Mathematical Foundations of Data Sciences

Gabriel Peyré
CNRS & DMA, École Normale Supérieure

CHAPTER 4

Wavelets

September 9, 2026

Signals contain broad trends and localized details that occur at different scales. Separating these components produces representations suited to compression and denoising, while allowing each scale to be processed efficiently. We build orthonormal wavelet bases from nested approximation spaces, derive fast decomposition and reconstruction algorithms in one and two dimensions, and explain how filter design controls localization, smoothness, and cancellation of polynomials.

The reference for this chapter is [1].

4.1 Multiresolution Approximation Spaces

A multiresolution approximation of $L^2(\mathbb{R})$ consists of nested closed subspaces $(V_j)_j$

$$L^2(\mathbb{R}) \supset \ldots \supset V_{j-1} \supset V_j \supset V_{j+1} \supset \ldots \supset \{0\} \tag{4.1}$$

related by dyadic scaling and invariant under translations on the corresponding dyadic grid:

$$f \in V_j \quad \Longleftrightarrow \quad f(\cdot/2) \in V_{j+1} \quad \text{and} \quad f \in V_j \quad \Longleftrightarrow \quad \forall\, n \in \mathbb{Z},\ f(\cdot + n2^j) \in V_j$$

Large j corresponds to coarse approximation spaces, and 2^j is often called the "scale".

At the fine-scale limit on the left of (4.1), $\cup_j V_j$ is dense in $L^2(\mathbb{R})$. Equivalently, $P_{V_j}(f) \longrightarrow f$ as $j \longrightarrow -\infty$, where P_V denotes orthogonal projection onto V:

$$P_V(f) = \underset{f' \in V}{\operatorname{argmin}}\ \|f - f'\|.$$

The limit on the right of (4.1) means that $\cap_j V_j = \{0\}$, or equivalently that $P_{V_j}(f) \longrightarrow 0$ as $j \longrightarrow +\infty$.

The first example consists of piecewise constant functions on dyadic intervals

$$V_j = \left\{ f \in L^2(\mathbb{R}) \ ; \ \forall n, f \text{ is constant on } [2^j n, 2^j(n+1)[\right\}, \tag{4.2}$$

A second example is the space used for Shannon interpolation of bandlimited signals

$$V_j = \left\{ f \ ; \ \operatorname{Supp}(\hat f) \subset [-2^{-j}\pi, 2^{-j}\pi] \right\} \tag{4.3}$$

whose elements are recovered exactly from the samples $f(2^j n)$. As for piecewise constant signals, the sampling grid has spacing 2^j. Here, however, orthogonal projection is a bandlimiting operation:

$$P_{V_j}(f) = \mathcal{F}^{-1}(\hat f \odot 1_{[-2^{-j}\pi, 2^{-j}\pi]}).$$

Scaling functions. We also require a scaling function $\varphi \in L^2(\mathbb{R})$ such that

$$\{\varphi(\cdot - n)\}_n \text{ is a Hilbertian orthonormal basis of } V_0.$$

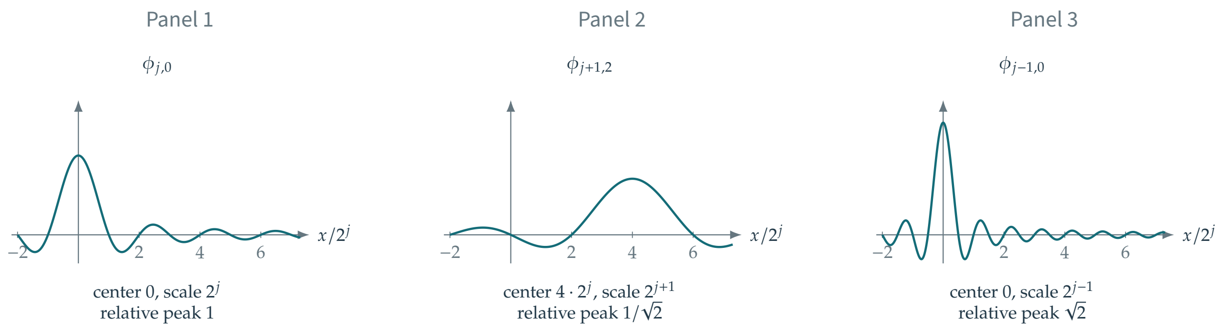


Figure 4.1. Translations and dilations of a scaling function generate the approximation spaces. The Shannon scaling function $\varphi(t) = \sin(\pi t)/(\pi t)$ is illustrated, with horizontal coordinate $x/2^j$ and heights relative to $2^{-j/2}$.

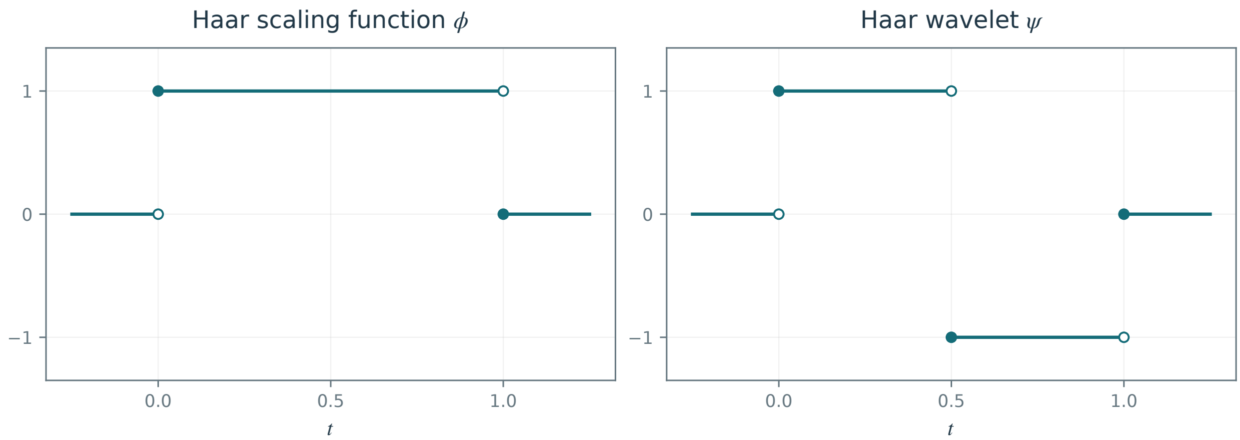


Figure 4.2. Haar scaling and wavelet functions. Both have amplitude $2^{-j/2}$ and unit L^2 norm.

By the dilation property, this implies that

$$\{\varphi_{j,n}\}_n \text{ is a Hilbertian orthonormal basis of } V_j \quad \text{where} \quad \varphi_{j,n} \stackrel{\text{def.}}{=} \frac{1}{2^{j/2}}\varphi\left(\frac{\cdot - 2^j n}{2^j}\right).$$

The normalization ensures $\|\varphi_{j,n}\| = 1$.

Note that one then has

$$P_{V_j}(f) = \sum_n \langle f, \varphi_{j,n}\rangle \varphi_{j,n}.$$

Figure 4.1 illustrates the effects of translation and scaling.

For the case of piecewise constant signals (4.2), one can use

$$\varphi = 1_{[0,1[} \quad \text{and} \quad \varphi_{j,n} = 2^{-j/2} 1_{[2^j n, 2^j(n+1)[}.$$

For the case of Shannon multiresolution (4.3), one can use $\varphi(t) = \sin(\pi t)/(\pi t)$ and one verifies

$$\langle f, \varphi_{j,n}\rangle = 2^{j/2} f(2^j n) \quad (f \in V_j) \quad \text{and} \quad f = \sum_n 2^{j/2} f(2^j n)\varphi_{j,n} \quad (f \in V_j).$$

Spectral orthogonalization. In many applications, V_0 is generated by translates of a function that are not orthogonal: $V_0 = \overline{\text{Span}}\{\theta(\cdot - n)\}_{n\in\mathbb{Z}}$. The next proposition orthogonalizes this family in the Fourier domain. Figure 4.3 shows the orthogonal cardinal spline functions obtained by this construction.

Proposition 4.1. *Let $\theta \in L^2(\mathbb{R})$. Its translates $\{\theta(\cdot - n)\}_{n\in\mathbb{Z}}$ are orthonormal if and only if*

$$A(\omega) \stackrel{\text{def.}}{=} \sum_k |\hat{\theta}(\omega - 2\pi k)|^2 = 1 \quad \textit{for almost every } \omega.$$

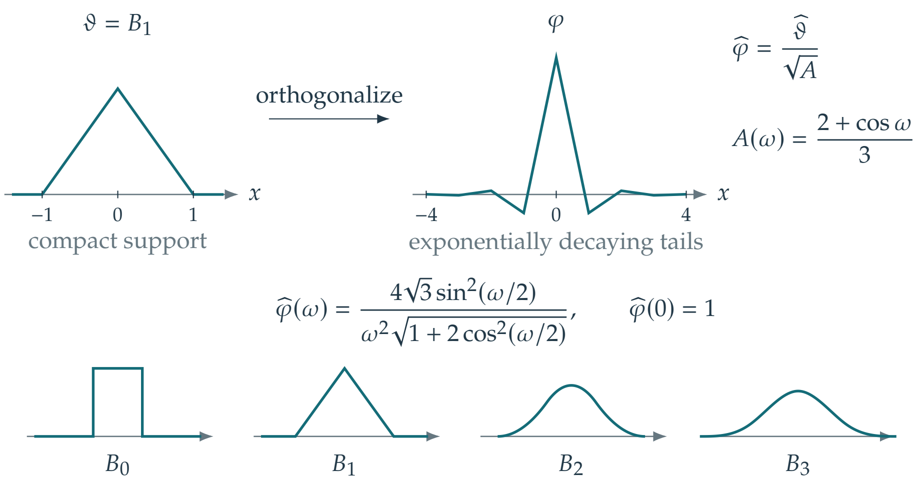


Figure 4.3. Orthogonalization of B-splines to define cardinal orthogonal spline functions.

If there exist $0 < a \leqslant b < +\infty$ *such that* $a \leqslant A(\omega) \leqslant b$ *almost everywhere, then* φ *defined by*

$$\hat{\varphi}(\omega) = \frac{\hat{\theta}(\omega)}{\sqrt{A(\omega)}}$$

is such that $\{\varphi(\cdot - n)\}_{n\in\mathbb{Z}}$ *is an orthonormal basis of* $\overline{\mathrm{Span}}\{\theta(\cdot - n)\}_{n\in\mathbb{Z}}$.

Proof. The family $\{\theta(\cdot - n)\}_{n\in\mathbb{Z}}$ is orthonormal if and only if

$$\langle \theta,\, \theta(\cdot - n)\rangle = (\theta \star \bar{\theta})(n) = \delta_{0,n} \overset{\text{def.}}{=} \begin{cases} 1 & \text{if} \quad n = 0, \\ 0 & \text{otherwise.} \end{cases}$$

where $\bar{\theta}(x) = \overline{\theta(-x)}$ and δ_0 is the discrete Dirac vector. By Tonelli's theorem and Plancherel's identity, $A \in L^1(\mathbb{T})$. Its Fourier coefficients are

$$\begin{aligned} \frac{1}{2\pi}\int_0^{2\pi} A(\omega)e^{-in\omega}\,\mathrm{d}\omega &= \frac{1}{2\pi}\int_{\mathbb{R}} |\hat{\theta}(\omega)|^2 e^{-in\omega}\,\mathrm{d}\omega \\ &= \langle \theta,\, \theta(\cdot + n)\rangle. \end{aligned}$$

Thus the translates are orthonormal exactly when the Fourier coefficients of A are those of the constant function 1. Uniqueness of Fourier coefficients in $L^1(\mathbb{T})$ gives $A = 1$ almost everywhere. Since A is 2π-periodic, normalization by the bounded function $1/\sqrt{A(\omega)}$ gives $\sum_k |\hat{\varphi}(\omega - 2\pi k)|^2 = 1$. It remains to check equality of the closed spans. Approximate the periodic multipliers $1/\sqrt{A}$ and $\sqrt{A}$ by trigonometric polynomials in $L^2(\mathbb{T})$. The bounds $a \leqslant A \leqslant b$ show that the corresponding products with $\hat{\theta}$ and $\hat{\varphi}$ converge in $L^2(\mathbb{R})$. Inverting the Fourier transform expresses each generator as an L^2 limit of finite linear combinations of translates of the other. □

Spline interpolation provides a typical application. For example, cubic splines are generated by translates of a box-spline function θ, which is a compactly supported piecewise polynomial.

4.2 Multiresolution Detail Spaces

The detail spaces are the orthogonal complements associated with the inclusions $V_j \subset V_{j-1}$. Since these subspaces are closed,

$$\forall\, j, \quad W_j \text{ is such that } V_{j-1} = V_j \oplus^{\perp} W_j.$$

The following diagram shows these nested approximation spaces and their orthogonal complements:

$$\begin{tikzcd}
\cdots \arrow[r] \arrow[dr] & V_{j-1} \arrow[r] \arrow[dr] & V_j \arrow[r] \arrow[dr] & V_{j+1} \arrow[r] & \cdots \\
& W_{j-1} & W_j & W_{j+1} &
\end{tikzcd}$$

$$V_{j-1} = V_j \oplus^{\perp} W_j$$

Suppose that W_0 admits an orthonormal basis of translates $\{\psi(\cdot - n)\}_n$. Scaling then gives

$$\{\psi_{j,n}\}_n \text{ is a Hilbertian orthonormal basis of } W_j \quad \text{where} \quad \psi_{j,n} \overset{\text{def.}}{=} \frac{1}{2^{j/2}}\psi\left(\frac{\cdot - 2^j n}{2^j}\right).$$

Orthogonal decomposition across all scales yields

$$L^2(\mathbb{R}) = \bigoplus_{j=-\infty}^{j=+\infty} W_j = V_{j_0} \oplus^{\perp} \bigoplus_{j \leqslant j_0} W_j.$$

For every $f \in L^2(\mathbb{R})$, the corresponding expansion converges in $L^2(\mathbb{R})$:

$$f = \lim_{(j_-,j_+)\to(-\infty,+\infty)} \sum_{j=j_-}^{j_+} P_{W_j} f = \lim_{j_- \to -\infty} \left(P_{V_{j_0}} f + \sum_{j=j_-}^{j_0} P_{W_j} f \right).$$

This decomposition shows that

$$\left\{\psi_{j,n} \; ; \; (j,n) \in \mathbb{Z}^2\right\}$$

is an orthonormal basis of $L^2(\mathbb{R})$, which is called a wavelet basis. Stopping at a coarsest scale gives the orthonormal basis

$$\left\{\psi_{j,n} \; ; \; j \leqslant j_0, n \in \mathbb{Z}\right\} \cup \left\{\varphi_{j_0,n} \; ; \; n \in \mathbb{Z}\right\}.$$

The forward wavelet transform computes all inner products of a function f with the elements of this basis.

Haar wavelets. For the Haar multiresolution (4.2), one has

$$W_j = \left\{ f \; ; \; \forall\, n \in \mathbb{Z}, \; f \text{ constant on } [2^{j-1}n, 2^{j-1}(n+1)) \quad \text{and} \quad \int_{n2^j}^{(n+1)2^j} f = 0 \right\}. \tag{4.4}$$

One choice of mother wavelet is

$$\psi(t) = \begin{cases} 1 & \text{for} \quad 0 \leqslant t < 1/2, \\ -1 & \text{for} \quad 1/2 \leqslant t < 1, \\ 0 & \text{otherwise}, \end{cases}$$

as shown in Figure 4.2.

Figure 4.5 illustrates how subtracting consecutive approximation-space projections isolates the details at each scale.

Shannon and splines. Figure 4.6 compares the sharp frequency partitions of Shannon wavelets with the smooth transitions of spline wavelets. Shannon wavelets can be viewed as a limiting case as the spline degree increases.

4.3 On Bounded Domains

On the periodic domain $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, we use period 1 rather than the 2π-period convention of the Fourier chapter. To obtain an orthonormal wavelet basis of $L^2(\mathbb{T})$, periodize the wavelets and restrict their translation points $2^j n$ to $[0,1]$, choosing $0 \leqslant n < 2^{-j}$. As in (1.6), the periodization of $f \in L^1(\mathbb{R})$ is

$$f^P = \sum_{n \in \mathbb{Z}} f(\cdot - n) \in L^1(\mathbb{T}).$$

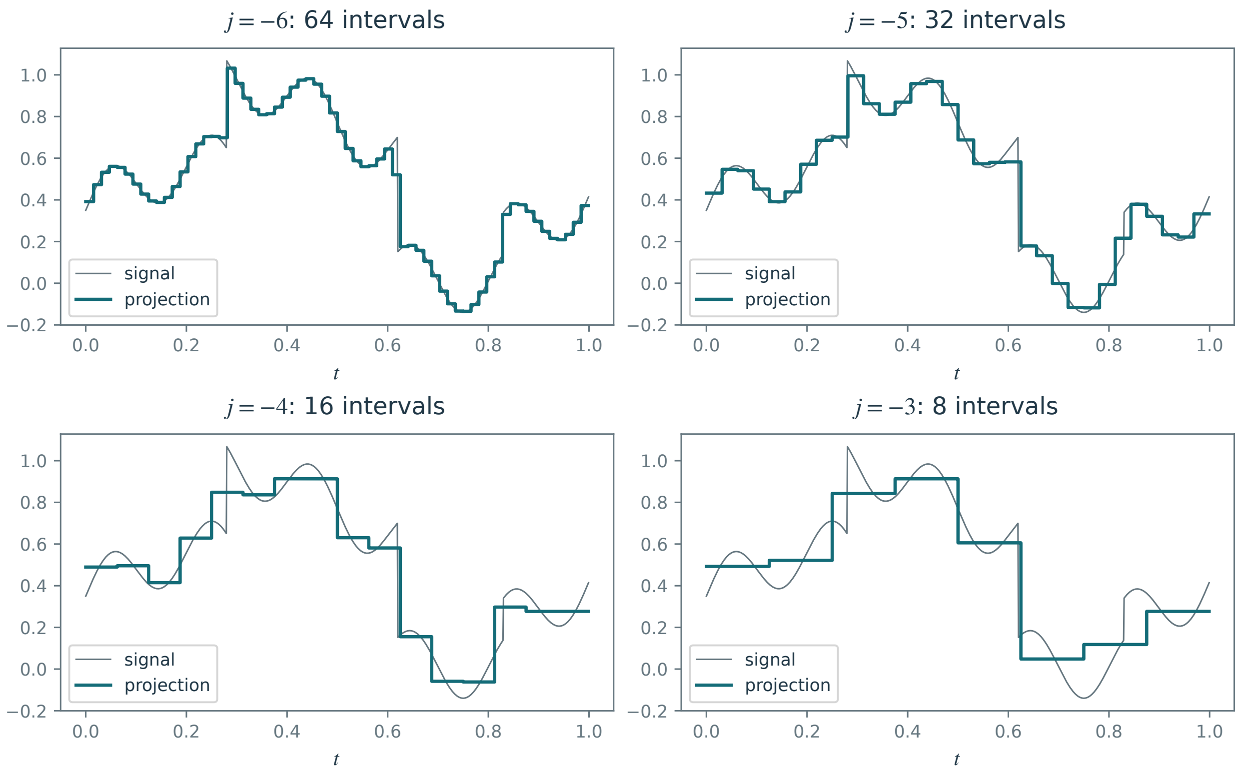


Figure 4.4. 1-D Haar multiresolution projection $P_{V_j} f$ of a function f.

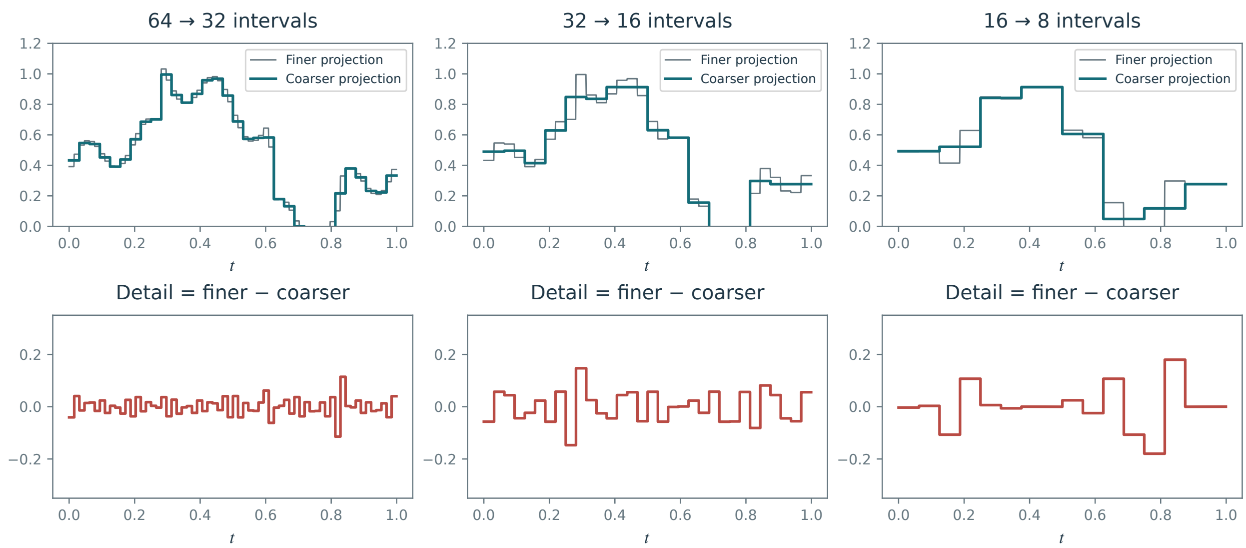


Figure 4.5. Three successive Haar approximations (top) and their extracted details (bottom). Coarser approximations use fewer intervals.

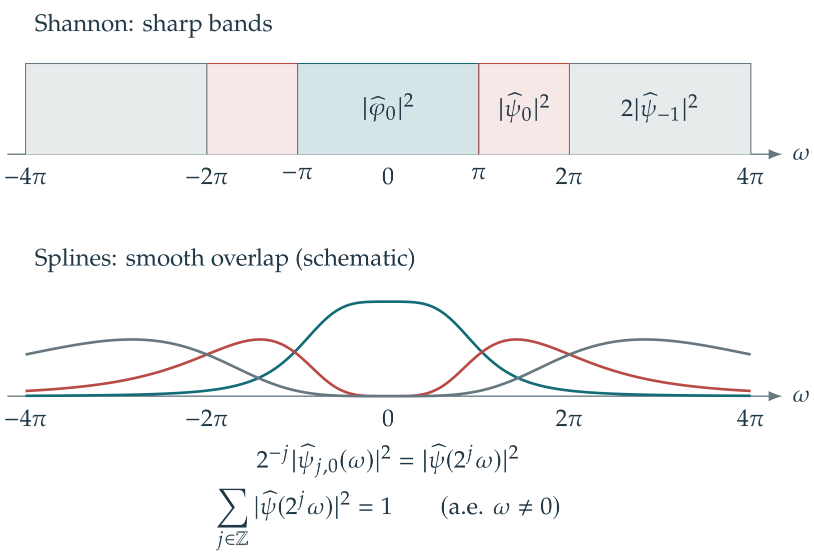


Figure 4.6. Spline and Shannon wavelet segmentation of the frequency axis.

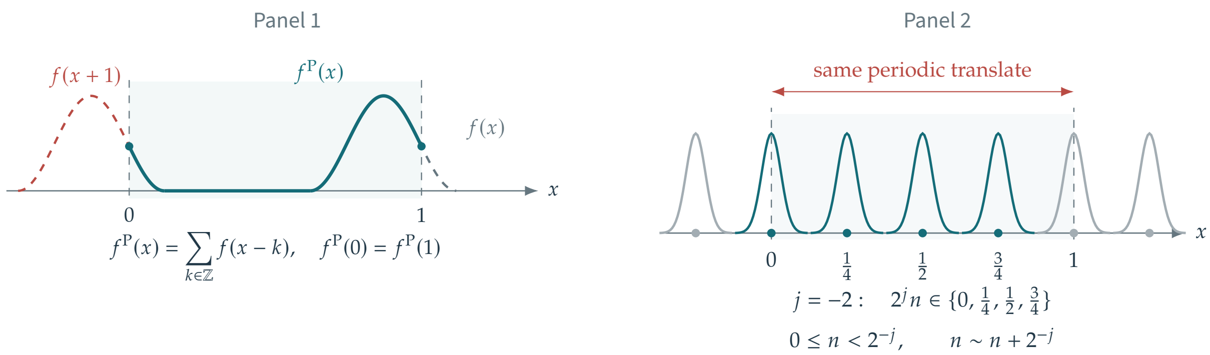


Figure 4.7. Periodized basis functions. Only the translations $0 \leqslant n < 2^{-j}$ are distinct, and $j_0 = 0$ is a common choice of coarsest scale.

For an integer coarsest level $j_0 \leqslant 0$, this gives the family

$$\left\{\psi^P_{j,n} \; ; \; j \leqslant j_0, 0 \leqslant n < 2^{-j}\right\} \cup \left\{\varphi^P_{j_0,n} \; ; \; 0 \leqslant n < 2^{-j_0}\right\}.$$

which is an orthonormal basis of $L^2(\mathbb{T})$, illustrated in Figure 4.7. One can also construct wavelet bases using Neumann (mirror) boundary conditions, but this is more involved.

4.4 Fast Wavelet Transform

4.4.1 Discretization

We now work over $\mathbb{R}/\mathbb{Z}$. We assume access to a discrete signal $a_J \in \mathbb{R}^N$ with $N = 2^{-J}$ at a fixed scale 2^J, whose entries are inner products with the scaling functions, i.e.

$$\forall\, n \in \{0, \dots, N-1\}, \quad a_{J,n} = \langle f, \varphi^P_{J,n} \rangle \approx 2^{J/2} f(2^J n), \tag{4.5}$$

for the underlying continuous function f. The approximation uses the normalization $\int \varphi = 1$. The equality assumes that acquisition gives exact access to $P_{V_J} f$, much as Shannon sampling assumes bandlimiting. This model is reasonable when the scaling functions $(\varphi_{J,n})_n$ approximate the acquisition device's point-spread function. Preprocessing the measurements can sometimes improve agreement with (4.5).

Starting from a_J, the discrete wavelet transform computes the coefficients

$$\forall\, j \in \{J+1, J+2, \dots, 0\}, \quad \forall\, n \in [\![0, 2^{-j}-1]\!], \quad a_{j,n} \overset{\text{def.}}{=} \langle f, \varphi^P_{j,n} \rangle, \quad \text{and} \quad d_{j,n} \overset{\text{def.}}{=} \langle f, \psi^P_{j,n} \rangle$$

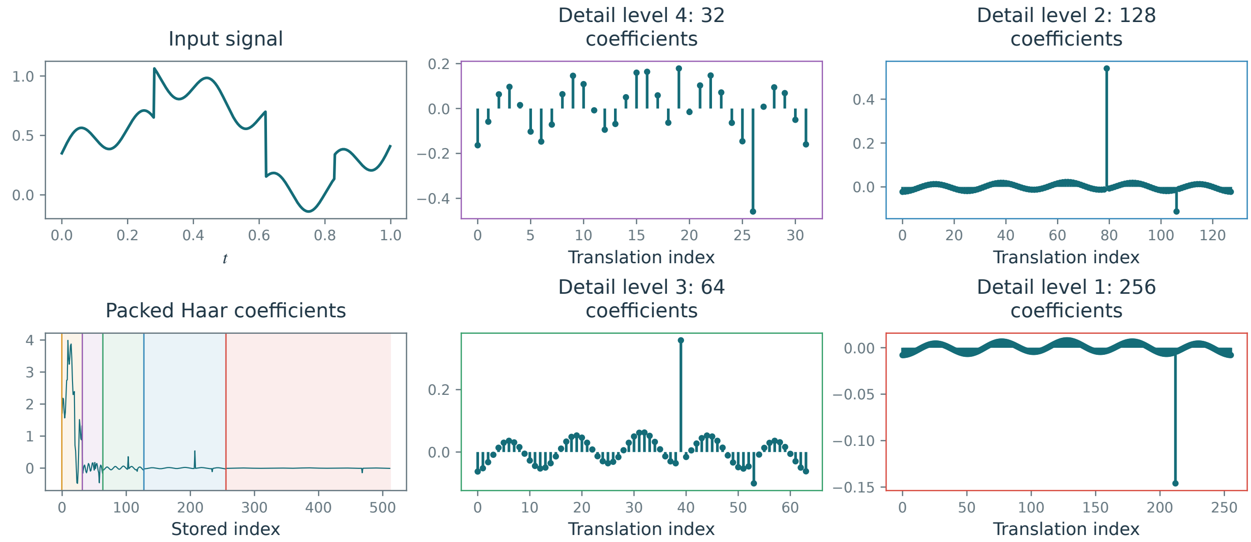


Figure 4.8. Wavelet coefficients. The first column shows the input signal and its packed decomposition; the other panels show individual detail scales.

in order of increasing j. After an approximation vector a_j has been used to compute the next scale, it can be discarded; the detail vectors d_j are retained.

The forward discrete wavelet transform on a bounded domain is thus the orthogonal finite-dimensional map

$$a_J \in \mathbb{R}^N \longmapsto \left\{d_{j,n} \; ; \; J < j \leqslant 0, \; 0 \leqslant n < 2^{-j}\right\} \cup \{a_0 \in \mathbb{R}\}.$$

By orthogonality, the inverse transform is the adjoint.

Figure 4.8 shows examples of wavelet coefficients. For each scale 2^j, there are 2^{-j} coefficients.

4.4.2 Forward Fast Wavelet Transform (FWT)

The algorithm applies a sequence of elementary operators

$$\forall j = J+1, \dots, 0, \quad (a_j, d_j) = \mathcal{W}_j(a_{j-1}) \quad \text{where} \quad \mathcal{W}_j : \mathbb{R}^{2^{-j+1}} \to \mathbb{R}^{2^{-j}} \times \mathbb{R}^{2^{-j}} \tag{4.6}$$

There are $|J| = \log_2(N)$ steps. Each $\mathcal{W}_j$ changes coordinates between orthonormal bases and is therefore orthogonal.

To describe the algorithm for $\mathcal{W}_j$, we introduce the filter coefficients $h, g \in \mathbb{R}^{\mathbb{Z}}$

$$h_n \stackrel{\text{def.}}{=} \frac{1}{\sqrt{2}} \langle \varphi(\cdot/2), \, \varphi(\cdot - n) \rangle \quad \text{and} \quad g_n \stackrel{\text{def.}}{=} \frac{1}{\sqrt{2}} \langle \psi(\cdot/2), \, \varphi(\cdot - n) \rangle. \tag{4.7}$$

Figure 4.9 illustrates the computation of these weights for the Haar system.

We denote by $\downarrow_2 \colon \mathbb{R}^K \to \mathbb{R}^{K/2}$ the subsampling operator by a factor of 2, i.e.

$$u \downarrow_2 \stackrel{\text{def.}}{=} (u_0, u_2, u_4, \dots, u_{K-2}).$$

Throughout the following derivation, the scaling function, wavelet, and filters are real, and the filters decay sufficiently fast. The notation $\bar{h}_n = h_{-n}$ and $\bar{g}_n = g_{-n}$ denotes reflection. Complex filters would require conjugation as well.

Proposition 4.2. *One has*

$$\mathcal{W}_j(a_{j-1}) = ((\bar{h} \star a_{j-1}) \downarrow_2, (\bar{g} \star a_{j-1}) \downarrow_2), \tag{4.8}$$

where $\star$ *denotes periodic convolutions on* $\mathbb{R}^{2^{-j+1}}$.

Proof. We note that $\varphi(\cdot/2)$ and $\psi(\cdot/2)$ belong to V_0, so one has the decompositions

$$\frac{1}{\sqrt{2}}\varphi(t/2) = \sum_n h_n \varphi(t-n) \quad \text{and} \quad \frac{1}{\sqrt{2}}\psi(t/2) = \sum_n g_n \varphi(t-n) \tag{4.9}$$

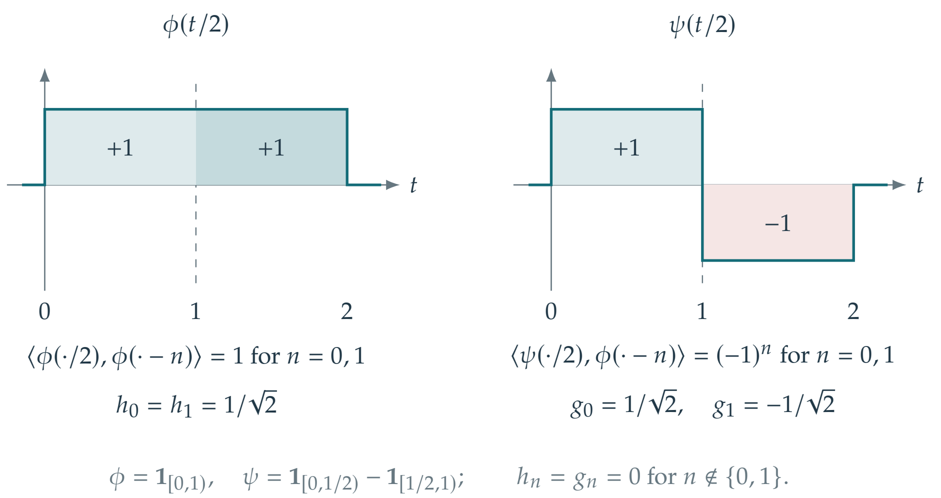


Figure 4.9. Haar filter weights as inner products.

Making the substitution $t \mapsto \frac{t-2^j p}{2^{j-1}}$ in (4.9), one obtains

$$\frac{1}{\sqrt{2}}\varphi\left(\frac{t-2^j p}{2^j}\right) = \sum_n h_n \varphi\left(\frac{t}{2^{j-1}} - (n+2p)\right)$$

(similarly for ψ) and then reindexing $n \mapsto n - 2p$, one obtains

$$\varphi_{j,p} = \sum_{n\in\mathbb{Z}} h_{n-2p}\varphi_{j-1,n} \quad \text{and} \quad \psi_{j,p} = \sum_{n\in\mathbb{Z}} g_{n-2p}\varphi_{j-1,n}.$$

For periodized functions $(\varphi^P_{j,n}, \psi^P_{j,p})$, these identities remain valid after periodizing the filters modulo 2^{-j+1}. Taking inner products with f on both sides gives the fundamental recursion below; the decay of h, g justifies exchanging the sum and inner product.

$$a_{j,p} = \sum_{n\in\mathbb{Z}} h_{n-2p} a_{j-1,n} = (\bar{h} \star a_{j-1})_{2p} \quad \text{and} \quad d_{j,p} = \sum_{n\in\mathbb{Z}} g_{n-2p} a_{j-1,n} = (\bar{g} \star a_{j-1})_{2p} \tag{4.10}$$

where $\bar{u}_n \overset{\text{def.}}{=} u_{-n}$. On the periodic domain $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, the same recursion holds with $\star$ interpreted as convolution on $\mathbb{Z}/2^{-j+1}\mathbb{Z}$. □

Figure 4.10 shows two successive decomposition steps, each separating a coarse approximation from its details.

The FWT thus operates as follows. The same real filters also act on complex-valued input signals:

- **Input:** signal $f \in \mathbb{C}^N$.
- **Initialization:** $a_J = f$.
- **For** $j = J, \dots, j_0 - 1$.

$$a_{j+1} = (a_j \star \bar{h}) \downarrow 2 \quad \text{and} \quad d_{j+1} = (a_j \star \bar{g}) \downarrow 2$$

- **Output:** the coefficients $\{d_j\}_{J<j\leqslant j_0} \cup \{a_{j_0}\}$.

If the supports of h and g each contain at most C indices, then evaluating each $\mathcal{W}_j$ requires $(2C)2^{-j}$ operations, so that the complexity of the whole wavelet transform is

$$\sum_{j=J+1}^{0} (2C)2^{-j} = 2C(2^{-J} - 1) < 2CN.$$

This shows that the fast wavelet transform is a linear-time algorithm. Figure 4.11 shows the iterative extraction of the wavelet coefficients. Figure 4.12 illustrates the storage layout: at each step, the new vectors a_j and d_j occupy the left portion of the output array.

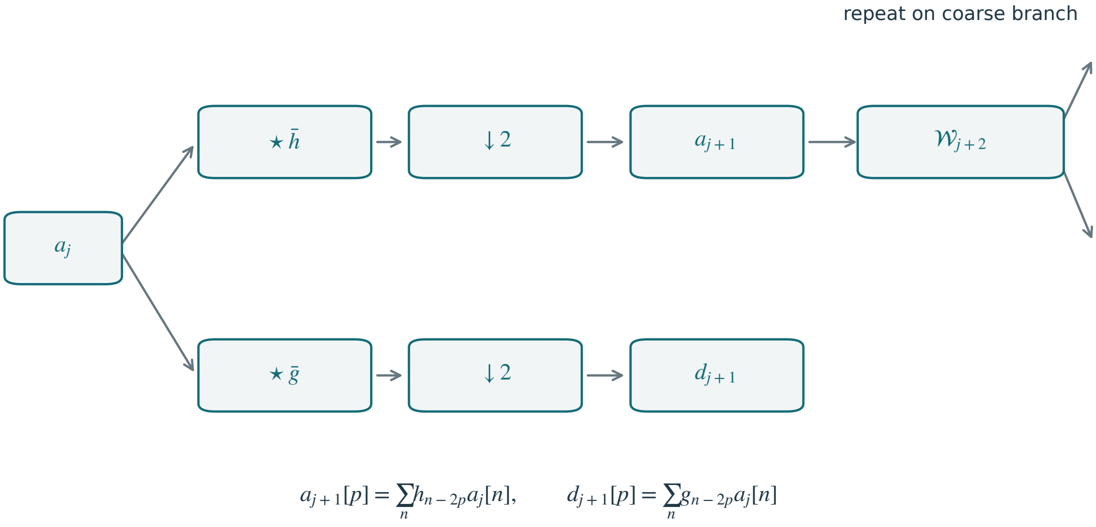


Figure 4.10. Forward filter bank decomposition.

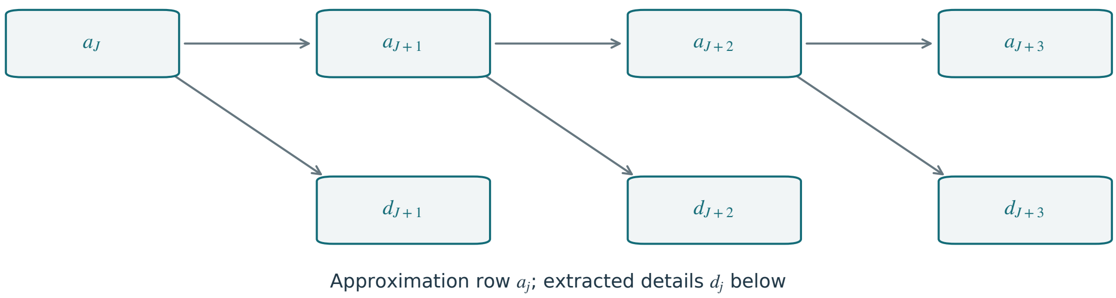


Figure 4.11. Pyramidal computation of wavelet coefficients: approximation vectors a_j on the top row and detail vectors d_j below.

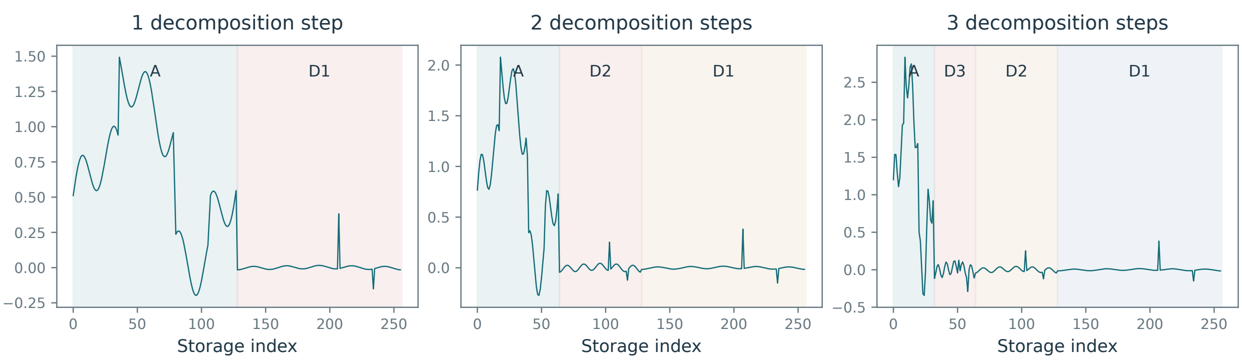


Figure 4.12. Packed coefficient arrays after one, two, and three wavelet decomposition steps.

Fast Haar transform. For the Haar wavelets, one has

$$\varphi_{j,n} = \frac{1}{\sqrt{2}}(\varphi_{j-1,2n} + \varphi_{j-1,2n+1}),$$

$$\psi_{j,n} = \frac{1}{\sqrt{2}}(\varphi_{j-1,2n} - \varphi_{j-1,2n+1}).$$

This corresponds to the filters

$$h = [\ldots,\ 0,\ h[0] = \frac{1}{\sqrt{2}},\ \frac{1}{\sqrt{2}},\ 0, \ldots],$$

$$g = [\ldots,\ 0,\ g[0] = \frac{1}{\sqrt{2}},\ -\frac{1}{\sqrt{2}},\ 0, \ldots].$$

The Haar transform iterates normalized sums and differences:

– **Input:** signal $f \in \mathbb{C}^N$.

– **Initialization:** $a_J = f$.

– **For** $j = J, \ldots, j_0 - 1$.

$$a_{j+1,n} = \frac{1}{\sqrt{2}}(a_{j,2n} + a_{j,2n+1}) \quad \text{and} \quad d_{j+1,n} = \frac{1}{\sqrt{2}}(a_{j,2n} - a_{j,2n+1}).$$

– **Output:** the coefficients $\{d_j\}_{J<j\leqslant j_0} \cup \{a_{j_0}\}$.

4.4.3 Inverse Fast Transform (iFWT)

The inverse algorithm proceeds by inverting each step (4.6)

$$\forall\, j = 0, -1, \ldots, J+1, \quad a_{j-1} = \mathcal{W}_j^{-1}(a_j, d_j) = \mathcal{W}_j^*(a_j, d_j), \tag{4.11}$$

where $\mathcal{W}_j^*$ is the adjoint for the canonical inner product on $\mathbb{R}^{2^{-j+1}}$; in matrix form, it is the transpose.

Let $\uparrow_2 \colon \mathbb{R}^{K/2} \to \mathbb{R}^K$ denote up-sampling by inserting zeros:

$$a \uparrow_2 = (a_0, 0, a_1, 0, \ldots, a_{K/2-1}, 0) \in \mathbb{R}^K.$$

Proposition 4.3. *One has*

$$\mathcal{W}_j^{-1}(a_j, d_j) = (a_j \uparrow_2) \star h + (d_j \uparrow_2) \star g.$$

Proof. Since $\mathcal{W}_j$ is orthogonal, $\mathcal{W}_j^{-1} = \mathcal{W}_j^*$. We write the whole transform as

$$\mathcal{W}_j = S_2 \circ C_{\bar{h},\bar{g}} \circ \mathcal{D} \quad \text{where} \quad \begin{cases} \mathcal{D}(a) = (a, a), \\ C_{\bar{h},\bar{g}}(a, b) = (\bar{h} \star a, \bar{g} \star b), \\ S_2(a, b) = (a \downarrow_2, b \downarrow_2). \end{cases}$$

The adjoints are

$$\mathcal{D}^*(a, b) = a + b, \quad C^*_{\bar{h},\bar{g}}(a, b) = (h \star a, g \star b), \quad \text{and} \quad S_2^*(a, b) = (a \uparrow_2, b \uparrow_2).$$

To verify the convolution adjoint, assume the sequences are summable; in the periodic setting, the sums are finite:

$$\langle f \star \bar{h},\ g \rangle = \sum_n (f \star \bar{h})_n g_n = \sum_n \sum_k f_k h_{k-n} g_n = \sum_k f_k \sum_n h_{k-n} g_n = \langle f,\ h \star g \rangle,$$

For the copying operator,

$$\langle \mathcal{D}(a),\ (u, v) \rangle = \langle a,\ u \rangle + \langle a,\ v \rangle = \langle a,\ u + v \rangle = \langle a,\ \mathcal{D}^*(u, v) \rangle,$$

and for down-sampling,

$$\langle f \downarrow_2,\ g \rangle = \sum_n f_{2n} g_n = \sum_n (f_{2n} g_n + f_{2n+1} 0) = \langle f,\ g \uparrow_2 \rangle.$$

This is shown using matrix notation in Figure 4.13. Composing these adjoints in reverse order gives the reconstruction formula. ◻

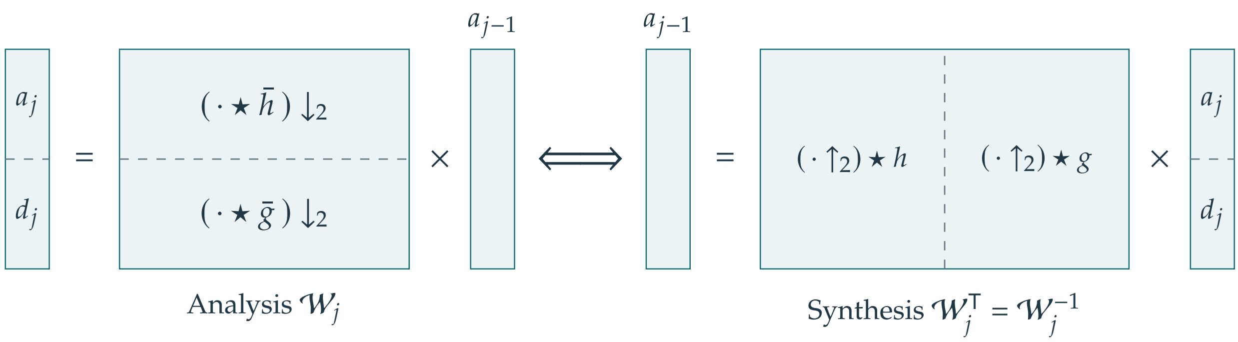


Figure 4.13. Analysis and inverse wavelet transforms in block-matrix form. Synthesis uses the transpose of the orthogonal analysis operator.

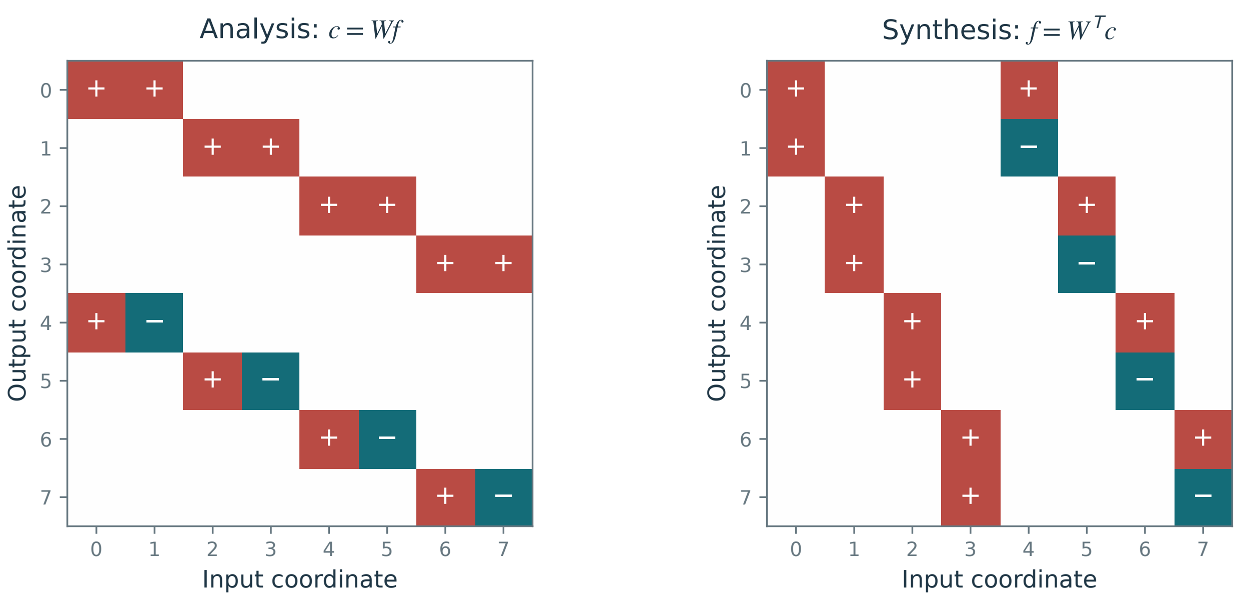
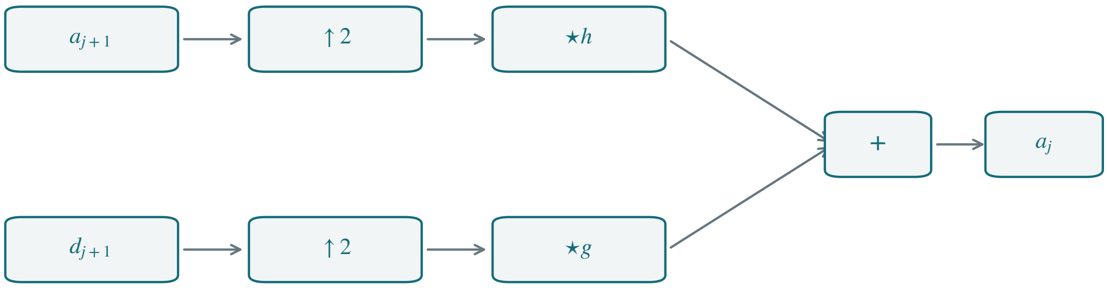


Figure 4.14. One-step Haar analysis matrix W and synthesis matrix W^{T} for eight samples.



$$a_j = (a_{j+1} \uparrow 2) \star h + (d_{j+1} \uparrow 2) \star g$$

Figure 4.15. Inverse filter-bank reconstruction.

Figure 4.14 makes this transpose relation concrete for one Haar step.

The inverse fast wavelet transform iteratively applies this elementary step

- **Input:** $\{d_j\}_{J<j\leqslant j_0} \cup \{a_{j_0}\}$.
- **For** $j = j_0, j_0 - 1, \ldots, J+1$.

$$a_{j-1} = (a_j \uparrow 2) \star h + (d_j \uparrow 2) \star g.$$

- **Output:** $f = a_J$.

The block diagram in Figure 4.15 reverses the operations of the forward transform in Figure 4.10.

4.5 2-D Wavelets

4.5.1 Anisotropic Wavelets

The anisotropic basis on $\mathbb{T}^2$ is the tensor product of the complete one-dimensional periodized basis $\mathcal{B}$ (including its coarsest scaling functions):

$$\{b_1 \otimes b_2 : b_1, b_2 \in \mathcal{B}\}.$$

Its wavelet–wavelet elements have the form

$$\psi_{(j_1,j_2),(n_1,n_2)}(x_1, x_2) = \psi^P_{j_1,n_1}(x_1)\psi^P_{j_2,n_2}(x_2). \tag{4.12}$$

The basis also includes wavelet–scaling, scaling–wavelet, and scaling–scaling products. The anisotropic 2-D transform follows the same separable construction as the 2-D FFT in Section 2.5.2. Treat the image $a_J \in \mathbb{R}^{2^{-J}\times 2^{-J}}$ as a matrix, apply the 1-D FWT to each row, and then apply it to each column. The total cost is $O(N)$ for $N = 2^{-2J}$ pixels.

4.5.2 Isotropic Wavelets

The anisotropic atoms (4.12) have independent scales in the two coordinate directions. For compactly supported wavelets, their supports lie in axis-aligned rectangles of size proportional to $2^{j_1} \times 2^{j_2}$. This directional bias can produce visible artifacts in denoising and compression when it does not match the image geometry.

An alternative uses a common scale in both directions. Tensor products of one-dimensional approximation spaces give a two-dimensional multiresolution

$$L^2(\mathbb{R}^2) \supset \ldots \supset V_{j-1} \otimes V_{j-1} \supset V_j \otimes V_j \supset V_{j+1} \otimes V_{j+1} \supset \ldots \supset \{0\}.$$

We write

$$V^O_j \overset{\text{def.}}{=} V_j \otimes V_j$$

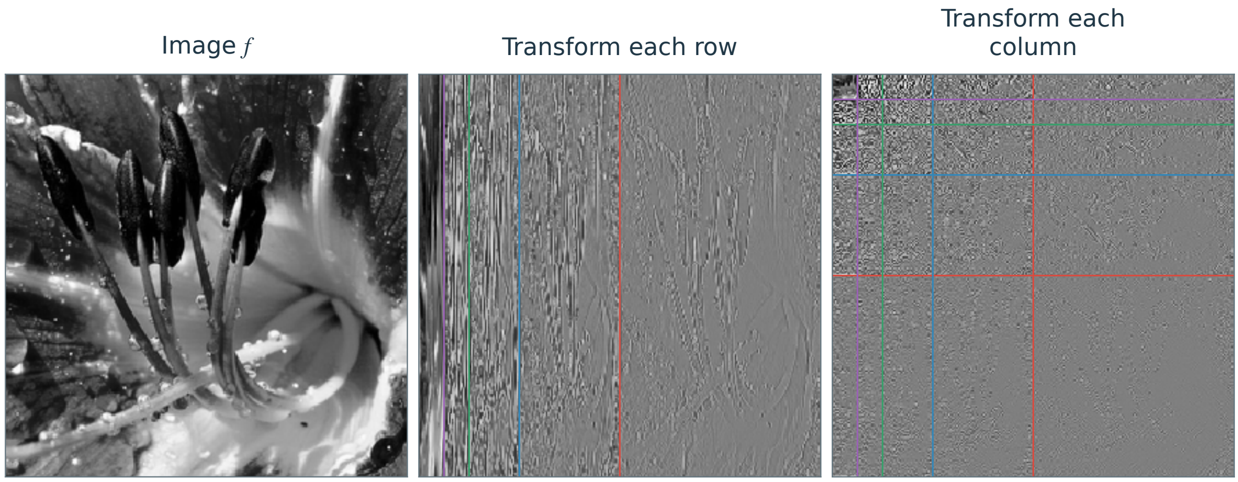


Figure 4.16. Steps of the anisotropic wavelet transform.

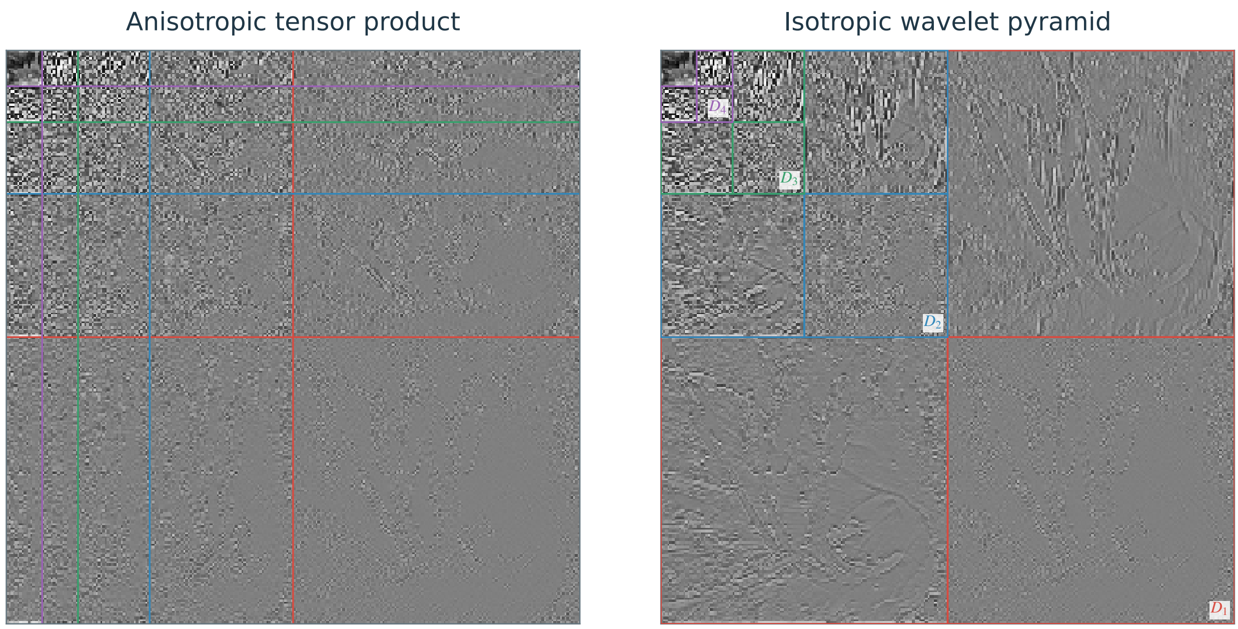


Figure 4.17. Anisotropic (left) versus isotropic (right) wavelet coefficients.

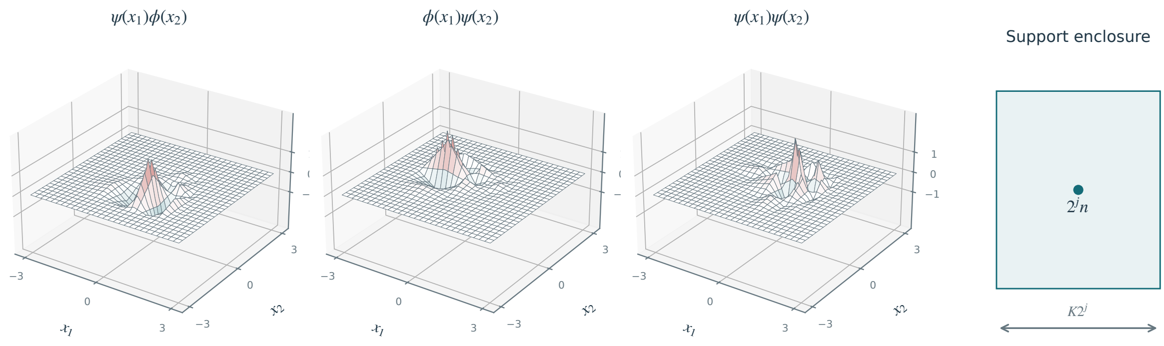


Figure 4.18. 2-D wavelets and a schematic support centered at $(2^j n_1, 2^j n_2)$, with width $K2^j$ (right).

for this isotropic 2-D approximation space.

Recall that the tensor product of two closed subspaces $V_1, V_2 \subset L^2(\mathbb{R})$ is

$$V_1 \otimes V_2 = \text{Closure}\left(\text{Span}\left\{f_1(x_1)f_2(x_2) \in L^2(\mathbb{R}^2)\ ;\ f_1 \in V_1, f_2 \in V_2\right\}\right).$$

If $(\varphi^s_k)_k$ are orthonormal bases of V_s, their tensor products $(\varphi^1_{k_1}(x_1)\varphi^2_{k_2}(x_2))_{k_1,k_2}$ form an orthonormal basis of $V_1 \otimes V_2$.

The tensor product distributes over these orthogonal sums

$$(V_j \oplus^\perp W_j) \otimes (V_j \oplus^\perp W_j) = V_j^O \oplus^\perp W_j^V \oplus^\perp W_j^H \oplus^\perp W_j^D \quad \text{where} \quad \begin{cases} W_j^V \overset{\text{def.}}{=} (V_j \otimes W_j), \\ W_j^H \overset{\text{def.}}{=} (W_j \otimes V_j), \\ W_j^D \overset{\text{def.}}{=} (W_j \otimes W_j). \end{cases}$$

The labels $\{V, H, D\}$ stand for *Vertical*, *Horizontal*, *Diagonal* detail spaces. The following diagram summarizes the nested spaces and their detail components:

$$\begin{tikzcd} \cdots \arrow[r] \arrow[dr] & V_{j-1} \otimes V_{j-1} \arrow[r] \arrow[dr] & V_j \otimes V_j \arrow[r] \arrow[dr] & V_{j+1} \otimes V_{j+1} \arrow[r] & \cdots \\ & W_{j-1}^{(2)} & W_j^{(2)} & W_{j+1}^{(2)} & \end{tikzcd}$$

$$V_{j-1} \otimes V_{j-1} = (V_j \otimes V_j) \oplus^\perp W_j^{(2)}$$

For $j \in \mathbb{Z}$, translates of a mother wavelet span each detail space. We index them by $n = (n_1, n_2) \in \mathbb{Z}^2$, or by indices in $\{0, \ldots, 2^{-j} - 1\}^2$ on the torus $\mathbb{T}^2$:

$$\forall\, \omega \in \{V, H, D\}, \quad W_j^\omega = \text{Span}\{\psi^\omega_{j,n_1,n_2}\}_{n_1,n_2}$$

where

$$\forall\, \omega \in \{V, H, D\}, \quad \psi^\omega_{j,n_1,n_2}(x) = \frac{1}{2^j}\psi^\omega\left(\frac{x_1 - 2^j n_1}{2^j}, \frac{x_2 - 2^j n_2}{2^j}\right)$$

and where the three mother wavelets are

$$\psi^H(x) = \psi(x_1)\varphi(x_2), \quad \psi^V(x) = \varphi(x_1)\psi(x_2), \quad \text{and} \quad \psi^D(x) = \psi(x_1)\psi(x_2).$$

Here the span is closed in L^2 on an unbounded domain; on the torus, the wavelets are periodized. Figure 4.18 displays examples of these wavelets.

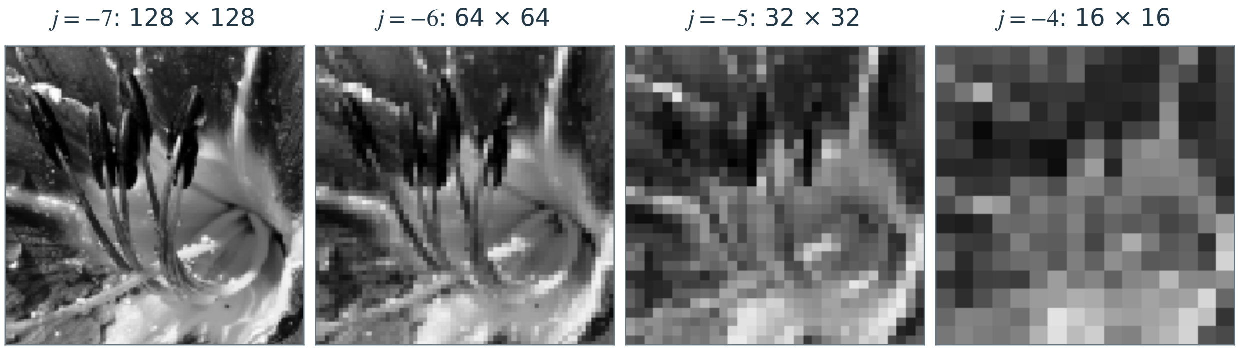


Figure 4.19. 2-D Haar approximation $P_{V_j^O} f$ for increasing j.

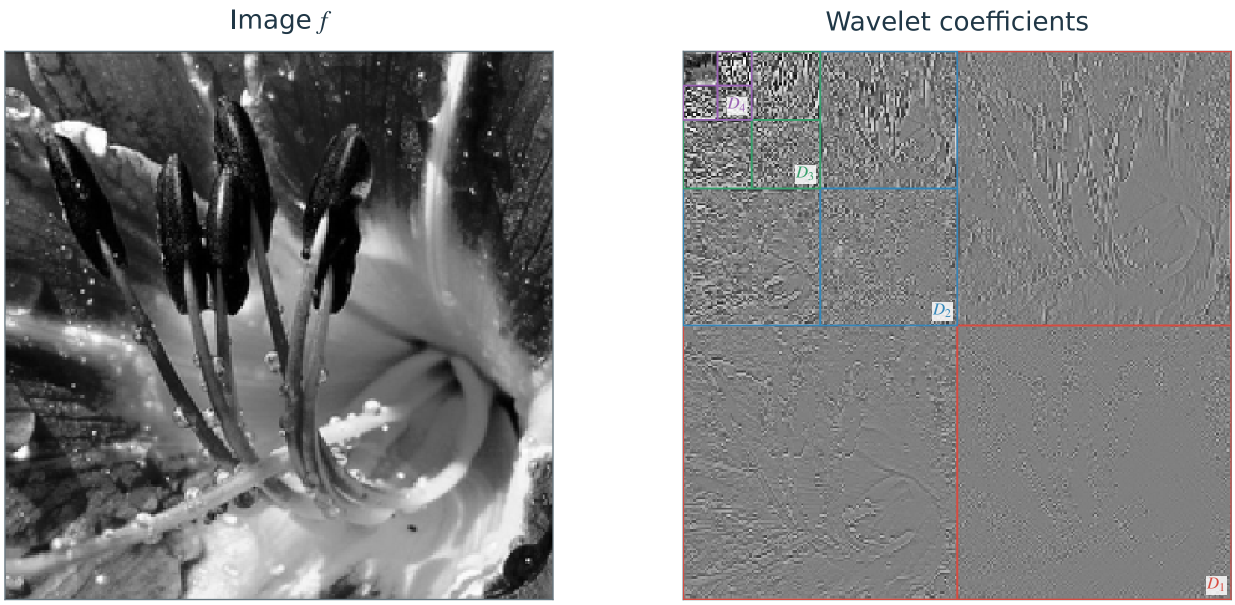


Figure 4.20. 2-D wavelet coefficients.

Haar 2-D multiresolution. The 2-D Haar construction gives piecewise constant approximations: functions in $V_j \otimes V_j$ are constant on squares of size $2^j \times 2^j$. Figure 4.19 shows the corresponding projections of an image.

Discrete 2-D wavelet coefficients. As in (4.5), assume that acquisition provides inner products of the continuous signal f with scaling functions at scale 2^J, with $N = 2^{-J}$ samples per direction:

$$\forall n \in \{0, \ldots, N-1\}^2, \quad a_{J,n} = \langle f, (\varphi_{J,n_1} \otimes \varphi_{J,n_2})^P \rangle$$

Discrete wavelet coefficients are defined as

$$\forall \omega \in \{V, H, D\}, \ \forall J < j \leqslant 0, \ \forall 0 \leqslant n_1, n_2 < 2^{-j}, \quad d_{j,n}^{\omega} = \langle f, \psi_{j,n}^{\omega} \rangle.$$

Here the wavelets are periodized. Approximation coefficients are defined as

$$a_{j,n} = \langle f, (\varphi_{j,n_1} \otimes \varphi_{j,n_2})^P \rangle.$$

Figure 4.20 shows how the wavelet coefficients are arranged in an image of N^2 pixels. Figure 4.21 shows other examples of wavelet decompositions.

One step of the forward 2-D transform. Each step separates fine-scale approximation coefficients into three detail arrays and a coarse approximation:

$$a_{j-1} \longmapsto (a_j, d_j^H, d_j^V, d_j^D).$$

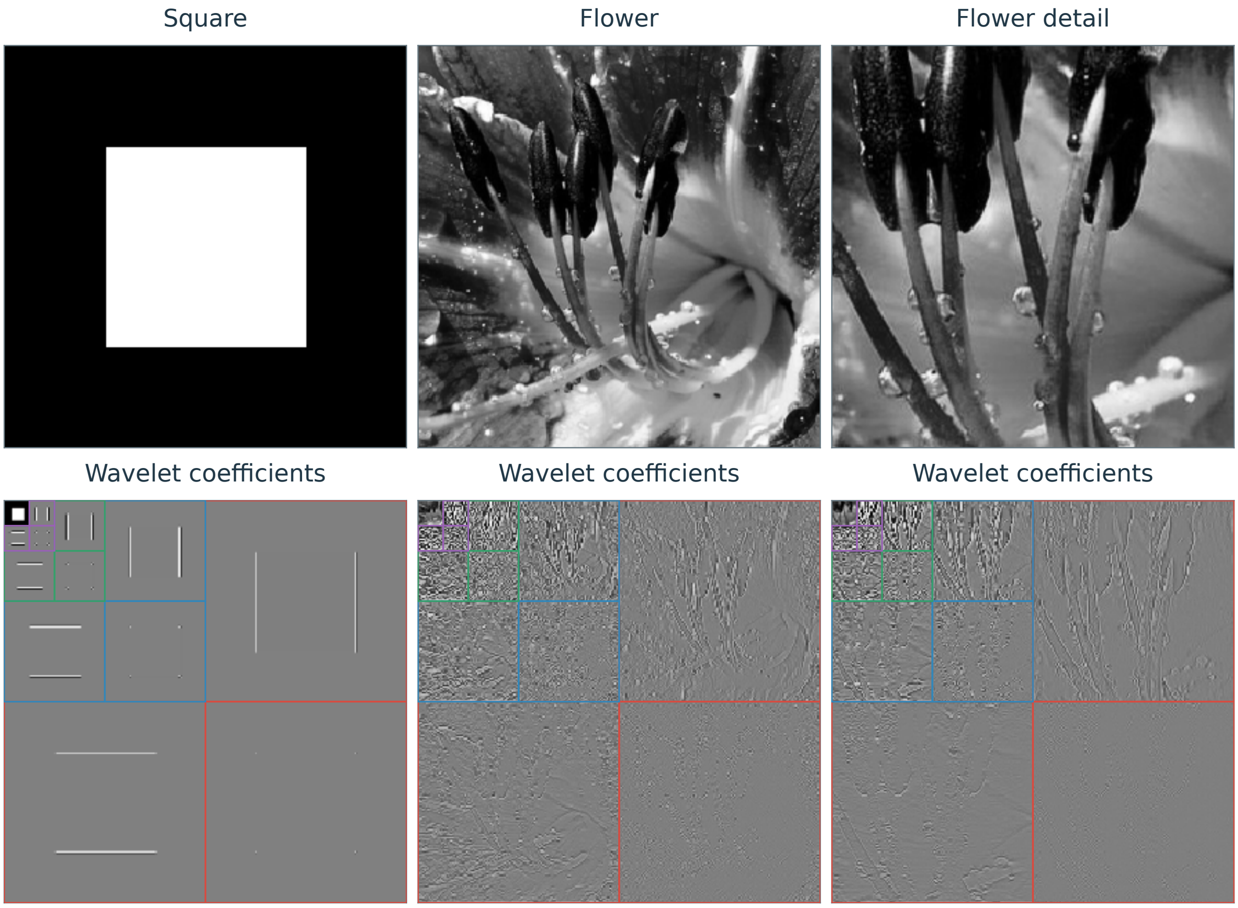


Figure 4.21. Images (top) and their wavelet coefficients (bottom).

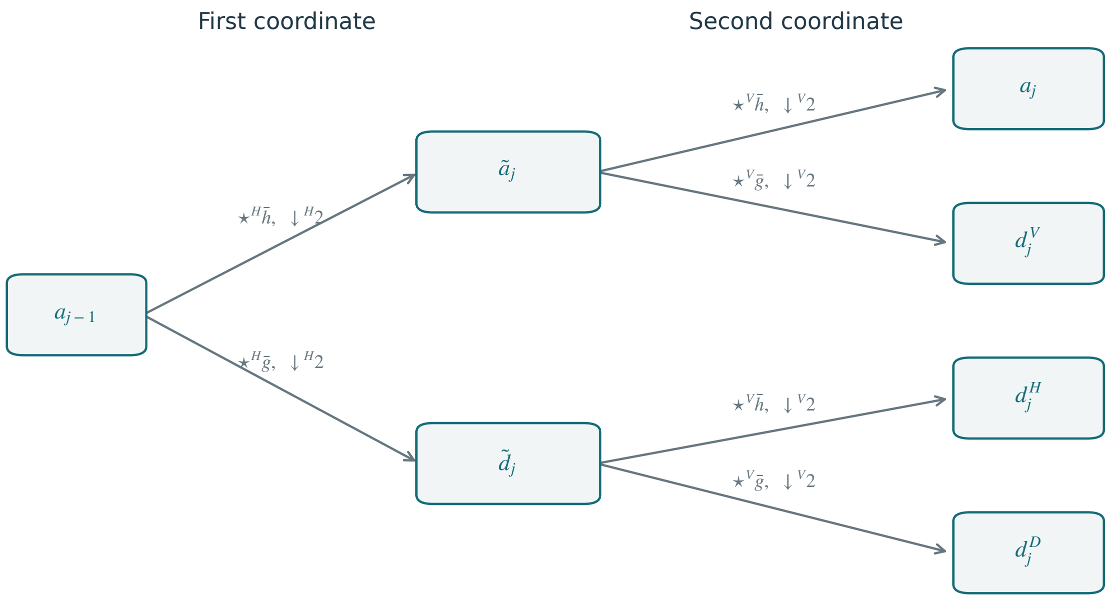


Figure 4.22. Forward 2-D filterbank step.

As in one dimension, this map is orthogonal. It is implemented by applying the filtering and sub-sampling formula (4.8) in each coordinate direction.

One first applies 1-D horizontal filtering and sub-sampling

$$\tilde{a}_j = (a_{j-1} \star^H \bar{h}) \downarrow^H 2$$
$$\tilde{d}_j = (a_{j-1} \star^H \bar{g}) \downarrow^H 2,$$

where $\star^H$ denotes convolution along the first coordinate (called horizontal here), applied to each column of the array

$$a \star^H b_{n_1,n_2} = \sum_{m_1=0}^{P-1} a_{n_1-m_1,n_2} b_{m_1}$$

Here $a \in \mathbb{C}^{P\times P}$ is a matrix and $b \in \mathbb{C}^P$ is a filter. The notation $\downarrow^H 2$ denotes sub-sampling in the horizontal direction

$$(a \downarrow^H 2)_{n_1,n_2} = a_{2n_1,n_2}.$$

One then applies 1-D vertical filtering and sub-sampling to $\tilde{a}_j$ and $\tilde{d}_j$ to obtain

$$\begin{aligned} a_j &= (\tilde{a}_j \star^V \bar{h}) \downarrow^V 2, & d_j^H &= (\tilde{d}_j \star^V \bar{h}) \downarrow^V 2, \\ d_j^V &= (\tilde{a}_j \star^V \bar{g}) \downarrow^V 2, & d_j^D &= (\tilde{d}_j \star^V \bar{g}) \downarrow^V 2, \end{aligned}$$

where the vertical operators are defined analogously to the horizontal operators, acting on rows.

These two forward steps are shown in the block diagram in Figure 4.22. The updates can be performed in place, storing all coefficients in an image of N^2 pixels, as shown in Figure 4.23. This produces the familiar multiscale arrangement used in Figure 4.21.

Fast 2-D wavelet transform. The 2-D FWT algorithm iterates these steps through the scales:

– **Input:** image $f \in \mathbb{C}^{N\times N}$.

– **Initialization:** $a_J = f$.

– **For** $j = J+1, \dots, j_0$.

$$\begin{aligned} \tilde{a}_j &= (a_{j-1} \star^H \bar{h}) \downarrow^H 2, & d_j^V &= (\tilde{a}_j \star^V \bar{g}) \downarrow^V 2, \\ \tilde{d}_j &= (a_{j-1} \star^H \bar{g}) \downarrow^H 2, & d_j^H &= (\tilde{d}_j \star^V \bar{h}) \downarrow^V 2, \\ a_j &= (\tilde{a}_j \star^V \bar{h}) \downarrow^V 2, & d_j^D &= (\tilde{d}_j \star^V \bar{g}) \downarrow^V 2. \end{aligned}$$

– **Output:** the coefficients $\{d_j^\omega\}_{J<j\leqslant j_0,\omega} \cup \{a_{j_0}\}$.

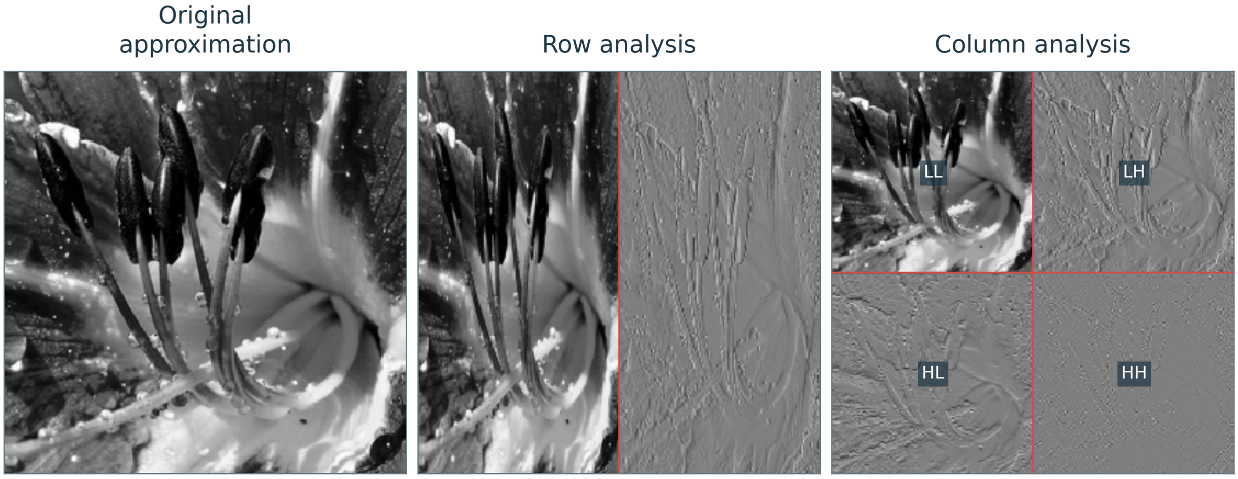


Figure 4.23. One step of the 2-D wavelet transform algorithm.

Fast 2-D inverse wavelet transform. The inverse transform undoes the horizontal and vertical filtering steps. The first step computes

$$\tilde{a}_j = (a_j \uparrow^V 2) \star^V h + (d_j^V \uparrow^V 2) \star^V g,$$
$$\tilde{d}_j = (d_j^H \uparrow^V 2) \star^V h + (d_j^D \uparrow^V 2) \star^V g,$$

where the vertical up-sampling is

$$(a \uparrow^V 2)_{n_1,n_2} = \begin{cases} a_{n_1,k} & \text{if} \quad n_2 = 2k, \\ 0 & \text{if} \quad n_2 = 2k+1. \end{cases}$$

The second inverse step computes

$$a_{j-1} = (\tilde{a}_j \uparrow^H 2) \star^H h + (\tilde{d}_j \uparrow^H 2) \star^H g.$$

Figure 4.24 gives the inverse filter-bank diagram corresponding to Figure 4.22.

The inverse fast wavelet transform iteratively applies these elementary steps

– **Input:** $\{d_j^\omega\}_{J<j\leqslant j_0,\omega} \cup \{a_{j_0}\}$.
– **For** $j = j_0, j_0 - 1, \ldots, J+1$.

$$\tilde{a}_j = (a_j \uparrow^V 2) \star^V h + (d_j^V \uparrow^V 2) \star^V g,$$
$$\tilde{d}_j = (d_j^H \uparrow^V 2) \star^V h + (d_j^D \uparrow^V 2) \star^V g,$$
$$a_{j-1} = (\tilde{a}_j \uparrow^H 2) \star^H h + (\tilde{d}_j \uparrow^H 2) \star^H g.$$

– **Output:** $f = a_J$.

4.6 Wavelet Design

Implementing the FWT requires the filters h and g, rather than explicit formulas for the scaling function φ and wavelet ψ. Most useful wavelets have no elementary closed form; their refinement equations define them implicitly, and the cascade algorithm approximates them.

This section derives constraints on the filters and gives examples. Under the usual quadrature-mirror construction, the low-pass filter h determines the high-pass filter g.

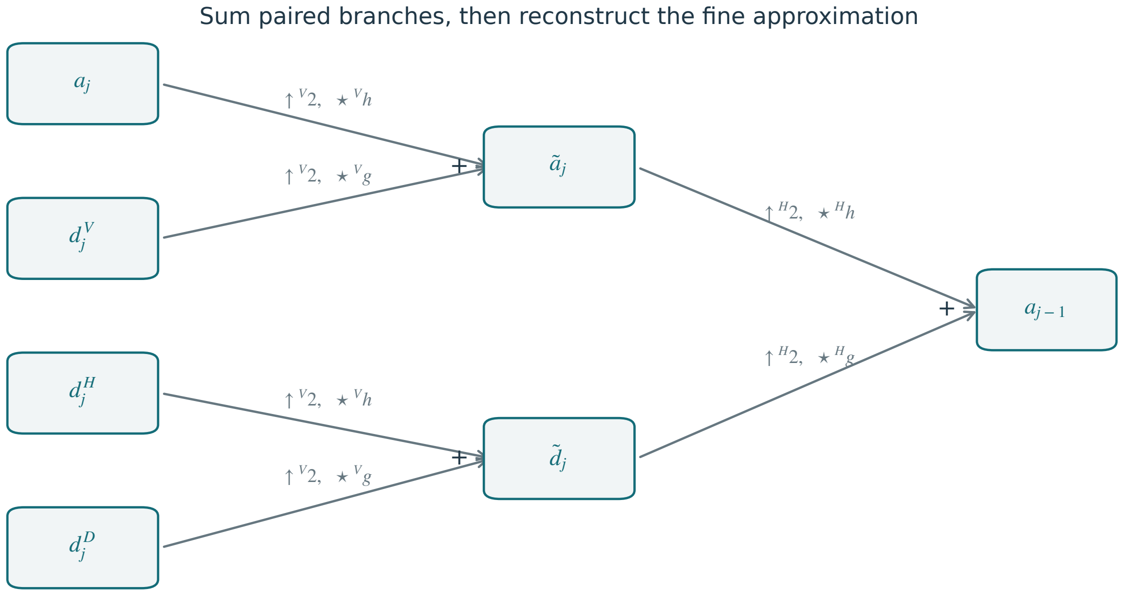


Figure 4.24. Backward 2-D filterbank step.

4.6.1 Low-pass Filter Constraints

We introduce the following three conditions on a filter h

$$\hat{h}(0) = \sqrt{2} \tag{C_1}$$

$$|\hat{h}(\omega)|^2 + |\hat{h}(\omega+\pi)|^2 = 2, \tag{C_2}$$

$$\inf_{\omega\in[-\pi/2,\pi/2]} |\hat{h}(\omega)| > 0. \tag{C^*}$$

These conditions use the Fourier series of the filter $h \in \mathbb{R}^{\mathbb{Z}}$:

$$\forall \omega \in \mathbb{R}/2\pi\mathbb{Z}, \quad \hat{h}(\omega) \stackrel{\text{def.}}{=} \sum_{n\in\mathbb{Z}} h_n e^{-in\omega}. \tag{4.13}$$

If $h \in \ell^1(\mathbb{Z})$, this defines a continuous periodic function $\hat{h} \in C^0(\mathbb{R}/2\pi\mathbb{Z})$, and the definition extends to $h \in \ell^2(\mathbb{Z})$ to give $\hat{h} \in L^2(\mathbb{R}/2\pi\mathbb{Z})$.

Theorem 4.4. *If an integrable real scaling function φ defines a multiresolution analysis and its filter h is summable, then (C_1) and (C_2) hold for h defined in (4.7). Conversely, if a finitely supported real filter h satisfies (C_1), (C_2), and (C^*), then there exists a φ defining multi-resolution approximation spaces whose associated filter is h as defined in (4.7).*

Proof. We prove the forward implication. The converse requires a separate construction.

We now prove condition (C_1). The refinement equation convolves a function on the real line with a discrete filter, represented by a sum of Dirac masses:

$$\frac{1}{\sqrt{2}}\varphi\left(\frac{t}{2}\right) = \sum_{n\in\mathbb{Z}} h_n \varphi(t-n). \tag{4.14}$$

Writing $h \star \varphi$ for this convolution and assuming $h \in \ell_1(\mathbb{Z})$ and $\varphi \in L^1(\mathbb{R})$, Fubini's theorem shows that $h \star \varphi \in L^1(\mathbb{R})$ and gives

$$\begin{aligned}\mathcal{F}(\sum_{n\in\mathbb{Z}} h_n\varphi(t-n))(\omega) &= \int_{\mathbb{R}} \sum_{n\in\mathbb{Z}} h_n \varphi(t-n)e^{-\mathrm{i}\omega t}\mathrm{d}t = \sum_{n\in\mathbb{Z}} h_n \int_{\mathbb{R}} \varphi(t-n)e^{-\mathrm{i}\omega t}\mathrm{d}t \\ &= \sum_{n\in\mathbb{Z}} h_n e^{-\mathrm{i}n\omega} \int_{\mathbb{R}} \varphi(x)e^{-\mathrm{i}\omega x}\mathrm{d}x = \hat{\varphi}(\omega)\hat{h}(\omega)\end{aligned}$$

where we substituted $x = t - n$. Here $\hat{h}(\omega)$ is the 2π-periodic Fourier series of the filter, defined in (4.13), while $\hat{\varphi}(\omega)$ is the continuous Fourier transform of the scaling function. The product therefore combines a 2π-periodic factor $\hat{h}$ with a generally nonperiodic factor $\hat{\varphi}$. We recall that $\mathcal{F}(f(\cdot/s)) = s\hat{f}(s\cdot)$. In the Fourier domain, equation (4.14) thus reads

$$\hat{\varphi}(2\omega) = \frac{1}{\sqrt{2}}\hat{h}(\omega)\hat{\varphi}(\omega). \tag{4.15}$$

One can show that $\hat{\varphi}(0) \neq 0$ (actually, $|\hat{\varphi}(0)| = 1$), so that this relation implies the first condition (C_1).

We now prove condition (C_2). The orthogonality of $\{\varphi(\cdot - n)\}_n$ can be expressed as a continuous convolution (see also Proposition 4.1)

$$\forall\, n \in \mathbb{Z}, \quad (\varphi \star \bar{\varphi})(n) = \delta_{n,0}$$

where $\bar{\varphi}(x) = \varphi(-x)$, and thus in the Fourier domain, using (4.15) which shows $\hat{\varphi}(\omega) = \frac{1}{\sqrt{2}}\hat{h}(\omega/2)\hat{\varphi}(\omega/2)$

$$1 = \sum_k |\hat{\varphi}(\omega + 2k\pi)|^2 = \frac{1}{2}\sum_k |\hat{h}(\omega/2 + k\pi)|^2 |\hat{\varphi}(\omega/2 + k\pi)|^2.$$

Since $\hat{h}$ is 2π-periodic, one can separate even and odd indices k and obtain

$$2 = |\hat{h}(\omega/2)|^2 \sum_k |\hat{\varphi}(\omega/2 + 2k\pi)|^2 + |\hat{h}(\omega/2 + \pi)|^2 \sum_k |\hat{\varphi}(\omega/2 + 2k\pi + \pi)|^2$$

To obtain condition (C_2), apply $\sum_k |\hat{\varphi}(\omega + 2k\pi)|^2 = 1$ at $\omega' = \omega/2$ instead of ω:

$$|\hat{h}(\omega')|^2 + |\hat{h}(\omega' + \pi)|^2 = 2.$$

The converse requires constructing a scaling function φ from the filter h; we do not prove it here. Iterating (4.15) suggests the infinite-product construction

$$\hat{\varphi}(\omega) = \prod_{k=1}^{\infty} \frac{\hat{h}(\omega/2^k)}{\sqrt{2}}. \tag{4.16}$$

Condition (C^*) can be shown to imply that this infinite product converges and defines a (non-periodic) function in $L^2(\mathbb{R})$. □

Condition (C^*) prevents zeros of $\hat{h}$ on the central half-period. It is a sufficient nondegeneracy condition for the converse construction used here.

4.6.2 High-pass Filter Constraints

We now introduce the following two conditions on a pair of filters (g, h)

$$|\hat{g}(\omega)|^2 + |\hat{g}(\omega + \pi)|^2 = 2 \tag{C_3}$$
$$\hat{g}(\omega)\hat{h}(\omega)^* + \hat{g}(\omega + \pi)\hat{h}(\omega + \pi)^* = 0. \tag{C_4}$$

Figure 4.25 illustrates these filter constraints.

Theorem 4.5. *If (φ, ψ) defines a multi-resolution analysis, then (C_3) and (C_4) hold for (h, g) defined in (4.7). Conversely, if the finitely supported real low-pass filter satisfies (C_1)–(C^*), and a summable real filter g satisfies (C_3)–(C_4), then there is a multiresolution wavelet pair (φ, ψ) whose associated filters are (h, g) as defined in (4.7). Furthermore,*

$$\hat{\psi}(\omega) = \frac{1}{\sqrt{2}}\hat{g}(\omega/2)\hat{\varphi}(\omega/2). \tag{4.17}$$

Proof. We prove the forward implication, beginning with condition (C_3). The refinement equation for the wavelet reads

$$\frac{1}{\sqrt{2}}\psi\left(\frac{t}{2}\right) = \sum_{n\in\mathbb{Z}} g_n \varphi(t - n)$$

and thus in the Fourier domain

$$\hat{\psi}(2\omega) = \frac{1}{\sqrt{2}}\hat{g}(\omega)\hat{\varphi}(\omega). \tag{4.18}$$

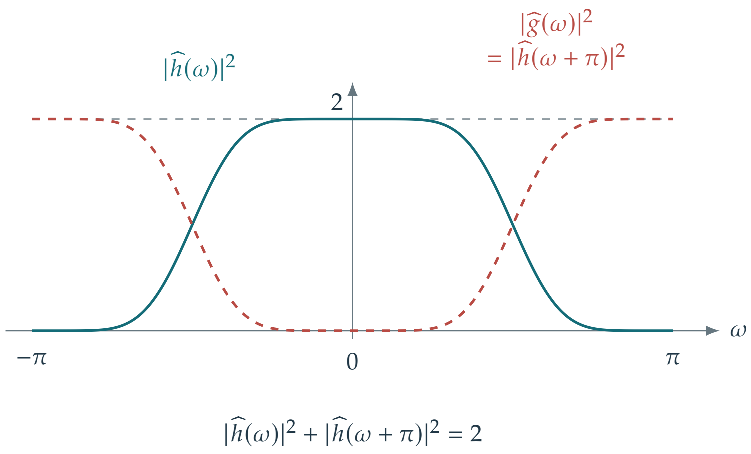


Figure 4.25. Constraints on low-pass and high-pass filters.

Orthogonality of the translates $\{\psi(\cdot - n)\}_n$ gives

$$\forall\, n \in \mathbb{Z}, \quad (\psi \star \bar{\psi})(n) = \delta_{n,0}$$

and thus in the Fourier domain (using Poisson's formula, see also Proposition 4.1)

$$\sum_k |\hat{\psi}(\omega + 2k\pi)|^2 = 1.$$

Apply the Fourier refinement equation (4.18) and separate the even and odd terms, as in the proof of condition (C_2) in Theorem 4.4. This gives condition (C_3). Figure 4.25 shows the squared Fourier magnitudes of a pair of filters satisfying this complementarity condition.

We now prove condition (C_4). The orthogonality between $\{\psi(\cdot - n)\}_n$ and $\{\varphi(\cdot - n)\}_n$ is written as

$$\forall\, n \in \mathbb{Z}, \quad \psi \star \bar{\varphi}(n) = 0$$

and hence in the Fourier domain (using Poisson's formula, similarly to Proposition 4.1)

$$\sum_k \hat{\psi}(\omega + 2k\pi)\hat{\varphi}^*(\omega + 2k\pi) = 0.$$

Using the Fourier domain refinement equations (4.15) and (4.18), this is equivalent to condition (C_4). □

Quadrature mirror filters. Quadrature mirror filters (QMF) define g as a function of h so that the conditions of Theorem 4.5 are automatically satisfied. Most wavelet constructions implicitly use this choice (other phase choices can yield different wavelets spanning the same detail spaces).

Proposition 4.6. *For a real filter $h \in \ell^1(\mathbb{Z})$ satisfying (C_2), the filter $g \in \ell^2(\mathbb{Z})$ defined by*

$$\forall\, n \in \mathbb{Z}, \quad g_n = (-1)^{1-n} h_{1-n} \tag{4.19}$$

satisfies conditions (C_3) and (C_4).

Proof. Taking the Fourier series of the defining relation gives

$$\hat{g}(\omega) = e^{-i\omega}\hat{h}(\omega + \pi)^*, \tag{4.20}$$

so that

$$|\hat{g}(\omega)|^2 + |\hat{g}(\omega+\pi)|^2 = |e^{-i\omega}\hat{h}(\omega+\pi)^*|^2 + |e^{-i(\omega+\pi)}\hat{h}(\omega+2\pi)^*|^2 = |\hat{h}(\omega+\pi)|^2 + |\hat{h}(\omega+2\pi)|^2 = 2.$$

where we used the fact that $\hat{h}$ is 2π-periodic, and also

$$\begin{aligned}\hat{g}(\omega)\hat{h}(\omega)^* + \hat{g}(\omega+\pi)\hat{h}(\omega+\pi)^* &= e^{-i\omega}\hat{h}(\omega+\pi)^*\hat{h}(\omega)^* + e^{-i(\omega+\pi)}\hat{h}(\omega+2\pi)^*\hat{h}(\omega+\pi)^* \\ &= (e^{-i\omega} + e^{-i(\omega+\pi)})\hat{h}(\omega+\pi)^*\hat{h}(\omega)^* = 0.\end{aligned}$$

□

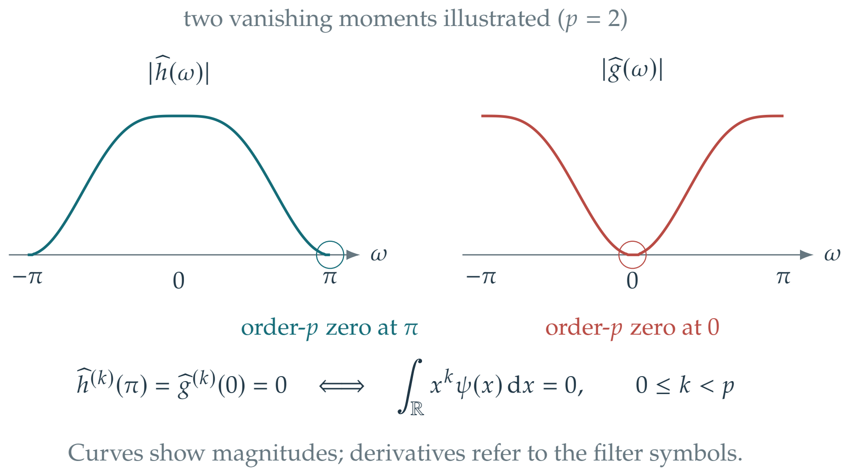


Figure 4.26. Vanishing moments correspond to vanishing derivatives of the Fourier transform.

4.6.3 Wavelet Design Constraints

One can construct a multiresolution wavelet pair by designing a finite filter h satisfying (C_1)–(C^*). The scaling function is obtained through the infinite product (4.16). Except in special cases such as Haar wavelets, it has no elementary closed form. The QMF relation (4.19) then defines g, and (4.17) defines ψ.

Fourier atoms are fixed exponentials, whereas mother wavelets can be designed to meet several requirements:
– Size of the support.
– Cancellation of polynomials, measured by the number p of vanishing moments.
– Symmetry (for compactly supported scalar orthogonal wavelets, the Haar system is the exceptional symmetric case).
– Smoothness (number of derivatives).

We now explain how these design requirements interact with conditions (C_1)–(C_4).

Vanishing moments. A wavelet ψ has p vanishing moments if

$$\forall\, k = 0, \ldots, p-1, \qquad \int_{\mathbb{R}} \psi(x)\, x^k \mathrm{d}x = 0. \tag{4.21}$$

Vanishing moments make $\langle f, \psi_{j,n} \rangle$ small when f has regularity C^α with $\alpha < p$ on $\mathrm{Supp}(\psi_{j,n})$. A Taylor expansion of f about $2^j n$ on the wavelet's support explains this cancellation. The vanishing-moment condition has the following equivalent Fourier characterization; see Figure 4.26.

Proposition 4.7. *Assuming the required moments of ψ and derivatives of the filter symbols exist, a wavelet obtained by the QMF construction (4.19) has p vanishing moments if and only if*

$$\forall\, k = 0, \ldots, p-1, \qquad \frac{\mathrm{d}^k \hat{h}}{\mathrm{d}\omega^k}(\pi) = \frac{\mathrm{d}^k \hat{g}}{\mathrm{d}\omega^k}(0) = 0. \tag{4.22}$$

Proof. Integrability of the moments allows differentiation under the Fourier integral: $\hat{\psi}^{(k)} = \mathcal{F}((-\mathrm{i}\cdot)^k \psi(\cdot))$. Thus

$$(-\mathrm{i})^k \int_{\mathbb{R}} x^k \psi(x) \mathrm{d}x = \hat{\psi}^{(k)}(0).$$

Relations (4.17) and (4.19) imply

$$\hat{\psi}(2\omega) = \hat{h}(\omega + \pi)^* \rho(\omega) \quad \text{where} \quad \rho(\omega) \overset{\text{def.}}{=} \frac{1}{\sqrt{2}} e^{-i\omega} \hat{\varphi}(\omega).$$

and differentiating this relation shows

$$2\hat{\psi}^{(1)}(2\omega) = \hat{h}(\omega + \pi)^* \rho^{(1)}(\omega) + \hat{h}^{(1)}(\omega + \pi)^* \rho(\omega)$$

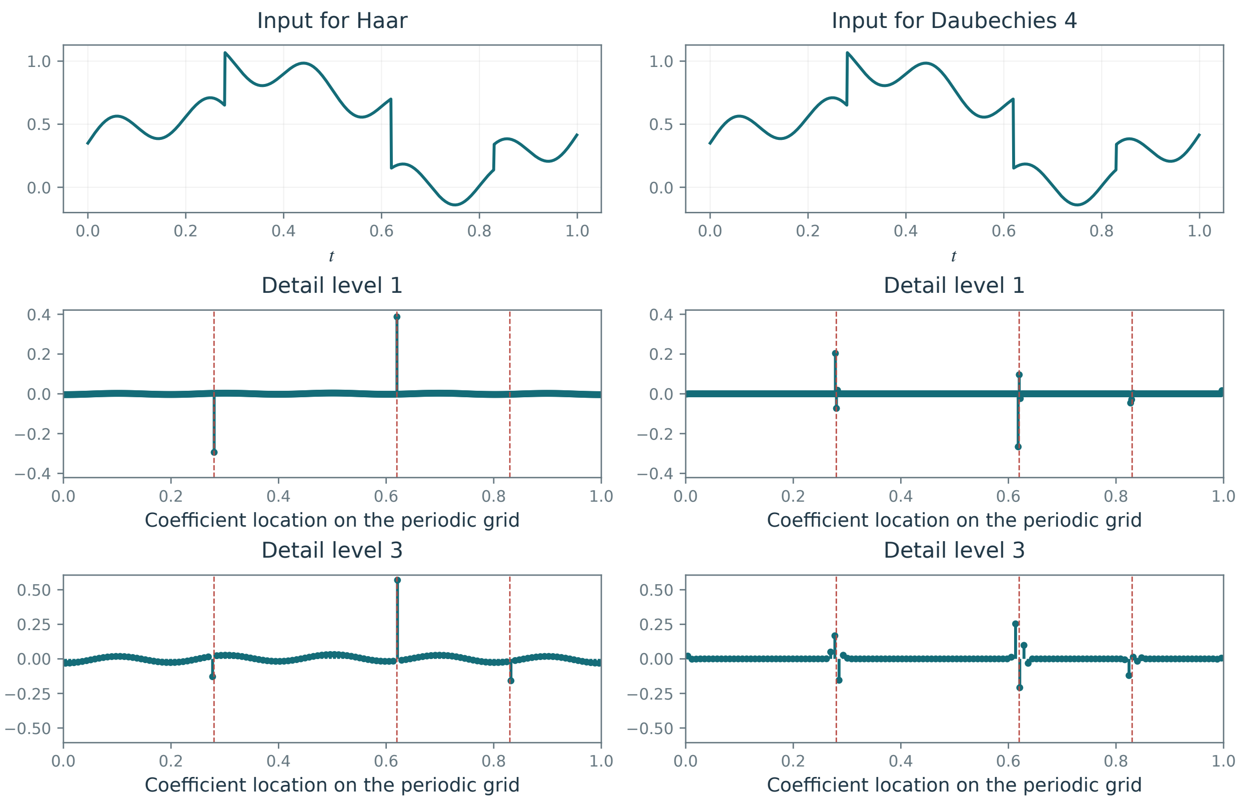


Figure 4.27. Location of large coefficients for Haar (left) and Daubechies-4 (right) wavelets applied to the same piecewise-smooth signal.

Since $\hat{h}(\pi) = 0$ and $\rho(0) = \hat{\varphi}(0)/\sqrt{2} \neq 0$, this shows that $\hat{\psi}^{(1)}(0) = 0 \Leftrightarrow \hat{h}^{(1)}(\pi)^* = 0$. Induction gives the higher-order equivalence $\hat{\psi}^{(k)}(0) = 0 \Leftrightarrow \hat{h}^{(k)}(\pi)^* = 0$, provided all derivatives of lower order vanish. □

Conditions (C_1) and (C_2) give $\hat{h}(\pi) = 0$, so every admissible wavelet has at least 1 vanishing moment: $\int \psi = 0$. By (4.22), additional vanishing moments require the Fourier transform $\hat{h}$ to vanish to higher order at $\omega = \pi$.

Support. Figure 4.27 shows the wavelet coefficients of a piecewise smooth signal. The cancellation provided by ψ suppresses coefficients in smooth regions, leaving the largest coefficients near singularities.

Choosing ψ with short support limits how many wavelets intersect each singularity. One can show that the size of the support of ψ is proportional to the size of the support of h. Short support competes with the vanishing-moment requirement (4.21). Indeed, one can prove that for an orthogonal wavelet basis with p vanishing moments

$$|\mathrm{Supp}(\psi)| \geqslant 2p - 1,$$

where the left-hand side denotes the length of the smallest closed interval containing the support of a compactly supported orthogonal mother wavelet.

Chapter 5 studies in detail the tradeoff between support size and vanishing moments in nonlinear approximation of piecewise smooth signals.

Smoothness. In compression or denoising applications, an approximate signal is recovered from a partial set I_M of coefficients,

$$f_M = \sum_{(j,n)\in I_M} \langle f, \psi_{j,n} \rangle \psi_{j,n}.$$

This finite sum is at least as smooth as ψ.

A smooth wavelet ψ helps avoid visible reconstruction artifacts. Smoothness controls the regularity of the reconstructed signal, while vanishing moments and localization primarily determine the approximation

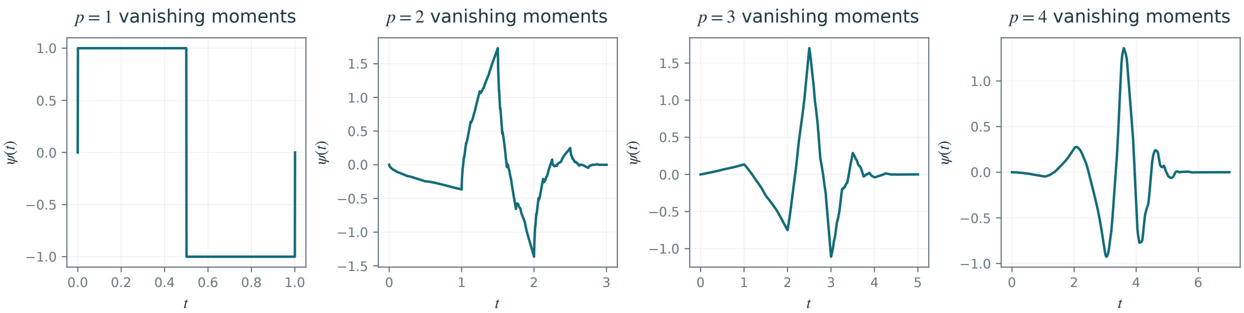


Figure 4.28. Daubechies mother wavelets ψ with increasing numbers p of vanishing moments.

rates studied below. These properties are often linked: in many families, including the Daubechies wavelets introduced next, more vanishing moments also bring greater smoothness.

4.6.4 Daubechies Wavelets

To build a wavelet ψ with a fixed number p of vanishing moments, one designs the filter h, and uses the quadrature mirror filter relation (4.19) to compute g. One therefore looks for h such that

$$|\hat{h}(\omega)|^2 + |\hat{h}(\omega+\pi)|^2 = 2, \quad \hat{h}(0) = \sqrt{2}, \quad \text{and} \quad \forall\, k = 0, \dots, p-1, \ \frac{\mathrm{d}^k \hat{h}}{\mathrm{d}\omega^k}(\pi) = 0.$$

These conditions give algebraic relations between the coefficients of h, which can be solved explicitly using the Euclidean division algorithm for polynomials.

The resulting Daubechies wavelets have p vanishing moments and attain the minimum support length $2p - 1$ among compactly supported orthogonal wavelets.

The following filter coefficients are rounded to four decimal places. For $p = 1$, the construction gives the Haar system (up to the overall sign of the wavelet), with

$$h = [h_0 = 0.7071; 0.7071].$$

For $p = 2$, one obtains the celebrated Daubechies 4 filter

$$h = [0.4830; h_0 = 0.8365; 0.2241; -0.1294],$$

and for $p = 3$,

$$h = [0; 0.3327; 0.8069; h_0 = 0.4599; -0.1350; -0.0854; 0.0352].$$

Wavelet display. Figure 4.28 compares Daubechies mother wavelets with increasing numbers of vanishing moments. Each continuous wavelet is approximated by a discrete wavelet $\bar{\psi}_{j,n}$ computed on a fine grid of N samples. Applying the inverse transform to the following coefficient vector produces that discrete wavelet: $d_{j',n'} = \delta_{j-j'}\delta_{n-n'}$.

References

[1] Stephane Mallat. *A wavelet tour of signal processing: the sparse way*. Academic press, 2008.