\infty}^{\infty} \phi(2x-k)e^{-i(2x-k)\omega/2}d(2x-k) \\
&= \sum_k \frac{h_k}{\sqrt{2}} e^{-ik\omega/2}\ \Phi\left(\frac{\omega}{2}\right) \\
&= m_0\left(\frac{\omega}{2}\right) \Phi\left(\frac{\omega}{2}\right).
\end{aligned}$$

By iterating (13), one gets

$$\Phi(\omega) = \prod_{n=1}^{\infty} m_0\left(\frac{\omega}{2^n}\right), \tag{14}$$

which is convergent under very mild conditions on rates of decay of the scaling function ϕ. There are several sufficient conditions for convergence of the product in (14). For instance, the uniform convergence on compact sets is assured if (i) $m_0(\omega) = 1$ and (ii) $|m_0(\omega) - 1| < C|\omega|^\epsilon$, for some positive C and ϵ. See also Theorem 1.2.

Next, we prove two important properties of wavelet filters associated with an orthogonal multiresolution analysis, *normalization* and *orthogonality.*

Normalization.

$$\sum_{k \in \mathbb{Z}} h_k = \sqrt{2}. \tag{15}$$

Proof:

$$\begin{aligned}
\int \phi(x)dx &= \sqrt{2}\sum_k h_k \int \phi(2x-k)dx \\
&= \sqrt{2}\sum_k h_k \, \frac{1}{2}\int \phi(2x-k)d(2x-k) \\
&= \frac{\sqrt{2}}{2}\sum_k h_k \int \phi(x)dx.
\end{aligned}$$

Since $\int \phi(x)dx \neq 0$ by assumption, (15) follows.

This result also follows from $m_0(0) = 1$.

Orthogonality. For any $l \in \mathbb{Z}$,

$$\sum_k h_k h_{k-2l} = \delta_l. \tag{16}$$

Proof: Notice first that from the scaling equation (10) it follows that

$$\begin{aligned}
\phi(x)\phi(x-l) &= \sqrt{2}\,\sum_k h_k \phi(2x-k)\phi(x-l) \qquad (17) \\
&= \sqrt{2}\,\sum_k h_k \phi(2x-k)\,\sqrt{2}\,\sum_m h_m \phi(2(x-l)-m).
\end{aligned}$$

By integrating the both sides in (17) we obtain

$$\begin{aligned}
\delta_l &= 2\sum_k h_k \left[\sum_m h_m \frac{1}{2}\int \phi(2x-k)\phi(2x-2l-m)\,d(2x)\right] \\
&= \sum_k \sum_m h_k h_m \delta_{k,2l+m} \\
&= \sum_k h_k h_{k-2l}.
\end{aligned}$$

The last line is obtained by taking $k = 2l + m$.

An important special case is $l = 0$ for which (16) becomes

$$\sum_k h_k^2 = 1. \tag{18}$$

One consequence of the orthogonality condition (16) is the following: the convolution of filter $\mathbf{h}$ with itself, $\mathbf{f} = \mathbf{h} \star \mathbf{h}$, is an *à trous.*[2]

The fact that the system $\{\phi(\bullet - k),\ k \in \mathbb{Z}\}$ constitutes an orthonormal basis for V_0 can be expressed in the Fourier domain in terms of either $\Phi(\omega)$ or $m_0(\omega)$.

(**a**) In terms of $\Phi(\omega)$:

$$\sum_{l=-\infty}^{\infty} |\Phi(\omega + 2\pi l)|^2 = 1. \tag{19}$$

By the [PAR] property of the Fourier transformation and the 2π-periodicity of $e^{i\omega k}$ one has

$$\begin{aligned} \delta_k &= \int_{\mathbb{R}} \phi(x)\overline{\phi(x-k)}dx \\ &= \frac{1}{2\pi}\int_{\mathbb{R}} \Phi(\omega)\overline{\Phi(\omega)}e^{i\omega k}d\omega \\ &= \frac{1}{2\pi}\int_0^{2\pi} \sum_{l=-\infty}^{\infty} |\Phi(\omega+2\pi l)|^2 e^{i\omega k}d\omega. \end{aligned} \tag{20}$$

The last line in (20) is the Fourier coefficient a_k in the Fourier series decomposition of

$$f(\omega) = \sum_{l=-\infty}^{\infty} |\Phi(\omega + 2\pi l)|^2.$$

Due to the uniqueness of Fourier representation, $f(\omega) = 1$. As a side results, we obtain that $\Phi(2\pi n) = 0, n \neq 0$, and $\sum_n \phi(x - n) = 1$. The last result follows from inspection of coefficients c_k in the Fourier decomposition of $\sum_n \phi(x - n)$, the series $\sum_k c_k e^{2\pi i k x}$. Since this function is 1-periodic,

$$c_k = \int_0^1 \left(\sum_n \phi(x-n)\right) e^{-2\pi i k x}dx = \int_{-\infty}^{\infty} \phi(x)e^{-2\pi i k x}dx = \Phi(2\pi k) = \delta_{0,k}.$$

Remark 1.1 Utilizing the identity (19), any set of independent functions spanning V_0, $\{\phi(x-k), k \in \mathbb{Z}\}$, can be orthogonalized in the Fourier domain. The orthonormal basis is generated by integer-shifts of the function

$$\mathcal{F}^{-1}\left[\frac{\Phi(\omega)}{\sqrt{\sum_{l=-\infty}^{\infty} |\Phi(\omega+2\pi l)|^2}}\right]. \tag{21}$$

This normalization in the Fourier domain is used in constructing of some wavelet bases.

[2]The attribute *à trous (Fr.)* ($\equiv$ with holes) comes from the property $f_{2n} = \delta_n$, i.e., each tap on even position in $\mathbf{f}$ is 0, except the tap f_0. Such filters are also called half-band filters.

(**b**) In terms of m_0 :

$$|m_0(\omega)|^2 + |m_0(\omega+\pi)|^2 = 1. \tag{22}$$

Since $\sum_{l=-\infty}^{\infty} |\Phi(2\omega + 2l\pi)|^2 = 1$, then by (13)

$$\sum_{l=-\infty}^{\infty} |m_0(\omega + l\pi)|^2 |\Phi(\omega + l\pi)|^2 = 1. \tag{23}$$

Now split the sum in (23) into two sums – one with odd and the other with even indices, i.e.,

$$\begin{aligned} 1 &= \sum_{k=-\infty}^{\infty} |m_0(\omega + 2k\pi)|^2 |\Phi(\omega + 2k\pi)|^2 + \\ &\quad \sum_{k=-\infty}^{\infty} |m_0(\omega + (2k+1)\pi)|^2 |\Phi(\omega + (2k+1)\pi)|^2. \end{aligned}$$

To simplify the above expression, we use relation (19) and the 2π-periodicity of $m_0(\omega)$.

$$\begin{aligned} 1 &= |m_0(\omega)|^2 \sum_{k=-\infty}^{\infty} |\Phi(\omega + 2k\pi)|^2 + |m_0(\omega+\pi)|^2 \sum_{k=-\infty}^{\infty} |\Phi((\omega+\pi) + 2k\pi)|^2 \\ &= |m_0(\omega)|^2 + |m_0(\omega+\pi)|^2. \end{aligned}$$

Whenever a sequence of subspaces satisfies MRA properties, there exists (though not unique) an orthonormal basis for $\mathbb{L}_2(\mathbb{R})$,

$$\{\psi_{jk}(x) = 2^{j/2}\psi(2^j x - k),\ j,k \in \mathbb{Z}\} \tag{24}$$

such that $\{\psi_{jk}(x),\ j\text{-fixed},\ k \in \mathbb{Z}\}$ is an orthonormal basis of the "difference space" $W_j = V_{j+1} \ominus V_j$. The function $\psi(x) = \psi_{00}(x)$ is called a *wavelet function* or informally *the mother wavelet*.

Next, we detail the derivation of a wavelet function from the scaling function. Since $\psi(x) \in V_1$ (because of the containment $W_0 \subset V_1$), it can be represented as

$$\psi(x) = \sum_{k \in \mathbb{Z}} g_k \sqrt{2}\phi(2x - k), \tag{25}$$

for some coefficients $g_k,\ k \in \mathbb{Z}$.

Define

$$m_1(\omega) = \frac{1}{\sqrt{2}} \sum_k g_k e^{-ik\omega}. \tag{26}$$

By mimicking what was done with m_0, we obtain the Fourier counterpart of (25),

$$\Psi(\omega) = m_1(\frac{\omega}{2})\Phi(\frac{\omega}{2}). \tag{27}$$

The spaces W_0 and V_0 are orthogonal by construction. Therefore,

$$\begin{aligned} 0 = \int \psi(x)\phi(x-k)dx &= \frac{1}{2\pi}\int \Psi(\omega)\overline{\Phi(\omega)}e^{i\omega k}d\omega \\ &= \frac{1}{2\pi}\int_0^{2\pi}\sum_{l=-\infty}^{\infty}\Psi(\omega+2l\pi)\overline{\Phi(\omega+2l\pi)}e^{i\omega k}d\omega. \end{aligned}$$

By repeating the Fourier series argument, as in (19), we conclude

$$\sum_{l=-\infty}^{\infty}\Psi(\omega+2l\pi)\overline{\Phi(\omega+2l\pi)} = 0.$$

By taking into account the definitions of m_0 and m_1, and by mimicking the derivation of (22), we find

$$m_1(\omega)\overline{m_0(\omega)} + m_1(\omega+\pi)\overline{m_0(\omega+\pi)} = 0. \tag{28}$$

From (28), we conclude that there exists a function $\lambda(\omega)$ such that

$$(m_1(\omega),\ m_1(\omega+\pi)\ ) = \lambda(\omega)\left(\overline{m_0(\omega+\pi)},\ -\overline{m_0(\omega)}\ \right). \tag{29}$$

By substituting $\xi = \omega + \pi$ and by using the 2π-periodicity of m_0 and m_1, we conclude that

$$\begin{aligned} \lambda(\omega) &= -\lambda(\omega+\pi), \text{ and} \\ \lambda(\omega) & \quad \text{is } 2\pi\text{-periodic.} \end{aligned} \tag{30}$$

Any function $\lambda(\omega)$ of the form $e^{\pm i\omega}S(2\omega)$, where S is an $\mathbb{L}_2([0, 2\pi])$, 2π-periodic function, will satisfy (28); however, only the functions for which $|\lambda(\omega)| = 1$ will define an orthogonal basis ψ_{jk} of $\mathbb{L}_2(\mathbb{R})$.

To summarize, we choose $\lambda(\omega)$ such that

(i) $\lambda(\omega)$ is 2π-periodic,

(ii) $\lambda(\omega) = -\lambda(\omega+\pi)$, and

(iii) $|\lambda(\omega)|^2 = 1$.

Standard choices for $\lambda(\omega)$ are $-e^{-i\omega}$, $e^{-i\omega}$, and $e^{i\omega}$; however, any other function satisfying (i)-(iii) will generate a valid m_1. We choose to define $m_1(\omega)$ as

$$m_1(\omega) = -e^{-i\omega}\overline{m_0(\omega+\pi)}. \tag{31}$$

since it leads to a convenient and standard connection between the filters $\mathbf{h}$ and $\mathbf{g}$.

The form of m_1 and the equation (19) imply that $\{\psi(\bullet - k),\ k \in \mathbb{Z}\}$ is an orthonormal basis for W_0.

Since $|m_1(\omega)| = |m_0(\omega+\pi)|$, the orthogonality condition (22) can be rewritten as

$$|m_0(\omega)|^2 + |m_1(\omega)|^2 = 1. \tag{32}$$

By comparing the definition of m_1 in (26) with

$$\begin{aligned} m_1(\omega) &= -e^{-i\omega}\frac{1}{\sqrt{2}}\sum_k h_k e^{i(\omega+\pi)k} \\ &= \frac{1}{\sqrt{2}}\sum_k (-1)^{1-k} h_k e^{-i\omega(1-k)} \\ &= \frac{1}{\sqrt{2}}\sum_n (-1)^n h_{1-n} e^{-i\omega n}, \end{aligned}$$

we relate g_n and h_n as

$$g_n = (-1)^n\ h_{1-n}. \tag{33}$$

In signal processing literature, the relation (33) is known as the *quadrature mirror relation* and the filters $\mathbf{h}$ and $\mathbf{g}$ as *quadrature mirror filters.*

Remark 1.2 Choosing $\lambda(\omega) = e^{i\omega}$ leads to the rarely used high-pass filter $g_n = (-1)^{n-1}\ h_{-1-n}$. It is sometimes convenient to define g_n as $(-1)^n h_{1-n+M}$, where M is a "shift constant." Such re-indexing of $\mathbf{g}$ affects only the shift-location of the wavelet function.

1.2 Haar Wavelets

In addition to their simplicity and formidable applicability, Haar wavelets have tremendous educational value. Here we illustrate some of the relations discussed in the Section 1.1 using the Haar wavelet. We start with $\phi(x) = \mathbf{1}(0 \le x \le 1)$ and pretend that everything else is unknown.

The scaling equation (10) is very simple for the Haar case. By inspection of simple graphs of two scaled Haar wavelets $\phi(2x)$ and $\phi(2x+1)$ stuck to each other, we conclude that the scaling equation is

$$\begin{aligned} \phi(x) &= \phi(2x) + \phi(2x-1) \\ &= \frac{1}{\sqrt{2}}\sqrt{2}\phi(2x) + \frac{1}{\sqrt{2}}\sqrt{2}\phi(2x-1), \end{aligned} \tag{34}$$

which yields the wavelet filter coefficients:

$$h_0 = h_1 = \frac{1}{\sqrt{2}}.$$

Now, the transfer functions become

$$m_0(\omega) = \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}e^{-i\omega 0}\right) + \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}e^{-i\omega 1}\right) = \frac{1+e^{-i\omega}}{2}.$$

and

$$m_1(\omega) = -e^{-i\omega}\,\overline{m_0(\omega+\pi)} = -e^{-i\omega}\left(\frac{1}{2} - \frac{1}{2}e^{i\omega}\right) = \frac{1-e^{-i\omega}}{2}.$$

Notice that $m_0(\omega) = |m_0(\omega)|e^{i\varphi(\omega)} = \cos\frac{\omega}{2}\cdot e^{-i\omega/2}$ (after $\cos x = \frac{e^{ix}+e^{-ix}}{2}$). Since $\varphi(\omega) = -\frac{\omega}{2}$, Haar's wavelet has *linear phase,* i.e., the scaling function is symmetric in the time domain. The orthogonality condition $|m_0(\omega)|^2 + |m_1(\omega)|^2 = 1$ is easily verified, as well.

Relation (27) becomes

$$\Psi(\omega) = \frac{1-e^{-i\omega/2}}{2}\Phi\left(\frac{\omega}{2}\right) = \frac{1}{2}\Phi\left(\frac{\omega}{2}\right) - \frac{1}{2}\Phi\left(\frac{\omega}{2}\right)e^{-i\omega/2},$$

and by applying the inverse Fourier transformation we obtain

$$\psi(x) = \phi(2x) - \phi(2x-1)$$

in the time-domain. Therefore we "discovered" the Haar wavelet function ψ. From the expression for m_1 or by inspecting the representation of $\psi(x)$ by $\phi(2x)$ and $\phi(2x-1)$, we "conclude" that $g_0 = -g_{-1} = \frac{1}{\sqrt{2}}$.

The Haar basis is not an appropriate basis for all applications for several reasons. The building blocks in Haar's decomposition are discontinuous functions that obviously are not effective in approximating smooth functions. Although the Haar wavelets are well localized in the time domain, in the frequency domain they decay at the slow rate of $O(\frac{1}{n})$.

1.3 Daubechies' Compactly Supported Wavelets

Daubechies was first to construct compactly supported orthogonal wavelets with a preassigned degree of smoothness. Here we present the idea of Daubechies, omitting some technical details. Detailed treatment of this topic can be found in the monograph Daubechies (1992), Chapters 6 and 7.

Suppose that ψ has N (≥ 2) vanishing moments, i.e., $\int x^n\psi(x)dx = 0,\ n = 0,1,\ldots,N-1$. The $m_0(\omega)$ has the form:

$$m_0(\omega) = \left(\frac{1+e^{-i\omega}}{2}\right)^N \mathcal{L}(\omega), \tag{35}$$

where $\mathcal{L}(\omega)$ is a trigonometric polynomial. Indeed, since ψ has N vanishing moments, then $\Psi^{(n)}(0) = 0, n = 0,1,\ldots,N-1$. By differentiating (27), we get

$$[m_1(\omega)\Phi(\omega)]^{(n)}_{\omega=0} = 0,\ n = 0,1,\ldots,N-1.$$

Since $\Phi(0) = 1$, it follows that $m_1^{(n)}(0)$ or, equivalently, $m_0^{(n)}(\pi) = 0$, for $n = 0,1,\ldots,N-1$. Thus, $m_0(\omega)$ has to be as in (35).

In terms of

$$M_0(\omega) = |m_0(\omega)|^2 = \left(\cos^2 \frac{\omega}{2}\right)^N \cdot |\mathcal{L}(\omega)|^2,$$

the orthogonality condition (22) becomes

$$M_0(\omega) + M_0(\omega + \pi) = 1. \tag{36}$$

$|\mathcal{L}(\omega)|^2$ is a polynomial in $\cos\omega$. It can be re-expressed as a polynomial in $y = \sin^2 \frac{\omega}{2}$ since $\cos\omega = 1-2\sin^2 \frac{\omega}{2}$. This re-expression is beneficial since we can use some standard results in theory of polynomials, to specify $|\mathcal{L}(\omega)|^2$. Denote this polynomial by $P(\sin^2 \frac{\omega}{2})$. In terms of the polynomial P the orthogonality condition (36) becomes

$$(1-y)^N P(y) + y^N P(1-y) = 1, \quad (y = \sin^2 \frac{\omega}{2}). \tag{37}$$

By Bezout's result (outlined below), there exists a unique solution of the functional equation (37). It can be found by the Euclidean algorithm since the polynomials $(1-y)^N$ and y^N are relatively prime.

Lemma 1.1 *(Bezout) If p_1 and p_2 are two polynomials of degree n_1 and n_2, respectively, with no common zeroes, then there exist unique polynomials q_1 and q_2 of degree $n_2 - 1$ and $n_1 - 1$, respectively, so that*

$$p_1(x)q_1(x) + p_2(x)q_2(x) = 1.$$

For the proof of the lemma, we direct the reader to Daubechies ([**?**], 169-170). The unique solution of (37) with degree $deg(P(y)) \leq N-1$ is

$$\sum_{k=0}^{N-1} \binom{N+k-1}{k} y^k, \quad y = \sin^2 \frac{\omega}{2}, \tag{38}$$

and since it is positive for $y \in [0,1]$, it does not contradict the positivity of $|\mathcal{L}(\omega)|^2$.

Remark: If the degree of a solution is not required to be minimal then any other polynomial $Q(y) = P(y) + y^N R(\frac{1}{2} - y)$ where R is an odd polynomial preserving the positivity of Q, will lead to a different solution for $m_0(\omega)$. By choosing $R \neq 0$, one can generalize the standard Daubechies family, to construct symmlets, complex Daubechies wavelets, coiflets, etc.

The function $|m_0(\omega)|^2$ is now completely determined. To finish the construction we have to find its square root. A result of Riesz, known as the *spectral factorization lemma*, makes this possible.

Lemma 1.2 *(Riesz) Let A be a positive trigonometric polynomial with the property $A(-x) = A(x)$. Then, A is necessarily of the form*

$$A(x) = \sum_{m=1}^{M} u_m \cos mx.$$

In addition, there exists a polynomial B of the same order $B(x) = \sum_{m=1}^{M} v_m e^{imx}$ such that $|B(x)|^2 = A(x)$. If the coefficients u_m are real, then B can be chosen so that the coefficients v_m are also real.

We first represent $|\mathcal{L}(\omega)|^2$ as the polynomial

$$\frac{a_0}{2} + \sum_{k=1}^{N-1} a_k \cos^k \omega,$$

by replacing $\sin^2 \frac{\omega}{2}$ in (38) by $\frac{1-\cos\omega}{2}$.

An auxiliary polynomial P_A, such that $|\mathcal{L}(e^{-i\omega})|^2 = |P_A(e^{-i\omega})|$, is formed.

If $z = e^{-i\omega}$, then $\cos\omega = \frac{z+z^{-1}}{2}$ and one such auxiliary polynomial is

$$P_A(z) = \frac{1}{2} \sum_{k=1-N}^{N-1} a_{|k|} z^{N-1+k}. \tag{39}$$

Since $P_A(z) = z^{2N-2} P_A(\frac{1}{z})$, the zeroes of $P_A(z)$ appear in reciprocal pairs if real, and quadruples $(z_i, \bar{z}_i, z_i^{-1}, \bar{z}_i^{-1})$ if complex. Without loss of generality we assume that $z_j, \bar{z}_j$ and r_j lie outside the unit circle in the complex plane. Of course, then $z_j^{-1}, \bar{z}_j^{-1}$ and r_j^{-1} lie inside the unit circle. The factorized polynomial P_A can be written as

$$\begin{aligned} P_A(z) &= \frac{1}{2} a_{N-1} \left[\prod_{i=1}^{I} (z - r_i)(z - \frac{1}{r_i}) \right] \\ &\quad \left[\prod_{j=1}^{J} (z - z_j)(z - \bar{z}_j)(z - z_j^{-1})(z - \bar{z}_j^{-1}) \right]. \end{aligned} \tag{40}$$

Here $r_1, r_2, \ldots, r_I$ are real and non zero, and $z_1, \ldots, z_J$ are complex; $I + 2J = N - 1$.

The goal is to take a square root from $|P_A(z)|$ and the following simple substitution puts $|P_A(z)|$ in a convenient form.

Since $z = e^{-i\omega}$, we replace $|(z - z_j)(z - \bar{z}_j^{-1})|$ by $|z_j|^{-1}|z - z_j|^2$, and the polynomial $|P_A|$ becomes

$$\frac{1}{2}|a_{N-1}| \prod_{i=1}^{I} |r_i^{-1}| \prod_{j=1}^{J} |z_j|^{-2} \cdot |\prod_{i=1}^{I} (z - r_i) \prod_{j=1}^{J} (z - z_j)(z - \bar{z}_j)|^2.$$

Now, $\mathcal{L}(\omega)$ becomes

$$\begin{aligned} &\pm(\frac{1}{2}|a_{N-1}| \prod_{i=1}^{I} |r_i^{-1}| \prod_{j=1}^{J} |z_j|^{-2})^{\frac{1}{2}} \\ &\cdot |\prod_{i=1}^{I} (z - r_i) \prod_{j=1}^{J} (z - z_j)(z - \bar{z}_j)|, \; z = e^{-i\omega}, \end{aligned} \tag{41}$$

where the sign is chosen so that $m_0(0) = \mathcal{L}(0) = 1$. Note that $deg[P_A(z)] = deg[|\mathcal{L}(z)|^2] = N - 1$.

Finally, the coefficients $h_0, h_1, \ldots, h_{2N-1}$ in the polynomial $\sqrt{2}\, m_0(\omega)$ are the desired wavelet filter coefficients.

Example 1.3 We will find m_0 for $N = 2$.

$|\mathcal{L}(\omega)|^2 = \sum_{k=0}^{2-1} \binom{2+k-1}{k} \sin^2 \frac{k\omega}{2} = 1 + 2\frac{1-\cos\omega}{2} = \frac{1}{2}4 - 1 \cdot \cos\omega$ gives $a_0 = 4$ and $a_1 = -1$.

The auxiliary polynomial P_A is

$$\begin{aligned} P_A(z) &= \frac{1}{2}\sum_{k=-1}^{1} a_{|k|} z^{1+k} \\ &= \frac{1}{2}(-1 + 4z - z^2) \\ &= -\frac{1}{2}\left(z - (2+\sqrt{3})\right)\left(z - (2-\sqrt{3})\right). \end{aligned}$$

One square root from the above polynomial is

$$\begin{aligned} \sqrt{\frac{1}{2}\left(|-1|\right)\frac{1}{2+\sqrt{3}}}\left(z - (2+\sqrt{3})\right) &= \frac{1}{\sqrt{2}}\sqrt{2-\sqrt{3}}\left(z - (2+\sqrt{3})\right) \\ &= \frac{1}{2}\left((\sqrt{3}-1)z - (1+\sqrt{3})\right). \end{aligned}$$

The change in sign in the expression above is necessary, since the expression should have the value of 1 at $z = 1$ or equivalently at $\omega = 0$. Finally,

$$\begin{aligned} m_0(\omega) &= \left(\frac{1+e^{-i\omega}}{2}\right)^2 \frac{1}{2}\left((1-\sqrt{3})e^{-i\omega} + (1+\sqrt{3})\right) \\ &= \frac{1}{\sqrt{2}}\left(\frac{1+\sqrt{3}}{4\sqrt{2}} + \frac{3+\sqrt{3}}{4\sqrt{2}}e^{-i\omega} + \frac{3-\sqrt{3}}{4\sqrt{2}}e^{-2i\omega} + \frac{1-\sqrt{3}}{4\sqrt{2}}e^{-3i\omega}\right). \end{aligned}$$

Table 1 gives **h**-filters for DAUB2 - DAUB10 wavelets.

1.4 Regularity of Wavelets

There is at least continuum many different wavelet bases. An appealing property of wavelets is diversity in their properties. One can construct wavelets with different smoothness, symmetry, oscillatory, support, etc. properties. Sometimes the requirements can be conflicting since some of the properties are exclusive. For example, there is no symmetric real-valued wavelet with a compact support. Similarly, there is no $\mathbb{C}^\infty$-wavelet function with an exponential decay, etc.

Scaling functions and wavelets can be constructed with desired degree of smoothness. The regularity (smoothness) of wavelets is connected with the rate of decay of scaling functions and ultimately with the number of vanishing moments of scaling and wavelet functions. For instance, the Haar wavelet has only the "zeroth" vanishing moment (as a consequence of the admissibility condition) resulting in a discontinuous wavelet function.

Theorem 1.2 is important in connecting the regularity of wavelets, the number of vanishing moments, and the form of the transfer function $m_0(\omega)$. The proof is based on the Taylor series argument and the scaling properties of wavelet functions. For details, see Daubechies (1992), pp 153–155. Let

$$\mathcal{M}_k = \int x^k \phi(x) dx \text{ and } \mathcal{N}_k = \int x^k \psi(x) dx,$$

Table 1: The $\mathbf{h}$ filters for Daubechies' wavelets for $N = 2, \ldots, 10$ vanishing moments.

k	DAUB2	DAUB3	DAUB4
0	0.4829629131445342	0.3326705529500827	0.2303778133088966
1	0.8365163037378080	0.8068915093110930	0.7148465705529161
2	0.2241438680420134	0.4598775021184915	0.6308807679298592
3	-0.1294095225512604	-0.1350110200102548	-0.0279837694168604
4		-0.0854412738820267	-0.1870348117190935
5		0.0352262918857096	0.0308413818355607
6			0.0328830116668852
7			-0.0105974017850690
k	DAUB5	DAUB6	DAUB7
0	0.1601023979741926	0.1115407433501095	0.0778520540850092
1	0.6038292697971887	0.4946238903984531	0.3965393194819173
2	0.7243085284377723	0.7511339080210954	0.7291320908462351
3	0.1384281459013216	0.3152503517091976	0.4697822874051931
4	-0.2422948870663808	-0.2262646939654398	-0.1439060039285650
5	-0.0322448695846383	-0.1297668675672619	-0.2240361849938750
6	0.0775714938400454	0.0975016055873230	0.0713092192668303
7	-0.0062414902127983	0.0275228655303057	0.0806126091510831
8	-0.0125807519990819	-0.0315820393174860	-0.0380299369350144
9	0.0033357252854738	0.0005538422011615	-0.0165745416306669
10		0.0047772575109455	0.0125509985560998
11		-0.0010773010853085	0.0004295779729214
12			-0.0018016407040475
13			0.0003537137999745
k	DAUB8	DAUB9	DAUB10
0	0.0544158422431070	0.0380779473638881	0.0266700579005487
1	0.3128715909143165	0.2438346746126514	0.1881768000776480
2	0.6756307362973218	0.6048231236902548	0.5272011889316280
3	0.5853546836542239	0.6572880780514298	0.6884590394535462
4	-0.0158291052563724	0.1331973858249681	0.2811723436606982
5	-0.2840155429615815	-0.2932737832793372	-0.2498464243271048
6	0.0004724845739030	-0.0968407832230689	-0.1959462743773243
7	0.1287474266204823	0.1485407493381040	0.1273693403356940
8	-0.0173693010018109	0.0307256814793158	0.0930573646035142
9	-0.0440882539307979	-0.0676328290613591	-0.0713941471663802
10	0.0139810279173996	0.0002509471148278	-0.0294575368218849
11	0.0087460940474065	0.0223616621236844	0.0332126740593155
12	-0.0048703529934519	-0.0047232047577528	0.0036065535669515
13	-0.0003917403733769	-0.0042815036824646	-0.0107331754833277
14	0.0006754494064506	0.0018476468830567	0.0013953517470513
15	-0.0001174767841248	0.0002303857635232	0.0019924052951842
16		-0.0002519631889428	-0.0006858566949593
17		0.0000393473203163	-0.0001164668551292
18			0.0000935886703200
19			-0.0000132642028945

be the kth moments of the scaling and wavelet functions, respectively.

Theorem 1.2 *Let $\psi_{jk}(x) = 2^{j/2}\psi(2^j x - k),\ j,k \in \mathbb{Z}$ be an orthonormal system of functions in $\mathbb{L}_2(\mathbb{R})$,*

$$|\psi(x)| \leq \frac{C_1}{(1+|x|)^\alpha},\ \alpha > N,$$

and $\psi \in \mathbb{C}^{N-1}(\mathbb{R})$, where the derivatives $\psi^{(k)}(x)$ are bounded for $k \leq N-1$.

Then, ψ has N vanishing moments,

$$\mathcal{N}_k = 0,\ \ 0 \leq k \leq N-1.$$

If, in addition,

$$|\phi(x)| \leq \frac{C_2}{(1+|x|)^\alpha},\ \alpha > N$$

then, the associated function $m_0(\omega)$ is necessarily of the form

$$m_0(\omega) = \left(\frac{1+e^{-i\omega}}{2}\right)^N \cdot \mathcal{L}(\omega), \tag{42}$$

where $\mathcal{L}$ is a 2π-periodic, $\mathbb{C}^{N-1}$-function.

The following definition of regularity is often used,

Definition 1.1 *The multiresolution analysis (or, the scaling function) is said to be r-regular if, for any $\alpha \in \mathbb{Z}$,*

$$|\phi^{(k)}(x)| \leq \frac{C}{(1+|x|)^\alpha},$$

for $k = 0, 1, \ldots, r$.

The requirement that ψ possesses N vanishing moments can be expressed in terms of Ψ, m_0, or equivalently, in terms of the filter $\mathbf{h}$.

Assume that a wavelet function $\psi(x)$ has N vanishing moments, i.e.,

$$\mathcal{N}_k = 0,\ k = 0, 1, \ldots, N-1. \tag{43}$$

By basic property of Fourier transformations, the requirement (43) corresponds to

$$\left.\frac{d^k \Psi(\omega)}{d\omega^k}\right|_{\omega=0} = 0,\ k = 0, 1, \ldots, N-1,$$

which implies

$$m_1^{(k)}(\omega)\,|_{\omega=0} = m_1^{(k)}(0) = 0,\ k = 0, 1, \ldots, N-1. \tag{44}$$

It is easy to check that in terms of m_0, relation (44) becomes

$$m_0^{(k)}(\omega)\,|_{\omega=\pi} = m_0^{(k)}(\pi) = 0,\ k = 0, 1, \ldots, N-1. \tag{45}$$

The argument is inductive. The case $k = 0$ follows from $\Psi(0) = m_1(0)\Phi(0)$ [(27) evaluated at $\omega = 0$] and the fact that $\Phi(0) = 1$. Since $\Psi'(0) = \frac{1}{2}m_1'(0)\Psi(0) + \frac{1}{2}m_1(0)\Psi'(0)$ it follows that $m_1'(0) = 0$, as well. Then, $m_1^{(N-1)}(0) = 0$ follows by induction.

The condition $m_1^{(k)}(0) = 0,\ k = 0, 1, \ldots, N-1$ translates to a constraint on the wavelet-filter coefficients

$$\sum_{n\in\mathbb{Z}} n^k g_n = \sum_{n\in\mathbb{Z}} (-1)^n n^k h_n = 0,\ \ k = 0, 1, \ldots, N-1. \tag{46}$$

How smooth are the wavelets from the Daubechies family? There is an apparent trade-off between the length of support and the regularity index of scaling functions. Daubechies (1988) and Daubechies and Lagarias (1991, 1992), obtained regularity exponents for wavelets in the Daubechies family.

Let ϕ be the DAUBN scaling function. There are two popular measures of regularity of ϕ: Sobolev and Hölder regularity exponents. Let α_N^* be the supremum of β such that

$$\int (1 + |\omega|)^\beta |\Phi(\omega)| d\omega < \infty,$$

and let α_N be the exponent of the Hölder space $\mathbb{C}^{\alpha_N}$ to which the scaling function ϕ belongs.

Table 2: Sobolev α_N^* and Hölder α_N regularity exponents of Daubechies' scaling functions.

N	1	2	3	4	5	6	7	8	9	10
α_N^*	0.5	1	1.415	1.775	2.096	2.388	2.658	2.914	3.161	3.402
α_N		0.550	0.915	1.275	1.596	1.888	2.158	2.415	2.661	2.902

The following result describes the limiting behavior of α_N.

Theorem 1.3

$$\lim_{N\to\infty} \alpha_N = N\left(1 - \frac{\log 3}{2\log 2}\right) + O(\frac{\ln N}{N}).$$

From Table 2, we see that DAUB4 is the first differentiable wavelet, since $\alpha > 1$. More precise bounds on α_N yield that ϕ from the DAUB3 family is, in fact, the first differentiable scaling function ($\alpha_3 = 1.0878$), even though it seams to have a peak at 1. See also Daubechies (1992), page 239, for the discussion.

Remark 1.3 the Sobolev and Hölder regularities are related, thus, Theorem 1.3 holds for the exponent α_N^*, as well.

1.5 Approximations and Characterizations of Functional Spaces

Any $\mathbb{L}_2(\mathbb{R})$ function f can be represented as

$$f(x) = \sum_{j,k} d_{jk}\psi_{jk}(x),$$

and this unique representation corresponds to a multiresolution decomposition $\mathbb{L}_2(\mathbb{R}) = \bigoplus_{j=-\infty}^{\infty} W_j$. Also, for any fixed j_0 the decomposition $\mathbb{L}_2(\mathbb{R}) = V_{j_0} \oplus \bigoplus_{j=j_0}^{\infty} W_j$ corresponds to the representation

$$f(x) = \sum_k c_{j_0,k}\phi_{j_0,k}(x) + \sum_{j \ge j_0}\sum_k d_{jk}\psi_{j,k}(x). \tag{47}$$

The first sum in (47) is an orthogonal projection $\mathbb{P}_{j_0}$ of f on V_{j_0}.

In general, it is possible to bound $||\mathbb{P}_{j_0}f - f|| = ||(\boldsymbol{I} - \mathbb{P}_{j_0})f||$ if the regularities of functions f and ϕ are known.

When both f and ϕ have n continuous derivatives, there exists a constant C such that

$$||(\boldsymbol{I} - \mathbb{P}_{j_0})f||_{\mathbb{L}_2} \le C \cdot 2^{-nj_0}||f||_{\mathbb{L}_2}.$$

A range of important function spaces can be fully characterized by wavelets. We list a few characterizations. For example, a function f belongs to the Hölder space $\mathbb{C}^s$ if and only if there is a constant C such that in an r-regular MRA ($r > s$) the wavelet coefficients satisfy

$$\begin{aligned} (i) \qquad & |c_{j_0,k}| \le C, \\ (ii) \qquad & |d_{j,k}| \le C \cdot 2^{-j(s+\frac{1}{2})},\ j \ge j_0, k \in \mathbb{Z}\,. \end{aligned} \tag{48}$$

A function f belongs to the Sobolev $\mathbb{W}_2^s(\mathbb{R})$ space if and only if

$$\sum_{j,k} |d_{jk}|^2 \cdot (1 + 2^{2js}) < \infty.$$

Even the general (non-homogeneous) Besov spaces, can be characterized by moduli of the wavelet coefficients of its elements. For a given r-regular MRA with $r > \max\{\sigma, 1\}$, the following result (see Meyer 1992, page 200) holds

Theorem 1.4 *Let I_j be a set of indices so that $\{\psi_i,\ i \in I_j\}$ constitutes an o.n. basis of the detail space W_j. There exist two constants $C' \ge C > 0$ such that, for every exponent $p \in [1, \infty]$, for each $j \in \mathbb{Z}$ and for every element $f(x) = \sum_{i \in I_j} d_i\psi_i(x)$ in W_j,*

$$C||f||_p \le 2^{j/2}2^{-j/p}\left(\sum_{i \in I_j}|d_i|^p\right)^{1/p} \le C'||f||_p.$$

The following characterization of Besov $\mathbb{B}^{\sigma}_{p,q}$ spaces can be obtained directly from this result. If the MRA has regularity $r > s$, then wavelet bases are Riesz bases for all $1 \leq p, q \leq \infty,\ 0 < \sigma < r$.

The function $f = \sum_k c_{j_0 k}\phi_{j_0 k}(x) + \sum_{j \geq j_0}\sum_k d_{jk}\psi_{jk}(x)$ belongs to $\mathbb{B}^{\sigma}_{p,q}$ space if its wavelet coefficients satisfy

$$\left(\sum_k |c_{j_0,k}|^p\right)^{1/p} < \infty,$$

and

$$\left\{\left(\sum_{i\in I_j} 2^{j(\sigma+1/2-1/p)}|d_i|^p\right)^{1/p},\ j \geq j_0\right\}$$

is an ℓ_q sequence, i.e, $\left[\sum_{j\geq j_0}\left(2^{j(\sigma+1/2-1/p)}(\sum_k |d_{j,k}|^p)^{1/p}\right)^q\right]^{1/q} < \infty.$

The results listed are concerned with global regularity. The local regularity of functions can also be studied by inspecting the magnitudes of their wavelet coefficients. For more details, we direct the reader to the work of Jaffard (1991) and Jaffard and Laurencot (1992).

1.5.1 Daubechies-Lagarias Algorithm

As a nice calculational example, we describe an algorithm for fast numerical calculation of wavelet values at a given point, based on the Daubechies-Lagarias (Daubechies and Lagarias, 1991, 1992) *local pyramidal algorithm*. The `matlab` routines `Phi.m` and `Psi.m` in the Matlab sesction implement the algorithm.

The scaling function and wavelet function in Daubechies' families have no explicit representations (except for the Haar wavelet). Sometimes, it necessary to find values of DAUB functions at arbitrary points; examples include calculation of coefficients in density estimation and non-equally spaced regression.

The Daubechies-Lagarias algorithm enables us to evaluate ϕ and ψ at a point with preassigned precision. We illustrate the algorithm on wavelets from the Daubechies family; however, the algorithm works for all orthoginal wavelet filters.

Let ϕ be the scaling function of the DAUBN wavelet. The support of ϕ is $[0, 2N-1]$. Let $x \in (0,1)$, and let $dyad(x) = \{d_1, d_2, \ldots d_n, \ldots\}$ be the set of 0-1 digits in the dyadic representation of x $(x = \sum_{j=1}^{\infty} d_j 2^{-j})$. By $dyad(x,n)$, we denote the subset of the first n digits from $dyad(x)$, i.e., $dyad(x,n) = \{d_1, d_2, \ldots d_n\}$.

Let $\mathbf{h} = (h_0, h_1, \ldots, h_{2N-1})$ be the wavelet filter coefficients. We build two $(2N-1) \times (2N-1)$ matrices as:

$$T_0 = (\sqrt{2}\cdot h_{2i-j-1})_{1\leq i,j\leq 2N-1} \text{ and } \quad T_1 = (\sqrt{2}\cdot h_{2i-j})_{1\leq i,j\leq 2N-1}. \tag{49}$$

Then the local pyramidal algorithm can be constructed based on Theorem 1.5.

Theorem 1.5

$$\lim_{n\to\infty} \quad T_{d_1} \cdot T_{d_2} \cdot \cdots \cdot T_{d_n} \tag{50}$$

$$= \begin{bmatrix} \phi(x) & \phi(x) & \ldots & \phi(x) \\ \phi(x+1) & \phi(x+1) & \ldots & \phi(x+1) \\ \vdots & & & \\ \phi(x+2N-2) & \phi(x+2N-2) & \ldots & \phi(x+2N-2) \end{bmatrix}.$$

The convergence of $||T_{d_1} \cdot T_{d_2} \cdot \cdots \cdot T_{d_n} - T_{d_1} \cdot T_{d_2} \cdot \cdots \cdot T_{d_{n+m}}||$ to zero, for fixed m, is exponential and constructive, i.e., effective decreasing bounds on the error can be established.

Example 1.4 Consider the DAUB2 scaling function ($N = 2$). The corresponding filter is $\mathbf{h} = \left(\frac{1+\sqrt{3}}{4\sqrt{2}}, \frac{3+\sqrt{3}}{4\sqrt{2}}, \frac{3-\sqrt{3}}{4\sqrt{2}}, \frac{1-\sqrt{3}}{4\sqrt{2}} \right)$. According to (49) the matrices T_0 and T_1 are given as

$$T_0 = \begin{bmatrix} \frac{1+\sqrt{3}}{4} & 0 & 0 \\ \frac{3-\sqrt{3}}{4} & \frac{3+\sqrt{3}}{4} & \frac{1+\sqrt{3}}{4} \\ 0 & \frac{1-\sqrt{3}}{4} & \frac{3-\sqrt{3}}{4} \end{bmatrix} \quad \text{and} \quad T_1 = \begin{bmatrix} \frac{3+\sqrt{3}}{4} & \frac{1+\sqrt{3}}{4} & 0 \\ \frac{1-\sqrt{3}}{4} & \frac{3-\sqrt{3}}{4} & \frac{3+\sqrt{3}}{4} \\ 0 & 0 & \frac{1-\sqrt{3}}{4} \end{bmatrix}.$$

Let us evaluate the scaling function at an arbitrary point, say $x = 0.45$. Twenty "decimals" in the dyadic representation of 0.45 are $dyad(0.45, 20) = \{\ 0,\ 1,\ 1,\ 1,\ 0,\ 0,\ 1,\ 1,\ 0,\ 0,\ 1,\ 1,\ 0,\ 0,\ 1,\ 1,\ 0,\ 0,\ 1,\ 1\ \}$. In addition to the value at 0.45, we get (for free) the values at 1.45 and 2.45 (the values 0.45, 1.45, and 2.45 are in the domain of ϕ, the interval [0,3]). The values $\phi(0.45), \phi(1.45)$, and $\phi(2.45)$ may be approximated as averages of the first, second, and third row, respectively in the matrix

$$\prod_{i \in dyad(0.45,20)} T_i = \begin{bmatrix} 0.86480582 & 0.86480459 & 0.86480336 \\ 0.08641418 & 0.08641568 & 0.08641719 \\ 0.04878000 & 0.04877973 & 0.04877945 \end{bmatrix}.$$

The Daubechies-Lagarias algorithm gives only the values of the scaling function. In applications, most of the evaluation needed involves the wavelet function. It turns out that another algorithm is unnecessary, due to the following result.

Theorem 1.6 *Let x be an arbitrary real number, let the wavelet be given by its filter coefficients, and let $\mathbf{u}$ with $2N-1$ be a vector defined as*

$$\mathbf{u}(x) = \{(-1)^{1-\lfloor 2x \rfloor} h_{i+1-\lfloor 2x \rfloor}, i = 0, \ldots, 2N-2\}.$$

If for some i the index $i + 1 - \lfloor 2x \rfloor$ is negative or larger than $2N - 1$, then the corresponding component of $\mathbf{u}$ is equal to 0.

Let the vector $\mathbf{v}$ be

$$\mathbf{v}(x,n) = \frac{1}{2N-1}\mathbf{1}' \prod_{i \in dyad(\{2x\},n)} T_i,$$

where $\mathbf{1}' = (1,1,\ldots,1)$ *is the row-vector of ones. Then,*

$$\psi(x) = \lim_{n\to\infty} \mathbf{u}(x)'\mathbf{v}(x,n),$$

and the limit is constructive.

Proof of the theorem is a straightforward but somewhat tedious re-expression of (25).

1.5.2 Moment Conditions Determine Filters

We saw that the requirement that the wavelet function possesses N-vanishing moments was expressed in terms of Φ, m_0, or $\mathbf{h}$.

Suppose that we wish to design a wavelet filter $\mathbf{h} = \{h_0,\ldots,h_{2N-1}\}$ only by considering properties of its filter taps. Assume that

$$\mathcal{N}_k = \int_{\mathbb{R}} x^k \psi(x)dx = 0, \text{ for } k = 0,1,\ldots,N-1. \tag{51}$$

As it was discussed in Section 1.1, some relevant properties of a multiresolution analysis can be expressed as relations involving coefficients of the filter $\mathbf{h}$.

For example, the normalization property gave

$$\sum_{i=0}^{2N-1} h_i = \sqrt{2},$$

the requirement for vanishing moments on ψ led to

$$\sum_{i=0}^{2N-1} (-1)^i i^k h_i = 0,\ k = 0,1,\ldots,N-1,$$

and, finally, the orthogonality property reflected to

$$\sum_{i=0}^{2N-1} h_i h_{i+2k} = \delta_k,\ k = 0,1,\ldots,N-1.$$

That defines $2N+1$ equations with $2N$ unknowns; however the system is solvable since the equations are not linearly independent. For example, the equation

$$h_0 - h_1 + h_2 - \cdots - h_{2N-1} = 0,$$

can be expressed as a linear combination of the others.

Example 1.5 For $N = 2$, we obtain the system:

$$\begin{cases} h_0 + h_1 + h_2 + h_3 = \sqrt{2} \\ h_0^2 + h_1^2 + h_2^2 + h_3^2 = 1 \\ -h_1 + 2h_2 - 3h_3 = 0 \\ h_0\ h_2 + h_1\ h_3 = 0 \end{cases},$$

which has the familiar solution $h_0 = \frac{1+\sqrt{3}}{4\sqrt{2}}, h_1 = \frac{3+\sqrt{3}}{4\sqrt{2}}, h_2 = \frac{3-\sqrt{3}}{4\sqrt{2}}$, and $h_3 = \frac{1-\sqrt{3}}{4\sqrt{2}}$.

For $N = 4$, the system is

$$\begin{cases} h_0 + h_1 + h_2 + h_3 + h_4 + h_5 + h_6 + h_7 = \sqrt{2} \\ h_0^2 + h_1^2 + h_2^2 + h_3^2 + h_4^2 + h_5^2 + h_6^2 + h_7^2 = 1 \\ h_0 - h_1 + h_2 - h_3 + h_4 - h_5 + h_6 - h_7 = 0 \\ h_0h_2 + h_1h_3 + h_2h_4 + h_3h_5 + h_4h_6 + h_5h_7 = 0 \\ h_0h_4 + h_1h_5 + h_2h_6 + h_3h_7 = 0 \\ h_0h_6 + h_1h_7 = 0 \\ 0h_0 - 1h_1 + 2h_2 - 3h_3 + 4h_4 - 5h_5 + 6h_6 - 7h_7 = 0 \\ 0h_0 - 1h_1 + 4h_2 - 9h_3 + 16h_4 - 25h_5 + 36h_6 - 49h_7 = 0 \\ 0h_0 - 1h_1 + 8h_2 - 27h_3 + 64h_4 - 125h_5 + 216h_6 - 343h_7 = 0. \end{cases}$$

The above systems can easily be solved by a symbolic software package such as `Maple` or `Mathematica.`

1.6 Discrete Wavelet Transformations

Discrete wavelet transformations (DWT) are applied to the discrete data sets to produce discrete outputs. Transforming signals and data vectors by DWT is a process that resembles the fast Fourier transformation (FFT), the Fourier method applied to a set of discrete measurements.

Table 3: The analogy between Fourier and wavelet methods

Fourier Methods	Fourier Integrals	Fourier Series	Discrete Fourier Transformations
Wavelet Methods	Continuous Wavelet Transformations	Wavelet Series	Discrete Wavelet Transformations

Discrete wavelet transformations map data from the time domain (the original or input data, signal vector) to the wavelet domain. The result is a vector of the same size. Wavelet transformations are linear and they can be defined by matrices of dimension $n \times n$ if they are applied to inputs of size n. Depending on boundary conditions, such matrices can be either orthogonal or "close" to orthogonal. When the matrix is orthogonal, the corresponding transformation is a rotation in $\mathbb{R}^n$ space in which the signal vectors represent coordinates of a single point. The coordinates of the point in the new, rotated space comprise the discrete wavelet transformation of the original coordinates.

Example 1.6 Let the vector be $\{1, 2\}$ and let $M(1, 2)$ be the point in $\mathbb{R}^2$ with coordinates given by the data vector. The rotation of the coordinate axes by an angle of $\frac{\pi}{4}$ can be interpreted as a DWT in the Haar wavelet basis. The rotation matrix is

$$W = \begin{pmatrix} \cos\frac{\pi}{4} & \sin\frac{\pi}{4} \\ \cos\frac{\pi}{4} & -\sin\frac{\pi}{4} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix},$$

and the discrete wavelet transformation of $(1,2)'$ is $W \cdot (1,2)' = (\frac{3}{\sqrt{2}}, -\frac{1}{\sqrt{2}})'$. Notice that *the energy* (squared distance of the point from the origin) is preserved, $1^2 + 2^2 = (\frac{1}{2})^2 + (\frac{\sqrt{3}}{2})^2$, since W is a rotation.

Example 1.7 Let $\mathbf{y} = (1, 0, -3, 2, 1, 0, 1, 2)$. If Haar wavelet is used, the values $f(n) = y_n,\ n = 0, 1, \ldots, 7$ are interpolated by the father wavelet, the vector represent the sampled piecewise constant function. It is obvious that such defined f belongs to Haar's multiresolution space V_0.

The following matrix equation gives the connection between $\mathbf{y}$ and the wavelet coefficients (data in the wavelet domain).

$$\begin{bmatrix} 1 \\ 0 \\ -3 \\ 2 \\ 1 \\ 0 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2} & 0 & \frac{1}{\sqrt{2}} & 0 & 0 & 0 \\ \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2} & 0 & -\frac{1}{\sqrt{2}} & 0 & 0 & 0 \\ \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & -\frac{1}{2} & 0 & 0 & \frac{1}{\sqrt{2}} & 0 & 0 \\ \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & -\frac{1}{2} & 0 & 0 & -\frac{1}{\sqrt{2}} & 0 & 0 \\ \frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} & 0 & \frac{1}{2} & 0 & 0 & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} & 0 & \frac{1}{2} & 0 & 0 & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} & 0 & -\frac{1}{2} & 0 & 0 & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} & 0 & -\frac{1}{2} & 0 & 0 & 0 & -\frac{1}{\sqrt{2}} \end{bmatrix} \cdot \begin{bmatrix} c_{00} \\ d_{00} \\ d_{10} \\ d_{11} \\ d_{20} \\ d_{21} \\ d_{22} \\ d_{23} \end{bmatrix}.$$

The solution is

$$\begin{bmatrix} c_{00} \\ d_{00} \\ d_{10} \\ d_{11} \\ d_{20} \\ d_{21} \\ d_{22} \\ d_{23} \end{bmatrix} = \begin{bmatrix} \sqrt{2} \\ -\sqrt{2} \\ 1 \\ -1 \\ \frac{1}{\sqrt{2}} \\ -\frac{5}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}.$$

Thus,

$$\begin{aligned} f &= \sqrt{2}\phi_{-3,0} - \sqrt{2}\psi_{-3,0} + \psi_{-2,0} - \psi_{-2,1} \\ &\quad + \frac{1}{\sqrt{2}}\psi_{-1,0} - \frac{5}{\sqrt{2}}\psi_{-1,1} + \frac{1}{\sqrt{2}}\psi_{-1,2} - \frac{1}{\sqrt{2}}\psi_{-1,3}. \end{aligned} \tag{52}$$

The solution is easy to verify. For example, when $x \in [0, 1)$,

$$f(x) = \sqrt{2} \cdot \frac{1}{2\sqrt{2}} - \sqrt{2} \cdot \frac{1}{2\sqrt{2}} + 1 \cdot \frac{1}{2} + \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} = 1/2 + 1/2 = 1\ (= y_0).$$

Performing wavelet transformations by multiplying the input vector with an appropriate orthogonal matrix is conceptually straightforward, but of limited practical value. Storing and manipulating transformation matrices when inputs are long (> 2000) may not even be feasible.

In the context of image processing, Burt and Adelson (1982a,b) developed orthogonal and biorthogonal pyramid algorithms. Pyramid or cascade procedures process an image at different scales, ranging from fine to coarse, in a tree-like algorithm. The images can be denoised, enhanced or compressed by appropriate scale-wise treatments.

Mallat (1989a,b) was the first to link wavelets, multiresolution analyses and cascade algorithms in a formal way. Mallat's cascade algorithm gives a constructive and efficient recipe for performing the discrete wavelet transformation. It relates the wavelet coefficients from different levels in the transformation by filtering with $\mathbf{h}$ and $\mathbf{g}$. Mallat's algorithm can be viewed as a wavelet counterpart of Danielson-Lanczos algorithm in fast Fourier transformations.

It is convenient to link the original signal with the space coefficients from the space V_J, for some J. Such link is exact for interpolating wavelets (Haar, Shannon, some biorthogonal and halfband-filter wavelets) and close to exact for other wavelets, notably coiflets. Then, coarser smooth and complementing detail spaces are (V_{J-1}, W_{J-1}), (V_{J-2}, W_{J-2}), etc. Decreasing the index in V-spaces is equivalent to coarsening the approximation to the data.

By a straightforward substitution of indices in the scaling equations (10) and (25), one obtains

$$\phi_{j-1,l}(x) = \sum_{k \in \mathbb{Z}} h_{k-2l} \phi_{jk}(x) \;\text{ and }\; \psi_{j-1,l}(x) = \sum_{k \in \mathbb{Z}} g_{k-2l} \phi_{jk}(x). \tag{53}$$

The relations in (53) are fundamental in developing the cascade algorithm.

Consider a multiresolution analysis $\cdots \subset V_{j-1} \subset V_j \subset V_{j+1} \subset \ldots$. Since $V_j = V_{j-1} \oplus W_{j-1}$, any function $v_j \in V_j$ can be represented uniquely as $v_j(x) = v_{j-1}(x) + w_{j-1}(x)$, where $v_{j-1} \in V_{j-1}$ and $w_{j-1} \in W_{j-1}$. It is customary to denote the coefficients associated with $\phi_{jk}(x)$ and $\psi_{jk}(x)$ by c_{jk} and d_{jk}, respectively.

Thus,

$$\begin{aligned} v_j(x) &= \sum_k c_{j,k} \phi_{j,k}(x) \\ &= \sum_l c_{j-1,l} \phi_{j-1,l}(x) + \sum_l d_{j-1,l} \psi_{j-1,l}(x) \\ &= v_{j-1}(x) + w_{j-1}(x). \end{aligned}$$

By using the general scaling equations (53), orthogonality of $w_{j-1}(x)$ and $\phi_{j-1,l}(x)$ for any j and l, and additivity of inner products, we obtain

$$\begin{aligned} c_{j-1,l} &= \langle v_j, \phi_{j-1,l} \rangle \\ &= \langle v_j, \sum_k h_{k-2l} \phi_{j,k} \rangle \\ &= \sum_k h_{k-2l} \langle v_j, \phi_{j,k} \rangle \\ &= \sum_k h_{k-2l} c_{j,k}. \end{aligned} \tag{54}$$

Similarly $d_{j-1,l} = \sum_k g_{k-2l} c_{j,k}$.

The cascade algorithm works in the reverse direction as well. Coefficients in the next finer scale corresponding to V_j can be obtained from the coefficients corresponding to V_{j-1} and W_{j-1}. The relation

$$\begin{aligned} c_{j,k} &= \langle v_j, \phi_{j,k} \rangle \\ &= \sum_l c_{j-1,l} \langle \phi_{j-1,l}, \phi_{j,k} \rangle + \sum_l d_{j-1,l} \langle \psi_{j-1,l}, \phi_{j,k} \rangle \\ &= \sum_l c_{j-1,l} h_{k-2l} + \sum_l d_{j-1,l} g_{k-2l}, \end{aligned} \tag{55}$$

describes a single step in the reconstruction algorithm.

Example 1.8 For DAUB2, the scaling equation at integers is

$$\phi(n) \;=\; \sum_{k=0}^{3} h_k\sqrt{2}\phi(2n-k).$$

Recall that $\mathbf{h} = \{h_0, h_1, h_2, h_3\} = \{\frac{1+\sqrt{3}}{4\sqrt{2}}, \frac{3-\sqrt{3}}{4\sqrt{2}}, \frac{3+\sqrt{3}}{4\sqrt{2}}, \frac{1-\sqrt{3}}{4\sqrt{2}}\}$.

Since $\phi(0) = \sqrt{2}h_0\phi(0)$ and $\sqrt{2}h_0 \neq 1$, it follows that $\phi(0) = 0$. Also, $\phi(3) = 0$. For $\phi(1)$ and $\phi(2)$ we obtain the system

$$\begin{bmatrix} \phi(1) \\ \phi(2) \end{bmatrix} = \sqrt{2}\cdot \begin{bmatrix} h_1 & h_0 \\ h_3 & h_2 \end{bmatrix} \cdot \begin{bmatrix} \phi(1) \\ \phi(2) \end{bmatrix}.$$

From $\sum_k \phi(x-k) = 1$ it follows that $\phi(1) + \phi(2) = 1$. Solving for $\phi(1)$ and $\phi(2)$ we obtain

$$\phi(1) = \frac{1+\sqrt{3}}{2} \text{ and } \phi(2) = \frac{1-\sqrt{3}}{2}.$$

Now, one can refine ϕ,

$$\begin{aligned}
\phi\left(\frac{1}{2}\right) &= \sum_k h_k\sqrt{2}\phi(1-k) = h_0\sqrt{2}\phi(1) = \frac{2+\sqrt{3}}{4}, \\
\phi\left(\frac{3}{2}\right) &= \sum_k h_k\sqrt{2}\phi(3-k) = h_1\sqrt{2}\phi(2) + h_2\sqrt{2}\phi(1) \\
&= \frac{3+\sqrt{3}}{4}\cdot\frac{1-\sqrt{3}}{2} + \frac{3-\sqrt{3}}{4}\cdot\frac{1+\sqrt{3}}{2} = 0, \\
\phi\left(\frac{5}{2}\right) &= \sum_k h_k\sqrt{2}\phi(5-k) = h_3\sqrt{2}\phi(2) = \frac{2-\sqrt{3}}{4},
\end{aligned}$$

or ψ,

$$\begin{aligned}
\psi(-1) &= \psi(2) = 0, \\
\psi\left(-\frac{1}{2}\right) &= \sum_k g_k\sqrt{2}\phi(-1-k) = h_1\sqrt{2}\phi(1) = -\frac{1}{4}, \qquad [g_n = (-1)^n h_{1-n}] \\
\psi(0) &= \sum_k g_k\sqrt{2}\phi(0-k) = g_{-2}\sqrt{2}\phi(2) + g_{-1}\sqrt{2}\phi(1) \\
&= -h_2\sqrt{2}\phi(1) = -\frac{\sqrt{3}}{4},
\end{aligned}$$

etc.

1.6.1 Discrete Wavelet Transformations as Linear Transformations

The change of basis in V_1 from $\mathcal{B}_1 = \{\phi_{1k}(x), k \in Z\}$ to $\mathcal{B}_2 = \{\phi_{0k}, k \in Z\} \cup \{\psi_{0k}, k \in Z\}$ can be performed by matrix multiplication, therefore, it is possible to define discrete wavelet transformation by matrices. We have already seen a transformation matrix corresponding to Haar's inverse transformation in Example 1.7.

Let the length of the input signal be 2^J, and let $\mathbf{h} = \{h_s,\ s \in \mathbb{Z}\}$ be the wavelet filter and let N be an appropriately chosen constant.

Denote by H_k is a matrix of size $(2^{J-k} \times 2^{J-k+1}),\ k = 1, \dots$ with entries

$$h_s,\ \ s = (N-1) + (i-1) - 2(j-1) \text{ modulo } 2^{J-k+1}, \tag{56}$$

at the position (i, j).

Note that H_k is a circulant matrix, its ith row is 1st row circularly shifted to the right by $2(i-1)$ units. This circularity is a consequence of using the *modulo* operator in (56).

By analogy, define a matrix G_k by using the filter $\mathbf{g}$. A version of G_k corresponding to the already defined H_k can be obtained by changing h_i by $(-1)^i h_{N+1-i}$. The constant N is a shift parameter and affects the position of the wavelet on the time scale. For filters from the Daubechies family, standard choice for N is the number of vanishing moments. See also Remark 1.2.

The matrix $\begin{bmatrix} H_k \\ G_k \end{bmatrix}$ is a basis-change matrix in 2^{J-k+1} dimensional space; consequently, it is unitary.

Therefore,

$$I_{2^{J-k}} = [H_k'\ G_k'] \begin{bmatrix} H_k \\ G_k \end{bmatrix} = H_k' \cdot H_k + G_k' \cdot G_k.$$

and

$$I = \begin{bmatrix} H_k \\ G_k \end{bmatrix} \cdot [H_k'\ G_k'] = \begin{bmatrix} H_k \cdot H_k' & H_k \cdot G_k' \\ G_k \cdot H_k' & G_k \cdot G_k' \end{bmatrix}.$$

That implies,

$$H_k \cdot H_k' = I,\ G_k \cdot G_k' = I,\ G_k \cdot H_k' = H_k \cdot G_k' = 0, \text{ and } H_k' \cdot H_k + G_k' \cdot G_k = I.$$

Now, for a sequence y the J-step wavelet transformation is $\mathbf{d} = W_J \cdot \mathbf{y}$, where

$$W_1 = \begin{bmatrix} H_1 \\ G_1 \end{bmatrix},\ \ W_2 = \begin{bmatrix} \begin{bmatrix} H_2 \\ G_2 \end{bmatrix} \cdot H_1 \\ G_1 \end{bmatrix},$$

$$W_3 = \begin{bmatrix} \begin{bmatrix} \begin{bmatrix} H_3 \\ G_3 \end{bmatrix} \cdot H_2 \\ G_2 \end{bmatrix} \cdot H_1 \\ G_1 \end{bmatrix}, \dots$$

Example 1.9 Suppose that $\mathbf{y} = \{1, 0, -3, 2, 1, 0, 1, 2\}$ and filter is $\mathbf{h} = (h_0, h_1, h_2, h_3) = \left(\frac{1+\sqrt{3}}{4\sqrt{2}}, \frac{3+\sqrt{3}}{4\sqrt{2}}, \frac{3-\sqrt{3}}{4\sqrt{2}}, \frac{1-\sqrt{3}}{4\sqrt{2}}\right)$. Then, $J = 3$ and matrices H_k and G_k are of dimension $2^{3-k} \times 2^{3-k+1}$.

$$H_1 = \begin{bmatrix} h_1 & h_2 & h_3 & 0 & 0 & 0 & 0 & h_0 \\ 0 & h_0 & h_1 & h_2 & h_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & h_0 & h_1 & h_2 & h_3 & 0 \\ h_3 & 0 & 0 & 0 & 0 & h_0 & h_1 & h_2 \end{bmatrix}$$

$$G_1 = \begin{bmatrix} -h_2 & h_1 & -h_0 & 0 & 0 & 0 & 0 & h_3 \\ 0 & h_3 & -h_2 & h_1 & -h_0 & 0 & 0 & 0 \\ 0 & 0 & 0 & h_3 & -h_2 & h_1 & -h_0 & 0 \\ -h_0 & 0 & 0 & 0 & 0 & h_3 & -h_2 & h_1 \end{bmatrix}.$$

Since,

$$\begin{aligned} H_1 \cdot \mathbf{y} &= \{2.19067, -2.19067, 1.67303, 1.15539\} \\ G_1 \cdot \mathbf{y} &= \{0.96593, 1.86250, -0.96593, 0.96593\}. \end{aligned}$$

$W_1\mathbf{y} = \{2.19067, -2.19067, 1.67303, 1.15539 \mid 0.96593, 1.86250, -0.96593, 0.96593\}.$

$$H_2 = \begin{bmatrix} h_1 & h_2 & h_3 & h_0 \\ h_3 & h_0 & h_1 & h_2 \end{bmatrix} \qquad G_2 = \begin{bmatrix} -h_2 & h_1 & -h_0 & h_3 \\ -h_0 & h_3 & -h_2 & h_1 \end{bmatrix}.$$

In this example, due to lengths of the filter and data, we can perform discrete wavelet transformation for two steps only, W_1 and W_2.

The two-step DAUB2 discrete wavelet transformation of $\mathbf{y}$ is
$W_2 \cdot \mathbf{y} = \{1.68301, 0.31699 \mid -3.28109, -0.18301 \mid 0.96593, 1.86250, -0.96593, 0.96593\}$, because

$$\begin{aligned} H_2 \cdot H_1 \cdot \mathbf{y} &= H_2 \cdot \{2.19067, -2.19067, 1.67303, 1.15539\} \\ &= \{1.68301, 0.31699\} \\ G_2 \cdot H_1 \cdot \mathbf{y} &= G_1 \cdot \{2.19067, -2.19067, 1.67303, 1.15539\} \\ &= \{-3.28109, -0.18301\}. \end{aligned}$$

For quadrature mirror wavelet filters $\mathbf{h}$ and $\mathbf{g}$, we define recursively up-sampled filters $\mathbf{h}^{[r]}$ and $\mathbf{g}^{[r]}$

$$\begin{aligned} \mathbf{h}^{[0]} &= \mathbf{h}, \ \ \mathbf{g}^{[0]} = \mathbf{g} \\ \mathbf{h}^{[r]} &= [\uparrow 2] \ \ \mathbf{h}^{[r-1]}, \ \ \ \mathbf{g}^{[r]} = [\uparrow 2] \ \ \mathbf{g}^{[r-1]}. \end{aligned}$$

In practice, the dilated filter $\mathbf{h}^{[r]}$ is obtained by inserting zeroes between the taps in $\mathbf{h}^{[r-1]}$. Let $\mathbf{H}^{[r]}$ and $\mathbf{G}^{[r]}$ be convolution operators with filters $\mathbf{h}^{[r]}$ and $\mathbf{h}^{[r]}$, respectively. A non-decimated wavelet transformation, NDWT, is defined as a sequential application of operators (convolutions) $\mathbf{H}^{[j]}$ and $\mathbf{G}^{[j]}$ on a given time series.

Definition 1.2 *Let* $\mathbf{a}^{(J)} = \mathbf{c}^{(J)}$ *and*

$$\begin{aligned} \mathbf{a}^{(j-1)} &= \mathbf{H}^{[J-j]}\mathbf{a}^{(j)}, \\ \mathbf{b}^{(j-1)} &= \mathbf{G}^{[J-j]}\mathbf{a}^{(j)}. \end{aligned}$$

The non-decimated wavelet transformation of $\mathbf{c}^{(J)}$ *is* $\mathbf{b}^{(J-1)}$, $\mathbf{b}^{(J-2)}$, $\dots$, $\mathbf{b}^{(J-j)}$, $\mathbf{a}^{(J-j)}$, *for some* $j \in \{1, 2, \dots, J\}$ *the depth of the transformation.*

If the length of an input vector $\mathbf{c}^{(J)}$ is 2^J, then for any $0 \le m < J$, $\mathbf{a}^{(m)}$ and $\mathbf{b}^{(m)}$ are of the same length. Let $\phi_j(x) = \phi_{j,0}(x)$ and $\psi_j(x) = \psi_{j,0}(s)$. If the measurement sequence $\mathbf{c}^{(J)}$ is associated with the function $f(x) = \sum_k c_k^{(J)} \phi_J(x - 2^{-J}k)$ then the kth coordinate of $\mathbf{b}^{(j)}$ is equal to

$$b_{jk} = \int \psi_j(x - 2^{-J}k) f(x) dx.$$

Thus, the coefficient b_{jk} provides information at scale 2^{J-j} and location k. One can think of a nondecimated wavelet transformation as sampled continuous wavelet transformation $\langle f(x), \frac{1}{\sqrt{a}}\psi\left(\frac{x-b}{a}\right)\rangle$ for $a = 2^{-j}$, and $b = k$.