

**FIGURE 5.2** Autocorrelation functions of the acoustic backscatter data in Table 5.1. The thick line is for acoustic Beam 1, the dashed line for acoustic Beam 2.

for  $N' \leq N - 1$  gives a measure of the dominant correlation timescale within a data series; for times longer than  $T^*$ , the data become decorrelated. There are roughly  $\Delta\tau N/T^*$  actual degrees of freedom (DoF) within the time series. In reality, the summation typically is limited to  $N' \ll N$  since low frequency components within the time series prevent the summation from converging to a constant value over the finite length of the record. In general, one should continue the summation until it reaches a near-constant value, which we take as the value for  $T^*$ . If no plateau is reached within a reasonable number of lags, no integral timescale exists. In that case, the integral timescale can be approximated by integrating only to the first zero crossing of the autocorrelation function (cf. Poulain and Niiler, 1989).

### 5.3.6 Correlation Analysis vs Linear Regression

Geophysical data are typically obtained from random temporal sequences or spatial fields that cannot be regarded as mutually independent. Because the data series depend on time and/or spatial coordinates, the use of linear regression to study relationships between data

series may lead to incomplete or erroneous conclusions. As an example, consider two time series: A white-noise series, consisting of identically distributed and mutually independent random variables, and the same series but with a time shift. As the values of the time series are statistically independent, the cross-correlation coefficient will be zero at zero lag, even though the time series are strictly linearly related. Regression analysis would show no relationship between the two series. However, cross-correlation analysis would reveal the linear relationship (a coefficient of unity) for a lag equal to the time shift. Correlation analysis is often a better way to study relations among time series than traditional regression analysis.

## 5.4 SPECTRAL ANALYSIS

Spectral analysis is used to partition the variance of a time series as a function of frequency. For stochastic time series such as wind waves, contributions from the different frequency components are measured in terms of the *power spectral density* (PSD). For deterministic waveforms such as surface tides, either the PSD or the *energy*

spectral density (ESD) can be used. Here, power is defined as energy per unit time. The need for two different spectral definitions lies in the boundedness of the integral of signal variance for increasing record length. In practice, the term *spectrum* is applied to all spectral functions including commonly used terms such as auto-spectrum and power spectrum. The term *cross-spectrum* is reserved for the “shared” power between two coincident time series. We also distinguish between *nonparametric* and *parametric* spectral methods. Nonparametric methods, which are based on conventional Fourier transforms, are not data-specific while parametric techniques are data-specific and assign a predetermined model to the time series. In general, we use parametric methods for short time series (few cycles of the oscillations of interest) and nonparametric methods for long time series (many cycles of the oscillations of interest).

The word spectrum is a carryover from optics. The “colors” red, white, and blue of the electromagnetic spectrum are often used to describe the frequency distribution of oceanographic spectra. A spectrum whose spectral density decreases with increasing frequency is called a “red” spectrum, by analogy to visible light where red corresponds to longer wavelengths (lower frequencies). Similarly, a spectrum whose magnitude increases with frequency is called a “blue” spectrum. A “white” spectrum is one in which the spectral constituents have near-equal amplitude throughout the frequency range. In the ocean, long-period variability (periods greater than several days) tends to have red spectra while instrument noise tends to have white spectra. Blue spectra are confined to certain frequency bands such as the low-frequency portion of wind–wave spectra and within the “weather band” ( $2 < \text{period} < 10$  days) for deep wind-generated currents. Spectra of wind-generated inertial currents in the deep ocean are often “blue-shifted” to frequencies a few percent higher than the local inertial frequency (Fu, 1981; Thomson et al., 1990).

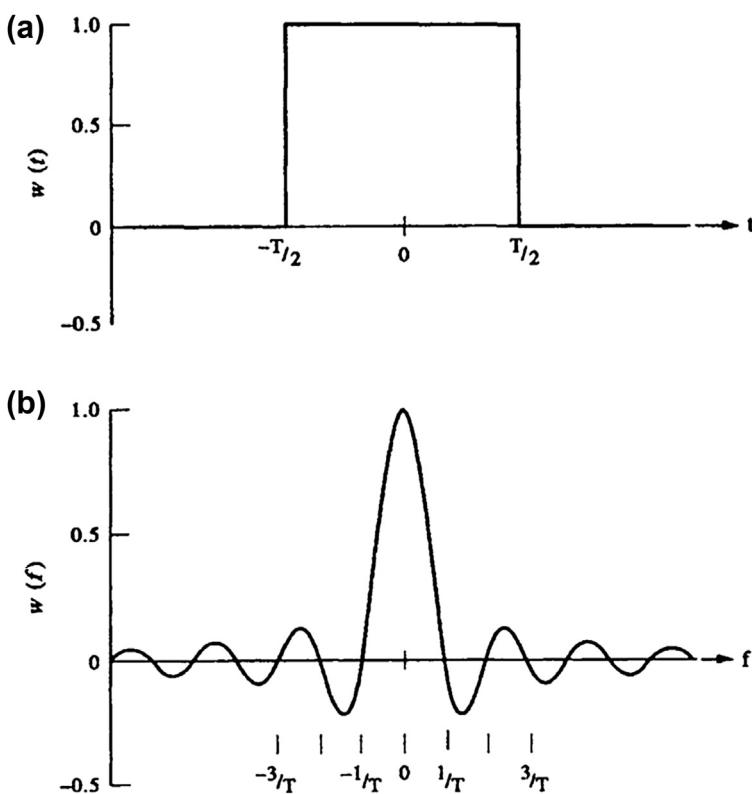
In the days before modern computers, it was customary to compute the spectrum of discrete oceanic data from the Fourier transform of the autocorrelation function using a small number of lag intervals, or “lags”. First formalized by Blackman and Tukey (1958), the autocorrelation method lacks the wide range of optional improvements to the computations and generalized “tinkering” permitted by more modern techniques. From a historical perspective, the autocorrelation approach has importance for the direct mathematical link it provides through the Wiener–Khinchin relations that link variance functions in the time domain to those in the frequency domain. Today, it is the spectral *periodogram* generated using the FFT or the Singleton Fourier transform that is most commonly used to estimate oceanic spectra. (We assume that the reader has a basic understanding of Fourier analysis and FFTs. Those unfamiliar with these concepts can proceed to [Section 5.8](#) where the topics are discussed in considerable detail.)

Other methods have been developed over the years as a result of fundamental performance limitations with the periodogram method. These limitations are: (1) restricted frequency resolution when distinguishing between two or more signals, with frequency resolution dictated by the available record length independent of the characteristics of the data or its signal-to-noise ratio (SNR); (2) energy “leakage” between the main lobe of a spectral estimate and adjacent side-lobes, with a resulting distortion and smearing of the spectral estimates, suppression of weak signals, and the need to use smoothing windows; (3) an inability to adequately determine the spectral content of short time series; and (4) an inability to adjust to rapid changes in signal amplitude or phase. Other techniques such as the maximum entropy method (MEM, best suited to short time series) and the wavelet transform (best suited to event-like signals and signals whose frequency content changes over time) are addressed in this chapter.

**Fundamental concepts:** Several basic concepts are woven into the fabric of this chapter. First of all, the sample data we collect are subsets of either stochastic or deterministic processes. Deterministic processes are predictable, stochastic ones are not. Secondly, the very act of sampling to generate a time series of finite duration is analogous to viewing an infinitely long time series through a narrow “window” in the shape of a rectangular box-car function (Figure 5.3(a)). The characteristics of this window in the frequency domain can severely distort the frequency content of the original data series from which the sample has been drawn. As illustrated by Figure 5.3(b), the sampling process results in spectral energy being “rippled” away from one

frequency (the central lobe of the response function) to a wide number range of adjacent frequencies. The large side-lobes of the rectangular window are responsible for the leakage of spectral energy from the central frequency to nearby frequencies.

A third point is that the spectra of random processes are themselves random processes. Therefore, if we are to determine the frequency content of a data series with some degree of statistical reliability (i.e., to be able to put confidence intervals on spectral peaks), we need to precondition the time series and average the raw periodogram estimates. Averaging can be done in the time domain by using specially designed windows or in the frequency domain



**FIGURE 5.3** The box-car (rectangular) window, which creates a sample time series from a “long” time series. (a) The box-car window in the time ( $t$ ) domain. Here,  $w(t) = 1$ ,  $-T/2 \leq t \leq T/2$ , and  $w = 0$  otherwise. (b) Frequency  $W(f)$  response of the box-car window in (a). The central lobe straddles each spectral (frequency) component within the time series and has a width,  $\Delta f = 2/T$ . Zeros occur at  $f = \pm m/T$ , where  $m = 1, 2, \dots$

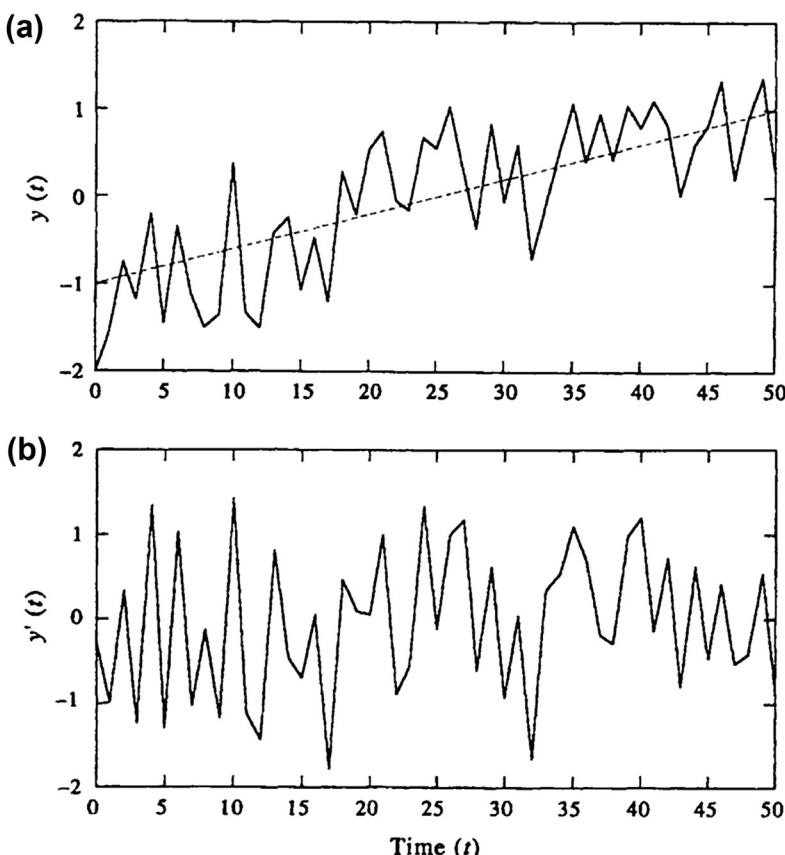
by averaging together adjacent spectral estimates. Windows (which are discussed in detail in [Section 5.4.6](#)) suppress Gibbs' phenomenon associated with finite length data series and enable us to increase the number of *degrees of freedom* (DoF) used in each spectral estimate. (Here, the term "degrees of freedom" refers to the number of statistically independent variables or values used in a particular estimate. We note that the spectral values are chi-square variables and that the DoF now apply to that PDF.) We can also improve spectral estimates by partitioning a time series into a series of segments and then conducting spectral analysis on the separate pieces. Spectral values in each frequency band for each piece are then averaged as a block to improve statistical reliability. This is similar to averaging adjacent spectral values in the periodogram, which will give a similar increase in the DoF of the resulting spectral estimate. The penalty for doing this is a loss in frequency resolution. The alternative—calculating a single periodogram and then smoothing in the frequency domain—suffers the same loss of frequency resolution for a smoothing that gives the same DoF.

Regardless of which averaging approach we choose, the results will be tantamount to viewing the data through another window in the frequency domain. Any smoothing window used to improve the reliability of the spectral estimates will again distort the results and impose structure on the data, such as periodic behavior, when no such structure may exist in the original time series. In addition, conventional methods make the implicit assumption that the unobserved data or correlation lag values situated outside the measurement interval are zero, which is generally not the case. The smoothing window results in smeared spectral estimates. The more modern parametric methods allow us to make more realistic assumptions about the nature of the process outside the measurement interval, other than to assume it is zero or cyclic. This eliminates the need for window

functions. The improvement over conventional FFT spectral estimates can be quite dramatic, especially for short records. However, even then, there remain pitfalls, which have tended to detract from the usefulness of these methods to oceanography. Each new method has its own advantages and disadvantages that must be weighed in context of the particular data set and the way it has been collected. For time series with low SNR, most of the modern methods are no better than the conventional FFT approach.

*Means and trends:* Prior to spectral analysis, the record mean and trend are generally removed from any time series ([Figure 5.4](#)). Unless stated otherwise, we will assume that the time series  $y(t)$  we wish to process has the form  $y'(t) = y(t) - \bar{y}(t)$ , where  $\bar{y}(t) = y_0 + \alpha t$  is the mean value and  $\alpha t$  is the linear trend ( $y_0$  and  $\alpha$  are constants). If the mean and trend are not removed prior to spectral analysis, they can distort the low-frequency components of the spectrum (see [Section 5.4.12](#)). Packaged spectral programs often include record mean and linear trend removal as part of the data preconditioning. Nonlinear trends are more difficult to remove, especially since a single function may not be appropriate for the entire data domain. The latter may apply also to linear trends.

The mean value removed from a record is not always the average for the entire record. For example, to examine interannual variability in the monthly time series of sea-level height, alongshore current velocity, or any other scalar,  $\eta(t_m)$ , we first calculate the mean monthly values  $\bar{\eta}(t_m)$  for each month separately over the entire record (e.g., the individual means for all Januaries, all Februaries, etc.). These mean monthly values for  $m = 1, 2, \dots, 12$ , rather than the simple average value for all values over the entire record, are then subtracted from the original data for the appropriate month to obtain monthly anomalies,  $\eta'(t_m) = \eta(t_m) - \bar{\eta}(t_m)$ . As with other averaging processes, the user will need to determine how many missing data values will be



**FIGURE 5.4** Mean and trend removal for an artificial time series  $y(t)$ . Here,  $y_0 = -1.0$ , the trend,  $\alpha = 0.025$ , and the fluctuating component,  $y'(t)$ , was obtained using a uniformly distributed random number generator. (a) Original time series, showing the linear trend; (b) Time series with the mean and linear trend removed.

permitted for a given month before the monthly value is considered “missing” or not available. Trend removal can then be applied to the monthly anomalies to obtain the final anomaly record. As a final comment, we note that certain records, such as those from moored bottom pressure recorders and near-surface transmissometers or dissolved oxygen sensors, will contain long-term nonlinear trends that should be removed from the data record prior to spectral analysis. However, this must be done cautiously. Unless one has a justified physical model for a particular trend (including a linear trend), removal of the trend may itself

add spurious frequency components to the detrended signal.

#### 5.4.1 Spectra of Deterministic and Stochastic Processes

Time series data can originate with deterministic or stochastic processes, or a mixture of the two. Turbulence arising from eddy-like motions generated by strong tidal currents in a narrow coastal channel provides an example of mixed deterministic and stochastic processes. To see the difference between the two types of processes in terms of conventional spectral estimation,

consider the case of a continuous *deterministic* signal,  $y(t)$ . If the total signal energy,  $E$ , is finite

$$E = \int_{-\infty}^{\infty} |y(t)|^2 dt < \infty \quad (5.17)$$

then  $y(t)$  is absolute-integrable over the entire domain and the Fourier transform  $Y(f)$  of  $y(t)$  exists. This leads to the standard transform pair

$$Y(f) = \int_{-\infty}^{\infty} y(t)e^{-i2\pi ft} dt \quad (5.18a)$$

$$y(t) = \int_{-\infty}^{\infty} Y(f)e^{i2\pi ft} df = \frac{1}{2\pi} \int_{-\infty}^{\infty} Y(\omega)e^{i\omega t} d\omega \quad (5.18b)$$

where  $e^{\pm i2\pi ft} = \cos(2\pi ft) \pm i\sin(2\pi ft)$ ,  $f$  is the frequency in cycles per unit time, and  $\omega = 2\pi f$  is the angular frequency in radians per unit time. The square of the modulus of the Fourier transform for all frequencies

$$S_E(f) = Y(f)Y^*(f) = |Y(f)|^2 \quad (5.19)$$

is then the Energy Spectral Density (ESD),  $S_E(f)$ , of  $y(t)$ . (As usual, the asterisk denotes the complex conjugate.) To show that Eqn (5.19) is an energy density, we use Parseval's theorem

$$\int_{-\infty}^{\infty} |y(t)|^2 dt = \int_{-\infty}^{\infty} |Y(f)|^2 df \quad (5.20)$$

which states that the total energy,  $E$ , of the signal in the time domain is equal to the total energy,  $\int S_E(f)df$ , of the signal in the frequency domain. Thus,  $S_E(f)$ , is an energy density (energy per unit frequency) which, when multiplied by  $df$ , yields a measure of the total signal energy in the frequency band centered near frequency  $f$ . The "power" of a deterministic signal,  $E/T$ , is zero in the limit of very long time series ( $T \rightarrow \infty$ ).

Now, suppose that  $y(t)$  is a stationary random process rather than a deterministic waveform. Unlike the case for the finite energy deterministic

signal, the total energy in the stochastic process is unbounded (the characteristics of the process remain unchanged over time) and functions of the form (Eqn (5.18)) do not exist. In other words, the Fourier transform method introduced earlier fails in the sense that the total energy, as defined by Eqn (5.17), does not decrease as the length of the time series increases without bound. To get around this problem, we must deal with the frequency distribution of the signal *power* (the time average of energy or energy per unit time,  $E/T$ ), which is a bounded function. The basis for spectral analysis of random processes is the autocorrelation function  $R_{yy}(\tau) = E[y(t)y(t+\tau)]$ . Using the Wiener–Khinchin relation, the PSD,  $S(f)$ , becomes

$$S(f) = \int_{-\infty}^{\infty} R_{yy}(\tau)e^{-i2\pi f\tau} d\tau \quad (5.21a)$$

For an ergodic random process, for which ensemble averages can be replaced by time averages,  $R_{yy}$  has the from

$$R_{yy}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} [y(t)y^*(t+\tau)]dt \quad (5.21b)$$

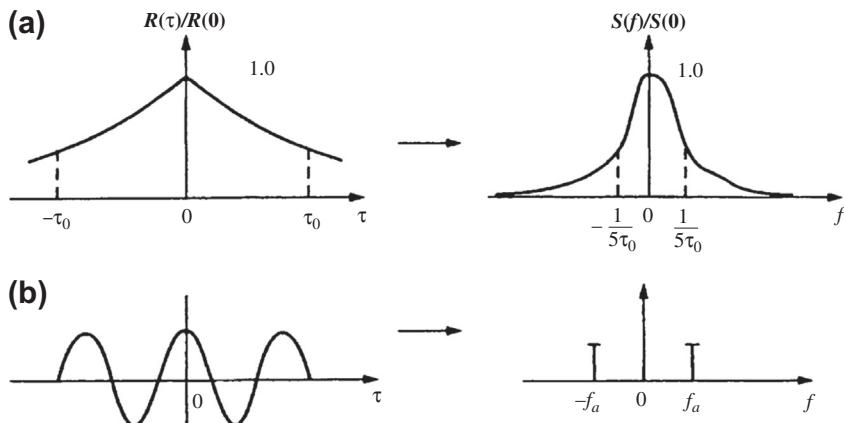
By definition, the energy and PSD functions quantify the signal variance per unit frequency. For example, in the case of a stationary random process, integration of  $S(f)$  gives the relation

$$s^2 = \int_{f-\Delta f/2}^{f+\Delta f/2} S(f)df \quad (5.22)$$

where  $s^2$  is the integrated signal variance in the narrow frequency range  $\Delta f = [f - \frac{1}{2}\Delta f, f + \frac{1}{2}\Delta f]$ . If we assume that the spectrum is nearly uniform over this frequency range, we find

$$S(f) \approx \frac{s^2}{\Delta f} \quad (5.23)$$

which defines the spectrum for a stochastic processes in terms of a power density, or variance



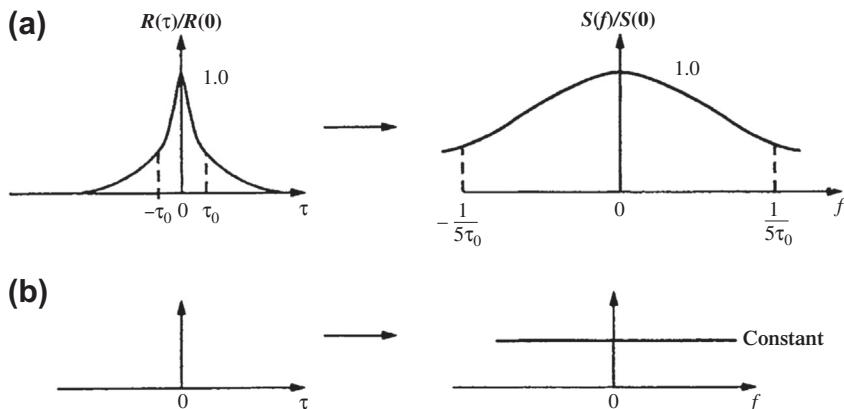
**FIGURE 5.5** Examples of slowly decaying autocorrelation functions,  $R(\tau)$ , as a function of time lag,  $\tau$ . Functions are normalized by their peak values. (a) The correlation function for a highly correlated signal leads to a relatively narrow power spectra distribution,  $S(f)$ ; (b) The case for autocorrelation,  $R(\tau) \approx \cos(2\pi f_a \Delta t)$  for a single frequency component,  $f_a$ , and corresponding line spectra at frequencies  $\pm f_a$ . (From Konyaev (1990).)

per unit frequency. The product  $S(f) \cdot \Delta f$  is the total signal variance within the frequency band  $\Delta f$  centered at frequency  $f$ .

At this point, there are several other basic concepts worth mentioning. First of all, a waveform whose autocorrelation function  $R(\tau)$  attenuates slowly with time lag,  $\tau$ , will have a narrow spectral distribution (Figure 5.5(a)) indicating that there are relatively few frequency components to destructively interfere with one another as  $\tau$  increases from zero. In the limiting case of only one frequency component,  $f_a$ , we find  $R(\tau) \approx \cos(2\pi f_a \tau)$  and Fourier *line spectra* appear at frequencies  $\pm f_a$  (Figure 5.5(b)). Because they consist of near monotone signals, tidal motions are highly autocorrelated and produce sharp spectral lines. In contrast, a rapidly decaying autocorrelation function implies a broad spectral distribution (Figure 5.6(a)) and a large number of frequency components in the original waveform. In the limit  $R(\tau) \rightarrow \delta(\tau)$  (Figure 5.6(b)), there is an infinite number of equal-amplitude frequency components in the waveform and the spectrum  $S(f) \rightarrow \text{constant}$  (white spectrum).

Figure 5.7 provides an example of time series data generated by the relation  $y(k) = \text{Acos}(2\pi nk/N)$

+  $\epsilon(k)$ , where  $k = 0, \dots, N$  is time in units of  $\Delta t = 1$ ,  $n/(N\Delta t) = 0.25$  is the frequency in units of  $\Delta t^{-1}$ , and  $\epsilon(k)$  is a random number between -1 and +1. (We will often use this type of generic example rather than a specific example from the oceanographic literature. That way, readers can directly compare their computational results with ours.) In the present case, if we set  $\Delta t = 1$  day, then the time series  $y(k)$  could represent east–west current velocity oscillations of a synoptic (3- to 10-day) period associated with wind-forced motions (cf. Cannon and Thomson, 1996). Here, we set  $A = 1$  and  $\epsilon(k) \neq 0$  for mostly deterministic data (Figure 5.7(a)) and  $A = 0$  for random data (Figure 5.7(b)). In the analysis, the record has been padded with zeroes up to time  $k = 2N$ . For the mostly deterministic case, the noise causes partial decorrelation of the signal with lag, but the spectral peak remains prominent. For the purely random case, the spectrum resembles white noise but with isolated spectral peaks that one might mistake as originating with some physical process. The latter result is a good example of why we need to attach confidence limits to the peaks of spectral estimates (see Section 5.4.8). It is disconcerting to see the number



**FIGURE 5.6** As for Figure 5.5 but for rapidly decaying autocorrelation functions,  $R(\tau)$ . (a) Correlation function for a weakly correlated signal leading to a broad power spectra density distribution. (b) The limiting case,  $R(\tau) \approx (2\pi f_a \Delta t)$ , and the related spectrum,  $S(f) = \text{constant}$  (a white spectrum). (From Konyaev (1990).)

of papers that are published in reputable journals that present spectra without including confidence intervals. At the same time, one must be cognizant of the meaning of the confidence intervals being presented. Use of a very low significance level might suggest a high degree of confidence in the results. However, this confidence can be rendered meaningless if a low significance level is selected. Significance levels of 95 or 99% are commonly accepted as “meaningful”.

#### 5.4.2 Spectra of Discrete Series

Consider an infinitely long time series  $y(t_n) = y_n$  sampled at equally spaced time increments  $t_n = n\Delta t$ , where  $\Delta t$  is the sampling interval and  $n$  is an integer,  $-\infty < n < \infty$ . From sampling theory, we know that a continuous representation of the discrete times series  $y_s(t)$ , can be represented as the product of the continuous time series  $y(t)$  with an infinite set of delta functions,  $\delta(t)$ , such that

$$\begin{aligned} y_s(t) &= y(t) \sum_{n=-\infty}^{\infty} \delta(t - n\Delta t) \\ &= y(t) \frac{\Xi(t/\Delta t)}{\Delta t} \end{aligned} \quad (5.24a)$$

where  $\Xi$  is the “sampling function”, for which the Fourier transform is

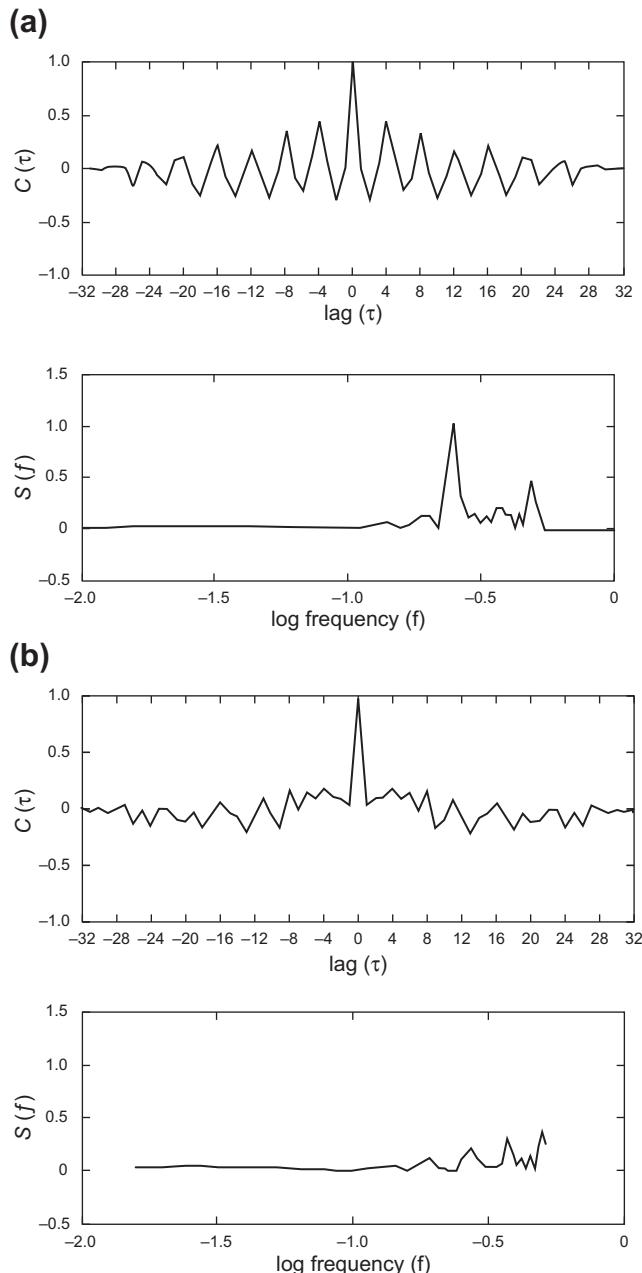
$$\begin{aligned} Y(f) &= \int_{-\infty}^{\infty} \left[ \sum_{n=-\infty}^{\infty} y(t) \delta(t - n\Delta t) \Delta t \right] e^{-i2\pi ft} dt \\ &= \Delta t \sum_{n=-\infty}^{\infty} y_n e^{-i2\pi f n} \end{aligned} \quad (5.24b)$$

In effect, the original time series is multiplied by a “picket fence” of delta functions  $\Xi(t/\Delta t) \approx \sum_{n=-\infty}^{\infty} \delta(t - n\Delta t)$ , which are zero everywhere except for the infinitesimal rectangular region occupied by each delta function (Figure 5.8(a) and (b)). Comparison of the above expression with Eqn (5.18) shows that retention of the time step  $\Delta t$  ensures conservation of the rectangular area in the two expressions as  $\Delta t \rightarrow 0$ . Provided that the time series  $y(t)$  has a limited number of frequencies (i.e., is band-limited), whereby all frequencies are contained in the Nyquist interval

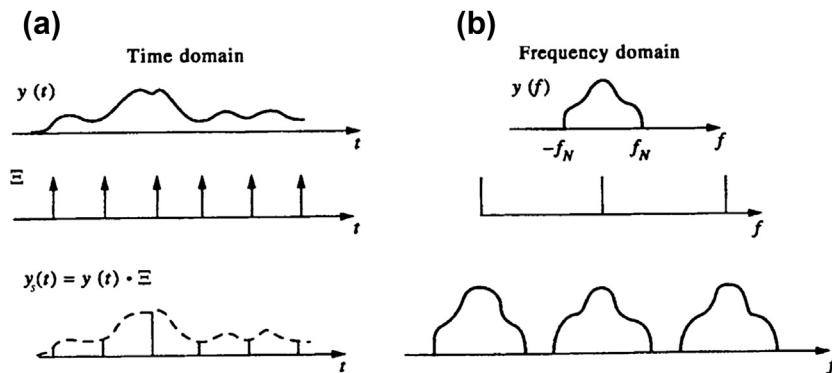
$$-f_N \leq f_k \leq f_N \quad (5.25)$$

in which  $f_N \equiv f_{\text{Nyquist}} = 1/(2\Delta t)$  is the Nyquist frequency, the ESD

$$S_E(f) = |Y(f)|^2 \quad (5.26)$$



**FIGURE 5.7** Autocovariance function,  $C(\tau)$ , and corresponding spectrum,  $S(f)$ , for the time series,  $y(k) = A \cos(2\pi nk/N) + \epsilon(k)$ ;  $k = 0, \dots, N$ ,  $\Delta t = 1$ ,  $n/N = 0.25$  is the frequency, and  $\epsilon(k)$  is a random number between  $-1$  and  $+1$ . (a)  $C(\tau)$  and  $S(f)$  for  $A = 1$  and  $\epsilon \neq 0$  (mostly deterministic data); and (b) for  $A = 0$  (purely random data). Records have been padded with zeros up to time  $k = 2N = 32$ .



**FIGURE 5.8** (a) A “picket fence” of delta functions,  $\delta(t - n\Delta t)$ , used to generate a discrete data series from a continuous time series. (b) The Fourier transform (schematic only) of the different functions. Here,  $y(f)$  denotes  $Y(f)$ .

is identical to that for a continuous function. Conversely, if  $Y(f) \neq 0$  for  $|f| > f_N$  then the sampled and original times series do not have the same spectrum for  $|f| < f_N$ . The spectrum Eqn (5.26) obtained by Fourier analysis of discrete time series is called a *periodogram* spectral estimate, a term first coined by Schuster (1898) in a study of sunspot cycles. (Note that we always use  $f_N$  for the Nyquist frequency; the subscript  $N$  for Nyquist should never be confused with the subscript  $N$  used in summations or the  $N$  used for degrees of freedom (DoF).)

Real oceanographic time series data are discrete and have finite duration,  $T = N\Delta t$ . Returning to Eqn (5.24), this means that the summation is over a limited range  $n = 1$  to  $N$ , and the spectral amplitude for the sample must be defined in terms of the discrete Fourier transform (DFT)

$$\begin{aligned} Y_k &= \Delta t \sum_{n=1}^N y_n e^{-i2\pi f_k n \Delta t} \\ &= \Delta t \sum_{n=1}^N y_n e^{-i2\pi k n / N}; \\ f_k &= k/N\Delta t, \quad k = 0, \dots, N \end{aligned} \quad (5.27)$$

The frequencies  $f_k$  are confined to the Nyquist interval, with positive frequencies,  $0 \leq f_k \leq f_N$ , corresponding to the range  $k = 0, \dots, N/2$  and

negative frequencies,  $-f_N \leq f_k \leq 0$ , to the range  $k = N/2, \dots, N$ . Since  $f_{N-k} = f_k$ , only the first  $N/2$  Fourier transform values are unique. Specifically,  $Y_k = Y_{N-k}$  so that we will generally confine our attention to the positive interval only.

The inverse Fourier transform (IFT) is defined as

$$y_n = \frac{1}{N\Delta t} \sum_{k=0}^{N-1} Y_k e^{i2\pi k n / N}, \quad n = 1, \dots, N \quad (5.28)$$

As indicated by Eqn (5.27), the Fourier transforms,  $Y_k$ , are specified for the discretized frequencies  $f_k$ , where  $f_k = kf_1$  and  $f_1 = 1/(N\Delta t) = 1/T$  characterizes both the fundamental frequency and the bandwidth,  $\Delta f$ , for the time series. The ESD for a discrete, finite-duration time series is then

$$S_E(f_k) = |Y_k|^2, \quad k = 0, \dots, N-1 \quad (5.29)$$

and Parseval's energy conservation theorem (5.20) becomes

$$\Delta t \sum_{n=1}^N |y_n|^2 = \Delta f \sum_{k=0}^{N-1} |Y_k|^2$$

where we have used  $\Delta f = 1/(N\Delta t)$ . A plot of  $|Y_k|^2$  vs frequency,  $f_k$ , gives the discrete form of the periodogram spectral estimate.

Any geophysical data set we collect is subject to discrete sampling and windowing. As noted earlier, a time series of geophysical data,  $y(t_n)$ , sampled at time steps  $\Delta t$  can be considered the product of an infinitely long time series with a rectangular window that spans the duration ( $T = N\Delta t$ ) of the measured data. The discrete spectrum  $S(f_k)$  is then the convolution of the true spectrum,  $S(f)$ , with the Fourier transform of the rectangular window (Figure 5.3(b)). Since the window allows us to see only a segment of the infinite time series, the spectrum  $S(f_k)$  provides a distorted picture of the actual underlying spectrum. This distortion, created during the Fourier transform of the rectangular window, consists of a broadening of the central lobe and leakage of power from the central lobe into the side lobes. (The “ripples” on either side of the central lobe in Figure 5.3(b) are side lobes.) A further problem is that the function  $Y_k$  and its Fourier transform now become periodic with period  $N$ , although the original infinite time series  $y(t)$ , of which our sample data are a subset, may have been nonperiodic.

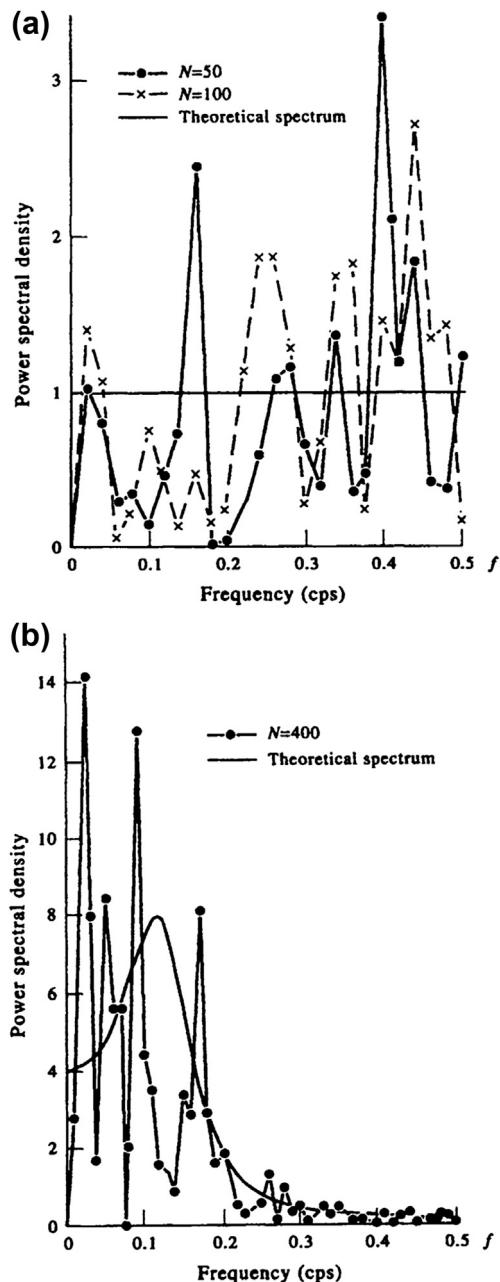
As noted in the previous section, the convergence of  $|Y(f)|^2$  to  $S(f)$  is smooth for deterministic functions in that the function  $|Y'(f)|^2$ , obtained by increasing the sample record length from  $T$  to  $T'$ , would be a smoother version of  $|Y(f)|^2$ . For stochastic signals, the function  $|Y'(f)|^2$  obtained from the longer time series ( $T'$ ) is just as erratic as the function for the shorter series. The sample spectra of a stochastic process do not converge in any statistical sense to a limiting value as  $T$  tends to infinity. Thus, the sample spectrum is not a consistent estimator in the sense that its PDF does not tend to cluster more closely about the true spectrum as the sample size increases. To show what we mean, consider the spectrum of a process consisting of  $N = 400$  random, normally distributed deviates (Gaussian white noise) sampled at 1 s intervals. (True white noise is a mathematical construct and is as physically impossible as the spike of an impulse function.) The highest frequency we can hope to

measure with these data is the Nyquist frequency,  $f_N = 0.5$  cps (cycles per second). The spectra computed from 50 and then from 100 values of the fully white-noise signal are presented in Figure 5.9(a). Also shown is the theoretical sample spectrum, corresponding to a uniform amplitude of 1.0. The shorter the sample used for the discrete spectral estimates, the greater the amplitude spikes in the power spectrum. This same tendency also is apparent in Table 5.2, which lists the means, variances, and MSEs (Mean Square Errors) computed from various subsamples of the white-noise signal. Here, MSE is defined as the variance plus bias of an estimator  $\hat{y}(t)$  of the true signal  $y(t)$ ; that is

$$\text{MSE} = E[(\hat{y} - y)^2] = V[\hat{y}] + B^2 \quad (5.30)$$

where  $B = E[\hat{y}] - y$  is the bias of the estimator. The mean is lower in both the  $N = 50$  and  $N = 400$  cases while it is greater in the case where  $N = 100$  and is exactly 1.0 for  $N = 200$ . The variance increases as  $N$  increases, as does the MSE. However, if this were a purely random discrete process (discrete white noise), the sample spectral estimator of the variance would be independent of the number of observations.

Now consider the spectrum of a second-order autoregressive process for a sample of  $N = 400$  measured at 1 s increments (Figure 5.9(b)). (An autoregressive process of order  $p$  is one in which the present value of  $y$  depends on a linear combination of the previous  $p$  values of  $y$ . See Section 5.5.2.) The Nyquist frequency is again 0.5 cps and the maximum bandwidth of the spectral resolution,  $\Delta f = 1/(N\Delta t)$ , is equal to 0.0025 cps. At higher frequencies, the sample spectrum appears to be a good estimator of the theoretical spectrum (the smooth solid line), while for the lower frequencies there are large spikes in the sample spectrum that are not characteristic of the true spectrum. This misleading appearance is largely a consequence of the fact that the theoretical spectrum has most of its energy at the lower frequencies. In reality, the computed raw spectrum



**FIGURE 5.9** Power spectra of discrete signals and their theoretical values. Frequency in cycles per second (cps); spectra are in units of amplitude-squared/cps. (a) Power spectrum for the first half ( $N = 50$ ) and full ( $N = 100$ ) realization of a discrete

**TABLE 5.2** Behavior of Sample Spectra of White Noise as the Record Length,  $N$ , is Increased

Record Length ( $N$ )	50	100	200	400
Mean	0.85	1.07	1.00	0.95
Variance	0.630	0.777	0.886	0.826
Mean square error	0.652	0.782	0.886	0.828

Units are arbitrary. (After Jenkins and Watts (1968).)

(i.e., with no smoothing) can fluctuate by 100% about the mean spectrum. The fluctuations are much smaller at higher frequencies simply because the actual spectral level is correspondingly smaller.

The basic reason that Fourier analysis breaks down when applied to real time series is that it is based on the assumption of fixed (stationary) amplitudes, frequencies, and phases (see Section 5.8). Stochastic series are instead characterized by random changes in frequency, amplitude, and phase. Thus, our treatment must be a statistical approach that makes it possible to accommodate these types of changes in our computation of the power spectrum.

#### 5.4.3 Conventional Spectral Methods

The two spectral estimation techniques founded on Fourier transform operations are the indirect autocorrelation approach popularized by Blackman and Tukey in the 1950s and the direct periodogram approach presently favored by the oceanographic community. The FFT is the most common algorithm for determining the periodogram. The autocorrelation approach is mainly discussed for completeness. These

normal white-noise process measured at 1-s intervals. (b) Power spectrum for one realization of a second-order autoregressive process of  $N = 400$  values measured at 1-s increments.  $f_N = 0.5$  cps is the Nyquist frequency and the maximum bandwidth of the spectral resolution,  $\Delta f = 1/N\Delta t = 0.0025$ /s. (From Jenkins and Watts (1968).)

methods fall into the category of nonparametric techniques, which are defined independently of any specific time series. Parametric techniques, described later in this chapter, make assumptions about the variability of the time series and rely on the series for parameter determination.

The following sections first describe the two conventional spectral analysis methods without providing details on how to improve spectral estimates. We wish to first outline the procedures for calculating spectra before describing how to improve the statistical reliability of the spectral estimates. Once this is done, we give a thorough description of windowing, frequency-band averaging, and other spectral improvement techniques.

#### 5.4.3.1 The Autocorrelation Method

In the Blackman–Tukey method, the autocovariance function,  $C_{yy}(\tau)$  (which equals the autocorrelation function,  $R_{yy}(\tau)$ , if the record mean has been removed), is first computed as a function of lag,  $\tau$ , and the Fourier transform of  $C_{yy}(\tau)$  used to obtain the PSD as a function of frequency. An unbiased estimator for the autocovariance function for a data set consisting of  $N$  equally spaced values  $\{y_1, y_2, \dots, y_N\}$  is

$$C_{yy}(\tau_m; N - m) = \frac{1}{N - m} \sum_{n=1}^{N-m} y_n y_{n+m} \quad (5.31a)$$

where  $m = 0, \dots, M$  is the number of lags ( $\tau_m = m\Delta t$ ) and  $M < N$ . In place of this estimator, some authors (cf. Kay and Marple, 1981) argue for the use of

$$C_{yy}(\tau_m; N) = \frac{1}{N} \sum_{n=1}^{N-m} y_n y_{n+m} \quad (5.31b)$$

which typically has a lower MSE than  $C_{yy}(\tau_m; N - m)$  for most finite data sets. Because  $E[C_{yy}(\tau_m; N)] = [(N - m)/N]C_{yy}(\tau_m; N - m)$ , the function  $C_{yy}(\tau; N)$  is a biased estimator for the autocovariance function. Despite this, we will often use the relation Eqn (5.31b) for the autocovariance function since it yields a PSD that is equivalent

to the PSD obtained from the direct application of the FFT, as discussed in the next section. Moreover, the weighting  $(N - m)/N$  acts like a triangular (Bartlett) smoothing window to help reduce spectral leakage. We will use Eqn (5.31a) when we want a “stand-alone” unbiased estimator of the covariance function, keeping in mind that this formulation gives largest weight to the most poorly determined components in the Fourier analysis, which is often not desirable despite the reduction in bias.

The one-sided PSD,  $G_k$ , for an autocovariance function with a total of  $M$  lags is found from the Fourier transform of the autocovariance function

$$G_k = 2\Delta t \sum_{m=0}^M C_{yy}(\tau_m) e^{-i2\pi km/M}, \quad k = 0, \dots, \frac{M}{2} \quad (5.32a)$$

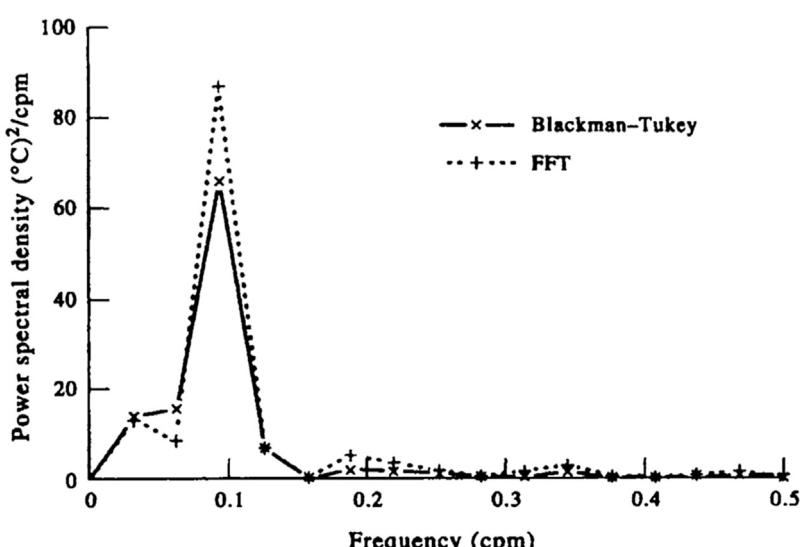
where  $\tau_m = m\Delta t$  and  $2\Delta t = 1/f_N$ . Since  $C_{yy}(\tau_m)$  is an even function, the spectrum of  $\{y_n\}$  can be calculated from the cosine transform

$$G_k = 2\Delta t \left[ C_{yy}(0) + 2 \sum_{m=1}^{\frac{M}{2}} C_{yy}(\tau_m) \cos\left(\frac{2\pi km}{N}\right) \right], \\ k = 0, \dots, \frac{M}{2} \quad (5.32b)$$

where  $G_k = 2S_k$  is centered at positive frequencies  $f_k = k/N\Delta t$  and the Nyquist interval  $0 \leq f_k \leq f_{Nyquist}$  is divided into  $N/2$  segments ( $N$  is even). For the two-sided spectrum,  $S_k$ , the first  $(N/2) + 1$  frequencies are identical to those for the one-sided spectrum and correspond to positive frequencies in the range  $0 \leq f_k \leq f_N$ . The last  $(N/2) - 1$  spectral values for the two-sided spectral density, defined for  $k = (N/2) + 1, (N/2) + 2, \dots, N - 1$ , correspond to spectral density estimates for negative frequencies in the range  $-f_N \leq f_k \leq 0$ .

The solid line in Figure 5.10 shows the spectrum of monthly mean SSTs derived from the cosine transform using the Blackman–Tukey autocorrelation method for the version (Eqn (5.31b)) of the autocovariance function. The temperature data span the 36-month period from

**FIGURE 5.10** Spectra ( $^{\circ}\text{C}$ ) $^2/\text{cpm}$  ( $\text{cpm} = \text{cycles per month}$ ) vs frequency (per month) for monthly mean sea surface temperatures collected at a coastal station in the northeast Pacific for the period January 1982–December 1984 (cf. Table 5.3). The solid line is the unsmoothed spectrum from the Blackman–Tukey autocorrelation method (the cosine transform of the autocovariance function Eqn (5.31b)); dashed line is the unsmoothed spectrum from the fast Fourier transform (FFT) method based on the first  $2^5$  (=32) data values. Spectral peaks span the annual period ( $f=0.083/\text{month}$ ).



January 1982 to December 1984 for Amphitrite Point (Table 5.3). Since we wish to compare the Blackman–Tukey spectrum in Figure 5.10 with that derived from the data series using a packaged FFT routine (the dashed line in Figure 5.10), the lags used to generate the Blackman–Tukey were computed for the first 32 ( $2^5$ ) points only, which is four fewer points than normally would be used in the Blackman–Tukey approach (see

Section 5.8 for a discussion of FFTs). In this case, artificially extending the lag correlation beyond 10–20% of the data, as recommended earlier, is a necessity if we are to obtain reasonable estimates of the spectra using the autocorrelation method. As expected, results reveal a strong spectral peak centered near, but not at, the annual frequency ( $f=1.0$  cycles per year = 0.083 cycles per month). There are too

**TABLE 5.3** Monthly Mean Sea Surface Temperatures SST ( $^{\circ}\text{C}$ ) at Amphitrite Point Lightstation ( $48^{\circ}55.16'\text{ N}$ ,  $125^{\circ}32.17'\text{ W}$ ) on the West Coast of Canada for January 1982 through December 1984

<b>YEAR 1982</b>												
<i>n</i>	1	2	3	4	5	6	7	8	9	10	11	12
SST	7.6	7.4	8.2	9.2	10.2	11.5	12.4	13.4	13.7	11.8	10.1	9.0
<b>YEAR 1983</b>												
<i>n</i>	13	14	15	16	17	18	19	20	21	22	23	24
SST	8.9	9.5	10.6	11.4	12.9	12.7	13.9	14.2	13.5	11.4	10.9	8.1
<b>YEAR 1984</b>												
<i>n</i>	25	26	27	28	29	30	31	32	33	34	35	36
SST	7.9	8.4	9.3	9.9	11.0	11.1	12.6	14.0	13.0	11.7	9.8	8.0

few data to enable us to accurately resolve the location of the frequency peak. In the present example, all spectral estimates are positive. However, the autocorrelation method can yield erroneous negative spectra for weak frequency components when there are gaps in the data record.

We emphasize that the spectra in Figure 5.10 have been constructed without any averaging or windowing. This means that each spectral estimate has the minimum possible two DoF (corresponding to the orthogonal sine and cosine components obtained from the Fourier transform) so that the error in each estimate is equal to the value of the estimate itself. Some form of averaging is needed if we are to place confidence limits on the spectra (see Sections 5.4.6 and 5.4.7). The two spectra are slightly different because the record used for the FFT method is shorter than that used for the autocovariance method.

#### 5.4.3.2 The Periodogram Method

The preferred method for estimating the PSD of a discrete sample  $\{y_1, y_2, \dots, y_N\}$  is the direct or periodogram method. Instead of first calculating the autocorrelation function, the data are transformed directly to obtain the Fourier components  $Y(f)$  using Eqn (5.27). To help avoid end effects (Gibbs' phenomenon) and wrap-around problems, the original time series can be padded with  $K \leq N$  zeroes after the mean has been removed from the time series. The padding will also increase the frequency resolution of the periodogram (see Section 5.4.9). Although use of  $K = N$  zeroes is not recommended for computational reasons, it has one advantage: The  $N$ -lag covariance function obtained from the IFT of the  $2N$ -point PSD is identical to the  $N$ -lag covariance function (Eqn (5.31b)), as noted previously in Section 5.4.3.1. As with the autocorrelation method, improvements in the statistical reliability of the spectral estimates would be attained by “windowing” the time series prior

to spectral estimation or by averaging the raw periodogram estimates over several adjacent frequency bands (see Sections 5.4.6 and 5.4.7).

The two-sided PSD (or autospectral density) for frequency  $f$  in the Nyquist interval  $-1/(2\Delta t) \leq f \leq 1/(2\Delta t)$  (i.e.,  $-f_N \leq f \leq f_N$ ) and a padding of  $K$  zeroes is

$$\begin{aligned} S_{yy}(f) &= \frac{1}{(N+K)\Delta t} \left| \Delta t \sum_{n=0}^{N+K-1} y_n e^{-i2\pi f n \Delta t} \right|^2 \\ &= \frac{1}{(N+K)\Delta t} |Y(f)|^2 \end{aligned} \quad (5.33a)$$

while the one-sided PSD for the positive frequency interval only,  $0 \leq f \leq 1/(2\Delta t)$ , is

$$G_{yy}(f) = 2S_{yy}(f) = \frac{2}{(N+K)\Delta t} |Y(f)|^2 \quad (5.33b)$$

Division by  $\Delta t$  transforms the ESD of Eqn (5.29) into a PSD,  $S_{yy}(f)$ .

Evaluation of Eqn (5.33a) using the FFT defines  $Y(f)$  in terms of the DFT estimates,  $Y(f_k) = Y_k$ , where the  $f_k$  forms a discrete set of  $(N+K)/2$  equally spaced frequencies  $f_k = \pm k / [(N+K)\Delta t]$ ,  $k = 0, 1, \dots, [(N+K)/2] - 1$  in the Nyquist interval,  $-1/2\Delta t \leq f_k \leq 1/2\Delta t$ . The case  $k = 0$  represents the mean component. The two-sided PSD is then

$$\begin{aligned} S_{yy}(0) &= \frac{1}{(N+K)\Delta t} |Y_0|^2, \quad k = 0 \\ S_{yy}(f_k) &= \frac{1}{(N+K)\Delta t} \left[ |Y_k|^2 + |Y_{N+K-k}|^2 \right], \\ k &= 1, \dots, \frac{(N+K)}{2} - 1 \end{aligned} \quad (5.34a)$$

$$\begin{aligned} S_{yy}(f_N) &= S_{yy}\left(f_{(N+K)/2-k}\right) \\ &= \frac{1}{(N+K)\Delta t} |Y_{(N+K)/2}|^2, \quad k = \frac{(N+K)}{2} \end{aligned}$$

and the one-sided PSD is

$$\begin{aligned} G_{yy}(0) &= \frac{1}{(N+K)\Delta t} |Y_0|^2, \quad k = 0 \\ G_{yy}(f_k) &= \frac{2}{(N+K)\Delta t} |Y_k|^2, \\ k &= 1, \dots, \frac{(N+K)}{2} - 1 \end{aligned} \quad (5.34b)$$

$$\begin{aligned} G_{yy}(f_N) &= G_{yy}\left(f_{(N+K)/2-k}\right) \\ &= \frac{1}{(N+K)\Delta t} |Y_{(N+K)/2}|^2, \\ k &= \frac{(N+K)}{2} \end{aligned}$$

Multiplication of  $S_{yy}(f) \equiv S_k$  (or  $G_k$ ) by the bandwidth of the signal  $\Delta f = 1/[(N+K)\Delta t]$  gives the estimated signal variance,  $\sigma_k^2$ , in the  $k$ th frequency band; i.e.,  $\sigma_k^2 = S'_k = S_k \Delta f$ . The summation

$$\sum_{n=0}^{N+K-1} S'_k = \sum_{n=0}^{N+K-1} S_k \Delta f \quad (5.35)$$

gives the variance and total power of the signal. The quantity

$$\begin{aligned} S'_k &= \frac{1}{[(N+K)\Delta t]^2} [ |Y_k|^2 + |Y_{N+K-k}|^2 ] \\ &= \frac{1}{(N+K)^2} \sum_{n=0}^{N+K-1} |y_n e^{-i2\pi f n \Delta t}|^2 \end{aligned} \quad (5.36)$$

is often computed as the periodogram. However, this is not correctly scaled as a PSD but represents the “peak” in the spectral plot rather than the “area” under the plot of  $S_k$  vs  $\Delta f$ . The representation Eqn (5.36) is sometimes useful although most oceanographers are more familiar with the PSD form of the periodogram.

It bears repeating that the use of Fourier transforms assumes a periodic structure to the sampled data when no periodic structure may actually exist in the time series. That is, the FFT of a finite length data record is equivalent to

assuming that the record is periodic. We again note that autospectral functions are always real so that  $S'_{yy}(f_k) = S'_{yy}(2f_N - f_k)$ , and the one-sided autospectral periodogram estimate becomes

$$G'_{yy}(f_k) = 2S'_k = \frac{2}{[(N+K)\Delta t]^2} |Y(f_k)|^2 \quad (5.37)$$

Until the 1960s, the direct transform method first used by Schuster (1898) to study “hidden periodicities” in measured sunspot numbers was seldom used due to difficulties with statistical reliability and extensive computational time. The introduction of the first practical FFT algorithms for spectral analysis (Cooley and Tukey, 1965) greatly reduced the computational time by taking advantage of patterns in DFT functions (see Section 5.8). Problems with the statistical reliability of the spectral estimates are resolved through appropriate windowing and averaging techniques, which we discuss in Sections 5.4.6 and 5.4.7. Figure 5.10 compares the unsmoothed periodogram spectral estimate for the monthly mean SST data at Amphitrite Point (Table 5.3) with the corresponding spectrum obtained from the Blackman–Tukey method. As mentioned earlier, the FFT requires data lengths equal to powers of two so that we have shortened the series to  $2^5 = 32$  months. As we found with the Blackman–Tukey autocorrelation method, the FFT spectrum of coastal temperatures has a strong peak near the annual period, albeit with a slightly different spectral amplitude.

#### 5.4.3.3 The PSD for Periodic Data

For a strictly periodic digital time series  $y(t)$  having an exact integer number of oscillations over the interval  $[0, T]$ , we can use the Fourier series expansion Eqn (5.256) and write

$$\begin{aligned} y(t) &= \frac{1}{2} A_0 + \sum_{n=1}^N [A_n \cos(\omega_n t) + B_n \sin(\omega_n t)] \\ &= \frac{1}{2} C_0 + \sum_{n=1}^N [C_n \cos(\omega_n t + \phi_n)] \end{aligned} \quad (5.38)$$

in which the constants  $A_n$ ,  $B_n$  are given by Eqn (5.258) and where

$$\begin{aligned} C_n &= (A_n^2 + B_n^2)^{1/2} \\ \phi_n &= \tan^{-1}(B_n/A_n) \end{aligned} \quad (5.39)$$

are the amplitude and phase of the complex Fourier coefficient for the  $n$ th frequency component,  $\omega_n = 2\pi f_n$ . Since the data record contains periodic components only, a plot of  $2|C_n|^2$  against  $n$  ( $n = 0, \dots, N - 1$ ) yields a series of distinct "spikes" or line spectra,  $S_n$ , with the variance divided equally between negative and positive frequencies

$$\begin{aligned} S_n &= \frac{(\Delta t)^2}{T} [ |C_n|^2 + |C_{N-n}|^2 ] \\ &= \frac{2\Delta t}{N} |C_n|^2 \end{aligned} \quad (5.40)$$

where the record mean value  $C_0$  has been subtracted from the record  $y(t)$ . Here we have assumed that  $y(t)$  is a real function. The squared Fourier components  $|C_n|^2$  give the contribution of the  $n$ th frequency component to the total variance and the various frequency components contribute additively to the total power of the time series. The contribution from each component is assumed to be independent of that from all other components.

#### 5.4.3.4 Variance-Preserving Spectra

Because the PSD,  $S_{yy}(f)$ , and frequency,  $f$ , of a time series often range over several orders of magnitude, spectral distributions are usually plotted as the logarithm of  $S_{yy}(f)$  vs the logarithm of frequency; i.e.,  $\log[S_{yy}(f)]$  vs  $\log(f)$ . This format allows the user to provide a compact presentation of the spectral distribution. The latter is also useful where a spectrum has a power law dependence of the form  $S_{yy}(f) \sim f^{-p}$ . In this case, the slope of the spectrum is given as  $p = -\log[S_{yy}(f)]/\log(f)$ . An example of a more narrowly focused format of  $\log[S_{yy}(f)]$  vs  $f$  (a log-linear plot) is presented in Figure 5.11(a) where we have used time series data generated

by the relation  $y(k) = A\cos(2\pi nk/N) + \varepsilon(k)$  from Section 5.4.1 (Figure 5.7). Spectral density has units of energy/frequency for the same units used for  $f$ . For example, the PSD of a current velocity record are typically in units of  $(\text{cm/s})^2/\text{cph}$  or  $(\text{cm/s})^2/\text{cpd}$  plotted against  $\log(\text{frequency})$  or frequency in cph (cycles per hour) or cpd (cycles per day), respectively. (Sometimes, m/s are used in place of cm/s, and vice versa.)

In the log-linear format, the integration proceeds over frequency bands of width  $\Delta f$  centered at frequency  $f_c$  (where the "c", in this case, stands for center of the frequency band), so that the area under each small rectangular segment of the spectral curve is equal to a pseudo-variance

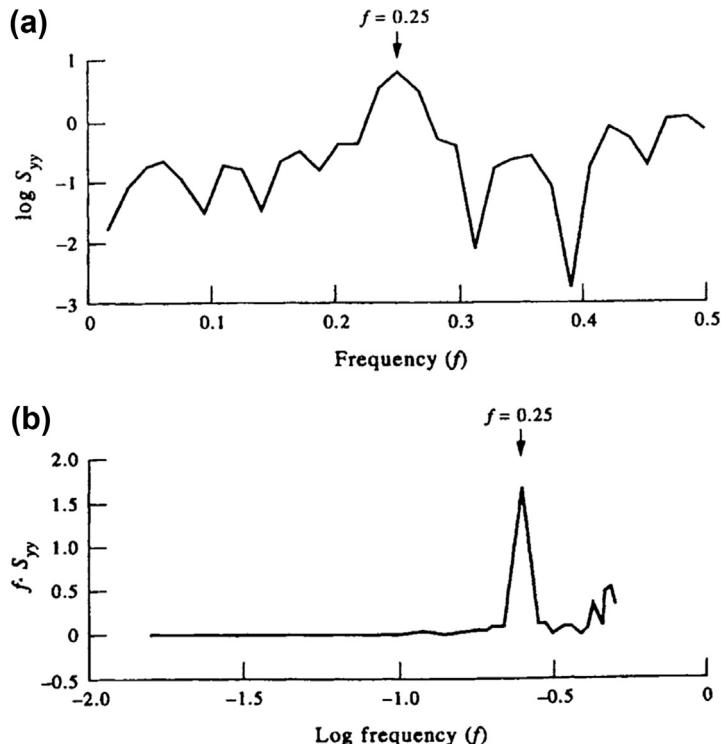
$$\sigma_*^2(f_c) = \int_{f_c-\Delta f/2}^{f_c+\Delta f/2} \log [S_{yy}(f)] df \quad (5.41)$$

Although log spectra plots have an appealing shape, the integral Eqn (5.41) is certainly not variance preserving. To preserve the signal variance,  $\sigma^2(f_c)$ , under the spectral curve, we need to plot  $fS_{yy}(f)$  vs  $\log(f)$ , as in Figure 5.11(b). Replacing  $df$  in Eqn (5.41) with  $d[\log(f)]$ , the true *variance-preserving* form of the spectrum becomes

$$\begin{aligned} \sigma^2(f_c) &= \int_{f_c-\Delta f/2}^{f_c+\Delta f/2} fS_{yy}(f) d[\log(f)] \\ &= \int_{f_c-\Delta f/2}^{f_c+\Delta f/2} S_{yy}(f) df \end{aligned} \quad (5.42)$$

where we have used the fact that  $d[\log(f)] = df/f$ . Equations (5.42) gives the true signal variance within the band  $\Delta f$ . In particular, if  $S_{yy}(f) \approx S_c$  is nearly constant over the frequency increment  $\Delta f$ , then  $\sigma^2(f_c) \approx S_c \Delta f$  is the signal variance in band  $\Delta f$  centered at frequency  $f_c$ . In this format,

**FIGURE 5.11** Two common types of spectral plot derived for the time series  $y(k) = A \cos(2\pi nk/N) + \epsilon(k)$  (see Figure 5.7). (a) A plot of log power spectral density,  $\log [S_{yy}(f)]$ , vs frequency,  $f$ ; (b) A variance-preserving plot in which  $f[S_{yy}(f)]$  is plotted against  $\log(f)$ .



there is a clear spectral peak at  $f = 0.25$  cycles per unit time that is associated with the term  $\cos(2\pi nk/N)$  in the original analytical expression.

#### 5.4.3.5 The Chi-Squared Property of Spectral Estimators

Throughout this chapter, we have claimed that each spectral estimate for maximum frequency resolution,  $1/T$ , obtained from Fourier transforms of stochastic time series has two DoF. We now present a more formal justification for that claim for discrete spectral estimators by showing that each estimate is a stochastic chi-square (pronounced “ki-square”) variable with two DoF (i.e., there are two independent squares entering the expression for the chi-square variable). Consider any stochastic white-noise process  $\eta(t)$ , for which  $E[\eta(t)] = 0$ . The Fourier components are

$$A(f) = \sum_{n=-N}^{N-1} \eta(n\Delta t) \cos(2\pi fn\Delta t) \quad (5.43)$$

$$B(f) = \sum_{n=-N}^{N-1} \eta(n\Delta t) \sin(2\pi fn\Delta t)$$

where as usual,  $-1/(2\Delta t) \leq f \leq 1/(2\Delta t)$  is the Nyquist interval, and it follows that  $E[A(f)] = 0 = E[B(f)]$ . Thus, at the harmonic frequencies  $f_k = k/N\Delta t$ , the variance is

$$\begin{aligned} V[A(f_k)] &= E[A^2(f_k)] = \sigma_\eta^2 \sum_{n=-N}^{N-1} \cos^2(2\pi f_k n \Delta t) \\ &= \frac{1}{2} N \sigma_\eta^2, \quad k = \pm 1, \pm 2, \dots, \pm(N-1) \\ &= N \sigma_\eta^2, \quad k = 0, -N \end{aligned} \quad (5.44a)$$

Similarly

$$\begin{aligned} V[B(f_k)] &= \frac{1}{2}N\sigma_\eta^2, \quad k = \pm 1, \pm 2, \dots, \pm(N-1) \\ &= 0, \quad k = 0, -N \end{aligned} \quad (5.44b)$$

When  $k \neq j$ , the covariance is

$$\begin{aligned} C[A(f_k), A(f_j)] &= \sigma_\eta^2 \sum_{n=-N}^{N-1} \cos(2\pi f_k n \Delta t) \\ &\quad \times \cos(2\pi f_j n \Delta t) = 0 \end{aligned} \quad (5.45a)$$

and

$$C[A(f_k), B(f_j)] = 0 \quad (\text{orthogonality condition}) \quad (5.45b)$$

Because  $A(f_k)$  and  $B(f_k)$  are linear functions of normal random variables,  $A(f_k)$  and  $B(f_k)$  are also distributed normally. Hence, the random variables

$$\begin{aligned} \frac{A^2(f_k)}{V[A(f_k)]} &= \frac{2A^2(f_k)}{N\sigma_\eta^2} \\ \frac{B^2(f_k)}{V[B(f_k)]} &= \frac{2B^2(f_k)}{N\sigma_\eta^2} \end{aligned} \quad (5.46)$$

are each distributed as  $\chi_1^2$ , which is a chi-square variable with one DoF.

Since the normal distributions  $A(f_k)$  and  $B(f_k)$  are independent random variables, the sum of their squares

$$\frac{2}{\sigma_\eta^2} [A^2(f_k) + B^2(f_k)] = \frac{2}{\Delta t \sigma_\eta^2} S_{yy}(f_k) \quad (5.47)$$

is distributed as  $\chi_2^2$ , which is chi-square variable with two DoF. Here,  $S_{yy}(f_k)$  is the sample spectrum. Thus

$$\frac{E[2S_{yy}(f_k)]}{\Delta t \sigma_\eta^2} = 2 \quad (5.48)$$

and

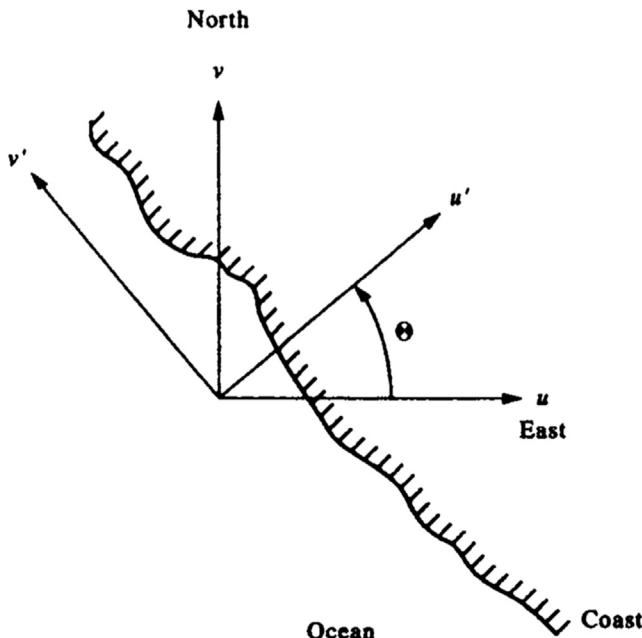
$$E[S_{yy}(f_k)] = \sigma_\eta^2 \Delta t \quad (5.49)$$

which is the spectrum. At the harmonic frequencies (set by the record length), the sample spectrum is an unbiased estimator of the white-noise spectrum of  $\eta(t)$ . Also, at these frequencies, the variance of the estimate is constant and independent of sample size. This explains the failure of the sample estimates of the variance to decrease with increasing sample size. We remark further that, even if  $\eta(t)$  is not normally distributed, the random variables  $A(f_k)$  and  $B(f_k)$  are very nearly normally distributed by the central limit theorem. Hence, the distribution of the  $S_{yy}(f)$  will be very nearly distributed as  $\chi_2^2$  regardless of the PDF of the  $\eta(t)$  process.

#### 5.4.4 Spectra of Vector Series

To calculate the spectra of vector time series such as current and wind, we first need to resolve the data into orthogonal components. Spectral analysis is then applied to the combined series of components and the results stored as a complex quantity in the computer. Raw data are recorded as speed and direction by rototype meters and as orthogonal components by acoustic and electromagnetic meters. The usual procedure is to convert recorded time series to an earth-referenced Cartesian coordinate system consisting of two orthogonal horizontal components and a vertical component (cf. Section 4.3.5). In the open ocean, horizontal velocities typically are resolved into components of eastward (zonal;  $u$ ) and northward (meridional;  $v$ ) time series, whereas in the coastal ocean it is preferable to resolve the vector components into cross-shore ( $u'$ ) and longshore ( $v'$ ) components through the rotation

$$\begin{pmatrix} u' \\ v' \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} \quad (5.50a)$$



**FIGURE 5.12** Cross-shore ( $u'$ ) and longshore ( $v'$ ) velocity components in a Cartesian coordinate system rotated through a positive (counterclockwise) angle from the eastward ( $u$ ) and northward ( $v$ ) directions.

$$\begin{aligned} u' &= u \cos \theta + v \sin \theta \\ v' &= -u \sin \theta + v \cos \theta \end{aligned} \quad (5.50b)$$

where the angle  $\theta$  is the orientation of the coastline (or the local bottom contours) measured counterclockwise from the eastward direction (Figure 5.12). Thus, in the case where the coastline is rotated counterclockwise to lie along a parallel of latitude (i.e.,  $\theta = \pi/2$ ), we find  $u' = v$  and  $v' = -u$ . Alternatively, one can let the current velocity observations define  $\theta$  as the direction of the major axis obtained from principal component analysis; that is, the axis which maximizes the variance in a scatter plot of  $u$  vs  $v$  (see Figure 4.14).

In coastal regions, the principal axis is usually closely parallel to the coastline. For studies of highly circularly polarized motions, such as inertial waves and tidal currents, resolution into clockwise and counterclockwise rotary components is

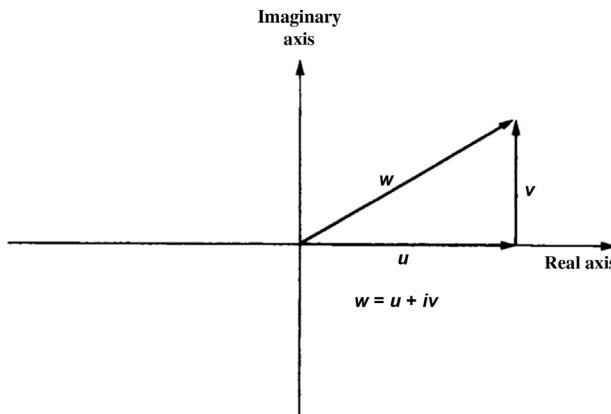
often more useful. The choice of representation depends on the preference of the investigator and the type of process being investigated. More is said on this subject in Section 5.4.4.2.

#### 5.4.4.1 **Cartesian Component Rotary Spectra**

The horizontal velocity vector can be represented in Cartesian coordinates as a complex function  $w(t)$  whose real part,  $u(t)$ , is the projection of the vector on the zonal (or cross-shelf) axis and whose imaginary part,  $v(t)$ , is the projection of the vector on the meridional (or longshelf) axis (Figure 5.13).

$$w(t) = u(t) + iv(t) \quad (5.51)$$

(The use of vector  $w(t)$  follows the convention of Gonella (1972), Mooers (1973) and others in their discussion of rotary spectral analysis and is not to be confused with the weights  $w(t)$ , generally



**FIGURE 5.13** Horizontal velocity represented as a complex vector,  $w = u + iv$ , with components ( $u, v$ ) along the real and imaginary axes, respectively.

written as  $w(t_n)$ , used in the sections on data windowing, or the vertical component of velocity,  $w$ . Gonella (1972) used  $u_1$  and  $u_2$  for the two horizontal velocity components.) A complete description of the time variability of a three-dimensional vector at a single point consists of six functions of frequency: three autospectra for the three velocity components,  $(u, v, w)$  and three cross-spectra. For the two-dimensional vectors considered in this section, there are two auto-spectra and one cross-spectrum. The DFT,  $W(f_k) = U(f_k) + iV(f_k)$ , ( $f_k = k/N\Delta t$ ,  $k = 1, \dots, N$ ;  $k = 0$  is the mean flow) is

$$\begin{aligned} W(f_k) &= \Delta t \sum_{n=0}^{N-1} w(t) e^{-i2\pi kn/N} \\ &= \Delta t \sum_{n=0}^{N-1} [u(t) + iv(t)] e^{-i2\pi kn/N} \end{aligned} \quad (5.52)$$

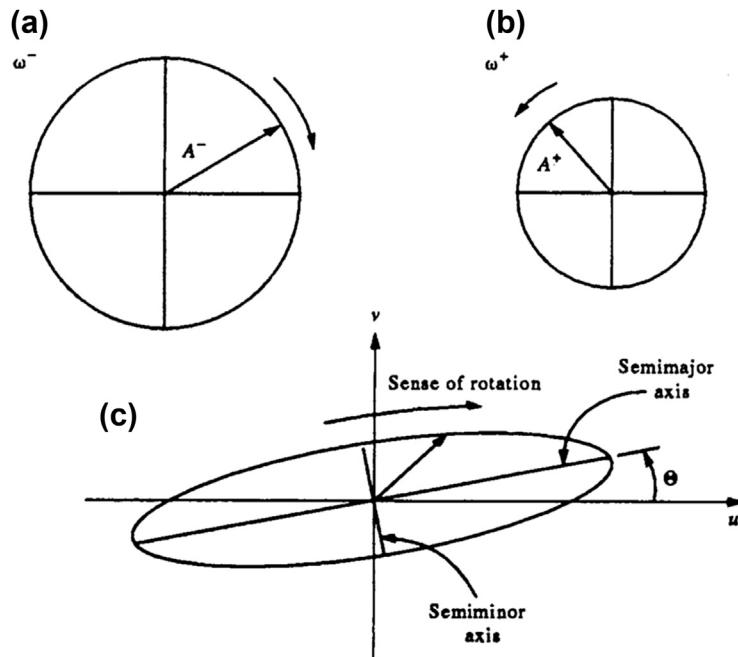
where  $U(f_k)$  and  $V(f_k)$  are the Fourier transforms of  $u(t)$  and  $v(t)$ , respectively. If the original record is separated into  $M$  blocks of length  $N'$ , where  $N = MN'$  is the total record length if no overlapping of segments is used, the spectral density function is given in terms of the number of segments used to form the block-averaged, one-sided autospectrum ( $0 \leq f'_k < \infty$ )

$$\begin{aligned} G_{ww}(f'_k) &= \frac{2}{N\Delta t} \sum_{m=1}^M |W_m(f'_k)|^2 \\ &= \frac{2}{N\Delta t} \sum_{m=1}^M \left\{ [W_{Rm}(f'_k)]^2 + [W_{Im}(f'_k)]^2 \right\} \\ &= \frac{2}{N\Delta t} \sum_{m=1}^M \left\{ [U_{Rm}(f'_k) - V_{Im}(f'_k)]^2 \right. \\ &\quad \left. + [U_{Im}(f'_k) + V_{Rm}(f'_k)]^2 \right\} \end{aligned} \quad (5.53)$$

where  $f'_k = k/N'\Delta t$ ,  $k = 0, 1, \dots, N'/2$  ( $k = 0$  is the mean flow) and for FFT analysis,  $N' = 2p$  (positive integer  $p$ ), and where the subscripts  $R$  and  $I$  stand for the real and imaginary parts of the given Fourier components.

#### 5.4.4.2 Rotary Component Spectra

Rotary analysis of currents involves the separation of the velocity vector for a specified frequency,  $\omega$ , into clockwise and counterclockwise rotating circular components with amplitudes  $A^-$ ,  $A^+$ , and relative phases  $\theta^-$ ,  $\theta^+$ , respectively. Thus, instead of dealing with two Cartesian components ( $u, v$ ) we deal with two circular components ( $A^-, \theta^-$ ;  $A^+, \theta^+$ ). Several reasons



**FIGURE 5.14** Current ellipses formed by the vector addition of two oppositely rotating vectors. (a) Clockwise component ( $\omega^-$ ) and (b) counterclockwise component ( $\omega^+$ ) with amplitudes,  $A^-$  and  $A^+$ , respectively. (c) General case of elliptical motion with major axis tilted at an angle  $\theta$  counterclockwise from east.  $\epsilon^-$  and  $\epsilon^+$  (not shown) are the angles of the two circular components at time  $t = 0$ .

can be given for using this approach: (1) The separation of a velocity vector into oppositely rotating components can reveal important aspects of the wave field at the specified frequencies. The method has proven especially useful for investigating currents over abrupt topography, wind-generated inertial motions, diurnal frequency continental shelf waves, and other forms of narrow-band oscillatory flow; (2) in many cases, one of the rotary components (typically, the clockwise component in the northern hemisphere and counterclockwise component in the southern hemisphere) dominates the currents so that we need to deal with one scalar quantity rather than two. Inertial motions, for example, are almost entirely clockwise (counterclockwise) rotary in the northern (southern) hemisphere so that the counterclockwise

(clockwise) component can be ignored for most applications; and (3) many of the rotary properties, such as spectral energy  $S^-(\omega)$  and  $S^+(\omega)$  and rotary coefficient,  $r(\omega)$ , are invariant under coordinate rotation so that local steering of the currents by bottom topography or the coastline are not factors in the analysis.

The vector addition of the two oppositely rotating circular vectors (Figure 5.14(a) and (b)) causes the tip of the combined vector (Figure 5.14(c)) to trace out an ellipse over one complete cycle. The eccentricity,  $e$ , of the ellipse is determined by the relative amplitudes of the two rotary components. Motions at frequency  $\omega$  are circularly polarized if one of the two components is zero; motions are rectilinear (back-and-forth along the same line) if both circularly polarized components have the same magnitude. In rotary

spectral format, the current vector  $w(t)$  can be written as the Fourier series

$$\begin{aligned} w(t) &= \overline{u(t)} + \sum_{k=1}^N U_k \cos(\omega_k t - \phi_k) \\ &\quad + i \left[ \overline{v(t)} + \sum_{k=1}^N V_k \cos(\omega_k t - \theta_k) \right] \\ &= \left[ \overline{u(t)} + i\overline{v(t)} \right] + \sum_{k=1}^N [U_k \cos(\omega_k t - \phi_k) \\ &\quad + iV_k \cos(\omega_k t - \theta_k)] \end{aligned} \quad (5.54)$$

in which  $\overline{u(t)} + i\overline{v(t)}$  is the mean velocity,  $\omega_k = 2\pi f_k = 2\pi k/N\Delta t$  is the angular frequency,  $t (=n\Delta t)$  is the time, and  $(U_k, V_k)$  and  $(\phi_k, \theta_k)$  are the amplitudes and phases, respectively, of the Fourier constituents for each frequency for the real and imaginary components. Subtracting the mean velocity and expanding the trigonometric functions, we find

$$\begin{aligned} w'(t) &= w(t) - [\overline{u(t)} + i\overline{v(t)}] \\ &= \sum_{k=1}^N \{U_{1k} \cos(\omega_k t) + U_{2k} \sin(\omega_k t) \\ &\quad + i[V_{1k} \cos(\omega_k t) + V_{2k} \sin(\omega_k t)]\} \end{aligned} \quad (5.55)$$

in which we have defined the even  $(U_{1k}, V_{1k})$  and odd  $(U_{2k}, V_{2k})$  functions as

$$U_{1k} = U_k \cos \phi_k, \quad U_{2k} = U_k \sin \phi_k \quad (5.56a)$$

$$V_{1k} = V_k \cos \theta_k, \quad V_{2k} = V_k \sin \theta_k \quad (5.56b)$$

Dropping the prime notation for  $w'(t)$  and following some reorganization, we can write the  $k$ th frequency component of the series as the sum of counterclockwise (+) and clockwise (-) components

$$\begin{aligned} w_k(t) &= w_k^+(t) + w_k^-(t) \\ &= A_k^+ \exp(i\epsilon_k^+) \exp(i\omega_k t) + A_k^- \exp(i\epsilon_k^-) \exp(-i\omega_k t) \\ &= \exp\left[\frac{i(\epsilon_k^+ + \epsilon_k^-)}{2}\right] \left\{ [A_k^+ + A_k^-] \cos\left[\frac{\epsilon_k^+ - \epsilon_k^-}{2} + \omega_k t\right] \right. \\ &\quad \left. + i[A_k^+ - A_k^-] \sin\left[\frac{\epsilon_k^+ - \epsilon_k^-}{2} + \omega_k t\right]\right\} \end{aligned} \quad (5.57)$$

where the counterclockwise and clockwise rotary component amplitudes are given by

$$A_k^+ = \frac{1}{2} \left\{ [(U_{1k} + V_{2k})]^2 + [(U_{2k} - V_{1k})]^2 \right\}^{1/2} \quad (5.58a)$$

$$A_k^- = \frac{1}{2} \left\{ [(U_{1k} - V_{2k})]^2 + [(U_{2k} + V_{1k})]^2 \right\}^{1/2} \quad (5.58b)$$

and the corresponding phase angles for time  $t = 0$ , by

$$\epsilon_k^+ = \tan^{-1}[(V_{1k} - U_{2k})/(U_{1k} + V_{2k})] \quad (5.59a)$$

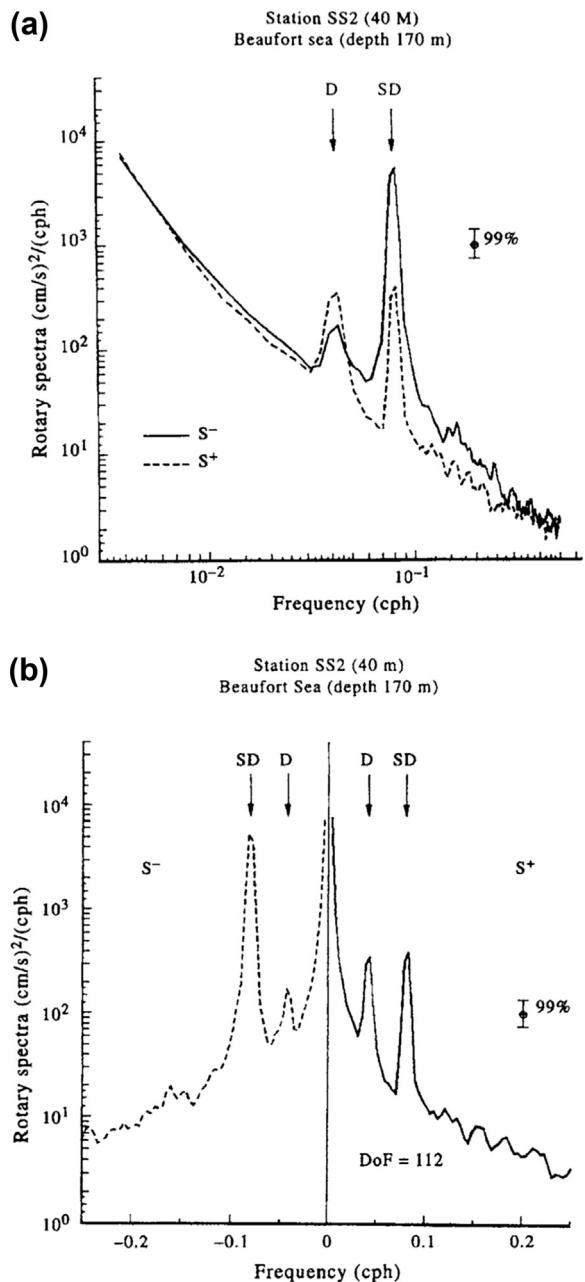
$$\epsilon_k^- = \tan^{-1}[(U_{2k} + V_{1k})/(U_{1k} - V_{2k})] \quad (5.59b)$$

Each of the constituents contributing to Eqn (5.55) has the form of an ellipse with major semiaxis of length  $L_M = (A_k^+ + A_k^-)$  and minor semiaxis of length  $L_m = |A_k^+ - A_k^-|$  (Figure 5.14(c)). The ellipse is tilted at an angle of  $\theta = \frac{1}{2}(\epsilon_k^+ + \epsilon_k^-)$  from the  $u$ -axis and the vector is along the major axis of the ellipse at time  $t = (\epsilon_k^+ - \epsilon_k^-)/(4\pi f_k)$ . The one-sided spectra  $(G_k^+, G_k^-) = (S_k^+, S_k^-)$  for the two oppositely rotating components for frequencies  $f_k = \omega_k/2\pi$  are

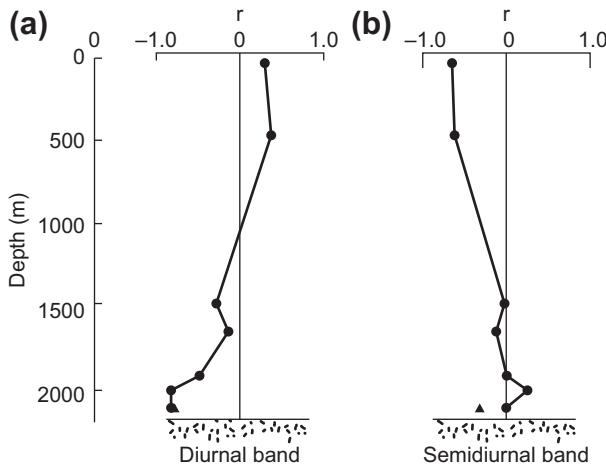
$$S(f_k^+) = S_k^+ = \frac{(A_k^+)^2}{N\Delta t}, \quad f_k = 0, \dots, 1/(2\Delta t) \quad (5.60a)$$

$$S(f_k^-) = S_k^- = \frac{(A_k^-)^2}{N\Delta t}, \quad f_k = -1/(2\Delta t), \dots, 0 \quad (5.60b)$$

Plots of rotary spectra are generally presented in two ways. In Figure 5.15(a), both  $S^-$  and  $S^+$  are plotted as functions of frequency magnitude,  $|f| \geq 0$ , with solid and dashed lines used for the clockwise and counterclockwise spectra, respectively. In Figure 5.15(b), we use the fact that clockwise spectra are defined for negative frequencies and counterclockwise spectra for positive frequencies. The spectra  $S(f_k^+)$  and  $S(f_k^-)$



**FIGURE 5.15** Rotary current spectra for hourly currents measured at 40-m depth in the Beaufort Sea, Arctic Ocean (water depth = 170 m). Peaks are at the diurnal (D) and semidiurnal (SD) tidal frequencies. Frequency resolution is 0.0005 cph and there are 112 degrees of freedom (DoF) per spectral band. Vertical bar gives the 99% level of confidence, (a) One-sided rotary spectra,  $S^-(f)$  and  $S^+(f)$ , vs  $f$  for positive frequency,  $f$ ; (b) Two-sided rotary spectra,  $S(f_k^+) = S^+$  and  $S(f_k^-) = S^-$  vs  $\log f$  for positive and negative frequencies,  $f_{\pm k}$ . (Courtesy E. Carmack, A. Rabinovich, and E. Kulikov.)



**FIGURE 5.16** Rotary coefficient,  $r(\omega)$ , as a function of depth for current oscillations in (a) the diurnal frequency band ( $\omega/2\pi \approx 0.04$  cph) and (b) the semidiurnal band ( $\omega/2\pi \approx 0.08$  cph). (From Allen and Thomson (1993).)

used in Figure 5.15(a) are then plotted on opposite sides of zero frequency. In these spectra, peak energy occurs at the diurnal and semidiurnal periods. The predominantly clockwise rotary motions at semidiurnal periods suggest a combination of tidal and near-inertial motions (at this latitude the inertial period is close to the semidiurnal tidal period).

Another useful property is the rotary coefficient

$$r(\omega) = \frac{S_k^+ - S_k^-}{S_k^+ + S_k^-} \quad (5.61)$$

which ranges from  $r = -1$  for clockwise motion, to  $r = 0$  for unidirectional flow, to  $r = +1$  for counterclockwise motion. The rotary nature of the flow can change considerably with position, depth, and time. As indicated by Figure 5.16, the observed diurnal tidal currents over Endeavour Ridge in the northeast Pacific change from moderately positive to strongly negative rotation with depth. In contrast, the semidiurnal currents change from strongly negative near the surface to strongly rectilinear at depth. (Data, in this case, are from a string

of current meters moored for a period of 9 months.) We remark that the definition Eqn (5.61) differs in sign from that of Gonella (1972), who used  $S_k^- - S_k^+$  rather than  $S_k^+ - S_k^-$  in the numerator. Because many types of oceanic flow are predominantly clockwise rotary in the northern hemisphere, Gonella's definition has the advantage that clockwise rotating currents have positive rotary coefficients. However, we find Gonella's definition a bit awkward since clockwise motions, which are linked to *negative* frequencies, then have *positive* rotary coefficients.

#### 5.4.4.3 Rotary Spectra (via Cartesian Components)

Gonella (1972) and Mooers (1973) present the rotary spectra in terms of their Cartesian counterparts and provide a number of rotational invariants for analyzing current and wind vectors at specified frequencies. Specifically, the one-side autospectra for the counterclockwise (CCW) and clockwise (CW) rotary components of the vector  $w(t) = u(t) + iv(t)$  are, in terms of their Cartesian components

$$G(f_k^+) = \frac{1}{2} [G_{uu}(f_k) + G_{vv}(f_k) + Q_{uv}(f_k)], \\ f_k \geq 0 \text{ (CCW component)} \quad (5.62a)$$

$$G(f_k^-) = \frac{1}{2} [G_{uu}(f_k) + G_{vv}(f_k) - Q_{uv}(f_k)], \\ f_k \leq 0 \text{ (CW component)} \quad (5.62b)$$

where  $G_{uu}(f_k)$  and  $G_{vv}(f_k)$  are the one-sided auto-spectra of the  $u$  and  $v$  Cartesian components of velocity and  $Q_{uv}(f_k)$  is the quadrature spectrum between the two components, where

$$Q_{uv}(f_k) = -Q_{uv}(-f_k) = (U_{1k}V_{2k} - V_{1k}U_{2k}) \quad (5.63)$$

As defined in [Section 5.6](#), the spectrum can be written in terms of cospectrum (real part) and quadrature spectrum (imaginary part)

$$G_{uv}(f_k) = C_{uv}(f_k) - iQ_{uv}(f_k) \quad (5.64)$$

### 5.4.5 Effect of Sampling on Spectral Estimates

Spectral estimates derived by conventional techniques are limited by two fundamental problems: (1) the finite length,  $T$ , of the time series; and (2) the discretization associated with the sampling interval,  $\Delta t$ . The first problem is inherent to all real data sets while the second is associated with finite instrument response times and/or the need to digitize the time series for the purposes of analysis.

Irrespective of the method used to calculate the power spectrum of a waveform, the record duration,  $T = N\Delta t$ , and sampling increment,  $\Delta t$ , impose severe limitations on the information that can be extracted. Ideally, we would like to have sensors that can sample rapidly enough (small  $\Delta t$ ) that no significant frequency component goes unresolved. This also eliminates aliasing problems in which unresolved spectral energy at frequencies higher than the Nyquist frequency is folded back into lower frequencies.

At the same time, we wish to record for a sufficiently long period (large  $N$ ) that we capture many cycles of the lowest frequency of interest. Long-term sampling also enables us to better resolve frequencies that are close together and to improve the statistics (confidence intervals) for spectral estimates. In reality, most data series are a compromise based on the frequencies of interest, the response limitations of the sensor, and cost. The choices of the sampling rate and the record duration are tailored to best meet the task at hand.

#### 5.4.5.1 Effect of Finite Record Length

As noted earlier, we can think of a data sample  $\{y(t)\}$  of duration  $T = N\Delta t$  as the output from an infinite physical process  $\{y'(t)\}$  viewed through a finite length window ([Figure 5.3](#)). The window has the shape of a “box-car” function,  $w(t_n) = w_n = w(n\Delta t)$ , which has unit amplitude and zero phase lag over the duration of the data sequence but is zero elsewhere. That is  $y(t_n) = w(t_n) \cdot y'(t_n)$  where

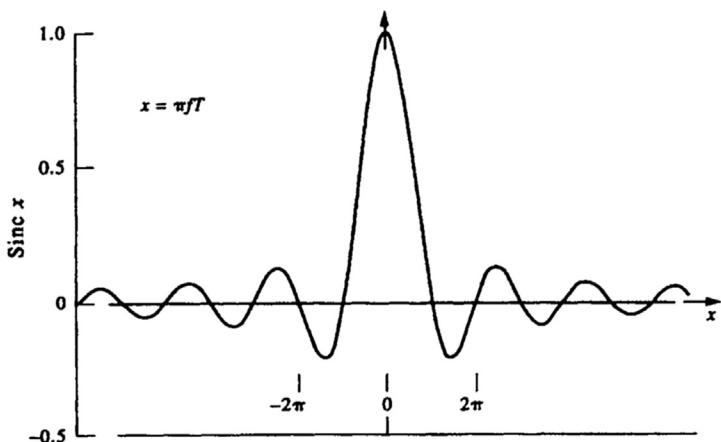
$$w_n = 1, \quad n = 0, \dots, N-1 \\ w_n = 0, \quad \text{for } n \geq N, \quad n < 0 \quad (5.65)$$

Since it is truncated, the data set has endpoint discontinuities, which lead to Gibbs’ phenomena (“ringing”) and ripple effects in the frequency domain. The DFT  $Y(f)$  of the truncated series  $y_n = y(n\Delta t)$  is

$$Y(f) = \sum_{n=-\infty}^{\infty} w_n y'_n e^{-i2\pi f n \Delta t} \quad (5.66)$$

In frequency space,  $Y(f)$  is the convolution (written as  $*$ ) of the Fourier transform of the infinite data set,  $Y'(f)$ , with the Fourier transform  $W(f)$  of the function  $w(t)$ . That is

$$Y(f) = \int_{-\infty}^{\infty} Y'(f') W'(f - f') df' \\ \equiv Y'(f) * W(f) \quad (5.67)$$



**FIGURE 5.17** The function  $\text{sinc}(x) = \sin(x)/x$  showing the large side-lobes, which are responsible for leakage of spectral power from a given frequency to adjacent frequencies.

where for a box-car function

$$\begin{aligned} W(f) &= T \exp(i\pi fT) \frac{\sin(\pi fN\Delta t)}{(\pi fN\Delta t)} \\ &\equiv T \exp(i\pi fT) \text{sinc}(\pi fN\Delta t) \end{aligned} \quad (5.68)$$

and  $\text{sinc}(x) \equiv \sin(x)/x$ . It is the large side-lobes or ripples of the sinc function (Figure 5.17), which are responsible for the leakage of spectral power from the main frequency components into neighboring frequency bands (and vice versa). In particular,  $Y(f)$  for a specific frequency  $f = f_0$  is spread to other frequencies according to the phase and amplitude weighting of the window function. Leakage has the effect of both reducing the spectral power in the central frequency component and contaminating it with spectral energy from adjacent frequency bands. Those familiar with the various mathematical forms for the Dirac delta function,  $\delta(f)$ , will recognize the formulation

$$\delta(f) = \lim_{f \rightarrow 0} \left[ \frac{\sin(\pi f\Delta t)}{\pi f\Delta t} \right] = \lim_{f \rightarrow 0} [\text{sinc}(\pi f\Delta t)]$$

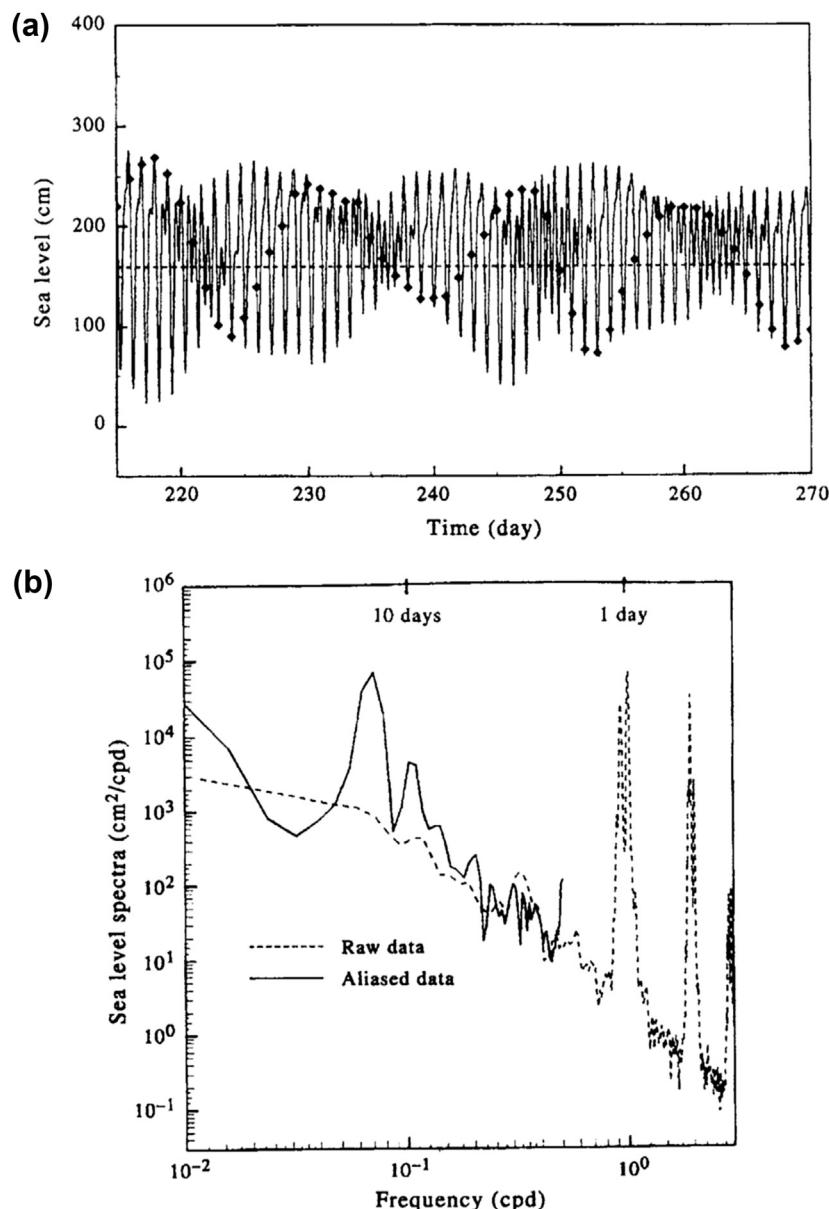
Thus, as the frequency resolution increases (i.e.,  $f \rightarrow 0$ ),  $Y(f) \rightarrow Y'(f)$ .

In addition to distorting the spectrum, the box-car window limits the frequency resolution

of the periodogram, independently of the data. The convolution  $Y'(f)^*W(f)$  means that the narrowest spectral response of the resultant transform is confined to the main-lobe width of the window transform. For a given window, the main-lobe width (the width between the  $-3 \text{ dB} = 10 \log(1/2)$  levels of the main lobe) determines the frequency resolution,  $\Delta f$ , of a particular window. For most windows, including the box-car window, this resolution is roughly the inverse of the observation time;  $\Delta f \approx 1/T = 1/(N\Delta t)$ .

#### 5.4.5.2 Aliasing

Poor discretization of time series data due to limitations in the response time of the sensor, limitations in the recording and data storage rates, or through postprocessing methods may cause *aliasing* of certain frequency components in the original waveform (Figure 5.18(a)). An aliased frequency is one that masquerades as another frequency. In Figures 5.18(b), for example, the considerable tidal energy at diurnal and semidiurnal periods (1 and 2 cpd) that is well resolved by the hourly sampled record is folded back to lower frequencies of roughly 0.065, 0.10, and 0.15 cpd (periods of 14.8, 9.6, and 7.4 days, respectively) when the original sea-level record is subsampled at daily



**FIGURE 5.18** The origin of aliasing. (a) The solid line is the tide height recorded at Victoria, British Columbia over a 60-day period from July 29 to September 27, 1975 (time in Julian days). The diamonds are the sea-level values one would obtain by only sampling once per day. (b) The power spectrum obtained from the two data series in (a). In this case, the high frequency energy (dashed curve) gets folded back into the spectrum at lower (aliased) frequencies (solid curve).

( $\Delta t = 24$  h) intervals. The aliased signals are nowhere near the original higher frequency tidal signals. If we knew nothing about the true spectrum, and were presented only with the aliased spectrum in Figure 5.18(b), we would be hard pressed to provide a physical explanation for the strong fortnightly and weather-band cycles in the sea-level time series.

As illustrated by Figure 5.18, it becomes impossible, for a specific sampling interval, to tell with certainty which frequency out of a large number of possible aliases is actually contributing to the signal variability. This leads to differences in the spectra between the continuous and discrete time series. Since we use the spectra of the discrete series to estimate the spectrum of the continuous series, the sampling interval must be properly selected to minimize the effect of the aliasing. If we know from previous analysis that there is little likelihood of significant energy at the disguised frequencies, then aliasing is not a problem. Otherwise, a degree of smoothing may be required to ensure that higher frequencies do not contaminate the lower frequencies. This smoothing must be performed prior to sampling or digitizing since aliased contributions cannot be recognized once they are present in the discrete data series.

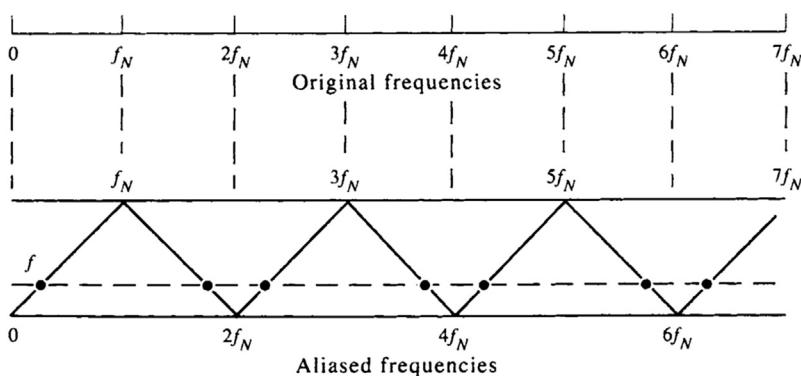
The aliasing problem can be illustrated in a number of ways. To begin with, we note that for discrete data at equally spaced intervals,  $\Delta t$ ,

we can measure only those frequency components lying within the principal frequency range,

$$-\omega_N \leq \omega \leq -\omega_0, \quad \omega_0 \leq \omega \leq \omega_N, \quad \omega_N \geq 0 \quad (5.69a)$$

$$-f_N \leq f \leq -f_0, \quad f_0 \leq f \leq f_N, \quad f_N \geq 0 \quad (5.69b)$$

in which  $\omega_N = \pi/\Delta t$  and  $f_N = 1/(2\Delta t)$  are the usual Nyquist frequencies in radians and cycles per unit time, respectively, and  $\omega_0 = 2\pi/T$  and  $f_0 = 1/T$  are corresponding fundamental frequencies for a time series of duration  $T$ . The Nyquist frequency is the highest frequency that can be extracted from a time series having a sampling rate of  $1/\Delta t$ . Clearly, if the original time series has spectral power at frequencies for which  $|f| \geq f_N$ , these spectral contributions are unresolved and will contaminate power associated with frequencies within the principal range (Figure 5.19). The unresolved variance becomes lumped together with other frequency components. Familiar examples of aliasing are the slow reverse rotation of stage-coach wheels in classic western movies due to the undersampling by the frame rate of the movie camera. Even in modern TV commercials or movies, distinguishable features on moving automobile tires or wheel frames often can be seen to rotate rapidly backwards, slow to a stop, then turn forward at the correct rotation speed as the vehicle gradually comes to a stop. Automobile commercials can



**FIGURE 5.19** The spectral energies of all frequencies,  $f = \omega/2\pi$ , at the nodes (•) located along the dotted line are folded back to the left, accordion style, into the spectral estimate for the spectrum,  $S(f)$  for the principal range,  $0 \leq f \leq f_N$  ( $0 \leq \omega \leq \omega_N$ ). (Adapted from Bendat and Piersol (1986).)

avoid this problem by equipping the wheels with featureless hubcaps, spokes, and tires— or by digitizing to a higher resolution.

If  $\omega, f \geq 0$  are frequencies inside the principal intervals (Eqn (5.69)); the frequencies outside the interval, which form aliases with these frequencies are (in sequence)

$$2\omega_N \pm \omega, 4\omega_N \pm \omega, \dots, 2p\omega_N \pm \omega \quad (5.70a)$$

$$2f_N \pm f, 4f_N \pm f, \dots, 2pf_N \pm f \quad (5.70b)$$

where  $p$  is a positive integer. These results lead to the alternate term *folding* frequency for the Nyquist frequency since spectral power outside the principal range is folded back, accordion style, into the principal interval. As illustrated by Figure 5.19, folding the power spectrum about  $f_N$  produces aliasing of frequencies  $2f_N - f$  with frequencies  $f$ ; folding the spectrum at  $2f_N$  produces aliasing of frequencies  $2f_N + f$  with frequencies  $2f_N - f$ , which are then folded back about  $f_N$  into frequency  $f$ , and so forth. For example, if  $f_N = 5$  rad/h, the observations at 2 rad/h are aliased with spectral contributions having frequencies of 8 and 12 rad/h, 18 and 22 rad/h, and so on.

We can verify that oscillations of frequency  $2p\omega_N \pm \omega$  (or  $2pf_N \pm f$ ) are indistinguishable from frequency  $\omega$  (or  $f$ ) by considering the data series  $x_\omega(t)$  created by the single frequency component  $x_\omega(t) = \cos(\omega t)$ . Using the transformation  $\omega \rightarrow (2p\omega_N \pm \omega)$ , together with  $t_n = n\Delta t$  and  $\omega_N = \pi/\Delta t$ , yields

$$\begin{aligned} x_\omega(t_n) &= \cos [(2p\omega_N \pm \omega)t_n] \\ &= \operatorname{Re} \{\exp[i(2p\omega_N \pm \omega)t_n]\} \\ &= \operatorname{Re} \{\exp[i2p\omega_N t_n] \exp[\pm i\omega t_n]\} \quad (5.71) \\ &= (+1)^{pn} \operatorname{Re} [\exp(\pm i\omega t_n)] \\ &= \cos(\omega t_n) = x_\omega(t_n) \end{aligned}$$

In other words, the spectrum of  $x(t)$  at frequency  $\omega$  will be a superposition of spectral contributions from frequencies  $\omega, 2p\omega_N \pm \omega, 4p\omega_N \pm \omega$ , and so forth. More specifically, it can

be shown that the aliased spectrum  $S_a(\omega)$  for discrete data is given by

$$S_a(\omega) = \sum_{n=-\infty}^{\infty} S(\omega + 2n\omega_N) \quad (5.72a)$$

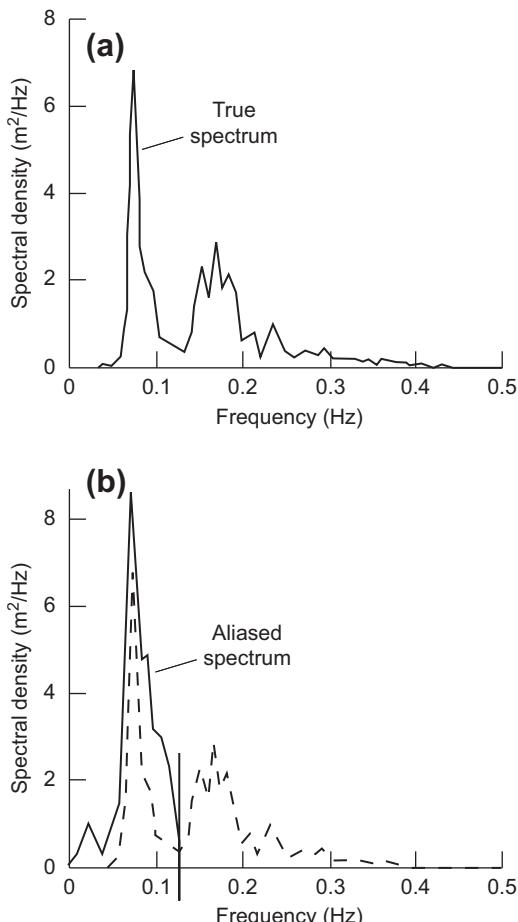
$$= S(\omega) + \sum_{n=1}^{\infty} [S(2n\omega_N - \omega) + S(2n\omega_N + \omega)] \quad (5.72b)$$

The true spectrum,  $S$ , gives the distorted spectrum,  $S_a$ , caused by the summation of overlapping copies of measured spectra in the principal interval. Only if the original record is devoid of spectral power at frequencies outside the principal frequency range will the spectrum of the observed record equal that of the actual oceanic variability. To avoid aliasing problems, one has no choice but to sample the data as frequently as justifiably possible (i.e., up to frequencies beyond which energy levels become small) or to filter the sampled data before they are recorded (as in the case of a stilling well used to eliminate gravity waves from a tidal record). A further example of spectral contamination by aliased frequencies is illustrated in Figure 5.20(a) and (b). In Figure 5.20(b), we have assumed that the wave recorder was inadvertently programmed to record at 0.13 Hz, corresponding to a limiting wave period of 7.69 s. The energy from the shorter-period waves was not measured but contaminate the energy of the longer-period waves when folded back about the Nyquist frequency.

#### 5.4.5.3 Nyquist Frequency Sampling

Sampling time series that has significant variability at the Nyquist frequency affords its own set of problems. Suppose we wish to represent  $y(t)$  through the usual Fourier relation

$$y(t) = \int_{-\omega_N}^{\omega_N} Y(\omega) e^{i\omega t} d\omega \quad (5.73)$$



**FIGURE 5.20** An aliased autospectrum. (a) The true spectrum,  $S(f)$  ( $\text{m}^2/\text{cps}$ ), of wind-generated waves as a function of frequency (Hz = cycles per second); (b) Aliased spectrum,  $S_a(f)$ , that would arise from folding about a hypothetical Nyquist frequency,  $f_N = 0.13 \text{ Hz}$ .

where we have assumed that  $Y(\omega) = 0$  for  $|\omega| > \omega_N$ . In this case, there is no aliasing problem since there is no power at frequencies greater than  $\omega_N$ . The function  $y(t)$  can be constructed from frequency components strictly in the interval  $(-\omega_N, \omega_N)$ . In discrete form for infinite length data

$$y(t) = \frac{1}{2\omega_N} \sum_{n=-\infty}^{\infty} \left[ y_n \int_{-\omega_N}^{\omega_N} e^{i\omega(t-n\Delta t)} d\omega \right] \quad (5.74a)$$

where the integral has the form of a sinc function such that

$$y(t) = \sum_{n=-\infty}^{\infty} y_n \frac{\sin [\omega_N(t - n\Delta t)]}{\omega_N(t - n\Delta t)} \quad (5.74b)$$

Given the data  $\{y_n\}$ , we can construct  $y(t)$ . However, suppose that  $y(t)$  fluctuates with the Nyquist frequency  $\omega_N$  such that

$$y(t) = y_0 \cos (\omega_N t + \theta) \quad (5.75)$$

where, for the sake of generality, the phase angle is arbitrary,  $0 \leq \theta \leq 2\pi$ . Then, using  $\sin(n\pi) = 0$  for all  $n$  (an integer)

$$\begin{aligned} y_n &= y(n\Delta t) = y_0 \cos (n\pi + \theta) \\ &= y_0[\cos (n\pi) \cos \theta] \\ &= y_0(-1)^n \cos \theta \end{aligned} \quad (5.76)$$

This leads to a component with amplitude  $y_n = y_0(-1)^n \cos \theta$ , which fluctuates in sign because of the term  $(-1)^n$ ,  $-\infty \leq n \leq \infty$ . If  $\theta$  is unknown, the function  $y(t)$  cannot be constructed. If  $\theta = k\pi/2$ , so that  $\cos(\omega_N t + \theta) = \sin(\omega_N t)$ , the observer will find no signal at all. In general,  $0 \leq |\cos \theta| \leq 1$  and the magnitude will always be less than  $y_0$ , resulting in biased data.

According to the above analysis, we should sample slightly more frequently than  $\Delta t$  if we are to fully resolve oscillations at the maximum frequency of interest (assumed to be the Nyquist frequency). A sampling rate of 2.5 samples per cycle of the frequency of interest appears to be acceptable whereby  $\Delta t = 1/(2.5f_N) = (2/5)(1/f_N) = (4/5)\pi/\omega_N$ .

#### 5.4.5.4 Frequency Resolution

The need to resolve spectral estimates in neighboring frequency bands is an important requirement of time series analysis. Without

sufficient resolution, it is not possible to determine whether a given spectral peak is associated with a single frequency, or is a smeared response containing a number of separate spectral peaks. A good example of this for tides is presented by Munk and Cartwright (1966), who show that for long records, the main constituents in the diurnal and semidiurnal frequency bands can be resolved into a multitude of other tidal frequencies. How well the peaks can be resolved depends on the frequency differences,  $\Delta f$ , between the peaks and the length,  $T$ , of the data set used in the analysis. For an unsmoothed periodogram, the frequency resolution in hertz is roughly the reciprocal of the time duration in seconds of the data.

The distinction between well-resolved and poorly resolved spectral estimates is somewhat subjective and depends on how we wish to define "resolution." As illustrated by diffraction patterns in classical optics, we can follow the "Rayleigh criterion" for the separation of spectral peaks (Jenkins and White, 1957). Recall that the diffraction pattern for a given frequency,  $f$ , of light varies as  $\text{sinc}(\phi) = \sin[(\phi - \phi_f)]/(\phi - \phi_f)$ , where  $\phi$  is the angle of the incident light beam to the grating. This also is the functional form for the spectral peak of a truncated time series (see *windowing* in the next section). Two spectral lines are said to be "well resolved" if the separation between peaks exceeds the difference in frequency between the center frequency to the maximum at the first side-lobe and "just resolved" if the spectral peak of one pattern coincides with the first zero of the second pattern (Figure 5.21(a)–(c)). Here, the separation in frequency is equal to the difference in frequency between the peak of one spectrum and the first zero of the function  $\sin(\phi)/\phi$  of the second (where  $\phi = \omega T/2$ ). The spectral peaks are "not resolved" if this separation is less than that between the center frequency and the first zero of the  $\sin(\phi)/\phi$  functions (Figure 5.21(d)).

Consider an oceanic record consisting of two sinusoidal components, both having amplitude  $y_0$  and constant phase lags such that

$$y(t) = y_0[\cos(\omega_1 t + \theta_1) + \cos(\omega_2 t + \theta_2)], \\ -T/2 \leq t \leq T/2 \quad (5.77)$$

where as usual  $\omega = 2\pi f$ . The one-sided, unsmoothed PSD,  $S(\omega)$ , for these data is then found from the Fourier transform

$$S(\omega) = \frac{1}{2} Ty_0^2 \left\{ \frac{\sin\left[\frac{1}{2}T(\omega - \omega_1)\right]}{\left[\frac{1}{2}T(\omega - \omega_1)\right]} + \frac{\sin\left[\frac{1}{2}T(\omega - \omega_2)\right]}{\left[\frac{1}{2}T(\omega - \omega_2)\right]} \right\}$$

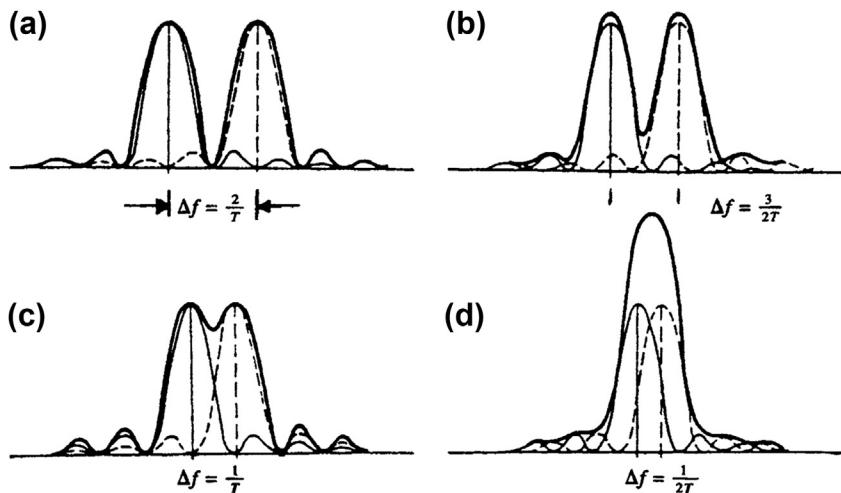
The power spectrum consists of two terms of the form  $\sin(\phi)/\phi$  centered at frequencies  $\omega_1$  and  $\omega_2$ . Using the Rayleigh criterion, we can just resolve the two peaks (i.e., determine if there is one or two sinusoids contributing to the spectrum) provided that the frequency separation  $\Delta\omega = |\omega_1 - \omega_2|$  ( $\Delta f = |f_1 - f_2|$ ) is equal to the frequency difference for the peak of one frequency and the first zero of  $\sin(\phi)/\phi$  for the other frequency. Since zeroes of  $\sin(\phi)/\phi$  occur at frequencies  $f$  equal to  $\pm 1/T, \pm 2/T, \dots, \pm p/T$ , the frequencies are just resolved when

$$\Delta\omega = \frac{2\pi}{T}; \quad \Delta f = \frac{1}{T} \quad (5.78a)$$

and well resolved for

$$\Delta\omega > \frac{3\pi}{T}; \quad \Delta f > \frac{3}{2T} \quad (5.78b)$$

In summary, resolution of two frequencies  $f_k$  and  $f_{k+1}(=f_k \pm \Delta f)$  using an unsmoothed periodogram or equivalently a rectangular window, requires a record of length  $T$ , where  $\Delta f = 1/T$  frequency units. Note also that  $1/T$  is equal to the fundamental frequency,  $f_1$ , which is the lowest frequency that we can calculate for the record. For some nonrectangular windows, the length of the data set must be increased to about  $2T = 2/\Delta f$  to achieve the same frequency separation.



**FIGURE 5.21** Resolution of spectral lines. (a, b) Well resolved; (c) just resolved; and (d) not resolved. (From Jenkins and White (1957).)

In a related study, Munk and Hasselman (1964) discuss the “super-resolution” of tidal frequency variability. The fact that time series of tidal heights vary at precise frequencies and have relatively large SNRs suggests that the traditional requirement (that a minimum record length  $T$  is required to separate tidal constituents separated by frequency difference  $\Delta f = 1/T$ ) is “grossly incomplete.” The modified resolvable frequency difference is

$$\Delta f = \frac{1}{rT}; \quad \Delta\omega = \frac{2\pi}{rT} \quad (5.79)$$

in which  $r \equiv (\text{signal level/noise level})^{1/2}$ . On this basis, the Rayleigh criterion must be considered a conservative measure of the resolution requirement for deterministic processes.

#### 5.4.6 Smoothing Spectral Estimates (Windowing)

The need for statistical reliability of spectral estimates brings us to the topic of spectral averaging or smoothing. As we have seen, DFTs (Discrete Fourier Transforms) provide an elegant method for decomposing a data sequence into a set of

discrete spectral estimates. For a data sequence of  $N$  values, the periodogram estimate of the spectrum can have a maximum of  $N/2$  Fourier components. If we use all  $N/2$  components to generate the periodogram, there are only two DoF per spectral estimate, corresponding to the coefficients  $A_n$ ,  $B_n$  of the sine and cosine functions for each Fourier component (see Sections 5.4.3.1 and 5.4.3.5) or, alternatively, to the magnitude and phase of each Fourier component (see Section 5.4.3.3). Based on the assumption that data are drawn from a normally distributed random sample, we can define the confidence limits for the spectrum in terms of a chi-squared distribution,  $\chi_n^2$ , where for  $n$  DoF

$$E[\chi_n^2] = \mu^2 = n, \quad E[(\chi_n^2 - \mu^2)] = \sigma^2 = 2n \quad (5.80)$$

Substituting  $n = 2$  into these expressions, we find that the standard deviation,  $\sigma$ , is equal to the mean,  $\mu$ , of the estimate, indicating that results based on two DoF are not statistically reliable. It is for this reason that some sort of ensemble averaging or smoothing of spectral estimates is required. The smoothing can be (1) applied directly to the time series through

convolution with a sliding averaging function or by (2) averaging adjacent spectral estimates. A one-shot smoothing applied to the entire data record marginally increases the number of DoF per spectral estimate. In most practical applications, the full time series is broken into a series of short overlapping segments and smoothing is applied to each of the overlapping segments. The analyst then ensemble averages the smoothed spectra from each segment to increase the number of DoF per spectral estimate. The greater the smoothing, the greater the number of DoF per spectral band, the narrower the confidence limits, and the greater the reliability of any observed spectral peaks. The trade-off is a longer processing time and a loss of spectral resolution that can remove smaller peaks that may or may not be indicative of real processes (see Figure 5.22).

A window is a smoothing function applied to finite observations or their Fourier transforms to minimize "leakage" in the spectral domain. Convolution in the time domain and multiplication in the frequency domain are adjoint Fourier functions (see Appendix G regarding convolution). A practical window is one which allows little of the energy in the main spectral lobe to leak into the side-lobes, where it can obscure and distort other spectral estimates that are present. In fact, weak signal spectral responses can be masked by higher side-lobes from stronger spectral responses. Skillful selection of tapered data windows can reduce the side-lobe leakage, although always at the expense of reduced resolution. Thus, we want a window that minimizes the side-lobes and maximizes (concentrates) the energy near the frequency of interest in the main lobe. These two performance limitations are rather troublesome when analyzing short data records. Short data occur in practice because many measured processes are event-like (of short duration) or have slowly time-varying spectra that may be considered constant over only short record segments. The window is applied to data to reduce the order of the discontinuity at the boundary of

the periodic extension since few harmonics will fit exactly into the length of the time series.

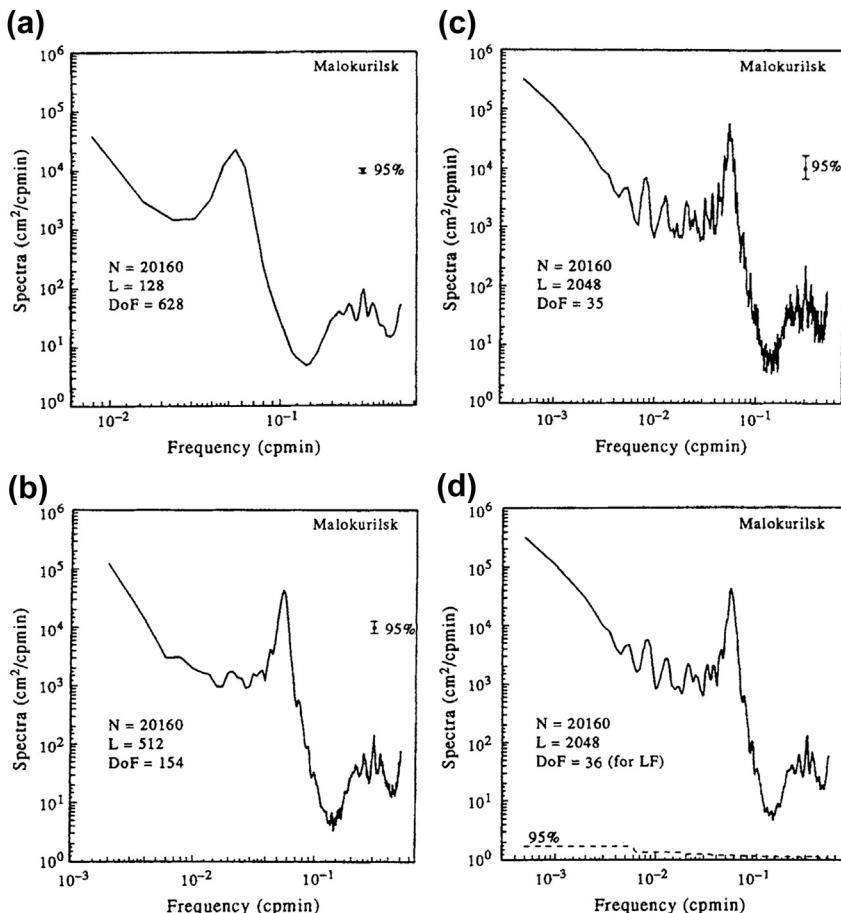
Signals with frequencies other than those of the basis set are not periodic in the observation window. The periodic extension of a signal, not commensurate with its natural period, exhibits discontinuities at the boundaries of the observational period. Such discontinuities are responsible for spectral contributions or leakage over the entire basis set. In the time domain, the windows are applied to the data as a multiplicative weighting (convolution) to reduce the order of the discontinuities at the boundary of the periodic extensions. The windowed data are brought to zero smoothly at the boundaries so that the periodic extensions of the data are continuous in many orders of the derivatives. The value of  $Y(f)$  at a particular frequency  $f$ , say  $f_0$ , is the sum of all the spectral contributions at each  $f$  weighted by the window centered at  $f_0$  and measured at  $f$

$$Y(f) = Y'(f) * W(f) \quad (5.81)$$

There exist a multitude of data windows or tapers with different shapes and characteristics ranging from the rectangular (box-car) window discussed in the previous section, to the classic Hanning and Hamming windows, to more sophisticated windows such as the Dolph–Chebyshev window. The type of window used for a given application depends on the required degree of side-lobe suppression, the allowable widening of the central lobe, and the amount of computing one is willing to endure. We will briefly discuss several of the conventional windows plus the Kaiser–Bessel window recommended by Harris (1978). Additional details on the Kaiser–Bessel window and filter are provided in Section 6.9.

#### 5.4.6.1 Desired Window Qualities

Windows affect the attributes of a given spectral analysis method, including its ability to detect and resolve periodic waveforms, its



**FIGURE 5.22** Spectra of sea-level oscillations recorded by a bottom-pressure gauge in Malokurilsk Bay on the west coast of Shikotan Island, Russia. Time series length,  $T = N\Delta t$ , where  $N = 20,160$  and  $\Delta t = 1$  min. Segment lengths are  $T_s = M\Delta t$ ,  $M \ll N$ . Each time series segment has been smoothed with a Kaiser–Bessel window with 50% overlap between segments. Block averaging has been used to smooth the spectral estimates. (a) Highly smoothed spectrum with  $M = 128$  ( $2^7$ ), degrees of freedom (DoF) = 628; (b) Moderately smoothed spectrum with  $M = 512$  ( $2^9$ ), DoF = 154; (c) Weakly smoothed spectrum with  $M = 2048$  ( $2^{11}$ ), DoF = 36; (d) Same as (c) except that DoF = 36 applies to the lowest frequency range only. For  $f \geq 6 \times 10^{-2}$  cycles/min, the number of spectral estimates averaged together increases as  $3 \times 36$ ,  $5 \times 36$ , and  $7 \times 36$ , for each of the next three frequency ranges. (Courtesy of Alexander Rabinovich.)

dynamic range, confidence intervals, and ease of implementation. Spectral estimates are affected not only by the broadband noise spectrum of the data but also by narrow-band signals that fall within the bandwidth of the window. Leakage of spectral power from a narrow-band spectral component,  $f_o$ , to another

frequency component,  $f_a$ , produces a bias in the amplitude and position of a spectral estimate. This bias is especially disruptive for the detection of weak signals in the presence of nearby strong signals. To reduce the bias, we need a “good” window. Although there are no universal standards for a good window, we would

like it to possess the following characteristics in Fourier transform space:

1. The central main lobe of the window (which is centered on the frequency of interest) should be as narrow as possible to improve the frequency resolution of adjacent spectral peaks in the data set, and the first side-lobes should be greatly attenuated relative to the main lobe to avoid contamination from other frequency components. Here, the narrowness of the central lobe is measured by the positions of the  $-3$  dB (half amplitude points,  $10 \log^{1/2}$ ) on either side of the lobe. Retention of a narrow central lobe, while suppressing the side-lobes, is not as easy as it sounds since suppression of the side-lobes invariably leads to a broadening of the central lobe;
2. The window should suppress the amplitudes of side-lobes at frequencies far removed from the central lobe. That is, the side-lobes should have a rapid asymptotic fall-off rate with frequency so that they leak relatively little energy into the spectral estimate at the central lobe (i.e., into the frequency of interest);
3. The coefficients of the window should be easy to generate for multiplication in the time domain and convolution in the Fourier transform domain.

A good performance indicator (PI) for the time domain window  $w(t)$  can be defined as the difference between the equivalent noise bandwidth (ENBW) and the bandwidth (BW), located between the  $-3$  dB levels of the central lobe (Harris, 1978)

$$\text{PI} = \frac{\text{ENBW} - \text{BW}}{\text{BW}} = \frac{\frac{1}{\text{BW}} \sum_n w^2(n\Delta t)}{\left[ \sum_n w(n\Delta t) \right]^2} - 1 \quad (5.82)$$

where we have normalized by the BW. The lower the value, the better the performance of

the filter; windows that perform well have values for this ratio ( $\times 100\%$ ) of between 4.0 and 5.5%. A summary of the figures of merit for several well-known windows is presented in [Table 5.4](#). PI values are obtained using columns four and five. For example, for the weakly performing box-car (rectangular) window,  $\text{PI} = 0.124$  (12.4%), while for the strongly performing Kaiser window,  $\text{PI} = 0.049$  (4.9%). The choice of window can be daunting; Harris lists more than 44 windows for smoothing spectral estimates.

#### **5.4.6.2 Rectangular (Box-Car) and Triangular Windows**

As discussed at the beginning of [Section 5.4](#), a rectangular window has an amplitude of unity throughout the observation interval of duration  $T = N\Delta t$ , with the weighting given by

$$\begin{aligned} w(n\Delta t) &= 1, \quad n = 0, 1, \dots, \\ &\quad N-1 \quad (\text{or } -N/2 \leq n \leq N/2) \quad (5.83) \\ &= 0, \quad \text{elsewhere} \end{aligned}$$

([Figure 5.23\(a\)](#)). Using the relation  $\omega T = N\theta$ , where  $\theta = \omega\Delta t$  and  $T = N\Delta t$ , the spectral window from the DFT is

$$W(\theta) = Te^{-i(N-1)\theta/2} \frac{\sin(N\theta/2)}{N\theta/2} \quad (5.84a)$$

$$|W(\theta)|^2 = T^2 \left[ \frac{\sin(N\theta/2)}{N\theta/2} \right]^2 \quad (5.84b)$$

([Figure 5.23\(b\)](#)) where the exponential term in [Eqn \(5.84a\)](#) gives the phase shift of the window as a function of the frequency  $\omega = \theta/\Delta t$ . The function  $W$ , the Dirichlet kernel, has strong side-lobes, with the power of the first side-lobe down only  $-13$  dB (factor of 0.22) from the main lobe. The remaining side-lobes fall off weakly at  $-6$  dB per octave, which is the functional rate for a discontinuity (an “octave” corresponds to a factor of 2 change in frequency). Zeros of  $W(\theta)$  occur at integer multiples of the frequency resolution,

**TABLE 5.4** Windows, Figures of Merit and Performance Indicator (PI)

Window	Highest Side-Lobe Level (dB)	Side-Lobe Attenuation (dB/octave)	ENBW (Bins)	3 dB BW (Bins)	PI	Overlap Correlation 75%	Overlap Correlation 50%
Rectangle	-13	-6	1.00	0.89	0.124	0.750	0.500
Triangle	-27	-12	1.33	1.28	0.031	0.719	0.250
Hanning	-32	-18	1.50	1.44	0.042	0.659	0.167
Hamming	-43	-6	1.36	1.30	0.046	0.707	0.235
Parzen	-21	-12	1.20	1.16	0.035	0.765	0.344
Tukey $\alpha = 0.5$	-15	-18	1.22	1.15	0.061	0.727	0.364
Kaiser $\alpha = 2.0$	-46	-6	1.50	1.43	0.049	0.657	0.169
Bessel							
$\alpha = 2.5$	-57	-6	1.65	1.57	0.051	0.595	0.112
$\alpha = 3.0$	-69	-6	1.80	1.71	0.052	0.539	0.074
$\alpha = 3.5$	-82	-6	1.93	1.83	0.054	0.488	0.048

The last column gives the correlation between adjacent data segments for the specified percentage segment overlap. For completeness, we include the Tukey and Parzen Windows. (Adapted from Harris (1978).)

$f_1 = 1/T$ , for which  $N\theta/2 = \omega T/2 = \pm p\pi$ . That is, where  $f = \pm p/T(\pm 1/T, \pm 2/T, \dots)$ .

The triangular (Bartlett) window

$$w(n\Delta t) = \begin{cases} \frac{n}{(N/2)}, & n = 0, 1, \dots, N/2 \\ \frac{N-n}{(N/2)}, & n = N/2, \dots, N-1 \end{cases} \quad (5.85a)$$

$$= \frac{N/2 - |n|}{(N/2)}, \quad 0 \leq |n| \leq N/2 \quad (5.85b)$$

(Figure 5.24(a)) has the DFT

$$W(\theta) = \frac{2T}{N} e^{-i(N-1)\theta/2} \left[ \frac{\sin(N\theta/2)}{N\theta/2} \right]^2 \quad (5.86a)$$

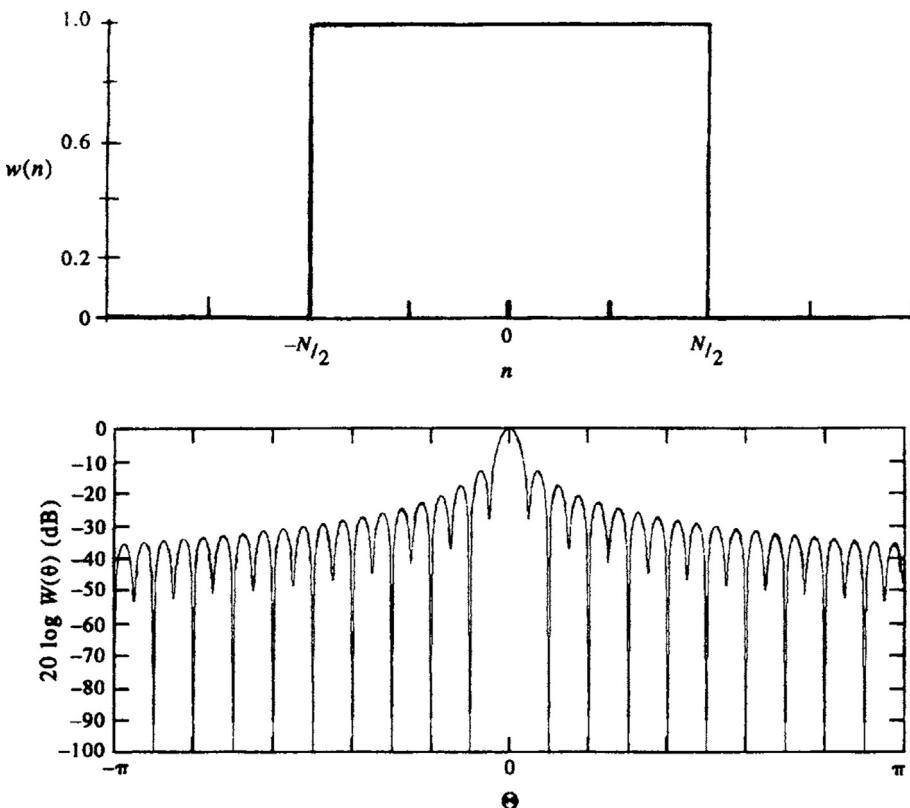
$$|W(\theta)|^2 = \frac{4T^2}{N^2} \left[ \frac{\sin(N\theta/2)}{N\theta/2} \right]^4 \quad (5.86b)$$

(Figure 5.24(b)) which we recognize as the square of the sinc function for the rectangular window. The main lobe between zero crossings has twice the width of the rectangular window but the level of the first side-lobe is down by -26 dB, twice that of the rectangular window. Despite the improvement over the box-car window, the side-lobes of the triangular window are still extensive and use of this window is not recommended if other windows are available.

The Parzen window

$$w(n\Delta t) = 1 - |n/(N/2)|^2, \quad 0 \leq |n| \leq N/2 \quad (5.87)$$

is the squared counterpart to the Bartlett window. This is the simplest of the continuous polynomial windows and has first side-lobes down by -22 dB and falls off with frequency as  $1/\omega^2$ .



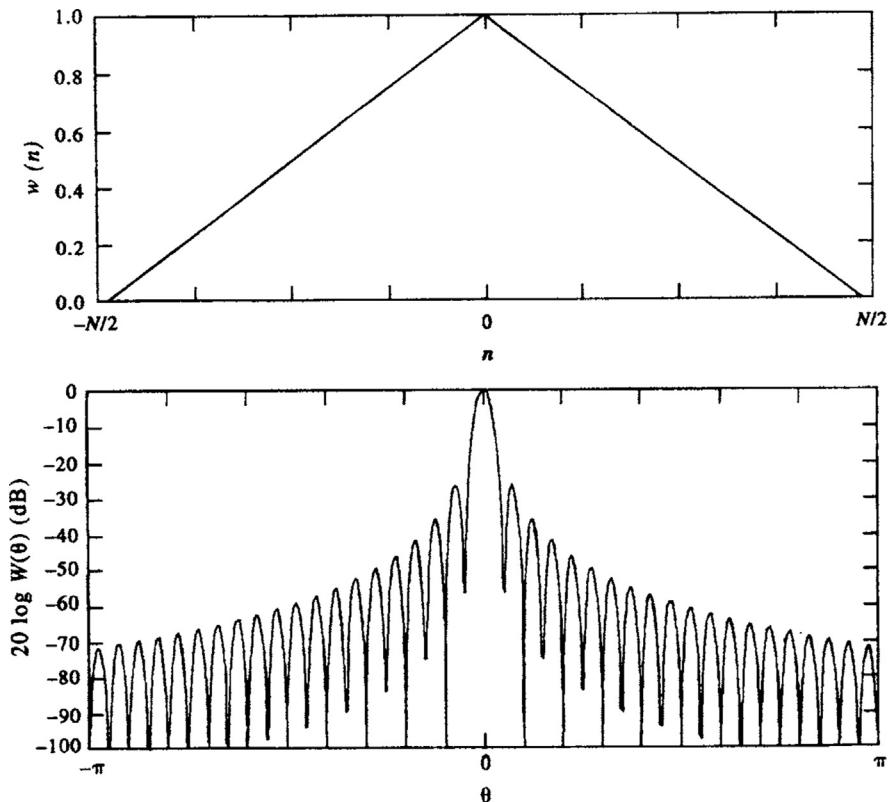
**FIGURE 5.23** A box-car (rectangular) window for  $N=41$  weights. (a) Weights,  $w(n)=1.0$  in the time domain ( $-20 \leq n \leq 20$ ). (b) Fourier transform of the weights,  $W(\theta)$ , plotted as  $20 \log|W(\theta)|$ , where  $\theta = \omega\Delta t/N = 40\pi/N$  is the frequency span of the window.

#### 5.4.6.3 Hanning and Hamming Windows (50% Overlap)

The Hann window, or *Hanning window* as it is most commonly known, is named after the Austrian meteorologist Julius von Hann and is part of a family of trigonometric windows having the generic form  $\cos^\alpha(n)$ , where the exponent,  $\alpha$ , is typically an integer from 1 through 4. The case  $\alpha=1$  leads to the *Tukey* (or *cosine-tapered*) window (Harris, 1978). As  $\alpha$  becomes larger, the window becomes smoother, the side-lobes fall off faster, and the main lobe widens. The Hanning window

( $\alpha=2$ ), also known as the *raised cosine* and *sine-squared* window, is defined in the time domain as

$$w(n\Delta t) = \begin{cases} \sin^2(\pi n/N) = \frac{1}{2}[1 - \cos(2\pi n/N)], & n = 0, 1, \dots, N-1 \\ \sin^2[\pi(n+N/2)/N] = \frac{1}{2}[1 - \cos[2\pi(n+N/2)/N]], & n = -N/2, \dots, N/2 \end{cases} \quad (5.88)$$



**FIGURE 5.24** The triangular (Bartlett) window for  $N = 51$  weights. (a) Weights,  $w(n)$  in the time domain ( $-20 \leq n \leq 20$ ). (b) Fourier transform of the weights,  $W(\theta)$ , plotted as  $20 \log|W(\theta)|$  (cf. Figure 5.22).

(Figure 5.25(a)), which is a continuous function with a continuous first derivative. The DFT of this weighting function is

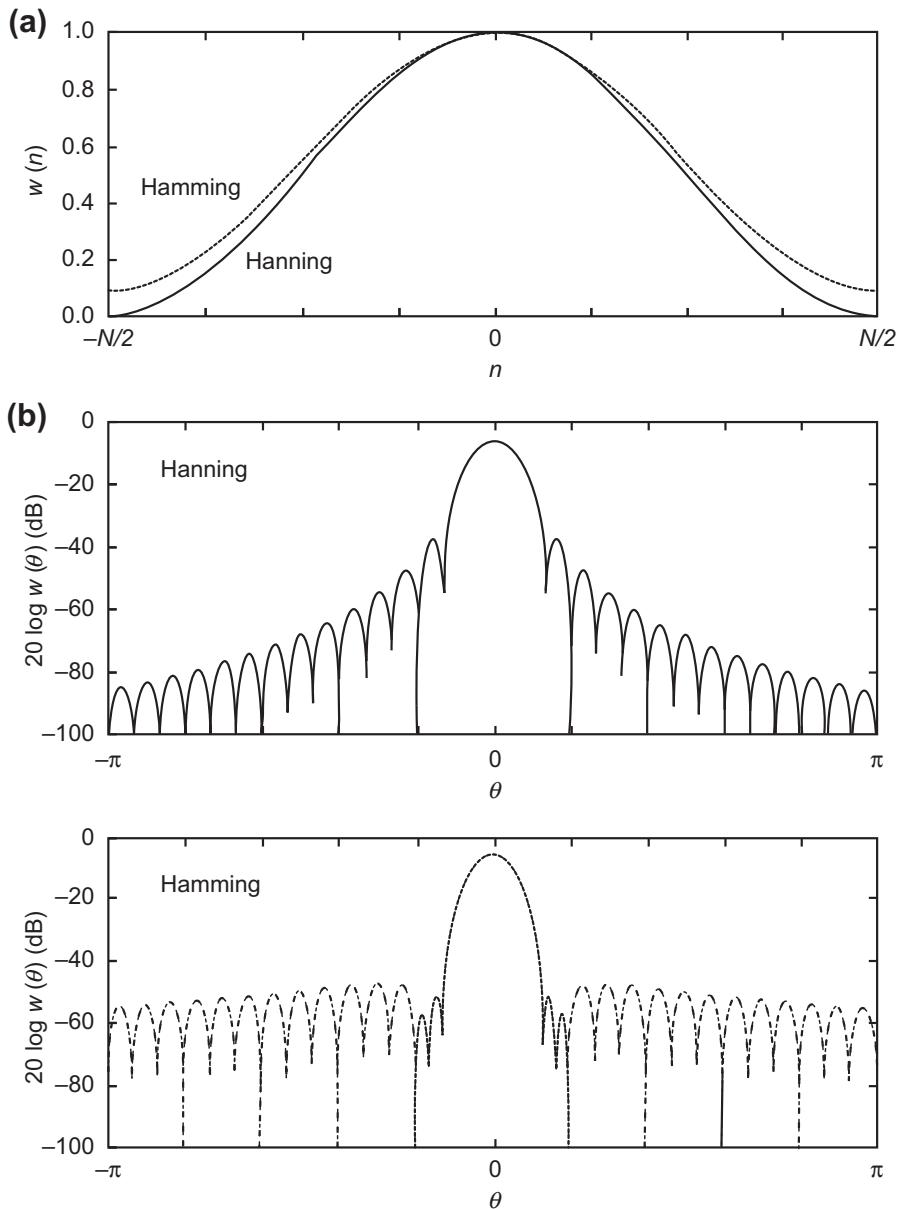
$$W(\theta) = \frac{1}{2}D(\theta) + \frac{1}{4}[D(\theta - \theta_1) + D(\theta + \theta_1)] \quad (5.89)$$

(Figure 5.25(b)), where  $\theta_1 = 2\pi/N$  and

$$D(\theta) = Te^{i\theta/2} \frac{\sin(N\theta/2)}{N\theta/2} \quad (5.90)$$

is the standard function (Dirichlet kernal) obtained for the rectangular and triangular

windows. Thus, the window consists of the summation of three sinc functions (Figure 5.25(c)), one centered at the origin,  $\theta = 0$ , and two other translated Dirichlet kernels having half the amplitude of the main kernal and offset by  $\theta = \pm 2\pi/N$  from the central lobe. There are several important features of the window response  $W(\theta)$ . First of all, the functions  $D$  are discrete and defined only at points that are multiples of  $2\pi/N$ , which also correspond to the zero crossings of the central function,  $D(\theta)$ . Secondly, for all of zero crossings except those at  $\theta_{\pm 1} = \pm 2\pi/N$ , the translated functions also have zero crossings at multiples of  $2\pi/N$ . As a result, only values at  $-2\pi/N$ ,  $0$ , and  $+2\pi/N$



**FIGURE 5.25** The Hanning and Hamming windows for  $N=41$  weights. (a) Weights,  $w(n)$ ,  $(-20 \leq n \leq 20)$ . (b) Fourier transform of the weights,  $W(\theta)$ , plotted as  $20 \log|W(\theta)|$  (cf. Figure 5.22). The response functions have not been re-scaled.

contribute to the window response. It is the widening of the main lobes of the translated functions that causes them to be nonzero at the first zero crossings of the central function. Lastly, because the translated functions are out of phase with the central function, they tend to cancel the side-lobe structure. The first side-lobe is down by  $-32$  dB (factor of  $0.025$ ) from the main lobe. The remaining side-lobes diminish as  $1/\omega^3$  or at about  $-18$  dB per octave.

An attractive aspect of the Hanning window is that smoothing in the frequency domain can be accomplished using only three convolution terms corresponding to  $\theta_0$ ,  $\theta_{\pm 1}$ . The Hanning-windowed Fourier transform,  $Y_H(f_k)$ , representing the spectrum for the frequency,  $f_k$ , is then obtained from the raw spectra  $Y$  for the frequencies,  $f_k$  and the two adjoining frequencies,  $f_{k-1}$  and  $f_{k+1}$ ; that is

$$Y_H(f_k) = \frac{1}{2} \left\{ Y(f_k) - \frac{1}{2} [Y(f_{k-1}) + Y(f_{k+1})] \right\} \quad (5.91)$$

The transform  $Y(f_k)$  has already been rectangular-windowed by the very act of collecting the data but is “raw” in the sense that no additional smoothing has been applied. Other processing advantages of the Hanning window are discussed by Harris (1978). Since the squares of the weighting terms  $(1/2)^2 + (1/4)^2 + (1/4)^2 = 3/8$ , the total energy will be reduced following the application of the Hanning window. To compensate, the amplitudes of the Fourier transforms,  $Y_H(f)$  should be multiplied by  $\sqrt{8/3}$  prior to computation of the spectra. Specifically

$$Y_H(f_k) = \Delta t (8/3)^{1/2} \sum_{n=0}^{N-1} y_n \times [1 - \cos(2\pi n/N)] e^{-i2\pi kn/N} \quad (5.92)$$

where  $f_k = k/(N\Delta t)$ .

The *Hamming window* is a variation on the Hanning window designed to cancel the first side-lobes. To accomplish this, the relative sizes of the three Dirichlet kernels are adjusted through a parameter,  $\gamma$  where

$$\begin{aligned} w(n\Delta t) &= \gamma + (1 - \gamma) \cos(2\pi n/N), \\ n &= -N/2, \dots, N/2 \end{aligned} \quad (5.93a)$$

$$\begin{aligned} W(\theta) &= \gamma D(\theta) + \frac{1}{2}(1 - \gamma)[D(\theta - 2\pi/N) \\ &\quad + D(\theta + 2\pi/N)] \end{aligned} \quad (5.93b)$$

Perfect cancellation of the first side-lobes (located at  $\theta_1 = 2.5\pi/N$ ) occurs when  $\gamma = 25/46 \approx 0.543478$ . Taking  $\gamma = 0.54$  leads to near-perfect cancellation at  $\theta_1 = 2.6\pi/N$  and a marked improvement in side-lobe level. The Hamming window is defined as

$$\begin{aligned} w(n\Delta t) &= 0.54 + 0.46 \cos(2\pi n/N), \\ n &= -N/2, \dots, N/2 \end{aligned} \quad (5.94)$$

and has a spectral distribution similar to that of the Hanning window with more “efficient” side-lobe attenuation. The highest side-lobe levels of the Hanning window occur at the first side-lobes and are down by  $-32$  dB from the main lobe. For the Hamming window, the first side-lobe is highly attenuated and the highest side-lobe level (the third side-lobe) is down by  $-43$  dB. To compensate for the filter, the amplitudes of the Fourier transforms  $Y_{\text{Ham}}(f)$  should be multiplied by  $\sqrt{5/2}$  prior to computation of the spectra. On a similar note, anyone using any of the windows in this section to calculate running mean time series should make sure each estimated value is divided by the sum of the weights used,  $\sum_N w_n$ .

#### 5.4.6.4 Kaiser–Bessel Window

Harris (1978) identifies the Kaiser–Bessel window as the “top performer” among the

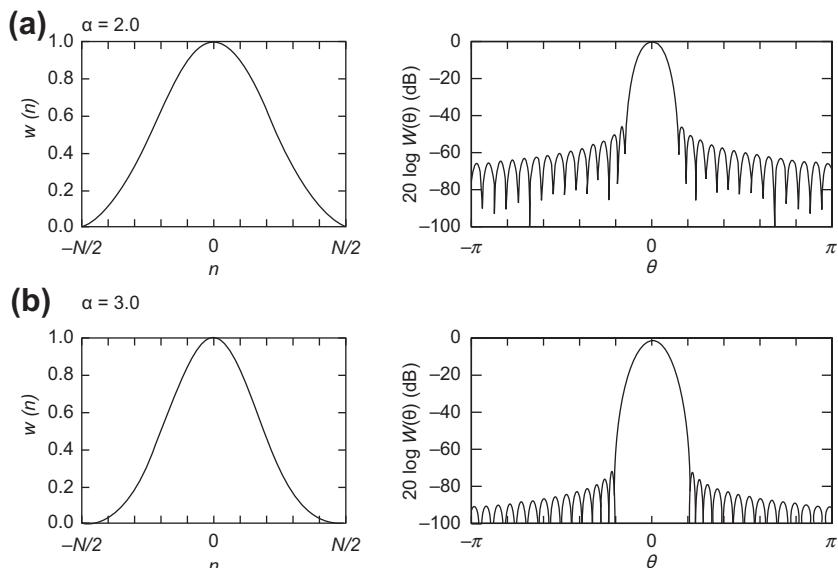
many different types of windows he considered. Among other factors, the coefficients of the window are easy to generate and it has a high ENBW, one of the criteria used to separate good and bad windows. The trade-off is an increased main-lobe width for reduced side-lobe levels. In the time domain the filter is defined in terms of the zeroth-order modified Bessel functions of the first kind.

$$w(n\Delta t) = \frac{I_0(\pi\alpha\Omega)}{I_0(\pi\alpha)}, \quad 0 \leq |n| \leq N/2 \quad (5.95)$$

where the argument  $\Omega = [1 - (2n/N)^2]^{1/2}$  and

$$I_0(x) = \sum_{k=0}^{\infty} \left[ \frac{(-1)^k (x/2)^k}{\Gamma(k+1)k!} \right]^2 \quad (5.96)$$

The parameter  $\pi\alpha$  is half of the time-bandwidth product, with  $\alpha$  typically having values 2.0, 2.5, 3.0, and 3.5. The transform is approximated by



**FIGURE 5.26** The Kaiser–Bessel window for  $N=51$  weights and  $\alpha=2.0$  and 3.0. (a) Weights,  $w(n)$ ,  $(-20 \leq n \leq 20)$ . (b) Fourier transform of the weights,  $W(\theta)$ , plotted as  $20 \log|W(\theta)|$  (cf. Figure 5.22). (From Harris (1978).)

$$W(\theta) \approx [N/I_0(\pi\alpha)] \frac{\sinh \left\{ \left[ \pi^2 \alpha^2 - (N\theta/2)^2 \right]^{1/2} \right\}}{\left\{ \left[ \pi^2 \alpha^2 - (N\theta/2)^2 \right]^{1/2} \right\}} \quad (5.97)$$

Plots of the weighting function  $w$  and the DFT for  $W$  are presented in Figure 5.26 for two values of the parameter  $\alpha$  (=2.0, 3.0). The modified Bessel function  $I_0$  is defined as follows:

For  $|x| \leq 3.75$

$$\begin{aligned} I_0(x) = & \{ \{ [(4.5813 \times 10^{-3}Z + 3.60768 \times 10^{-2}) \\ & + 2.659732 \times 10^{-1}]Z + 1.2067492 \}Z \\ & + 3.0899424 \}Z + 3.5156229 \}Z + 1.0 \end{aligned} \quad (5.98a)$$

where for real  $x$

$$Z = (x/3.75)^2 \quad (5.98b)$$

For  $|x| > 3.75$

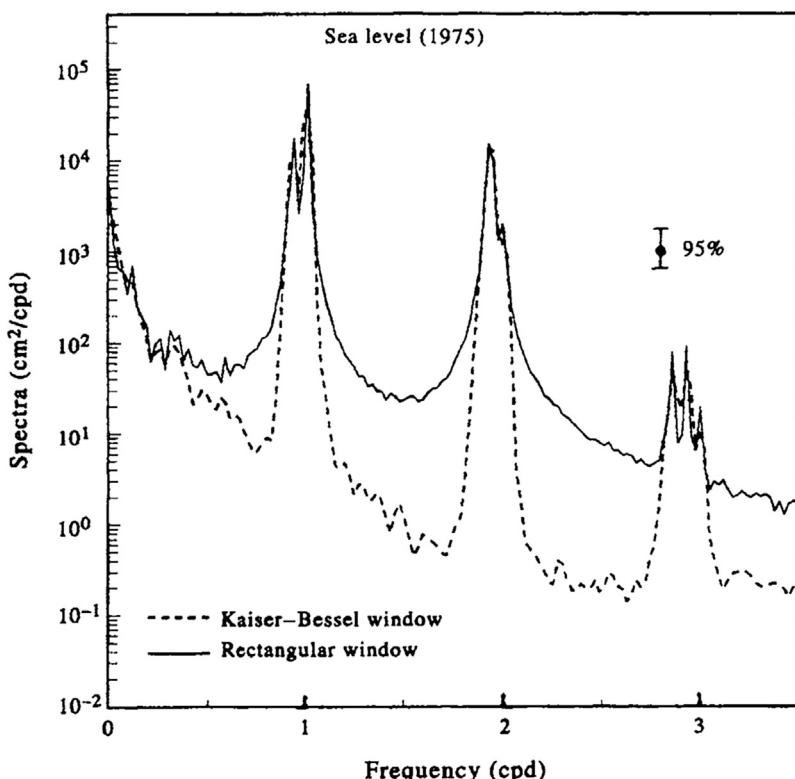
The usefulness of the Kaiser–Bessel window is nicely illustrated by Figure 5.27. Here, we compare the average spectra (in  $\text{cm}^2/\text{cpd}$ )

$$I_o(x) = \exp(|x|)/|x|^{1/2} \left\{ \left[ \left( \left[ \left( \left[ (3.92377 \times 10^{-3})Z - 1.647633 \times 10^{-2} \right)Z + 2.635537 \times 10^{-2} \right]Z - 2.057706 \times 10^{-2} \right)Z + 9.16281 \times 10^{-3} \right]Z - 1.57565 \times 10^{-3} \right)Z + 2.25319 \times 10^{-3} \right]Z + 1.328592 \times 10^{-2} \right]Z + 3.9894228 \times 10^{-1} \right\} \quad (5.98c)$$

where

$$Z = 3.75/|x| \quad (5.98d)$$

obtained from a year-long record of hourly coastal sea level following application of a rectangular window (the worst possible window)



**FIGURE 5.27** Spectra ( $\text{cm}^2/\text{cpd}$ ) of the hourly coastal sea-level height recorded at Victoria, British Columbia during 1975 following windowing (number of hourly samples,  $N=8750$ ). Linear frequency. Solid line: Rectangular window. Dashed line: Kaiser–Bessel window with  $\alpha=3$ . Both windows have a length of 1024 h (=42.67 days) and there are 32 degrees of freedom, using a total of 16 50% overlapping data segments. The tidal peak centered at 3 cpd results from nonlinear interactions within the semidiurnal frequency band. Vertical line is the 95% level of confidence. (Courtesy, Alexander Rabinovich.)

and a Kaiser–Bessel window (the best possible window) to a series of overlapping data segments. In each case, the window length is 42.7 days and there are  $K = 32$  DoF per spectral estimate, corresponding to roughly 16 separate spectral estimates derived using 50% window overlaps. Both windows preserve the strong spectra peaks within the tidal frequency bands centered at 1, 2, and 3 cpd. However, unlike the rectangular window, application of the Kaiser–Bessel window results in little energy leakage from the tidal bands to adjacent frequency bands. The high spectral levels at periods shorter than about 2 days ( $f > 0.5$  cpd) in the nontidal portion of the rectangular-windowed spectra is an artifact of the window. The slightly better ability of the rectangular window to resolve frequency components within the various tidal bands is outweighed by the high contamination of the spectrum at nontidal frequencies.

#### 5.4.7 Smoothing Spectra in the Frequency Domain

As we noted earlier, each spectral estimator for a random process is a chi-squared function with only two DoF. Because of this minimal number of DoF, some sort of smoothing or filtering is needed to increase the statistical significance of a given spectral estimate. The windowing approach described in the previous section, in which we partitioned the time series into a series of shorter overlapping segments, is one of a number of computational methods used to smooth (average) spectral estimates.

##### 5.4.7.1 Band Averaging

For a time series consisting of  $N$  data points, one of the simplest forms of smoothing is to use the DFT or FFT to calculate individual spectral estimates for the maximum number of frequency bands ( $N/2$ ) and then average together adjacent spectral estimates. The resultant spectral estimate is assigned to the midpoint of the

average. Thus, we could average bands 1, 2, and 3, to form a single spectral estimate centered at band 2, then bands 4, 5, and 6 to form an estimate centered at band 5, and so on. It is often useful in this type of *frequency band averaging* to use an odd-numbered smoother so that the center point is easily defined. In particular, if we were to average groups of three adjacent (and different) bands to form each estimate, the number of DoF per estimate would increase from two to six. In the case of the Blackman–Tukey autocovariance method, the equivalent procedure would be to use larger lag steps in the computation of the autocovariance function before its transform is taken. This is functionally equivalent to smoothing by averaging together the individual spectral estimates.

##### 5.4.7.2 Block Averaging

As we remarked earlier, a common smoothing technique is to segment the time series (of length  $N$ ) into a series of short, equal-length segments of length  $N_s$  (where  $N_s = N/K$ , and  $K$  is a positive integer). Spectra are then computed for each of the  $K$  segments and the spectral values for each frequency band then *block averaged* to form the final spectral estimates for each frequency band. If there is no overlap between segments, the resulting DoF for the composite spectrum will be  $2K$ . This assumes that the individual sample spectra have not been windowed and that each spectral estimate is a chi-squared variable with two DoF. Since the frequency resolution of a time series is inversely proportional to its length, the major difficulty with this approach is that the shorter time series have fewer spectral values than the original record over the same Nyquist frequency range. In other words, the maximum resolvable frequency  $1/(2\Delta t)$  remains the same since  $\Delta t$  is unchanged, but the frequency spacing between adjacent spectral estimates is increased for the short segments because of the reduced record lengths.

However, by not overlapping adjacent segments, we could be overly conservative in our estimate of the number of degrees of freedom (DoF). For that reason, most analysts overlap adjacent segments by 30–50% so that more uniform weighting is given to individual data points. The need for overlapping segments is necessary when a window is applied to each individual segment prior to calculation of the spectra. The effect of the window is to reduce the effective length of each segment in the time domain so that, for some sharply defined windows such as the Kaiser–Bessel window, even adjoining segments with 50% overlap can be considered independent time series for spectral analysis. As in Figure 5.27, the DoF of the periodograms averaged together is 4K, rather than 2K for the nonoverlapping segments. Consideration must be given to the correlation among individual estimates (the greater the overlap the higher the correlation). Nuttall and Carter (1980) report that 92% of the maximum number of equivalent degrees of freedom (EDOF) can be achieved for a Hanning window, which uses 50% overlap. Clearly, we must sacrifice something to gain improved statistical reliability. That “something” is a loss of frequency resolution due to the broad central lobe that accompanies windows with negligible side-lobes.

As an example, consider the spectrum of a 1-min sampled time series  $y(t) = A\cos(2\pi ft) + \epsilon(t)$  of length 512 min composed of Gaussian white noise  $\epsilon(t)$  ( $|\epsilon| \leq 1$ ) and a single cosine component of amplitude,  $A$ , and frequency  $f = 0.23$  cpm (period  $T = 1/f = 4.3$  min). The magnitude of the deterministic component,  $A$ , is five times the standard deviation of the white-noise signal and  $V[\epsilon] = (1/\sqrt{2})\text{cm}^2$ . The raw periodogram (Figure 5.28(a)) reveals a large narrow peak at the frequency (0.23 cpm) of the single cosine term plus a large number of smaller peaks associated with the white-noise oscillations. In this case, there has been no spectral smoothing and the resultant spectral estimates are chi-squared functions with two DoF. The variances of the spectral peaks are as large as the peaks themselves. If we

average together three adjacent spectral components (Figure 5.28(b)), we obtain a much smoother spectrum,  $S(f)$ . Here,  $S_i = S(f_i)$  is defined by  $S_i = 1/3[S(f_{i-1}) + S(f_i) + S(f_{i+1})]$ ,  $S_{i+3} = 1/3[S(f_{i+2}) + S(f_{i+3}) + S(f_{i+4})]$ , and so on. Each of the new spectral estimates now has six DoF instead of only two. The bottom two panels in this figure show what happens if we increase the number of frequency bands averaged together to seven (Figure 5.28(c)) and then to 15 (Figure 5.28(d)). Note that, with increasing DoF, our confidence in the existence of a spectral peak increases but delineation of the peak frequency decreases. With increasing DoF, there is increased smoothing of all spectral peaks (see also Figure 5.22). The same effect can be achieved by operating on the autocovariance function rather than on the Fourier spectral estimates. In particular, a spectrum similar to Figure 5.28(a) is obtained using the autocovariance transform method on the time series  $y(t)$  for a time lag of 1 min (the sampling interval). If we apply a lag of 3 min in computing the autocovariance transform, we obtain a spectrum similar to Figure 5.28(b), and so on. Any differences between the two methods will be due to computational uncertainties.

To determine the number of DoF for any block averaging, we define the normalized standard error  $\epsilon(G)$  of the one-sided spectrum,  $\tilde{G}_{yy}(f)$ , of the time series  $y(t)$  of finite length  $T = N\Delta t$ , as

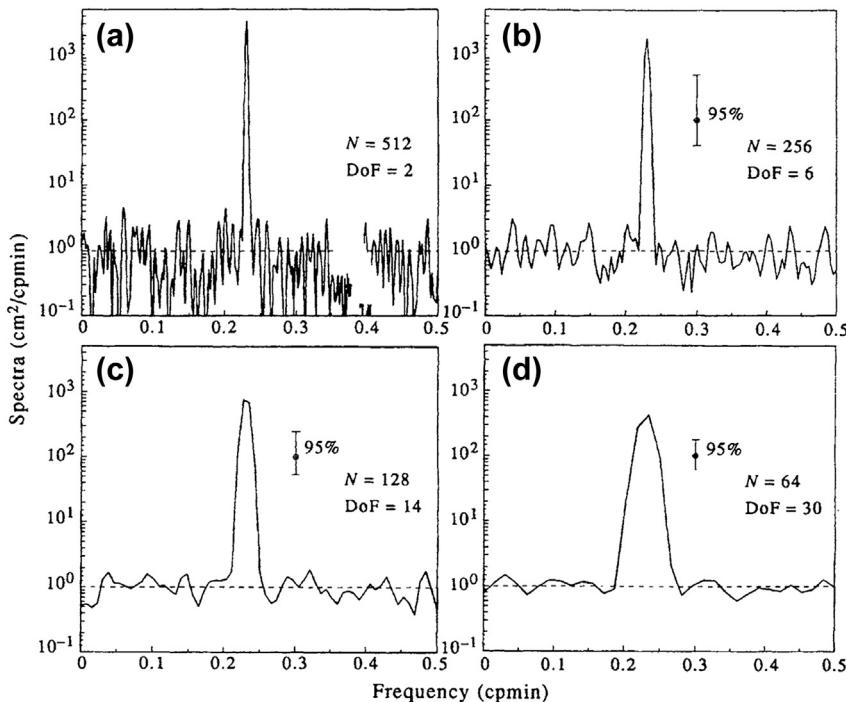
$$\epsilon[\tilde{G}_{yy}(f)] = \frac{V[\tilde{G}_{yy}(f)]^{1/2}}{G_{yy}(f)} \quad (5.99)$$

where  $G_{yy}$  is the true spectrum,  $V[\tilde{G}]$  is the variance of  $\tilde{G}$ , the tilde (~) denotes the raw estimate of the observed time series, and

$$\tilde{G}_{yy}(f)/G_{yy}(f) = \chi_2^2/2 \quad (5.100)$$

is a chi-square variable with  $n = 2$  DoF. For the narrowest possible resolution  $\Delta f = 1/T$ , we have

$$\epsilon[\tilde{G}_{yy}(f)] = \frac{(2n)^{1/2}}{n} = (2/n)^{1/2} \quad (5.101)$$



**FIGURE 5.28** Periodogram power spectral estimates for a time series composed of Gaussian white noise and a single cosine constituent with a frequency of 0.23 cpmmin and amplitude five times that of the white-noise component.  $N$  = number of spectral bands and vertical lines are the 95% confidence intervals. (a) Raw (unsmoothed) periodogram, with degrees of freedom (DoF) = 2; (b) Smoothed periodogram, by averaging three adjacent spectral estimates such that DoF = 6; (c) As with (b) but for seven frequency bands, and DoF = 14; (d) As with (b) but for 15 frequency bands, DoF = 30.

For maximum resolution,  $n = 2$  and so  $\epsilon(\tilde{G}) = 1$ , giving the not-so-useful result that the standard deviation of the estimate is as large as the estimate itself. If, on the other hand, we average the spectral estimates for each frequency for the maximum resolution spectra using a total of  $N_s$  separate and independent record segments of length  $T_s$  (where  $T = N_s \cdot T_s$ ) we find

$$\tilde{G}_{yy}(f) = \frac{2}{N_s T_s} \sum_{i=1}^{N_s} |Y_i(f_i, T_s)|^2 \quad (5.102)$$

so that

$$\epsilon[\tilde{G}_{yy}(f)] = (2n/2N_s)^{1/2} = (1/N_s)^{1/2} \quad (5.103)$$

The resolution (effective) bandwidth is  $b_e = N_s/T = 1/T_s$ . Since the first estimate, Eqn (5.101), gives two DoF per spectral band, the spectral averaging expressed by Eqn (5.103) gives  $2N_s$  DoF per frequency band.

#### 5.4.8 Confidence Intervals on Spectra

We can generalize Eqn (5.101) by noting that the ratio of the estimated spectrum and the expected values of the true spectrum

$$\frac{\nu \tilde{G}_{yy}(f)}{G_{yy}(f)} = \chi_\nu^2 \quad (5.104)$$

is distributed as a chi-square variable with  $\nu$  DoF. It then follows that

$$P\left[\chi_{\alpha/2,\nu}^2 < \frac{\nu\tilde{G}_{yy}(f)}{G_{yy}(f)} < \chi_{1-\alpha/2,\nu}^2\right] = 1 - \alpha \quad (5.105)$$

where

$$P\left[\chi_{\nu}^2 \leq \chi_{\alpha/2,\nu}^2\right] = \alpha/2 \quad (5.106)$$

Thus, the true spectrum,  $G_{yy}(f)$ , is expected to fall into the interval

$$\frac{\nu\tilde{G}_{yy}(f)}{\chi_{1-\alpha/2,\nu}^2} < G_{yy}(f) < \frac{\nu\tilde{G}_{yy}(f)}{\chi_{\alpha/2,\nu}^2} \quad (5.107)$$

with  $(1 - \alpha)100\%$  confidence. In this form, the confidence limit applies only to the frequency  $f$  and not to other spectral estimates. We further point out that the DoF,  $\nu$ , in the above expressions are different for windowed and nonwindowed time series. For windowed time series, we need to use the “equivalent” degrees of freedom (DoF), as presented in Table 5.5 for some of the more commonly used windows.

Another way to view these arguments is to equate  $\tilde{G}_{yy}(f)$  with the measured standard

deviation,  $s^2(f)$ , of the spectrum and  $G_{yy}(f)$  with the true variance,  $\sigma^2(f)$ . Then

$$\frac{(\nu - 1)s^2(f)}{\chi_{1-\alpha/2,\nu}^2} < \sigma^2(f) < \frac{(\nu - 1)s^2(f)}{\chi_{\alpha/2,\nu}^2} \quad (5.108)$$

If spectral peaks fall outside the range Eqn (5.108) then to the  $(1 - \alpha)100\%$  confidence level they are unlikely to have occurred by chance. The confidence levels are found by looking up the values for  $\chi_{1-\alpha/2,\nu}^2$  and  $\chi_{\alpha/2,\nu}^2$  in a chi-square table, then calculating the intervals based on the observed standard deviation,  $s$ . (Confidence limits on spectral coherency functions are given in Section 5.6.6.1.)

#### 5.4.8.1 Confidence Intervals on a Logarithmic Scale

The confidence intervals derived above apply only to individual frequencies,  $f$ . This results from the fact that the confidence interval is determined by the value  $\tilde{G}_{yy}(f)$  of the spectral estimate and will be different for each spectral estimate. It would be convenient if we could have a single confidence interval that applies to all of the spectral values at all frequencies. To obtain such a confidence interval, we transform the spectrum using the  $\log_{10}$  function. Transforming the above confidence limits we have

$$\begin{aligned} \log [\tilde{G}_{yy}(f)] + \log \left[ \nu / \chi_{1-\alpha/2,\nu}^2 \right] &\leq \log [G_{yy}(f)] \\ &\leq \log [\tilde{G}_{yy}(f)] + \log \left[ \nu / \chi_{\alpha/2,\nu}^2 \right] \end{aligned} \quad (5.109)$$

or

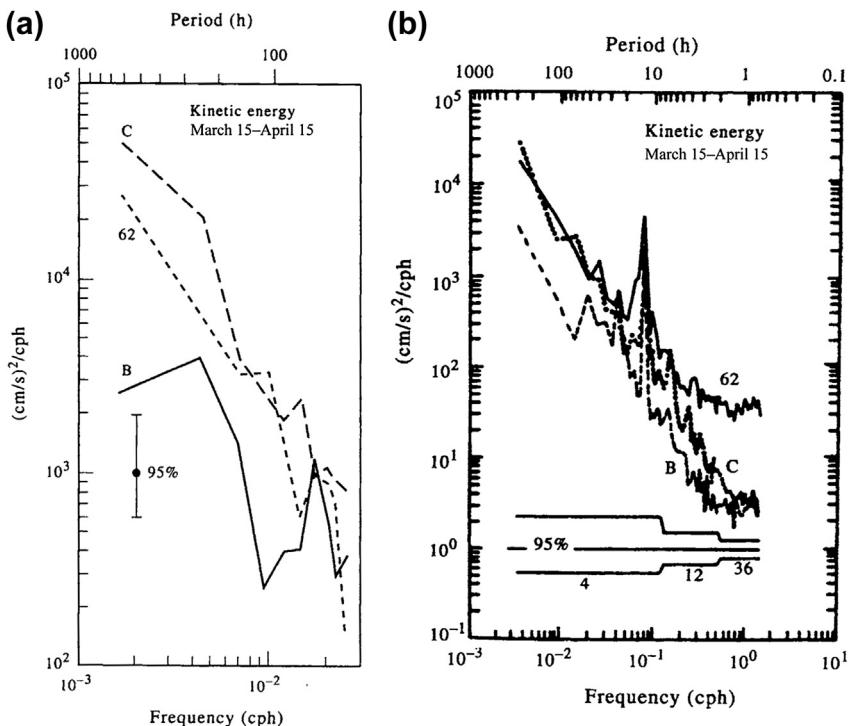
$$\begin{aligned} \log \left[ \nu / \chi_{1-\alpha/2,\nu}^2 \right] &\leq \log [G_{yy}(f)] - \log [\tilde{G}_{yy}(f)] \\ &\leq \log \left[ \nu / \chi_{\alpha/2,\nu}^2 \right] \end{aligned} \quad (5.110a)$$

$$\begin{aligned} \log \left[ \nu / \chi_{1-\alpha/2,\nu}^2 \right] &\leq \log [G_{yy}(f) / \tilde{G}_{yy}(f)] \\ &\leq \log \left[ \nu / \chi_{\alpha/2,\nu}^2 \right] \end{aligned} \quad (5.110b)$$

**TABLE 5.5** Equivalent Degrees of Freedom for Spectra Calculated Using Different Windows

Type of Window	Equivalent Degrees of Freedom
Truncated periodogram	$(N/M)$
Bartlett window	$3(N/M)$
Daniell window	$2(N/M)$
Parzen window	$3.708614(N/M)$
Hanning window	$(8/3)(N/M)$
Hamming window	$2.5164(N/M)$

*N* is the number of data points in the time series and *M* is the half-width of the window in the time (or spatial) domain. (From Priestley (1981).  $N \neq M$  for the truncated periodogram.



**FIGURE 5.29** Confidence intervals for current velocity spectra at 50-m depth for three locations (B, C, and 62) on the northeast Gulf of Alaska shelf ( $59.5^\circ \text{N}$ ,  $142.2^\circ \text{W}$ ), March 15–April 15, 1976. (a) 95% interval for the low-pass filtered currents. The single vertical bar applies to all frequencies; (b) 95% interval for unfiltered records. Confidence interval narrows at higher frequencies with the increased number of degrees of freedom (4–36) used in selected frequency ranges. (Adapted from Muench and Schumacher (1979).)

where  $\log [G_{yy}(f)/\tilde{G}_{yy}(f)] \rightarrow 0$  as the estimated spectrum approaches the real spectrum; i.e.,  $\tilde{G}_{yy}(f) \rightarrow G_{yy}(f)$ . When the estimated spectrum is plotted on a log scale, a single vertical confidence interval is determined for all frequencies by the upper and lower bounds in the above expressions (Figure 5.29(a)). The spectral estimate  $\tilde{G}_{yy}(f)$  itself is no longer a part of the confidence interval. This aspect, together with the fact that most spectral amplitudes span many orders of magnitude, is a principal reason for presenting spectra as log values. If larger numbers of spectral estimates are averaged together at higher frequencies (i.e.,  $\nu$  is increased), the confidence interval narrows with increasing frequency

(Figure 5.29(b)). Note that the length of the confidence interval is longer above the central point than below.

#### 5.4.8.2 Fidelity and Stability

The general objective of all spectral analysis is to estimate the function  $G_{yy}(f)$  as accurately as possible. This involves two basic requirements:

1. The mean smoothed spectrum,  $\tilde{G}_{yy}(f)$ , be as close as possible to the actual spectrum  $G_{yy}(f)$ . That is, the bias

$$B(f) = G_{yy}(f) - \tilde{G}_{yy}(f) \quad (5.111)$$

should be small. If this is true for all frequencies, then  $\tilde{G}_{yy}$  is said to reproduce  $G_{yy}(f)$  with high *fidelity*.

- For a time series of length  $T$  that has been segmented into  $M$  pieces for spectral estimation, the variance of the smoothed spectral estimator for bandwidth  $b_1$  is

$$V[\tilde{G}_{yy}(f)] \approx \frac{(M/b_1)}{T} [G_{yy}(f)]^2 \quad (5.112)$$

and should be small. If this is true, the spectral estimator is said to have high *stability*.

#### 5.4.9 Zero-Padding and Prewitthing

For logistical reasons, many of the time series that oceanographers collect are too short for accurate definition of certain spectral peaks. The frequency resolution  $\Delta f = 1/T$  for a record of length  $T$  may not be sufficient to resolve closely spaced spectral components. Also, discrete points in the computed spectrum may be too widely spaced to adequately delineate the actual frequency of the spectral peaks. Unfortunately, the first problem—that of trying to distinguish waveforms with nearly the same frequency—can only be solved by collecting a longer time series; i.e., by increasing  $T$  to sharpen up the frequency resolution  $f$  of the periodogram. However, the second problem—that of locating the frequency of a spectral peak more precisely—can be addressed by padding (extending) the time series with zeros prior to Fourier transforming. Transforming the data with zeros serves to refine the frequency scale through interpolation between PSD estimates within the Nyquist interval  $-f_N \leq f \leq f_N$ . That is, additional frequency components are added between those that would be obtained with a nonzero-padded transform. Adding zeros helps fill in the shape of the spectrum but in no case is there an improvement in the fundamental frequency resolution. *Zero-padding* is useful for: (1) smoothing the appearance of the periodogram estimates via

interpolation; (2) resolving potential ambiguities where the frequency difference between line spectra is greater than the fundamental frequency resolution; (3) helping define the exact frequency of spectral peaks by reducing the “quantization” accuracy error; and (4) extending the number of samples to an integer power of two for FFT analysis. An example of how zero-padding improves the spectral resolution of a simple digitized data set is provided in [Figure 5.30](#). We again emphasize that increased zero-padding helps locate the frequency of discernible spectral peaks, in this case the peaks of the  $\sin x/x$  function, but cannot help distinguish closely spaced frequency components that were unresolved by the original time series prior to padding.

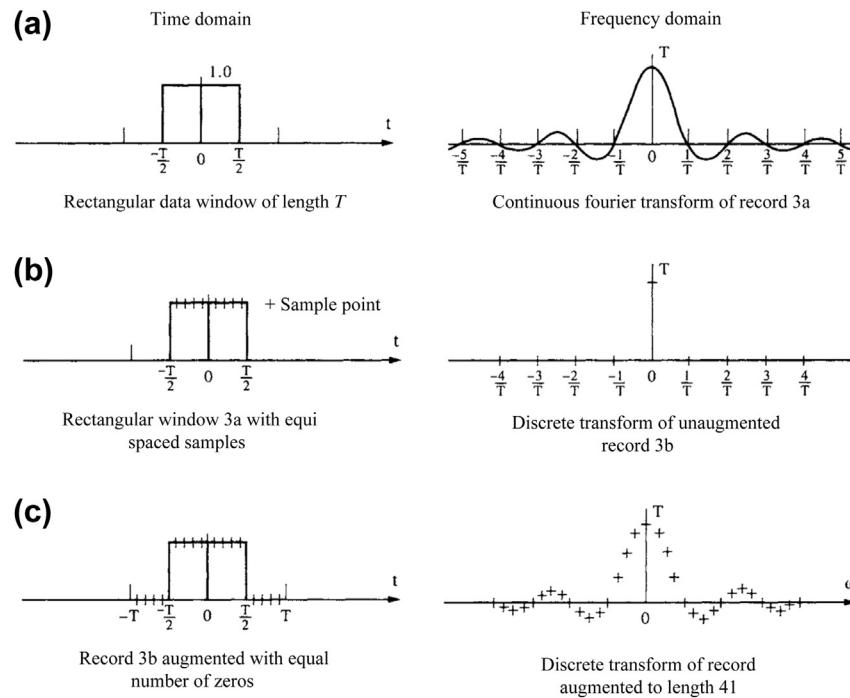
*Prewitthing* is a filtering or smoothing technique used to improve the statistical reliability of spectral estimates by reducing the leakage from the most intense spectral components and low-frequency components of the time series that are poorly resolved. To reduce the biasing of these components, the data are smoothed by a window whose spectrum is inversely proportional to the unknown spectrum being considered. Within certain frequency bands, the spectrum becomes more uniformly distributed and approaches that of white noise. Information on the form of the window necessary to construct the white spectrum must be available prior to the application of the smoothing. In effect, the time series,  $y(n\Delta t)$  is filtered with the weighting function,  $w(n\Delta t)$  such that the output is

$$y'(n\Delta t) = w(n\Delta t) \cdot (n\Delta t) \quad (5.113)$$

has a nearly white spectrum. Once the spectrum  $S'_{yy}(\omega)$  is determined, the desired spectrum is derived directly as

$$S_{yy}(\omega) = \frac{S'_{yy}(\omega)}{|W(\omega)|^2} \quad (5.114)$$

The best aspects of the parametric and nonparametric spectral techniques can be combined if a parametric model is used to prewhiten the time series prior to the application of a



**FIGURE 5.30** Use of zero-padding to improve the delineation of spectral peaks. (a) A continuous box-car window of length,  $T$  and its continuous Fourier transform; (b) a discrete sample of (a) at equally spaced sampling intervals and its discrete Fourier transform; (c) same as (b) but with zero-padding of  $2T$  data points. Note that the middle panel on the right is not a misprint. Transform values lie on the horizontal axis at the points  $-4/T$ ,  $-3/T$ , and so on. (From Henry and Graefe (1971).)

smoothed periodogram analysis. In most pre-whitening situations, one is limited to using the first-difference filter in which the current data value has subtracted from it the next value multiplied by some weighting coefficient,  $0 \leq \alpha \leq 1$ . That is  $y'(t) = y(t) - \alpha y(t + \Delta t)$ . The weighting coefficient can be taken as equal to the correlation coefficient of the initial data series with a shift of one time step,  $\Delta t$ . The filter suppresses low frequencies and stresses high frequencies and has a frequency response

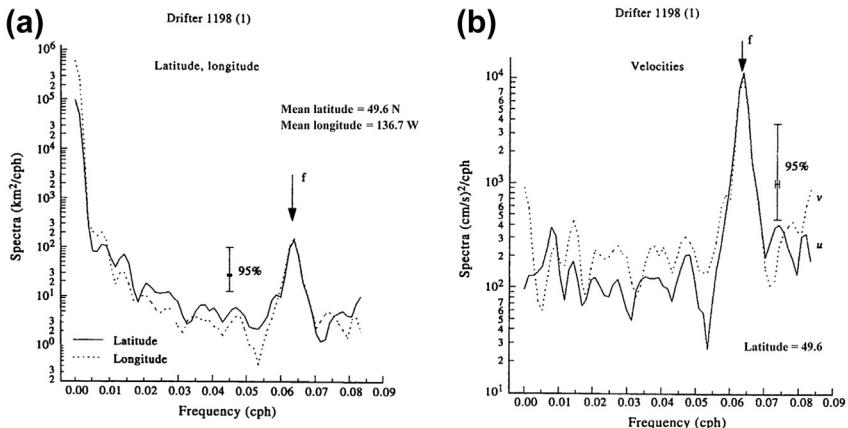
$$\begin{aligned} W(f) &= \left[ 1 - \alpha e^{-i2\pi f \Delta t} \right]^2 \\ &= 1 - 2\alpha \cos(2\pi f \Delta t) + \alpha^2 \end{aligned} \quad (5.115)$$

Prewhitenning reduces leakage and increases the effectiveness of frequency averaging of the spectral estimate (reduces the random error).

The reduced leakage gives rise to a greater dynamic range of the analysis and allows us to examine weak spectral components. Notice that, if  $Y(f)$  is the Fourier transform of  $y(t)$ , then the Fourier transform of  $y'(t)$  is

$$Y'(\omega) = \int_t y'(t) e^{-i\omega t} dt \approx \omega Y(\omega) \quad (5.116)$$

so that *first differencing* is like a linear high-pass filter with amplitude  $|W(\omega)| = |\omega|$ . This effect shows up quite well in the processing of satellite-tracked drifter data. Spectra of the drifter positions (longitude,  $x(t)$ ; latitude,  $y(t)$ ) as functions of time,  $t$ , are generally “red” whereas the spectra of the corresponding drifter velocities (zonal,  $u = \Delta x / \Delta t$ ; meridional,  $v = \Delta y / \Delta t$ ) are considerably “whiter” (Figure 5.31).



**FIGURE 5.31** Effect of a first-difference (high-pass) filter on resulting spectra. (a) Spectra of longitude ( $\Delta x$ ) and latitude ( $\Delta y$ ) displacements of a satellite-tracked drifter launched in the northeast Pacific in September 1990 ( $\Delta t = 3$  h; duration,  $T = 90$  days); (b) as with (a) but for the zonal ( $u = \Delta x/\Delta t$ ) and meridional velocity ( $v = \Delta y/\Delta t$ ). Mean position of the drifter was  $49.6^\circ$  N,  $136.7^\circ$  W;  $f$  denotes the mean inertial frequency; vertical line is the 95% confidence interval.

#### 5.4.10 Spectral Analysis of Unevenly Spaced Time Series

Most discrete oceanographic time series data are recorded at equally spaced time increments. However, some situations arise where the recorded data are spaced unevenly in time or space. For example, positional data obtained from satellite-tracked drifters are sampled at irregular time intervals due to the eastward progression in the swaths of polar-orbiting satellites and to the advection of the drifters by surface currents. Repeated time series oceanic transects are typically spaced at irregular intervals due to the vagaries of ship scheduling and weather. In addition, instrumental problems and data dropouts generally lead to “gappy,” irregularly spaced time series.

As noted in Section 3.17, a common technique for dealing with irregularly sampled or gappy data is to interpolate data values to a regular grid. This works well as long as there are not too many gaps and the gaps are of short duration relative to the signals of interest. Long data gaps can lead to the creation of erroneous low-frequency oscillations in the data at periods

comparable to the gap lengths. Only for the least-squares (LS) method for harmonic analysis described in Section 5.9 is unevenly sampled data perfectly acceptable. Vaníček (1971), Lomb (1976) and others have devised an LS spectral analysis method for unevenly spaced time series. The Lomb method described by Press et al. (1992) evaluates data, and associated sines and cosines, at the times,  $t_n$ , that the data are measured. For the  $N$  data values  $x(t_n) = x_n$ ,  $i = 1, \dots, N$ , the Lomb-normalized periodogram is defined as

$$P(\omega) = \frac{1}{2\sigma^2} \left\{ \frac{\left[ \sum_{n=1}^N (x_n - \bar{x}) \cos[\omega(t_n - \tau)] \right]^2}{\sum_{n=1}^N \cos^2[\omega(t_n - \tau)]} + \frac{\left[ \sum_{n=1}^N (x_n - \bar{x}) \sin[\omega(t_n - \tau)] \right]^2}{\sum_{n=1}^N \sin^2[\omega(t_n - \tau)]} \right\} \quad (5.117)$$

where as usual

$$\bar{x} = \frac{1}{N} \sum_{n=1}^N x_n; \quad \sigma^2 = \frac{1}{N-1} \sum_{n=1}^N (x_n - \bar{x})^2 \quad (5.118)$$

are the mean and standard deviation of the time series, and the time offset,  $\tau$ , is defined by

$$\tan(2\omega\tau) = \frac{\sum_{n=1}^N \sin(2\omega t_n)}{\sum_{n=1}^N \cos(2\omega t_n)} \quad (5.119)$$

The offset  $\tau$  renders Eqn (5.117) identical to the equation we would derive if we attempted to estimate the harmonic content of a data set at frequency  $\omega$  using the linear LS model

$$x(t) = A \cos(\omega t) + B \sin(\omega t) \quad (5.120)$$

In fact, Vaníček's founding paper on the technique refers to it as an LS spectral analysis method. The method, which gives superior results to FFT methods, weights the data on a per point basis rather than on a time-interval basis. By not using weights that span a constant time interval, the method reduces errors introduced by unevenly sampled data. For further details on the Lomb periodogram, including the introduction of significance testing of spectral peaks, the reader is referred to Press et al. (1992; pp. 569–577).

#### 5.4.11 General Spectral Bandwidth and Q of the System

Once the PSD,  $S(\omega)$ , has been computed, the general spectral bandwidth BW may be determined from the three moments,  $m_k$ , of the spectra

$$\begin{aligned} m_k &= \int_0^\infty \omega^k S(\omega) d\omega, \quad k = 0, 1, 2 \\ &\approx \sum_{i=0}^{N/2} \omega_i^k S(\omega_i) \Delta\omega \end{aligned} \quad (5.121)$$

where  $N/2$  is the number of spectral estimates and  $\Delta\omega$  is the frequency resolution of the spectral estimates (cf. Masson, 1996). In particular

$$BW = [(m_2 m_0 / m_1^2) - 1]^{1/2} \quad (5.122)$$

The bandwidth,  $\Delta\omega_{BW}$ , of a particular spectral peak within an oscillatory system can be used to estimate the dissipation of the system at the peak (resonant) frequency,  $\omega_r$ . Specifically, the "Q" or *Quality factor* of the system measures the amount of energy,  $E$ , stored in a linear oscillator compared to the amount of energy lost per cycle through frictional dissipation,  $\omega^{-1} dE/dt$  (Rabinovich, 2009). The Q-factor characterizes the sharpness of the resonant frequency and is commonly used as a direct measure of tidal dissipation in the ocean. Suppose that the energy of a simple linear system passes through a maximum at resonance frequency and that the energy of the system falls to 50% of its maximum value at frequencies  $\omega \approx \omega_r \pm \Delta\omega_{BW}/2$ . The Q of the system is then given by

$$Q = \frac{\omega E}{dE/dt} = \frac{\omega_r}{\Delta\omega_{BW}} = \alpha^{-1} \quad (5.123)$$

where  $E = E_o e^{-\alpha\omega t}$  is the system energy as it decays from an initial value  $E_o$  with a dimensionless damping coefficient,  $\alpha$ . For example, Wunsch (1972) finds  $Q \approx 3.3$  for an apparent resonant period of 14.8 h for the North Atlantic Ocean while Garrett and Munk (1971) obtain a global-wide lower bound of 25 for normal modes near the semidiurnal frequency.

#### 5.4.12 Summary of the Standard Spectral Analysis Approach

In summary, PSD estimates for time series  $y(t)$  can be obtained as follows using the standard autocorrelation and periodogram approaches:

1. Remove the mean and trend from the time series. Failure to remove the trend can lead to spurious energy (power) at low

frequencies. Remove *obvious* “spikes” caused by errant sensor responses or other forms of recording glitch, and also try to adjust the data series for discontinuities caused by internal offsets in the instrument or to sudden changes in sensor position or depth (Figure 5.32(a)). Removing spikes and adjusting for offsets is not as easy as it sounds. However, if not taken into account in the original time series, spikes and offsets can lead to erroneous spectral distributions (Figure 5.32(b)).

2. If block averaging is to be used to improve the statistical reliability of the spectral estimates (i.e., to increase the number of DoF), divide the data series into  $M$  sequential blocks of  $N'$  data values each, where  $N' = N/M$  (see Section 5.4.7). Depending on which type of window is to be applied, the sequential blocks can have up to 50% overlap.
3. To partially reduce end effects (Gibbs’ phenomenon) or to increase the series length to a power of two for FFT analysis, pad the data with  $K \leq N$  zeroes. Also pad the record with zeroes if you wish to increase the frequency resolution or center spectral estimates in specific frequency bands. To further reduce end effects and side-lobe leakage, taper the time series using a Hanning (raised cosine) window, Kaiser–Bessel window, or other appropriate window (see Section 5.4.6).
4. Compute the Fourier transforms,  $Y(f_k)$ ,  $k = 0, 1, 2, \dots, N - 1$ , for the time series (for convenience, we have taken the number of padded values as  $K = 0$ ). For block-segmented data, calculate the Fourier transforms,  $Y_m(f_k)$ , for each of the  $M$  blocks ( $m = 1, \dots, M$ ) where  $k = 0, 1, \dots, N' - 1$  and  $N' < N$ . To reduce the variance associated with the tapering in step 3, the transforms can be computed for overlapping segments.
5. Rescale the spectra to account for the loss of “energy” during application of the window.

That is, adjust the scale factor of  $Y(f_k)$  (or  $Y_m(f_k)$  in the case of smaller block size partitioning) to account for the reduction in spectral energy due to the tapering in step 3. For the Hanning window, multiply the amplitudes of the Fourier transforms by  $\sqrt{8/3}$ . The rescaling factors for other windows are listed in the right-hand column of Table 5.5.

6. Compute the raw PSD for the time series (or for each block) where for the two-sided spectral density estimates:

$$S_{yy}(f_k) = \frac{1}{N\Delta t} [Y^*(f_k)Y(f_k)], \\ k = 0, 1, 2, \dots, N - 1$$

(no block averaging)

$$S_{yy}(f_k; m) = \frac{1}{N\Delta t} [Y_m^*(f_k)Y_m(f_k)], \\ k = 0, 1, 2, \dots, N' - 1 \quad (5.124a)$$

(block averaging) and for the one-sided spectral density estimates

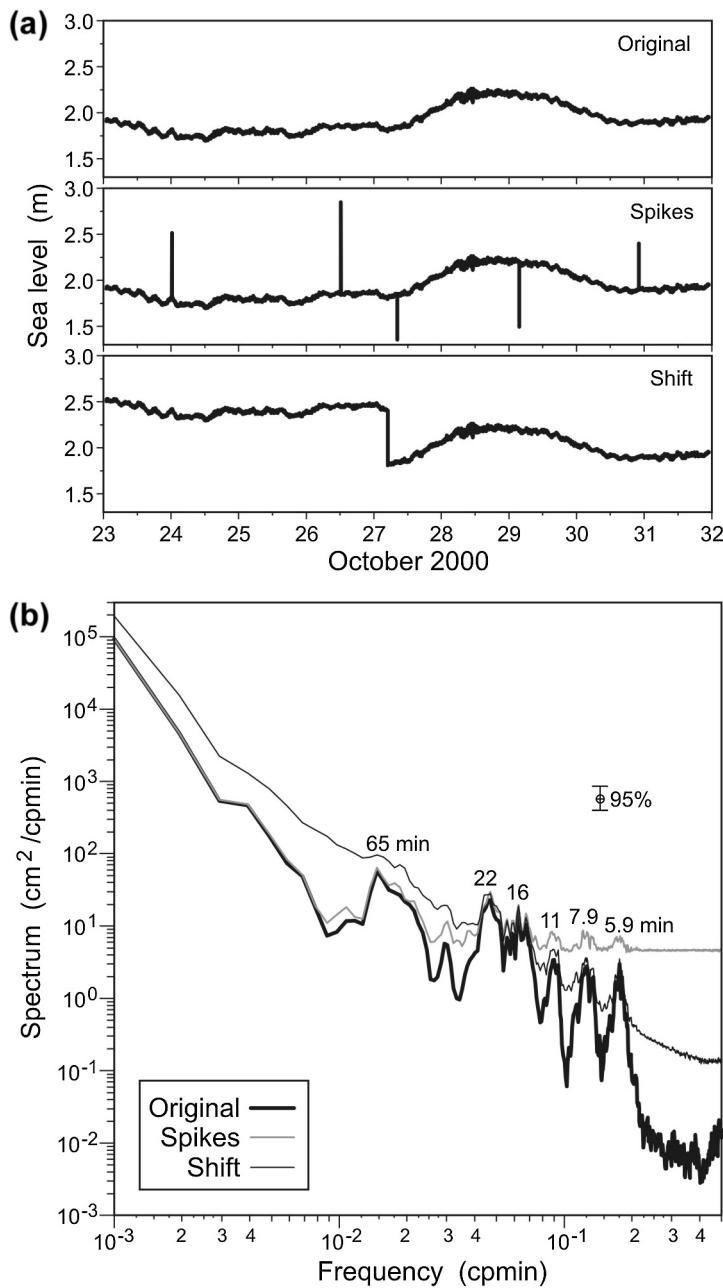
$$G_{yy}(f_k) = \frac{2}{N\Delta t} [Y^*(f_k)Y(f_k)], \\ k = 0, 1, 2, \dots, N/2$$

(no block averaging)

$$G_{yy}(f_k; m) = \frac{2}{N\Delta t} [Y_m^*(f_k)Y_m(f_k)], \\ k = 0, 1, 2, \dots, N'/2 \quad (5.124b)$$

(block averaging)

7. In the case of the block-segmented data, average the raw spectral density estimates from the  $M$  blocks of data, frequency-band by frequency-band, to obtain the smoothed periodogram for  $S_{yy}(f_k)$  or  $G_{yy}(f_k)$ . Remember, the trade-off for increased smoothing (more DoF) is a decrease in frequency resolution.



**FIGURE 5.32** Effects of data spikes and offsets on spectral estimates. (a) The top line shows a 1-min sampled sea-level times series for Victoria, British Columbia. The middle panel is the same series but in which five data values have been converted into “spikes” (single data points with anomalously high values). In the bottom panel, we have inserted a single negative offset of 0.5 m midway through the original time series; (b) spectra of the three time series in (a). There is considerable loss of high frequency information compared to the original time series and the addition of erroneous low frequency energy in the case of the offset time series. Numbers denote periods in minutes of selected spectral peaks. (Courtesy of Alexander Rabinovich.)

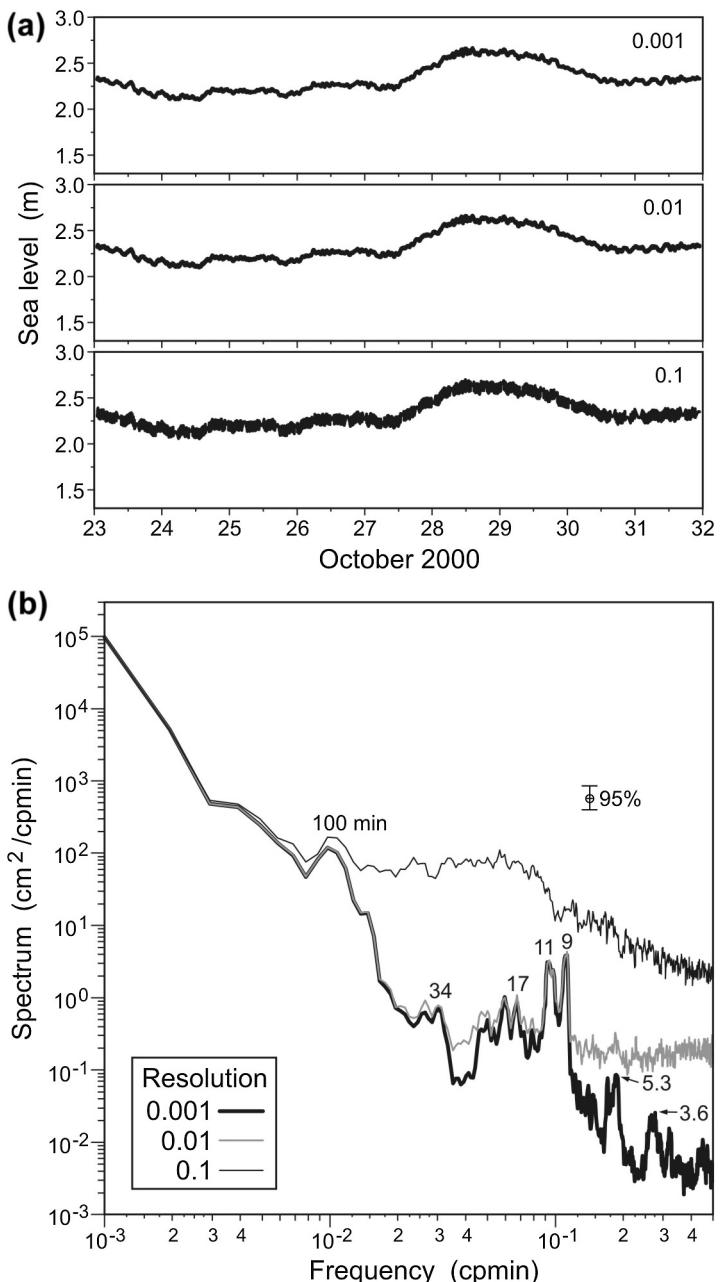
- Incorporate 80, 90, and/or 95% confidence limits in spectral plots to indicate the statistical reliability of spectral peaks. Most authors use the 95% confidence intervals.

We can illustrate some additional points in the above summary using the log–log spectra of sea-level oscillations (Figure 5.22) recorded over 14 days (20,160 min) in 1991 at Malokurilsk Bay on the west coast of Shikotan Island in the western Pacific. The main spectral peak is centered at a period of 18.6 min and corresponds to a wind-generated seiche amplitude of about 25 cm (Rabinovich and Levyant, 1992). All spectra have been obtained using segmented versions of the 14-day time series. Each time series segment has been smoothed using a Kaiser–Bessel window with 50% overlap between segments, and each segment has been treated as an independent time series. An FFT algorithm was used to calculate the spectrum for each segment. The smoothest spectrum (Figure 5.22(a)) is based on block averaged spectral estimates from roughly 157 overlapping segments ( $\sim 20,160 \text{ min}/128 \text{ min}$ ), the moderately smooth spectrum (Figure 5.22(b)) from the average of 39 overlapping segments, and the noisiest spectrum (Figure 5.22(c)) from the average of 10 overlapping segments. Taking into account the 50% overlap between segments and the fact that there are two DoF per raw spectral estimate, there are 628 ( $=157 \times 4$ ), 154, and 36 DoF for the three spectra, respectively. The smoothed spectrum in Figure 5.22(d) is derived using a slightly different approach. Although the segment lengths are the same as those in Figure 5.22(c) (i.e., 2048 min), the number of DoF is increased with increasing frequency,  $\omega$ . In this sliding scale, the lowest frequency range uses 36 DoF (as with Figure 5.22(c)), the next frequency band averages together the spectra for three adjacent frequencies to give 108 DoF, the next averages together the spectra for five adjacent frequencies to give 180 DoF, and so on.

As indicated by Figure 5.22, increasing the number of frequency bands averaged in each spectral estimate enhances the overall smoothness of the spectrum and improves the statistical reliability for specific spectral peaks. The number of degrees of freedom (DoF) increases and the confidence interval narrows. The penalty we pay for improved statistical confidence is reduced resolution of the spectral peaks. As in Figure 5.22(a), too much smoothing diminishes our ability to specify the frequency of spectral peaks and washes out peaks linked to some of the weaker seiches. Because each time series segment is so short, we also lose definition at the low-frequency end of the spectrum. As indicated by Figure 5.22(c), too little smoothing leads to a noisy spectrum for which few spectral peaks are associated with any physical processes. The sliding DoF scale in Figure 5.22(d) is a useful compromise.

One last point. Up until now, we have assumed that the sensors being used to collect the data have the sensitivity to record all of the variations of interest. If this is not the case, then no form of spectral analysis can extract information from the signal, regardless of the temporal resolution. Consider, for example, Figure 5.33(a), which shows a 9-day time series of bottom pressure (sea-level height) collected at 1-min intervals in Saanich Inlet on Vancouver Island. The top line shows the raw bottom pressure data sampled at 0.001 m (1 mm) equivalent vertical resolution. This is followed by time series generated by rounding off the 1-min data values to vertical resolutions that are factors of 10 and 100 lower than that of the original record. The impact of the lower vertical resolution is clearly displayed by the spectra in Figure 5.33(b). As would be the case for inadequate sensor resolution, the spectra of vertical displacements becoming increasingly degraded at higher frequencies. Although the sampling interval is the same for all time series, the spectral details

**FIGURE 5.33** The importance of sensor resolution to the detection of physical signals using spectral analysis. (a) 1-min sea-level record collected by a modern pressure gauge at Patricia Bay, Saanich Inlet, British Columbia. The top panel shows with original time series at 0.001 m (1 mm) vertical resolution, followed by time series formed by degrading (using decimal runoff) the original series to 0.01 and 0.1 m vertical resolution; (b) Spectra for the three time series showing the loss of information with increased degradation in vertical resolution. Numbers refer to spectral peaks in periods of minutes. The sampling rate is the same in all cases. (*Courtesy, Alexander Rabinovich.*)



of the sea-level signal are lost, including the background roll-off as a function of frequency.

*Covariance function:* Since the covariance function,  $C_{yy}(\tau)$ , and the autospectrum are Fourier transform pairs, the above analysis can be used to obtain a smoothed or unsmoothed estimate of the covariance function. To do this, first calculate the Fourier transform,  $Y(f)$ , of the time series, and determine the product  $S_{yy}(f) = N^{-1} \Delta t [Y^*(f)Y(f)]$ . Then take the IFT of the autospectrum,  $S_{yy}(f)$ , to obtain the covariance function,  $C_{yy}(\tau)$ . If the spectrum is unsmoothed prior to the IFT (or inverse fast Fourier transform (IFFT) if the FFT was used), we obtain the raw covariance function. If, on the other hand, the autospectrum is smoothed prior to the above integral using one of the spectral windows, such as the Hanning window, the covariance function also will be a smoothed function.

*A word of caution:* Although everyone agrees on the basic formulation for the DFT and the inverse discrete Fourier transform (IDFT), there are several ways to normalize the relations using the number of records,  $N$ . In our definitions, [Eqns \(5.26\) and \(5.28\)](#),  $N$  appears in the denominator of the IDFT. Some authors normalize using  $1/N$  in the DFT only, while others insist on symmetry by using  $1/\sqrt{N}$  in both DFT and its inverse. When using “canned” programs to obtain DFTs and IDFTs, ensure that you know how the transforms are defined and adapt your analysis to fit the appropriate processing routines.

## 5.5 SPECTRAL ANALYSIS (PARAMETRIC METHODS)

If the analytical model for a time series was known exactly, a sensible spectral estimation method would be to fit the model spectrum to the observed spectrum and determine any unknown parameters. In general, however,

oceanic variability is too complex to admit simple analytical models and parametric spectral estimates over the full frequency range of the data series. In addition, the imposition of an overly simplified spectral model could seriously degrade any estimation. On the other hand, it is reasonable that relatively simple spectral models might adequately reflect the system dynamics over limited frequency bands. Under some very general conditions, any stationary series can be represented in closed form by a statistical model in which the corresponding spectrum is a rational function of frequency (i.e., a ratio of two polynomials in  $\omega$ ).

If the time series under investigation is long relative to the timescales of interest, and if the spectrum is not overly complicated and does not have a too large dynamic range, the simple smoothed periodogram technique will probably yield adequate results. At a minimum, it will identify the major features in the spectrum. For shorter time series or in studies of fine spectral structure, other techniques may be more applicable. One such spectral analysis technique was developed by Burg (1967, 1972), who showed that it was possible to obtain the power spectrum by requiring the spectral estimate to be the most random (i.e., to have the maximum entropy) of any power spectrum, which is consistent with the measured data. This leads to a spectral estimate with a high frequency resolution since the method uses the available lags in the autocovariance function without modification and makes a nonzero estimate (prediction) of the autocorrelation function beyond those, which are routinely calculated from the data. Because the spectral values are computed using a maximum entropy condition, the resulting spectral estimates are not accurate in terms of spectral amplitude.

The most popular of the “modern” parametric techniques is the *autoregressive power spectral density* (AR PSD) model whose origins