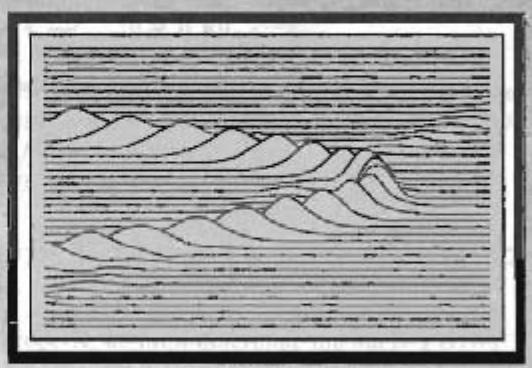
10 Fourier Analysis of Signals Using the Discrete Fourier Transform



10.0 INTRODUCTION

In Chapter 8, we developed the discrete Fourier transform (DFT) as a Fourier representation of finite-length signals. Because the DFT can be computed efficiently, it plays a central role in a wide variety of signal-processing applications, including filtering and spectrum analysis. In this chapter, we take an introductory look at Fourier analysis of signals using the DFT.

In applications and algorithms based on explicit evaluation of the Fourier transform, it is ideally the discrete-time Fourier transform (DTFT) that is desired, although it is the DFT that can actually be computed. For finite-length signals, the DFT provides frequency-domain samples of the DTFT, and the implications of this sampling must be clearly understood and accounted for. For example, as considered in Section 8.7, in linear filtering or convolution implemented by multiplying DFTs rather than DTFTs, a circular convolution is implemented, and special care must be taken to ensure that the results will be equivalent to a linear convolution. In addition, in many filtering and spectrum analysis applications, the signals do not inherently have finite length. As we will discuss, this inconsistency between the finite-length requirement of the DFT and the reality of indefinitely long signals can be accommodated exactly or approximately through the concepts of *windowing*, *block processing*, and the *time-dependent Fourier transform*.

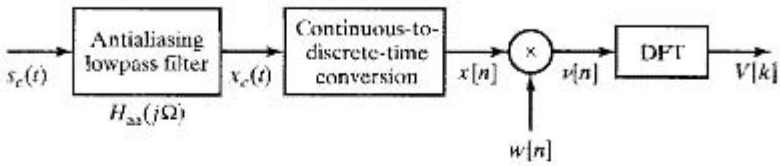


Figure 10.1 Processing steps in the discrete-time Fourier analysis of a continuous-time signal.

10.1 FOURIER ANALYSIS OF SIGNALS USING THE DFT

One of the major applications of the DFT is in analyzing the frequency content of continuous-time signals. For example, as we describe in Section 10.4.1, in speech analysis and processing, frequency analysis is particularly useful in identifying and modeling the resonances of the vocal cavity. Another example, introduced in Section 10.4.2, is Doppler radar, in which the velocity of a target is represented by the frequency shift between the transmitted and received signals.

The basic steps in applying the DFT to continuous-time signals are indicated in Figure 10.1. The antialiasing filter is incorporated to eliminate or minimize the effect of aliasing when the continuous-time signal is converted to a sequence by sampling. The need for multiplication of $x[n]$ by $w[n]$, i.e., windowing, is a consequence of the finite-length requirement of the DFT. In many cases of practical interest, $s_c(t)$ and, consequently, $x[n]$ are very long or even indefinitely long signals (such as with speech or music). Therefore, a finite-duration window $w[n]$ is applied to $x[n]$ prior to computation of the DFT. Figure 10.2 illustrates the Fourier transforms of the signals in Figure 10.1. Figure 10.2(a) shows a continuous-time spectrum that tapers off at high frequencies but is not bandlimited. It also indicates the presence of some narrowband signal energy, represented by the narrow peaks. The frequency response of an antialiasing filter is illustrated in Figure 10.2(b). As indicated in Figure 10.2(c), the resulting continuous-time Fourier transform $X_c(j\Omega)$ contains little useful information about $S_c(j\Omega)$ for frequencies above the cutoff frequency of the filter. Since $H_{aa}(j\Omega)$ cannot be ideal, the Fourier components of the input in the passband and the transition band also will be modified by the frequency response of the filter.

The conversion of $x_c(t)$ to the sequence of samples $x[n]$ is represented in the frequency domain by periodic replication, frequency normalization, and amplitude scaling i.e.,

$$X(e^{j\omega}) = \frac{1}{T}\sum_{r=-\infty}^{\infty} X_c\left(j\frac{\omega}{T} + j\frac{2\pi r}{T}\right). \tag{10.1}$$

This is illustrated in Figure 10.2(d). In a practical implementation, the antialiasing filter cannot have infinite attenuation in the stopband. Therefore, some nonzero overlap of the terms in Eq. (10.1), i.e., aliasing, can be expected; however, this source of error can be made negligibly small either with a high-quality continuous-time filter or through the use of initial oversampling followed by more effective discrete-time lowpass filtering and decimation, as discussed in Section 4.8.1. If $x[n]$ is a digital signal, so that A/D conversion is incorporated in the second system in Figure 10.1, then quantization error is also introduced. As we have seen in Section 4.8.2, this error can be modeled as a

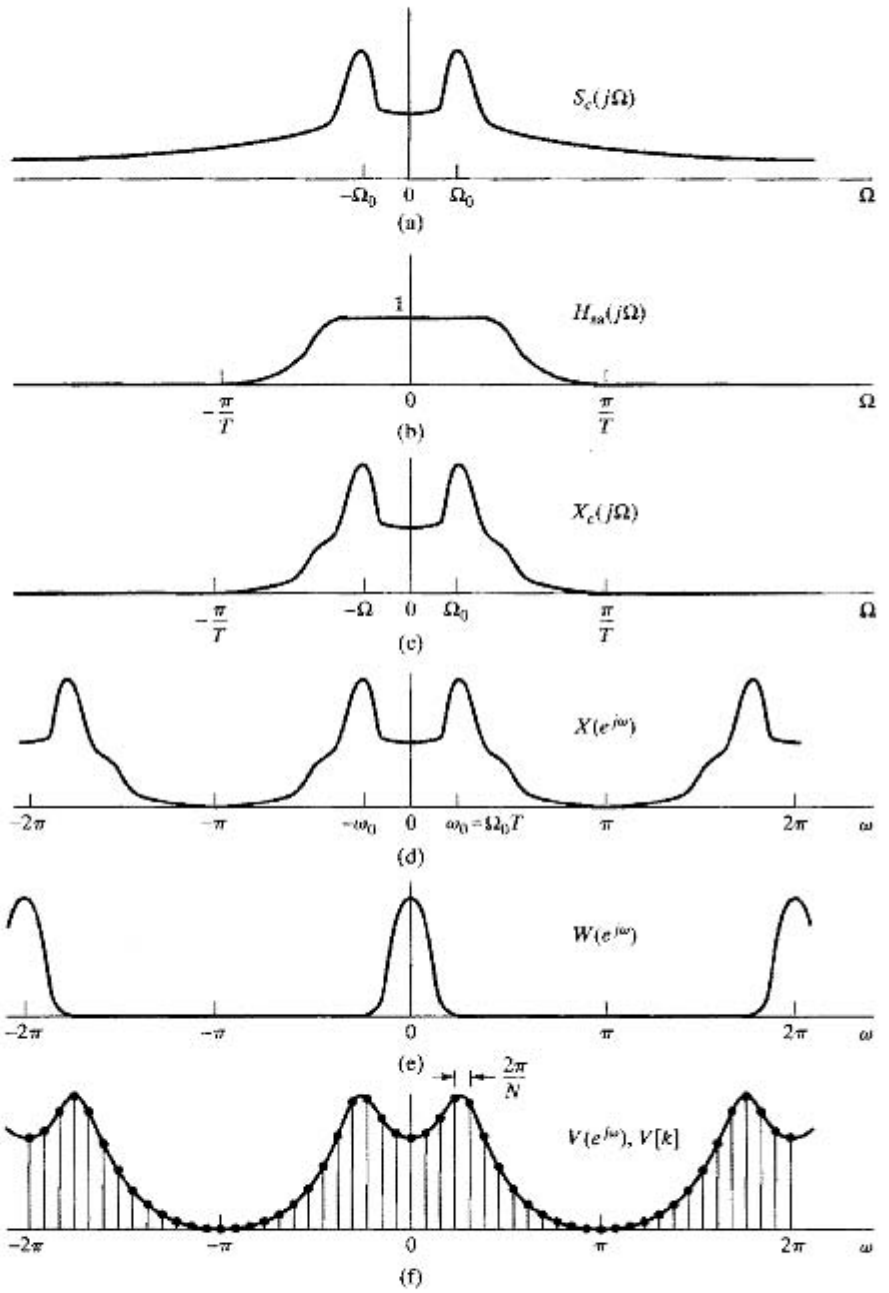


Figure 10.2 Illustration of the Fourier transforms of the system of Figure 10.1. (a) Fourier transform of continuous-time input signal. (b) Frequency response of antialiasing filter. (c) Fourier transform of output of antialiasing filter. (d) Fourier transform of sampled signal. (e) Fourier transform of window sequence. (f) Fourier transform of windowed signal segment and frequency samples obtained using DFT samples.

noise sequence added to $x[n]$. The noise can be made negligible through the use of fine-grained quantization.

The sequence $x[n]$ is typically multiplied by a finite-duration window $w[n]$, since the input to the DFT must be of finite duration. This produces the finite-length sequence $v[n] = w[n]x[n]$. The effect in the frequency domain is a periodic convolution, i.e.,

$$V(e^{j\omega}) = \frac{1}{2\pi}\int_{-\pi}^{\pi} X(e^{j\theta})W(e^{j(\omega-\theta)})d\theta. \tag{10.2}$$

Figure 10.2(e) illustrates the Fourier transform of a typical window sequence. Note that the main lobe is assumed to be concentrated around $\omega = 0$, and, in this illustration, the side lobes are very small, suggesting that the window tapers at its edges. The properties of windows such as the Bartlett, Hamming, Hanning, Blackman, and Kaiser windows are discussed in Chapter 7 and in Section 10.2. At this point, it is sufficient to observe that convolution of $W(e^{j\omega})$ with $X(e^{j\omega})$ will tend to smooth sharp peaks and discontinuities in $X(e^{j\omega})$. This is depicted by the continuous curve plotted in Figure 10.2(f).

The final operation in Figure 10.1 is the computation of the DFT. The DFT of the windowed sequence $v[n] = w[n]x[n]$ is

$$V[k] = \sum_{n=0}^{N-1} v[n]e^{-j(2\pi/N)kn}, \qquad k = 0, 1, \ldots, N-1, \tag{10.3}$$

where we assume that the window length L is less than or equal to the DFT length N. $V[k]$, the DFT of the finite-length sequence $v[n]$, corresponds to equally spaced samples of the DTFT of $v[n]$; i.e.,

$$V[k] = V(e^{j\omega})\big|_{\omega=2\pi k/N}. \tag{10.4}$$

Figure 10.2(f) also shows $V[k]$ as the samples of $V(e^{j\omega})$. Since the spacing between DFT frequencies is $2\pi/N$, and the relationship between the normalized discrete-time frequency variable and the continuous-time frequency variable is $\omega = \Omega T$, the DFT frequencies correspond to the continuous-time frequencies

$$\Omega_k = \frac{2\pi k}{NT}. \tag{10.5}$$

The use of this relationship between continuous-time frequencies and DFT frequencies is illustrated by Examples 10.1 and 10.2.

Example 10.1 Fourier Analysis Using the DFT

Consider a bandlimited continuous-time signal $x_c(t)$ such that $X_c(j\Omega) = 0$ for $|\Omega| \geq 2\pi(2500)$. We wish to use the system of Figure 10.1 to estimate the continuous-time spectrum $X_c(j\Omega)$. Assume that the antialiasing filter $H_{aa}(j\Omega)$ is ideal, and the sampling rate for the C/D converter is $1/T = 5000$ samples/s. If we want the DFT samples $V[k]$ to be equivalent to samples of $X_c(j\Omega)$ that are at most $2\pi(10)$ rad/s or 10 Hz apart, what is the minimum value that we should use for the DFT size N?

From Eq. (10.5), we see that adjacent samples in the DFT correspond to continuous-time frequencies separated by $2\pi/(NT)$. Therefore, we require that

$$\frac{2\pi}{NT} \leq 20\pi,$$

which implies that

$$N \geq 500$$

satisfies the condition. If we wish to use a radix-2 FFT algorithm to compute the DFT in Figure 10.1, we would choose $N = 512$ for an equivalent continuous-time frequency spacing of $\Delta\Omega = 2\pi(5000/512) = 2\pi(9.77)$ rad/s.

Example 10.2 Relationship Between DFT Values

Consider the problem posed in Example 10.1, in which $1/T = 5000$, $N = 512$, and $x_c(t)$ is real-valued and is sufficiently bandlimited to avoid aliasing with the given sampling rate. If it is determined that $V[11] = 2000(1 - j)$, what can be said about other values of $V[k]$ or about $X_c(j\Omega)$?

Referring to the symmetry properties of the DFT given in Table 8.2, $V[k] = V^*[((-k))_N]$, $k = 0, 1, \ldots, N - 1$, and consequently, $V[N - k] = V^*[k]$, so it follows in this case that

$$V[512 - 11] = V[501] = V^*[11] = 2000(1 - j).$$

We also know that the DFT sample $k = 11$ corresponds to the continuous-time frequency $\Omega_{11} = 2\pi(11)(5000)/512 = 2\pi(107.4)$, and similarly, $k = 501$ corresponds to the frequency $-2\pi(11)(5000)/512 = -2\pi(107.4)$. Although windowing smooths the spectrum, we can say that

$$X_c(j\Omega_{11}) = X_c(j2\pi(107.4)) \approx T \cdot V[11] = 0.4(1 + j).$$

Note that the factor T is required to compensate for the factor $1/T$ introduced by sampling, as in Eq. (10.1). We can again exploit symmetry to conclude that

$$X_c(-j\Omega_{11}) = X_c(-j2\pi(107.4)) \approx T \cdot V^*[11] = 0.4(1 - j).$$

Many commercial real-time spectrum analyzers are based on the principles embodied in Figures 10.1 and 10.2. It should be clear from the preceding discussion, however, that numerous factors affect the interpretation the DFT of a windowed segment of the sampled signal in terms of the continuous-time Fourier transform of the original input $s_c(t)$. To accommodate and mitigate the effects of these factors, care must be taken in filtering and sampling the input signal. Furthermore, to interpret the results correctly, the effects of the time-domain windowing and of the frequency-domain sampling inherent in the DFT must be clearly understood. For the remainder of the discussion, we will assume that the issues of antialiasing filtering and continuous-to-discrete-time conversion have been satisfactorily handled and are negligible. In the next section, we concentrate specifically on the effects of windowing and of the frequency-domain sampling imposed by the DFT. We choose sinusoidal signals as the specific class of examples to discuss, because sinusoids are perfectly bandlimited and they are easily computed. However, most of the issues raised by the examples apply more generally.

10.2 DFT ANALYSIS OF SINUSOIDAL SIGNALS

The DTFT of a sinusoidal signal $A\cos(\omega_0 n + \phi)$ (existing for all n) is a pair of impulses at $+\omega_0$ and $-\omega_0$ (repeating periodically with period 2π). In analyzing sinusoidal signals using the DFT, windowing and spectral (frequency-domain) sampling have important effects. As we will see in Section 10.2.1, windowing smears or broadens the impulses of the Fourier representation, thus, the exact frequency is less sharply defined. Windowing also reduces the ability to resolve sinusoidal signals that are close together in frequency. The spectral sampling inherent in the DFT has the effect of potentially giving a misleading or inaccurate picture of the true spectrum of the sinusoidal signal. This effect is discussed in Section 10.2.3.

10.2.1 The Effect of Windowing

Consider a continuous-time signal consisting of the sum of two sinusoidal components; i.e.,

$$s_c(t) = A_0\cos(\Omega_0 t + \theta_0) + A_1\cos(\Omega_1 t + \theta_1), \qquad -\infty < t < \infty. \tag{10.6}$$

Assuming ideal sampling with no aliasing and no quantization error, we obtain the discrete-time signal

$$x[n] = A_0\cos(\omega_0 n + \theta_0) + A_1\cos(\omega_1 n + \theta_1), \qquad -\infty < n < \infty, \tag{10.7}$$

where $\omega_0 = \Omega_0 T$ and $\omega_1 = \Omega_1 T$. The windowed sequence $v[n]$ in Figure 10.1 is then

$$v[n] = A_0 w[n]\cos(\omega_0 n + \theta_0) + A_1 w[n]\cos(\omega_1 n + \theta_1). \tag{10.8}$$

To obtain the DTFT of $v[n]$, we can expand Eq. (10.8) in terms of complex exponentials and use the frequency-shifting property of Eq. (2.158) in Section 2.9.2. Specifically, we rewrite $v[n]$ as

$$\begin{aligned} v[n] = {} & \frac{A_0}{2}w[n]e^{j\theta_0}e^{j\omega_0 n} + \frac{A_0}{2}w[n]e^{-j\theta_0}e^{-j\omega_0 n} \\ & + \frac{A_1}{2}w[n]e^{j\theta_1}e^{j\omega_1 n} + \frac{A_1}{2}w[n]e^{-j\theta_1}e^{-j\omega_1 n}, \end{aligned} \tag{10.9}$$

from which, with Eq. (2.158), it follows that the Fourier transform of the windowed sequence is

$$\begin{aligned} V(e^{j\omega}) = {} & \frac{A_0}{2}e^{j\theta_0}W(e^{j(\omega-\omega_0)}) + \frac{A_0}{2}e^{-j\theta_0}W(e^{j(\omega+\omega_0)}) \\ & + \frac{A_1}{2}e^{j\theta_1}W(e^{j(\omega-\omega_1)}) + \frac{A_1}{2}e^{-j\theta_1}W(e^{j(\omega+\omega_1)}). \end{aligned} \tag{10.10}$$

According to Eq. (10.10), the Fourier transform of the windowed signal consists of the Fourier transform of the window, shifted to the frequencies $\pm\omega_0$ and $\pm\omega_1$ and scaled by the complex amplitudes of the individual complex exponentials that make up the signal.

Example 10.3 Effect of Windowing on Fourier Analysis of Sinusoidal Signals

In this example, we consider the system of Figure 10.1 and, in particular, $W(e^{j\omega})$ and $V(e^{j\omega})$ for $s_c(t)$ of the form of Eq. (10.6), a sampling rate $1/T = 10$ kHz and a rectangular window $w[n]$ of length 64. The signal amplitude and phase parameters are $A_0 = 1$, $A_1 = 0.75$, and $\theta_0 = \theta_1 = 0$, respectively. To illustrate the essential features, we specifically display only the magnitudes of the Fourier transforms.

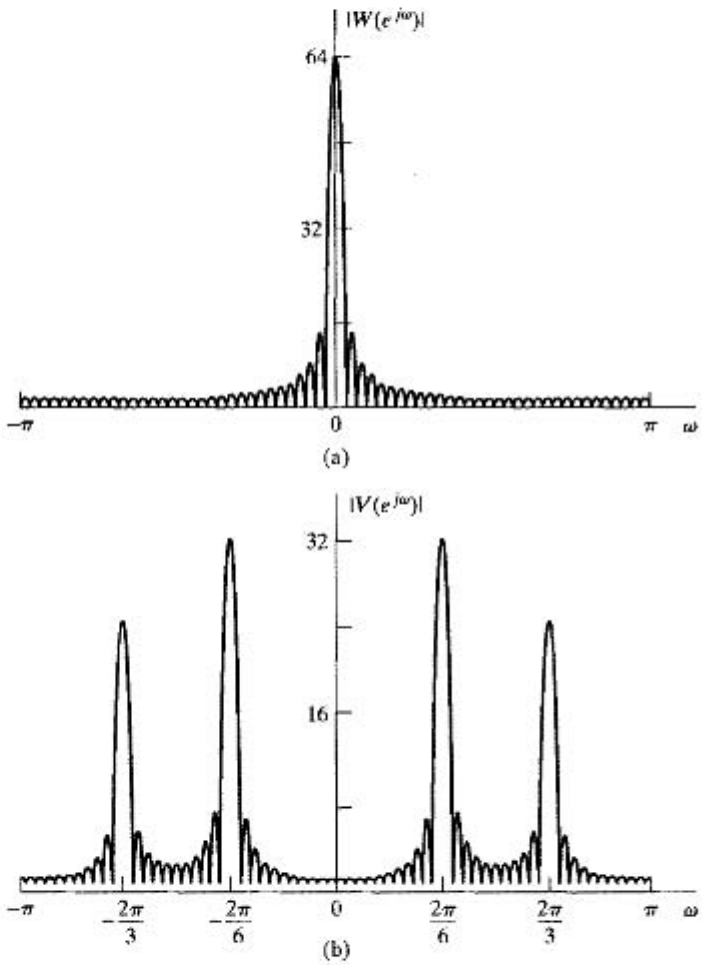


Figure 10.3 Illustration of Fourier analysis of windowed cosines with a rectangular window. (a) Fourier transform of window. (b)–(e) Fourier transform of windowed cosines as $\Omega_1 - \Omega_0$ becomes progressively smaller. (b) $\Omega_0 = (2\pi/6) \times 10^4$, $\Omega_1 = (2\pi/3) \times 10^4$.

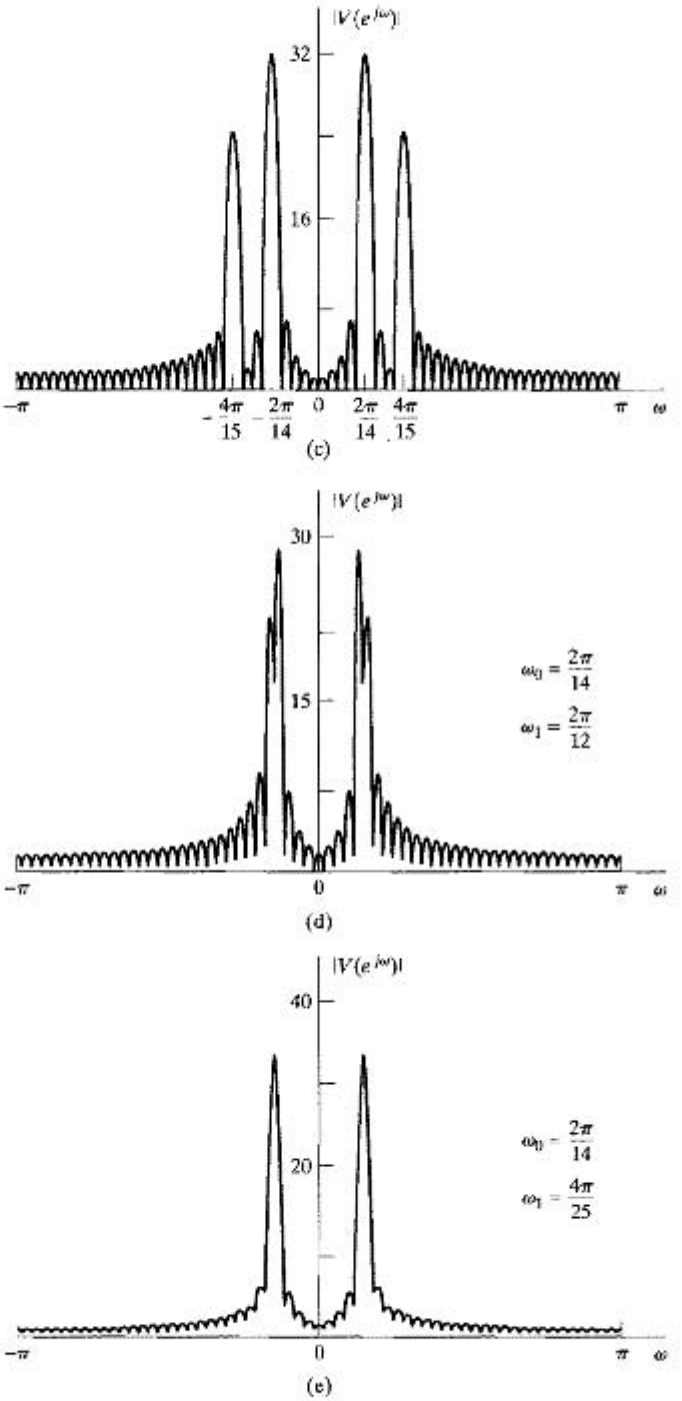


Figure 10.3 *(continued)* (c) $\Omega_0 = (2\pi/14) \times 10^4$, $\Omega_1 = (4\pi/15) \times 10^4$. (d) $\Omega_0 = (2\pi/14) \times 10^4$, $\Omega_1 = (2\pi/12) \times 10^4$. (e) $\Omega_0 = (2\pi/14) \times 10^4$, $\Omega_1 = (4\pi/25) \times 10^4$.

In Figure 10.3(a), we show $W(e^{j\omega})$, and in Figures 10.3(b), (c), (d), and (e), we show $|V(e^{j\omega})|$ for several choices of Ω_0 and Ω_1 in Eq. (10.6) or, equivalently, ω_0 and ω_1 in Eq. (10.7). In Figure 10.3(b), $\Omega_0 = (2\pi/6) \times 10^4$ and $\Omega_1 = (2\pi/3) \times 10^4$, or, equivalently, $\omega_0 = 2\pi/6$ and $\omega_1 = 2\pi/3$. In Figure 10.3(c)–(e), the frequencies become progressively closer. For the parameters in Figure 10.3(b), the frequency and amplitude of the individual components are evident. Specifically, Eq. (10.10) suggests that, with no overlap between the replicas of $W(e^{j\omega})$ at ω_0 and ω_1, there will be a peak of height $32A_0$ at ω_0 and $32A_1$ at ω_1, since $W(e^{j\omega})$ has a peak height of 64. In Figure 10.3(b), the two peaks are at approximately $\omega_0 = 2\pi/6$ and $\omega_1 = 2\pi/3$, with peak amplitudes in the correct ratio. In Figure 10.3(c), there is more overlap between the window replicas at ω_0 and ω_1, and while two distinct peaks are present, the amplitude of the spectrum at $\omega = \omega_0$ is affected by the amplitude of the sinusoidal signal at frequency ω_1 and vice versa. This interaction is called *leakage*: The component at one frequency leaks into the vicinity of another component owing to the spectral smearing introduced by the window. Figure 10.3(d) shows the case where the leakage is even greater. Notice how side lobes adding out of phase can *reduce* the heights of the peaks. In Figure 10.3(e), the overlap between the spectrum windows at ω_0 and ω_1 is so significant that the two peaks visible in (b)–(d) have merged into one. In other words, with this window, the two frequencies corresponding to Figure 10.3(e) will not be resolved in the spectrum.

10.2.2 Properties of the Windows

Reduced resolution and leakage are the two primary effects on the spectrum as a result of applying a window to the sinusoidal signal. The resolution is influenced primarily by the width of the main lobe of $W(e^{j\omega})$, whereas the degree of leakage depends on the relative amplitude of the main lobe to the side lobes of $W(e^{j\omega})$. In Chapter 7, in a filter design context, we showed that the width of the main lobe and the relative side-lobe amplitude depend primarily on the window length L and the shape (amount of tapering) of the window. The rectangular window, which has Fourier transform

$$W_r(e^{j\omega}) = \sum_{n=0}^{L-1} e^{-j\omega n} = e^{-j\omega(L-1)/2} \frac{\sin(\omega L/2)}{\sin(\omega/2)}, \tag{10.11}$$

has the narrowest main lobe for a given length ($\Delta_{ml} = 4\pi/L$), but it has the largest side lobes of all the commonly used windows. Other windows discussed in Chapter 7 include the Bartlett, Hann, and Hamming windows. The DTFTs of all these windows have main-lobe width $\Delta_{ml} = 8\pi/(L-1)$, which is approximately twice that of the rectangular window, but they have significantly smaller side-lobe amplitudes. The problem with all these windows is that there is no possibility of trade-off between main-lobe width and side-lobe amplitude, since the window length is the only variable parameter.

As we saw in Chapter 7, the Kaiser window is defined by

$$w_K[n] = \begin{cases} \dfrac{I_0[\beta(1 - [(n-\alpha)/\alpha]^2)^{1/2}]}{I_0(\beta)}, & 0 \le n \le L-1, \\ 0, & \text{otherwise}, \end{cases} \tag{10.12}$$

where $\alpha = (L-1)/2$ and $I_0(\cdot)$ is the zero$^{\text{th}}$-order modified Bessel function of the first kind. (Note that the notation of Eq. (10.12) differs slightly from that of Eq. (7.72) in

that L denotes the length of the window in Eq. (10.12), whereas the length of the filter design window in Eq. (7.72) is denoted $M + 1$.) We have already seen in the context of the filter design problem that this window has two parameters, β and L, which can be used to trade between main-lobe width and relative side-lobe amplitude. (Recall that the Kaiser window reduces to the rectangular window when $\beta = 0$.) The main-lobe width Δ_{ml} is defined as the symmetric distance between the central zero-crossings. The relative side-lobe level A_{sl} is defined as the ratio in dB of the amplitude of the main lobe to the amplitude of the largest side lobe. Figure 10.4, which is a duplicate of Figure 7.32, shows Fourier transforms of Kaiser windows for different lengths and different values of β. In designing a Kaiser window for spectrum analysis, we want to specify a desired value of A_{sl} and determine the required value of β. Figure 10.4(c) shows that the relative side-lobe amplitude is essentially independent of the window length and thus depends only on β. This was confirmed by Kaiser and Schafer (1980), who obtained the following least squares approximation to β as a function of A_{sl}:

$$\beta = \begin{cases} 0, & A_{\text{sl}} \leq 13.26, \\ 0.76609(A_{\text{sl}} - 13.26)^{0.4} + 0.09834(A_{\text{sl}} - 13.26), & 13.26 < A_{\text{sl}} \leq 60, \\ 0.12438(A_{\text{sl}} + 6.3), & 60 < A_{\text{sl}} \leq 120. \end{cases} \tag{10.13}$$

Using values of β from Eq. (10.13) gives windows with actual side-lobe values that differ by less than 0.36 from the value of A_{sl} used in Eq. (10.13) for the entire range of $13.26 < A_{\text{sl}} < 120$. (Note that the value 13.26 is the relative side-lobe amplitude of the rectangular window, to which the Kaiser window reduces for $\beta = 0$.)

Figure 10.4(c) also shows that the main-lobe width is inversely proportional to the length of the window. The trade-off between main-lobe width, relative side-lobe amplitude, and window length is displayed by the approximate relationship

$$L \simeq \frac{24\pi(A_{\text{sl}} + 12)}{155\Delta_{\text{ml}}} + 1, \tag{10.14}$$

which was also given by Kaiser and Schafer (1980).

Equations (10.12), (10.13), and (10.14) are the necessary equations for determining a Kaiser window with desired values of main-lobe width and relative side-lobe amplitude. To design a window for prescribed values of A_{sl} and Δ_{ml} requires simply the computation of β from Eq. (10.13), the computation of L from Eq. (10.14), and the computation of the window using Eq. (10.12). Many of the remaining examples of this chapter use the Kaiser window. Other spectrum analysis windows are considered by Harris (1978).

10.2.3 The Effect of Spectral Sampling

As mentioned previously, the DFT of the windowed sequence $v[n]$ provides samples of $V(e^{j\omega})$ at the N equally spaced discrete-time frequencies $\omega_k = 2\pi k/N$, $k = 0, 1, \ldots, N - 1$. These are equivalent to the continuous-time frequencies $\Omega_k = (2\pi k)/(NT)$, for $k = 0, 1, \ldots, N/2$ (assuming that N is even). The indices $k = N/2 + 1, \ldots, N - 1$ correspond to the negative continuous-time frequencies $-2\pi(N - k)/(NT)$. Spectral sampling, as imposed by the DFT, can sometimes produce misleading results. This effect is best illustrated by example.

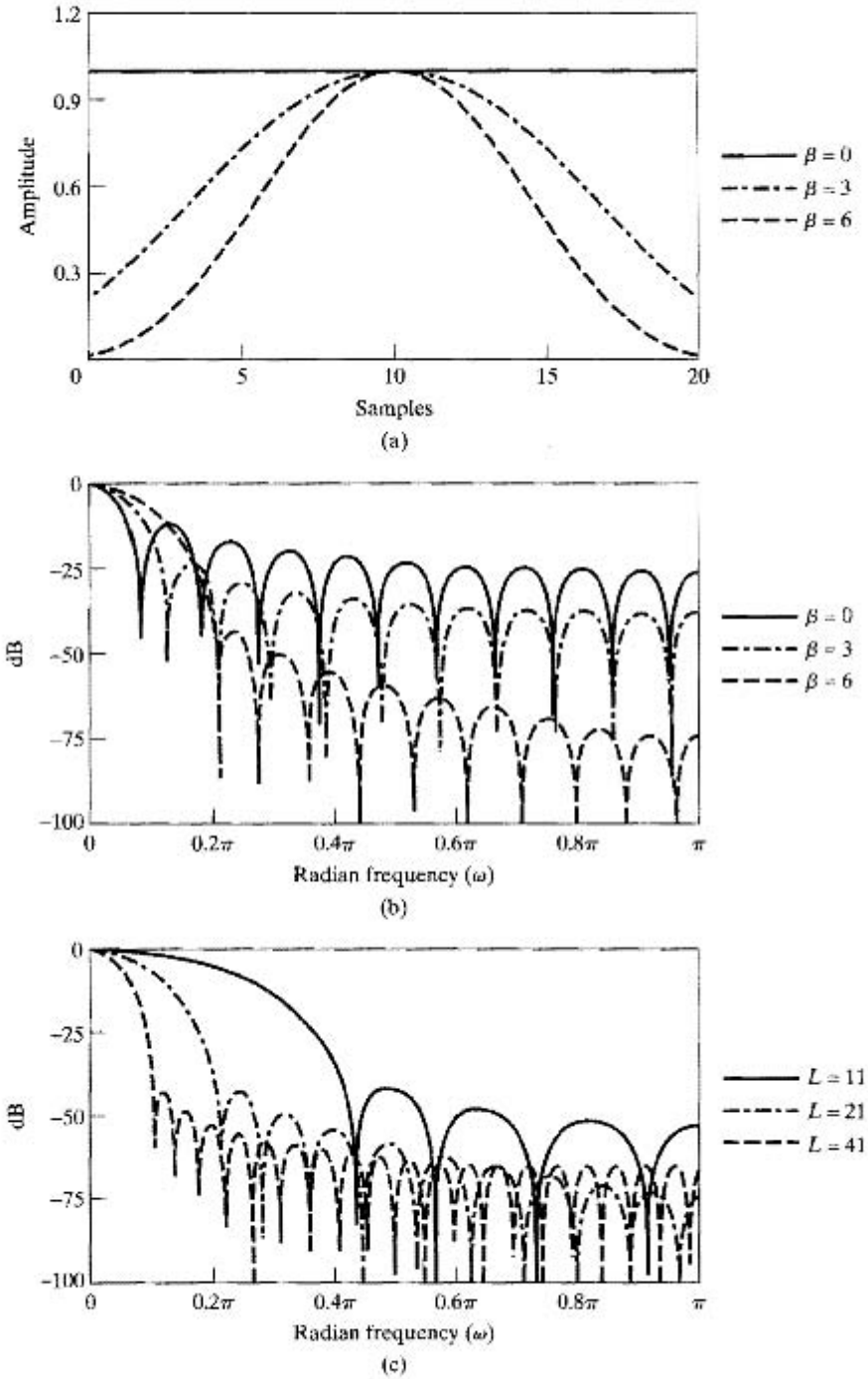


Figure 10.4 (a) Kaiser windows for $\beta = 0$, 3, and 6 and $L = 21$. (b) Fourier transform corresponding to windows in (a). (c) Fourier transforms of Kaiser windows with $\beta = 6$ and $L = 11$, 21, and 41.

Example 10.4 Illustration of the Effect of Spectral Sampling

Consider the same parameters as in Figure 10.3(c) in Example 10.3, i.e., $A_0 = 1$, $A_1 = 0.75, \omega_0 = 2\pi/14$, $\omega_1 = 4\pi/15$, and $\theta_1 = \theta_2 = 0$ in Eq. (10.8). $w[n]$ is a rectangular window of length 64. Then

$$v[n] = \begin{cases} \cos\left(\frac{2\pi}{14}n\right) + 0.75\cos\left(\frac{4\pi}{15}n\right), & 0 \le n \le 63, \\ 0, & \text{otherwise.} \end{cases} \tag{10.15}$$

Figure 10.5(a) shows the windowed sequence $v[n]$. Figures 10.5(b), (c), (d), and (e) show the corresponding real part, imaginary part, magnitude, and phase, respectively, of the DFT of length $N = 64$. Observe that since $x[n]$ is real, $X[N-k] = X^*[k]$ and

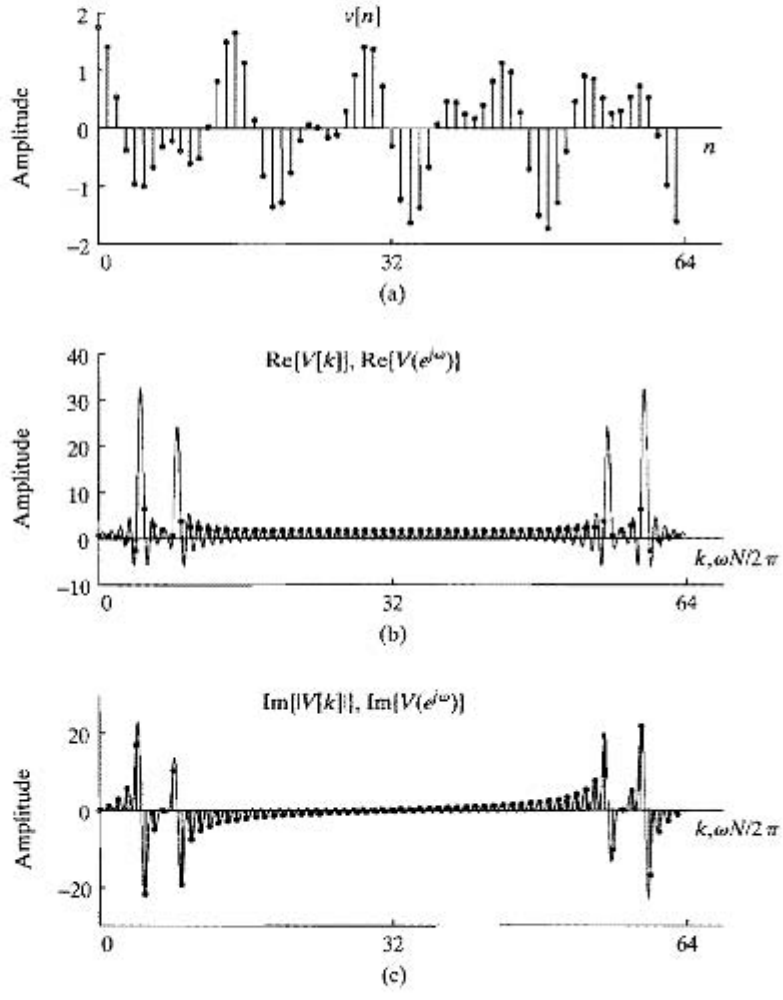


Figure 10.5 Cosine sequence and DFT with a rectangular window for $N = 64$. (a) Windowed signal. (b) Real part of DFT. (c) Imaginary part of DFT. Note that the DTFT is superimposed as the light continuous line.

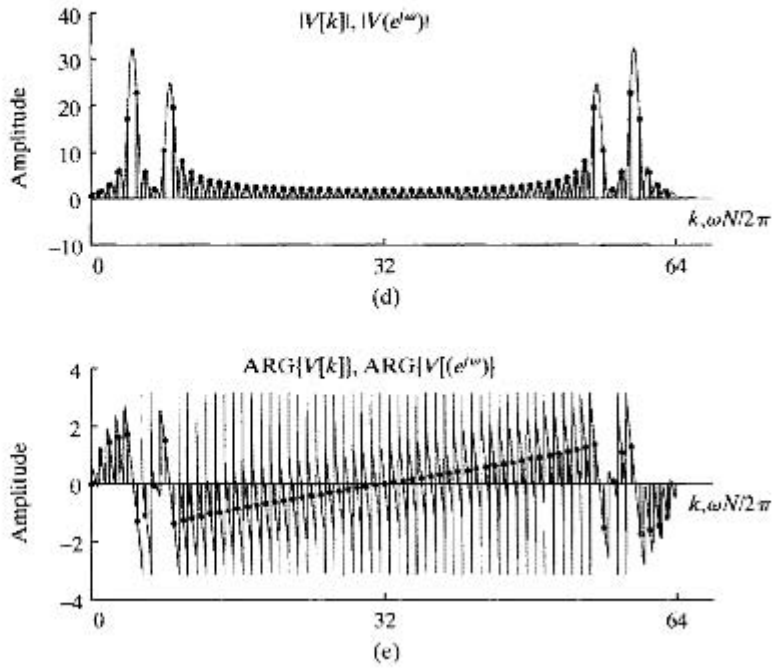


Figure 10.5 (*continued*) (d) Magnitude of DFT. (e) Phase of DFT.

$X(e^{j(2\pi-\omega)}) = X^*(e^{j\omega})$; i.e., the real part and the magnitude are even functions and the imaginary part and phase are odd functions of k and ω.

In Figures 10.5(b)–(e), the horizontal (frequency) axis is labeled in terms of the DFT index or frequency sample number k. The value $k = 32$ corresponds to $\omega = \pi$ or, equivalently, $\Omega = \pi/T$. As is the usual convention in displaying the DFT of a time sequence, we display the DFT values in the range from $k = 0$ to $k = N - 1$, corresponding to displaying samples of the DTFT in the frequency range 0 to 2π. Because of the inherent periodicity of the DTFT, the first half of this range corresponds to the positive continuous-time frequencies, i.e., Ω between zero and π/T, and the second half of the range to the negative frequencies, i.e., Ω between $-\pi/T$ and zero. Note the even periodic symmetry of the real part and the magnitude and the odd periodic symmetry of the imaginary part and the phase.

Recall that the DFT $V[k]$ is a sampled version of the DTFT $V(e^{j\omega})$. Superimposed on each DFT with a light gray line in Figures 10.5(b)–(e) is the corresponding DTFT, i.e., $\mathcal{R}e\{V(e^{j\omega})\}$, $\mathcal{I}m\{V(e^{j\omega})\}$, $|V(e^{j\omega})|$, and $\text{ARG}\{V(e^{j\omega})\}$ respectively. The frequency scale for these functions is the specially defined normalized scale denoted $\omega N/(2\pi)$; i.e., N on the DFT index scale corresponds to $\omega = 2\pi$ on the conventional frequency scale of the DTFT. We also follow this convention of superimposing the DTFT in Figures 10.6, 10.7, 10.8, and 10.9.

The magnitude of the DFT in Figure 10.5(d) corresponds to samples of $|V(e^{j\omega})|$ (the light continuous line), which shows the expected concentration around $\omega_1 = 2\pi/7.5$ and $\omega_0 = 2\pi/14$, the frequencies of the two sinusoidal components of the input. Specifically, the frequency $\omega_1 = 4\pi/15 = 2\pi(8.533\ldots)/64$ lies between the DFT samples corresponding to $k = 8$ and $k = 9$. Likewise, the frequency $\omega_0 = 2\pi/14 = 2\pi(4.5714\ldots)/64$ lies between the DFT samples corresponding to $k = 4$ and $k = 5$. Note that the frequency locations of the peaks of the gray curve in Figure 10.5(d)

are between spectrum samples obtained from the DFT. In general, the locations of peaks in the DFT values do not necessarily coincide with the exact frequency locations of the peaks in the DTFT, since the true spectrum peaks can lie between spectrum samples. Correspondingly, as evidenced in Figure 10.5(d), the relative amplitudes of peaks in the DFT will not necessarily reflect the relative amplitudes of the spectrum peaks of $|V(e^{j\omega})|$.

Example 10.5 Signal Frequencies Matching DFT Frequencies Exactly

Consider the sequence

$$v[n] = \begin{cases} \cos\left(\frac{2\pi}{16}n\right) + 0.75\cos\left(\frac{2\pi}{8}n\right), & 0 \le n \le 63, \\ 0, & \text{otherwise}, \end{cases} \tag{10.16}$$

as shown in Figure 10.6(a). Again, a rectangular window is used with $N = L = 64$. This is very similar to the previous example, except that in this case, the frequencies of the cosines coincide exactly with two of the DFT frequencies. Specifically, the frequency $\omega_1 = 2\pi/8 = 2\pi 8/64$ corresponds exactly to the DFT sample $k = 8$ and the frequency $\omega_0 = 2\pi/16 = 2\pi 4/64$ to the DFT sample $k = 4$.

The magnitude of the 64-point DFT of $v[n]$ for this example is shown in Figure 10.6(b) and corresponds to samples of $|V(e^{j\omega})|$ (which again is superimposed

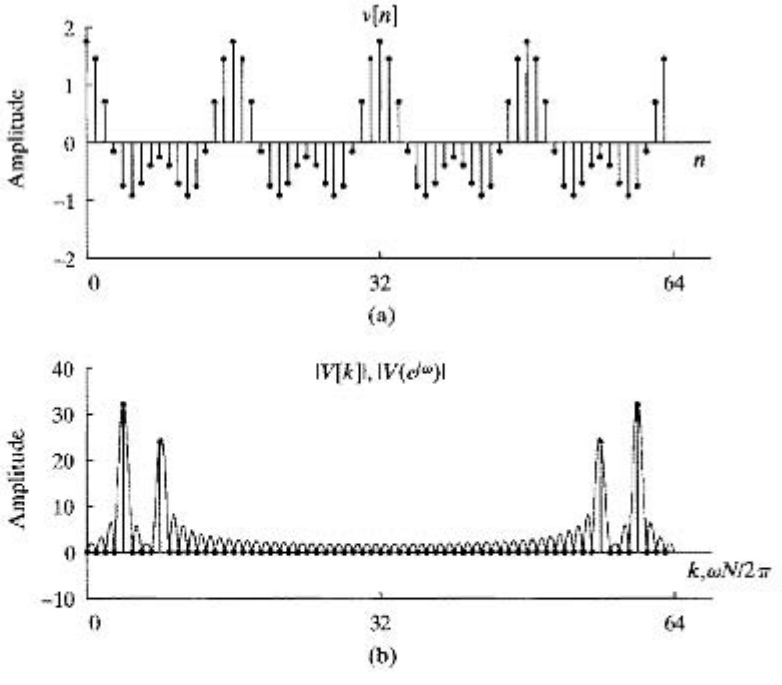


Figure 10.6 Discrete Fourier analysis of the sum of two sinusoids for a case in which the Fourier transform is zero at all DFT frequencies except those corresponding to the frequencies of the two sinusoidal components. (a) Windowed signal. (b) Magnitude of DFT. Note that ($|V(e^{j\omega})|$) is superimposed as the light continuous line.

with a light line) at a frequency spacing of $2\pi/64$. Although the signal parameters in Example 10.4 are very similar, the appearance of the DFT is for this example and strikingly different. In particular, for this example, the DFT has two strong spectral lines at the samples corresponding to the frequencies of the two sinusoidal components in the signal and no frequency content at the other DFT values. In fact, this clean appearance of the DFT in Figure 10.6(b) is largely an illusion resulting from the sampling of the spectrum. Comparing Figures 10.6(b) and (c), we can see that the reason for the clean appearance of Figure 10.6(b) is that for this choice of parameters, the Fourier transform is exactly zero at the frequencies that are sampled by the DFT, except those corresponding to $k = 4, 8, 64 - 8$, and $64 - 4$. Although the signal of Figure 10.6(a) has significant content at almost all frequencies, as evidenced by the gray curve in Figure 10.6(b), we do not see this in the DFT, because of the sampling of the spectrum. Another way of viewing this is to note that the 64-point rectangular window selects exactly an integer number of periods of the two sinusoidal components in Eq. (10.16). The 64-point DFT then corresponds to the DFS of this signal replicated with period 64. This replicated signal will have only four nonzero DFS coefficients corresponding to the two sinusoidal components on Eq. (10.16). This is an example of how the inherent assumption of periodicity gives a correct answer to a different problem. We are interested in the finite-length case and the results are quite misleading from that point of view.

To illustrate this point further, we can extend $v[n]$ in Eq. (10.16) by zero-padding to obtain a 128-point sequence. The corresponding 128-point DFT is shown in Figure 10.7. With this finer sampling of the spectrum, the presence of significant content at other frequencies becomes apparent. In this case, the windowed signal is *not* naturally periodic with period 128.

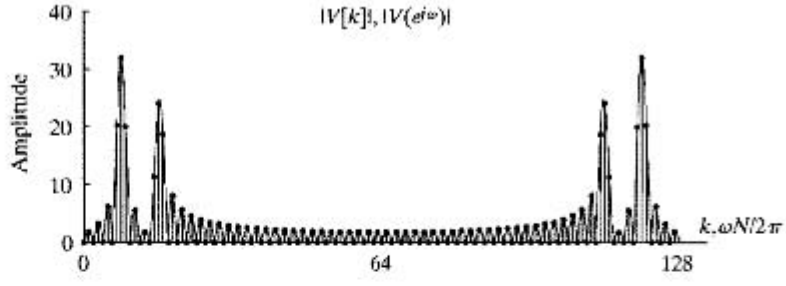


Figure 10.7 DFT of the signal as in Figure 10.6(a), but with twice the number of frequency samples used in Figure 10.6(b).

In Figures 10.5, 10.6, and 10.7, the windows were rectangular. In the next set of examples, we illustrate the effect of different choices for the window.

Example 10.6 DFT Analysis of Sinusoidal Signals Using a Kaiser Window

In this example we return to the frequency, amplitude, and phase parameters of Example 10.4, but now with a Kaiser window applied, so that

$$v[n] = w_K[n]\cos\left(\frac{2\pi}{14}n\right) + 0.75w_K[n]\cos\left(\frac{4\pi}{15}n\right). \tag{10.17}$$

where $w_K[n]$ is the Kaiser window as given by Eq. (10.12). We will select the Kaiser window parameter β to be equal to 5.48, which, according to Eq. (10.13), results in a window for which the relative side-lobe amplitude is $A_s = 40$ dB. Figure 10.8(a) shows

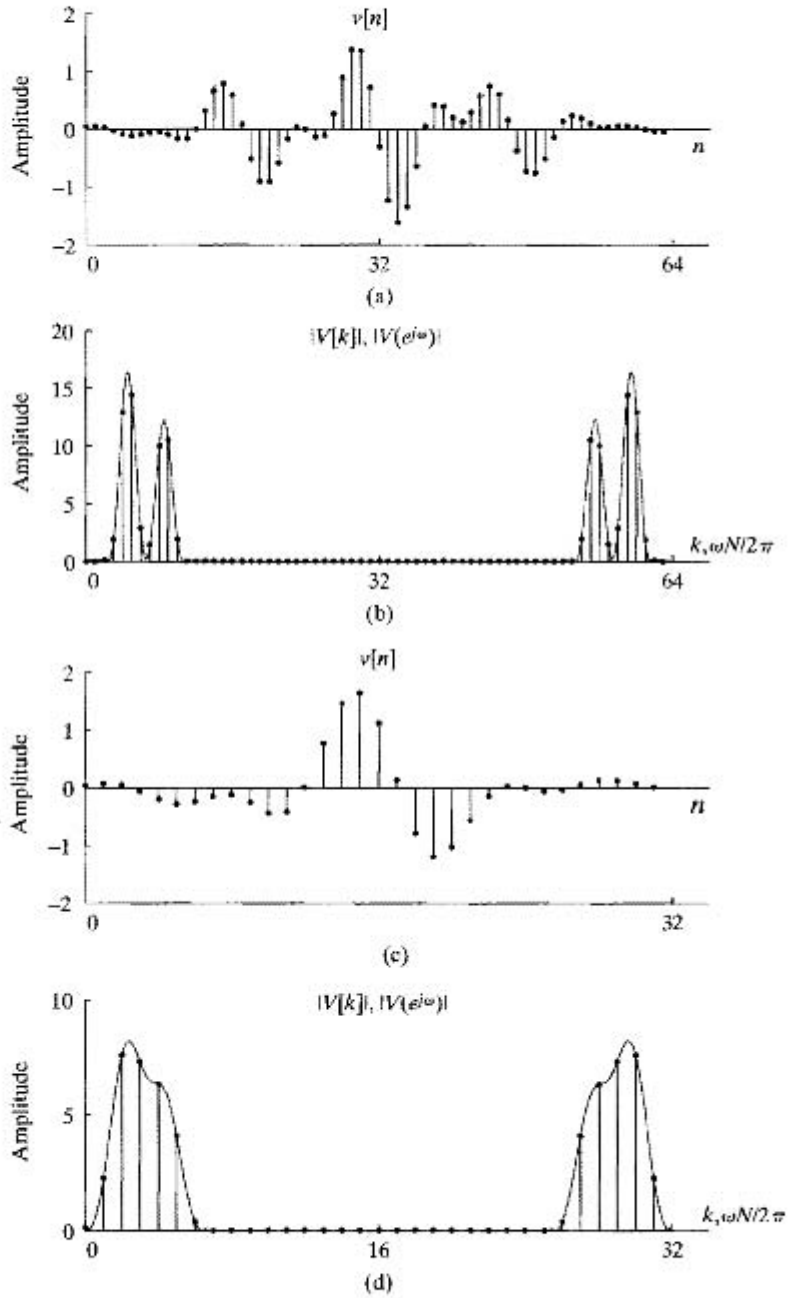


Figure 10.8 Discrete Fourier analysis with Kaiser window. (a) Windowed sequence for $L = 64$. (b) Magnitude of DFT for $L = 64$. (c) Windowed sequence for $L = 32$. (d) Magnitude of DFT for $L = 32$.

the windowed sequence $v[n]$ for a window length of $L = 64$, and Figure 10.8(b) shows the magnitude of the corresponding DFT. From Eq. (10.17), we see that the difference between the two frequencies is $\omega_1 - \omega_0 = 2\pi/7.5 - 2\pi/14 = 0.389$. From Eq. (10.14), it follows that the width of the main lobe of the Fourier transform of the Kaiser window with $L = 64$ and $\beta = 5.48$ is $\Delta_{\text{ml}} = 0.401$. Thus, the main lobes of the two replicas of $W_K(e^{j\omega})$ centered at ω_0 and ω_1 will just slightly overlap in the frequency interval between the two frequencies. This is evident in Figure 10.8(b), where we see that the two frequency components are clearly resolved.

Figure 10.8(c) shows the same signal, multiplied by a Kaiser window with $L = 32$ and $\beta = 5.48$. Since the window is half as long, we expect the width of the main lobe of the Fourier transform of the window to double, and Figure 10.8(d) confirms this. Specifically, Eqs. (10.13) and (10.14) confirm that for $L = 32$ and $\beta = 5.48$, the main-lobe width is $\Delta_{\text{ml}} = 0.815$. Now, the main lobes of the two copies of the Fourier transform of the window overlap throughout the region between the two cosine frequencies, and we do not see two distinct peaks.

In all the previous examples except in Figure 10.7, the DFT length N was equal to the window length L. In Figure 10.7, zero-padding was applied to the windowed sequence before computing the DFT to obtain the Fourier transform on a more finely divided set of frequencies. However, we must realize that this zero-padding will not improve the ability to resolve close frequencies, which depends on the length and shape of the window. This is illustrated by the next example.

Example 10.7 DFT Analysis with 32-point Kaiser Window and Zero-Padding

In this example, we repeat Example 10.6 using the Kaiser window with $L = 32$ and $\beta = 5.48$, and with the DFT length varying. Figure 10.9(a) shows the DFT magnitude for $N = L = 32$ as in Figure 10.8(d), and Figures 10.9(b) and (c) show the DFT magnitude again with window length $L = 32$, but with DFT lengths $N = 64$ and $N = 128$, respectively. As with Example 10.5, this zero-padding of the 32-point sequence results in finer spectral sampling of the DTFT. As shown by the light continuous curve, the underlying envelope of each DFT magnitude in Figure 10.9 is the same. Consequently, increasing the DFT size by zero-padding does not change the ability to resolve the two sinusoidal frequency components, but it does change the spacing of the frequency samples. If N were increased beyond 128, the dots denoting the DFT sample

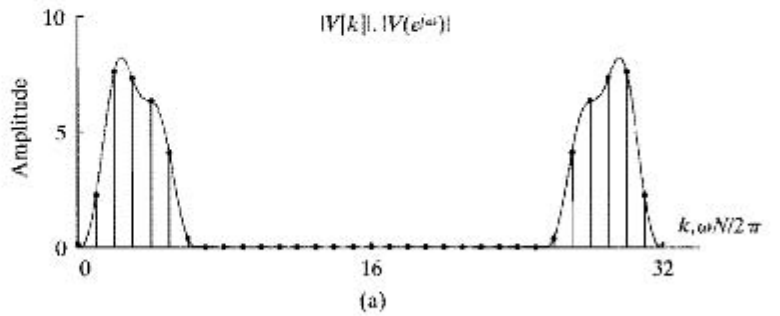


Figure 10.9 Illustration of effect of DFT length for Kaiser window of length $L = 32$. (a) Magnitude of DFT for $N = 32$.

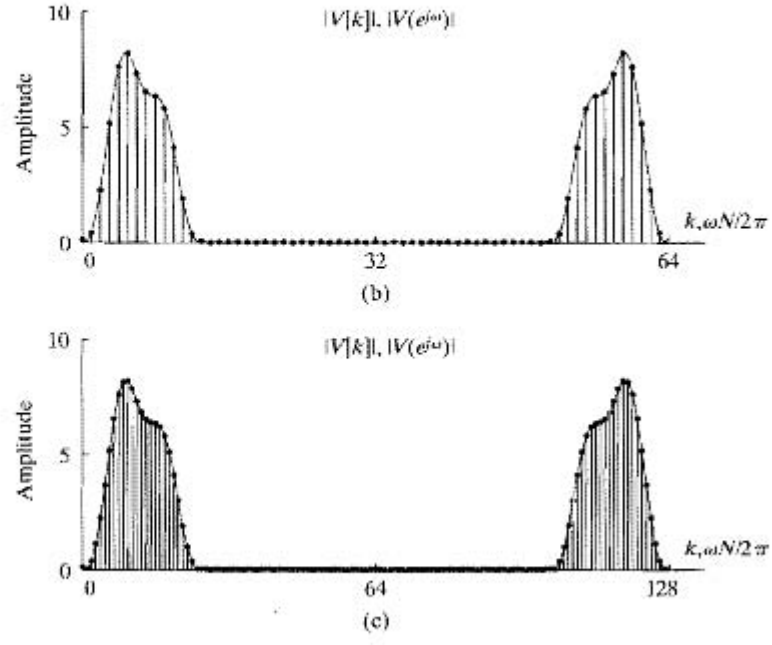


Figure 10.9 (*continued*) (b) Magnitude of DFT for $N = 64$. (c) Magnitude of DFT for $N = 128$.

values would tend to merge together and become indistinct. Consequently, DFT values are often plotted by connecting consecutive points by straight-line segments without indicating each individual point. For example, in Figures 10.5 through 10.8, we have shown a light continuous line as the DTFT $|V(e^{j\omega})|$ of the finite-length sequence $v[n]$. In fact, this curve is a plot of the DFT of the sequence after zero-padding to $N = 2048$. In these examples, this sampling of the DTFT is sufficiently dense so as to be indistinguishable from the function of the continuous variable ω.

For a complete representation of a sequence of length L, the L-point DFT is sufficient, since the original sequence can be recovered exactly from it. However, as we saw in the preceding examples, simple examination of the L-point DFT can result in misleading interpretations. For this reason, it is common to apply zero-padding, so that the spectrum is sufficiently oversampled and important features are therefore readily apparent. With a high degree of time-domain zero-padding or frequency-domain oversampling, simple interpolation (e.g., linear interpolation) between the DFT values provides a reasonably accurate picture of the Fourier spectrum, which can then be used, for example, to estimate the locations and amplitudes of spectrum peaks. This is illustrated in the following example.

Example 10.8 Oversampling and Linear Interpolation for Frequency Estimation

Figure 10.10 shows how a 2048-point DFT can be used to obtain a finely spaced evaluation of the Fourier transform of a windowed signal and how increasing the window width improves the ability to resolve closely spaced sinusoidal components. The signal

of Example 10.6 having frequencies $2\pi/14$ and $4\pi/15$ was windowed with Kaiser windows of lengths $L = 32, 42, 54$, and 64 with $\beta = 5.48$. First, note that in all cases, the 2048-point DFT gives a smooth result when the points are connected by straight lines. In Figure 10.10(a), where $L = 32$, the two sinusoidal components are not resolved, and, of course, increasing the DFT length will only result in a smoother curve. As the window length increases from $L = 32$ to $L = 42$, however, we see improvement in our ability to distinguish the two frequencies and the approximate relative amplitudes of each sinusoidal component. The dashed lines in all the figures indicate the DFT indices $k_0 = 146 \approx 2048/14$ and $k_1 = 273 \approx 4096/15$, which correspond to the nearest DFT frequencies ($N = 2048$) for the cosine components. Note that the 2048-point DFT in Figure 10.10(c) would be much more effective for precisely locating the peak of the windowed Fourier transform than the coarsely sampled DFT in Figure 10.8(b), which is also computed with a 64-point Kaiser window. Note also that the amplitudes of the two peaks in Figure 10.10 are very close to being in the correct ratio of 0.75 to 1.

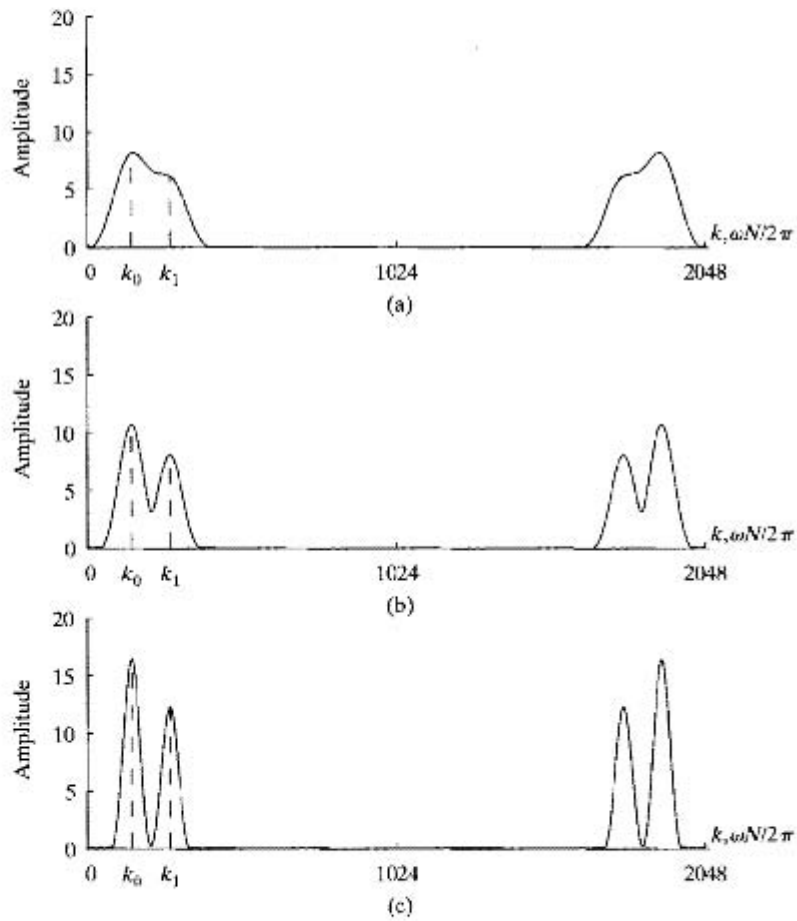


Figure 10.10 Illustration of the computation of the DFT for $N \gg L$ with linear interpolation to create a smooth curve: (a) $N = 1024$, $L = 32$. (b) $N = 1024$, $L = 42$. (c) $N = 1024$, $L = 64$. (The values $k_0 = 146 \approx 2048/14$ and $k_1 = 273 \approx 4096/15$ are the closest DFT frequencies to $\omega_0 = 2\pi/14$ and $\omega_1 = 4\pi/15$ when the DFT length is $N = 2048$.)

10.3 THE TIME-DEPENDENT FOURIER TRANSFORM

In Section 10.2, we illustrated the use of the DFT for obtaining a frequency-domain representation of a signal composed of sinusoidal components. In that discussion, we assumed that the frequencies of the cosines did not change with time, so that no matter how long the window, the signal properties (amplitudes, frequencies, and phases) would be the same from the beginning to the end of the window. Long windows give better frequency resolution, but in practical applications of sinusoidal signal models, the signal properties (e.g., amplitude, frequency) often change with time. For example, nonstationary signal models of this type are required to describe radar, sonar, speech, and data communication signals. This conflicts with the use of long analysis windows. A single DFT estimate is not sufficient to describe such signals, and as a result, we are led to the concept of the *time-dependent Fourier transform,* also referred to as the short-time Fourier transform.[1]

We define the time-dependent Fourier transform of a signal $x[n]$ as

$$X[n, \lambda] = \sum_{m=-\infty}^{\infty} x[n+m]w[m]e^{-j\lambda m}, \tag{10.18}$$

where $w[n]$ is a window sequence. In the time-dependent Fourier representation, the one-dimensional sequence $x[n]$, a function of a single discrete variable, is converted into a two-dimensional function of the time variable n, which is discrete, and the frequency variable λ, which is continuous.[2] Note that the time-dependent Fourier transform is periodic in λ with period 2π; therefore, we need consider only values of λ for $0 \leq \lambda < 2\pi$ or any other interval of length 2π.

Equation (10.18) can be interpreted as the DTFT of the shifted signal $x[n+m]$, as viewed through the window $w[m]$. The window has a stationary origin, and as n changes, the signal slides past the window, so that at each value of n, a different portion of the signal is extracted by the window for Fourier analysis. As an illustration, consider the following example.

Example 10.9 Time-Dependent Fourier Transform of a Linear Chirp Signal

A continuous-time linear chirp signal is defined as

$$x_c(t) = \cos(\theta(t)) = \cos(A_0 t^2), \tag{10.19}$$

[1] Further discussion of the time-dependent Fourier transform can be found in a variety of references, including Allen and Rabiner (1977), Rabiner and Schafer (1978), Crochiere and Rabiner (1983) and Quatieri (2002).

[2] We denote the frequency variable of the time-dependent Fourier transform by λ to maintain a distinction from the frequency variable of the conventional DTFT, which we always denote by ω. We use the mixed bracket–parenthesis notation $X[n, \lambda)$ as a reminder that n is a discrete variable, and λ is a continuous variable.

where A_0 has units of radians/s^2. (Such signals are called chirps because, in the auditory frequency range, short pulses sound like bird chirps.) The signal $x_c(t)$ in Eq. (10.19) is a member of the more general class of frequency modulation (FM) signals for which the *instantaneous frequency* is defined as the time derivative of the cosine argument $\theta(t)$. Therefore, in this case, the instantaneous frequency is

$$\Omega_i(t) = \frac{d\theta(t)}{dt} = \frac{d}{dt}\left(A_0 t^2\right) = 2A_0 t, \tag{10.20}$$

which varies in proportion to time; hence, the designation as a *linear* chirp signal. If we sample $x_c(t)$, we obtain the discrete-time linear chirp signal[3]

$$x[n] = x_c(nT) = \cos(A_0 T^2 n^2) = \cos(\alpha_0 n^2), \tag{10.21}$$

where $\alpha_0 = A_0 T^2$ has units of radians. The instantaneous frequency of the sampled chirp signal is a frequency-normalized, sampled version of the instantaneous frequency of the continuous-time signal; i.e.,

$$\omega_i[n] = \Omega_i(nT) \cdot T = 2A_0 T^2 n = 2\alpha_0 n, \tag{10.22}$$

which displays the same proportional increase with sample index n, with α_0 controlling the rate of increase. Figure 10.11 shows two 1201-sample segments of the sampled chirp signal in Eq. (10.21) with $\alpha_0 = 15\pi \times 10^{-6}$. (The samples are connected by straight lines for plotting.) Observe that over a short interval, the signal looks sinusoidal, but the spacing between peaks becomes smaller and smaller as time progresses, indicating increasing frequency with time.

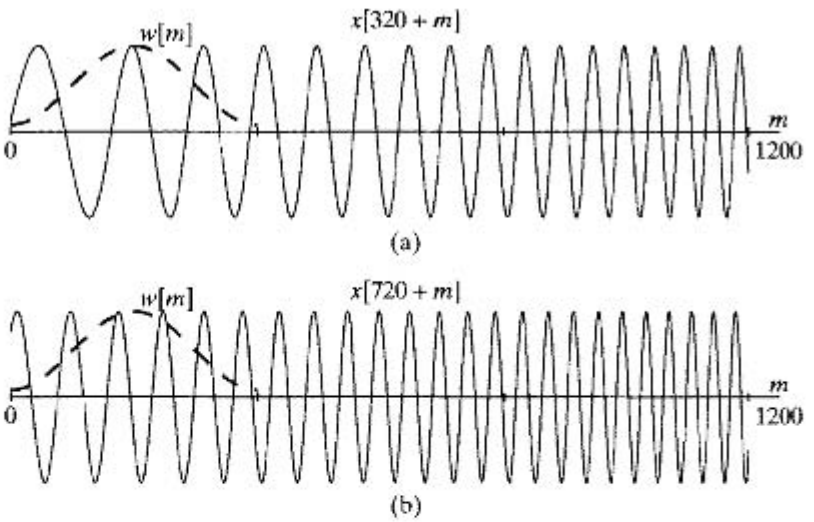


Figure 10.11 Two segments of the linear chirp signal $x[n] = \cos(\alpha_0 n^2)$ for $\alpha_0 = 15\pi \times 10^{-6}$ with a 400-sample Hamming window superimposed. (a) $X[n, \lambda)$ at $n = 320$ would be the DTFT of the top trace multiplied by the window. (b) $X[720, \lambda)$ would be the DTFT of the bottom trace multiplied by the window.

[3]We have seen discrete-time linear complex exponential chirp signals in Chapter 9 in the context of the chirp transform algorithm.

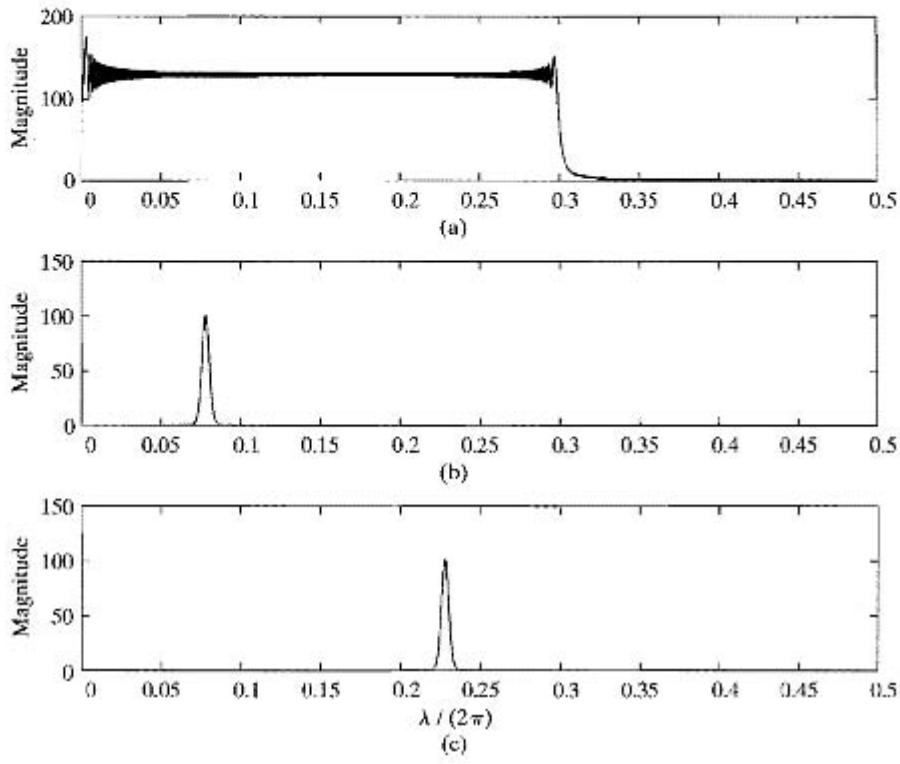


Figure 10.12 DTFTs of segments of a linear chirp signal: (a) DTFT of 20,000 samples of the signal $x[n] = \cos(\alpha_0 n^2)$. (b) DTFT of $x[5000 + m]w[m]$ where $w[m]$ is a Hamming window of length $L = 401$; i.e., $X[5000, \lambda)$. (c) DTFT of $x[15,000 + m]w[m]$ where $w[m]$ is a Hamming window of length $L = 401$; i.e., $X[15,000, \lambda)$.

The relationship of the shifted signal to the window in time-dependent Fourier analysis is also illustrated in Figure 10.11. Typically, $w[m]$ in Eq. (10.18) has finite length around $m = 0$, so that $X[n, \lambda)$ displays the frequency characteristics of the signal around time n. Figure 10.11(a) shows $x[320 + m]$ as a function of m for $0 \leq m \leq 1200$ together with a Hamming window $w[m]$ of length $L = 401$ samples. The time-dependent transform at time $n = 320$ is the DTFT of $w[m]x[320 + m]$. Similarly, Figure 10.11(b) shows the window and a later segment of the chirp signal beginning at sample $n = 720$.

Figure 10.12 illustrates the importance of the window in discrete-time Fourier analysis of time-varying signals. Figure 10.12(a) shows the DTFT of 20,000 samples (with a rectangular window) of the discrete-time chirp. Over this interval, the normalized instantaneous frequency of the chirp,

$$f_i[n] = \omega_i[n]/(2\pi) = 2\alpha_0 n/(2\pi),$$

goes from 0 to $0.00003\pi(20{,}000)/(2\pi) = 0.3$. This variation of the instantaneous frequency forces the DTFT representation, which involves only fixed frequencies acting over all n, to include all frequencies in that range and beyond as is evident in Figure 10.12(a). Thus, the DTFT of a long segment of the signal shows only that the

signal has a wide bandwidth in the conventional DTFT sense. On the other hand, Figures 10.12(b) and (c) show DTFTs using a 401 sample Hamming window for segments of the chirp waveform at $n = 5000$ and 15,000, respectively. Thus, Figures 10.12(b) (c) are plots [as functions of $\lambda/(2\pi)$] of the time-dependent Fourier transform values $|X[5000, \lambda)|$ and $|X[15,000, \lambda)|$, respectively. Since the window length $L = 401$ is such that the signal does not change frequency very much across the window interval, the time-dependent Fourier transform tracks the frequency variation very well. Note that at samples 5000 and 15,000, we would expect a peak in the time-dependent transform at $\lambda/(2\pi) = 0.00003\pi(5000)/(2\pi) = 0.075$ and $\lambda/(2\pi) = 0.00003\pi(15,000)/(2\pi) = 0.225$, respectively. This is confirmed by examination of Figures 10.12(b) and (c).

Example 10.10 Plotting $X[n,\lambda]$: The Spectrogram

In Figure 10.13, we show a display as a function of both time index n and frequency $\lambda/(2\pi)$ of the magnitude of the time-dependent Fourier transform, $|Y[n, \lambda)|$, for the signal

$$y[n] = \begin{cases} 0 & n < 0 \\ \cos(\alpha_0 n^2) & 0 \le n \le 20{,}000 \\ \cos(0.2\pi n) & 20{,}000 < n \le 25{,}000 \\ \cos(0.2\pi n) + \cos(0.23\pi n) & 25{,}000 < n. \end{cases} \tag{10.23}$$

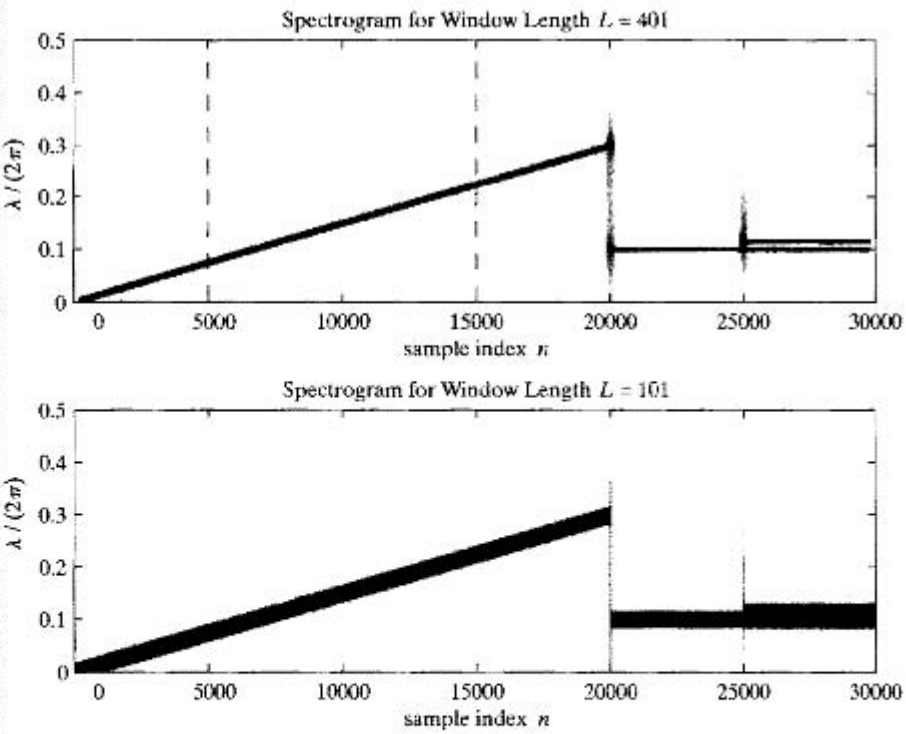


Figure 10.13 The magnitude of the time-dependent Fourier transform of $y[n]$ in Eq. (10.23): (a) Using a Hamming window of length $L = 401$. (b) Using a Hamming window of length $L = 101$.

Note that the signal $y[n]$ is equal to $x[n]$ in Eq. (10.21) in Example 10.9 for $0 \leq n \leq 20{,}000$, and then it abruptly changes to cosine components with fixed frequencies for $n > 20{,}000$. This signal was designed to make several important points about time-dependent Fourier analysis. First, consider Figure 10.13(a), which shows the time-dependent Fourier transform of $y[n]$ over the interval $0 \leq n \leq 30{,}000$ with a Hamming window of length $L = 401$. This display, which shows $20\log_{10}|Y[n, \lambda)|$ as a function of $\lambda/2\pi$ in the vertical dimension, and the time index n in the horizontal dimension is called a *spectrogram*. The value $20\log_{10}|Y[n, \lambda)|$ over a restricted range of 50 dB is represented by the darkness of the marking at $[n, \lambda)$. The plots in Figures 10.12(b) and (c) are vertical slices (shown in Figure 10.12 as magnitude) through the image at $n = 5000$ and $n = 15{,}000$ respectively at the locations of the dashed lines in Figure 10.13(a). Note the linear progression during the chirp interval. Also, note that during the constant-frequency intervals, the dark line remains horizontal. The width of the dark features in Figure 10.13(a) is dependent on the width of the main lobe Δ_{ml} of the DTFT of the window. Table 7.2 indicates that for the Hamming window, this width is approximately $\Delta_{\text{ml}} = 8\pi/M$ wherein $M+1$ is the window length. For a 401-point window, $\Delta_{\text{ml}}/(2\pi) = 0.01$. Thus, the two close-in-frequency cosines are clearly resolved in the interval $25{,}000 < n \leq 30{,}000$, because their normalized frequency difference is $(0.23\pi - 0.2\pi)/(2\pi) = 0.015$, which is significantly greater than the main-lobe width 0.01. Note that the vertical width of the dark sloping bar for the chirp interval is wider than the horizontal bars representing the constant-frequency intervals. This extra broadening is caused by the frequency variation across the window and is a small-scale version of the effect seen in Figure 10.12(a), wherein the variation across the 20,000-sample window is much greater.

The image in Figure 10.13(a) illustrates another important aspect of time-dependent Fourier analysis. The 401-sample window provides good frequency resolution at almost all points in time. However, note that at $n = 20{,}000$ and 25,000 the signal properties change abruptly, so that for an interval of about 401 samples around these times, the window contains samples from both sides of the change. This leads to the fuzzy area wherein the signal properties are much less clearly represented by the spectrogram. We can improve the ability to resolve events in the time dimension by shortening the window. This is illustrated in Figure 10.13(b) wherein the window length is $L = 101$. The points of change are much better resolved with this window. However, the normalized main-lobe frequency width of a 101-sample Hamming window is $\Delta_{\text{ml}}/(2\pi) = 0.04$, and the two constant-frequency cosines after $n = 25{,}000$ are only separated by 0.015 in normalized frequency. Thus, as is clear from Figure 10.13(b), the two frequencies are not resolved with the 101-sample window, although the location of the abrupt changes in the signal are much more accurately resolved in time.

Examples 10.9 and 10.10 illustrate how the principles of discrete-time Fourier analysis that were discussed in Sections 10.1 and 10.2 can be applied to signals whose properties vary with time. Time-dependent Fourier analysis is widely used both as an analysis tool for displaying signal properties and as a representation for signals. In the latter use, it is important to develop a deeper understanding of the two-dimensional representation in Eq. (10.18).

10.3.1 Invertibility of $X[n,\lambda)$

Since $X[n, \lambda)$ is the DTFT of $x[n + m]w[m]$, the time-dependent Fourier transform is invertible if the window has at least one nonzero sample. Specifically, from the Fourier

transform synthesis equation (2.130),

$$x[n+m]w[m] = \frac{1}{2\pi}\int_0^{2\pi} X[n,\lambda)e^{j\lambda m}d\lambda, \qquad -\infty < m < \infty, \tag{10.24}$$

or equivalently,

$$x[n+m] = \frac{1}{2\pi w[m]}\int_0^{2\pi} X[n,\lambda)d\lambda \tag{10.25}$$

if $w[m] \neq 0$.[4] Thus with m chosen as any one value for which $w[m] \neq 0$, $x[n]$ for all values of n can be recovered from $X[n,\lambda)$ using Eq. (10.25).

While the above discussion shows that the time-dependent Fourier transform is an invertible transformation, Eq. (10.24) and (10.25) do not provide a computable inverse, since evaluating them requires knowing $X[n,\lambda)$ at all λ and also requires evaluating an integral. However, the inverse transform becomes a DFT when $X[n,\lambda)$ is sampled in both the time and frequency dimensions. We will discuss this matter more fully in Section 10.3.4.

10.3.2 Filter Bank Interpretation of $X[n,\lambda)$

A rearrangement of the sum in Eq. (10.18) leads to another useful interpretation of the time-dependent Fourier transform. If we make the substitution $m' = n + m$ in Eq. (10.18), then $X[n,\lambda)$ can be written as

$$X[n,\lambda) = \sum_{m'=-\infty}^{\infty} x[m']w[-(n-m')]e^{j\lambda(n-m')}. \tag{10.26}$$

Equation (10.26) can be interpreted as the convolution

$$X[n,\lambda) = x[n] * h_\lambda[n], \tag{10.27a}$$

where

$$h_\lambda[n] = w[-n]e^{j\lambda n}. \tag{10.27b}$$

From Eq. (10.27a), we see that the time-dependent Fourier transform as a function of n with λ fixed can be interpreted as the output of an LTI filter with impulse response $h_\lambda[n]$ or, equivalently, with frequency response

$$H_\lambda(e^{j\omega}) = W(e^{j(\lambda-\omega)}). \tag{10.28}$$

In general, a window that is nonzero for positive time will be called a *noncausal window*, since the computation of $X[n,\lambda)$ using Eq. (10.18) requires samples that follow sample n in the sequence. Equivalently, in the linear-filtering interpretation, the impulse response $h_\lambda[n] = w[-n]e^{j\lambda n}$ is noncausal if $w[n] = 0$ for $n < 0$. That is, a window that is nonzero for $n \geq 0$ gives a noncausal impulse response $h_\lambda[n]$ in Eq. (10.27b), whereas if the window is nonzero for $n \leq 0$, the linear filter is causal.

[4]Since $X[n,\lambda)$ is periodic in λ with period 2π, the integration in Eqs. (10.24) and (10.25) can be over any interval of length 2π.

In the definition of Eq. (10.18), the time origin of the window is held fixed, and the signal is considered to be shifted past the interval of support of the window. This effectively redefines the time origin for Fourier analysis to be at sample n of the signal. Another possibility is to shift the window as n changes, keeping the time origin for Fourier analysis fixed at the original time origin of the signal. This leads to a definition for the time-dependent Fourier transform of the form

$$\check{X}[n, \lambda) = \sum_{m=-\infty}^{\infty} x[m]w[m-n]e^{-j\lambda m}. \tag{10.29}$$

The relationship between the definitions of Eqs. (10.18) and (10.29) is easily shown to be

$$\check{X}[n, \lambda) = e^{-j\lambda n}X[n, \lambda). \tag{10.30}$$

The definition of Eq. (10.18) is particularly convenient when we consider using the DFT to obtain samples in λ of the time-dependent Fourier transform, since, if $w[m]$ is of finite length in the range $0 \leq m \leq (L-1)$, then so is $x[n+m]w[m]$. On the other hand, the definition of Eq. (10.29) has some advantages for the interpretation of Fourier analysis in terms of filter banks. Since our primary interest is in applications of the DFT, we will base most of our discussions on Eq. (10.18).

10.3.3 The Effect of the Window

The primary purpose of the window in the time-dependent Fourier transform is to limit the extent of the sequence to be transformed, so that the spectral characteristics are approximately constant over the duration of the window. The more rapidly the signal characteristics change, the shorter the window should be. We saw in Section 10.2 that as the window becomes shorter, frequency resolution decreases. The same effect is true, of course, for $X[n, \lambda)$. On the other hand, as the window length decreases, the ability to resolve changes with time increases. Consequently, the choice of window length becomes a trade-off between frequency resolution and time resolution. This trade-off was illustrated in Example 10.10.

The effect of the window on the properties of the time-dependent Fourier transform can be seen by assuming that the signal $x[n]$ has a conventional DTFT $X(e^{j\omega})$. First, let us assume that the window is unity for all m; i.e., assume that there is no window at all. Then, from Eq. (10.18),

$$X[n, \lambda) = X(e^{j\lambda})e^{j\lambda n}. \tag{10.31}$$

Of course, a typical window for time-dependent spectrum analysis tapers to zero, so as to select only a portion of the signal for analysis. On the other hand, as discussed in Section 10.2, the length and shape of the window are chosen so that the Fourier transform of the window is narrow in λ compared with variations in λ of the Fourier transform of the signal. Thus, the need for good time resolution and good frequency resolution often requires compromise. The Fourier transform of a typical window is illustrated in Figure 10.14(a).

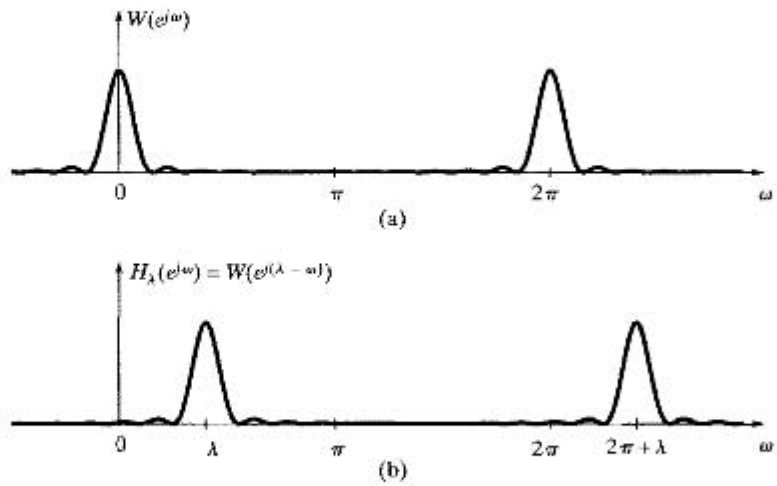


Figure 10.14 (a) Illustration of the Fourier transform of a Bartlett window for time-dependent Fourier analysis. (b) Equivalent bandpass filter for time-dependent Fourier analysis.

If we consider the time-dependent Fourier transform for fixed n, then it follows from the properties of DTFTs that

$$X[n, \lambda) = \frac{1}{2\pi} \int_0^{2\pi} e^{j\theta n} X(e^{j\theta}) W(e^{j(\lambda-\theta)}) d\theta; \tag{10.32}$$

i.e., the Fourier transform of the shifted signal is convolved with the Fourier transform of the window. This is similar to Eq. (10.2), except that in Eq. (10.2), we assumed that the signal was not successively shifted relative to the window. Here, we compute a Fourier transform for each value of n. In Section 10.2, we saw that the ability to resolve two narrowband signal components depends on the width of the main lobe of the Fourier transform of the window, whereas the degree of leakage of one component into the vicinity of the other depends on the relative side-lobe amplitude. The case of no window at all corresponds to $w[n] = 1$ for all n. In this case, $W(e^{j\omega}) = 2\pi\delta(\omega)$ for $-\pi \leq \omega \leq \pi$, which gives precise frequency resolution but no time resolution.

In the linear-filtering interpretation of Eqs. (10.27a), (10.27b), and (10.28), $W(e^{j\omega})$ typically has the lowpass characteristics depicted in Figure 10.14(a), and consequently, $H_\lambda(e^{j\omega})$ is a bandpass filter whose passband is centered at $\omega = \lambda$, as depicted in Figure 10.14(b). Clearly, the width of the passband of this filter is approximately equal to the width of the main lobe of the Fourier transform of the window. The degree of rejection of adjacent frequency components depends on the relative side-lobe amplitude.

The preceding discussion suggests that if we are using the time-dependent Fourier transform to obtain a time-dependent estimate of the frequency spectrum of a signal, it is desirable to taper the window to lower the side lobes and to use as long a window as feasible to improve the frequency resolution. This has already been illustrated in Examples 10.9 and 10.10, and we will consider other examples in Section 10.4. However, before doing so, we discuss the use of the DFT in explicitly evaluating the time-dependent Fourier transform.

10.3.4 Sampling in Time and Frequency

Explicit computation of $X[n, \lambda)$ can be done only on a finite set of values of λ, corresponding to sampling the time-dependent Fourier transform in the domain of its frequency variable λ. Just as finite-length signals can be exactly represented through samples of the DTFT, signals of indeterminate length can be represented through samples of the time-dependent Fourier transform, if the window in Eq. (10.18) has finite length. As an example, suppose that the window has length L with samples beginning at $m = 0$; i.e.,

$$w[m] = 0 \qquad \text{outside the interval } 0 \leq m \leq L-1. \tag{10.33}$$

If we sample $X[n, \lambda)$ at N equally spaced frequencies $\lambda_k = 2\pi k/N$, with $N \geq L$, then we can recover the original windowed sequence from the sampled time-dependent Fourier transform. Specifically, if we define $X[n, k]$ to be

$$X[n, k] = X[n, 2\pi k/N) = \sum_{m=0}^{L-1} x[n+m]w[m]e^{-j(2\pi/N)km}, \qquad 0 \leq k \leq N-1, \tag{10.34}$$

then $X[n, k]$ with n fixed is the DFT of the windowed sequence $x[n+m]w[m]$. Using the inverse DFT, we obtain

$$x[n+m]w[m] = \frac{1}{N}\sum_{k=0}^{N-1} X[n, k]e^{j(2\pi/N)km}, \qquad 0 \leq m \leq L-1. \tag{10.35}$$

Since we assume that the window $w[m] \neq 0$ for $0 \leq m \leq L-1$, the sequence values can be recovered in the interval from n through $(n+L-1)$ using the equation

$$x[n+m] = \frac{1}{Nw[m]}\sum_{k=0}^{N-1} X[n, k]e^{j(2\pi/N)km}, \qquad 0 \leq m \leq L-1. \tag{10.36}$$

The important point is that the window has finite length and that we can take at least as many samples in the λ dimension as there are nonzero samples in the window; i.e., $N \geq L$. While Eq. (10.33) corresponds to a noncausal window, we could have used a causal window with $w[m] \neq 0$ for $-(L-1) \leq m \leq 0$ or a symmetric window such that $w[m] = w[-m]$ for $|m| \leq (L-1)/2$, with L an odd integer. The use of a noncausal window in Eq. (10.34) is simply more convenient for our analysis, since it leads very naturally to the interpretation of the sampled time-dependent Fourier transform as the DFT of the windowed block of samples beginning with sample n.

Since Eq. (10.34) corresponds to sampling Eq. (10.18) in λ, it also corresponds to sampling Eqs. (10.26), (10.27a), and (10.27b) in λ. Specifically, Eq. (10.34) can be rewritten as

$$X[n, k] = x[n] * h_k[n], \qquad 0 \leq k \leq N-1, \tag{10.37a}$$

where

$$h_k[n] = w[-n]e^{j(2\pi/N)kn}. \tag{10.37b}$$

Equations (10.37a) and (10.37b) can be viewed as a bank of N filters, as depicted in Figure 10.15, with the k^{th} filter having frequency response

$$H_k(e^{j\omega}) = W(e^{j[(2\pi k/N)-\omega]}). \tag{10.38}$$

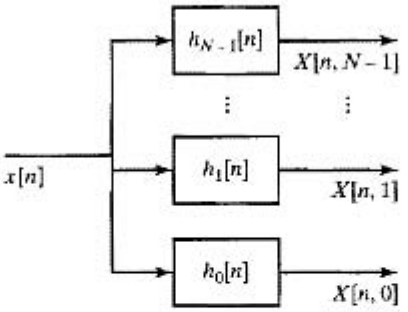


Figure 10.15 Filter bank representation of the time-dependent Fourier transform.

Our discussion suggests that $x[n]$ for $-\infty < n < \infty$ can be reconstructed if $X[n, \lambda)$ or $X[n, k]$ is sampled in the time dimension, as well. Specifically, using Eq. (10.36), we can reconstruct the signal in the interval $n_0 \leq n \leq n_0 + L - 1$ from $X[n_0, k]$, and we can reconstruct the signal in the interval $n_0 + L \leq n \leq n_0 + 2L - 1$ from $X[n_0 + L, k]$, and so on. Thus, $x[n]$ can be reconstructed exactly from the time-dependent Fourier transform sampled in both the frequency and the time dimension. In general, for the region of support of the window as specified in Eq. (10.33), we define this sampled time-dependent Fourier transform as

$$X[rR, k] = X[rR, 2\pi k/N] = \sum_{m=0}^{L-1} x[rR + m]w[m]e^{-j(2\pi/N)km}, \qquad (10.39)$$

where r and k are integers such that $-\infty < r < \infty$ and $0 \leq k \leq N - 1$. To further simplify our notation, we define

$$X_r[k] = X[rR, k] = X[rR, \lambda_k), \qquad -\infty < r < \infty, \quad 0 \leq k \leq N - 1, \qquad (10.40)$$

where $\lambda_k = 2\pi k/N$. This notation denotes explicitly that the sampled time-dependent Fourier transform is simply a sequence of N-point DFTs of the windowed signal segments

$$x_r[m] = x[rR + m]w[m], \qquad -\infty < r < \infty, \quad 0 \leq m \leq L - 1, \qquad (10.41)$$

with the window position moving in jumps of R samples in time. Figure 10.16 shows lines in the $[n, \lambda)$-plane corresponding to the region of support of $X[n, \lambda)$ and the grid of sampling points in the $[n, \lambda)$-plane for the case $N = 10$ and $R = 3$. As we have shown, it is possible to uniquely reconstruct the original signal from such a two-dimensional discrete representation for appropriate choice of L.

Equation (10.39) involves the following integer parameters: the window length L; the number of samples in the frequency dimension, or the DFT length N; and the sampling interval in the time dimension, R. Although not all choices of these parameters will permit exact reconstruction of the signal, numerous combinations of N, R, and $w[n]$ and L can be used. The choice $L \leq N$ guarantees that it is possible to reconstruct the windowed segments $x_r[m]$ from the block transforms $X_r[k]$. If $R < L$, the segments overlap, but if $R > L$, some of the samples of the signal are not used and therefore cannot be reconstructed from $X_r[k]$. Thus, as one possibility, if the three sampling parameters satisfy the relation $R \leq L \leq N$, then we can (in principle) recover R samples of $x[n]$ block-by-block for all n from $X_r[k]$. Notice that each block of R samples of the

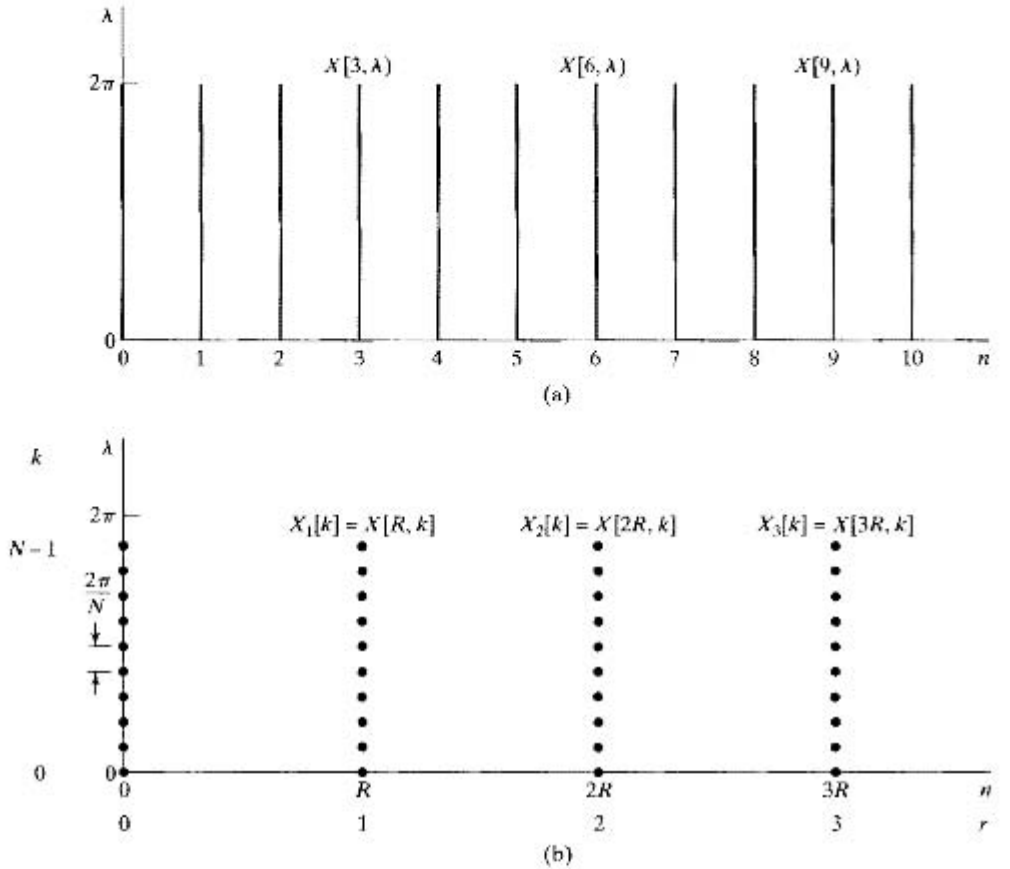


Figure 10.16 (a) Region of support for $X[n, \lambda)$. (b) Grid of sampling points in the $[n, \lambda)$-plane for the sampled time-dependent Fourier transform with $N = 10$ and $R = 3$.

signal is represented by N complex numbers in the sampled time-dependent Fourier representation; or, if the signal is real, only N real numbers are required, due to the symmetry of the DFT.

As a specific example, the signal can be reconstructed exactly from the sampled time-dependent Fourier transform for the special case $R = L = N$. In this case, N samples of a real signal are represented by N real numbers, and this is the minimum that we could expect to achieve for an arbitrarily chosen signal. For $R = L = N$ we can recover $x_r[m] = x[rR + m]w[m]$ for $0 \leq m \leq N - 1$ by computing the inverse DFT of $X_r[k]$. Therefore, we can express $x[n]$ for $rR \leq n \leq [(r + 1)R - 1]$ in terms of the windowed segments $x_r[m]$ as

$$x[n] = \frac{x_r[n - rR]}{w[n - rR]} \qquad rR \leq n \leq [(r + 1)R - 1], \tag{10.42}$$

i.e., we recover the N-point windowed segments, remove the effect of the window, and then abut the segments together to reconstruct the original sequence.

10.3.5 The Overlap–Add Method of Reconstruction

While the previous discussion verifies the possibility of theoretically exact reconstruction of the signal from its time- and frequency-sampled time-dependent Fourier transform, the demonstration proof is not a viable reconstruction algorithm when modifications are made to the time-dependent Fourier transform as is common, for example, in applications such as audio coding and noise reduction. In these applications, division by a tapering window as required in Eq. (10.42) can greatly enhance errors at the edges; therefore, the signal blocks may not fit together smoothly. In such applications, it is helpful to make R smaller than L and N so that the blocks of samples overlap. Then, if the window is properly chosen, it will not be necessary to undo the windowing as in Eq. (10.42).

Suppose that $R \leq L \leq N$. Then we can write

$$x_r[m] = x[rR + m]w[m] = \frac{1}{N}\sum_{k=0}^{N-1} X_r[k]e^{j(2\pi k/N)m} \quad 0 \leq m \leq L-1. \tag{10.43}$$

The recovered segments are shaped by the window, and their time origin is at the beginning of the window. A different approach to putting the signal back together that is more robust to changes in $X_r[k]$ is to shift the windowed segments to their original time locations rR and then simply add them together; i.e.,

$$\hat{x}[n] = \sum_{r=-\infty}^{\infty} x_r[n - rR]. \tag{10.44}$$

If we can show that $\hat{x}[n] = x[n]$ for all n, then Eqs. (10.43) and (10.44) together comprise a method for *time-dependent Fourier synthesis* having the capability of perfect reconstruction. Substituting Eq. (10.43) into Eq. (10.44) leads to the following representation of $\hat{x}[n]$:

$$\begin{aligned}\hat{x}[n] &= \sum_{r=-\infty}^{\infty} x[rR + n - rR]w[n - rR] \\ &= x[n]\sum_{r=-\infty}^{\infty} w[n - rR]\end{aligned} \tag{10.45}$$

If we define

$$\tilde{w}[n] = \sum_{r=-\infty}^{\infty} w[n - rR], \tag{10.46a}$$

then the reconstructed signal in Eq. (10.45) can be expressed as

$$\hat{x}[n] = x[n]\tilde{w}[n]. \tag{10.46b}$$

It follows from Eq. (10.46b) that the condition for perfect reconstruction is

$$\tilde{w}[n] = \sum_{r=-\infty}^{\infty} w[n - rR] = C \quad -\infty < n < \infty, \tag{10.47}$$

i.e., the shifted-by-R copies of the window must add to a constant reconstruction gain C for all n.

Note that the sequence $\tilde{w}[n]$ is a periodic sequence (with period R) comprised of time-aliased window sequences. As a simple example, consider a rectangular window $w_{\text{rect}}[n]$ of length L samples. If $R = L$, the windowed segments simply fit together block-by-block with no overlap. In this case, the condition of Eq. (10.47) is satisfied with $C = 1$, because the shifted windows fit together with no overlap and no gaps. (A simple sketch will confirm this.) If L for the rectangular window is even, and $R = L/2$ a simple analysis or sketch will again verify that the condition of Eq. (10.47) is satisfied with $C = 2$. In fact, if $L = 2^v$, the signal $x[n]$ can be perfectly reconstructed from $X_r[k]$ by the overlap–add method of Eq. (10.44) when $L \leq N$ and $R = L, L/2, \ldots, 1$. The corresponding reconstruction gains would be $C = 1, 2, \ldots, L$. While this demonstrates that the overlap–add method can perfectly reconstruct the original signal for some rectangular windows, and some window spacings R, the rectangular window is rarely used in time-dependent Fourier analysis/synthesis because of its poor leakage properties. Other tapered windows such as the Bartlett, Hann, Hamming, and Kaiser windows are more commonly used. Fortunately, these windows with their superior spectral isolation properties, can also produce perfect or near-perfect reconstruction from the time-dependent Fourier transform.

Two windows with which perfect reconstruction can be achieved are the Bartlett and Hann windows, which were introduced in Chapter 7 in the context of FIR filter design. They are defined again here in Eqs (10.48) and (10.49), respectively:

Bartlett (triangular)

$$w_{\text{Bart}}[n] = \begin{cases} 2n/M, & 0 \leq n \leq M/2, \\ 2 - 2n/M, & M/2 < n \leq M, \\ 0, & \text{otherwise} \end{cases} \tag{10.48}$$

Hann

$$w_{\text{Hann}}[n] = \begin{cases} 0.5 - 0.5\cos(2\pi n/M), & 0 \leq n \leq M, \\ 0, & \text{otherwise} \end{cases} \tag{10.49}$$

As these windows are defined, the window length is $L = M + 1$ with the two end samples equal to zero.[5] With M even and $R = M/2$, then it is easily shown for the Bartlett window that the condition of Eq. (10.47) is satisfied with $C = 1$. Figure 10.17(a) shows overlapping Bartlett windows of length $M + 1$ (first and last samples zero) when $R = M/2$. It is clear that these shifted windows add up to the reconstruction gain constant $C = 1$. Figure 10.17(b) shows the same choice of $L = M + 1$ and $R = M/2$ for the Hann window. Although it is less obvious from this plot, it is also true that these shifted windows add up for all n to the constant $C = 1$. A similar statement is also true for the Hamming window and many other windows.

[5]With these definitions, the actual number of nonzero samples is $M - 1$ for both the Bartlett and Hann windows, but the inclusion of the zero samples leads to convenient mathematical simplifications.

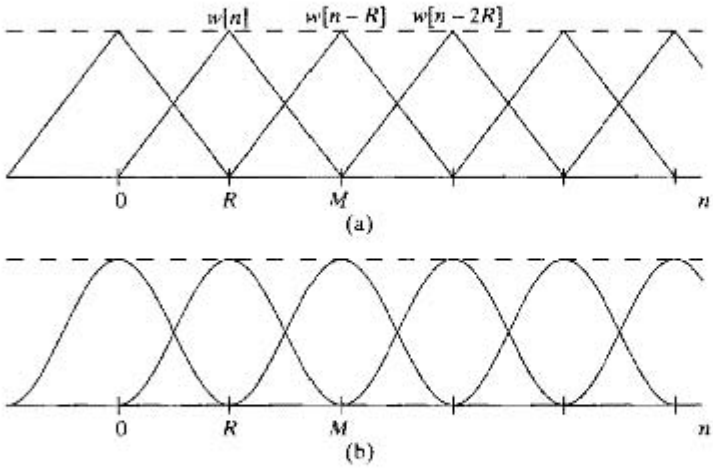


Figure 10.17 (a) Shifted $M + 1$-point Bartlett windows with $R = M/2$. (b) Shifted $M + 1$-point Hann windows with $R = M/2$. The dashed line is the periodic sequence $\tilde{w}[n]$.

Figure 7.30 gives a comparison of the DTFT of the rectangular, Bartlett and Hann windows. Note that the main-lobe width of the Bartlett and Hann windows is twice that of the rectangular window of the same length L, but the side lobes are significantly lower in amplitude for both the Bartlett and Hann windows. Thus, they and the other windows in Figure 7.30 are much preferred over the rectangular window for time-dependent Fourier analysis/synthesis.

While Figure 10.17 is intuitively plausible, it is less obvious that the Bartlett and Hann windows for $M = 2^v$ can provide perfect reconstruction for values of $R = M/2, M/4, \ldots, 1$ with corresponding reconstruction gains of $M/(2R)$. To see this, it is helpful to recall that the envelope sequence $\tilde{w}[n]$ is inherently periodic with period R, so it can be represented by an inverse DFT as

$$\tilde{w}[n] = \sum_{r=-\infty}^{\infty} w[n - rR] = \frac{1}{R} \sum_{k=0}^{R-1} W(e^{j(2\pi k/R)}) e^{j(2\pi k/R)n}, \tag{10.50}$$

where $W(e^{j(2\pi k/R)})$ is the DTFT of $w[n]$ sampled at frequencies $(2\pi k/R)$, $k = 0, 1, \ldots, R - 1$. From Eq. (10.50) it is clear that a condition for perfect reconstruction is

$$W(e^{j(2\pi k/R)}) = 0 \qquad k = 1, 2, \ldots, R - 1, \tag{10.51a}$$

and if Eq. (10.51a) holds, then it follows from Eq. (10.50) that the reconstruction gain is

$$C = \frac{W(e^{j0})}{R}. \tag{10.51b}$$

Problem 7.43 of Chapter 7 explores the notion that the commonly used Bartlett, Hann, Hamming, and Blackman windows can be represented in terms of rectangular windows for which it is relatively easy to obtain a closed-form expression for the DTFT of the window. In particular, Problem 7.43 gives the result that for M even, the Bartlett window defined as in Eq. (10.48) has DTFT

$$W_{\text{Bart}}(e^{j\omega}) = \left(\frac{2}{M}\right) \left(\frac{\sin(\omega M/4)}{\sin(\omega/2)}\right)^2 e^{-j\omega M/2}. \tag{10.52}$$

From Eq. (10.52) it follows that the Bartlett window Fourier transform has equally spaced zeros at frequencies $4\pi k/M$, for $k = 1, 2, \ldots, M - 1$. Therefore, if we choose R

so that $2\pi k/R = 4\pi k/M$ or $R = M/2$, the condition Eq. (10.51a) is satisfied. Substituting $\omega = 0$ into Eq. (10.52) gives $W_{\text{Bart}}(e^0) = M/2$, so it follows that perfect reconstruction results with $C = M/(2R) = 1$ if $R = M/2$. Choosing $R = M/2$ aligns the frequencies $2\pi k/R$ with all the zeros of $W_{\text{Bart}}(e^{j\omega})$. If M is divisible by 4, we can use $R = M/4$ and the frequencies $2\pi k/R$ will still align with zeros of $W_{\text{Bart}}(e^{j\omega})$, and the reconstruction gain will be $C = M/(2R) = 2$. If M is a power of two, R can be smaller with concomitant increase in C.

The DTFT $W_{\text{Hann}}(e^{j\omega})$ also has zeros equally spaced at integer multiples of $4\pi/M$, so exact reconstruction is also possible with the Hann window defined as in Eq. (10.49). The equally spaced zeros of $W_{\text{Bart}}(e^{j\omega})$ and $W_{\text{Hann}}(e^{j\omega})$ are evident in the plots in Figure 7.30(b) and (c), respectively. Figure 7.30(d) shows the Hamming window, which is a version of the Hann window that is optimized to minimize the side-lobe levels. As a result of the adjustment of the coefficients from 0.5 and 0.5 to 0.54 and 0.46, the zeros of $W_{\text{Hamm}}(e^{j\omega})$ are slightly displaced, so it is no longer possible to choose R so that the frequencies $2\pi k/R$ fall precisely on zeros of $W_{\text{Hamm}}(e^{j\omega})$. However, as shown in Table 7.2, the maximum side-lobe level for frequencies above $4\pi/M$ is -41 dB. Thus, the condition of Eq. (10.51a) is satisfied approximately at each of the frequencies $2\pi k/R$. Equation 10.50 shows that if Eq. (10.51a) is not satisfied exactly, $\tilde{w}[n]$ will tend to oscillate around C with period R imparting a slight amplitude modulation to the reconstructed signal.

10.3.6 Signal Processing Based on the Time-Dependent Fourier Transform

A general framework for signal processing based on the time-dependent Fourier transform is depicted in Figure 10.18. This system is based on the fact that a signal $x[n]$ can be reconstructed exactly from its time- and frequency-sampled time-dependent Fourier transform $X_r[k]$ if the window and sampling parameters are appropriately chosen, as discussed above. If the processing shown in Figure 10.18 is done so that $Y_r[k]$ maintains its integrity as a time-dependent Fourier transform, then a processed signal $y[n]$ can be reconstructed by a technique of time-dependent Fourier synthesis, such as the overlap–add method or a technique involving a bank of bandpass filters. For example, if $x[n]$ is an audio signal, $X_r[k]$ can be quantized for signal compression. The time-dependent Fourier representation provides a natural and convenient framework, wherein auditory masking phenomena can be exploited to "hide" the quantization noise. (See, for example, Bosi and Goldberg, 2003 and Spanias, Painter and Atti, 2007.) Time-dependent Fourier synthesis is then used to reconstruct a signal $y[n]$ for listening. This is the basis for MP3 audio coding, for example. Another application is audio noise suppression,

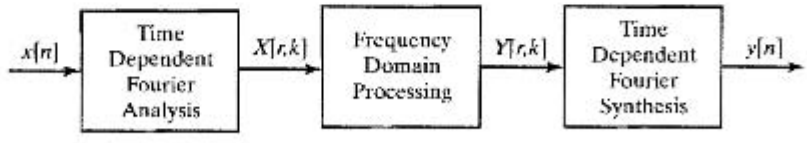


Figure 10.18 Signal processing based on time-dependent Fourier analysis/synthesis.

wherein the acoustic noise spectrum is estimated and then either subtracted from the time-dependent Fourier spectrum of the input signal or used as the basis for Wiener filtering applied to the $X_r[k]$. (See Quatieri, 2002.) These and many other applications are greatly facilitated by the FFT algorithms that are available for efficient computation of the discrete-time-dependent Fourier transform.

A discussion of applications of this type would take us too far afield; however, these kinds of block-processing techniques for discrete-time signals were also introduced in Chapter 8, when we discussed the use of the DFT for implementing the convolution of a finite-length impulse response with an input signal of indefinite length. This method of implementation of LTI systems has a useful interpretation in terms of the definitions and concepts of time-dependent Fourier analysis and synthesis, as discussed so far.

Specifically, assume that $x[n] = 0$ for $n < 0$, and suppose that we compute the time-dependent Fourier transform for $R = L$ and a rectangular window. In other words, the sampled time-dependent Fourier transform $X_r[k]$ consists of a set of N-point DFTs of segments of the input sequence

$$x_r[m] = x[rL + m], \qquad 0 \le m \le L - 1. \tag{10.53}$$

Since each sample of the signal $x[n]$ is included, and the blocks do not overlap, it follows that

$$x[n] = \sum_{r=0}^{\infty} x_r[n - rL]. \tag{10.54}$$

Now, suppose that we define a new time-dependent Fourier transform

$$Y_r[k] = H[k]X_r[k], \qquad 0 \le k \le N - 1, \tag{10.55}$$

where $H[k]$ is the N-point DFT of a finite-length unit sample sequence $h[n]$ such that $h[n] = 0$ for $n < 0$ and for $n > P - 1$. If we compute the inverse DFT of $Y_r[k]$, we obtain

$$y_r[m] = \frac{1}{N}\sum_{k=0}^{N-1} Y_r[k]e^{j(2\pi/N)km} = \sum_{\ell=0}^{N-1} x_r[\ell]h[((m - \ell))_N]. \tag{10.56}$$

That is, $y_r[m]$ is the N-point circular convolution of $h[m]$ and $x_r[m]$. Since $h[m]$ has length P samples and $x_r[m]$ has length L samples, it follows from the discussion of Section 8.7 that if $N \ge L + P - 1$, then $y_r[m]$ will be identical to the linear convolution of $h[m]$ with $x_r[m]$ in the interval $0 \le m \le L + P - 2$, and it will be zero, otherwise. Thus, it follows that if we construct an output signal

$$y[n] = \sum_{r=0}^{\infty} y_r[n - rL], \tag{10.57}$$

then $y[n]$ is the output of an LTI system with impulse response $h[n]$. The procedure just described corresponds exactly to the *overlap–add* method of block convolution. The overlap–save method discussed in Section 8.7 can also be applied within the framework of the time-dependent Fourier transform.

10.3.7 Filter Bank Interpretation of the Time-Dependent Fourier Transform

Another way to see that the time-dependent Fourier transform can be sampled in the time dimension is to recall that for fixed λ (or for fixed k if the analysis frequencies are $\lambda_k = 2\pi k/N$) the time-dependent Fourier transform is a one-dimensional sequence in time that is the output of a bandpass filter with frequency response as in Eq. (10.28).

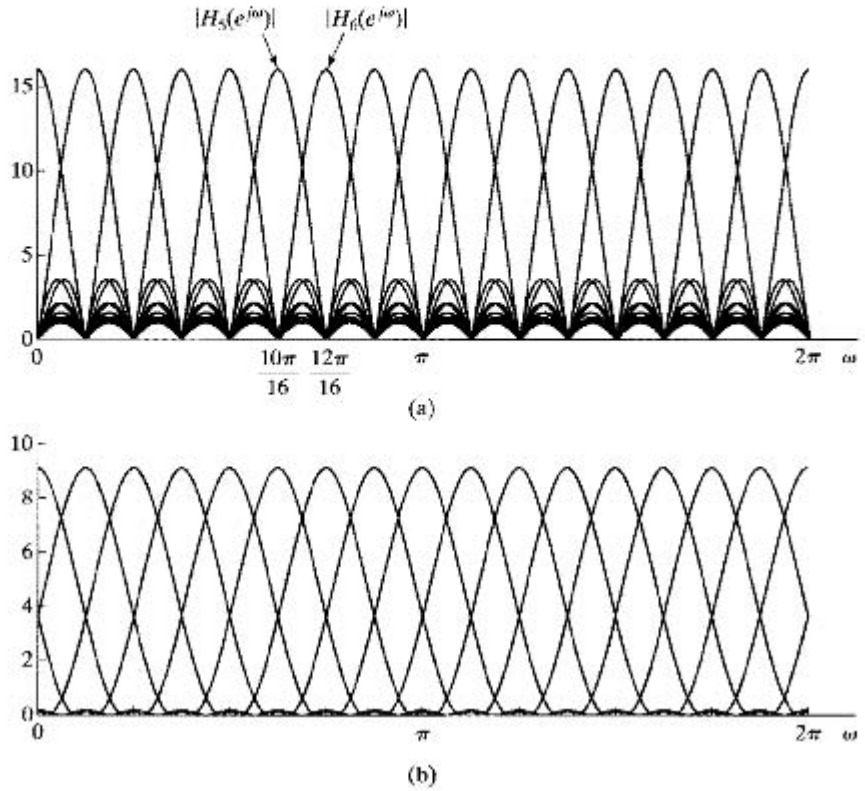


Figure 10.19 Filterbank frequency response. (a) Rectangular window. (b) Kaiser window.

This is illustrated in Figure 10.19. Figure 10.19(a) shows the equivalent set of bandpass filters corresponding to a rectangular window with $L = N = 16$. Figure 10.19 illustrates the filter bank interpretation, even for the case where L and N are much larger. When N increases, the filter bands become narrower, and the side lobes overlap with adjacent channels in the same way. Note that the passbands of the filters corresponding to the rectangular window overlap significantly, and their frequency selectivity is not good by any standard. In fact, the side lobes of any one of the bandpass filters overlap completely with several of the passbands on either side. This suggests that, in general, we might encounter a problem with aliasing in the time dimension, since the Fourier transform of any other finite-length tapering window will not be an ideal filter response either. Our discussion in Section 10.3.5, however, shows that even the rectangular window can provide perfect reconstruction with overlapping windows, in spite of the poor frequency selectivity. Although aliasing occurs in the individual bandpass filter outputs, it can be argued that the aliasing distortion cancels out when all channels are recombined in the overlap–add synthesis. This notion of alias cancellation is one of the important concepts to result from a detailed investigation of the filter bank interpretation.

If a tapering window is used, the side lobes are greatly reduced. Figure 10.19(b) shows the case for a Kaiser window of the same length as the rectangular window used

in Figure 10.19(a), i.e., $L = N = 16$. The side lobes are much smaller, but the main lobe is much wider, so the filters overlap even more. Again, the previous argument based on block processing ideas shows conclusively that we can reconstruct the original signal almost exactly from the time- and frequency-sampled time-dependent Fourier transform if R is small enough. Thus, for a Kaiser window such as in Figure 10.19(b), the sampling rate of the sequences representing each of the bandpass analysis channels could be $2\pi/R = \Delta_{\text{ml}}$, where Δ_{ml} is the width of the main lobe of the Fourier transform of the window.[6] In the example of Figure 10.19(b), the main lobe width is approximately $\Delta_{\text{ml}} = 0.4\pi$, which implies that the time sampling interval could be $R = 5$ for nearly perfect reconstruction of the signal from $X[rR, \lambda_k)$ by the overlap–add method. More generally, in the case of the Hamming window of length $L = M + 1$ samples, for example, $\Delta_{\text{ml}} = 8\pi/M$ so nominally, the time sampling interval should be $R = M/4$. With this sampling rate in time, our discussion above shows that the signal $x[n]$ could be reconstructed nearly perfectly from $X[rR, \lambda_k)$ using a Hamming window and the overlap–add method of synthesis with $R = L/4$ and $L \leq N$.

When using the overlap–add method of analysis/synthesis, the parameters generally satisfy the relation $R \leq L \leq N$. This implies that (taking account of symmetries) the effective total number of samples (numbers) per second of the time-dependent Fourier representation $X[rR, \lambda_k)$ is a factor of N/R greater than the sample rate of $x[n]$ itself. This may not be an issue in some applications, but it presents a significant problem in data compression applications, such as audio coding. Fortunately, the filter bank point of view is the basis for showing that it is possible to choose these parameters to satisfy $R = N < L$ and still achieve nearly perfect reconstruction of the signal from its time-dependent Fourier transform. An example of such an analysis/synthesis system was discussed in Section 4.7.6, where $R = N = 2$, and the lowpass and highpass filters have impulse responses of length L, which can be as large as desired to achieve sharp cutoff filters. The two-channel filter bank can be generalized to a higher number of channels with $R = N$, and, as in the example of Section 4.7.6, polyphase techniques can be employed to increase computational efficiency. The advantage of requiring $R = N$ is that the total number of samples/s remains the same as for the input $x[n]$. As an example, Figure 10.20 shows the first few bandpass channels of the basic analysis filter bank specified by the MPEG-II audio coding standard. This filter bank performs time-dependent Fourier analysis with offset center frequencies $\lambda_k = (2k + 1)\pi/64$ using 32 real bandpass filters. Since the real bandpass filters have a pair of passbands centered at frequencies $\pm\lambda_k$, this is equivalent to 64 complex bandpass filters. In this case, the length of the impulse responses (equivalent to the window length) is $L = 513$ with the first and last samples being equal to zero. The downsampling factor is $R = 32$. Observe that the filters overlap significantly at their band edges, and downsampling by $R = 32$ causes significant aliasing distortion. However, a more detailed analysis of the complete analysis/synthesis system shows that the aliasing distortion due to the nonideal frequency responses cancels in the reconstruction process.

[6] Since, for our definition, the time-dependent Fourier transform channel signals, $X[n, \lambda_k)$, are bandpass signals centered on frequency λ_k, they can be frequency-downshifted by λ_k, so that the result is a lowpass signal in the band $\pm\Delta_{\text{ml}}$. The resulting lowpass signals have highest frequency $\Delta_{\text{ml}}/2$, so the lowest sampling rate would be $2\pi/R = \Delta_{\text{ml}}$. If $R = N$, the frequency-downshifting occurs automatically as a result of the downsampling operation.

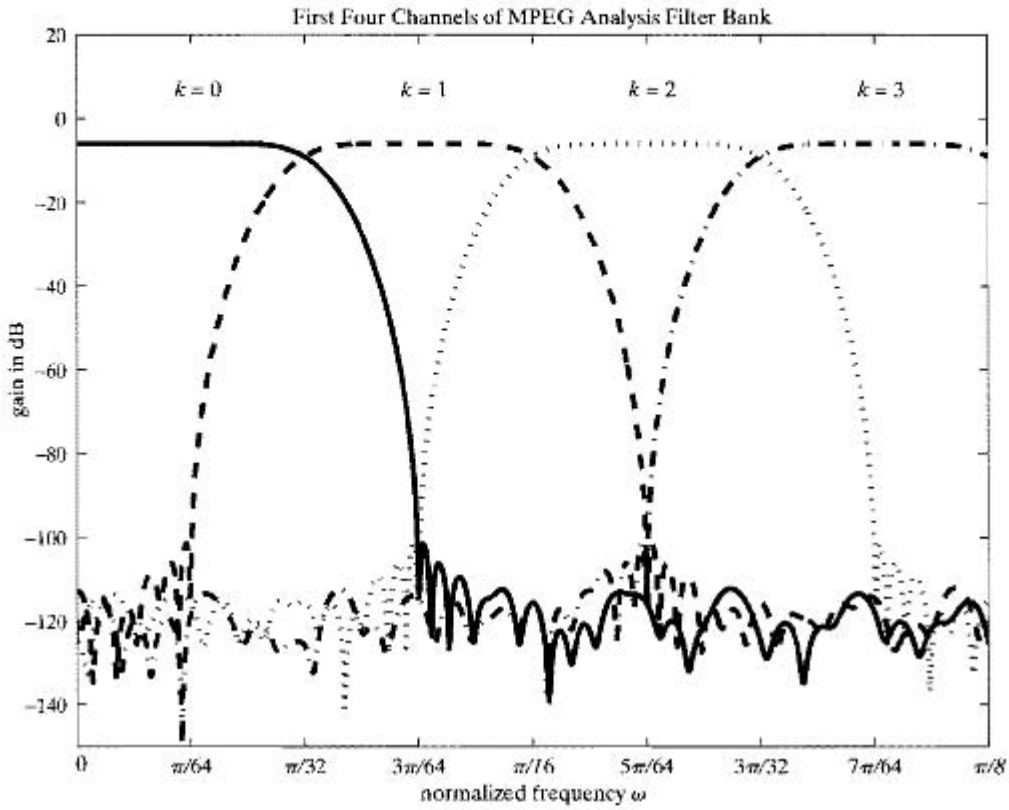


Figure 10.20 Several bandpass channels for the MPEG-II analysis filter bank.

A full-scale discussion of analysis and synthesis filter banks is beyond our scope in this chapter. An outline of such a discussion is given as the basis for Problem 10.46, and detailed discussions can be found in Rabiner and Schafer (1978), Crochiere and Rabiner (1983) and Vaidyanathan (1993).

10.4 EXAMPLES OF FOURIER ANALYSIS OF NONSTATIONARY SIGNALS

In Section 10.3.6, we considered a simple example of how the time-dependent Fourier transform can be used to implement linear filtering. In such applications, we are not so much interested in spectral resolution as in whether it is possible to reconstruct a modified signal from the modified time-dependent Fourier transform. On the other hand, the concept of the time-dependent Fourier transform is often used as a framework for a variety of techniques for obtaining spectrum estimates for nonstationary discrete-time signals, and in these applications spectral resolution, time variation, and other issues are the most important.

A nonstationary signal is a signal whose properties vary with time, for example, a sum of sinusoidal components with time-varying amplitudes, frequencies, or phases. As we will illustrate in Section 10.4.1 for speech signals and in Section 10.4.2 for Doppler

radar signals, the time-dependent Fourier transform often provides a useful description of how the signal properties change with time.

When we apply time-dependent Fourier analysis to a sampled signal, the entire discussion of Section 10.1 holds for each DFT that is computed. In other words, for each segment $x_r[n]$ of the signal, the sampled time-dependent Fourier transform $X_r[k]$ would be related to the Fourier transform of the original continuous-time signal by the processes described in Section 10.1. Furthermore, if we were to apply the time-dependent Fourier transform to sinusoidal signals with constant (i.e., nontime-varying) parameters, the discussion of Section 10.2 should also apply to each of the DFTs that we compute. When the signal frequencies do not change with time, it is tempting to assume that the time-dependent Fourier transform would vary only in the frequency dimension in the manner described in Section 10.2, but this would be true only in very special cases. For example, the time-dependent Fourier transform will be constant in the time dimension if the signal is periodic with period N_p and $L = \ell_0 N_p$ and $R = r_0 N_p$, where ℓ_0 and r_0 are integers; i.e., the window includes exactly ℓ_0 periods, and the window is moved by exactly r_0 periods between computations of the DFT. In general, even if the signal is exactly periodic, the varying phase relationships that would result as different segments of the waveform are shifted into the analysis window would cause the time-dependent Fourier transform to vary in the time dimension. However, for stationary signals, if we use a window that tapers to zero at its ends, the magnitude $|X_r[k]|$ will vary only slightly from segment to segment, with most of the variation of the complex time-dependent Fourier transform occurring in the phase.

10.4.1 Time-Dependent Fourier Analysis of Speech Signals

Speech is produced by excitation of an acoustic tube, the *vocal tract*, which is terminated on one end by the lips and on the other end by the glottis. There are three basic classes of speech sounds:

- *Voiced sounds* are produced by exciting the vocal tract with quasi-periodic pulses of airflow caused by the opening and closing of the glottis.
- *Fricative sounds* are produced by forming a constriction somewhere in the vocal tract and forcing air through the constriction so that turbulence is created, thereby producing a noise-like excitation.
- *Plosive sounds* are produced by completely closing off the vocal tract, building up pressure behind the closure, and then abruptly releasing the pressure.

Detailed discussions of models for the speech signal and applications of the time-dependent Fourier transform are found in texts such as Flanagan (1972), Rabiner and Schafer (1978), O'Shaughnessy (1999), Parsons (1986) and Quatieri (2002).

With a constant vocal tract shape, speech can be modeled as the response of an LTI system (the vocal tract) to a quasiperiodic pulse train for voiced sounds or wideband noise for unvoiced sounds. The vocal tract is an acoustic transmission system characterized by its natural frequencies, called *formants*, which correspond to resonances in its frequency response. In normal speech, the vocal tract changes shape relatively slowly with time as the tongue and lips perform the gestures of speech, and thus it can be

modeled as a slowly time-varying filter that imposes its frequency-response properties on the spectrum of the excitation. A typical speech waveform is shown in Figure 10.21.

From this brief description of the process of speech production and from Figure 10.21, we see that speech is definitely a nonstationary signal. However, as illustrated

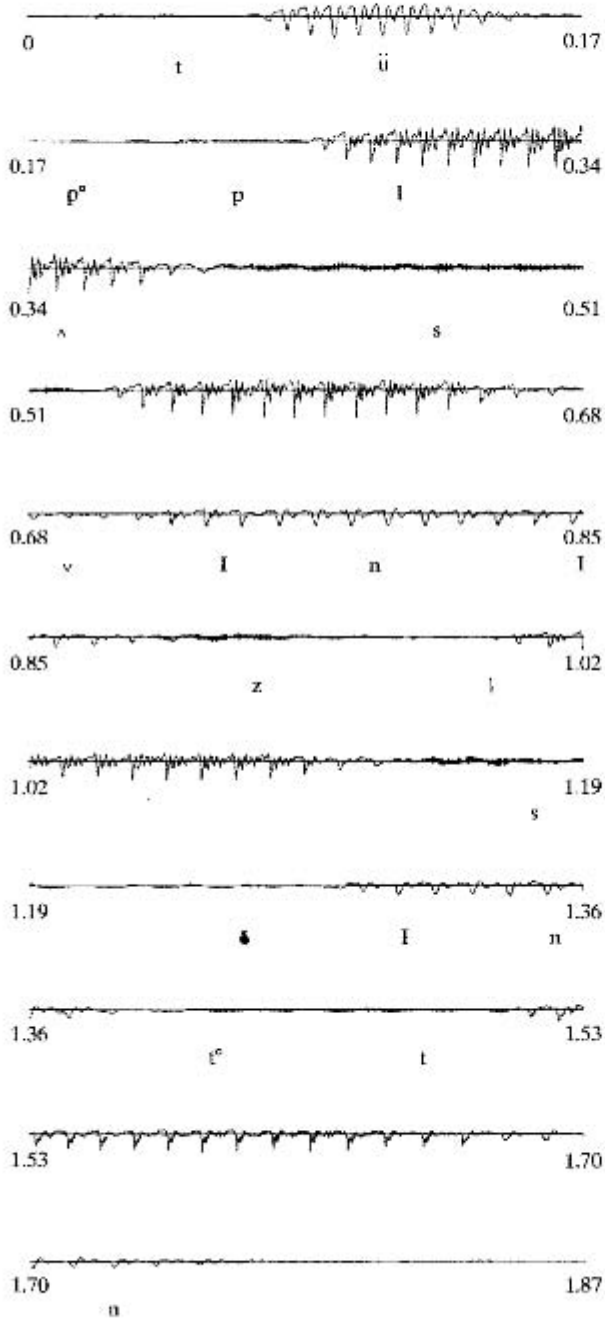


Figure 10.21 Waveform of the speech utterance "Two plus seven is less than ten." Each line is 0.17 s in duration. The time-aligned phonemic transcript is indicated below the waveform. The sampling rate is 16,000 samples/s, so each line represents 2720 samples.

in the figure, the characteristics of the signal can be assumed to remain essentially constant over time intervals on the order of 30 or 40 ms. The frequency content of the speech signal may range up to 15 kHz or higher, but speech is highly intelligible even when bandlimited to frequencies below about 3 kHz. Commercial telephone systems, for example, typically limit the highest transmitted frequency to about 3 kHz. A standard sampling rate for digital telephone communication systems is 8000 samples/s.

Figure 10.21 shows that the waveform consists of a sequence of quasiperiodic *voiced* segments interspersed with noise-like *unvoiced* segments. This figure suggests that if the window length L is not too long, the properties of the signal will not change appreciably from the beginning of the segment to the end. Thus, the DFT of a windowed speech segment should display the frequency-domain properties of the signal at the time corresponding to the window location. For example, if the window length is long enough so that the fundamental frequency and its harmonics are resolved, the DFT of a windowed segment of voiced speech should show a series of peaks at integer multiples of the fundamental frequency of the signal in that interval. This would normally require that the window span several periods of the waveform. If the window is too short, then the harmonics will not be resolved, but the general spectrum shape will still be evident. This is typical of the trade-off between frequency resolution and time resolution that is required in the analysis of nonstationary signals. We saw this before in Example 10.9. If the window is too long, the signal properties may change too much across the window; if the window is too short, resolution of narrowband components will be sacrificed. This trade-off is illustrated in the following example.

Example 10.11 Spectrogram Display of the Time-Dependent Fourier Transform of Speech

Figure 10.22(a) shows a spectrogram display of the time-dependent Fourier transform of the speech signal in Figure 10.21. The time waveform is also shown on the same time scale, below the spectrogram. More specifically, Figure 10.22(a) is a *wideband spectrogram*. A wideband spectrogram representation results from a window that is relatively short in time and is characterized by poor resolution in the frequency dimension and good resolution in the time dimension. The frequency axis is labeled in terms of continuous-time frequency. Since the sampling rate of the signal was 16,000 samples/s, it follows that the frequency $\lambda = \pi$ corresponds to 8 kHz. The specific window used in Figure 10.22(a) was a Hamming window of duration 6.7 ms, corresponding to $L = 108$. The value of R was 16, representing 1-ms time increments.[7] The broad, dark bars that move horizontally across the spectrogram correspond to the resonance frequencies of the vocal tract, which, as we see, change with time. The vertically striated appearance of the spectrogram is due to the quasiperiodic nature of voiced portions of the waveform, as is evident by comparing the variations in the waveform display and the spectrogram. Since the length of the analysis window is on the order of the length of a period of the waveform, as the window slides along in time, it alternately covers high-energy segments of the waveform and then lower energy segments in between, thereby producing the vertical striations in the plot during voiced intervals.

In a *narrowband* time-dependent Fourier analysis, a longer window is used to provide higher frequency resolution, with a corresponding decrease in time resolution.

[7] In plotting spectrograms, it is common to use relatively small values of R so that a smoothly varying display is obtained.

Such a narrowband analysis of speech is illustrated by the display in Figure 10.22(b). In this case, the window was a Hamming window of duration 45 ms. This corresponds to $L = 720$. The value of R was again 16.

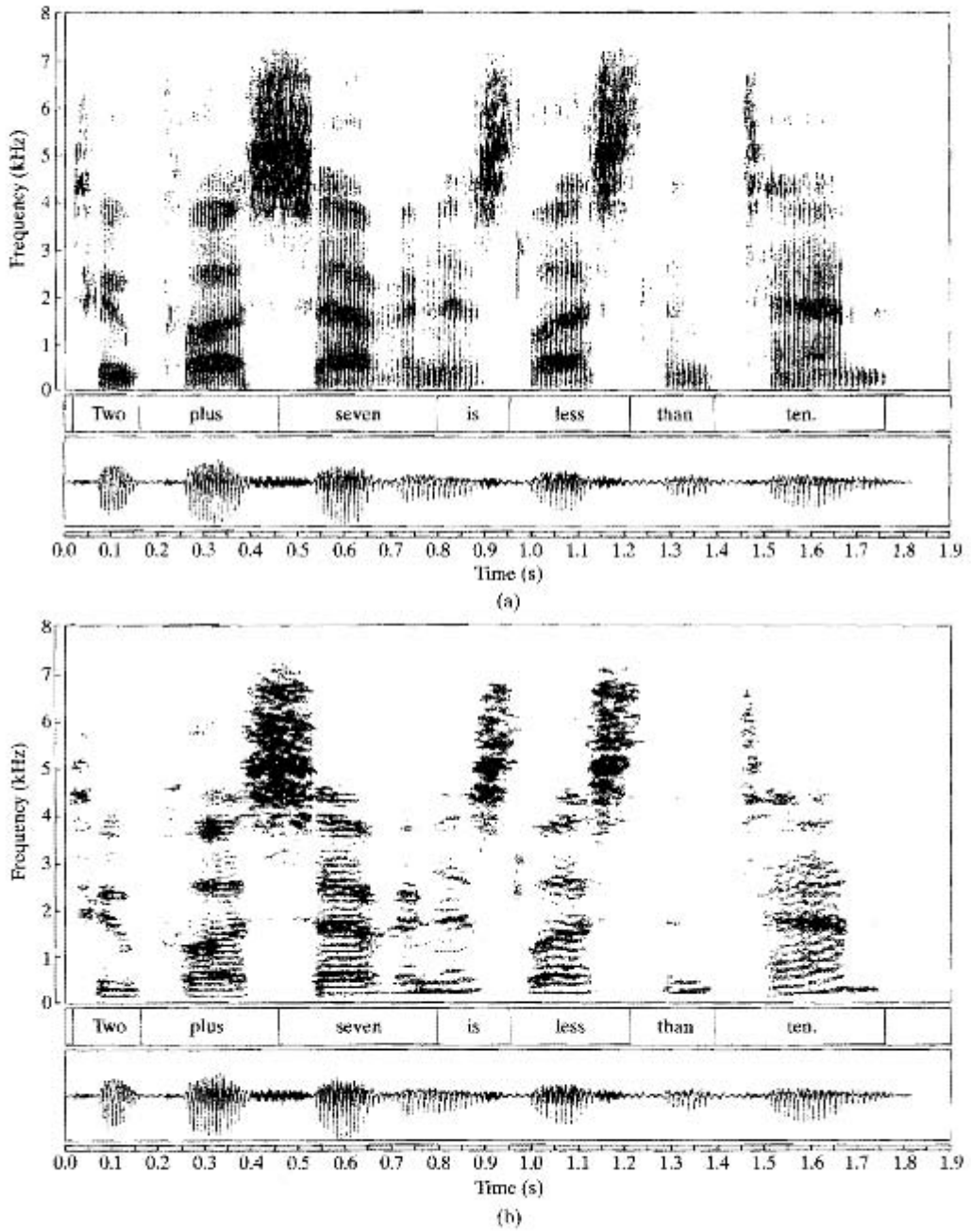


Figure 10.22 (a) Wideband spectrogram of waveform of Figure 10.21. (b) Narrowband spectrogram.

This example only hints at the many reasons that the time-dependent Fourier transform is so important in speech analysis and processing. Indeed, the concept is used directly and indirectly as the basis for acoustic–phonetic analysis and for many fundamental speech-processing applications, such as digital coding, noise and reverberation removal, speech recognition, speaker verification, and speaker identification. For present purposes, our discussion simply serves as an introductory illustration.

10.4.2 Time-Dependent Fourier Analysis of Radar Signals

Another application area in which the time-dependent Fourier transform plays an important role is radar signal analysis. The following are elements of a typical radar system based on the Doppler principle:

- Antennas for transmitting and receiving (often the same).
- A transmitter that generates an appropriate signal at microwave frequencies. In our discussion, we will assume that the signal consists of sinusoidal pulses. While this is often the case, other signals may be used, depending on the specific radar objectives and design.
- A receiver that amplifies and detects echoes of the transmitted pulses that have been reflected from objects illuminated by the antenna.

In such a radar system, the transmitted sinusoidal signal propagates at the speed of light, reflects off the object, and returns at the speed of light to the antenna, thereby undergoing a time delay of the round-trip travel time from the antenna to the object. If we assume that the transmitted signal is a sinusoidal pulse of the form $\cos(\Omega_0 t)$ and the distance from the antenna to the object is $\rho(t)$, then the received signal is a pulse of the form

$$s(t) = \cos[\Omega_0(t - 2\rho(t)/c)], \tag{10.58}$$

where c is the velocity of light. If the object is not moving relative to the antenna, then $\rho(t) = \rho_0$, where ρ_0 is the *range*. Since the time delay between the transmitted and received pulses is $2\rho_0/c$, a measurement of the time delay may be used to estimate the range. If, however, $\rho(t)$ is not constant, the received signal is an angle-modulated sinusoid and the phase difference contains information about both the range and the relative motion of the object with respect to the antenna. Specifically, let us represent the time-varying range in a Taylor's series expansion as

$$\rho(t) = \rho_0 + \dot{\rho}_0 t + \frac{1}{2!}\ddot{\rho}_0 t^2 + \cdots. \tag{10.59}$$

where ρ_0 is the nominal range, $\dot{\rho}_0$ is the velocity, $\ddot{\rho}_0$ is the acceleration, and so on. Assuming that the object moves with constant velocity (i.e., $\ddot{\rho}_0 = 0$), and substituting Eq. (10.59) into Eq. (10.58), we obtain

$$s(t) = \cos[(\Omega_0 - 2\Omega_0\dot{\rho}_0/c)t - 2\Omega_0\rho_0/c]. \tag{10.60}$$

In this case, the frequency of the received signal differs from the frequency of the transmitted signal by the *Doppler frequency*, defined as

$$\Omega_d = -2\Omega_0\dot{\rho}_0/c. \tag{10.61}$$

Thus, the time delay can still be used to estimate the range, and we can determine the speed of the object relative to the antenna if we can determine the Doppler frequency.

In a practical setting, the received signal is generally very weak, and thus a noise term should be added to Eq. (10.60). We will neglect the effects of noise in the simple analysis of this section. Also, in most radar systems, the signal of Eq. (10.60) would be frequency shifted to a lower nominal frequency in the detection process. However, the Doppler shift will still satisfy Eq. (10.61), even if $s(t)$ is demodulated to a lower center frequency.

To apply time-dependent Fourier analysis to such signals, we first bandlimit the signal to a frequency band that includes the expected Doppler frequency shifts and then sample the resulting signal with an appropriate sampling period T, thereby obtaining a discrete-time signal of the form

$$x[n] = \cos[(\omega_0 - 2\omega_0\dot{\rho}_0/c)n - 2\omega_0\rho_0/c], \tag{10.62}$$

where $\omega_0 = \Omega_0 T$. In many cases, the object's motion would be more complicated than we have assumed, requiring the incorporation of higher order terms from Eq. (10.59) and thereby producing a more complicated angle modulation in the received signal. Another way to represent this more complicated variation of the frequency of the echoes is to use the time-dependent Fourier transform with a window that is short enough, so that the assumption of constant Doppler-shifted frequency is valid across the entire window interval, but not so short as to sacrifice adequate resolution when two or more moving objects create Doppler-shifted return signals that are superimposed at the receiver.

Example 10.12 Time-Dependent Fourier Analysis of Doppler Radar Signals

An example of time-dependent Fourier analysis of Doppler radar signals is shown in Figure 10.23. (See Schaefer, Schafer and Mersereau, 1979.) The radar data had been preprocessed to remove low-velocity Doppler shifts, leaving the variations displayed in the figure. The window for the time-dependent Fourier transform was a Kaiser window with $N = L = 64$ and $\beta = 4$. In the figure, $|X_r[k]|$ is plotted with time as the vertical dimension (increasing upward) and frequency as the horizontal dimension.[8] In this case, the successive DFTs are plotted close together. A hidden-line elimination algorithm is used to create a two-dimensional view of the time-dependent Fourier transform. To the left of the center line is a strong peak that moves in a smooth path through the time-frequency plane. This corresponds to a moving object whose velocity varies in a regular manner. The other broad peaks in the time-dependent Fourier transform are due to noise and spurious returns called *clutter* in radar terminology. An example of motion that might create such a variation of the Doppler frequency is a rocket moving at constant velocity but rotating about its longitudinal axis. A peak moving through the time-dependent Fourier transform might correspond to reflections from a fin on the rocket that is alternately moving toward and then away from the antenna because of the spinning of the rocket. Figure 10.23(b) shows an estimate of the Doppler frequency as a function of time. This estimate was obtained simply by locating the highest peak in each DFT.

[8] The plot shows the negative frequencies on the left of the line through the center of the plot and positive frequencies on the right. This can be achieved by computing the DFT of $(-1)^n x_r[n]$ and noting that the computation effectively shifts the origin of the DFT index to $k = N/2$. Alternatively, the DFT of $x_r[n]$ can be computed and then reindexed.

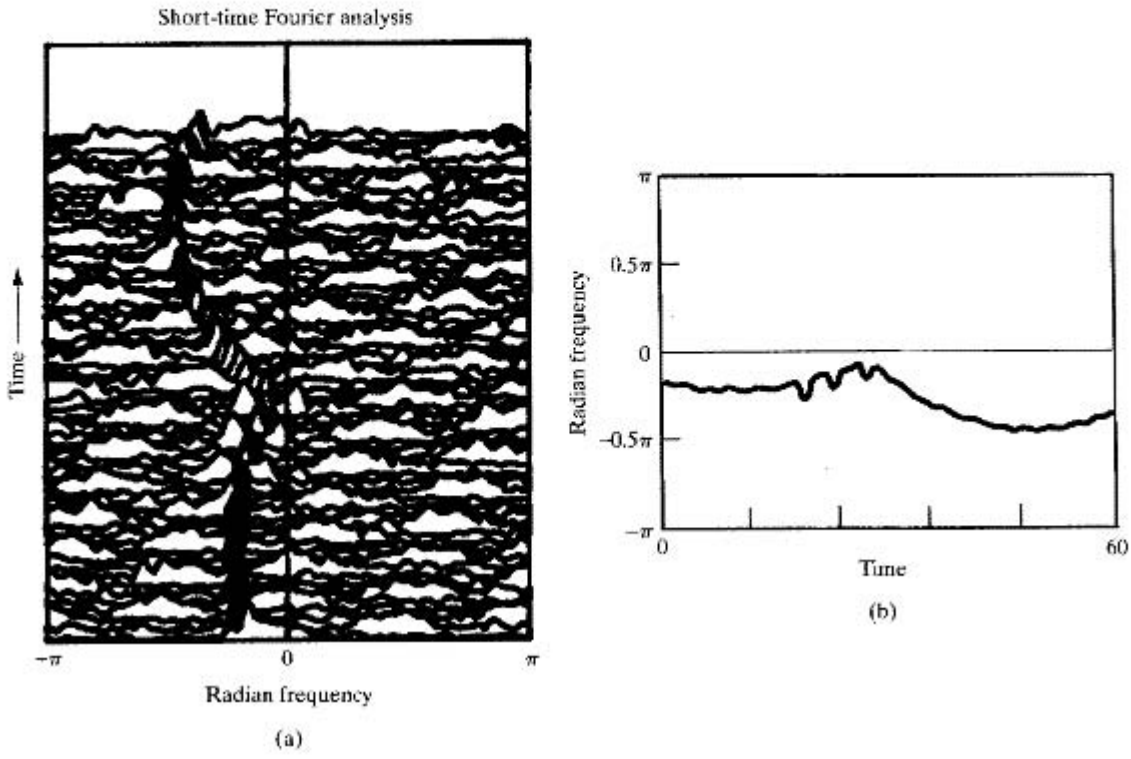


Figure 10.23 Illustration of time-dependent Fourier analysis of Doppler radar signal. (a) Sequence of Fourier transforms of Doppler radar signal. (b) Doppler frequency estimated by picking the largest peak in the time-dependent Fourier transform.

10.5 FOURIER ANALYSIS OF STATIONARY RANDOM SIGNALS: THE PERIODOGRAM

In previous sections, we discussed and illustrated Fourier analysis for sinusoidal signals with stationary (nontime-varying) parameters and for nonstationary signals such as speech and radar. In cases where the signal can be modeled by a sum of sinusoids or a linear system excited by a periodic pulse train, the Fourier transforms of finite-length segments of the signal have a convenient and natural interpretation in terms of Fourier transforms, windowing, and linear system theory. However, more noise-like signals, such as the example of unvoiced speech in Section 10.4.1, are best modeled as random signals.

As we discussed in Section 2.10 and as shown in Appendix A, random processes are often used to model signals when the process that generates the signal is too complex for a reasonable deterministic model. Typically, when the input to an LTI system is modeled as a stationary random process, many of the essential characteristics of the input and output are adequately represented by averages, such as the mean value (dc level), variance (average power), autocorrelation function, or power density spectrum.

Consequently, it is of particular interest to estimate these for a given signal. As discussed in Appendix A, an estimate of the mean value of a stationary random process from a finite-length segment of data is the *sample mean*, defined as

$$\hat{m}_x = \frac{1}{L}\sum_{n=0}^{L-1} x[n]. \tag{10.63}$$

Similarly, an estimate of the variance is the *sample variance*, defined as

$$\hat{\sigma}_x^2 = \frac{1}{L}\sum_{n=0}^{L-1}(x[n] - \hat{m}_x)^2. \tag{10.64}$$

The sample mean and the sample variance, which are themselves random variables, are *unbiased* and *asymptotically unbiased* estimators, respectively; i.e., the expected value of $\hat{m}_x$ is the true mean m_x and the expected value of $\hat{\sigma}_x^2$ approaches the true variance σ_x^2 as L approaches ∞. Furthermore, they are both *consistent* estimators; i.e., they improve with increasing L, since their variances approach zero as L approaches ∞.

In the remainder of this chapter, we study the estimation of the power spectrum[9] of a random signal using the DFT. We will see that there are two basic approaches to estimating the power spectrum. One approach, which we develop in this section, is referred to as *periodogram analysis* and is based on direct Fourier transformation of finite-length segments of the signal. The second approach, developed in Section 10.6, is to first estimate the autocovariance sequence and then compute the Fourier transform of this estimate. In either case, we are typically interested in obtaining unbiased consistent estimators. Unfortunately, the analysis of such estimators is very difficult, and generally, only approximate analyses can be accomplished. Even approximate analyses are beyond the scope of this text, and we refer to the results of such analyses only in a qualitative way. Detailed discussions are given in Blackman and Tukey (1958), Hannan (1960), Jenkins and Watts (1968), Koopmans (1995), Kay and Marple (1981), Marple (1987), Kay (1988) and Stoica and Moses (2005).

10.5.1 The Periodogram

Let us consider the problem of estimating the power density spectrum $P_{ss}(\Omega)$ of a continuous-time signal $s_c(t)$. An intuitive approach to the estimation of the power spectrum is suggested by Figure 10.1 and the associated discussion in Section 10.1. Based on that approach, we now assume that the input signal $s_c(t)$ is a stationary random signal. The antialiasing lowpass filter creates a new stationary random signal whose power spectrum is bandlimited, so that the signal can be sampled without aliasing. Then, $x[n]$ is a stationary discrete-time random signal whose power density spectrum $P_{xx}(\omega)$ is proportional to $P_{ss}(\Omega)$ over the bandwidth of the antialiasing filter; i.e.,

$$P_{xx}(\omega) = \frac{1}{T}P_{ss}\left(\frac{\omega}{T}\right), \qquad |\omega| < \pi, \tag{10.65}$$

where we have assumed that the cutoff frequency of the antialiasing filter is π/T and that T is the sampling period. (See Problem 10.39 for a further consideration of sampling of

[9] The term *power spectrum* is commonly used interchangeably with the more precise term *power density spectrum*.

random signals.) Consequently, a good estimate of $P_{xx}(\omega)$ will provide a useful estimate of $P_{ss}(\Omega)$. The window $w[n]$ in Figure 10.1 selects a finite-length segment (L samples) of $x[n]$, which we denote $v[n]$, the Fourier transform of which is

$$V(e^{j\omega}) = \sum_{n=0}^{L-1} w[n]x[n]e^{-j\omega n}. \tag{10.66}$$

Consider as an estimate of the power spectrum the quantity

$$I(\omega) = \frac{1}{LU}|V(e^{j\omega})|^2, \tag{10.67}$$

where the constant U anticipates a need for normalization to remove bias in the spectrum estimate. When the window $w[n]$ is the rectangular window sequence, this estimator for the power spectrum is called the *periodogram*. If the window is not rectangular, $I(\omega)$ is called the *modified periodogram*. Clearly, the periodogram has some of the basic properties of the power spectrum. It is nonnegative, and for real signals, it is a real and even function of frequency. Furthermore, it can be shown (Problem 10.33) that

$$I(\omega) = \frac{1}{LU} \sum_{m=-(L-1)}^{L-1} c_{vv}[m]e^{-j\omega m}, \tag{10.68}$$

where

$$c_{vv}[m] = \sum_{n=0}^{L-1} x[n]w[n]x[n+m]w[n+m]. \tag{10.69}$$

We note that the sequence $c_{vv}[m]$ is the aperiodic correlation sequence for the finite-length sequence $v[n] = w[n]x[n]$. Consequently, the periodogram is in fact the Fourier transform of the aperiodic correlation of the windowed data sequence.

Explicit computation of the periodogram can be carried out only at discrete frequencies. From Eqs. (10.66) and (10.67), we see that if the DTFT of $w[n]x[n]$ is replaced by its DFT, we will obtain samples at the DFT frequencies $\omega_k = 2\pi k/N$ for $k = 0, 1, \ldots, N-1$. Specifically, samples of the periodogram are given by

$$I[k] = I(\omega_k) = \frac{1}{LU}|V[k]|^2, \tag{10.70}$$

where $V[k]$ is the N-point DFT of $w[n]x[n]$. If we want to choose N to be greater than the window length L, appropriate zero-padding would be applied to the sequence $w[n]x[n]$.

If a random signal has a nonzero mean, its power spectrum has an impulse at zero frequency. If the mean is relatively large, this component will dominate the spectrum estimate, causing low-amplitude, low-frequency components to be obscured by leakage. Therefore, in practice the mean is often estimated using Eq. (10.63), and the resulting estimate is subtracted from the random signal before computing the power spectrum estimate. Although the sample mean is only an approximate estimate of the zero-frequency component, subtracting it from the signal often leads to better estimates at neighboring frequencies.

10.5.2 Properties of the Periodogram

The nature of the periodogram estimate of the power spectrum can be determined by recognizing that, for each value of ω, $I(\omega)$ is a random variable. By computing the mean and variance of $I(\omega)$, we can determine whether the estimate is biased and whether it is consistent.

From Eq. (10.68), the expected value of $I(\omega)$ is

$$\mathcal{E}\{I(\omega)\} = \frac{1}{LU}\sum_{m=-(L-1)}^{L-1}\mathcal{E}\{c_{vv}[m]\}e^{-j\omega m}. \tag{10.71}$$

The expected value of $c_{vv}[m]$ can be expressed as

$$\begin{aligned}\mathcal{E}\{c_{vv}[m]\} &= \sum_{n=0}^{L-1}\mathcal{E}\{x[n]w[n]x[n+m]w[n+m]\}\\ &= \sum_{n=0}^{L-1}w[n]w[n+m]\mathcal{E}\{x[n]x[n+m]\}.\end{aligned} \tag{10.72}$$

Since we are assuming that $x[n]$ is stationary,

$$\mathcal{E}\{x[n]x[n+m]\} = \phi_{xx}[m], \tag{10.73}$$

and Eq. (10.72) can then be rewritten as

$$\mathcal{E}\{c_{vv}[m]\} = c_{ww}[m]\phi_{xx}[m], \tag{10.74}$$

where $c_{ww}[m]$ is the aperiodic autocorrelation of the window, i.e.,

$$c_{ww}[m] = \sum_{n=0}^{L-1}w[n]w[n+m]. \tag{10.75}$$

That is, the mean of the aperiodic autocorrelation of the windowed signal is equal to the aperiodic autocorrelation of the window multiplied by the true autocorrelation function; i.e., *in an average sense*, the autocorrelation function of the data window appears as a window on the true autocorrelation function.

From Eq. (10.71), Eq. (10.74), and the modulation–windowing property of Fourier transforms (Section 2.9.7), it follows that

$$\mathcal{E}\{I(\omega)\} = \frac{1}{2\pi LU}\int_{-\pi}^{\pi}P_{xx}(\theta)C_{ww}(e^{j(\omega-\theta)})d\theta, \tag{10.76}$$

where $C_{ww}(e^{j\omega})$ is the Fourier transform of the aperiodic autocorrelation of the window, i.e.,

$$C_{ww}(e^{j\omega}) = |W(e^{j\omega})|^2. \tag{10.77}$$

According to Eq. (10.76), both the periodogram and the modified periodogram are biased estimates of the power spectrum, since $\mathcal{E}\{I(\omega)\}$ is not equal to $P_{xx}(\omega)$. Indeed, we see that the bias arises as a result of convolution of the true power spectrum with the Fourier transform of the aperiodic autocorrelation of the data window. If we increase the window length, we expect that $W(e^{j\omega})$ should become more concentrated around $\omega = 0$, and thus $C_{ww}(e^{j\omega})$ should look increasingly like a periodic impulse train. If the scale factor $1/(LU)$ is correctly chosen, then $\mathcal{E}\{I(\omega)\}$ should approach $P_{xx}(\omega)$ as

$C_{ww}(e^{j\omega})$ approaches a periodic impulse train. The scale can be adjusted by choosing the normalizing constant U so that

$$\frac{1}{2\pi LU}\int_{-\pi}^{\pi}|W(e^{j\omega})|^2 d\omega = \frac{1}{LU}\sum_{n=0}^{L-1}(w[n])^2 = 1, \tag{10.78}$$

or

$$U = \frac{1}{L}\sum_{n=0}^{L-1}(w[n])^2. \tag{10.79}$$

For the rectangular window, we should choose $U = 1$, while other data windows would require a value of $0 < U < 1$ if $w[n]$ is normalized to a maximum value of 1. Alternatively, the normalization can be absorbed into the amplitude of $w[n]$. Therefore, if properly normalized, the periodogram and modified periodogram are both asymptotically unbiased; i.e., the bias approaches zero as the window length increases.

To examine whether the periodogram is a consistent estimate or becomes a consistent estimate as the window length increases, it is necessary to consider the behavior of the variance of the periodogram. An expression for the variance of the periodogram is very difficult to obtain even in the simplest cases. However, it has been shown (see Jenkins and Watts, 1968) that over a wide range of conditions, as the window length increases,

$$\text{var}[I(\omega)] \simeq P_{xx}^2(\omega). \tag{10.80}$$

That is, the variance of the periodogram estimate is approximately the same size as the square of the power spectrum that we are estimating. Therefore, since the variance does not asymptotically approach zero with increasing window length, the periodogram is not a consistent estimate.

The properties of the periodogram estimate of the power spectrum just discussed are illustrated in Figure 10.24, which shows periodogram estimates of white noise using rectangular windows of lengths $L = 16, 64, 256$, and 1024. The sequence $x[n]$

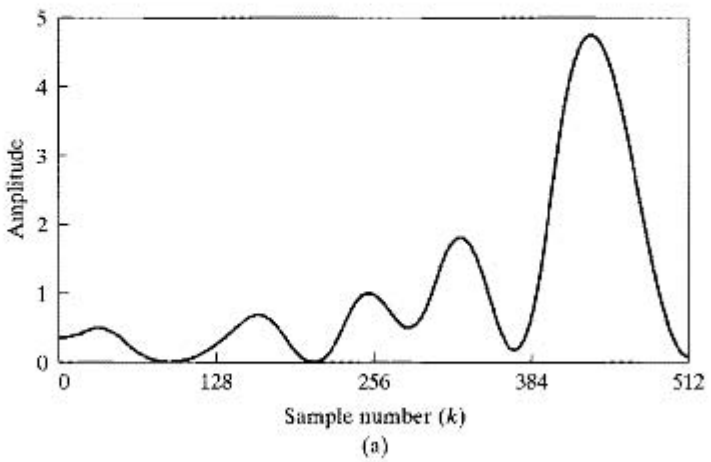
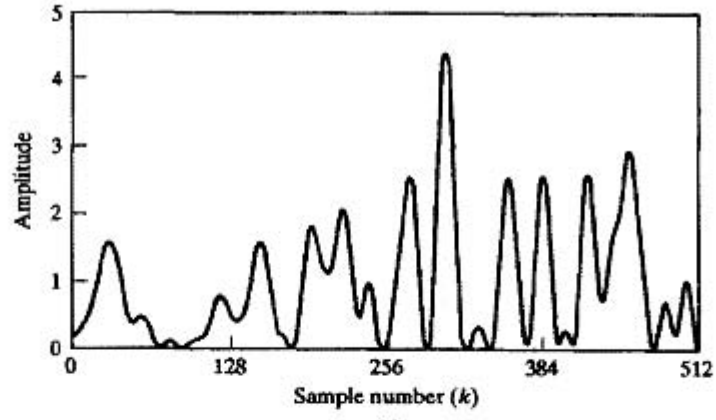
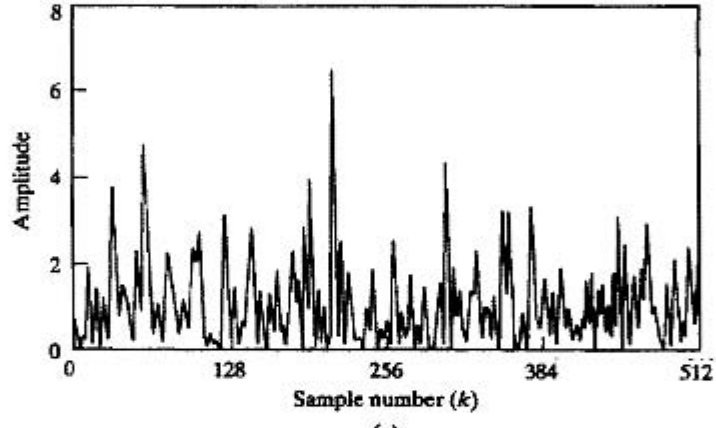


Figure 10.24 Periodograms of pseudorandom white-noise sequence. (a) Window length $L = 16$ and DFT length $N = 1024$.





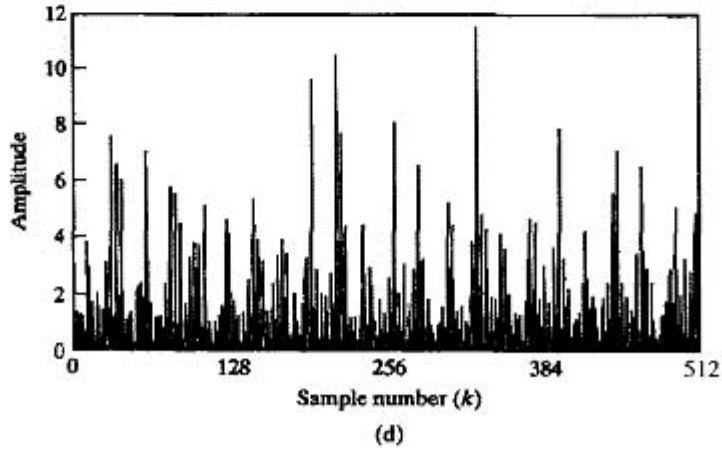


Figure 10.24 (*continued*) (b) $L = 64$ and $N = 1024$. (c) $L = 256$ and $N = 1024$. (d) $L = 1024$ and $N = 1024$.

was obtained from a pseudorandom-number generator whose output was scaled so that $|x[n]| \leq \sqrt{3}$. A good random-number generator produces a uniform distribution of amplitudes, and the measured sample-to-sample correlation is small. Thus, the power spectrum of the output of the random-number generator could be modeled in this case by $P_{xx}(\omega) = \sigma_x^2 = 1$ for all ω. For each of the four rectangular windows, the periodogram was computed with normalizing constant $U = 1$ and at frequencies $\omega_k = 2\pi k/N$ for $N = 1024$ using the DFT. That is,

$$I[k] = I(\omega_k) = \frac{1}{L}|V[k]|^2 = \frac{1}{L}\left|\sum_{n=0}^{L-1} w[n]x[n]e^{-j(2\pi/N)kn}\right|^2. \tag{10.81}$$

In Figure 10.24, the DFT values are connected by straight lines for purposes of display. Recall that $I(\omega)$ is real and an even function of ω, so we only need to plot $I[k]$ for $0 \leq k \leq N/2$ corresponding to $0 \leq \omega \leq \pi$. Note that the spectrum estimate fluctuates more rapidly as the window length L increases. This behavior can be understood by recalling that, although we view the periodogram method as a direct computation of the spectrum estimate, we have seen that the underlying correlation estimate of Eq. (10.69) is, in effect, Fourier transformed to obtain the periodogram. Figure 10.25 illustrates a windowed sequence, $x[n]w[n]$, and a shifted version, $x[n+m]w[n+m]$, as required in Eq. (10.69). From this figure, we see that $(L-m)$ signal values are involved in computing a particular correlation lag value $c_{vv}[m]$. Thus, when m is close to L, only a few values of $x[n]$ are involved in the computation, and we expect that the estimate of the correlation sequence will be considerably more inaccurate for these values of m and consequently will also show considerable variation between adjacent values of m. On the other hand, when m is small, many more samples are involved, and the variability of $c_{vv}[m]$ with m should not be as great. The variability at large values of m manifests itself in the Fourier transform as fluctuations at all frequencies, and thus, for large L, the periodogram estimate tends to vary rapidly with frequency. Indeed, it can be shown (see Jenkins

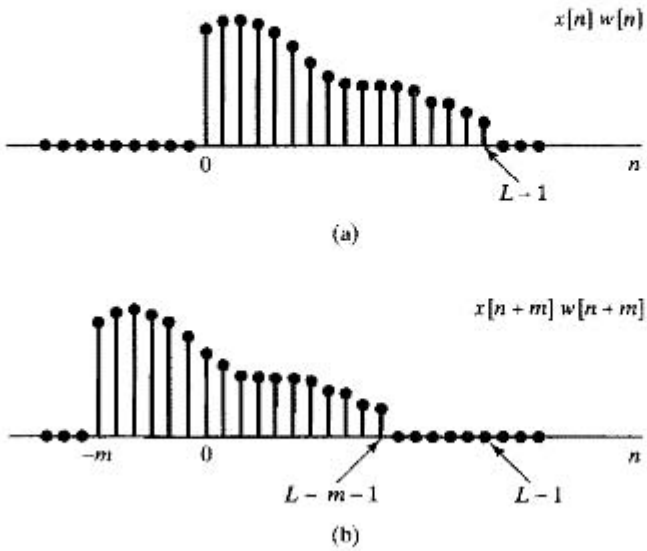


Figure 10.25 Illustration of sequences involved in Eq. (10.69). (a) A finite-length sequence. (b) Shifted sequence for $m > 0$.

and Watts, 1968) that if $N = L$, the periodogram estimates at the DFT frequencies $2\pi k/N$ become uncorrelated. Since, as N increases, the DFT frequencies get closer together, this behavior is inconsistent with our goal of obtaining a good estimate of the power spectrum. We would prefer to obtain a smooth spectrum estimate without random variations resulting from the estimation process. This can be accomplished by averaging multiple independent periodogram estimates to reduce the fluctuations.

10.5.3 Periodogram Averaging

The averaging of periodograms in spectrum estimation was first studied extensively by Bartlett (1953); later, after fast algorithms for computing the DFT were developed, Welch (1970) combined these computational algorithms with the use of a data window $w[n]$ to develop the method of averaging modified periodograms. In periodogram averaging, a data sequence $x[n], 0 \le n \le Q-1$, is divided into segments of length-L samples, with a window of length L applied to each; i.e., we form the segments

$$x_r[n] = x[rR + n]w[n], \qquad 0 \le n \le L-1. \tag{10.82}$$

If $R < L$ the segments overlap, and for $R = L$ the segments are contiguous. Note that Q denotes the length of the available data. The total number of segments depends on the values of, and relationship among, R, L, and Q. Specifically, there will be K full-length segments, where K is the largest integer for which $(K-1)R + (L-1) \le Q-1$. The periodogram of the r^{th} segment is

$$I_r(\omega) = \frac{1}{LU}|X_r(e^{j\omega})|^2, \tag{10.83}$$

where $X_r(e^{j\omega})$ is the DTFT of $x_r[n]$. Each $I_r(\omega)$ has the properties of a periodogram, as described previously. Periodogram averaging consists of averaging the K periodogram estimates $I_r(\omega)$; i.e., we form the time-averaged periodogram defined as

$$\bar{I}(\omega) = \frac{1}{K}\sum_{r=0}^{K-1} I_r(\omega). \tag{10.84}$$

To examine the bias and variance of $\bar{I}(\omega)$, let us take $L = R$, so that the segments do not overlap, and assume that $\phi_{xx}[m]$ is small for $m > L$; i.e., signal samples more than L apart are approximately uncorrelated. If we assume that the periodograms $I_r(\omega)$ are identically distributed independent random variables, then the expected value of $\bar{I}(\omega)$ is

$$\mathcal{E}\{\bar{I}(\omega)\} = \frac{1}{K}\sum_{r=0}^{K-1} \mathcal{E}\{I_r(\omega)\}, \tag{10.85}$$

or, since we assume that the periodograms are independent and identically distributed,

$$\mathcal{E}\{\bar{I}(\omega)\} = \mathcal{E}\{I_r(\omega)\} \qquad \text{for any } r. \tag{10.86}$$

From Eq. (10.76), it follows that

$$\mathcal{E}\{\bar{I}(\omega)\} = \mathcal{E}\{I_r(\omega)\} = \frac{1}{2\pi LU}\int_{-\pi}^{\pi} P_{xx}(\theta)C_{ww}(e^{j(\omega-\theta)})d\theta, \tag{10.87}$$

where L is the window length. When the window $w[n]$ is the rectangular window, the method of averaging periodograms is called *Bartlett's procedure*, and in this case it can be shown that

$$c_{ww}[m] = \begin{cases} L - |m|, & |m| \leq (L-1), \\ 0 & \text{otherwise}, \end{cases} \tag{10.88}$$

and, therefore,

$$C_{ww}(e^{j\omega}) = \left(\frac{\sin(\omega L/2)}{\sin(\omega/2)} \right)^2. \tag{10.89}$$

That is, the expected value of the average periodogram spectrum estimate is the convolution of the true power spectrum with the Fourier transform of the triangular sequence $c_{ww}[n]$ that results as the autocorrelation of the rectangular window. Thus, the average periodogram is also a biased estimate of the power spectrum.

To examine the variance, we use the fact that, in general, the variance of the average of K independent identically distributed random variables is $1/K$ times the variance of each individual random variable. (See Bertsekas and Tsitsiklis, 2008.) Therefore, the variance of the average periodogram is

$$\text{var}[\bar{I}(\omega)] = \frac{1}{K} \text{var}[I_r(\omega)], \tag{10.90}$$

or, with Eq. (10.80), it follows that

$$\text{var}[\bar{I}(\omega)] \simeq \frac{1}{K} P_{xx}^2(\omega). \tag{10.91}$$

Consequently, the variance of $\bar{I}(\omega)$ is inversely proportional to the number of periodograms averaged, and as K increases, the variance approaches zero.

From Eq. (10.89), we see that as L, the length of the segment $x_r[n]$, increases, the main lobe of $C_{ww}(e^{j\omega})$ decreases in width, and consequently, from Eq. (10.87), $\mathcal{E}\{\bar{I}(\omega)\}$ more closely approximates $P_{xx}(\omega)$. However, for fixed total data length Q, the total number of segments (assuming that $L = R$) is Q/L; therefore, as L increases, K decreases. Correspondingly, from Eq. (10.91), the variance of $\bar{I}(\omega)$ will increase. Thus, as is typical in statistical estimation problems, for a fixed data length there is a trade off between bias and variance. However, as the data length Q increases, both L and K can be allowed to increase, so that as Q approaches ∞, the bias and variance of $\bar{I}(\omega)$ can approach zero. Consequently, periodogram averaging provides an asymptotically unbiased, consistent estimate of $P_{xx}(\omega)$.

The preceding discussion assumed that nonoverlapping rectangular windows were used in computing the time-dependent periodograms. Welch (1970) showed that if a different window shape is used, the variance of the average periodogram still behaves, as in Eq. (10.91). Welch also considered the case of overlapping windows and showed that if the overlap is one-half the window length, the variance is further reduced by almost a factor of 2, due to the doubling of the number of sections. Greater overlap does not continue to reduce the variance, because the segments become decreasingly independent as the overlap increases.

10.5.4 Computation of Average Periodograms Using the DFT

As with the periodogram, the average periodogram can be explicitly evaluated only at a discrete set of frequencies. Because of the availability of the FFT algorithms for computing the DFT, a particularly convenient and widely used choice is the set of frequencies $\omega_k = 2\pi k/N$ for an appropriate choice of N. From Eq. (10.84), we see that if the DFT of $x_r[n]$ is substituted for the Fourier transform of $x_r[n]$ in Eq. (10.83), we obtain samples of $\bar{I}(\omega)$ at the DFT frequencies $\omega_k = 2\pi k/N$ for $k = 0, 1, \ldots, N-1$. Specifically, with $X_r[k]$ denoting the DFT of $x_r[n]$,

$$I_r[k] = I_r(\omega_k) = \frac{1}{LU}|X_r[k]|^2, \tag{10.92a}$$

$$\bar{I}[k] = \bar{I}(\omega_k) = \frac{1}{K}\sum_{r=0}^{K-1} I_r[k]. \tag{10.92b}$$

It is worthwhile to note the relationship between periodogram averaging and the time-dependent Fourier transform as discussed in detail in Section 10.3. Equation (10.92a) shows that, except for the introduction of the normalizing constant $1/(LU)$, each individual periodogram is simply the magnitude-squared of the time-dependent Fourier transform at time rR and frequency $2\pi k/N$. Thus, for each frequency index k, the average power spectrum estimate at frequency corresponding to k is the time average of the time-sampled time-dependent Fourier transform. This can be visualized by considering the spectrograms in Figure 10.22. The value $\bar{I}[k]$ is simply the average along a horizontal line at frequency $2\pi k/N$ (or $2\pi k/(NT)$ in analog frequency).[10] Averaging the wideband spectrogram implies that the resulting power spectrum estimate will be smooth when considered as a function of frequency, while the narrowband condition corresponds to longer time windows and thus, less smoothness in frequency.

We denote $I_r(2\pi k/N)$ as the sequence $I_r[k]$ and $\bar{I}(2\pi k/N)$ as the sequence $\bar{I}[k]$. According to Eqs. (10.92a) and (10.92b), the average periodogram estimate of the power spectrum is computed at N equally spaced frequencies by averaging the magnitude of the DFTs of the windowed data segments with the normalizing factor LU. This method of power spectrum estimation provides a very convenient framework within which to trade off between resolution and variance of the spectrum estimate. It is particularly simple and efficient to implement using the FFT algorithms discussed in Chapter 9. An important advantage of the method over those to be discussed in Section 10.6 is that the spectrum estimate is always nonnegative.

10.5.5 An Example of Periodogram Analysis

Power spectrum analysis is a valuable tool for modeling signals, and it also can be used to detect signals, particularly when it comes to finding hidden periodicities in sampled

[10]Note that the spectrogram is normally computed such that the windowed segments overlap considerably as r varies, while in periodogram averaging R is normally equal to the window length or half the window length.

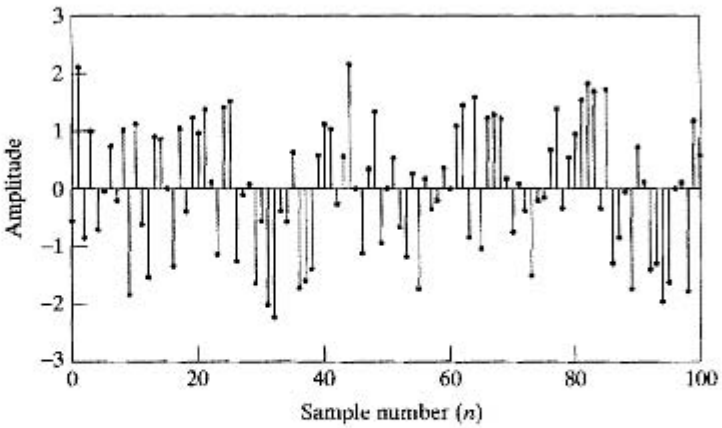


Figure 10.26 Cosine sequence with white noise, as in Eq. (10.93).

signals. As an example of this type of application of the average periodogram method, consider the sequence

$$x[n] = A\cos(\omega_0 n + \theta) + e[n], \tag{10.93}$$

where θ is a random variable uniformly distributed between 0 and 2π, is independent of $e[n]$, and $e[n]$ is a zero-mean white-noise sequence that has a constant power spectrum; i.e., $P_{ee}(\omega) = \sigma_e^2$ for all ω. In signal models of this form, the cosine is generally the desired component and $e[n]$ is an undesired noise component. Often, in practical signal detection problems, we are interested in the case for which the power in the cosine signal is small compared with the noise power. It can be shown (see Problem 10.40) that over the base period of frequency $|\omega| \leq \pi$, the power spectrum for this signal is

$$P_{xx}(\omega) = \frac{A^2\pi}{2}[\delta(\omega - \omega_0) + \delta(\omega + \omega_0)] + \sigma_e^2 \qquad \text{for } |\omega| \leq \pi. \tag{10.94}$$

From Eqs. (10.87) and (10.94), it follows that the expected value of the average periodogram is

$$\mathcal{E}\{\bar{I}(\omega)\} = \frac{A^2}{4LU}[C_{WW}(e^{j(\omega-\omega_0)}) + C_{WW}(e^{j(\omega+\omega_0)})] + \sigma_e^2. \tag{10.95}$$

Figures 10.26 and 10.27 show the use of the averaging method for a signal of the form of Eq. (10.93), with $A = 0.5$, $\omega_0 = 2\pi/21$, and random phase $0 \leq \theta < 2\pi$. The noise was uniformly distributed in amplitude such that $-\sqrt{3} < e[n] \leq \sqrt{3}$. Therefore, it is easily shown that $\sigma_e^2 = 1$. The mean of the noise component is zero. Figure 10.26 shows 101 samples of the sequence $x[n]$. Since the noise component $e[n]$ has a maximum amplitude $\sqrt{3}$, the cosine component in the sequence $x[n]$ (having period 21) is not visually apparent.

Figure 10.27 shows average periodogram estimates of the power spectrum for rectangular windows with amplitude 1, so that $U = 1$, and of lengths $L = 1024, 256, 64$, and 16, with the total record length $Q = 1024$ in all cases. Except for Figure 10.27(a), the windows overlap by one-half the window length. Figure 10.27(a) is the periodogram of the entire record, and Figures 10.27(b), (c), and (d) show the average periodogram

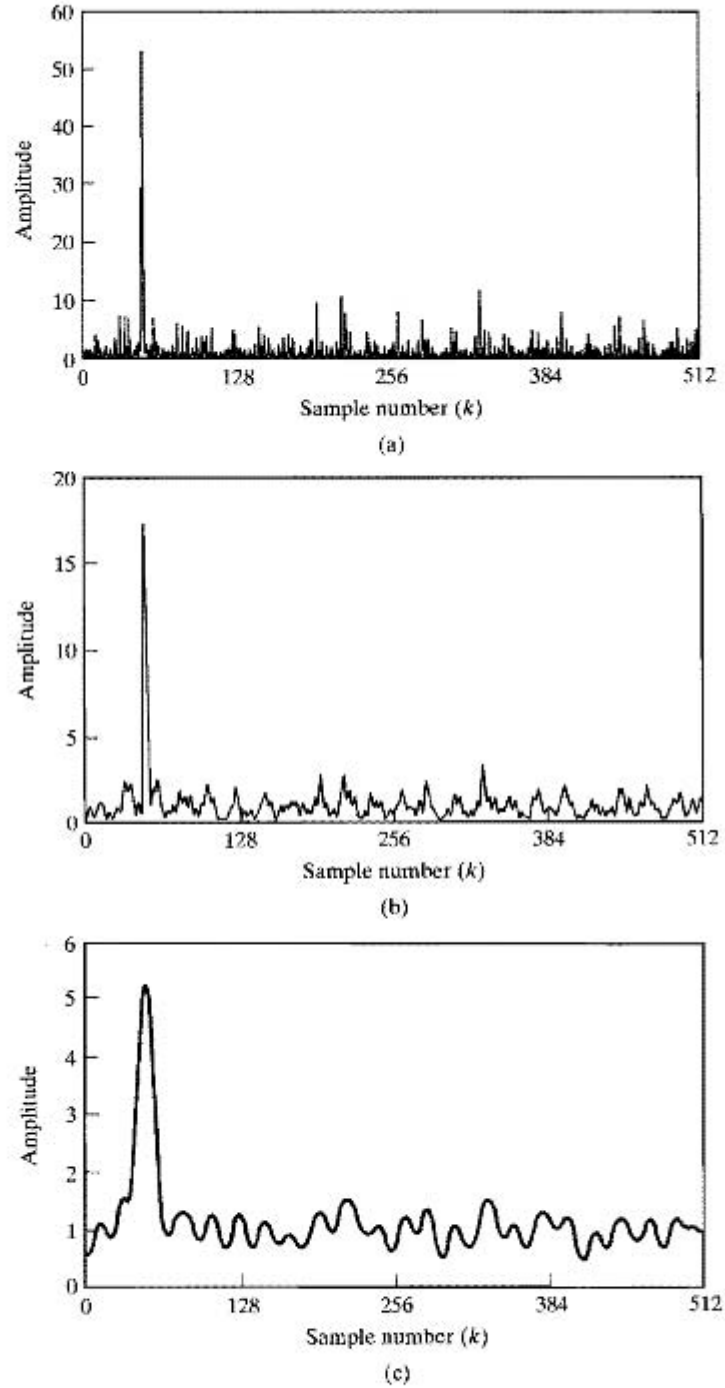


Figure 10.27 Example of average periodogram for signal of length $Q = 1024$. (a) Periodogram for window length $L = Q = 1024$ (only one segment). (b) $K = 7$ and $L = 256$ (overlap by $L/2$). (c) $K = 31$ and $L = 64$.

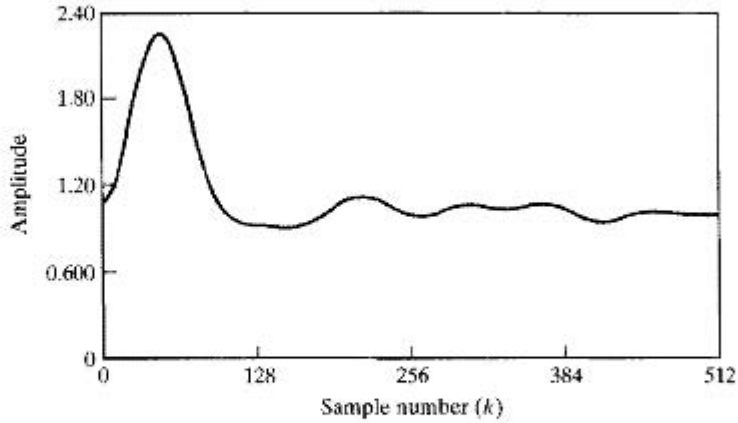


Figure 10.27 (*continued*) (d) $K = 127$ and $L = 16$.

for $K = 7$, 31, and 127 segments, respectively. In all cases, the average periodogram was evaluated using 1024-point DFTs at frequencies $\omega_k = 2\pi k/1024$. (For window lengths $L < 1024$, the windowed sequence was augmented with zero-samples before computing the DFT.) Therefore, the frequency $\omega_0 = 2\pi/21$ lies between DFT frequencies $\omega_{48} = 2\pi 48/1024$ and $\omega_{49} = 2\pi 49/1024$.

In using such estimates of the power spectrum to detect the presence and/or the frequency of the cosine component, we might search for the highest peaks in the spectrum estimate and compare their size with that of the surrounding spectrum values. From Eqs. (10.89) and (10.95), the expected value of the average periodogram at the frequency ω_0 is

$$\mathcal{E}\{\bar{I}(\omega_0)\} = \frac{A^2 L}{4} + \sigma_e^2. \tag{10.96}$$

Thus, if the peak due to the cosine component is to stand out against the variability of the average periodogram, then in this special case, we must choose L so that $A^2 L/4 \gg \sigma_e^2$. This is illustrated by Figure 10.27(a), where L is as large as it can be for the record length Q. We see that $L = 1024$ gives a very narrow main lobe of the Fourier transform of the autocorrelation of the rectangular window, so it would be possible to resolve very closely spaced sinusoidal signals. Note that for the parameters of this example ($A = 0.5$, $\sigma_e^2 = 1$) and with $L = 1024$, the peak amplitude in the periodogram at frequency $2\pi/21$ is close, but not equal, to the expected value of 65. We also observe additional peaks in the periodogram with amplitudes greater than 10. Clearly, if the cosine amplitude A had been smaller by only a factor of 2, it is possible that its peak would have been confused with the inherent variability of the periodogram.

We have seen that the only sure way to reduce the variance of the spectrum estimate is to increase the record length of the signal. This is not always possible, and even if it is possible, longer records require more processing. We can reduce the variability

of the estimate while keeping the record length constant if we use shorter windows and average over more sections. The cost of doing this is illustrated by parts (b), (c), and (d) of Figure 10.27. Note that as more sections are used, the variance of the spectrum estimate decreases, but in accordance with Eq. (10.96), so does the amplitude of the peak as a result of the cosine. Thus, we again face a trade-off. That the shorter windows reduce variability is clear, especially if we compare the high-frequency regions away from the peak in parts (a), (b) and (c) of Figure 10.27. Recall that the idealized power spectrum of the model for the pseudorandom-noise generator is a constant ($\sigma_e^2 = 1$) for all frequencies. In Figure 10.27(a) there are peaks as high as about 10 when the true spectrum is 1. In Figure 10.27(b), the variation away from 1 is less than about 3, and in Figure 10.27(c), the variation around 1 is less than 0.5. However, shorter windows also reduce the peak amplitude of any narrowband component, and they also degrade our ability to resolve closely spaced sinusoids. This reduction in peak amplitude is also clear from Figure 10.27. Again, if we were to reduce A by a factor of 2 in Figure 10.27(b), the peak height would be approximately 4, which is not much different from many of the other peaks in the high-frequency region. In Figure 10.27(c) a reduction of A by a factor of 2 would make the peak approximately 1.25, which would be indistinguishable from the other ripples in the estimate. In Figure 10.27(d), the window is very short, and thus the fluctuations of the spectrum estimate are greatly reduced, but the spectrum peak due to the cosine is very broad and barely above the noise even for $A = 0.5$. If the length were any smaller, spectral leakage from the negative-frequency component would cause there to be no distinct peak in the low-frequency region.

This example confirms that the average periodogram provides a straightforward method of trading off between spectral resolution and reduction of the variance of the spectrum estimate. Although the theme of the example was the detection of a sinusoid in noise, the average periodogram could also be used in signal modeling. The spectrum estimates of Figure 10.27 clearly suggest a signal model of the form of Eq. (10.93), and most of the parameters of the model could be estimated from the average periodogram power spectrum estimate.

10.6 SPECTRUM ANALYSIS OF RANDOM SIGNALS USING ESTIMATES OF THE AUTOCORRELATION SEQUENCE

In the previous section, we considered the periodogram as a direct estimate of the power spectrum of a random signal. The periodogram or the average periodogram is a direct estimate in the sense that it is obtained directly by Fourier transformation of the samples of the random signal. Another approach, based on the fact that the power density spectrum is the Fourier transform of the autocorrelation function, is to first obtain an estimate of the autocorrelation function $\hat{\phi}_{xx}[m]$ for a finite set of lag values $-M \leq m \leq M$, and then apply a window $w_c[m]$ before computing the DTFT of this estimate. This approach to power spectrum estimation is often referred to as the *Blackman–Tukey method*. (See Blackman and Tukey, 1958.) In this section, we explore some of the important facets of this approach and show how the DFT can be used to implement it.

Let us assume, as before, that we are given a finite record of a random signal $x[n]$. This sequence is denoted

$$v[n] = \begin{cases} x[n] & \text{for } 0 \le n \le Q-1, \\ 0 & \text{otherwise.} \end{cases} \tag{10.97}$$

Consider an estimate of the autocorrelation sequence as

$$\hat{\phi}_{xx}[m] = \frac{1}{Q} c_{vv}[m], \tag{10.98a}$$

where, since $c_{vv}[-m] = c_{vv}[m]$,

$$c_{vv}[m] = \sum_{n=0}^{Q-1} v[n]v[n+m] = \begin{cases} \displaystyle\sum_{n=0}^{Q-|m|-1} x[n]x[n+|m|], & |m| \le Q-1, \\ 0 & \text{otherwise,} \end{cases} \tag{10.98b}$$

corresponding to the aperiodic correlation of a rectangularly windowed segment of $x[n]$ of length Q.

To determine the properties of this estimate of the autocorrelation sequence, we consider the mean and variance of the random variable $\hat{\phi}_{xx}[m]$. From Eqs. (10.98a) and (10.98b), it follows that

$$\mathcal{E}\{\hat{\phi}_{xx}[m]\} = \frac{1}{Q} \sum_{n=0}^{Q-|m|-1} \mathcal{E}\{x[n]x[n+|m|]\} = \frac{1}{Q} \sum_{n=0}^{Q-|m|-1} \phi_{xx}[m], \tag{10.99}$$

and since $\phi_{xx}[m]$ does not depend on n for a stationary random process,

$$\mathcal{E}\{\hat{\phi}_{xx}[m]\} = \begin{cases} \left(\dfrac{Q-|m|}{Q}\right)\phi_{xx}[m], & |m| \le Q-1, \\ 0 & \text{otherwise.} \end{cases} \tag{10.100}$$

From Eq. (10.100), we see that $\hat{\phi}_{xx}[m]$ is a biased estimate of $\phi_{xx}[m]$, since $\mathcal{E}\{\hat{\phi}_{xx}[m]\}$ is not equal to $\phi_{xx}[m]$, but the bias is small if $|m| \ll Q$. We see also that an unbiased estimator of the autocorrelation sequence for $|m| \le Q-1$ is

$$\check{\phi}_{xx}[m] = \left(\frac{1}{Q-|m|}\right) c_{vv}[m]; \tag{10.101}$$

i.e., the estimator is unbiased if we divide by the number of nonzero terms in the sum of lagged products involved in computing each value of $c_{vv}[m]$, rather than by the total number of samples in the data record.

The variance of the autocorrelation function estimates is difficult to compute, even with simplifying assumptions. However, approximate formulas for the variance of both $\hat{\phi}_{xx}[m]$ and $\check{\phi}_{xx}[m]$ can be found in Jenkins and Watts (1968). For our purposes here, it is sufficient to observe from Eq. (10.98b) that as $|m|$ approaches Q, fewer and fewer samples of $x[n]$ are involved in the computation of the autocorrelation estimate; therefore, the variance of the autocorrelation estimate can be expected to increase with increasing $|m|$. In the case of the periodogram, this increased variance affects the spectrum estimate at all frequencies, because all the autocorrelation lag values are implicitly involved in the computation of the periodogram. However, by explicitly computing the autocorrelation estimate, we are free to choose which correlation lag

values to include when estimating the power spectrum. Thus, we define the power spectrum estimate

$$S(\omega) = \sum_{m=-(M-1)}^{M-1} \hat{\phi}_{xx}[m] w_c[m] e^{-j\omega m}, \tag{10.102}$$

where $w_c[m]$ is a symmetric window of length $(2M-1)$ applied to the estimated autocorrelation function. We require that the product of the autocorrelation sequence and the window be an even sequence when $x[n]$ is real, so that the power spectrum estimate will be a real, even function of ω. Therefore, the correlation window must be an even sequence. By limiting the length of the correlation window so that $M \ll Q$, we include only autocorrelation estimates for which the variance is low.

The mechanism by which windowing the autocorrelation sequence reduces the variance of the power spectrum estimate is best understood in the frequency domain. From Eqs. (10.68), (10.69), and (10.98b), it follows that, with $w[n] = 1$ for $0 \le n \le (Q-1)$, i.e., a rectangular window, the periodogram is the Fourier transform of the autocorrelation estimate $\hat{\phi}_{xx}[m]$; i.e.,

$$\hat{\phi}_{xx}[m] = \frac{1}{Q} c_{vv}[m] \overset{\mathcal{F}}{\longleftrightarrow} \frac{1}{Q} |V(e^{j\omega})|^2 = I(\omega). \tag{10.103}$$

Therefore, from Eq. (10.102), the spectrum estimate obtained by windowing of $\hat{\phi}_{xx}[m]$ is the convolution

$$S(\omega) = \frac{1}{2\pi} \int_{-\pi}^{\pi} I(\theta) W_c(e^{j(\omega-\theta)}) d\theta. \tag{10.104}$$

From Eq. (10.104), we see that the effect of applying the window $w_c[m]$ to the autocorrelation estimate is to convolve the periodogram with the Fourier transform of the autocorrelation window. This will smooth the rapid fluctuations of the periodogram spectrum estimate. The shorter the correlation window, the smoother the spectrum estimate will be, and vice versa.

The power spectrum $P_{xx}(\omega)$ is a nonnegative function of frequency, and the periodogram and the average periodogram automatically have this property by definition. In contrast, from Eq. (10.104), it is evident that nonnegativity is not guaranteed for $S(\omega)$, unless we impose the further condition that

$$W_c(e^{j\omega}) \ge 0 \qquad \text{for } -\pi < \omega \le \pi. \tag{10.105}$$

This condition is satisfied by the Fourier transform of the triangular (Bartlett) window, but it is not satisfied by the rectangular, Hanning, Hamming, or Kaiser windows. Therefore, although these latter windows have smaller side lobes than the triangular window, spectral leakage may cause negative spectrum estimates in low-level regions of the spectrum.

The expected value of the smoothed periodogram is

$$\begin{aligned} \mathcal{E}\{S(\omega)\} &= \sum_{m=-(M-1)}^{M-1} \mathcal{E}\{\hat{\phi}_{xx}[m]\} w_c[m] e^{-j\omega m} \\ &= \sum_{m=-(M-1)}^{M-1} \phi_{xx}[m] \left(\frac{Q-|m|}{Q}\right) w_c[m] e^{-j\omega m}. \end{aligned} \tag{10.106}$$

If $Q \gg M$, the term $(Q - |m|)/Q$ in Eq. (10.106) can be neglected,[11] so we obtain

$$\mathcal{E}\{S(\omega)\} \cong \sum_{m=-(M-1)}^{M-1} \phi_{xx}[m] w_c[m] e^{-j\omega m} = \frac{1}{2\pi} \int_{-\pi}^{\pi} P_{xx}(\theta) W_c(e^{j(\omega-\theta)}) d\theta. \quad (10.107)$$

Thus, the windowed autocorrelation estimate leads to a biased estimate of the power spectrum. Just as with the average periodogram, it is possible to trade spectral resolution for reduced variance of the spectrum estimate. If the length of the data record is fixed, we can have lower variance if we are willing to accept poorer resolution of closely spaced narrowband spectral components, or we can have better resolution if we can accept higher variance. If we are free to observe the signal for a longer time (i.e., increase the length Q of the data record), then both the resolution and the variance can be improved. The spectrum estimate $S(\omega)$ is asymptotically unbiased if the correlation window is normalized so that

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} W_c(e^{j\omega}) d\omega = 1 = w_c[0]. \quad (10.108)$$

With this normalization, as we increase Q together with the length of the correlation window, the Fourier transform of the correlation window approaches a periodic impulse train and the convolution of Eq. (10.107) duplicates $P_{xx}(\omega)$.

The variance of $S(\omega)$ has been shown (see Jenkins and Watts, 1968) to be of the form

$$\text{var}[S(\omega)] \simeq \left(\frac{1}{Q} \sum_{m=-(M-1)}^{M-1} w_c^2[m] \right) P_{xx}^2(\omega). \quad (10.109)$$

Comparing Eq. (10.109) with the corresponding result in Eq. (10.80) for the periodogram leads to the conclusion that, to reduce the variance of the spectrum estimate, we should choose M and the window shape, possibly subject to the condition of Eq. (10.105), so that the factor

$$\left(\frac{1}{Q} \sum_{m=-(M-1)}^{M-1} w_c^2[m] \right) \quad (10.110)$$

is as small as possible. Problem 10.37 deals with the computation of this variance reduction factor for several commonly used windows.

Estimation of the power spectrum based on the Fourier transform of an estimate of the autocorrelation function is a clear alternative to the method of averaging periodograms. It is not necessarily better in any general sense; it simply has different features, and its implementation would be different. In some situations, it may be desirable to compute estimates of both the autocorrelation sequence and the power spectrum, in which case it would be natural to use the method of this section. Problem 10.43 explores the issue of determining an autocorrelation estimate from the average periodogram.

[11] More precisely, we could define an effective correlation window $w_e[m] = w_c[m](Q - |m|)/Q$.

10.6.1 Computing Correlation and Power Spectrum Estimates Using the DFT

The autocorrelation estimate

$$\hat{\phi}_{xx}[m] = \frac{1}{Q}\sum_{n=0}^{Q-|m|-1} x[n]x[n+|m|] \tag{10.111}$$

is required for $|m| \leq M-1$ in the method of power spectrum estimation that we are considering. Since $\hat{\phi}_{xx}[-m] = \hat{\phi}_{xx}[m]$, it is necessary to compute Eq. (10.111) only for nonnegative values of m, i.e., for $0 \leq m \leq M-1$. The DFT and its associated fast computational algorithms can be used to advantage in the computation of $\hat{\phi}_{xx}[m]$, if we observe that $\hat{\phi}_{xx}[m]$ is the aperiodic discrete convolution of the finite-length sequence $x[n]$ with $x[-n]$. If we compute $X[k]$, the N-point DFT of $x[n]$, and multiply by $X^*[k]$, we obtain $|X[k]|^2$, which corresponds to the circular convolution of the finite-length sequence $x[n]$ with $x[((-n))_N]$, i.e., a *circular autocorrelation*. As our discussion in Section 8.7 suggests, and as developed in Problem 10.34, it should be possible to augment the sequence $x[n]$ with zero-valued samples and force the circular autocorrelation to be equal to the desired aperiodic autocorrelation over the interval $0 \leq m \leq M-1$.

To see how to choose N for the DFT, consider Figure 10.28. Figure 10.28(a) shows the two sequences $x[n]$ and $x[n+m]$ as functions of n for a particular positive value of m. Figure 10.28(b) shows the sequences $x[n]$ and $x[((n+m))_N]$ that are involved in the circular autocorrelation corresponding to $|X[k]|^2$. Clearly, the circular autocorrelation will be equal to $Q\hat{\phi}_{xx}[m]$ for $0 \leq m \leq M-1$ if $x[((n+m))_N]$ does not wrap around and overlap $x[n]$ when $0 \leq m \leq M-1$. From Figure 10.28(b), it follows that this will be the case whenever $N-(M-1) \geq Q$ or $N \geq Q+M-1$.

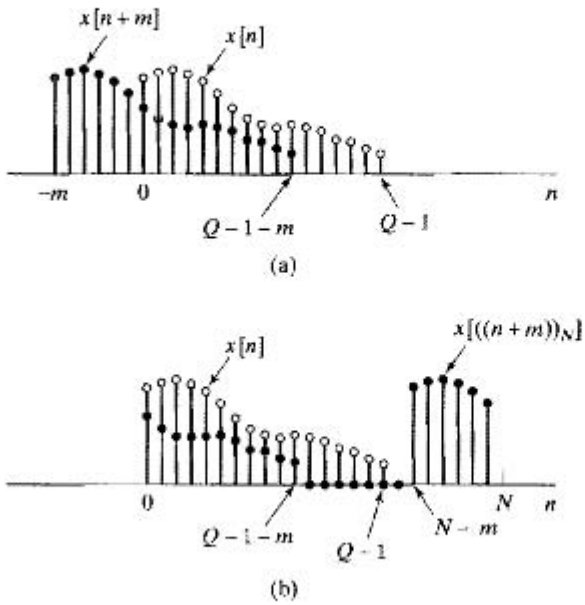


Figure 10.28 Computation of the circular autocorrelation. (a) $x[n]$ and $x[n+m]$ for a finite-length sequence of length Q. (b) $x[n]$ and $x[((n+m))_N]$ as in circular correlation.

In summary, we can compute $\hat{\phi}_{xx}[m]$ for $0 \le m \le M-1$ by the following procedure:

1. Form an N-point sequence by augmenting $x[n]$ with $(M-1)$ zero-samples.
2. Compute the N-point DFT,
$$X[k] = \sum_{n=0}^{N-1} x[n]e^{-j(2\pi/N)kn} \qquad \text{for } k = 0, 1, \ldots, N-1.$$
3. Compute
$$|X[k]|^2 = X[k]X^*[k] \qquad \text{for } k = 0, 1, \ldots, N-1.$$
4. Compute the inverse DFT of $|X[k]|^2$ to obtain
$$\tilde{c}_{vv}[m] = \frac{1}{N}\sum_{k=0}^{N-1} |X[k]|^2 e^{j(2\pi/N)km} \qquad \text{for } m = 0, 1, \ldots, N-1.$$
5. Divide the resulting sequence by Q to obtain the autocorrelation estimate
$$\hat{\phi}_{xx}[m] = \frac{1}{Q}\tilde{c}_{vv}[m] \qquad \text{for } m = 0, 1, \ldots, M-1.$$
This is the desired set of autocorrelation values, which can be extended symmetrically for negative values of m.

If M is small, it may be more efficient simply to evaluate Eq. (10.111) directly. In this case, the amount of computation is proportional to $Q \cdot M$. In contrast, if the DFTs in this procedure are computed using one of the FFT algorithms discussed in Chapter 9 with $N \ge Q + M - 1$, the amount of computation will be approximately proportional to $N \log_2 N$ for N a power of 2. Consequently, for sufficiently large values of M, use of the FFT is more efficient than direct evaluation of Eq. (10.111). The exact break-even value of M will depend on the particular implementation of the DFT computations; however, as shown by Stockham (1966), this value would probably be less than $M = 100$.

To reduce the variance of the estimate of the autocorrelation sequence or the power spectrum estimated from it, we must use large values of the record length Q. This is not generally a problem with computers having large memories and fast processors. However, since M is generally much less than Q, it is possible to section the sequence $x[n]$ in a manner similar to the procedures that were discussed in Section 8.7.3 for convolution of a finite-length impulse response with an indefinitely long input sequence. Rader (1970) presented a particularly efficient and flexible procedure that uses many of the properties of the DFT of real sequences to reduce the amount of computation required. The development of this technique is the basis for Problem 10.44.

Once the autocorrelation estimate has been computed, samples of the power spectrum estimate $S(\omega)$ can be computed at frequencies $\omega_k = 2\pi k/N$ by forming the finite-length sequence

$$s[m] = \begin{cases} \hat{\phi}_{xx}[m]w_c[m], & 0 \le m \le M-1, \\ 0, & M \le m \le N-M, \\ \hat{\phi}_{xx}[N-m]w_c[N-m], & N-M+1 \le m \le N-1, \end{cases} \tag{10.112}$$

where $w_c[m]$ is the symmetric correlation window. Then the DFT of $s[m]$ is

$$S[k] = S(\omega)|_{\omega=2\pi k/N}, \qquad k = 0, 1, \ldots, N-1. \tag{10.113}$$

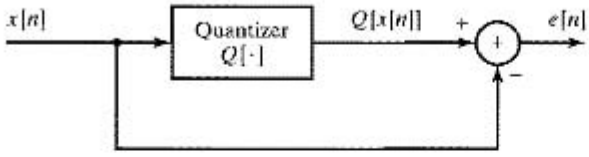


Figure 10.29 Procedure for obtaining quantization noise sequence.

where $S(\omega)$ is the Fourier transform of the windowed autocorrelation sequence as defined by Eq. (10.102). Note that N can be chosen as large as is convenient and practical, thereby providing samples of $S(\omega)$ at closely spaced frequencies. However, as our discussions in this chapter have consistently shown, the frequency resolution is always determined by the length and shape of the window $w_c[m]$.

10.6.2 Estimating the Power Spectrum of Quantization Noise

In Chapter 4, we assumed that the error introduced by quantization has the properties of a white-noise random process. The techniques discussed so far in this chapter were used to compute the power spectrum estimates of Figure 4.60 that were used to suggest the validity of this approximation. In this section, we provide additional examples of the use of estimates of the autocorrelation sequence and power spectrum estimation in studying the properties of quantization noise. The discussion will reinforce our confidence in the white-noise model, and it will also offer an opportunity to point out some practical aspects of power spectrum estimation.

Consider the experiment depicted in Figure 10.29. A lowpass-filtered speech signal $x_c(t)$ was sampled at a 16-KHz rate, yielding the sequence of samples $x[n]$ that were plotted in Figure 10.21.[12] These samples were quantized with a 10-bit linear quantizer ($B = 9$), and the corresponding error sequence $e[n] = Q[x[n]] - x[n]$ was computed. Figure 10.30 shows 2000 consecutive samples of the speech signal plotted on the first and third lines of the graph. The second and fourth lines show the corresponding quantization error sequence. Visual inspection and comparison of these two plots tends to strengthen our belief in the previously assumed model; i.e., that the noise appears to vary randomly throughout the range $-2^{-(B+1)} < e[n] \leq 2^{-(B+1)}$. However, such qualitative observations can be misleading. The flatness of the quantization noise spectrum can be verified only by estimating the autocorrelation sequence and power spectrum of the quantization noise $e[n]$.

Figure 10.31 shows estimates of the autocorrelation and power spectrum of the quantization noise for a record length of $Q = 3000$ samples. The autocorrelation sequence estimate was calculated over the range of lags $|m| \leq 100$ using Eqs. (10.98a) and (10.98b). The resulting estimate is shown in Figure 10.31(a). Over this range, $-1.45 \times 10^{-8} \leq \hat{\phi}[m] \leq 1.39 \times 10^{-8}$ except for $\hat{\phi}[0] = 3.17 \times 10^{-7}$. The autocorrelation estimate suggests that the sample-to-sample correlation of the noise sequence is quite low. The resulting autocorrelation estimate was multiplied by Bartlett windows

[12] Although the samples were originally quantized to 12 bits by the A/D converter, for purposes of this experiment, they were scaled to a maximum value of 1, and a small amount of random noise was added to the samples. We assume that these samples are "unquantized," i.e., we consider the 12-bit samples to effectively be unquantized relative to the subsequent quantization that we are applying in this discussion.

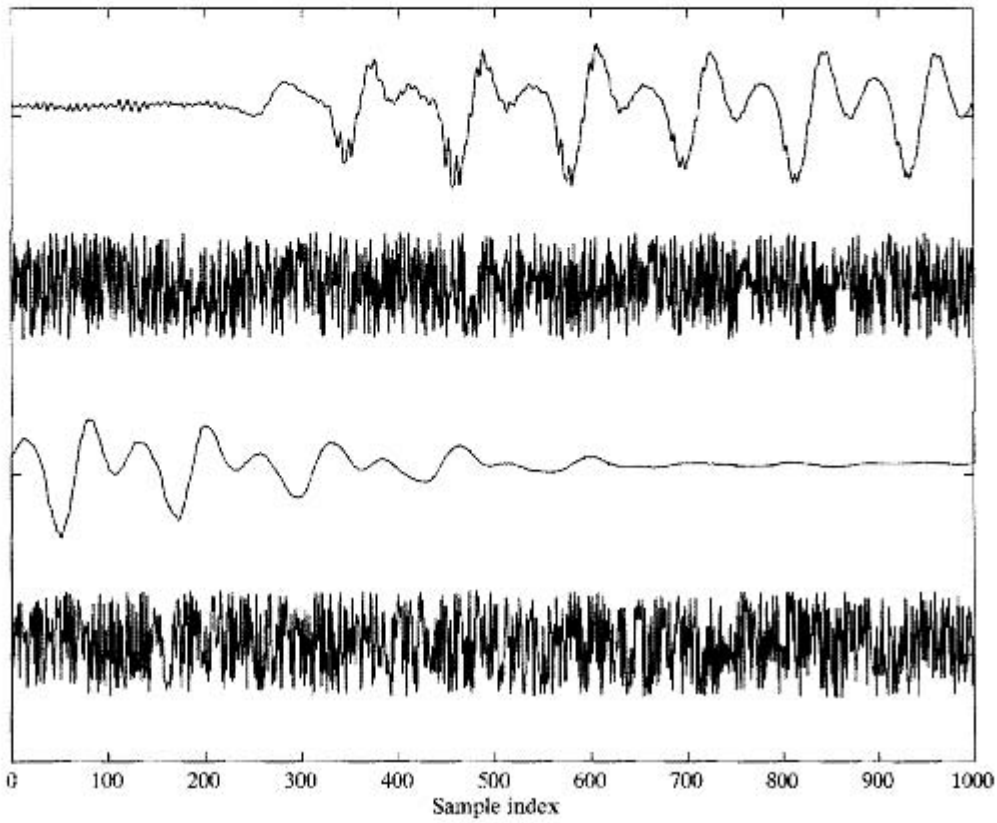


Figure 10.30 Speech waveform (first and third lines) and the corresponding quantization error (second and fourth lines) for 10-bit quantization (magnified 2^9 times). Each line corresponds to 1000 consecutive samples connected by straight lines for convenience in plotting.

with $M = 100$ and $M = 50$. The windows are shown in Figure 10.31 superimposed on $\hat{\phi}[m]$ (with scaling so that they can be plotted on the same axes) and the corresponding spectrum estimates, computed as discussed in Section 10.6.1, are shown in Figure 10.31(b).

As seen in Figure 10.31(b), the Blackman–Tukey spectrum estimate for $M = 100$ (the thin continuous line) shows somewhat erratic fluctuations about the dashed line plotted at the spectrum level $10\log_{10}(2^{-18}/12) = -64.98$ dB (the value of the white power spectrum with $\sigma_e^2 = 2^{-2B}/12$ for $B = 9$). The heavy line shows the power spectrum estimate for $M = 50$. We see from Figure 10.31(b) that the spectrum estimate is within ± 2 dB of the spectrum of the white-noise approximation for $B + 1 = 10$ for all frequencies. As discussed in Section 10.6, the shorter window gives smaller variance and a smoother spectrum estimate resulting from the lower frequency resolution of the shorter window. In either case, the spectrum estimate seems to support the validity of the white-noise model for quantization noise.

Although we have computed quantitative estimates of the autocorrelation and the power spectrum, our interpretation of these measurements has been only qualitative. It

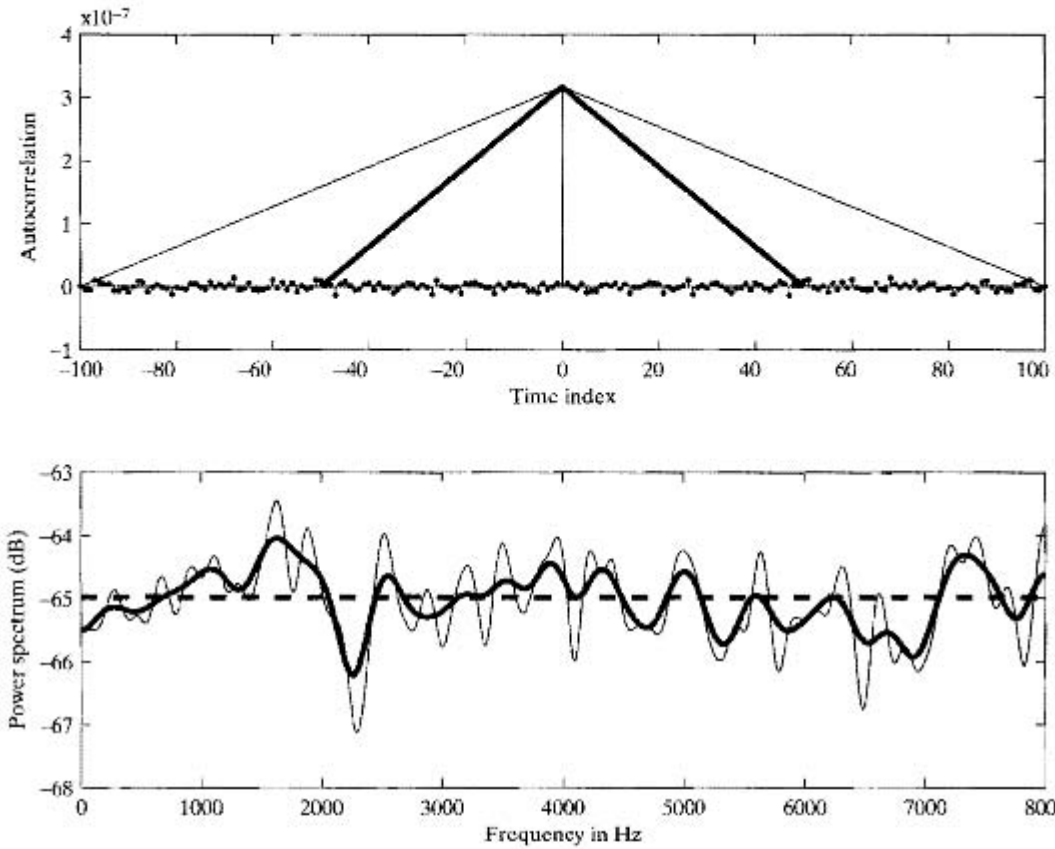


Figure 10.31 (a) Autocorrelation estimate for 10-bit quantization noise for $|m| \leq 100$ with record length $Q = 3,000$. (b) Power spectrum estimates by the Blackman–Tukey method using Bartlett windows with $M = 100$ and $M = 50$. (Dashed line shows level of $10 \log_{10}(2^{-18}/12)$.)

is reasonable now to wonder how small the autocorrelation would be if $e[n]$ were really a white-noise process. To give quantitative answers to such questions, confidence intervals for our estimates could be computed and statistical decision theory applied. (See Jenkins and Watts (1968), for some tests for white noise.) In many cases, however, this additional statistical treatment is not necessary. In a practical setting, we are often comfortable and content simply with the observation that the normalized autocorrelation is very small everywhere, except at $m = 0$.

Among the many important insights of this chapter is that the estimate of the autocorrelation and power spectrum of a stationary random process should improve if the record length is increased. This is illustrated by Figure 10.32, which corresponds to Figure 10.31, except that Q was increased to 30,000 samples. Recall that the variance of the autocorrelation estimate is proportional to $1/Q$. Thus, increasing Q from 3000 to 30,000 should bring about a tenfold reduction in the variance of the estimate. A comparison of Figures 10.31(a) and 10.32(a) seems to verify this result. For $Q = 3000$, the estimate

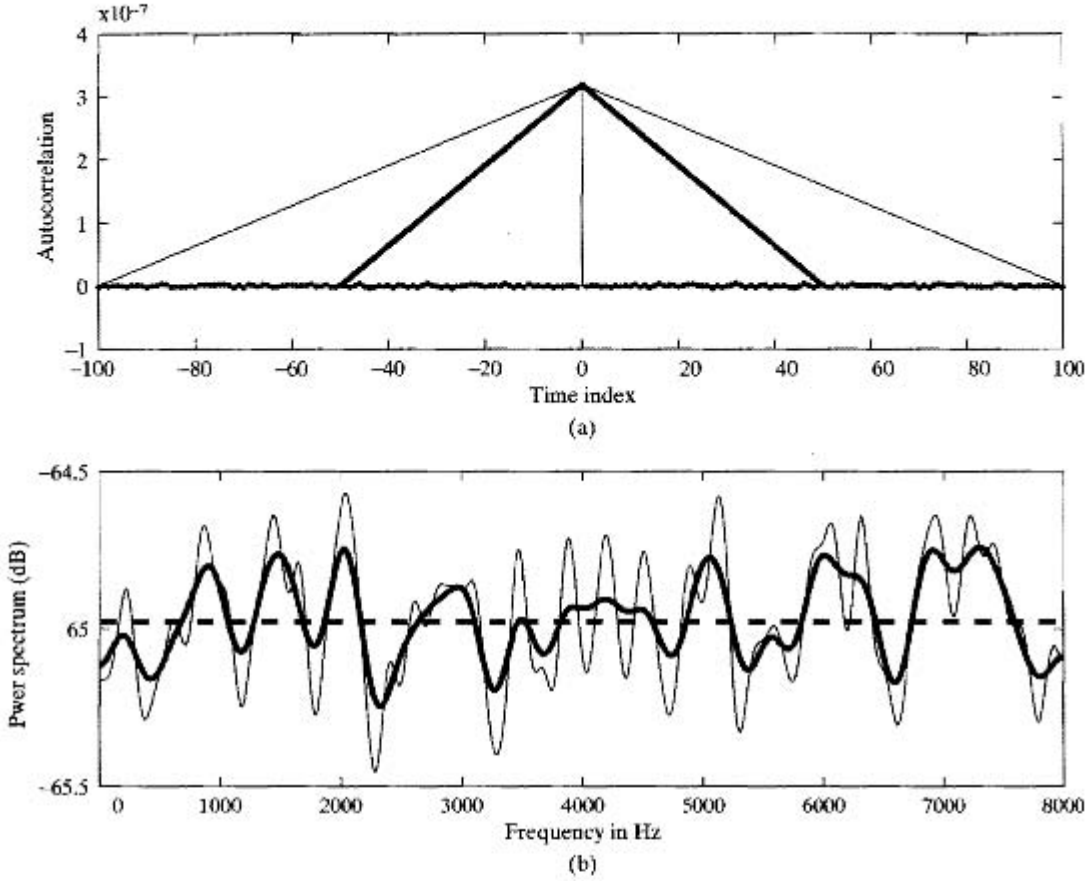


Figure 10.32 (a) Autocorrelation estimate for 10-bit quantization noise; record length $Q = 30{,}000$. (b) Power spectrum estimates by the Blackman–Tukey method using Bartlett windows with $M = 100$ and $M = 50$.

falls between the limits $-1.45 \times 10^{-8} \leq \hat{\phi}[m] \leq 1.39 \times 10^{-8}$, while for $Q = 30{,}000$, the limits are $-4.5 \times 10^{-9} \leq \hat{\phi}[m] \leq 4.15 \times 10^{-9}$. Comparing the range of variation for $Q = 3000$ with the range for $Q = 30{,}000$ indicates that the reduction is consistent with the tenfold reduction in variance that we expected.[13] We note from Eq. (10.110) that a similar reduction in variance of the spectrum estimate is also expected. This is again evident in comparing Figure 10.31(b) with Figure 10.32(b). (Be sure to note that the scales are different between the two sets of plots.) The variation about the white-noise approximate spectrum level is only ± 0.5 dB in the case of the longer record length. Note that the spectrum estimates in Figure 10.32(b) display the same trade off between variance and resolution.

In Chapter 4 we argued that the white-noise model was reasonable, as long as the quantization step size was small. When the number of bits is small, this condition does

[13]Recall that a reduction in variance by a factor of 10 corresponds to a reduction in amplitude by a factor of $\sqrt{10} \approx 3.16$.

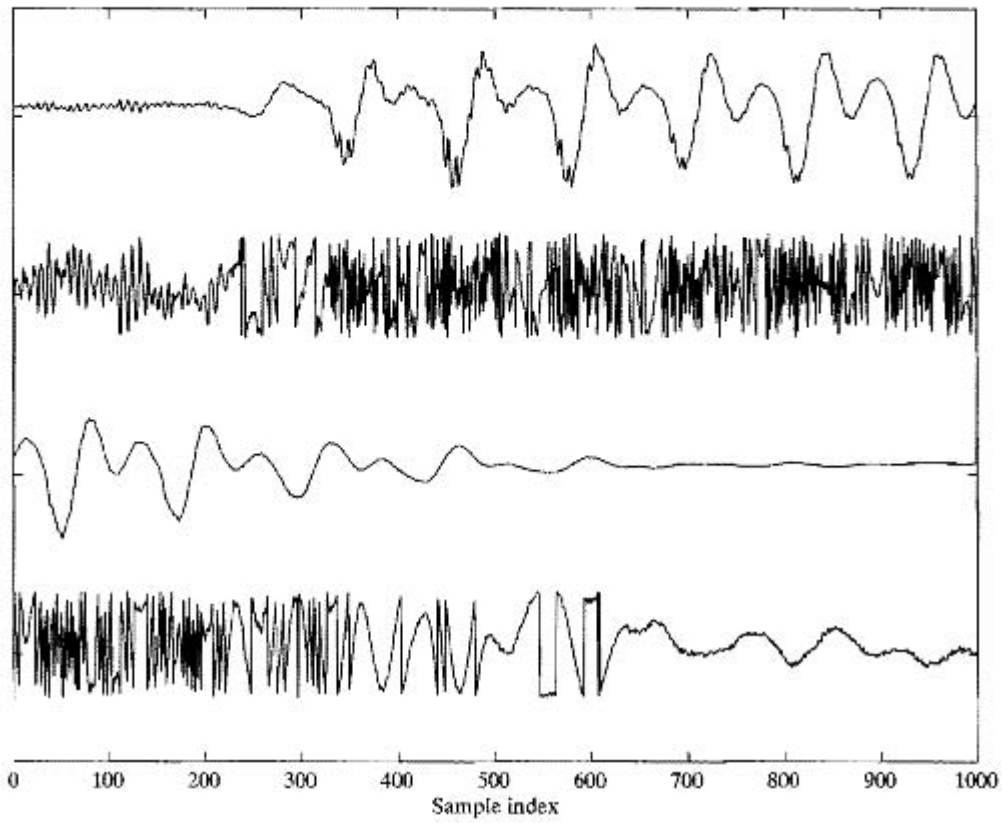


Figure 10.33 Speech waveform (first and third lines) and the corresponding quantization error (second and fourth lines) for 4-bit quantization (magnified 2^3 times). Each line corresponds to 1000 consecutive samples connected by straight lines for convenience in plotting.

not hold. To see the effect on the quantization noise spectrum, the previous experiment was repeated using only 16 quantization levels, or 4 bits. Figure 10.33 shows the speech waveform and quantization error for 4-bit quantization. Note that some portions of the error waveform tend to look very much like the original speech waveform. We would expect this to be reflected in the estimate of the power spectrum.

Figure 10.34 shows the autocorrelation and power spectrum estimates of the error sequence for 4-bit quantization for a record length of 30,000 samples. In this case, the autocorrelation shown in Figures 10.34(a) is much less like the ideal autocorrelation for white noise. This is not surprising in view of the obvious correlation between the signal and noise displayed in Figure 10.33. Figure 10.34(b) shows the power spectrum estimates for Bartlett windows with $M = 100$ and $M = 50$, respectively. Clearly, the spectrum is not flat, although the general level reflects the average noise power. In fact, as we shall see, the quantization noise tends to have the general shape of the speech spectrum. Thus, the white-noise model for quantization noise can be viewed only as a rather crude approximation in this case, and it would be less valid for coarser quantization.

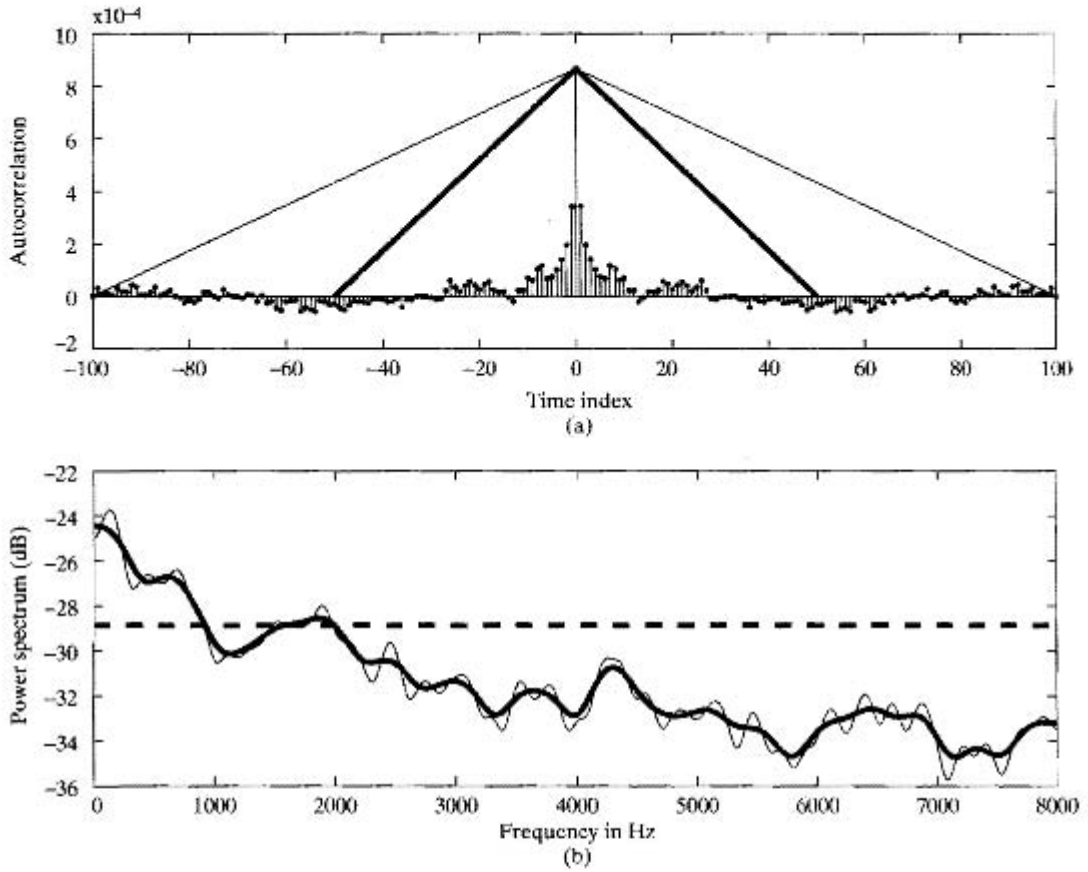


Figure 10.34 (a) Autocorrelation estimate for 4-bit quantization noise; record length $Q = 30{,}000$. (b) Power spectrum estimates by the Blackman–Tukey method using Bartlett windows with $M = 100$ and $M = 50$. (Dashed line shows level of $10 \log_{10}(2^{-6}/12)$.)

The example of this section illustrates how autocorrelation and power spectrum estimates can be used to support theoretical models. Specifically, we have demonstrated the validity of some of our basic assumptions in Chapter 4, and we have given an indication of how these assumptions break down for very crude quantization. This is only a rather simple, but useful, example that shows how the techniques of the current chapter can be applied in practice.

10.6.3 Estimating the Power Spectrum of Speech

We have seen that the time-dependent Fourier transform is particularly well-suited to the representation of speech signals, since it can track the time-varying nature of the speech signal. However, in some cases, it is useful to take a different point of view. In particular, even though the waveform of speech as in Figure 10.21 shows significant variability in time, as does its time-dependent Fourier transform in Figure 10.22, it is nevertheless possible to *assume* that it is a stationary random signal and apply our

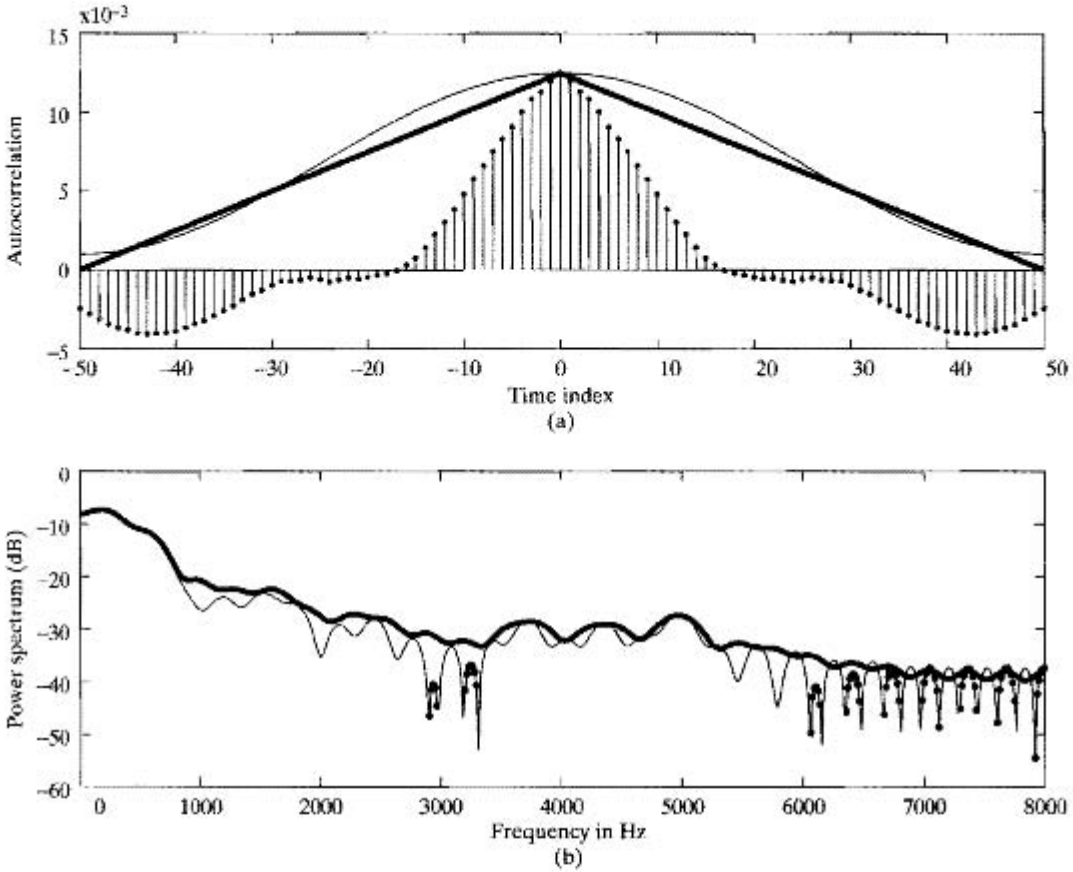


Figure 10.35 (a) Autocorrelation estimate for speech signal of Figure 10.21; record length $Q = 30{,}000$. (b) Power spectrum estimates by the Blackman–Tukey method using Bartlett window (heavy line) and Hamming window (light line) with $M = 50$.

long-term spectrum analysis techniques. These methods average over a time interval that is much longer than the changing events of speech. This gives a general spectrum shape that can be useful in designing speech coders and in determining the bandwidth requirements for speech transmission.

Figure 10.35 shows an example of estimating the power spectrum of speech using the Blackman–Tukey method. The autocorrelation sequence estimated from $Q = 30{,}000$ samples of the speech signal in Figure 10.21 is shown in Figure 10.35(a), together with Bartlett and Hamming windows of length $2M + 1 = 101$. Figure 10.35(b) shows the corresponding power spectrum estimates. The two estimates are grossly similar but dramatically different in detail. This is because of the nature of the DTFTs of the windows. Both have the same main-lobe width $\Delta\omega_m = 8\pi/M$, however their side lobes are very different. The side lobes of the Bartlett window are strictly nonnegative, while those of the symmetric Hamming window (which are smaller than those of the Bartlett window) are negative for some frequencies. When convolved with the periodogram

corresponding to the autocorrelation estimate, this yields the dramatically different results shown.

The Bartlett window guarantees a positive spectrum estimate for all frequencies. However, this is not true for the Hamming window. The effect of this is particularly pronounced in regions of rapid variability of the periodogram, where side lobes due to adjacent frequencies can cancel or interfere to produce negative spectrum estimates. The dots in Figure 10.35(b) show the frequencies where the spectrum estimate was negative. When plotting in dB, it is necessary to take the absolute value of the negative estimates. Thus, while the Bartlett window and the Hamming window have the same main-lobe width, the positive side lobes of the Bartlett window tend to fill in gaps between relatively strong frequencies, while the lower side lobes of the Hamming window lead to less leakage between frequencies, but the danger of negative spectrum estimates as positive and negative side lobes interact.

The Hamming window (or other windows such as the Kaiser window) can be used in spectrum estimation without danger of negative estimates if they are used in the method of averaging periodograms that are discussed in Section 10.5.3. This method guarantees positive estimates, because positive periodograms are averaged. Figure 10.36 shows a comparison of the Blackman–Tukey estimates of Figure 10.35(b) with an estimate obtained by the Welch method of averaging modified periodograms. The heavy dashed line is the Welch estimate. Note that it follows the general shape of the other two estimates, but it differs significantly in the high frequency region, where the speech spectrum is naturally small, and where the frequency response of the analog antialiasing filter causes the spectrum to be very small. Because of its superior ability to deliver consistent resolution for spectra with wide dynamic range, and because it is easily implemented using the DFT, the method of averaging periodograms is widely used in many practical applications of spectrum estimation.

All the spectrum estimates in Figure 10.36 show that the speech signal is characterized by a peak below 500 Hz and a fall-off with increasing frequency by 30 to 40 dB at 6 KHz. Several prominent peaks between 3 KHz and 5 KHz could be due to higher vocal tract resonances that do not vary with time. A different speaker or different speech material would certainly produce a different spectrum estimate, but the general nature of the spectrum estimates would be similar to those of Figure 10.36.

10.7 SUMMARY

One of the important applications of signal processing is spectrum analysis of signals. Because of the computational efficiency of the FFT, many of the techniques for spectrum analysis of continuous-time or discrete-time signals use the DFT either directly or indirectly. In this chapter, we explored and illustrated some of these techniques.

Many of the issues associated with spectrum analysis are best understood in the context of the analysis of sinusoidal signals. Since the use of the DFT requires finite-length signals, windowing must be applied in advance of the analysis. For sinusoidal signals, the width of the spectral peak observed in the DFT is dependent on the window length, with an increasing window length resulting in the sharpening of the peak. Consequently, the ability to resolve closely spaced sinusoids in the spectrum estimate

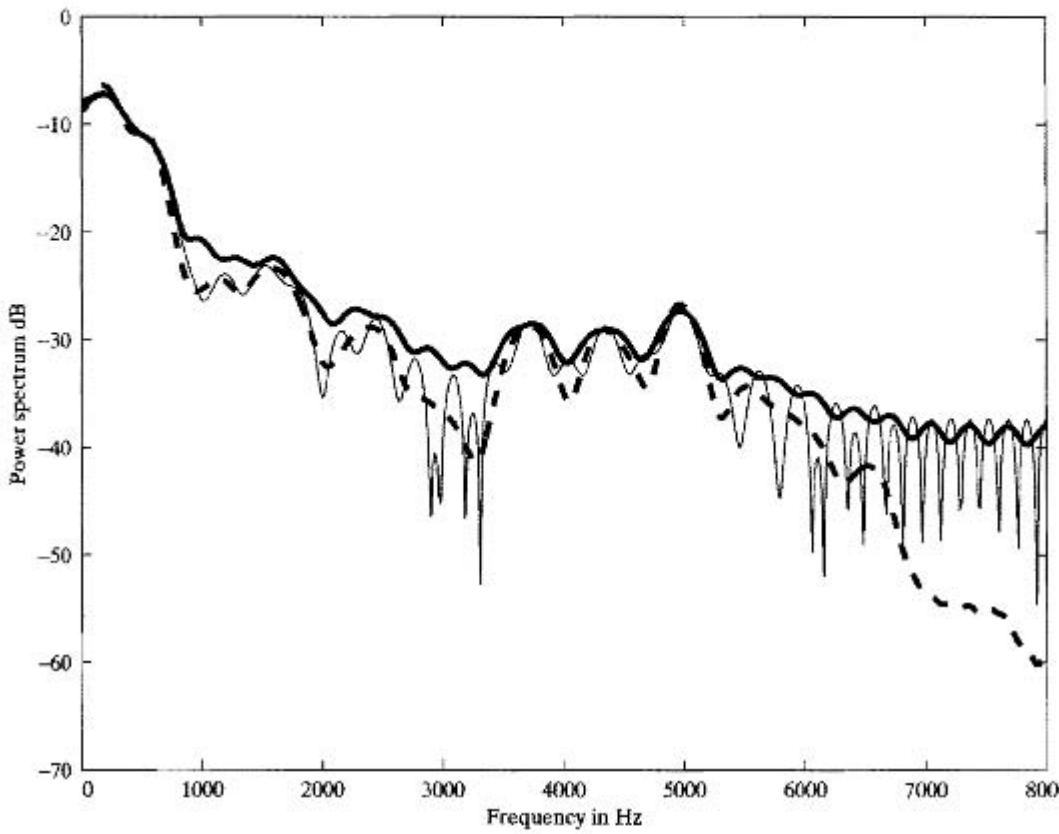


Figure 10.36 Power spectrum estimates by the Blackman–Tukey method using Bartlett window (heavy line) and Hamming window (light line) with $M = 50$. The dashed line shows the power spectrum obtained by averaging overlapping periodograms using a Hamming window with $M = 50$.

decreases as the window becomes shorter. A second, independent effect inherent in spectrum analysis using the DFT is the associated spectral sampling. Specifically, since the spectrum can be computed only at a set of sample frequencies, the observed spectrum can be misleading if we are not careful in its interpretation. For example, important features in the spectrum may not be directly evident in the sampled spectrum. To avoid this, the spectral sample spacing can be reduced by increasing the DFT size in one of two ways. One method is to increase the DFT size while keeping the window length fixed (requiring zero-padding of the windowed sequence). This does not increase resolution. The second method is to increase both the window length and the DFT size. In this case, spectral sample spacing is decreased, and the ability to resolve closely spaced sinusoidal components is increased.

While increased window length and resolution are typically beneficial in the spectrum analysis of stationary data, for time-varying data, it is generally preferable to keep the window length sufficiently short, so that over the window duration, the signal characteristics are approximately stationary. This leads to the concept of the time-dependent

Fourier transform, which, in effect, is a sequence of Fourier transforms obtained as the signal slides past a finite-duration window. A common and useful interpretation of the time-dependent Fourier transform is as a bank of filters, with the frequency response of each filter corresponding to the transform of the window, frequency shifted to one of the DFT frequencies. The time-dependent Fourier transform has important applications both as an intermediate step in filtering signals and for analyzing and interpreting time-varying signals, such as speech and radar signals. Spectral analysis of nonstationary signals typically involves a trade-off between time and frequency resolution. Specifically, our ability to track spectral characteristics in time increases as the length of the analysis window decreases. However, a shorter analysis window results in decreased frequency resolution.

The DFT also plays an important role in the analysis of stationary random signals. An intuitive approach to estimating the power spectrum of random signals is to compute the squared magnitude of the DFT of a segment of the signal. The resulting estimate, called the periodogram, is asymptotically unbiased. The variance of the periodogram estimate, however, does not decrease to zero as the length of the segment increases; consequently, the periodogram is not a good estimate. However, by dividing the available signal sequence into shorter segments and averaging the associated periodograms, we can obtain a well-behaved estimate. An alternative approach is to first estimate the autocorrelation function. This can be done either directly or with the DFT. If a window is then applied to the autocorrelation estimates followed by the DFT, the result, referred to as the smoothed periodogram, is a good spectrum estimate.

Problems

Basic Problems with Answers

10.1. A real continuous-time signal $x_c(t)$ is bandlimited to frequencies below 5 kHz; i.e., $X_c(j\Omega) = 0$ for $|\Omega| \geq 2\pi(5000)$. The signal $x_c(t)$ is sampled with a sampling rate of 10,000 samples per second (10 kHz) to produce a sequence $x[n] = x_c(nT)$ with $T = 10^{-4}$. Let $X[k]$ be the 1000-point DFT of $x[n]$.

(a) To what continuous-time frequency does the index $k = 150$ in $X[k]$ correspond?
(b) To what continuous-time frequency does the index $k = 800$ in $X[k]$ correspond?

10.2. A continuous-time signal $x_c(t)$ is bandlimited to 5 kHz; i.e., $X_c(j\Omega) = 0$ for $|\Omega| \geq 2\pi(5000)$. $x_c(t)$ is sampled with period T, producing the sequence $x[n] = x_c(nT)$. To examine the spectral properties of the signal, we compute the N-point DFT of a segment of N samples of $x[n]$ using a computer program that requires $N = 2^v$, where v is an integer.

Determine the *minimum* value for N and the range of sampling rates

$$F_{\min} < \frac{1}{T} < F_{\max}$$

such that aliasing is avoided, and the effective spacing between DFT values is *less* than 5 Hz; i.e., the equivalent continuous-time frequencies at which the Fourier transform is evaluated are separated by less than 5 Hz.

10.3. A continuous-time signal $x_c(t) = \cos(\Omega_0 t)$ is sampled with period T to produce the sequence $x[n] = x_c(nT)$. An N-point rectangular window is applied to $x[n]$ for $0, 1, \ldots, N-1$, and $X[k]$, for $k = 0, 1, \ldots, N-1$, is the N-point DFT of the resulting sequence.

(a) Assuming that Ω_0, N, and k_0 are fixed, how should T be chosen so that $X[k_0]$ and $X[N-k_0]$ are nonzero, and $X[k] = 0$ for all other values of k?

(b) Is your answer unique? If not, give another value of T that satisfies the conditions of part (a).

10.4. Let $x_c(t)$ be a real-valued, bandlimited signal whose Fourier transform $X_c(j\Omega)$ is zero for $|\Omega| \geq 2\pi(5000)$. The sequence $x[n]$ is obtained by sampling $x_c(t)$ at 10 kHz. Assume that the sequence $x[n]$ is zero for $n < 0$ and $n > 999$.

Let $X[k]$ denote the 1000-point DFT of $x[n]$. It is known that $X[900] = 1$ and $X[420] = 5$. Determine $X_c(j\Omega)$ for as many values of Ω as you can in the region $|\Omega| < 2\pi(5000)$.

10.5. Consider estimating the spectrum of a discrete-time signal $x[n]$ using the DFT with a Hamming window applied to $x[n]$. A conservative rule of thumb for the frequency resolution of windowed DFT analysis is that the frequency resolution is equal to the width of the main lobe of $W(e^{j\omega})$. You wish to be able to resolve sinusoidal signals that are separated by as little as $\pi/100$ in ω. In addition, your window length L is constrained to be a power of 2. What is the minimum length $L = 2^v$ that will meet your resolution requirement?

10.6. The following are three different signals $x_i[n]$ that are the sum of two sinusoids:

$$x_1[n] = \cos(\pi n/4) + \cos(17\pi n/64),$$

$$x_2[n] = \cos(\pi n/4) + 0.8\cos(21\pi n/64),$$

$$x_3[n] = \cos(\pi n/4) + 0.001\cos(21\pi n/64).$$

We wish to estimate the spectrum of each of these signals using a 64-point DFT with a 64-point rectangular window $w[n]$. Indicate which of the signals' 64-point DFTs you would expect to have two distinct spectral peaks after windowing.

10.7. Let $x[n]$ be a 5000-point sequence obtained by sampling a continuous-time signal $x_c(t)$ at $T = 50\ \mu$s. Suppose $X[k]$ is the 8192-point DFT of $x[n]$. What is the equivalent frequency spacing in continuous time of adjacent DFT samples?

10.8. Assume that $x[n]$ is a 1000-point sequence obtained by sampling a continuous-time signal $x_c(t)$ at 8 kHz and that $X_c(j\Omega)$ is sufficiently bandlimited to avoid aliasing. What is the minimum DFT length N such that adjacent samples of $X[k]$ correspond to a frequency spacing of 5 Hz or less in the original continuous-time signal?

10.9. $X_r[k]$ denotes the time-dependent Fourier transform (TDFT) defined in Eq. (10.40). For this problem, consider the TDFT when both the DFT length $N = 36$ and the sampling interval $R = 36$. The window $w[n]$ is a rectangular window of length $L = 36$. Compute the TDFT $X_r[k]$ for $-\infty < r < \infty$ and $0 \leq k \leq N-1$ for the signal

$$x[n] = \begin{cases} \cos(\pi n/6), & 0 \leq n \leq 35, \\ \cos(\pi n/2), & 36 \leq n \leq 71, \\ 0, & \text{otherwise.} \end{cases}$$

10.10. Figure P10.10 shows the spectrogram of a chirp signal of the form

$$x[n] = \sin\left(\omega_0 n + \frac{1}{2}\lambda n^2\right).$$

Note that the spectrogram is a representation of the magnitude of $X[n, k]$, as defined in Eq. (10.34), where the dark regions indicate large values of $|X[n, k]|$. Based on the figure, estimate ω_0 and λ.

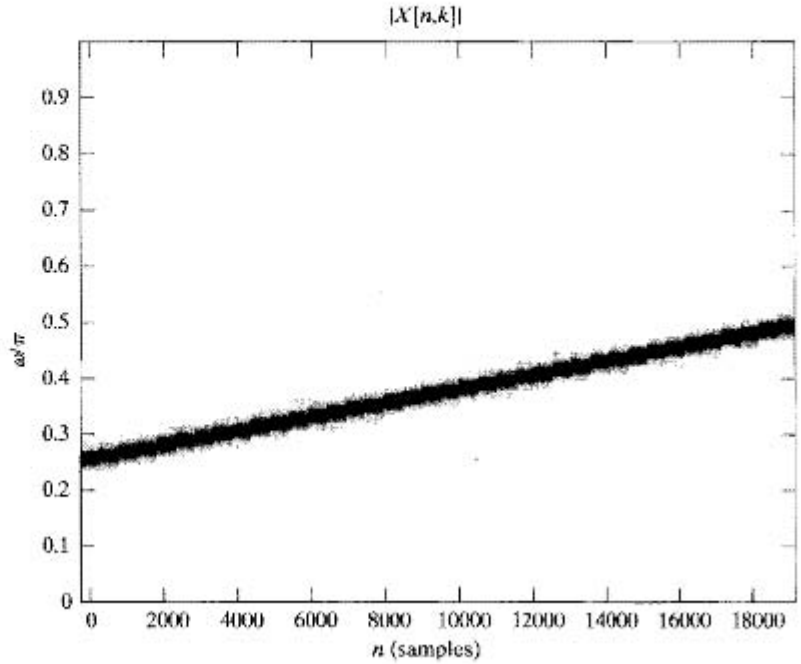


Figure P10.10

10.11. A continuous-time signal is sampled at a sampling rate of 10 kHz, and the DFT of 1024 samples is computed. Determine the continuous-time frequency spacing between spectral samples. Justify your answer.

10.12. Let $x[n]$ be a signal with a single sinusoidal component. The signal $x[n]$ is windowed with an L-point Hamming window $w[n]$ to obtain $v_1[n]$ before computing $V_1(e^{j\omega})$. The signal $x[n]$ is also windowed with an L-point rectangular window to obtain $v_2[n]$, which is used to compute $V_2(e^{j\omega})$. Will the peaks in $|V_2(e^{j\omega})|$ and $|V_1(e^{j\omega})|$ have the same height? If so, justify your answer. If not, which should have a larger peak?

10.13. It is desired to estimate the spectrum of $x[n]$ by applying a 512-point Kaiser window to the signal before computing $X(e^{j\omega})$.

(a) The requirements for the frequency resolution of the system specify that the largest allowable main lobe for the Kaiser window is $\pi/100$. What is the best side-lobe attenuation expected under these constraints?

(b) Suppose that you know that $x[n]$ contains two sinusoidal components at least $\pi/50$ apart, and that the amplitude of the stronger component is 1. Based on your answer to part (a), give a threshold on the smallest value of the weaker component you would expect to see over the side lobe of the stronger sinusoid.

10.14. A speech signal is sampled with a sampling rate of 16,000 samples/s (16 kHz). A window of 20-ms duration is used in time-dependent Fourier analysis of the signal, as described in Section 10.3, with the window being advanced by 40 samples between computations of the DFT. Assume that the length of each DFT is $N = 2^v$.

(a) How many samples are there in each segment of speech selected by the window?

(b) What is the "frame rate" of the time-dependent Fourier analysis; i.e., how many DFT computations are done per second of input signal?

(c) What is the minimum size N of the DFT such that the original input signal can be reconstructed from the time-dependent Fourier transform?

(d) What is the spacing (in Hz) between the DFT samples for the minimum N from part (c)?

10.15. A real-valued continuous-time segment of a signal $x_c(t)$ is sampled at a rate of 20,000 samples/s, yielding a 1000-point finite-length discrete-time sequence $x[n]$ that is nonzero in the interval $0 \leq n \leq 999$. It is known that $x_c(t)$ is also bandlimited such that $X_c(j\Omega) = 0$ for $|\Omega| \geq 2\pi(10{,}000)$; i.e., assume that the sampling operation does not introduce any distortion due to aliasing.

$X[k]$ denotes the 1000-point DFT of $x[n]$. $X[800]$ is known to have the value $X[800] = 1 + j$.

(a) From the information given, can you determine $X[k]$ at any other values of k? If so, state which value(s) of k and what the corresponding value of $X[k]$ is. If not, explain why not.

(b) From the information given, state the value(s) of Ω for which $X_c(j\Omega)$ is known and the corresponding value(s) of $X_c(j\Omega)$.

10.16. Let $x[n]$ be a discrete-time signal whose spectrum you wish to estimate using a windowed DFT. You are required to obtain a frequency resolution of at least $\pi/25$ and are also required to use a window length $N = 256$. A safe estimate of the frequency resolution of a spectral estimate is the main-lobe width of the window used. Which of the windows in Table 7.2 will satisfy the criteria given for frequency resolution?

10.17. Let $x[n]$ be a discrete-time signal obtained by sampling a continuous-time signal $x_c(t)$ with some sampling period T so that $x[n] = x_c(nT)$. Assume $x_c(t)$ is bandlimited to 100 Hz, i.e. $X_c(j\Omega) = 0$ for $|\Omega| \geq 2\pi(100)$. We wish to estimate the continuous-time spectrum $X_c(j\Omega)$ by computing a 1024-point DFT of $x[n]$, $X[k]$. What is the smallest value of T such that the equivalent frequency spacing between consecutive DFT samples $X[k]$ corresponds to 1 Hz or less in continuous-time frequency?

10.18. Figure P10.18 shows the magnitude $|V[k]|$ of the 128-point DFT $V[k]$ for a signal $v[n]$. The signal $v[n]$ was obtained by multiplying $x[n]$ by a 128-point rectangular window $w[n]$; i.e., $v[n] = x[n]w[n]$. Note that Figure P10.18 shows $|V[k]|$ only for the interval $0 \leq k \leq 64$. Which of the following signals could be $x[n]$? That is, which are consistent with the information shown in the figure?

$$x_1[n] = \cos(\pi n/4) + \cos(0.26\pi n),$$

$$x_2[n] = \cos(\pi n/4) + (1/3)\sin(\pi n/8),$$

$$x_3[n] = \cos(\pi n/4) + (1/3)\cos(\pi n/8),$$

$$x_4[n] = \cos(\pi n/8) + (1/3)\cos(\pi n/16),$$

$$x_5[n] = (1/3)\cos(\pi n/4) + \cos(\pi n/8),$$

$$x_6[n] = \cos(\pi n/4) + (1/3)\cos(\pi n/8 + \pi/3).$$

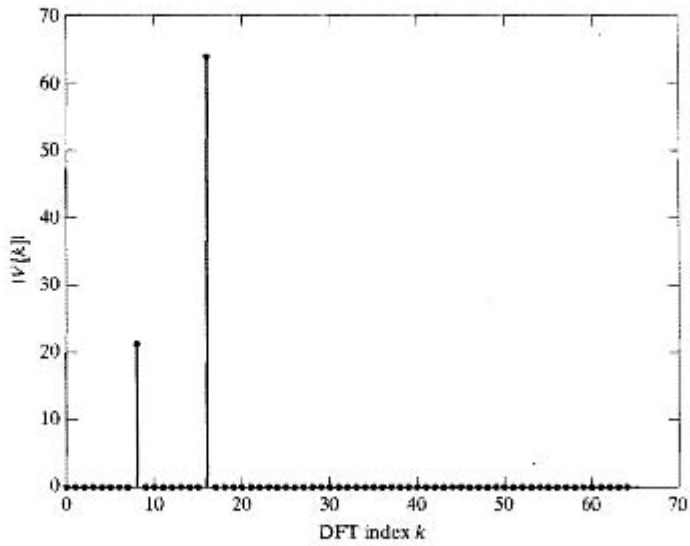


Figure P10.18

10.19. A signal $x[n]$ is analyzed using the time-dependent Fourier transform $X_r[k]$, as defined in Eq. (10.40). Initially, the analysis is performed with an $N = 128$ DFT using an $L = 128$-point Hamming window $w[n]$. The time-domain sampling of adjacent blocks is $R = 128$; i.e., the windowed segments are offset by 128 samples in time. The frequency resolution obtained with this analysis is not sufficient, and it is desired to improve the resolution. Several methods of modifying the analysis are suggested to accomplish this goal. Which of the following methods will improve the frequency resolution of the time-dependent Fourier transform $X_r[k]$?

METHOD 1: Increase N to 256 while maintaining L and R at the same values.
METHOD 2: Increase both N and L to 256, while maintaining R the same.
METHOD 3: Decrease R to 64 while maintaining the same N and L.
METHOD 4: Decrease L to 64 while maintaining the same N and R.
METHOD 5: Maintain N, R and L the same, but change $w[n]$ to be a rectangular window.

10.20. Assume that you wish to estimate the spectrum of $x[n]$ by applying a Kaiser window to the signal before computing the DTFT. You require that the side lobe of the window be 30 dB below the main lobe and that the frequency resolution be $\pi/40$. The width of the main lobe of the window is a safe estimate of the frequency resolution. Estimate the minimum window length L that will meet these requirements.

Basic Problems

10.21. Let $x[n] = \cos(2\pi n/5)$ and $v[n]$ be the sequence obtained by applying a 32-point rectangular window to $x[n]$ before computing $V(e^{j\omega})$. Sketch $|V(e^{j\omega})|$ for $-\pi \leq \omega \leq \pi$, labeling the frequencies of all peaks and the first nulls on either side of the peak. In addition, label the amplitudes of the peaks and the strongest side lobe of each peak.

10.22. In this problem we are interested in estimating the spectra of three very long real-valued data sequences $x_1[n]$, $x_2[n]$, and $x_3[n]$, each consisting of the sum of two sinusoidal components. However, we only have a 256-point segment of each sequence available for analysis.

Let $\bar{x}_1[n]$, $\bar{x}_2[n]$, and $\bar{x}_3[n]$ denote the 256-point segments of $x_1[n]$, $x_2[n]$, and $x_3[n]$, respectively. We have some information about the nature of the spectra of the infinitely long sequences, as indicated in Eqs. (P10.22-1) through (P10.22-3). Two different spectral analysis procedures are being considered for use, one using a 256-point rectangular window and the other a 256-point Hamming window. These procedures are described below. In the descriptions, the signal $\mathcal{R}_N[n]$ denotes the N-point rectangular window and $\mathcal{H}_N[n]$ denotes the N-point Hamming window. The operator $\mathrm{DFT}_{2048}\{\cdot\}$ indicates taking the 2048-point DFT of its argument after zero-padding the end of the input sequence. This will give a good interpolation of the DTFT from the frequency samples of the DFT.

$$X_1(e^{j\omega}) \approx \delta(\omega + \frac{17\pi}{64}) + \delta(\omega + \frac{\pi}{4}) + \delta(\omega - \frac{\pi}{4}) + \delta(\omega - \frac{17\pi}{64}) \quad \text{(P10.22-1)}$$

$$X_2(e^{j\omega}) \approx 0.017\delta(\omega + \frac{11\pi}{32}) + \delta(\omega + \frac{\pi}{4}) + \delta(\omega - \frac{\pi}{4}) + 0.017\delta(\omega - \frac{11\pi}{32}) \quad \text{(P10.22-2)}$$

$$X_3(e^{j\omega}) \approx 0.01\delta(\omega + \frac{257\pi}{1024}) + \delta(\omega + \frac{\pi}{4}) + \delta(\omega - \frac{\pi}{4}) + 0.01\delta(\omega - \frac{257\pi}{1024}) \quad \text{(P10.22-3)}$$

Based on Eqs. (P10.22-1) through (P10.22-3), indicate which of the spectral analysis procedures described below would allow you to conclude responsibly whether the anticipated frequency components were present. A good justification at a minimum will include a quantitative consideration of both resolution and side-lobe behavior of the estimators. Note that it is possible that both or neither of the algorithms will work for any given data sequence. Table 7.2 may be useful in deciding which algorithm(s) to use with which sequence.

Spectral Analysis Algorithms

Algorithm 1: Use the entire data segment with a rectangular window.

$$v[n] = \mathcal{R}_{256}[n]x[n]$$

$$\left|V(e^{j\omega})\right|_{\omega=\frac{2\pi k}{2048}} = \left|\mathrm{DFT}_{2048}\{v[n]\}\right|.$$

Algorithm 2: Use the entire data segment with a Hamming window.

$$v[n] = \mathcal{H}_{256}[n]x[n]$$

$$\left|V(e^{j\omega})\right|_{\omega=\frac{2\pi k}{2048}} = \left|\mathrm{DFT}_{2048}\{v[n]\}\right|.$$

10.23. Sketch the spectrogram obtained by using a 256-point rectangular window and 256-point DFTs with no overlap ($R = 256$) on the signal

$$x[n] = \cos\left[\frac{\pi n}{4} + 1000\sin\left(\frac{\pi n}{8000}\right)\right]$$

for the interval $0 \le n \le 16{,}000$.

10.24. **(a)** Consider the system of Figure P10.24-1 with input $x(t) = e^{j(3\pi/8)10^4 t}$, sampling period $T = 10^{-4}$, and

$$w[n] = \begin{cases} 1, & 0 \le n \le N-1, \\ 0, & \text{otherwise.} \end{cases}$$

What is the smallest nonzero value of N such that $X_w[k]$ is nonzero at exactly one value of k?

(b) Suppose now that $N = 32$, the input signal is $x(t) = e^{j\Omega_0 t}$, and the sampling period T is chosen such that no aliasing occurs during the sampling process. Figures P10.24-2 and P10.24-3 show the magnitude of the sequence $X_w[k]$ for $k = 0, \ldots, 31$ for the following two different choices of $w[n]$:

$$w_1[n] = \begin{cases} 1, & 0 \le n \le 31, \\ 0, & \text{otherwise}, \end{cases}$$

$$w_2[n] = \begin{cases} 1, & 0 \le n \le 7, \\ 0, & \text{otherwise}. \end{cases}$$

Indicate which figure corresponds to which choice of $w[n]$. State your reasoning clearly.

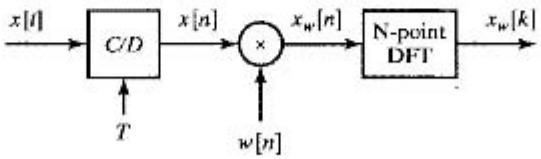


Figure P10.24-1

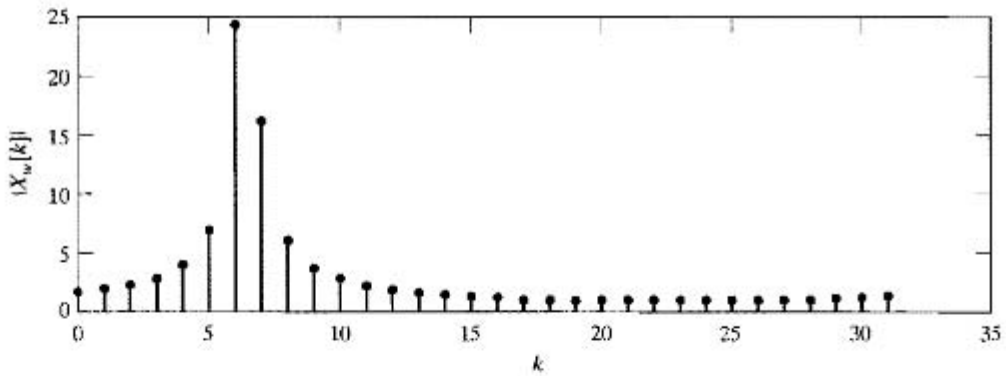


Figure P10.24-2

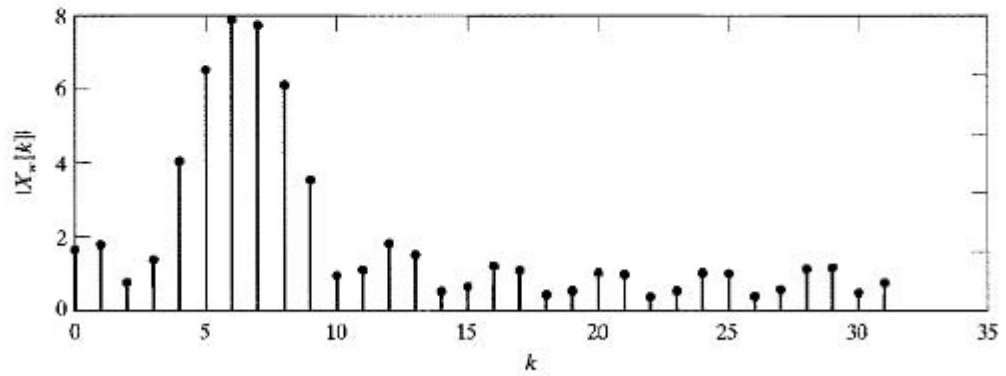


Figure P10.24-3

(c) For the input signal and system parameters of part (b), we would like to estimate the value of Ω_0 from Figure P10.24-3 when the sampling period is $T = 10^{-4}$. Assuming that the sequence

$$w[n] = \begin{cases} 1, & 0 \le n \le 31, \\ 0, & \text{otherwise}, \end{cases}$$

and that the sampling period is sufficient to ensure that no aliasing occurs during sampling, estimate the value of Ω_0. Is your estimate exact? If it is not, what is the maximum possible error of your frequency estimate?

(d) Suppose you were provided with the exact values of the 32-point DFT $X_w[k]$ for the window choices $w_1[n]$ and $w_2[n]$. Briefly describe a procedure to obtain a precise estimate of Ω_0.

Advanced Problems

10.25. In Figure P10.25, a filter bank is shown for which

$$h_0[n] = 3\delta[n+1] + 2\delta[n] + \delta[n-1],$$

and

$$h_q[n] = e^{j\frac{2\pi qn}{M}} h_0[n], \qquad \text{for } q = 1, \ldots, N-1.$$

The filter bank consists of N filters, modulated by a fraction $1/M$ of the total frequency band. Assume M and N are both greater than the length of $h_0[n]$.

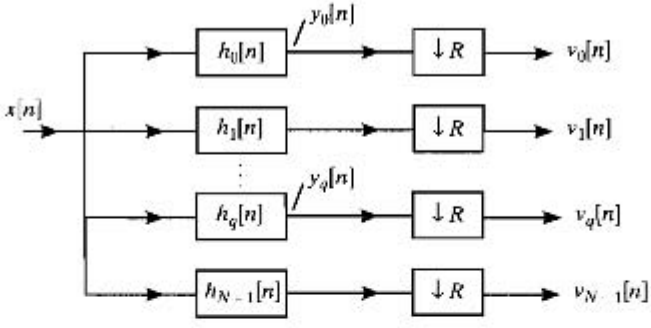


Figure P10.25 Filter bank

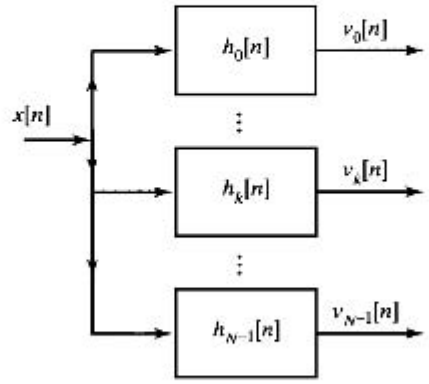
(a) Express $y_q[n]$ in terms of the time-dependent Fourier transform $X[n, \lambda)$ of $x[n]$, and sketch and label explicitly the values for the associated window in the time-dependent Fourier transform.

For parts (b) and (c), assume that $M = N$. Since $v_q[n]$ depends on the two integer variables q and n, we alternatively write it as the two-dimensional sequence $v[q, n]$.

(b) For $R = 2$, describe a procedure to recover $x[n]$ for all values of n if $v[q, n]$ is available for all integer values of q and n.

(c) Will your procedure in (b) work if $R = 5$? Clearly explain.

10.26. The system in Figure P10.26-1 uses a modulated filter bank for spectral analysis. (For further illustration, Figure P10.26-2 shows how the frequency responses $H_k(e^{j\omega})$ relate.) The impulse response of the prototype filter $h_0[n]$ is sketched in Figure P10.26-3.



$h_k[n] = e^{j\omega_k n} h_0[n]$, $\omega_k = \dfrac{2\pi k}{N}$, where $k = 0, 1, \ldots, N-1$

$h_0[n]$ = lowpass prototype filter $H_k(z) = H_0(e^{-j2\pi k/N} z)$

Figure P10.26-1

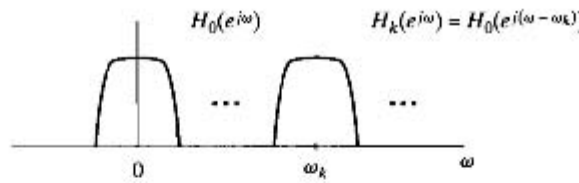


Figure P10.26-2

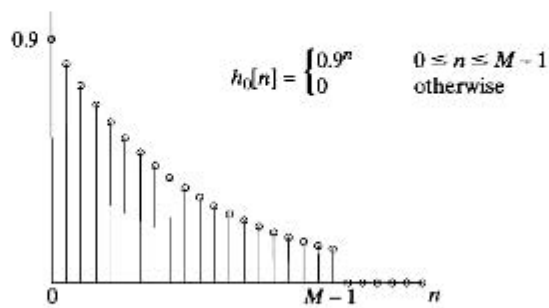


Figure P10.26-3

An alternative system for spectral analysis is shown in Figure P10.26-4. Determine $w[n]$ so that $G[k] = v_k[0]$, for $k = 0, 1, \ldots, N-1$.

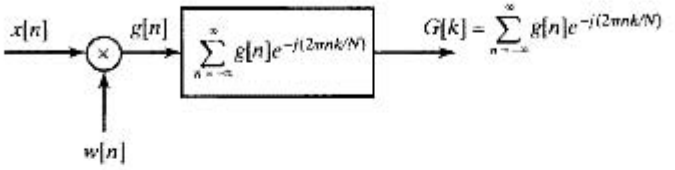


Figure P10.26-4

10.27. We are interested in obtaining 256 equally spaced samples of the z-transform of $x_w[n]$. $x_w[n]$ is a windowed version of an arbitrary sequence $x[n]$ where $x_w[n] = x[n]w[n]$ and $w[n] = 1, 0 \le n \le 255$ and $w[n] = 0$ otherwise. The z-transform of $x_w[n]$ is defined as

$$X_w(z) = \sum_{n=0}^{255} x[n]z^{-n}.$$

The samples $X_w[k]$ that we would like to compute are

$$X_w[k] = X_w(z)\Big|_{z=0.9e^{j\frac{2\pi}{256}k}} \qquad k = 0, 1, \ldots, 255.$$

We would like to process the signal $x[n]$ with a modulated filter bank, as indicated in Figure P10.27.

Each filter in the filter bank has an impulse response that is related to the prototype *causal* lowpass filter $h_0[n]$ as follows:

$$h_k[n] = h_0[n]e^{-j\omega_k n} \qquad k = 1, 2, \ldots, 255.$$

Each output of the filter bank is sampled once, at time $n = N_k$, to obtain $X_w[k]$, i.e.,

$$X_w[k] = v_k[N_k].$$

Determine $h_0[n]$, ω_k and N_k so that

$$X_w[k] = v_k[N_k] = X_w(z)\Big|_{z=0.9e^{j\frac{2\pi}{256}k}} \qquad k = 0, 1, \ldots, 255.$$

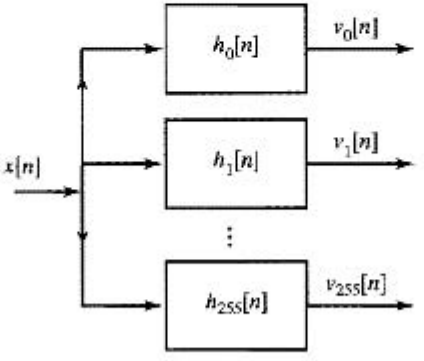


Figure P10.27

10.28. **(a)** In Figure P10.28-1, we show a system for spectral analysis of a signal $x_c(t)$, where

$$G_k[n] = \sum_{l=0}^{N-1} g_l[n] e^{-j\frac{2\pi}{N}lk},$$

$$N = 512, \quad \text{and} \quad LR = 256.$$

For the most general choice of the multiplier coefficient a_l, determine the choice for L and R which will result in the smallest number of multiplies per second.

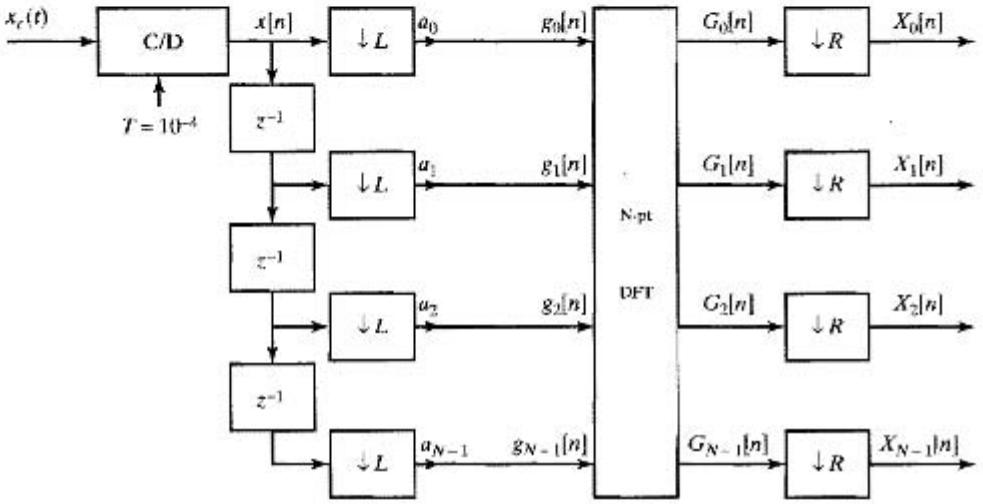


Figure P10.28-1

(b) In Figure P10.28-2, we show another system for spectral analysis of a signal $x_c(t)$, where

$$h[n] = \begin{cases} (0.93)^n & 0 \le n \le 255 \\ 0 & \text{otherwise} \end{cases},$$

$$h_k[n] = h[n]e^{-j\omega_k n}, \quad k = 0, 1, \cdots, N-1, \quad \text{and} \quad N = 512.$$

Listed below are **two** possible choices for M, **four** possible choices for ω_k, and **six** possible choices for the coefficients a_l. From this set identify all combinations for which $Y_k[n] = X_k[n]$, i.e., for which both systems will provide the same spectral analysis. There may be more than one.

M: (a) 256 (b) 512

ω_k: (a) $\frac{2\pi k}{256}$ (b) $\frac{2\pi k}{512}$ (c) $\frac{-2\pi k}{256}$ (d) $\frac{-2\pi k}{512}$

a_l:				
	(a)	$(0.93)^l$	l=0, 1, ⋯, 255,	zero otherwise
	(b)	$(0.93)^{-l}$	l=0, 1, ⋯, 511	
	(c)	$(0.93)^l$	l=0, 1, ⋯, 511	
	(d)	$(0.93)^{-l}$	l=0, 1, ⋯, 255,	zero otherwise
	(e)	$(0.93)^l$	l=256, 257, ⋯, 511,	zero otherwise
	(f)	$(0.93)^{-l}$	l=256, 257, ⋯, 511,	zero otherwise

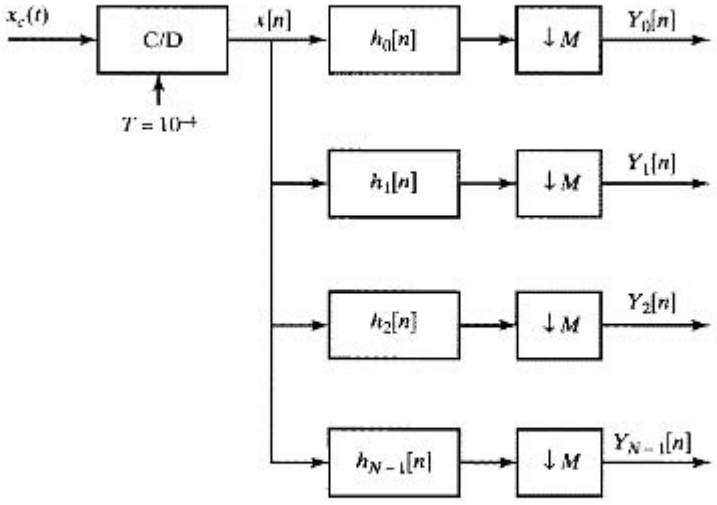


Figure P10.28-2

10.29. The system shown in Figure P10.29 is proposed as a spectrum analyzer. The basic operation is as follows: The spectrum of the sampled input is frequency-shifted; the lowpass filter selects the lowpass band of frequencies; the downsampler "spreads" the selected frequency band back over the entire range $-\pi < \omega < \pi$; and the DFT samples that frequency band uniformly at N frequencies.

Assume that the input is bandlimited so that $X_c(j\Omega) = 0$ for $|\Omega| \geq \pi/T$. The LTI system with frequency response $H(e^{j\omega})$ is an ideal lowpass filter with gain of one and cutoff frequency π/M. Furthermore, assume that $0 < \omega_1 < \pi$ and the data window $w[n]$ is a rectangular window of length N.

(a) Plot the DTFTs, $X(e^{j\omega})$, $Y(e^{j\omega})$, $R(e^{j\omega})$, and $R_d(e^{j\omega})$ for the given $X_c(j\Omega)$ and for $\omega_1 = \pi/2$ and $M = 4$. Give the relationship between the input and output Fourier transforms for each stage of the process; e.g., in the fourth plot, you would indicate $R(e^{j\omega}) = H(e^{j\omega})Y(e^{j\omega})$.

(b) Using your result in part (a), generalize to determine the band of continuous-time frequencies in $X_c(j\Omega)$ that falls within the passband of the lowpass discrete-time filter. Your answer will depend on M, ω_1 and T. For the specific case of $\omega_1 = \pi/2$ and $M = 4$, indicate this band of frequencies on the plot of $X_c(j\Omega)$ given for part (a).

(c) (i) What continuous-time frequencies in $X_c(j\Omega)$ are associated with the DFT values $V[k]$ for $0 \leq k \leq N/2$?

(ii) What continuous-time frequencies in $X_c(j\Omega)$ do the values for $N/2 < k \leq N-1$ correspond to? In each case, give a formula for the frequencies Ω_k.

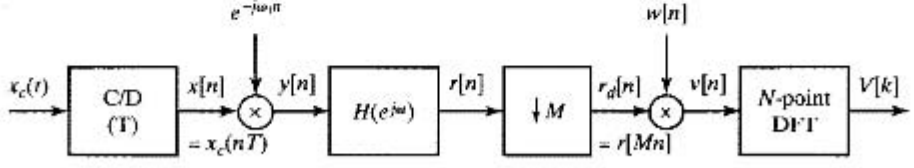


Figure P10.29

10.30. Consider a real time-limited continuous-time signal $x_c(t)$ whose duration is 100 ms. Assume that this signal has a bandlimited Fourier transform such that $X_c(j\Omega) = 0$ for $|\Omega| \geq 2\pi(10,000)$ rad/s; i.e., assume that aliasing is negligible. We want to compute samples of $X_c(j\Omega)$ with 5-Hz spacing over the interval $0 \leq \Omega \leq 2\pi(10{,}000)$. This can be done with a 4000-point DFT. Specifically, we want to obtain a 4000-point sequence $x[n]$ for which the 4000-point DFT is related to $X_c(j\Omega)$ by

$$X[k] = \alpha X_c(j2\pi \cdot 5 \cdot k), \qquad k = 0, 1, \ldots, 1999,$$

where α is a known scale factor. Three methods are proposed to obtain a 4000-point sequence whose DFT gives the desired samples of $X_c(j\Omega)$.

METHOD 1: $x_c(t)$ is sampled with a sampling period $T = 25\ \mu$s; i.e., we compute $X_1[k]$, the DFT of the sequence

$$x_1[n] = \begin{cases} x_c(nT), & n = 0, 1, \ldots, 3999, \\ 0, & \text{otherwise.} \end{cases}$$

Since $x_c(t)$ is time limited to 100 ms, $x_1[n]$ is a finite-length sequence of length 4000 (100 ms/25 μs).

METHOD 2: $x_c(t)$ is sampled with a sampling period of $T = 50\ \mu$s. Since $x_c(t)$ is time limited to 100 ms, the resulting sequence will have only 2000 (100 ms/50 μs) nonzero samples; i.e.,

$$x_2[n] = \begin{cases} x_c(nT), & n = 0, 1, \ldots, 1999, \\ 0, & \text{otherwise.} \end{cases}$$

In other words, the sequence is padded with zero-samples to create a 4000-point sequence for which the 4000-point DFT $X_2[k]$ is computed.

METHOD 3: $x_c(t)$ is sampled with a sampling period of $T = 50\ \mu$s, as in Method 2. The resulting 2000-point sequence is used to form the sequence $x_3[n]$ as follows:

$$x_3[n] = \begin{cases} x_c(nT), & 0 \leq n \leq 1999, \\ x_c((n - 2000)T), & 2000 \leq n \leq 3999, \\ 0, & \text{otherwise.} \end{cases}$$

The 4000-point DFT $X_3[k]$ of this sequence is computed.

For each of the three methods, determine how each 4000-point DFT is related to $X_c(j\Omega)$. Indicate this relationship in a sketch for a "typical" Fourier transform $X_c(j\Omega)$. State explicitly which method(s) provide the desired samples of $X_c(j\Omega)$.

10.31. A continuous-time finite-duration signal $x_c(t)$ is sampled at a rate of 20,000 samples/s, yielding a 1000-point finite-length sequence $x[n]$ that is nonzero in the interval $0 \leq n \leq 999$. Assume for this problem that the continuous-time signal is also bandlimited such that $X_c(j\Omega) = 0$ for $|\Omega| \geq 2\pi(10{,}000)$; i.e., assume that negligible aliasing distortion occurs in sampling. Assume also that a device or program is available for computing 1000-point DFTs and inverse DFTs.

(a) If $X[k]$ denotes the 1000-point DFT of the sequence $x[n]$, how is $X[k]$ related to $X_c(j\Omega)$? What is the effective continuous-time frequency spacing between DFT samples?

The following procedure is proposed for obtaining an expanded view of the Fourier transform $X_c(j\Omega)$ in the interval $|\Omega| \leq 2\pi(5000)$, starting with the 1000-point DFT $X[k]$.

Step 1. Form the new 1000-point DFT

$$W[k] = \begin{cases} X[k], & 0 \leq k \leq 250, \\ 0, & 251 \leq k \leq 749, \\ X[k], & 750 \leq k \leq 999. \end{cases}$$

Step 2. Compute the inverse 1000-point DFT of $W[k]$, obtaining $w[n]$ for $n = 0, 1, \ldots, 999$.

Step 3. Decimate the sequence $w[n]$ by a factor of 2 and augment the result with 500 consecutive zero samples, obtaining the sequence

$$y[n] = \begin{cases} w[2n], & 0 \le n \le 499, \\ 0, & 500 \le n \le 999. \end{cases}$$

Step 4. Compute the 1000-point DFT of $y[n]$, obtaining $Y[k]$.

(b) The designer of this procedure asserts that

$$Y[k] = \alpha X_c(j2\pi \cdot 10 \cdot k), \qquad k = 0, 1, \ldots, 500,$$

where α is a constant of proportionality. Is this assertion correct? If not, explain why not.

10.32. An analog signal consisting of a sum of sinusoids was sampled with a sampling rate of $f_s = 10000$ samples/s to obtain $x[n] = x_c(nT)$. Four spectrograms showing the time-dependent Fourier transform $|X[n, \lambda)|$ were computed using either a rectangular or a Hamming window. They are plotted in Figure P10.32. (A log amplitude scale is used, and only the top 35 dB is shown.)

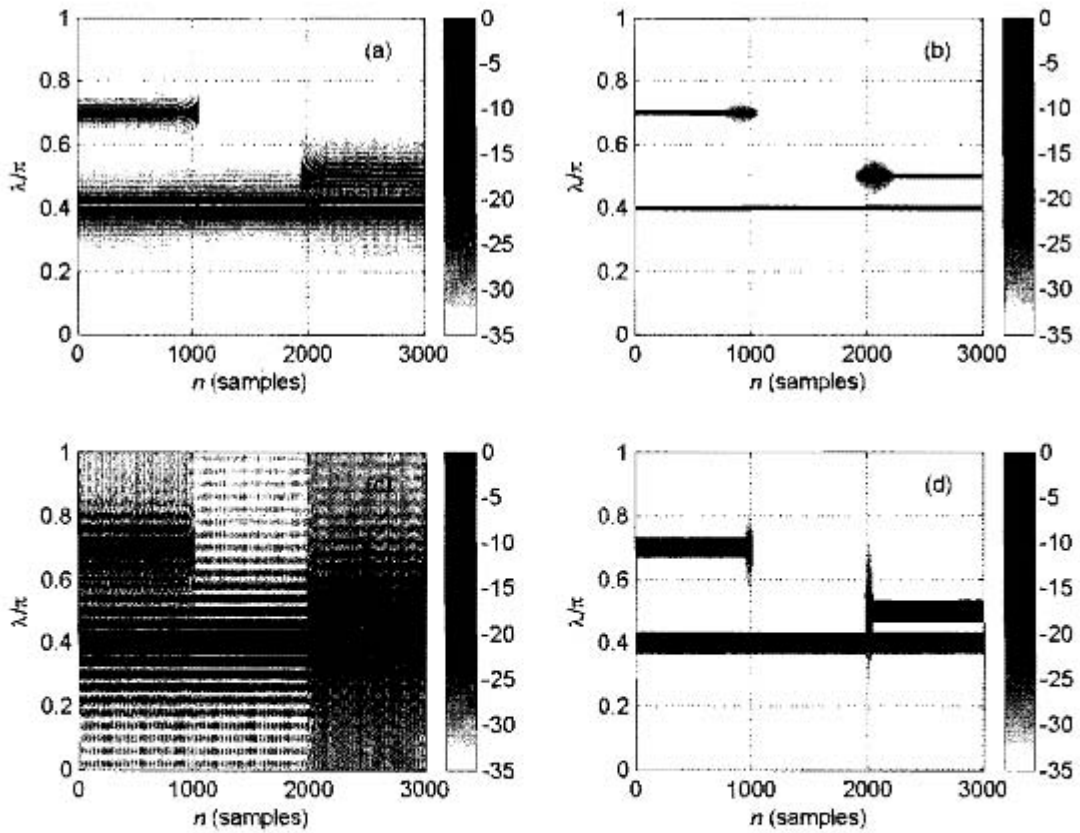


Figure P10.32

(a) Which spectrograms were computed with a rectangular window?
(a) (b) (c) (d)

(b) Which pair (or pairs) of spectrograms have approximately the same frequency resolution?
(a&b) (b&d) (c&d) (a&d) (b&c)

(c) Which spectrogram has the shortest time window? (a) (b) (c) (d)

(d) To the nearest 100 samples, estimate the window length L (in samples) of the window in spectrogram (b).

(e) Use the spectrographic data in Figure P10.32 to assist you in writing an equation (or equations) for an analog sum of sinusoids $x_c(t)$, which when sampled at a sampling rate of $f_s = 10000$, would produce the above spectrograms. Be as complete as you can in your description of the signal. Indicate any parameters that cannot be obtained from the spectrogram.

10.33. The periodogram $I(\omega)$ of a discrete-time random signal $x[n]$ was defined in Eq. (10.67) as

$$I(\omega) = \frac{1}{LU}|V(e^{j\omega})|^2,$$

where $V(e^{j\omega})$ is the DTFT of the finite-length sequence $v[n] = w[n]x[n]$, with $w[n]$ a finite-length window sequence of length L, and U is a normalizing constant. Assume that $x[n]$ and $w[n]$ are real.

Show that the periodogram is also equal to $1/LU$ times the Fourier transform of the aperiodic autocorrelation sequence of $v[n]$; i.e.,

$$I(\omega) = \frac{1}{LU}\sum_{m=-(L-1)}^{L-1} c_{vv}[m]e^{-j\omega m},$$

where

$$c_{vv}[m] = \sum_{n=0}^{L-1} v[n]v[n+m].$$

10.34. Consider a finite-length sequence $x[n]$ such that $x[n] = 0$ for $n < 0$ and $n \geq L$. Let $X[k]$ be the N-point DFT of the sequence $x[n]$, where $N > L$. Define $c_{xx}[m]$ to be the aperiodic autocorrelation function of $x[n]$; i.e.,

$$c_{xx}[m] = \sum_{n=-\infty}^{\infty} x[n]x[n+m].$$

Define

$$\tilde{c}_{xx}[m] = \frac{1}{N}\sum_{m=0}^{N-1} |X[k]|^2 e^{j(2\pi/N)km}, \qquad m = 0, 1, \ldots, N-1.$$

(a) Determine the minimum value of N that can be used for the DFT if we require that

$$c_{xx}[m] = \tilde{c}_{xx}[m], \qquad 0 \leq m \leq L-1.$$

(b) Determine the minimum value of N that can be used for the DFT if we require that

$$c_{xx}[m] = \tilde{c}_{xx}[m], \qquad 0 \leq m \leq M-1,$$

where $M < L$.

10.35. The symmetric Bartlett window, which arises in many aspects of power spectrum estimation, is defined as

$$w_B[m] = \begin{cases} 1 - |m|/M, & |m| \leq M-1, \\ 0, & \text{otherwise.} \end{cases} \tag{P10.35-1}$$

The Bartlett window is particularly attractive for obtaining estimates of the power spectrum by windowing an estimated autocorrelation function, as discussed in Section 10.6. This is because its Fourier transform is nonnegative, which guarantees that the smoothed spectrum estimate will be nonnegative at all frequencies.

(a) Show that the Bartlett window as defined in Eq. (P10.35-1) is $(1/M)$ times the aperiodic autocorrelation of the sequence $(u[n] - u[n-M])$.

(b) From the result of part (a), show that the Fourier transform of the Bartlett window is

$$W_B(e^{j\omega}) = \frac{1}{M}\left[\frac{\sin(\omega M/2)}{\sin(\omega/2)}\right]^2, \tag{P10.35-2}$$

which is clearly nonnegative.

(c) Describe a procedure for generating other finite-length window sequences that have nonnegative Fourier transforms.

10.36. Consider a signal

$$x[n] = \left[\sin\left(\frac{\pi n}{2}\right)\right]^2 u[n]$$

whose time-dependent discrete Fourier transform is computed using the analysis window

$$w[n] = \begin{cases} 1, & 0 \leq n \leq 13, \\ 0, & \text{otherwise.} \end{cases}$$

Let $X[n,k] = X[n, 2\pi k/7)$ for $0 \leq k \leq 6$, where $X[n,\lambda)$ is defined as in Section 10.3.

(a) Determine $X[0,k]$ for $0 \leq k \leq 6$.

(b) Evaluate $\sum_{k=0}^{6} X[n,k]$ for $0 \leq n < \infty$.

Extension Problems

10.37. In Section 10.6, we showed that a smoothed estimate of the power spectrum can be obtained by windowing an estimate of the autocorrelation sequence. It was stated (see Eq. (10.109)) that the variance of the smoothed spectrum estimate is

$$\text{var}[S(\omega)] \simeq FP_{xx}^2(\omega),$$

where F, the *variance ratio* or *variance reduction factor*, is

$$F = \frac{1}{Q}\sum_{m=-(M-1)}^{M-1}(w_c[m])^2 = \frac{1}{2\pi Q}\int_{-\pi}^{\pi}|W_c(e^{j\omega})|^2 d\omega.$$

As discussed in Section 10.6, Q is the length of the sequence $x[n]$ and $(2M-1)$ is the length of the symmetric window $w_c[m]$ that is applied to the autocorrelation estimate. Thus, if Q is fixed, the variance of the smoothed spectrum estimate can be reduced by adjusting the shape and duration of the window applied to the correlation function.

In this problem we will show that F decreases as the window length decreases, but we also know from the previous discussion of windows in Chapter 7 that the width of the main lobe of $W_c(e^{j\omega})$ increases with decreasing window length, so that the ability to

resolve two adjacent frequency components is reduced as the window width decreases. Thus, there is a trade-off between variance reduction and resolution. We will study this trade-off for the following commonly used windows:

Rectangular

$$w_R[m] = \begin{cases} 1, & |m| \le M-1, \\ 0, & \text{otherwise.} \end{cases}$$

Bartlett (triangular)

$$w_B[m] = \begin{cases} 1 - |m|/M, & |m| \le M-1, \\ 0, & \text{otherwise.} \end{cases}$$

Hanning/Hamming

$$w_H[m] = \begin{cases} \alpha + \beta\cos[\pi m/(M-1)], & |m| \le M-1, \\ 0, & \text{otherwise.} \end{cases}$$

($\alpha = \beta = 0.5$ for the Hanning window, and $\alpha = 0.54$ and $\beta = 0.46$ for the Hamming window.)

(a) Find the Fourier transform of each of the foregoing windows; i.e., compute $W_R(e^{j\omega})$, $W_B(e^{j\omega})$, and $W_H(e^{j\omega})$. Sketch each of these Fourier transforms as functions of ω.

(b) For each of the windows, show that the entries in the following table are approximately true when $M \gg 1$:

Window Name	Approximate Main-lobe Width	Approximate Variance Ratio (F)
Rectangular	$2\pi/M$	$2M/Q$
Bartlett	$4\pi/M$	$2M/(3Q)$
Hanning/Hamming	$3\pi/M$	$2M(\alpha^2 + \beta^2/2)/Q$

10.38. Show that the time-dependent Fourier transform, as defined by Eq. (10.18), has the following properties:

(a) *Linearity:*

$$\text{If} \quad x[n] = ax_1[n] + bx_2[n], \quad \text{then} \quad X[n,\lambda) = aX_1[n,\lambda) + bX_2[n,\lambda).$$

(b) *Shifting:* If $y[n] = x[n - n_0]$, then $Y[n,\lambda) = X[n - n_0,\lambda)$.

(c) *Modulation:* If $y[n] = e^{j\omega_0 n}x[n]$, then $Y[n,\lambda) = e^{j\omega_0 n}X[n,\lambda - \omega_0)$.

(d) *Conjugate Symmetry:* If $x[n]$ is real, then $X[n,\lambda) = X^*[n,-\lambda)$.

10.39. Suppose that $x_c(t)$ is a real, continuous-time stationary random signal with autocorrelation function

$$\phi_c(\tau) = \mathcal{E}\{x_c(t)x_c(t+\tau)\}$$

and power density spectrum

$$P_c(\Omega) = \int_{-\infty}^{\infty} \phi_c(\tau)e^{-j\Omega\tau}d\tau.$$

Consider a discrete-time stationary random signal $x[n]$ that is obtained by sampling $x_c(t)$ with sampling period T; i.e., $x[n] = x_c(nT)$.

(a) Show that $\phi[m]$, the autocorrelation sequence for $x[n]$, is

$$\phi[m] = \phi_c(mT).$$

(b) What is the relationship between the power density spectrum $P_c(\Omega)$ for the continuous-time random signal and the power density spectrum $P(\omega)$ for the discrete-time random signal?

(c) What condition is necessary such that

$$P(\omega) = \frac{1}{T} P_c\left(\frac{\omega}{T}\right), \qquad |\omega| < \pi?$$

10.40. In Section 10.5.5, we considered the estimation of the power spectrum of a sinusoid plus white noise. In this problem, we will determine the true power spectrum of such a signal. Suppose that

$$x[n] = A\cos(\omega_0 n + \theta) + e[n],$$

where θ is a random variable that is uniformly distributed on the interval from 0 to 2π and $e[n]$ is a sequence of zero-mean random variables that are independent of each other and also independent of θ. In other words, the cosine component has a randomly selected phase, and $e[n]$ represents white noise.

(a) Show that for the preceding assumptions, the autocorrelation function for $x[n]$ is

$$\phi_{xx}[m] = \mathcal{E}\{x[n]x[m+n]\} = \frac{A^2}{2}\cos(\omega_0 m) + \sigma_e^2\delta[m],$$

where $\sigma_e^2 = \mathcal{E}\{(e[n])^2\}$.

(b) From the result of part (a), show that over one period in frequency, the power spectrum of $x[n]$ is

$$P_{xx}(\omega) = \frac{A^2\pi}{2}[\delta(\omega - \omega_0) + \delta(\omega + \omega_0)] + \sigma_e^2, \qquad |\omega| \le \pi.$$

10.41. Consider a discrete-time signal $x[n]$ of length N samples that was obtained by sampling a stationary, white, zero-mean continuous-time signal. It follows that

$$\mathcal{E}\{x[n]x[m]\} = \sigma_x^2\delta[n-m],$$

$$\mathcal{E}\{x[n]\} = 0.$$

Suppose that we compute the DFT of the finite-length sequence $x[n]$, thereby obtaining $X[k]$ for $k = 0, 1, \ldots, N-1$.

(a) Determine the approximate variance of $|X[k]|^2$ using Eqs. (10.80) and (10.81).

(b) Determine the cross-correlation between values of the DFT; i.e., determine $\mathcal{E}\{X[k]X^*[r]\}$ as a function of k and r.

10.42. A bandlimited continuous-time signal has a bandlimited power spectrum that is zero for $|\Omega| \ge 2\pi(10^4)$ rad/s. The signal is sampled at a rate of 20,000 samples/s over a time interval of 10 s. The power spectrum of the signal is estimated by the method of averaging periodograms as described in Section 10.5.3.

(a) What is the length Q (number of samples) of the data record?

(b) If a radix-2 FFT program is used to compute the periodograms, what is the minimum length N if we wish to obtain estimates of the power spectrum at equally spaced frequencies no more than 10 Hz apart?

(c) If the segment length L is equal to the FFT length N in part (b), how many segments K are available if the segments do not overlap?

(d) Suppose that we wish to reduce the variance of the spectrum estimates by a factor of 10 while maintaining the frequency spacing of part (b). Give two methods of doing this. Do these two methods give the same results? If not, explain how they differ.

10.43. Suppose that an estimate of the power spectrum of a signal is obtained by the method of averaging periodograms, as discussed in Section 10.5.3. That is, the spectrum estimate is

$$\bar{I}(\omega) = \frac{1}{K}\sum_{r=0}^{K-1} I_r(\omega),$$

where the K periodograms $I_r(\omega)$ are computed from L-point segments of the signal using Eqs. (10.82) and (10.83). We define an estimate of the autocorrelation function as the inverse Fourier transform of $\bar{I}(\omega)$; i.e.,

$$\bar{\phi}[m] = \frac{1}{2\pi}\int_{-\pi}^{\pi} \bar{I}(\omega)e^{j\omega m}\,d\omega.$$

(a) Show that

$$\mathcal{E}\{\bar{\phi}[m]\} = \frac{1}{LU}c_{ww}[m]\phi_{xx}[m],$$

where L is the length of the segments, U is a normalizing factor given by Eq. (10.79), and $c_{ww}[m]$, given by Eq. (10.75), is the aperiodic autocorrelation function of the window that is applied to the signal segments.

(b) In the application of periodogram averaging, we normally use an FFT algorithm to compute $\bar{I}(\omega)$ at N equally spaced frequencies; i.e.,

$$\bar{I}[k] = \bar{I}(2\pi k/N), \qquad k = 0, 1, \ldots, N-1,$$

where $N \geq L$. Suppose that we compute an estimate of the autocorrelation function by computing the inverse DFT of $\bar{I}[k]$, as in

$$\bar{\phi}_p[m] = \frac{1}{N}\sum_{k=0}^{N-1} \bar{I}[k]e^{j(2\pi/N)km}, \qquad m = 0, 1, \ldots, N-1.$$

Obtain an expression for $\mathcal{E}\{\bar{\phi}_p[m]\}$.

(c) How should N be chosen so that

$$\mathcal{E}\{\bar{\phi}_p[m]\} = \mathcal{E}\{\bar{\phi}[m]\}, \qquad m = 0, 1, \ldots, L-1?$$

10.44. Consider the computation of the autocorrelation estimate

$$\hat{\phi}_{xx}[m] = \frac{1}{Q}\sum_{n=0}^{Q-|m|-1} x[n]x[n+|m|], \tag{P10.44-1}$$

where $x[n]$ is a real sequence. Since $\hat{\phi}_{xx}[-m] = \hat{\phi}_{xx}[m]$, it is necessary only to evaluate Eq. (P10.44-1) for $0 \leq m \leq M-1$ to obtain $\hat{\phi}_{xx}[m]$ for $-(M-1) \leq m \leq M-1$, as is required to estimate the power density spectrum using Eq. (10.102).

(a) When $Q \gg M$, it may not be feasible to compute $\hat{\phi}_{xx}[m]$ using a single FFT computation. In such cases, it is convenient to express $\hat{\phi}_{xx}[m]$ as a sum of correlation estimates based on shorter sequences. Show that if $Q = KM$,

$$\hat{\phi}_{xx}[m] = \frac{1}{Q}\sum_{i=0}^{K-1} c_i[m],$$

where

$$c_i[m] = \sum_{n=0}^{M-1} x[n+iM]x[n+iM+m],$$

for $0 \leq m \leq M-1$.

(b) Show that the correlations $c_i[m]$ can be obtained by computing the N-point *circular* correlations

$$\tilde{c}_i[m] = \sum_{n=0}^{N-1} x_i[n]y_i[((n+m))_N],$$

where the sequences

$$x_i[n] = \begin{cases} x[n+iM], & 0 \le n \le M-1, \\ 0, & M \le n \le N-1, \end{cases}$$

and

$$y_i[n] = x[n+iM], \qquad 0 \le n \le N-1. \tag{P10.44-2}$$

What is the *minimum* value of N (in terms of M) such that $c_i[m] = \tilde{c}_i[m]$ for $0 \le m \le M-1$?

(c) State a procedure for computing $\hat{\phi}_{xx}[m]$ for $0 \le m \le M-1$ that involves the computation of $2K$ N-point DFTs of real sequences and *one* N-point inverse DFT. How many complex multiplications are required to compute $\hat{\phi}_{xx}[m]$ for $0 \le m \le M-1$ if a radix-2 FFT is used?

(d) What modifications to the procedure developed in part (c) would be necessary to compute the cross-correlation estimate

$$\hat{\phi}_{xy}[m] = \frac{1}{Q}\sum_{n=0}^{Q-|m|-1} x[n]y[n+m], \qquad -(M-1) \le m \le M-1,$$

where $x[n]$ and $y[n]$ are real sequences known for $0 \le n \le Q-1$?

(e) Rader (1970) showed that, for computing the autocorrelation estimate $\hat{\phi}_{xx}[m]$ for $0 \le m \le M-1$, significant savings of computation can be achieved if $N = 2M$. Show that the N-point DFT of a segment $y_i[n]$ as defined in Eq. (P10.44-2) can be expressed as

$$Y_i[k] = X_i[k] + (-1)^k X_{i+1}[k], \qquad k = 0, 1, \ldots, N-1.$$

State a procedure for computing $\hat{\phi}_{xx}[m]$ for $0 \le m \le M-1$ that involves the computation of K N-point DFTs and one N-point inverse DFT. Determine the total number of complex multiplications in this case if a radix-2 FFT is used.

10.45. In Section 10.3 we defined the time-dependent Fourier transform of the signal $x[m]$ so that, for fixed n, it is equivalent to the regular DTFT of the sequence $x[n+m]w[m]$, where $w[m]$ is a window sequence. It is also useful to define a time-dependent autocorrelation function for the sequence $x[n]$ such that, for fixed n, its regular Fourier transform is the magnitude squared of the time-dependent Fourier transform. Specifically, the time-dependent autocorrelation function is defined as

$$c[n, m] = \frac{1}{2\pi}\int_{-\pi}^{\pi} |X[n, \lambda)|^2 e^{j\lambda m} d\lambda,$$

where $X[n, \lambda)$ is defined by Eq. (10.18).

(a) Show that if $x[n]$ is real

$$c[n, m] = \sum_{r=-\infty}^{\infty} x[n+r]w[r]x[m+n+r]w[m+r];$$

i.e., for fixed n, $c[n, m]$ is the aperiodic autocorrelation of the sequence $x[n+r]w[r]$, $-\infty < r < \infty$.

(b) Show that the time-dependent autocorrelation function is an even function of m for n fixed, and use this fact to obtain the equivalent expression

$$c[n, m] = \sum_{r=-\infty}^{\infty} x[r]x[r-m]h_m[n-r],$$

where

$$h_m[r] = w[-r]w[-(m+r)]. \tag{P10.45-1}$$

(c) What condition must the window $w[r]$ satisfy so that Eq. (P10.45-1) can be used to compute $c[n, m]$ for fixed m and $-\infty < n < \infty$ by causal operations?

(d) Suppose that

$$w[-r] = \begin{cases} a^r, & r \geq 0, \\ 0, & r < 0. \end{cases} \tag{P10.45-2}$$

Find the impulse response $h_m[r]$ for computing the m^{th} autocorrelation lag value, and find the corresponding system function $H_m(z)$. From the system function, draw the block diagram of a causal system for computing the m^{th} autocorrelation lag value $c[n, m]$ for $-\infty < n < \infty$ for the window of Eq. (P10.45-2).

(e) Repeat part (d) for

$$w[-r] = \begin{cases} ra^r, & r \geq 0, \\ 0, & r < 0. \end{cases}$$

10.46. Time-dependent Fourier analysis is sometimes implemented as a bank of filters, and even when FFT methods are used, the filter bank interpretation may provide useful insight. This problem examines that interpretation, the basis of which is the fact that when λ is fixed, the time-dependent Fourier transform $X[n, \lambda)$, defined by Eq. (10.18), is simply a sequence that can be viewed as the result of a combination of filtering and modulation operations.

(a) Show that $X[n, \lambda)$ is the output of the system of Figure P10.46-1 if the impulse response of the LTI system is $h_0[n] = w[-n]$. Show also that if λ is fixed, the overall system in Figure P10.46-1 behaves as an LTI system, and determine the impulse response and frequency response of the equivalent LTI system.

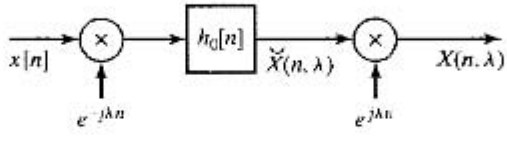


Figure P10.46-1

(b) Assuming λ fixed in Figure P10.46-1, show that, for typical window sequences and for fixed λ, the sequence $s[n] = \breve{X}[n, \lambda)$ has a lowpass DTFT. Show also that, for typical window sequences, the frequency response of the overall system in Figure P10.46 is a bandpass filter centered at $\omega = \lambda$.

(c) Figure P10.46-2 shows a bank of N bandpass filter channels, where each channel is implemented as in Figure P10.46-1. The center frequencies of the channels are $\lambda_k = 2\pi k/N$, and $h_0[n] = w[-n]$ is the impulse response of a lowpass filter. Show that the individual outputs $y_k[n]$ are samples (in the λ-dimension) of the time-dependent Fourier transform. Show also that the overall output is $y[n] = Nw[0]x[n]$; i.e., show that the system of Figure P10.46-2 reconstructs the input exactly (within a scale factor) from the sampled time-dependent Fourier transform.

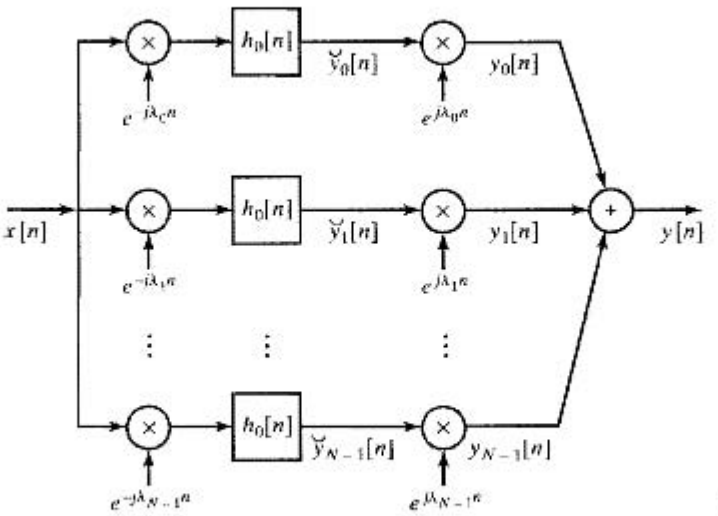


Figure P10.46-2

The system of Figure P10.46-2 converts the single input sequence $x[n]$ into N sequences, thereby increasing the total number of samples per second by the factor N. As shown in part (b), for typical window sequences, the channel signals $\breve{y}_k[n]$ have lowpass Fourier transforms. Thus, it should be possible to reduce the sampling rate of these signals, as shown in Figure P10.46-3. In particular, if the sampling rate is reduced by a factor $R = N$, the total number of samples per second is the same as for $x[n]$. In this case, the filter bank is said to be *critically sampled*. (See Crochiere and Rabiner, 1983.) Reconstruction of the original signal from the decimated channel signals requires interpolation as shown. Clearly, it is of interest to determine how well the original input $x[n]$ can be reconstructed by the system.

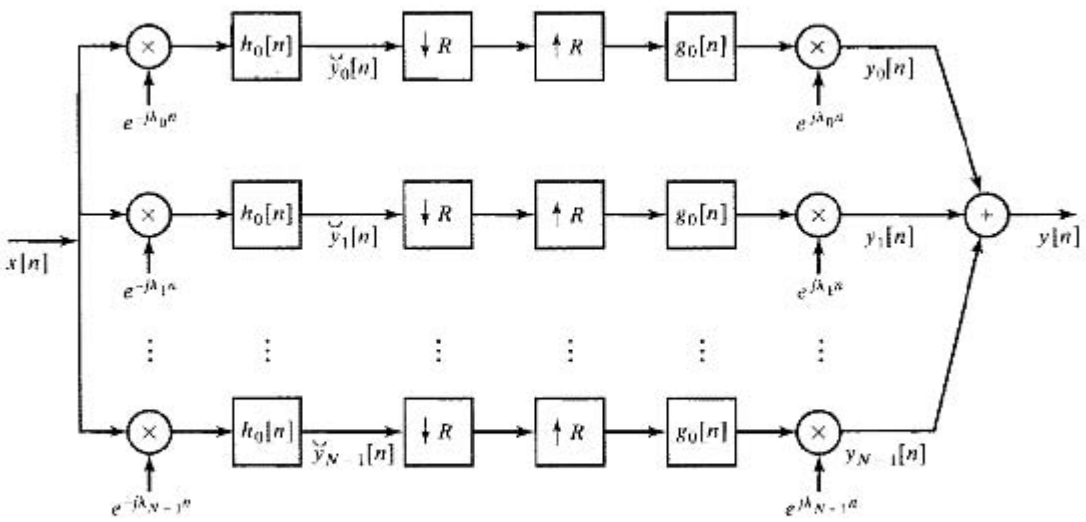


Figure P10.46-3

(d) For the system of Figure P10.46-3, show that the regular DTFT of the output is given by the relation

$$Y(e^{j\omega}) = \frac{1}{R}\sum_{\ell=0}^{R-1}\sum_{k=0}^{N-1} G_0(e^{j(\omega-\lambda_k)})H_0(e^{j(\omega-\lambda_k-2\pi\ell/R)})X(e^{j(\omega-2\pi\ell/R)}),$$

where $\lambda_k = 2\pi k/N$. This expression clearly shows the aliasing resulting from the decimation of the channel signals $\hat{y}[n]$. From the expression for $Y(e^{j\omega})$, determine a relation or set of relations that must be satisfied jointly by $H_0(e^{j\omega})$ and $G_0(e^{j\omega})$ such that the aliasing cancels and $y[n] = x[n]$.

(e) Assume that $R = N$ and the frequency response of the lowpass filter is an ideal lowpass filter with frequency response

$$H_0(e^{j\omega}) = \begin{cases} 1, & |\omega| < \pi/N, \\ 0, & \pi/N < |\omega| \le \pi. \end{cases}$$

For this frequency response $H_0(e^{j\omega})$, determine whether it is possible to find a frequency response of the interpolation filter $G_0(e^{j\omega})$ such that the condition derived in part (d) is satisfied. If so, determine $G_0(e^{j\omega})$.

(f) *Optional:* Explore the possibility of exact reconstruction when the frequency response of the lowpass filter $H_0(e^{j\omega})$ (the Fourier transform of $w[-n]$) is nonideal and nonzero in the interval $|\omega| < 2\pi/N$.

(g) Show that the output of the system of Figure P10.46-3 is

$$y[n] = N\sum_{r=-\infty}^{\infty} x[n-rN]\sum_{\ell=-\infty}^{\infty} g_0[n-\ell R]h_0[\ell R + rN - n].$$

From this expression, determine a relation or set of relations that must be satisfied jointly by $h_0[n]$ and $g_0[n]$ such that $y[n] = x[n]$.

(h) Assume that $R = N$ and the impulse response of the lowpass filter is

$$h_0[n] = \begin{cases} 1, & -(N-1) \le n \le 0, \\ 0, & \text{otherwise.} \end{cases}$$

For this impulse response $h_0[n]$, determine whether it is possible to find an impulse response of the interpolation filter $g_0[n]$ such that the condition derived in part (g) is satisfied. If so, determine $g_0[n]$.

(i) *Optional:* Explore the possibility of exact reconstruction when the impulse response of the lowpass filter $h_0[n] = w[-n]$ is a tapered window with length greater than N.

10.47. Consider a stable LTI system with a real input $x[n]$, a real impulse response $h[n]$, and output $y[n]$. Assume that the input $x[n]$ is white noise with zero mean and variance σ_x^2. The system function is

$$H(z) = \frac{\sum_{k=0}^{M} b_k z^{-k}}{1 - \sum_{k=1}^{N} a_k z^{-k}},$$

where we assume the a_ks and b_ks are real for this problem. The input and output satisfy the following difference equation with constant coefficients:

$$y[n] = \sum_{k=1}^{N} a_k y[n-k] + \sum_{k=0}^{M} b_k x[n-k].$$

If all the a_ks are zero, $y[n]$ is called a *moving-average* (MA) linear random process. If all the b_ks are zero, except for b_0, then $y[n]$ is called an *autoregressive* (AR) linear random process. If both N and M are nonzero, then $y[n]$ is an *autoregressive moving-average* (ARMA) linear random process.

(a) Express the autocorrelation of $y[n]$ in terms of the impulse response $h[n]$ of the linear system.

(b) Use the result of part (a) to express the power density spectrum of $y[n]$ in terms of the frequency response of the system.

(c) Show that the autocorrelation sequence $\phi_{yy}[m]$ of an MA process is nonzero only in the interval $|m| \leq M$.

(d) Find a general expression for the autocorrelation sequence for an AR process.

(e) Show that if $b_0 = 1$, the autocorrelation function of an AR process satisfies the difference equation

$$\phi_{yy}[0] = \sum_{k=1}^{N} a_k \phi_{yy}[k] + \sigma_x^2,$$

$$\phi_{yy}[m] = \sum_{k=1}^{N} a_k \phi_{yy}[m-k], \qquad m \geq 1.$$

(f) Use the result of part (e) and the symmetry of $\phi_{yy}[m]$ to show that

$$\sum_{k=1}^{N} a_k \phi_{yy}[|m-k|] = \phi_{yy}[m], \qquad m = 1, 2, \ldots, N.$$

It can be shown that, given $\phi_{yy}[m]$ for $m = 0, 1, \ldots, N$, we can always solve uniquely for the values of the a_ks and σ_x^2 for the random-process model. These values may be used in the result in part (b) to obtain an expression for the power density spectrum of $y[n]$. This approach is the basis for a number of parametric spectrum estimation techniques. (For further discussion of these methods, see Gardner, 1988; Kay, 1988; and Marple, 1987.)

10.48. This problem illustrates the basis for an FFT-based procedure for interpolating the samples (obtained at a rate satisfying the Nyquist theorem) of a periodic continuous-time signal. Let

$$x_c(t) = \frac{1}{16} \sum_{k=-4}^{4} \left(\frac{1}{2}\right)^{|k|} e^{jkt}$$

be a periodic signal that is processed by the system in Figure P10.48.

(a) Sketch the 16-point sequence $G[k]$.

(b) Specify how you would change $G[k]$ into a 32-point sequence $Q[k]$ so that the 32-point inverse DFT of $Q[k]$ is a sequence

$$q[n] = \alpha x_c\left(\frac{n2\pi}{32}\right), \qquad 0 \leq n \leq 31,$$

for some nonzero constant α. You need not specify the value of α.

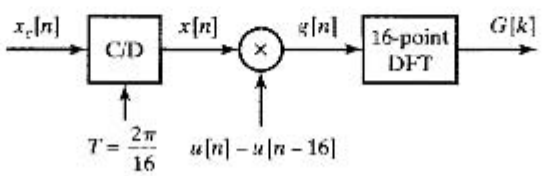


Figure P10.48

10.49. In many real applications, practical constraints do not allow long time sequences to be processed. However, significant information can be gained from a windowed section of the sequence. In this problem, you will look at computing the Fourier transform of an infinite-duration signal $x[n]$, given only a block of 256 samples in the range $0 \le n \le 255$. You decide to use a 256-point DFT to estimate the transform by defining the signal

$$\hat{x}[n] = \begin{cases} x[n], & 0 \le n \le 255, \\ 0, & \text{otherwise,} \end{cases}$$

and computing the 256-point DFT of $\hat{x}[n]$.

(a) Suppose the signal $x[n]$ came from sampling a continuous-time signal $x_c(t)$ with sampling frequency $f_s = 20$ kHz; i.e.,

$$x[n] = x_c(nT_s),$$

$$1/T_s = 20 \text{ kHz}.$$

Assume that $x_c(t)$ is bandlimited to 10 kHz. If the DFT of $\hat{x}[n]$ is written $\hat{X}[k]$, $k = 0, 1, \ldots, 255$, what are the continuous-time frequencies corresponding to the DFT indices $k = 32$ and $k = 231$? Be sure to express your answers in Hertz.

(b) Express the DTFT of $\hat{x}[n]$ in terms of the DTFT of $x[n]$ and the DTFT of a 256-point rectangular window $w_R[n]$. Use the notation $X(e^{j\omega})$ and $W_R(e^{j\omega})$ to represent the DTFTs of $x[n]$ and $w_R[n]$, respectively.

(c) Suppose you try an averaging technique to estimate the transform for $k = 32$:

$$X_{\text{avg}}[32] = \alpha \hat{X}[31] + \hat{X}[32] + \alpha \hat{X}[33].$$

Averaging in this manner is equivalent to multiplying the signal $\hat{x}[n]$ by a new window $w_{\text{avg}}[n]$ before computing the DFT. Show that $W_{\text{avg}}(e^{j\omega})$ must satisfy

$$W_{\text{avg}}(e^{j\omega}) = \begin{cases} 1, & \omega = 0, \\ \alpha, & \omega = \pm 2\pi/L, \\ 0, & \omega = 2\pi k/L, \quad \text{for } k = 2, 3, \ldots, L-2, \end{cases}$$

where $L = 256$.

(d) Show that the DTFT of this new window can be written in terms of $W_R(e^{j\omega})$ and two shifted versions of $W_R(e^{j\omega})$.

(e) Derive a simple formula for $w_{\text{avg}}[n]$, and sketch the window for $\alpha = -0.5$ and $0 \le n \le 255$.

10.50. It is often of interest to zoom in on a region of a DFT of a signal to examine it in more detail. In this problem, you will explore two algorithms for implementing this process of obtaining additional samples of $X(e^{j\omega})$ in a frequency region of interest.

Suppose $X_N[k]$ is the N-point DFT of a finite-length signal $x[n]$. Recall that $X_N[k]$ consists of samples of $X(e^{j\omega})$ every $2\pi/N$ in ω. Given $X_N[k]$, we would like to compute N samples of $X(e^{j\omega})$ between $\omega = \omega_c - \Delta\omega$ and $\omega = \omega_c + \Delta\omega$ with spacing $2\Delta\omega/N$, where

$$\omega_c = \frac{2\pi k_c}{N}$$

and

$$\Delta\omega = \frac{2\pi k_\Delta}{N}.$$

This is equivalent to zooming in on $X(e^{j\omega})$ in the region $\omega_c - \Delta\omega < \omega < \omega_c + \Delta\omega$. One system used to implement the zoom is shown in Figure P10.50-1. Assume that $x_z[n]$ is zero-padded as necessary before the N-point DFT and $h[n]$ is an ideal lowpass filter with a cutoff frequency $\Delta\omega$.

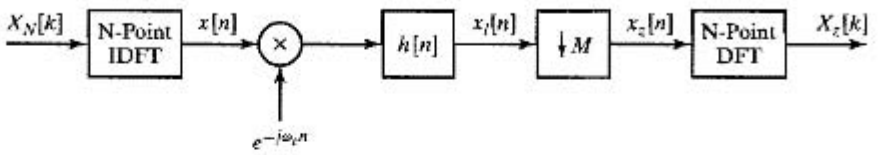


Figure P10.50-1

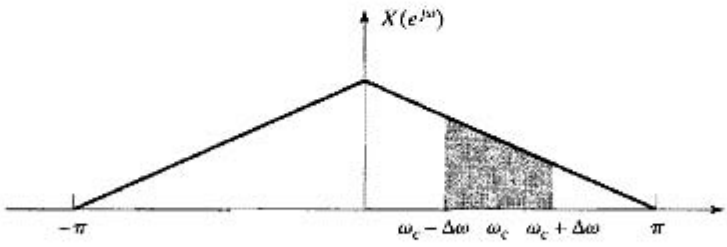


Figure P10.50-2

(a) In terms of k_Δ and the transform length N, what is the largest (possibly noninteger) value of M that can be used if aliasing is to be avoided in the downsampler?

(b) Consider $x[n]$ with the Fourier transform shown in Figure P10.50-2. Using the maximum value of M from part (a), sketch the Fourier transforms of the intermediate signals $x_1[n]$ and $x_2[n]$ when $\omega_c = \pi/2$ and $\Delta\omega = \pi/6$. Demonstrate that the system provides the desired frequency samples.

Another procedure for obtaining the desired samples can be developed by viewing the finite-length sequence $X_N[k]$ indexed on k as a discrete-time data sequence to be processed as shown in Figure P10.50-3. The impulse response of the first system is

$$p[n] = \sum_{r=-\infty}^{\infty} \delta[n + rN],$$

and the filter $h[n]$ has the frequency response

$$H(e^{j\omega}) = \begin{cases} 1, & |\omega| \le \pi/M, \\ 0, & \text{otherwise.} \end{cases}$$

The zoomed output signal is defined as

$$X_z[n] = \tilde{X}_{NM}[Mk_c - Mk_\Delta + n], \quad 0 \le n \le N - 1,$$

for appropriate values of k_c and k_Δ. Assume that k_Δ is chosen so that M is an integer in the following parts.

(c) Suppose that the ideal lowpass filter $h[n]$ is approximated by a causal Type I linear-phase filter of length 513 (nonzero for $0 \le n \le 512$). Indicate which samples of $\tilde{X}_{NM}[n]$ provide the desired frequency samples.

(d) Using sketches of a typical spectrum for $X_N[k]$ and $X(e^{j\omega})$, demonstrate that the system in Figure P10.50-3 produces the desired samples.

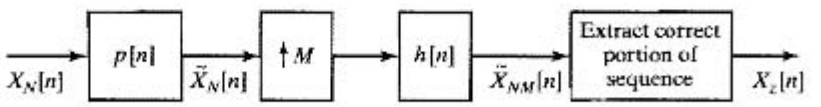


Figure P10.50-3