Chapter 3

Frequency analysis

> If our ear uses a certain technique to analyze a signal, then if you use that same mathematical technique, you will be doing something like our ear. You might miss important things, but you would miss things that our ear would miss too.
>
> *Ingrid Daubechies*[1]

It is well known that human hearing responds to the frequencies of sound. Evolution has shaped our sense of hearing into a superb frequency analyzer, whose full characteristics we do not yet completely understand. Nevertheless, it is clear that we perceive tones of different frequency as having different pitch, and musical notes are called higher or lower depending on whether they have a corresponding higher or lower frequency.

The mathematical analysis of the frequency content of signals is called *Fourier analysis.* In this chapter we shall sketch the basic outlines of Fourier analysis as it applies to discrete signals and use it to analyze the frequency content of wavelets. A deeper understanding of wavelets can be gained from studying their frequency content, and by examining how this frequency content relates to wavelet transforms of signals. To keep the mathematics as simple as possible we shall focus on 1D signals, although in Section 3.5 we shall describe some 2D applications.

3.1 Discrete Fourier analysis

The frequency content of a discrete signal is revealed by its *discrete Fourier transform* (DFT). The DFT of a discrete signal is usually per-

[1]Daubechies' quote is from [BUR].

formed via a computer using an algorithm called a *fast Fourier transform* (FFT). No attempt will be made in this primer to prove any of the results about the DFT that we shall be using. The emphasis instead will be on how the DFT is applied in signal analysis, especially in wavelet analysis. There are many excellent treatments of the DFT and the FFT available for the reader who desires further discussion and proofs; we list a few of these references at the end of this chapter. In any case, the results we shall be using are all standard ones that are quite well known. Our approach will be to illustrate some fundamental ideas via a few examples rather than via a theoretical derivation.

The DFT reveals the frequency content of a discrete signal; so we shall begin by reviewing the notion of frequency. The analog signals $\cos 2\pi\nu x$ and $\sin 2\pi\nu x$, where x denotes time, both have a fundamental time period of $1/\nu$. Consequently, these signals repeat their basic form ν times in a unit-length time interval, i.e., they have a *frequency* of ν cycles/unit-time. The notion of frequency of a discrete signal is closely related to frequency of analog signals, as the following examples show. In the next section we shall make this connection more precise, but it may help to first examine some illustrative examples.

As a first example of a DFT, consider the discrete signal, Signal I, shown in Figure 3.1(a). Signal I consists of 1024 samples of the analog signal

$$g_1(x) = 2\cos 4\pi x + 0.5\sin 24\pi x. \tag{3.1}$$

It has frequencies of 2 and 12 for its two terms. In Figure 3.1(b) we show the DFT of Signal I. The two inner spikes are located at ± 2, corresponding to the frequency 2 of the cosine term of Signal I, and the two outer spikes are located at ± 12, corresponding to the frequency 12 of the sine term of Signal I. Because of the way in which the DFT is defined (see the next section), the spikes appear at $\pm\nu$ instead of just at ν when the signal contains a sine or cosine term of frequency ν. Notice that the lengths of the spikes in the DFT of Signal I that correspond to the cosine term in Signal I appear to be about 4 times greater than the lengths of the spikes that correspond to the sine term, and 4 is also the ratio of the amplitude of the cosine term to the amplitude of the sine term in Signal I. This example shows how the DFT can identify the frequencies—and the relative sizes of their contributions in terms of amplitude—that make up the frequency content of a signal.

As a second example of a DFT, consider the discrete signal, Signal II, shown in Figure 3.1(c). Signal II consists of 1024 samples of the analog signal

$$g_2(x) = \frac{1 + \cos 24\pi x}{1 + 4x^2}\,. \tag{3.2}$$

Its DFT is shown in Figure 3.1(d). The significant values in the DFT are clustered around 0 and ± 12. These values of ± 12 clearly correspond to

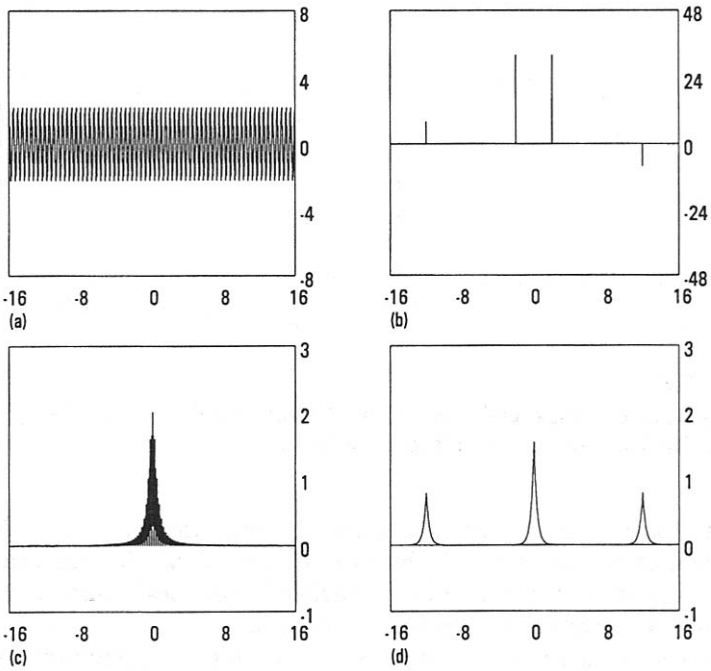

FIGURE 3.1
(a) Signal I. (b) DFT of Signal I. (c) Signal II. (d) DFT of Signal II.

the frequency 12 in the cosine term in Formula (3.2), but the cause of the cluster of significant DFT values around 0 is less clear. We shall see what these very low frequency values have to do with Signal II in Section 3.3, when we examine the frequency content of averaged signals in wavelet multiresolution analysis.

Frequency content of wavelets

As a third example of DFTs, we consider the frequency content of scaling signals and wavelets. This will enable us to comprehend the effects of wavelet MRAs on the frequency content of a signal, which we shall examine in Section 3.3.

As a typical case of scaling signals and wavelets, consider the Coif12 scaling signal $\mathbf{V}_1^1$ and wavelet $\mathbf{W}_1^1$. In Figure 3.2(a) we show the squares of the magnitudes of the DFT of the Coif12 scaling signal $\mathbf{V}_1^1$, which is called the *spectrum* of this signal. Notice that the significant, non-zero values of this spectrum are concentrated in the central half of the graph near the origin, which are the lower frequencies. The higher frequencies lie in the left and right quarters of the graph, and for these higher frequencies the

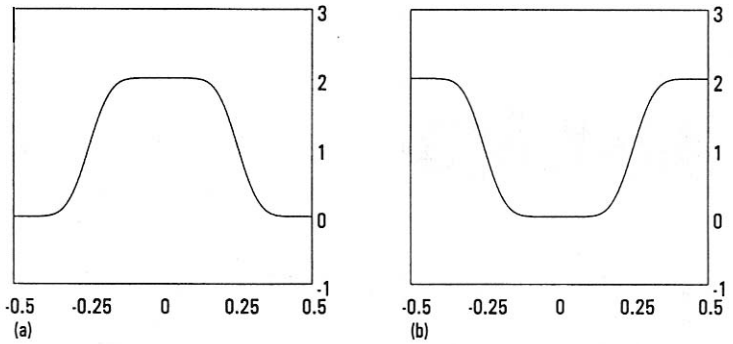

FIGURE 3.2
(a) Squares of magnitudes of DFT of Coif12 scaling signal $\mathbf{V}_1^1$, its spectrum. (b) Spectrum of Coif12 wavelet $\mathbf{W}_1^1$.

values of the spectrum are approximately zero. In contrast, consider the spectrum of the wavelet $\mathbf{W}_1^1$, shown in Figure 3.2(b). Its significant, nonzero values are concentrated in the higher frequencies, while for the lower frequencies its values are approximately zero.

Notice that the graphs in Figure 3.2 seem complementary to one another. In fact, it is the case that the sum of their values equals the constant 2. This is one of the fundamental properties of the spectrums of first-level scaling signals and wavelets.

3.2 Definition of the DFT and its properties

In this section, we shall state the definition of the DFT and some of its principal properties. We shall be fairly brief; more thorough discussions can be found in many references, some of which are listed at the end of this chapter.

To state the definition of the DFT in a succinct form, we shall make use of the Σ-notation for summation. A sum $g_1 + g_2 + \cdots + g_M$ can be more compactly expressed as

$$\sum_{m=1}^{M} g_m.$$

For instance, the following sum

$$a + ar + ar^2 + \cdots + ar^{M-1} = \sum_{m=1}^{M} ar^{m-1}$$

should be familiar from algebra.

Using this summation notation, we can state the definition of the DFT of a signal $\mathbf{f}$ of length N. We shall denote this DFT by $\mathcal{F}\mathbf{f}$, and its values $(\mathcal{F}\mathbf{f})_n$ are defined by

$$(\mathcal{F}\mathbf{f})_n = \sum_{m=1}^{N} f_m e^{-i2\pi(n-1)(m-1)/N}. \tag{3.3}$$

Although the variable n in (3.3) can be any integer, we shall see that the periodicity property stated below implies that the values $(\mathcal{F}\mathbf{f})_n$ for $n = -N/2$ to $n = N/2 - 1$ are sufficient for describing $\mathcal{F}\mathbf{f}$.

In Formula (3.3) we use the complex exponentials $e^{-i2\pi(n-1)(m-1)/N}$, which are defined via *Euler's formulas:*

$$e^{i\theta} = \cos\theta + i\sin\theta \tag{3.4}$$

and

$$e^{-i\theta} = \cos\theta - i\sin\theta. \tag{3.5}$$

One consequence of (3.5) is that the modulus (or magnitude), $\left|e^{-i\theta}\right|$, of the complex number $e^{-i\theta}$ is equal to 1. That is, $e^{-i\theta}$ lies on the unit-circle in the complex plane.

Another consequence of (3.4) and (3.5) is that

$$\cos 2\pi\nu x = \frac{1}{2}\, e^{-i2\pi\nu x} + \frac{1}{2}\, e^{i2\pi\nu x}. \tag{3.6}$$

This equation shows that $\cos 2\pi\nu x$ can be expressed as a sum of two complex exponentials having frequencies $\pm\nu$, which should remind the reader of the DFT of Signal I discussed in the previous section. Similarly,

$$\sin 2\pi\nu x = \frac{i}{2}\, e^{-i2\pi\nu x} - \frac{i}{2}\, e^{i2\pi\nu x} \tag{3.7}$$

which shows that $\sin 2\pi\nu x$ can also be expressed as a sum of complex exponentials having frequencies of $\pm\nu$.

The following four properties are the principal ones satisfied by the DFT. We use the notation $\mathbf{f} \overset{\mathcal{F}}{\longmapsto} \mathcal{F}\mathbf{f}$ to symbolize the DFT operation.

Properties of the DFT

1. *Linearity.* For all constants α and β, and all signals $\mathbf{f}$ and $\mathbf{g}$ of length N,

$$\alpha\mathbf{f} + \beta\mathbf{g} \overset{\mathcal{F}}{\longmapsto} \alpha\mathcal{F}\mathbf{f} + \beta\mathcal{F}\mathbf{g}.$$

2. *Periodicity.* If $\mathbf{f}$ is a signal of length N, then $\mathcal{F}\mathbf{f}$ has period N; that is,

$$(\mathcal{F}\mathbf{f})_{n+N} = (\mathcal{F}\mathbf{f})_n \tag{3.8}$$

holds for all integers n.

3. *Inversion.* The signal $\mathbf{f}$ can be obtained from $\mathcal{F}\mathbf{f}$ by

$$f_m = \frac{1}{N} \sum_{n=-N/2}^{N/2-1} (\mathcal{F}\mathbf{f})_n e^{i2\pi(n-1)(m-1)/N} \tag{3.9}$$

for $m = 1, 2, \ldots, N$.

4. *Parseval's Equality.* The signal $\mathbf{f}$ and its DFT $\mathcal{F}\mathbf{f}$ satisfy

$$\sum_{m=1}^{N} |f_m|^2 = \frac{1}{N} \sum_{n=-N/2}^{N/2-1} |(\mathcal{F}\mathbf{f})_n|^2 . \tag{3.10}$$

Because of periodicity, the N values $(\mathcal{F}\mathbf{f})_n$ for $n = -N/2$ to $n = N/2-1$ are sufficient for uniquely determining a DFT $\mathcal{F}\mathbf{f}$. It is for this reason that all of our figures of DFTs have the same number of values as the signal being transformed; and all software packages that compute DFTs follow this same convention.

The Inversion Formula (3.9) is particularly important. For one thing, it ensures that distinct signals must have distinct DFTs. For another, it allows for a useful interpretation of DFT values. Each DFT value $(\mathcal{F}\mathbf{f})_n$ is an amplitude for a discrete complex exponential signal

$$e^{i2\pi(n-1)(m-1)/N}, \quad m = 1, 2, \ldots, N, \tag{3.11}$$

which is a sampled version of $e^{i2\pi(n-1)x}$, a complex exponential analog signal of frequency $n-1$. The sample points are $x_m = (m-1)/N$. Or, since

$$\frac{(n-1)(m-1)}{N} = \frac{n-1}{\Omega} \cdot \frac{(m-1)\Omega}{N} ,$$

one can also view (3.11) as sample values of $e^{i2\pi(n-1)x/\Omega}$, a complex exponential of frequency $(n-1)/\Omega$. In this latter case, the sample points are $x_m = (m-1)\Omega/N$. This latter case is important when signals are obtained from measured samples of analog signals over a time interval of length Ω. In any case, the Inversion Formula (3.9) shows that the signal $\mathbf{f}$ can be realized by summing these discrete complex exponential signals with amplitudes given by the DFT values $(\mathcal{F}\mathbf{f})_n$, and multiplying by the scale factor $1/N$.

Parseval's Equality (3.10) can be regarded as a Conservation of Energy property, provided we include a scale factor of $1/N$ again. This Conservation of Energy, as with wavelet transforms, is useful for understanding how to make applications of the DFT operation. In fact, in some instances, the DFT, or transforms closely related to it, can be used for compression and noise removal. While these are fascinating topics, there is insufficient space

in this primer for discussing it any further. We do give some references, however, in the last section of this chapter.

Another way of interpreting Parseval's Equality is to observe that the left side of (3.10) is equal to the energy $\mathcal{E}_{\mathbf{f}}$ of the signal $\mathbf{f}$, and that the right side of (3.10) is equal to the average, or mean, value of the spectrum $|\mathcal{F}\mathbf{f}|^2$ of $\mathbf{f}$. That is, Parseval's Equality states that *the energy of a signal* $\mathbf{f}$ *is equal to the mean value of its spectrum* $|\mathcal{F}\mathbf{f}|^2$.

One final remark needs to be made about periodicity and inversion. They imply, since the right side of (3.9) also has period N in the integer variable m, that the finite signal $\mathbf{f}$ should be assumed to be a subsignal of a periodic signal having period N. That is,

$$f_{m+N} = f_m \tag{3.12}$$

for all integers m. Notice that (3.12) is the wrap-around property of signals that we made use of for scaling signals and wavelets in Chapter 2.

z-transforms *

The z-transform provides a more flexible way of expressing the values of the DFT, which is especially helpful in wavelet theory. Since some readers may find the theory of z-transforms to be difficult, we shall treat it as optional material. This material will be used later only in optional sections.

The z-transform of $\mathbf{f}$ will be denoted by $\mathbf{f}[z]$ and is defined by

$$\mathbf{f}[z] = \sum_{m=1}^{N} f_m z^{m-1}. \tag{3.13}$$

The variable z takes its values on the unit-circle of the complex plane.[2]

If we set z equal to $e^{-i2\pi(n-1)/N}$, then

$$\mathbf{f}[e^{-i2\pi(n-1)/N}] = (\mathcal{F}\mathbf{f})_n. \tag{3.14}$$

Formula (3.14) shows that the DFT, $\mathcal{F}\mathbf{f}$, consists of the values of $\mathbf{f}[z]$ at the points $z = e^{-i2\pi(n-1)/N}$ which lie uniformly spaced around the unit-circle in the complex plane.

One application of the z-transform is to the computation of DFTs of scaling signals and wavelets. To do this, we must first define the *cyclic translation* of a signal $\mathbf{f}$. The cyclic translation forward by 1 unit is denoted by $\mathcal{T}_1\mathbf{f}$ and is defined by

$$\mathcal{T}_1\mathbf{f} = (f_N, f_1, f_2, \ldots, f_{N-1}). \tag{3.15}$$

[2]There is another definition of z-transform that is also used in signal processing. The one we are working with is most useful for finite length signals.

Notice the wrap-around at the end of $\mathcal{T}_1\mathbf{f}$; because of this wrap-around we use the adjective *cyclic* in describing this translation. The cyclic translation backward by 1 unit is denoted by $\mathcal{T}_{-1}\mathbf{f}$ and is defined by

$$\mathcal{T}_{-1}\mathbf{f} = (f_2, f_3, \ldots, f_N, f_1). \tag{3.16}$$

All other cyclic translations are defined via compositions, e.g., $\mathcal{T}_2 = \mathcal{T}_1 \circ \mathcal{T}_1$, $\mathcal{T}_3 = \mathcal{T}_1 \circ \mathcal{T}_1 \circ \mathcal{T}_1$, and so on. The most important property of these cyclic translations is that $\mathcal{T}_k \circ \mathcal{T}_m = \mathcal{T}_{k+m}$ for all integers k and m.

The key property of the z-transform is that

$$\mathcal{T}_k\mathbf{f}[z] = \mathbf{f}[z]z^k \tag{3.17}$$

which holds *provided z is equal to $e^{-i2\pi(n-1)/N}$ for some integer n.* To see why (3.17) holds, we demonstrate it for $k = 1$, and higher powers of k follow by repeating the argument for $k = 1$. If $k = 1$, then

$$\mathcal{T}_1\mathbf{f}[z] = f_N + f_1 z + \cdots + f_{N-1}z^{N-1}.$$

But, if $z = e^{-i2\pi(n-1)/N}$, then $z^N = 1$. Consequently

$$\begin{aligned}\mathcal{T}_1\mathbf{f}[z] &= f_1 z + \cdots + f_{N-1}z^{N-1} + f_N z^N \\ &= (f_1 + f_2 z + \cdots + f_N z^{N-1})z \\ &= \mathbf{f}[z]z\end{aligned}$$

which proves (3.17) for $k = 1$.

One application of (3.17) is to the relationship between the frequency content of $\mathbf{V}_1^1$ and all the other first-level scaling signals. Since $\mathbf{V}_m^1 = \mathcal{T}_{2m-2}\mathbf{V}_1^1$, it follows from (3.17) that

$$\mathbf{V}_m^1[z] = \mathbf{V}_1^1[z]z^{2m-2}. \tag{3.18}$$

Similarly, we have

$$\mathbf{W}_m^1[z] = \mathbf{W}_1^1[z]z^{2m-2} \tag{3.19}$$

for the first-level wavelets. These last two formulas elucidate the relationship between the frequency contents of scaling signals and wavelets because of Formula (3.14), which relates z-transforms to DFTs. For example, because $|z| = 1$, it follows that

$$\left|\mathcal{F}\mathbf{V}_m^1\right|^2 = \left|\mathcal{F}\mathbf{V}_1^1\right|^2, \quad \left|\mathcal{F}\mathbf{W}_m^1\right|^2 = \left|\mathcal{F}\mathbf{W}_1^1\right|^2$$

which shows that the spectrum of each first-level scaling signal is equal to the spectrum of $\mathbf{V}_1^1$ and the spectrum of each first-level wavelet is equal to the spectrum of $\mathbf{W}_1^1$. We shall make further use of Formulas (3.18) and (3.19) in the next section.

3.3 Frequency description of wavelet analysis

In this section we shall examine how the frequency content of a signal is changed when the signal undergoes a wavelet analysis. To be more precise, we shall compare the frequency contents of the averaged signals $\mathbf{A}^k$ and the detail signals $\mathbf{D}^k$ created in a k-level MRA with the frequency content of the original signal $\mathbf{f}$.

Let's examine a 1-level Coif12 MRA of Signal II, shown in Figure 3.1(c). This example illustrates the fundamental aspects of the effect on frequency content of an MRA using a Daubechies wavelet. At the end of this section we shall discuss the mathematical details for those readers who are interested in them.

When a first averaged signal $\mathbf{A}^1$ is created, it consists of a sum of multiples of the first level scaling signals [see Formula (2.13b)]. In Figure 3.3(a) we show the first averaged signal, using Coif12 scaling signals, for Signal II. The DFT of this first averaged signal is shown in Figure 3.3(b). It is interesting to compare this DFT with the DFT of Signal II in Figure 3.1(d) and the spectrum of the Coif12 scaling signal $\mathbf{V}_1^1$ in Figure 3.2(a). In order to make this comparison, the values of the spectrum must be graphed over $[-16, 16]$ instead of $[-0.5, 0.5]$, but this can be easily done by a change of scale of the horizontal (frequency) axis. The values of this spectrum are approximately equal to the constant 2 near the origin (for the lower frequencies), and are approximately equal to the constant 0 at the left and right (for the higher frequencies).

The relation between the DFT of $\mathbf{A}^1$ and the spectrum $\left|\mathcal{F}\mathbf{V}_1^1\right|^2$ of $\mathbf{V}_1^1$ is

$$\mathcal{F}\mathbf{A}^1 \approx \frac{1}{2}\left|\mathcal{F}\mathbf{V}_1^1\right|^2 \mathcal{F}\mathbf{f}. \tag{3.20}$$

On the right side of (3.20) is the product of the two signals $\left|\mathcal{F}\mathbf{V}_1^1\right|^2$ and $\mathcal{F}\mathbf{f}$; hence this approximation says that each value $(\mathcal{F}\mathbf{A}^1)_n$ of the DFT of $\mathbf{A}^1$ is approximately equal to 1/2 times $\left|(\mathcal{F}\mathbf{V}_1^1)_n\right|^2 (\mathcal{F}\mathbf{f})_n$. Thus 1/2 times the spectrum of $\mathbf{V}_1^1$ acts as a *low-pass filter* on the values of the DFT of $\mathbf{f}$, allowing through only the low frequency values (since $\left|\mathcal{F}\mathbf{V}_1^1\right|^2 \approx 0$ for high frequency values).

It is interesting to examine the way in which this low-pass filtering affects the averaged signal. By comparing the averaged signal $\mathbf{A}^1$ in Figure 3.3(a) with the original signal $\mathbf{f}$ in Figure 3.1(c) we can see that there is much less rapid oscillation in the averaged signal; this is due to the suppression of the high frequencies by the low-pass filtering.

In contrast to the scaling signal, 1/2 times the spectrum of the Coif12 wavelet $\mathbf{W}_1^1$ acts as a *high-pass filter,* allowing through only the high frequency portions of the DFT of $\mathbf{f}$. In fact, the following approximation

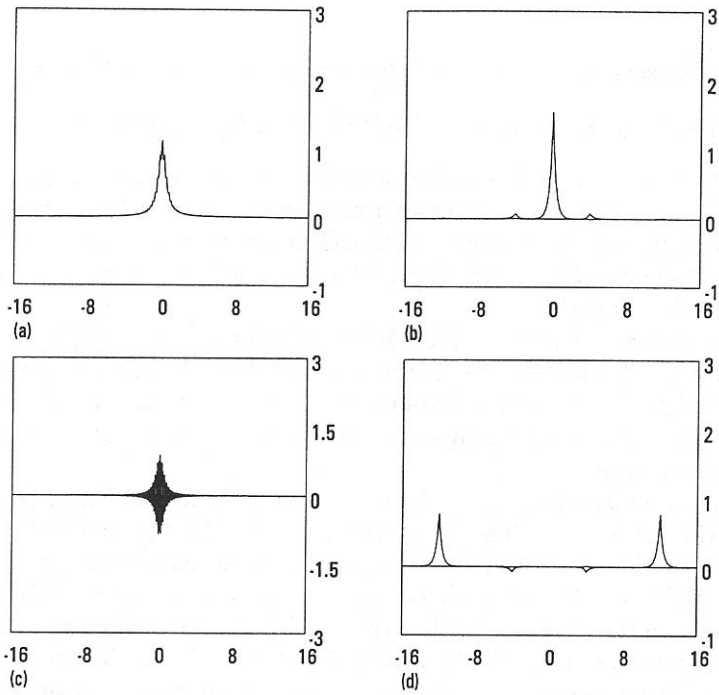

FIGURE 3.3
Frequency decomposition of 1-level Coif12 MRA. (a) First averaged signal for Signal II. (b) DFT of first averaged signal [c.f., Figures 3.1(d) and 3.2(a)]. (c) First detail signal for Signal II. (d) DFT of first detail signal [c.f., Figures 3.1(d) and 3.2(b)].

holds

$$\mathcal{F}\mathbf{D}^1 \approx \frac{1}{2}\left|\mathcal{F}\mathbf{W}_1^1\right|^2 \mathcal{F}\mathbf{f}. \tag{3.21}$$

As can be seen from Figure 3.2(a), the factor $\left|\mathcal{F}\mathbf{W}_1^1\right|^2$ is approximately zero for the low frequency values, and is approximately 2 for the high frequency values. It then follows from (3.21)—and we can check it by examining Figure 3.3(d)—that the DFT of the first detail signal mostly consists of the higher frequency values of the DFT of the signal $\mathbf{f}$. The effect that this has on the detail signal $\mathbf{D}^1$ is that it contains the most rapid, high frequency oscillations from the original signal.

This example of a 1-level Coif12 MRA of Signal II is typical of all wavelet MRAs. The frequency content of the first averaged signal $\mathbf{A}^1$ consists of low frequency values resulting from a low-pass filtering of the frequency content of the signal by 1/2 times the spectrum of the 1-level scaling signal $\mathbf{V}_1^1$. And the first-level detail signal $\mathbf{D}^1$ has a frequency content obtained from a high-pass filtering of the frequency content of the signal by 1/2 times

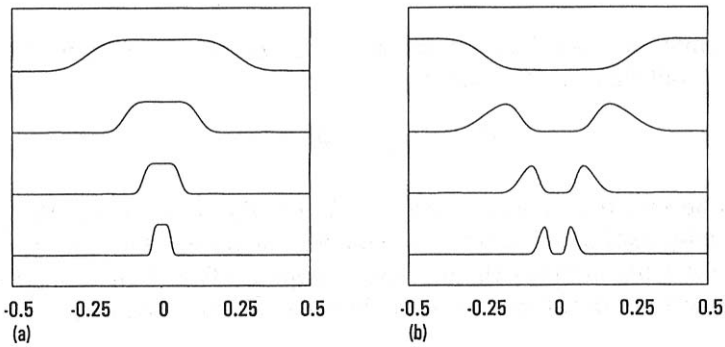

FIGURE 3.4
(a) Spectra of Coif12 scaling signals, from $k = 1$ at top to $k = 4$ at bottom. (b) Spectra of Coif12 wavelets from $k = 1$ at top to $k = 4$ at bottom. The spectra have been multiplied by constants 2^{-k} for each k in order to make their graphs of similar size.

the spectrum of the 1-level wavelet $\mathbf{W}_1^1$.

As with the first level, the DFTs of higher level averaged signals and fluctuation signals can be obtained by multiplying the signal's DFT by the spectra of higher level scaling signals and wavelets. For example, the DFT of the second averaged signal $\mathbf{A}^2$ satisfies

$$\mathcal{F}\mathbf{A}^2 \approx \frac{1}{4}\left|\mathcal{F}\mathbf{V}_1^2\right|^2 \mathcal{F}\mathbf{f} \tag{3.22}$$

and the DFT of the second detail signal $\mathbf{D}^2$ satisfies

$$\mathcal{F}\mathbf{D}^2 \approx \frac{1}{4}\left|\mathcal{F}\mathbf{W}_1^2\right|^2 \mathcal{F}\mathbf{f}. \tag{3.23}$$

In Formula (3.22), 1/4 times the spectrum $\left|\mathcal{F}\mathbf{V}_1^2\right|^2$ of the 2-level scaling signal $\mathbf{V}_1^2$ acts as a low-pass filter. This low-pass filter is graphed in Figure 3.4(a) as the second graph from the top. Notice that this low-pass filter allows through only much lower frequencies than the first-level filtering due to $\left|\mathcal{F}\mathbf{V}_1^1\right|^2$. The spectrum $\left|\mathcal{F}\mathbf{W}_1^2\right|^2$ in Formula (3.23), however, does not act as a high-pass filter. As can be seen from the second graph at the top of Figure 3.4(b), 1/4 times $\left|\mathcal{F}\mathbf{W}_1^2\right|^2$ acts as a *band-pass* filter, in the sense that only two small isolated bands of frequency values are non-zero for this spectrum.

Higher level MRAs follow this same pattern. The k-level averaged signal $\mathbf{A}^k$ has a DFT which satisfies

$$\mathcal{F}\mathbf{A}^k \approx \frac{1}{2^k}\left|\mathcal{F}\mathbf{V}_1^k\right|^2 \mathcal{F}\mathbf{f}. \tag{3.24}$$

And the signal $\mathbf{L}_k = 2^{-k}\left|\mathcal{F}\mathbf{V}_1^k\right|^2$ acts as a low-pass filter. In Figure 3.4(a) we show these low-pass filters $\mathbf{L}_k$ for $k = 3$ and $k = 4$. Notice that only

a very small interval of values around the origin are non-zero for $\mathbf{L}_4$. The k-level detail signal's DFT satisfies

$$\mathcal{F}\mathbf{D}^k \approx \frac{1}{2^k} \left|\mathcal{F}\mathbf{W}_1^k\right|^2 \mathcal{F}\mathbf{f}. \tag{3.25}$$

As can be seen from Figure 3.4(b), for $k > 1$, the signal $2^{-k} \left|\mathcal{F}\mathbf{W}_1^k\right|^2$ acts as a band-pass filter. The bands of non-zero values for this filter consist of the values lying between the downward sloping left and right sides of the central portions of the graphs of the low-pass filters $\mathbf{L}_k$ and $\mathbf{L}_{k-1}$.

Low-pass and high-pass filtering *

In this optional subsection we shall discuss the mathematical formulation of the low-pass and high-pass filtering interpretation of wavelet MRAs. Our tool will be the z-transform described at the end of Section 3.2.

From Formula (2.13b), we have

$$\mathbf{A}^1 = a_1\mathbf{V}_1^1 + a_2\mathbf{V}_2^1 + \cdots + a_{N/2}\mathbf{V}_{N/2}^1 \tag{3.26}$$

where $(a_1, a_2, \ldots, a_{N/2})$ is the first trend subsignal $\mathbf{a}^1$. Formula (3.14) tells us that to obtain the DFT, $\mathcal{F}\mathbf{A}^1$, of $\mathbf{A}^1$ it suffices to obtain its z-transform $\mathbf{A}^1[z]$. By Formulas (3.26) and (3.18) we have

$$\begin{aligned}\mathbf{A}^1[z] &= a_1\mathbf{V}_1^1[z] + a_2\mathbf{V}_2^1[z] + \cdots + a_{N/2}\mathbf{V}_{N/2}^1[z] \\ &= a_1\mathbf{V}_1^1[z] + a_2\mathbf{V}_1^1[z]z^2 + \cdots + a_{N/2}\mathbf{V}_1^1[z]z^{N-2}.\end{aligned}$$

Thus $\mathbf{A}^1[z]$ satisfies

$$\mathbf{A}^1[z] = \mathbf{V}_1^1[z]\left(a_1 + a_2z^2 + \cdots + a_{N/2}z^{N-2}\right). \tag{3.27}$$

Because of Formula (3.27) we see that the DFT of $\mathbf{A}^1$ is obtained as a product of the DFT of $\mathbf{V}_1^1$ with the DFT obtained from the polynomial

$$a_1 + a_2z^2 + \cdots + a_{N/2}z^{N-2} = \mathbf{a}^1[z^2].$$

It now remains to examine the connection between the polynomial $\mathbf{a}^1[z^2]$ and the z-transform of the signal $\mathbf{f}$.

The polynomial $\mathbf{a}^1[z^2]$ satisfies

$$\begin{aligned}\mathbf{a}^1[z^2] &= a_1 + a_2z^2 + \cdots + a_{N/2}z^{N-2} \\ &= (\mathbf{f}\cdot\mathbf{V}_1^1) + (\mathbf{f}\cdot\mathbf{V}_2^1)z^2 + \cdots + (\mathbf{f}\cdot\mathbf{V}_{N/2}^1)z^{N-2} \\ &= \sum_{m=1}^{N/2} (\mathbf{f}\cdot\mathbf{V}_m^1)z^{2m-2} \\ &= \sum_{m=1}^{N/2} (\mathbf{f}\cdot\mathcal{T}_{2m-2}\mathbf{V}_1^1)z^{2m-2}.\end{aligned} \tag{3.28}$$

This last polynomial in z consists of the even powered terms from the polynomial on the left side of the following identity:

$$\sum_{k=1}^{N}(\mathbf{f}\cdot\mathcal{T}_{k-1}\mathbf{V}_1^1)z^{k-1}=\mathbf{f}[z]\mathbf{V}_1^1[z^{-1}]. \tag{3.29}$$

The fact that (3.29) holds is a consequence of the definition of multiplication of polynomials; we leave its proof as an exercise for the reader.

The sum of even powered terms can be extracted from the polynomial on the right side of (3.29) by the following identity:

$$\mathbf{f}[z]\mathbf{V}_1^1[z^{-1}]+\mathbf{f}[-z]\mathbf{V}_1^1[-z^{-1}]=2\sum_{m=1}^{N/2}(\mathbf{f}\cdot\mathcal{T}_{2m-2}\mathbf{V}_1^1)z^{2m-2}. \tag{3.30}$$

Combining (3.30) with (3.28) yields

$$\mathbf{a}^1[z^2]=\frac{1}{2}\mathbf{f}[z]\mathbf{V}_1^1[z^{-1}]+\frac{1}{2}\mathbf{f}[-z]\mathbf{V}_1^1[-z^{-1}]. \tag{3.31}$$

Formula (3.31) is the desired relation between the polynomial $\mathbf{a}^1[z^2]$ and the z-transform of $\mathbf{f}$. Combining it with (3.27) yields

$$\mathbf{A}^1[z]=\frac{1}{2}\mathbf{f}[z]\mathbf{V}_1^1[z]\mathbf{V}_1^1[z^{-1}]+\frac{1}{2}\mathbf{f}[-z]\mathbf{V}_1^1[z]\mathbf{V}_1^1[-z^{-1}]. \tag{3.32}$$

In order to interpret Equation (3.32) correctly, we need to understand the effects of substituting z^{-1} and $-z$ into z-transforms. Substituting z^{-1} into the z-transform $\mathbf{f}[z]$ yields

$$\mathbf{f}[z^{-1}]=\sum_{m=1}^{N}f_m(z^{-1})^{m-1}.$$

We shall use the overbar notation, $\overline{\ }$, to denote the conjugation of complex numbers. Since z is on the unit-circle, we have $z^{-1}=\overline{z}$. And, because each value f_m of $\mathbf{f}$ is a real number, we have $f_m=\overline{f_m}$. Therefore

$$\begin{aligned}\mathbf{f}[z^{-1}]&=\sum_{m=1}^{N}\overline{f_m}\,\overline{z^{m-1}}\\&=\overline{\mathbf{f}[z]}.\end{aligned}$$

Thus, for example, $\mathbf{V}_1^1[z]\mathbf{V}_1^1[z^{-1}]=\left|\mathbf{V}_1^1[z]\right|^2$.

To interpret the substitution of $-z$, we observe that $-z$ is the reflection of z through the origin of the complex plane. Consequently, the low frequency values that lie near $e^{i0}=1$ are reflected into high frequency values

that lie near $e^{\pm i\pi} = -1$, and vice versa. For example, in Figure 3.2, the spectrum $\left|\mathbf{V}_1^1[-z]\right|^2$ has a graph that is similar to the graph of the spectrum $\left|\mathbf{W}_1^1[z]\right|^2$ shown in Figure 3.2(b). The graph of $\left|\mathbf{V}_1^1[-z^{-1}]\right|^2$ is identical to the graph of $\left|\mathbf{V}_1^1[-z]\right|^2$, because the substitution of z^{-1} produces a complex conjugate which is then eliminated by the modulus-square operation. These considerations show that for the Coif12 scaling function, we have

$$\left|\mathbf{V}_1^1[z]\right| \left|\mathbf{V}^1[-z^{-1}]\right| \approx 0 \tag{3.33}$$

except for two small intervals of values of z centered on $e^{\pm i\pi/2}$. The approximation in (3.33) is true for all of the Daubechies scaling functions.

Based on (3.32) and (3.33) we have the following approximation:

$$\mathbf{A}^1[z] \approx \frac{1}{2}\left|\mathbf{V}_1^1[z]\right|^2 \mathbf{f}[z]. \tag{3.34}$$

Using the connection between DFTs and z-transforms, the approximation (3.20) follows from the approximation (3.34).

Similar calculations for the first-level detail signal $\mathbf{D}^1$ and fluctuation subsignal $\mathbf{d}^1$ yield

$$\mathbf{D}^1[z] = \frac{1}{2}\mathbf{f}[z]\mathbf{W}_1^1[z]\mathbf{W}_1^1[z^{-1}] + \frac{1}{2}\mathbf{f}[-z]\mathbf{W}_1^1[z]\mathbf{W}_1^1[-z^{-1}] \tag{3.35}$$

and

$$\mathbf{d}^1[z] = \frac{1}{2}\mathbf{f}[z]\mathbf{W}_1^1[z^{-1}] + \frac{1}{2}\mathbf{f}[-z]\mathbf{W}_1^1[-z^{-1}]. \tag{3.36}$$

And we also have the approximation

$$\mathbf{D}^1[z] \approx \frac{1}{2}\left|\mathbf{W}_1^1[z]\right|^2 \mathbf{f}[z], \tag{3.37}$$

which implies the approximation (3.21).

The other approximations, (3.22) through (3.25), can be proved in the same way as (3.20) and (3.21).

3.4 Correlation and feature detection

In this section we shall describe a standard method for detecting a short-lived feature within a more complicated signal. This method, known as correlation, is a fundamental part of Fourier analysis. We shall also describe some of the ways in which wavelet analysis can be used to enhance the basic correlation method for feature detection.

Let's begin by examining feature detection for 1D signals. Feature detection is important in seismology, where there is a need to identify characteristic features that indicate, say, earthquake tremors within a long seismological signal. Or, in an electrocardiogram (ECG), it might be necessary to identify portions of the ECG that indicate an abnormal heartbeat.

At the top of Figure 3.5(a) we show a simulated ECG, which we shall refer to as Signal C. The feature that we wish to locate within Signal C is shown in the middle of Figure 3.5(a); this feature is meant to simulate an abnormal heartbeat. It is, of course, easy for us to visually locate the abnormal heartbeat within Signal C, but that is a far cry from an algorithm that a computer could use for automatic detection.

As noted above, the standard method used for feature detection is correlation. The *correlation* of a signal $\mathbf{f}$ with a signal $\mathbf{g}$, both having lengths of N values, will be denoted by $(\mathbf{f} : \mathbf{g})$. It is also a signal of length N, and its k^{th} value $(\mathbf{f} : \mathbf{g})_k$ is defined by

$$(\mathbf{f} : \mathbf{g})_k = f_1 g_k + f_2 g_{k+1} + \cdots + f_N g_{k+N-1}. \tag{3.38}$$

In order for the sum in (3.38) to make sense, the signal $\mathbf{g}$ needs to be periodically extended, i.e., we assume that $g_{k+N} = g_k$ for each k. When computing the correlation $(\mathbf{f} : \mathbf{g})$, the signal $\mathbf{f}$ is the feature that we wish to detect within the signal $\mathbf{g}$. Usually the signal $\mathbf{f}$ is similar to the abnormal heartbeat signal shown in the middle of Figure 3.5(a), in the sense that the values of $\mathbf{f}$ are 0 except near the central portion of the signal. This reduces the distortion that results from assuming that $\mathbf{g}$ is periodic. We will show later how correlations are related to Fourier analysis, but first we shall describe their use in feature detection.

The rationale behind using the correlation $(\mathbf{f} : \mathbf{g})$ to detect the location of $\mathbf{f}$ within $\mathbf{g}$ is the following. If a portion of $\mathbf{g}$ matches the form of the central portion of $\mathbf{f}$—where the significant, non-zero values are concentrated—then, for a certain value of k, the terms in (3.38) will all be squares. This produces a positive sum which is generally larger than the sums for the other values of $(\mathbf{f} : \mathbf{g})$. In order to normalize this largest value so that it equals 1, we shall divide the values of $(\mathbf{f} : \mathbf{g})$ by the energy of $\mathbf{f}$. That is, we denote the *normalized correlation* of $\mathbf{f}$ with $\mathbf{g}$ by $\langle \mathbf{f} : \mathbf{g} \rangle$, and the k^{th} value of $\langle \mathbf{f} : \mathbf{g} \rangle$ is

$$\langle \mathbf{f} : \mathbf{g} \rangle_k = \frac{f_1 g_k + f_2 g_{k+1} + \cdots + f_N g_{k+N-1}}{\mathcal{E}_{\mathbf{f}}}. \tag{3.39}$$

At the end of this section we shall discuss why, under the right conditions, the maximum value for $\langle \mathbf{f} : \mathbf{g} \rangle$ is approximately 1.

As an example of these ideas, we show at the bottom of Figure 3.5(a) the graph of the squares of the positive values of the normalized correlation $\langle \mathbf{f} : \mathbf{g} \rangle$ for the abnormal heartbeat and Signal C. Notice how *the maximum for this graph clearly locates the position of the heartbeat within Signal C.*

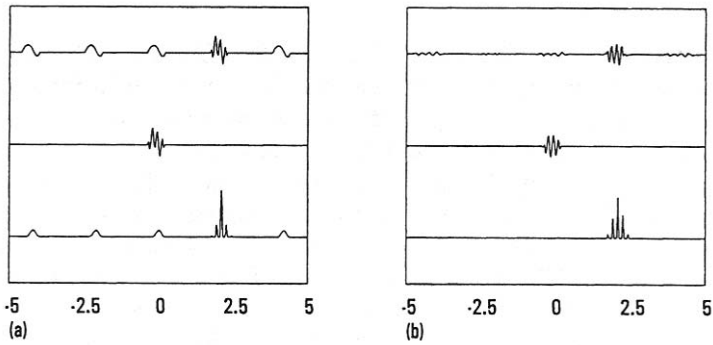

FIGURE 3.5
(a) Top: Signal C. Middle: Abnormal heartbeat. Bottom: Squares of positive values of normalized correlation. (b) Top: Signal C minus its fourth Coif30 averaged signal. Middle: Abnormal heartbeat minus its fourth Coif30 averaged signal. Bottom: Squares of positive values of normalized correlation.

The value of this maximum is 1, thus providing the following simple criterion for locating an abnormal heartbeat: *if a normalized correlation value is near* 1, *then an abnormal heartbeat is probably present at the location of this value.* We have ignored the negative values of the normalized correlation because a negative value of $\langle \mathbf{f} : \mathbf{g} \rangle$ indicates a preponderance of oppositely signed values, which is a clear indication that the values of $\mathbf{f}$ and $\mathbf{g}$ are not matched. The squaring serves to emphasize the maximum value near 1. It is not necessary, but produces a more easily interpretable graph—a graph for which the maximum value near 1 more clearly stands out from smaller values.

Notice that there are smaller peaks in the bottom graph of Figure 3.5(a) that mark the locations of the normal heartbeats in Signal C. These smaller peaks are present because the abnormal heartbeat was created by forming a sum of a normal heartbeat plus a high frequency "noise" term. Consequently, these peaks reflect a partial correspondence between the normal heartbeat term and each of the normal heartbeats in Signal C.

We shall now describe a wavelet based method for suppressing these peaks in the detection signal resulting from the normal heartbeats. While this may not be necessary for the case of Signal C, it might be desired for other signals. Furthermore, the method is a simpler 1D version of the 2D method that we shall describe in the next section.

Each normal heartbeat in Signal C has a spectrum that has significant values only for very low frequencies in comparison to the high frequency oscillations that are clearly visible in the abnormal heartbeat. Our method consists, therefore, of subtracting away an averaged signal from Signal C. As we saw in the previous section, the higher the level k of the averaged signal

$\mathbf{A}^k$, the nearer to zero are the low frequencies which make up the non-zero values of the averaged signal's spectrum $\left|\mathcal{F}\mathbf{A}^k\right|^2$. Hence, by subtracting away from Signal C an averaged signal $\mathbf{A}^k$ for high enough k, we can remove the low frequency values from the spectrum of $\mathbf{f}$ that arise from the normal heartbeats.

For example, at the top of Figure 3.5(b) we show the signal that is the difference between Signal C and its fourth averaged signal $\mathbf{A}^4$. Comparing this signal with Signal C we can see that the normal heartbeats have been removed, but there is still a residue corresponding to the abnormal heartbeat. In the middle of Figure 3.5(b) we show the signal equal to the difference between the abnormal heartbeat and its fourth averaged signal, which is a match to the residue of the abnormal heartbeat in the signal at the top of the figure. Finally, at the bottom of Figure 3.5(b) we show the graph of all the squares of the positive values of the normalized correlation of the middle signal with the top signal. This signal clearly locates the position of the abnormal heartbeat, at the same location as before, but without the secondary peaks for the normal heartbeats.

This wavelet method may strike some readers as contrived; but, in the next section we shall use this same method on 2D images of real scenes and achieve the same successful results. We introduced the method here because it is easier to understand its frequency interpretation in the 1D case.

DFT method of computing correlations

There is a simple relationship between correlations and the DFTs of signals. We shall now give a brief sketch of this relationship. Since the normalized correlation $\langle \mathbf{f} : \mathbf{g} \rangle$ simply consists of dividing $(\mathbf{f} : \mathbf{g})$ by the energy of $\mathbf{f}$, we shall concentrate on the problem of computing $(\mathbf{f} : \mathbf{g})$.

The following formula, which we shall prove later, describes the relationship between the DFT of $(\mathbf{f} : \mathbf{g})$ and the DFTs of $\mathbf{f}$ and $\mathbf{g}$:

$$(\mathbf{f} : \mathbf{g}) \stackrel{\mathcal{F}}{\longmapsto} \mathcal{F}\mathbf{f}\,\overline{\mathcal{F}\mathbf{g}} \tag{3.40}$$

where $\mathcal{F}\mathbf{f}\,\overline{\mathcal{F}\mathbf{g}}$ is the product signal with values $\left(\mathcal{F}\mathbf{f}\,\overline{\mathcal{F}\mathbf{g}}\right)_n = (\mathcal{F}\mathbf{f})_n\,\overline{(\mathcal{F}\mathbf{g})_n}$. This formula shows that the frequency content of $(\mathbf{f} : \mathbf{g})$ simply consists of the product of the DFT of $\mathbf{f}$ with the complex conjugate of the DFT of $\mathbf{g}$. Formula (3.40) also gives us the following three-step method for computing the correlation $(\mathbf{f} : \mathbf{g})$.

DFT calculation of correlation $(\mathbf{f} : \mathbf{g})$

Step 1. Compute DFTs of $\mathbf{f}$ and $\mathbf{g}$.

Step 2. Multiply the values of $\mathcal{F}\mathbf{f}$ and $\overline{\mathcal{F}\mathbf{g}}$ to produce $\mathcal{F}\mathbf{f}\,\overline{\mathcal{F}\mathbf{g}}$.

Step 3. Compute the DFT inverse of $\mathcal{F}\mathbf{f}\,\overline{\mathcal{F}\mathbf{g}}$ to produce $(\mathbf{f}:\mathbf{g})$.

Although this method may appear convoluted at first sight, it is actually much more efficient than a direct calculation of $(\mathbf{f}:\mathbf{g})$ based on Formula (3.38).

The reason that the DFT calculation of correlations is more efficient than Formula (3.38) is because an FFT algorithm is employed for performing the DFTs. Formula (3.38) requires N multiplications and $N-1$ additions in order to calculate each of the N values of $(\mathbf{f}:\mathbf{g})$. That amounts to $2N^2-N$ operations in total to compute $(\mathbf{f}:\mathbf{g})$. Computing the correlation $(\mathbf{f}:\mathbf{g})$ by this direct method is said to require $\mathrm{O}(N^2)$ operations; meaning that a bounded multiple of N^2 operations is needed. An FFT algorithm requires only $\mathrm{O}(N\log_2 N)$ operations, hence the DFT calculation of correlation also requires only $\mathrm{O}(N\log_2 N)$ operations. When N is large, say $N \geq 1024$, then $N\log_2 N$ is significantly smaller than N^2.

The DFT calculation of correlation is even more efficient when 2D images are involved, as in the next section. For 2D images, say both N by N, the correlation defined in (3.48) in the next section requires $\mathrm{O}(N^4)$ operations if a direct calculation is performed. A DFT calculation, however, requires only $\mathrm{O}(N^2\log_2 N)$ operations. The value of N need not be very large at all in order for $N^2\log_2 N$ to be significantly smaller than N^4.

Proof of (3.40) *

Formula (3.40) can be proved easily using z-transforms. Making use of cyclic translation [see Section 3.2], we can rewrite (3.38) in the form

$$(\mathbf{f}:\mathbf{g})_k = \mathbf{f}\cdot\mathcal{T}_{k-1}\mathbf{g}. \tag{3.41}$$

Consequently, the z-transform of $(\mathbf{f}:\mathbf{g})$ is

$$\begin{aligned}(\mathbf{f}:\mathbf{g})[z] &= \sum_{k=1}^{N}(\mathbf{f}\cdot\mathcal{T}_{k-1}\mathbf{g})\,z^{k-1}\\ &= \mathbf{f}[z]\,\mathbf{g}[z^{-1}].\end{aligned} \tag{3.42}$$

The second equality in (3.42) holds for the same reason that Equality (3.29) holds. Since $\mathbf{g}$ is a real-valued signal, we have $\mathbf{g}[z^{-1}] = \overline{\mathbf{g}[z]}$; so (3.42) becomes

$$(\mathbf{f}:\mathbf{g})[z] = \mathbf{f}[z]\,\overline{\mathbf{g}[z]}. \tag{3.43}$$

From (3.43) and the relation between z-transforms and DFTs, we obtain Formula (3.40).

Normalized correlations and feature detection *

In this subsection we shall briefly examine the mathematical justification for using normalized correlations to detect the presence of one signal within another, more complicated, signal. This discussion will make use of concepts from linear algebra—in particular, Cauchy's inequality for scalar products. Those readers who are not conversant with linear algebra should feel free to skip this subsection; we shall not be referring to it in the sequel.

Let $\mathbf{f}$ be a signal of positive energy, $\mathcal{E}_{\mathbf{f}} > 0$. We assume a positive energy in order to force $\mathbf{f} \neq (0, 0, \ldots, 0)$, because it is clearly pointless to try to detect the signal $(0, 0, \ldots, 0)$ within any signal. Suppose that the signal $\mathbf{g}$ contains the signal $\mathbf{f}$ in the sense that, for some integer m between 1 and N,

$$\mathbf{g} = \mathcal{T}_{m-1}\mathbf{f} + \mathbf{n} \tag{3.44}$$

where $\mathcal{T}_{m-1}\mathbf{f}$ is a cyclic translate of $\mathbf{f}$ and $\mathbf{n}$ is a *noise term.* By noise term we mean an undesired portion of the signal. The signal $\mathbf{n}$ is certainly undesired because we want to detect the presence of $\mathbf{f}$; however, the detection method based on normalized correlation works best when $\mathbf{n}$ is, in fact, random noise that is completely uncorrelated to $\mathbf{f}$. By completely uncorrelated to $\mathbf{f}$ we mean that

$$\langle \mathbf{f} : \mathbf{n} \rangle_j = \frac{\mathbf{f} \cdot \mathcal{T}_{j-1}\mathbf{n}}{\mathcal{E}_{\mathbf{f}}} \approx 0 \tag{3.45}$$

holds for each integer j. Of course, as we saw with the example of Signal C, it is not absolutely necessary that (3.45) holds; but it makes a more general derivation easier, and this derivation provides some insight into the case of Signal C as well.

In any case, assuming that (3.45) holds, we now let $k = N - m + 2$ and note that $\mathcal{T}_{k-1} \circ \mathcal{T}_{m-1} = \mathcal{T}_N = \mathcal{T}_0$. Since $\mathcal{T}_0$ is the identity mapping, we then have

$$\begin{aligned} \langle \mathbf{f} : \mathbf{g} \rangle_k &= \frac{\mathbf{f} \cdot (\mathcal{T}_{k-1} \circ \mathcal{T}_{m-1}\mathbf{f})}{\mathcal{E}_{\mathbf{f}}} + \frac{\mathbf{f} \cdot \mathcal{T}_{k-1}\mathbf{n}}{\mathcal{E}_{\mathbf{f}}} \\ &= \frac{\mathbf{f} \cdot \mathbf{f}}{\mathcal{E}_{\mathbf{f}}} + \frac{\mathbf{f} \cdot \mathcal{T}_{k-1}\mathbf{n}}{\mathcal{E}_{\mathbf{f}}} \,. \end{aligned}$$

Because $\mathbf{f} \cdot \mathbf{f} = \mathcal{E}_{\mathbf{f}}$ and (3.45) holds, we then have

$$\langle \mathbf{f} : \mathbf{g} \rangle_k = 1 \tag{3.46}$$

except for a small error that (3.45) allows us to ignore.

On the other hand, if n is any of the integers between 1 and N, then letting $j = n + m - 2$ yields

$$\begin{aligned} \langle \mathbf{f} : \mathbf{g} \rangle_n &= \frac{\mathbf{f} \cdot \mathcal{T}_j\mathbf{f}}{\mathcal{E}_{\mathbf{f}}} + \frac{\mathbf{f} \cdot \mathcal{T}_{n-1}\mathbf{n}}{\mathcal{E}_{\mathbf{f}}} \\ &\approx \frac{\mathbf{f} \cdot \mathcal{T}_j\mathbf{f}}{\mathcal{E}_{\mathbf{f}}} \,. \end{aligned}$$

By the Cauchy inequality, we obtain

$$|\mathbf{f} \cdot \mathcal{T}_j \mathbf{f}| \leq \sqrt{\mathcal{E}_\mathbf{f}} \sqrt{\mathcal{E}_{\mathcal{T}_j \mathbf{f}}} = \mathcal{E}_\mathbf{f}$$

and equality holds if and only if $\mathcal{T}_j \mathbf{f} = \mathbf{f}$ or $\mathcal{T}_j \mathbf{f} = -\mathbf{f}$. Therefore, except for a small error which (3.45) allows us to ignore, we have

$$-1 \leq \langle \mathbf{f} : \mathbf{g} \rangle_n \leq 1. \tag{3.47}$$

Equality holds on the right side of (3.47) only when $\mathbf{f} = \mathcal{T}_j \mathbf{f}$.

This discussion shows that we can detect the presence of cyclic translates of $\mathbf{f}$ within the signal $\mathbf{g}$ by the location of maximum value(s) of 1 among the positive values of $\langle \mathbf{f} : \mathbf{g} \rangle$. The method works best when those translates of $\mathbf{f}$ which are not equal to $\mathbf{f}$ produce scalar products that are much smaller than the energy of $\mathbf{f}$; this is the case, for instance, with the abnormal heartbeat considered above. Furthermore, although (3.45) does not hold for Signal C, the value of $\mathbf{f} \cdot \mathcal{T}_{k-1} \mathbf{n}$ is 0; hence (3.46) still holds. And, for the other index values j, the values of $(\mathbf{f} \cdot \mathcal{T}_j \mathbf{f})/\mathcal{E}_\mathbf{f}$ and $(\mathbf{f} \cdot \mathcal{T}_{n-1} \mathbf{n})/\mathcal{E}_\mathbf{f}$ are small enough that they add up to less than 1.

The discussion in this subsection can be extended to 2D images and will apply to the wavelet based method of object detection in images that we shall describe in the next section.

3.5 Object detection in 2D images

In the previous section we described a basic method of feature detection, and a wavelet based enhancement of it, for 1D signals. In this section we shall describe how objects can be detected within 2D images. We shall discuss some examples related to character detection and the more difficult problem of locating small objects within complicated scenes.

Our first example shows how to identify the image of the character, P, shown in Gr 1 of Figure 3.6, within the sequence of three characters, PQW, shown in Gr 2 of Figure 3.6. One method for doing this is to compute a normalized correlation. For 2D images $\mathbf{f}$ and $\mathbf{g}$, a normalized correlation of $\mathbf{f}$ with $\mathbf{g}$ is defined as a correlation divided by the energy of $\mathbf{f}$, as we did for 1D signals in the previous section. To understand the 2D definition of correlation, we rewrite the 1D definition in (3.38) in the following form:

$$(\mathbf{f} : \mathbf{g})_k = \sum_{n=1}^{N} f_n g_{n+k-1}.$$

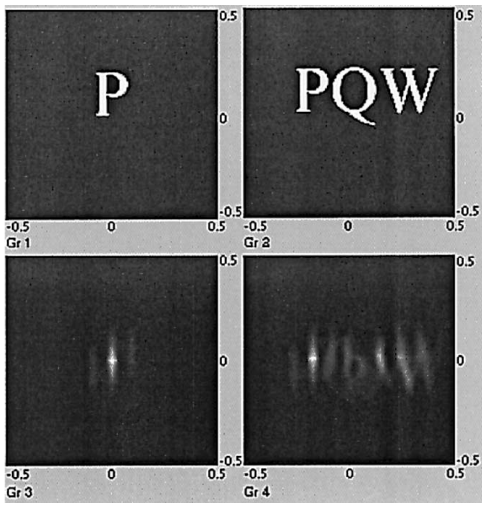

FIGURE 3.6
Gr 1: Image of P. Gr 2: Image of PQW. Gr 3: Squares of positive values of normalized correlation of Gr 1 with itself. Gr 4: Squares of positive values of normalized correlation of Gr 1 with Gr 2.

The 2D definition of the correlation $(\mathbf{f} : \mathbf{g})$ of two M by N images $\mathbf{f}$ and $\mathbf{g}$ is defined by

$$(\mathbf{f} : \mathbf{g})_{k,j} = \sum_{n=1}^{N} \sum_{m=1}^{M} f_{n,m} g_{n+k-1,m+j-1}. \tag{3.48}$$

Formula (3.48) requires a periodic extension of $\mathbf{g}$, i.e., $g_{n+N,m+M} = g_{n,m}$ is assumed to hold for all integers m and n. Based on (3.48), we define the normalized correlation $\langle \mathbf{f} : \mathbf{g} \rangle$ of two M by N images $\mathbf{f}$ and $\mathbf{g}$ by

$$\langle \mathbf{f} : \mathbf{g} \rangle_{k,j} = \frac{(\mathbf{f} : \mathbf{g})_{k,j}}{\mathcal{E}_{\mathbf{f}}} . \tag{3.49}$$

As with 1D signals, the identification of an object $\mathbf{f}$ within an image $\mathbf{g}$ is achieved by locating maximum values that are approximately 1 from among the positive values of $\langle \mathbf{f} : \mathbf{g} \rangle$. For example, if the image of P in Gr 1 of Figure 3.6 is the object $\mathbf{f}$ and the image of PQW in Gr 2 is the image $\mathbf{g}$, then we show the squares of the positive values of $\langle \mathbf{f} : \mathbf{g} \rangle$ in Gr 4. The maximum value of 1 is clearly visible and locates the position of P within the image of PQW.

Our second example involves locating a small object within a complicated scene. This problem reveals the limitations of the correlation method; nevertheless, with some assistance from wavelet analysis, we shall outline an

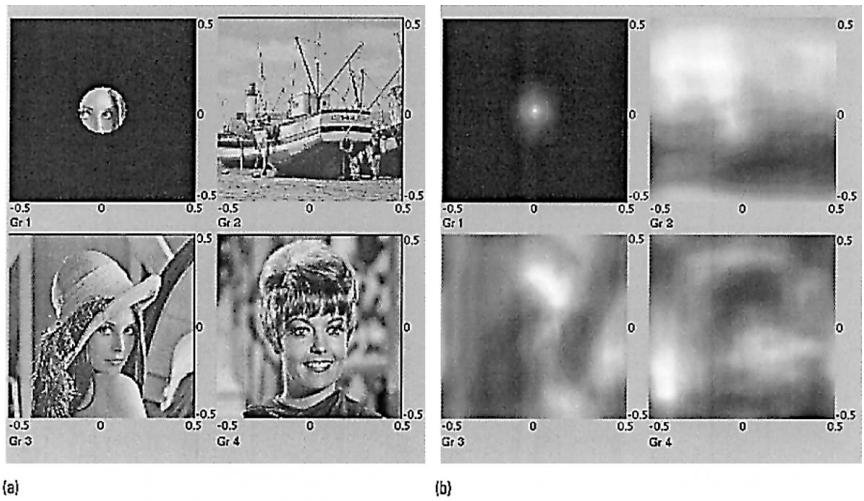

FIGURE 3.7
(a) Four images, Gr 1 is the object that we wish to locate within the other three images. (b) Squares of positive values of normalized correlations of Gr 1 from (a) with each of the images in (a).

effective solution. In Gr 1 of Figure 3.7(a), we show a small image. We wish to identify the location, or the absence, of this object within the other three images in Gr 2 through Gr 4 of Figure 3.7(a). In Figure 3.7(b) we show the squares of the positive values of the normalized correlations of the desired object with each of the images in Figure 3.7(a). It is clear from these Figures that the method fails; it successfully locates Gr 1 within Gr 1, but does not correctly locate Gr 1 within Gr 3. Furthermore, for Gr 2 and Gr 3, the maximum values of the normalized correlations are 1.14 and 1.13. The fact that these maximum values are slightly greater than 1 indicates a partial breakdown of the method, but the worst result is the fact that Gr 1 is not present at all within Gr 2.

We shall now describe a wavelet based approach to identifying the location of this object. This approach rests on the fact that objects can often be identified via their edges; for example, there is some evidence that our visual system works on this basis. We saw at the end of the last chapter that the first-level detail signal $\mathbf{D}^1$ can provide us with an image consisting of edges. Based on these observations, we shall use the following three-step method for locating the object.

Edge Correlation Method of Object Location

Step 1. Compute a first-level detail image $\mathbf{D}^1$ for the object.

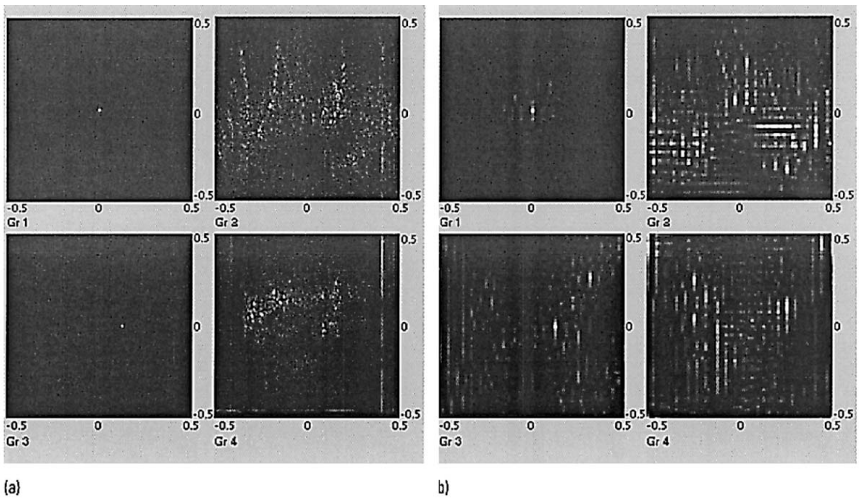

FIGURE 3.8
(a) Squares of positive values of the normalized correlations produced by the Edge Correlation Method applied to the images in Figure 3.7(a). (b) Similarly produced images resulting from the third trends of the images in Figure 3.7(a).

Step 2. Eliminate from this detail image any extraneous edge effects resulting from the outer edges in the object. This is not always necessary, but is needed with the object in Gr 1 in Figure 3.7(a). Removing its outer edge effects can be done by only retaining values that lie within a small enough radius.

Step 3. Compute normalized correlations of the image from Step 2 with first-level detail signals for each of the other images. Determine if there are any values near 1 and where they are located.

In Figure 3.8(a) we show the results of this Edge Correlation Method, using Coif12 wavelets to produce the detail images. We have graphed the squares of the positive values of the normalized correlations, which emphasizes a maximum value of 1. The location of the object within Gr 3 in Figure 3.7(a) is clearly indicated. The value of the normalized correlation at this point is 1. For Gr 2 and Gr 4 in Figure 3.8(a), there are no clear indications that the object is located within the corresponding images in Figure 3.7(a). In fact, the maximum values of the normalized correlations are 0.267 and 0.179, respectively, which are significantly less than 1.

We close this section with another example. In the last example, the object was in the same orientation within the image where we located it.

Clearly this would not often be the case. For example, the object might be rotated through some angle, or it might be reflected about some line through the origin, or some combination of these last two operations. Let's refer to these operations as *symmetry operations* on the object. One solution to the problem of locating the object when it is in a different orientation is to perform the Edge Correlation Method on a finite set of objects obtained from performing a large number of symmetry operations on the initial object. The problem with this approach is that it is prohibitively time consuming.

We shall now outline a less time consuming solution to the problem just described. The gist of this solution is to perform the Edge Correlation Method on trend subimages for some specific level of wavelet transforms of the images. For example, in Figure 3.8(b), we show the squares of the positive values of the normalized correlation images produced by the Edge Correlation Method applied to the third trend subimages $\mathbf{a}^3$ of each of the images in Figure 3.7(a). The location of the object within Gr 3 in Figure 3.7(a) is clearly indicated. The value of the normalized correlation at this point is 0.852, which is not 1 but is close to it. For Gr 2 and Gr 4 in Figure 3.8(b), there are no clear indications that the object is located within the corresponding images in Figure 3.7(a). In fact, the maximum values of the normalized correlations are 0.431 and 0.402, respectively, which are significantly less than 1.

The advantage of the method just described is that, by working with third level trend subimages, the sizes of the images are reduced by a factor of 64. This makes it feasible to perform a large number of symmetry operations on the object in order to search for rotated, or reflected, versions of it.

In illustrating the Edge Correlation Method, we ignored one important point. In order for the method to work effectively, the location of the object within an image must be centered on even index values. This is because when the first level subimages are computed, the scaling signals and wavelets are all shifts of $\mathbf{V}_1^1$ and $\mathbf{W}_1^1$ by even integers. Therefore, in general, it is necessary to perform the Edge Correlation Method on three other objects obtained by shifting the initial object by one index value in the horizontal direction, by one index value in the vertical direction, and by one index value in both directions.

3.6 Creating scaling signals and wavelets *

We conclude this chapter with an outline of the way in which scaling signals and wavelets are created from properties of their z-transforms. This material is more difficult than the other material in this chapter, and makes use of some earlier optional material; so those readers who wish to skip over

it may certainly do so. No use will be made subsequently of this material.

The heart of wavelet theory is MRA, and so we begin by expressing the 1-level MRA equation $\mathbf{f} = \mathbf{A}^1 + \mathbf{D}^1$ in terms of z-transforms:

$$\mathbf{f}[z] = \mathbf{A}^1[z] + \mathbf{D}^1[z]. \tag{3.50}$$

Using Formulas (3.32) and (3.35) in the right side of (3.50), we obtain

$$\begin{aligned}\mathbf{f}[z] = \mathbf{f}[z] &\left\{ \frac{1}{2}\mathbf{V}_1^1[z]\mathbf{V}_1^1[z^{-1}] + \frac{1}{2}\mathbf{W}_1^1[z]\mathbf{W}_1^1[z^{-1}] \right\} \\ &+ \mathbf{f}[-z] \left\{ \frac{1}{2}\mathbf{V}_1^1[z]\mathbf{V}_1^1[-z^{-1}] + \frac{1}{2}\mathbf{W}_1^1[z]\mathbf{W}_1^1[-z^{-1}] \right\}. \end{aligned} \tag{3.51}$$

By comparing the two sides of (3.51) we see that the following two equations must hold:

$$\mathbf{V}_1^1[z]\mathbf{V}_1^1[z^{-1}] + \mathbf{W}_1^1[z]\mathbf{W}_1^1[z^{-1}] = 2 \tag{3.52a}$$

$$\mathbf{V}_1^1[z]\mathbf{V}_1^1[-z^{-1}] + \mathbf{W}_1^1[z]\mathbf{W}_1^1[-z^{-1}] = 0. \tag{3.52b}$$

In order to make (3.52b) hold, we define $\mathbf{W}_1^1[z]$ by

$$\mathbf{W}_1^1[z] = -z^{2k+1}\mathbf{V}_1^1[-z^{-1}] \tag{3.53}$$

where the exponent $2k+1$ is an odd integer that we shall specify later.

Before we go further, it is interesting to observe that (3.53) implies that

$$\left|\mathbf{W}_1^1[z]\right| = \left|\mathbf{V}_1^1[-z^{-1}]\right| \tag{3.54}$$

because z is on the unit-circle (so $|z| = 1$). Formula (3.54) implies the approximation (3.33) for the Coif12 case that we considered in Section 3.3, given the graphs of $\left|\mathbf{V}_1^1\right|^2$ and $\left|\mathbf{W}_1^1\right|^2$ shown in Figure 3.2(a).

We now return to our derivation of the z-transforms of scaling functions and wavelets. Combining (3.53) and (3.52a), and the identity

$$\mathbf{V}_1^1[z]\mathbf{V}_1^1[z^{-1}] = \left|\mathbf{V}_1^1[z]\right|^2,$$

we conclude that

$$\left|\mathbf{V}_1^1[z]\right|^2 + \left|\mathbf{V}_1^1[-z]\right|^2 = 2. \tag{3.55}$$

In order to satisfy (3.55) it is easier to work with the function $P(\theta)$ defined by

$$P(\theta) = \frac{1}{\sqrt{2}}\,\mathbf{V}_1^1[e^{i2\pi\theta}], \tag{3.56}$$

where θ is a real variable. Using this function $P(\theta)$, Equation (3.55) becomes

$$|P(\theta)|^2 + |P(\theta + 1/2)|^2 = 1. \tag{3.57}$$

We will now show how the Daub4 scaling numbers in Equation (2.3) can be obtained by solving Equation (3.57). We begin by observing that the following trigonometric identity

$$|\cos \pi\theta|^2 + |\cos \pi(\theta + 1/2)|^2 = 1 \tag{3.58}$$

resembles the form of (3.57). In fact, if we were to set $P(\theta) = e^{i\pi\theta} \cos \pi\theta$, then we would be led to the two scaling numbers $\alpha_1 = \alpha_2 = 1/\sqrt{2}$ for the Haar scaling function $\mathbf{V}_1^1$. We leave the details to the reader; the reasoning involved is a simplified version of the argument that we shall now use to obtain the Daub4 scaling numbers.

The first step is to cube both sides of (3.58) obtaining

$$\begin{aligned} 1 &= \left(\cos^2 \pi\theta + \sin^2 \pi\theta\right)^3 \\ &= \cos^6 \pi\theta + 3\cos^4 \pi\theta \sin^2 \pi\theta + 3\cos^2 \pi\theta \sin^4 \pi\theta + \sin^6 \pi\theta. \end{aligned} \tag{3.59}$$

We now require that $|P(\theta)|^2$ satisfies

$$|P(\theta)|^2 = \cos^6 \pi\theta + 3\cos^4 \pi\theta \sin^2 \pi\theta, \tag{3.60}$$

which are the first two terms on the right side of (3.59). The remaining two terms on the right side of (3.59) are equal to $|P(\theta + 1/2)|^2$; so (3.57) holds.

Our final task is to obtain a function $P(\theta)$ which satisfies (3.60). In general, this is done via a result known as the *Riesz lemma.* This approach is described in the references on wavelets given at the end of the chapter. For the case of the Daub4 scaling functions, however, we can find $P(\theta)$ by a more direct, though still somewhat tricky, argument. We observe that

$$|P(\theta)|^2 = \cos^4 \pi\theta \left[\cos^2 \pi\theta + 3\sin^2 \pi\theta\right]; \tag{3.61}$$

so we could set $P(\theta) = [\cos \pi\theta]^2 \left[\cos \pi\theta - i\sqrt{3} \sin \pi\theta\right]$. For reasons that will be clear at the end, however, we instead define $P(\theta)$ by

$$P(\theta) = e^{i3\pi\theta} [\cos \pi\theta]^2 \left[\cos \pi\theta - i\sqrt{3} \sin \pi\theta\right]. \tag{3.62}$$

Notice that this formula for $P(\theta)$ implies that $P(\theta)$ satisfies (3.61), because $\left|e^{i3\pi\theta}\right|^2 = 1$. Since (3.61) holds, it follows that (3.60) does as well.

Our final task is to convert (3.62) into a form that allows us to read off the z-transform of $\mathbf{V}_1^1[z]$. To do this, we observe that

$$\begin{aligned} P(\theta) &= e^{i6\pi\theta} \left[e^{-i\pi\theta} \cos \pi\theta\right]^2 \left[e^{-i\pi\theta} \cos \pi\theta - i\sqrt{3} e^{-i\pi\theta} \sin \pi\theta\right] \\ &= e^{i6\pi\theta} \frac{1}{4} \left[1 + e^{-i2\pi\theta}\right]^2 \left[\frac{1 + e^{-i2\pi\theta}}{2} + \frac{\sqrt{3}}{2} \left(e^{-i2\pi\theta} - 1\right)\right]. \end{aligned}$$

Multiplying out the last expression and simplifying yields the formula we need:

$$P(\theta) = \frac{1+\sqrt{3}}{8} + \frac{3+\sqrt{3}}{8}\, e^{i2\pi\theta} + \frac{3-\sqrt{3}}{8}\, e^{i4\pi\theta} + \frac{1-\sqrt{3}}{8}\, e^{i6\pi\theta}. \quad (3.63)$$

Since $\mathbf{V}_1^1[e^{i2\pi\theta}] = \sqrt{2}\, P(\theta)$, we obtain the z-transform $\mathbf{V}_1^1[z]$ from Formula (3.63) by setting $z = e^{i2\pi\theta}$:

$$\mathbf{V}_1^1[z] = \frac{1+\sqrt{3}}{4\sqrt{2}} + \frac{3+\sqrt{3}}{4\sqrt{2}}\, z + \frac{3-\sqrt{3}}{4\sqrt{2}}\, z^2 + \frac{1-\sqrt{3}}{4\sqrt{2}}\, z^3. \quad (3.64)$$

Formula (3.64) tells us that $\mathbf{V}_1^1$ is defined by

$$\mathbf{V}_1^1 = (\alpha_1, \alpha_2, \alpha_3, \alpha_4, 0, 0, \ldots, 0)$$

where $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ are the Daub4 scaling numbers defined in Formula (2.3). We have thus shown how those scaling numbers can be obtained via z-transform theory.

The definition of the Daub4 wavelet numbers now follows easily. If we set $k = 1$ in Formula (3.53), then $\mathbf{W}_1^1[z] = -z^3 \mathbf{V}_1^1[-z^{-1}]$. This equation combined with (3.64) yields

$$\mathbf{W}_1^1[z] = \frac{1-\sqrt{3}}{4\sqrt{2}} + \frac{\sqrt{3}-3}{4\sqrt{2}}\, z + \frac{3+\sqrt{3}}{4\sqrt{2}}\, z^2 + \frac{-1-\sqrt{3}}{4\sqrt{2}}\, z^3. \quad (3.65)$$

Formula (3.65) implies that

$$\mathbf{W}_1^1 = (\beta_1, \beta_2, \beta_3, \beta_4, 0, 0, \ldots, 0)$$

where $\beta_1, \beta_2, \beta_3, \beta_4$ are the Daub4 wavelet numbers defined in Formula (2.8).

We end this section by noting that other properties of wavelets can be obtained by requiring that the function $P(\theta)$ satisfies certain identities. For instance, the condition that the Daub4 scaling numbers satisfy

$$\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 = \sqrt{2}$$

is equivalent to the requirement that

$$P(0) = 1. \quad (3.66)$$

Notice that the function $P(\theta)$ defined above does satisfy this requirement. Furthermore, the conditions on the Daub4 wavelet numbers, stated in Equations (2.11) and (2.12), can be easily seen to be equivalent to the two equations

$$\mathbf{W}_1^1[1] = 0, \quad \frac{d}{dz}\mathbf{W}_1^1[1] = 0. \quad (3.67)$$

Tracing back through the definitions, it is not hard to show that these last two equations are equivalent to

$$P(1/2) = 0, \quad \frac{dP}{d\theta}(1/2) = 0. \tag{3.68}$$

These equations are, indeed, satisfied by the function $P(\theta)$ defined above.

Equations (3.66) and (3.68) are important because they show how crucial identities involving the scaling numbers and wavelets can be expressed in terms of values of $P(\theta)$ and its derivatives at $\theta = 0$ and $\theta = 1/2$. For example, the function $P(\theta)$ for the Coif6 case is required to satisfy

$$P(0) = 1, \quad \frac{dP}{d\theta}(0) = 0, \quad \frac{d^2P}{d\theta^2}(0) = 0 \tag{3.69}$$

and

$$P(1/2) = 0, \quad \frac{dP}{d\theta}(1/2) = 0. \tag{3.70}$$

The equations in (3.69) correspond to Equations (2.33a) through (2.33c), while the equations in (3.70) correspond to Equations (2.32a) and (2.32b).

3.7 Notes and references

The DFT and the FFT are described in [BRH], [BRI], [WA1], and [WA2]. Using the DFT for noise removal is described in [WA2] and [WA3], and the use of related transforms for compression is discussed in [WAN].

The detection of abnormal heart signals in ECGs is considerably more complicated than our discussion indicates. Further details on wavelet based methods can be found in the papers [STC] and [AKA].

The construction of the Daub4 scaling numbers and wavelet numbers described in Section 3.6 is adapted from a discussion in [STR], where the Daub6 case is examined. The construction of scaling numbers and wavelet numbers in general, based on the Riesz lemma, is described in [DAU] and [VEK]. Newer methods are described in [SW1] and [SW2]. There is also an excellent discussion of the construction of the Coif$I$ scaling numbers and wavelet numbers in [BGG].

A rather complete characterization of wavelets from the standpoint of their frequency content is given in [HEW].