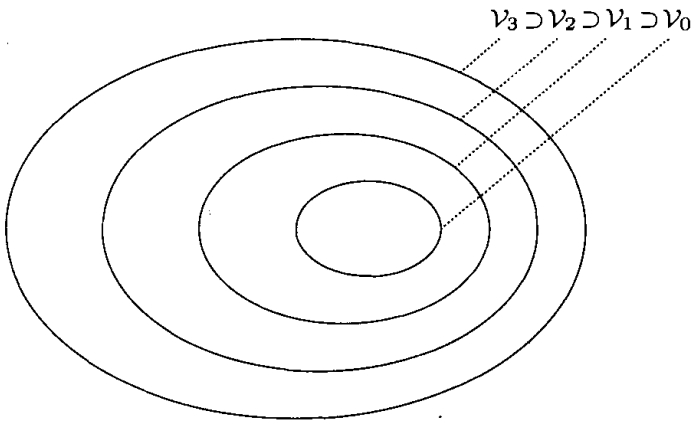


Figure 2.1. Nested Vector Spaces Spanned by the Scaling Functions

where the coefficients $h(n)$ are a sequence of real or perhaps complex numbers called the scaling function coefficients (or the scaling filter or the scaling vector) and the $\sqrt{2}$ maintains the norm of the scaling function with the scale of two.

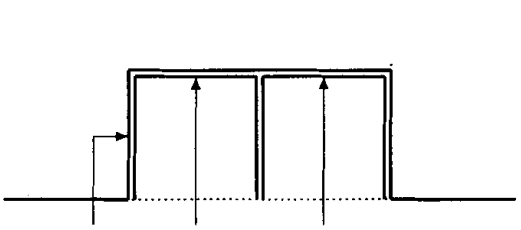
This recursive equation is fundamental to the theory of the scaling functions and is, in some ways, analogous to a differential equation with coefficients $h(n)$ and solution $\varphi(t)$ that may or may not exist or be unique. The equation is referred to by different names to describe different interpretations or points of view. It is called the refinement equation, the multiresolution analysis (MRA) equation, or the dilation equation.

The Haar scaling function is the simple unit-width, unit-height pulse function $\varphi(t)$ shown in Figure 2.2, and it is obvious that $\varphi(2t)$ can be used to construct $\varphi(t)$ by

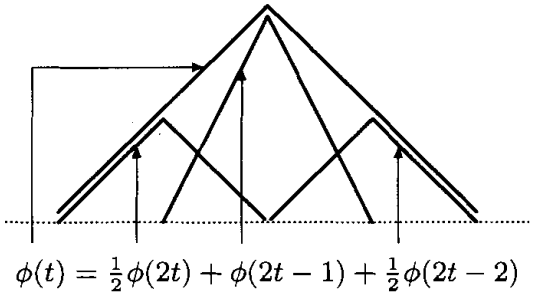
$$\varphi(t) = \varphi(2t) + \varphi(2t-1) \tag{2.14}$$

which means (2.13) is satisfied for coefficients $h(0) = 1/\sqrt{2}$, $h(1) = 1/\sqrt{2}$.

The triangle scaling function (also a first order spline) in Figure 2.2 satisfies (2.13) for $h(0) = \frac{1}{2\sqrt{2}}$, $h(1) = \frac{1}{\sqrt{2}}$, $h(2) = \frac{1}{2\sqrt{2}}$, and the Daubechies scaling function shown in the first part of



(a) Haar (same as φ_{D2})

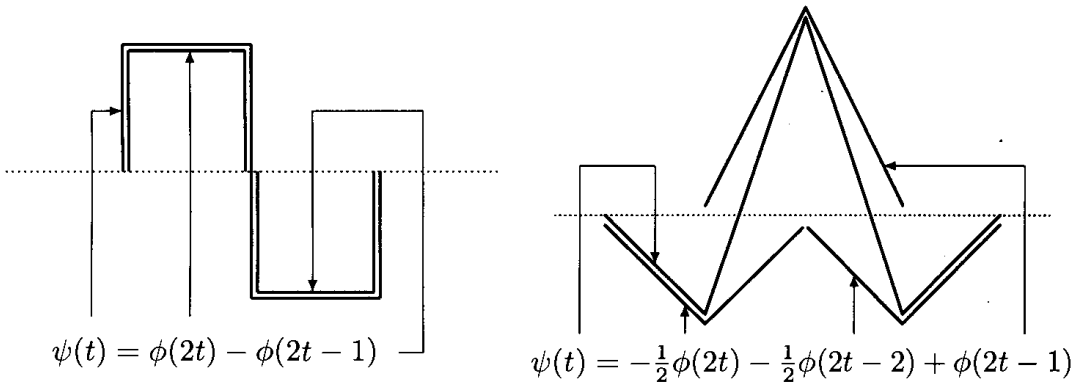


(b) Triangle (same as φ_{S1})

Figure 2.2. Haar and Triangle Scaling Functions

where 2^j is the scaling of t (j is the $\log_2$ of the scale), $2^{-j}k$ is the translation in t, and $2^{j/2}$ maintains the (perhaps unity) L^2 norm of the wavelet at different scales.

The Haar and triangle wavelets that are associated with the scaling functions in Figure 2.2 are shown in Figure 2.4. For the Haar wavelet, the coefficients in (2.24) are $h_1(0) = 1/\sqrt{2}$, $h_1(1) = -1/\sqrt{2}$ which satisfy (2.25). The Daubechies wavelets associated with the scaling functions in Figure 6.1 are shown in Figure 6.2 with corresponding coefficients given later in the book in Tables 6.1 and 6.2.



(a) Haar (same as ψ_{D2}) (b) Triangle (same as ψ_{S1})

Figure 2.4. Haar and Triangle Wavelets

We have now constructed a set of functions $\varphi_k(t)$ and $\psi_{j,k}(t)$ that could span all of $L^2(\mathbf{R})$. According to (2.19), any function $g(t) \in L^2(\mathbf{R})$ could be written

$$g(t) = \sum_{k=-\infty}^{\infty} c(k)\,\varphi_k(t) + \sum_{j=0}^{\infty}\sum_{k=-\infty}^{\infty} d(j,k)\,\psi_{j,k}(t) \tag{2.28}$$

as a series expansion in terms of the scaling function and wavelets.

In this expansion, the first summation in (2.28) gives a function that is a low resolution or coarse approximation of $g(t)$. For each increasing index j in the second summation, a higher or finer resolution function is added, which adds increasing detail. This is somewhat analogous to a Fourier series where the higher frequency terms contain the detail of the signal.

Later in this book, we will develop the property of having these expansion functions form an orthonormal basis or a tight frame, which allows the coefficients to be calculated by inner products as

$$c(k) = c_0(k) = \langle g(t), \varphi_k(t)\rangle = \int g(t)\,\varphi_k(t)\,dt \tag{2.29}$$

and

$$d_j(k) = d(j,k) = \langle g(t), \psi_{j,k}(t)\rangle = \int g(t)\,\psi_{j,k}(t)\,dt. \tag{2.30}$$

The coefficient $d(j,k)$ is sometimes written as $d_j(k)$ to emphasize the difference between the time translation index k and the scale parameter j. The coefficient $c(k)$ is also sometimes written as $c_j(k)$ or $c(j,k)$ if a more general "starting scale" other than $j = 0$ for the lower limit on the sum in (2.28) is used.

2.7 Examples of Wavelet Expansions

In this section, we will try to show the way a wavelet expansion decomposes a signal and what the components look like at different scales. These expansions use what is called a length-8 Daubechies basic wavelet (developed in Chapter 6), but that is not the main point here. The local nature of the wavelet decomposition is the topic of this section.

These examples are rather standard ones, some taken from David Donoho's papers and web page. The first is a decomposition of a piecewise linear function to show how edges and constants are handled. A characteristic of Daubechies systems is that low order polynomials are completely contained in the scaling function spaces $\mathcal{V}_j$ and need no wavelets. This means that when a section of a signal is a section of a polynomial (such as a straight line), there are no wavelet expansion coefficients $d_j(k)$, but when the calculation of the expansion coefficients overlaps an edge, there is a wavelet component. This is illustrated well in Figure 2.6 where the high resolution scales gives a very accurate location of the edges and this spreads out over k at the lower scales. This gives a hint of how the DWT could be used for edge detection and how the large number of small or zero expansion coefficients could be used for compression.

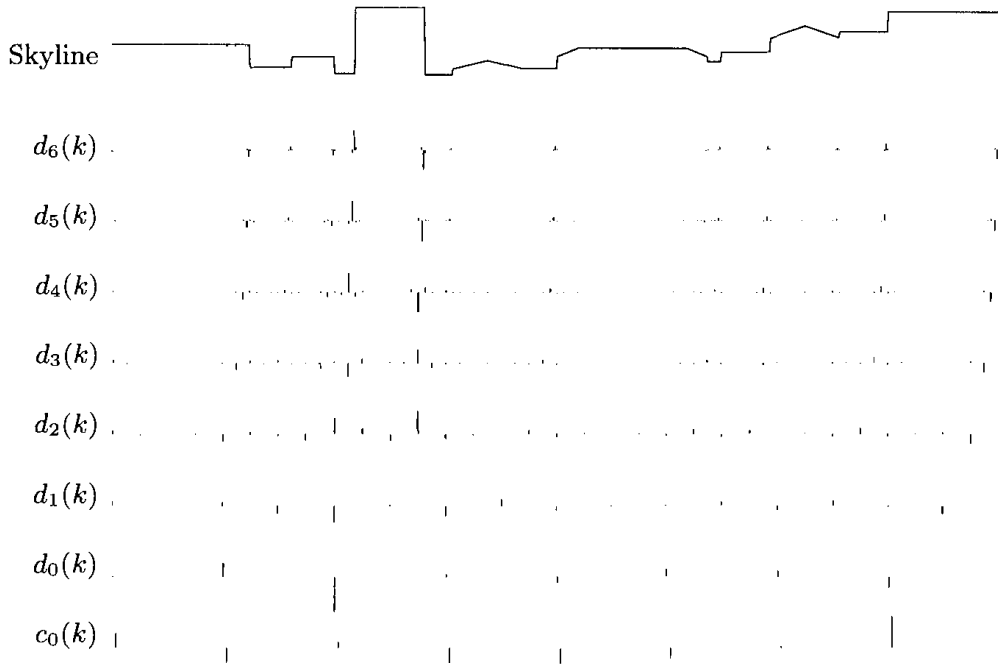
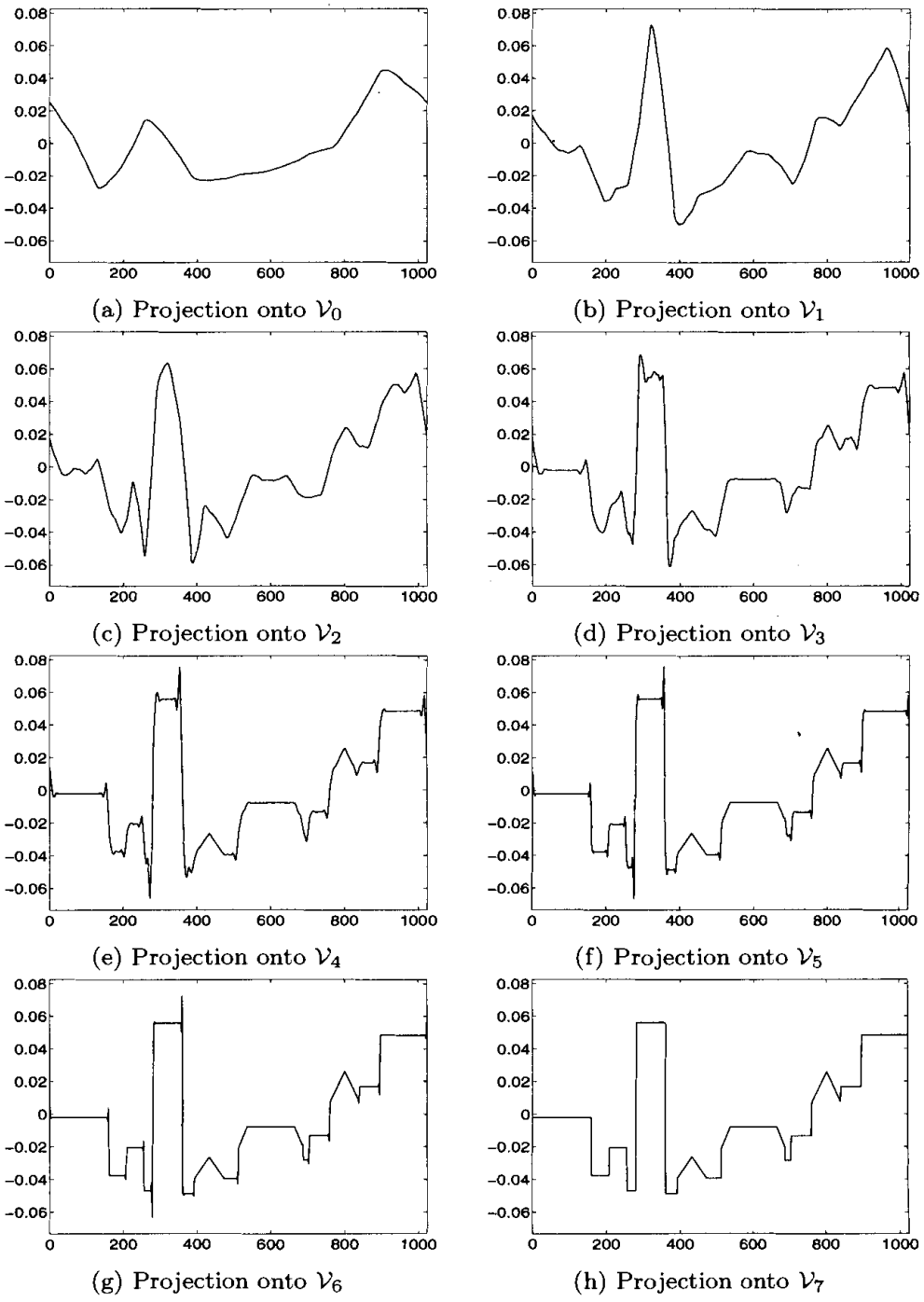


Figure 2.5. Discrete Wavelet Transform of the Houston Skyline, using $\psi_{D8'}$ with a Gain of $\sqrt{2}$ for Each Higher Scale

Figure 2.6 shows the approximations of the skyline signal in the various scaling function spaces $\mathcal{V}_j$. This illustrates just how the approximations progress, giving more and more resolution at higher scales. The fact that the higher scales give more detail is similar to Fourier methods, but the localization is new. Figure 2.7 illustrates the individual wavelet decomposition by showing

the components of the signal that exist in the wavelet spaces $\mathcal{W}_j$ at different scales j. This shows the same expansion as Figure 2.6, but with the wavelet components given separately rather than being cumulatively added to the scaling function. Notice how the large objects show up at the lower resolution. Groups of buildings and individual buildings are resolved according to their width. The edges, however, are located at the higher resolutions and are located very accurately.



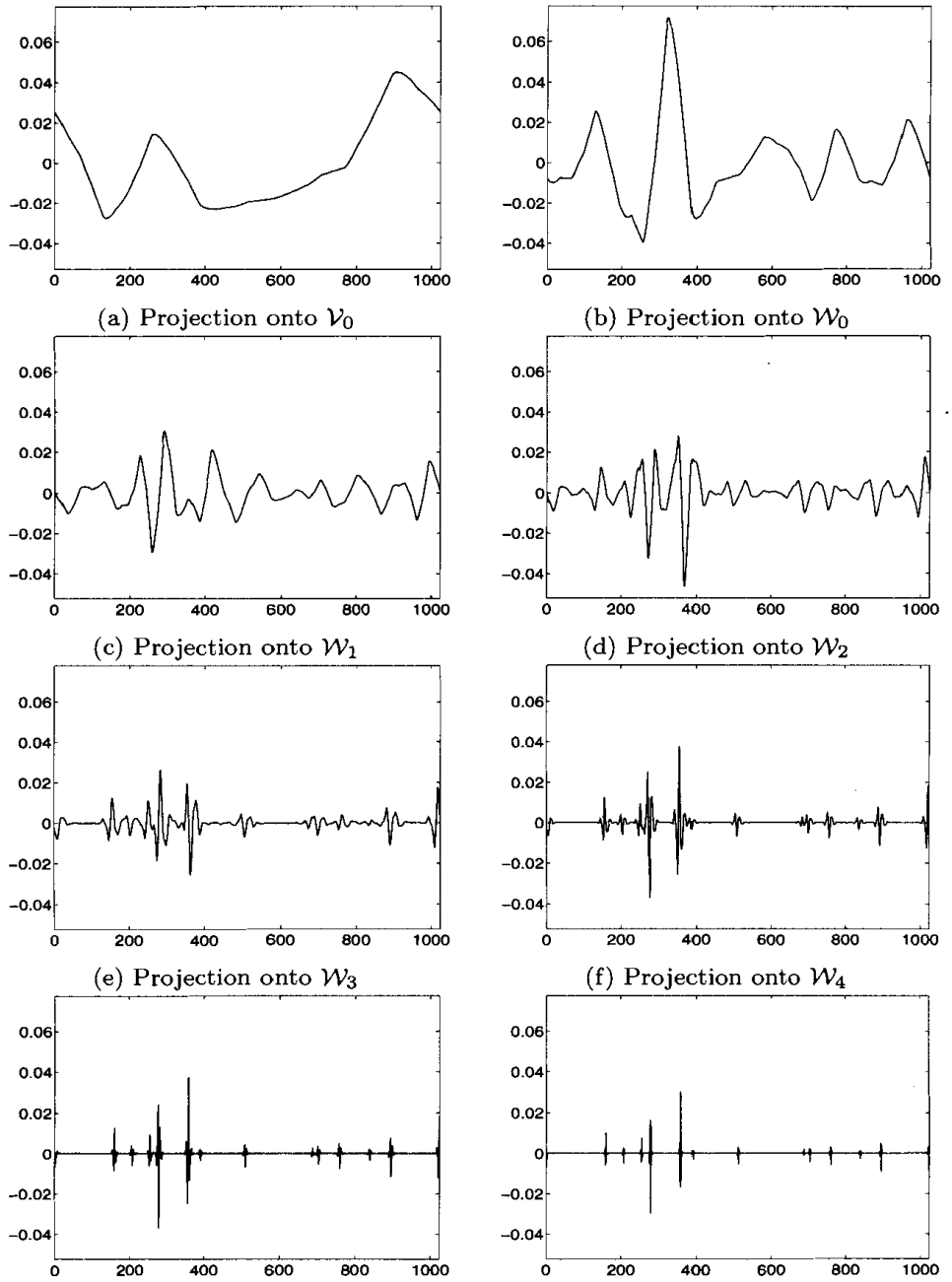
(a) Projection onto $\mathcal{V}_0$ (b) Projection onto $\mathcal{V}_1$

(c) Projection onto $\mathcal{V}_2$ (d) Projection onto $\mathcal{V}_3$

(e) Projection onto $\mathcal{V}_4$ (f) Projection onto $\mathcal{V}_5$

(g) Projection onto $\mathcal{V}_6$ (h) Projection onto $\mathcal{V}_7$

Figure 2.6. Projection of the Houston Skyline Signal onto $\mathcal{V}$ Spaces using ϕ_{D8}.



(a) Projection onto $\mathcal{V}_0$ (b) Projection onto $\mathcal{W}_0$

(c) Projection onto $\mathcal{W}_1$ (d) Projection onto $\mathcal{W}_2$

(e) Projection onto $\mathcal{W}_3$ (f) Projection onto $\mathcal{W}_4$

(g) Projection onto $\mathcal{W}_5$ (h) Projection onto $\mathcal{W}_6$

Figure 2.7. Projection of the Houston Skyline Signal onto $\mathcal{W}$ Spaces using $\psi_{D8'}$

The second example uses a chirp or doppler signal to illustrate how a time-varying frequency is described by the scale decomposition. Figure 2.8 gives the coefficients of the DWT directly as a function of j and k. Notice how the location in k tracks the frequencies in the signal in a way the Fourier transform cannot. Figures 2.9 and 2.10 show the scaling function approximations and the wavelet decomposition of this chirp signal. Again, notice in this type of display how the "location" of the frequencies are shown.

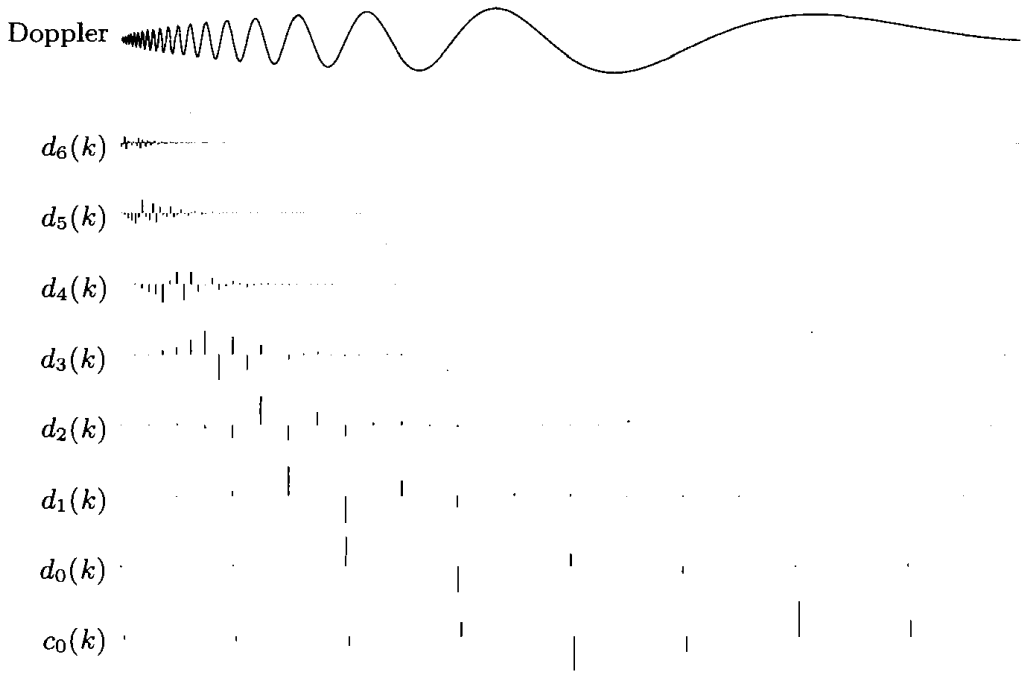


Figure 2.8. Discrete Wavelet Transform of a Doppler, using $\psi_{D8'}$ with a gain of $\sqrt{2}$ for each higher scale.

2.8 An Example of the Haar Wavelet System

In this section, we can illustrate our mathematical discussion with a more complete example. In 1910, Haar [Haa10] showed that certain square wave functions could be translated and scaled to create a basis set that spans L^2. This is illustrated in Figure 2.11. Years later, it was seen that Haar's system is a particular wavelet system.

If we choose our scaling function to have compact support over $0 \leq t \leq 1$, then a solution to (2.13) is a scaling function that is a simple rectangle function

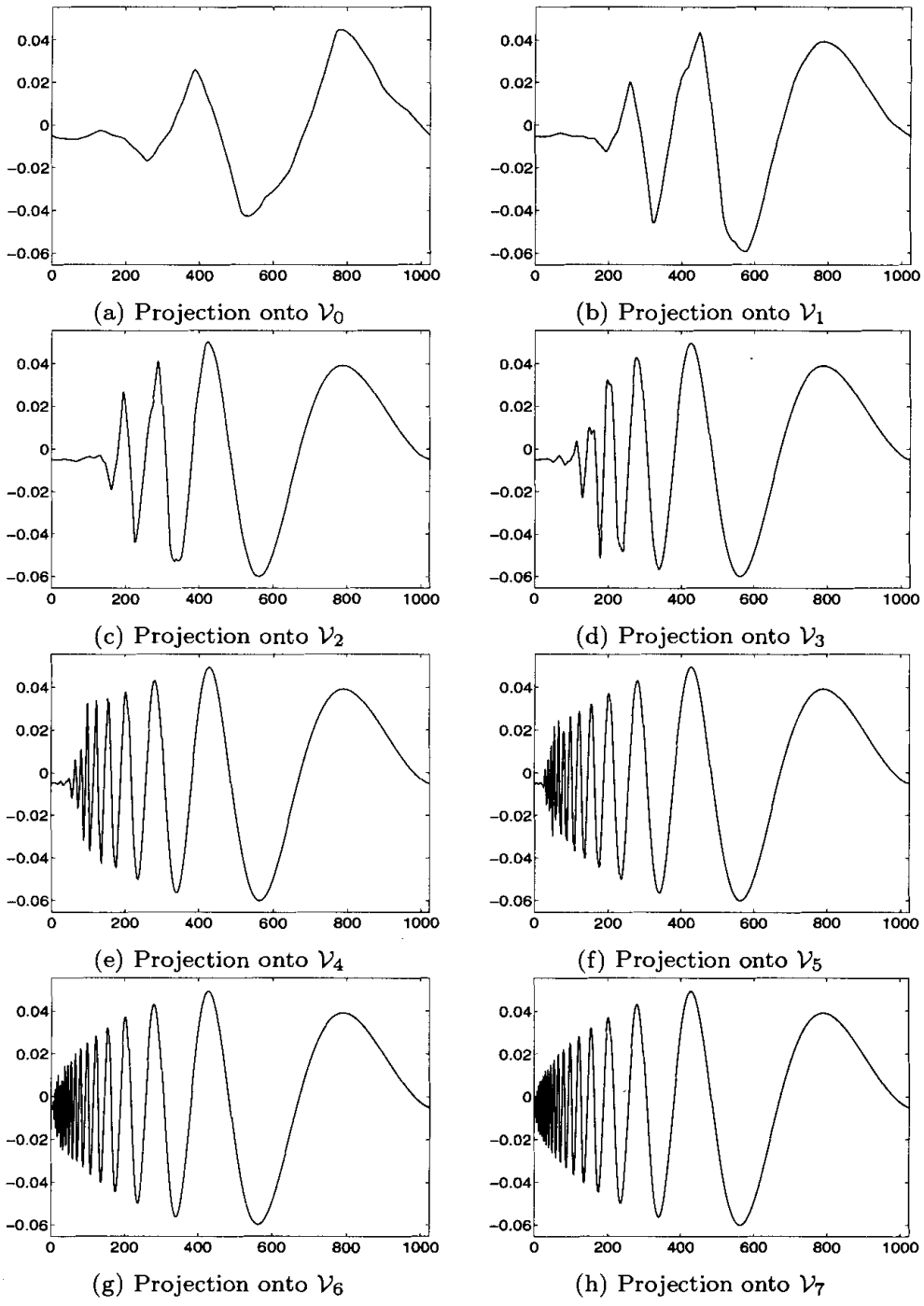
$$\varphi(t) = \begin{cases} 1 & \text{if } 0 < t < 1 \\ 0 & \text{otherwise} \end{cases} \tag{2.42}$$

with only two nonzero coefficients $h(0) = h(1) = 1/\sqrt{2}$ and (2.24) and (2.25) require the wavelet to be

$$\psi(t) = \begin{cases} 1 & \text{for } 0 < t < 0.5 \\ -1 & \text{for } 0.5 < t < 1 \\ 0 & \text{otherwise} \end{cases} \tag{2.43}$$

with only two nonzero coefficients $h_1(0) = 1/\sqrt{2}$ and $h_1(1) = -1/\sqrt{2}$.

$\mathcal{V}_0$ is the space spanned by $\varphi(t-k)$ which is the space of piecewise constant functions over integers, a rather limited space, but nontrivial. The next higher resolution space $\mathcal{V}_1$ is spanned by $\varphi(2t-k)$ which allows a somewhat more interesting class of signals which does include $\mathcal{V}_0$. As we consider higher values of scale j, the space $\mathcal{V}_j$ spanned by $\varphi(2^j t - k)$ becomes better able to approximate arbitrary functions or signals by finer and finer piecewise constant functions.



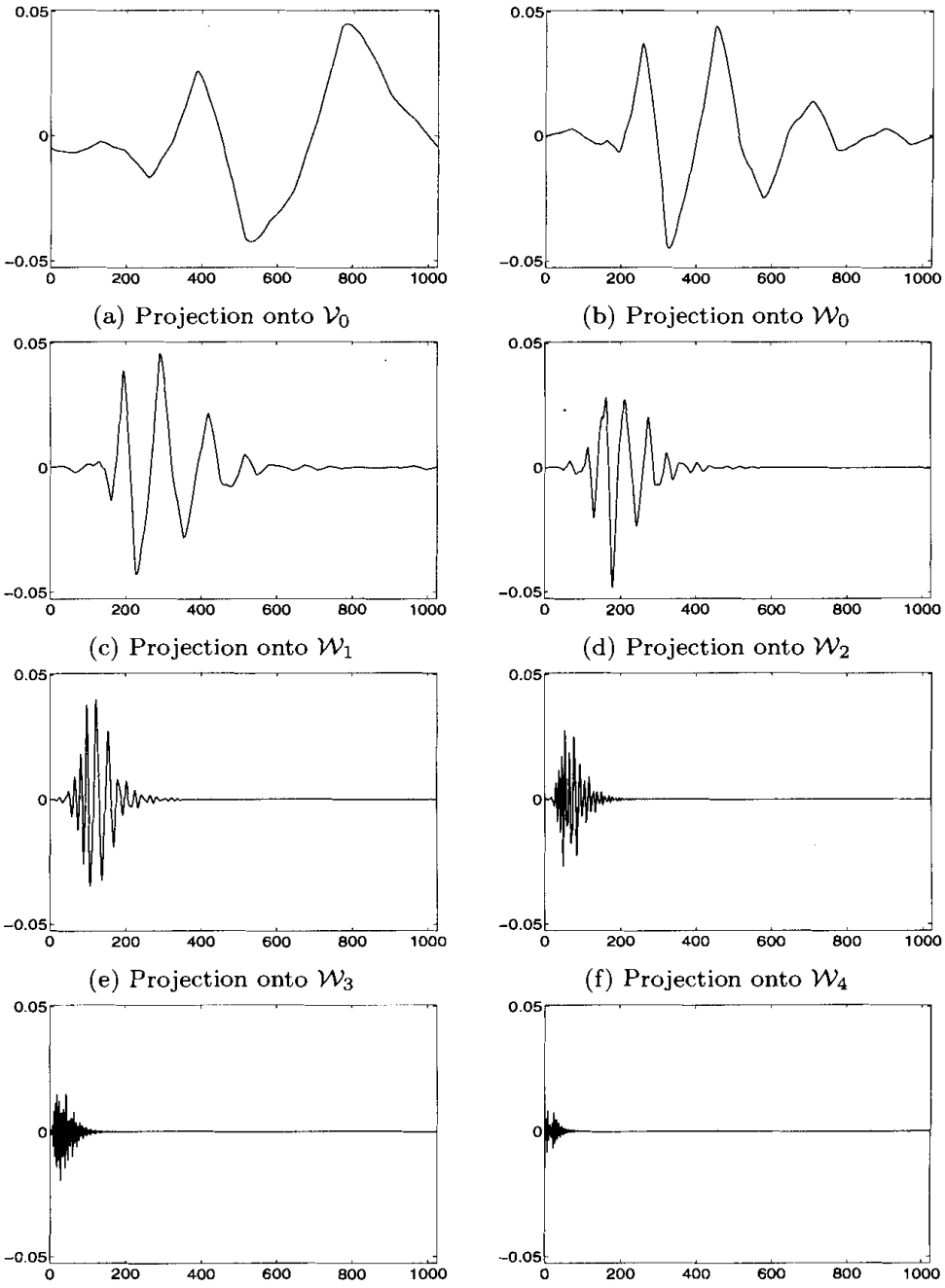
(a) Projection onto $\mathcal{V}_0$ (b) Projection onto $\mathcal{V}_1$

(c) Projection onto $\mathcal{V}_2$ (d) Projection onto $\mathcal{V}_3$

(e) Projection onto $\mathcal{V}_4$ (f) Projection onto $\mathcal{V}_5$

(g) Projection onto $\mathcal{V}_6$ (h) Projection onto $\mathcal{V}_7$

Figure 2.9. Projection of the Doppler Signal onto $\mathcal{V}$ Spaces using $\phi_{D8'}$



(a) Projection onto $\mathcal{V}_0$

(b) Projection onto $\mathcal{W}_0$

(c) Projection onto $\mathcal{W}_1$

(d) Projection onto $\mathcal{W}_2$

(e) Projection onto $\mathcal{W}_3$

(f) Projection onto $\mathcal{W}_4$

(g) Projection onto $\mathcal{W}_5$

(h) Projection onto $\mathcal{W}_6$

Figure 2.10. Projection of the Doppler Signal onto $\mathcal{W}$ Spaces using $\psi_{D8'}$

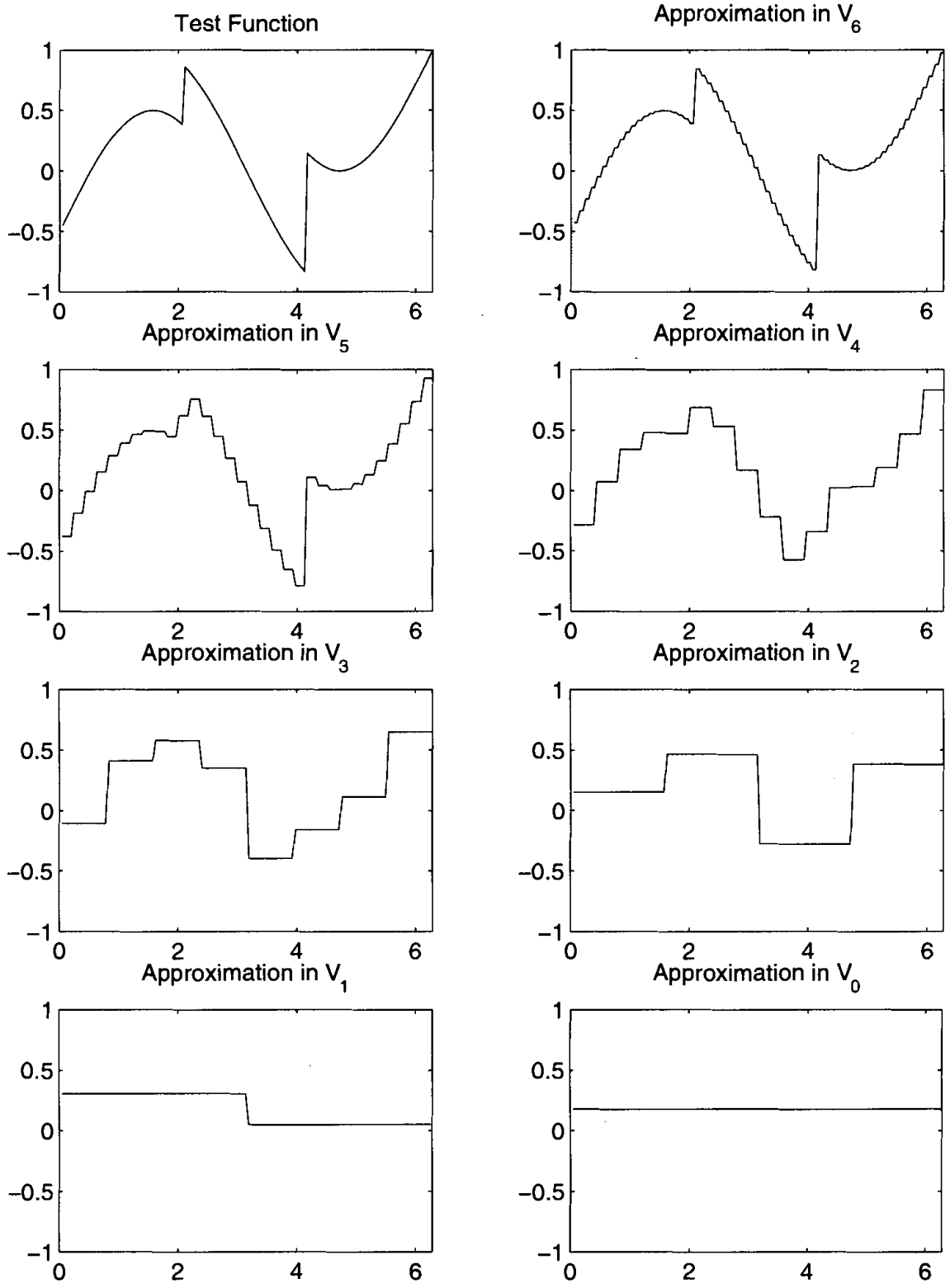


Figure 2.15. Haar Function Approximation in $\mathcal{V}_j$

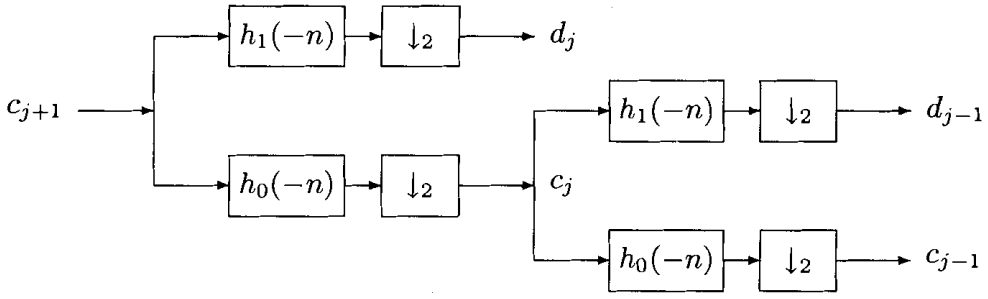


Figure 3.3. Two-Stage Two-Band Analysis Tree

As we will see in Chapter 5, the FIR filter implemented by $h(-n)$ is a lowpass filter, and the one implemented by $h_1(-n)$ is a highpass filter. Note the average number of data points out of this system is the same as the number in. The number is doubled by having two filters; then it is halved by the decimation back to the original number. This means there is the possibility that no information has been lost and it will be possible to completely recover the original signal. As we shall see, that is indeed the case. The aliasing occurring in the upper bank can be "undone" or cancelled by using the signal from the lower bank. This is the idea behind perfect reconstruction in filter bank theory [Vai92, Fli94].

This splitting, filtering, and decimation can be repeated on the scaling coefficients to give the two-scale structure in Figure 3.3. Repeating this on the scaling coefficients is called *iterating the filter bank.* Iterating the filter bank again gives us the three-scale structure in Figure 3.4.

The frequency response of a digital filter is the discrete-time Fourier transform of its impulse response (coefficients) $h(n)$. That is given by

$$H(\omega) = \sum_{n=-\infty}^{\infty} h(n)\, e^{i\omega n}. \tag{3.12}$$

The magnitude of this complex-valued function gives the ratio of the output to the input of the filter for a sampled sinusoid at a frequency of ω in radians per seconds. The angle of $H(\omega)$ is the phase shift between the output and input.

The first stage of two banks divides the spectrum of $c_{j+1}(k)$ into a lowpass and highpass band, resulting in the scaling coefficients and wavelet coefficients at lower scale $c_j(k)$ and $d_j(k)$. The second stage then divides that lowpass band into another lower lowpass band and a bandpass band. The first stage divides the spectrum into two equal parts. The second stage divides the lower half into quarters and so on. This results in a logarithmic set of bandwidths as illustrated in Figure 3.5. These are called "constant-Q" filters in filter bank language because the ratio of the band width to the center frequency of the band is constant. It is also interesting to note that a musical scale defines octaves in a similar way and that the ear responds to frequencies in a similar logarithmic fashion.

For any practical signal that is bandlimited, there will be an upper scale $j = J$, above which the wavelet coefficients, $d_j(k)$, are negligibly small [GOB94]. By starting with a high resolution description of a signal in terms of the scaling coefficients c_J, the analysis tree calculates the DWT

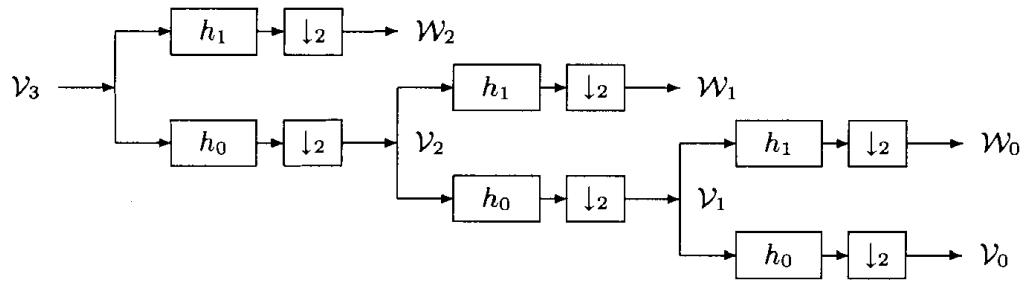


Figure 3.4. Three-Stage Two-Band Analysis Tree

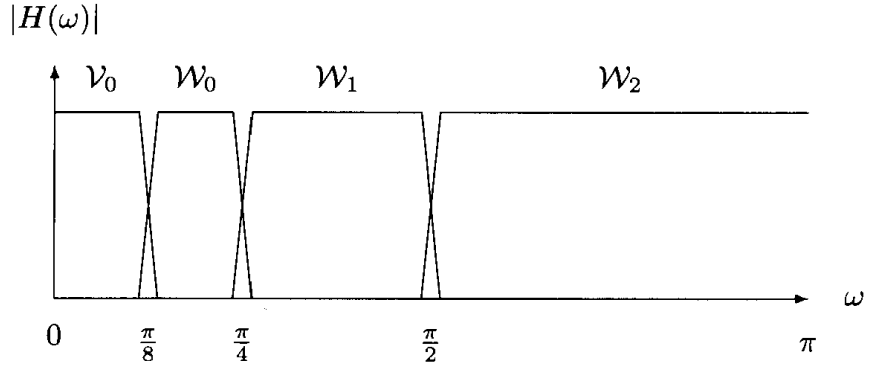


Figure 3.5. Frequency Bands for the Analysis Tree

down to as low a resolution, $j = j_0$, as desired by having $J - j_0$ stages. So, for $f(t) \in \mathcal{V}_J$, using (2.8) we have

$$f(t) = \sum_k c_J(k)\, \varphi_{J,k}(t) \tag{3.13}$$

$$= \sum_k c_{J-1}(k)\, \varphi_{J-1,k}(t) + \sum_k d_{J-1}(k)\, \psi_{J-1,k}(t) \tag{3.14}$$

$$f(t) = \sum_k c_{J-2}(k)\, \varphi_{J-2,k}(t) + \sum_k \sum_{j=J-2}^{J-1} d_j(k)\, \psi_{j,k}(t) \tag{3.15}$$

$$f(t) = \sum_k c_{j_0}(k)\, \varphi_{j_0,k}(t) + \sum_k \sum_{j=j_0}^{J-1} d_j(k)\, \psi_{j,k}(t) \tag{3.16}$$

which is a finite scale version of (2.33). We will discuss the choice of j_0 and J further in Chapter 9.

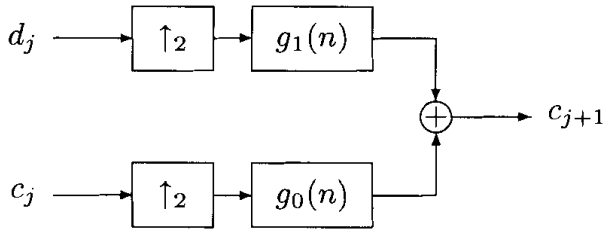


Figure 3.6. Two-Band Synthesis Bank

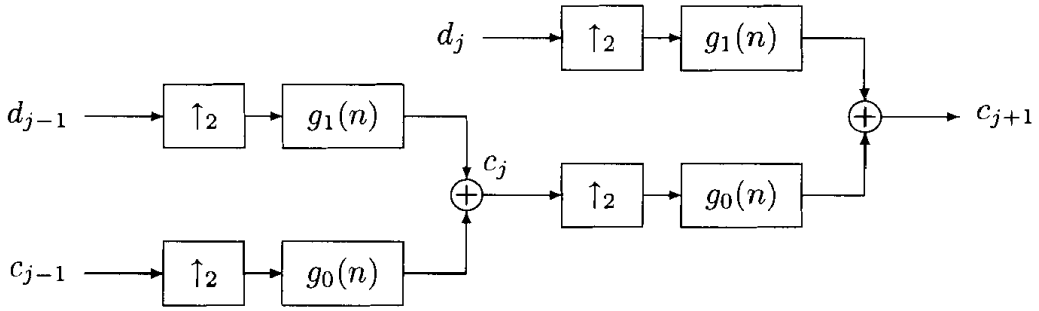


Figure 3.7. Two-Stage Two-Band Synthesis Tree

3.3 Input Coefficients

One might wonder how the input set of scaling coefficients c_{j+1} are obtained from the signal to use in the systems of Figures 3.2 and 3.3. For high enough scale, the scaling functions act as "delta functions" with the inner product to calculate the high scale coefficients as simply a sampling of $f(t)$ [GOB92, OGB92]. If the samples of $f(t)$ are above the Nyquist rate, they are good approximations to the scaling coefficients at that scale, meaning no wavelet coefficients are necessary at that scale. This approximation is particularly good if moments of the scaling function are zero or small. These ideas are further explained in Section 6.8 and Chapter 9.

An alternative approach is to "prefilter" the signal samples to make them a better approximation to the expansion coefficients. This is discussed in [Str86].

This set of analysis and synthesis operations is known as Mallat's algorithm [Mal89b, Mal89c]. The analysis filter bank efficiently calculates the DWT using banks of digital filters and downsamplers, and the synthesis filter bank calculates the inverse DWT to reconstruct the signal from the transform. Although presented here as a method of calculating the DWT, the filter bank description also gives insight into the transform itself and suggests modifications and generalizations that would be difficult to see directly from the wavelet expansion point of view. Filter banks will be used more extensively in the remainder of this book. A more general development of filter banks is presented in Section 7.2.