11

The wavelet transform and its relation to multirate filter banks

11.0 INTRODUCTION

In this chapter we study the wavelet transform, which has received a great deal of attention since the middle eighties. The mathematical aspects of wavelet transforms were introduced in Grossmann and Morlet [1984]. The topic has since been treated in considerable detail by several authors in the mathematics literature [Meyer, 1986], [Daubechies, 1988], [Mallat, 1989a,b], and [Strang, 1989], and a number of books have appeared [Coifman, et al., 1990], [Chui, 1991], [Daubechies, 1992]. The fundamental papers by Daubechies and by Mallat were influential in generating an unprecedented amount of activity in this area. Daubechies developed a systematic technique for generating finite-duration orthonormal wavelets, and also established the connection between continuous-time 'orthonormal wavelets' and the digital filter bank studied in Sec. 5.3.6 (FIR power symmetric filter bank). This result on finite-duration orthonormal wavelets triggered considerable interest in the mathematics as well as the signal processing communities.

Wavelet transforms are closely related to tree structured digital filter banks, and hence to multiresolution analysis described in Sec. 5.8. We know that tree structured filter banks give rise to nonuniform filter bandwidths (Fig. 5.8-4) and nonuniform decimation ratios in the subbands. These two nonuniformities can be considered to be the fundamental ingredients of the wavelet transform.

Even before the wavelet transform was formally introduced, such non uniform filter banks were already employed in the speech processing literature. See Nelson, et al. [1972], Schafer, et al. [1975], and pp. 301–303 of the text by Rabiner and Schafer [1978]. Also see McGee and Zhang [1990] for the design and use of such filter banks in music. Using nonuniform filter banks, one can exploit the decreasing resolution of the human ear at higher

frequencies [Flanagan, 1972]. This nonuniform nature of the ear also explains the evolution of the musical scale. Figure 11.1-1 demonstrates this, by showing the locations of the notes c and g in the major diatonic scale, for several octaves. On a logarithmic scale, these would appear to be nearly equispaced. The notes in between c and g are not shown, but it is clear that they become sparser and sparser as the frequency increases.

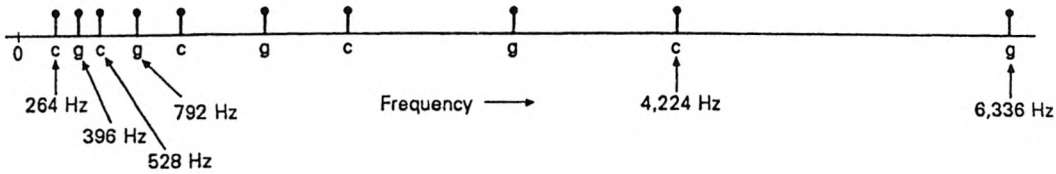

Figure 11.1-1 Pitch-frequencies corresponding to the keys 'c' and 'g' in a piano. These correspond to the major diatonic scale of western music. The spacing is very nonuniform, and will appear to be almost uniform on a logarithmic scale.

The literature on wavelet transforms is extensive, but most of it requires a level of mathematical preparation which is perhaps unsuitable for many signal-processing experts. In the signal processing literature, a number of authors have explored the relation between wavelets and multirate filter banks. Tutorial treatments can be found in the magazine articles by Rioul and Vetterli [1991], and by Hlawatsch and Boudreaux-Bartels [1992]. Further references in the signal processing literature are Evangelista [1989], Wornell [1990], Gopinath and Burrus [1991], Vaidyanathan [1991c], Soman and Vaidyanathan [1991 and 1992a,b], Tewfik and Kim [1992], Akansu and Liu [1991], Akansu, et al. [1992], Wornell and Oppenheim [1992], and Vetterli and Herley [1992]. In this chapter we will develop the basic ideas of wavelet transforms in a manner suitable for the signal processing person who understands the traditional Fourier transform, and has some familiarity with filter banks.

11.1 BACKGROUND AND OUTLINE

The conventional discrete-time Fourier transform pair is given by

$$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x(n)e^{-j\omega n}, \tag{11.1.1}$$

$$x(n) = \frac{1}{2\pi}\int_0^{2\pi} X(e^{j\omega})e^{j\omega n}d\omega. \tag{11.1.2}$$

If $x(n)$ is a single frequency signal,† that is, $x(n) = e^{j\omega_0 n}$ then

$$X(e^{j\omega}) = 2\pi\delta_a(\omega - \omega_0), \quad 0 \le \omega < 2\pi. \tag{11.1.3}$$

† A signal of the form $\cos(\omega_0 n + \theta)$ is sometimes referred to as a signal

We say that the transform is completely localized at ω_0. In contrast, the time domain plot of $e^{j\omega_0 n}$ is infinite in extent (in fact its magnitude is unity for *all* n). This is consistent with the *uncertainty principle* which says (heuristically) that if $x(n)$ has a 'wide' support, then $X(e^{j\omega})$ has 'narrow' support in $-\pi \leq \omega < \pi$.

In Sec. 11.2.4 a more quantitative statement of the uncertainty principle will be presented. The principle is most easily demonstrated for the Fourier transform pair in Fig. 11.1-2. As N increases, the signal $x(n)$ becomes less localized, but the main lobe of the Fourier transform gets narrower. As N approaches infinity, the transform looks more and more like the impulse (Dirac delta) function.

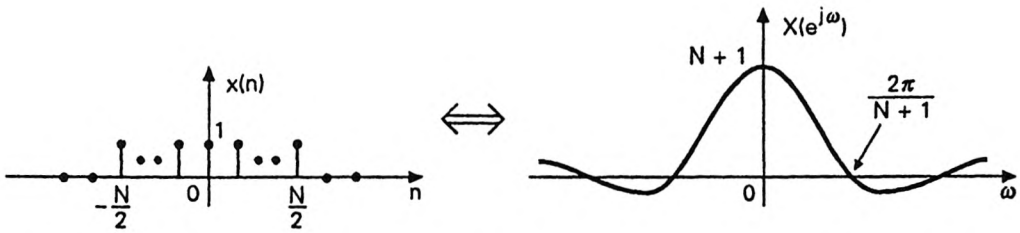

Figure 11.1-2 A Fourier transform pair, demonstrating the uncertainty principle.

The above localization property of the Fourier transform rejects the notion of "frequency that varies with time." But such a notion is often useful. For example, as the musician passes from a low to a high note, the 'frequency' (more accurately the 'pitch') is said to change in real time. According to Fourier transform theory this is meaningless because a single frequency is always associated with *infinite* time duration. If we apply Fourier analysis to the signal shown in Fig. 11.1-3(a) (where "frequency" makes an abrupt transition), we find that it is composed of an infinite number of frequencies. Another example is shown in Fig. 11.1-3(b) (a rising signal). In the regions $t < 0$ and $t > 1$ we naturally wish to think of this as a 'zero frequency signal', whereas in $0 < t < 1$ the signal has high frequency components. For signals of the above kind, it is desirable to find a time-frequency representation where the notion of 'frequency changing with time' can be formally accomodated.

The short-time Fourier transform (or time dependent Fourier trans form), abbreviated as STFT, is a tool that fills this need [Gabor, 1946], [Rabiner and Schafer, 1978], [Portnoff, 1980], [Oppenheim and Schafer, 1989]. Here the signal $x(n)$ is multiplied with a window (typically of finite duration), and then the Fourier transform computed. The window is then shifted in uniform amounts, and the above computation repeated. We will see that

---

with frequency ω_0. However, this is a superposition of two single-frequency signals, viz., $e^{j\omega_0 n}$ and $e^{-j\omega_0 n}$.

the computation of the STFT is equivalent to the implementation of a filter bank where all the filters have the same bandwidth, and each filter is followed by a decimator. The duration of the window governs the time localization of the analysis, the bandwidths of the filters govern the frequency resolution, and the decimator governs the stepsize for window movement.

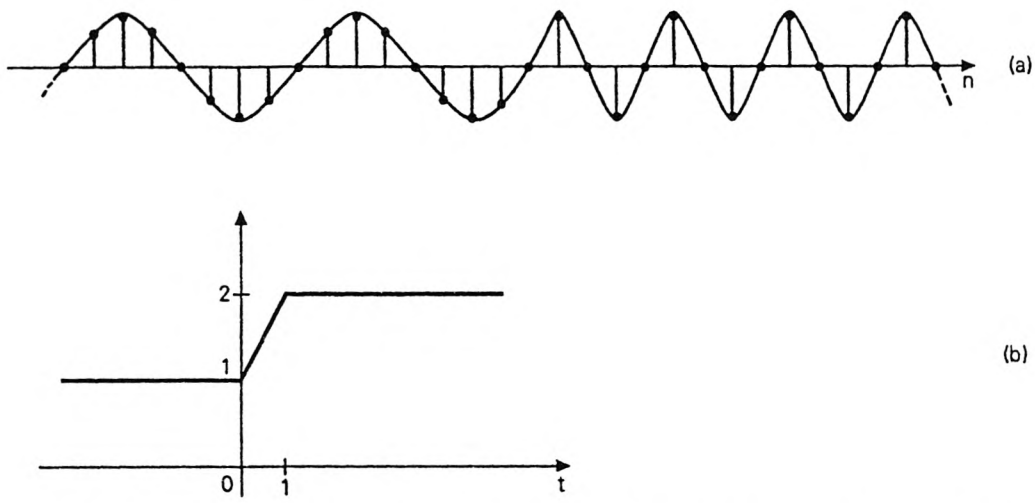

Figure 11.1-3 (a) A signal whose 'frequency' changes abruptly, and (b) a rising signal.

The wavelet transform, which is a more recent advance, generalizes the STFT by incorporating two novel features into its definition. First it permits nonuniform bandwidths (as in a tree structured filter bank system; see Fig. 5.8-4), so that the resolution is higher (i.e., bandwidth smaller) at lower frequencies. This makes the "fractional resolution" identical for all center-frequencies. Second, the nonuniform bandwidths automatically lead us to use different decimators for the different filter outputs. Such nonuniform systems are well-suited for the processing of sound signals, because of the decreasing resolution of the human ear at higher frequencies.

The traditional Fourier transform representation (2.1.21) can be regarded as an expansion of $x_a(t)$ in terms of the basis functions $e^{j\Omega t}$. (It will be easier to switch our discussion to continuous-time for a while.) The basis functions (which are functions of time) are parameterized by the frequency variable Ω. We will see that the short-time Fourier transform is a representation of a signal in terms of a different class of basis functions, indexed by *two variables*, namely *time* as well as *frequency.*

The wavelet transform is a further modification of this, which allows nonuniform frequency resolution. We will see that, in the wavelet transformation, the basis functions have a very unusual property, namely, all

the basis functions are generated by dilation and shift of a single function $\psi(t)$ (called a mother wavelet). Thus, instead of representing $x_a(t)$ as a linear combination of the functions $e^{j\Omega t}$ (as does the Fourier transform representation), the wavelet transform attempts to represent $x_a(t)$ as a linear combination of

$$\psi_{k\ell}(t) \triangleq 2^{-k/2}\psi(2^{-k}t - \ell) \quad \text{(wavelet basis functions)}$$

where k and ℓ are integers. (This is only an example.) That is,

$$x_a(t) = \sum_k \sum_\ell X_{WT}(k,\ell) \underbrace{2^{-k/2}\psi(2^{-k}t - \ell)}_{\text{basis } \psi_{k\ell}(t)} \quad \text{(wavelet expansion).}$$

This *double* summation should be compared with the traditional Fourier transform representation

$$x_a(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} X_a(j\Omega)e^{j\Omega t}d\Omega,$$

which is a *single* integral, in the frequency variabe Ω. In the wavelet expansion, k plays the role of 'frequency' and ℓ plays the role of 'time'. The index variables (k,ℓ) are integers, unlike 'Ω' which is a continuous variable in the Fourier expansion.

In a manner analogous to the orthonormality of the basis functions $\{e^{j\Omega t}\}$ in the Fourier representation, we will later discuss orthonormality of the basis functions $\{\psi_{k\ell}(t)\}$. As a preview example, the functions in Fig. 11.5-7 represent an orthonormal basis (the Haar basis) for the set of finite energy functions.

Time-frequency representations. The tools we develop in this chapter come under the general class of "time-frequency representations." In these representations, the signal is represented in a domain which is a hybrid between time and frequency, for example, a time-localized Fourier transform with the center of localization shifted uniformly. We will see later that the double index "$k\ell$" in $\psi_{k\ell}(t)$ above represents time-frequency.

The use of time-frequency representations reflects the philosophy that some aspects of a signal are most conveniently represented in the time domain, whereas there are certain other aspects which are best represented in terms of frequency. Consider, for example, the two signals shown in Fig. 11.1-4. Both of these have an underlying periodic waveform $p(t)$, for example a musical note of fixed pitch. The envelopes of the signals [$e_1(t)$ and $e_2(t)$] are, however, different. The first signal has a rapid rise followed by a steady state, and then a slow decay. The second envelope has a totally different behavior. (The envelope for a given note is typically determined by the source, for example, the musical instrument chosen). While it is useful to regard the pitch of the note in terms of impulses in the Fourier transform

domain, it is more convenient to describe the envelope in the time domain, as directly preceived by the ear.

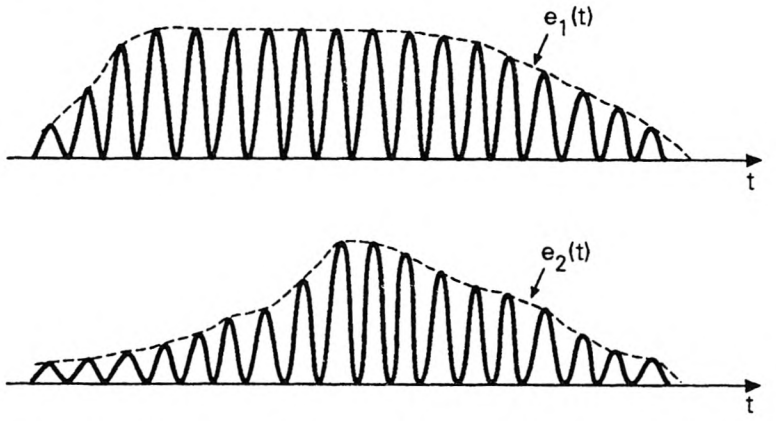

Figure 11.1-4 Two signal pulses, with nearly the same pitch, but different envelopes.

Chapter Outline

In view of the philosophy outlined above, we will first develop the short-time Fourier transform (STFT) and then the wavelet transform. In Sec. 11.2 we introduce the discrete time STFT and inverse STFT, and then develop the filter bank interpretation. In Sec. 11.2.4 we will briefly summarize the main features of the continuous-time STFT. We will show in Sec. 11.3.1 that (continuous-time) wavelet transforms can be obtained by performing two simple modifications to the STFT. This will also place in evidence the relation to filter banks.

We then introduce, in Sec. 11.3.3, the discrete-time wavelet transform (which *cannot* be obtained merely by sampling the continuous time version) and establish the connection to tree-structured, maximally decimated filter banks. The fundamental relation between paraunitary filter banks (Chap. 6) and the so-called "orthonormal wavelet decomposition" will be developed in Sec. 11.4.

In Sec. 11.5 we return to continuous time wavelets, and establish the relation between *continuous* time orthonormal wavelets and *discrete* time paraunitary filter banks. This section will show how to systematically generate finite duration orthonormal basis functions $\psi_{k\ell}(t)$ for the representation of finite energy functions. Finally, in Sec. 11.5.4 we will describe a technique for generating orthonormal wavelets with deeper properties, such as *regularity*.

We will see that the continuous time wavelet functions satisfy some very interesting mathematical properties (e.g., self-similarity) which are not relevant in the discrete time case. For such reasons, the discrete time counterpart is sometimes not regarded as "wavelets" but merely as "filter banks." In any case, both the continuous- and discrete-time versions share several

common properties (sufficient for many applications), when viewed in terms of the frequency domain quantities and subsampling operations.

Notice that our presentation in this chapter has the following order: (a) discrete-time STFT, (b) continuous-time STFT, (c) continuous-time wavelets, (d) discrete-time wavelets, and (e) connection between continuous time wavelets and discrete-time filter banks. We have opted for this ordering because there appears to be a natural "flow" between two successive items on this list. Historically, however, items (b) and (a) should have been switched, as the continuous-time STFT has been known since the middle forties [Gabor, 1946].

11.2 THE SHORT-TIME FOURIER TRANSFORM

In Sec. 4.1.2 we introduced the DFT filter bank (Fig. 4.1-16). This system merely computes the DFT of M successive samples of the input, and then repeats the operation for the next set of M samples. We saw that this could be interpreted in terms of a window that slides past the data. This could also be interpreted in terms of a bank of bandpass filters with 13 dB stopband attenuation. We also generalized this to the case of a uniform DFT filter bank (Fig. 4.3-5), with the advantage that the filters could now offer sharper cutoff and higher attenuation. These systems are essentially implementations of the so called short-time Fourier transform (STFT) to be described now.

In short-time Fourier transformation, a signal $x(n)$ is multiplied with a window $v(n)$ (typically finite in duration). See Fig. 11.2-1. The Fourier transform of the product $x(n)v(n)$ is computed, and then the window $v(n)$ is shifted in time, and the Fourier transform of the product computed again. This operation results in a separate Fourier transformation for each location m of the center of the window. In other words, we obtain a function $X_{STFT}(e^{j\omega}, m)$ of two variables ω and m. The frequency variable ω is continuous, and takes the usual range $-\pi \leq \omega < \pi$. The shift-variable m is typically an integer multiple of some fixed integer K. Figure 11.2-2 shows a two dimensional plot which represents the idea. These are often called *spectrograms.* See Oppenheim and Schafer [1989] for real-time examples of spectrograms. (Also notice the cover picture in that reference!)

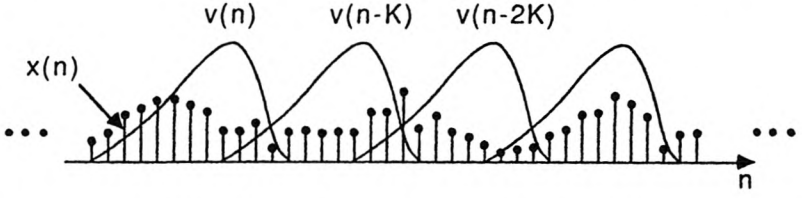

Figure 11.2-1 Pertaining to the short-time Fourier transform.

Essentially, for any fixed m, the window captures the features of the signal $x(n)$ in the local region around m. The window therefore helps to localize

the time domain data, before obtaining the frequency domain information. From the above discussions it is clear that the short time Fourier transform can be written mathematically as

$$X_{STFT}(e^{j\omega}, m) = \sum_{n=-\infty}^{\infty} x(n)v(n-m)e^{-j\omega n}. \qquad (11.2.1)$$

If we set $v(n) = 1$ for all n, this reduces to the traditional Fourier transform for any choice of m.

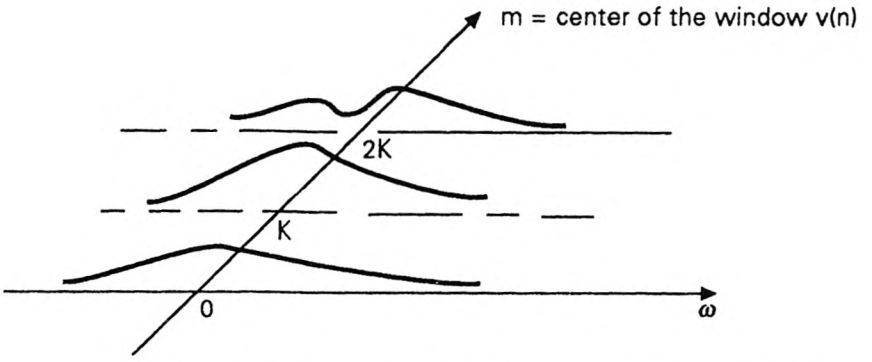

Figure 11.2-2 Demonstrating $X_{STFT}(e^{j\omega}, m)$ for $m = 0, K, 2K \ldots$

Existence of the STFT. The traditional Fourier transform (11.1.1) exists only if the signal $x(n)$ satisfies some subtle mathematical conditions [Oppenheim, Willsky and Young, 1983]. In a practical STFT system, however, $v(n)$ has finite duration, so that the above summation always converges (i.e., the STFT exists). Thus many of the subtle mathematical questions which are raised in connection with the Fourier transform do not arise in the STFT regime.

11.2.1 Interpretation Using Bandpass Filters

For a variety of reasons, it is convenient to interpret the STFT using the notion of filter banks. In addition to enhancing insight, this also gives a practical scheme for implementation. Furthermore this interpretation helps us to generalize the STFT to obtain more flexibility (Sec. 11.2.3). Finally, the theory of perfect reconstruction filter banks can be used to obtain practical "inversion" formulas for STFT.

Traditional Fourier Transform as a Bank of Filters

We will begin by presenting a filter bank interpretation for the traditional Fourier transform (11.1.1). The evaluation of $X(e^{j\omega})$ at a fixed frequency ω_0 can be pictorially represented as in Fig. 11.2-3(a). This is a cascade of two systems.

1. *The modulator* $e^{-j\omega_0 n}$. This performs a frequency-shift. More specifically, it shifts the Fourier transform towards the left by the amount ω_0, so that the zero-frequency value of $S(e^{j\omega})$ is equal to $X(e^{j\omega_0})$ (Fig. 11.2-3(b)).
2. *The LTI system* $H(e^{j\omega})$. This has impulse response $h(n) = 1$ for all n. This system is evidently unstable (Sec. 2.1.2), but let us ignore these fine details for the moment. Its "frequency response" is

$$H(e^{j\omega}) = \sum_{n=-\infty}^{\infty} h(n)e^{-j\omega n} = 2\pi\delta_a(\omega), \quad -\pi \le \omega < \pi,$$

where $\delta_a(.)$ is the Dirac delta function which was defined at the beginning of Sec. 2.1. † Thus, $H(e^{j\omega})$ is an ideal "lowpass" filter, which "passes" only the zero-frequency signal. Every other frequency is completely suppressed. This filter can be regarded as the limit, as $\Delta\omega \to 0$, of an ideal filter with response $2\pi/\Delta\omega$ for $|\omega| \le \Delta\omega/2$ and zero elsewhere.

The output $y(n)$ of the system is therefore a zero-frequency signal with $Y(e^{j\omega}) = 2\pi X(e^{j\omega_0})\delta_a(\omega)$ for $-\pi \le \omega < \pi$, that is,

$$y(n) = X(e^{j\omega_0}), \quad \text{for all } n.$$

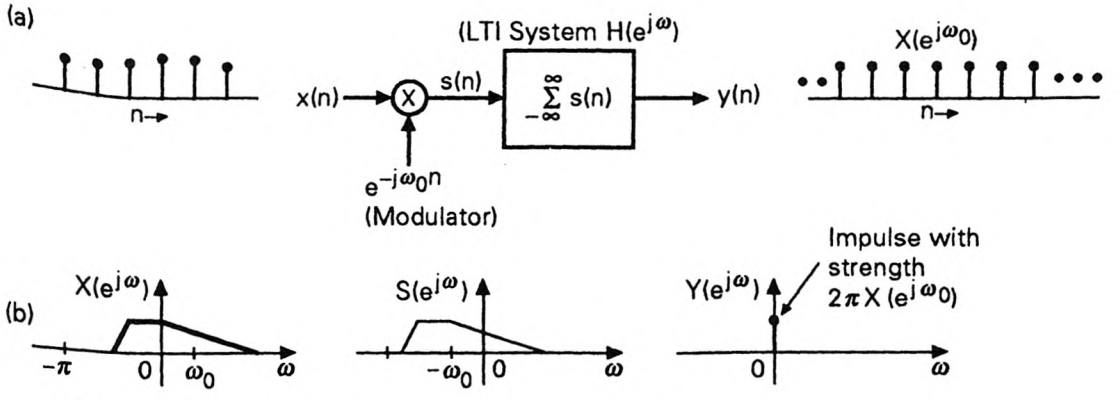

Figure 11.2-3 (a) Representation of Fourier transformation in terms of linear systems. (b) Frequency domain quantitites, sketched for an arbitrary example.

Summarizing, the process of evaluating $X(e^{j\omega_0})$ can be looked upon as a linear system, which takes the input $x(n)$ and produces a *constant* output

† Of course, the Fourier transform $H(e^{j\omega})$ repeats periodically with period 2π, but we will not explicity show it in the formulas.

$y(n)$ whose value is equal to $X(e^{j\omega_0})$ for all time n. Thus, any sample of $y(n)$ can be taken to be the value of $X(e^{j\omega_0})$. The Fourier transform operator which evaluates $X(e^{j\omega})$ for *all* ω is, therefore, a *bank* of modulators followed by filters. This system has an uncountably infinite number of channels.

The STFT as a Bank of Filters

From its definition it is clear that the STFT can be represented as in Fig. 11.2-4(a). In this figure, ω_0 and m are constants. So $y(n)$ is constant for all n, with $y(n) = X_{STFT}(e^{j\omega_0}, m)$. To gain further insight, let us rearrange the definition to obtain

$$X_{STFT}(e^{j\omega}, m) = e^{-j\omega m} \sum_{n=-\infty}^{\infty} x(n)v(n-m)e^{j\omega(m-n)}. \qquad (11.2.2)$$

Figure 11.2-4(b) shows this interpretation, where the indices k and m have been replaced with n, to be consistent with standard notations (Sec. 2.1.2). This is a linear system with two parts. The first is an LTI filter with impulse response $v(-n)e^{j\omega_0 n}$. This is followed by the modulator $e^{-j\omega_0 n}$ (linear time varying device). The output $y_0(n)$ of this system is now a function of n [unlike in Fig. 11.2-4(a)]. For any specific value of n, say $n = m$, this output represents the Fourier transform of $x(.)$ "in the neighbourhood of m," because m represents the location of the window $v(k)$ in the time domain. For the special case where $v(k) = 1$ for all k, this output becomes a constant [equal to the traditional Fourier transform $X(e^{j\omega_0})$] for all n.

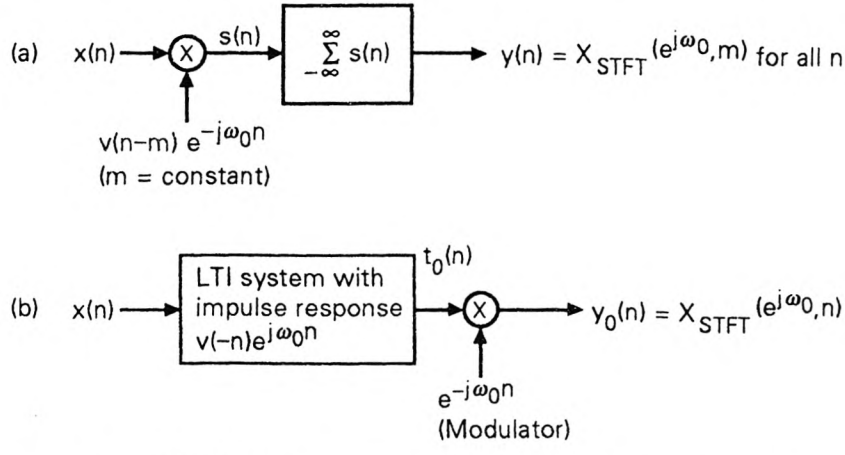

Figure 11.2-4 (a) The STFT represented in terms of a linear system and (b) a rearrangement.

In most applications, $v(n)$ has a lowpass transform $V(e^{j\omega})$. So $v(-n)$ has the lowpass transform $V(e^{-j\omega})$. The modulated version $v(-n)e^{j\omega_0 n}$ represents a bandpass filter $V(e^{-j(\omega-\omega_0)})$. See Fig. 11.2-5. The output sequence

$t_0(n)$ in Fig. 11.2-4(b) is, therefore, the output of a bandpass filter, whose passband is centered around ω_0. The effect of the modulator $e^{-j\omega_0 n}$ is merely to re-center this around zero frequency.

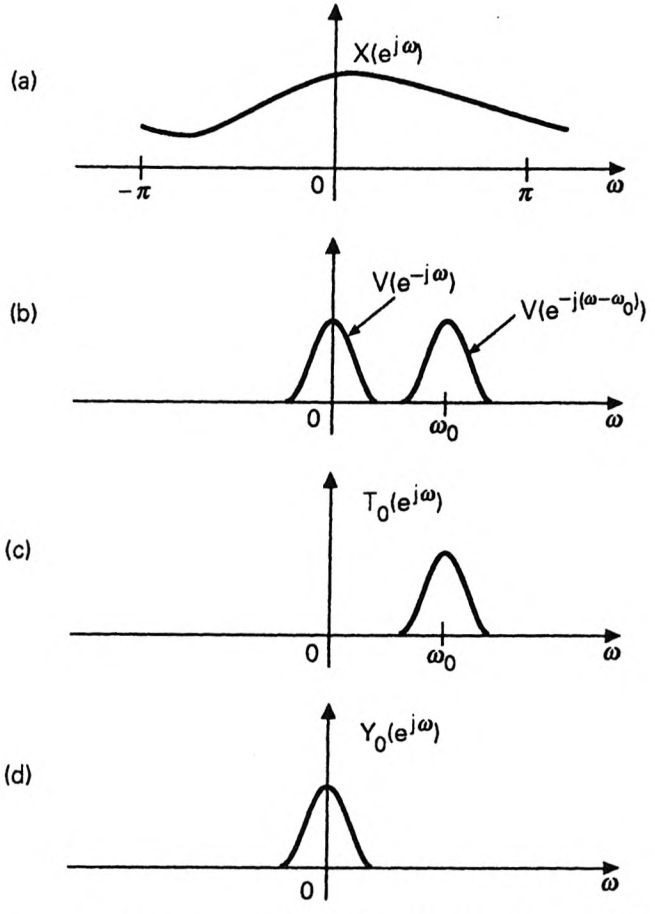

Figure 11.2-5 Demonstration of how the STFT works. (a) $X(e^{j\omega})$, (b) the window-transform and its shifted version, (c) output of LTI filter and (d) traditional Fourier transform of $X_{STFT}(e^{j\omega_0}, n)$.

For every frequency ω_0 the STFT performs the filtering operation of Fig. 11.2-4(b) to produce an output sequence $X_{STFT}(e^{j\omega_0}, n)$. So the STFT can be looked upon as a *filter bank*, with infinite number of filters (one 'per frequency'). In practice, we are interested in computing the Fourier transform at a discrete set of frequencies

$$0 \leq \omega_0 < \omega_1 < \ldots < \omega_{M-1} < 2\pi. \tag{11.2.3}$$

In this case the STFT reduces to a filter bank with M bandpass filters $H_k(z)$ with responses $H_k(e^{j\omega}) = V(e^{-j(\omega-\omega_k)})$ (and followed by modulators). This is shown in Fig. 11.2-6. The passband of $H_k(e^{j\omega})$ is centered around ω_k. The output signals $y_k(n)$ represent the STFT coefficients.

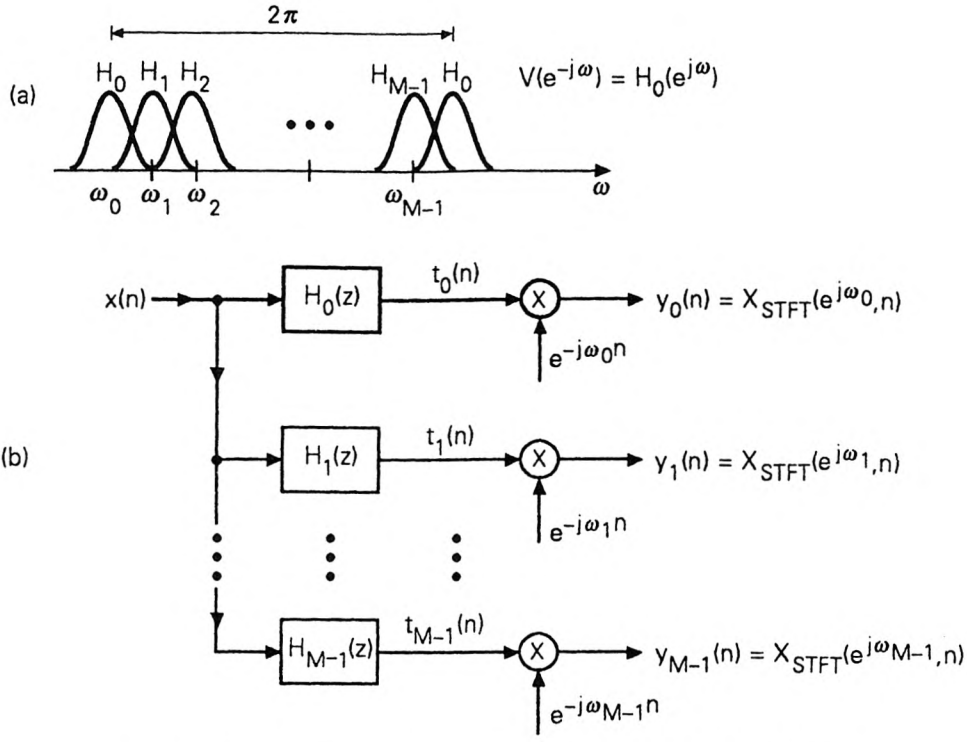

Figure 11.2-6 The STFT operation viewed as a filter bank. (a) $V(e^{-j\omega})$ and shifted versions, and (b) the filter bank.

The uniform DFT bank. If the frequencies ω_k are uniformly spaced, then the above system becomes the uniform DFT bank (Sec. 4.1.2). In this system the M filters are related as

$$H_k(z) = H_0(zW^k), \quad 0 \le k \le M-1, \tag{11.2.4}$$

where $W = e^{-j2\pi/M}$. This means that the frequency responses are uniformly shifted versions of $H_0(e^{j\omega})$, i.e.,

$$H_k(e^{j\omega}) = H_0(e^{j(\omega-\frac{2\pi k}{M})}), \quad 0 \le k \le M-1. \tag{11.2.5}$$

The unshifted filter is $H_0(e^{j\omega}) = V(e^{-j\omega})$. In Fig. 4.3-5 we saw an implementation of this set of filters, in terms of the polyphase components

$E_\ell(z)$ (Sec. 4.3) and the DFT matrix $\mathbf{W}$. The window $v(n)$ is completely determined by $E_\ell(z)$ because

$$V(z^{-1}) = H_0(z) = E_0(z^M) + z^{-1}E_1(z^M) + \ldots + z^{-(M-1)}E_{M-1}(z^M). \tag{11.2.6}$$

It is now clear that the uniform DFT bank is a device to compute the short-time Fourier transform at uniformly spaced frequencies. In particular if the polyphase components are replaced with unity [i.e., $E_k(z) = 1$ for all k], then the system merely computes the DFT of a block of M samples (Fig. 4.1-16). In this case $v(n)$ is a rectangular window of length M.

Choice of v(n), and Time-Frequency Tradeoff

Unlike the Fourier transform, the STFT is not uniquely defined unless we specify the window $v(n)$. The fact that $v(n)$ can be chosen by the user offers a great deal of flexibility. The choice of $v(n)$ essentially governs the tradeoff between 'time localization' and 'frequency resolution' as explained next.

The signal $y_0(m)$ can be considered to represent the change or 'evolution' of the Fourier transform of $x(n)$, evaluated around frequency ω_0. Thus $y_0(m)$ represents the local information, *around* time m (since m represents the location of the window $v(k)$ in the time domain) *around* frequency ω_0. It is clear that as $V(e^{j\omega})$ becomes narrower, the bandpass filters in Fig. 11.2-6(b) get narrower, and $y_k(n)$ gets closer to $X(e^{j\omega_k})$. This means that the information in the frequency domain is becoming more and more localized. However, as $V(e^{j\omega})$ gets narrower, the window $v(n)$ gets wider (uncertainity principle) so that the localization of information in the time domain is compromised. Fig. 11.2-7 demonstrates the tradeoff between time localization and frequency resolution.

The fact that time localization and frequency resolution are conflicting requirements has given rise to interesting theoretical questions. For example, what is the choice of window that will give the best frequency resolution for a given time localization? To answer this question, one has to define the term "best" using a mathematical measure. To take a specific case, suppose we constrain $v(n)$ to be a symmetric window of finite length $N+1$. What is the best choice of the coefficients $v(n)$ so that the energy of $V(e^{j\omega})$ is most concentrated in a specified region $|\omega| \leq \alpha$? This problem was in fact addressed in Sec. 3.2.2, where we found the solution to be the *prolate spheroidal sequence,* obtainable from an eigenvector of a positive definite matrix. See Sec. 11.2.4 for another measure of "best localization."

Time-Frequency Representation, and Decimation

It is often stated that $X_{STFT}(e^{j\omega}, m)$ is a time-frequency representation, because it is a function of time m as well as frequency ω. If the passband width of $V(e^{j\omega})$ [hence that of $V(e^{-j\omega})$] is narrow, then the signal $y_0(n)$ in Fig. 11.2-4(b) is a narrowband lowpass signal. This means that $y_0(n)$ varies slowly with the time index n. An extreme case is when $V(e^{j\omega})$ is an impulse (traditional Fourier transform) so that $y_0(n) = X(e^{j\omega_0})$ for all n.

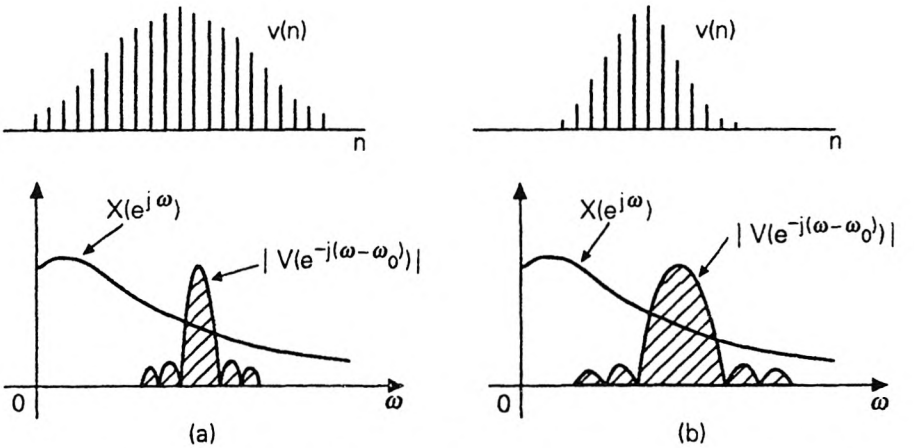

Figure 11.2-7 Demonstrating the tradeoff between time localization and frequency resolution. (a) Wide window $v(n)$; poor time localization and good frequency resolution. (b) Narrow window $v(n)$; good time localization and poor frequency resolution.

The slowly-varying nature of $X_{STFT}(e^{j\omega}, n)$, [i.e., $y_0(n)$ in Fig. 11.2-4(b)] can be exploited to decimate it, thereby resulting in a more economical time-frequency representation. If the decimation ratio is M, then this is equivalent to moving the window $v(k)$ by M samples at a time. (That is, M is the 'step size' for window movement). If $y_0(n)$ were not varying at all (as with traditional Fourier transform), then we would have to retain only one sample, and its value is $X(e^{j\omega_0})$.

Figure 11.2-8 shows a decimated STFT system, where the modulators have been moved past the decimators. Since the filters have equal bandwidth, the decimation ratios n_k can be taken to be equal. With $n_k = M$ this represents a maximally decimated analysis bank. In a more general system n_k could be different for different k, and moreover $H_k(z)$ may not be derived from one prototype by modulation. Such a system, however, does not represent the STFT obtainable by moving a single window $v(k)$ across the data $x(n)$. When we introduce the wavelet transforms in Sec. 11.3, we will admit such generalized systems.

The time-frequency representation offers a whole family of tradeoffs ('time localization' versus 'frequency resolution' tradeoff) between the two extremes, viz., $x(n)$ (time domain representation), and $X(e^{j\omega})$ (frequency domain representation). The filter bank system performs an operation analogous to Fourier transformation, yet the outputs of the transform are time varying. After performing the maximal decimation, the time-frequency representation has the same number of samples per unit time as does $x(n)$. There is no redundancy in the representation.

The time-frequency grid. Figure 11.2-9 demonstrates a grid in the two-dimensional time-frequency plane. The vertical lines represent the 'frequencies' where the STFT is computed (i.e., center frequencies of the filters). The horizontal lines represent the sample locations in the time domain, for the decimated filter outputs. The intersections of the lines represent the location of a sample of $X_{STFT}(e^{j\omega_k}, m)$. This grid represents uniform sampling of both 'frequency' ω and 'time' n. The fact that the time spacing is M corresponds to the fact that the window is moved in steps of M units at a time. The frequency spacing (spacing between center frequencies of adjacent filters) is $2\pi/M$ because there are M filters of identical bandwidths.

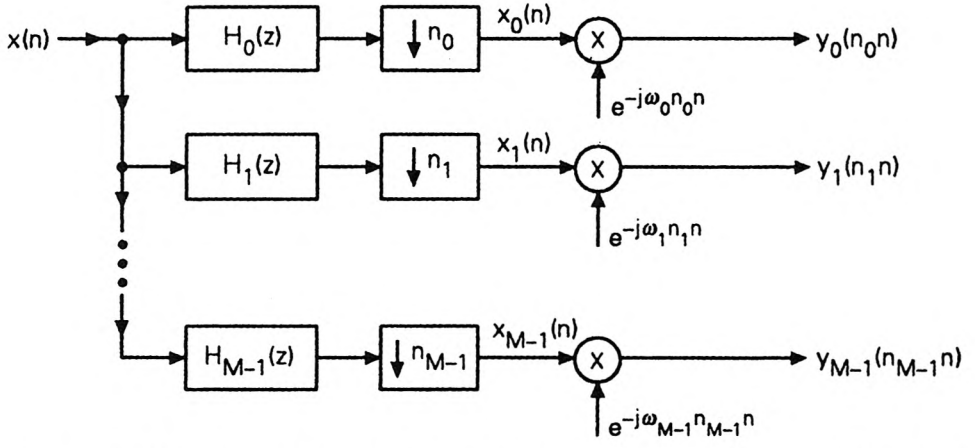

Figure 11.2-8 An analysis bank with decimators and modulators. The signal $y_k(n_k n)$ represents the decimated version of $X_{STFT}(e^{j\omega_k}, n)$ where ω_k is the center frequency of $H_k(e^{j\omega})$. Usually $n_k = M$ for all k.

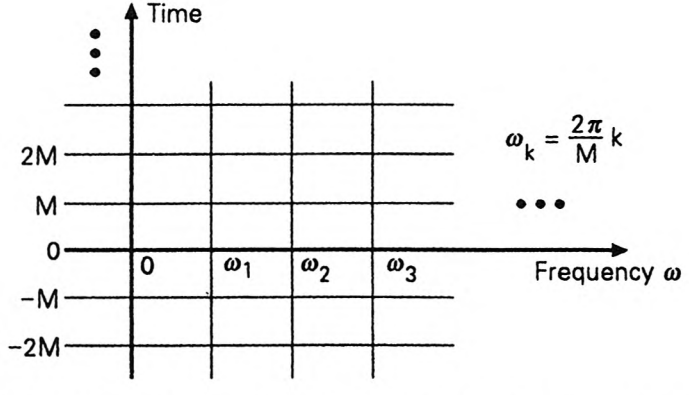

Figure 11.2-9 The two-dimensional time-frequency grid for evaluating the short-time Fourier transform.

11.2.2 Inversion of the STFT

Starting from the definition (11.2.1) it is easy to derive a number of inversion formulas that recover $x(n)$ from $X_{STFT}(e^{j\omega}, m)$. As $X_{STFT}(e^{j\omega}, m)$ is the traditional Fourier transform of $x(n)v(n-m)$, it is clear that

$$x(n)v(n-m) = \frac{1}{2\pi}\int_0^{2\pi} X_{STFT}(e^{j\omega}, m)e^{j\omega n}d\omega. \tag{11.2.7}$$

For example if we set $n = m$ we obtain the STFT inversion formula

$$x(m)v(0) = \frac{1}{2\pi}\int_0^{2\pi} X_{STFT}(e^{j\omega}, m)e^{j\omega m}d\omega, \tag{11.2.8}$$

so that we can recover $x(m)$ for all m as long as $v(0) \neq 0$. [If $v(0) = 0$, pick some other value of m in (11.2.7)]. Notice that it is not necessary to know $v(n)$ for all n, in order to recover $x(n)$ from its STFT.

A second inversion formula is given by

$$x(n) = \frac{1}{2\pi}\int_0^{2\pi}\Big(\sum_{m=-\infty}^{\infty} X_{STFT}(e^{j\omega}, m)v^*(n-m)\Big)e^{j\omega n}d\omega, \tag{11.2.9}$$

provided $\sum_m |v(m)|^2 = 1$. To prove this we mererly substitute (11.2.1) into the RHS of (11.2.9), which then reduces to $x(n)\sum_m |v(m)|^2$. There are some subtleties about the formula, which we mention here. Problem 11.3 covers details.

1. If $\sum_m |v(m)|^2 \neq 1$ but finite, we can divide the right side of (11.2.9) by $\sum_m |v(m)|^2$ and obtain the inversion. However, if the window $v(n)$ has infinite energy (as in the important special case of $v(m) = 1$), this inversion formula cannot be applied.
2. Suppose we replace $v^*(n-m)$ in (11.2.9) with $w^*(n-m)$ where $w(n)$ is an arbitrary sequence with the restriction that $\sum_n v(n)w^*(n) = 1$. Then the inversion formula still works!
3. The function $G(e^{j\omega}, m)$ satisfying

$$x(n) = \frac{1}{2\pi}\int_0^{2\pi}\Big(\sum_{m=-\infty}^{\infty} G(e^{j\omega}, m)v^*(n-m)\Big)e^{j\omega n}d\omega, \tag{11.2.10}$$

is not unique. For example, suppose z_0 is a zero of the z-transform $\sum_k v^*(k)z^{-k}$. Then $G(e^{j\omega}, m) \triangleq X_{STFT}(e^{j\omega}, m) + z_0^m$ satisfies (11.2.10) for the same sequence $x(n)$. This is unlike the case of traditional Fourier transform, where (11.1.2) is not satisfied if we replace $X(e^{j\omega})$ with something else.

Filter Bank Interpretation of the Inverse Transform

It is valuable to express the inverse transform using filter bank notation. Recall that Fig. 11.2-6(b) offers a practical means of implementing the STFT. Viewed like this, the STFT is a transformation of a one dimensional sequence $x(n)$ into a two dimensional sequence $y_k(m)$ (i.e., function of two integer variables k and m). The number and locations of the frequencies ω_k might appear to be arbitrary, and so might the shapes of the filters. In fact they *are* somewhat arbitrary, subject primarily to the requirement that the signal $x(n)$ be reconstructible from the STFT coefficients $y_k(n)$ with reasonable accuracy in reasonable amount of time. It turns out that, as long as the filters $H_k(z)$ are chosen properly, we can find stable synthesis filters $F_k(z)$ to recover $x(n)$ perfectly. On the other hand, inversion of the traditional Fourier transformation (11.1.1) requires the computation of an *integral.*

With the STFT implemented as in Fig. 11.2-8, the reconstruction is done by using a synthesis bank as shown in Fig. 11.2-10. Typically $n_k = M$ for all k, but we will use n_k for generality. The z-transform of $\hat{x}(n)$ is given by

$$\hat{X}(z) = \sum_{k=0}^{M-1} X_k(z^{n_k}) F_k(z). \tag{11.2.11}$$

In the time domain this is equivalent to

$$\begin{aligned} \hat{x}(n) &= \sum_{k=0}^{M-1} \sum_{m=-\infty}^{\infty} x_k(m) f_k(n - n_k m) \\ &= \sum_{k=0}^{M-1} \sum_{m=-\infty}^{\infty} y_k(n_k m) e^{j\omega_k (n_k m)} f_k(n - n_k m). \end{aligned} \tag{11.2.12}$$

If the synthesis filters $F_k(z)$ are such that $\hat{x}(n) = x(n)$, we can say that (11.2.12) is the representation of $x(n)$ in terms of the decimated STFT coefficients $y_k(n_k m)$ just as (11.1.2) is the representation of $x(n)$ in terms of the traditional Fourier transform 'coefficients' $X(e^{j\omega})$. While (11.1.2) is an integral in terms of the single variable ω, the new representation is a *double* summation (in the integer variables k and m.) The reconstruction is stable if the filters $F_k(z)$ are stable.

Example 11.2.1: Reconstructing x(n) from STFT Coefficients

Assume that there are no decimators, that is, the STFT is as in Fig. 11.2-6. Let the window $v(n)$ satisfy the Nyquist property, viz., $v(Mn) = 0$ for $n \neq 0$. Then, the M filters $H_k(z)$ defined in (11.2.4) satisfy the property $\sum_{k=0}^{M-1} H_k(z) = c_0$ for constant $c_0 \neq 0$ (Sec. 4.6.1). This means

that we can reconstruct $x(n)$ from the STFT coefficients $y_k(n)$ simply by adding them after demodulation, that is,

$$x(n) = c_1 \sum_{k=0}^{M-1} y_k(n) e^{j\omega_k n} \tag{11.2.13}$$

for some constant c_1.

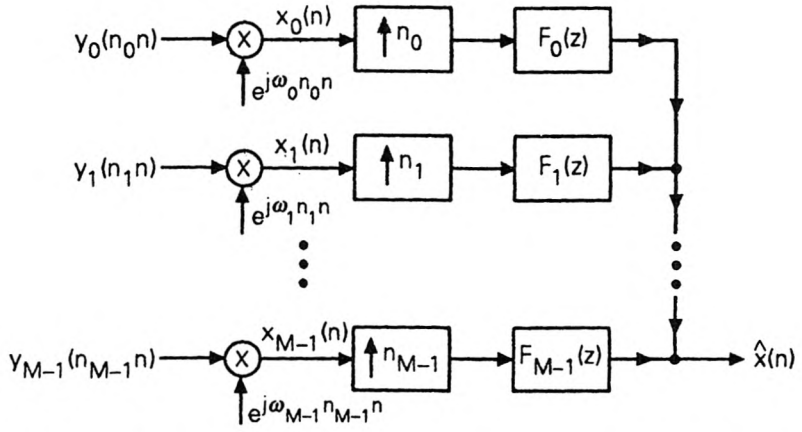

Figure 11.2-10 The synthesis bank used to reconstruct $x(n)$ from its STFT coefficients. Usually $n_k = M$ for all k.

11.2.3 Generalizations of the STFT

We know that we can recover an arbitrary signal $x(n)$ from its (decimated) STFT coefficients, provided we can design a synthesis bank with perfect reconstruction (PR) property. However, since the analysis filters are derived from a single prototype by modulation, the PR requirement will in turn restrict the coefficients of $v(n)$ severely (Example 5.7.2).

If we relax the requirement that all analysis filters be derived from one prototype $v(n)$, we can obtain more flexibility. For this we generalize the STFT idea by viewing any analysis filter bank as a *generalized Fourier transformer*. The outputs of the filters are narrowband signals, and represent the localized Fourier transform as described above. This generalized system, however, is not derivable from a traditional single sliding window system as in Fig. 11.2-1, i.e., the simple description (11.2.1) does not hold. These are more appropriately called "spectrum analyzers" rather than short time Fourier transformers. From Chap. 5–8 we know that there exist many techniques to perfectly reconstruct $x(n)$ from the (possibly decimated) filter outputs.

We will say that the quantities

$$x_k(n), \quad 0 \le k \le M-1, \quad -\infty \le n \le \infty, \tag{11.2.14}$$

in Fig. 11.2-8 are the generalized STFT coefficients of the signal $x(n)$. The transform domain is characterized by *two* integer variables k and n. The STFT pair can be written as

$$x_k(n) = \sum_{m=-\infty}^{\infty} x(m)h_k(n_k n - m) \quad \text{(STFT)}, \tag{11.2.15}$$

and

$$x(n) = \sum_{k=0}^{M-1} \sum_{m=-\infty}^{\infty} x_k(m)f_k(n - n_k m) \quad \text{(Inverse STFT)}. \tag{11.2.16}$$

The decimators n_k should be chosen to be inversely proportional to the passband widths of the filters $H_k(z)$. If these bandwidths are equal, then $n_k = M$ for all k. Notice that the modulators and demodulators have not been included in the definition for simplicity; they cancel each other anyway, and do not provide further insights. We can regard (11.2.15) and (11.2.16) as the generalized STFT pair or the "spectrum analyzer/synthesizer pair". It is also called the "filter-bank transform pair." For the special case where the analysis filters are as in (11.2.5) with $H_0(z) = V(z^{-1})$, this reduces to the traditional STFT which is computed from a single window $v(n)$.

Comments.

1. The above STFT/inverse STFT definition assumes that the filters $H_k(z)$ and $F_k(z)$ are such that the filter bank system (Fig. 11.2-8 followed by Fig. 11.2-10) has perfect reconstruction property. It is easy to ensure that $H_k(z)$ are stable. If $F_k(z)$ are also stable, we have a stable reconstruction scheme for performing the inverse transform.
2. *Nonuniqueness.* In Fig. 11.2-6(b), suppose α is a zero of all the analysis filters, that is, $H_k(\alpha) = 0$ for all k. This means that if we replace $x(n)$ with $x(n) + \alpha^n$, the STFT coefficients do not change. In other words, the sequence $x(n)$ producing the transform domain coefficients is not unique. Such a situation is easily avoided in practice. For example, in a perfect reconstruction system this situation will not arise. (Because, it would imply that the input $x(n) = \alpha^n$ produces zero output).

Basis Functions and Orthonormality

Consider the conventional Fourier transform representation (11.1.2). Here $x(n)$ is a linear combination of the sequences $e^{j\omega n}$, and the set $\{e^{j\omega n}\}$ is said to be a basis for the space of sequences representable as in (11.1.2). The basis functions $e^{j\omega n}$, are orthonormal in following sense:

$$\sum_{n=-\infty}^{\infty} (e^{j\omega_1 n})^* e^{j\omega_2 n} = 2\pi\delta_a(\omega_1 - \omega_2), \tag{11.2.17}$$

that is, the "inner-product" of $e^{j\omega_1 n}$ and $e^{j\omega_2 n}$ is equal to zero unless $\omega_1 = \omega_2$.

By the above analogy we see that the quantities

$$\eta_{km}(n) \triangleq f_k(n - n_k m) \tag{11.2.18}$$

play the role of "basis functions" in (11.2.16). Notice that this is a doubly indexed family of functions. The first index is the filter number k, and the second index determines the time shift. Such basis functions will be called the "filter-bank like basis".

It is of interest to impose the orthonormality property on the basis functions $\{\eta_{km}(n)\}$. This property means that

$$\sum_{n=-\infty}^{\infty} f^*_{k_1}(n - n_{k_1} m_1) f_{k_2}(n - n_{k_2} m_2) = \delta(k_1 - k_2)\delta(m_1 - m_2). \tag{11.2.19}$$

In other words, the above summation should be zero except for the case where $k_1 = k_2$ and $m_1 = m_2$ (and reduces to unity in that case). How should we design the filters $F_k(z)$ in order to ensure this? We will return to this very interesting issue in Sec. 11.4, and show that the paraunitary property of the polyphase matrix (Chap. 6) is sufficient!

Relation between $\mathbf{h_k(n)}$ and $\mathbf{f_k(n)}$. In the traditional Fourier transform pair, the basis function $e^{j\omega n}$ appears in (11.1.2) whereas its conjugate appears in (11.1.1). Inspection of the STFT pair reveals no obvious analogy of this relation. The only requirement is that the functions $h_k(n_k n - m)$ and $f_k(n - n_k m)$ be related in such a way as to ensure perfect reconstruction. We will return to this later, and show that if the basis functions are orthonormal then $f_k(n) = h^*_k(-n)$. This is very similar to the relation between analysis and synthesis filters in a paraunitary perfect reconstruction system (Sec. 6.2.1).

Table 11.2.1 provides a summary of the discrete-time STFT.

11.2.4 The Continuous-Time Case

Historically, the STFT idea was first developed for the continuous-time case even though our presentation here started with the discrete case. In 1946 Dennis Gabor † considered windowed versions of the continuous-time Fourier transform. Gabor used a Gaussian window, that is, a function of the form $v(t) = v(0)e^{-bt^2}, b > 0$. The corresponding continuous-time STFT is called the *Gabor transform* [Gabor, 1946]. Notice that this window does not have finite duration.

† He received the Nobel Prize in 1971 for contributions to the principles of holography.

TABLE 11.2.1 Short-Time Fourier Transform (STFT), discrete-time.

Here are the key equations governing the discrete-time traditional and short-time Fourier transforms. Note that the subscript on $\eta_\omega(n)$ is the real-valued continuous variable ω. For the STFT, there are two subscripts, as in $\eta_{\omega,m}(n)$ and $\eta_{km}(n)$. In the former, ω is a real (continuous) variable. In the latter, k is integer-valued (*center-frequency number* or *filter number*). The double subscripts arise because the transform domain is *time-frequency* rather than *frequency,* as would be the case for the traditional Fourier transform. Read the text, particularly in the neighbourhood of the following equations, to fully appreciate assumptions and conditions.

Traditional Discrete-time Fourier Transform

$$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x(n)e^{-j\omega n} \qquad \text{(transform)} \qquad (11.1.1)$$

$$x(n) = \frac{1}{2\pi}\int_0^{2\pi} X(e^{j\omega}) \underbrace{e^{j\omega n}}_{\text{basis } \eta_\omega(n)} d\omega \qquad \text{(inverse transform)} \qquad (11.1.2)$$

Discrete-Time STFT

$$X_{STFT}(e^{j\omega}, m) = \sum_{n=-\infty}^{\infty} x(n)v(n-m)e^{-j\omega n} \qquad \text{(STFT)} \qquad (11.2.1)$$

$$x(m)v(0) = \frac{1}{2\pi}\int_0^{2\pi} X_{STFT}(e^{j\omega}, m)e^{j\omega m}d\omega \qquad \text{(inverse STFT)} \qquad (11.2.8)$$

$$x(n) = \frac{1}{2\pi}\int_0^{2\pi} \sum_{m=-\infty}^{\infty} X_{STFT}(e^{j\omega}, m) \underbrace{v^*(n-m)e^{j\omega n}}_{\text{basis } \eta_{\omega,m}(n)} d\omega \qquad \text{(inverse STFT)} \qquad (11.2.9)$$

Generalized Discrete-Time STFT (Filter-Bank Transformer)

$$x_k(n) = \sum_{m=-\infty}^{\infty} x(m)h_k(n_k n - m), \quad 0 \le k \le M-1 \qquad \text{(STFT)}, \qquad (11.2.15)$$

$$x(n) = \sum_{k=0}^{M-1} \sum_{m=-\infty}^{\infty} x_k(m) \underbrace{f_k(n - n_k m)}_{\text{basis } \eta_{km}(n)} \qquad \text{(inverse STFT)}. \qquad (11.2.16)$$

It is assumed that the filter bank with analysis filters $h_k(n)$, synthesis filters $f_k(n)$ and decimation ratios n_k (Fig. 11.2-8, 11.2-10) has perfect reconstruction ($\hat{x}(n) = x(n)$). For orthonormal basis, $f_k(n) = h_k^*(-n)$.

Because of the close resemblance to the discrete-time case, we only summarize the main points. Given a signal $x(t)$ we define the STFT as

$$X_{STFT}(j\Omega,\tau)=\int_{-\infty}^{\infty}x(t)v(t-\tau)e^{-j\Omega t}dt \quad \text{(STFT)}, \tag{11.2.20}$$

where $v(t)$ is an appropriate window function, typically with lowpass Fourier transform $V(j\Omega)$. (Proper choice of $v(t)$ ensures existence of the integral). Once again we can find an inversion formula similar to (11.2.8), and obtain

$$x(\tau)v(0)=\frac{1}{2\pi}\int_{-\infty}^{\infty}X_{STFT}(j\Omega,\tau)e^{j\Omega\tau}d\Omega \quad \text{(inv. STFT)}, \tag{11.2.21}$$

which works as long as $v(0)\neq 0$. The inverse transform analogous to (11.2.9) is given by the double integral

$$x(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\Big(\int_{-\infty}^{\infty}X_{STFT}(j\Omega,\tau)v^*(t-\tau)d\tau\Big)e^{j\Omega t}d\Omega \quad \text{(inv. STFT)}, \tag{11.2.22}$$

and this is associated with similar subtleties as itemized after (11.2.9).

Choice of "Best Window" to Optimize Localization

We know that if the window $v(t)$ is 'narrow' in the time domain, its Fourier transform is 'broad' and vice versa. This means that there is a tradeoff between time localization and frequency resolution. To make this idea more precise, the rms (root-mean squared) duration of a signal is introduced in the literature ([Gabor, 1946], [Papoulis, 1977a]. Thus consider the two nonnegative quantities D_t and D_f defined by

$$D_t^2=\frac{1}{E}\int_{-\infty}^{\infty}t^2v^2(t)dt, \quad D_f^2=\frac{1}{2\pi E}\int_{-\infty}^{\infty}\Omega^2|V(j\Omega)|^2d\Omega. \tag{11.2.23}$$

where E is the window energy, that is, $E=\int v^2(t)dt$. (For this discussion $v(t)$ is real.) We say that D_t is the rms time domain duration and D_f the rms frequency domain duration of $v(t)$. Figure 11.2-11 shows the rms duration D_t for a number of signals (the reader is requested to verify these in Problem 11.5). It is interesting that a triangular waveform has a smaller rms duration than a rectangular waveform, even though they have identical 'traditional duration'! This is because the factor t^2 in the definition of D_t^2 increases the penalty on nonzero values of $v(t)$, as t increases.

Uncertainty principle. It turns out that the product D_tD_f cannot be arbitrarily small. Here is a quantitative statement of uncertainty principle: $D_tD_f\geq 0.5$, with equality if and only if $v(t)=Ae^{-\alpha t^2},\alpha>0$ (Problem 11.6). Thus the optimal window is a Gaussian waveform, and its 'traditional' duration is infinite.

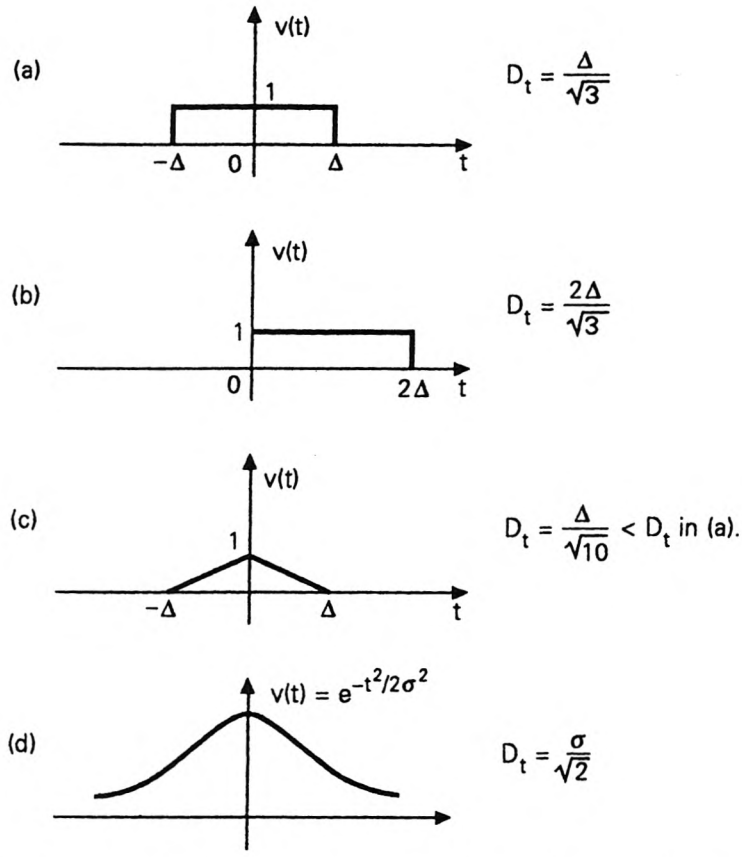

Figure 11.2-11 RMS duration of some typical signals. (a) Rectangular window, (b) one-sided rectangular window, (c) triangular window, and (d) Gaussian window.

Filter Bank Interpretation

To obtain further insight we rewrite the STFT for fixed frequency Ω_k as

$$X_{STFT}(j\Omega_k, \tau) = e^{-j\Omega_k \tau} \int_{-\infty}^{\infty} x(t) v(t-\tau) e^{j\Omega_k (\tau - t)} dt. \tag{11.2.24}$$

Defining $h(t) = v(-t)$, we see that the integral above is a convolution of $x(t)$ with the filter having impulse response

$$h_k(t) = h(t) e^{j\Omega_k t} = v(-t) e^{j\Omega_k t}. \tag{11.2.25}$$

Thus

$$X_{STFT}(j\Omega_k, \tau) = e^{-j\Omega_k \tau} \int_{-\infty}^{\infty} x(t) h_k(\tau - t) dt. \tag{11.2.26}$$

Figure 11.2-12 shows this "filtering" interpretation. We have replaced τ with t everywhere to conform with traditional notations. The output $y_k(t)$ represents the STFT of $x(t)$, evaluated with window $v(t)$, at the frequency Ω_k. Assuming $v(-t)$ is narrowband lowpass, the output of the filter in the figure is narrowband bandpass, centered at Ω_k. So the signal $y_k(\tau)$ is narrowband lowpass, that is, it is "slowly varing in τ." It represents an estimate of the Fourier transform of $x(t)$ 'localized around time τ' and around the frequency Ω_k. With the system of Fig. 11.2-12 repeated for several values of Ω_k, the complete system is equivalent to a bank of bandpass filters (Fig. 11.2-13). All filter responses are shifted versions of the prototype response $H(j\Omega)$. Finally, if $v(t) = 1$ for all t, then $y_k(t)$ is constant for all t, that is, $y_k(t) = X(j\Omega_k)$ (traditional Fourier transform).

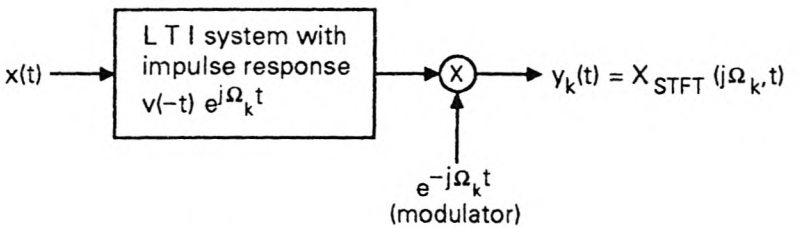

Figure 11.2-12 The continuous-time STFT as an LTI filter followed by modulator. For each frequency of interest Ω_k, we have one such filter, resulting in a filter bank.

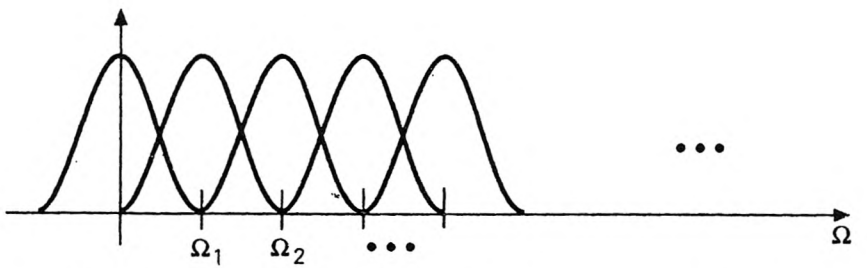

Figure 11.2-13 Performing the STFT at a discrete set of frequencies is equivalent to the use of a bank of bandpass filters.

Uniformly Sampled Version of the STFT

Since $y_k(t)$ is narrowband lowpass, we can sample it with appropriate sampling period, say T, to obtain $X_{STFT}(j\Omega_k, nT)$. See Fig. 11.2-14. (With nonideal filters, aliasing due to sampling is unavoidable, and must somehow be canceled later). One special choice of Ω_k is particularly illuminating, viz.,

$\Omega_k = k\Omega_0$, for integer k and fixed Ω_0. In this case we have

$$\mathcal{X}(k,n) \triangleq X_{STFT}(jk\Omega_0, nT) = e^{-j(\Omega_0 T)kn} \int_{-\infty}^{\infty} x(t)v(t-nT)e^{jk\Omega_0(nT-t)}dt. \tag{11.2.27}$$

Essentially, $\mathcal{X}(k,n)$ is a uniformly sampled version of the two-dimensional function $X_{STFT}(j\Omega,\tau)$. The set of sample points described by $(k\Omega_0, nT)$ is called the Gabor lattice (or Von Neumann lattice in quantum mechanics literature; [Von Neumann, 1955]). The question now is, can we reconstruct $x(t)$ from this sampled transform? If the product $\Omega_0 T$ is sufficiently small, the answer is in the affirmative. It is shown in the literature (also see Problem 11.4) that if $\Omega_0 T = 2\pi$, the function $x(t)$ can indeed be written explicitly in terms of the samples $\mathcal{X}(k,n)$. However, the reconstruction procedure itself is unstable; one requires the condition $\Omega_0 T < 2\pi$ for stability. See Daubechies [1990] for elaboration on this point.

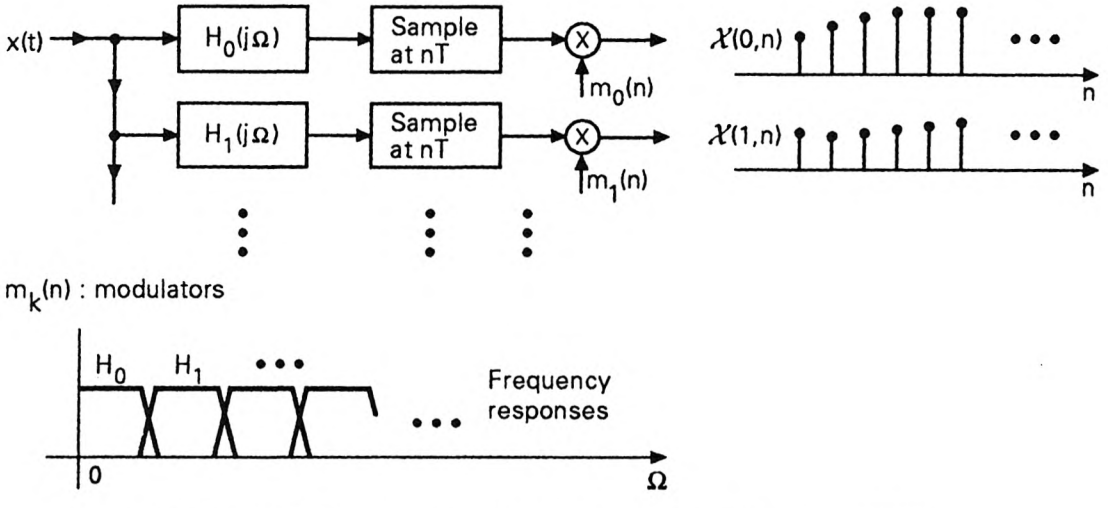

Figure 11.2-14 Sampled version of the continuous-time STFT.

The fact that we can reconstruct $x(t)$ from the two dimensional sampled version might occassion an initial surprise, since we have not assumed any bandlimited property of $x(t)$. Consider, however, the situation where we increase Ω_0 for a fixed value of the product $\Omega_0 T < 2\pi$. In the extreme case where $\Omega_0 \to \infty$, we have $T \to 0$. This is equivalent to having a single filter with infinite bandwidth, whose output is $x(t)$ itself; this output is sampled with samples spaced infinitesimally close together. In other words, the transformed version is essentially $x(t)$ itself!

11.3 THE WAVELET TRANSFORM

While the short-time Fourier transform is a convenient generalization of the Fourier transform, it still has some disadvantages. To appreciate this,

consider Fig. 11.3-1 which shows two cases. For the first case $x(t)$ is a high-frequency signal, and many cycles are captured by the window. For the second, $x(t)$ is of low frequency, so that very few cycles are within the window. Thus the accuracy of the estimate of the Fourier transform is poor at low frequencies, and improves as the frequency increases. This can also be seen from the fact that the bandpass filters in Fig. 11.2-13 have equal bandwidths, rather than bandwidth increasing with center-frequency.

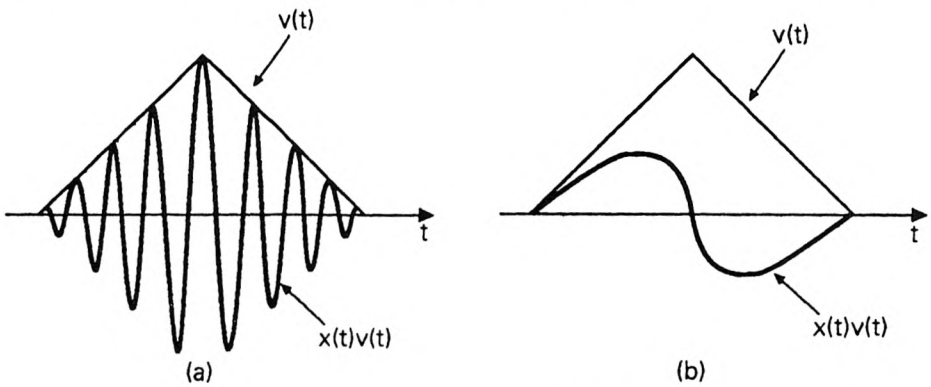

Figure 11.3-1 The windowed function $x(t)v(t)$ for (a) high-frequency signal $x(t)$, and (b) low-frequency signal $x(t)$.

Another issue is revealed by considering the rising signal of Fig. 11.1-3(b). We see that if the window is narrow, it helps to localize the rising portion very well, as compared to a wide window. With a narrow window, however, the information in the steady part of the signal changes very slowly. It will be appropriate here to have a window whose width adjusts itself with 'frequency'. This can be accomplished by using a filter bank where the lowpass filter has a narrower bandwidth (wider time-width) than the bandpass and highpass filters.

One (conceptual) way to do this is to replace the window $v(t)$ with a function of both frequency and time, so that the time domain plot of the window gets wider (i.e., bandwidth gets narrower) as frequency decreases. In this way, the window captures nearly the same number of zero-crossings for any sinusoidal input with arbitrary frequency. Furthermore, as the window gets wider, it is also desirable to have wider step sizes for moving the window (equivalently larger decimation ratio n_k in Fig. 11.2-8).

These goals are nicely accomplished by the wavelet transform. We begin by developing the continuous-time wavelet transform which is conceptually easier.

11.3.1 Passing from STFT to Wavelets

Step 1. Nonuniform filter banks

The bandpass filters in Fig. 11.2-6 have equal bandwidth because they

are obtained by modulation of a single filter. As a first step, we give up this modulation scheme, and obtain the filters $h_k(t)$ as

$$h_k(t) = a^{-k/2} h(a^{-k}t), \quad a > 1, \; k = \text{integer}. \tag{11.3.1}$$

Equivalently, in the frequency domain,

$$H_k(j\Omega) = a^{k/2} H(ja^k\Omega). \tag{11.3.2}$$

Thus all the responses are obtained by *frequency-scaling* of a prototype response $H(j\Omega)$. This is unlike the case of STFT, where all filters were obtained by *frequency-shift* of a prototype.

The scale factor $a^{-k/2}$ in (11.3.1) is meant to ensure that the energy $\int_{-\infty}^{\infty} |h_k(t)|^2 dt$ is independent of k. This can be regarded as a normalizing convention or height convention.

Example 11.3.1

Assuming that $H(j\Omega)$ is bandpass with cutoff frequencies α and β, we obtain the responses shown in Fig. 11.3-2(a). Note that $H_0(j\Omega) = H(j\Omega)$. We have assumed $a = 2$, and $\beta = 2\alpha$. The bandedges of adjacent filters overlap, as indicated. The passband gets narrower as the center frequency decreases. Quantitatively, we define the center frequency to be the geometric mean of the two cutoff edges, that is,

$$\Omega_k = 2^{-k}\sqrt{\alpha\beta} = \alpha 2^{-k}\sqrt{2}. \tag{11.3.3}$$

These are nonuniformly located, and appear to be uniform if the frequency axis is represented on a logarithmic scale. Notice that $H(j\Omega)$ is bandpass rather than highpass. One often restricts k to be nonnegative, so that there are no filters to the right of the bandpass filter $H_0(j\Omega)$. This is acceptable if the input signal has no information beyond this filter, that is, if it is bandlimited. From Fig. 11.3-2(a) we see that the ratio

$$\frac{\text{bandwidth}}{\text{center-frequency } \Omega_k} = \frac{2^{-k}(\beta - \alpha)}{2^{-k}\sqrt{\alpha\beta}} = \frac{\beta - \alpha}{\sqrt{\alpha\beta}} = \frac{1}{\sqrt{2}} \tag{11.3.4}$$

is independent of the filter number k. † Notice a slight change in convention here: as k increases, the center frequency *decreases*. This happens to be more convenient.

† In the language of electrical filter theory [Sedra and Brackett, 1978], such a system is often said to be a 'constant Q' system. The quantity 'Q' (Quality factor) is usually defined as (center-frequency/bandwidth), i.e., the reciprocal of (11.3.4).

Nonuniform bandwidths as in the above example are very useful for the analysis of sound signals. See, for example, [Nelson, et al., 1972], and pp. 301-303 of Rabiner and Schafer [1978]. This is because of the decreasing frequency-resolution of the human ear with increasing frequency [Flanagan, 1972].

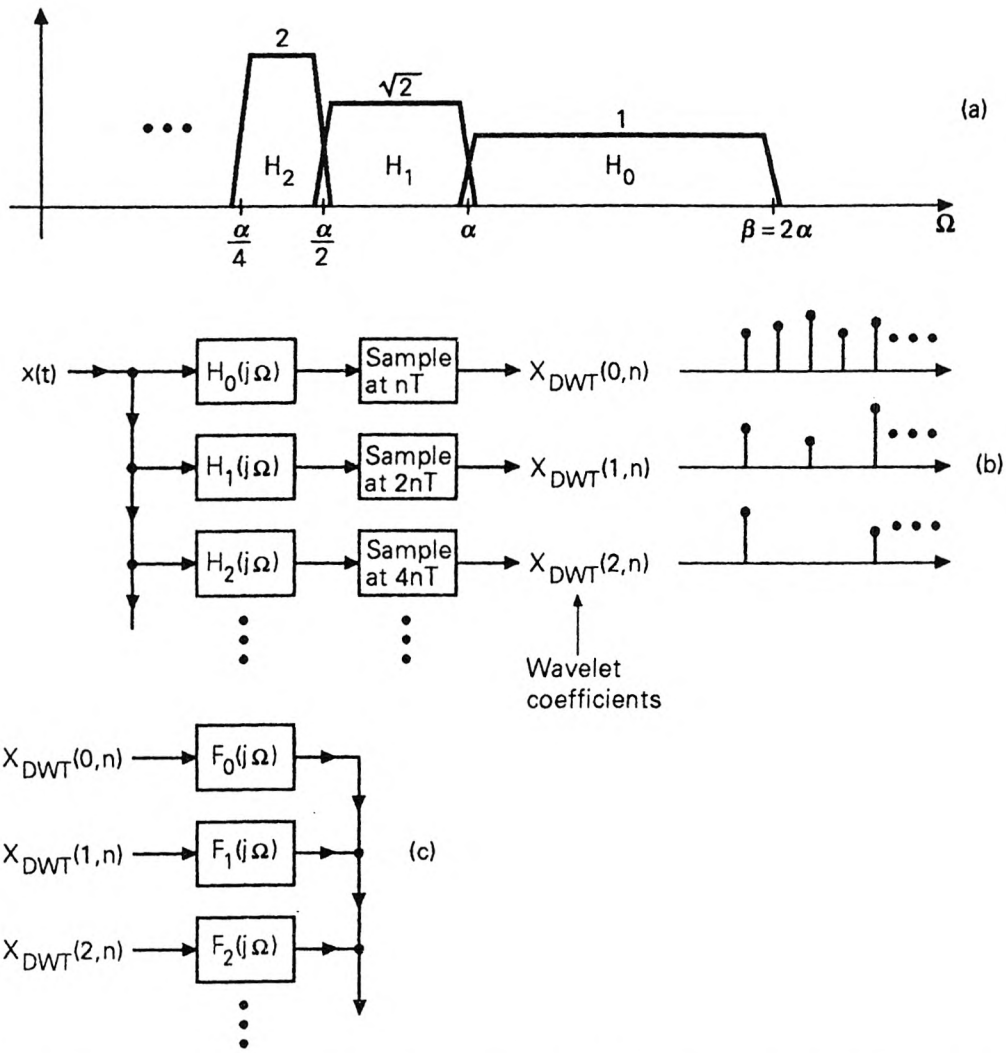

Figure 11.3-2 (a) Frequency responses obtained by the scaling process (11.3.2) with $a = 2$. (b) Analysis bank representation of discrete wavelet transform. (c) Synthesis bank which would reconstruct $x(t)$ from the set of wavelet coefficients $X_{DWT}(k,n)$. In this figure, the signal $X_{DWT}(k,n)$ indicates a continuous-time impulse train $\sum_{n=-\infty}^{\infty} X_{DWT}(k,n)\delta(t-2^k nT)$.

With the filters redefined as in (11.3.1), the filter outputs can be obtained by modifying the right hand side of (11.2.26) with

$$a^{-k/2}e^{-j\Omega_k\tau}\int_{-\infty}^{\infty}x(t)h\Big(a^{-k}(\tau-t)\Big)dt. \tag{11.3.5}$$

This is the first of the two modifications of the STFT, which will lead to wavelet transforms.

Step 2. Nonuniform decimation

Since the bandwidth of $H_k(j\Omega)$ is smaller for large k, we can sample its output at a correspondingly lower rate. Equivalently, viewed in the time domain, the width of $h_k(t)$ is larger so that we can afford to move the window by a larger step size. We will do this by replacing the continuous variable τ with na^kT in (11.3.5), where n is an integer. This means that the step size for window movement is a^kT and it increases with k, that is, increases as the center-frequency Ω_k (hence bandwidth) of the filter *decreases.* Thus the quantity $h[a^{-k}(\tau-t)]$ in (11.3.5) is now replaced with

$$h\Big(a^{-k}(na^kT-t)\Big) \;=\; h(nT-a^{-k}t). \tag{11.3.6}$$

Summarizing, we are computing

$$X_{DWT}(k,n)=a^{-k/2}\int_{-\infty}^{\infty}x(t)h(nT-a^{-k}t)dt, \quad k,n \text{ integers}, \tag{11.3.7a}$$

i.e.,

$$X_{DWT}(k,n)=\int_{-\infty}^{\infty}x(t)h_k(na^kT-t)dt, \quad k,n \text{ integers}. \tag{11.3.7b}$$

We have omitted the inconsequential factor $e^{-j\Omega_k\tau}$ which appeared earlier in (11.3.5). The above integral represents the convolution between $x(t)$ and $h_k(t)$, evaluated at a discrete set of points na^kT. In other words, the output of the convolution (a continuous-time function) is sampled with spacing a^kT. Fig. 11.3-2(b) is a schematic of this for $a=2$. The kth sampler merely retains the samples at the locations $(2^kT)n$, where n =integer.

The subscript DWT above stands for *discrete wavelet transform.* Also, the dependence on a and $h(t)$ is not explicitly indicated in the notation $X_{DWT}(k,n)$.

Time-frequency grid. Figure 11.3-3 shows the time-frequency diagrams for the STFT and the wavelet transform, and this summarizes the fundamental difference between these two. In the former, the frequency spacing and time spacing are uniform. In the latter, the frequency spacing

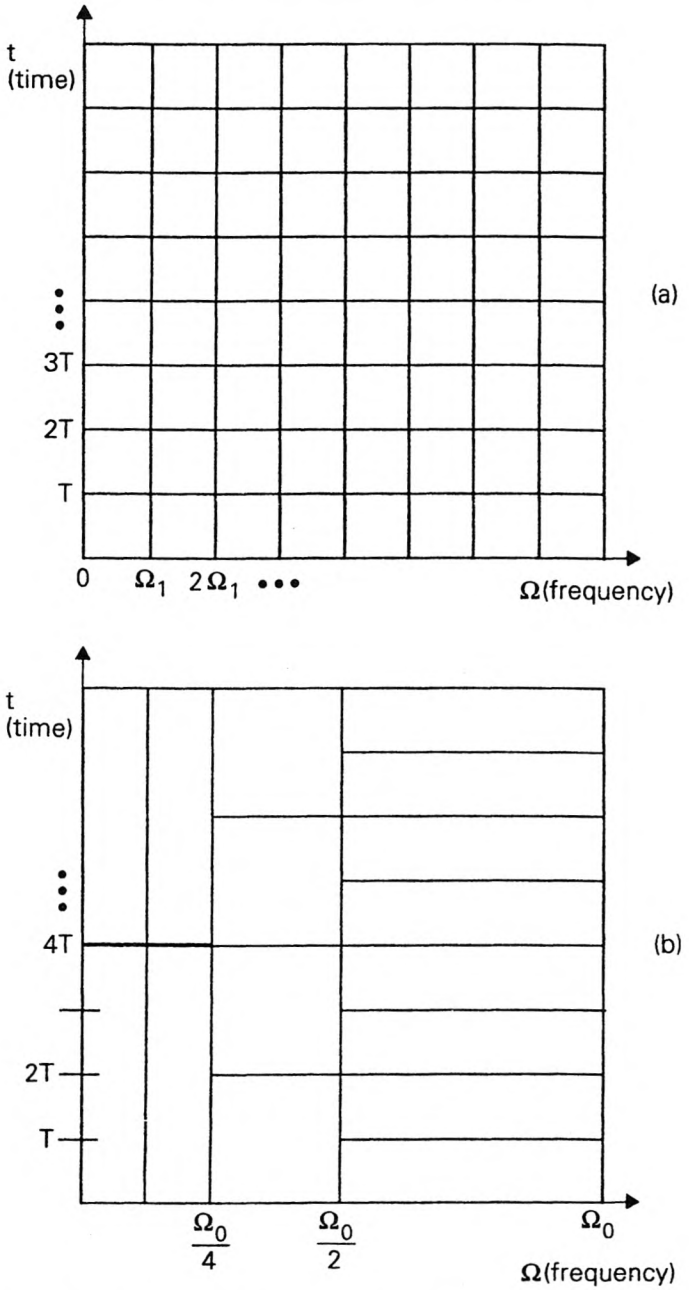

Figure 11.3-3 Fundamental differences between the STFT and the wavelet transform. (a) In the STFT, time and frequency axes are typically uniformly divided. (b) In the wavelet transform, the frequency samples are spaced closer together at lower frequencies, and the corresponding time samples are spaced wider apart.

is smaller at lower frequencies, and the corresponding time-spacing is larger. Notice that 'frequency' spacing refers to the spacing between adjacent filters, and time-spacing refers to the sampling period used for the filtered outputs. The wavelet transform is not explicitly implemented by a moving window because there is in reality no unique window here. The system is in essence a filter bank, and is somewhat analogous to a family of windows (wider for low frequencies, etc.) as explained above.

General Definition of the Wavelet Transform

Equation (11.3.7a) is a special case of the more general definition of the continuous wavelet transformation (CWT) given in the literature, viz.,

$$X_{CWT}(p,q) = \frac{1}{\sqrt{|p|}} \int_{-\infty}^{\infty} x(t) f\Big(\frac{t-q}{p}\Big) dt, \tag{11.3.8}$$

where p and q are real-valued continuous variables. This reduces to (11.3.7a) if we identify

$$p = a^k, \quad q = a^k Tn, \quad \text{and} \quad f(t) = h(-t). \tag{11.3.9}$$

This choice is equivalent to evaluating (11.3.8) at a discrete set of points in the (p,q) plane, hence the name DWT for (11.3.7). DWT is different from discrete-time wavelet transforms (DTWT) to be discussed in Sec. 11.3.3. Quantities such as $X_{CWT}(p,q)$ and $X_{DWT}(k,n)$ are also called *wavelet coefficients.*

The CWT is a mapping of the function $x(t)$ into a two dimensional function $X_{CWT}(p,q)$ of the continuous variables p,q. The DWT is a mapping of $x(t)$ (t being still continuous-time) into a two dimensional *sequence* $X_{DWT}(k,n)$. The computation of $X_{DWT}(k,n)$ is equivalent to the implementation of the bank of filters $H_k(j\Omega)$, followed by sampling of their outputs at rates proportional to the filter bandwidths.

11.3.2 Inversion of the Wavelet Transform

The "inverse wavelet transform", if it exists, reconstructs the signal $x(t)$ from the wavelet coefficients. A direct inversion formula for (11.3.8) can be found in Daubechies [1990].

We will consider only the discretized case. Whether we can reconstruct $x(t)$ from the discretized version $X_{DWT}(k,n)$ depends on the prototype filter $h(t)$, and the discretizing parameters a and T which completely characterize the transformation. If the inverse transform exists, it has the appearance

$$x(t) = \sum_k \sum_n X_{DWT}(k,n)\psi_{kn}(t) \quad \text{(inverse DWT)}, \tag{11.3.10}$$

where $\psi_{kn}(t)$ are the basis functions.

Filter Bank Interpretation of Inversion

Suppose the wavelet coefficients $X_{DWT}(k,n)$ are generated using the analysis bank in Fig. 11.3-2(b). The reconstruction of $x(t)$ from these coefficients can be visualized as a problem of designing the synthesis filters $F_k(j\Omega)$ shown in Fig. 11.3-2(c). If the analysis/synthesis system has the perfect reconstruction property, then the recovery is perfect.

We have to be careful with the interpretation of Fig. 11.3-2(c). Since $X_{DWT}(k,n)$ is a *sequence,* the signal which is input to the *continuous-time* filter $F_k(j\Omega)$ is actually an impulse train of the form

$$v_k(t) = \sum_n X_{DWT}(k,n)\delta_a(t - a^k nT). \tag{11.3.11}$$

The output of the synthesis filter bank is therefore

$$\hat{x}(t) = \sum_k \sum_n X_{DWT}(k,n) f_k(t - a^k nT). \tag{11.3.12a}$$

Since the synthesis filters $F_k(j\Omega)$ have to retain only the frequency region passed by the analysis filters, it is reasonable that they be all generated from a fixed prototype synthesis filter $f(t)$, similar to (11.3.1). That is

$$f_k(t) = a^{-k/2} f(a^{-k}t). \tag{11.3.12b}$$

Substituting this into the preceding equation, and assuming perfect reconstruction, we get

$$x(t) = \sum_k \sum_n X_{DWT}(k,n) \underbrace{a^{-k/2} f(a^{-k}t - nT)}_{\text{basis } \psi_{kn}(t)}. \tag{11.3.12c}$$

Thus

$$\psi_{kn}(t) = a^{-k/2}\psi(a^{-k}t - nT) = a^{-k/2}\psi\left[a^{-k}(t - na^k T)\right], \tag{11.3.12d}$$

where we have defined

$$\psi(t) = f(t). \tag{11.3.12e}$$

Thus, Eqn. (11.3.12c) expresses $x(t)$ as a linear combination of a set of basis functions $\psi_{kn}(t)$ which are obtained by *dilations* (i.e., $t \to a^{-k}t$) and *shifts* (i.e., $t \to t - na^k T$) of a single *wavelet function* $\psi(t)$ or *mother wavelet.* The DWT coefficients $X_{DWT}(k,n)$ are the weights of these basis functions.

Using the relations $\psi(t) = f(t)$ and $f_k(t) = a^{-k/2} f(a^{-k}t)$, we can express each basis function $\psi_{kn}(t)$ in terms of the filter $f_k(t)$. Thus,

$$\begin{aligned} \psi_{kn}(t) &= a^{-k/2} f(a^{-k}t - nT) \\ &= a^{-k/2} f(a^{-k}(t - na^k T)) \\ &= f_k(t - na^k T), \end{aligned} \tag{11.3.12f}$$

which is a shifted version of the synthesis filter $f_k(t)$.

Orthonormal Basis

Of particular interest is the case where $\{\psi_{kn}(t)\}$ is a set of orthonormal functions. Such functions satisfy

$$\int_{-\infty}^{\infty} \psi_{kn}^*(t)\psi_{\ell m}(t)dt = \delta(k-\ell)\delta(n-m). \tag{11.3.13a}$$

Using Parseval's theorem, this becomes

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} \Psi_{kn}^*(j\Omega)\Psi_{\ell m}(j\Omega)d\Omega = \delta(k-\ell)\delta(n-m). \tag{11.3.13b}$$

By using the orthonormality property in (11.3.10) we obtain

$$X_{DWT}(k,n) = \int_{-\infty}^{\infty} x(t)\psi_{kn}^*(t)dt. \tag{11.3.13c}$$

Comparing with (11.3.7a) we conclude

$$\begin{aligned}\psi_{kn}(t) &= a^{-k/2}h^*(nT - a^{-k}t)\\ &= a^{-k/2}h^*\big(a^{-k}(na^kT - t)\big)\\ &= h_k^*(a^knT - t).\end{aligned} \tag{11.3.14a}$$

In particular, $\psi_{00}(t) = \psi(t) = h^*(-t)$. But we have $\psi(t) = f(t)$ so that, in the orthonormal case, $f(t) = h^*(-t)$. Thus,

$$f_k(t) = h_k^*(-t) \qquad \text{(orthonormal case)}. \tag{11.3.14b}$$

This is very similar to the relation (6.2.6) for the perfect reconstruction paraunitary QMF banks! (We could let $c = 1$ and $L = 0$ in (6.2.6) without affecting any significant conclusions of Sec. 6.2).

Table 11.3.1 summarizes the definition and the main features of the wavelet transform.

Completeness, uniqueness, and so forth

Given an arbitrary function $x(t)$, suppose we compute $X_{DWT}(k,n)$ using (11.3.7a). Can we then express $x(t)$ as in (11.3.10)? That is, can we invert the transformation? The answer depends on $\psi(t)$, and the discretization parameters T and a. If this is possible for a specified class of functions $\{x(t)\}$, then the wavelet basis $\{\psi_{kn}(t)\}$ is said to be *complete* over this class. In Sec. 11.5 we will see how to generate a complete orthonormal basis for the class of finite energy functions.

Another practical requirement in addition to completeness is that, the transformation and reconstruction formulas (11.3.7a) and (11.3.10) should be 'stable' so that the computations do not "blow up." In the next section where we study discrete time wavelets, we will see that the issues of completeness and stability are much easier to address.

TABLE 11.3.1 The continuous-time wavelet transform

Traditional Continuous-Time Fourier Transform

$$X(j\Omega) = \int_{-\infty}^{\infty} x(t)e^{-j\Omega t}dt \qquad \text{(Fourier transform)} \tag{2.1.20}$$

$$x(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} X(j\Omega)e^{j\Omega t}d\Omega \qquad \text{(inverse transform)} \tag{2.1.21}$$

Wavelet transform, general

$$X_{CWT}(p,q) = \frac{1}{\sqrt{|p|}}\int_{-\infty}^{\infty} x(t)f\Big(\frac{t-q}{p}\Big)dt \tag{11.3.8}$$

Discrete Wavelet Transform, DWT (still continuous-time)

Obtained by setting $p = a^k$, $q = a^kTn$, $f(t) = h(-t)$ in (11.3.8), for integer k, n. (We can take $T = 1$ for simplicity.)

$$X_{DWT}(k,n) = a^{-k/2}\int_{-\infty}^{\infty} x(t)h(nT - a^{-k}t)dt \quad \text{(DWT)}, \tag{11.3.7a}$$

$$x(t) = \sum_k \sum_n X_{DWT}(k,n)\underbrace{a^{-k/2}f(a^{-k}t - nT)}_{\text{basis } \psi_{kn}(t)} \quad \text{(inverse DWT)}. \tag{11.3.12c}$$

The inversion formula assumes that the filter bank of Fig. 11.3-2 has the perfect reconstruction property, with the filters chosen as

$$\underbrace{h_k(t) = a^{-k/2}h(a^{-k}t)}_{\text{analysis filters}}, \qquad \underbrace{f_k(t) = a^{-k/2}f(a^{-k}t)}_{\text{synthesis filters}}$$

Thus, the functions $h(t)$ and $f(t)$ in (11.3.7a) and (11.3.12c) play the role of prototype filters in a filter bank where all the filters are derived by *dilation* of a single filter. The basis functions in (11.3.12c) are dilated ($t \to a^{-k}t$) and shifted ($t \to t - na^kT$) versions of $f(t)$, that is,

$$\psi_{kn}(t) = a^{-k/2}\psi(a^{-k}t - nT), \tag{11.3.12d}$$

where $\psi(t) = f(t)$ = wavelet function or mother wavelet.

Continued →

Special case of orthonormal basis functions

$$\int_{-\infty}^{\infty} \psi_{kn}^*(t)\psi_{\ell m}(t)dt = \delta(k-\ell)\delta(n-m) \quad \text{(orthonormality)}. \tag{11.3.13a}$$

Under this condition,

$$X_{DWT}(k,n) = \int_{-\infty}^{\infty} x(t) \underbrace{a^{-k/2} f^*(a^{-k}t - nT)}_{\psi_{kn}^*(t)} dt, \tag{11.3.13c}$$

$$x(t) = \sum_k \sum_n X_{DWT}(k,n) \underbrace{a^{-k/2} f(a^{-k}t - nT)}_{\text{basis } \psi_{kn}(t)}. \tag{11.3.12c}$$

Again, $\psi_{00}(t) = f(t) = \psi(t) =$ wavelet function.

Filter Bank Properties in the Orthonormal Case

1. Synthesis filters for perfect reconstruction: $f_k(t) = h_k^*(-t)$.
2. Relation between prototype filters: $f(t) = h^*(-t)$.

11.3.3 Discrete-Time Wavelet Transforms

We now extend the wavelet transformation to the case of discrete-time signals. The starting point again is a set of filters with frequency responses having an appearance similar to Fig. 11.3-2(a). In the continuous-time case these filters were related as in (11.3.1). If we attempt to mimic this by replacing t with the discrete-time index n, the quantity $a^{-k}n$ does not in general remain an integer. Let us, therefore, try to imitate the frequency domain relation (11.3.2) rather than the time domain relation.

Consider the example $a = 2$. The equivalent of (11.3.2) for digital filters would be

$$H_k(e^{j\omega}) = H(e^{j2^k\omega}), \tag{11.3.15}$$

that is, $H_k(z) = H(z^{2^k})$ where k is a nonnegative integer. For highpass $H(e^{j\omega})$, the responses of $H_k(e^{j\omega})$ for $k = 1$ and $k = 2$ are shown in Fig. 11.3-4. This shows that in general $H_k(z)$ is a multiband (rather than bandpass) filter, and further modification is required to obtain bandpass responses. For this we cascade $H_k(z)$ with appropriate filters. Thus, let $G(z)$ be a lowpass

filter with response as in Fig. 11.3-5(a). Then the responses of

$$H(z),\ G(z)H(z^2),\ G(z)G(z^2)H(z^4),\ \ldots \tag{11.3.16}$$

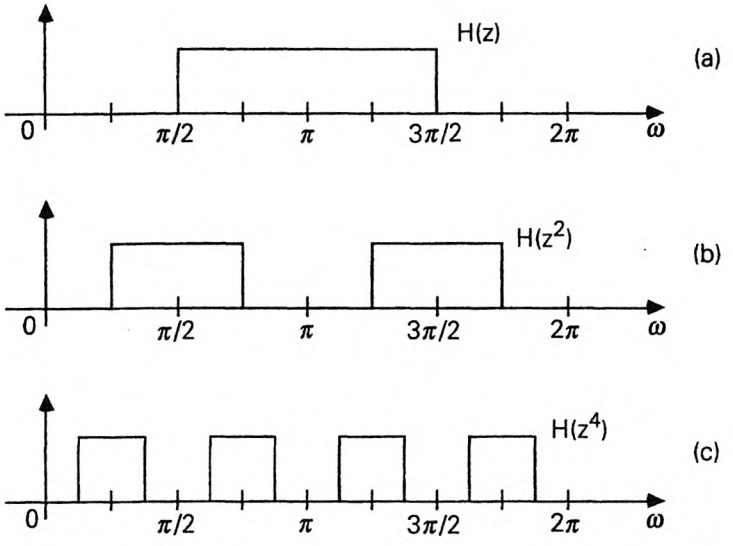

Figure 11.3-4 Magnitude response plots of (a) $H(z)$, (b) $H(z^2)$, and (c) $H(z^4)$.

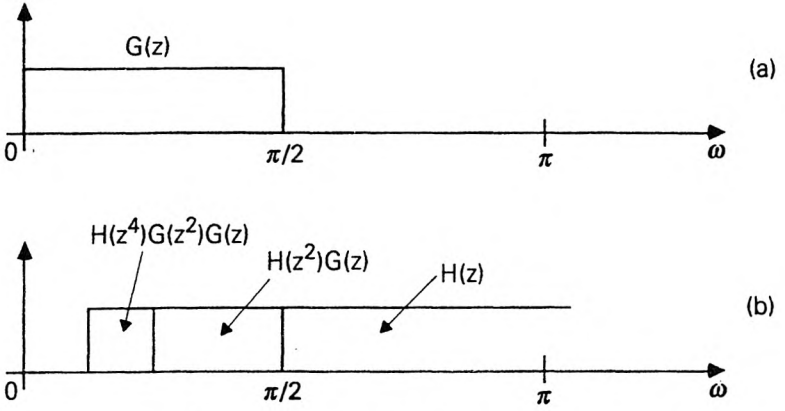

Figure 11.3-5 Magnitude responses of (a) lowpass $G(z)$, and (b) combinations of $H(z)$ and $G(z)$.

are as shown in Fig. 11.3-5(b). These resemble Fig. 11.3-2(a) (except for the heights, which can be adjusted to make all the filter energies equal; see Fig. 11.3-8 later). The plots are shown only for $0 \leq \omega < \pi$ as we assume, for simplicity, magnitude symmetry with respect to π. The filters are bandpass, with center frequencies

$$\omega_k = c \times 2^{-k}\pi, \quad 0 \leq k \leq M-1, \tag{11.3.17}$$

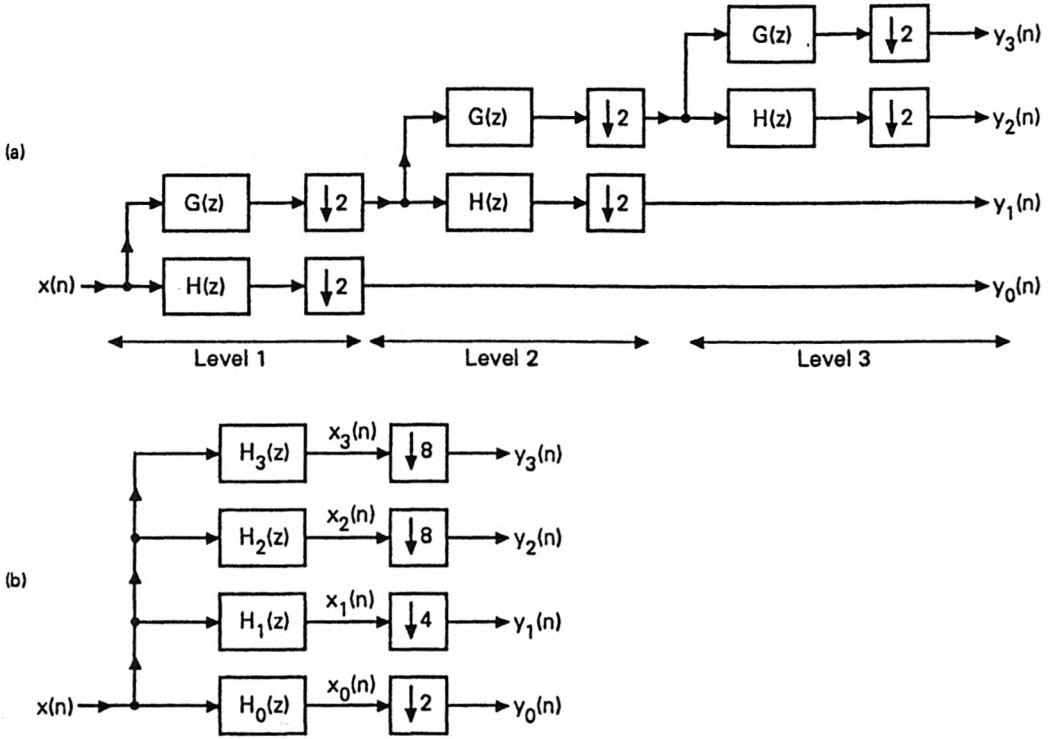

Figure 11.3-6 (a) A 3-level binary tree-structured QMF bank and (b) the equivalent four-channel system.

for appropriate c, and passband widths $BW_k = 2^{-k}\pi/2$ (measured only in the range $[0, \pi]$). Thus, the ratio BW_k/ω_k is independent of k.

From our experience with QMF banks (Sec. 5.8) we already know that filters of the form (11.3.16) can be generated with the help of a binary tree structure. Fig. 11.3-6 shows a three-level tree structure along with the equivalent nontree form, which has four filters. More generally if the tree has L levels, the number of channels is $M = L+1$. The signals $x_k(n)$ can be

decimated by the amounts shown, in order to obtain a maximally decimated system precisely as in a QMF bank. The decimation ratios are also consistent with the bandwidths of the signals $x_k(n)$. As in any multirate system with nonideal filters, decimation introduces aliasing. By using the techniques of Chap. 5 this can be canceled with an appropriate synthesis bank.

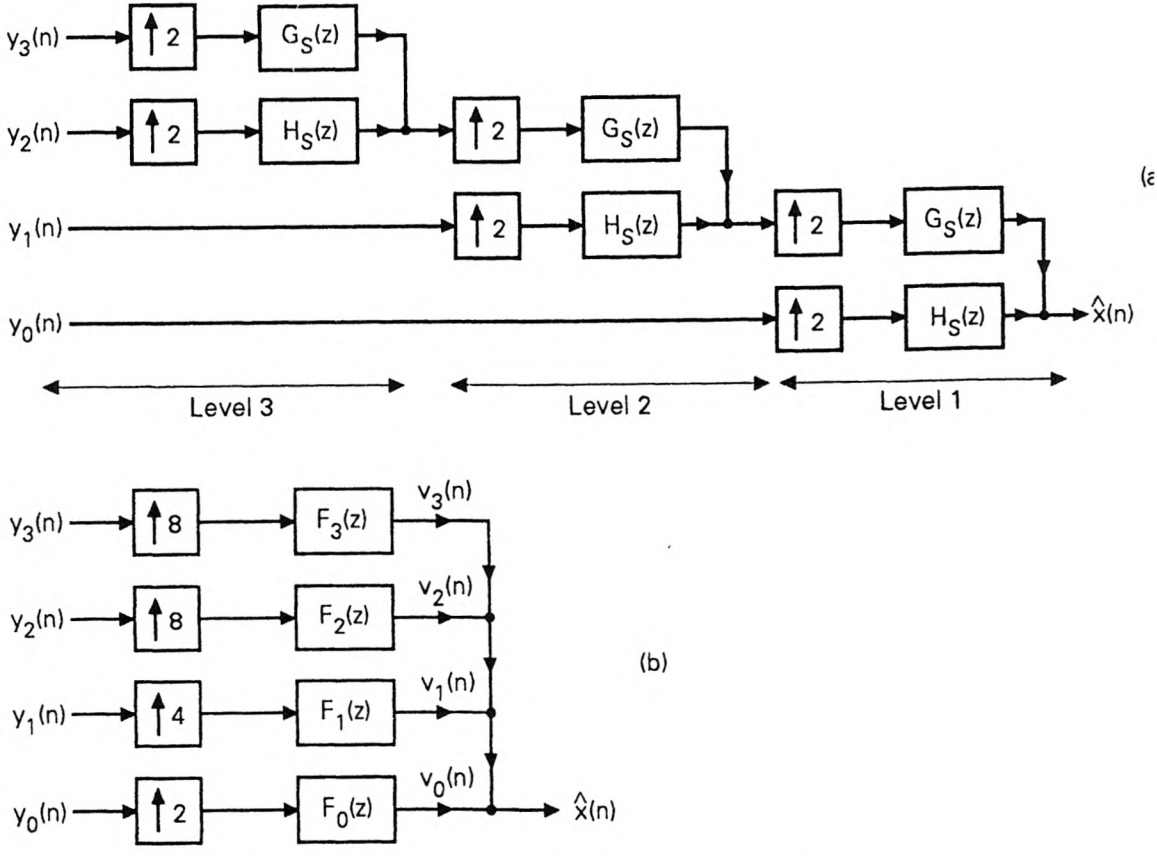

Figure 11.3-7 (a) The synthesis bank corresponding to Fig. 11.3-6 and (b) equivalent four-channel system.

Defining the Discrete-Time Wavelet Transform

In the z-domain we have $X_k(z) = H_k(z)X(z)$ so that

$$x_k(n) = \sum_m x(m)h_k(n-m). \tag{11.3.18}$$

The decimated signals $y_k(n)$ are the *wavelet coefficients*, and the wavelet

transform is given by

$$y_k(n) = \sum_{m=-\infty}^{\infty} x(m)h_k(2^{k+1}n - m), \quad 0 \le k \le M-2,$$
$$y_{M-1}(n) = \sum_{m=-\infty}^{\infty} x(m)h_{M-1}(2^{M-1}n - m) \quad \text{(DTWT)}. \tag{11.3.19}$$

This is analogous to the situation in Fig. 11.2-8 where we obtained the STFT coefficients using a multirate filter bank. Equation (11.3.19) is the discrete-time wavelet transform (DTWT).

The inverse transform. The inversion of the above transform can be performed by designing an appropriate synthesis bank. Consider the synthesis filter bank of Fig. 11.3-7(a). This is equivalent to Fig. 11.3-7(b) with filters expressible entirely in terms of $G_s(z)$ and $H_s(z)$. For example,

$$\begin{aligned} F_0(z) &= H_s(z), \\ F_1(z) &= H_s(z^2)G_s(z), \\ F_2(z) &= H_s(z^4)G_s(z^2)G_s(z), \end{aligned} \tag{11.3.20}$$

and so on. We know from Sec. 5.8 that if the filters $G(z), H(z), G_s(z)$ and $H_s(z)$ are appropriately designed, the tree structured system produces perfect reconstruction, that is, $\widehat{x}(n) = x(n)$. Under this condition we can express

$$\begin{aligned} X(z) = F_0(z)Y_0(z^2) + F_1(z)Y_1(z^4) + \ldots + F_{M-2}(z)Y_{M-2}(z^{2^{M-1}}) \\ + F_{M-1}(z)Y_{M-1}(z^{2^{M-1}}) \qquad \text{(inverse DTWT)}. \end{aligned} \tag{11.3.21}$$

In Eqn. (4.1.22) we showed how to express the output of an interpolation filter in the time domain. Using similar principles, the above expression for inverse DTWT can be written in the time domain as

$$\begin{aligned} x(n) = \sum_{k=0}^{M-2} \sum_{m=-\infty}^{\infty} y_k(m) \underbrace{f_k(n - 2^{k+1}m)}_{\eta_{km}(n)} \\ + \sum_{m=-\infty}^{\infty} y_{M-1}(m) \underbrace{f_{M-1}(n - 2^{M-1}m)}_{\eta_{M-1,m}(n)}, \end{aligned} \tag{11.3.22a}$$

that is,

$$x(n) = \sum_{k=0}^{M-1} \sum_{m=-\infty}^{\infty} y_k(m)\eta_{km}(n) \quad \text{(inverse DTWT)}. \tag{11.3.22b}$$

Here $\eta_{km}(n)$ are the wavelet basis functions, and the weights $y_k(m)$ are the wavelet coefficients of $x(n)$ with respect to the above basis. Notice that $\eta_{k0}(n) = f_k(n) =$ synthesis filters. The basis function $\eta_{km}(n)$ is the impulse response $f_k(n)$ *shifted* by an appropriate amount.

Relation to Multiresolution Components

Refer to the analysis/synthesis system in Figs. 11.3-6 and 11.3-7. Assuming that $H_3(z)$ and $F_3(z)$ are good lowpass filters, the signal $v_3(n)$ is a lowpass filtered version of $x(n)$. (This is only approximately so, because of aliasing and imaging caused by the decimation and interpolation operations). The signal $v_2(n)$, on the other hand is a bandpass filtered version, and adds finer high frequency details. Thus, we can regard $v_3(n)$ to be a lowpass (i.e., smoothed) approximation of $x(n)$, whereas the sum $v_3(n) + v_2(n)$ is a 'higher resolution' approximation. By adding the component $v_1(n)$ we get a further refined approximation. Finally when $v_0(n)$ (the finest 'detail signal') is added, we obtain perfect recovery of $x(n)$. The tree structure (or wavelet decomposition) can therefore be used to transmit information (e.g., a picture in video conferencing) in various installments, with successively improved fine details.

Some Practical Requirements

1. *Stability.* In practice we require the filters $H_k(z)$ and $F_k(z)$ to be stable (equivalently $H(z), G(z), H_s(z)$ and $G_s(z)$ stable). This ensures that the procedure (11.3.19) to construct the wavelet coefficients, as well as the inversion procedure (11.3.22b) are stable.
2. *Orthonormality.* It is also desirable to have an orthonormal set of basis functions $\eta_{km}(n)$. This means

$$\sum_{n=-\infty}^{\infty} \eta_{km}(n)\eta^*_{\ell i}(n) = \delta(k-\ell)\delta(m-i). \tag{11.3.23}$$

By using this in (11.3.22) we verify that orthonormality implies

$$y_k(m) = \sum_{n=-\infty}^{\infty} x(n)\eta^*_{km}(n). \tag{11.3.24}$$

By comparing (11.3.19), (11.3.22a), and (11.3.24) we can eliminate $\eta_{km}(n)$ and obtain the relation

$$f_k(n) = h^*_k(-n). \tag{11.3.25}$$

Thus, the analysis and synthesis filters are related as above when the basis is orthonormal. This is similar to the relation between filters in a *paraunitary* PR QMF bank! (Sec. 6.2). In the next section, we

will present the precise relation between paraunitary QMF banks and orthonormal wavelets.

3. *Height conventions.* The increasing heights in Fig. 11.3-8 are chosen such that the energies of the filters are equal. This, however, is only a convention. It should be realized that the output $y_k(n)$ of the filter $H_k(z)$ is an estimate of $X(e^{j\omega_k})$ (with no scale-factor discrepancies) only if the heights of $|H_k(e^{j\omega})|$ in their passbands are inversely proportional to the bandwidths (i.e., heights are $c, 2c, 4c, \ldots$). See Problem 11.1. This requirement is *not consistent* with Fig. 11.3-8 (where the heights are $c, \sqrt{2}c, 2c \ldots$). In this chapter, conventions for heights will be flexible, depending on the particular context.

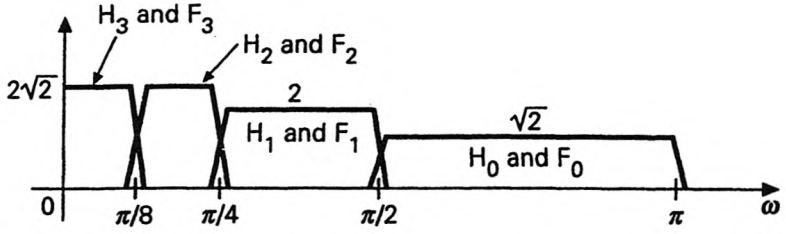

Figure 11.3-8 Typical appearances of magnitude responses of filters in the 3-level tree.

11.3.4 Summary

The STFT

The short time Fourier transform system is reproduced in Fig. 11.3-9. This is an M channel filter bank with equal-bandwidth filters having equispaced center frequencies. All the filters are generated from a prototype $H_0(z)$ as in (11.2.5). $E_k(z)$ are the polyphase components of this prototype as shown by (11.2.6). The system can be viewed as the sliding window system (Fig. 11.2-1) with $H_0(e^{j\omega}) = V(e^{-j\omega})$.

The quantity $|x_k(n)|$ represents the estimate of the magnitude of the transform of $x(i)$, *around* the center-frequency of $H_k(e^{j\omega})$, with the data $x(i)$ 'localized' *around* time n (which is the window position). As shown in earlier figures, each of the outputs $x_k(n)$ can be decimated by a factor $\leq M$. For any fixed n, the set of values

$$|x_k(n)|, \quad 0 \leq k \leq M-1, \tag{11.3.26}$$

i.e., the vector $[\,x_0(n) \quad \ldots \quad x_{M-1}(n)\,]^T$, provides a 'snapshot' of the magnitude of the Fourier transform of $x(i)$, localized around time n. The snapshot is delivered as a *uniformly* sampled version (in the frequency domain). The figure demonstrates this sampling for $n = -1, 3$, and 7. In this demonstration, we see that the signal changes slowly from lowpass to highpass. Thus, the STFT keeps track of the evolution of the Fourier transform. As shown in

the figure the set of filters can be either sharp-cutoff, or highly overlapping. The latter happens, for example, when $E_k(z) = 1$ for all k; this corresponds to the computation of DFT of blocks of the input.

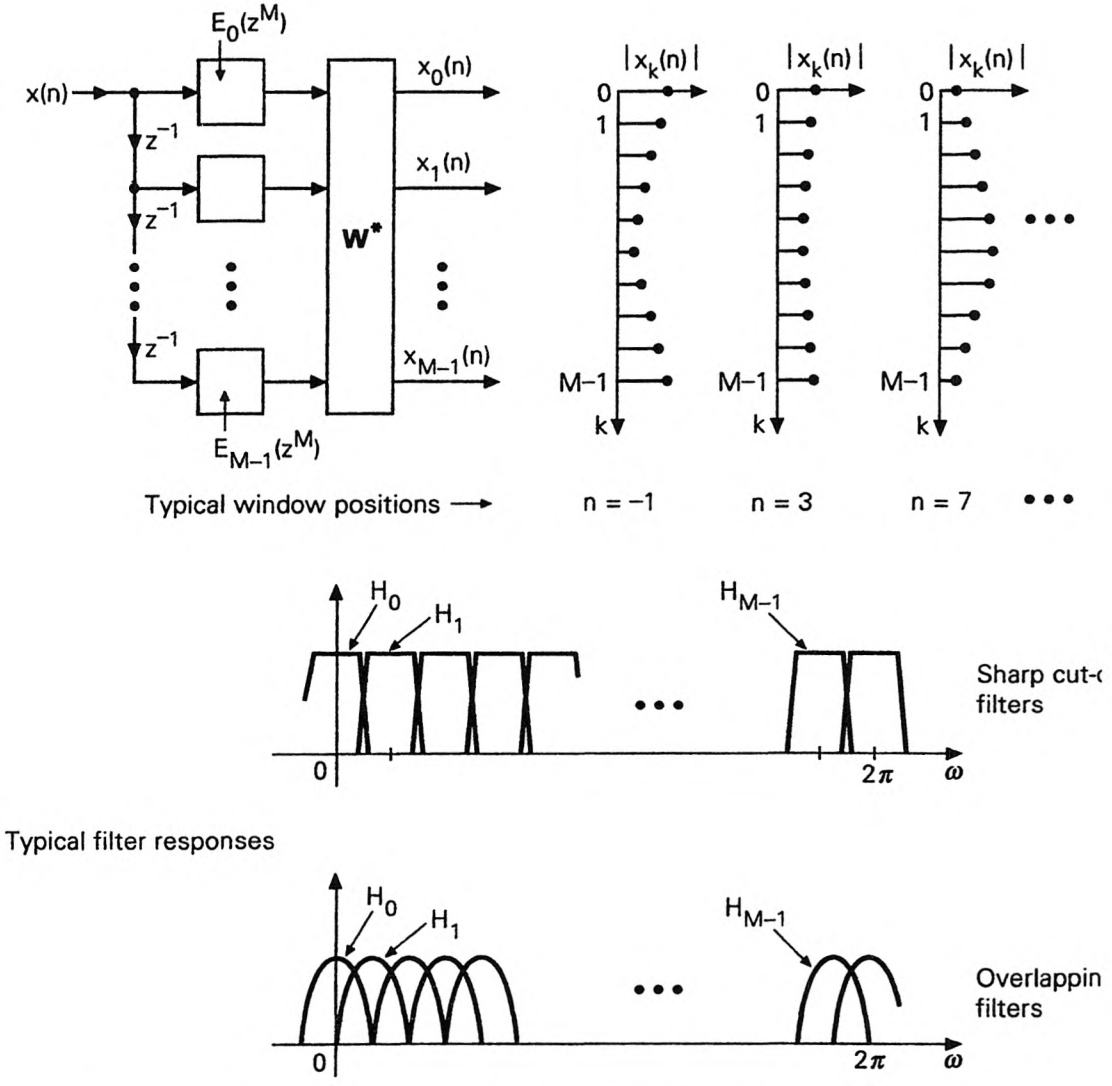

Figure 11.3-9 Uniform DFT filter banks as short-time Fourier transformers. For each window position n, we obtain a 'snapshot' of localized Fourier transform.

The Discrete-Time Wavelet Transform

The wavelet system (nonuniform filter bank) is reproduced in Fig. 11.3-10. This is an M-channel filter bank with nonuniform bandwidths for the filters. All the filters are generated from a tree structure as in Fig. 11.3-6.

The system cannot be viewed as the sliding window system, but one can imagine that there is an underlying window whose width is adjusted according to frequency. The outputs of the filters (bandpass signals) are modulated to obtain $x_k(n)$ which are *lowpass* signals, with *increasing* bandwidths as k decreases.

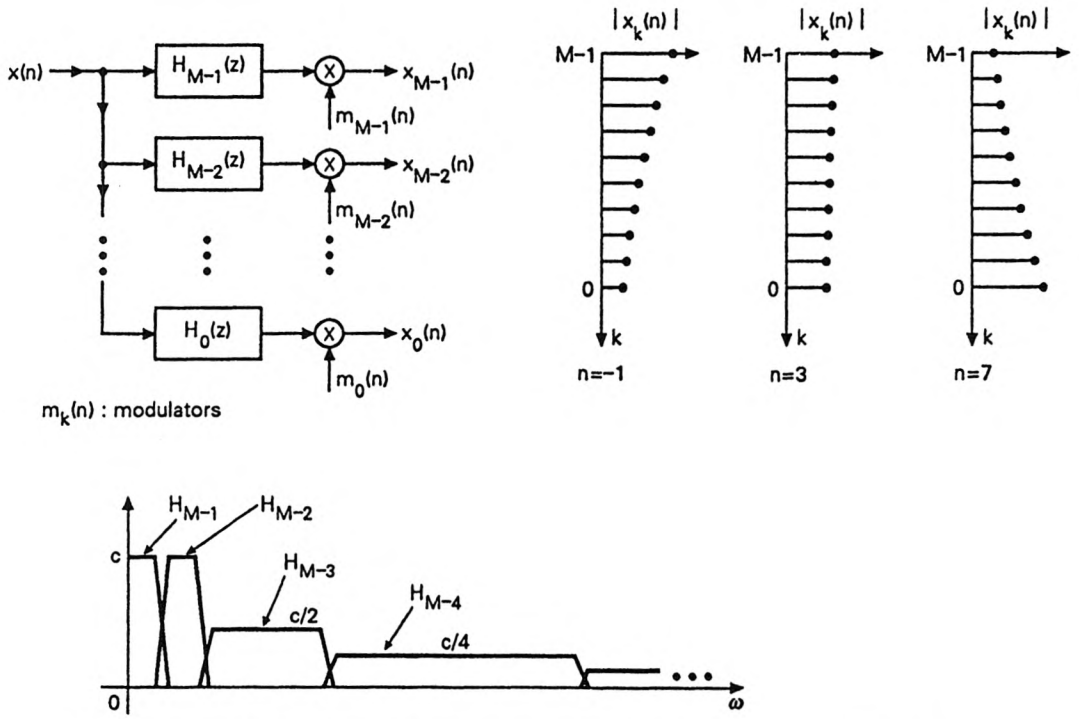

Figure 11.3-10 Summarizing the operation of a wavelet filter bank.

Figure 11.3-10 also shows the responses of the nonuniformly spaced filters. The quantity $x_k(n)$ represents the estimate of the transform of $x(i)$, *around* the center-frequency of $H_k(e^{j\omega})$, with the data $x(i)$ "localized" *around* time n. Even though all the signals $x_k(n)$ are lowpass, their bandwidths are different. Thus $x_1(n)$ varies "more slowly" than $x_0(n)$, and so forth. So the signals are decimated by unequal amounts (in fact by a factor inversely proportional to the filter bandwidth).

For any fixed n, the set of values

$$x_k(n), \quad 0 \leq k \leq M-1, \tag{11.3.27}$$

provides a 'snapshot' of the nonuniformly sampled version of the Fourier transform of $x(i)$, localized around n. (This assumes that the heights of the

filters are inversely proportional to the bandwidths, see Problem 11.1). The figure demonstrates this for $n = -1, 3$, and 7.

General Comments

1. If the analysis/synthesis system (i.e., Fig. 11.3-6 followed by Fig. 11.3-7) has the perfect reconstruction property $\hat{x}(n) = x(n)$, then (11.3.22b) represents the expansion of $x(n)$ in terms of the wavelet basis functions $\eta_{km}(n)$. If (11.3.23) holds, then the basis is orthonormal.
2. *FIR wavelet basis.* If the filters $H_s(z)$ and $G_s(z)$ are FIR, then the basis functions $\eta_{km}(n)$ are finite-length sequences.
3. When the QMF system satisfies the perfect reconstruction property, the basis is *complete* in the sense that any $x(n)$ can be expressed in this manner. Since we know that there exist FIR perfect reconstruction systems, this shows how to obtain a complete FIR wavelet basis (though $x(n)$ may not be FIR).
4. *The disk-partition diagram.* Figure 11.3-11(a) explains how the frequency domain is partitioned when performing the wavelet analysis. For simplicity we have shown only the upper half of the unit circle in the z-plane. (If the filters have real coefficients, the lower half need not be shown). The first level of the tree partitions the half-circle into two quarter circles. The second level splits the low-frequency quarter circle into two equal halfs. This process is repeated L times in an L-level tree. For comparison, Fig. 11.3-11(b) shows the frequency partition for the case of uniform bandwidth filter banks, (e.g., STFT). Here the circle is divided into wedges of equal size. In both methods, the output of each filter is decimated in inverse proportion to the angular width of the wedge.

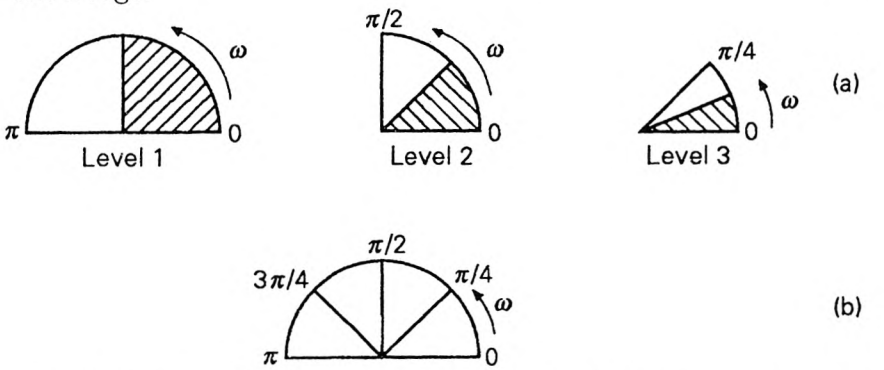

Figure 11.3-11 The disk partitioning diagram. (a) Wavelet transform, and (b) short-time Fourier transform.

11.4 DISCRETE-TIME ORTHONORMAL WAVELETS

From the previous sections we know that we can obtain a wavelet decomposition of a sequence $x(n)$ by using a tree structured perfect reconstruction

QMF bank of Figs. 11.3-6, 11.3-7. The coefficients $f_k(n)$ of the synthesis filters $F_k(z)$ govern the basis functions (see eqn. (11.3.22a)) whereas the decimated outputs $y_k(n)$ are the wavelet coefficients. Since the perfect reconstruction property holds for any $x(n)$, the expansion (11.3.22a) holds for any $x(n)$.

♠**Main points of this section.** We will show that if the filters $G_s(z)$ and $H_s(z)$ in the synthesis bank have the paraunitary property, then the basis functions $\{\eta_{km}(n)\}$ are orthonormal. Our development is in two steps. In Sec. 11.4.1 we will prove this result for a one-level tree (i.e., just a two channel QMF bank). In Sec. 11.4.2 this will be extended to an arbitrary number of levels. Finally in Sec. 11.4.3 a similar result will be proved for an M channel system with identical decimation ratio in all channels.

11.4.1 Two Channel Paraunitary QMF Banks

The fundamental building block in Figs. 11.3-6, 11.3-7 is the two-channel QMF bank reproduced in Fig. 11.4-1. Here $H_0(z)$ and $H_1(z)$ are typically lowpass and highpass, respectively, like $G(z)$ and $H(z)$ in Fig. 11.3-5. [In Fig. 11.3-6(a) we used the notations $G(z)$ and $H(z)$ instead of $H_0(z)$ and $H_1(z)$ in order to avoid confusion with Fig. 11.3-6(b)]. Here the two filters have equal bandwidth. We can think of this system as a simple special case of wavelet decomposition. Thus the wavelet coefficients are

$$y_k(n) = \sum_{m=-\infty}^{\infty} x(m) h_k(2n-m), \quad k = 0, 1. \tag{11.4.1}$$

Assuming that the synthesis bank gives perfect reconstruction, we can express $x(n)$ as

$$x(n) = \sum_{m=-\infty}^{\infty} y_0(m) \underbrace{f_0(n-2m)}_{\eta_{0m}(n)} + \sum_{m=-\infty}^{\infty} y_1(m) \underbrace{f_1(n-2m)}_{\eta_{1m}(n)}. \tag{11.4.2}$$

The wavelet bases $\eta_{km}(m)$ are indicated. These are stable if $F_k(z)$ are stable, and FIR if $F_k(z)$ are FIR.

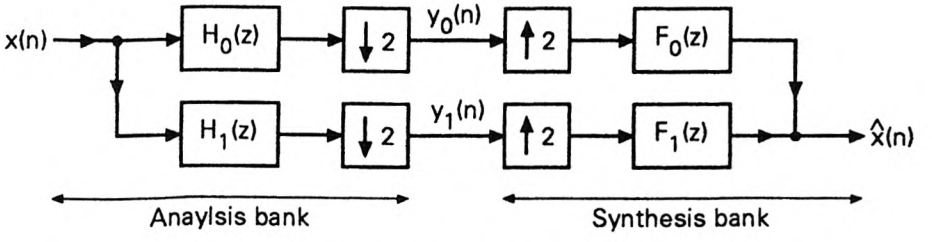

Figure 11.4-1 The two-channel QMF bank.

Orthonormality of Wavelet Basis

The wavelet basis is said to be orthonormal if (11.3.23) holds. Substituting for $\eta_{km}(n)$ from (11.4.2), this can be rewritten as

$$\sum_{n=-\infty}^{\infty} f_k(n-2m) f_\ell^*(n-2i) = \delta(k-\ell)\delta(m-i), \tag{11.4.3}$$

that is, after a change of variables as

$$\sum_{n=-\infty}^{\infty} f_k(n) f_\ell^*(n-2m) = \delta(k-\ell)\delta(m). \tag{11.4.4}$$

This means that $f_k(n)$ is orthogonal to the even-shifted versions of $f_\ell(n)$. Now the left side of this equation is the cross-correlation between $f_k(n)$ and $f_\ell(n)$ evaluated for even-lag $2m$ (Problem 2.14). Since the z-transform of the cross-correlation function is $F_k(z)\widetilde{F}_\ell(z)$, we can rewrite (11.4.4) in the z-domain as follows:

$$F_k(z)\widetilde{F}_\ell(z)\Big|_{\downarrow 2} = \delta(k-\ell). \tag{11.4.5}$$

As explained in Sec. 4.1, the notation $A(z)\Big|_{\downarrow M}$ is an abbreviation which indicates decimation by M, for example,

$$A(z)\Big|_{\downarrow 2} = \frac{1}{2}[A(z^{1/2}) + A(-z^{1/2})]. \tag{11.4.6}$$

We can rewrite (11.4.5) as

$$\widetilde{F}_k(z)F_\ell(z) + \widetilde{F}_k(-z)F_\ell(-z) = 2\delta(k-\ell), \quad 0 \le k,\ell \le 1, \tag{11.4.7}$$

that is,

$$\widetilde{\mathbf{F}}(z)\mathbf{F}(z) = 2\mathbf{I} \quad \text{(wavelet orthonormality condition)}, \tag{11.4.8}$$

where

$$\mathbf{F}(z) \triangleq \begin{bmatrix} F_0(z) & F_1(z) \\ F_0(-z) & F_1(-z) \end{bmatrix}. \tag{11.4.9}$$

Thus orthonormality of the wavelet basis is equivalent to the paraunitary condition (11.4.8). Notice, in particular, that this implies the power complementary property $|F_0(e^{j\omega})|^2 + |F_1(e^{j\omega})|^2 = 2$.

Paraunitariness of R(z) and Wavelet Orthonormality

Recall (Sec. 5.5) that the two synthesis filters $F_0(z)$ and $F_1(z)$ can be expressed in terms of their 2×2 polyphase matrix $\mathbf{R}(z)$ in the form

$$[F_0(z) \quad F_1(z)] = [z^{-1} \quad 1]\mathbf{R}(z^2). \tag{11.4.10a}$$

Using this relation, we can express the matrix $\mathbf{F}(z)$ as

$$\mathbf{F}(z) = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} z^{-1} & 0 \\ 0 & 1 \end{bmatrix} \mathbf{R}(z^2). \tag{11.4.10b}$$

The above equation implies

$$\widetilde{\mathbf{F}}(z)\mathbf{F}(z) = 2\widetilde{\mathbf{R}}(z^2)\mathbf{R}(z^2). \tag{11.4.11}$$

Thus, the discrete-time wavelet orthonormality condition holds if and only if

$$\widetilde{\mathbf{R}}(z)\mathbf{R}(z) = \mathbf{I}, \tag{11.4.12}$$

which of course is the well-known (normalized-) paraunitary condition (Sec. 6.1.1)!

Designs Which Satisfy Orthonormality

In Sec. 5.3.6 we introduced an FIR perfect reconstruction QMF bank (invented independently in Smith and Barnwell [1984] and Mintzer [1985]). In this system, the FIR filter $H_0(z)$ is power symmetric, that is, satisfies

$$\widetilde{H}_0(z)H_0(z) + \widetilde{H}_0(-z)H_0(-z) = 2.$$

Under this condition we showed that if the remaining filters are chosen according to

1. $H_1(z) = -z^{-N}\widetilde{H}_0(-z)$, where N = order of $H_0(z)$, and
2. $F_0(z) = \widetilde{H}_0(z)$ and $F_1(z) = \widetilde{H}_1(z)$,

then we have perfect reconstruction, with $\hat{x}(n) = x(n)$. In Sec. 6.3.2 we saw that the above choice of filters ensures that the polyphase matrices $\mathbf{E}(z)$ and $\mathbf{R}(z)$ are paraunitary. This, therefore, ensures that the wavelet basis is orthonormal! Summarizing, the procedure to obtain a finite duration (FIR) orthonormal wavelet basis is as follows.

1. Design the Nth order FIR power symmetric filter $H_0(z)$. This is done either by starting from a zero-phase half band filter and computing a spectral factor $H_0(z)$ (Sec. 5.3.6), or equivalently by optimizing the lattice structure of Fig. 6.4-1. (Sec. 6.4.3).
2. Define the second analysis filter $H_1(z)$ and the synthesis filters $F_0(z)$ and $F_1(z)$ as above. Then the polyphase matrix $\mathbf{R}(z)$ satisfies the paraunitary condition $\widetilde{\mathbf{R}}(z)\mathbf{R}(z) = \mathbf{I}$. Notice that the synthesis filters are noncausal, but this is consistent with the delay-free reconstruction.

3. Define the wavelet basis as indicated in (11.4.2). Then the wavelet orthonormality condition is satisfied.
4. The analysis/synthesis system has the perfect reconstruction property, that is, $\hat{x}(n) = x(n)$. So, (11.4.2) holds, and represents the expansion of the arbitrary input sequence $x(n)$ in terms of the orthonormal wavelet basis functions $\eta_{km}(n)$.

Completeness. Since any FIR power symmetric filter can be generated using the lattice mentioned in Step 1, any two-channel FIR orthonormal basis can be generated using this lattice.

11.4.2 Orthonormal Wavelets from Tree-Structured Paraunitary QMF Banks

Now consider an L-level tree-structured QMF bank (as demonstrated in Figs. 11.3-6, 11.3-7 for $L = 3$). The wavelet basis functions $\eta_{km}(n)$ are indicated in (11.3.22a) in terms of the filter coefficients $f_k(i)$. We now show that these basis functions are orthonormal if each of the two channel systems $[G_s(z), H_s(z)]$ has polyphase matrix $\mathbf{R}(z)$ satisfying $\widetilde{\mathbf{R}}(z)\mathbf{R}(z) = \mathbf{I}$.

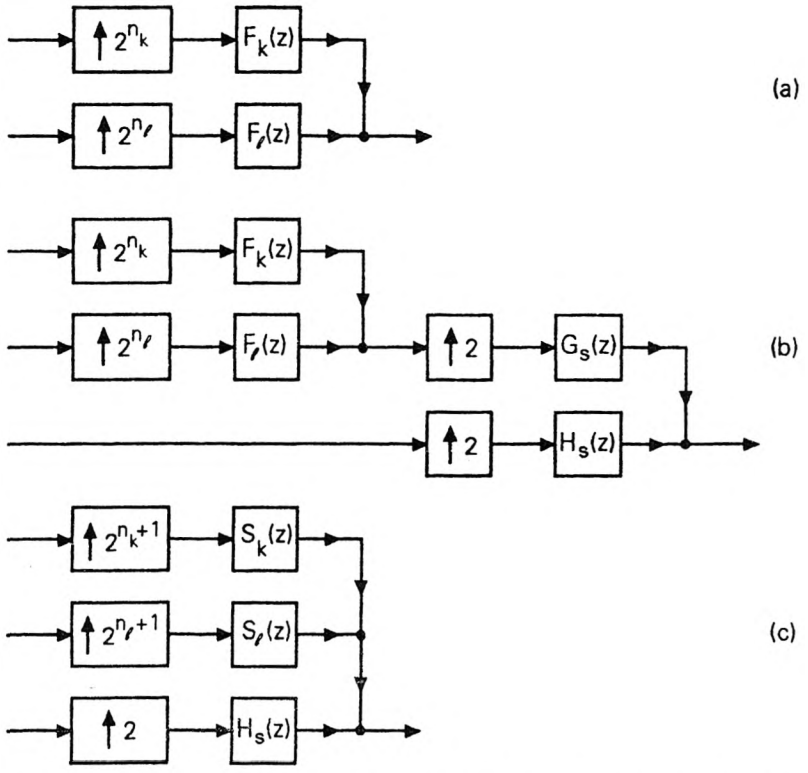

Figure 11.4-2 (a) Two of the $L+1$ branches in the synthesis bank of an L-level tree; (b) adding the $(L+1)$th level; and (c) redrawing the three branches.

From the previous section we know this to be true for $L = 1$. For arbitrary L we use an inductive reasoning. The tree structure with L levels has $L+1$ branches. Fig. 11.4-2(a) shows two of these branches, with $n_k \geq n_\ell$. Suppose we add another level to the tree. This adds a new branch, and modifies the existing branches as shown in Fig. 11.4-2(b). Assuming that

1. the wavelet bases are orthonormal for the L-level tree, and that
2. the new set of filters $[G_s(z), H_s(z)]$ has polyphase matrix $\mathbf{R}(z)$ satisfying $\widetilde{\mathbf{R}}(z)\mathbf{R}(z) = \mathbf{I}$,

we prove that the wavelet bases for the $(L+1)$-level tree are orthonormal.

From Sec. 6.3.1 recall that the paraunitary relation $\widetilde{\mathbf{R}}(z)\mathbf{R}(z) = \mathbf{I}$ is equivalent to the conditions

$$G_s(z)\widetilde{G}_s(z)\Big|_{\downarrow 2} = 1, \quad H_s(z)\widetilde{H}_s(z)\Big|_{\downarrow 2} = 1, \quad G_s(z)\widetilde{H}_s(z)\Big|_{\downarrow 2} = 0. \tag{11.4.13}$$

Instead of saying that $\mathbf{R}(z)$ is paraunitary, we will often say that the filter pair $[G_s(z), H_s(z)]$ is paraunitary (i.e., (11.4.13) holds).

Expressing Wavelet Orthonormality in z-Domain

Orthonormality of wavelets for the L-level tree implies

$$\sum_{n=-\infty}^{\infty} f_k(n - 2^{n_k}m) f_\ell^*(n - 2^{n_\ell}i) = \delta(k-\ell)\delta(m-i). \tag{11.4.14}$$

After a change of variables this can be rearranged as

$$\sum_{n=-\infty}^{\infty} f_k(n) f_\ell^*(n - 2^{n_\ell}m) = \delta(k-\ell)\delta(m), \tag{11.4.15}$$

using $2^{n_k} \geq 2^{n_\ell}$ (see Problem 11.13).

The summation on the left hand side of (11.4.15) is the cross-correlation between $f_k(n)$ and $f_\ell(n)$, evaluated at lags $2^{n_\ell}m$. Since the z-transform of the cross-correlation sequence is $F_k(z)\widetilde{F}_\ell(z)$, we can rephrase the above as

$$F_k(z)\widetilde{F}_\ell(z)\Big|_{\downarrow 2^{n_\ell}} = \delta(k-\ell), \quad 0 \leq k, \ell \leq L. \tag{11.4.16}$$

This is therefore another way of saying that the wavelet basis obtained from the L-level tree is orthonormal.

The inductive reasoning. The three branches of the $(L+1)$-level tree, shown in Fig. 11.4-2(b), can be redrawn as in Fig. 11.4-2(c) where

$$S_k(z) \triangleq F_k(z^2)G_s(z), \quad S_\ell(z) \triangleq F_\ell(z^2)G_s(z). \tag{11.4.17}$$

By using the identity

$$\Big(A(z^2)B(z)\Big)\Big|_{\downarrow 2^{K+1}} = \Big(A(z)(B(z)|_{\downarrow 2})\Big)\Big|_{\downarrow 2^K} \tag{11.4.18}$$

(Problem 11.9), we can prove that (11.4.13) and (11.4.16) imply

$$S_k(z)\widetilde{S}_\ell(z)\Big|_{\downarrow 2^{(n_\ell+1)}} = \delta(k-\ell),\ S_k(z)\widetilde{H}_s(z)\Big|_{\downarrow 2} = 0,\ S_\ell(z)\widetilde{H}_s(z)\Big|_{\downarrow 2} = 0, \tag{11.4.19}$$

which is sufficient to prove that the wavelet basis generated at the $(L+1)$th level remains orthonormal! The above reasoning does not assume that the pair $[G_s(z), H_s(z)]$ has to be the same for all levels of the tree. Notice finally that the FIR nature of the filters $G_s(z), H_s(z)$ ensures that the wavelet basis functions are FIR as well. The FIR nature also means, in particular, that the wavelet transformation as well as the inverse transformation are stable. Summarizing, we have proved:

♠**Theorem 11.4.1. Wavelet orthonormality.** Consider the L-level tree structure demonstrated in Figs. 11.3-6 and 11.3-7 for $L = 3$. Let the filters $G(z)$, $H(z)$, $G_s(z)$, and $H_s(z)$ be such that this is a perfect reconstruction system, that is, $\hat{x}(n) = x(n)$ so that $x(n)$ has the wavelet expansion (11.3.22a), where $M = L + 1$. Let $\mathbf{R}(z)$ be the 2×2 polyphase matrix of $[G_s(z), H_s(z)]$. Then the discrete-time wavelet basis $\{\eta_{km}(n)\}$ is orthonormal if and only if $\mathbf{R}(z)$ is paraunitary, that is, if and only if $[G_s(z), H_s(z)]$ forms a paraunitary pair. ♢

Unit energy property. Equation (11.4.15) implies, in particular, that $\sum_n |f_k(n)|^2 = 1$, that is, all the filters have unit energy. This is consistent with the increasing heights shown in Fig. 11.3-8. Note that the unit energy property holds *regardless* of the quality of the frequency responses (e.g., stopband attenuation, sharpness of cutoff, etc.), and is a direct consequence of the paraunitary property of the pair $[G_s(z), H_s(z)]$.

Use of lattice structure. As a converse, it can be shown that essentially all orthonormal wavelet bases can be generated using the lattice. See [Soman and Vaidyanathan, 1993] for precise statements.

Design example 11.4.1: STFT and Wavelet Filter Banks

A. Generalized STFT with orthonormal basis. Consider a four-channel maximally decimated filter bank system, designed using the tree structure of Fig. 5.8-1. For simplicity we take the filters at various levels to be the same, that is,

$$H_0^{(k)}(z) = G(z), H_1^{(k)}(z) = H(z), F_0^{(k)}(z) = G_s(z), F_1^{(k)}(z) = H_s(z).$$

We design this to be a perfect reconstruction system by designing $G(z)$ to be a causal Nth order FIR power symmetric filter (Sec. 5.3.6) and taking the remaining filters to be

$$H(z) = -z^{-N}\widetilde{G}(-z),\ G_s(z) = \widetilde{G}(z),\ H_s(z) = \widetilde{H}(z).$$

$G_s(z)$ and $H_s(z)$ are noncausal FIR filters. Assuming that $G(z)$ has been properly scaled, the power symmetric property implies

$$\widetilde{G}_s(z)G_s(z) + \widetilde{G}_s(-z)G_s(-z) = 2,$$

which in turn ensures $\widehat{x}(n) = x(n)$. The above choice of filters ensures the paraunitary condition $\widetilde{\mathbf{R}}(z)\mathbf{R}(z) = \mathbf{I}$.

For our demonstration we take $G(z)$ to be a 23rd order filter, designed as in Sec. 6.3.2. Fig. 11.4-3(a) shows the responses of $G_s(z)/\sqrt{2}$ and $H_s(z)/\sqrt{2}$. Fig. 11.4-3(b) shows the responses of the four analysis filters $H_k(z)$. These filters have equal passband bandwidths. Note that $|G_s(e^{j\omega})| = |G(e^{j\omega})|$ and $|H_s(e^{j\omega})| = |H(e^{j\omega})|$. Also $|F_k(e^{j\omega})| = |H_k(e^{j\omega})|$ for each k, according to the construction of these filters. Unlike in traditional STFT, the filters are not obtained by modulation of a prototype, hence the name "generalized" STFT.

B. Wavelet filter bank with orthonormal basis. We take a three level tree as in Fig. 11.3-6. Let $G(z), H(z)$, $G_s(z)$ and $H_s(z)$ be as above so that we again have a perfect reconstruction system. The responses of the four analysis filters are now as in Fig. 11.4-3(c). In these plots, we have normalized $|H_k(e^{j\omega})|_{max}$ to be unity for convenience. Notice again that $|F_k(e^{j\omega})| = |H_k(e^{j\omega})|$ according to construction.

Since $\mathbf{R}(z)$ is paraunitary, the basis functions for the above generalized STFT as well as the wavelet filter bank are orthonormal.

11.4.3 Orthonormal Wavelets From M-channel Uniform-Decimation Paraunitary QMF Banks

We now generalize the result of Sec. 11.4.1 for the case of an M channel maximally decimated QMF bank with uniform decimation ratio for all channels (Fig. 5.4-1). If this system has the perfect reconstruction property, then $\widehat{x}(n) = x(n)$ so that we can express $x(n)$ as

$$x(n) = \sum_{k=0}^{M-1} \sum_{m=-\infty}^{\infty} y_k(m) \underbrace{f_k(n - Mm)}_{\eta_{km}(n)}. \tag{11.4.20}$$

We have used the notation $y_k(n)$, instead of $v_k(n)$ as in Fig. 5.4-1, to be consistent with the rest of this chapter. We can regard $y_k(m)$ as the wavelet coefficients of $x(n)$ with respect to the wavelet basis defined by $\eta_{km}(n)$ indicated above. Note that $\eta_{k0}(n) = f_k(n)$, whereas $\eta_{km}(n)$ for arbitrary m is obtained by shifting $f_k(n)$ by *multiples of* M.

In the binary tree structure (Sec. 11.3.3) the bandwidths of the filters become smaller and smaller as the center frequency decreases (Fig. 11.3-8). But in the QMF bank under consideration the analysis filters typically tend to have equal bandwidths because the channels have equal decimation

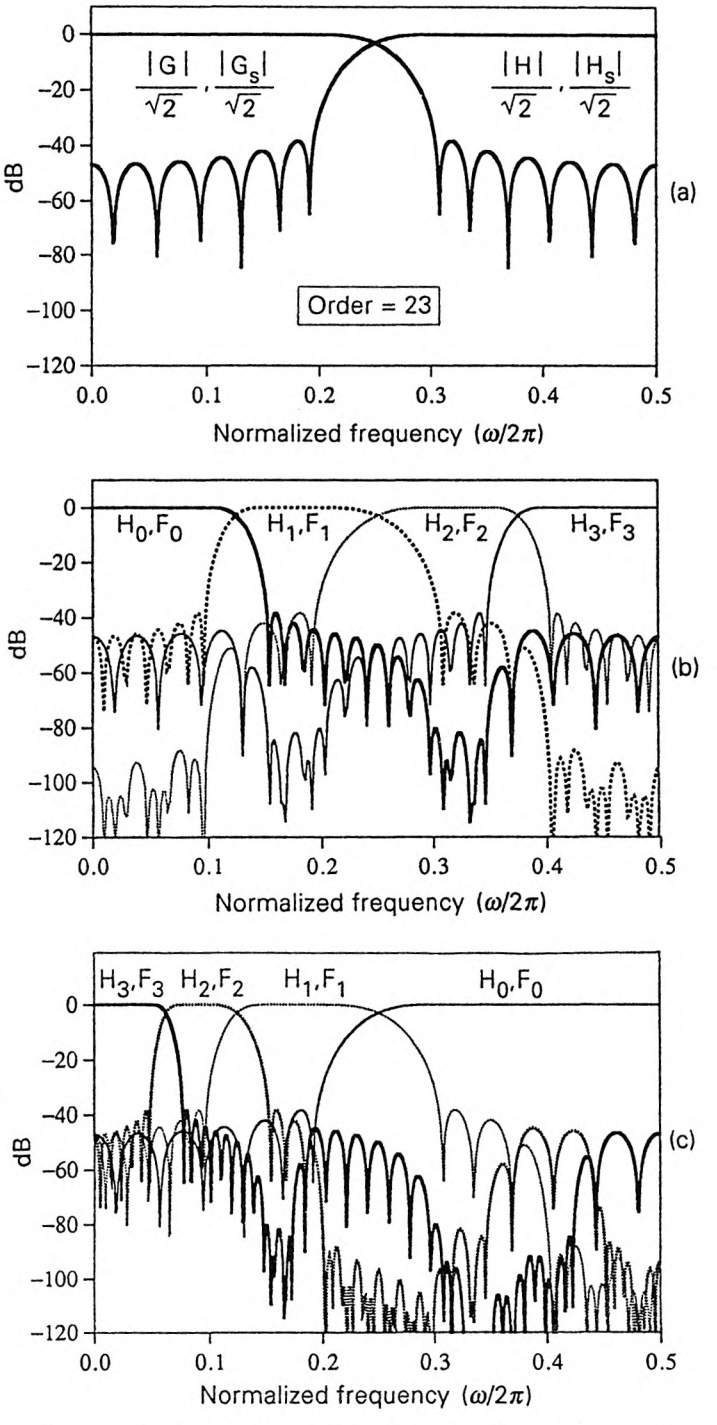

Figure 11.4-3 Design example 11.4-1. Magnitude responses of various filters. (a) $G_s(z)$, $G(z)$, $H_s(z)$ and $H(z)$, (b) STFT filters and (c) wavelet filters.

ratio M. This system is therefore closer to traditional (uniform bandwidth) spectrum analyzers rather than wavelet transformers.

Review of perfect reconstruction. Recall that the above QMF bank can always be redrawn as in Fig. 5.5-3(b), where $\mathbf{E}(z)$ and $\mathbf{R}(z)$ are the $M \times M$ polyphase matrices defined in Sec. 5.5. We also know that if $\mathbf{E}(z)$ is FIR and paraunitary then the choice of synthesis filters according to $f_k(n) = h_k^*(-n)$ gives rise to an FIR perfect reconstruction system. In this case, the expansion (11.4.20) holds. The above choice of synthesis filters implies, in particular, that $\mathbf{R}(z)$ is also paraunitary.

Paraunitariness of R(z) Implies Wavelet Orthonormality

We now show that the discrete time wavelet basis functions $\{\eta_{km}(n)\}$ are orthonormal if, and only if, $\widetilde{\mathbf{R}}(z)\mathbf{R}(z) = \mathbf{I}$.

Proof. Proceeding as in the previous subsection, we see that orthonormality of the basis implies

$$\sum_{n=-\infty}^{\infty} f_k(n) f_\ell^*(n - Mm) = \delta(k - \ell)\delta(m). \tag{11.4.21}$$

Notice that this implies, in particular, that all filters have unit energy [regardless of the quality of the responses $F_k(e^{j\omega})$]. The above equation can be re-expressed in the z-domain as

$$\widetilde{F}_\ell(z)F_k(z)\Big|_{\downarrow M} = \delta(k - \ell) \quad \text{(wavelet orthonormality condition).} \tag{11.4.22}$$

Using the z-domain expression for decimation (Sec. 4.1) this becomes

$$\sum_{m=0}^{M-1} \widetilde{F}_\ell(zW^m)F_k(zW^m) = M\delta(k - \ell), \quad 0 \le k,l \le M - 1, \tag{11.4.23}$$

with $W = e^{-j2\pi/M}$. This can be expressed compactly by defining

$$\mathbf{F}(z) \triangleq \begin{bmatrix} F_0(z) & F_1(z) & \dots & F_{M-1}(z) \\ F_0(zW) & F_1(zW) & \dots & F_{M-1}(zW) \\ \vdots & \vdots & \ddots & \vdots \\ F_0(zW^{M-1}) & F_1(zW^{M-1}) & \dots & F_{M-1}(zW^{M-1}) \end{bmatrix}. \tag{11.4.24}$$

Condition (11.4.23) can be expressed in terms of $\mathbf{F}(z)$ as

$$\widetilde{\mathbf{F}}(z)\mathbf{F}(z) = M\mathbf{I} \quad \text{(wavelet orthonormality condition).}$$

In a manner analogous to (11.4.10b), we can show (Problem 11.10) that $\mathbf{F}(z)$ is related to the $M \times M$ polyphase matrix $\mathbf{R}(z)$ according to

$$\mathbf{F}(z) = \mathbf{\Gamma W \Lambda}(z)\mathbf{R}(z^M). \tag{11.4.25}$$

Here $\mathbf{\Gamma}$ and $\mathbf{\Lambda}(z)$ are diagonal matrices with $[\mathbf{\Gamma}]_{ii} = e^{-j2\pi i/M}$, $[\mathbf{\Lambda}(z)]_{ii} = z^{-(M-1-i)}$, and $\mathbf{W}$ is the $M \times M$ DFT matrix. Using this we can show that the condition $\widetilde{\mathbf{F}}(z)\mathbf{F}(z) = M\mathbf{I}$ is equivalent to $\widetilde{\mathbf{R}}(z)\mathbf{R}(z) = \mathbf{I}$. Thus the wavelet basis functions are orthonormal if, and only if, the matrix $\mathbf{R}(z)$ is normalized paraunitary. $\nabla\nabla\nabla$

This is an important result. From Sec. 6.4 and 6.5 we know how to generate the complete class of paraunitary matrices (of given degree and size). Using this we can, therefore, generate all discrete-time finite duration orthonormal wavelets.

11.4.4 Generalizations

The tree structured system of Fig. 11.3-6 can be generalized in several ways. First, instead of splitting a signal into two bands at a time, we can split into several bands. Second, signals such as $y_0(n)$ which are not split any further can themselves be decomposed into further subbands. In this way we obtain a very general tree structure. By modifying the synthesis bank appropriately, we can retain the perfect reconstruction property. The wavelet expansion equation (11.3.22a) can be modified for this case. This is called the *wavelet packet expansion* and was introduced in Coifman et al. [1990]. It can be shown that the paraunitary property of the filters ensures orthonormality of the basis functions of this expansion [Soman and Vaidyanathan, 1992].

The maximally decimated filter bank with arbitrary decimators n_k in the subbands (Fig. P5-32, Chap. 5 Problems section) is clearly a generalization of the filter banks discussed above. In this case, the transform domain coefficients are given by (11.2.15) and the inverse transform by (11.2.16). This is called the generalized STFT pair, the "filter bank transform" pair, or the "general discrete-time wavelet transform pair." Some of the properties of this system are summarized in Table 11.4.1. The nonuniform filter bank has been studied in a number of references, for example, Hoang and Vaidyanathan [1989], Kovačević and Vetterli [1991a], Nayebi et. al. [1991a], and Soman and Vaidyanathan [1993]. Also see Problem 11.22

11.5 CONTINUOUS-TIME ORTHONORMAL WAVELET BASIS

Consider the wavelet decomposition (11.3.7a) for a continuous time signal $x(t)$. Since k and n are integers, this is the discrete-wavelet transform (DWT) (but not discrete-time, as t is continuous). The case where $a = 2$ is commonly known as the diadic (or binary) wavelet decomposition. We will consider this special case, and further assume $T = 1$. For this case we show how, under certain conditions, an orthonormal basis can be generated starting from the discrete-time QMF bank.

For the discrete-time case (Figs. 11.3-6, 11.3-7) all the synthesis filters $F_k(z)$ (hence the wavelet basis functions $\eta_{km}(n)$) were constructed in terms of the basic filters $G_s(z)$ and $H_s(z)$. In a similar way we wish to construct a continuous-time wavelet basis starting from two basic functions. For this

TABLE 11.4.1 The discrete-time wavelet transform (DTWT)

The most general form is a nonuniform maximally decimated perfect reconstruction filter bank (Fig. P5-32, in Chap. 5 Problems section). Here $\sum_i 1/n_i = 1$ (maximal decimation). Let n_i = integers for simplicity.

$$x_k(n) = \sum_{m=-\infty}^{\infty} x(m) h_k(n_k n - m), \quad 0 \le k \le M-1 \quad \text{(DTWT)} \quad (11.2.15)$$

$$x(n) = \sum_{k=0}^{M-1} \sum_{m=-\infty}^{\infty} x_k(m) \underbrace{f_k(n - n_k m)}_{\text{basis } \eta_{km}(n)} \quad \text{(inverse DTWT)} \quad (11.2.16)$$

This is identical to the *generalized* STFT, i.e., the 'filter bank' transform.

Orthonormal basis

1. *Definition:* $\sum_{n=-\infty}^{\infty} \eta_{km}(n) \eta_{\ell i}^*(n) = \delta(k-\ell)\delta(m-i)$.
2. Equivalently $\sum_n f_k(n) f_m^*(n + n_{k,m}p) = \delta(k-m)\delta(p)$. That is, $\left[F_k(z)\widetilde{F}_m(z)\right]_{\downarrow n_{k,m}} = \delta(k-m)$. Here $n_{k,m} = \gcd(n_k, n_m)$ (Problem 11.22).
3. *Perfect reconstruction condition:* $f_k(n) = h_k^*(-n)$.
4. *Paraunitariness.* A nonuniform filter bank can be converted into an equivalent uniform system with larger number of channels. If the larger system is paraunitary, the original system is orthonormal.

Special cases

1. *Uniform filter bank.* Here $n_k = M$ for all k. (Uniform bandwidth filters.) Now orthonormality is equivalent to the paraunitary property of the $M \times M$ polyphase matrix $\mathbf{R}(z)$.
2. *Octave filter bank.* Here $n_k = 2^{k+1}$ for $0 \le k \le M-2$ and $n_{M-1} = n_{M-2}$. Typical frequency responses have octave spacing and bandwidth, as in Fig. 11.3-8. Can be designed using the binary tree structure (Fig. 11.3-6, 11.3-7). $[G_s(z), H_s(z)]$ need not be the same at all levels of the tree. Orthonormality can be achieved by forcing $[G_s(z), H_s(z)]$ to be paraunitary at each level [i.e., forcing (11.4.13)]. Equivalently,
   a) $|G_s(e^{j\omega})|^2 + |H_s(e^{j\omega})|^2 = 2$ and
   b) $H_s(z) = z^{-N}\widetilde{G}_s(-z)$ for some odd N.
3. *Wave packets.* If the binary tree (Fig. 11.3-6, 11.3-7) is replaced with a more general version (with maximal decimation at each level), we obtain wavelet packets. The basis functions are orthonormal if the filter bank at each level has the paraunitary property.

we introduce the two infinite products

$$\Phi(\omega) = \frac{1}{\sqrt{2}} G_s(e^{j\omega/2}) \prod_{m=2}^{\infty} \frac{1}{\sqrt{2}} G_s(e^{j2^{-m}\omega}), \tag{11.5.1a}$$

$$\Psi(\omega) = \frac{1}{\sqrt{2}} H_s(e^{j\omega/2}) \prod_{m=2}^{\infty} \frac{1}{\sqrt{2}} G_s(e^{j2^{-m}\omega}), \tag{11.5.1b}$$

where

$$G_s(z) = \sum_n g_s(n) z^{-n}, \quad H_s(z) = \sum_n h_s(n) z^{-n}. \tag{11.5.2}$$

We denote the inverse Fourier tranforms of $\Phi(\omega)$ and $\Psi(\omega)$ as $\phi(t)$ and $\psi(t)$ respectively. These are functions of the continuous time variable t, since $\Phi(\omega)$ and $\Psi(\omega)$ are not periodic in ω. † The function $\psi(t)$ (called the *wavelet function*) will play a crucial role in our discussions. The function $\phi(t)$ (called the *scaling function*) will enter many equations involving $\psi(t)$ [e.g., (11.5.15b) and (11.5.18b)].

To demonstrate how the infinite products work, assume that $G_s(e^{j\omega})$ and $H_s(e^{j\omega})$ are ideal lowpass and highpass filters with cutoff $\pi/2$. Fig. 11.5-1 shows a number of stretched (or dilated) versions of $G_s(e^{j\omega})$. From these we can judge that the infinite products $\Phi(\omega)$ and $\Psi(\omega)$ are lowpass and bandpass as shown in Fig. 11.5-2.

♠**Main points of this section.** Our aim is to generate wavelet basis functions $\psi_{k\ell}(t)$ by dilations and shifts of the wavelet function $\psi(t)$, that is,

$$\psi_{k\ell}(t) = 2^{-k/2} \psi(2^{-k}t - \ell). \tag{11.5.3}$$

1. We will see that if $G_s(z)$ and $H_s(z)$ are FIR, then $\phi(t)$ and $\psi(t)$ are of finite duration. Thus, for each finite k and ℓ the function $\psi_{k\ell}(t)$ has finite duration.
2. Suppose $G_s(z)$ and $H_s(z)$ are FIR and form a paraunitary pair, that is, their polyphase matrix $\mathbf{R}(z)$ is paraunitary (i.e., (11.4.13) holds). Under some further mild conditions (to be made precise in Theorem 11.5.1), we will see that the set of functions $\{\psi_{k\ell}(t)\}$ is orthonormal. In fact, the set $\{\psi_{k\ell}(t)\}$ can be used to represent any finite energy function $x(t)$ in the form (11.3.10). That is, $\{\psi_{k\ell}(t)\}$ is complete over the so-called $L^2(R)$ class of functions.
3. In general, the function $\psi(t)$ obtained from the digital filters $G_s(z)$ and $H_s(z)$ as above is not "smooth". We will see that by constraining

† For convenience we continue to use ω rather than Ω in this section, even though $\phi(t)$ and $\psi(t)$ are continuous-time functions.

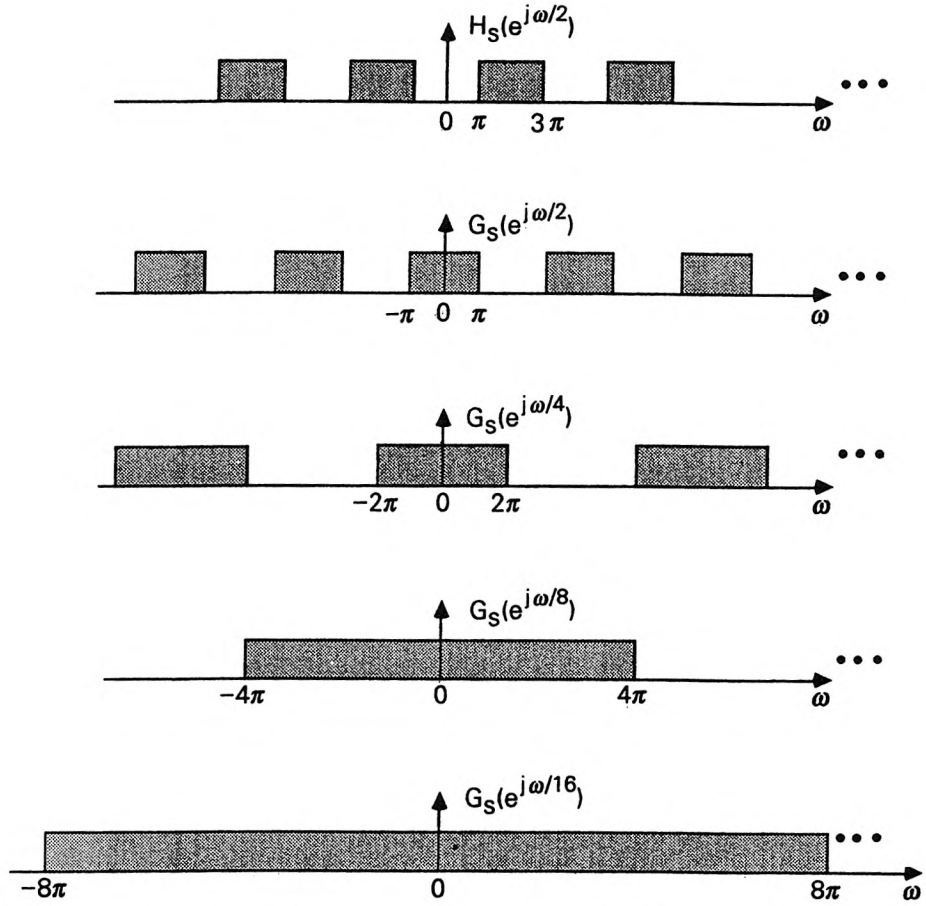

Figure 11.5-1 Various dilated versions of the basic filters, which take part in the infinite product.

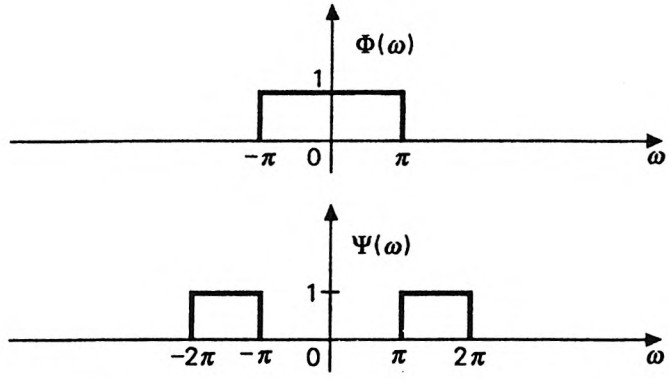

Figure 11.5-2 The magnitudes of $\Phi(\omega)$ and $\Psi(\omega)$ when $G_s(e^{j\omega})$ and $H_s(e^{j\omega})$ are ideal filters as in Fig. 11.5-1.

$G_s(e^{j\omega})$ to have multiple zeros at $\omega = \pi$, the function $\psi(t)$ can be made "smooth" or "regular." We will also see how to design the Nth order FIR filter $G_s(z)$ in such a way that $G_s(z)$ has "as many zeros as possible" at $\omega = \pi$, under the additional constraint that $[G_s(z), H_s(z)]$ be a paraunitary pair. This maximizes "regularity" under the FIR orthonormal constraint.

The goal of this section is to provide a detailed derivation of the above results. We will also consider several illustrative examples to demonstrate these ideas. Some of the deeper mathematical issues (such as "completeness") will not be proved, but references will be provided.

11.5.1 Study of the Functions $\Phi(\omega)$ and $\Psi(\omega)$

As a first step towards the above goal, we now study some of the key properties of the infinite products $\Phi(\omega)$ and $\Psi(\omega)$. We begin with some examples.

Example 11.5.1: Infinite Products

Convergence of infinite products is tricky [Apostol, 1974]. Let $S = \prod_{m=1}^{\infty} s_m$. If there is an $\epsilon > 0$ such that $|s_m| < 1 - \epsilon$ for all m exceeding some integer m_0, then this product vanishes. On the other hand if $|s_m| > 1 + \epsilon$ for all $m > m_0$, the product does not converge to a finite value at all.

There do exist infinite products that converge to finite nonzero values. For example, let $s_m = a^{b^m}$, with $|b| < 1$. Then

$$S = \prod_{m=1}^{\infty} s_m = \prod_{m=1}^{\infty} a^{b^m} = a^{(\sum_{m=1}^{\infty} b^m)} = a^{b/(1-b)}. \tag{11.5.4}$$

The above result is clearly finite. Readers who do not feel comfortable with the transition from infinite products to infinite sums in (11.5.4) must see Problem 11.14. For theorems on convergence of infinite products, see [Apostol, 1974] and pp.11-12 of Gradshteyn and Ryzhik [1980].

As a second example, let $s_m = \cos(2^{-m}\omega)$. We will show that the product converges:

$$\prod_{m=1}^{\infty} \cos(2^{-m}\omega) = \frac{\sin\omega}{\omega}. \tag{11.5.5}$$

Note that $|\cos(2^{-m}\omega)| < 1$ for almost all ω, and yet the product converges to a nonzero value for almost all ω. This does not violate the statement in the preceding paragraph because there is no $\epsilon > 0$ such that $|\cos(2^{-m}\omega)| < 1 - \epsilon$ for all m greater than some m_0.

A simple proof of (11.5.5) is as follows: by using the identity $\cos\alpha = (\sin 2\alpha)/(2\sin\alpha)$ we can simplify the partial product of the first K factors as

$$\prod_{m=1}^{K} \cos(2^{-m}\omega) = \frac{\sin\omega}{2^K \sin(\omega/2^K)}$$

As $K \to \infty$ we have $\sin(\omega/2^K) \to \omega/2^K$, so that

$$\lim_{K\to\infty} \prod_{m=1}^{K} \cos(2^{-m}\omega) = \frac{\sin\omega}{\omega}$$

indeed. Thus, the partial product converges to $(\sin\omega/\omega)$ pointwise, for each ω.

To obtain a 'system theoretic feeling' for the identity (11.5.5), recall that a product in the ω-domain translates to a convolution in the time domain. The inverse transform of $\cos(\omega/2)$ is a sum of two impulses, and that of $\cos(\omega/4)$ is a sum of two impulses placed closer together, and so on [Fig. 11.5-3(a)]. If these impulse trains are convolved, we obtain the rectangular window function [Fig. 11.5-3(b)] which indeed is the inverse transform of $(\sin\omega/\omega)$.

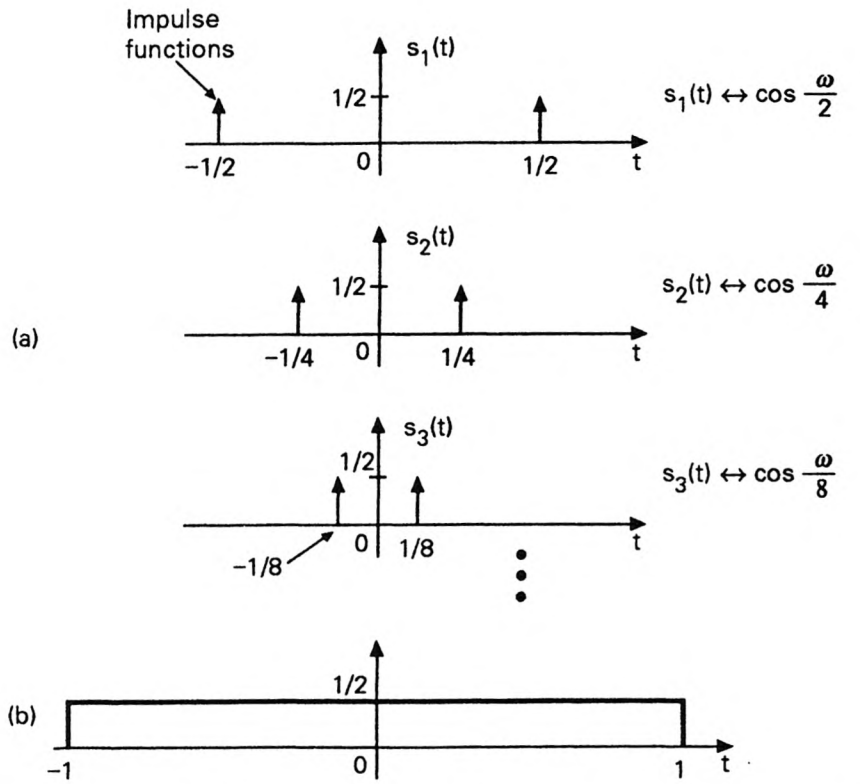

Figure 11.5-3 (a) Impulse functions whose Fourier transforms have the form $\cos(\omega/2^m)$ and (b) results of convolution of an infinite number of these.

Convergence of the Infinite Products

Consider the partial products

$$\Phi_L(\omega) = \frac{1}{\sqrt{2}} G_s(e^{j\omega/2}) \prod_{m=2}^{L} \frac{1}{\sqrt{2}} G_s(e^{j2^{-m}\omega}), \tag{11.5.6a}$$

$$\Psi_L(\omega) = \frac{1}{\sqrt{2}} H_s(e^{j\omega/2}) \prod_{m=2}^{L} \frac{1}{\sqrt{2}} G_s(e^{j2^{-m}\omega}), \tag{11.5.6b}$$

with

$$\Phi_1(\omega) = \frac{1}{\sqrt{2}} G_s(e^{j\omega/2}), \quad \Psi_1(\omega) = \frac{1}{\sqrt{2}} H_s(e^{j\omega/2}). \tag{11.5.6c}$$

The infinite products $\Phi(\omega)$ and $\Psi(\omega)$ are, by definition, limits of these partial products as $L \to \infty$. Since $G_s(e^{j\omega})$ and $H_s(e^{j\omega})$ have period 2π, $\Phi_L(2^L\omega)$ and $\Psi_L(2^L\omega)$ have period 2π. In other words, the partial products $\Phi_L(\omega)$ and $\Psi_L(\omega)$ have period $2^{L+1}\pi$, and the infinite products $\Phi(\omega)$ and $\Psi(\omega)$ are nonperiodic. The inverse transforms therefore represent continuous-time functions $\phi(t)$ and $\psi(t)$. Notice that the quantities

$$\begin{aligned} \Phi_L(2^L\omega) &= (\frac{1}{\sqrt{2}})^L G_s(e^{j\omega}) \ldots G_s(e^{j2^{L-2}\omega}) G_s(e^{j2^{L-1}\omega}), \\ \Psi_L(2^L\omega) &= (\frac{1}{\sqrt{2}})^L G_s(e^{j\omega}) \ldots G_s(e^{j2^{L-2}\omega}) H_s(e^{j2^{L-1}\omega}), \end{aligned} \tag{11.5.7}$$

are related to the synthesis filters $F_L(z)$ and $F_{L-1}(z)$ of the L-level tree structure (Fig. 11.3-7) as follows:

$$\Phi_L(2^L\omega) = \frac{F_L(e^{j\omega})}{(\sqrt{2})^L}, \quad \Psi_L(2^L\omega) = \frac{F_{L-1}(e^{j\omega})}{(\sqrt{2})^L}. \tag{11.5.8}$$

Types of convergence. It is said that $\Psi_L(\omega)$ converges to $\Psi(\omega)$ *pointwise*, if for any ω in the range $-2^L\pi \le \omega < 2^L\pi$, the quantity $|\Psi_L(\omega) - \Psi(\omega)|$ tends to zero as $L \to \infty$. We say that $\Psi_L(\omega)$ converges to $\Psi(\omega)$ in the mean square sense if $\int_{-2^L\pi}^{2^L\pi} |\Psi_L(\omega) - \Psi(\omega)|^2 d\omega$ tends to zero as $L \to \infty$. Finally the convergence is said to be 'uniform' in an interval if some deeper requirements are satisfied [Kreyszig, 1972]. In this chapter, 'convergence' stands for pointwise convergence.

As mentioned above, pointwise convergence is not always guaranteed. However, if $|G_s(\omega)|/\sqrt{2} \le 1$ for all ω, then we have pointwise convergence. To see this note that when $|G_s(\omega)|/\sqrt{2} \le 1$ we have

$$|\Phi_{L+1}(\omega)| \le |\Phi_L(\omega)|. \tag{11.5.9}$$

Thus $|\Phi_L(\omega)|$ is a monotone nonincreasing sequence in L, with a lower bound equal to zero. As a result, $\Phi_L(\omega)$ converges pointwise to a limit. The infinite-product $\Phi(\omega)$ denotes this limiting function. Similar comments hold for $\Psi(\omega)$. Notice, in particular, that if $[G_s(z), H_s(z)]$ is a paraunitary pair satisfying (11.4.13), we have

$$|G_s(e^{j\omega})|^2 + |H_s(e^{j\omega})|^2 = 2, \tag{11.5.10}$$

so that $|G_s(e^{j\omega})|/\sqrt{2} \leq 1$ indeed, and pointwise convergence holds.

Example 11.5.2

Let $G_s(e^{j\omega})$ and $H_s(e^{j\omega})$ be ideal lowpass and highpass filters with cutoff $\pi/2$. From Fig. 11.5-2 we know that $\Phi(\omega)$ and $\Psi(\omega)$ are ideal lowpass and bandpass filters respectively. Their impulse responses are

$$\phi(t) = \frac{\sin(\pi t)}{\pi t}, \quad \psi(t) = \frac{\sin(\pi t/2)}{\pi t/2}\cos(3\pi t/2). \tag{11.5.11}$$

Example 11.5.3.

Next assume that $G_s(z)$ and $H_s(z)$ are FIR with

$$G_s(z) = \frac{1+z^{-1}}{\sqrt{2}}, \quad H_s(z) = \frac{1-z^{-1}}{\sqrt{2}}. \tag{11.5.12}$$

Then

$$G_s(e^{j\omega}) = \sqrt{2}e^{-j\omega/2}\cos(\omega/2), \; H_s(e^{j\omega}) = j\sqrt{2}e^{-j\omega/2}\sin(\omega/2). \tag{11.5.13}$$

By using these in the definitions of $\Phi(\omega)$ and $\Psi(\omega)$ and applying the identity (11.5.5), we obtain

$$\Phi(\omega) = e^{-j\omega/2}\frac{\sin(\omega/2)}{\omega/2}, \quad \Psi(\omega) = je^{-j\omega/2}\frac{\sin^2(\omega/4)}{\omega/4}. \tag{11.5.14}$$

The inverse transforms $\psi(t)$ and $\phi(t)$ are shown in Fig. 11.5-4. These are also causal because of our causal choice of $G_s(z)$ and $H_s(z)$. We will use the above function $\psi(t)$ later to generate the so-called Haar-basis of wavelets (Example 11.5.5).

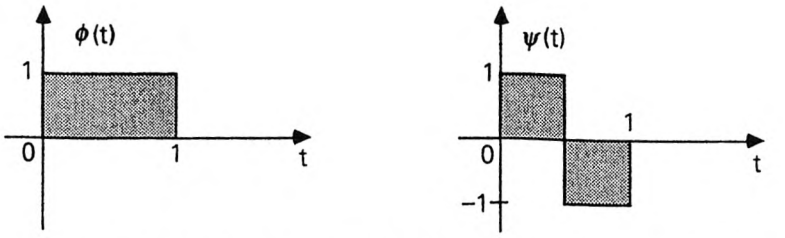

Figure 11.5-4 Example 11.5-3. The functions $\phi(t)$ and $\psi(t)$ which result from the choice of $G_s(z)$ and $H_s(z)$ as in (11.5.12).

Time Domain Interpretation of the Infinite Products

From the definition (11.5.1) we see that $\Phi(\omega)$ and $\Psi(\omega)$ can be expressed as

$$\Phi(\omega) = \frac{1}{\sqrt{2}} G_s(e^{j\omega/2})\Phi(\omega/2) \tag{11.5.15a}$$

$$\Psi(\omega) = \frac{1}{\sqrt{2}} H_s(e^{j\omega/2})\Phi(\omega/2). \tag{11.5.15b}$$

The above products can be expressed as convolutions in the time domain. To do this, note that the inverse transform of $G_s(e^{j\omega})$, viewed as a function of continuous time argument t, can be written as an impulse train:

$$\sum_{n=-\infty}^{\infty} g_s(n)\delta_a(t-n). \tag{11.5.16}$$

Coupled with the fact that the inverse transform of $\frac{1}{2}F(\omega/2)$ is $f(2t)$, we see that

$$\phi(t) = 2\sqrt{2}\Big(\sum_n g_s(n)\delta_a(2t-n)\Big) * \phi(2t), \tag{11.5.17}$$

where $*$ denotes convolution. Simplifying this we obtain the recursion

$$\phi(t) = \sqrt{2} \sum_{n=-\infty}^{\infty} g_s(n)\phi(2t-n). \tag{11.5.18a}$$

Summarizing, the infinite product (11.5.1a) is equivalent to the above recursive relation, where $g_s(n)$ is related to $G_s(z)$ as in (11.5.2). Similarly, from (11.5.15b) we obtain the recursion

$$\psi(t) = \sqrt{2} \sum_{n=-\infty}^{\infty} h_s(n)\phi(2t-n). \tag{11.5.18b}$$

Usually we take the filters $G_s(z)$ and $H_s(z)$ to be FIR, so that (11.5.18a) and (11.5.18b) are finite summations. For example if $G_s(z)$ and $H_s(z)$ are as in (11.5.12), the summations have only two terms. In this case $\phi(t)$ and $\psi(t)$ are as in Fig. 11.5-4, and the above recursions are easily verified.

Self-similarity. Figure 11.5-5 demonstrates the recursion (11.5.18a) for $N = 3$. Part (a) shows an example of $\phi(t)$. Part (b) shows the shifted and scaled versions $\sqrt{2}g_s(n)\phi(2t - n)$. Since $N = 3$, there are four curves, and these add up to $\phi(t)$ for all t. Notice that in the region $0 \leq t \leq 0.5$, the signal $\phi(t)$ is identical to a scaled version of $\phi(2t)$, which indicates the self-similar behavior of $\phi(t)$. Similar comments holds for $2.5 \leq t \leq 3.0$. It should be noticed that there is no simple equivalent of this elegant property, for the case of discrete-time wavelets.

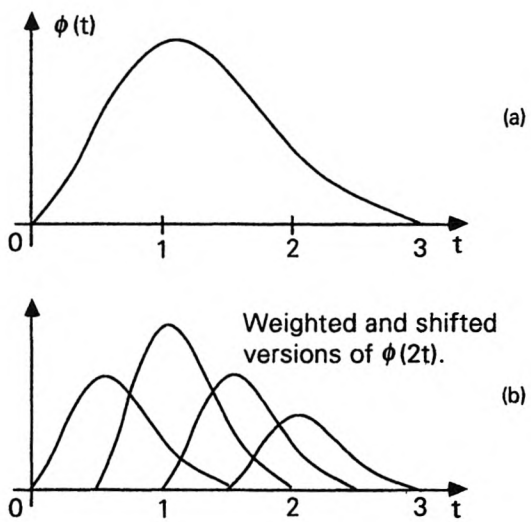

Figure 11.5-5 Demonstrating how a function is formed from superposition of shifted versions of a compressed version.

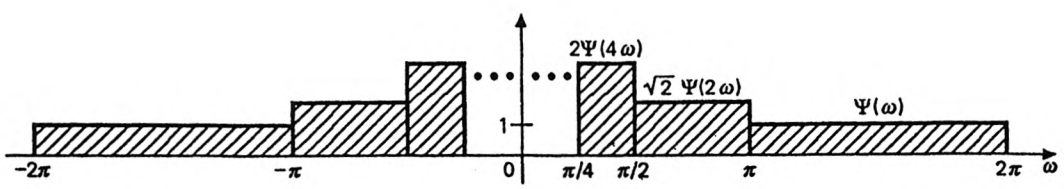

Figure 11.5-6 Various compressed versions of $\Psi(\omega)$. These are used to generate an orthonormal basis of wavelets.

Generation of the Wavelet Basis

Suppose $\Psi(\omega)$ is a bandpass response (e.g., approximating the one in

Fig. 11.5-2). Then the functions $2^{k/2}\Psi(2^k\omega)$ are bandpass with center frequencies as well as bandwidths reduced by the factor 2^k (e.g., see Fig. 11.5-6). (The scale factor $2^{k/2}$ serves to keep the energy of the filter same for all k.) Let these functions be used in the filter bank arrangement of Fig. 11.3-2(c), that is, $F_k(j\omega) = 2^{k/2}\Psi(2^k\omega)$. If this system has perfect reconstruction, then the reconstructed signal is

$$x(t) = \sum_k 2^{-k/2} \sum_\ell X_{DWT}(k,\ell)\psi(2^{-k}t - \ell). \tag{11.5.19}$$

Thus, the wavelet basis functions are

$$\psi_{k\ell}(t) \triangleq 2^{-k/2}\psi(2^{-k}t - \ell), \tag{11.5.20}$$

where k and ℓ are integers. The basis functions have Fourier transforms

$$\Psi_{k\ell}(\omega) = 2^{k/2}e^{-j2^k\ell\omega}\Psi(2^k\omega). \tag{11.5.21}$$

If $x(t)$ is appropriately bandlimited, we can restrict k to be nonnegative.

For appropriate choice of the FIR filters $G_s(z)$ and $H_s(z)$, the functions $\psi_{k\ell}(t)$ can be made orthonormal. The following examples will demonstrate this. We will see later that the orthonormality property can be induced by designing $[G_s(z), H_s(z)]$ to be a paraunitary pair.

Example 11.5.4

Recall that when $G_s(e^{j\omega})$ and $H_s(e^{j\omega})$ are ideal filters (as in Fig. 11.5-1), $\Psi(\omega)$ is the ideal bandpass filter shown in Fig. 11.5-2. The quantities $\Psi(2^k\omega)$ do not overlap for two different values of k (Fig. 11.5-6). From this it is clear (use Parseval's relation) that the functions $\psi(2^{-k_1}t)$ and $\psi(2^{-k_2}t)$ are orthogonal for $k_1 \neq k_2$. As a further step, it is easy to verify that any two functions $\psi_{k\ell}(t)$ and $\psi_{mn}(t)$ are orthonormal (Problem 11.15).

Example 11.5.5: The Haar Basis

In the previous example, orthonormality was easy to see by considering the frequency domain, whereas a time domain approach is more suited in some cases. We know that if $G_s(z)$ and $H_s(z)$ are as in (11.5.12), $\psi(t)$ is the function shown in Fig. 11.5-4. Some of the wavelet basis functions $\psi_{k\ell}(t)$ are sketched in Fig. 11.5-7, and are evidently orthonormal. These are called the Haar basis functions.

It is easily verified that any two members of the family $\{\psi_{k\ell}(t)\}$ are orthonormal. It can be shown (see references in [Daubechies, 1988]) that these functions form a basis for the class of all real finite energy functions [commonly called the class $L^2(R)$].

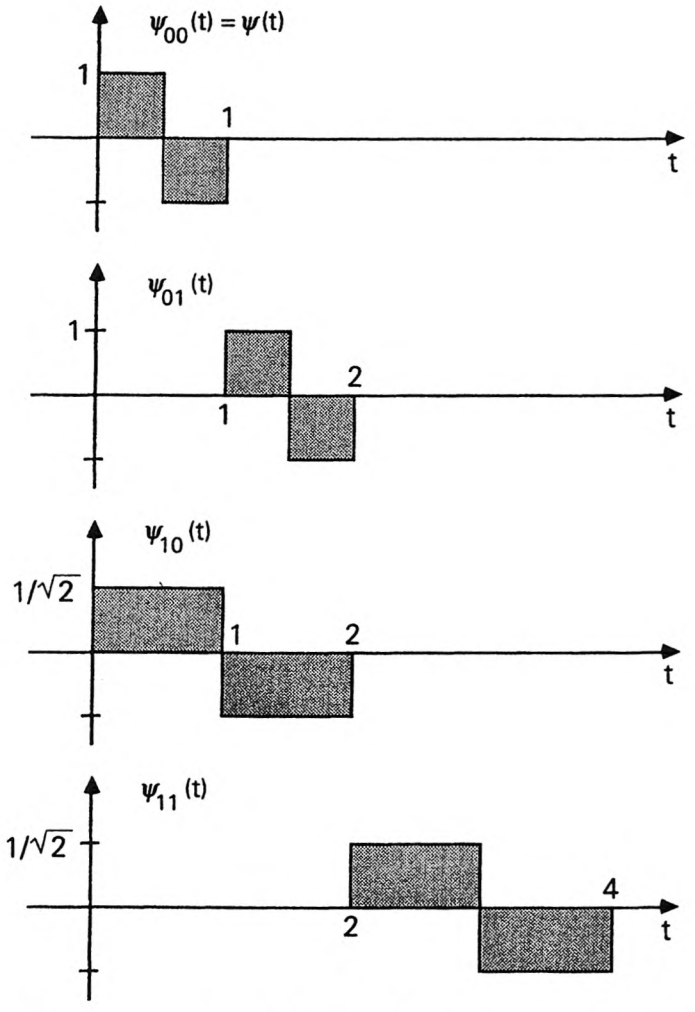

Figure 11.5-7 Example 11.5.5. Some of the basis functions belonging to the Haar basis.

Basis Functions with Finite Duration

We now show that if $G_s(z)$ and $H_s(z)$ are FIR, the functions $\phi(t)$, $\psi(t)$, and $\psi_{k\ell}(t)$ have finite duration. The infinite products in (11.5.1), are equivalent to infinite convolutions in the time domain. The inverse transform of $G_s(e^{j\omega})$, viewed as a function of the continuous variable t, can be expressed as the impulse-train function (11.5.16). The same comment holds for $H_s(e^{j\omega})$. Using this we obtain

$$\begin{aligned}\phi(t) &= \frac{2}{\sqrt{2}}\sum_n g_s(n)\delta_a(2t-n) * \frac{4}{\sqrt{2}}\sum_n g_s(n)\delta_a(4t-n) * \ldots,\\ \psi(t) &= \frac{2}{\sqrt{2}}\sum_n h_s(n)\delta_a(2t-n) * \frac{4}{\sqrt{2}}\sum_n g_s(n)\delta_a(4t-n) * \ldots,\end{aligned} \tag{11.5.22}$$

where $*$ indicates convolution. Assume that $G_s(z)$ and $H_s(z)$ are causal FIR with order N, that is,

$$G_s(z) = \sum_{n=0}^{N} g_s(n) z^{-n}, \quad H_s(z) = \sum_{n=0}^{N} h_s(n) z^{-n}. \tag{11.5.23}$$

The region where the inverse transform of $G_s(e^{j2^{-m}\omega})$ can be nonzero is given by $0 \leq t \leq 2^{-m}N$. As m increases, this quantity has smaller and smaller duration, and the impulses are squeezed tighter. The convolution of infinite number of these will result in a function whose duration is

$$\frac{N}{2} + \frac{N}{4} + \frac{N}{8} + \ldots = N. \tag{11.5.24}$$

Thus, $\phi(t)$ and $\psi(t)$ are causal functions with duration N. The wavelet basis functions $\psi_{k\ell}(t)$ are therefore of finite duration $2^k N$.

11.5.2 Generating Orthonormal Wavelets

From Example 11.5.4 we know that if the filters $G_s(e^{j\omega})$ and $H_s(e^{j\omega})$ are ideal then the wavelet basis $\psi_{k\ell}(t)$ is orthonormal. We also know from Example 11.5.5, that non ideal filters $G_s(z)$ and $H_s(z)$ can sometimes be used to get an orthonormal basis. In Sec. 11.5.3 we will derive a more general result, viz., if $[G_s(z), H_s(z)]$ is a paraunitary pair then $\{\psi_{k\ell}(t)\}$ is an orthonormal family, subject to some further mild conditions (Theorem 11.5.1). In this section we will prove a preliminary result required for this purpose.

Recall that the orthonormality of the wavelet basis is defined by the equation

$$\int_{-\infty}^{\infty} \psi_{k_1 \ell_1}(t) \psi^*_{k_2 \ell_2}(t) dt = \delta(k_1 - k_2)\delta(\ell_1 - \ell_2). \tag{11.5.25}$$

In view of Parseval's relation, this is equivalent to

$$\int_{-\infty}^{\infty} \Psi_{k_1 \ell_1}(\omega) \Psi^*_{k_2 \ell_2}(\omega) \frac{d\omega}{2\pi} = \delta(k_1 - k_2)\delta(\ell_1 - \ell_2). \tag{11.5.26}$$

Using (11.5.21) and making appropriate change of variables, this becomes

$$2^{i/2} \int_{-\infty}^{\infty} \Psi(\omega) \Psi^*(2^i \omega) e^{-j\omega(\ell_1 - 2^i \ell_2)} \frac{d\omega}{2\pi} = \delta(i)\delta(\ell_1 - \ell_2), \tag{11.5.27}$$

where $i = k_2 - k_1$.

Our aim is to establish the connection between the above orthonormality and the paraunitary property of $[G_s(z), H_s(z)]$. Recall that $[G_s(z), H_s(z)]$ is

a paraunitary pair if (11.4.13) holds. This condition means that if we define the polyphase matrix $\mathbf{R}(z)$ according to

$$[G_s(z) \quad H_s(z)] = [z^{-1} \quad 1]\,\mathbf{R}(z^2), \tag{11.5.28}$$

then $\mathbf{R}(z)$ is paraunitary, i.e., $\widetilde{\mathbf{R}}(z)\mathbf{R}(z) = \mathbf{I}$. In Sec. 11.4 we already used the paraunitary property to derive discrete-time orthonormal wavelets.

We will now show that the above paraunitariness implies the relation

$$2^{i/2}\int_{-2^L\pi}^{2^L\pi} \Psi_L(\omega)\Psi^*_{L+i}(2^i\omega)e^{-j\omega(\ell_1-2^i\ell_2)}\,\frac{d\omega}{2\pi} = \delta(i)\delta(\ell_1-\ell_2), \tag{11.5.29}$$

for all integers $L > 0$, where $\Psi_L(\omega)$ is the partial product defined earlier. In the next section we will derive further conditions under which we can let $L \to \infty$ and obtain (11.5.27) from this.

Proof that Paraunitary Property Implies (11.5.29)

For convenience we first make a change of variables and rewrite (11.5.29) as

$$2^{i/2}2^L\int_{-\pi}^{\pi} \Psi_L(2^L\omega)\Psi^*_{L+i}(2^{L+i}\omega)e^{-j2^L\omega(\ell_1-2^i\ell_2)}\frac{d\omega}{2\pi} = \delta(i)\delta(\ell_1-\ell_2). \tag{11.5.30}$$

Case when $i > 0$. From the definition of $\Psi_L(\omega)$ we have (11.5.8), and we can rewrite (11.5.30) as

$$\int_{-\pi}^{\pi} F_{L-1}(e^{j\omega})F^*_L(e^{j\omega})A(e^{j2^L\omega})e^{-j2^L\omega(\ell_1-2^i\ell_2)}\frac{d\omega}{2\pi} = 0, \tag{11.5.31}$$

where $F_L(z)$ and $F_{L-1}(z)$ are the top two filters in the L-level tree-structured synthesis bank (Sec. 11.3). Here $A(z)$ is a discrete-time transfer function whose details are irrelevant for the proof.

If the pair $[G_s(z), H_s(z)]$ satisfies (11.4.13), the L-level tree structure generates orthonormal wavelets (Sec. 11.4.2). This implies, in particular,

$$\left(F_{L-1}(z)\widetilde{F}_L(z)\right)\Big|_{\downarrow 2^L} A(z) = 0, \tag{11.5.32}$$

that is,

$$\left(F_{L-1}(z)\widetilde{F}_L(z)A(z^{2^L})\right)\Big|_{\downarrow 2^L} = 0. \tag{11.5.33}$$

In other words, the inverse transform of $F_{L-1}(z)\widetilde{F}_L(z)A(z^{2^L})$ is zero at locations of the form 2^Lm, where m = integer. This proves (11.5.31).

Case when $i = 0$. Now (11.5.30) reduces to

$$\int_{-\pi}^{\pi} F_{L-1}(e^{j\omega}) F_{L-1}^{*}(e^{j\omega}) e^{-j2^{L}\omega(\ell_1-\ell_2)} \frac{d\omega}{2\pi} = \delta(\ell_1 - \ell_2). \tag{11.5.34}$$

But $F_{L-1}(z)$ satisfies (11.4.16), that is,

$$\left(F_{L-1}(z)\widetilde{F}_{L-1}(z)\right)\Big|_{\downarrow 2^{L}} = 1. \tag{11.5.35}$$

In other words, if we evaluate the inverse transform of $F_{L-1}(z)\widetilde{F}_{L-1}(z)$ at the locations $2^{L}n$, the result is zero for $n \neq 0$ and unity for $n = 0$. This establishes (11.5.34) indeed. $\nabla\nabla\nabla$

Summarizing, we have proved that the paraunitary property of the filter bank $[G_s(z), H_s(z)]$ implies the orthonormality (11.5.29) for finite L. It has been shown [Mallat, 1989b] that this also implies (11.5.27) (i.e., it holds for $L \to \infty$) as long as $G_s(e^{j\omega}) \neq 0$ in $|\omega| \leq \pi/2$. This is only a mild requirement, since a lowpass filter in a QMF bank usually satisfies this. In Sec. 11.5.3 we will deal with the details of Mallat's conditions, and state the result more formally (Lemma 11.5.1 and Theorem 11.5.1).

Class of functions covered by the basis. Notice that $\psi_{k\ell}(t)$ has finite duration since $G_s(z)$ and $H_s(z)$ are FIR. The finite duration orthonormal wavelet basis $\psi_{k\ell}(t)$ generated in the above manner can be used to represent any *finite energy* signal $x(t)$ as in (11.3.10). See Daubechies [1988], Mallat [1989a,b], and references therein. Real signals with finite energy are said to belong to the class $L^2(R)$. So, we say that $\{\psi_{k\ell}(t)\}$ is complete over $L^2(R)$. With further restrictions on the FIR filters $G_s(z)$, the basis can be used to represent a wider class of functions, as elaborated in the above references. Proofs of these 'completeness statements' are beyond the scope of this chapter.

Design Example 11.5.1

Consider again the filters $G(z)$ and $H(z)$ used in Design example 11.4.1. With $G_s(z) = \widetilde{G}(z)$ and $H_s(z) = \widetilde{H}(z)$, the magnitude responses are as in Fig. 11.4-3(a). Since $[G_s(z), H_s(z)]$ is paraunitary, we can use these filters to obtain the orthonormal wavelet basis described above. Figure 11.5-8 shows the functions $\phi(t)$ and $\psi(t)$ as well as their Fourier transform magnitudes, verifying that they are lowpass and bandpass, respectively. Since the order of $G_s(z)$ is $N = 23$, the durations of $\phi(t)$ and $\psi(t)$ are also 23 as seen from the plots.

11.5.3 Orthonormality as L Approaches Infinity

The fact that (11.5.29) holds for all L is *not sufficient* to imply the orthonormality condition (11.5.27) of the continuous time wavelets, as demonstrated in the following example, shown to the author by Ingrid Daubechies.

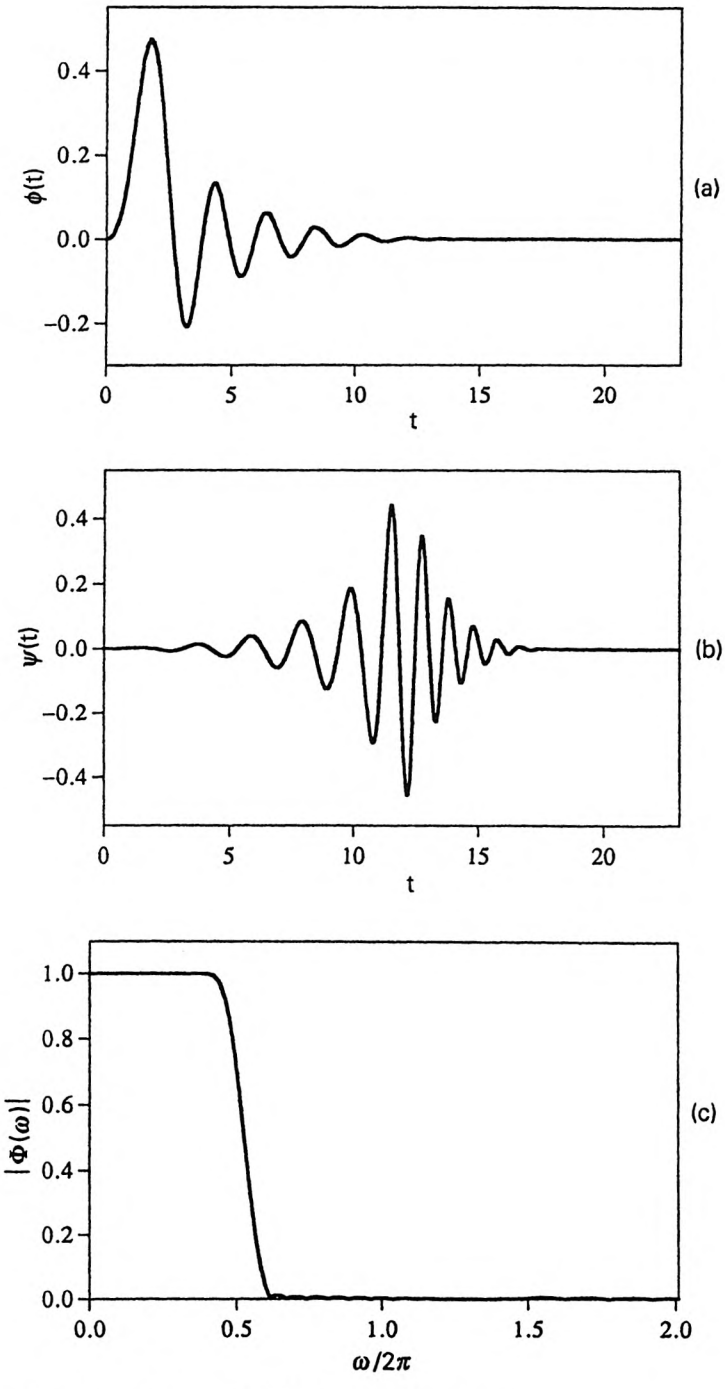

Figure 11.5-8 Design example 11.5.1. *Continued* →

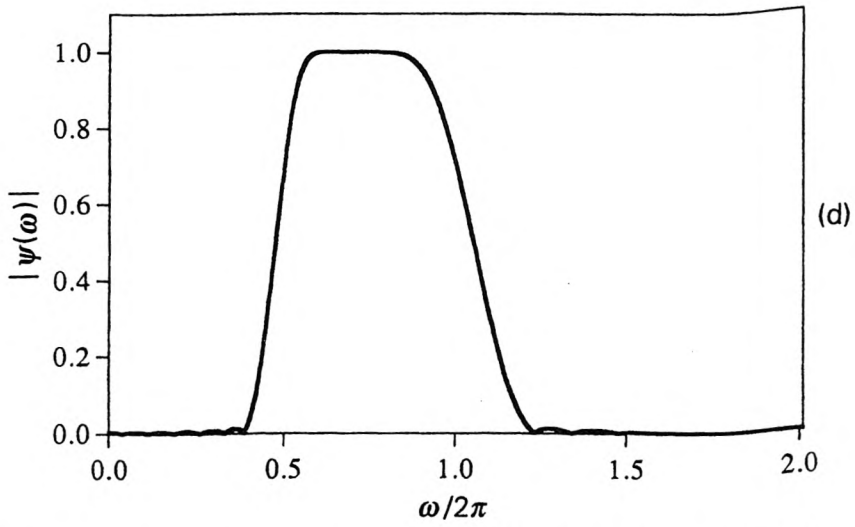

Figure 11.5-8 Design example 11.5.1. The functions $\phi(t)$ and $\psi(t)$, obtained by starting from $G_s(z)$ and $H_s(z)$. The magnitudes of $G_s(e^{j\omega})$ and $H_s(e^{j\omega})$ are as in Fig. 11.4-3(a).

Example 11.5.6

Consider the filter bank with filters

$$G_s(z) = \frac{1+z^{-3}}{\sqrt{2}}, \quad H_s(z) = \frac{1-z^{-3}}{\sqrt{2}}. \tag{11.5.36}$$

This is a modification of the Haar basis example, with z replaced by z^3. It can be verified that the polyphase matrix $\mathbf{R}(z)$ is still paraunitary [i.e., (11.4.13) holds]. The functions $\Psi(\omega)$ and $\Phi(\omega)$ are obtained by replacing ω with 3ω in (11.5.14). The inverse transforms $\phi(t)$ and $\psi(t)$ are shown in Fig. 11.5-9. From the plots of Fig. 11.5-9 we can verify, in particular, that $\psi_{00}(t)$ and $\psi_{02}(t)$ are not mutually orthogonal.

We will now derive the conditions under which the paraunitary property of $[G_s(z), H_s(z)]$ implies the orthonormality condition (11.5.27).

Case when $i = 0$. Eqn. (11.5.27) reduces to

$$\int_{-\infty}^{\infty} |\Psi(\omega)|^2 e^{-j\omega n} \frac{d\omega}{2\pi} = \delta(n) \tag{11.5.37}$$

for all integers n. We will express this in terms of $\Phi(\omega)$ for convenience of future discussion. The functions $\Phi(\omega)$ and $\Psi(\omega)$ are expressible as in (11.5.15).

Using this and the power complementary relation (11.5.10) (induced by paraunitariness), it can be shown that

$$|\Phi(\omega)|^2 + |\Psi(\omega)|^2 = |\Phi(\omega/2)|^2. \tag{11.5.38}$$

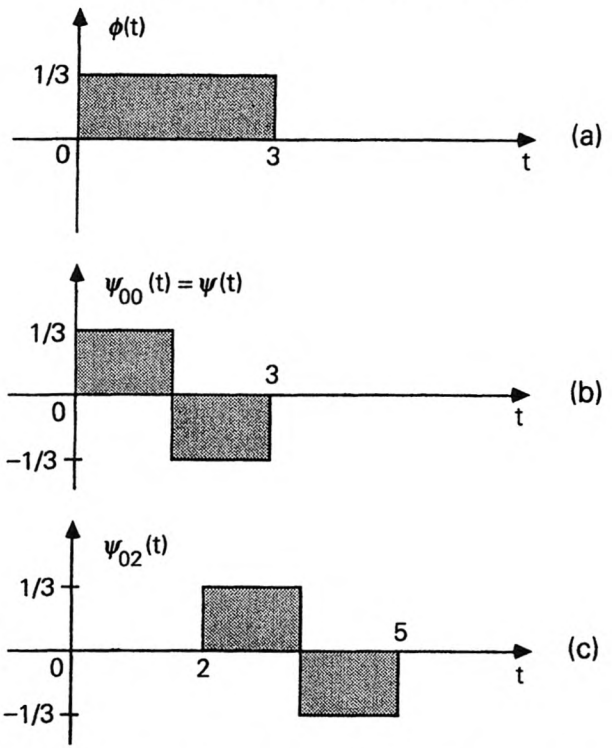

Figure 11.5-9 Example 11.5.6. Generation of basis functions using $G_s(z) = (1 + z^{-3})/\sqrt{2}$, $H_s(z) = (1 - z^{-3})/\sqrt{2}$. (a) Scaling function $\phi(t)$, (b) $\psi(t)$ and $\psi_{00}(t)$ and (c) $\Psi_{02}(t)$.

By using this we can verify that (11.5.37) is satisfied if

$$\int_{-\infty}^{\infty} |\Phi(\omega)|^2 e^{-j\omega n} \frac{d\omega}{2\pi} = \delta(n) \tag{11.5.39}$$

for all integers n. This equation has an interesting time domain interpretation. Thus, define $r_\phi(\tau)$ to be the deterministic autocorrelation function

$$r_\phi(\tau) = \int_{-\infty}^{\infty} \phi(t)\phi^*(t - \tau)\, dt \tag{11.5.40}$$

(as we did in Problem 2.14). Then, $|\Phi(\omega)|^2$ is the Fourier transform of $r_\phi(\tau)$. So (11.5.39) is equivalent to

$$r_\phi(n) = \delta(n), \qquad \text{for all integers } n. \tag{11.5.41}$$

That is, if we 'sample' the function $r_\phi(\tau)$ with sampling period $T = 1$, the result is the unit pulse function $\delta(n)$. In other words, $r_\phi(\tau)$ has periodic zero-crossings, at (nonzero) integer values of the argument τ. This is the continuous-time analog of the *Nyquist(1)* property (Sec. 4.6.1), and we say that $r_\phi(\tau)$ is a Nyquist(1) function.

Using this interpretation and the standard expression (2.1.22) for the Fourier transform of a sampled signal, we can re-express (11.5.39) as

$$\sum_{k=-\infty}^{\infty} |\Phi(\omega + 2\pi k)|^2 = 1. \tag{11.5.42}$$

Summarizing, the equations (11.5.39), (11.5.41) and (11.5.42) describe the same condition. If this condition is satisfied then (11.5.27) holds for $i = 0$.

Case when $i \neq 0$. Now (11.5.27) can be rewritten as

$$\int_{-\infty}^{\infty} \Psi(\omega)\Psi^*(2^i\omega)e^{-j\omega n}\, d\omega = 0, \quad \text{for all integers } n. \tag{11.5.43}$$

By using the definitions of the partial products (11.5.6), we can simplify this to the equivalent form

$$\int_{-\infty}^{\infty} \Psi_1(\omega)\Psi_{i+1}^*(2^i\omega)|\Phi(\omega/2)|^2 e^{-j\omega n}\, d\omega = 0, \quad \text{for all integers } n. \tag{11.5.44}$$

Using the fact that $\Psi_1(\omega)$ and $\Psi_{i+1}^*(2^i\omega)$ have a common period of 4π, this can be rewritten as

$$\sum_{k=-\infty}^{\infty} \int_{-2\pi}^{2\pi} \Psi_1(\omega)\Psi_{i+1}^*(2^i\omega)e^{-j\omega n}\left|\Phi(\frac{\omega}{2} + 2\pi k)\right|^2 d\omega = 0, \quad \text{for all integers } n. \tag{11.5.45}$$

Interchanging the integral with the summation, we can rewrite this as

$$\int_{-2\pi}^{2\pi} \Psi_1(\omega)\Psi_{i+1}^*(2^i\omega)e^{-j\omega n} \sum_{k=-\infty}^{\infty} \left|\Phi(\frac{\omega}{2} + 2\pi k)\right|^2 d\omega = 0, \quad \text{for all integers } n. \tag{11.5.46}$$

If $r_\phi(\tau)$ is Nyquist(1) that is, if (11.5.42) holds, then this is equivalent to

$$\int_{-2\pi}^{2\pi} \Psi_1(\omega)\Psi_{i+1}^*(2^i\omega)e^{-j\omega n} d\omega = 0, \quad \text{for all integers } n. \tag{11.5.47}$$

But we have already shown this to be true in the paraunitary case [set $L = 1$ and $i \neq 0$ in (11.5.29)]. Summarizing, we have proved the following result.

♠**Lemma 11.5.1. Continuous time orthonormal wavelets.** Let $\Psi(\omega)$ and $\Phi(\omega)$ be defined as in (11.5.1) where $G_s(z)$ and $H_s(z)$ satisfy the paraunitary conditions (11.4.13) [i.e., the polyphase matrix $\mathbf{R}(z)$ is paraunitary]. If the scaling function $\phi(t)$ is such that its autocorrelation $r_\phi(\tau)$ is Nyquist(1), [i.e., if any one of (11.5.39), (11.5.41) or (11.5.42) holds], then the wavelet basis functions $\psi_{k\ell}(t) \triangleq 2^{-k/2}\psi(2^{-k}t - \ell)$ are orthonormal. ◊

Example 11.5.7

Consider the Haar basis example again. The function $\phi(t)$ is as in Fig. 11.5-4, and its autocorrelation is a triangular waveform [Fig. 11.5-10(a)], which is zero for $|\tau| \geq 1$. So the Nyqusit(1) property is automatically satisfied, and the wavelet basis $\psi_{k\ell}(t)$ is orthonormal as seen earlier.

For Example 11.5.6 on the other hand, the function $\psi(t)$ [Fig. 11.5-9(a)] has the triangular autocorrelation $r_\phi(\tau)$, shown in Fig. 11.5-10(b). This does not satisfy the Nyquist(1) condition [e.g., $r_\phi(1) \neq 0$]. So we do not expect the basis functions $\psi_{k\ell}(t)$ to be orthonormal. Thus, as demonstrated earlier, $\psi_{00}(t)$ and $\psi_{02}(t)$ are not orthonormal.

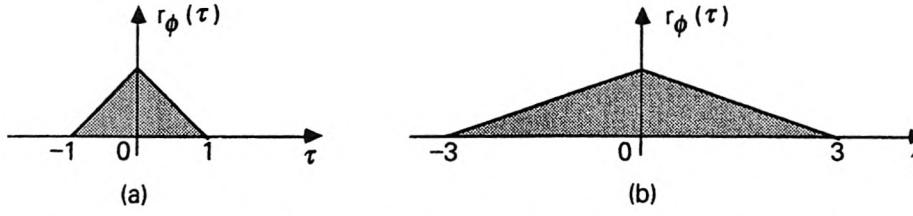

Figure 11.5-10 The autocorrealtion $r_\phi(t)$ of $\phi(t)$, generated from (a) $G_s(z) = (1 + z^{-1})/\sqrt{2}$ and (b) $G_s(z) = (1 + z^{-3})/\sqrt{2}$.

Satisfying the Nyquist(1) Condition

The natural question now is this: how should we design $G_s(z)$ so that the above Nyquist(1) condition is satisfied? The answer to this has been provided by Mallat. We will state this in a slightly modified form here: if $[G_s(z), H_s(z)]$ is a FIR paraunitary pair [i.e., if (11.4.13) holds] and if the following two conditions hold:

1. $|G_s(e^{j0})| = \sqrt{2}$ and
2. $G_s(e^{j\omega}) \neq 0$ for $|\omega| \leq \pi/2$,

then the Nyquist(1) property (11.5.42) is indeed satisfied. The proof is beyond the scope of this chapter, and can be found in Mallat [1989b]. Note that the second condition is trivially satisfied in most QMF designs.

To demonstrate this result, consider again the Haar-basis (example 11.5.3). We have $|G_s(e^{j\omega})| = \sqrt{2}|\cos(\omega/2)|$, and the above two conditions

are satisfied; so the Nyquist(1) condition holds, and the basis $\psi_{k\ell}(t)$ is orthonormal. In example 11.5.6, however, we have $|G_s(e^{j\omega})| = \sqrt{2}|\cos(3\omega/2)|$, so that $G_s(e^{\pm j\pi/3}) = 0$. This violates the second condition above. The Nyquist(1) property is not satisfied in this case, and the basis $\psi_{k\ell}(t)$ is not orthonormal as demonstrated earlier.

We can summarize the main points of the preceding discussions as follows.

♠Theorem 11.5.1. Generating continuous-time finite duration orthonormal wavelets. Suppose $\Psi(\omega)$ and $\Phi(\omega)$ are defined as in (11.5.1).

1. If $G_s(z)$ and $H_s(z)$ are causal FIR filters of order N, then $\phi(t)$ and $\psi(t)$ are causal with duration equal to N. So the functions defined according to $\psi_{k\ell}(t) \triangleq 2^{-k/2}\psi(2^{-k}t - \ell)$ have finite duration.
2. Suppose the FIR filter pair $[G_s(z), H_s(z)]$ is paraunitary i.e., (11.4.13) holds (i.e., the polyphase matrix $\mathbf{R}(z)$ is paraunitary). If in addition $|G_s(e^{j0})| = \sqrt{2}$ and $G_s(e^{j\omega}) \neq 0$ for $|\omega| \leq \pi/2$, the wavelet basis functions $\psi_{k\ell}(t)$ are orthonormal. ♢

Deeper Significance of the Nyquist(1) Condition (11.5.42)

Let $R(z)$ denote the z-transform of the sampled autocorrelation function $r_\phi(nT)$ with $T = 1$, that is, $R(z) = \sum_n r_\phi(n)z^{-n}$. It can be shown (Problem 11.19) that $R(z)$ satisfies the equation

$$\widetilde{G}_s(z)G_s(z)S(z)\Big|_{\downarrow 2} = S(z). \tag{11.5.48}$$

In other words if we substitute $S(z) = R(z)$ into the left side of (11.5.48), it reduces to $R(z)$. In the paraunitary case, we have $\widetilde{G}_s(z)G_s(z)\Big|_{\downarrow 2} = 1$ so that the function $S(z) = \textit{constant}$ is a solution to (11.5.48). If it turns out that the *only* solution to (11.5.48) is a constant, then this implies in particular that $R(z)$ is a constant, i.e., that $r_\phi(\tau)$ is Nyquist(1).

11.5.4 Regularity Considerations

In Fig. 11.5-11 we show the response $|G_s(e^{j\omega})|$ and the corresponding wavelet function $\psi(t)$ generated from the filter pair $[G_s(z), H_s(z)]$, for two cases. In both the examples, $G_s(z)$ and $H_s(z)$ are fifth order FIR filters, designed to satisfy the conditions of Theorem 11.5.1. So the basis functions $\psi_{k\ell}(t)$ are orthonormal.

We see that $\psi(t)$ is much more 'smooth' or 'regular' in part (a) than in part (b) of the figure. Qualitatively speaking, the latter function has strong 'high-frequency' components, so that it is less 'regular'. Smoothness of $\psi(t)$ can be obtained if $G_s(e^{j\omega})$ has a sufficient number of zeros (say K zeros) at $\omega = \pi$ (i.e., $z = -1$). In this case $G_s(e^{j2^{-m}\omega})$ has K zeros at each of the

frequencies

$$\omega = 2^m\pi + 2^m 2\pi\ell, \quad m = 2,3,4\ldots \quad \ell = 0,1,2\ldots \tag{11.5.49}$$

Using this it can be shown (Problem 11.23) that $\Psi(\omega)$ has K zeros at each of the frequencies $4\pi n$, for $n = 1,2,3\ldots$ If $\Psi(\omega)$ is plotted with a logarithmic scale for the frequency axis (Bode plot), these zero locations get more and more crowded as the frequency increases.

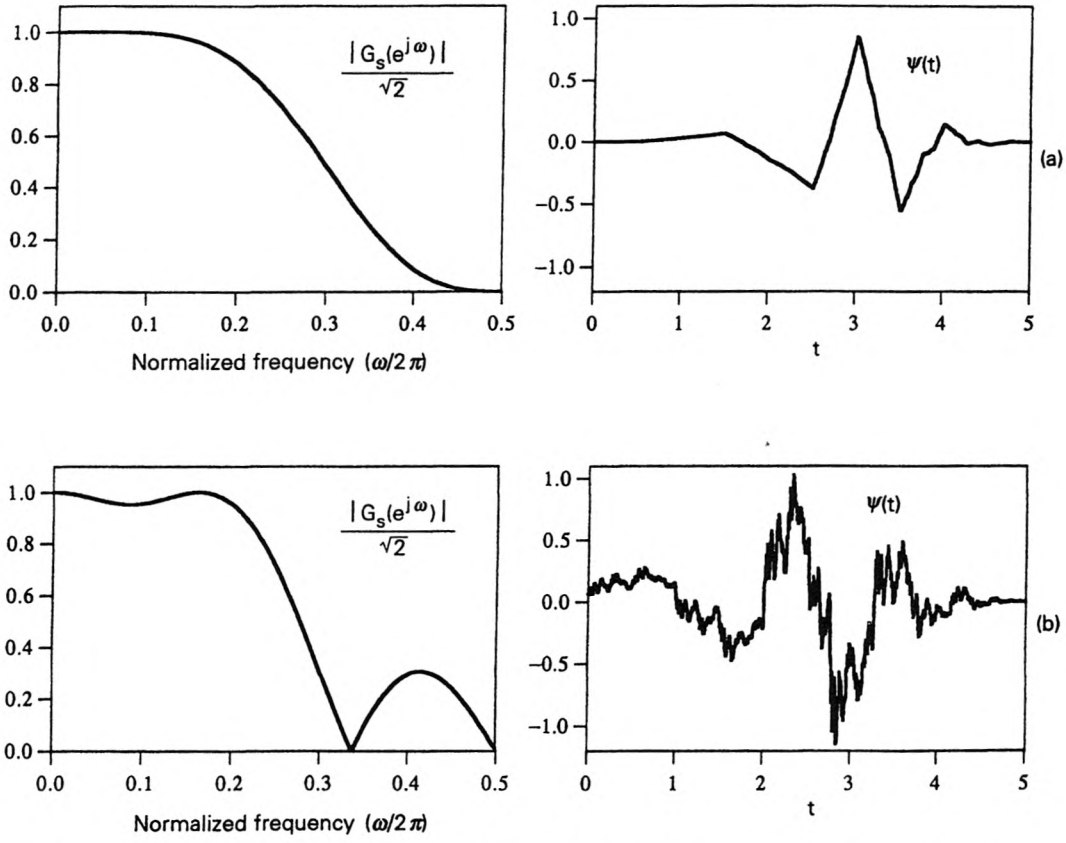

Figure 11.5-11 Two examples of plots of $|G_s(e^{j\omega})|$ and $\psi(t)$. (a) $\psi(t)$ looks regular or smooth and (b) $\psi(t)$ does not look smooth.

The smoothness or regularity of $\psi(t)$ improves with the number of zeros of $G_s(e^{j\omega})$ at $\omega = \pi$. For a more quantitative statement of this, see Daubechies [1988] where the author defines regularity in terms of the asymptotic decrease of the product (11.5.1) as ω increases, and relates this decrease to the number of zeros of $G_s(e^{j\omega})$ at $\omega = \pi$.

Let $G_s(z)$ and $H_s(z)$ be FIR with order N. For orthonormality of $\psi_{k\ell}(t)$, the filters have to satisfy (11.4.13). In particular, $\widetilde{G}_s(z)G_s(z)|_{\downarrow 2} = 1$, which

means that $\widetilde{G}_s(z)G_s(z)$ is a half-band filter. In other words, $G_s(z)$ should be a spectral factor of the half-band filter $F(z) \triangleq \widetilde{G}_s(z)G_s(z)$. Subject to this constraint we will show that, no more than $(N+1)/2$ zeros of $G_s(z)$ can be located at $\omega = \pi$. We will also show how to find the coefficients of $G_s(z)$ such that it has $(N+1)/2$ zeros at $\omega = \pi$. These results are based on the theory of maximally flat FIR filters, developed in [Herrmann, 1971].

Maximally Flat FIR Filters

In Chapter 3 we outlined a number of techniques for FIR filter design, but left out maximally flat (linear-phase) FIR filters. These are the FIR counterparts, in some sense, of IIR Butterworth filters. If we design a maximally flat FIR half band filter and take $G_s(z)$ to be one of its spectral factors, then the function $\psi(t)$ designed as in the previous section has a much smoother plot, as pointed out by Daubechies.

Meaning of maximal flatness. Refer to Fig. 11.5-12, which shows the response $F(e^{j\omega})$ of a zero-phase lowpass filter $F(z)$, with $F(e^{j\pi}) = 0$. Suppose the derivative $dF(e^{j\omega})/d\omega$ has N_0 zeros at $\omega = 0$ and N_π zeros at $\omega = \pi$. We say that the (degree of) flatness is N_0 at $\omega = 0$ and N_π at $\omega = \pi$. Let N_d be the largest possible number of zeros of the derivative in the range $0 \leq \omega \leq \pi$. [This number is determined by the order of $F(z)$.] If $N_d = N_0 + N_\pi$, we say that $F(e^{j\omega})$ is maximally flat. This means that, for a given flatness at $\omega = 0$, the flatness at $\omega = \pi$ has been maximized (or vice versa). Notice that unlike an IIR Butterworth filter, the flatness need not be the same at $\omega = 0$ and π.

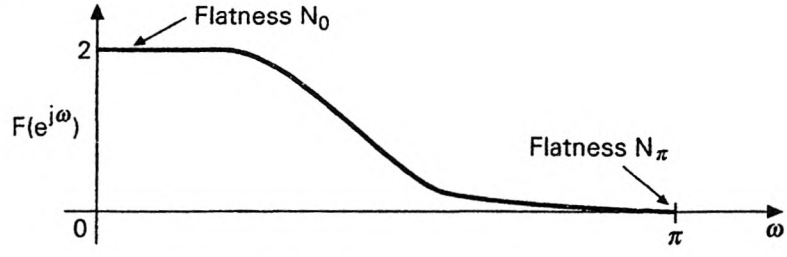

Figure 11.5-12 Pertaining to maximally flat FIR filters.

Closed form expression for the filter coefficients. We now show how to find the coefficients of maximally flat FIR filters. We discuss the problem in terms of a polynomial $P(y)$, and then make the change of variables $y \rightarrow \sin^2(\omega/2)$ to obtain the filter response $F(e^{j\omega})$. Here $P(y) = \sum_{n=0}^{N} p_n y^n$ (i.e., polynomial with order N). Let $P(0) = 2$, and let $K-1$ denote the degree of flatness at $y = 1$ (Fig. 11.5-13). This means $P(y)$ has K zeros at $y = 1$. Let the degree of flatness at $y = 0$ be $L-1$. Thus the derivative $dP(y)/dy$ has $K+L-2$ zeros. Maximal flatness implies that $dP(y)/dy$ has

no other zeros, so that the order of $P(y)$ is

$$N = K + L - 1. \tag{11.5.50}$$

Since $P(y)$ has K zeros at $y = 1$, we can write

$$P(y) = Q(y)(1 - y)^K, \tag{11.5.51}$$

where $Q(y)$ is a polynomial with order $L - 1$, that is,

$$Q(y) = \sum_{\ell=0}^{L-1} q_\ell y^\ell. \tag{11.5.52}$$

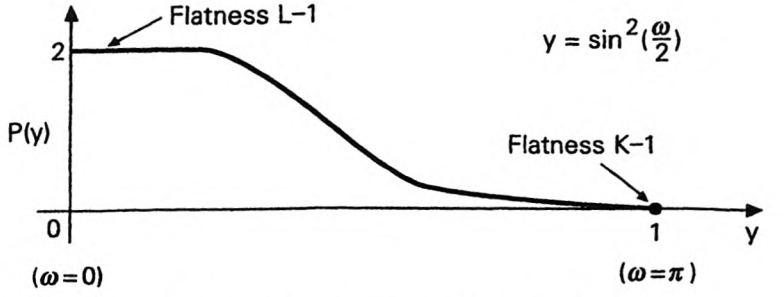

Figure 11.5-13 The maximally flat polynomial.

Our aim now is to find the L unknowns q_ℓ by imposing the following L constraints on $P(y)$:

$$P(0) = 2, \quad P^{(n)}(0) = 0, \; 1 \le n \le L - 1, \tag{11.5.53}$$

where $P^{(n)}(0)$ denotes the nth derivative evaluated at $y = 0$. The process is enormously simplified by rewriting (11.5.51) as

$$Q(y) = P(y)(1 - y)^{-K}. \tag{11.5.54}$$

By applying Leibnitz's rule for the derivative of a product (e.g., see p. 147 of Apostol [1961]), we obtain

$$Q^{(\ell)}(0) = \sum_{n=0}^{\ell} \binom{\ell}{n} P^{(n)}(0) \left(\frac{d^{\ell-n}(1 - y)^{-K}}{dy^{\ell-n}} \Big|_{y=0} \right). \tag{11.5.55}$$

In view of (11.5.53), only the term with $n = 0$ survives, and we obtain

$$Q^{(\ell)}(0) = 2 \frac{d^\ell (1 - y)^{-K}}{dy^\ell} \Big|_{y=0} = 2K(K + 1) \ldots (K + \ell - 1) = \frac{2(K + \ell - 1)!}{(K - 1)!} \tag{11.5.56}$$

Since $q_\ell = \left(Q^{(\ell)}(0)/\ell!\right)$, we finally obtain

$$q_\ell = 2\binom{K+\ell-1}{\ell}, \tag{11.5.57}$$

so that

$$P(y) = (1-y)^K \sum_{\ell=0}^{L-1} 2\binom{K+\ell-1}{\ell} y^\ell. \tag{11.5.58}$$

Summarizing, this function $P(y)$ represents the polynomial with smallest order $(= K+L-1)$ having the following properties: (a) $P(0) = 2, P(1) = 0$, (b) degree of flatness $L-1$ at $y = 0$, and (c) degree of flatness $K-1$ at $y = 1$. Substituting $y = \sin^2(\omega/2)$, we obtain the zero-phase maximally flat FIR filter

$$F(e^{j\omega}) = \cos^{2K}(\omega/2) \sum_{\ell=0}^{L-1} 2\binom{K+\ell-1}{\ell} \sin^{2\ell}(\omega/2). \tag{11.5.59}$$

The filter $F(z)$ has order $2N = 2(K+L-1)$. The transfer function $F(z)$ is verified to be (Problem 11.18)

$$F(z) = z^K \left(\frac{1+z^{-1}}{2}\right)^{2K} \underbrace{\sum_{\ell=0}^{L-1} 2(-z)^\ell \binom{K+\ell-1}{\ell} \left(\frac{1-z^{-1}}{2}\right)^{2\ell}}_{\widehat{F}(z)}. \tag{11.5.60}$$

The filter indicated as $\widehat{F}(z)$ does not have zeros on the unit circle. Its purpose is to provide the passband flatness.

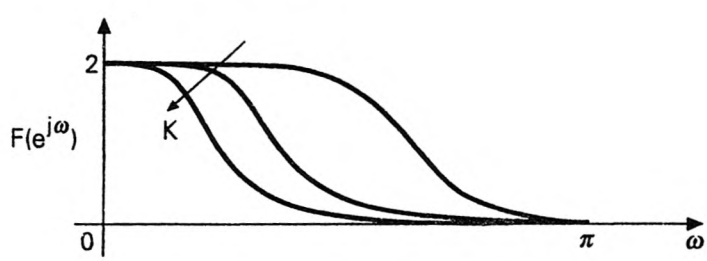

Figure 11.5-14 Effect of changing K, for fixed order $2(K+L-1)$.

Wavelets with Regularity, from Half-band Maximally Flat Filters

The effect of relative values of K and L is clearly demonstrated in Fig. 11.5-14. For fixed order $2(K+L-1)$, increase of K results in wider stopband.

If $K = L$, we obtain a response with symmetry around $\pi/2$. More precisely, in this case we have

$$F(z) + F(-z) = 2. \tag{11.5.61}$$

(See below for proof.) In other words, $F(z)$ is a half band filter.

Design procedure. With $F(z)$ so chosen if we define $G_s(z)$ to be a spectral factor of $F(z)$ and take $H_s(z)$ in the usual manner, that is, $H_s(z) = -z^{-N}\widetilde{G}_s(-z)$, the paraunitary conditions (11.4.13) are satisfied. Furthermore $G_s(e^{j\omega}) \neq 0$ for $|\omega| \leq \pi/2$ so that the conditions of Theorem 11.5.1 are satisfied. If the function $\Psi(\omega)$ is now constructed as in (11.5.1b), then $\{\psi_{k\ell}(t)\}$ is a set of finite duration orthonormal functions with maximum regularity. Notice that even though $G_s(z)$ is a spectral factor of $F(z)$, only a spectral factor $S(z)$ of the function $\widehat{F}(z)$ [indicated in (11.5.60)] has to be computed. We then have

$$G_s(z) = \left(\frac{1+z^{-1}}{2}\right)^K S(z). \tag{11.5.62}$$

Proof that (11.5.61) holds when $K = L$. When $K = L$, the polynomial $P(y)$ has flatness $K-1$ at each of the points $y = 0$ and $y = 1$. Defining $R(y) = P(1-y)$, we see that this polynomial has the same flatness at these two points, but satisfies $R(0) = 0$ and $R(1) = 2$. Thus $P(y) + R(y) = 2$ at $y = 0$ as well as $y = 1$. This sum has the same flatness $K-1$ at $y = 0$ and $y = 1$, i.e., its derivative has a total of $2K-2$ zeros if we count the points $y = 0$ and $y = 1$. Since the order of $P(y) + R(y)$ is $2K-1$, the derivative cannot have any further zeros in $0 < y < 1$. This shows that $P(y) + R(y) = 2$ for all y. Substituting $y = \sin^2(\omega/2)$, this gives $F(e^{j\omega}) + F(-e^{j\omega}) = 2$ which is equivalent to (11.5.61). ▽▽▽

From the above discussion it is also clear that the maximum possible number of zeros of $G_s(e^{j\omega})$ at the frequency $\omega = \pi$ is given by $K = (N+1)/2$.

Order estimation. In traditional signal processing applications, maximally flat FIR filters are not used as commonly as equiripple filters. This is because, for a given set of specifications (Fig. 3.1-1), the filter order is much higher. For example if we let $2\delta_1 = \delta_2 = 0.05$, then the filter order for a given transition width Δf is estimated to be [Kaiser, 1979]

$$2N \approx \frac{1}{2(\Delta f)^2}. \tag{11.5.63}$$

The order grows as $1/(\Delta f)^2$, and not as $1/\Delta f$ as in the equiripple case!

Design Example 11.5.2: Wavelet Regularity

Let $K = L = 3$ so that the maximally flat filter $F(z)$ has order $2(K + L - 1) = 10$. The spectral factor $G_s(z)$ has order $N = 5$, with three zeros

at $\omega = \pi$ (since $K = 3$). The response $|G_s(e^{j\omega})|/\sqrt{2}$ for this example was shown in Fig. 11.5-11(a), along with the function $\psi(t)$ derived from this filter.

Table 11.5.1 provides a summary of the key concepts and equations, pertaining to the generation of continuous time wavelets.

11.6 CONCLUDING REMARKS

Wavelets were thoroughly studied in the mathematics literature only after the mid 1980s. Most of the analysis was confined to the continuous time case, which has a wider scope for deeper mathematial issues. Discrete time wavelet transforms, on the other hand, are equivalent to tree structured digital filter banks, and can be understood with the help of elementary signal processing theory, and a fair amount of matrix theory.

Even before the development of wavelets and paraunitary filter banks, nonuniform filter banks have been used in speech processing literature [Nelson, et al., 1972], [Schafer, et al., 1975]. The motivation at that time was that, nonuniform bandwidths could be used to exploit the nonuniform frequency resolution of the human ear [Flanagan, 1972]. The more recent results on wavelets in the mathematical and signal processing literature enhance our understanding, and enable us to perform orthonormal decomposition.

As we saw in this chapter, discrete time orthonormal wavelet transforms are very easy to implement, simply by designing a two channel paraunitary QMF bank and then building the tree structure. If we use the cascaded lattice structure (Sec. 6.4), then the paraunitary property is retained in spite of multiplier quantization. This means that both the perfect reconstruction and the wavelet orthonormality properties can be retained in spite of multiplier quantization. Furthermore, the lattice structure generates the complete class of FIR orthonormal basis functions.

It does not appear to be appropriate to obtain discrete time wavelets by discretizing the continuous time version. Such an approach would typically result in the loss of many of the desirable properties such as orthonormality and perfect reconstruction. Furthermore, as we found in Sec. 11.5, continous time wavelets are often generated by starting from discrete time filter banks anyway.

Thus, if the wavelet application is already in the digital domain, it is really not necessary to understand the deeper results [e.g., the fundamental functions $\phi(t), \psi(t)$, self similarity, and so forth] developed in Sec. 11.5. In this case it is sufficient to understand the results of Sec. 11.4; the nonuniform nature of the digital filter responses (as well as the nonuniform decimation) already provides the nonuniform time-frequency grid [Fig. 11.3-3(b)], which is the key to many of the advantages of wavelet transformation.

Symmetric wavelet basis. In Chap. 7 we saw that two-channel FIR perfect reconstruction QMF banks cannot simultaneously have linear phase and paraunitary properties, unless the filters have fairly trivial forms. As a

TABLE 11.5.1 Generation of continuous-time wavelet basis

Definition of the infinite products:

$$\Phi(\omega) = \frac{1}{\sqrt{2}} G_s(e^{j\omega/2}) \prod_{m=2}^{\infty} \frac{1}{\sqrt{2}} G_s(e^{j2^{-m}\omega}), \tag{11.5.1a}$$

$$\Psi(\omega) = \frac{1}{\sqrt{2}} H_s(e^{j\omega/2}) \prod_{m=2}^{\infty} \frac{1}{\sqrt{2}} G_s(e^{j2^{-m}\omega}), \tag{11.5.1b}$$

where $G_s(z) = \sum_n g_s(n) z^{-n}$ and $H_s(z) = \sum_n h_s(n) z^{-n}$. Then

$$\Phi(\omega) = \frac{1}{\sqrt{2}} G_s(e^{j\omega/2}) \Phi(\omega/2), \quad \Psi(\omega) = \frac{1}{\sqrt{2}} H_s(e^{j\omega/2}) \Phi(\omega/2). \tag{11.5.15}$$

Equivalently

$$\phi(t) = \sqrt{2} \sum_{n=-\infty}^{\infty} g_s(n) \phi(2t-n), \quad \psi(t) = \sqrt{2} \sum_{n=-\infty}^{\infty} h_s(n) \phi(2t-n). \tag{11.5.18}$$

If $G_s(z)$ and $H_s(z)$ are causal FIR (order N), the functions $\psi(t)$ and $\phi(t)$ are causal, and of finite duration N.

Definition of the basis functions: $\psi_{k\ell}(t) = 2^{-k/2} \psi(2^{-k}t - \ell)$

Orthonormality

Suppose $G_s(z)$ and $H_s(z)$ are causal and FIR, satisfying:

1. $|G_s(e^{j\omega})|^2 + |H_s(e^{j\omega})|^2 = 2$
2. $H_s(z) = z^{-N} \widetilde{G}_s(-z)$ for some odd N.

Equivalently,

$$G_s(z)\widetilde{G}_s(z)\Big|_{\downarrow 2} = 1, \quad H_s(z)\widetilde{H}_s(z)\Big|_{\downarrow 2} = 1, \quad G_s(z)\widetilde{H}_s(z)\Big|_{\downarrow 2} = 0. \tag{11.4.13}$$

In other words, $[G_s(z), H_s(z)]$ is an FIR paraunitary pair, i.e., the polyphase matrix $\mathbf{R}(z)$ satisfies $\widetilde{\mathbf{R}}(z)\mathbf{R}(z) = \mathbf{I}$. If we further have

$$|G_s(e^{j0})| = \sqrt{2}, \quad \text{and} \quad G_s(e^{j\omega}) \neq 0, \ |\omega| \leq \pi/2,$$

then $\{\psi_{k\ell}(t)\}$ forms an orthonormal set. In particular, each function $\psi_{k\ell}(t)$ has unit energy. The Haar basis (Example 11.5.5) is a familiar example of this. Also see Lemma 11.5.1 and Theorem 11.5.1.

Regularity. If the above $G_s(z)$ is designed to be a spectral factor of a maximally flat FIR half-band filter $F(z)$ [i.e., $F(z)$ as in (11.5.60) with $K = L$] then $\psi(t)$ exhibits very smooth behavior. See Fig. 11.5-11 for demonstration. Thus $\{\psi_{k\ell}(t)\}$ are *regular, orthonormal,* and of *finite* duration for finite k.

result, if we confine the orthonormal wavelet basis functions $\eta_{km}(n)$ in Sec. 11.4.2 to have linear phase, then the functions $F_k(z)$ are severely restricted. If we give up the paraunitary property (hence wavelet orthonormality), it is possible to design linear phase perfect reconstruction systems (i.e., symmetric wavelet basis functions) with greater flexibility of coefficients. This was demonstrated in Chap. 7. Further discussion of symmetric nonorthonormal wavelets can be found in Vetterli and Herley [1992]. For the case of *symmetric orthonormal* wavelets, see [Soman, Vaidyanathan and Nguyen, 1992]. For the case of nonbinary tree strutures (as used in wavelet packets) the simultaneous imposition of paraunitary and linear phase properties does not severely restrict the filter responses, and it is possible to obtain nontrivial, symmetric, orthonormal wavelet basis functions. Also see Problem 11.17.

PROBLEMS

11.1. Consider the following system where $H(e^{j\omega})$ is an ideal filter, with center frequency ω_0.

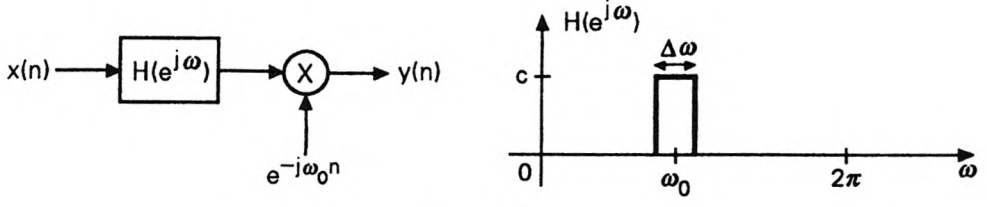

Figure P11-1

Let $c > 0$ be the height of the passband response. Assume that the Fourier transform $X(e^{j\omega})$ is constant in the passband of the filter. We know that $y(n)$ is a slowly varying signal for small $\Delta\omega$ (slower for smaller $\Delta\omega$).

a) Find an expression for $y(n)$.
b) For what value of n is $|y(n)|$ maximum?
c) Let $|y(n)|$ be maximum for $n = n_0$. Suppose we wish to have $X(e^{j\omega_0}) = y(n_0)$. Show then that $c = 2\pi/\Delta\omega$.

This shows that, in order for the outputs of a filter bank to deliver a snap-shot of the (time-varying) Fourier transform, the filter *heights* should be inversely proportional to the bandwidths.

11.2. Consider the STFT system in (11.2.2) where $v(n)$ is of finite duration N. Suppose we implement this as a uniform DFT system (11.2.5), where $H_0(z)$ is as in (11.2.6). If we wish to have an FIR inversion system, show that $v(n)$ cannot have more than M nonzero coefficients.

11.3. Consider the discrete-time STFT and its inverse given in (11.2.1) and (11.2.9) respectively.

a) Show that if the relation (11.2.1) is used in (11.2.9) this results in $x(n) = x(n)\sum_m |v(m)|^2$.
b) Let z_0 be a zero of $\sum_k v^*(k)z^{-k}$. Show that if we substitute z_0^m in place of $X_{STFT}(e^{j\omega}, m)$ on the right hand side of (11.2.9), it reduces to zero. This shows that if $X_{STFT}(e^{j\omega}, m)$ satisfies the inverse relation (11.2.9), then so does $X_{STFT}(e^{j\omega}, m) + cz_0^m$ for any constant c.

11.4. Consider the sampled STFT given in (11.2.27). Here the STFT is evaluated on a two-dimensional grid with samples located at $\Omega = k\Omega_0$ and $\tau = nT$. Define the quantities

$$Y(\tau, \Omega) = \frac{1}{T}\sum_{k=-\infty}^{\infty}\sum_{n=-\infty}^{\infty} \mathcal{X}(k, n)e^{-j\Omega(nT)}e^{j\tau(k\Omega_0)}, \qquad (P11.4a)$$

$$X_1(\tau, \Omega) = \sum_{m=-\infty}^{\infty} x(\tau + mT)e^{-j\Omega(mT)}, \qquad (P11.4b)$$

$$V_1(\tau,\Omega) = \sum_{m=-\infty}^{\infty} v^*(\tau + mT)e^{-j\Omega(mT)}. \qquad (P11.4c)$$

Assume that the sample spacings Ω_0 and T are related as $\Omega_0 T = 2\pi$.

a) Verify that $Y(\tau,\Omega)$ is periodic in τ with period T, and periodic in Ω with period Ω_0. Also verify that $X_1(\tau,\Omega)$ and $V_1(\tau,\Omega)$ are periodic in Ω with period Ω_0.

b) Assuming that summations and integrals can be interchanged when necessary, show that

$$Y(\tau,\Omega) = X_1(\tau,\Omega)V_1^*(\tau,\Omega). \qquad (P11.4d)$$

Thus, given the sampled STFT $\mathcal{X}(k,n)$ and the window $v(t)$, we can recover $x(t)$ as follows:

a) Compute $Y(\tau,\Omega)$ for $0 \le \tau < T$ and $0 \le \Omega < \Omega_0$.

b) Compute $V_1(\tau,\Omega)$ for $0 \le \tau < T$ and $0 \le \Omega < \Omega_0$.

c) Using (P11.4d), compute $X_1(\tau,\Omega)$ for $0 \le \tau < T$ and $0 \le \Omega < \Omega_0$. (This assumes $V_1(\tau,\Omega) \neq 0$ in this range).

d) Find $x(\tau + mT)$ by using inverse Fourier transform relation corresponding to (P11.4b). This gives $x(\tau + mT)$ for any integer m and for any τ in the range $0 \le \tau < T$. So $x(t)$ can be recovered for any t.

11.5. For the signals given in Fig. 11.2-11 (a), (b), and (c), verify the indicated expressions for the RMS duration D_t.

11.6. Recall the RMS durations defined in (11.2.23). We now show $D_t D_f \ge 0.5$. For simplicity, assume $x(t)$ is a real function of the real variable t. Let $X(j\Omega)$ denote its Fourier transform. For the purpose of this problem, it is useful to review Cauchy-Schwartz inequality [Appendix C, Problem C.7(c)]. All integrals are in the range $-\infty$ to ∞. Assume that all relevant integrals exist, and that $x(t)$ has 'sufficient decay' so that $tx^2(t) \to 0$ as $t \to \pm\infty$.

a) Prove the inequality

$$\left(\int tx(t)\frac{dx(t)}{dt}dt\right)^2 \le \int t^2x^2(t)dt \int \left(\frac{dx(t)}{dt}\right)^2 dt. \qquad (P11.6)$$

b) Using the fact that $j\Omega X(j\Omega)$ is the Fourier transform of $dx(t)/dt$, prove that the right hand side of the above inequality is equal to $E^2D_t^2D_f^2$, where $E = \int x^2(t)dt$.

c) Show that $\int tx(t)\frac{dx(t)}{dt}dt = -0.5E$. (Use integration by parts, i.e., $\int udv = uv - \int vdu$.)

d) Combine these results to prove $D_tD_f \ge 0.5$. From Cauchy-Schwartz inequality we know that equality occurs if and only if $tx(t) = cdx(t)/dt$ for some constant c. By integrating this, show that $D_tD_f = 0.5$ if and only if $x(t) = Ae^{-\alpha t^2}$ for some real A and $\alpha > 0$. In other words, $x(t)$ is a Gaussian pulse.

11.7. The continuous-time Fourier transform pair is given by (2.1.20) and (2.1.21). From this one can verify the Fourier transform relations: (i) $dx(t)/dt \leftrightarrow j\Omega X(j\Omega)$, and (ii) $tx(t) \leftrightarrow jdX(j\Omega)/d\Omega$. Now suppose $x(t) = e^{-t^2/(2\sigma^2)}$. (Note that the subscripts a have been dropped for simplicity.)

a) Using the above Fourier transform properties, show that $X(j\Omega)$ satisfies the differential equation $[dX(j\Omega)/d\Omega] = -\sigma^2\Omega X(j\Omega)$.

b) By appropriate integration show that $X(j\Omega) = ce^{-\sigma^2\Omega^2/2}$.

c) By using the following facts:

$$c = X(0) = \int e^{-t^2/(2\sigma^2)}dt, \quad x(0) = 1 = \frac{c}{2\pi}\int e^{-\sigma^2\Omega^2/2}d\Omega,$$

show that $c = \sigma\sqrt{2\pi}$. Thus $X(j\Omega) = \sigma\sqrt{2\pi}e^{-\sigma^2\Omega^2/2}$, a well-known result.

11.8. Let $x(t) = e^{-t^2/(2\sigma^2)}$ where $\sigma > 0$. We know from Problem 11.7 that $X(j\Omega) = \sigma\sqrt{2\pi}e^{-\sigma^2\Omega^2/2}$. Denote the RMS time duration by $D_t(\sigma)$.

a) Show that the RMS frequency duration is given by

$$D_f(\sigma) = \frac{D_t(\sigma)}{\sigma^2} \qquad (P11.8)$$

b) By using the fact that $D_t(\sigma)D_f(\sigma) = 0.5$ for Gaussian $x(t)$, show that $D_t(\sigma) = \sigma/\sqrt{2}$.

11.9. Let $A(z)$ and $B(z)$ be rational transfer functions, and let m_1 and m_2 be positive integers. Show that

$$\left(A(z^{m_1})B(z)\right)\Big|_{\downarrow m_1 m_2} = \left(A(z)(B(z)|_{\downarrow m_1})\right)\Big|_{\downarrow m_2}. \qquad (P11.9)$$

11.10. Show that the Type 2 polyphase matrix $\mathbf{R}(z)$ of a synthesis bank is related to the filter matrix (11.4.24) according to (11.4.25).

11.11. Consider the synthesis bank shown below.

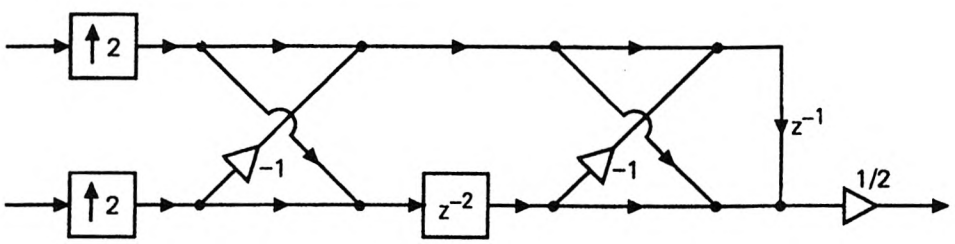

Figure P11-11

We know how to associate a polyphase matrix $\mathbf{R}(z)$ with this system. Verify that $\mathbf{R}(z)$ satisfies $\tilde{\mathbf{R}}(z)\mathbf{R}(z) = \mathbf{I}$. From Sec. 11.4.1 we know how to generate an orthonormal set of discrete-time wavelet basis functions $\eta_{km}(n)$ from this filter bank. List the four sequences $\eta_{km}(n)$ for $k = 0, 1$, and $m = 0, 1$. Verify that these are indeed orthonormal.

11.12. Consider the two-level tree structured synthesis bank shown in Fig. P11-12.

Let $G_s(z) = c(1 + z^{-1})$ and $H_s(z) = c(1 - z^{-1})$, where c is some positive constant. From Sec. 11.3.3 we know how to generate discrete-time wavelet basis functions $\eta_{km}(n)$ from this filter bank.

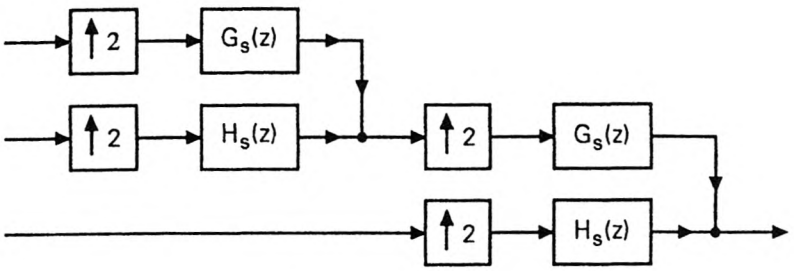

Figure P11-12

a) List the six sequences $\eta_{km}(n)$ for $k = 0, 1$, and 2, and $m = 0, 1$.
b) Pick any two sequences from the above list as you wish, and verify that they are orthonormal for appropriate choice of c.

11.13. Show that the condition (11.4.14) can be rewritten as (11.4.15). Clearly explain where the inequality $2^{n_k} \geq 2^{n_\ell}$ is used in this rewriting.

11.14. Consider the infinite product $\prod_{m=1}^{\infty} a^{b^m}$, where $|b| < 1$. We show that this converges to the quantity $S = a^{b/(1-b)}$. Note that, since $|b| < 1$, we can write $S = a^{(\sum_{m=1}^{\infty} b^m)}$. Define the partial product $S_L = \prod_{m=1}^{L} a^{b^m}$. Show that $S_L - S = S \times (a^{-b^{L+1}/(1-b)} - 1)$. Hence show that $S_L \to S$ as $L \to \infty$.

11.15. Let $G_s(e^{j\omega/2})$ and $H_s(e^{j\omega/2})$ be as in Fig. 11.5-1. Then $\Psi(\omega)$ is the ideal bandpass function shown in Fig. 11.5-2. Prove then that the wavelet basis functions $\psi_{k\ell}(t)$ [defined in (11.5.3)] form an orthonormal set. [*Hint.* It might help to note that $\int_0^{2\pi} e^{j\omega n} d\omega = 2\pi\delta(n)$.]

11.16. Consider the Harr basis generated in Example 11.5.5. Sketch $\psi_{3,0}(t)$. Also sketch the functions $\psi_{1,\ell}(t)$ for $0 \leq \ell \leq 3$, and verify that these are orthogonal to $\psi_{3,0}(t)$.

11.17. In Sec. 11.5 we generated the functions $\phi(t)$ and $\psi(t)$ starting from the discrete-time transfer functions $G_s(z)$ and $H_s(z)$. Suppose the filters $G_s(z)$ and $H_s(z)$ are as in (11.5.23), that is, causal FIR. From Chap. 7 we know that $G_s(z)$ and $H_s(z)$ cannot have linear phase unless we give up the paraunitary property of the polyphase matrix. (The only exception to this gives trivial frequency responses, as seen in Chap. 7). In Chap. 7 we also found that we can obtain two-channel FIR perfect reconstruction systems with linear phase analysis and synthesis filters if we give up the paraunitary property. Under this condition, the functions $\phi(t)$ and $\psi(t)$ exhibit some kind of symmetry, which you prove in this problem. We say that a function $f(t)$ is symmetric if $f(t) = f(2t_0 - t)$ for some finite t_0 [antisymmetric if $f(t) = -f(2t_0 - t)$] and t_0 is said to be the center of symmetry.

a) Assume $g_s(n) = g_s(N - n)$ and $h_s(n) = -h_s(N - n)$ (which implies lin-

ear phase in the real coefficient case). Does $\phi(t)$ exhibit symmetry or antisymmetry? If so, what is the center of symmetry or antisymmetry?

b) Repeat part (a) with $\psi(t)$ instead of $\phi(t)$.

By using the FIR perfect reconstruction QMF bank with the above synthesis filters, it is possible to obtain the so-called bi-orthonormal symmetric wavelet functions with finite duration. For further details, see Vetterli and Herley [1992], and Soman, Vaidyanathan and Nguyen [1992].

11.18. Starting from the frequency response (11.5.59), verify that the transfer function $F(z)$ is indeed given by (11.5.60).

11.19. Let $r_\phi(\tau)$ be the autocorrelation function of the scaling function $\phi(t)$, defined as in (11.5.40). Define the z-transform $R(z) = \sum_n r_\phi(n)z^{-n}$. It can then be shown that $R(z)$ satisfies

$$\tilde{G}_s(z)G_s(z)R(z)\Big|_{\downarrow 2} = R(z). \qquad (P11.19)$$

a) Verify the above equation for the function $\phi(t)$ in Fig. 11.5-4. Repeat the same for Fig. 11.5-9.

b) More generally, give a proof of (P11.19).

11.20. Consider the functions $\Phi(\omega)$ and $\Psi(\omega)$ defined in Sec. 11.5, and assume that $[G_s(z), H_s(z)]$ is a paraunitary pair. Show then that

$$|\Psi(\omega)|^2 + |\Psi(2\omega)|^2 + |\Psi(4\omega)|^2 + \ldots + |\Psi(2^L\omega)|^2 + |\Phi(2^L\omega)|^2 = |\Phi(\omega/2)|^2,$$

for any integer $L \geq 0$. This is pictorially demonstrated as shown in Fig. P11-20. if the "squeezed" bandpass filters generated from $|\Psi(\omega)|^2$ and the 'squeezed' lowpass filter $|\Phi(2^L\omega)|^2$ are added, the result is precisely the *stretched* lowpass filter $|\Phi(\omega/2)|^2$ (The figure assumes magnitude symmetry with respect to $\omega = 0$.)

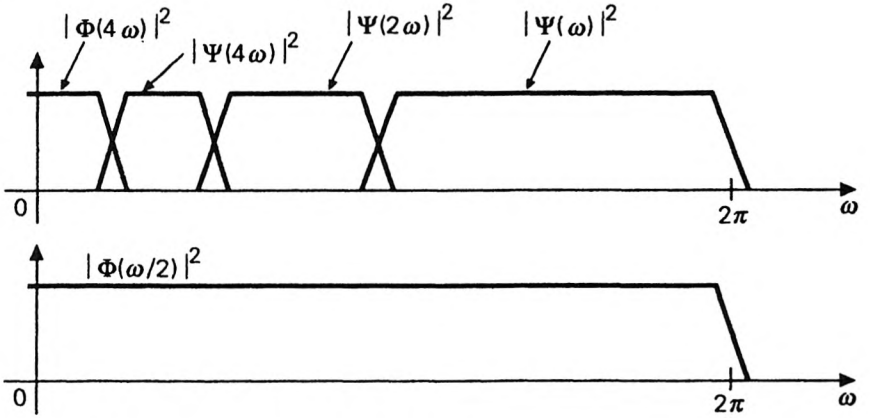

Figure P11-20

11.21. Consider the maximally decimated filter bank of Fig. 5.4-1, with a wide sense stationary input $x(n)$. Suppose the analysis bank is FIR and paraunitary (i.e.,

the polyphase matrix $\mathbf{E}(z)$ (Fig. 5.5-3(b)) is paraunitary). Let $x(n)$ be zero-mean and white. Show that the decimated subband output signals $y_k(n)$ are uncorrelated, i.e., $E[y_k(n)y_m^*(i)] = 0$ for $k \neq m$. (For consistency, we have used $y_k(n)$ in place of the notation $v_k(n)$ used in Fig. 5.4-1.) Show also that all signals $y_k(n)$ are white, with equal variance for all k. (*Note.* These properties would not, in general, be true if $\mathbf{E}(z)$ were not paraunitary.)

11.22. *Most general, nonuniform, orthonormal, discrete-time wavelets.* In Problem 5.32 we introduced the filter bank with nonuniform integer decimators n_k. Assume that this system is maximally decimated, that is, $\sum 1/n_k = 1$. Let $y_k(n)$, $0 \leq k \leq M-1$ be the decimated subband outputs. Assuming perfect reconstruction (i.e., $\hat{x}(n) = x(n)$) show that we can express $x(n)$ in terms of the synthesis filters as

$$x(n) = \sum_{k=0}^{M-1} \sum_{m} y_k(m) f_k(n - n_k m). \qquad (P11.22a)$$

Thus, $x(n)$ has been expanded in terms of the basis functions $\eta_{km}(n) = f_k(n - n_k m)$, and $y_k(m)$ can be regarded as the generalized wavelet coefficients.

a) Recall that the basis $\eta_{km}(n)$ is orthonormal if the synthesis filters satisfy (11.2.19). Show that this orthonormality condition can be rewritten as

$$\sum_{n=-\infty}^{\infty} f_{k_1}^*(n) f_{k_2}(n - g_{k_1 k_2} i) = \delta(k_1 - k_2)\delta(i), \qquad (P11.22b)$$

where $g_{k_1 k_2}$ is the greatest common divisor (gcd) of n_{k_1} and n_{k_2}.

b) How would you extend the result of Problem 11.21 to this nonuniform orthonormal case?

11.23. Assume that $G_s(e^{j\omega})$ has K zeros at the frequency $\omega = \pi$, and that the infinite product defining $\Psi(\omega)$ converges. Show that $\Psi(\omega)$ has K zeros at each frequency of the form $4\pi n$, $n =$ positive integer.