CHAPTER 11

Time-Scale Signal Transforms

Petroleum seismologists discovered the modern form of the continuous wavelet transform in the mid-1980s. For some time, researchers had been using time-frequency transforms—such as the Gabor transform and its broader family of short-time Fourier transforms—for analyzing signals containing localized frequency components. Speech and seismic waveforms are representative examples. Windowed Fourier analysis becomes problematic, however, when the structure of the signal involves transients of varying scale. Then the short-time tools behave more like the global Fourier transform, and their approximations converge poorly. One idea put forward to improve decomposition convergence was to replace the frequency variable with a scale parameter in the transform relation. The basis functions for this new method were shifted and dilated versions of each other. So they looked like little waves: *wavelets*.

This research caught the eye of mathematicians who found that the new technique held a wealth of special properties. It could be discretized. Wavelets were close kin to theoretical tools used in the study of singular integral operators (mathematical physics), the frame signal decomposition structure (harmonic analysis), quadrature mirror filters (communication theory), and the scale space representation (signal and image analysis). And against the intuition of all theoreticians of the time, there were found orthonormal bases for the L^2 Hilbert space that consisted of smooth, rapidly decaying, similarly shaped elements: *orthonormal wavelets*.

This chapter develops both continuous and discrete scale-based transforms. The topics include the continuous wavelet transform; further development of the idea of frames, which we covered in Chapters 3 and 10; the concept of multiresolution analysis; orthogonal wavelets; discrete wavelet transforms; and, finally, the construction of multiresolution analyses and orthogonal wavelets. Wavelet decomposition furnishes an alternative approach for describing signal structure.

There is a rich research literature on wavelet transforms, including a history of the discipline [1] and many excellent introductory treatments [2–12].

Signal Analysis: Time, Frequency, Scale, and Structure, by Ronald L. Allen and Duncan W. Mills
ISBN: 0-471-23441-9

11.1 SIGNAL SCALE

In a variety of signal analysis applications we have taken note of the problems that arise due to the scale of signal features. The idea is that of the extent of recognizable portions of the signal or the width of regions of interest within the signal may vary over time. The scale of signal features affects the behavior of such signal analysis elements as edge detectors, shape detectors, and local frequency identification algorithms.

From several standpoints we have attempted to interpret signals, and from each of them we had to deal with the issue of scale in a rather informal manner. For instance, in Chapter 4 we pursued time domain techniques for understanding signals. Scale issues affect edge and peak detection, obviously, and even the decision about how wide noise removal filters should be must take into account the size of objects sought within the signal. In Chapter 9 we designed filters to find periodicities within signals. But when such oscillatory components are localized—and so they often are in natural signals—then the extent of the oscillatory phenomenon affects the outcome of the analysis.

The dilation of a function has the same basic shape. For scale-based signal analysis we generally use translations and *dilations* or *scalings* of a basic signal $\psi(t)$:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right). \tag{11.1}$$

In the next section, we shall show, following Grossmann and Morlet [13], that dilations (variations in parameter a) and translations (variations in b) support a decomposition of a general function $\psi(t)$.

In the previous chapter we covered time-frequency transforms, which combine time and frequency information in the transformed signal. These time-frequency transforms achieve local frequency estimation, but the windowed Fourier transforms suffer from a fixed window size. It turns out, as a consequence, that they do not effectively handle signals with transients and components whose pitch changes rapidly. Making the time-domain window more localized (narrower) makes the frequency-domain window less localized (wider) and vice versa. Time-scale transforms can deal with these last problems; indeed we can mark this insight by petroleum geologists as the grand opening of modern wavelet theory. But time-scale transforms too have deficiencies. One such is the lack of translation-invariance. The final chapter explores some signal analysis applications and examines the tradeoffs between pure time-domain, time-frequency, and time-scale methods.

11.2 CONTINUOUS WAVELET TRANSFORMS

This section presents the continuous wavelet transform. The wavelet representation for one-dimensional signals was developed by Grossmann and Morlet to overcome the deficiencies of the Gabor transform for seismic applications [13]. Wavelets are special functions whose translations and dilations can be used for expansions of

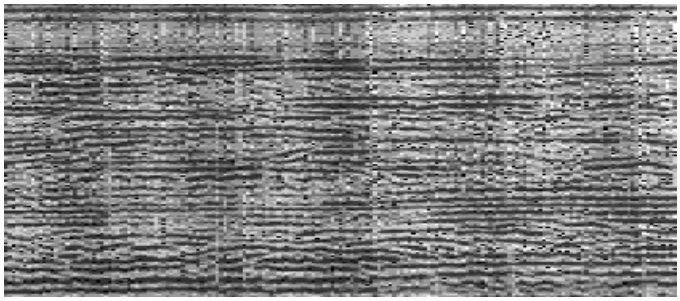


Fig. 11.1. Typical seismic section.

square-integrable functions. In the discussion of the fixed window size implicit in the Gabor representation in Section 2.4, it was noted that the Gabor representation is burdened by the problem of high-magnitude, high-frequency coefficients that is so typical of the Fourier transform.

11.2.1 An Unlikely Discovery

Seismic signals contain many irregular and isolated transients (Figure 11.1). The drawback of the Fourier transform is that it represents signal frequencies as present for all time, when in many situations, and in seismic signal interpretation in particular, the frequencies are localized. The Gabor transform and its more general variant, the short-time Fourier transform (STFT), provide local frequency analysis. One feature of the short-time transforms is that the window size remain fixed. This is acceptable as long as the signal frequency bursts are confined to regions approximating the size of the transform window.

However, in seismic applications, even the STFT becomes problematic. The problem is that seismic signals have many transients, and Grossmann and Morlet found the windowed Fourier algorithms to be numerically unstable. That is, a slight change in the input seismic trace results in a quite pronounced change in the decomposition coefficients. Grossmann and Morlet identified the fixed window size as contributing to the difficulty. Their solution was to keep the same basic filter shape, but to shrink its time-domain extent. That is, they resorted to a transform based on *signal scale*.

11.2.2 Basic Theory

This section introduces the fundamental ideas behind continuous-domain wavelet transforms.

11.2.2.1 Definition and Motivation. Let us begin with a formal definition of a wavelet. The idea is rather recent, and this special signal type passes through the scientific and engineering literature by means of a variety of monikers.

Definition (Analyzing Wavelet). The square-integrable signal $\psi(t)$ is an *analyzing wavelet* if it satisfies the *admissibility condition*

$$C_\psi = \int_{-\infty}^{\infty} \frac{|\Psi(\omega)|^2}{|\omega|} d\omega < \infty, \tag{11.2}$$

where $\Psi(\omega)$ is the radial Fourier transform of $\psi(t)$. The quantity in (11.2) is called the *admissibility factor.* Other names for analyzing wavelets are *basic wavelet*, *continuous wavelet*, *admissible wavelet,* and *mother wavelet.* In some analyses it is convenient to normalize the analyzing wavelet,

$$\sqrt{\langle \psi(t), \psi(t) \rangle} = 1, \tag{11.3}$$

but normalization is not a necessary condition for generating useful time-scale transforms or performing the inverse wavelet transform.

The admissibility condition makes possible the inversion relation for the transform. There are some further consequences, however: Wavelets are bandpass filters with a quick frequency cutoff characteristic and have zero mean in the time domain.

The *wavelet transform* is a time-scale transform that uses a scaled and translated version of the analyzing wavelet in a Hilbert space inner product to convert one-dimensional time-varying signals to a two-dimensional scale and translations space:

Definition (Wavelet Transform). Let

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right) \tag{11.4}$$

Let $f(t)$ be square-integrable. The wavelet transform of $f(t)$ is defined as the inner product

$$F_\psi(a,b) = \mathcal{W}[f(t)](a,b) = \int_{-\infty}^{\infty} f(t) \overline{\psi_{a,b}(t)} dt \equiv (f(t), \psi_{a,b}(t)). \tag{11.5}$$

The wavelet transform is a mapping from the one-dimensional time domain to a two-dimensional space consisting of a scale a and a translation b (Figure 11.2).

An inverse wavelet transform synthesizes $f(t)$ from the two-dimensional $W[f(t)](a, b)$:

Definition (Inverse Wavelet Transform). The inverse wavelet transform is the two-dimensional integral,

$$f(t) = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W[f(t)](a,b) \psi(t) \, d\mu, \tag{11.6}$$

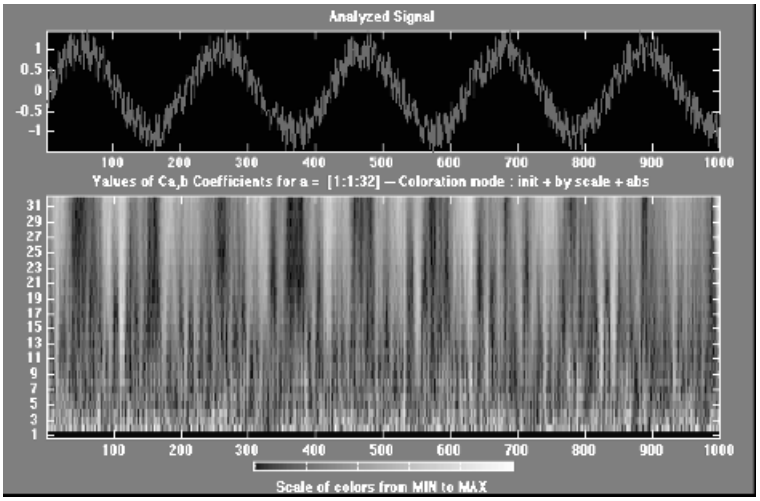


Fig. 11.2. Example of continuous wavelet transform of a noisy sine wave.

where

$$d\mu = \frac{dadb}{a^2} \tag{11.7}$$

and

$$C_\psi = \int_{-\infty}^{\infty} \frac{|\Psi(\omega)|^2}{|\omega|}\, d\omega. \tag{11.8}$$

Remarks. The definition of a wavelet is fundamentally quite simple, consisting of a *time* criterion (square integrability) and a *frequency* criterion expressed by the admissibility condition. Some references include normalization in the definition of a wavelet, but we emphasize that unit energy is an option, not a necessity. At the present level of development, (11.6) suggests that the admissibility condition (11.2) allows the inverse wavelet transform to be carried out. On cursory inspection the admissibility condition would suggest that the spectrum of an analyzing wavelet should decay rapidly for large $|\omega|$, and since wavelets are defined to be square-integrable, this is automatically fulfilled. On the other hand, the presence of $|\omega|$ in the denominator of (11.8) imposes two further requirements on the time domain behavior of the analyzing wavelet. One is obvious, the other is a bit more subtle, but both are relatively easy to satisfy, as the following discussion demonstrates.

Proposition. Let $\psi(t)$ be an analyzing wavelet as previously defined. Then the admissibility criterion is satisfied if

(i) the analyzing wavelet is of zero mean, that is,

$$\int_{-\infty}^{\infty} \psi(t)\, dt = 0, \tag{11.9}$$

and

(ii)

$$t\psi(t) \in L^1(\mathrm{R}). \tag{11.10}$$

Proof: The first condition is obvious: We require $\Psi(0) = 0$ to ensure that the integrand in (11.2) remains finite at $\omega = 0$. Equation (11.9) simply restates this in terms of the Fourier transform,

$$\lim_{\omega \to 0} \int_{-\infty}^{\infty} \psi(t) e^{-j\omega t} dt = \int_{-\infty}^{\infty} \psi(t)\, dt = 0. \tag{11.11}$$

The significance of the second criterion is best demonstrated by dividing the real line into three segments and examining the integral (11.8),

$$C_\psi = \int_{-\infty}^{-1} \frac{|\Psi(\omega)|^2}{|\omega|}\, d\omega + \int_{-1}^{1} \frac{|\Psi(\omega)|^2}{|\omega|}\, d\omega + \int_{1}^{\infty} \frac{|\Psi(\omega)|^2}{|\omega|}\, d\omega. \tag{11.12}$$

Our primary interest is the integral over the interval $t \in [-1, 1]$. According to the moment theorem developed in Chapter 5, if $t\psi(t) \in L^1(\mathbb{R})$, then the first derivative of $\Psi(\omega)$ exists and is bounded. Designate the maximum value of this derivative in the interval $|\omega| \le 1$:

$$\frac{d}{d\omega}\Psi(\omega) \le M. \tag{11.13}$$

According to the mean value theorem of differential calculus, if a function $g(\omega)$ is bounded and continuous on an interval $[a, b]$, then

$$\int_a^b g(\omega)\, d\omega \le M(b - a). \tag{11.14}$$

Designating $\frac{d}{d\omega}|\Psi(\omega)| = g(\omega)$, (11.14) implies

$$|\Psi(\omega)| \le M \cdot 2|\omega| \tag{11.15}$$

for $|\omega| \le 1$. This bound is actually tighter than implied by (11.15). Since $\Psi(0) = 0$, the relevant interval is effectively halved so that

$$|\Psi(\omega)| \le M \cdot |\omega|. \tag{11.16}$$

Returning to the admissibility condition, we now have

$$\int_{-1}^{1} \frac{|\Psi(\omega)|^2}{|\omega|}\, d\omega \le \int_{-1}^{1} M^2 |\omega|\, d\omega \le M, \tag{11.17}$$

thus bounding one portion of the admissibility condition. The remaining two integrals along their respective semi-infinite intervals are easily handled. Since $\frac{|\Psi(\omega)|^2}{|\omega|} \ge 0$ over $|\omega| \le 1$, it follows that

$$\int_{-\infty}^{-1} \frac{|\Psi(\omega)|^2}{|\omega|}\, d\omega + \int_{1}^{\infty} \frac{|\Psi(\omega)|^2}{|\omega|}\, d\omega \le \int_{-\infty}^{\infty} \frac{|\Psi(\omega)|^2}{|\omega|}\, d\omega < \int_{-\infty}^{\infty} |\Psi(\omega)|^2 d\omega. \tag{11.18}$$

This is bounded by virtue of the L^2 Fourier transform. In summary, the overall proposition is proved by virtue of (11.11), (11.17), and (11.18). ■

Remark. The conditions (11.9) and (11.10) are not difficult to satisfy. The first criterion,

$$\int_{-\infty}^{\infty} \psi(t)\, dt = 0 \tag{11.19}$$

implies that a wavelet must oscillate about the time axis—it puts the wave into a *wave*let. The stipulation $t\psi(t) \in L^1(d\mathbb{R})$ can be met if (for example) $\psi(t) \in L^1(\mathbb{R})$ and *has compact support*. By definition, this would imply that any function $\psi(t) \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ with zero mean is a wavelet.

Proposition (Fourier Representation of Wavelet Transform). Let $f(t) \in L^2(\mathbb{R})$. Then for a given scale a, the wavelet transform is proportional to a Fourier transform into the space of translations:

$$W[f(t)](a, b) = \frac{1}{\sqrt{2\pi}} \mathcal{F}[F(\gamma)](-b), \tag{11.20}$$

where

$$F(\gamma) = \sqrt{|a|} \cdot \mathcal{F}[f(t)](\gamma) \cdot \overline{\mathcal{F}[\psi(t)](a\gamma)}. \tag{11.21}$$

By definition and by Parseval's relation, it readily follows that

$$W[f(t)](a, b) = \langle f(t), \psi_{a,b}(t) \rangle = \frac{1}{\sqrt{2\pi}} \langle \mathcal{F}[f(t)](\gamma), \overline{\mathcal{F}[\psi_{a,b}(t)](\gamma)} \rangle. \tag{11.22}$$

However,

$$\mathcal{F}[\psi_{a,b}(t)](\gamma) = e^{j\gamma b} \mathcal{F}[\psi_{a,0}(t)](\gamma) = e^{j\gamma b} \sqrt{|a|}\, \mathcal{F}[\psi(t)](\gamma a) \tag{11.23}$$

so that (11.22) can be expressed in the desired form,

$$W[f(t)](a, b) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \sqrt{|a|} \cdot \mathcal{F}[f(t)](\gamma) \cdot \mathcal{F}[\psi_{a,b}(t)](\gamma) \cdot e^{-j\gamma b} d\gamma \tag{11.24}$$

and the proposition is proven. ■

This intermediate result is useful for establishing the more important Parseval relation for the wavelet transform.

Theorem (Wavelet Transform Parseval Relations). Let $f(t) \in L^2(\mathbb{R})$ and $g(t) \in L^2(\mathbb{R})$, and let C_ψ be the admissibility coefficient as previously defined. Then

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W[f(t)](a, b) \overline{W[g(t)](a, b)}\, d\mu = C_\psi \langle f(t), g(t) \rangle. \tag{11.25}$$

Proof: Let $F(\gamma)$ be d efined as in (11.21) and define

$$G(\gamma) \equiv \sqrt{|a|} \cdot \mathcal{F}[g(t)](\gamma) \cdot \overline{\mathcal{F}[\psi(t)](a\gamma)}. \tag{11.26}$$

Then according to the previous proposition,

$$\begin{aligned} &\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W[f(t)](a, b) \overline{W[g(t)](a, b)}\, d\mu \\ &\quad = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \mathcal{F}[F(\gamma)](-b) \overline{\frac{1}{\sqrt{2\pi}} \mathcal{F}[G(\gamma)](-b)}\, d\mu \end{aligned} \tag{11.27}$$

Using the Parseval relation to convert the b-space Fourier transforms back to γ space, the above integral takes the form

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathcal{F}[f(t)](\gamma) \overline{\mathcal{F}[g(t)](\gamma)} \cdot |\mathcal{F}[\psi(t)](a\gamma)|^2 d\gamma\, d\alpha, \tag{11.28}$$

where $d\alpha \equiv \dfrac{da}{a}$.

The integrals over a and γ can be separated so that (11.28) becomes

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |\mathcal{F}[\psi(t)](a\gamma)|^2 d\alpha \int_{-\infty}^{\infty} \mathcal{F}[f(t)](\gamma) \overline{\mathcal{F}[g(t)](\gamma)}\, d\gamma. \tag{11.29}$$

Applying Parseval's relation to the second of these integrals gives a time-domain inner product:

$$\sqrt{2\pi} \langle f(t), g(t) \rangle. \tag{11.30}$$

The substitution of variables $\omega = a\gamma$ into the first integral implies $\frac{da}{a} = \frac{d\omega}{|\omega|}$ so that (11.29) takes the desired form,

$$C_\psi \langle f(t), g(t) \rangle, \tag{11.31}$$

completing the proof. ■

Theorem (Inverse Wavelet Transform). Let $f(t)$ be square-integrable. The synthesis problem for the continuous wavelet transform takes the form

$$f(t) = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W[f(t)](a, b) \psi_{a,b}(t)\, d\mu. \tag{11.32}$$

Proof: This inversion formula follows directly from the Parseval relation (11.25), which can be written

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W[f(t)](a, b) \overline{\int_{-\infty}^{\infty} g(t) \overline{\psi_{a,b}(t)}\, dt} d\mu = C_\psi \langle f(t), g(t) \rangle. \tag{11.33}$$

This can be rearranged in the more suggestive form:

$$\int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W[f(t)](a, b) \psi_{a,b}(t)\, d\mu \right] \overline{g(t)}\, dt = C_\psi \langle f(t), g(t) \rangle. \tag{11.34}$$

Since $g(t)$ is an arbitrary function in $L^2(\mathbb{R})$, (11.34) implies (11.25), and the proposition is proven. ■

Remark. Note that the wavelet $\psi_{a,b}(t)$ is not conjugated when taking the inverse transform (11.25), in contrast to the forward wavelet transform (11.5).

Since $\Psi(\omega) \in L^2(\mathbb{R})$, we must obtain $\Psi(\omega) \to 0$ as $\omega \to \infty$; hence $y(t)$ is a band-pass filter. The details are left as an exercise.

11.2.2.2 Algebraic Properties. As in the case of the Fourier transform, operations such as scaling, translation, and linear combination can be applied to both the analyzing wavelet and the signal waveform. The proofs are straightforward, some are given explicitly in the text, and others are left as exercises. In the following discussion, we assume all signals are square-integrable.

Let us first cover operations on the analyzing wavelet.

Proposition. Let α, β be complex constants and $\psi(t)$, $\phi(t)$ are wavelets. If we define $\theta(t) = \alpha\psi(t) + \beta\phi(t)$, then

$$W_\theta[f(t)](a, b) = \bar{\alpha} W_\psi[f(t)](a, b) + \bar{\beta} W_\phi[f(t)](a, b). \tag{11.35}$$

Proof: Follows trivially from the linearity of the integral (exercise). ■

Proposition (Translation of Analyzing Wavelet). Let γ be a real constant and $\psi(t)$ be a wavelet. If we define $\theta(t) = \psi(t-\gamma)$, then

$$W_\theta[f(t)](a,b) = W_\psi[f(t)](a, b+\gamma a). \tag{11.36}$$

Proof: By definition,

$$W_\theta[f(t)](a,b) = \int_{-\infty}^{\infty} f(t)\frac{1}{\sqrt{|a|}}\overline{\psi\left(\frac{t-b}{a}-\gamma\right)}\,dt = \int_{-\infty}^{\infty} f(t)\frac{1}{\sqrt{|a|}}\overline{\psi\left(\frac{t-(b+\gamma a)}{a}\right)}\,dt, \tag{11.37}$$

which proves the theorem. ■

Proposition (Scaling of Analyzing Wavelet). Let $\eta > 0$ and $\psi(t)$ be a wavelet. If $\theta(t) = \frac{1}{\eta}\psi\left(\frac{t}{\eta}\right)$, then

$$W_\theta[f(t)](a,b) = \frac{1}{\sqrt{\eta}} W_\psi[f(t)](a\eta, b). \tag{11.38}$$

Proof: Exercise. ■

Now let us turn to signal operations and the resulting wavelet transformations.

Proposition (Linearity). Let α, β be complex constants. If we define $\theta(t) = \alpha\psi(t) + \beta\phi(t)$, then

$$W_\theta[\alpha f(t) + \beta g(t)](a,b) = \alpha W_\psi[f(t)](a,b) + \beta W_\phi[f(t)](a,b). \tag{11.39}$$

Proof: The proof is straightforward and left as an exercise. Note the similarity to, and subtle difference between, this case and the similar operation on the analyzing wavelet. ■

Proposition (Translation). Let γ be a real constant. Then

$$W[f(t-\gamma)](a,b) = W[f(t)](a, b-\gamma). \tag{11.40}$$

Proof: Exercise. ■

Proposition (Scaling of Signal). Let $\eta > 0$. Then

$$W\left[\frac{1}{\eta} f\left(\frac{1}{\eta}\right)\right](a,b) = \frac{1}{\sqrt{\eta}} W[f(t)]\left(\frac{a}{\eta}, \frac{b}{\eta}\right). \tag{11.41}$$

Proof: By change of variables $\tau = t/\eta$, it follows that

$$W\left[\frac{1}{\eta}f\left(\frac{1}{\eta}\right)\right](a,b) = \int_{-\infty}^{\infty} f(\tau)\frac{1}{\sqrt{|a|}}\overline{\psi\left(\frac{\eta\tau - b}{a}\right)}\,d\tau = \int_{-\infty}^{\infty} f(\tau)\frac{1}{\sqrt{|a|}}\overline{\psi\left(\frac{\tau - (b/\eta)}{(a/\eta)}\right)}\,d\tau. \tag{11.42}$$

Since

$$W[f(t)]\left(\frac{a}{\eta},\frac{b}{\eta}\right) = \int_{-\infty}^{\infty} f(\tau)\sqrt{\frac{|\eta|}{|a|}}\overline{\psi\left(\frac{\tau - (b/\eta)}{(a/\eta)}\right)}\,d\tau \tag{11.43}$$

the desired relation (11.41) follows. ■

11.2.2.3 Synthesis with Positive Scale. One final set of properties follows when we restrict the dilation parameter a to positive values.

Practical signal analysis and synthesis algorithms benefit from the elimination of redundant data. We now demonstrate a condition under which the reconstruction (11.32) (and by inference, the forward wavelet transform) requires only positive values of the dilation. We show that this condition is met by *all real-valued wavelets*, which comprise the vast majority of continuous and discrete wavelets.

Proposition (Positive Dilation Values). If

$$\int_0^{\infty}\frac{|\Psi(\omega)|^2}{|\omega|}\,d\omega = \int_{-\infty}^{0}\frac{|\Psi(\omega)|^2}{|\omega|}\,d\omega, \tag{11.44}$$

then (note the limit on the domain of a)

$$f(t) = \frac{1}{C_\psi}\int_0^{\infty}\int_{-\infty}^{\infty} W[f(t)](a,b)\psi_{a,b}(t)\,d\mu, \tag{11.45}$$

where

$$C_\psi = \int_0^{\infty}\frac{|\Psi(\omega)|^2}{|\omega|}\,d\omega = \int_{-\infty}^{0}\frac{|\Psi(\omega)|^2}{|\omega|}\,d\omega. \tag{11.46}$$

Proof: First, note

$$\int_{-\infty}^{\infty}\frac{|\Psi(\omega)|^2}{|\omega|}\,d\omega = 2\int_0^{\infty}\frac{|\Psi(\omega)|^2}{|\omega|}\,d\omega = 2\int_{-\infty}^{0}\frac{|\Psi(\omega)|^2}{|\omega|}\,d\omega. \tag{11.47}$$

The overall proof is best carried out by reconsidering the steps leading up to the Parseval relation (11.25). Note that if (11.44) holds, the two auxiliary functions,

$$G(\gamma) \equiv \sqrt{|a|}\cdot\mathcal{F}[g(t)](\gamma)\cdot\overline{\mathcal{F}[\psi(t)](a\gamma)} \tag{11.48}$$

and

$$F(\gamma) \equiv \sqrt{|a|} \cdot \mathcal{F}[f(t)](\gamma) \cdot \overline{\mathcal{F}[\psi(t)](a\gamma)}, \tag{11.49}$$

display the necessary symmetry in a so that (11.27) can be reformulated as an integral over positive dilations only:

$$2\int_0^\infty \int_{-\infty}^\infty W[f(t)](a,b)\overline{W[g(t)](a,b)}\, d\mu = 2\left(\int_0^\infty \frac{|\Psi(\omega)|^2}{|\omega|}\, d\omega\right)\langle f(t),\, g(t)\rangle. \tag{11.50}$$

The factors of 2 cancel and, starting from (11.50), it is straightforward to reproduce the wavelet inversion formula, leading to the desired result (11.45). ■

This proposition is of more than passing interest, as indicated by our next observation.

Theorem (Real Wavelets). If $\psi(t)$ is a real-valued function, then (11.44) is satisfied.

Proof: This is easily established from the Fourier transform of $\psi(t)$,

$$\Psi(\omega) = \int_{-\infty}^{\infty} \psi(t) e^{-j\omega t} dt. \tag{11.51}$$

If $\psi(t) \in \mathbb{R}$, then

$$\Psi(-\omega) = \int_{-\infty}^{\infty} \psi(t) e^{j\omega t} dt = \overline{\Psi(\omega)}. \tag{11.52}$$

From here is easy to establish condition (11.44), since

$$\int_{-\infty}^{0} \frac{|\Psi(\omega)|^2}{|\omega|}\, d\omega = -\int_0^{-\infty} \frac{|\Psi(\omega)|^2}{|\omega|}\, d\omega. \tag{11.53}$$

With a simple substitution of variables $\eta = -\omega$, this can be rearranged to the desired result,

$$-\int_0^{-\infty} \frac{|\Psi(\omega)|^2}{|\omega|}\, d\omega = -\int_\infty^0 \frac{|\Psi(\gamma)|^2}{|\gamma|}\, d\gamma = \int_0^\infty \frac{|\Psi(\gamma)|^2}{|\gamma|}\, d\gamma. \tag{11.54}$$

■

Remarks. Note how the condition (11.52) is explicitly used to establish the first equality in (11.54). Also, the importance of this theorem lies in the implication that all real-valued wavelets can lead to reconstruction on the half-plane $a \in [0, \infty]$.

Note that some authors *define* synthesis to occur over this restricted domain, but they are often tacitly restricting the discussion to real-valued $\psi(t)$, which form the overwhelming majority of practical wavelets. Selected complex-valued wavelets (to be considered later) may also satisfy (11.45) with a suitable redefinition of C_ψ, but whenever complex-valued wavelets are under consideration, the reader should exercise caution when performing reconstruction.

Table 11.1 summarizes our results so far.

11.2.2.4 Wavelets by Convolution. Convolution is a smoothing operation which preserves any existing localized properties of the functions involved. It is simple to show that under certain reasonable conditions, wavelets generate other wavelets through the convolution operation.

Theorem (Wavelets Through Convolution). If $\psi(t)$ is a wavelet and $\lambda(t) \in L^1$, then

$$\phi \equiv \psi * \lambda \tag{11.55}$$

is a wavelet.

Proof: We first need to establish that $\phi \in L^2$. This can be carried out in the time domain, but it is simpler to consider the frequency domain where

$$\mathcal{F}[\phi(t)](\omega) = \Psi(\omega)\Lambda(\omega). \tag{11.56}$$

TABLE 11.1. Wavelet Transform Properties[a]

Signal Expression	Wavelet Transform or Property
$\psi_{a,b}(t) = \frac{1}{\sqrt{\lvert a\rvert}}\psi\left(\frac{t-b}{a}\right)$	Dilation and translation of $\psi(t)$
$C_\psi = \int_{-\infty}^{\infty} \frac{\lvert\Psi(\omega)\rvert^2}{\lvert\omega\rvert}\, d\omega$	Admissibility factor
$f(t)$	$W[f(t)](a,b) = \int_{-\infty}^{\infty} f(t)\overline{\psi_{a,b}(t)}\, dt$
$W[f(t)](a,b) = \frac{1}{\sqrt{2\pi}}\mathcal{F}[F(\gamma)](-b)$	Fourier transform representation
$f(t) = \frac{1}{C_\psi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W[f(t)](a,b)\psi(t)\frac{da\,db}{a^2}$	Inverse
$\theta(t) = \alpha\psi(t) + \beta\phi(t)$	$W_\theta[f(t)] = \bar{\alpha}W_\psi[f(t)] + \bar{\beta}W_\phi[f(t)]$

[a]In the table, $\psi(t)$ is square-integrable.

It is easy to establish that this spectrum is L^2. While we cannot assert that $\Lambda(\omega)$ is integrable, it is certainly bounded and

$$\int_{-\infty}^{\infty} |\Psi(\omega)\Lambda(\omega)|^2 d\omega = \int_{-\infty}^{\infty} |\Psi(\omega)|^2 |\Lambda(\omega)|^2 d\omega < |\Lambda(\omega)|_{\max}^2 \int_{-\infty}^{\infty} |\Psi(\omega)|^2 d\omega < \infty, \tag{11.57}$$

which proves $\mathcal{F}[\phi(t)](\omega) \in L^2$. The inverse Fourier transform maps L^2 to L^2 so that

$$\phi(t) \in L^2. \tag{11.58}$$

The admissibility condition on $\phi(t)$ follows in a similar manner:

$$\int_{-\infty}^{\infty} \frac{|\Phi(\omega)|^2}{|\omega|} d\omega = \int_{-\infty}^{\infty} \frac{|\Psi(\omega)|^2}{|\omega|} |\Lambda(\omega)|^2 d\omega < |\Lambda(\omega)|_{\max}^2 \int_{-\infty}^{\infty} \frac{|\Psi(\omega)|^2}{|\omega|} d\omega < \infty. \tag{11.59}$$

Conditions (11.58) and (11.59) establish that $\phi(t)$ is a wavelet. ■

Now let us turn to some examples of analyzing wavelets.

11.2.3 Examples

Continuous analyzing wavelets are atomic functions with imposed oscillations. For example, we have seen that the Gaussian time-scale atom is not a wavelet, but operating on a Gaussian by taking one or more derivatives can impose the necessary waviness to ensure that the zero-mean condition (11.19) is satisfied.

11.2.3.1 First derivative of a Gaussian. Let us first consider the analyzing wavelet. A bona fide wavelet is created by applying the first derivative to a Gaussian,

$$\psi(t) = A_0 \cdot \left[-\frac{d}{dt} e^{-t^2} \right] = 2A_0 t e^{-t^2}. \tag{11.60}$$

The normalization constant A_0 can be determined by solving a straightforward Gaussian integral,

$$\int_{-\infty}^{\infty} |\psi(t)|^2 dt = 4A_0^2 \int_{-\infty}^{\infty} t^2 e^{-2t^2} dt \equiv 1, \tag{11.61}$$

which leads to

$$A_0 = \sqrt[4]{2/\pi}. \tag{11.62}$$

The normalization verifies that $\psi(t)$ is in fact square-integrable and by inspection, due to the odd symmetry of (11.60), the zero-mean condition

$$\int_{-\infty}^{\infty} \psi(t)\, dt = 0 \tag{11.63}$$

is assured.

Next, we check the admissibility criteria. The Fourier domain is easily handled by applying the time differentiation property,

$$\mathcal{F}[\psi(t)](\omega) = j\omega A_o \sqrt{\pi} e^{-\omega^2/4}. \tag{11.64}$$

Then for $\omega < 0$

$$\frac{|\psi(\omega)|^2}{|\omega|} = -A_o^2 \pi \omega e^{-\omega^2/2} \tag{11.65}$$

and for positive frequencies

$$\frac{|\psi(\omega)|^2}{|\omega|} = A_o^2 \pi \omega e^{-\omega^2/2}. \tag{11.66}$$

The coefficient (11.8) takes the form

$$C_\psi = -A_o^2 \pi \int_{-\infty}^{0} \omega e^{-\omega^2/2} d\omega + A_o^2 \pi \int_{0}^{\infty} \omega e^{-\omega^2/2} d\omega. \tag{11.67}$$

These integrals defined along the half-line can be evaluated by noting that each integrand can be represented as a derivative, so (11.67) now reads

$$C_\psi = A_o^2 \pi \int_{-\infty}^{0} \frac{d}{d\omega} e^{-\omega^2/2} d\omega - A_o^2 \pi \int_{0}^{\infty} \frac{d}{d\omega} e^{-\omega^2/2} d\omega \tag{11.68}$$

so

$$C_\psi = 2\pi A_0^2. \tag{11.69}$$

Remark. Note that (11.54) holds for this real-valued analyzing wavelet, as expected. If reconstruction uses only positive values of scale (as per (11.45)), then

$$C_\psi = \pi A_0^2 \tag{11.70}$$

should be used in place of (11.69).

Example (Gaussian Transient). We will generate and discuss the wavelet transform of a Gaussian pulse

$$f(t) = e^{-\alpha t^2}, \tag{11.71}$$

where α is a positive factor. The analyzing wavelet (11.60) with an applied scale a and translation b reads

$$\psi\left(\frac{t-b}{a}\right) = 2A_0 e^{-b^2/a^2} \cdot \left(\frac{t-b}{a}\right) e^{-\left(\frac{t^2-2bt}{a^2}\right)} \tag{11.72}$$

and the wavelet transform integral breaks down conveniently,

$$W[f(t)](a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t)\psi\left(\frac{t-b}{a}\right)dt = C(a,b)[I_1 - bI_2], \tag{11.73}$$

where

$$C(a,b) = \frac{2A_0}{a\sqrt{a}} e^{-b^2/a^2} \tag{11.74}$$

and the integrals

$$I_1 \equiv \int_{-\infty}^{\infty} te^{-(\alpha+(1/a^2))t^2+(2b/a^2)t}dt = \frac{a\sqrt{\pi}}{\sqrt{a^2\alpha+1}}\left(\frac{b}{a^2\alpha+1}\right)e^{\left(\frac{b^2}{a^2}\right)\left(\frac{1}{a^2\alpha+1}\right)} \tag{11.75}$$

and

$$I_2 \equiv \int_{-\infty}^{\infty} e^{-(\alpha+(1/a^2))t^2+(2b/a^2)t}dt = \frac{a\sqrt{\pi}}{\sqrt{a^2\alpha+1}} e^{\left(\frac{b^2}{a^2}\right)\left(\frac{1}{a^2\alpha+1}\right)} \tag{11.76}$$

are evaluated using standard Gaussian integration. The result

$$W[f(t)](a,b) = \frac{-a^2b\alpha\sqrt{\alpha}}{\sqrt{a(a^2\alpha+1)}} e^{\left(\frac{b^2}{a^2}\right)\left(\frac{1}{a^2\alpha+1}\right)} \tag{11.77}$$

is a two-dimensional function of scale and translation shown in Figure 11.3.

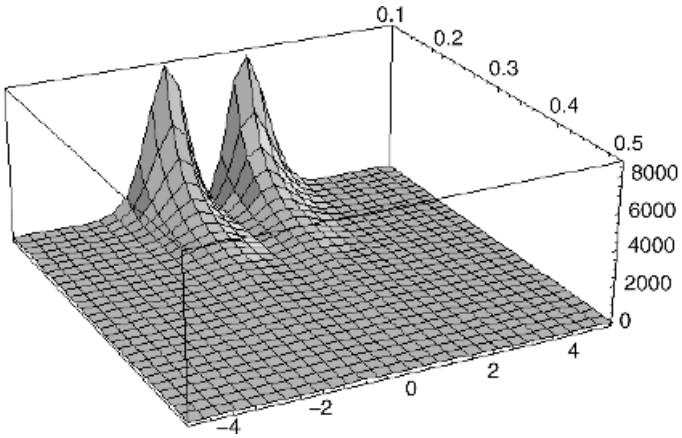


Fig. 11.3. Wavelet transform of Gaussian pulse.

Example (Rectangular Pulse). Consider a rectangular pulse of width D analyzed by the same wavelet as above. The wavelet transform again takes the form

$$W[f(t)](a,b) = \frac{1}{\sqrt{a}}\int_{-\infty}^{\infty} f(t)\psi\left(\frac{t-b}{a}\right)dt = C(a,b)[I_1 - bI_2], \tag{11.78}$$

where as before

$$C(a,b) = \frac{2A_0}{a\sqrt{a}}e^{-b^2/a^2}. \tag{11.79}$$

The integrals now read

$$I_1 \equiv \int_{-D/2}^{D/2} te^{-(1/a^2)t^2 + (2b/a^2)t}dt \tag{11.80}$$

and

$$I_2 \equiv \int_{-D/2}^{D/2} e^{-(1/a^2)t^2 + (2b/a^2)t}dt\,. \tag{11.81}$$

The central feature of each integral is the exponential

$$e^{-[f(a)t^2 + g(b)t]}, \tag{11.82}$$

where $f(a) \equiv 1/a^2$ and $g(a) \equiv -2b^2/a^2$. With proper manipulation, (11.81) and (11.82) can be handled analytically. It is a simple matter to complete the square on the argument of the above exponential, transforming it:

$$f(a)t^2 + g(b)t \rightarrow (f(a)t^2 + g(b)t + x) - x, \tag{11.83}$$

where $x = \frac{1}{4}\frac{g^2(b)}{f(a)}$. If we let $y \equiv \sqrt{f(a)}t + \frac{1}{2}\frac{g(b)}{\sqrt{f(a)}}$ and make a substitution of variables, then

$$I_1 \rightarrow \frac{1}{\sqrt{f(a)}}\int_{L_1}^{L_2}\left(y - \frac{1}{2}\frac{g(b)}{\sqrt{f(a)}}\right)e^{-y^2}e^{x}\cdot\frac{1}{\sqrt{f(a)}}dy, \tag{11.84}$$

where the limits $L_1 \equiv \frac{-d\sqrt{f(a)}}{2} + \frac{1}{2}\frac{g(b)}{\sqrt{f(a)}}$ and $L_2 \equiv \frac{d\sqrt{f(a)}}{2} + \frac{1}{2}\frac{g(b)}{\sqrt{f(a)}}$. This conveniently breaks into two terms,

$$I_1 = \frac{e^x}{f(a)}\left[\int_{L_1}^{L_2} ye^{-y^2}dy - \frac{1}{2}\frac{g(b)}{\sqrt{f(a)}}\int_{L_1}^{L_2} e^{-y^2}dy\right]. \tag{11.85}$$

Now $ye^{-y^2} = \left(-\frac{1}{2}\right)\frac{d}{dy}\left[e^{-y^2}\right]$ so

$$I_1 = \frac{e^x}{f(a)}\left[\left(-\frac{1}{2}\right)\left(e^{L_2^2} - e^{L_1^2}\right) - \frac{1}{2}\frac{g(b)}{\sqrt{f(a)}}[erf(L_2) + erf(L_1)]\right]. \tag{11.86}$$

With similar operations, it is easy to show

$$I_2 = \frac{e^x}{\sqrt{f(a)}}[erf(L_2) + erf(L_1)]. \tag{11.87}$$

11.2.3.2 Second Derivative of a Gaussian ("Mexican Hat"). Taking a further derivative provides an analyzing wavelet

$$\psi(t) = B_0 \cdot \left[\frac{de^{-t^2}}{dt^2}\right] = -2B_0[1 - 2t^2]e^{-t^2}. \tag{11.88}$$

It is readily shown that

$$\int_{-\infty}^{\infty} |\psi(t)|^2 dt = 3B_0^2\sqrt{\frac{\pi}{2}}, \tag{11.89}$$

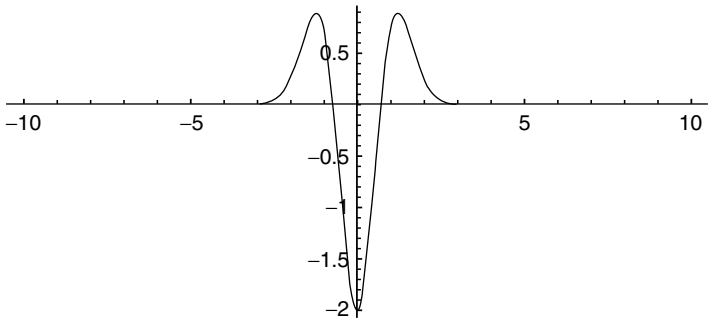


Fig. 11.4. Mexican hat wavelet.

so the normalization constant is

$$B_0 = \frac{1}{\sqrt{3}} \sqrt[4]{2/\pi}. \tag{11.90}$$

This wavelet is shown in Figure 11.4. It has an even symmetry, but equal area above and below the axis, so that (11.9) is satisfied.[1] These details are left as an exercise.

In the Fourier domain, the Mexican hat provides a spectrum

$$\mathcal{F}[\psi(t)](\omega) = -B_0 \omega^2 \sqrt{\pi} e^{-\frac{\omega^2}{4}}, \tag{11.91}$$

and the admissibility coefficient is

$$C_\psi \equiv \int_{-\infty}^{\infty} \frac{|\psi(\omega)|^2}{|\omega|}\, d\omega = \frac{2}{3}\sqrt{2\pi}\,. \tag{11.92}$$

Example (Gaussian Transient). Consider the Mexican hat applied to the Gaussian transient of (11.71). The scaled and translated analyzing wavelet is easily found:

$$\psi\left(\frac{t-b}{a}\right) = [d_0 - d_1 t + d_2 t^2] e^{-\frac{b^2}{a^2}} e^{-\left(\frac{t^2 - 2bt}{a^2}\right)}, \tag{11.93}$$

where

$$d_0 = -2B_0\left[1 - 2\frac{b^2}{a^2}\right], \tag{11.94}$$

$$d_1 = \frac{8B_0}{a^2}, \tag{11.95}$$

[1]It might resemble a traditional Mexican *sombrero* in cross section—hence the name.

and

$$d_2 = \frac{4B_0}{a^2}. \tag{11.96}$$

It is left as an exercise to show that the wavelet transform takes the form

$$\mathcal{W}[f(t)](a,b) = \frac{e^{-b^2/a^2}}{\sqrt{a}}[d_0 I_0 - d_1 I_1 + d_2 I_2] \tag{11.97}$$

with

$$I_0 \equiv \int_{-\infty}^{\infty} e^{-(\alpha + (1/a^2))t^2 + (2b/a^2)t} dt = \frac{a\sqrt{\pi}}{\sqrt{a^2\alpha + 1}} e^{\left(\frac{b^2}{a^2}\right)\left(\frac{1}{a^2\alpha + 1}\right)} \tag{11.98}$$

and I_1 as in (11.75), and

$$\begin{aligned} I_2 &\equiv \int_{-\infty}^{\infty} t^2 e^{-(\alpha + (1/a^2))t^2 + (2b/a^2)t} dt \\ &= \frac{a\sqrt{\pi}}{\sqrt{a^2\alpha + 1}} \left(\frac{a^2}{2(a^2\alpha + 1)}\right)\left(1 + \frac{4(b^2/a^2)}{a^2\alpha + 1}\right) e^{\left(\frac{b^2}{a^2}\right)\left(\frac{1}{a^2\alpha + 1}\right)}. \end{aligned} \tag{11.99}$$

11.3 FRAMES

It is much easier to construct frames based upon wavelets than upon the short-time Fourier transform. Building computer applications requires us to work with discrete rather than continuous signal representations. One requirement for signal analysis is that our discrete representation be capable of representing any signal; this is a *completeness* or *spanning* condition. If we also ask that our discrete representation also be *numerically stable*—that is, a small change in a signal results in a small change in its decomposition coefficients—then we must use a *frame* representation.

As a generalization of orthonormal bases, Chapter 3 introduced frames. We remember that $F = \{f_n : n \in \mathbb{Z}\}$ from a Hilbert space H is a *frame* if there are $A, B \in \mathbb{R}$ such that $A > 0$, $B > 0$, and for all $x \in H$,

$$A\|x\|^2 \leq \sum_{n=-\infty}^{\infty} \left|\langle x, f_n\rangle\right|^2 \leq B\|x\|^2. \tag{11.100}$$

The frame F is *tight* if its lower and upper bounds—A and B, respectively—are equal. Any frame $F \subset H$ spans H. A frame F is *exact* if, when an element is removed from it, it ceases to be a frame. If F is orthonormal, then F is tight; in fact, $A = B = 1$, and F is exact.

Recall from Chapter 10 that the Balian–Low theorem imposes strict constraints on the time- and frequency-domain sampling intervals for a frame of windowed Fourier atoms. The time- and frequency-domain sampling intervals, T and Ω, respectively are critical:

(i) If $T\Omega < 2\pi$, then the time-frequency density $(T\Omega)^{-1}$ exceeds the Nyquist density $(2\pi)^{-1}$, and frames of windowed Fourier atoms are possible.
(ii) If $T\Omega > 2\pi$, then we are sampling below the Nyquist density and there are no windowed Fourier frames.
(iii) If we sample at precisely the Nyquist density, $T\Omega = 2\pi$, and $F = \{w_{m,n}(t) = e^{\Omega jnt}w(t - mT)$: $m, n \in \mathbb{Z}\}$ is a frame, then either $w(t)$ or its Fourier transform $W(\omega)$ is not well-localized (i.e., not a window function).

In this section we shall see that the wavelet transform is not so restrictive; one can find wavelets $\psi(t)$ that allow tight frames aslong as $\Omega \neq 0$, 1 and $T \neq 0$ [14, 15].

11.3.1 Discretization

The wavelet discretization procedure is analogous to discretization of time-frequency transforms. Instead of applying a time and frequency increment, we use a time and scale increment on a signal model. The signal model is an admissible wavelet $\psi(t)$.

We have noted that the continuous wavelet transform is an inner product. It measures similarity of $x(t)$ and $\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right)$ as follows:

$$X_{\psi}(a, b) = \left\langle x, \psi_{a,b} \right\rangle = \int_{-\infty}^{\infty} x(t)\overline{\psi_{a,b}(t)}\, dt\ , \tag{11.101}$$

where $a, b \in \mathbb{R}$. The wavelet $\psi(t)$ must satisfy the admissibility condition (11.2). For simplicity, let us consider only the case $a > 0$ and assume $\psi(t) \in \mathbb{R}$. The inner product (11.101) measures the similarity of $x(t)$ and $a^{-1/2}\psi_{a,b}(t)$, which is a dilated version of $\psi(t)$, shifted so that it centers at time $t = b$.

Suppose we are searching a candidate signal for a prototype shape $\psi(t)$. This is a typical signal analysis problem. Perhaps the shape $\psi(t)$ resembles the signal trace we are trying to detect, or it maybe it responds significantly to some feature—such as an edge—that we can use in a structural description to identify the candidate. If we know the exact location and time-domain extent, we can fix $a, b \in \mathbb{R}$ and perform the inner product computation. If $x(t)$ happens to be a scalar multiple (an attenuated or amplified replica) of $\psi_{a,b}(t)$, then the Schwarz inequality

$$|\langle x(t), \psi_{a,b}(t)\rangle| \leq \|x\|\, \|\psi_{a,b}\| \tag{11.102}$$

will be an equality. Thus, we threshold the inner product (11.102) as a percentage of $\|x\| \times \|\psi_{a,b}\|$ to obtain a measure of the match between prototype and candidate signals.

One the other hand, if we do not know the location and time extent—and this is the more common and daunting signal recognition problem—then the task of performing many inner products in (11.102) becomes a computational burden. We can correlate $\psi_{a,b}(t)$ with local values of $x(t)$, say restricted to $[b - c, b + c]$, for some $c > 0$. But then our inner product varies with the L^2 norm of $x(t)$ restricted to $[b - c, b + c]$. This is conventional normalized cross-correlation, where we divide the inner product by the norm of the candidate signal in a region of interest. Nevertheless, there is in principle a continuous range of scale factors, offsets, and (perhaps) window widths—a, b, and c, respectively. To make the analysis practical, we must choose a discrete set of locations and signal prototype sizes against which we compare the candidate waveform.

Let us start discretization with scale increment $a_0 > 0$. Our discussion closely follows [3]. Dyadic decomposition remains the most common. In this case $a_0 = 2$, and we have dilation steps $\psi(t/2)$, $\psi(t)$, $\psi(2t)$, $\psi(4t)$, and so on. These signal models are, respectively, twice as large, exactly the same, half as large, and one quarter as large in time-domain extent as the root scale element $\psi(t)$. If we let $a = {a_0}^m$, then $\psi(t{a_0}^{-m})$ is ${a_0}^m$ times wider than $\psi(t)$.

Now let us decide how to discretize the time domain. A moment's thought shows that we cannot just take $b = nb_0$ for some $b_0 > 0$ and $n \in \mathbb{Z}$. Note that if $a_0 = 2$, then signal prototypes at the scale $a = {a_0}^1$ have the shape of $\psi(t/2)$ and occupy twice the time-domain extent as at unit scale $a = 1$. Thus, we should cover the time-domain with step increments that are twice as far apart as at unit scale. That way, the time-domain coverage and overlap between prototypes at unit and double scale is proportional. Similarly, if $a = {a_0}^{-1}$, then models at this scale look like $\psi(2t)$ and take only half the time-domain width as at unit scale. We could repeat this logic at quadruple and quarter scales, but the point is that time-domain steps for scale $a = {a_0}^m$ should be in increments of the product $b_0{a_0}^m$. For wavelet transform discretization, we employ wavelet atoms of the form

$$\psi_{m,n}(t) = {a_0}^{-\frac{m}{2}} \psi\left(\frac{t - nb_0{a_0}^m}{{a_0}^m}\right) = {a_0}^{-\frac{m}{2}} \psi({a_0}^{-m}t - nb_0). \tag{11.103}$$

Note that—in accord with other established notations [3, 10]—we use the first discrete index for the scale variable and use the second for the time variable.

As with the short-time Fourier transform, discretization implies a structural description of a signal. Windowed Fourier transforms produce a tiling of the time-frequency plane by signal atoms that occupy equally sized regions. In contrast, time-scale discretizations, as with wavelets, tile the plane with regions of varying size. Signal atoms tuned to higher frequencies have a more restricted time-domain support (Figure 11.5).

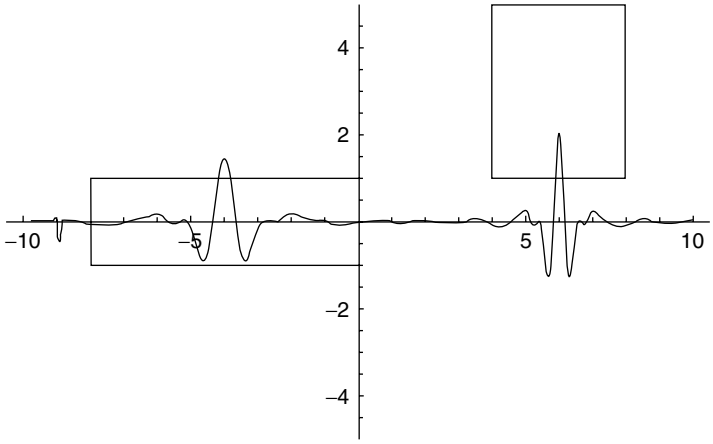


Fig. 11.5. Tiling of the time-frequency plane by a discretized wavelet transform.

11.3.2 Conditions on Wavelet Frames

In order for a discretization based on translations and dilations to constitute a frame, certain necessary conditions must obtain. We quote the following theorem.

Theorem (Necessity of Admissible Wavelet). Suppose $\psi(t) \in L^2(\mathbb{R})$, $a_0 > 0$, and

$$F = \left\{ \psi_{m,n}(t) = a_0^{-\frac{m}{2}} \psi(a_0^{-m} t - n b_0) \,\middle|\, m, n \in \mathbb{Z} \right\} \tag{11.104}$$

constitutes a frame with lower and upper bounds A and B, respectively. Then

$$A b_0 \ln\, a_0 \leq \int_0^\infty \frac{|\Psi(\omega)|^2}{\omega}\, d\omega \leq B b_0 \ln\, a_0 \tag{11.105a}$$

and

$$A b_0 \ln\, a_0 \leq \int_0^\infty \frac{|\Psi(\omega)|^2}{\omega}\, d\omega \leq B b_0 \ln\, a_0\,, \tag{11.105b}$$

where $\Psi(\omega)$ is the (radial) Fourier transform of $\psi(t)$.

Proof: Due to Daubechies [3, 15]. ■

Remark. Interestingly, for a family of translates and dilates to be a frame, $\psi(t)$ must be admissible. One might think that the admissibility condition (11.2) is a

technicality, concocted just to make the wavelet transform inversion work. We see now that it is essential for signal analysis using families of scaled, translated atoms—that is, for wavelet frames.

11.3.3 Constructing Wavelet Frames

This section covers one method for constructing tight wavelet frames [3, 14, 15]. Let $v(t)$ be real-valued, k times continuously differentiable, and approximate the unit step as follows:

$$v(t) = \begin{cases} 0 & \text{if } t \le 0, \\ 1 & \text{if } t \ge 1. \end{cases} \tag{11.106}$$

An example (Figure 11.6) of $v \in C^1$ is the following

$$v(t) = \begin{cases} 0 & \text{if } t \le 0, \\ \sin^2\left(\dfrac{\pi t}{2}\right) & \text{if } t \le 0 \le 1, \\ 1 & \text{if } t \ge 1. \end{cases} \tag{11.107}$$

Now let $a_0 > 1$ and $b_0 > 0$. We will specify two square-integrable signals, $\psi^+(t)$ and $\psi^+(t)$, by their *normalized* radial Fourier transforms, $\Psi^+(\omega)$ and $\Psi^-(\omega)$, respectively. Let $L = 2\pi[b_0(a_0^2 - 1)]^{-1}$, define

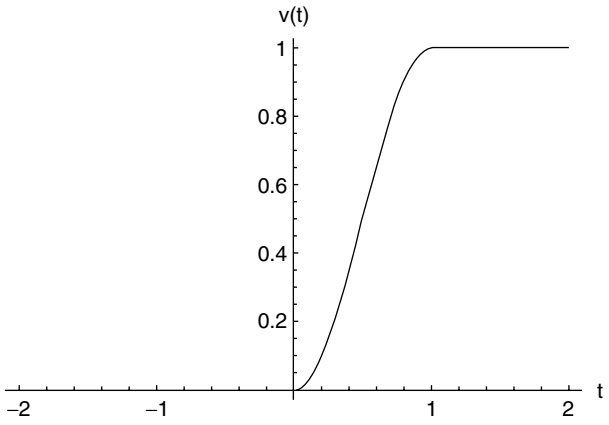
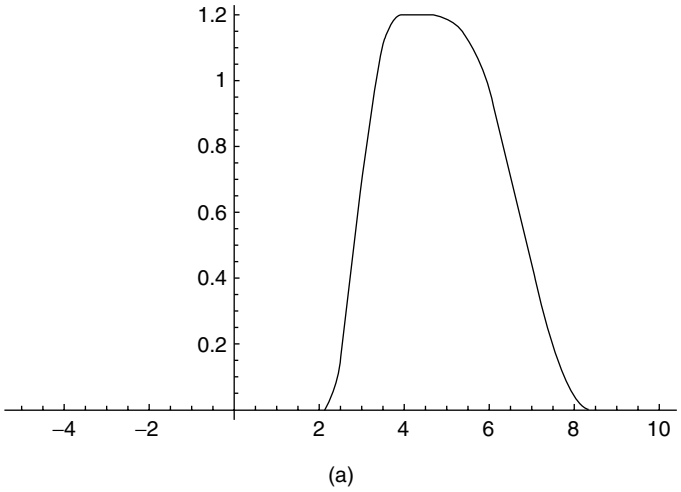
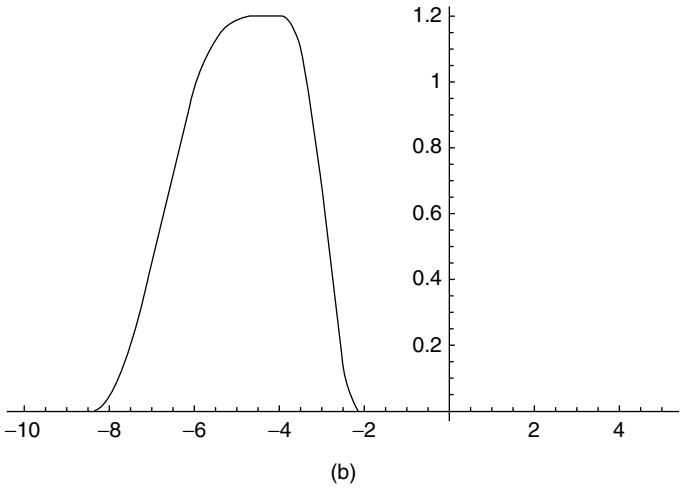


Fig. 11.6. Continuously differentiable approximation to the unit step.

$$\Psi^{+}(\omega) = (\ln a_0)^{-\frac{1}{2}} \begin{cases} 0 & \text{if } \omega \le L \text{ or } \omega \ge La_0^2, \\ \sin\left(\frac{\pi}{2} v\left(\frac{\omega - L}{L(a_0 - 1)}\right)\right) & \text{if } L \le \omega \le La_0, \\ \cos\left(\frac{\pi}{2} v\left(\frac{\omega - La_0}{La_0(a_0 - 1)}\right)\right) & \text{if } La_0 \le \omega \le La_0^2 \end{cases} \tag{11.108}$$



(a)



(b)

Fig. 11.7. Fourier transforms of atoms used for a tight frame based on translation and dilation: (a) $\Psi^{+}(\omega)$ and (b) $\Psi^{-}(\omega)$.

and set $\Psi^{-}(\omega) = \Psi^{+}(-\omega)$. Then—as Figure 11.7 illustrates for the choices $a_0 = 2$, $b_0 = 1$, and $v(t)$ given by (11.107)—$\Psi^{+}(\omega)$ is finitely supported on $[L, La_0^2]$.

The special construction (11.108) guarantees

$$\sum_{m=-\infty}^{\infty} \left|\Psi^{+}(a_0^m \omega)\right|^2 = \begin{cases} \dfrac{1}{\ln(a_0)} & \text{if } 0 < \omega, \\ 0 & \text{if } \omega \le 0. \end{cases} \tag{11.109}$$

To see this, note that the sequence $\{a_0{}^m \mid m \in \mathbb{Z}\}$ decreases toward zero as $m \to -\infty$ and increases toward ∞ as $m \to \infty$. If $\omega > 0$, then there must be exactly one $m \in \mathbb{Z}$ such that $\omega a_0{}^m \in [L, La_0]$, the interval of the sin() term in (11.108). Then the next summand's argument $\omega a_0{}^{m+1}$ falls in the interval supporting the cos() term, $[La_0, La_0{}^2]$. The consequence is that exactly two summands from (11.109) are non-zero: one sin() term for $\omega a_0{}^m$ and one cos() term for $\omega a_0{}^{m+1}$. The sum of their squares is unity justifying (11.109).

We turn to the frame condition. If $x(t) \in L^2(\mathbb{R})$ and $X(\omega)$ is its normalized radial Fourier transform, $X(\omega) = (2\pi)^{-(1/2)} \int x(t) \exp(-j\omega t)\, dt$, then following Ref. 3 we find

$$\begin{aligned} &\sum_{m,n=-\infty}^{\infty} \left|\langle x, \psi^{+}{}_{m,n}\rangle\right|^2 \\ &= \sum_{m,n=-\infty}^{\infty} \left|\langle X, \Psi^{+}{}_{m,n}\rangle\right|^2 \\ &= \sum_{m,n=-\infty}^{\infty} a_0^m \left| \int_{-\infty}^{\infty} e^{j\omega n b_0 a_0^m} X(\omega) \Psi^{+}(a_0^m \omega)\, d\omega \right|^2 \\ &= \sum_m a_0^m \sum_n \left| \sum_k \int_{k\frac{2\pi}{\Omega}}^{(k+1)\frac{2\pi}{\Omega}} X(\omega) \Psi^{+}(a_0^m \omega) e^{j\omega n b_0 a_0^m}\, d\omega \right|^2 . \end{aligned} \tag{11.110}$$

Our strategy has been to break down the integral over $\mathbb{R}$ in (11.110) into an infinite sum of integrals over a finite interval. Note that we have used the fact that the normalized radial Fourier transform is an isometry: $\langle x, y\rangle = \langle X, Y\rangle$. By resorting to the normalized frequency transform, we economize on 2π factors. We recall from basic Hilbert space theory that $\{e_n(t) = (\Omega/2\pi)^{1/2} \exp(jn\Omega t) \mid n \in \mathbb{Z}\}$ is

an orthonormal basis for the Hilbert space $H = L^2[0, 2\pi/\Omega]$. Thus, by Chapter 2's abstract version of the Pythagorean theorem, $\|x\|_2^2 = \Sigma_n |\langle x, e_n\rangle|^2$. Above, we set $\Omega = b_0 a_0{}^m$, break up the integral over $\mathbb{R}$ into sections $2\pi/\Omega$ wide, and interchange summation and integral. With the substitution $\omega = \theta + 2\pi\frac{k}{\Omega}$, (11.110) continues as follows:

$$\sum_m a_0^m \sum_n \left| \int_0^{\frac{2\pi}{\Omega}} e^{j\theta n\Omega} \sum_k X\left(\theta + \frac{2\pi k}{\Omega}\right) \Psi^+\left(a_0^m \theta + \frac{2\pi k}{b_0}\right) d\theta \right|^2$$
$$= \sum_m a_0^m \sum_n \left| \int_0^{\frac{2\pi}{\Omega}} e^{j\theta n\Omega} Y(\theta) d\theta \right|^2 = \sum_m a_0^m \frac{2\pi}{\Omega} \|Y\|_2^2, \tag{11.111}$$

where $Y(\theta)$ is the summation in the integrand at the top of (11.111). Having disposed of one summation, we now backtrack, writing the $\|Y\|_2^2$ term (which is the L^2 norm over H) as an integral, first as a sum over a finite interval, and then as over all of $\mathbb{R}$:

$$\sum_m a_0^m \frac{2\pi}{\Omega} \|Y\|_2^2 = \frac{2\pi}{b_0} \sum_m \int_0^{\frac{2\pi}{\Omega}} Y(\theta)\overline{Y(\theta)}\, d\theta$$
$$= \frac{2\pi}{b_0} \sum_m \int_{-\infty}^{\infty} |X(\omega)|^2 \left|\Psi^+(a_0^m \omega)\right|^2 d\omega = \frac{2\pi}{b_0 \ln a_0} \int_0^{\infty} |X(\omega)|^2 d\omega. \tag{11.112}$$

A similar argument—with a little analytical discomfort but still no pain—gives

$$\sum_{m,\, n = -\infty}^{\infty} \left|\langle x, \psi^-_{m,n}\rangle\right|^2 = \frac{1}{2\pi b_0 \ln a_0} \sum_m \int_{-\infty}^{0} |X(\omega)|^2 d\omega, \tag{11.113}$$

where $\psi^-(t)$ has normalized Fourier transform $\Psi^-(\omega)$. Now we claim that $F = \{\psi_{m,n}{}^+(t)\} \cup \{\psi_{m,n}{}^-(t)\}$ is a frame. Indeed, (11.110) through (11.112) and (11.113) imply

$$\sum_{\substack{m,\, n = -\infty \\ \varepsilon = +,-}}^{\infty} \left|\langle x, \psi^{\varepsilon}_{m,n}\rangle\right|^2 = \frac{2\pi}{b_0 \ln a_0} \int_{-\infty}^{\infty} |X(\omega)|^2 d\omega = \frac{2\pi}{b_0 \ln a_0} \int_{-\infty}^{\infty} |x(t)|^2 dt; \tag{11.114}$$

we see that (11.114) is a tight frame with bounds $\frac{2\pi}{b_0 \ln a_0}$.

This construction—albeit clearly contrived—shows how readily we can construct frames for signal analysis based on translations and dilations. We can actually loosen the provisions $a_0 > 1$ and $b_0 > 0$ by reviewing the argument. We see that positive and negative integer powers of a_0 were used to justify (11.109). Hence, as long as $0 < a_0$ and $a_0 \neq 1$, the same argument applies (exercise).

Now, it turns out that this frame is not a particularly stellar choice for signal analysis [3]. For one thing, its frame elements do not consist entirely of translations and dilations of a single element; rather, there are two prototype patterns from which the others derive. Worse, however, is the fact that the elements in our special tight frame have poor time-domain decay. Their spectrum has finite support. The consequence is that inner products will have nonzero responses even when the analyzing frame elements are offset some distance away from the candidate signals. The exercises invite the reader to explore these ideas further.

11.3.4 Better Localization

We can construct frames based on translations and dilations of a Gaussian root signal. Such frames offer satisfactory time-domain decay. The drawback, however, is that the precise explication of the frame condition is not as elegant as in the previous section's special construction. There is a sufficient condition for wavelet frames [3], which, despite its ponderous formulation, allows one to estimate frame bounds for $\psi(t)$ having good time and frequency-domain decay.

11.3.4.1 Sufficient Conditions for Wavelet Frames. Let us state some theorems due to Daubechies [3].

Theorem (Sufficiency). Suppose $\psi(t) \in L^2(\mathbb{R})$, $a_0 > 1$, $\Psi(\omega)$ is the (radial) Fourier transform of $\psi(t)$, and

$$\beta(s) = \sup_{1 \leq |\omega| \leq a_0} \left(\sum_{m=-\infty}^{\infty} \left|\Psi(a_0^m \omega)\right| \left|\Psi(a_0^m \omega + s)\right| \right). \tag{11.115}$$

Further assume that

$$\inf_{1 \leq |\omega| \leq a_0} \left(\sum_{m=-\infty}^{\infty} \left|\Psi(a_0^m \omega)\right|^2 \right) > 0; \tag{11.116a}$$

$$\sup_{1 \leq |\omega| \leq a_0} \left(\sum_{m=-\infty}^{\infty} \left|\Psi(a_0^m \omega)\right|^2 \right) < \infty; \tag{11.116b}$$

and, for some $\varepsilon > 0$, $\beta(s)$ decays as fast as $(1 + |s|)^{-1-\varepsilon}$. Then there is $B_0 > 0$ such that for any $b_0 < B_0$, $F = \{\psi_{m,n}(t) = a_0^{-(m/2)}\psi(a_0^{-m}t - nb_0) | m, n \in \mathbb{Z}\}$ is a frame.

Corollary (Bounds). With the theorem's assumptions and notation, let

$$C = \sum_{\substack{k=-\infty \\ k \neq 0}}^{\infty} \left[\beta\left(\frac{2\pi k}{b_0}\right)\beta\left(-\frac{2\pi k}{b_0}\right)\right]^{\frac{1}{2}}, \tag{11.117}$$

and suppose $b_0 < B_0$. Then $F = \{\psi_{m,n}(t) \mid m, n \in \mathbb{Z}\}$ has lower and upper frame bounds, A and B, respectively:

$$A = \frac{1}{b_0}\left[\inf_{1 \leq |\omega| \leq a_0}\left(\sum_{m=-\infty}^{\infty} \left|\Psi(a_0^m\omega)\right|^2\right) - C\right] \tag{11.118a}$$

and

$$B = \frac{1}{b_0}\left[\sup_{1 \leq |\omega| \leq a_0}\left(\sum_{m=-\infty}^{\infty} \left|\Psi(a_0^m\omega)\right|^2\right) + C\right]. \tag{11.118b}$$

Proofs: Again due to Daubechies [3, 15].

Remark. These technical conditions will be met, for example, if $\psi(t)$ obeys the following:

- Its time- and frequency-domain decay rates are not too slow.
- Its spectrum is zero for $\omega = 0$: $\Psi(0) = \int \psi(t)\, dt = 0$.

The conditions do imply that $\psi(t)$ is admissible (11.2). Moreover, under these mild constraints, there will be many combinations of scale and time steps for which F comprises a frame [3].

11.3.4.2 Example: Mexican Hat. Let us consider the Mexican hat function, introduced in Section 11.2.3.2 (Figure 11.4). This signal is the second derivative of the Gaussian: $\psi(t) = \exp(-t^2/2)$. Normalizing, $\|\psi\|_2 = 1$, gives

$$\psi(t) = \frac{2\pi^{-\frac{1}{4}}}{\sqrt{3}}(1 - t^2)\exp\left(-\frac{t^2}{2}\right). \tag{11.119}$$

Table 11.2 repeats some estimates for the lower and upper frame bounds, A and B, respectively, for frames based on translations and dilations of the Mexican hat [3].

Notice that as the time domain increment b_0 increases, then frame lower bound A decreases much faster toward zero; we might interpret this as indicating that there are finite energy signals that are more and more orthogonal to the frame elements. Decreasing the scale domain increment $a = a_0^{1/k}$, $k = 1, 2, 3, 4$, etc., mitigates this tendency.

TABLE 11.2. Lower and Upper Bound Estimates for Frames Based on the Mexican Hat

b_0	A	B	$a = 2^1$
0.25	13.091	14.183	
0.50	6.546	7.092	
0.75	4.364	4.728	
1.0	3.223	3.596	
1.25	2.001	3.454	
1.50	0.325	4.221	
1.75	—	—	No frame

b_0	A	B	$a = 2^{1/2}$
0.25	27.273	27.278	Nearly exact
0.50	13.673	13.676	
0.75	9.091	9.093	
1.0	6.768	6.870	
1.25	4.834	6.077	
1.50	2.609	6.483	
1.75	0.517	7.276	

b_0	A	B	$a = 2^{1/3}$
0.25	40.914	40.914	Nearly exact
0.50	20.457	20.457	
0.75	13.638	13.638	
1.0	10.178	10.279	
1.25	7.530	8.835	
1.50	4.629	9.009	
1.75	1.747	9.942	

b_0	A	B	$a = 2^{1/4}$
0.25	55.552	55.552	Nearly exact
0.50	27.276	27.276	
0.75	18.184	18.184	
1.0	13.586	13.690	
1.25	10.205	11.616	
1.50	6.594	11.590	
1.75	2.928	12.659	

What time- and scale-domain increments make the best choices? To answer this question, we recall the formula for reconstructing a signal $x(t)$ from frame elements (Section 3.3.4). Let $\{\psi_k(t): k \in \mathbb{Z}\}$ enumerate the doubly indexed frame F of translations and dilations of $y(t)$, $F = \{\psi_{m,n}(t)\}$. Then,

$$x = \sum_k \langle x, S^{-1}\psi_k\rangle \psi_k = \sum_k \langle x, \psi_k\rangle S^{-1}\psi_k, \tag{11.120}$$

where $S = T^*T$; T is the frame operator, $T_F(x)(k) = \langle x, \psi_k\rangle$; and T^* is the frame operator adjoint, $T^*(s) = \sum_{k=-\infty}^{\infty} s(k)\psi_k$, where $s(k) \in l^2(\mathbb{Z})$. Now by the Frame Characterization Theorem of Section 3.3.4.3, we can write the frame condition as $AI \le S \le BI$, where I is the identity operator on $L^2(\mathbb{R})$.

Suppose that the lower and upper frame bounds are almost equal, a condition that several of the alternatives in Table 11.2 allow [3]. As $B \to A$, $\varepsilon = B/A - 1 \to 0$, and the operator $S = T^*T$ is close to a midpoint operator between AI and BI: $S \approx \frac{A+B}{2} I$. Thus, $S^{-1} \approx \frac{2}{A+B} I$, and (11.120) becomes

$$x = \sum_k \langle x, S^{-1}\psi_k\rangle \psi_k \approx \frac{2}{A+B} \sum_k \langle x, \psi_k\rangle \psi_k. \tag{11.121}$$

Equation (11.121) is a simple, approximate reconstruction formula for $x(t)$ that is valid when the frame F is almost exact. Thus, choosing time and scale dilation factors that provide an almost exact frame facilitates reconstruction of $x(t)$ from its frame coefficients.

The next section develops wavelet theory that provides not just frames, but orthonormal bases for finite energy signals based on translations and dilations of a single prototype signal.

11.4 MULTIRESOLUTION ANALYSIS AND ORTHOGONAL WAVELETS

After the publication of Grossmann and Morlet's paper in 1984, wavelet methods attracted researchers—including the present authors—from a broad range of scientific and engineering disciplines.[2] The new scale-based transform posed an alternative to short-time Fourier techniques for seismic applications [13, 16]. It facilitated the construction of frames (Section 11.3), which are necessary for numerically stable signal modeling [14, 15]. Wavelets were used for analyzing sound waves [17] and adapted to multiscale edge detection [18]. Academic meetings were exciting. It was still unclear how powerful this tool could become.

[2]"Such a portentous and mysterious monster roused all my curiosity" (Melville).

The wavelet transform was but one of several mixed domain signal transforms known in the mid-1980s. Among the others were time-frequency techniques such as the short-time Fourier transform and the Wigner distribution (Chapter 10). Wavelet analysis, in contrast, represents a scale-based transform.

Recall that if we set $\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right)$, then the wavelet transform

$$X_w(a,b) = \langle x, \psi_{a,b}\rangle = \int_{-\infty}^{\infty} x(t)\overline{\psi_{a,b}(t)}\, dt \tag{11.122}$$

measures the similarity of $x(t)$ and the scaled, shifted wavelet $\psi_{a,b}(t)$. This makes it a multiscale shape detection technique.

Because Grossmann and Morlet's wavelets are also bandpass filters, the convolution (11.122) effects a frequency selection from the source signal $x(t)$. Assuming that $\psi(t) \in L^2(\mathbb{R})$ is an analyzing wavelet and $\Psi(\omega)$ is its radial Fourier transform, the inverse wavelet transform is given by

$$x(t) = \frac{1}{C_\Psi}\int_0^\infty\int_{-\infty}^{\infty} X_w(a,b)\frac{\psi_{a,b}(t)}{a^2}\, dbda\,, \tag{11.123}$$

where C_ψ is the admissibility factor, $C_\psi = \int_{-\infty}^{\infty}\frac{|\Psi(\omega)|^2}{|\omega|}\, d\omega < \infty$. Thus, the transform (11.122) characterizes the signal $x(t)$ and can be the basis for signal comparisons, matching, and interpretation.

To accomplish wavelet-based signal analysis on a computer requires, of course, that the transform be discretized. For example, we might study transform coefficients of the form $x_w(m,n) = X_w(m\Delta, nT) = \langle x(t), \psi_{m\Delta,nT}(t)\rangle$. This leads to the construction of wavelet frames, which support signal characterization and numerically stable representation. These benefits would be all the stronger if a wavelet frame could be somehow refined into an orthonormal basis.

Nonetheless, it was the intuition of pioneering researchers that—just as they had shown for windowed Fourier expansions—a Balian–Low type of result would hold for wavelets, precluding orthonormal bases. As exciting as the developments of the 1980s had been, the prospects for well-localized short-time Fourier bases appeared quite bleak. The critical time-frequency sampling density, $T\Omega = 2\pi$, does not permit frames let alone orthogonal windowed Fourier bases, unless either the windowing function $w(t)$ or its Fourier transform $W(\omega)$ fails to be well-localized: $\|tw(t)\|_2 = \infty$ or $\|\omega W(\omega)\|_2 = \infty$ (Section 10.5.1). Anticipating equally negative results for the new scale-based transforms too, Meyer [19] tried to prove a version of Balian–Low for wavelets. To his own and everyone else's surprise, he failed and instead found an orthonormal wavelet basis!

Meyer's basis [4,19] proceeds from a wavelet $\psi(t)$ whose normalized Fourier transform is given by

$$\Psi(\omega) = \begin{cases} \dfrac{e^{\frac{j\omega}{2}}}{\sqrt{2\pi}} \sin\left[\dfrac{\pi}{2}\upsilon\left(\dfrac{3|\omega|}{2\pi} - 1\right)\right] & \text{for} \quad \dfrac{2\pi}{3} \le |\omega| \le \dfrac{4\pi}{3}, \\ \dfrac{e^{\frac{j\omega}{2}}}{\sqrt{2\pi}} \cos\left[\dfrac{\pi}{2}\upsilon\left(\dfrac{3|\omega|}{2\pi} - 1\right)\right] & \text{for} \quad \dfrac{4\pi}{3} \le |\omega| \le \dfrac{8\pi}{3}, \\ 0 & \text{otherwise}. \end{cases} \tag{11.124}$$

In (11.124) $\upsilon(t)$ is a C^k signal, where $v(t) \approx u(t)$, except on (0, 1). It specializes the $v(t)$ used earlier; the extra proviso is $v(t) + v(1 - t) = 1$. Figure 11.8 shows Meyer's wavelet $\psi(t) = \mathcal{F}^{-1}[\Psi(\omega)]](t)$.

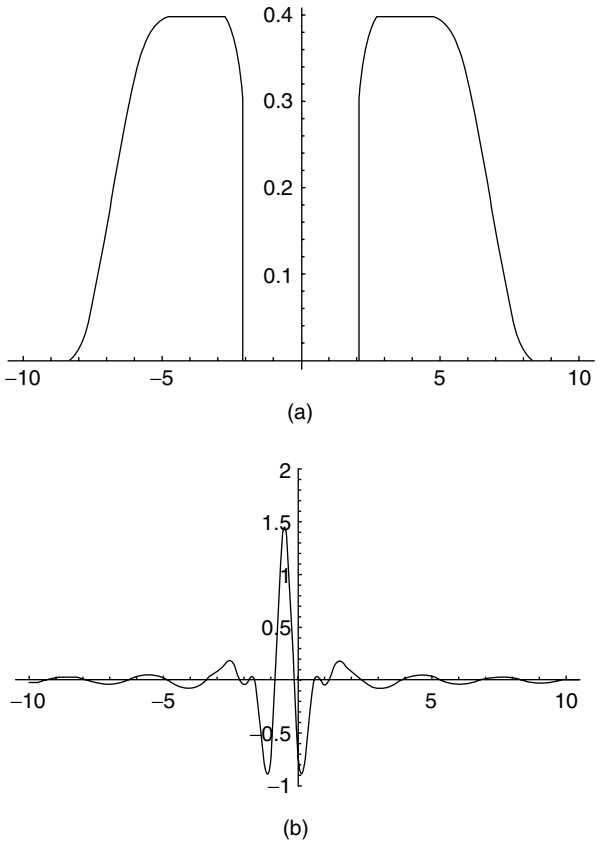


Fig. 11.8. Its normalized Fourier transform (a) and the Meyer wavelet (b). Its translations and dilations form an orthonormal basis for L^2 signals.

Until this discovery, mathematicians had more or less given up on finding orthonormal expansions for $L^2(\mathbb{R})$ using smooth basis elements. It was assumed that there had to be discontinuities in the time domain such as with Haar's basis, which was discovered some 75 years earlier [20], or in the frequency domain, such as with Shannon's $\text{sinc}(t) = \sin(t)/t$ basis of Ref. 21 (Section 3.3.3). A later construction of Strömberg provides another orthonormal basis of continuous functions [22]; it predated Meyer's function by a few years and had been largely overlooked.

From the desks of several other mathematicians, more examples of orthonormal wavelet bases soon issued. But—except for the intricate calculations, carefully concocted estimations, and miraculous cancellations—there seemed to be no connection between these diverse constructions. Could there be no rules for building them? It was a fascinating mess.

11.4.1 Multiresolution Analysis

The unifying breakthrough came when Mallat [23] and Meyer elaborated the concept of a *multiresolution analysis* (MRA) for square-integrable signals. A computer vision researcher, Mallat was especially inspired by the similarities between some of the recent wavelet basis developments and work in pyramid decompositions for signal and image analysis such as the Laplacian pyramid [24], quadrature mirror filter banks employed in communication engineering [25], and scale space decompositions [26–28].

The MRA concept leads to a rich theory of the scale-based structure of signals. As a bonus, the MRA establishes a set of rules for constructing a wide range of orthonormal wavelet bases. Mallat and Meyer found the rules for building orthonormal wavelet bases in a quite unexpected place: the very applied areas of multiscale signal decomposition, image analysis, and efficient communication engineering. The discovery of important theoretical concepts out of utterly practical problems seems to be a distinct characteristic within the new discipline of wavelet analysis.

11.4.1.1 Definition. A multiresolution analysis of $L^2(\mathbb{R})$ is an abstract structure, but it has close links to several signal analysis ideas that we have already covered. We present the formal definition and develop some of the theory of bases made up of translations of a single root signal. The presentation follows closely and may be considered a tutorial on the classic papers [23, 29].

Definition (Multiresolution Analysis). A *multiresolution analysis* (or *multiresolution approximation*, MRA) is a chain of closed subspaces $\{V_i : i \in \mathbb{Z}\}$ in $L^2(\mathbb{R})$ such that the following conditions hold:

(i) The V_i are nested within one another: $\ldots \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \ldots$.

(ii) The union of the V_i is dense in $L^2(\mathbb{R})$: $\overline{\bigcup_{n=-\infty}^{\infty} V_i} = L^2(\mathbb{R})$.

(iii) The intersection of the V_i is the signal of zero norm (zero almost everywhere), which we write $\bigcap_{i=-\infty}^{\infty} V_i = 0$.

(iv) Elements of the spaces are dyadically scaled (more precisely, *dilated*) versions of one another: $x(t) \in V_i \Leftrightarrow x(2t) \in V_{i+1}$.

(v) For any $x(t) \in V_0$ and any $k \in \mathbb{Z}$, $x(t-k) \in V_0$.

(vi) There is an isomorphism from V_0 onto the Hilbert space of square-summable discrete signals I: $V_0 \to l^2$ such that for any $k \in \mathbb{Z}$, if $I(x(t)) = s(n) \in l^2$, then $I(x(t-k)) = s(n-k)$.

Remark. We give the classic definition and notation for the MRA [23]. It has become common to index the V_i in the other direction: $V_i \supset V_{i+1}$. So readers must pay close attention to an author's V_i indexing.

Nowadays, many treatments (for instance, Refs. 3, 8, and 9) replace (vi) with the requirement that V_0 has an orthonormal basis of translates of a single finite-energy signal: $\{\phi(t-n) \mid n \in \mathbb{Z}\}$. This works. But so early on, it also seems incredible; we prefer to proceed from the apparently weaker criterion. In Section 11.4.2 we demonstrate that there is indeed a special function in V_0, called a *scaling function*, whose translates comprise an orthonormal basis of V_0.

Finally, note that by an *isomorphism* in (vi) we mean only a bounded, one-to-one, linear map, with a bounded inverse. Some mathematics texts, for example [30], define the term to mean also $\langle Ix, Iy\rangle = \langle x, y\rangle$, which implies an *isometry*; this we do not assume herein. The last MRA property (vi) is very strong, though. The isomorphism is a bounded linear map: There is an M such that $\|Ix\| \le M\|x\|$ for all $x \in V_0$. For linear maps this is equivalent to continuity. Range(I) is all of l^2. If it were an isometry, then we would be easily able to show that V_0 has a scaling function; but with our weaker assumption, this requires quite a bit more work.

11.4.1.2 Examples. Although it is rich with mathematical conditions, which might appear difficult to satisfy, we can offer some fairly straightforward instances of multiresolution analyses. Here are three examples where the root spaces consist of:

- Step functions;
- Piecewise linear functions;
- Cubic splines.

Note that the root spaces in these examples contain increasingly smooth signals.

Example (Step Functions). It is easiest to begin with a root space V_0 comprised of step functions and define the spaces of non-unit scale by dilation of V_0 elements. Let $u(t)$ be the analog unit step signal and set $V_0 = \{x(t) \in L^2(\mathbb{R}) \mid$ for all $n \in \mathbb{Z}$, there is a $c_n \in \mathbb{R}$ such that $x(t) = c_n[u(t-n) - u(t-n-1)]$ for $t \in (n, n+1)\}$. So elements of V_0 are constant on the open unit intervals $(n, n+1)$. The boundary values of the

signals do not matter, since $\mathbb{Z}$ is a countable set and thus has Lebesgue measure zero. We define $V_i = \{y(t) \in L^2(\mathbb{R}) \mid \text{for some } x(t) \in V_0,\ y(t) = x(2^i t)\}$. Thus, V_i signals are constant on intervals $(n2^{-i}, (n+1)2^{-i})$, where $n \in \mathbb{Z}$. Let us show that each of the MRA properties holds.

(i) Signals that are constant on $(n2^{-i}, (n+1)2^{-i})$ for all $n \in \mathbb{Z}$ will also be constant on subintervals $(n2^{-i-1}, (n+1)2^{-i-1})$, so the first property holds.

(ii) From Chapter 3, we know that the step functions are dense in $L^2(\mathbb{R})$; since arbitrarily narrow steps are contained in V_i for i sufficiently large, we know that the V_i are dense in $L^2(\mathbb{R})$.

(iii) For a nonzero signal to be in all of the V_i, it would have to have arbitrarily wide steps, so the intersection property must be satisfied.

(iv) This is how we define the V_i for $i \neq 0$.

(v) Integral translates of signals in V_0 are obviously in V_0, since an integral translate is still constant on unit intervals.

(vi) If $x(t) \in V_0$, and $x(t) = c_n(u(t-n) - u(t-n-1)]$ for $t \in (n, n+1)$, then we set $I(x(t)) = s(n)$, where $s(n) = c_n$ for all $n \in \mathbb{Z}$; then $I(x(t-k)) = s(n-k)$, and I is an isomorphism. This is left as an exercise.

This MRA is the orthonormal basis of Haar in modern guise [20]. For analyzing blocky signals, this simple MRA is appropriate (Figure 11.9a). But when studying smoother signals, decompositions on the Haar set require quite a few coefficients in

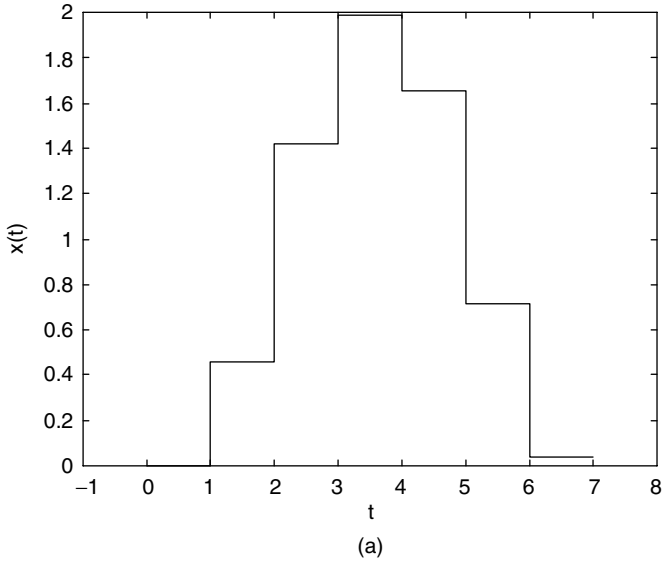
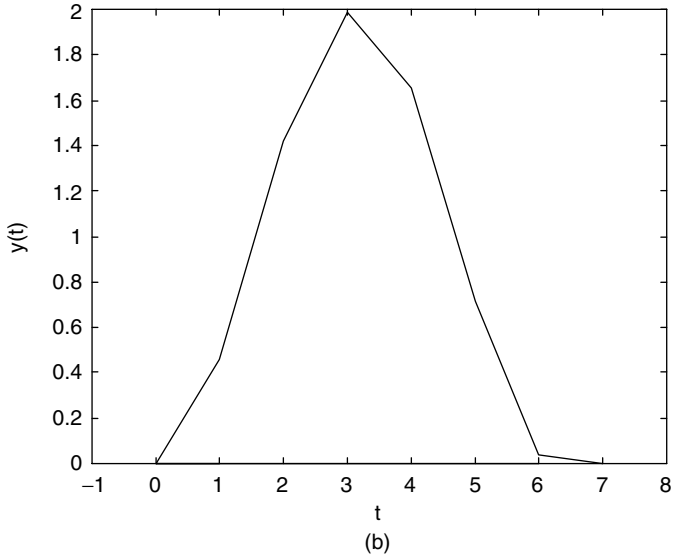


Fig. 11.9. Typical elements of an MRA built by step functions (a), piecewise linear functions (b), and cubic splines (c).



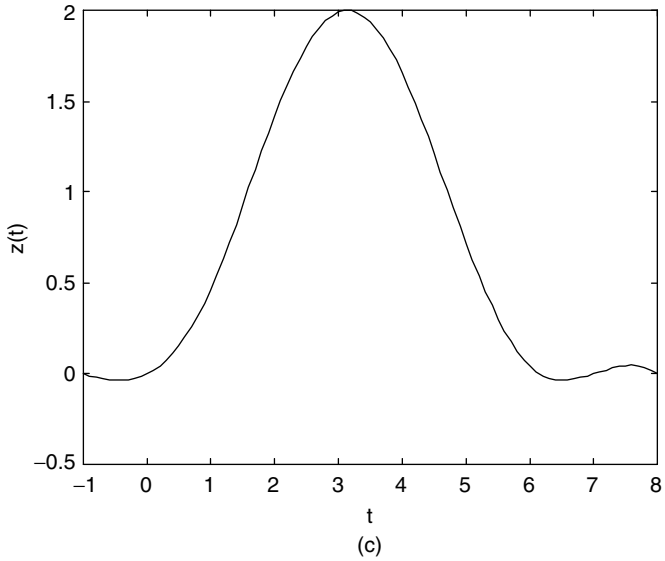


Fig. 11.9 (*Continued*)

order to smooth out the discontinuities present in the basis elements. So the approximations are often quite inefficient.

Example (Piecewise Linear Functions). Let us refine the multiscale structure so that it employs piecewise continuous signals. MRAs using such functions are better

for signal analysis purposes when we have to interpret signals that do not contain abrupt jumps in value (Figure 11.9b). Let us define V_0 to be the $L^2(\mathbb{R})$ continuous functions that are piecewise linear on integral intervals $[n, n+1]$. We define the remaining spaces via the MRA's dilation property (iv): $x(t) \in V_i \Leftrightarrow x(2t) \in V_{i+1}$. The six MRA properties are clear, except perhaps for density (ii) and the isomorphism (vi). Note, however, that a piecewise continuous $x(t) \in V_0$ is determined uniquely by its values on integers. We define $I(x(t)) = s(n)$, where $s(n) = x(n)$ for all $n \in \mathbb{Z}$. We can approximate a step function to arbitrary precision with piecewise linear functions, and the step functions are dense in $L^2(\mathbb{R})$, so the second MRA property holds. Strömberg [22] elaborated this MRA's theory.

Example (Cubic Splines). A more interesting example relies on cubic splines. Here, the root space V_0 consists of all functions that are twice continuously differentiable and equal to a cubic polynomial on integral intervals $[n, n+1]$. Again, dilation defines the other V_i: $x(t) \in V_i \Leftrightarrow x(2t) \in V_{i+1}$. Numerical analysis texts, (e.g., Ref. 31, show that a continuous function can be approximated to any precision with a cubic spline, so the V_i are dense in $L^2(\mathbb{R})$.

As we continue to develop MRA theory, we shall return to these examples.

11.4.1.3 Links to Signal Analysis Legacy. Before expounding more theory, let us reflect on how the multiresolution analysis concept echoes many ideas from previous multiscale signal analysis techniques. In fact, stripped of their mathematical formalism, several of the multiresolution analysis properties (i)–(vi) have conceptual precedents in prior multiscale representations. Others embody ideas that were only implicit in the intuitive constructs of earlier methods of interpretation.

For example, the nested sequence of subspaces in (i) embodies the concept of the representations becoming ever finer in resolution. The inclusion property indicates that every coarse representation of a signal may also be considered to be a fine resolution version of some other waveform that has an even coarser shape. The subspaces are closed; each V_i contains the limit of its convergent function sequences. Coarse resolution representations are useful because:

- Using them can reduce the time required for pattern searches, such as in elementary edge and feature detection applications [32–34].
- Some signal features appear only at certain scales [35].

We need closure to guarantee that given a finite-energy signal $x(t)$ and an approximation error, there is some V_i that approximates $x(t)$ to within the tolerance.

What does the union property (ii) mean? If one looks in a sufficiently fine resolution space, then there is a finite energy signal that is arbitrarily close to any given signal. V_i signals are only approximations of real signals, but we can choose them to be very good approximations. Notice that (i) and (ii) encapsulate the intuitive notion of earlier researchers that scale is critical for structural decompositions of signals.

To interpret a signal we have to determine either the specific time extents of signal shapes, or we must search for shapes across all scales. This insight is the motivation behind scale space analysis [26–28], which we first considered in Chapter 4.

The intersection property (iii) tells us that, from a scale-based signal analysis standpoint, any meaningful signal must be visible to the MRA at some scale. That is, if a signal is composed of structures that have such a fine scale that it must appear in all the subspaces, then this signal must be the almost everywhere zero signal.

The next property concern resolution. Next, if a function is in space V_i, then its dilation by a factor of 2 is in the next higher resolution space V_{i+1}. Furthermore, its dilation by a factor of 1/2 is in the lower resolution space V_{i-1}. Thus, the implication of (iv) is that one ascends and descends in resolution by means of dilations, such as in the classic Laplacian pyramid construction [24].

The subspace V_0 contains signals that resolve features to the unit of measure. Signals in V_0 may be translated by integral amounts, and we are assured that the result remains in the root space (v).

Lastly, discrete samples characterize the MRA functions that model the signals. Property (vi) formalizes this by requiring an isomorphism between V_0 and the square-summable sequences of real numbers. This discretization has the further property that the discrete representations of functions are translated when the function is translated. Without this invariance, the discrete samples associated with a waveform $x(t)$ in V_0 might change with the translation of $x(t)$ by integral steps. For the lower resolution spaces V_i, $i < 0$, translation invariance does not hold. It will become apparent in applications that the overall situation for translation invariance is far from satisfactory; indeed, it is problematic.

11.4.1.4 Bases of Translates: Theory. Our first theoretical result on MRAs comes right out of the sixth criterion. It shows that the root space V_0 has a basis, consisting of integral translates of a single square-integrable signal. Multiresolution analysis exploits the special properties of such bases.

Proposition (Basis). If $\{V_i \mid i \in \mathbb{Z}\}$ is an MRA in $L^2(\mathbb{R})$, then there is $e(t) \in V_0$ such that $\{e(t-k) \mid k \in \mathbb{Z}\}$ is a basis for V_0.

Proof: Let $\delta(n-k)$ be the discrete delta signal delayed by $k \in \mathbb{Z}$. $\{\delta(n-k) \mid k \in \mathbb{Z}\}$ is an orthonormal basis for l^2. Let I: $V_0 \to l^2$ be the isomorphism of MRA property (vi). Since I is onto, we may set $e_k(t) = I^{-1}(\delta(n-k))$. Then $\{e_k(t) \mid k \in \mathbb{Z}\}$ is a basis for V_0. However, $I(e_0(t)) = \delta(n)$, so that the translation-invariance provision of (vi) also implies $I(e(t-k)) = \delta(n-k)$, whence $e_k(t) = e(t-k)$. ■

The proposition guarantees a basis that is useful for those pattern matching applications where we expect candidate signals containing the shape of the root element. The basis elements comprise our model or prototype signals. For computational purposes, we prefer such bases to be orthonormal, since that simplifies expansion coefficient computations. But, again, orthogonality requires more work. Let us explore some of the theory of such bases.

Definition (Riesz Basis). $E = \{e_n \mid n \in \mathbb{Z}\}$ is an *unconditional* or *Riesz basis* in a Hilbert space H if

(i) E spans H: For any $x \in H$, there is $s \in l^2$ such that

$$x = \sum_{n=-\infty}^{\infty} s(n) e_n . \tag{11.125a}$$

(ii) There are $0 < A \le B < \infty$ such that for any $s \in l^2$,

$$A\|s\| \le \left\| \sum_{n=-\infty}^{\infty} s(n) e_n \right\| \le B\|s\| . \tag{11.125b}$$

The constants A and B are called the *lower* and *upper Riesz basis bounds*, respectively.

Notice how (11.125b) cross-couples the norms of H and l^2. An orthonormal basis is also a Riesz basis. Note too that the lower bound condition in (11.125b) implies that a Riesz basis is linearly independent (exercises). The next result shows that property (vi) of an MRA is equivalent to V_0 having a Riesz basis.

Theorem (Riesz Basis Characterization). $E = \{e_k \mid k \in \mathbb{Z}\}$ is a Riesz basis in a Hilbert space H if and only if:

(i) There is an isomorphism I from H onto l^2 such that $I(e_k) = \delta(n-k)$, where $\delta(n)$ is the discrete delta.
(ii) I^{-1} is bounded.

Proof: Let $E = \{e_k \mid k \in \mathbb{Z}\}$ be a Riesz basis in H. Since for any $x \in H$, there is $s \in l^2$ such that (11.125a) holds, we may set $Ix = s$. This map is well-defined by the linear independence of E; that is, for each x, the coefficient sequence $\{s(n)\}$ is unique. The linearity follows from the properties of the square-summable sequences, and clearly $I(e_k) = \delta(n-k)$. The map is also onto, because any $s \in l^2$ defines an element of H as in (11.125a). This sum converges in H by the upper bound inequality in (11.125b). The boundedness of I follows as well: $\|Ix\| = \|s\| \le A^{-1}\|x\|$. The inverse $J = I^{-1}$ exists because I is one-to-one and onto (a *bijection*). $\|Js\| = \|x\| = \|\Sigma s(k) e_k\| \le B\|s\|$ by (11.125b), so J is bounded too.

Conversely, let I: $H \to l^2$ be an isomorphism obeying (i) and (ii). Let $x \in H$. We need to show that x is in the closure of Span(E). Let $Ix = s \in l^2$. In general, it does not necessarily follow that if an isomorphism is bounded, then its inverse is bounded. However, we are assuming $J = I^{-1}$ is bounded, so it is also continuous [30]. Thus,

$$\lim_{N\to\infty} J\left(\sum_{k=-N}^{N} s(k)\delta(n-k) \right) = \sum_{k=-\infty}^{\infty} s(k) J(\delta(n-k)) = \sum_{k=-\infty}^{\infty} s(k) e_k, \tag{11.126}$$

and x is a limit of elements in Span(E). Let $\|I\| = A^{-1}$ and $\|J\| = B$. Then A and B are the lower and upper Riesz bounds for E (11.125b). ■

Corollary. Property (vi) of an MRA is equivalent to V_0 having a Riesz basis of translates.

Proof: Since the isomorphism I in (vi) has a bounded inverse and I is onto, we may find $e_k(t)$ such that $I(e_k(t)) = \delta(n - k)$, where $\delta(n)$ is the discrete delta signal. The theorem tells us that $\{e_k(t) \mid k \in \mathbb{Z}\}$ is a Riesz basis. The translation invariance for V_0 implies that $e_k(t) = e(t - k)$. ■

The conditions (11.125a) and (11.125b) for a Riesz basis resemble the criterion for a frame, which we studied in Chapters 3 and 10. Indeed, the following corollary shows that a Riesz basis is a frame. Of course, the converse is not true; a Riesz basis must be linearly independent, while frames can be overcomplete.

Corollary (Frame). If $E = \{e_k \mid k \in \mathbb{Z}\}$ is a Riesz basis in a Hilbert space H, then E is a frame.

Proof: Let let $I: H \to l^2$ be the isomorphism promised in the Riesz basis characterization: $I(e_k) = \delta(n - k)$. Let I^* be the Hilbert space adjoint operator for I. We introduced the adjoint operator in Chapter 3 and therein applied it to the study of frames (Section 3.3.4). The adjoint cross-couples the inner product relations of H and l^2 so that if $s(n)$ is square-summable, then $\langle Ix, s\rangle = \langle x, I^*s\rangle$. Note that $I^*: l^2 \to H$ is an isomorphism, bounded, and in fact $\|I^*\| = \|I\|$. For example, to show I^* is one-to-one, let $I^*v = I^*w$ for some $v(n)$, $w(n) \in l^2$. Then for any $h \in H$, $\langle h, I^*v\rangle = \langle h, I^*w\rangle$. But this implies $\langle Ih, v\rangle = \langle Ih, w\rangle$. Since I is onto and h is arbitrary, Ih could be any finite-energy discrete signal. In other words, $\langle s, v\rangle = \langle s, w\rangle$ for all $s \in l^2$. But this means $v = w$ too. We leave the remaining I^* details as exercises.

Now, let $f_k = I^*(\delta(n - k))$. Since $\delta(n - k) = (I^*)^{-1}(f_k)$ and $((I^*)^{-1})^{-1} = I^*$ is bounded, the proposition says that $F = \{f_k \mid k \in \mathbb{Z}\}$ is a Riesz basis in H. If $x \in H$, then by (11.125a) there is $s \in l^2$ such that $x = \sum_{k=-\infty}^{\infty} s(k) f_k$. We see that $\langle x, e_k\rangle = s(k)$ by calculating $\langle f_i, e_k\rangle = \langle I^*(\delta(n - i)), e_k\rangle = \langle \delta(n - i), I(e_k)\rangle = \langle \delta(n - i), \delta(n - k)\rangle$. So

$$\sum_{k=-\infty}^{\infty} \left|\langle x, e_k\rangle\right|^2 = \sum_{k=-\infty}^{\infty} |s(k)|^2 = \|s\|_2^2. \tag{11.127}$$

Since I^* and $(I^*)^{-1}$ are bounded and $x = I^*s$, we have

$$\frac{\|x\|^2}{\|I^*\|^2} \le \sum_{k=-\infty}^{\infty} \left|\langle x, e_k\rangle\right|^2 = \sum_{k=-\infty}^{\infty} |s(k)|^2 \le \|x\|^2 \left\|(I^*)^{-1}\right\|^2, \tag{11.128}$$

which is precisely a frame condition on E. ■

Theorem (Orthonormal Translates). Let $\phi(t) \in L^2(\mathbb{R})$ and $\Phi(\omega) = \mathcal{F}[\phi(t)](\omega)$ be its (radial) Fourier transform. The family $F = \{\phi(t-k) \mid k \in \mathbb{Z}\}$ is orthonormal if and only if

$$\sum_{k=-\infty}^{\infty} \left|\Phi(\omega + 2\pi k)\right|^2 = 1 \tag{11.129}$$

for almost all $\omega \in \mathbb{R}$.

Proof: An interesting application of Fourier transform properties makes this proof work. Let us define $a_k = \langle \phi(t), \phi(t-k) \rangle$. Note that—by a simple change of variable in the inner product integral—F is orthonormal if and only if a_k is zero when $k \neq 0$ and unity when $k = 0$. We calculate

$$a_k = \langle \phi(t), \phi(t-k) \rangle = \frac{1}{2\pi} \langle \Phi(\omega), \Phi(\omega) e^{-jk\omega} \rangle = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi(\omega) \overline{\Phi(\omega)} e^{jk\omega} d\omega. \tag{11.130}$$

by the Parseval and shift properties. The right-hand integrand in is $|\Phi(\omega)|^2 e^{jk\omega}$. We break up the integral into 2π-wide pieces, invoke the exponential's periodicity, and swap the order of summation and integration to get

$$a_k = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \int_0^{2\pi} \left|\Phi(\omega + 2\pi n)\right|^2 e^{jk\omega} \, d\omega = \frac{1}{2\pi} \int_0^{2\pi} e^{jk\omega} \sum_{n=-\infty}^{\infty} \left|\Phi(\omega + 2\pi n)\right|^2 d\omega. \tag{11.131}$$

We move the sum into the integral, since $\Phi = \mathcal{F}\phi \in L^2(\mathbb{R})$, so that $|\Phi(\omega)|^2 \in L^1(\mathbb{R})$. Let us define

$$P_\Phi(\omega) = \sum_{n=-\infty}^{\infty} \left|\Phi(\omega + 2\pi n)\right|^2. \tag{11.132}$$

Now observe

$$\int_0^{2\pi} \sum_{n=-\infty}^{\infty} \left|\Phi(\omega + 2\pi n)\right|^2 d\omega = \int_0^{2\pi} P_\Phi(\omega) \, d\omega = \int_{-\infty}^{\infty} \left|\Phi(\omega)\right|^2 d\omega = \left\|\Phi(\omega)\right\|_2^2, \tag{11.133}$$

so that $P_\Phi(\omega)$ is finite for almost all $\omega \in \mathbb{R}$. This allows us to interchange summation and integration with the Lebesgue integral (Chapter 3) in (11.131). We can say more about $P_\Phi(\omega)$: It is 2π-periodic, and the right-hand side of (11.131) is precisely the expression for its Fourier series coefficient, which is a_k. Within the inner product there hides nothing less than a Fourier series analysis equation for the special periodic function $P_\Phi(\omega)$! We use this periodization argument a lot.

Let us check our claim that $P_\Phi(\omega) = 1$ almost everywhere if and only if the family of translates F is orthonormal. First, if $P_\Phi(\omega) = 1$, then (11.131) becomes

$$a_k = \frac{1}{2\pi}\int_0^{2\pi} e^{jk\omega} d\omega = \langle \phi(t), \phi(t-k)\rangle. \tag{11.134}$$

The integral in (11.134) is 2π if $k = 0$ and zero if $k \neq 0$. So $F = \{\phi(t-k)\}$ must be orthonormal. Conversely, suppose F is orthonormal so that $a_k = 1$ if $k = 0$ and $a_k = 0$ if $k \neq 0$. Because (11.131) gives the Fourier series coefficients for the 2π-periodic function $P_\Phi(\omega)$, we know that $P_\Phi(\omega)$ has all zero Fourier series coefficients except for its DC term, which is one. In other words,

$$P_\Phi(\omega) = \sum_{k=-\infty}^{\infty} a_k e^{jk\omega} = 1e^0 = 1. \tag{11.135}$$

■

The following corollary shows that when the translates of $\phi(t)$ are orthonormal, then its spectrum, as given by the support of $\Phi(\omega)$, cannot be too narrow [8]. Scaling functions cannot have simple frequency components. This result uses the Lebesgue measure of a set, an idea introduced in Section 3.4.1.

Corollary (Spectral Support). Let $\phi(t) \in L^2(\mathbb{R})$, $\Phi(\omega) = \mathcal{F}[\phi(t)](\omega)$ be its (radial) Fourier transform, let $\text{Support}(\Phi) = \{\omega \in \mathbb{R} \,|\, \Phi(\omega) \neq 0\}$, and let $\mu(A)$ be the Lebesgue measure of a measurable set A. If the family $F = \{\phi(t-k) \,|\, k \in \mathbb{Z}\}$ is orthonormal, then $\mu(\text{Support}(\Phi)) \geq 2\pi$. Under these assumptions, moreover, $\mu(\text{Support}(\Phi)) = 2\pi$ if and only if $|\Phi(\omega)| = \chi_A$, for some Lebesgue measurable $A \subset \mathbb{R}$ with $\mu(A) = 2\pi$.

Proof: Since $\|\phi\|_2 = 1$, we know $\|\Phi\|_2 = (2\pi)^{1/2}$, by Plancherel's formula. The theorem then implies $|\Phi(\omega)| \leq 1$ for almost all $\omega \in \mathbb{R}$. Consequently,

$$\mu(\text{Support}(\Phi)) = \int_{\text{Support}(\Phi)} 1\, d\omega \geq \int_{-\infty}^{\infty} |\Phi(\omega)|^2 d\omega = 2\pi. \tag{11.136}$$

Now suppose (11.136) is an equality, but $0 < |\Phi(\omega)| < 1$ on some set $B \subset \mathbb{R}$ with $\mu(B) > 0$. Then

$$\int_B |\Phi(\omega)|^2 d\omega < \int_B 1\, d\omega = \mu(B) \tag{11.137}$$

and

$$\begin{aligned} \|\Phi\|_2^2 = 2\pi &= \int_{\text{Support}(\Phi)} |\Phi(\omega)|^2 d\omega < \mu(\text{Support}(\Phi)\backslash B) + \mu(B) \\ &= \mu(\text{Support}(\Phi)) = 2\pi \end{aligned} \tag{11.138}$$

a contradiction. Conversely, assume F is orthonormal and $|\Phi(\omega)| = \chi_A$, for some Lebesgue measurable $A \subset \mathbb{R}$ with $\mu(A) = 2\pi$. Then we quickly see that

$$\mu(A) = \mu(\text{Support}(\Phi)) = \|\Phi\|_2^2 = 2\pi. \tag{11.139}$$

■

Here is a second characterization of unconditional bases. In the next section, we use this result to find the scaling function for an MRA. We begin with a lemma [9].

Lemma. Suppose $\phi(t) \in L^2(\mathbb{R})$, $\Phi(\omega) = \mathcal{F}[\phi(t)](\omega)$ is its (radial) Fourier transform, $F = \{\phi(t-k) \mid k \in \mathbb{Z}\}$, and we define $P_\Phi(\omega)$ as above (11.132). If $s \in l^2$, then

$$\left\| \sum_{k=-\infty}^{\infty} s(k)\phi(t-k) \right\|_2^2 = \frac{1}{2\pi} \int_0^{2\pi} |S(\omega)|^2 P_\Phi(\omega)\, d\omega, \tag{11.140}$$

where $S(\omega)$ is the discrete-time Fourier transform of $s(k)$.

Proof: Let us consider a linear combination of the $\phi(t-k)$, $\sum_{k=p}^{q} s(k)\phi(t-k)$, where $s \in l^2$. Using the above periodization technique, we compute

$$\begin{aligned} \left\| \sum_{k=p}^{q} s(k)\phi(t-k) \right\|_2^2 &= \frac{1}{2\pi} \left\| \sum_{k=p}^{q} s(k) e^{-jk\omega} \Phi(\omega) \right\|_2^2 \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left| \sum_{k=p}^{q} s(k) e^{-jk\omega} \right|^2 |\Phi(\omega)|^2 d\omega = \frac{1}{2\pi} \int_0^{2\pi} \left| \sum_{k=p}^{q} s(k) e^{-jk\omega} \right|^2 P_\Phi(\omega)\, d\omega \end{aligned} \tag{11.141}$$

By assumption, $s(k)$ is square-summable. Hence its discrete-time Fourier transform exists (Chapter 7), and we may pass to the double summation limit in (11.141). Indeed, as $p, q \to \infty$, the last integrand in (11.141) becomes $|S(\omega)|^2 P_\Phi(\omega)$, where $S(\omega)$ is the DTFT of $s(k)$. ■

Theorem (Riesz Translates Basis Characterization). Suppose $\phi(t) \in L^2(\mathbb{R})$, $\Phi(\omega) = \mathcal{F}[\phi(t)](\omega)$ is its (radial) Fourier transform, $F = \{\phi(t-k) \mid k \in \mathbb{Z}\}$, $0 < A \le B < \infty$, and we define $P_\Phi(\omega)$ as above (11.132). Then the following are equivalent:

(i) F is a Riesz basis with lower and upper bounds $\sqrt{A}$ and $\sqrt{B}$, respectively.
(ii) $A \le P_\Phi(\omega) \le B$ for almost all $\omega \in \mathbb{R}$.

Proof: Suppose (ii), and let $s \in l^2$. Then $A|S(\omega)|^2 \le |S(\omega)|^2 P_\Phi(\omega) \le B|S(\omega)|^2 < \infty$ almost everywhere, too, where $S(\omega)$ is the DTFT of $s(k)$. Integrating on $[0, 2\pi]$, we see

$$\frac{A}{2\pi} \int_0^{2\pi} |S(\omega)|^2 d\omega \le \frac{1}{2\pi} \int_0^{2\pi} |S(\omega)|^2 P_\Phi(\omega)\, d\omega \le \frac{B}{2\pi} \int_0^{2\pi} |S(\omega)|^2 d\omega\,. \tag{11.142}$$

We know that $2\pi\|s\|^2 = \|S\|^2$, and so the lemma implies

$$A\|s\|_2^2 \leq \left\| \sum_{k=-\infty}^{\infty} s(k)\phi(t-k) \right\|_2^2 \leq B\|s\|_2^2 \; ; \tag{11.143}$$

this is precisely the Riesz basis condition for lower and upper bounds $\sqrt{A}$ and $\sqrt{B}$, respectively.

Now let us assume (i) and try to show $A \leq P_\Phi(\omega)$ almost everywhere on $[0, 2\pi]$. Following [9], we set $Q_{\Phi,a} = \{\omega \in [0, 2\pi] \colon P_\Phi(\omega) < a\}$. If the Lebesgue measure of $Q_{\Phi,a}$, $\mu(Q_{\Phi,a})$, is zero for almost all $a \in \mathbb{R}$, then $P_\Phi(\omega)$ diverges almost everywhere, and, in particular, $A \leq P_\Phi(\omega)$. We can thus suppose that there is some $a \in \mathbb{R}$ such that $\mu(Q_{\Phi,a}) > 0$. Let χ_a be the characteristic function on $Q_{\Phi,a}$:

$$\chi_a(\omega) = \begin{cases} 1 & \text{if } \omega \in Q_{\Phi,\,a}, \\ 0 & \text{if } \omega \notin Q_{\Phi,\,a}. \end{cases} \tag{11.144}$$

By the theory of Lebesgue measure, if a set is measurable, then so is its characteristic function. This entitles us to compute the inverse discrete-time Fourier transform of $\chi_a(\omega)$:

$$x_a(n) = \frac{1}{2\pi} \int_0^{2\pi} \chi_a(\omega) e^{j\omega n} d\omega, \tag{11.145}$$

where $x_a \in l^2$. From (i) and the lemma we see

$$\begin{aligned} A\|x_a\|_2^2 &\leq \left\| \sum_{k=-\infty}^{\infty} x_a(k)\phi(t-k) \right\|_2^2 = \frac{1}{2\pi} \int_0^{2\pi} |\chi_a(\omega)|^2 P_\Phi(\omega)\, d\omega \\ &= \frac{1}{2\pi} \int_0^{2\pi} \chi_a(\omega) P_\Phi(\omega) d\omega = \frac{1}{2\pi} \int_{Q_a} P_\Phi(\omega) d\omega. \end{aligned} \tag{11.146}$$

By our choice of $Q_{\Phi,a}$, $P_\Phi(\omega) < a$ for $\omega \in Q_{\Phi,a}$, and (11.146) entails

$$A\|x_a\|_2^2 \leq \frac{a}{2\pi} \mu(Q_{\Phi,\,a}). \tag{11.147}$$

But $\|x_a\|^2 = (2\pi)^{-1}\|\chi_a\|^2 = (2\pi)^{-1}\mu(Q_{\Phi,a})$, and, by (11.147), $A \leq a$. This gives us a contradiction by the following observation. If $A \leq P_\Phi(\omega)$ almost everywhere on $[0, 2\pi]$, then we are done. Otherwise, there must be some $U \subseteq \mathbb{R}$ such that $\mu(U) > 0$ and $P_\Phi(\omega) < A$. But then there must also be some $a > 0$ such that $P_\Phi(\omega) < a < A$ and $\mu(Q_{\Phi,a}) > 0$. But our argument above has just proven that $A \leq a$, a contradiction.

Let us continue to assume (i) and try to show $P_\Phi(\omega) \le B$ for almost all $\omega \in [0, 2\pi]$. Define $P_{\Phi,a} = \{\omega \in [0, 2\pi]: P_\Phi(\omega) > a\}$. Much like before, if $\mu(P_{\Phi,a}) = 0$ almost everywhere, then $P_\Phi(\omega) = 0$ for almost all $\omega \in [0, 2\pi]$, and thus $P_\Phi(\omega) \le B$. Assume that some $a > 0$ gives $\mu(P_{\Phi,a}) > 0$. Now the argument parallels the one just given and is left as an exercise [9]. For an alternative proof, see [8]. ■

11.4.2 Scaling Function

From our study of bases of translates and Riesz expansions, we can show that every multiresolution analysis $V = \{V_i\}$ has a special *scaling function*, whose translates form an orthonormal basis for V_0. The MRA structure is appropriately named; the scaling function property further implies that signals in every V_i look like combinations of dilated versions of V_0 elements.

11.4.2.1 Existence. In the following result, the equalities in (11.148a) and (11.148b) are assumed to hold almost everywhere.

Proposition (Spanning Translates). If $x(t), y(t) \in L^2(\mathbb{R})$, $X(\omega)$ and $Y(\omega)$ are their respective radial Fourier transforms, and $s(k) \in l^2$, then the following are equivalent:

$$y(t) = \sum_{k=-\infty}^{\infty} s(k)x(t-k), \tag{11.148a}$$

$$Y(\omega) = S(\omega)X(\omega), \tag{11.148b}$$

where $S(\omega)$ is the discrete-time Fourier transform of $s(k)$.

Proof: Now, assuming (11.148b), we have

$$y(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} S(\omega)X(\omega)e^{j\omega t}d\omega = \frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\sum_{k=-\infty}^{\infty} s(k)e^{-jk\omega}\right)X(\omega)e^{j\omega t}d\omega. \tag{11.149}$$

Hence,

$$y(t) = \sum_{k=-\infty}^{\infty} s(k)\frac{1}{2\pi}\int_{-\infty}^{\infty} X(\omega)e^{j\omega(t-k)}d\omega = \sum_{k=-\infty}^{\infty} s(k)x(t-k). \tag{11.150}$$

To show the converse, we work backwards through the equalities in (11.150) to the front of (11.149), a Fourier transform synthesis equation for $y(t)$. We must have (11.148b) except on a set of Lebesgue measure zero. ■

This section's main result comes from the classic source papers on wavelets and multiresolution analysis [23, 29].

Theorem (Scaling Function). If $\{V_i: i \in \mathbb{Z}\}$ is an MRA, then there is some $\phi(t) \in V_0$ such that $\{\phi(t-k): k \in \mathbb{Z}\}$ is an orthonormal basis of V_0.

Proof: By the Riesz Basis Characterization (Section 11.4.1.3), there is some $g(t) \in V_0$ such that $F = \{g(t-k) \mid k \in \mathbb{Z}\}$ is a Riesz basis for V_0. Let us say it has lower and upper bounds $\sqrt{A}$ and $\sqrt{B}$, respectively. The Riesz Translates Basis Characterization implies

$$A \leq \sum_{k=-\infty}^{\infty} \left|G(\omega + 2\pi k)\right|^2 \leq B \tag{11.151}$$

for almost all $\omega \in \mathbb{R}$, where $G(\omega)$ is the (radial) Fourier transform of $g(t)$. Note that the sum in (11.151) is the 2π-periodic function $P_G(\omega)$, defined in (11.132). The Riesz bounds on $P_G(\omega)$ allow us to define (almost everywhere) the $L^2(\mathbb{R})$ function

$$\Phi(\omega) = \frac{G(\omega)}{\sqrt{P_G(\omega)}}. \tag{11.152}$$

$\Phi(\omega)$ is the Fourier transform of $\phi(t) \in L^2(\mathbb{R})$, and our claim is that $\phi(t)$ works. The previous proposition implies $\phi(t) \in \overline{\text{Span}\{g(t-k)\}}$. Since V_0 is closed, $\phi(t) \in V_0$, and so $\phi(t-k) \in V_0$, by MRA property (v). Equation (11.152) works both ways, and we see that $F = \{g(t-k) \mid k \in \mathbb{Z}\}$—which is dense in V_0—is in the closure of $\{\phi(t-k\} \mid k \in \mathbb{Z}\}$. Thus, $\overline{\text{Span}\{\phi(t-k)\}} = V_0$. It remains to show that the $\phi(t-k)$ are orthonormal. We calculate

$$\sum_{k=-\infty}^{\infty} \left|\Phi(\omega + 2\pi k)\right|^2 = \sum_{k=-\infty}^{\infty} \left|\frac{G(\omega + 2\pi k)}{\sqrt{P_G(\omega + 2\pi k)}}\right|^2 = \sum_{k=-\infty}^{\infty} \frac{\left|G(\omega + 2\pi k)\right|^2}{P_G(\omega)} = 1. \tag{11.153}$$

By the Orthonormal Translates criterion (11.129), $\{\phi(t-k\} \mid k \in \mathbb{Z}\}$ is an orthonormal set.

Corollary. Let $\{V_i: i \in \mathbb{Z}\}$ be an MRA, and $\phi(t) \in V_0$ be given by the theorem. Then, $\{2^{i/2}\phi(2^i t - k): k \in \mathbb{Z}\}$ is an orthonormal basis of V_i.

Proof: By properties (iv) and (v) of the MRA, the scaled versions of an orthonormal basis for V_0 will constitute an orthonormal basis for V_i. ■

Definition (Scaling Function). Let $V = \{V_i: i \in \mathbb{Z}\}$ be an MRA and $\phi(t) \in V_0$ such that $\{\phi(t-k): k \in \mathbb{Z}\}$ is an orthonormal basis of V_0. Then $\phi(t)$, known from the theorem, is called a *scaling function* of the MRA.

Any translate $\phi(t-k)$ of a scaling function $\phi(t)$ is still a scaling function. The next corollary [9] characterizes scaling functions for an MRA.

Corollary (Uniqueness). Let $V = \{V_i : i \in \mathbb{Z}\}$ be an MRA and $\phi(t) \in V_0$ be the scaling function found in the proof. Then $\theta(t) \in V_0$ is a scaling function for V if and only if there is a 2π-periodic function $P(\omega)$ such that

(i) $\Theta(\omega) = P(\omega)\Phi(\omega)$;

(ii) $|P(\omega)| = 1$ almost everywhere on $[0, 2\pi]$.

Proof: Exercise. ■

11.4.2.2 Examples. Let us look at some examples of scaling functions for the three multiresolution analyses that we know.

Example (Step Functions). The scaling function for the Haar MRA, for which V_0 consists of constant functions on unit intervals $(n, n-1)$, is just the unit square pulse $\phi(t) = u(t) - u(t-1)$.

Example (Piecewise Continuous Functions). Finding this scaling function is not so direct. The continuity of V_0 elements forces us to reflect on how a possible scaling function $\phi(t)$ might be orthogonal to its translates $\phi(t-k)$. It is clear that $\phi(t)$ cannot be finitely supported. For then we could take the last interval $(n, n+1)$ to the right over which $\phi(t)$ is nonzero, the last interval $(m, m+1)$ proceeding to the left over which $\phi(t)$ is nonzero, and compute the inner product $\langle \phi(t), \phi(t-(n-m)) \rangle$. A simple check of cases shows that it is never zero. Evidently, $\phi(t) \neq 0$ on $(n, n+1)$ for arbitrarily large $|n|$, and the inner products $\langle \phi(t), \phi(t-k) \rangle$ involve an infinite number of terms.

But rather than stipulating from the start that V_0 must have a scaling function, we have elected to define our MRAs using the apparently weaker isomorphism condition (vi). The existence of this isomorphism $I: V_0 \to l^2$, which commutes with translations by integral amounts, is equivalent to V_0 having a Riesz basis. This facilitates our study of the Strömberg MRA. If we can find a Riesz basis $F = \{g(t-k) \mid k \in \mathbb{Z}\}$ for V_0, then the Scaling Function Theorem (11.152) readily gives the Fourier transform $\Phi(\omega)$ of $\phi(t)$. Let $g(t) \in V_0$ be the simple triangular pulse with $g(0) = 1$, $g(t) = t+1$ on $(-1, 0)$, $g(t) = 1-t$ on $(0, 1)$, and $g(t) = 0$ otherwise. Then $x(t) = \Sigma a_k g(t-k)$ is piecewise linear, continuous, and $x(k) = a_k$ for all $k \in \mathbb{Z}$. We can define the isomorphism I by $(Ix)(k) = a_k$. This map commutes with integer translations. The Riesz Basis Characterization implies that $\{g(t-k) \mid k \in \mathbb{Z}\}$ is a Riesz basis. In fact, $g(t) = I^{-1}(\delta(n))$, where I is the isomorphism from V_0 to l^2, and $\delta(n)$ is the discrete delta signal. We compute

$$G(\omega) = \int_{-\infty}^{\infty} g(t) e^{-j\omega t} dt = \left[\frac{\sin\left(\frac{\omega}{2}\right)}{\frac{\omega}{2}} \right]^2 = \operatorname{sinc}^2\left(\frac{\omega}{2}\right). \tag{11.154}$$

Define

$$\Phi(\omega) = \frac{G(\omega)}{\sqrt{P_G(\omega)}} = \frac{\operatorname{sinc}^2\left(\frac{\omega}{2}\right)}{\sqrt{\sum_{k=-\infty}^{\infty} \operatorname{sinc}^4\left(\frac{\omega + 2\pi k}{2}\right)}} = \frac{\operatorname{sinc}^2\left(\frac{\omega}{2}\right)}{4\sin^2(\omega)\sqrt{\sum_{k=-\infty}^{\infty} (\omega + 2\pi k)^{-4}}}. \tag{11.155}$$

We define the utility the function $\Sigma_n(\omega)$ as follows:

$$\Sigma_n(\omega) = \sum_{k=-\infty}^{\infty} (\omega + 2\pi k)^{-n}. \tag{11.156}$$

Though it tests our competence in differential calculus, it is possible [23] to develop closed form expressions for the $\Sigma_n(\omega)$, beginning with the standard summation [36]:

$$\Sigma_2(\omega) = \sum_{k=-\infty}^{\infty} (\omega + 2\pi k)^{-2} = \frac{1}{4}\sin^{-2}\left(\frac{\omega}{2}\right). \tag{11.157}$$

Twice differentiating (11.157) gives

$$\frac{d^2}{d\omega^2}\Sigma_2(\omega) = 6\sum_{k=-\infty}^{\infty} (\omega + 2\pi k)^{-4} = 6\Sigma_4(\omega) = \frac{1}{4}\cot^2\left(\frac{\omega}{2}\right)\csc^2\left(\frac{\omega}{2}\right) + \frac{1}{8}\csc^4\left(\frac{\omega}{2}\right). \tag{11.158}$$

Finally,

$$\Phi(\omega) = \frac{G(\omega)}{\sqrt{P_G(\omega)}} = \frac{\sqrt{6}}{\omega^2\sqrt{\frac{1}{4}\cot^2\left(\frac{\omega}{2}\right)\csc^2\left(\frac{\omega}{2}\right) + \frac{1}{8}\csc^4\left(\frac{\omega}{2}\right)}}. \tag{11.159}$$

Taking the inverse Fourier transform gives $\phi(t)$ (Figure 11.10). Notice from its magnitude spectrum that $\phi(t)$ is an analog low-pass filter.

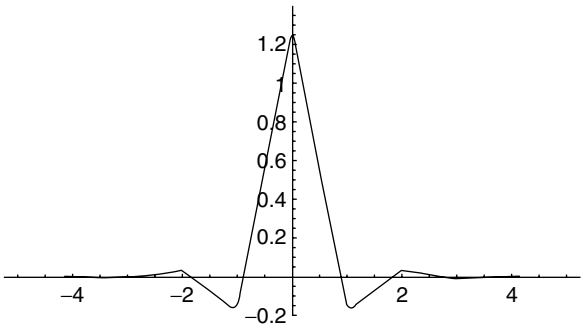


Fig. 11.10. Scaling function for the MRA consisting of continuous piecewise signals.

Example (Spline Functions). The procedure for finding this scaling function is similar to the one we use for the piecewise continuous MRA. The root space V_0 contains the $L^2(\mathbb{R})$ signals that are continuously differentiable and equal to cubic polynomials on each interval $[n, n+1]$, where $n \in \mathbb{Z}$ [37, 38]. To find a scaling function we need to find a Riesz basis. Let $g(t) \in V_0$ be the cubic spline that satisfies $g(0) = 1$ and $g(n) = 0$ otherwise. This is a rounded tent function. Then $x(t) = \Sigma a_k g(t-k)$ is a cubic spline on intervals $[n, n+1]$, continuously differentiable, and $x(k) = a_k$ for all $k \in \mathbb{Z}$. Once again we set $(Ix)(k) = a_k$ and invoke the Riesz Basis Characterization toward showing $\{g(t-k) \mid k \in \mathbb{Z}\}$ to be a Riesz basis. We compute the radial Fourier transform of $g(t)$:

$$G(\omega) = \left(1 - \frac{2}{3}\sin^2\left(\frac{\omega}{2}\right)\right)^{-1} \operatorname{sinc}^4\left(\frac{\omega}{2}\right). \tag{11.160}$$

We can derive a cubic spline scaling function by the standard formula (11.152):

$$\Phi(\omega) = \frac{G(\omega)}{\sqrt{P_G(\omega)}} = \frac{[\Sigma_8(\omega)]^{-\frac{1}{2}}}{\omega^4}, \tag{11.161}$$

where $\Sigma_n(\omega)$ is given by (11.156). Again, we can compute the $\Sigma_n(\omega)$ by taking successive derivatives—six actually—of $\Sigma_2(\omega)$. With either resolute patience or a symbolic computation software package, we calculate

$$\begin{aligned}\frac{d^6}{d\omega^6}\Sigma_2(\omega) = {} & \frac{1}{4}\cot^6\left(\frac{\omega}{2}\right)\csc^2\left(\frac{\omega}{2}\right) + \frac{57}{8}\cot^4\left(\frac{\omega}{2}\right)\csc^4\left(\frac{\omega}{2}\right) \\ & + \frac{45}{4}\cot^2\left(\frac{\omega}{2}\right)\csc^6\left(\frac{\omega}{2}\right) + \frac{17}{16}\csc^8\left(\frac{\omega}{2}\right).\end{aligned} \tag{11.162}$$

Consequently,

$$\Sigma_8(\omega) = \frac{1}{5040}\frac{d^6}{d\omega^6}(\Sigma_2(\omega)) \tag{11.163}$$

and

$$\Phi(\omega) = \frac{1}{\omega^4\sqrt{\Sigma_8(\omega)}}. \tag{11.164}$$

So once again, we find a scaling function using frequency-domain analysis. The key relationship is (11.152). Inverse Fourier transformation gives $\phi(t)$ (Figure 11.11). For the cubic spline MRA too, the scaling function is a low-pass filter. In comparison to the MRA for piecewise linear signals, observe the flatter reject band of $\Phi(\omega)$ for the spline MRA.

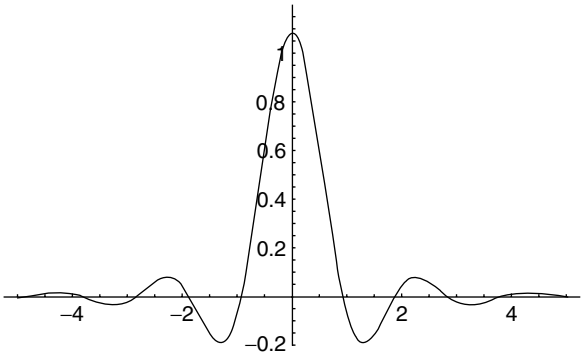


Fig. 11.11. Scaling function for the MRA consisting of cubic splines.

11.4.3 Discrete Low-Pass Filter

The scaling function is not the only special function connected with the multiresolution analysis structure. This section shows that for every MRA of finite energy signals we can find a special discrete filter [10, 23, 29]. This filter will prove useful when we discretize our theory and use it in signal analysis applications. In fact, we shall show that it is a low-pass filter. Mathematically thorough introductions to this material include [8, 9].

Suppose $V = \{V_i : i \in \mathbb{Z}\}$ is an MRA and $\phi(t) \in V_0$ is its scaling function. Because $\phi(t/2) \in V_{-1} \subset V_0$ and since integral translates of $\phi(t)$ span V_0, we see that

$$\frac{1}{2}\phi\left(\frac{t}{2}\right) = \sum_{n=-\infty}^{\infty} h_n \phi(t-n), \tag{11.165}$$

where the sequence $\{h_n \mid n \in \mathbb{Z}\}$ is square-summable. Hilbert space theory tells us that

$$h_n = \left\langle \frac{1}{2}\phi\left(\frac{t}{2}\right), \phi(t-n) \right\rangle. \tag{11.166}$$

These observations lead to the following definition.

Definition (Associated Filter). The $\phi(t)$ be the scaling function of an MRA,$V = \{V_i : i \in \mathbb{Z}\}$. If H_ϕ is the discrete filter with impulse response $h_\phi(n) = h_n$, where h_n is given by (11.166), then H_ϕ is called the *associated filter* to V (and to $\phi(t)$).

As we develop the properties of the discrete filter associated to an MRA, we shall see that it is indeed a low-pass filter. When there is no ambiguity, we drop the subscript: $H = H_\phi$. The following proposition gives a formula for the discrete-time Fourier transform $H(\omega)$ of the associated filter [23].

Proposition. Let $\phi(t)$ be the scaling function of an MRA, $V = \{V_i : i \in \mathbb{Z}\}$; let $\Phi(\omega)$ be its Fourier transform; and let H be the associated discrete filter with impulse response, $h(n) = h_n$, given by (11.166). Then,

$$\Phi(2\omega) = \Phi(\omega)H(\omega), \tag{11.167}$$

where $H(\omega)$ is the DTFT of $h(n)$: $H(\omega) = \sum_n h(n)e^{-j\omega n}$.

Proof: Apply the radial Fourier transform to both sides of (11.165). ■

Remark. The relation shows that $H(\omega)$ has a low-pass filter characteristic (Figure 11.12). The dilation $\Phi(2\omega)$ looks just like $\Phi(\omega)$, except that it is contracted with respect to the independent frequency-domain variable ω by a factor of two. The relation (11.167) shows that a multiplication by $H(\omega)$ accomplishes this, and the

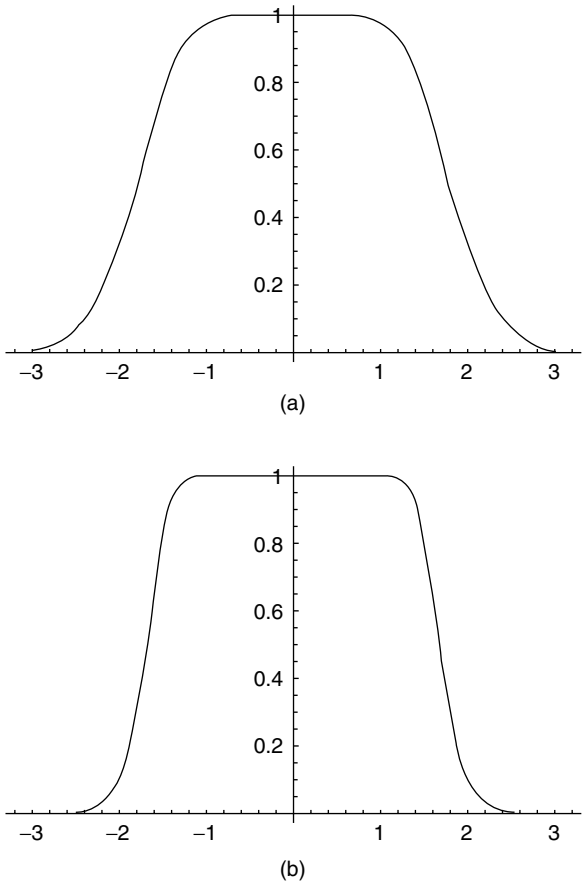


Fig. 11.12. Associated low-pass filters for the MRAs consisting of piecewise linear functions (a) and cubic splines (b). Note the better cutoff behavior of the filter based on cubic spline approximations.

only this can be the case, intuitively, is that $H(\omega)$ is approximately unity around the DC value $\omega = 0$, and it falls off to essentially zero halfway through the passband of the spectrum of $\Phi(\omega)$. This is not a rigorous argument, but it can be made so by assuming some decay constraints on the scaling function $\phi(t)$ [23, 29].

Proposition. Let $\phi(t)$ be the scaling function of an MRA, $V = \{V_i : i \in \mathbb{Z}\}$; $\Phi(\omega)$ its Fourier transform; $h(n)$ the associated discrete low-pass filter (11.166), and $H(\omega)$ its DTFT. Then,

$$|H(\omega)|^2 + |H(\omega + \pi)|^2 = 1, \tag{11.168}$$

for almost all $\omega \in \mathbb{R}$.

Proof: By the Orthonormal Translates Theorem, $\sum_{k=-\infty}^{\infty} |\Phi(2\omega + 2\pi k)|^2 = 1$. Inserting (11.167) into this identity gives

$$\sum_{k=-\infty}^{\infty} |\Phi(\omega + \pi k)|^2 |H(\omega + \pi k)|^2 = 1 \tag{11.169}$$

almost everywhere. Next, we split left-hand side of (11.169) into a sum over even integers and odd integers:

$$|H(\omega)|^2 \sum_{k=-\infty}^{\infty} |\Phi(\omega + 2\pi k)|^2 + |H(\omega + \pi)|^2 \sum_{k=-\infty}^{\infty} |\Phi(\omega + 2\pi k + \pi)|^2 = 1, \tag{11.170}$$

where we have used $H(\omega) = H(\omega + 2\pi)$. The Orthonormal Translates Theorem tells us that the infinite sums in (11.170) are unity, and (11.168) follows. ■

Remarks. Discrete filters satisfying (11.168) are *conjugate* filters, familiar from Chapter 9. Conjugate filters are used to filter a signal in such a way that it can be subsampled and exactly reconstructed later. The Laplacian pyramid technique provides decomposition by filtering and subsampling as well as exact reconstruction [24]. Various efficient signal compression and transmission techniques rely on this idea [25, 39, 40]. What is particularly striking about the MRA structure is that quite disparate signal theoretic techniques such as scale space analysis [26–28], orthogonal bases, and pyramid decompositions all meet together in one place.

The following results [8] link the low-pass filter H of an MRA, $V = \{V_i : i \in \mathbb{Z}\}$, with the low-resolution subspace $V_{-1} \subset V_0 \in V$.

Lemma. Let $\phi(t)$ be the scaling function of an MRA, let $\Phi(\omega)$ be its Fourier transform, and let $C(\omega) \in L^2[0, 2\pi]$ be 2π-periodic. Then $C(\omega)\Phi(\omega) \in L^2(\mathbb{R})$.

Proof: We compute

$$\int_{-\infty}^{\infty} |\Phi(\omega)|^2 |C(\omega)|^2 d\omega = \sum_{n=-\infty}^{\infty} \int_0^{2\pi} |\Phi(\omega + 2\pi n)|^2 |C(\omega + 2\pi n)|^2 d\omega$$
$$= \sum_{n=-\infty}^{\infty} \int_0^{2\pi} |\Phi(\omega + 2\pi n)|^2 |C(\omega)|^2 d\omega, \qquad (11.171)$$

where we have invoked $C(\omega) = C(\omega + 2\pi)$. Interchanging summation and integration in (11.171), and again using $\sum |\Phi(\omega + 2\pi n)|^2 = 1$, we get

$$\int_{-\infty}^{\infty} |\Phi(\omega)|^2 |C(\omega)|^2 d\omega = \int_0^{2\pi} \sum_{n=-\infty}^{\infty} |\Phi(\omega + 2\pi n)|^2 |C(\omega)|^2 d\omega$$
$$= \int_0^{2\pi} |C(\omega)|^2 d\omega = \|C\|^2_{2,\, L^2[0,\, 2\pi]}. \qquad (11.172)$$

■

Proposition (V_0 Characterization). Let $V = \{V_i : i \in \mathbb{Z}\}$ be an MRA, let $\phi(t)$ be its scaling function, and let $\Phi = \mathcal{F}(\phi)$ be its (radial) Fourier transform. Then the root space $V_0 \in V$ contains precisely those $x(t) \in L^2(\mathbb{R})$ such that $X(\omega) = C(\omega)\Phi(\omega)$ for some 2π-periodic $C(\omega) \in L^2[0, 2\pi]$, where $X(\omega) = \mathcal{F}(x)$.

Proof: Let $x(t) \in V_0$. Then $x(t) = \Sigma c_k \phi(t - k)$ for some $c(k) = c_k$, where $c \in l^2$. We compute $X(\omega)$ as follows:

$$\int_{-\infty}^{\infty} x(t) e^{-j\omega t} dt = \int_{-\infty}^{\infty} \sum_{k=-\infty}^{\infty} c_k \phi(t-k) e^{-j\omega t} dt = \sum_{k=-\infty}^{\infty} c_k \int_{-\infty}^{\infty} \phi(t-k) e^{-j\omega t} dt$$
$$= \sum_{k=-\infty}^{\infty} c_k e^{-j\omega k} \int_{-\infty}^{\infty} \phi(t) e^{-j\omega t} dt = C(\omega)\Phi(\omega). \qquad (11.173)$$

Now suppose $X(\omega) = C(\omega)\Phi(\omega)$ for some 2π-periodic $C(\omega) \in L^2[0, 2\pi]$. By the lemma, $X(\omega) \in L^2(\mathbb{R})$, and we can write $C(\omega) = \Sigma c_k e^{-j\omega k}$, where $c(k)$ is the inverse DTFT of $C(\omega)$. Whence the computation (11.173) shows that $x(t)$ is in the closure of the span of $\{\phi(t - k) \mid k \in \mathbb{Z}\}$; so $x \in V_0$. ■

Corollary. With the same notation, define the operator T: $V_0 \to L^2[0, 2\pi]$ by $Tx = C$, where $C(\omega)$ is the 2π-periodic function with $X(\omega) = C(\omega)\Phi(\omega)$ guaranteed by the proposition. Then:

(i) T is linear.
(ii) If $x \in V_0$, then $2\pi\|x\|_2^2 = \|C\|_{2, L^2[0, 2\pi]}^2$.
(iii) If $x, y \in V_0$, $C = Tx$, and $D = Ty$, then $2\pi\langle x, y\rangle_{L^2(\mathbb{R})} = \langle C, D\rangle_{L^2[0, 2\pi]}$.

Proof: Linearity (i) is left as an exercise. For (ii), let $x(t) \in V_0$, so that $X(\omega) = C(\omega)\Phi(\omega)$. Then, the Plancherel's formula for $L^2(\mathbb{R})$ and (11.172) entail

$$\|X\|_{2, L^2(\mathbb{R})}^2 = \int_{-\infty}^{\infty} |\Phi(\omega)|^2 |C(\omega)|^2 d\omega = \|C\|_{2, L^2[0, 2\pi]}^2 = 2\pi\|x\|_{2, L^2(\mathbb{R})}^2. \tag{11.174}$$

From (11.174) and the polarization identity [8, 30] for inner product spaces (Chapter 2), $4\langle x, y\rangle = \|x + y\|^2 - \|x - y\|^2 + j\|x + jy\|^2 - j\|x - jy\|^2$, (iii) follows. ■

Definition (Canonical Linear Map on V_0). The linear map $Tx = C$, where $C(\omega)$ is the 2π-periodic function with $X(\omega) = C(\omega)\Phi(\omega)$ guaranteed by the corollary, is called the *canonical map* from V_0 to $L^2[0, 2\pi]$.

Proposition (V_{-1} Characterization). Let $V = \{V_i : i \in \mathbb{Z}\}$ be an MRA, let $\phi(t)$ be its scaling function, $\Phi = \mathcal{F}(\phi)$, and let $H = H_\phi$ the associated low-pass filter. Then the first low resolution subspace $V_{-1} \in V$ contains precisely those $x(t) \in L^2(\mathbb{R})$ such that $X(\omega) = C(2\omega)H(\omega)\Phi(\omega)$ for some 2π-periodic $C(\omega) \in L^2[0, 2\pi]$, where $X(\omega) = \mathcal{F}(x)$.

Proof: Let $x(t) \in V_{-1}$, so that $2x(2t) \in V_0$. Then $2x(2t) = \Sigma c_k \phi(t - k)$ for some $c_k = c(k) \in l^2$. Thus,

$$2x(t) = \sum_{k=-\infty}^{\infty} c(k)\phi\left(\frac{t}{2} - k\right). \tag{11.175}$$

Taking Fourier transforms again [8]:

$$\begin{aligned} 2X(\omega) &= 2\sum_{k=-\infty}^{\infty} c_k \int_{-\infty}^{\infty} \phi(s) e^{-j\omega(2s + 2k)} ds \\ &= 2\sum_{k=-\infty}^{\infty} c_k e^{-2j\omega k} \int_{-\infty}^{\infty} \phi(s) e^{-j(2\omega)s} ds = 2C(2\omega)\Phi(2\omega), \end{aligned} \tag{11.176}$$

where we have made the substitution $s = t/2 - k$, and $C(\omega)$ is the DTFT of $c(k)$. From (11.167), $\Phi(2\omega) = H(\omega)\Phi(\omega)$; thus, $X(\omega) = C(2\omega)H(\omega)\Phi(\omega)$. For the converse, let $X(\omega) = C(2\omega)H(\omega)\Phi(\omega)$ for some 2π-periodic $C(\omega) \in L^2[0, 2\pi]$. Since $C(2\omega)H(\omega)$ is still 2π-periodic, the Lemma applies, and $X(\omega) \in L^2(\mathbb{R})$. Finally, (11.176) reverses to show $x(t) \in V_0$. ■

It is possible to generalize this result (exercise). The next section explains a mathematical surprise that arises from MRA theory.

11.4.4 Orthonormal Wavelet

Besides the scaling function and the associated discrete low-pass filter, a third special function accompanies any multiresolution approximation of $L^2(\mathbb{R})$: the *orthonormal wavelet* [23, 29]. Our presentation has been guided by the mathematically complete introductions [8, 9].

Definition (Orthonormal Wavelet). Let $\psi(t) \in L^2(\mathbb{R})$. If its dilations and translations $\{2^{n/2}\psi(2^n t - m)\colon m, n \in \mathbb{Z}\}$ are an orthogonal basis of $L^2(\mathbb{R})$, then ψ is an *orthogonal wavelet*. If $\|\psi\| = 1$, then ψ is an *orthonormal* wavelet.

At the beginning of this chapter we considered an extension of the Fourier transform based on scale and location as transform parameters. The transform inversion required a special signal, the *admissible wavelet*, in order to succeed, and we found that admissible wavelets had to be analog band-pass filters. Now, for MRAs it turns out that the special associated orthogonal wavelet too is a band-pass filter. To discover how it is that an MRA supports this extraordinary function, we examine the orthogonal complements of the component spaces V_i of the multiresolution analysis $V = \{V_i : i \in \mathbb{Z}\}$.

11.4.4.1 Existence. Consider first $V_{-1} \subset V_0$. From Hilbert space theory (Chapters 2 and 3), we know that every element of V_0 can be written uniquely as a sum $x = v + w$, where $v \in V_{-1}$ and $w \perp v$. The set of all such $w \in V_0$ constitute a Hilbert subspace of V_0; let us denote it by W_{-1}. We say that V_0 is the *direct sum* of V_{-1} and W_{-1}: $V_0 = V_{-1} \oplus W_{-1}$. In general, every V_{i+1} is the direct sum of V_i and W_i, where W_i is the *orthogonal complement* of V_i in V_{i+1}. We know already that the V_i have orthonormal bases made up of translations and dilations of the scaling function $\phi(t)$. We can also find an orthonormal basis of W_i by the Gram–Schmidt orthonormalization procedure, of course [31]. But this does not imply that the basis elements are translates of one another, and it exposes no relation between the basis elements so found and the rest of the MRA structure. We want a more enlightening theory.

Lemma (First W_{-1} Characterization). Let $V = \{V_i\colon i \in \mathbb{Z}\}$ be an MRA; let $\phi(t)$ be its scaling function; let $\Phi = \mathcal{F}(\phi)$ be the (radial) Fourier transform of $\phi(t)$; let $H = H_\phi$ be the associated low-pass filter; and let $Tx = C$ be the canonical linear map from V_0 to $L^2[0, 2\pi]$. Then $x(t) \in W_{-1} \subset V_0$ if and only if

$$C(\omega)\overline{H(\omega)} + C(\omega + \pi)\overline{H(\omega + \pi)} = 0 \tag{11.177}$$

for almost all $\omega \in [0, 2\pi]$.

Proof: Let $y(t) \in V_{-1}$ and let $Ty = D$. Then by the V_{-1} characterization, $Y(\omega) = A(2\omega)H(\omega)\Phi(\omega)$ for some 2π-periodic $A \in L^2[0, 2\pi]$. So $D(\omega) = A(2\omega)H(\omega)$ almost

everywhere on $[0, 2\pi]$. By the corollary to the V_0 characterization, $\langle x, y\rangle = \langle C, D\rangle$. Thus, $x(t) \in W_{-1}$ if and only if $\langle C(\omega), A(2\omega)H(\omega)\rangle = 0$, for almost all $\omega \in [0, 2\pi]$. Writing out the inner product integral [8], we see that this is further equivalent to

$$\int_0^{2\pi} C(\omega)\overline{A(2\omega)H(\omega)}\, d\omega = \int_0^{\pi} \overline{A(2\omega)}[C(\omega)\overline{H(\omega)} + C(\omega+\pi)\overline{H(\omega+\pi)}]\, d\omega = 0 \tag{11.178}$$

for almost all $\omega \in [0, 2\pi]$. Since $y(t)$ is any element of V_{-1}, the $A(2\omega)$ in the integrand on the right-hand side of (11.178) is an arbitrary π-periodic signal; evidently, $x(t) \in W_{-1}$ if and only if the π-periodic factor $C(\omega)\overline{H(\omega)} + C(\omega+\pi)\overline{H(\omega+\pi)} = 0$ almost everywhere on $[0, \pi]$. Finally, by the 2π-periodicity of $C(\omega)$ and $H(\omega)$, this same expression must be almost everywhere zero on $[0, 2\pi]$.

Lemma (Second W_{-1} Characterization). Let $V = \{V_i: i \in \mathbb{Z}\}$ be an MRA, let $\phi(t)$ be its scaling function, let $\Phi = \mathcal{F}(\phi)$ be the Fourier transform of ϕ, let $H = H_\phi$ be the associated low-pass filter, and let $Tx = C$ be the canonical map from V_0 to $L^2[0, 2\pi]$. Then, $x(t) \in W_{-1}$ if and only if $X(\omega) = e^{-j\omega}S(2\omega)\overline{H(\omega+\pi)}\Phi(\omega)$ for some 2π-periodic $S(\omega) \in L^2[0, 2\pi]$.

Proof: Resorting to some linear algebra tricks, we formulate the previous lemma's criterion as a determinant. Thus, $x(t) \in W_{-1}$ is equivalent to

$$\det\begin{bmatrix} \overline{H(\omega+\pi)} & C(\omega) \\ \overline{-H(\omega)} & C(\omega+\pi) \end{bmatrix} = 0 \tag{11.179}$$

for almost all $\omega \in [0, 2\pi]$. This means that the columns of the matrix (11.179) are linearly dependent—that is, they are proportional via a 2π-periodic function $R(\omega)$:

$$\begin{bmatrix} C(\omega) \\ C(\omega+\pi) \end{bmatrix} = R(\omega)\begin{bmatrix} \overline{H(\omega+\pi)} \\ \overline{-H(\omega)} \end{bmatrix} \tag{11.180}$$

for just as many $\omega \in [0, 2\pi]$. Now substitute $\omega + \pi$ for ω in (11.180) to see

$$\begin{bmatrix} C(\omega+\pi) \\ C(\omega) \end{bmatrix} = R(\omega+\pi)\begin{bmatrix} \overline{H(\omega)} \\ \overline{-H(\omega+\pi)} \end{bmatrix} \tag{11.181}$$

whence $C(\omega) = -R(\omega+\pi)\overline{H(\omega+\pi)}$. Further putting $\omega+\pi$ for ω in (11.181) gives $C(\omega) = R(\omega)\overline{H(\omega+\pi)}$. Evidently, $R(\omega) = -R(\omega+\pi)$ for almost all $\omega \in [0, 2\pi]$. Hence we have shown that $x(t) \in W_{-1}$ if and only if $X(\omega) = C(\omega)\Phi(\omega)$, where

$C(\omega) = R(\omega)\overline{H(\omega + \pi)}$ for some 2π-periodic $R(\omega) \in L^2[0, 2\pi]$ with $R(\omega) = -R(\omega + \pi)$. We define $S(\omega) = \exp(j\omega/2)R(\omega/2)$. Then $S(\omega + 2\pi) = S(\omega)$ almost everywhere, and

$$X(\omega) = C(\omega)\Phi(\omega) = R(\omega)\overline{H(\omega + \pi)}\Phi(\omega) = e^{-j\omega}S(2\omega)\overline{H(\omega + \pi)}\Phi(\omega). \tag{11.182}$$

■

Lemma ($\mathbf{W_0}$ Characterization). Let $V = \{V_i : i \in \mathbb{Z}\}$ be an MRA, $\phi(t)$ its scaling function, $\Phi = \mathcal{F}(\phi)$ the Fourier transform of ϕ, $H = H_\phi$ the associated low-pass filter, and $Tx = C$ the canonical map from V_0 to $L^2[0, 2\pi]$. Then, $x(t) \in W_0$ if and only if $X(2\omega) = e^{-j\omega}S(2\omega)\overline{H(\omega + \pi)}\,\Phi(\omega)$ for some 2π-periodic $S(\omega) \in L^2[0, 2\pi]$.

Proof: We note that $x(t) \in W_0$ if and only if $\langle x(t), v(t)\rangle = 0$ for all $v(t) \in V_0$, which is equivalent to $\langle x(t/2), v(t/2)\rangle = 0$. But any $f(t) \in V_{-1}$ is of the form $v(t/2)$, so this is also equivalent to $x(t/2) \perp V_{-1}$; in other words, $y(t) = x(t/2) \in W_{-1}$. The previous lemma says $Y(\omega) = e^{-j\omega}S(2\omega)\overline{H(\omega + \pi)}\,\Phi(\omega)$ for some 2π-periodic $S(\omega) \in L^2[0, 2\pi]$. But $Y(\omega) = 2X(2\omega)$. ■

The main result of this section—very probably the main result of this chapter, likely the main result of this book, and arguably the main result of Fourier analysis in the latter half of the twentieth century—is expressed in the following theorem and its corollary [23, 29].

Theorem (Orthonormal Basis of $\mathbf{W_0}$). Suppose $V = \{V_i : i \in \mathbb{Z}\}$ is an MRA, $\phi(t)$ its scaling function, $\Phi = \mathcal{F}(\phi)$ the Fourier transform of ϕ, and $H = H_\phi$ is the associated low-pass filter. If $\psi(t)$ is the finite energy signal whose Fourier transform is given by

$$\Psi(2\omega) = e^{-j\omega}\overline{H(\omega + \pi)}\,\Phi(\omega), \tag{11.183}$$

then $\{\psi(t - k) \mid k \in \mathbb{Z}\}$ is an orthonormal basis for W_0.

Proof: We know that the W_0 characterization lemma entails $\psi(t) \in W_0$. (In fact, $\psi(t)$ represents the very straightforward case where the lemma's $S(\omega) = 1$ almost everywhere on $[0, 2\pi]$.) Let $x(t) \in W_0$. Let us show that a linear combination of translates of $\psi(t)$ is arbitrarily close to $x(t)$. The previous lemma says that $X(2\omega) = e^{-j\omega}S(2\omega)\overline{H(\omega + \pi)}\,\Phi(\omega)$ for some 2π-periodic $S(\omega) \in L^2[0, 2\pi]$. Thus, $X(\omega) = S(\omega)\Psi(\omega)$, almost everywhere on $[0, 2\pi]$. But we know this condition already from the Spanning Translates Proposition of Section 11.4.2.1: It means that

$$x(t) = \sum_{n=-\infty}^{\infty} s(n)\psi(t - n), \tag{11.184}$$

where $s(n)$ is the inverse discrete-time Fourier transform of $S(\omega)$. Verily, the closure of $\{\psi(t - k) \mid k \in \mathbb{Z}\}$ is all of W_0.

What about orthonormality? If we attempt to apply the Orthonormal Translates criterion (11.129), then we have

$$\begin{aligned}\sum_{k=-\infty}^{\infty} |\Psi(\omega+2k\pi)|^2 &= \sum_{k=-\infty}^{\infty} \left|\Phi\left(\frac{\omega}{2}+k\pi\right)\right|^2 \left|H\left(\frac{\omega}{2}+(k+1)\pi\right)\right|^2 \\ &= \sum_{k=-\infty}^{\infty} \left|\Phi\left(\frac{\omega}{2}+2k\pi\right)\right|^2 \left|H\left(\frac{\omega}{2}+2k\pi+\pi\right)\right|^2 \\ &\quad + \sum_{k=-\infty}^{\infty} \left|\Phi\left(\frac{\omega}{2}+2k\pi+\pi\right)\right|^2 \left|H\left(\frac{\omega}{2}+2k\pi+2\pi\right)\right|^2 .\end{aligned} \tag{11.185}$$

By the orthonormality of $\{\phi(t-k)\}$, we have

$$\begin{aligned}\sum_{k=-\infty}^{\infty} |\Psi(\omega+2k\pi)|^2 &= \sum_{k=-\infty}^{\infty} \left|\Phi\left(\frac{\omega}{2}+k\pi\right)\right|^2 \left|H\left(\frac{\omega}{2}+(k+1)\pi\right)\right|^2 \\ &= \sum_{k=-\infty}^{\infty} \left|\Phi\left(\frac{\omega}{2}+2k\pi\right)\right|^2 \left|H\left(\frac{\omega}{2}+2k\pi+\pi\right)\right|^2 \\ &\quad + \sum_{k=-\infty}^{\infty} \left|\Phi\left(\frac{\omega}{2}+2k\pi+\pi\right)\right|^2 \left|H\left(\frac{\omega}{2}+2k\pi+2\pi\right)\right|^2 .\end{aligned} \tag{11.186}$$

By the 2π-periodicity of $H(\omega)$, this becomes

$$\begin{aligned}\sum_{k=-\infty}^{\infty} |\Psi(\omega+2k\pi)|^2 &= \left|H\left(\frac{\omega}{2}+\pi\right)\right|^2 \sum_{k=-\infty}^{\infty} \left|\Phi\left(\frac{\omega}{2}+2k\pi\right)\right|^2 \\ &\quad + \left|H\left(\frac{\omega}{2}\right)\right|^2 \sum_{k=-\infty}^{\infty} \left|\Phi\left(\frac{\omega}{2}+2k\pi+\pi\right)\right|^2\end{aligned} \tag{11.187}$$

and by Orthonormal Translates applied to the Φ summations on the bottom of (11.187), we get

$$\sum_{k=-\infty}^{\infty} |\Psi(\omega+2k\pi)|^2 = \left|H\left(\frac{\omega}{2}+\pi\right)\right|^2 + \left|H\left(\frac{\omega}{2}\right)\right|^2. \tag{11.188}$$

But this last expression is unity by (11.168), and thus $\{\psi(t-k)\,|\,k\in\mathbb{Z}\}$ is orthonormal.

Corollary (Existence of Orthonormal Wavelet). With the theorem's assumptions and notation, let $\psi(t) \in L^2(\mathbb{R})$ be defined as above. Then the dilations and translations $\{2^{n/2}\psi(2^n t - m)\colon m, n \in \mathbb{Z}\}$ are an orthonormal basis of $L^2(\mathbb{R})$, and so ψ is an orthonormal wavelet [8–10].

Proof: Since $V_1 = V_0 \oplus W_0$, dilations of $x(t) \in W_0$ by 2^i are in W_i: $x(2^i t) \in W_i$. In fact, $\{2^{i/2}\psi(2^i t - m)\colon m \in \mathbb{Z}\}$ is an orthonormal basis for W_i. But, $V_{i+1} = V_i \oplus W_i$, so

$B_{i+1} = \{2^{n/2}\psi(2^n t - m)\colon n < i+1 \text{ and } m \in \mathbb{Z}\}$ is an orthonormal set inside V_{i+1}. By the intersection property (iii) of the MRA, though, $\bigcap_{i=-\infty}^{\infty} V_i = 0$; B_{i+1} must be dense in V_{i+1}. By the MRA union property (ii), $\overline{\bigcup_{n=-\infty}^{\infty} V_i} = L^2(\mathbb{R})$; $L^2(\mathbb{R})$ must be the Hilbert space direct sum of the W_i:

$$\bigoplus_{i=-\infty}^{\infty} W_i = L^2(\mathbb{R}). \tag{11.189}$$

■

The wavelet $\psi(t)$ that we have found is essentially unique. The exercises outline an argument that any other W_0 function that is an orthogonal wavelet for square-integrable signals must have a Fourier transform that differs from the formula (11.183) by a factor that is unity almost everywhere on $[0, 2\pi]$.

11.4.4.2 Examples. Let us show some examples of orthonormal wavelets from the multiresolution analyses of square-integrable signals that we already know.

Example (Step Functions). In the Haar MRA [20], the root space V_0 consists of constant functions on unit intervals $(n, n-1)$, and so $\phi(t) = u(t) - u(t-1)$. We compute its Fourier transform by

$$\Phi(\omega) = \int_{-\infty}^{\infty} \phi(t) e^{-j\omega t} dt = e^{-\frac{j\omega}{2}} \frac{\sin(\omega/2)}{\omega/2} \tag{11.190}$$

and the relation $\Phi(2\omega) = \Phi(\omega)H(\omega)$ gives the associated low-pass filter:

$$H(\omega) = e^{-\frac{j\omega}{2}} \cos(\omega/2). \tag{11.191}$$

From $\Psi(\omega) = e^{-j\omega}\overline{H(\omega + \pi)}\Phi(\omega)$, we calculate

$$\Psi(\omega) = -je^{\frac{j\omega}{2}} \frac{\sin^2(\omega/4)}{\omega/4}. \tag{11.192}$$

But (11.192) is the radial Fourier transform of the function

$$\psi(t) = \begin{cases} -1 & \text{if } -1 \le t < \frac{1}{2}, \\ 1 & \text{if } -\frac{1}{2} \le t < 0, \\ 0 & \text{if otherwise.} \end{cases} \tag{11.193}$$

so $\psi(t)$ above is the orthogonal wavelet for the step function MRA.

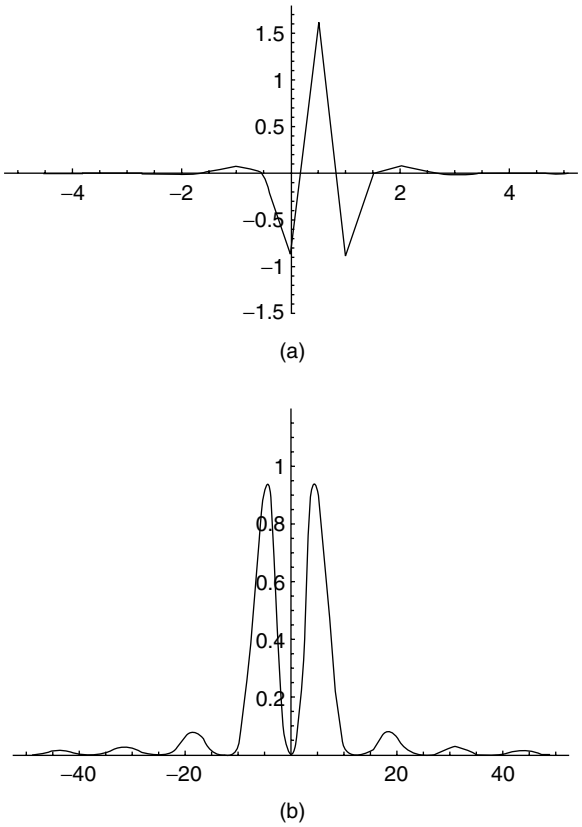


Fig. 11.13. For the Strömberg MRA: The orthogonal wavelet, (a) which we compute from a messy, but exact expression for its Fourier transform (b).

Example (Piecewise Continuous Functions). We found the scaling function for the Strömberg [22] MRA in Section 11.4.2.2 and the associated low-pass filter in Section 11.4.3. The expression $\Psi(2\omega) = e^{j\omega}\overline{H(\omega + \pi)}\,\Phi(\omega)$ as well as (11.159) and (11.167) give us the Fourier transform for the piecewise continuous MRA's wavelet. Then, we can compute $\psi(t)$ via the inverse transform (Figure 11.13).

Example (Spline Functions). In the third MRA we have studied, the root space V_0 contains continuously differentiable, finite-energy signals that are cubic polynomials on unit intervals $[n,\ n+1]$. This MRA was developed by Lemarié [37] and Battle [38]. The same strategy works once more. We know the scaling function's Fourier transform $\Phi(\omega)$ from (11.164). The discrete-time Fourier transform for the associated low-pass filter is the ratio $H(\omega) = \Phi(2\omega)/\Phi(\omega)$. Hence, we find $\Psi(\omega)$, and inverse transforming gives the orthogonal wavelet (Figure 11.14).

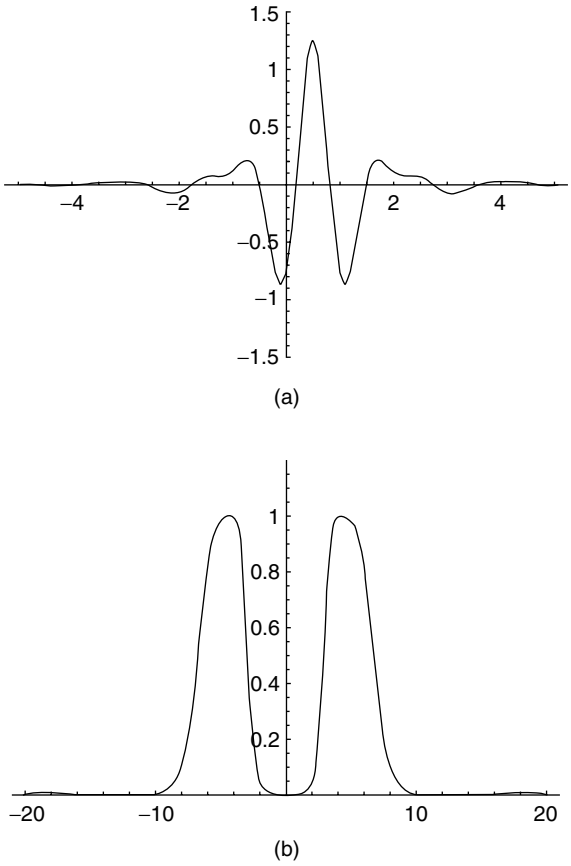


Fig. 11.14. For the cubic spline MRA of Lemarié and Battle: The orthogonal wavelet (a) and its Fourier transform (b).

The next chapter provides various examples of mixed domain signal analysis. In particular, it covers the use of the multiresolution analysis structure for combining time- and scale-domain analysis of signals that arise in typical applications.

11.5 SUMMARY

For signal analysis, both the continuous wavelet transform and the multiresolution analysis that leads to orthogonal wavelets are attractive and feasible. This chapter covered the signal analytic essentials of wavelet theory. The continuous transform finds more applications in signal understanding, as long as basic facts about input signals—location, scale, spectrum—are known beforehand. The orthogonal transforms, based on multiresolution analysis, tend to find more use in compression. We shall see in the last chapter that there are efficient algorithms for decomposition

using MRAs that lend themselves to efficient signal description, compression, as well as pattern recognition applications.

We omitted the proof of the theorem on the necessity of admissibility for frames based on translation and dilation (Section 11.3.2). The proof itself is somewhat technical, and, besides, in compensation, we later showed that an elegant construction, the MRA, leads to orthogonal wavelets from which efficient algorithms and straightforward frame constructions proceed.

There are three special functions that arise from an MRA, $V = \{V_i: i \in \mathbb{Z}\}$:

- The scaling function $\phi(t)$ whose translates form an orthonormal basis for the root space V_0;
- The associated low-pass filter H_ϕ;
- And, finally, the orthonormal wavelet $\psi(t)$ whose translations and dilations constitute and orthonormal basis for finite-energy signals.

Some conditions in the definition we give for an MRA are consequences of the others. For example, the third criterion in Section 11.4.1.1—that only the zero signal should appear in all V_i—follows from the rest. This and other interdependencies were only noted some years later [8, 41].

We provided only a few examples of MRAs, but these suffice for introductory signal analysis applications. Important extensions include compactly supported wavelets [3], approximation theory using wavelets [2], and multidimensional signal analysis using wavelets [10, 23].

Let us remark about how wavelets can be used in image processing and analysis. The principal focus of this book notwithstanding, we should also note that our MRA definition extends from one-dimensional functions (signals) to two-dimensional functions (images). Indeed there are MRA structures for the n-dimensional Hilbert space $L^2(\mathbb{R}^n)$. The technique is identical to that used to extend the Gaussian and Laplacian pyramid constructions to images [24].

We just repeat the definition of Section 11.4.1.1 for functions of two variables, $x(s, t) \in L^2(\mathbb{R}^2)$, denoting the subspace chain $V_{2,n}$. We can then derive a two-dimensional scaling function $\phi(s, t)$ as before (exercise). The most common approach, however, is to use a tensor product of one-dimensional MRAs to get an MRA for $L^2(\mathbb{R}^2)$: $V_{2,n} = V_{1,n} \otimes V_{1,n}$, where $\{V_n: n \in \mathbb{Z}\}$ is an MRA in $L^2(\mathbb{R})$. Then, there is a scaling function for the two-dimensional MRA of $L^2(\mathbb{R}^2)$: $\Phi(x, y) = \phi(x)\phi(y)$. While there is a single scaling function for the two-dimensional case, there are now three wavelets: $\Psi_1(x, y) = \phi(x)\psi(y)$, sensitive to *high vertical frequencies*; $\Psi_2(x, y) = \psi(x)\phi(y)$, *horizontal*; $\Psi_3(x, y) = \psi(x)\psi(y)$, *diagonal* or *corners.*

This gives an orientation selective decomposition, especially suitable for images with large x- and y-direction edge components. A number of applications involve such images: inspection of manufactured materials, remote sensing applications, and seismic signal processing among others.

For a history of wavelets, see Ref. [1]. References 10 and 42 list software resources and toolkits for signal analysis using wavelets.

REFERENCES

1. B. B. Hubbard, *The World According to Wavelets*, Wellesley, MA: A. K. Peters, 1996.
2. C. Chui, *Introduction to Wavelets*, San Diego, CA: Academic, 1992.
3. I. Daubechies, *Ten Lectures on Wavelets*, Philadelphia: SIAM, 1992.
4. Y. Meyer, *Wavelets: Algorithms and Applications*, Philadelphia: SIAM, 1993.
5. T. H. Koornwinder, ed., *Wavelets: An Elementary Treatment of Theory and Applications*, Singapore: World Scientific, 1993.
6. G. Kaiser, *A Friendly Guide to Wavelets*, Boston: Birkhauser, 1994.
7. M. Holschneider, *Wavelets: An Analysis Tool*, Oxford: Oxford University Press, 1995.
8. E. Hernandez and G. Weiss, *A First Course on Wavelets*, Boca Raton, FL: CRC Press, 1996.
9. P. Wojtaszczyk, *A Mathematical Introduction to Wavelets*, Cambridge: Cambridge University Press, 1997.
10. S. Mallat, *A Wavelet Tour of Signal Processing*, San Diego, CA: Academic, 1998.
11. H. L. Resnikoff and R. O. Wells, Jr., *Wavelet Analysis: The Scalable Structure of Information*, New York: Springer-Verlag, 1998.
12. C. Blatter, *Wavelets: A Primer*, Natick, MA: A. K. Peters, 1999.
13. A. Grossmann and J. Morlet, Decompositions of Hardy functions into square integrable wavelets of constant shape, *SIAM Journal of Mathematical Analysis*, vol. 15, pp. 723–736, 1984.
14. I. Daubechies, A. Grossmann, and Y. Meyer, Painless non-orthogonal expansions, *Journal of Mathematical Physics*, vol. 27, pp. 1271–1283, 1986.
15. I. Daubechies, The wavelet transform, time-frequency localization and signal analysis, *IEEE Transactions on Information Theory*, vol. 36, no. 5, pp. 961–1005, September 1990.
16. P. Goupillaud, A. Grossmann, and J. P. Morlet, Cycle-octave and related transforms in seismic signal analysis, *Geoexploration*, vol. 23, pp. 85–102, 1984.
17. R. Kronland-Martinet, J. P. Morlet, and A. Grossmann, Analysis of sound patterns through wavelet transforms, *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 1, no. 2, pp. 97–126, 1987.
18. A. Grossmann, Wavelet transforms and edge detection, in *Stochastic Processes in Physics and Engineering*, S. Albeverio et al., ed., Dordrecht, Holland: D. Reidel Publishing Company, pp. 149–157, 1988.
19. Y. Meyer, Principe d'incertitude, bases hilbertiennes et algèbres d'opérateurs, *Séminaire Bourbaki*, no. 662, 1985–1986.
20. A. Haar, Zur theorie der orthogonalen Functionensysteme, *Mathematische Annalen*, vol. 69, pp. 331–371, 1910.
21. C. E. Shannon, Communication in the presence of noise, *Proceedings of the Institute of Radio Engineers*, vol. 37, pp. 10–21, 1949.
22. J.-O. Strömberg, A modified Franklin system and higher order spline systems on $\mathbb{R}^n$ as unconditional bases for Hardy spaces, in *Proceedings of the Conference in Honor of Antoni Zygmund*, vol. II, W. Becker, A. P. Calderon, R. Fefferman, and P. W. Jones, eds., New York: Wadsworth, pp. 475–493, 1981.

23. S. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 674–693, July 1989.

24. P. J. Burt and E. H. Adelson, The Laplacian pyramid as a compact image code, *IEEE Transactions on Communications*, vol. COM-31, no. 4, pp. 532–540, April 1983.

25. R. E. Crochiere and L. R. Rabiner, *Multirate Digital Signal Processing*, Englewood Cliffs, NJ: Prentice-Hall, 1983.

26. A. P. Witkin, Scale-space filtering, *Proceedings of the 8th International Joint Conference on Artificial Intelligence*, Karlsruhe, W. Germany, 1983. Also in *From Pixels to Predicates*, A. P. Pentland, ed., Norwood, NJ: Ablex, 1986.

27. J. J. Koenderink, The structure of images, *Biological Cybernetics*, vol. 50, pp. 363–370, 1984.

28. T. Lindeberg, Scale space for discrete signals, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 3, pp. 234–254, March 1990.

29. S. G. Mallat, Multiresolution approximations and wavelet orthonormal bases of $L^2(\mathbb{R}^n)$, *Transactions of the American Mathematical Society*, vol. 315, no. 1, pp. 69–87, September 1989.

30. E. Kreyszig, *Introductory Functional Analysis with Applications*, New York: John Wiley & Sons, 1989.

31. J. Stoer and R. Bulirsch, *Introduction to Numerical Analysis*, 2nd ed., New York: Springer-Verlag, 1992.

32. A. Rosenfeld and M. Thurston, Edge and curve detection for visual scene analysis, *IEEE Transactions on Computers*, vol. 20, no. 5, pp. 562–569, May 1971.

33. S. Tanimoto and T. Pavlidis, A hierarchical data structure for picture processing, *Computer Graphics and Image Processing*, vol. 4, no. 2, pp. 104–119, June 1975.

34. R. Y. Wong and E. L. Hall, Sequential hierarchical scene matching, *IEEE Transactions on Computers*, vol. C-27, no. 4, pp. 359–366, April 1978.

35. D. Marr, *Vision*, New York: W. H. Freeman and Company, 1982.

36. I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, Orlando, FL: Academic, 1980.

37. P.-G. Lemarié, Ondelettes à localisation exponentielle, *Journal de Mathématiques Pures et Appliquées*, vol. 67, pp. 227–236, 1988.

38. G. Battle, A block spin construction of ondelettes. Part 1: Lemarié functions, *Communications in Mathematical Physics*, vol. 110, pp. 601–615, 1987.

39. P. P. Vaidyanathan, Multirate digital filters, filter banks, polyphase networks, and applications: A tutorial, *Proceedings of the IEEE*, vol. 78, pp. 56–93, January 1990.

40. O. Rioul and M. Vetterli, Wavelets and signal processing, *IEEE SP Magazine*, pp. 14–38, October 1991.

41. W. R. Madych, Some elementary properties of multiresolution analyses of $L^2(\mathbb{R}^n)$, in *Wavelets: A Tutorial in Theory and Applications*, C. K. Chui, ed., Boston: Academic, pp. 259–294, 1992.

42. A. Bruce, D. Donoho, and H. Y. Gao, Wavelet analysis, *IEEE Spectrum*, pp. 26–35, October 1996.

PROBLEMS

1. Let $\psi(t)$ be an analyzing wavelet.

(a) Show that $\Psi(0) = 0$.

(b) Also, $\int_{-\infty}^{\infty} \psi(t)\, dt = 0$.

(c) Show that $\psi(t)$ is a bandpass filter.

2. Provide an example of a bandpass filter that is not an analyzing wavelet.

3. Let $V = \{V_i : i \in \mathbb{Z}\}$ be a multiresolution analysis of $L^2(\mathbb{R})$ signals; let $\phi(t)$ be the scaling function for V; let $\Phi(\omega)$ be the (radial) Fourier transform of $\phi(t)$; and let H be the associated discrete filter with impulse response, $h(n) = h_n$, given by (11.166). Show in detail that $\Phi(2\omega) = \Phi(\omega)H(\omega)$, where $H(\omega)$ is the DTFT of $h(n)$.

4. Let $\alpha, \beta \in \mathbb{C}$ and $\psi(t), \phi(t)$ be (analyzing) wavelets. If $\theta(t) = \alpha\psi(t) + \beta\phi(t)$, then show

$$W_\theta[f(t)](a, b) = \bar{\alpha} W_\psi[f(t)](a, b) + \bar{\beta} W_\phi[f(t)](a, b). \tag{11.194}$$

5. Let $\eta > 0$, let $\psi(t)$ be a wavelet, and let $\theta(t) = \frac{1}{\eta}\psi\left(\frac{t}{\eta}\right)$. Show that

$$W_\theta[f(t)](a, b) = \frac{1}{\sqrt{\eta}} W_\psi[f(t)](a\eta, b). \tag{11.195}$$

6. Let $\alpha, \beta \in \mathbb{C}$ and define $\theta(t) = \alpha\psi(t) + \beta\phi(t)$. Show that

$$W_\theta[\alpha f(t) + \beta g(t)](a, b) = \alpha W_\psi[f(t)](a, b) + \beta W_\phi[f(t)](a, b) \tag{11.196}$$

7. If $\gamma \in \mathbb{R}$, then show

$$W[f(t - \gamma)](a, b) = W[f(t)](a, b - \gamma). \tag{11.197}$$

8. Let $m, n \in \mathbb{Z}$; $a_0, b_0 > 0$; and suppose that

$$\psi_{m,n}(t) = a_0^{-\frac{m}{2}} \psi\left(\frac{t - n b_0 a_0^m}{a_0^m}\right) = a_0^{-\frac{m}{2}} \psi(a_0^{-m} t - n b_0). \tag{11.198}$$

(a) Find the radial Fourier transform $\Psi_{m,n}(\omega)$ of $\psi_{m,n}(t)$.

(b) Defining $\Psi^-(\omega) = \Psi^+(-\omega)$, where $\Psi^+(\omega)$ is given by (11.108), and using the arguments of Section 11.3, show that

$$\sum_{m,\,n=-\infty}^{\infty} \left|\langle x, \bar{\psi}_{m,n}\rangle\right|^2 = \frac{1}{2\pi b_0 \ln a_0} \sum_m \int_{-\infty}^{0} |X(\omega)|^2 d\omega, \tag{11.199}$$

where $\bar{\psi}(t)$ is the inverse (radial) Fourier transform of $\bar{\Psi}(\omega)$.

9. Let H be a Hilbert space.

(a) Show that an orthonormal basis in H is also a Riesz basis with unit bounds.

(b) Show that a Riesz basis is linearly independent.

(c) Give an example of a frame in $L^2(\mathbb{R})$ that is not a Riesz basis.

10. Let $V = \{V_i\}$ be the Haar MRA which we defined in Section 11.4.1.1. Signals in V_i are constant on intervals $(n2^{-i}, (n+1)2^{-i})$, where $n \in \mathbb{Z}$. If $x(t) \in V_0$, and $x(t) = c_n[u(t-n) - u(t-n-1)]$ for $t \in [n, n+1)$, then we set $I(x(t)) = s(n)$, where $s(n) = c_n$ for all $n \in \mathbb{Z}$.

(a) Show that I is an isomorphism, a bounded linear operator that is one-to-one and onto.

(b) Show that if $I(x) = s(n)$ and $k \in \mathbb{Z}$, then $I(x(t-k)) = s(n-k)$.

(c) Is I an isometry? Explain.

(d) Are the elements of V_0 compactly supported? Explain.

11. Let W_0 be the continuous $L^2(\mathbb{R})$ signals that are piecewise linear on $[n, n+1]$, $n \in \mathbb{Z}$. Define $x(t) \in W_i \Leftrightarrow x(2t) \in W_{i+1}$.

(a) Verify MRA properties (i), (iii)–(v) for $W = \{W_i\}$.

(b) Let $V = \{V_i\}$ be the Haar MRA of the previous problem. Assuming that step functions are dense in $L^2(\mathbb{R})$, argue likewise for W by showing that given $v(t) \in V_0$, then some $w(t) \in W$ is arbitrarily close to $v(t)$. *Moral*: An approximation of an approximation is still an approximation.

(c) Let $w \in W_0$ and set $Iw = s$, where $s(n) = w(n)$. Show that I is an isomorphism.

12. Let $l^\infty(\mathbb{Z})$ be the normed linear space of bounded sequences, with $\|x\| = \sup\{|x(n)|: n \in \mathbb{Z}\}$. Define an operator $Tx = y$ as follows: $y(n) = x(n)(|n|+1)^{-1}$. Show that

(a) T is linear.

(b) T is one-to-one.

(c) T is onto.

(d) T is bounded.

(e) T^{-1} is not bounded [30].

13. Let I: $H \to K$ be a bounded Hilbert space isomorphism and let I^* be the Hilbert space adjoint of I.

(a) Show that $\|I^*\| = \|I\|$ [30].

(b) Supposing that I is onto, show that I^* is an isomorphism.

(c) Show that if $E = \{e_k \mid k \in \mathbb{Z}\}$ is a Riesz basis in H, then there is $F = \{f_k \mid k \in \mathbb{Z}\}$ such that F is a Riesz basis and [9]

$$\langle e_m, f_n \rangle = \begin{cases} 1 & \text{if } m = n, \\ 0 & \text{if } m \neq n. \end{cases} \tag{11.200}$$

14. Let $\phi(t) \in L^2(\mathbb{R})$, let $\Phi(\omega) = \mathcal{F}[\phi(t)](\omega)$ be its radial Fourier transform, let $s \in l^2$, $S(\omega)$ be its discrete-time Fourier transform, and set $P(\omega) = \sum_{n=-\infty}^{\infty} |\Phi(\omega + 2\pi n)|^2$. Show that

(a) $\displaystyle\sum_{k=-\infty}^{\infty} s(k)\Phi(t-k) \in L^2(\mathbb{R})$;

(b) $\displaystyle\left\| \sum_{k=-\infty}^{\infty} s(k)\phi(t-k) \right\|_2^2 = \frac{1}{2\pi} \int_0^{2\pi} |S(\omega)|^2 P(\omega)\, d\omega.$

15. This problem uses the concept of Lebesgue measure. Let $P(\omega)$ be defined as above. Define $P_a = \{\omega \in [0, 2\pi] : P(\omega) > a\}$ and $Q_a = \{\omega \in [0, 2\pi]: P(\omega) < a\}$. Referring to Ref. 9, show that:

(a) If the Lebesgue measure of P_a, $\mu(P_a)$, is zero for almost all $a \in \mathbb{R}$, then $P(\omega) = 0$ almost everywhere. *Hint*: Suppose not, so that $P(\omega) > 0$ on $U \subseteq \mathbb{R}$. Then $U = \cup P_{1/n}$, where $n > 0$ is a natural number. What is $\mu(P_{1/n})$? Apply the Lebesgue measure to the countable union.

(b) Similarly, if $\mu(Q_a) = 0$ for almost all $a \in \mathbb{R}$, then $P(\omega) = \infty$ almost everywhere.

16. Complete the proof of the second Riesz basis characterization theorem [9]. Suppose $\phi(t) \in L^2(\mathbb{R})$, $\Phi(\omega) = \mathcal{F}[\phi(t)](\omega)$ is its Fourier transform, $F = \{\phi(t-k) \mid k \in \mathbb{Z}\}$, $0 < A \leq B < \infty$, $P(\omega) = \Sigma|\Phi(\omega + 2\pi n)|^2$, and F is a Riesz basis with lower and upper bounds $\sqrt{A}$ and $\sqrt{B}$, respectively. Show that $P(\omega) \leq B$ for almost all $\omega \in \mathbb{R}$.

(a) Define $P_a = \{\omega \in [0, 2\pi]: P(\omega) > a\}$. Show that if $\mu(P_a) = 0$ almost everywhere, then $P(\omega) = 0$ for almost all $\omega \in [0, 2\pi]$, and $P(\omega) \leq B$. *Hint*: Consider the case $a = 1/(n+1)$ for $n \in \mathbb{N}$.

(b) Hence, assume that for some $a > 0$, $\mu(P_a) > 0$. Let χ_a be the characteristic function on P_a and $x_a(n)$ be the inverse discrete-time Fourier transform of $\chi_a(\omega)$. Show $\|x_a\|^2 = (2\pi)^{-1}\mu(P_a)$.

(c) Show $\|\Sigma x_a(k)\phi(t-k)\|^2 \geq a\|x_a\|^2$.

(d) Show $B \geq a$.

(e) Conclude that, unless $B \geq \Sigma|\Phi(\omega + 2\pi k)|^2$ almost everywhere on $[0, 2\pi]$, a contradiction arises.

17. This problem uses Lebesgue measure, but is straightforward. Suppose $\phi(t) \in L^2(\mathbb{R})$, $\Phi(\omega) = \mathcal{F}[\phi(t)](\omega)$ is its (radial) Fourier transform, and $F = \{\phi(t-k) \mid k \in \mathbb{Z}\}$ is an orthonormal set.

(a) Show that $\|\Phi\|_2 = (2\pi)^{1/2}$.

(b) Show that $|\Phi(\omega)| \leq 1$ for almost all $\omega \in \mathbb{R}$.

18. This problem studies Strömberg MRA [9, 22], wherein V_0 consists of continuous signals that are linear on integral segments $[n, n+1]$.

(a) Let $\phi(t)$ be a scaling function for $V = \{V_i\}$. Prove that $\phi(t)$ cannot be finitely supported. Assume that it is finitely supported and derive a contradiction as follows. Take the last interval $(n, n+1)$ to the right over which $\phi(t)$ is nonzero, take the last interval $(m, m+1)$ proceeding to the left over which $\phi(t)$ is nonzero, and compute the inner product $\langle \phi(t), \phi(t-(n-m)) \rangle$.

(b) Enumerate all of the cases for the inner product, and show that it is never zero.

(c) Show that $\phi(t) \neq 0$ on $(n, n+1)$ for arbitrarily large $|n|$.

(d) Conclude that the inner products $\langle \phi(t), \phi(t-k) \rangle$ involve an infinite number of terms.

19. Show that the scaling function for an MRA is essentially unique. More precisely, let $V = \{V_i : i \in \mathbb{Z}\}$ be an MRA and let $\phi(t) \in V_0$ be its scaling function. Show that $\theta(t) \in V_0$ is a scaling function for V if and only if there is a 2π-periodic function $P(\omega)$ such that $\Theta(\omega) = P(\omega)\Phi(\omega)$ and $|P(\omega)| = 1$ almost everywhere on $[0, 2\pi]$.

20. Let $V = \{V_i : i \in \mathbb{Z}\}$ be an MRA, $\phi(t)$ its scaling function, $\Phi = \mathcal{F}(\phi)$, and $H = H_\phi$ the associated low-pass filter. State and prove a generalization of the V_{-1} Characterization of Section 11.4.3 for any $V_N \subset V_0$, where $N < 0$.

21. With the same notation as in the previous problem, define the operator $T: V_0 \to L^2[0, 2\pi]$ by $Tx = C$, where $C(\omega)$ is the 2π-periodic function with $X(\omega) = C(\omega)\Phi(\omega)$ guaranteed by the V_0 characterization (Section 11.4.3).

(a) Show that T is linear: $T(x + y) = Tx + Ty$ and $T(ax) = aTx$ for $a \in \mathbb{C}$.

(b) If $c(n) \in l^2$ is the inverse DTFT of $C(\omega) = Tx$, show that $\|c\| = \|x\|$.

(c) Show that T is a bounded linear map.

22. This problem uses Lebesgue measure to show that the orthogonal wavelet for an MRA is essentially unique [8]. Suppose $V = \{V_i : i \in \mathbb{Z}\}$ is an MRA, $\phi(t)$ is its scaling function, $\Phi = \mathcal{F}(\phi)$ is the Fourier transform of ϕ, and $H = H_\phi$ is the associated low-pass filter. Let $\psi(t) \in W_0$ be an orthogonal wavelet for $L^2(\mathbb{R})$.

(a) Show that there is a 2π-periodic $v(\omega) \in L^2[0, 2\pi]$ such that

$$\Psi(2\omega) = v(2\omega)e^{-j\omega}\overline{H(\omega + \pi)}\,\Phi(\omega). \tag{11.201}$$

(b) Show that

$$\sum_{k=-\infty}^{\infty} |\Psi(\omega + 2\pi k)|^2 = 1 = |v(\omega)|^2 \sum_{k=-\infty}^{\infty} \left|H\left(\frac{\omega}{2} + \pi k + \pi\right)\right|^2 \left|\Phi\left(\frac{\omega}{2} + \pi k\right)\right|^2. \tag{11.202}$$

(c) Summing over even and odd integers separately, show that the final expression above becomes

$$|v(\omega)|^2 \left(\sum_{k=-\infty}^{\infty} \left|H\left(\frac{\omega}{2} + \pi\right)\right|^2 \left|\Phi\left(\frac{\omega}{2} + 2\pi k\right)\right|^2 + \sum_{k=-\infty}^{\infty} \left|H\left(\frac{\omega}{2}\right)\right|^2 \left|\Phi\left(\frac{\omega}{2} + 2\pi k + \pi\right)\right|^2 \right). \tag{11.203}$$

(d) Prove that

$$1 = |v(\omega)|^2 \left(\left|H\left(\frac{\omega}{2} + \pi\right)\right|^2 + \left|H\left(\frac{\omega}{2}\right)\right|^2 \right) = |v(\omega)|^2 \tag{11.204}$$

for almost all $\omega \in [0, 2\pi]$.

(e) Conclude that the Fourier transform of $\psi(t)$ has the form

$$\Psi(2\omega) = v(2\omega) e^{-j\omega} \overline{H(\omega + \pi)} \Phi(\omega), \tag{11.205}$$

where $v(\omega)$ is measurable, and has period 2π, and $v(\omega) = 1$ almost everywhere on $[0, 2\pi]$.

23. Show that to get an orthonormal basis for W_0 an alternative definition for the Fourier transform of $\psi(t)$ is

$$\Psi(2\omega) = e^{j\omega} \overline{H(\omega + \pi)} \Phi(\omega). \tag{11.206}$$

Show that with this change of the exponent's sign $\{\psi(t - k) \mid k \in \mathbb{Z}\}$ is still an orthonormal basis for W_0.

The following problems involve some extension of concepts in the text, may require some exploration of the research literature, and are generally more difficult than the preceding exercises.

24. Expand the construction of tight wavelet frames in Section 11.3.3 to include the case $0 < a_0 < 1$. Show that (11.109) continues to hold.

25. Investigate the application of frames of translations and dilations as constructed in Section 11.3.3. Assume that $a_0 = 2$ and $b_0 = 1$.

(a) Using a mathematical software package such as Mathematica or Matlab, or by developing your own Fourier transform software in a high-level programming language, find the inverse Fourier transforms for $\Psi^+(\omega)$ and $\Psi^-(\omega)$.

(b) As Daubechies remarks [3], this frame is not well-localized, and this becomes a problem for certain signal analysis tasks. By experiments of your own design, justify this claim.

(c) Continue your critique of this frame by considering the fact that it consists of translations and dilations of two distinct root elements, $\psi^+(t)$ and $\psi^-(t)$. In particular, explore the consequences of the definition $\Psi^-(\omega) = \Psi^+(-\omega)$. What difficulties does this impose on signal analysis applications? Develop experiments that justify your contention.

(d) Develop experiments using translations and dilations of the Mexican hat wavelet and compare the performance to the frame in part (c) based on $\psi^+(t)$ and $\psi^-(t)$.

26. Extend the idea of a multiresolution analysis to $L^2(\mathbb{R}^2)$.

(a) Reformulate the definition of Section 11.4.1.1 for functions of two variables, $x(s, t) \in L^2(\mathbb{R}^2)$.

(b) Let $\{V_{2,k}: k \in \mathbb{Z}\}$ be an MRA for $L^2(\mathbb{R}^2)$. Show that there is a unique image $\phi(s, t)$ such that $\{2^k\phi(2^k s - n, 2^k t - m): m, n \in \mathbb{Z}\}$ constitutes an orthonormal basis for $V_{2,k}$.