4

HAAR WAVELET ANALYSIS

4.1 WHY WAVELETS?

Wavelets were first applied in geophysics to analyze data from seismic surveys, which are used in oil and mineral exploration to get "pictures" of layering in subsurface rock. In fact, geophysicists rediscovered them; mathematicians had developed them to solve abstract problems some 20 years earlier, but had not anticipated their applications in signal processing.[1]

Seismic surveys are made up of many two-dimensional pictures or slices. These are sewn together to give a three-dimensional image of the structure of rock below the surface. Each slice is obtained by placing geophones—seismic "microphones"—at equally spaced intervals along a line, the seismic line. Dynamite is set off at one end of the line to create a seismic wave in the ground. Every geophone along the line records the movement of the earth due to the blast, from start to finish; this record is its *seismic trace* (see Figure 4.1).

The first wave recorded by the geophones is the direct wave, which travels along the surface. This is usually not important. Subsequent waves are reflected off rock layers below ground. These are the important ones. Knowledge of the time that the wave hits a geophone gives information about where the layer that reflected it is located. The "wiggles" that the wave produces tell something about

[1]See Meyer's book (Meyer, 1993) for an interesting, first-hand account of how wavelets developed
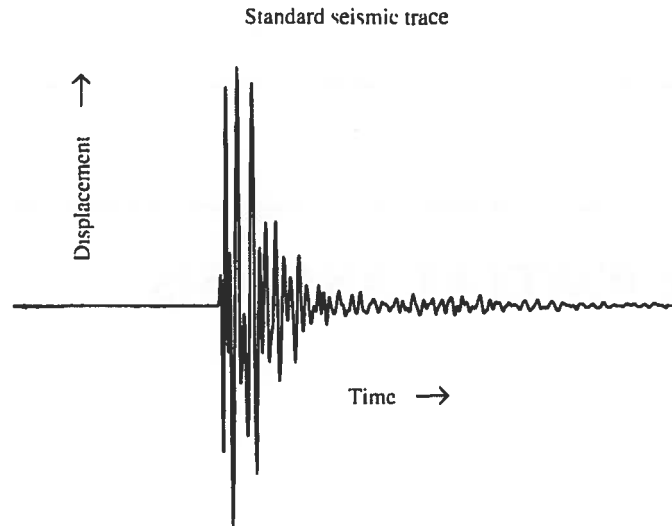
Standard seismic trace

Figure 4.1. The figure shows a typical seismic trace Displacement is plotted versus time. Both the oscillations and the time they occur are important

the fine details of the layer. The traces from the all the geophones on a line can be combined to give the slice for the ground directly beneath the line.

The key to an accurate seismic survey is proper analysis of each trace. The Fourier transform is not a good tool here. It can only provide frequency information (the oscillations that comprise the signal). It gives no direct information about *when* an oscillation occurred. Another tool, the short-time Fourier transform, is better. The full time interval is divided into a number of small, equal time intervals; these are individually analyzed using the Fourier transform. The result contains time and frequency information. However, there is a problem with this approach. The equal time intervals are not adjustable; the times when very short duration, high-frequency bursts occur are hard to detect.

Enter wavelets. Wavelets can keep track of time and frequency information. They can be used to "zoom in" on the short bursts mentioned previously, or to "zoom out" to detect long, slow oscillations.

4.2 HAAR WAVELETS

4.2.1 The Haar Scaling Function

There are two functions that play a primary role in wavelet analysis: the scaling function ϕ and the wavelet ψ. These two functions generate a family of functions that can be used to break up or reconstruct a signal. To emphasize the "marriage" involved in building this "family," ϕ is sometimes called the "father wavelet," and ψ is sometimes referred to as the "mother wavelet."
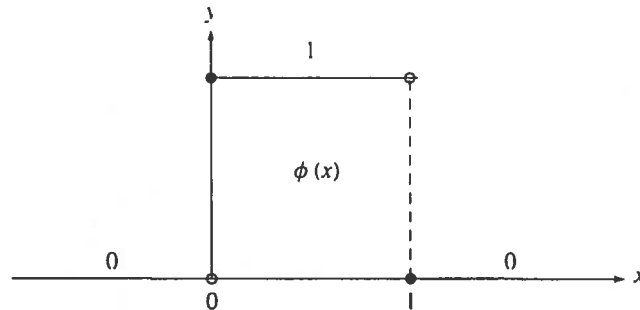
Figure 4.2. Graph of the Haar scaling function.

The simplest wavelet analysis is based on the *Haar scaling function*, whose graph is given in Figure 4.2. The building blocks are translations and dilations (both in height and width) of this basic graph.

We want to illustrate the basic ideas involved in such an analysis. Consider the signal shown in Figure 4.3. We may think of this as a measurement of some
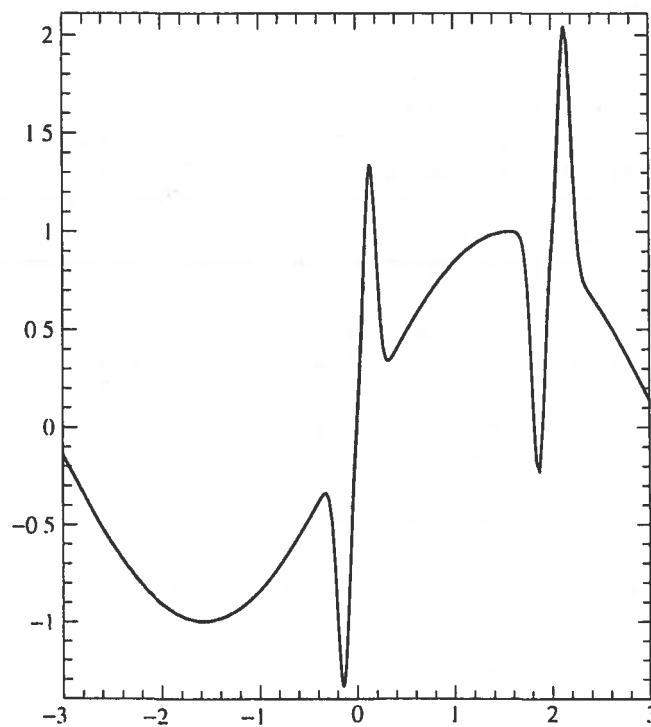

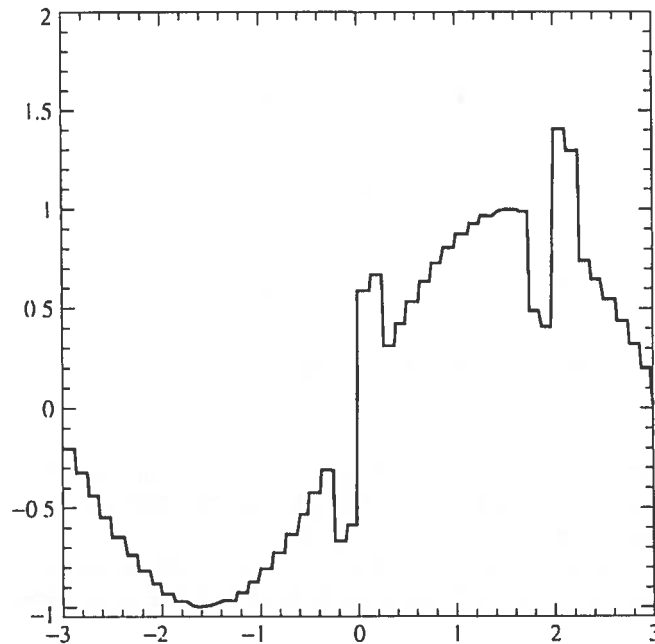**Figure 4.3.** Voltage from a faulty meter.

Figure 4.4. Approximation of voltage signal by Haar functions

physical quantity—perhaps line voltage over a single cycle—as a function of time. The two sharp spikes in the graph might represent noise coming from a loose connection in the volt meter, and we want to filter out this undesirable noise. The graph in Figure 4.4 shows one possible approximation to the signal using Haar building blocks. The high-frequency noise shows up as tall thin blocks. An algorithm that deletes the thin blocks will eliminate the noise and not disturb the rest of the signal.

The building blocks generated by the Haar scaling function are particularly simple and they illustrate the general ideas underlying a multiresolution analysis, which we will discuss in detail. The disadvantage of the Haar wavelets is that they are discontinuous and therefore do not approximate continuous signals very well (Figure 4.4 does not really approximate Figure 4.3 too well). In later sections we introduce other wavelets, due to Daubechies, that are continuous but still retain the localized behavior exhibited by the Haar wavelets.

Definition 4.1   The Haar scaling function is defined as

$$\phi(x) = \begin{cases} 1 & \text{if } 0 \le x < 1, \\ 0 & \text{elsewhere.} \end{cases}$$

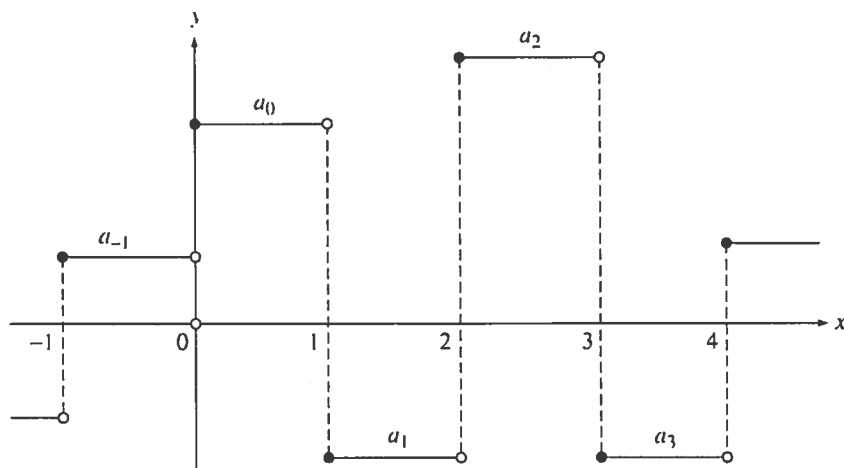The graph of the Haar scaling function is given in Figure 4.2.

Figure 4.5. Graph of typical element in V_0.

The function $\phi(x - k)$ has the same graph as ϕ but is translated to the right by k units (assuming k is positive). Let V_0 be the space of all functions of the form

$$\sum_{k \in \mathbb{Z}} a_k \phi(x - k), \qquad a_k \in R,$$

where k can range over any finite set of positive or negative integers. Since $\phi(x - k)$ is discontinuous at $x = k$ and $x = k + 1$, an alternative description of V_0 is that it consists of all piecewise constant functions whose discontinuities are contained in the set of integers. Since k ranges over a finite set, each element of V_0 is zero outside a bounded set. Such a function is said to have *finite or compact support*. The graph of a typical element of V_0 is given in Figure 4.5. Note that a function in V_0 may not have discontinuities at all the integers (for example, if $a_1 = a_2$, then the previous sum is continuous at $x = 2$).

Example 4.2 The graph of the function

$$f(x) = 2\phi(x) + 3\phi(x - 1) + 3\phi(x - 2) - \phi(x - 3) \in V_0$$

is given in Figure 4.6. This function has discontinuities at $x = 0, 1, 3$, and 4. ∎

We need blocks that are thinner to analyze signals of high frequency. The building block whose width is half that of the graph of ϕ is given by the graph of $\phi(2x)$ shown in Figure 4.7.

The function $\phi(2x - k) = \phi(2(x - k/2))$ is the same as the graph of the function of $\phi(2x)$ but shifted to the right by $k/2$ units. Let V_1 be the space of
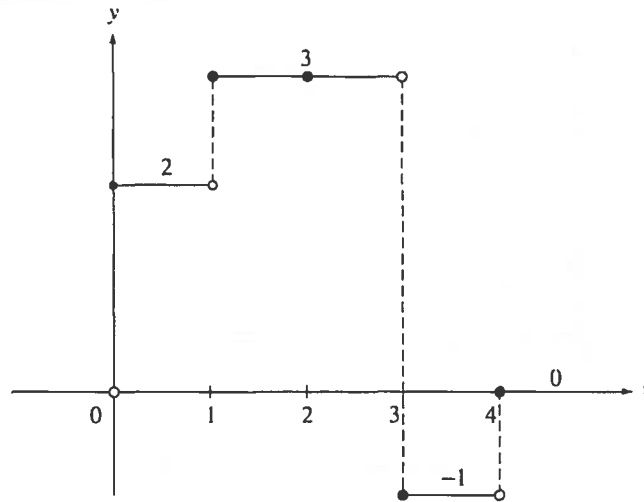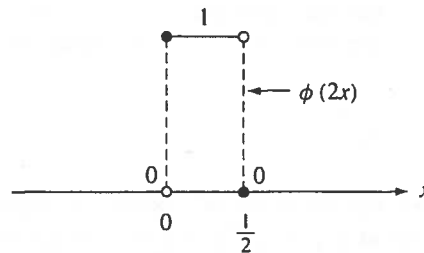
Figure 4.6. Plot of f in Example 4 2.

Figure 4.7. Graph of $\phi(2x)$.

functions of the form

$$\sum_{k \in Z} a_k \phi(2x - k), \qquad a_k \in R.$$

Geometrically, V_1 is the space of piecewise constant functions of finite support with possible discontinuities at the half integers $\{0, \pm 1/2, \pm 1, \pm 3/2 \ldots\}$.

Example 4.3 The graph of the function

$$f(x) = 4\phi(2x) + 2\phi(2x - 1) + 2\phi(2x - 2) - \phi(2x - 3) \in V_0$$

is given in Figure 4.8. This function has discontinuities at $x = 0, 1/2, 3/2$, and 2.
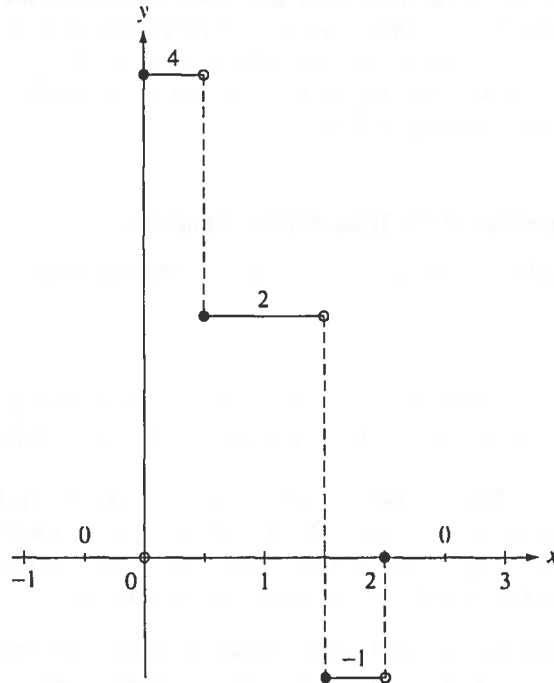
We make the following more general definition. ∎

Figure 4.8. Plot of f in Example 4.3.

Definition 4.4 Suppose j is any nonnegative integer. The space of step functions at level j, denoted by V_j, is defined to be the space spanned by the set

$$\{\ldots\phi(2^j x + 1), \phi(2^j x), \phi(2^j x - 1), \phi(2^j x - 2)\ldots\}$$

over the real numbers. V_j is the space of piecewise constant functions of finite support whose discontinuities are contained in the set

$$\{\ldots-1/2^j, 0, 1/2^j, 2/2^j, 3/2^j \ldots\}.$$

A function in V_0 is a piecewise constant function with discontinuities contained in the set of integers. Any function in V_0 is also contained in V_1, which consists of piecewise constant functions whose discontinuities are contained in the set of half integers $\{\ldots-1/2, 0, 1/2, 1, 3/2\ldots\}$. Likewise, $V_1 \subset V_2$ and so forth:

$$V_0 \subset V_1 \subset \cdots V_{j-1} \subset V_j \subset V_{j+1}\cdots.$$

This containment is strict. For example, the function $\phi(2x)$ belongs to V_1 but does not belong to V_0 (since $\phi(2x)$ is discontinuous at $x = 1/2$).

V_j contains all relevant information up to a resolution scale of order 2^{-j}. As j gets larger, the resolution gets finer. The fact that $V_j \subset V_{j+1}$ means that

no information is lost as the resolution gets finer. This containment relation is also the reason why V_j is defined in terms of $\phi(2^j x)$ instead of $\phi(ax)$ for some other factor a. If, for example, we had defined V_2 using $\phi(3x - j)$ instead of $\phi(4x - j)$, then V_2 would not contain V_1 (since the set of multiples of 1/2 is not contained in the set of multiples of 1/3).

4.2.2 Basic Properties of the Haar Scaling Function

The following theorem is an easy consequence of the definitions.

Theorem 4.5

- A function $f(x)$ belongs to V_0 if and only if $f(2^j x)$ belongs to V_j.
- A function $f(x)$ belongs to V_j if and only if $f(2^{-j} x)$ belongs to V_0.

Proof. To prove the first statement, if a function f belongs to V_0 then $f(x)$ is a linear combination of $\{\phi(x - k), k \in Z\}$. Therefore, $f(2^j x)$ is a linear combination of $\{\phi(2^j x - k), k \in Z\}$, which means that $f(2^j x)$ is a member of V_j. The proofs of the converse and the second statement are similar. ∎

The graph of the function $\phi(2^j x)$ is a spike of width $1/2^j$. When j is large, the graph of $\phi(2^j x)$ (appropriately translated) is similar to one of the spikes of a signal that we may wish to filter out Thus it is desirable to have an efficient algorithm to decompose a signal into its V_j components. One way to perform this decomposition is to construct an orthonormal basis for V_j (using the L^2 inner product).

Let's start with V_0. This space is generated by ϕ and its translates. The functions $\phi(x - k)$ each have unit norm in L^2, that is,

$$\|\phi(x - k)\|_{L^2}^2 = \int_{-\infty}^{\infty} \phi(x - k)^2 \, dx = \int_k^{k+1} 1 \, dx = 1.$$

If j is different from k, then $\phi(x - j)$ and $\phi(x - k)$ have disjoint supports (see Figure 4.9). Therefore

$$\langle \phi(x - j), \phi(x - k) \rangle_{L^2} = \int_{-\infty}^{\infty} \phi(x - j)\phi(x - k) \, dx = 0, \qquad j \neq k$$

and so the set $\{\phi(x - k), k \in Z\}$ is an orthonormal basis for V_0.

The same argument establishes the following more general result.

Theorem 4.6 The set of functions $\{2^{j/2}\phi(2^j x - k): k \in Z\}$ is an orthonormal basis of V_j

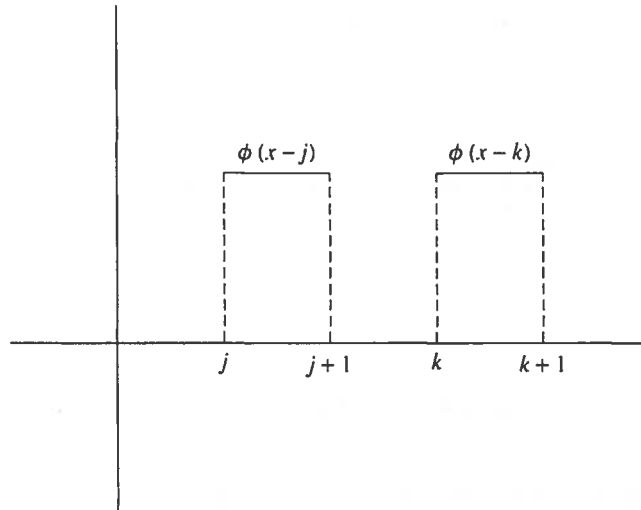(The factor $2^{j/2}$ is present since $\int_{-\infty}^{\infty}(\phi(2^j x))^2 \, dx = 1/2^j$.)

Figure 4.9. $\phi(x - j)$ and $\phi(x - k)$ have disjoint support

4.2.3 The Haar Wavelet

Having an orthonormal basis of V_j is only half of the picture. In order to solve our noise filtering problem, we need to have a way of isolating the "spikes" which belong to V_j but which are not members of V_{j-1}. This is where the wavelet ψ enters the picture.

The idea is to decompose V_j as an orthogonal sum of V_{j-1} and its complement. Again, let's start with $j = 1$ and identify the orthogonal complement of V_0 in V_1. Since V_0 is generated by ϕ and its translates, it is reasonable to expect that the orthogonal complement of V_0 is generated by the translates of some function ψ. Two key facts are needed to construct ψ.

1. ψ is a member of V_1 and so ψ can be expressed as $\psi(x) = \sum_l a_l \phi(2x - l)$ for some choice of $a_l \in R$ (and only a finite number of the a_l are non-zero).

2. ψ is orthogonal to V_0. This is equivalent to $\int \psi(x) \phi(x - k) \, dx = 0$ for all integers k.

The first requirement means that ψ is built from blocks of width $1/2$—that is, scalar multiples of Figure 4.7 and its translates. The second requirement with $k = 0$ implies $\int_{-\infty}^{\infty} \psi(x) \phi(x) \, dx = 0$. The simplest ψ satisfying both of these requirements is the function whose graph appears in Figure 4.10. This graph consists of two blocks of width one-half and can be written as

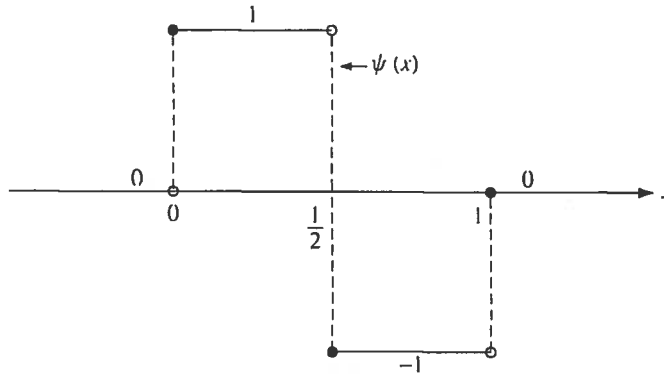$$\psi(x) = \phi(2x) - \phi(2(x - 1/2))$$
$$= \phi(2x) - \phi(2x - 1)$$

Figure 4.10. The Haar wavelet $\psi(x)$

thus satisfying the first requirement. In addition, we have

$$\int_{-\infty}^{\infty} \phi(x)\psi(x)\,dx = \int_{0}^{1/2} 1\,dx - \int_{1/2}^{1} 1\,dx$$
$$= 1/2 - 1/2$$
$$= 0.$$

Thus, ψ is orthogonal to ϕ. If $k \neq 0$, then the support of $\psi(x)$ and the support of $\phi(x - k)$ do not overlap and so $\int \psi(x)\phi(x - k)\,dx = 0$. Therefore, ψ belongs to V_1 and is orthogonal to V_0; ψ is called the *Haar wavelet*.

Definition 4.7 The Haar wavelet is the function

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

Its graph is given in Figure 4.10.

You can show (see exercise 5) that any function

$$f_1 = \sum_{k} a_k \phi(2x - k) \in V_1$$

is orthogonal to V_0 (i.e., orthogonal to each $\phi(x - l), l \in Z$) if and only if

$$a_1 = -a_0, \quad a_3 = -a_2 \cdots .$$

In this case,

$$f_1 = \sum_{k \in \mathbb{Z}} a_{2k} (\phi(2x - 2k) - \phi(2x - 2k - 1)) = \sum_{k \in \mathbb{Z}} a_{2k} \psi(x - k). \qquad (4.1)$$

In other words, a function in V_1 is orthogonal to V_0 if and only if it is of the form $\sum_k a_k \psi(x - k)$ (relabeling a_{2k} by a_k).

Let W_0 be the space of all functions of the form

$$\sum_{k \in Z} a_k \psi(x - k), \qquad a_k \in R,$$

where, again, we assume that only a finite number of the a_k are nonzero. What we have just shown is that W_0 is the orthogonal complement of V_0 in V_1; in other words, $V_1 = V_0 \oplus W_0$ (recall from Chapter 0 that $\oplus$ means that V_0 and W_0 are orthogonal to each other).

In a similar manner, the following, more general result can be established.

Theorem 4.8 Let W_j be the space of functions of the form

$$\sum_{k \in Z} a_k \psi(2^j x - k), \qquad a_k \in R,$$

where we assume that only a finite number of a_k are nonzero. W_j is the orthogonal complement of V_j in V_{j+1} and

$$V_{j+1} = V_j \oplus W_j.$$

Proof. To establish this theorem, we must show two facts:

1. Each function in W_j is orthogonal to every function in V_j.
2. Any function in V_{j+1} which is orthogonal to V_j must belong to W_j.

For the first requirement, suppose that $g = \sum_{k \in Z} a_k \psi(2^j x - k)$ belongs to W_j and suppose f belongs to V_j. We must show

$$\langle g, f \rangle_{L^2} = \int_{-\infty}^{\infty} g(x)\overline{f(x)}\, dx = 0.$$

Since $f(x)$ belongs to V_j, the function $f(2^{-j}x)$ belongs to V_0. So

$$0 = \int_{-\infty}^{\infty} \sum_{k \in Z} a_k \psi(x - k)\overline{f(2^{-j}x)}\, dx \qquad \text{(because ψ is orthogonal to V_0)}$$

$$= 2^j \int_{\infty}^{\infty} \sum_{k \in Z} a_k \psi(2^j y - k)\overline{f(y)}\, dy \qquad \text{(by letting $y = 2^{-j}x$)}$$

$$= 2^j \int_{-\infty}^{\infty} g(y)\overline{f(y)}\, dy$$

Therefore, g is orthogonal to any $f \in V_j$ and the first requirement has been established.

The discussion leading to Eq. (4.1) establishes the second requirement when $j = 0$ since we showed that any function in V_1 which is orthogonal to V_0 must be a linear combination of $\{\psi(x - k),\ k \in Z\}$. The case for general j is analogous to the case when $j = 0$. ∎

By successively decomposing V_j, V_{j-1}, and so on, we have

$$
\begin{aligned}
V_j &= W_{j-1} \oplus V_{j-1} \\
    &= W_{j-1} \oplus W_{j-2} \oplus V_{j-2} \\
    &\quad\cdots \\
    &= W_{j-1} \oplus W_{j-2} \oplus \cdots \oplus W_0 \oplus V_0.
\end{aligned}
$$

So each f in V_j can be decomposed uniquely as a sum

$$
f = w_{j-1} + w_{j-2} + \cdots + w_0 + f_0,
$$

where each w_l belongs to W_l, $0 \le l \le j - 1$ and f_0 belongs to V_0. Intuitively, w_l represents the "spikes" of f of width $1/2^{l-1}$ that cannot be represented as linear combinations of spikes of other widths.

What happens when j goes to infinity? The answer is provided in the following theorem.

Theorem 4.9 The space $L^2(R)$ can be decomposed as an infinite orthogonal direct sum

$$
L^2(R) = V_0 \oplus W_0 \oplus W_1 \oplus \cdots.
$$

In particular, each $f \in L^2(R)$ can be written uniquely as

$$
f = f_0 + \sum_{j=0}^{\infty} w_j,
$$

where f_0 belongs to V_0 and w_j belongs to W_j.

The infinite sum should be thought of as a limit of finite sums. In other words,

$$
f = f_0 + \lim_{N \to \infty} \sum_{j=0}^{N} w_j. \tag{4.2}
$$

where the limit is taken in the sense of L^2. Although the proof of this result is beyond the scope of this text, some intuition can be given. There are two key ideas. The first is that any function in $L^2(R)$ can be approximated by continuous functions (for an intuitive explanation of this idea, see Lemma 1.38). The second
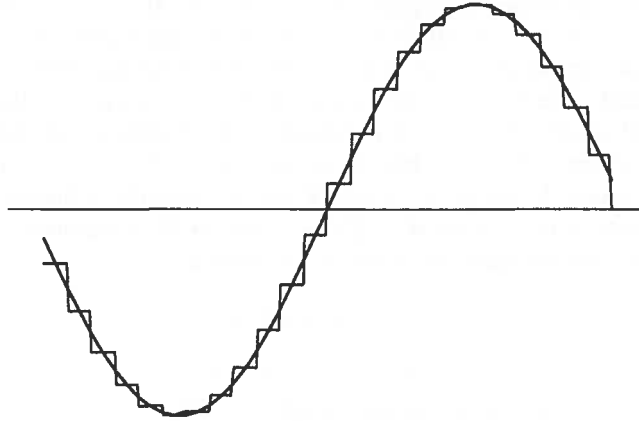
Figure 4.11. Approximating by step functions.

is that any continuous function can be approximated as closely as desired by a step function whose discontinuities are multiples of 2^{-j} for suitably large j (see Figure 4.11). Such a step function, by definition, belongs to V_j. The theorem is then established by putting both ideas together.

4.3 HAAR DECOMPOSITION AND RECONSTRUCTION ALGORITHMS

4.3.1 Decomposition

Now that V_j has been decomposed as a direct sum of V_0 and W_l for $0 \leq l < j$, the solution to our noise filtering problem is theoretically easy. First, we approximate f by a step function $f_j \in V_j$ (for j suitably large) using Theorem 4.9 Then we decompose f_j into its components

$$f_j = f_0 + w_1 + \cdots + w_{j-1}, \qquad w_l \in W_l.$$

The component, w_l, represents a "spike" of width $1/2^{l+1}$. For l sufficiently large, these spikes are thin enough to represent noise. For example, suppose that spikes of width less than .01 represent noise. Since $2^{-6} > .01 > 2^{-7}$, any w_j with $j \geq 6$ represents noise. To filter out this noise, these components can be set equal to zero. The rest of the sum represents a signal that is still relatively close to f and which is noise free.

In order to implement this theoretical algorithm, an efficient way of performing the decomposition given in Theorem 4 9 is needed. The first step is to approximate the original signal f by a step function of the form

$$f_j(x) = \sum_{l \in 7} a_l \phi(2^j x - l). \tag{4.3}$$

The procedure is to sample the signal at $x = \ldots - 1/2^j, 0, 1/2^j, \ldots$, which leads to $a_l = f(l/2^j)$ for $l \in Z$. An illustration of this procedure is given in Figure 4.11, where f is the continuous signal and f_j is the step function. Here, j is chosen so that the mesh size 2^{-j} is small enough so that $f_j(x)$ captures the essential features of the signal. The range of l depends on the domain of the signal. If the signal is defined on $0 \le x \le 1$, then the range of l is $0 \le l \le 2^j - 1$. In general, we will not specify the range of l unless a specific example is discussed.

Now, the task is to decompose $\phi(2^j x - l)$ into its W_l components for $l < j$. The following relations between ϕ and ψ are needed.

$$\phi(2x) = (\psi(x) + \phi(x))/2, \qquad (4.4)$$

$$\phi(2x - 1) = (\phi(x) - \psi(x))/2, \qquad (4.5)$$

which follow easily by looking at their graphs (see Figures 4.2 and 4.10). More generally, we have the following lemma.

Lemma 4.10   The following relations hold for all $x \in R$.

$$\phi(2^j x) = (\psi(2^{j-1}x) + \phi(2^{j-1}x))/2, \qquad (4.6)$$

$$\phi(2^j x - 1) = (\phi(2^{j-1}x) - \psi(2^{j-1}x))/2. \qquad (4.7)$$

This lemma follows by replacing x by $2^{j-1}x$ in Eqs. (4.4) and (4.5).

This lemma can be used to decompose $\phi(2^j x - l)$ into its W_l components for $l < j$. An example will help illustrate the process.

Example 4.11   Let f be given by the graph in Figure 4.12. Notice that the mesh size needed to capture all the features of f is 2^{-2}. A description of f in terms of $\phi(2^2 x - l)$ is given by

$$f(x) = 2\phi(4x) + 2\phi(4x - 1) + \phi(4x - 2) - \phi(4x - 3) \qquad (4.8)$$

We wish to decompose f into its W_1, W_0, and V_0 components. The following equations follow from Eqs. (4.6) and (4.7) with $j = 2$.

$$\phi(4x) = (\psi(2x) + \phi(2x))/2,$$

$$\phi(4x - 1) = (\phi(2x) - \psi(2x))/2,$$

$$\phi(4x - 2) = \phi(4(x - 1/2)) = (\psi(2(x - 1/2)) + \phi(2(x - 1/2)))/2,$$

$$\phi(4x - 3) = \phi(4(x - 1/2) - 1) = (\phi(2(x - 1/2)) - \psi(2(x - 1/2)))/2.$$

Using these equations together with Eq. (4.8) and collecting terms yield

$$\begin{aligned}
f(x) &= [\psi(2x) + \phi(2x)] + [\phi(2x) - \psi(2x)] \\
&\quad + [\psi(2x - 1) + \phi(2x - 1)]/2 - [\phi(2x - 1) - \psi(2x - 1)]/2 \\
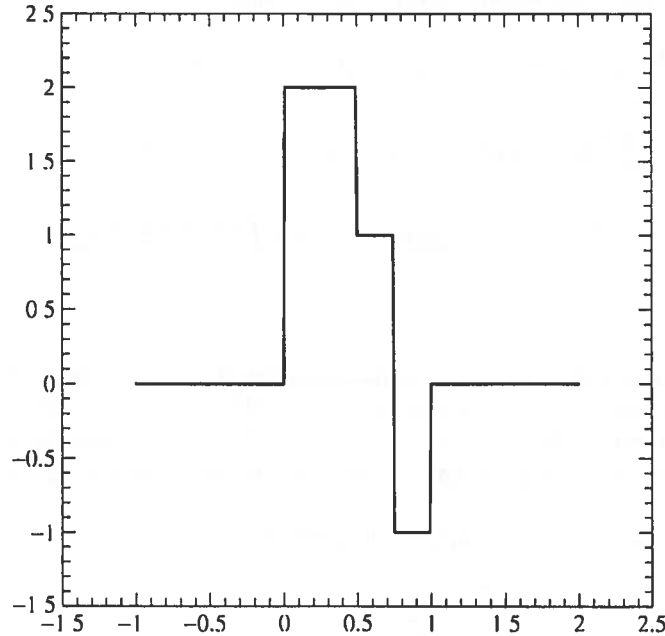&= \psi(2x - 1) + 2\phi(2x)
\end{aligned}$$

Figure 4.12. Graph of Example 4.11.

The W_1 component of $f(x)$ is $\psi(2x - 1)$, since W_1 is the linear span of $\{\psi(2x - k);\ k \in Z\}$. The V_1 component of $f(x)$ is given by $2\phi(2x)$. This component can be further decomposed into a V_0 component and a W_0 component by using the equation $\phi(2x) = (\phi(x) + \psi(x))/2$. The result is

$$f(x) = \psi(2x - 1) + \psi(x) + \phi(x).$$

This equation can also be verified geometrically by examining the graphs of the functions involved. The terms in the expression at right are the components of f in W_1, W_0, and V_0, respectively. ■

Using this example as a guide, we can proceed with the general decomposition scheme as follows. First divide up the sum $f_j(x) = \sum_k a_k \phi(2^j x - k)$ into even and odd terms:

$$f_j(x) = \sum_{k \in Z} a_{2k} \phi(2^j x - 2k) + \sum_{k \in Z} a_{2k+1} \phi(2^j x - 2k - 1). \tag{4.9}$$

Next, we use Eqs. (4.6) and (4.7) with x replaced by $x - k2^{1-j}$:

$$\phi(2^j x - 2k) = (\psi(2^{j-1} x - k) + \phi(2^{j-1} x - k))/2, \tag{4.10}$$

$$\phi(2^j x - 2k - 1) = (\phi(2^{j-1} x - k) - \psi(2^{j-1} x - k))/2. \tag{4.11}$$

Substituting these expressions into Eq. (4 9) yields

$$f_j(x) = \sum_{k \in Z} a_{2k} \left(\psi(2^{j-1}x - k) + \phi(2^{j-1}x - k) \right) / 2$$

$$+ \sum_{k \in /} a_{2k+1} \left(\phi(2^{j-1}x - k) - \psi(2^{j-1}x - k) \right) / 2$$

$$= \sum_{k \in /} \left(\frac{a_{2k} - a_{2k+1}}{2} \right) \psi(2^{j-1}x - k) + \left(\frac{a_{2k} + a_{2k+1}}{2} \right) \phi(2^{j-1}x - k)$$

$$= w_{j-1} + f_{j-1}.$$

The first term on the right, w_{j-1}, represents the W_{j-1} component of f_j since W_{j-1} is, by definition, the linear span of $\{\psi(2^{j-1}x - k), \ k \in Z\}$. Likewise, the second term on the right, f_{j-1}, represents the V_{j-1} component of f_j. We summarize the preceding decomposition algorithm in the following theorem.

Theorem 4.12 (Haar Decomposition) Suppose

$$f_j(x) = \sum_{k \in /} a_k^j \phi(2^j x - k) \in V_j.$$

Then f_j can be decomposed as

$$f_j = w_{j-1} + f_{j-1}. \qquad \blacklozenge$$

where

$$w_{j-1} = \sum_{k \in Z} b_k^{j-1} \psi(2^{j-1}x - k) \in W_{j-1}.$$

$$f_{j-1} = \sum_{k \in 7} a_k^{j-1} \phi(2^{j-1}x - k) \in V_{j-1}$$

with

$$b_k^{j-1} = \frac{a_{2k}^j - a_{2k+1}^j}{2}, \qquad a_k^{j-1} = \frac{a_{2k}^j + a_{2k-1}^j}{2}.$$

The preceding process can now be repeated with j replaced by $j - 1$ to decompose f_{j-1} as $w_{j-2} + f_{j-2}$. Continuing in this way, we achieve the decomposition

$$f_j = w_{j-1} + w_{j-2} + \cdots + w_0 + f_0.$$

To summarize the decomposition process, a signal is first discretized to produce an approximate signal $f_j \in V_j$ as in Theorem 4.9. Then the decomposition algorithm in Theorem 4 12 produces a decomposition of f_j into its various frequency components: $f_j = w_{j-1} + w_{j-2} + \cdots + w_0 + f_0$.
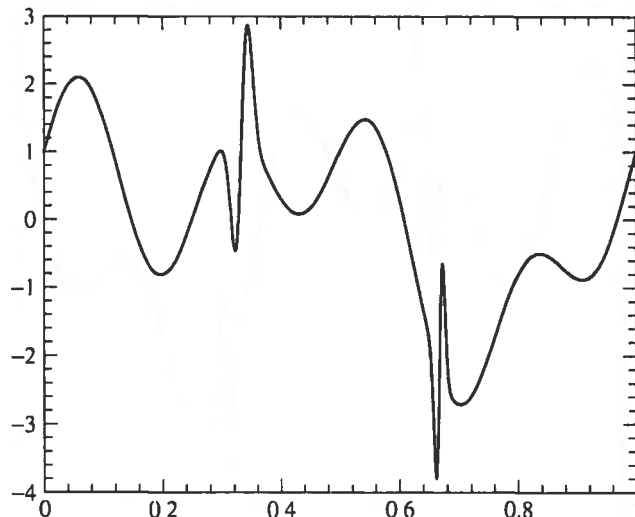
Figure 4.13. Sample signal

Example 4.13 Consider the signal, f, defined on the unit interval $0 \le x \le 1$ given in Figure 4.13. We discretize this signal over 2^8 nodes (so $a_k^8 = f(k/2^8)$) since a mesh size of width $1/2^8$ seems to capture the essential features of this signal. Thus

$$f_8(x) = \sum_{k=0}^{2^8-1} f(k/2^8)\phi(2^8 x - k)$$

approximates f well enough for our purposes. Using Theorem 4.12, we decompose f_8 into its components in V_j, for $j = 8, 7, \ldots, 0$. Plots of the components in $f_8 \in V_8$, $f_7 \in V_7$, $f_6 \in V_6$, and $f_4 \in V_4$ are given in Figures 4.14, 4.15, 4.16 and 4.17. A plot of the W_7 component is given in Figure 4.18. The W_7 component is small except where the original signal contains a sharp spike of width $\approx 2^{-8}$ (at $x \approx 0.3$ and $x \approx 0.65$).   ■

4.3.2 Reconstruction

Having decomposed f into its V_0 and $W_{j'}$ components for $0 \le j' < j$, then what? The answer depends on the goal. If the goal is to filter out noise, then the $W_{j'}$ components of f corresponding to the unwanted frequencies can be thrown out and the resulting signal will have much less noise. If the goal is data compression, the $W_{j'}$ components that are small can be thrown out, without appreciably changing the signal. Only the significant $W_{j'}$ components (the larger $b_k^{j'}$) need to be stored or transmitted and significant data compression can be achieved. Of course, what constitutes "small" depends on the tolerance for error for a particular application.
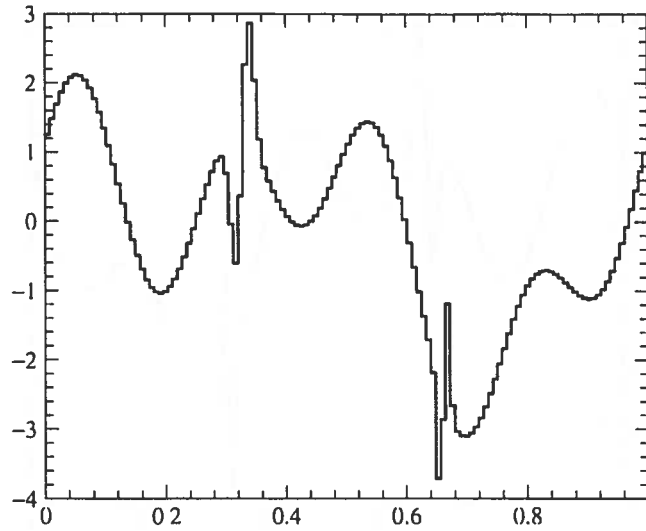
Figure 4.14. V_8 component.
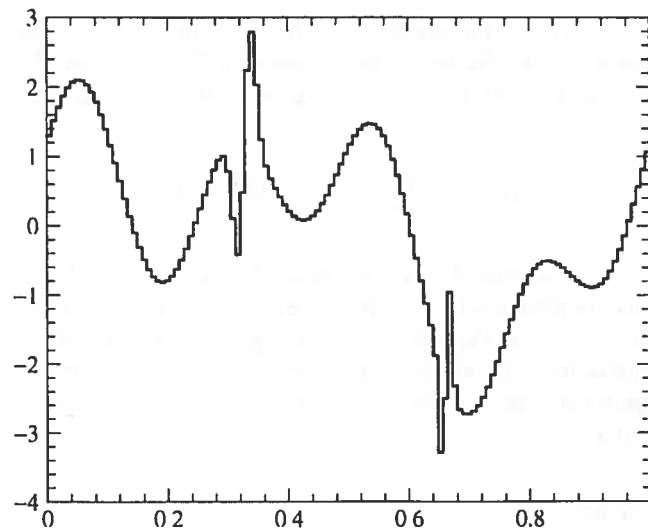
Figure 4.15. V_7 component.

In either case, since the $b_k^{j'}$ have been modified, we need a reconstruction algorithm (for the receiving end of the signal perhaps) so that the compressed or filtered signal can be rebuilt in terms of the basis elements $\phi(2^j x - l)$ of V_j; that is,

$$f(x) = \sum_{l \in Z} a_l^j \phi(2^j x - l).$$
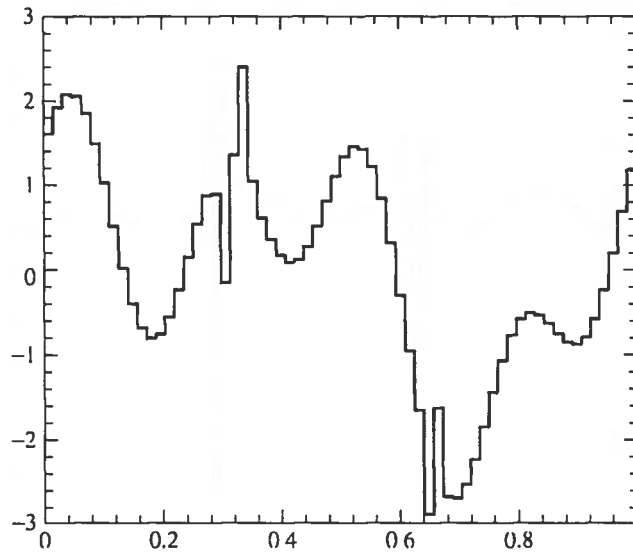
Figure 4.16. V_6 component.
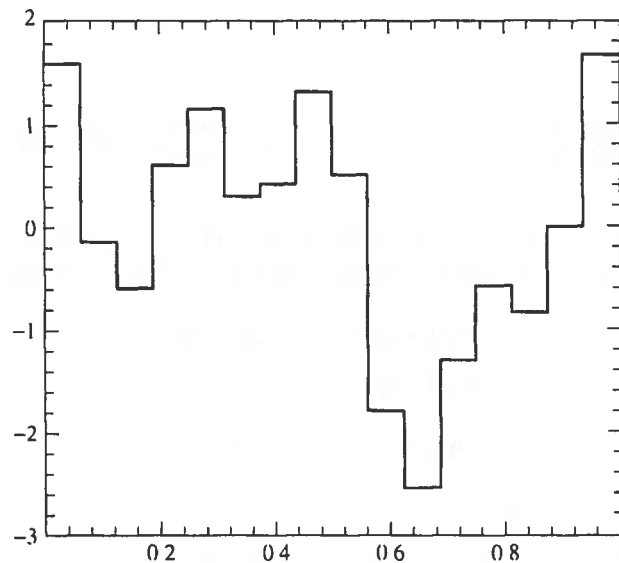
Figure 4.17. V_4 component.

Once this is done, the graph of the signal f is a step function of height a_l^j over the interval $l/2^j \le x \le (l+1)/2^j$.

We start with a signal of the form

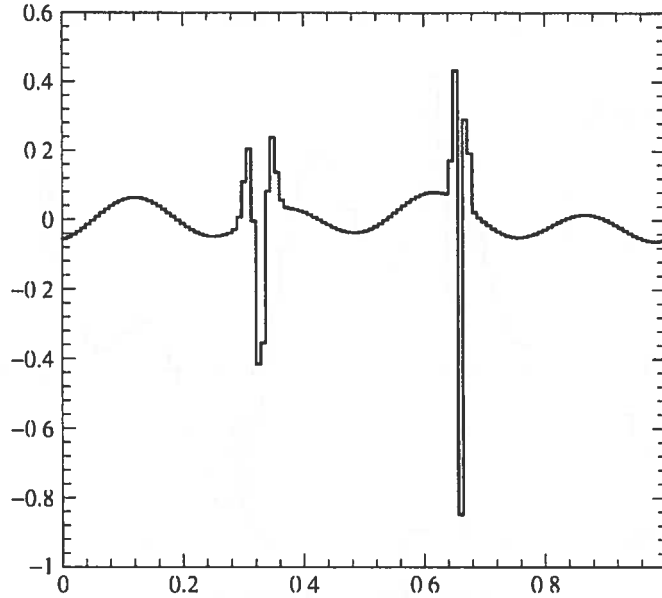$$f(x) = f_0(x) + w_0(x) + \cdots + w_{j-1}(x), \qquad w_l \in W_l,$$

Figure 4.18. W_7 component.

where

$$f_0(x) = \sum_{k \in \mathbb{Z}} a_k^0 \phi(x - k) \in V_0 \quad \text{and} \quad w_l = \sum_k b_k^l \psi(2^l x - k) \in W_l$$

for $0 \le l \le j - 1$. Our goal is to rewrite f as $f(x) = \sum_l a_l^j \phi(2^j x - l)$ and find an algorithm for the computation of the constants a_l^j. We use the equations

$$\phi(x) = \phi(2x) + \phi(2x - 1),\tag{4.12}$$

$$\psi(x) = \phi(2x) - \phi(2x - 1).\tag{4.13}$$

which follow from the definitions of ϕ and ψ. Replacing x by $2^{j-1}x$ yields

$$\phi(2^{j-1}x) = \phi(2^j x) + \phi(2^j x - 1),\tag{4.14}$$

$$\psi(2^{j-1}x) = \phi(2^j x) - \phi(2^j x - 1).\tag{4.15}$$

Using Eq. (4.12) with x replaced by $x - k$, we have

$$f_0(x) = \sum_{k \in \mathbb{Z}} a_k^0 \phi(x - k) \qquad \text{(the definition of } f_0)$$

$$= \sum_{k \in \mathbb{Z}} a_k^0 \phi(2x - 2k) + a_k^0 \phi(2x - 2k - 1) \qquad \text{from Eq (4.12).}$$

So

$$f_0(x) = \sum_{k \in \mathbb{Z}} \hat{a}_l^1 \phi(2x - l). \tag{4.16}$$

where

$$\hat{a}_l^1 = \begin{cases} a_k^0 & \text{if } l = 2k \text{ is even,} \\ a_k^0 & \text{if } l = 2k + 1 \text{ is odd.} \end{cases}$$

Similarly, $w_0 = \sum_k b_k^0 \psi(x - k)$ can be written [using Eq. (4.13) for $\psi(x - k)$] as

$$w_0(x) = \sum_{l \in \mathbb{Z}} \hat{b}_l^1 \phi(2x - l). \tag{4.17}$$

where

$$\hat{b}_l^1 = \begin{cases} b_k^0. & \text{if } l = 2k \text{ is even,} \\ -b_k^0. & \text{if } l = 2k + 1 \text{ is odd.} \end{cases}$$

Combining Eqs. (4.16) and (4.17) yields

$$f_0(x) + w_0(x) = \sum_{l \in \mathbb{Z}} a_l^1 \phi(2x - l),$$

where

$$a_l^1 = \hat{a}_l^1 + \hat{b}_l^1 = \begin{cases} a_k^0 + b_k^0 & \text{if } l = 2k, \\ a_k^0 - b_k^0 & \text{if } l = 2k + 1. \end{cases}$$

Next, $w_1 = \sum_k b_k^1 \psi(2x - k)$ is added to this sum in the same manner [using Eqs. (4.12) and (4.13) with x replaced by $2x - k$]:

$$f_0(x) + w_0(x) + w_1(x) = \sum_{l \in \mathbb{Z}} a_l^2 \phi(2^2 x - l),$$

where

$$a_l^2 = \begin{cases} a_k^1 + b_k^1 & \text{if } l = 2k, \\ a_k^1 - b_k^1. & \text{if } l = 2k + 1. \end{cases}$$

Note that the a_l^0 and b_l^0 coefficients determine the a_l^1 coefficients Then, the a_l^1 and b_l^1 coefficients determine the a_l^2 coefficients, and so on, in a recursive manner.

The preceding reconstruction algorithm is summarized in the following theorem.

Theorem 4.14 *(Haar Reconstruction)* If $f = f_0 + w_0 + w_1 + w_2 + \cdots + w_{j-1}$ with

$$f_0(x) = \sum_{k \in \mathbb{Z}} a_k^0 \phi(x - k) \in V_0 \quad \text{and} \quad w_{j'}(x) = \sum_{k \in \mathbb{Z}} b_k^{j'} \psi(2^{j'}x - k) \in W_{j'}$$

for $0 \le j' < j$, then

$$f(x) = \sum_{l \in \mathbb{Z}} a_l^j \phi(2^j x - l) \in V_j$$

where the $a_l^{j'}$ are determined recursively for $j' = 1$, then $j' = 2 \ldots$ until $j' = j$ by the algorithm

$$a_l^{j'} = \begin{cases} a_k^{j'-1} + b_k^{j'-1}, & \text{if } l = 2k \text{ is even,} \\ a_k^{j'-1} - b_k^{j'-1}, & \text{if } l = 2k + 1 \text{ is odd.} \end{cases}$$

Example 4.15 We apply the decomposition and reconstruction algorithms to compress the signal f that is shown in Figure 4.19; f is defined on the unit interval. (This is the same signal used in Example 4.13.)

We discretize this signal over 2^8 nodes (so $a_k^8 = f(k/2^8)$) and then decompose this signal (as in Theorem 4.12) to obtain $f = f_0 + w_0 + w_1 + w_2 + \cdots + w_7$
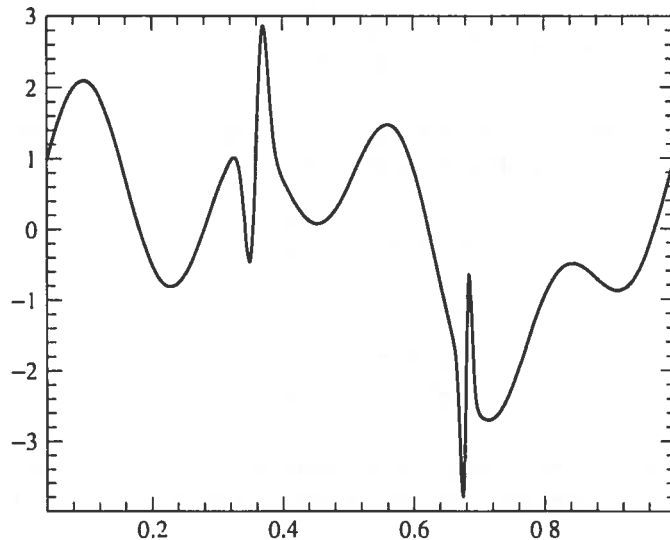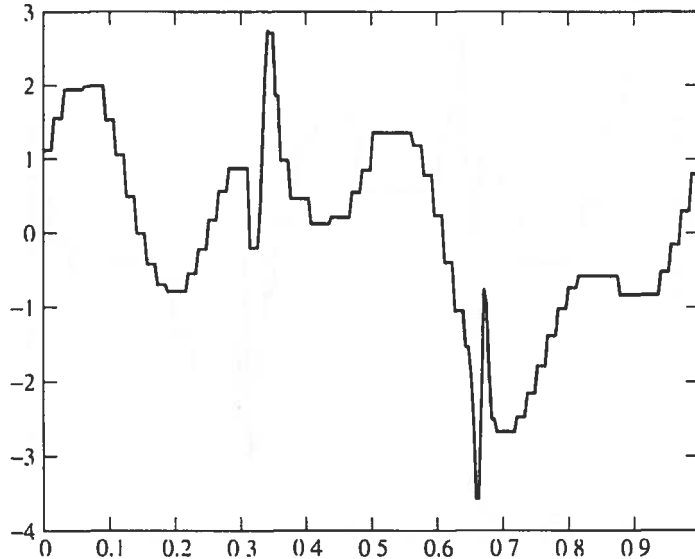

Figure 4.19. Sample signal

Figure 4.20. Graph showing 80% compression

with

$$f_0(x) = a_0^0 \phi(x) \in V_0 \quad \text{and} \quad w_{j'}(x) = \sum_{k \in Z} b_k^{j'} \psi(2^{j'} x - k) \in W_{j'}$$

for $0 \le j' \le 7$. Note that there is only one basis term $\phi(x)$ in V_0 since the interval of consideration is $0 \le x \le 1$. We first use 80% compression on this decomposed signal, which means that after ordering the $|b_k^j|$ by size, we set the smallest 80% equal to zero (retaining the largest 20%). Then we reconstruct the signal as in Theorem 4.14. The result is graphed in Figure 4.20. Figure 4.21 illustrates the same process with 90% compression. The relative L^2 error is 0.0895 for 80% compression and 0.1838 for 90% compression. ∎

4.3.3 Filters and Diagrams

The decomposition and reconstruction algorithms can be put in the language of discrete filters and simple operators acting on a sequence of coefficients. The algorithms can then be visualized in terms of block diagrams.

We will do the decomposition algorithm first. Define two discrete filters (convolution operators) H and L via their impulse responses, which are the sequences h and ℓ:

$$h = \left(\cdots 0 \cdots \underbrace{-\frac{1}{2} \quad \frac{1}{2}}_{k=-1 \quad 0} \cdot 0 \cdots \right), \qquad \ell = \left(\cdots 0 \cdots \underbrace{\frac{1}{2} \quad \frac{1}{2}}_{k=-1 \quad 0} \cdot 0 \cdot \right)$$
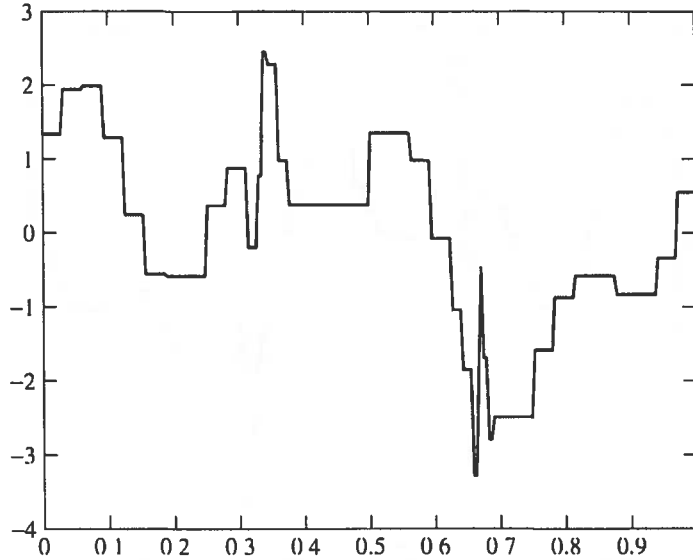
Figure 4.21. Graph showing 90% compression.

If $\{x_k\} \in \ell^2$, then $H(x) := h * x$ and $L(x) := \ell * x$. The resulting sequences are

$$H(x)_k = (h * x)_k = \frac{1}{2}x_k - \frac{1}{2}x_{k+1}, \qquad L(x)_k = (\ell * x)_k = \frac{1}{2}x_k + \frac{1}{2}x_{k+1}.$$

If we keep only even subscripts, then $H(x)_{2k} = (h * x)_{2k} = \frac{1}{2}x_{2k} - \frac{1}{2}x_{2k+1}$ and $L(x)_{2k} = (\ell * x)_{2k} = \frac{1}{2}x_{2k} + \frac{1}{2}x_{2k+1}$. This operation of discarding the odd coefficients in a sequence is called *down sampling*; we will denote the corresponding operator by D.

We now apply these ideas to go from level j scaling coefficients a_k^j to get the level $j - 1$ scaling and wavelet coefficients. Using Theorem 4.12 and replacing x by a_k^j,

$$b_k^{j-1} = DH(a^j)_k \quad \text{and} \quad a_k^{j-1} = DL(a^j)_k$$

Figure 4.22 illustrates the decomposition algorithm. The down-sampling operator D is replaced by the more suggestive symbol, "$2\downarrow$."

The reconstruction algorithm also requires defining two discrete filters $\widetilde{H}$ and $\widetilde{L}$ via their corresponding impulse responses,

$$\widetilde{h} = (\cdots 0 \cdots \underbrace{1 \; -1}_{k=0, \, 1} \cdots 0 \cdots), \qquad \widetilde{\ell} = (\cdots 0 \cdots \underbrace{1 \; \; 1}_{k=0, \, 1} \cdots 0 \cdots);$$

for a sequence $\{x_k\}$, we have $(\widetilde{h} * x)_k = x_k - x_{k-1}$ and $(\widetilde{\ell} * x)_k = x_k + x_{k-1}$. Here is an important observation. If x and y are sequences in which the *odd*
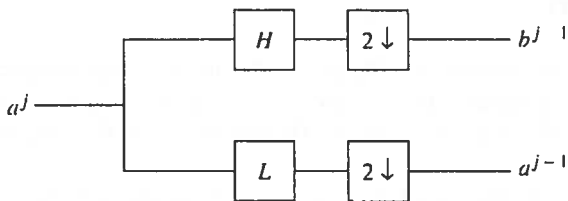
Figure 4.22. Haar decomposition diagram.

entries are all 0, then

$$(\widetilde{h} * x)_l = \begin{cases} x_{2k}, & l = 2k \text{ is even,} \\ -x_{2k}, & l = 2k + 1 \text{ is odd,} \end{cases} \qquad (\widetilde{\ell} * y)_l = \begin{cases} y_{2k}, & l = 2k \text{ is even,} \\ y_{2k}, & l = 2k + 1 \text{ is odd.} \end{cases}$$

Adding the two sequences $\widetilde{h} * x$ and $\widetilde{\ell} * y$ then gives us

$$(\widetilde{h} * x)_l + (\widetilde{\ell} * y)_l = \begin{cases} x_{2k} + y_{2k}, & l = 2k \text{ is even,} \\ y_{2k} - x_{2k}, & l = 2k + 1 \text{ is odd,} \end{cases}$$

This is almost the pattern for the reconstruction formula given in Theorem 4.14. Although the x_{2k+1}'s and y_{2k+1}'s are 0, the x_{2k}'s and y_{2k}'s are ours to choose, so we set $x_{2k} = b_k^{j-1}$ and $y_{2k} = a_k^{j-1}$; that is,

$$x = \left(\cdots 0 \quad b_{-1}^{j-1} \quad 0 \quad \underbrace{b_0^{j-1}}_{k=0} \quad 0 \quad b_1^{j-1} \quad 0 \quad b_2^{j-1} \quad 0 \cdots \right)$$

(all odd entries are zero) and similarly for y. The sequences x and y are called up samples of the sequences b^{j-1} and a^{j-1}. We use the U to denote the up-sampling operator, so $x = Ub^{j-1}$ and $y = Ua^{j-1}$. The reconstruction formula then takes the compact form

$$a^j = \widetilde{\ell} U a^{j-1} + \widetilde{H} U b^{j-1}$$

We illustrate the reconstruction step in Figure 4.23. The up-sampling operator is replaced by the symbol $2 \uparrow$.
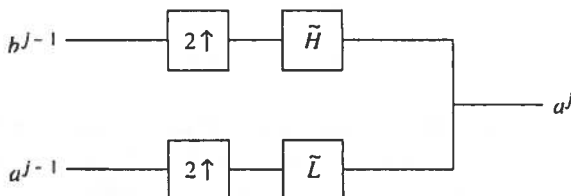

Figure 4.23. Haar reconstruction diagram

4.4 SUMMARY

In this section, we present a summary of the ideas of this chapter The format is a step-by-step procedure used to process (compress or denoise) a given signal $y = f(t)$. We will let ϕ and ψ be the Haar scaling function and wavelet.

Step 1. *Sample.* If the signal is continuous (analog), $y = f(t)$, where t represents time, choose the top level $j = J$ so that 2^J is larger than the Nyquist rate for the signal (see the discussion just before Theorem 2.23). Let

$$a_k^J = f(k/2^J).$$

In practice, the range of k is a finite interval determined by the duration of the signal. For example, if the duration of the signal is $0 \leq t \leq 1$, then the range for k is $0 \leq k \leq 2^J - 1$ (or perhaps $1 \leq l \leq 2^J$). If the signal is discrete to start with (i.e., a sequence of numbers), then this step is unnecessary. The top level a_k^J is set equal to the kth term in the sampled signal, and 2^J is taken to be the sampling rate. In any case, we have the highest-level approximation to f given by

$$f_J(x) = \sum_{k \in Z} a_k^J \phi(2^J x - k).$$

Step 2. *Decomposition.* The decomposition algorithm decomposes f_J into

$$f_J = w_{J-1} + \cdots + w_{j\ 1} + f_{j\ 1},$$

where

$$w_{j-1} = \sum_{l \in Z} b_l^{j-1} \psi(2^{j-1} x - l),$$

$$f_{j-1} = \sum_{l \in Z} a_l^{j-1} \phi(2^{j-1} x - l).$$

The coefficients, b_l^{j-1} and a_l^{j-1} are determined from the a^j recursively by the algorithm

$$a_l^{j-1} = DL(a^j)_k. \tag{4.18}$$

$$b_l^{j-1} = DH(a^j)_k. \tag{4.19}$$

where H and L are the high- and low-pass filters from Section 4.3 3. When $j = J$, a_k^{J-1} and $b_k^{J\ 1}$ are determined by a_k^J, which are the sampled signal values from Step 1. Then j becomes $J - 1$ and a_k^{J-2} and b_k^{J-2} are determined from a_k^{J-1}. Then j becomes $J - 2$, and so on, until

either the level reached is satisfactory for some purpose or there are too few coefficients to continue. Unless otherwise stated, the decomposition algorithm will continue until the $j = 0$ level is reached, that is,

$$f_J = w_{J-1} + \cdots + w_0 + f_0.$$

Step 3. *Processing.* After decomposition, the signal is now in the form

$$f_J(x) = \sum_{j=0}^{J-1} w_j + f_0 \tag{4.20}$$

$$= \sum_{j=0}^{J-1} \left(\sum_{k \in \mathbb{Z}} b_l^j \psi(2^j x - k) \right) + \sum_{k \in \mathbb{Z}} a_k^0 \phi(x - k). \tag{4.21}$$

The signal can now be filtered by modifying the wavelet coefficients b_k^j. How this is to be done depends on the problem at hand. To filter out all high frequencies, all the b_k^j would be set to zero for j above a certain value. Perhaps only a certain segment of the signal corresponding to particular values of k is to be filtered. If data compression is the goal, then the b_k^j that are below a certain threshold (in absolute value) would be set to zero. Whatever the goal, the process modifies the b_k^j.

Step 4. *Reconstruction.* Now the goal is to take the modified signal, f_J, in the form (4.21) (with the modified b_k^j) and reconstruct it as

$$f_J = \sum_{k \in Z} a_k^J \phi(2^J x - k).$$

This is accomplished by the reconstruction algorithm discussed in Section 4.3.3:

$$a^j = \tilde{L} U a^{j-1} + \tilde{H} U b^{j-1} \tag{4.22}$$

for $j = 1, \ldots, J$. When $j = 1$, the a_k^1 are computed from the a_k^0 and b_k^0. When $j = 2$, the a_k^2 are computed from the a_k^1 and b_k^1 and so forth. The range of k is determined by the time duration of the signal. When $j = J$ (the top level), a_k^J represents the approximate value of the processed signal at time $x = k/2^J$. Of course, these a_k^J are different from the original a_k^J due the modification of coefficients in Step 3.

EXERCISES

1. Let ϕ and ψ be the Haar scaling and wavelet functions, respectively Let V_j and W_j be the spaces generated by $\phi(2^j x - k)$, $k \in Z$, and $\psi(2^j x - k)$,

$k \in Z$, respectively. Consider the function defined on $0 \leq x < 1$ given by

$$f(x) = \begin{cases} -1, & 0 \leq x < 1/4, \\ 4, & 1/4 \leq x < 1/2, \\ 2, & 1/2 \leq x < 3/4, \\ -3, & 3/4 \leq x < 1. \end{cases}$$

Express f first in terms of the basis for V_2 and then decompose f into its component parts in W_1, W_0, and V_0. In other words, find the Haar wavelet decomposition for f. Sketch each of these components.

2. Repeat exercise 1 for the function

$$f(x) = \begin{cases} 2, & 0 \leq x < 1/4, \\ -3, & 1/4 \leq x < 1/2, \\ 1, & 1/2 \leq x < 3/4, \\ 3, & 3/4 \leq x < 1. \end{cases}$$

3. If A and B are finite-dimensional, orthogonal subspaces of an inner product space V, then show

$$\dim(A \oplus B) = \dim A + \dim B.$$

If A and B are not necessarily orthogonal, then what is the relationship between $\dim(A + B)$, $\dim A$, and $\dim B$?

4. (a) Let V_n be the spaces generated by $\phi(2^n x - k)$, $k \in Z$, where ϕ is the Haar scaling function. On the interval $0 \leq x < 1$, what are the dimensions of the spaces V_n and W_n for $n \geq 0$ (just count the number of basis elements)?

   (b) Using the result of problem 3, count the dimension of the space on the right side of the equality

$$V_n = W_{n-1} \oplus W_{n-2} \oplus \ldots \oplus W_0 \oplus V_0.$$

   Is your answer the same as the one you computed for $\dim V_n$ in part (a)?

5. Let ϕ and ψ be the Haar scaling and wavelet functions, respectively. Let V_j and W_j be the spaces generated by $\phi(2^j x - k)$, $k \in Z$, and $\psi(2^j x - k)$, $k \in Z$, respectively. Suppose $f(x) = \sum_k a_k \phi(2x - k)$ $(a_k \in R)$ belongs to V_1. Show explicitly that if f is orthogonal to each basis element $\phi(x - l) \in V_0$, for all integers l, then $a_{2l+1} = -a_{2l}$ for all integers l and hence show

$$f(x) = \sum_{l \in Z} a_{2l} \psi(x - l) \in W_0.$$

6. Reconstruct $g \in V_3$, given these coefficients in its Haar wavelet decomposition:

$$a^2 = [1/2, 2, 5/2, -3/2], \qquad b^2 = [-3/2, -1, 1/2, -1/2].$$

The first entry in each list corresponds to $k = 0$. Sketch g.

7. Reconstruct $h \in V_3$ over the interval $0 \leq x \leq 1$, given the following coefficients in its Haar wavelet decomposition:

$$a^1 = [3/2, -1], \quad b^1 = [-1, -3/2], \quad b^2 = [-3/2, -3/2, -1/2, -1/2].$$

The first entry in each list corresponds to $k = 0$. Sketch h.
The remaining problems require some programming in a language such as MATLAB, Maple, or C. The code in Appendix C may be useful.

8. (Haar wavelets on $[0, 1]$). Let n be a positive integer, and let f be a continuous function defined on $[0, 1]$. Let $h_k(t) = \sqrt{n}\phi(nt - k)$, where $\phi(t)$ is the Haar scaling function (which is 1 on the interval $[0, 1)$ and zero elsewhere). Form the L^2 projection of f onto the span of the h_k's,

$$f_n = \langle f, h_0 \rangle h_0 + \cdots + \langle f, h_{n-1} \rangle h_{n-1}$$

Show that f_n converges uniformly to f on $[0, 1]$. For $f(t) = 1 - t^2$, use MATLAB or Maple to find the Haar wavelet decomposition on $[0, 1]$ for $n = 4, 8$, and 16. Plot the results.

9. Let

$$f(t) = e^{-t^2/10} \left(\sin(2t) + 2\cos(4t) + 0.4\sin(t)\sin(50t) \right).$$

Discretize the function f over the interval $0 \leq t \leq 1$ as described in Step 1 of Section 4.4. Use $n = 8$ as the top level (so there are 2^8 nodes in the discretization). Implement the decomposition algorithm described in Step 2 of Section 4.4 using the Haar wavelets. Plot the resulting levels, $f_{j-1} \in V_{j-1}$ for $j = 8 \ldots 1$ and compare with the original signal.

10 (Continuation of exercise 9). Filter the wavelet coefficients computed in exercise 9 by setting to zero any wavelet coefficient whose absolute value is less than $tol = 0.1$. Then reconstruct the signal as described in Section 4.4. Plot the reconstructed f_8 and compare with the original signal. Compute the relative l^2 difference between the original signal and the compressed signal. Experiment with various tolerances. Keep track of the percentage of wavelet coefficients that have been filtered out.

11. Haar wavelets can be used to detect a discontinuity in a signal. Let $g(t)$ be
defined on $0 \leq t < 1$ via

$$
g(t) = \begin{cases} 0, & 0 \leq t < 7/17, \\ 1 - t^2, & 7/17 \leq t < 1. \end{cases}
$$

Discretize the function g over the interval $0 \leq t \leq 1$ as described in Step
1 of Section 4.4. Use $n = 7$ as the top level (so there are 2^7 nodes in the
discretization). Implement a 1-level decomposition. Plot the magnitudes of
the level 6 wavelet *coefficients*. Which wavelet has the largest coefficient?
What t corresponds to this wavelet? Try the method again with 7/17 replaced
by 8/9, and then by 2/7. Why do you think the method works?