4

Wavelets

4.1 Wavelets

Gabor functions and other time-frequency dictionaries correspond to decompositions of signals into elementary components that take up, in a sense, one unit of area in the time-frequency plane. In the case of Malvar functions the plane need not be divided uniformly. Wavelets correspond to one type of decomposition in which the elementary time-frequency rectangles have the form $[\ell/2^k, (\ell+1)/2^k) \times [2^k, 2^{k+1})$.

4.1.1 Time-scale representation, continuous wavelet transform, inversion

The continuous wavelet transform is defined in the following way. Suppose that $\psi \in L^2(\mathbb{R})$ is integrable and that $\int \psi = 0$. Define $\psi_s(x) = \frac{1}{s}\psi\left(\frac{x}{s}\right)$ and define $W(f)(x,s) = f * \tilde{\psi}_s(x)$ where $\tilde{\psi}(x) = \bar{\psi}(-x)$. Here is a somewhat miraculous looking fact. For appropriate ψ one has

$$f(x) = \int_{s=0}^{\infty} f * \psi_s * \bar{\psi}_s \, \frac{ds}{s} \tag{4.1}$$

Exercise 4.1.1. Show that (4.1) holds if

$$\int_0^\infty |\hat{\psi}(s\xi)|^2 \, \frac{ds}{s} \equiv 1.$$

By a change of variables the right hand side, if it exists, only depends on $\xi = \pm 1$. The integral exists if $\hat{\psi}$ is continuous and vanishes at zero and decays fast enough at infinity.

Exercise 4.1.2. Define a numerical continuous wavelet transform in terms of samples of the derivative of the Gaussian $\mathrm{e}^{-\pi x^2}$.

4.1.2 Discrete wavelet transform and Orthonormal bases

One of the motivations behind the discrete wavelet transform is based on the following question: is it possible that by taking samples of the *wavelet reproducing formula* 4.1 one can find an orthonormal basis for $L^2(\mathbb{R})$? Notice that when $s = 2^j$ and $x = k/2^j$ one has

$$W_\psi(f)(k/2^j, 1/2^j) = f * \psi_{1/2^j}(k/2^j) = \int_{-\infty}^{\infty} f(y) 2^j \tilde{\psi}(2^j(y - k/2^j))\, dy = \int_{-\infty}^{\infty} f(y) 2^j \bar{\psi}(2^j y - k)\, dy = 2^{j/2} \langle f,\, \psi_{jk} \rangle$$

where $\psi_{jk}(x) = 2^{j/2}\psi(2^j x - k)$. Also, if we replace the average of $f * \tilde{\psi}_s * \psi_s$ over the rectangle $[k/2^j, (k+1)/2^j] \times [1/2^j, 2/2^j]$ by the value at $(k/2^j, 1/2^j)$ one gets $\langle f,\, \psi_{jk}\rangle \psi_{jk}$. Now replacing the convolution with ψ_s and integral over s with the sum over the corresponding samples $(k/2^j, 1/2^j)$ one has the *Riemann* approximation of formula (4.1) by

$$f \approx \sum_{jk} \langle f,\, \psi_{jk} \rangle \psi_{jk} \tag{4.2}$$

Replacing "≈" by "=" in this last formula, for any f, would be tantamount to the statement that the functions $\psi_{jk}(x) = 2^{j/2}\psi(2^j x - k)$ form an orthonormal basis for $L^2(\mathbb{R})$.

We have already seen examples of this, namely the bandlimited wavelets discussed in Chapter **??**. Here we discuss a second example, in some ways dual to the bandlimited wavelets.

4.2 Haar Wavelets

The Haar wavelets were introduced by Alfred Haar in 1909 as an example of an orthonormal basis for L^2. Define $\psi(x) = \mathbb{1}_{[0,1/2)} - \mathbb{1}_{[1/2,1)}$. Note that $\|\psi\|_{L^2} = 1$. As before, let $\psi_{jk} = 2^{j/2}\psi(2^j x - k)$.

Exercise 4.2.1. Show that either ψ_{jk} and $\psi_{j'k'}$ have disjoint supports or else (if $j' > j$), ψ_{jk} is constant on the support $[k'/2^{j'}, (k'+1)/2^{j'})$ of ψ_{jk}. Show that, consequently, the functions ψjk are mutually orthogonal.

Next we make use of a property of functions in $L^2\mathbb{R}$.

Theorem 4.2.2. *Let $P_j f$ denote the orthogonal projection of f onto the space V_j of functions that are piecewise constant on intervals $[k/2^j, (k+1)/2^j)$. Then $\|P_j f - f\|_{L^2} \to 0$ as $j \to \infty$.*

The projection maps f to its average values on each of the intervals $I_{jk} = [k/2^j, (k+1)/2^j)$. This theorem is not obvious. It is proved first for the case when f is a continuous function that vanishes outside of a finite interval, then one bootstraps by showing that any function in $L^2(\mathbb{R})$ can be approximated by such continuous, compactly supported functions.

4.3 Scaling functions

The Haar function $H(x) = \mathbb{1}_{[0,1)}$ satisfies $H(x) = H(2x) + H(2x-1)$. This type of relationship is called a *scaling relationship* or a *two-scale dilation equation* or *refinement equation.*

Definition 4.3.1. *A function φ satisfies a two scale dilation equation on $\mathbb{R}$ is there is a sequence $\{h_k\}$ of numbers such that*

$$\varphi(x) = 2\sum_{k\in\mathbb{Z}} h_k \varphi(2x-k). \tag{4.3}$$

The φ is called a scaling function and $\{h_k\}$ is called the scaling filter.

With this particular normalization, in the Haar case one has $h_0 = 1/2 = h_1$ and all other $h_k = 0$.

Another example is the tent function $T(x) = (1 - |x-1|)_+$ which takes the value 0 if $|x-1| > 1$ and $(1 - |x-1|)$ if $|x-1| < 1$. This function satisfies

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

so $h_0 = 1/4 = h_2$, $h_1 = 1/2$ and $h_k = 0$ for all other k.

4.3.1 Convolution

The convolution of two sequences $\mathbf{c} = \{c_k\}$ and $\mathbf{d} = \{d_k\}$ defined for $k \in \mathbb{Z}$ is defined as $(\mathbf{c} * \mathbf{d})_k = \sum_{\ell\in\mathbb{Z}} c_{k-\ell} d_\ell$.

Exercise 4.3.2. Verify that if $\{c_k\}$ and $\{d_k\}$ are the coefficients of formal Laurent series $C(z) = \sum_{k\in\mathbb{Z}} c_k z^k$ and $D(z) = \sum_{k\in\mathbb{Z}} d_k z^k$ then $(\mathbf{c} * \mathbf{d})_k$ is the coefficient of z^k in the formal product $C(z)D(z)$. Consequently, explain why $\mathbf{c} * \mathbf{d} = \mathbf{d} * \mathbf{c}$, that is, convolution is commutative.

Suppose now that $\mathbf{h} = \{h_k\}$ and $\mathbf{g} = \{g_k\}$ are two scaling sequences for scaling functions φ and ψ respectively. We claim then that $\mathbf{h} * \mathbf{g}$ is the scaling filter for $\varphi * \psi$ which is also a scaling function.

To prove this it is actually easier to work in the Fourier domain.

Exercise 4.3.3. Show that if φ satisfies the refinement equation (4.3) then its Fourier transform satisfies

$$\hat{\varphi}(\xi) = \frac{1}{2}\sum_k h_k e^{-\pi i k\xi}\hat{\varphi}\big(\frac{\xi}{2}\big) = H\big(\frac{\xi}{2}\big)\hat{\varphi}\big(\frac{\xi}{2}\big)$$

Similarly, $\hat{\psi} = G(\xi/2)\psi(\xi/2)$.

With this exercise and the fact that the Fourier transform maps convolutions to products it follows that

$$\widehat{\varphi * \psi}(\xi) = \hat{\varphi}\hat{\psi} = H(\xi/2)\varphi(\xi/2)G(\xi/2)\psi(\xi/2) = H(\xi/2)G(\xi/2)\widehat{\varphi * \psi}(\xi/2).$$

Noticing that, up to factors of two, $H(\xi)G(\xi)$ is just the product $(\sum_k h_k z^k)(\sum_\ell g_\ell z^\ell$ evaluated at $z = e^{-\pi i\xi}$, it follows that HG is the Fourier series of the convolution $\mathbf{h} * \mathbf{k}$ so the claim follows from taking inverse Fourier transforms. This proves that the convolution of two scaling functions is another scaling function.

Example 4.3.4. The function $T(x) = H * H$ where $H = \mathbb{1}_{[0,1]}$ is the Haar function. The scaling sequence of the Haar function is $h_0 = 1/2 = h_1$ and all other $h_k = 0$. Convolving this sequence with itself gives $g = h * h$ where $g_k = \sum_\ell h_{k-\ell h_\ell} = 0$ unless there is some ℓ such that $k - \ell \in \{0, 1\}$ and $\ell \in \{0, 1\}$. Then $g_0 = h_0 h_0 = 1/4$, $g_1 = h_1 h_0 + h_0 h_1 = 1/2$ and $g_2 = h_1 h_1 = 1/4$ so g is the scaling sequence of $\varphi_2 = T(x) = H * H$. In general, $\varphi_N \varphi_{N-1} * H$, the N-fold convolution of H with itself, defines a sequence of scaling functions called *spline* scaling functions since the functions φ_N are known as *B-splines.*

Exercise 4.3.5. Show that the scaling coefficients of φ_N are given by $h_k^N = \frac{1}{2^N}\binom{N}{k}$ for $k =), \dots, N$ and, thus $\sum_k h_k = 1$. Plot φ_N up to $N = 5$. Use the matlab convolution command `conv` to generate the sample values of φ^N.

4.3.2 Conditions on scaling filters

Compact support

Suppose that in equation (4.3) only the coefficients $h_0, h_1, \dots, h_M$ are nonzero and suppose that we apply the *two-scale transition* operator

$$Tf(x) = 2\sum_{k\in\mathbb{Z}} h_k f(2x - k) \tag{4.4}$$

to a continuous function f supported in $[0, M]$. If $x \geq M$ then $2x - k \geq 2M - k \geq M$ whenever $h_k \neq 0$ and similarly if $x \leq 0$ then $2x - k \leq -k \leq 0$ when $h_k \neq 0$. Therefore $Tf(x) = 0$ outside of $[0, M]$. More generally, if $B > M$ and $A < 0$ then $Tf(x) = 0$ whenever $x > (B + M)/2$ or $x < A/2$ if h is supported in $\{0, 1, \dots, M\}$.

Proposition 4.3.6. *If $h_k = 0$ whenever $k < 0$ or $k > M$ then the operator T in (4.4) maps the space of continuous functions supported in $[0, M]$ into itself. Additionally, if $x > M$ or $x < 0$ and $f = 0$ outside of a finite interval $[A, B]$ then, for large enough n, $T^n f(x) = 0$.*

The second statement follows from iteration on the last observation above.

Corollary 4.3.7. *If the scaling equation (4.3) has a continuous solution φ where the scaling sequence $\mathbf{h}$ is supported in $\{0, 1, \dots, M\}$ then φ is supported in $[0, M]$.*

A more difficult question is whether φ can have compact support if $\mathbf{h}$ has infinite support. It turns out that, at least if the integral of φ makes sense, then

4.3.3 Integral one

Suppose that we want $\int \varphi \neq 0$ for the scaling function φ. Then, upon integrating both sides of (4.3) and using the fact that $\int \varphi(2x - k)\, dx = \frac{1}{2}\int \varphi$, upon dividing out by the integrals one is left with the equation $\sum h_k = 1$. That $\int \varphi \neq 0$ is needed to make sense of the scaling equation $\hat{\varphi}(0) = H(0)\hat{\varphi}(0)$ as a limiting case of $\hat{\varphi}(\xi) = H(\xi/2)\hat{\varphi}(\xi/2)$.

4.3.4 Biorthogonal pairs

Now we are going to impose a condition not on a single scaling function but instead on a pair of scaling functions.

Definition 4.3.8. *A pair φ and $\widetilde{\varphi}$ of L^2 functions is said to be shift biorthogonal if for each k and ℓ in $\mathbb{Z}$ one has*

$$\langle \varphi(\cdot - k),\, \widetilde{\varphi}(\cdot - \ell)\rangle = \delta_{k\ell}.$$

Let H and $\widetilde{H}$ be the scaling filters of φ and $\widetilde{\varphi}$, that is, $\widehat{\varphi}(2\xi) = H(\xi)\widehat{\varphi}(\xi)$ and $\widehat{\widetilde{\varphi}}(2\xi) = \widetilde{H}(\xi)\widehat{\widetilde{\varphi}}(\xi)$. Suppose that we want to impose the condition that φ is orthogonal to its integer shifts, namely,

$$\langle \varphi(x - k),\, \varphi(x - \ell)\rangle = \delta_{k\ell}.$$

Since the Fourier transform of a shift is a modulated Fourier transform, Plancherel's theorem tells us that if φ and $\widetilde{\varphi}$ are shift biorthogonal then

$$\begin{aligned}\langle \varphi(\cdot - k),\, \widetilde{\varphi}(\cdot - \ell)\rangle &= \delta_{k\ell}\\ &= \int \widehat{\varphi}(\xi)\overline{\widehat{\widetilde{\varphi}}}(\xi)\mathrm{e}^{-2\pi i(k-\ell)\xi}\, d\xi\\ &= \sum_{m\in\mathbb{Z}}\int_0^1 \widehat{\varphi}(\xi+m)\overline{\widehat{\widetilde{\varphi}}}(\xi+m)\mathrm{e}^{-2\pi i(k-\ell)\xi}\, d\xi.\end{aligned}$$

The function $\Phi(\xi) = \sum_{m\in\mathbb{Z}}\int_0^1 \widehat{\varphi}(\xi+m)\overline{\widehat{\widetilde{\varphi}}}(\xi+m)$ is periodic and the quantity above is its $k-m$-th Fourier coefficient. So orthogonality says that all Fourier coefficients are zero except for the zeroth one, which is one. The only function that has this property (because of the Fourier uniqueness theorem) is the constant function one so we have

Proposition 4.3.9. *The functions φ and $\widetilde{\varphi}$ are shift orthogonal if and only if*

$$\Phi(\xi) = \sum_{m\in\mathbb{Z}} \widehat{\varphi}(\xi+m)\overline{\widehat{\widetilde{\varphi}}}(\xi+m)$$

is identically equal to one.

Let's try to translate this into a condition on the scaling coefficients. Recall that $\widehat{\varphi}(\xi) = \widehat{\varphi}(\xi/2)H(\xi/2)$ where $H(\xi) = \frac{1}{2}\sum_k h_k \mathrm{e}^{-2\pi i k\xi}$. Therefore we can write

$$\begin{aligned}\Phi(\xi) &= \sum_{m\in\mathbb{Z}} \widehat{\varphi}(2\xi+m)\overline{\widehat{\widetilde{\varphi}}}(2\xi+m)\\ &= \Big(\sum_{m\text{ even}} + \sum_{m\text{ odd}}\Big)\widehat{\varphi}(2\xi+m)\overline{\widehat{\widetilde{\varphi}}}(2\xi+m)\\ &= \sum_k \widehat{\varphi}(2\xi+2k)\overline{\widehat{\widetilde{\varphi}}}(2\xi+2k) + \widehat{\varphi}(2\xi+2k+1)\overline{\widehat{\widetilde{\varphi}}}(2\xi+2k+1)\\ &= \sum_k H\overline{\widetilde{H}}(\xi+k)\widehat{\varphi}\overline{\widehat{\widetilde{\varphi}}}(\xi+k) + H\overline{\widetilde{H}}(\xi+k+1/2)\widehat{\varphi}\overline{\widehat{\widetilde{\varphi}}}(\xi+k+1/2)\\ &= H\overline{\widetilde{H}}(\xi)\sum_k \widehat{\varphi}\overline{\widehat{\widetilde{\varphi}}}(\xi+k) + H\overline{\widetilde{H}}(\xi+1/2)\sum_k \widehat{\varphi}\overline{\widehat{\widetilde{\varphi}}}(\xi+k+1/2)\\ &= H\overline{\widetilde{H}}(\xi)\Phi(\xi) + H\overline{\widetilde{H}}(\xi+1/2)\Phi(\xi+1/2) = H\overline{\widetilde{H}}(\xi) + H\overline{\widetilde{H}}(\xi+1/2)\end{aligned}$$

since $\Phi \equiv 1$. A pair $H, \widetilde{H}$ of periodic functions

$$H\overline{\widetilde{H}}(\xi) + H\overline{\widetilde{H}}(\xi+1/2) \equiv 1 \tag{4.5}$$

is called a *quadrature mirror filter* (QMF) pair. The condition can be reexpressed in terms of the coefficients h_k and $\tilde{h}_k$ by interpreting (4.5) in terms of Fourier coefficients. That is, the k-th coefficient of $H\overline{\widetilde{H}}(\xi)$ is the convolution of the sequences $\mathbf{h}$ and $\widetilde{\widetilde{\mathbf{h}}}$ while that of $H\overline{\widetilde{H}}(\xi+1/2)$ is $(-1)^k\sum_\ell h_\ell \overline{\widetilde{h}}_{\ell-k}$ so the orthogonality condition can be written

$$\sum_k h_k \overline{\widetilde{h}}_{k-2\ell} = \delta_{k\ell}. \tag{4.6}$$

Spline scaling functions

The N-th order B-spline ϕ^N is the $N+1$-st order autoconvolution of the Haar function $\phi^0 = \mathbb{1}_{[0,1)}$. The filter $H^N(\xi) = ((1+\mathrm{e}^{-2\pi i\xi})/2)^{N+1}$. Consider the polynomial

$$P_L(y) = \sum_{\ell=0}^{L-1} \binom{L-1+\ell}{\ell} y^\ell$$

which is a solution of the *Bezout equation*

$$(1-y)^L P_L(y) + y^L P_L(1-y) = 1. \tag{4.7}$$

The change of variables $y = \sin^2 \pi\xi$ yields

$$\cos^{2L} \pi\xi P_L(\sin^2 \pi\xi) + \sin^{2L} \pi\xi P_L(\cos^2 \pi\xi) = 1$$

or, equivalently,

$$\Big(\frac{1+\mathrm{e}^{2\pi i\xi}}{2}\Big)^{2L} \mathrm{e}^{-2\pi iL\xi} P_L(\sin^2 \pi\xi) + \Big(\frac{1+\mathrm{e}^{2\pi i\xi}}{2}\Big)^{2L} (-1)^L \mathrm{e}^{-2\pi iL\xi} P_L(\cos^2 \pi\xi) = 1.$$

This suggests that, if $N+1 \le 2L$ the filter

$$\widetilde{H}^{N,L} = \Big(\frac{1+\mathrm{e}^{2\pi i\xi}}{2}\Big)^{2L-N-1} P_L(\sin^2 \pi\xi)\, \mathrm{e}^{-2\pi iL\xi}$$

satisfies the quadrature mirror filter pair condition (4.5) together with H^N.

When $N=1$ the function φ^1 is the piecewise linear spline $\varphi^1(t) = (1-|t-1|)_+$ that is equal to t on $[0,1)$, to $2-t$ on $[1,2)$ and zero elsewhere. In this case the smallest $L \ge [(N+1)/2]$ is $L=1$. However, this filter does not give rise to a square integrable scaling function. The smallest L that does so is $L=2$ yielding the filter

$$\widetilde{H}^{2,2}(\xi) = \Big(\frac{1+\mathrm{e}^{2\pi i\xi}}{2}\Big)^2 \big(1+2\sin^2 \pi\xi\big)\, \mathrm{e}^{-4\pi i\xi}$$

If we write $z = \mathrm{e}^{-2\pi i\xi}$ then

$$\begin{aligned}
\widetilde{H}^{2,2}(z) &= \Big(\frac{1+z^{-1}}{2}\Big)^2 \Big(1+2\Big(\frac{z^{1/2}-z^{-1/2}}{2i}\Big)^2\Big) z^2 \\
&= \Big(\frac{1+z}{2}\Big)^2 \Big(1-\frac{1}{2}\Big(z-2+\frac{1}{z}\Big)\Big) \\
&= \frac{1}{8z}(1+2z+z^2)(2z-(z^2-2z+1)) \\
&= -\frac{1}{8z}(1+2z+z^2)(1-4z+z^2) = -\frac{1}{8z}(1-2z-6z^2-2z^3+z^4)
\end{aligned}$$

Therefore $\widetilde{\mathbf{h}}$ is the sequence such that $(h_{-1},\ldots,h_3) = \frac{1}{8}(-1,2,6,2,1)$.

Exercise 4.3.10. Writing $H^*(z) = \overline{H}(1/\overline{z})$, show that

$$H^1 \widetilde{H}^{2,2}(z) + H^1 \widetilde{H}^{2,2}(-z) = 1$$

which is equivalent to (4.5) when $z = \mathrm{e}^{-2\pi i\xi}$.

Orthogonality

In some cases it is possible to take $\widetilde{H} = H$ in the condition (4.5) and then H is called a quadrature mirror filter and the scaling function φ is said to be an orthogonal scaling function since $\widetilde{\varphi} = \varphi$ in this case.

4.3.5 Regularity

It is desirable for scaling functions to have some continuity and even differentiability. In fact, the combination of regularity and orthogonality is a real challenge in building wavelets. The pointwise behavior of the scaling function φ turns out to depend on eigenvalues of certain matrices associated with the scaling sequence when the two-scale refinement operator acts on sample sequences. For example, suppose that $y_1, \dots, y_{M-1}$ are thought of as the integer values of the function f fed into (4.4). We can extend $\mathbf{y}$ to all integers by setting $y_\ell = 0$ if $\ell \notin \{1, \dots, M-1\}$. Then for a half integer $x = \ell/2$ we get

$$Tf(\ell/2) = 2\sum_{k=0}^{M} h_k f(2x-k) = 2\sum_{k=0}^{M} h_k f(\ell-k) = 2\sum_{k=0}^{M} h_k y_{\ell-k} = 2\sum_{k=0}^{\ell-1} h_k y_{\ell-k} = 2\sum_{k=1}^{\ell} h_{\ell-k} y_k.$$

Thus we can think of inputting $\mathbf{y}$ into the matrices

$$T_0 = \begin{pmatrix} h_0 & 0 & 0 & 0 \cdots & 0 & 0 \\ h_2 & h_1 & h_0 & 0 \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 \cdots & h_M & h_{M-1} \end{pmatrix} \quad \text{and} \quad T_1 = \begin{pmatrix} h_1 & h_0 & 0 & 0 \cdots & 0 & 0 \\ h_3 & h_2 & h_1 & 0 \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 \cdots & 0 & h_M \end{pmatrix}$$

we output upsamples representing the values of Tf at the points $\ell/2$ where $\ell = 1, \dots, (M+1)$ for the outputs from T_0 and $\ell = M+2, \dots 2M-1$ for the outputs from T_1. These represent samples of Tf on the interval $[0, (M+1)/2]$ and on $[M+2/2, M]$ respectively. We can input the outputs of T_0 into T_0 to define samples of $T^2 f$ on $[0, (M+1)/4]$ and into T_1 to define samples on $[(M+2)/4, (M+1)/2]$ and into T_1 to define the corresponding samples at quarter integers on the right half of the supporting interval $[0, M]$. Very roughly speaking, the regularity of φ depends on the largest eigenvalues of the matrices T_i. Suppose, for example, that the input vector $\mathbf{y}$ is an eigenvector of T_0 with eigenvalue $\lambda > 0$. Then the output of $T_0\mathbf{y}$ will be $\lambda\mathbf{y}$ but, as samples, $T_0\mathbf{y}$ will now *live* on $[0, (M+1)/2]$. Repeating one would have $T_0^N\mathbf{y} = \lambda^N\mathbf{y}$ but now representing samples on $[0, (M+1)/2^N]$. The difference between the maximum and minimum value on $[0, (M+1)/2^N]$ will then grow like λ^N and the *local Hölder exponent* will be the largest $\alpha > 0$ such that $(2^\alpha\lambda)^N$ remains bounded, or $2^\alpha\lambda \le 1$ or $\alpha = \log_2(1/\lambda)$.

Exercise 4.3.11. Verify that the filter

$$H(z) = \frac{1}{2(1+\nu^2)}\Big(\nu(1+\nu) + (1+\nu)z + (1-\nu)z^2 + \nu(1-\nu)z^3\Big)$$

where $\nu \in (0,1)$ corresponds to a QMF when we set $z = \mathrm{e}^{-2\pi i\xi}$. Explain why, with h_k twice the coefficient of z^k, $k = 0,1,2,3$, the scaling function defined by this filter is supported in $[0,3]$. Find the eigenvalues of the matrices T_0 and T_1 in this case and plot approximate scaling functions for a few values of ν and numerically estimate the Hölder exponent.

Calculation of compactly supported scaling functions: a diversion

The technique outlined for considering regularity also gives us a method for computing sample values of the scaling function. Suppose that one wishes to compute the scaling function φ that satisfies

$$\varphi(x) = 2\sum_{k=0}^{M} h_k \varphi(2x-k)$$

and that the solution is a continuous function defined on $[0, M]$. This says that the vector $[\varphi(1), \dots, \varphi(M-1)]^T$ is a one-eigenvector of the matrix $T(\ell, m) = h_{2\ell-m}$. Thus T is the matrix T_1 except for the last row and column. Now we input this sequence into T_0 and into T_1 to get the samples on corresponding dyadic intervals.

4.3.6 Multiple scaling functions

There is no reason why a scaling relation has to be built on a single function. It can be built on a vector function instead. This is not generalization for generalizations sake. Multiple scaling functions can actually achieve tradeoffs that single component scaling functions cannot. The following theorem due to Daubechies illustrates limitations of singly generated scaling functions.

Theorem 4.3.12. *If φ satisfies the scaling relation 4.3 and is continuous, compactly supported and orthogonal to its integer shifts, that is, $\langle \varphi, \varphi(\cdot - k)\rangle = \delta_{0k}$ then φ cannot be a symmetric function.*

There is no reason why a scaling relation has to be built on a single function. It can be built on a vector function instead. This is not generalization for generalizations sake. Multiple scaling functions can actually achieve tradeoffs that single component scaling functions cannot. The following theorem due to Daubechies illustrates limitations of singly generated scaling functions.

Theorem 4.3.13. *If φ satisfies the scaling relation 4.3 and is continuous, compactly supported and orthogonal to its integer shifts, that is, $\langle \varphi, \varphi(\cdot - k)\rangle = \delta_{0k}$ then φ cannot be a symmetric function.*

A *scaling vector* is a vector-valued function that satisfies a matrix analogue of the scaling equation (4.3). It is also worth reiterating here that the dilation factor 2 is by no means the only possible one can base scaling functions on much more general *expansive matrices*. Our goal here is not to cover the theory in its utmost generality but, rather, to give a particular family of examples to illustrate what is possible when one is willing to consider scaling vectors. A two-scale, two component scaling vector is then a solution of

$$\begin{pmatrix} \phi^1(t) \\ \phi^2(t) \end{pmatrix} = 2\sum_k C_k \begin{pmatrix} \phi^1(2t-k) \\ \phi^2(2t-k) \end{pmatrix}$$

in which C_k are 2×2 matrices.

In table 4.1 we give scaling coefficient matrices for biorthogonal scaling functions known as DGHM scaling functions due to their inventors Donovan, Geronimo, Hardin and Massopust [**?**].

Table 4.1. DGHM scaling and wavelet coefficients

C_{-2}	$\frac{1}{24}\begin{bmatrix} 0 & -(1+2s)\sqrt{2} \\ 0 & 0 \end{bmatrix}$	D_{-2}	$\frac{1}{24}\begin{bmatrix} 0 & -(1+2s)\sqrt{2} \\ 0 & -2-4s \end{bmatrix}$
C_{-1}	$\frac{1}{24}\begin{bmatrix} 8s-2 & (5-2s)\sqrt{2} \\ 0 & 0 \end{bmatrix}$	D_{-1}	$\frac{1}{24}\begin{bmatrix} 8s-2 & (5-2s)\sqrt{2} \\ (8s-2)\sqrt{2} & 10-4s \end{bmatrix}$
C_0	$\frac{1}{24}\begin{bmatrix} 12 & (5-2s)\sqrt{2} \\ 0 & 8+4s \end{bmatrix}$	D_0	$\frac{1}{24}\begin{bmatrix} -12 & (5-2s)\sqrt{2} \\ 0 & 4s-10 \end{bmatrix}$
C_1	$\frac{1}{24}\begin{bmatrix} 8s-2 & -(1+2s)\sqrt{2} \\ (8-8s)\sqrt{2} & 8+4s \end{bmatrix}$	D_1	$\frac{1}{24}\begin{bmatrix} 8s-2 & -(1+2s)\sqrt{2} \\ (2-8s)\sqrt{2} & 2+4s \end{bmatrix}$

The scaling functions can be plotted with Strela's `MWMP` software. The utility `multiplot` plots components of scaling vectors and wavelets given the starting sequence of coefficient matrices. The coefficient matrices are stored in a file `coef.m`. If the DGHM coefficients are not already in this file, you can add the lines

```
elseif strcmp(flt,'dghm')
s=-0.2;     % specify value of parameter s
L=sqrt(2)*[0, (-1-2*s)/(12*sqrt(2)), (-2+8*s)/(24),(5-2*s)/(12*sqrt(2)), 1/2,
(5-2*s)/(12*sqrt(2)), (-2+8*s)/(24),  (-1-2*s)/(12*sqrt(2));
0, 0,  0, 0,  0, (8+4*s)/24,  8*(1-s)/(12*sqrt(2)), (8+4*s)/24];

t=(1+2*s)/(5*s-2);
```

```
H=sqrt(2)*[0, -(1+2*t)*sqrt(2)/(24), (-2+8*t)/(24), (5-2*t)*sqrt(2)/(24), -1/2,
(5-2*t)*sqrt(2)/(24), (-2+8*t)/(24), -(1+2*t)*sqrt(2)/(24);
0, -(1+2*t)/(12), (-2+8*t)*sqrt(2)/(24),  (5-2*t)/(12), 0,
 -(5-2*t)/(12), (2-8*t)*sqrt(2)/(24), (1+2*t)/(12)];
```

You can then call the multiplot routine by

```
multiplot('dghm');
```

which will reproduce Figure 4.1.

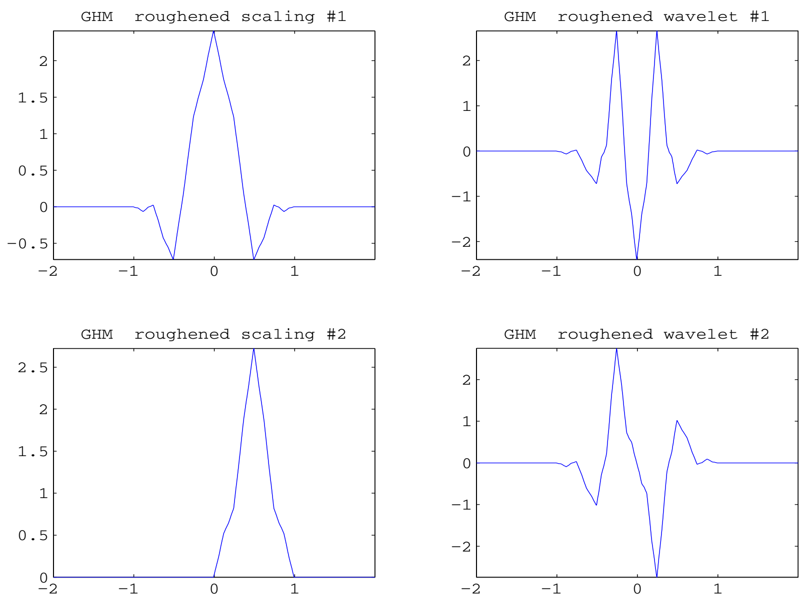

Fig. 4.1. Components of DGHM scaling function with $s = -0.2$ (left) and corresponding wavelet (right).

4.3.7 Regularity properties of DGHM scaling vectors

The scaling filters just discussed appear to come out of the blue. In fact, they arise in a somewhat natural way as we will discuss now. The tent function $T(x) = \mathbb{1}_{[0,1]} * \mathbb{1}_{[0,1]}$ is a scaling function since it is the autoconvolution of the Haar box function. But it is not orthogonal to its shifts and the known means of producing an alternative generator of $V(T)$ that is orthogonal to its shifts yields a generator that no longer has compact support. As an alternative, one considers building a larger shift invariant space that contains $V(T)$ by building a scaling vector whose components are both continuous, compactly supported functions in such a way that the scaling vector is orthogonal to its integer shifts. The tent function will be one component of the initial scaling vector. To build the second component w we consider a scaling equation

$$w(x) = \sum_k \big(a_k w(2x-k) + b_k T(2x-k)\big).$$

Suppose that we want w to be *minimally supported* in $[0,1]$. The only shifted tent function supported in $[0,1]$ is the $k=0$ term $T(2x)$ since so we should take $b_k = 0$ unless $k=0$. In addition, if w is supported in $[0,1]$ then only $w(2x)$ and $w(2x-1)$ will live inside $[0,1]$ in general so we should take $a_k = 0$ unless $k=0$ or $k=1$. This leaves us with

$$w(x) = a_0 w(2x) + a_1 w(2x-1) + bT(2x).$$

Since $T(2x)$ is symmetric with respect to $x = 1/2$, if we also want w to be symmetric with respect to $x = 1/2$ then we should take $a_0 = a_1$. This leaves us with

$$w(x) = a(w(2x) + w(2x-1)) + bT(2x).$$

Normalizing $b = 1$ we then will investigate the behavior of a solution to

$$w_s(x) = s(w(2x) + w(2x-1)) + T(2x) \tag{4.8}$$

as it depends on s. Consider the mapping

$$f \mapsto s(f(2x) + f(2x-1)).$$

A solution w of (4.8) can be reexpressed as

$$(I - \Gamma_s)w = h(2x-1).$$

The contraction mapping theorem [**?**] guarantees that a solution of this equation exists in a given metric space of functions provided that Γ_s is a contraction in that metric. The solution can then be computed by the Neumann series expansion $w = \sum_{n=0}^{\infty} \Gamma_s^n (h(2x-1))$ but we can compute the samples directly from the scaling equation.

The function space norm that we have in mind is the C^α Hölder (semi)-norm

$$\|f\|_\alpha = \sup_{0 \le x,y \le 1} \frac{|f(x) - f(y)|}{|x-y|^\alpha}$$

Proposition 4.3.14. *The mapping Γ_s is a contraction on the space $S_\alpha^{1/2}$, the subspace of $C^\alpha([0,1])$ consisting of those f symmetric with respect to $x = 1/2$ if $2^\alpha s < 1$.*

Proof. Symmetry plays an important role in proving the contraction property. First, by its very definition Γ_s maps $S_\alpha^{1/2}$ into itself. If x and y are both in $[0,1/2]$ then $2x$ and $2y$ are in $[0,1]$ so $f(2x-1) = 0 = f(2y-1$ and since $f \in C^\alpha$,

$$\begin{aligned}\frac{|\Gamma_s f(x) - \Gamma_s f(y)|}{|x-y|^\alpha} &= \frac{s(|f(2x) - f(2y)|}{|x-y|^\alpha} \\ &= s2^\alpha \frac{(|f(2x) - f(2y)|}{|2x-2y|^\alpha} \le s2^\alpha \|f\|_{C^\alpha}.\end{aligned}$$

A similar estimate applies when both $x > 1/2$ and $y > 1/2$. Consider then the case in which $0 \le x \le 1/2$ and $1/2 < y \le 1$. Then $f(2x-1) = 0 = f(2y)$. However, since $f \in S_\alpha^{1/2}$ one has $f(t) = f(1-t)$ for $t \in [0,1]$ so

$$\begin{aligned}\frac{|\Gamma_s f(x) - \Gamma_s f(y)|}{|x-y|^\alpha} &= \frac{s(|f(2x) - f(2y-1)|}{|x-y|^\alpha} \\ &= s\frac{(|f(2x) - f(2y-1)|2^\alpha|x+y-1|^\alpha}{|x-y|^\alpha|2x-2(1-y)|^\alpha} \\ &= 2^\alpha s \frac{(|f(2x) - f(2-2y)|}{|2x-2(1-y)|^\alpha}\frac{(|x+y-1|^\alpha}{|x-y|^\alpha} \le s2^\alpha \|f\|_{C^\alpha}\end{aligned}$$

where the last inequality follows since $2-2y \in [0,1]$ and $|x-(1-y)| = |x-1/2-(1/2-y)| \le |1/2-x| + |y-1/2| = |x-y|$ whenever $x \le 1/2$ and $y > 1/2$. This proves the proposition.

4.3.8 Biorthogonal generators

Now think of s as a parameter and consider the problem of finding generators $\{\varphi_s^1, \varphi_s^2\}$ and *dual pairs* s and $\tilde{s}$ such that the shifts of the functions φ_s^i and $\varphi_{\tilde{s}}^j$ are biorthogonal, that is,

$$\langle \varphi_s^i(\cdot - k),\, \varphi_{\tilde{s}}^j(\cdot - \ell)\rangle = \delta_{k\ell}\delta_{i,j}$$

The generators have to be linear combinations of shifts of T and w from which T and w can be recovered so, for the sake of simplicity we will seek numbers γ, β, and δ such that

$$\varphi_s^1(x) = \gamma(T(x) - \beta(w_s(x) + w_s(x+1))) \qquad \text{and} \qquad \varphi_s^2 = \delta w_s. \tag{4.9}$$

Solving for the parameters γ, β and δ leads one to the choice of coefficients in Table 4.1. The generators are biorthogonal when $\tilde{s} = \frac{5s-2}{1+2s}$.

Exercise 4.3.15. Let $H_s(z) = \sum C_k(s) z^k$ with $C_k(s)$ as in table 4.1. Show that the *conditions of biorthogonality* hold, that is, $H_s(z)H_{\tilde{s}}^*(z) + H_s(-z)H_{\tilde{s}}^*(-z) = I$. This identity is best verified using Maple or some other symbolic package.

Open Problem Find, if possible, a parametric family of shift invariant spaces with three (bi)orthogonal generators such that each generator is is symmetric and minimally supported and has two bounded derivatives.

Such generators would be desirable for numerical solution of certain boundary value problems.

4.3.9 Differentiation of DGHM spaces

If a scaling function is differentiable then its derivative is another scaling function. This is because of *distribution theory*, one of whose tenets is that differentiation is convolution with the derivative of the *Dirac delta* which is itself a scaling distribution. Since convolutions of scaling functions or distributions also satisfy a two-scale relation, the claim follows. But what does the derivative of a scaling function look like? In particular, what is the scaling filter? In the [?, **?**]

Table 4.2. Smoothed and roughened DGHM scaling coefficients

C_{-2}^-	$\frac{1+2s}{24}\begin{bmatrix}-1 & 1\\ -1 & 1\end{bmatrix}$		
C_{-1}^-	$\frac{1}{24}\begin{bmatrix}6 & 20s-8\\ 6 & 20s-8\end{bmatrix}$	C_{-1}^+	$\frac{1}{24}\begin{bmatrix}6 & 2(s-1)\\ 9 & 3(2s-1)\end{bmatrix}$
C_0^-	$\frac{1}{24}\begin{bmatrix}14+4s & 0\\ 0 & 14+4s\end{bmatrix}$	C_0^+	$\frac{1}{24}\begin{bmatrix}12 & 0\\ 0 & 6\end{bmatrix}$
C_1^-	$\frac{1}{24}\begin{bmatrix}6 & 8-20s\\ -6 & 20s-8\end{bmatrix}$	C_1^+	$\frac{1}{24}\begin{bmatrix}6 & 2(1-s)\\ -9 & 3(2s-1)\end{bmatrix}$
C_2^-	$\frac{1+2s}{24}\begin{bmatrix}-1 & -1\\ 1 & 1\end{bmatrix}$		

Exercise 4.3.16. Use the `MWMP` package to plot the multiple scaling functions and verify the relationship between the smoothened scaling functions and the original DGHM scaling functions.

Table 4.3. Smoothed and roughened DGHM wavelet coefficients

D^-_{-2}	$\frac{1+2s}{12}\begin{bmatrix} -\sqrt{2} & \sqrt{2} \\ -2 & 2 \end{bmatrix}$		
D^-_{-1}	$\frac{1}{12}\begin{bmatrix} 6\sqrt{2} & (20s-8)\sqrt{2} \\ 12 & 40s-16 \end{bmatrix}$	D^+_{-1}	$\frac{1}{48}\begin{bmatrix} \frac{3\sqrt{2}}{2} & (4s-1)\frac{\sqrt{2}}{2} \\ 3 & (4s-1) \end{bmatrix}$
D^-_0	$\frac{1}{12}\begin{bmatrix} 0 & (4s-34)\sqrt{2} \\ 8s-20 & 0 \end{bmatrix}$	D^+_0	$\frac{1}{48}\begin{bmatrix} 0 & -3\sqrt{2} \\ -6 & 0 \end{bmatrix}$
D^-_1	$\frac{1}{12}\begin{bmatrix} -6\sqrt{2} & (20s-8)\sqrt{2} \\ 12 & 16-40s \end{bmatrix}$	D^+_1	$\frac{1}{48}\begin{bmatrix} \frac{-3\sqrt{2}}{2} & (4s-1)\frac{\sqrt{2}}{2} \\ 3 & (1-4s) \end{bmatrix}$
D^-_2	$\frac{1+2s}{12}\begin{bmatrix} \sqrt{2} & \sqrt{2} \\ -2 & -2 \end{bmatrix}$		

Consider the matrix

$$R(z) = \begin{bmatrix} 0 & 2\sqrt{2} \\ 1-z & -1-z \end{bmatrix}. \tag{4.10}$$

Define the scaling filter

$$H(z) = \sum_k C_k z^k \tag{4.11}$$

Let $H_+(z)$ denote the corresponding filter of the smoothened wavelets and $H_-(z)$ the filter of the roughened wavelets.

Theorem 4.3.17. *Let $-1/2 < s < 1/7$ so that Φ_s is Lipschitz continuous while $-1 < \tilde{s} < 0$ and hence $\Phi_{\tilde{s}}$ is continuous. Let $R(z)$ be given by (4.10) and set:*

$$H_s^-(z) \;=\; 2\,R^{-1}(z^2)\,H_s(z)\,R(z), \quad H_{\tilde{s}}^+(z) \;=\; \frac{1}{2}\,R^*(z^2)\,H_{\tilde{s}}(z)\,(R^*)^{-1}(z).$$

Then the new filters satisfy the biorthogonality condition:

$$H_{\tilde{s}}^+(z)\,(H_s^-)^*(z) \,+\, H_{\tilde{s}}^+(-z)\,(H_s^-)^*(-z) \;=\; I.$$

4.3.10 From scaling functions to wavelets: multiresolution analysis

Definition 4.3.18. *A multiresolution analysis of $L^2(\mathbb{R})$ is a sequence of closed subspaces V_j, $j \in \mathbb{Z}$ having the following properties:*

- *MRA1* $V_j \subset V_{j+1}$
- *MRA2* $V_{j+1} = \{f(2\cdot) : f \in V_j\}$
- *MRA3* $\cup_j V_j$ *is dense in* $L^2(\mathbb{R})$ *and*
- *MRA4* $\cap_j V_j = \{0\}$.
- *MRA5 There is a $\varphi \in V_0$ whose shifts $\varphi(\cdot - k)$ form an orthonormal basis for V_0.*

This is the usual definition of an MRA. Conditions (MRA1), (MRA2) and (MRA5) imply that φ is an orthogonal scaling function. It is useful also to extend the multiresolution framework to include biorthogonal scaling functions such as the spline functions outlined above.

Definition 4.3.19. *A pair of sequences V_j and $\widetilde{V}_j$ of closed subspaces of $L^2(\mathbb{R})$ is said to form a biorthogonal pair of multiresolution analyses provided that both V_j and $\widetilde{V}_j$ satisfy (MRA1)–(MRA4) but condition (MRA5) is replaced by the existence of functions φ and $\widetilde{\varphi}$ whose shifts form Riesz bases for $V(\varphi)$ and $\widetilde{V}_0 = V(\widetilde{\varphi})$ respectively and such that φ and $\widetilde{\varphi}$ are shift biorthogonal.*

Now consider the problem of characterizing those functions in $V_1(\varphi)$ that are orthogonal to $V_0(\widetilde{\varphi})$. In other words, we wish to identify the orthogonal complement of $V_0(\widetilde{\varphi})$ inside of $V_1(\varphi)$. Define W_0 to be the orthogonal complement of $V_0(\widetilde{\varphi})$ inside of $V_1(\varphi)$. Then W_0 is shift invariant, that is, $g \in W_0$ if and only if $g(\cdot - 1) \in W_0$. If $g \in W_0$ then $g(x) = \sum_{k=-\infty}^{\infty} c_k \varphi(2x-k)$ so that $\hat{g}(\xi) = C(\xi/2)\hat{\varphi}(\xi/2)$ with $C(z) = \sum_k c_k z^k$ and

$$\begin{aligned}
\langle g(\cdot - \ell),\, \widetilde{\varphi}\rangle &= \frac{1}{2}\int \mathrm{e}^{-2\pi i \ell \xi} C\Big(\frac{\xi}{2}\Big)\overline{\widetilde{H}\Big(\frac{\xi}{2}\Big)}\hat{\varphi}\Big(\frac{\xi}{2}\Big)\overline{\widetilde{\varphi}\Big(\frac{\xi}{2}\Big)}\, d\xi \\
&= \int_0^1 \mathrm{e}^{-4\pi i \ell \xi} C(\xi)\overline{\widetilde{H}(\xi)} \sum_{k=-\infty}^{\infty} \hat{\varphi}\Big(\xi + k\Big)\overline{\widetilde{\varphi}\Big(\xi + k\Big)}\, d\xi \\
&= \int_0^{1/2} \Big(C(\xi)\overline{\widetilde{H}(\xi)}(\xi) + C(\xi + 1/2)\overline{\widetilde{H}(\xi)}(\xi + 1/2)\Big)\, \mathrm{e}^{-4\pi i \ell \xi}\, d\xi = 0.
\end{aligned}$$

Since this is true for any ℓ and since the functions $\mathrm{e}^{4\pi i \ell \xi}$ form a complete orthogonal set on $[0, 1/2)$ it must be that

$$C(\xi)\overline{\widetilde{H}(\xi)} + C(\xi + 1/2)\overline{\widetilde{H}(\xi + 1/2)} = 0$$

on $[0, 1/2)$. In other words,

$$\begin{matrix} C(\xi) \\ C(\xi + 1/2) \end{matrix} = M(\xi) \begin{matrix} \overline{\widetilde{H}(\xi + 1/2)} \\ -\overline{\widetilde{H}(\xi)} \end{matrix}$$

where $M(\xi)$ is a scalar function on $[0, 1/2)$. Replacing ξ by $\xi + 1/2$ it also follows that $M(\xi + 1/2) = -M(\xi)$ so that $M(\xi) = N(2\xi)\mathrm{e}^{2\pi i \xi}$ where N is otherwise an arbitrary element of $L^2(\mathbb{T})$. In summary, any element $g \in W_0$ can be written

$$\hat{g}(\xi) = \mathrm{e}^{\pi i \xi}\overline{\widetilde{H}(\xi/2 + 1/2)}N(\xi)\hat{\varphi}(\xi/2).$$

Example 4.3.20. Taking $N(\xi) = -1$ above we have that $\hat{\psi}(\xi) = -\mathrm{e}^{\pi i \xi}\overline{\widetilde{H}(\xi/2 + 1/2)}\hat{\varphi}(\xi/2)$ defines a function $\psi(x) = \sum_k (-1)^k \overline{\widetilde{h}}_{1-k}\varphi(2x - k)$ in W_0.

Exercise 4.3.21. Let ψ be defined as in (4.3.20) and let $\widetilde{\psi}$ be defined completely analogously as $\widetilde{\psi}(t) = \sum_k (-1)^k \overline{h}_{1-k}\widetilde{\varphi}(2t - k)$. Show that $\psi_{jk}(t) = 2^{j/2}\psi(2^j t - k)$ and $\widetilde{\psi}_{jk}(t) = 2^{j/2}\widetilde{\psi}(2^j t - k)$ form biorthogonal families in the sense that $\langle \psi_{jk},\, \widetilde{\psi}_{j'k'}\rangle = \delta_{jj'}\delta_{kk'}$.

In general, if filters $H, \widetilde{H}, G, \widetilde{G}$ satisfy

$$\begin{pmatrix} H(z) & G(z) \\ H(-z) & G(-z) \end{pmatrix} \begin{pmatrix} \widetilde{H}(1/z) & \widetilde{H}(-1/z) \\ \widetilde{G}(1/z) & \widetilde{G}(-1/z) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{4.12}$$

then the corresponding scaling and wavelet functions satisfy the corresponding conditions of biorthogonality

$$\begin{aligned}
&\langle \varphi(\cdot - k,\, \widetilde{\varphi}(\cdot - \ell)\rangle = \delta_{k\ell} \qquad &&\langle \psi(\cdot - k,\, \widetilde{\varphi}(\cdot - \ell)\rangle = 0 \\
&\langle \psi(\cdot - k,\, \widetilde{\psi}(\cdot - \ell)\rangle = \delta_{k\ell} \qquad &&\langle \phi(\cdot - k,\, \widetilde{\psi}(\cdot - \ell)\rangle = 0
\end{aligned}$$

Laurent polynomials

A Laurent polynomial has the form $L(z) = \sum_{k=N}^{M} a_k z^k$, that is, L is a polynomial in the variables z and $1/z$. When $z = \mathrm{e}^{2\pi i \xi}$ one can, with a slight abuse of notation, write $L(\xi) = L(z) = \sum_{k=N}^{M} a_k \mathrm{e}^{2\pi i k \xi}$. One then says that $L(\xi)$ is a finite filter. The Laurent degree of a Laurent polynomial L is $D(L) = M - N$ ($a_N \neq 0 \neq a_N$). This is different from the usual degree $\deg p$ of an ordinary polynomial. For example, a monomial z^N has Laurent degree zero. Nevertheless, just as with ordinary polynomials, there is a Euclidean algorithm for Laurent polynomials.

Theorem 4.3.22. *If p and q are Laurent polynomials with $D(p) \geq D(q)$ then starting with $p_0 = p$, $q_0 = q$ and letting $p_{m+1} = q_m$ and letting q_{m+1} denote the remainder from dividing q_m into p_m, let n be the smallest integer such that $q_n = 0$. Then p_n is the greatest common divisor of p and q.*

Writing $p_m = t_m q_m + q_{m+1}$, the Euclidean algorithm can be expressed in matrix form

$$\begin{pmatrix} p_1(z) \\ p_0(z) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ t_0(z) & 1 \end{pmatrix} \begin{pmatrix} q_0(z) \\ q_1(z) \end{pmatrix} \cdots \begin{pmatrix} p_n(z) \\ p_{n-1}(z) \end{pmatrix} \prod_{m=1}^{n-1} \begin{pmatrix} 1 & 0 \\ t_m(z) & 1 \end{pmatrix} \begin{pmatrix} q_m(z) \\ q_{m+1}(z) \end{pmatrix} \tag{4.13}$$

The factorization of a Laurent polynomial into lesser degree terms is not unique. For example,

$$z + \frac{1}{z} = \Big(z+1\Big)\Big(1+\frac{1}{z}\Big) - 2 = \Big(2z+1\Big)\Big(\frac{1}{2}+\frac{1}{z}\Big) - \frac{5}{2}.$$

We will return to the question of decompositions of Laurent polynomials and their role in wavelet theory momentarily.

Completeness: The τ-cycle condition

Not every quadrature mirror filter gives rise to an orthogonal scaling function. When H is a trigonometric polynomial, the product $\prod_{j=1}^{\infty} H(\xi/2^j)$ defines an orthogonal scaling function provided that H does not have a τ-cycle of zeros in $[0,1)$, that is, there is not a set of zeros $\{\xi_1,\dots,\xi_N\}$ that is invariant under the mapping $\tau : \xi \mapsto 2\xi \mod 1$.

Example 4.3.23. The filter $H(\xi) = (1 + \mathrm{e}^{-6\pi i\xi})/2$ has zeros at $\xi = 1/3$ and $\xi = 2/3$ which form a τ-cycle. The function $\phi = \frac{1}{\sqrt{3}}\mathbb{1}_{[0,3)}$ satisfies $\phi(t) = \phi(2t) + \phi(2t-3)$ so that $\hat{\phi}(\xi) = \Big(\frac{1}{2}\hat{\phi}(\xi/2) + \frac{1}{2}\mathrm{e}^{-3\pi i\xi}\Big)\hat{\phi}(\xi/2) = H(\xi/2)\hat{\varphi}(\xi/2)$.

There are other nontrivial features of the mathematical connection between multiresolution analysis and orthonormal or biorthogonal wavelets. For example, there exist orthonormal wavelet bases for $L^2(\mathbb{R})$ that are not generated by a multiresolution analysis. Such wavelets have been characterized in terms of a so-called dimension function [**?**].

Orthonormal wavelet bases

In the case of a standard multiresolution analysis, effectively one has $\varphi = \widetilde{\varphi}$ and the wavelet function ψ generates an orthonormal wavelet basis for $L^2(\mathbb{R})$. Property MRA2 implies that V_{j+1} consists of all dyadic dilates of functions in V_j. Since $V_0 \oplus W_0 = V_1$ we can rescale this as $W_1 = \{g(2x) : g(x) \in W_0\}$ where W_1 is the orthogonal complement of V_1 inside V_2 and concludes that $V_2 = V_1 \oplus W_1 = V_0 \oplus W_0 \oplus W_1$ and continuing in this way, $V_N = W_{N-1} \oplus \cdots \oplus W_0 \oplus W_{-N} \cdots \oplus W_{-M}$. Now using properties MRA3 and MRA4 we can conclude that $L^2(\mathbb{R}) = \oplus_{j=-\infty}^{\infty} W_j$. Since the function $\psi_{jk}(x) = 2^{j/2}\psi(2^j x - k)$ form an orthonormal basis for W_j as k ranges over $\mathbb{Z}$ it follows that ψ_{jk} forms an orthonormal basis for $L^2(\mathbb{R})$ as j,k range over $\mathbb{Z}$.

4.3.11 Fast wavelet transform

The *discrete wavelet transform* treats a discrete input sequence $\{c_k\}$ as though it is the sequence of coefficients of an element $f \in V(\varphi)$, namely, $f_{\mathbf{c}} = \sum_k c_k \varphi(\cdot - k)$ and proceeds to decompose $\{a_k\}$ into sequences c_k^j and d_k^j in which c_k^j is the projection of f onto $V_{-j}(\varphi)$ and d_k^j is the projection onto $W_{-j}(\varphi)$. From biorthogonality, the coefficients are those of shifts of φ with shifts and dilates of $\widetilde{\varphi}$ and $\widetilde{\psi}$ respectively. Letting $\{\widetilde{h}_k\}$ denote the filter sequence of $\widetilde{H}$ and $\{\widetilde{g}_k\}$ that of $\widetilde{G}$ where $\widehat{\widetilde{\psi}}(2\xi) = \widetilde{G}(\xi)\widehat{\widetilde{\psi}}(\xi)$, one has

$$(\widetilde{\mathcal{H}}\mathbf{c})_k = 2\sum_{\ell} \overline{\widetilde{h}}_{\ell-2k} c_k; \qquad (\widetilde{\mathcal{G}}\mathbf{c})_k = 2\sum_{\ell} \overline{\widetilde{g}}_{\ell-2k} c_k$$

These mappings correspond to coefficient mappings from $V_0(\varphi)$ to $V_{-1}(\varphi)$ and $W_{-1}(\varphi)$ respectively. To recover the original sequence one uses the mappings

$$(\mathcal{H}^*\mathbf{c})_k = 2\sum_{\ell} \overline{h}_{k-2\ell} c_k; \qquad (\mathcal{G}^*\mathbf{c})_k = 2\sum_{\ell} \overline{g}_{k-2\ell} c_k.$$

The discrete wavelet transform boils reiterates the filtering operators $\widetilde{\mathcal{H}}$ and $\widetilde{\mathcal{G}}$ respectively on the sequences c^1, c^2 etc. which are the outputs of corresponding powers of $\widetilde{\mathcal{H}}$.

The so-called *fast wavelet transform* actually corresponds to periodized versions of the discrete wavelet transform.

Scaling function

The scaling function can be computed by applying the cascade algorithm to the sequence $\mathbf{c} = \delta_0$. A java applet can be found at `Wim Swelden's wavelet cascade applet`.

Discrete implementations 1: FWT matrix

We are going to forego a discussion of practical implementations of the discrete wavelet transform because, in fact, there is a more general approach that also addresses some practical implementation issues that arise when trying to process long data sequences. The method is called *lifting.* For C programmers, there is a package called `liftpack` that implements lifting and can be applied to image processing, as does the package `waili` Nonetheless, the discussion of lifting does involve some technicalities and, for those who want to cut to the chase and apply wavelet transforms to data with very little effort, `Wavelab802`, which is discussed in the appendix to this chapter, is recommended.

Discrete implementations 2: Lifting

Given one finite filter pair satisfying (4.12) and fixing H, all other finite filter pairs satisfying (4.12) were characterized by Vetterli and Herley [**?**] as follows

Theorem 4.3.24. *Suppose that $H(z), \widetilde{H}(z)$ is a finite QMF pair as in (4.5) and suppose that $G(z), \widetilde{G}(z)$ are corresponding high pass filters such that (4.12). Then, up to a monomial in z, any other biorthogonal extension $\{H, G^{\mathrm{new}}, \widetilde{H}^{\mathrm{new}}, \widetilde{G}^{\mathrm{new}}\}$ has the form*

$$\begin{aligned} G^{\mathrm{new}}(z) &= G(z) + s(z^2)H(z) \\ \widetilde{H}^{\mathrm{new}} &= \widetilde{H}(z) - \overline{s(\bar{z}^{-2})}\widetilde{G}(z) \\ \widetilde{G}^{\mathrm{new}}(z) &= \widetilde{G}(z). \end{aligned}$$

The significance of z^2 is that it allows one to express, for example,

$$\psi^{\mathrm{new}} = \psi + \sum_k s_k \varphi(\cdot - k); \qquad s(z) = \sum s_k z^k.$$

which is orthogonal to $\widetilde{W_0} = W_0(\widetilde{\psi})$. Sweldens [**?**] refers to the mapping $G \mapsto G^{\mathrm{new}}$ as a *lifting step* in the sense that it replaces G by a Laurent polynomial of higher or *lifted* degree.

The discrete wavelet transform just encountered can be summarized in the following diagram, with the left part of the diagram corresponding to the wavelet transform/decomposition and the right hand side corresponding to the inverse wavelet transform/reconstruction. Processing done on the wavelet transform coefficients can be inserted in between.

We will look more closely at the conditions of biorthogonality (4.12), namely

$$\begin{aligned} H(z)\widetilde{H}(1/z) + G(z)\widetilde{G}(1/z) &= 1 \\ H(z)\widetilde{H}(-1/z) + G(z)\widetilde{G}(-1/z) &= 0 \end{aligned}$$

The first key step to *lifting* is the polyphase matrix

$$P(z) = \begin{pmatrix} H_{\mathrm{e}}(z) & H_{\mathrm{o}}(z) \\ G_{\mathrm{e}}(z) & G_{\mathrm{o}}(z) \end{pmatrix}$$

where $H(z) = H_{\mathrm{e}}(z) + zH_{\mathrm{o}}(z)$ is the decomposition into even and odd powers. Then the outputs λ and γ in Figure 4.2 can be expressed as

$$\begin{pmatrix} \lambda(z) \\ \gamma(z) \end{pmatrix} = \widetilde{P}(z) \begin{pmatrix} x_{\mathrm{e}}(z) \\ z^{-1}x_{\mathrm{o}}(z) \end{pmatrix}.$$

When the conditions of biorthogonality are satisfied one can then express the reconstruction as ($Y = X$ in Figure 4.2)

$$\begin{pmatrix} x_{\mathrm{e}}(z) \\ zx_{\mathrm{o}}(z) \end{pmatrix} = P(z) \begin{pmatrix} \lambda(z) \\ \gamma(z) \end{pmatrix}$$

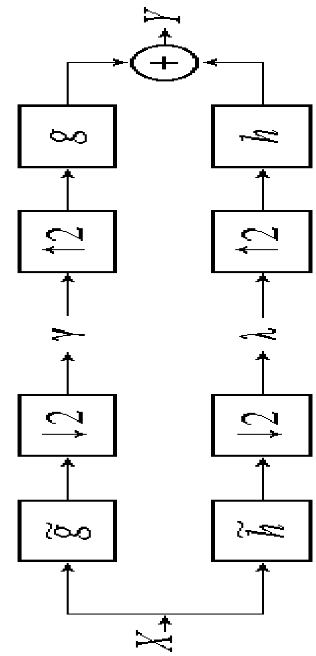

Fig. 4.2. Subband coding scheme equivalent to discrete wavelet transform (left) and inverse discrete wavelet transform (right).

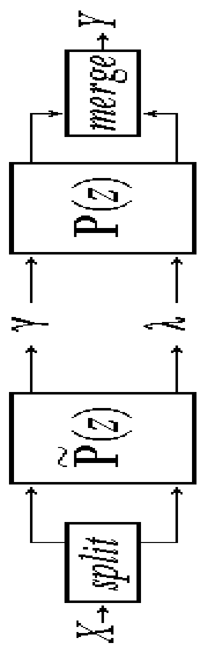

Fig. 4.3. Polyphase matrix version of subband coding (right).

Then one can replace Figure 4.2 by Figure 4.3. The significance of the polyphase approach is that no downsampling is required, hence only half the number of multiplications required for the convolution-decimation scheme of Figure 4.2 are need in the decomposition and reconstruction stages of Figure 4.3.

The multiplications by upper triangular matrices correspond to lifting steps. In particular, application of Theorem 4.3.22 allows one to express the polyphase matrix $P(z)$ in the form

$$P(z) = \begin{pmatrix} K_1 & 0 \\ 0 & K_2 \end{pmatrix} \prod_{m=1}^{n} \begin{pmatrix} 1 & s_m(z) \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ t_m(z) & 1 \end{pmatrix} = K \prod_{m=1}^{n} S_m(z) T_m(z). \tag{4.14}$$

where K_i are scalars. Each such lifting step – multiplication by S_m corresponds to an operation of the form in Herley and Vetterli's characterization of biorthogonal filters. Because the factorization of Laurent polynomials is not unique, some choices can be made in the lifting factorization.

Exercise 4.3.25. Given the filters $\widetilde{H}(z) = -z^{-2}/8 + z^{-1/4+3/4+z/4-z^2/8}$ and $\widetilde{G}(z) = z^{-2}/4 - z^{-1}/2 + 1/4$, write the polyphase matrix $\widetilde{P}(z)$ for this filter pair and factorize $\widetilde{P}$ into lifting matrices. Finally, identify the corresponding matrix P and, thus, corresponding dual filters.

Discrete implementations 3: Prefiltering

When $f \in V_0$, $f(t) = \sum_k c_k \varphi(t-k)$, typically it is not true that $c_k = f_k$, though it is true for certain biorthogonal scaling function pairs. The mapping SC from the sample sequence $\{f(k)\}$ to the coefficient sequence $\{c_k\}$ is linear and covariant in the sense that $SC(\{f(k-\ell)\}_k = SCf_{k-\ell}$. In principle, then, one just needs to know the image of δ_{0k}. It is easier to work in reverse and ask which combination $\{c_k\}$ maps to δ_{0k}. That is, $\sum c_k \varphi(-k) = 1$

BiMRAs, spline wavelets and dual wavelets

In some cases it is convenient to work with biorthogonal scaling functions. In the definition of an MRA one replaces (MRA5) – the condition that $\varphi(\cdot - k)$ are orthogonal – by the condition that they form a Riesz basis for their span. That is, there is a constant c such that $\frac{1}{C}\|\{c_k\}\|_{\ell^2}^2 \le \|\sum_k c_k \varphi(\cdot - k)\|_{L^2}^2 \le C\|\{c_k\}\|_{\ell^2}^2$. A pair of such MRAs with generating scaling functions $\varphi(t) = 2\sum h_k \varphi(2t-k)$ and $\tilde{\varphi}(t) = 2\sum_\ell \tilde{h}_\ell \tilde{\varphi}(\cdot - \ell)$ is said to form a biorthogonal pair provided that $\langle \varphi(\cdot - k), \tilde{\varphi}(\cdot - \ell)\rangle = \delta_{k\ell}$.

Exercise 4.3.26. Show that if a scaling function pair φ and $\tilde{\varphi}$ is a biorthogonal pair then, denoting its filters by H and $\tilde{H}$ respectively and setting $G(\xi) = \mathrm{e}^{-2\pi i \xi}\overline{H(\xi)}$ then the conditions of biorthogonality

$$\begin{pmatrix} H(\xi) & H(\xi+1/2) \\ G(\xi) & G(\xi+1/2) \end{pmatrix} \begin{pmatrix} \overline{\tilde{H}(\xi)} & \overline{\tilde{G}(\xi)} \\ \tilde{H}(\xi+1/2) & \tilde{G}(\xi+1/2)) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{4.15}$$

Orthonormal wavelets and compactly supported wavelets

Wavelet design

4.4 Wavelets in two dimensions and wavelet image compression

While several constructions of scaling functions and wavelets in two-variables are known, the one that is used the most is a tensor product construction, meaning that the multiresolution spaces for functions of two variables are constructed in terms of products of one dimensional wavelets. The scaling function then has the form $\varphi(x,y) = \phi(x)\phi(y)$ where ϕ is a scaling function of one variable. If ϕ is an orthogonal scaling function then the shifts $\varphi_{k\ell}(x,y) = \phi(x-k)\phi(y-\ell)$ form an orthogonal family of functions in $L^2(\mathbb{R}^2)$. Denote by $V(\varphi) = V_0(\varphi)$ the closed subspace of $L^2(\mathbb{R}^2)$ generated by the $\mathbb{Z}^2$ shifts $\varphi_{k\ell}$ of φ and define $V_1(\varphi) = \{f(2x) : f \in V_0(\varphi)\}$. In the one-dimensional case the space $V_1 \setminus V_0 = W_0$ was generated by a single function ψ but in the bivariate case $V_1 \setminus V_0$ is generated by the $\mathbb{Z}^2$ shifts of three functions, namely $\psi^1(x,y) = \phi(x)\psi(y)$, $\psi^2(x,y) = \psi(x)\phi(y)$ and $\psi^3 = \psi(x)\psi(y)$.

Exercise 4.4.1. Show that the functions $\psi^\nu(x,y)$, $\nu = 1,2,3$ and their $\mathbb{Z}^2$ shifts are orthogonal to one another with respect to the standard inner product

$$\langle f,\, g\rangle = \int\int f(x,y)\,\overline{g(x,y)}\,dx\,dy$$

on $L^2(\mathbb{R}^2)$.

Exercise 4.4.2. Explain why, with φ and ψ as above, the functions $\psi^\nu_{jk}(x,y) = 2^{j/2}\psi^\nu(2^j x - k_1, 2^j y - k_2)$ form an orthonormal basis for $L^2(\mathbb{R}^2$ as j ranges over $\mathbb{Z}$, $k \in \mathbb{Z}^2$ and $\nu = 1,2,3$.

4.5 Curvelets and other time-frequency tools