Tutorial paper

Biomedical signal processing (in four parts)

Part 3 The power spectrum and coherence function

R. E. Challis

Ultrasonics & Digital Signal Processing Laboratory, University of Keele, Newcastle under Lyme, Staffs. ST5 5BG, UK

R. I. Kitney

Signal Processing Group, Imperial College, South Kensington, London SW7 2AZ, UK

Abstract—*This is the third in a series of four tutorial papers on biomedical signal processing and concerns the estimation of the power spectrum (PS) and coherence function (CF) od biomedical data. The PS is introduced and its estimation by means of the discrete Fourier transform is considered in terms of the problem of resolution in the frequency domain. The periodogram is introduced and its variance, bias and the effects of windowing and smoothing are considered. The use of the autocovariance function as a stage in power spectral estimation is described and the effects of windows in the autocorrelation domain are compared with the related effects of windows in the original time domain. The concept of coherence is introduced and the many ways in which coherence functions might be estimated are considered.*

Keywords—*Biomedical, Coherence function, Fourier, Power spectrum, Signal processing.*

Med. & Biol. Eng. & Comput., 1991, 29, 225–241

1 Introduction

THE FIRST PAPER of this series (CHALLIS and KITNEY, 1990) introduced the reader to time-domain methods which could be applied to individual signals in order to estimate their basic properties. A signal recording was regarded in a statistical sense as a sample from an underlying population signal. In our second paper (CHALLIS and KITNEY, 1991) these ideas were extended to include the basic frequency-domain methods of signal description and comparison, and we reviewed the equivalence between signal operations in the time domain and corresponding manipulations in the frequency domain. In this paper we develop the basic concepts of CHALLIS and KITNEY (1990) and (1991) and consider in detail the classical methods of estimating a signal power spectrum. An impressive battery of appropriate computational techniques has been developed in the past by a number of workers in many different institutions; all of the methods are of necessity compromises in some form or another and are therefore prone to a range of uncertainties. We consider these in some detail, particularly in the context of the tradeoff between good frequency-domain resolution and problems of bias and variance in the frequency-domain estimates. We conclude our discussions in this paper with what might justifiably be regarded as the *pièce de résistance* of the frequency-domain methods—that is the coherence function as a frequency-domain implementation of the idea of continuous cross-correlation between two signals.

In the next and final paper of this series we introduce the ideas of so-called modern spectral analysis and show how these methods may sometimes be used in cases where acceptable statistical compromises cannot be obtained using classical methods.

2 The power spectrum

The power spectrum, or power spectral density function, of a continuous time signal is defined as the square modulus of the Fourier transform of the signal.

$$P(\omega) = F(\omega)F^*(\omega) = |F(\omega)|^2 \tag{1}$$

where $P(\omega)$ is the power spectrum and $F(\omega)$ the Fourier transform. The amplitude of $P(\omega)$ is always positive for a real signal and it represents the signal power which exists in the frequency interval ω to $\omega + d\omega$. Power over a finite frequency interval, ω_1 to ω_2 say, is obtained by integration, thus

$$P(\omega_1, \omega_2) = \int_{\omega_1}^{\omega_2} P(\omega)\, d\omega \tag{2}$$

From a statistical viewpoint $P(\omega)$ can be regarded as the distribution in frequency space of the variance in the time-

Correspondence should be addressed to R. E. Challis.

Received 1st August 1989

domain signal, an idea which follows from Parseval's theorem. Related to this is the Wiener-Khintchine theorem which leads to an alternative definition of the power spectrum on the basis of the time-domain autocovariance function, defined as

$$S_{xx}(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} [x(b) - \bar{x}][x(t+\tau) - \bar{x}]\, dt \quad (3)$$

and the power spectrum is then

$$P(\omega) = \int_{-\infty}^{\infty} S_{xx}(\tau) e^{-j\omega\tau}\, d\tau \quad (4)$$

Both definitions of $P(\omega)$ are valid and are exactly equivalent for a continuous time signal of infinite duration. If it were possible not only to record a signal of infinite duration, but also to calculate its power spectrum, we would obtain a result which was totally reliable in a statistical sense: there would be no uncertainties such as bias, variance and limited resolution in our calculation. However, in real situations we have data of limited duration and only finite time and computational resources to apply to the calculation of power spectra. We therefore have to make estimates of the power spectra on the basis of data recordings of finite duration, using the notion that each finite length recording is in fact a sample taken from an infinite duration recording. We regard this recording of infinite duration as an underlying process known as the *population signal.*

There are many different methods which can be used for the estimation of power spectra; the performance of each one is limited in the sense that the available resolution in frequency space has some finite lower limit and that the estimates themselves suffer from error due to bias and variance which relate both to the manner in which the estimation is carried out and to the statistical properties of the original data, reckoned to be a sample recording of the population signal. The resolution, bias and variance are frequently not independent of each other and some compromise between all three is normally, knowingly or unknowingly, adopted by the investigator.

The methods available for special estimation fall into two major groupings; the *classical methods* are generally based on Fourier transformation in some form or another and are the subject of this section of our tutorial. In contrast the methods of so-called *modern spectral analysis* are based on the idea of fitting a system model to the data followed by the evaluation of the spectral properties of the model; these will form the basis of the final paper in this series. The choice between the two groups of technique is made on the basis of many inter-related factors, and is often quite a complicated issue. It does however seem that the most consistent reason for choosing modern methods is in situations where the classical ones will not yield adequate resolution for the signal at hand. In the sections of this paper which remain we will discuss the main methods of classical spectral analysis in the context of machine computation of spectral estimates of one-dimensional signals evenly sampled in time. This is representative of very many of the analytical tasks which may face the bioengineer. Similar principles apply to the analysis of multidimensional signals but an adequate discussion of these would be of such complexity as to mask many of the essential concepts we are trying to clarify in these tutorials. We begin our discussion with a review of the frequency resolution properties of the discrete Fourier transform (DFT), as this algorithm is employed in most classical methods of spectral estimation.

3 Resolution of the DFT

We recall that the DFT is the name given to the calculation of FS coefficients for a discrete signal which is either periodic or assumed to be periodic with a period equal to the length of the recording we have for analysis. Calculation is most commonly achieved by means of the FFT algorithm and data are organised so that the time-domain recording which is input to the calculation consists of an integer power of two samples in time with sampling interval T_s. The constraint of Nyquist's theorem requires that the original signal is band-limited so that no part of it (including its noise) contains energy at frequencies greater than $1/2T_s$. In most practical situations the sampling frequency is arranged to be at least four times the maximum frequency contained in the input signal. For an N-point input sequence an FFT operation yields as output a frequency-domain sequence of N spectral lines, each of which is a complex number corresponding to the complex addition of real and imaginary parts of each spectral component, analogously to the cosine and sine components of the Fourier series. The frequency spacing between spectral lines is $1/T$ where T is the duration of the time domain record. This sets the frequency resolution of the FFT method.

Clearly, the longer the duration of the input signal the better will be the frequency resolution in the sense that the interval between calculated spectral lines becomes smaller. However, there are many limitations to the initial recording length, the most obvious of which is that for a given sampling frequency a long input recording may exceed the available memory space in the computer employed for the analysis. We now illustrate the resolution problem with a few examples.

The left hand side of Fig. 1 shows four time-domain signals each consisting of 128 points. The first three are single-frequency sinusoids and the bottom trace is the summation of the first three. The right hand side of the figure shows the corresponding FFTs; as we would expect we see single spectral lines for the single-frequency traces and all three lines present for the composite signal. Fig. 2 shows three different sinusoids consisting of $1\frac{1}{2}$, $3\frac{1}{2}$ and $5\frac{1}{2}$ cycles together with their sum; the spectral representation of these is indeed very confusing. Because one cycle in 128 points represents the minimum resolvable frequency interval there is no position in the FFT output sequence which could receive these components of $(M + \frac{1}{2})$ cycles in the recording. The result is that their spectral energy is spread into adjacent frequency positions in the FFT. The right hand side of Fig. 2 shows this spread and downward-facing arrows indicate the positions at which the components would have turned up had a spectral line been available. The bottom right hand trace on Fig. 2 illustrates just how difficult it can be to discern the main spectral components due to the limited resolution of the transform.

The phenomenon of spreading of spectral energy is of course the leakage effect which we discussed in our previous paper of this series (CHALLIS and KITNEY, 1990). The resulting FFT is merely the circular convolution of the 'true' spectrum of the time-domain signal with the Fourier transform of the window used to extract the signal from the (notionally) infinitely long recording of which it is part. In the case of Figs. 1 and 2 the data extraction window was of sampled rectangular form, M points long say. Its modulus Fourier transform is $\sin(\omega M/2)/\sin\omega/2$ and it consists of a series of lobes with zeros at intervals in ω of $2\pi/M$. We recall that for a continuous signal the same effect would have occurred but the window in the frequency domain would have been of the form $\sin(\omega T/2)/(\omega T/2)$. The positions of the zeros are the reason why the

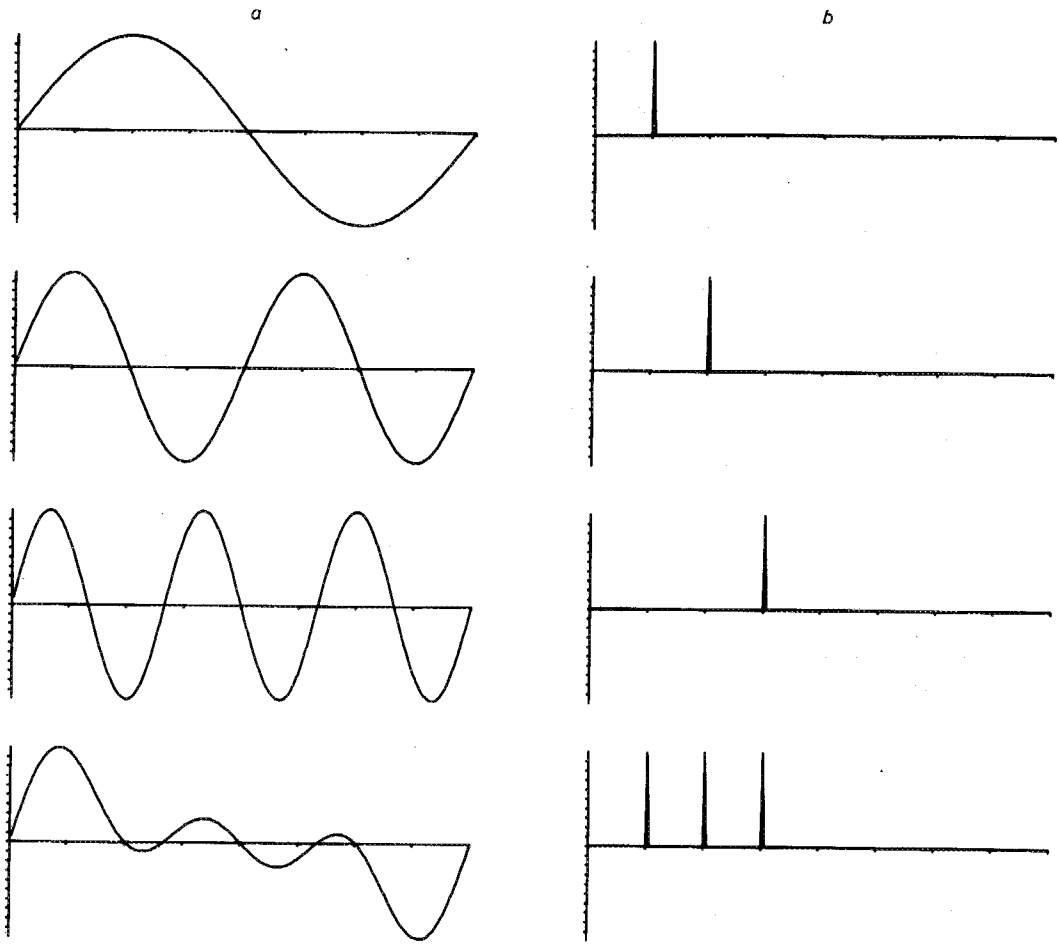

Fig. 1 *(a) Three single-frequency sinusoids of 128 samples each and their summation. (b) The FFTs of each of the four time-domain recordings*

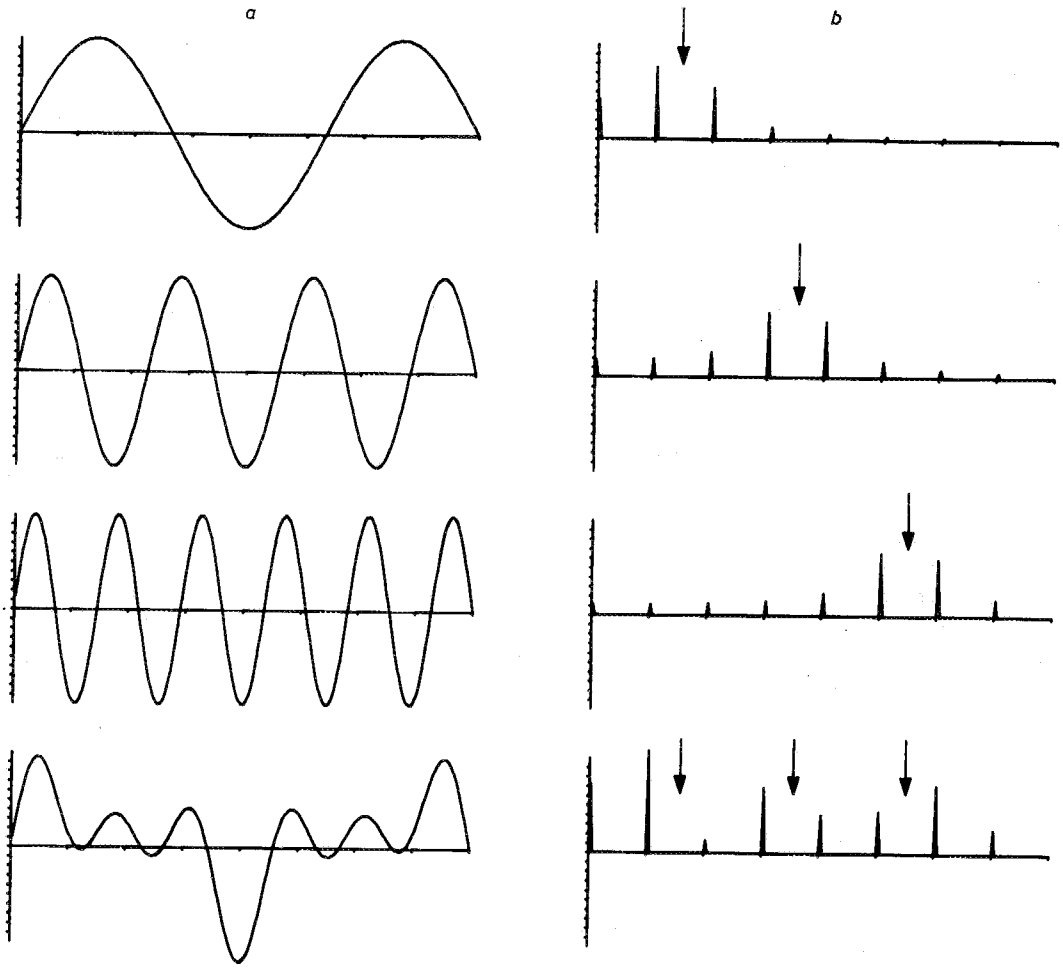

Fig. 2 *(a) Three single-frequency sinusoids which do not fit exactly into one data recording. (b) Their FFTs indicate leakage of energy on either side of their 'true' spectral positions, marked with downward arrows*

spectra on Fig. 1 appeared as we would have expected in the absence of leakage effects: the convolving window had zeros at every possible spectral line position except the centre one. With non-integer frequency components these zeros locate between possible spectral line positions and finite energy is spread to other spectral lines, as we have seen.

The basic resolution of the FFT can be increased by artificially increasing the length of the input recording by adding zero value signal to it. This method is known as *zero padding* and it is illustrated in Fig. 3. The original signal recording consists of two cycles in 128 data points, and this has been stretched to 1024 points by the addition of zeros. The FFT is then taken of this new 1024-point recording; it has a frequency resolution commensurate with its new recording length, the new interval between spectral lines being $\frac{128}{1024} = \frac{1}{8}$ times the original interval.

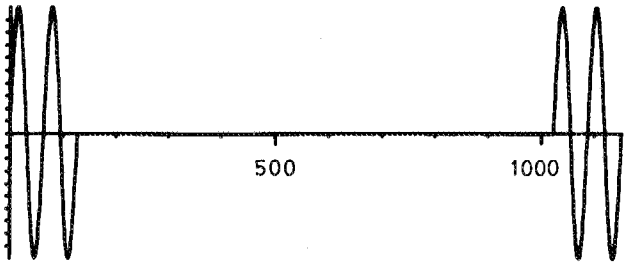

Fig. 3 *The method of zero padding. The 128-point two-cycle sine wave signal is stretched to 1024 data points by the addition of a sequence of zero-valued samples*

The application of this technique to our basic sinusoidal data is shown in Fig. 4. It is clear that resolution is increased in that we have more spectral lines in a given frequency interval. However, our ability to resolve particular frequency components is not much better than before and the convolution of the 'true' spectrum with the frequency domain window is clear. It is of course possible to separate a time-domain recording from the (notionally) long signal of which it is part by means of a window which has smoother edges than the simple stepwise on/off of the rectangular window. There are many such windows (see for instance ELLIOTT and RAO, 1982) and during the 1950s and 1960s a popular subject for research was known as 'window carpentry'! The most commonly known is the so-called Blackman-Tukey (or Hanning) window which weights an input signal recording with an inverted cosine according to

$$w(n) = \tfrac{1}{2}\left(1 - \cos\frac{2\pi n}{N}\right) \quad \text{for } n = 0, 1, 2, \ldots, (N-1) \quad (5)$$

When data are extracted with such a window the resulting frequency transform again consists of the 'true' spectrum in convolution with the Fourier transform of the window shape. Now the Fourier transform of the inverted cosine has a wider main lobe and much lower sidelobe amplitudes than the frequency-domain window function which corresponds to the rectangular data window. Fig. 5 shows spectral calculations carried out on the same zero padded sinusoidal data as before, only using the Hanning window over the non-zero part of the signal. The resulting FFT spectra show clearly the wider main lobe and lower sidelobes. Also illustrated is the possible 'loss' of a whole component (the two-cycle one) due to signal distortion brought about by the nonlinear windowing operation. It is also clear from Fig. 5 (bottom trace) just how difficult it can be to discern particular spectral components, even if they are present in the transformed data. It would be in signal conditions such as these that one would turn to the techniques of modern spectral analysis in order to separate closely spaced signal components which are easily 'losable' with Fourier methods.

4 The periodogram

The periodogram is an estimate of a signal power spectrum made on the basis of the modulus squared Fourier transform. Notionally it takes the form of a histogram whose bin width corresponds to the minimum resolvable frequency interval in the Fourier transform operation. For a discrete or line spectral plot based on a Fourier series computation this interval would be $1/T$, where T is the duration of the recording extracted for analysis. The definition of the periodogram can include continuous spectra based on the Fourier transform of a finite length signal; in this case the notional histogram bin width is infinitesimally small and equal to $d\omega$. The height of each bin is such that the bin area (width × height) is equal to the estimated spectral power in the corresponding frequency interval. The area under the whole periodogram is of course equal to the total estimated signal power. In practice we merely assume the notion of the histogram and regard the periodogram as the sequence of numbers corresponding to successive bin heights. Other definitions of the periodogram have been suggested from time to time and the interested reader is referred to CHATFIELD (1984) for further discussion.

The Fourier transform of a discrete time signal of finite length N is given by

$$X(\omega) = \sum_{n=0}^{N-1} x(n)e^{-j\omega n} \quad (6)$$

The periodogram is given by

$$I(\omega) = \frac{1}{N}|X(\omega)|^2 \quad (7)$$

This function is continuous in frequency space and will form the basis of our discussion of the periodogram method. It should be noted however that the majority of classical spectral computations are presently achieved by means of the FFT and the implication of this is that the resulting periodogram $I(k)$ will be a version of $I(\omega)$ sampled in frequency at intervals of $\omega = k2\pi/N$. Specifically, the DFT of $x(n)$ is

$$X(k) = \sum_{n=0}^{N-1} x(n)e^{-jkn(2\pi/N)} \quad (8)$$

and the discrete frequency periodogram is

$$I(k) = \frac{1}{N}|X(k)|^2 \quad (9)$$

The computational efficiency of the FFT has caused the sampled periodogram to find great favour as an estimator for power spectral density. It does of course suffer from the same limitation as the FFT in that the minimum resolvable frequency interval is $1/T$ Hz, where T is the original recording duration; this limit is equivalent to the digital frequency $(2\pi/N)$ which appears in eqn. 15 of our previous paper (CHALLIS and KITNEY, 1991). The effectiveness of the periodogram as an estimator of the underlying power spectrum is assessed in terms of the variance and bias of the resulting estimates. These are difficult to derive in a general form and the mathematics involved is beyond the scope of this tutorial. The interested reader is referred to JENKINS and WATTS (1968) and to discussions in OPPENHEIM and SCHAFER (1975) and CHATFIELD (1984). We

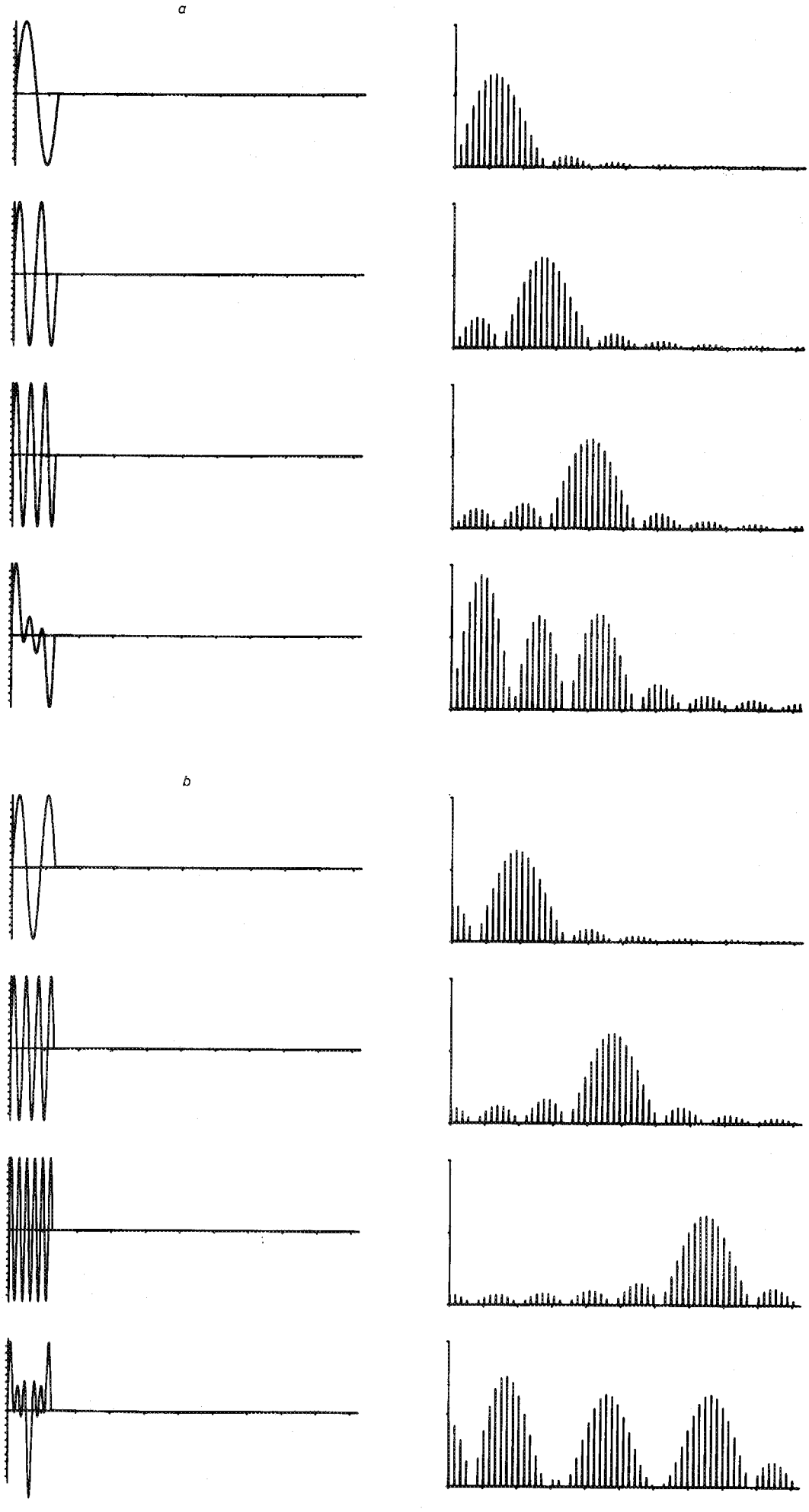

Fig. 4 *The effect of zero padding on the FFT. The sample sine wave signals are padded from 128 to 1024 points and their FFTs are shown on the right. (a) Integer number of periods; (b) integer + half number of periods*

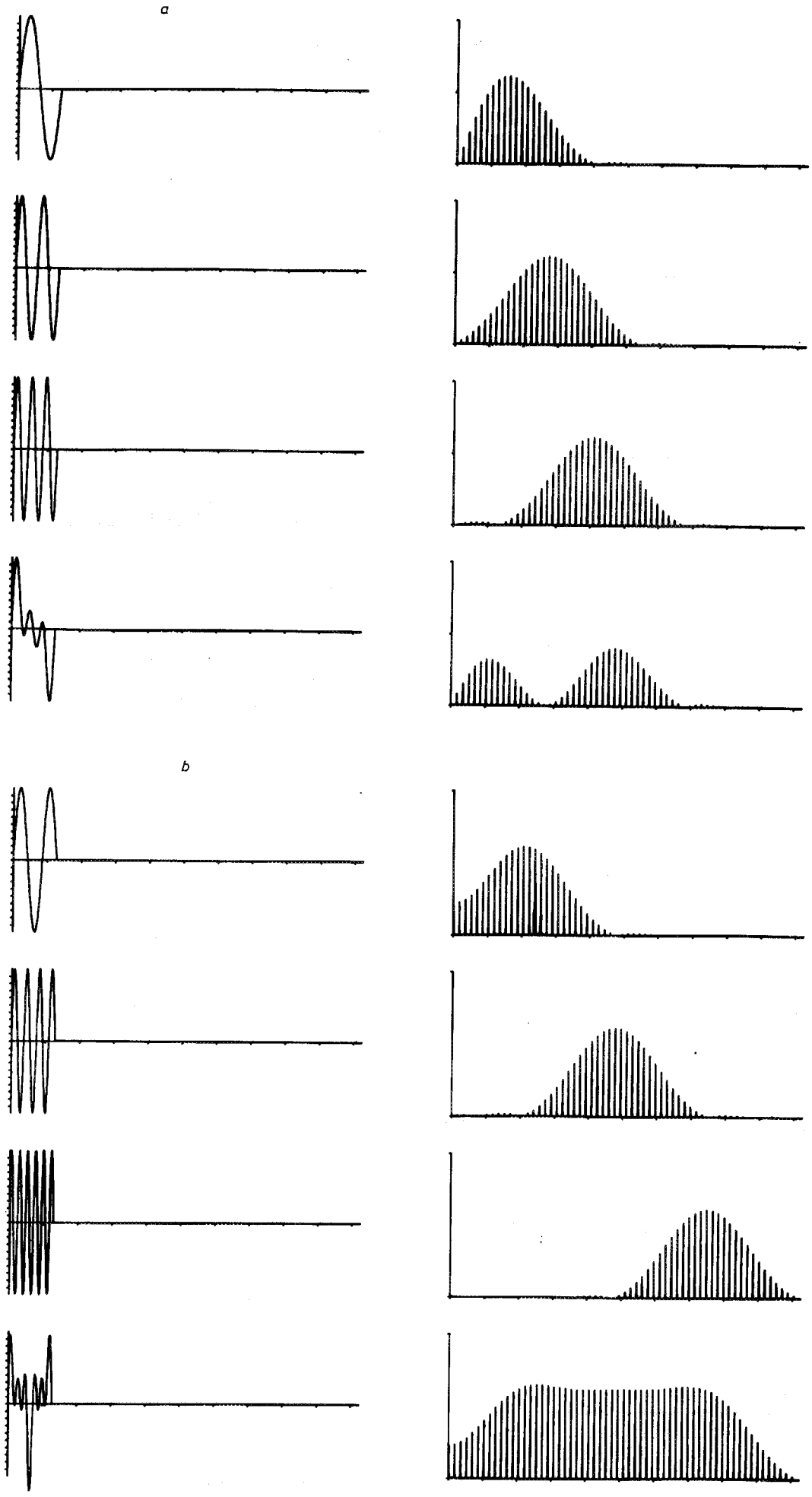

Fig. 5 *The same FFT computation carried out on simple sinusoidal data windowed with the inverted cosine function* $w(n) = \frac{1}{2}(1 - \cos 2\pi n/N)$

shall summarise the principal results here and discuss their implications for 'day to day' laboratory processing.

4.1 *Variance of the periodogram*

The simplest approach is to begin with a sampled time-domain signal N points long, each sample of which is representative of a wholly random process with a white spectrum, zero mean and variance σ_x^2. The real and imaginary components in the DFT are linear combinations of these samples and are also normally distributed and are independent of each other. On the basis of this it can be shown (JENKINS and WATTS, 1968; OPPENHEIM and SCHAFER, 1975) that the variance of the periodogram approaches σ_x^4 as N approaches infinity.

$$\text{var } [(I(\omega)] = \sigma_x^4 \quad \text{(10)}$$

$$\text{lt } N \rightarrow \infty$$

where

$$\text{var } [x(n)] = \sigma_x^2$$

The corresponding analysis for a non-white signal is rather more complicated. OPPENHEIM and SCHAFER (1975) take the approach that a non-white random sequence can be generated by starting with a random sequence and imposing a linear filtering process upon it. They derive an approximate expression for the variance of the periodogram and show that, if the squared magnitude of the frequency response of the linear system is $P_{xx}(\omega)$, then

$$\text{var } [I(\omega)] = P_{xx}^2(\omega) \quad \text{(11)}$$

$$\text{lt } N \rightarrow \infty$$

Thus as the recording length increases the variance of the periodogram approaches the square of the power spectrum value at each frequency. *This is a most important result.* It implies that the estimate $I(\omega)$ does not improve as more data points are brought into its calculation. Instead the variance approaches a constant value and $I(\omega)$ is thus not a consistent estimator (CHALLIS and KITNEY, 1990).

This property is perhaps to be expected when we compare the calculation of $I(\omega)$ with that of a consistent estimating process such as evaluation of the mean value of a sequence of numbers. The estimate of the mean improves (variance decreases) as more data points contribute to the single output quantity $\bar{x}$. With the periodogram there is not a single output quantity whose variance is to be controlled. In the discrete case $I(k)$ for example, N input data contribute to the estimation of N spectral lines.

The outcome of the inconsistency in $I(\omega)$ is that crude periodograms tend to be rather jagged curves with strong fluctuations about the true spectrum value. As a result the curves become difficult to interpret or even meaningless in some cases, and this may be true even when the data contain a proportion of deterministic energy mixed with random components. Fig. 6 illustrates the effect dramatically. The problem can be overcome in part by controlling

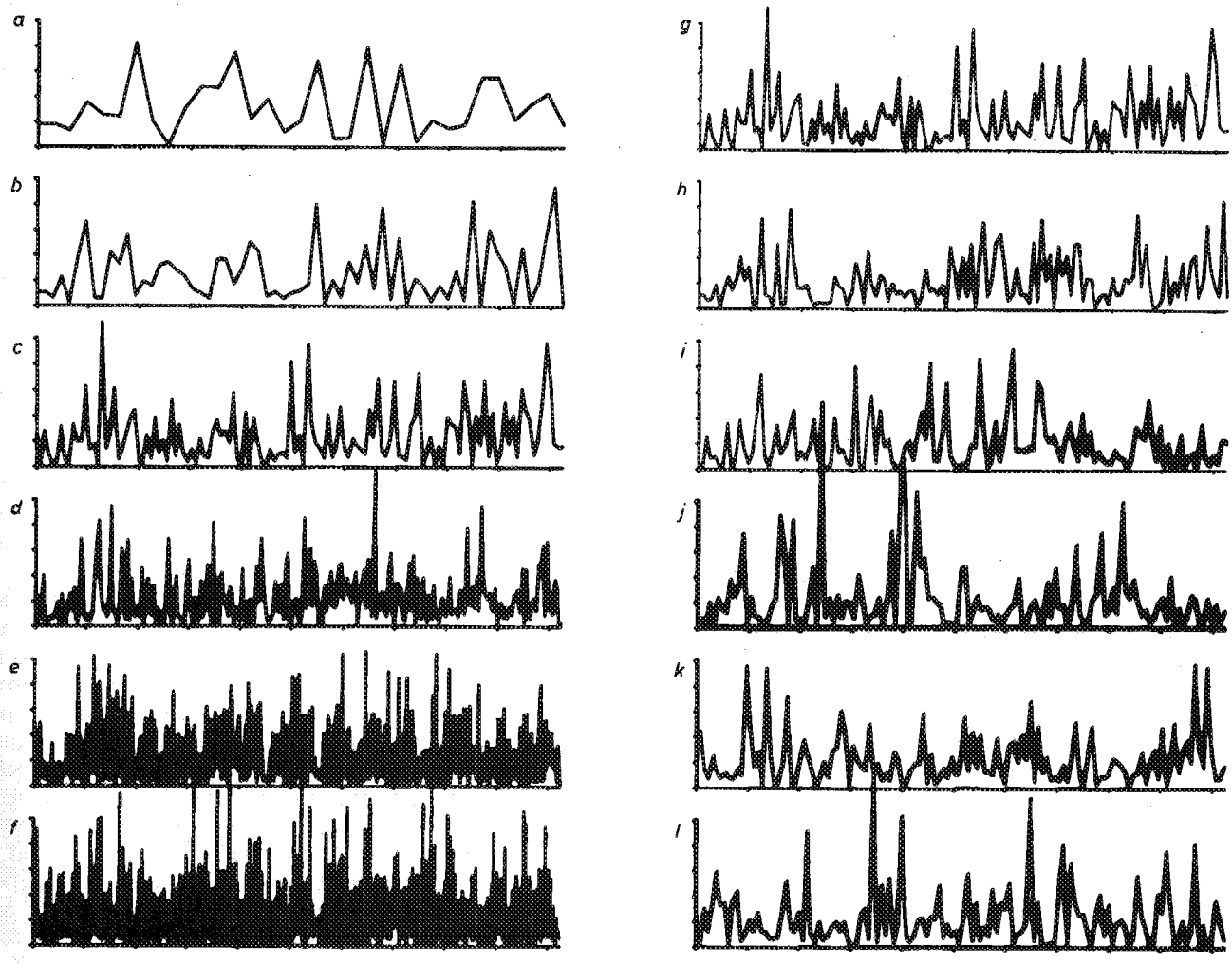

Fig. 6 *Experiment to demonstrate the variability of the periodogram. A 2048-point near Gaussian sequence (mean = 0, SD = 1·0) was generated and 12 periodograms were calculated. The left-hand plots show the first halves of six periodograms of differing lengths taken from the beginning of the data; lengths chosen were 64 point (upper), 128, 256, 512, 1024 and 2048 (lower). The right-hand side shows the first halves of six 256-point periodograms which were taken from six different starting points in the data (at the beginning (upper), point 64, 128, 256, 512, 1024 and 1280 (lower))*

the variance by means of averaging procedures, and we will consider this idea later.

4.2 *Bias in the periodogram*

Bias in an estimate means that the estimator is always likely to yield an erroneous result irrespective of the extent to which the variance is controlled. The estimate may be consistently too big or too small due to bias. Now the Fourier transform that underlies the periodogram is the convolution of the 'true' spectrum (subject to variance) with the Fourier transform of the window used to separate the recording of data. We have seen that this leads to leakage of energy from parts of the spectrum where it 'should' be to frequencies on either side; this occurs for both the Fourier transform and the DFT. Leakage causes the resulting transform to be in error and we interpret this error as bias. The periodogram, calculated on the basis of the Fourier transform or DFT, is thus a biased estimate of the underlying power spectrum and this bias is caused by the spectral leakage which results from the window process used to extract the recording of data from the (actual or assumed) underlying population signal.

The severity of the bias depends upon the spectral width of the data window, as does the frequency resolution of the periodogram (see Figs. 1 and 2). Now for good frequency resolution and low bias we require a data window which is narrow in the frequency domain, and this implies that the window will be wide in the time domain; i.e. the extracted data recording will be relatively long. We have already seen that long data sequences lead to periodograms of uncontrolled variance. A compromise thus exists between the need for short data sequences to give acceptable variance and the requirement for long data sequences to give acceptable frequency-domain resolution and adequately controlled bias.

Any solution to the compromise between resolution, bias and variance must be approached with great care and the method adopted will depend on the nature of the data it is required to analyse. In some instances the square modulus of a single FFT will suffice as a power spectral estimate and no further action will need to be taken. More frequently, and particularly with signals containing a high proportion of random energy, it will be necessary to employ more complex methods, most of which involve some type of averaging procedure to control the variance and smooth the spectrum.

4.3 *Smoothed spectral estimates*

There are various ways by which the variance of a spectral estimate can be reduced to effect a smoothing operation on the function and render it more intelligible. The simplest method consists of averaging adjacent samples in the periodogram according to some weighting function, a process known as windowing. More complicated methods such as the Bartlett and Welch (WELCH, 1967) methods involve segmenting the time-domain recording into sections, calculating a periodogram for each, and then forming the average of a number of segmental periodograms. Many of the more old-fashioned techniques involve calculation and modification of the autocovariance function followed by estimation of the power spectrum by means of the Wiener-Khintchine theorem (CHALLIS and KITNEY, 1991). There are an enormous number of possible permutations of technique, far too many to include in this brief tutorial. We shall therefore set out the principal details of each class of methods to the point where our reader can confidently begin experimentation with some idea of the results which should be expected.

4.4 *Windowing the periodogram*

The discrete periodogram, calculated by means of the FFT say, can be smoothed by averaging adjacent samples either with equal weight or according to some weighting function such as $(\frac{1}{4}, \frac{1}{2}, \frac{1}{4})$. The averaging process can be carried out rapidly by means of a recursive operation on the periodogram, using the methods of digital filtering (CHALLIS and KITNEY, 1982; 1983*a*; *b*) only now applied to discrete frequency domain recordings instead of time-domain sequences. The weighting function in frequency space is known as the window, generally denoted $W(\omega)$. The new periodogram which results from the windowing operation, the smoothed spectrum, is the convolution of the original periodogram with the window function, i.e.

$$I_s(\omega) = \frac{1}{2\pi} \int_{-\pi}^{\pi} I(\Omega) W(\omega - \Omega)\, d\Omega \tag{12}$$

Although it might be clear from an intuitive point of view that the smoother spectrum will have a reduced variance it is difficult to derive a quantitative expression for variance in a general way. OPPENHEIM and SCHAFER (1975), after making certain assumptions which include the condition that the window width is narrow by comparison with the spread of variations in the true power spectrum $P_{xx}(\omega)$, lead to the following result:

$$\text{var}\,[I_s(\omega)] \approx \left[\frac{1}{2\pi N} \int_{-\pi}^{\pi} W^2(\omega)\, d\omega\right] \text{var}\,[I(\omega)] \tag{13}$$

in the range $0 < \omega < \pi$. That is to say the variance of the periodogram is improved by a factor proportional to the mean square value of the window in frequency space. We note that the frequency-domain window has an inverse Fourier transform which is the discrete sequence given by

$$w(m) = \frac{1}{2\pi} \int_{-\pi}^{\pi} W(\omega) e^{j\omega m}\, d\omega \tag{14}$$

This leads to an alternative expression for the variance improvement:

$$\text{var}\,[I_s(\omega)] \approx \left[\frac{1}{N} \sum_{m=-(M-1)}^{M-1} w^2(m)\right] \text{var}\,[I(\omega)] \tag{15}$$

where the extent of the time domain function is $2M - 1$.

The variance improvement can therefore be expressed as a factor proportional to the mean square value of the inverse discrete Fourier transform of the frequency-domain window. This variance improvement is again quoted for the range $0 < \omega < \pi$. Derivations for $\omega = 0$ (DC) and $\omega = \pi$ (Nyquist frequency) yield different variance results and will not be included in our discussion here; we do not regard this as a serious omission because DC is removed from most signals and finite components at the Nyquist frequency would probably involve serious problems of alias.

An operation which averages over a number of spectral lines will clearly limit the frequency resolution in the periodogram as well as control its variance. The resolution worsens as the window is made wider in frequency and the *bandwidth* of the window enables us to quantify this effect. There are various definitions of bandwidth, and the simplest is the frequency interval between the innermost zeros of the function. Sometimes the half-power points of the function are used, in line with basic ideas of frequency response curves. Fig. 7 illustrates the most common window functions together with expressions for their bandwidth (based on the interval between inner zeros) and an

approximate measure of the improvement in variance we might expect from their use.

We should emphasise that in the majority of practical situations the investigator only need resort to fairly simple functions, the gains to be had from the more complicated windows being marginal in this method.

$$\text{var}\,[I_N(\omega)] = \frac{1}{K}\,\text{var}\,[I_M(\omega)] \tag{18}$$

The averaged periodogram is therefore a consistent estimate of the power spectrum because as the number of segments in the average is increased so the variance tends

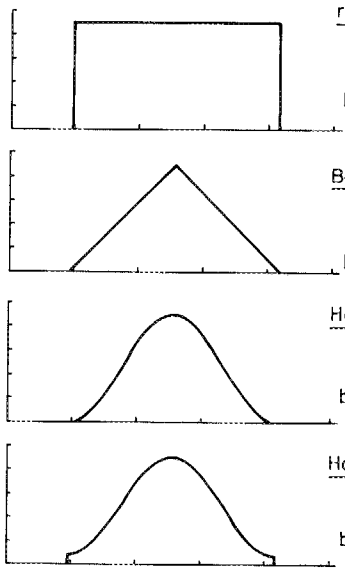

Fig. 7 *Some common window functions that can be used to control variance in periodograms. The annotation adjacent to each trace gives the window function w(m), the bandwidth as a measure of the frequency interval between main lobes of the Fourier transform of the window, and the factor VF, which is a measure of the extent to which we might expect the use of the window to improve the variance of the periodogram*

4.5 *Averaging the periodogram*

The idea behind this method is that the time-domain data recording is broken up into a number of equal length segments and the periodogram of each is calculated. The periodogram of the whole recording is then reckoned to be the average of all the segmental periodograms. The method is sometimes known as the Bartlett procedure after its supposed originator. The N-point data sequence is divided into K segments, each M samples long such that $N = KM$. If the individual periodograms are to be obtained by FFT then M must be an integer power of two and K and N will be constrained accordingly. The jth segment includes the data points $x(n + (j-1)M)$ for $0 \leqslant n \leqslant M-1$. The jth periodogram, normally calculated as the square modulus FFT divided by M, is

$$I_j(\omega) = \frac{1}{M}\left|\sum_{n=0}^{M-1} x(n + (j-1)M)e^{-j\omega n}\right|^2 \tag{16}$$

K successive periodograms are averaged to form the overall estimate of the power spectrum

$$I(\omega) = \frac{1}{K}\sum_{j=1}^{K} I_j(\omega) \tag{17}$$

The method assumes that the segment length M is such that no components exist in the spectrum which are of a low enough frequency to cause data in one segment to be correlated with data in another. This implies that the time-domain autocorvariance function is negligibly small for lags of the order of and greater than M and that successive segmental periodograms are independent of each other. The result of this assumption of independence is that the variance of the averaged periodogram varies inversely as the number of segmental periodograms brought into the average. For normally distributed white spectrum random signals for which the segments are independent as above it can be shown that:

towards zero. While this is a most useful result we must not neglect the fact that control of variance has been achieved at the expense of worsened frequency resolution, as this varies inversely as the recording length which, as we have seen, is reduced by factor K. The minimum resolvable frequency will be of the order of $2\pi/M$, compared with $2\pi/KM$ for the unsegmented recording. We therefore compromise with our choice of K; we first choose segment length M on the basis of our required resolution and then fit in as many averages, K, either as were possible from our initial recording length N or as were found necessary to give adequately controlled variance. This latter can be interpreted as a spectrum which is sufficiently smooth for us to discern whatever features are required for the purposes of any particular experiment. Some experimentation with K and M will generally show the investigator what are the best compromises.

The question of resolution in the frequency domain can be approached more rigorously by considering the frequency-domain window implied by the segmenting procedure. We note that each segment of data in the Bartlett procedure had been extracted by multiplying by a rectangular window in the time domain. We shall see that this implies that each segmented periodogram was in fact the convolution of the 'underlying' function with a frequency-domain window of the form

$$W(\omega) = \left[\frac{\sin \omega M/2}{\sin \omega/2}\right]^2 \tag{19}$$

The Bartlett procedure is the most simple case of the more general method first proposed by WELCH (1967) which combines the averaged periodogram idea with the use of non-rectangular windows for extraction of the segments in the time domain. We consider now some of the features of this method.

As before an N-point data recording is broken up into K segments each of length N. Each is multiplied by a window function, which in the simplest case is rectangular

but which would generally have more gradual transitions at its edges such as one of the cosine tapers. The Fourier transform of an M-point segment takes the form

$$X(\omega) = \sum_{n=0}^{M-1} x(n)w(n)e^{-j\omega n} \tag{20}$$

where $w(n)$ is the time-domain data extraction window; it is zero outside the range $0 \leqslant n \leqslant M-1$. As we would expect $X(\omega)$ is the frequency-domain convolution of the 'underlying' Fourier transform with the Fourier transform of the window

$$X(\omega) = \frac{1}{2\pi} \int_{2\pi} F(\Omega)W(\omega - \Omega)\, d\Omega \tag{21}$$

where

$$W(\omega) = \sum_{n=0}^{M-1} w(n)e^{-j\omega n} \tag{22}$$

and $F(\Omega)$ represents the amplitude spectrum of the population signal. The jth segmental periodogram is calculated by taking the modulus squared of the Fourier transform and dividing by M, as before, but with an additional scaling factor equal to the mean squared value of the time-domain window; this maintains the windowing operation at unity gain. We have

$$I_j(\omega) = \frac{1}{MU} \left| \sum_{n=0}^{M-1} x(n + (j-1)M)w(n + (j-1)M)e^{-j\omega n} \right|^2 \tag{23}$$

or equivalently

$$I_j(\omega) = \frac{1}{MU} |X_j(\omega)|^2 \tag{24}$$

where $X_j(\omega)$ is the Fourier transform of the jth segment.

The scaling factor U is simply

$$U = \frac{1}{M} \sum_{n=0}^{M-1} w^2(n) \tag{25}$$

and the overall spectral estimate is obtained by averaging over K segments as before

$$I(\omega) = \frac{1}{K} \sum_{j=1}^{K} I_j(\omega) \tag{26}$$

The improvement in variance which results from the method is shown by Welch to be approximately factor K, again for idealised random signals which are normally distributed, of white spectrum and within which no significant correlations exist between the segments.

The advantage of the non-rectangular windows of the Welch method is that some extra control of bias and frequency resolution can be had from careful tailoring of the time-domain window shape. Welch shows that the expected value of the averaged periodograms is the convolution of the 'true' power spectrum $P_{xx}(\omega)$ with a frequency-domain window which is the *scaled square modulus* of the Fourier transform of the time-domain window. That is

$$I(\omega) = \frac{1}{2\pi} \int_{-\pi}^{\pi} P_{xx}(\Omega)H(\omega - \Omega)\, d\Omega \tag{27}$$

where

$$H(\omega) = \frac{1}{MU} |W(\omega)|^2 \tag{27a}$$

and U is the scaling factor defined in eqn. 25.

The principal advantage of the Welch method is that the square modulus window $H(\omega)$ cannot take negative values. This prevents an anomalous biasing effect which occurs with windows implied by some other procedures and which sometimes results in negative values appearing at some frequencies in the power spectral estimate. Such a result would of course be meaningless and the spectrum of which it was part would not be usable as evidence of activity in any realisable system.

5 Power spectral estimates involving the autocovariance function

We saw earlier (CHALLIS and KITNEY, 1991) that the FT of the autocovariance function is the power spectrum of a continuous signal. This result is known as the Wiener-Khintchine relationship after its originators. For a discrete signal of unlimited duration a similar relationship applies. The resulting spectrum is continuous and repeats at intervals of the sampling frequency. Just as for discrete signals, we regard the lags in the discrete autocovariance function as being each of unit time and the sampling frequency is consequently 2π rad s^{-1}. We have also seen that the estimation of the discrete autocovariance function on the basis of a data recording of limited length leads to a biased estimate of the function. Correction for this bias by linear scaling leads to a serious increase in variance at the longer lag intervals. The FFT can be used to calculate the estimates of autocovariance and these can be achieved in either a circular or a linear computation.

The question obviously arises as to how the investigator should go about estimation of the power spectrum by means of the discrete autocovariance function and, indeed, should he/she bother at all with such a relatively complicated route. Why not merely use the periodogram calculated directly using the FFT using one of the methods we have already described? The autocovariance function was used as an intermediate step in the estimation of power spectra for many years before the advent of modern electronic computers and efficient algorithms such as the FFT. It was fairly simple to calculate and as the result of the calculation 'grew' lag by lag so the investigator could easily see when further lags would no longer yield a significantly useful contribution to the function. The number of arithmetic operations, and consequently the time taken on them, could be minimised.

The raw biased autocovariance estimate would give a fairly clear indication of the distribution of energy in random signal components and in the truly periodic deterministic components. The approximate bandwidth and power in the random components could be determined by the width and height of the peak at zero lag and a measure of the power in the deterministic components could be had from their amplitude in the *acvf*. The power spectrum could then be formally estimated by taking the cosine transform of the *acvf*. This is of course equivalent to the exponential Fourier transform because the *acvf* is an even function consisting of cosine terms only. If $s_{xx}(k)$ is the autocovariance estimate of the n-point sequence we have

$$s_{xx}(k) = \frac{1}{N} \sum_{n=-(N-1)}^{N-1} [x(n) - \bar{x}][x(n+k) - \bar{x}] \tag{28}$$

The periodogram is then

$$I(\omega) = \frac{1}{\pi} \sum_{k=-(N-1)}^{N-1} s_{xx}(k) \cos \omega k \tag{29}$$

There are circumstances when it is still of use to take a similar route to the periodogram. The most common reason is that one has a very noisy signal which is believed to contain genuinely periodic deterministic components. The autocovariance sequence (generally the linear estimate) is calculated by means of the zero-padded FFT as we have shown earlier (Fig. 8). Various windowing operations can now be carried out, depending on the requirements of the particular experiment at hand. For example some scaling could be carried out to correct for bias (Fig. 8*c*) and the function could then be windowed out at greater lags in order to de-emphasise those parts of it which suffered most from the increase in bias (Fig. 8*d*). Alternatively it may be required to use windows to separate out the deterministic and random components for separate analysis. In Figs. 8*e* and 8*f* this has been done

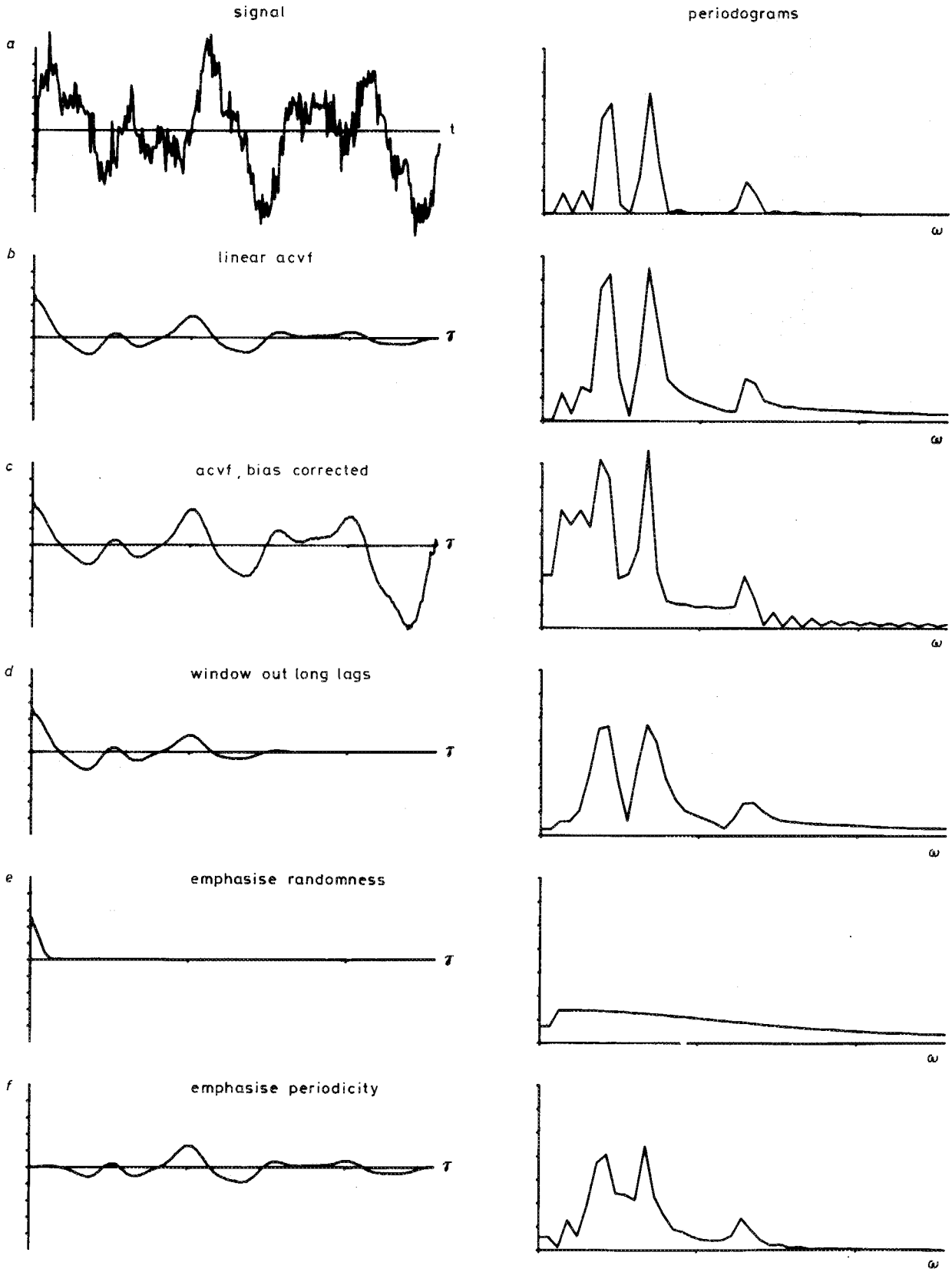

Fig. 8 *Calculation of the periodogram (a) directly by FFT applied to the time-domain signal; (b) from the linear acvf; (c) from the linear acvf corrected for bias; (d) from (c) with the most variant portion of the acvf windowed out; (e) with emphasis on the random components in the signal; (f) with emphasis on the deterministic components*

using a flat-topped cosine taper. The periodograms which correspond to these autocovariance functions are shown on the right-hand side of the figure. For convenience these have been calculated by means of the FFT rather than the cosine transform.

A few words on the detail of this step are appropriate here. The *acvf* is defined as an even function of time and it is centred on the zero lag point. On Fig. 8 only the right-hand halves of the *acvf*s are plotted. Now implicit in the DFT computation is the assumption that the time-domain recording repeats identically in time and the computer memory array containing the *acvf* must be organised to ensure that both positive and negative lag portions are processed. They may be placed anywhere in the array (Fig. 9) and the modulus of each complex output terms of the DFT calculated to give the periodogram. If the input to the DFT is organised so as to place the *acvf* with its positive lag portion at the beginning and its negative lag portion at the end of the array then the real part of the DFT output will provide the periodogram estimates and the imaginary parts should all be zero.

We now consider how the window applied to the *acvf* has effect on the resulting periodogram. We shall restrict our discussion to windows which have their maximum pass at zero lag and are of course symmetrical about zero lag. The result is quite simple; from our previous discussion on time and frequency domain equivalence we see that multiplication by a window function in the lag time domain implies that the resulting periodogram will be convolved with the Fourier transform of whatever window is applied. Note that in this case the frequency-domain window is not the square modulus function as was applicable to windows applied direct to time-domain data. It is therefore important to ensure that the window chosen does not have negative-valued sidelobes because these could impart to the resulting spectral estimate a negative bias sufficient to cause some negative power components.

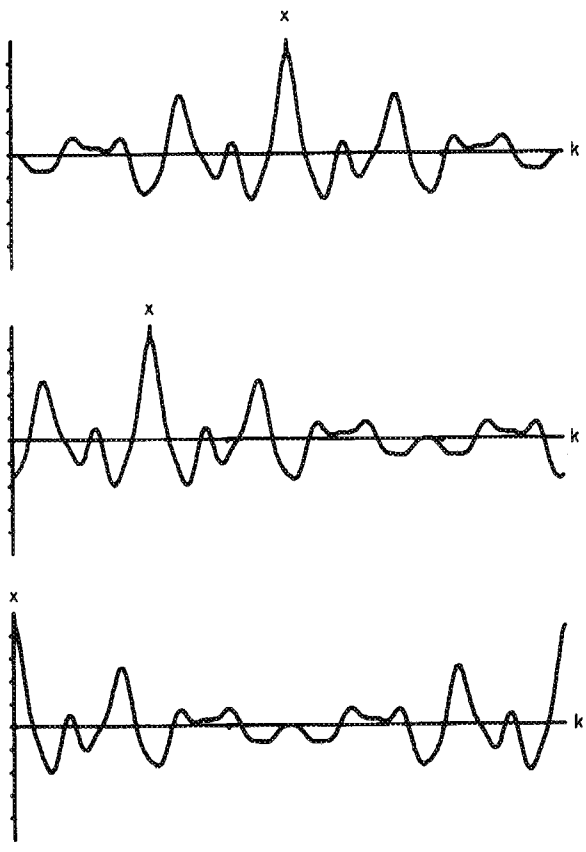

Fig. 9 *Possible positions of the acvf in the computer array before calculations of the FFT*

To conclude this discussion we consider what windows are implied when we do no more than compute the periodogram on the basis of the biased or unbiased *acvf*, with no extra shaping applied to it. We have shown earlier that the *acvf* estimate is a biased version of the underlying population *acvf*, according to

$$s_{xx}(m) = \left[1 - \frac{|m|}{N}\right]\gamma_{xx}(m) \qquad -(N-1) \leqslant m \leqslant N-1 \tag{30}$$

The resulting periodogram is

$$I(k) = \sum_{m=-(N-1)}^{N-1} \left[1 - \frac{|m|}{N}\right]\gamma_{xx}(m)e^{-j\omega m} \tag{31}$$

This is the Fourier transform of a windowed autocovariance sequence. The window is triangular (sometimes called a Bartlett window) and given by

$$\omega(n) = \left[1 - \frac{|m|}{N}\right] \qquad |m| < N$$
$$0 \qquad |m| \geqslant N \tag{32}$$

The convolution in the frequency domain is with the Fourier transform of this window which is given by

$$W(k) = \sum_{m=-(N-1)}^{N-1} \left[1 - \frac{|m|}{N}\right]e^{-j\omega m}$$

$$\therefore \quad W(k) = \frac{1}{N}\left[\frac{\sin\dfrac{\omega N}{2}}{\sin\dfrac{\omega}{2}}\right]^2 \tag{33}$$

We have seen that it is possible to correct for the bias in the autocovariance sequence by linear scaling, although at risk of increased variance at longer lags. The operation is

$$s'_{xx}(m) = \frac{N}{N-|m|}\,s_{xx}(m) \tag{34}$$

The Fourier transform of $s'_{xx}(m)$, while suffering from the increased variance in $s'_{xx}(m)$, is equivalent to the Fourier transform of a sample of the population autocovariance sequence $\gamma_{xx}(m)$, which has been extracted by means of a simple rectangular window, given by

$$w(m) = 1 \qquad |m| < N$$
$$0 \qquad |m| \geqslant N \tag{35}$$

The equivalent frequency-domain window is the Fourier transform of this, given by

$$W(k) = \frac{\sin\left[\dfrac{\omega}{2}(2N-1)\right]}{\sin\left[\dfrac{\omega}{2}\right]} \tag{36}$$

This window can take negative values and can lead to anomalous biasing in the final power spectral estimate. Finally, we note that the window implied by our use of the biased autocovariance sequence is identical in form to the frequency-domain window which would be implied by our calculation of the periodogram by direct Fourier transformation of the time-domain sequence, it being the square modulus of the Fourier transform of the discrete rectangular gate of N points duration. This we would expect because the two methods of arriving at the periodograms are exactly equivalent to each other.

This concludes our discussion of the classical methods of power spectral estimation. We emphasise that all estimates are a compromise between frequency resolution, bias and variance which result both from the statistical properties of the input data and from the methods used to make the estimate. Very many compromises between data length, window procedures and averaging methods are possible and many may yield equivalent results to each other. The best way through this apparent jungle of possibilities is probably to use the simplest methods to gain very approximate estimates of the general properties of the data and then to gradually refine the technique to identify specific features which are believed to be associated with the hypothesis under investigation.

This task is frequently made easier if the spectral analysis procedures are tested on simulated data with a known balance between deterministic and random components. In this connection it is useful to have programmed on one's computer system some basic tools which can be used to generate simulated data; these should include the facility to generate sequences of random data of known spectral properties, source distribution shape and variance. These can be used in conjunction with deterministic function generation routines to generate data which have similar (but better known) properties to those obtained in initial computations on the real test data. The performance of spectral estimation procedures can then be tested rigorously before being applied to experimental data. In this way the analyst will gain greatly increased confidence in his/her estimates of the spectral properties of unknown and frequently variant biological data.

There remains to consider frequency-domain methods by which two signals can be compared by means of their joint spectral power. In the final part of this paper we consider the coherence junction as a means of achieving this comparison.

6 The concept of coherence

In an earlier section of this tutorial (CHALLIS and KITNEY, 1991) we have seen that the cross-power spectrum is the frequency-domain equivalent to the cross-covariance function. Like its time-domain counterpart it is a measure of the similarity between two signals and it expressed this quantitatively as the distribution of joint spectral power in the frequency domain. The coherence function is also a frequency-domain measure of similarity between signals. It differs from the cross-power spectrum in that it is normalised to yield values between zero (signals uncorrelated) and unity (signals identical). The function in its basic form is complex and defined as

$$C_{xy}(\omega) = \frac{G_{xy}(\omega)}{\sqrt{G_{xx}(\omega)G_{yy}(\omega)}} \tag{37}$$

where $G_{xy}(\omega)$ is the cross-power spectrum at frequency ω between signals $x(t)$ and $y(t)$, and $G_{xx}(\omega)$ and $G_{yy}(\omega)$ are the power spectra (also known as autospectra) of $x(t)$ and $y(t)$, respectively. More commonly used is the magnitude-squared coherence (MSC) function, defined as

$$MSC = |C_{xy}(\omega)|^2 = \frac{|G_{xy}(\omega)|^2}{G_{xx}(\omega)G_{xy}(\omega)} \tag{38}$$

If we have an appropriately composite pair of data recordings we could estimate G_{xx}, G_{yy} and G_{xy} by Fourier methods to obtain, at a particular frequency ω

$$G_{xx} = (a + jb) \quad G_{yy} = (c + jd) \quad G_{xy} = (a + jb)(c - jd)$$

If these expressions are substituted into eqn. 38 we arrive at the (at first) confusing result that

$$|C_{xy}(\omega)|^2 = 1 \text{ for all } \omega \tag{39}$$

This is perhaps no surprise when we consider that the cross-spectrum is merely a complex product of the two constituent spectra, and that the coherence function is normalised by placing a not dissimilar product in the denominator. Why then are coherence functions of use to the analyst? The answer is that calculation of coherence by means of averaged estimates based on segments of the original signals indeed yields meaningful quantitative measures of similarity between them, and these measures take values between zero and unity.

In a frequency-domain measure of similarity between two signals we are interested in more than merely the joint presence of energy at the same frequency in the two recordings. This would of course be available from simple cross-spectral estimates. The signal content of two recordings may change considerably with time. Particular frequency components may wax and wane in amplitude and the relative phase between components in the two channels may alter with time. Noise may occur sporadically or continuously in one or both channels and new frequency components, not regarded as noise, may arrive by additive processes into the signals or as a result of nonlinear changes in channel properties. The coherence function provides an analytical tool by which all of these changes can be monitored and quantified in the frequency domain. Attention is more generally focused on the MSC function and we consider its calculation here.

The two data recordings which are to be compared are broken up into N segments of equal length. The auto- and cross-spectra are estimated on the basis of averages drawn from the individual spectra of the segments. Typically, the individual spectra are obtained by means of the FFT. If $X_n(\omega)$ and $Y_n(\omega)$ are the complex FFTs of the nth segment, then

$$\begin{aligned} G_{xx}(\omega) &= \frac{1}{N}\sum_{n=1}^{N}|X_n(\omega)|^2 \\ G_{yy}(\omega) &= \frac{1}{N}\sum_{n=1}^{N}|Y_n(\omega)|^2 \\ G_{xy}(\omega) &= \frac{1}{N}\sum_{n=1}^{N}X_n(\omega)Y_n^*(\omega) \end{aligned} \tag{40}$$

The segments may overlap (CARTER *et al.*, 1973) to a degree and the choice of segment length and number of segments depends on the statistical constraints which characterise the data and the analysis required of it. We have seen that estimates of the power spectrum made on the basis of the periodogram are inconsistent in that the bias and variance of the estimate do not necessarily reduce as longer data lengths are analysed. These uncertainties can be reduced by estimating the periodogram using an average of a larger number of short data segments. However, the attainment of good spectral resolution in an FFT operation requires fairly long data lengths as the spacing between spectral lines is inversely proportional to the (assumed) time-domain repetition period for the data. There is thus a compromise required in the choice of segment length: it must be long enough to provide adequate frequency resolution and short enough to allow sufficient segments in the average to control bias and variance. In addition each segment must be extracted using a time-domain window to control sidelobes in the

frequency-domain functions and thereby limit the bias which arises due to the resulting leakage.

The conflicting requirements of number of segments and segment length can be alleviated to some extent by overlapping segments as this permits more averages to be taken while maintaining the same frequency resolution. This is only effective up to a point because the improvement declines due to greater correlation between segments. Computational overheads increase with increased overlap and tradeoff exists between this overhead and the improvements obtainable. CARTER *et al.* (1973) have shown theoretically and experimentally that for Hanning windowed segments there is little point in employing overlap beyond about 50 per cent.

The statistical properties of the MSC estimator have been described by CARTER *et al.* (1973), who give expressions for bias and variance that result from the use of the segmented estimation procedure. Two approximations useful for a number of segments greater than 32 are reproduced here; they assume an ideal Hanning window applied to each segment and no overlap between segments. Both signals are assumed to be zero-mean, stationary Gaussian processes. If the MSC is γ then the bias and variance are

$$\text{bias} = B(|\gamma|^2) \approx (1 - |\gamma|^2)^2/N$$

$$\text{var} = V(|\gamma|^2) \approx \begin{cases} 1/N^2 & |\gamma|^2 = 0 \\ 2|\gamma|^2(1 - |\gamma|^2)/N & 0 < |\gamma^2| \leqslant 1 \end{cases} \quad (41)$$

These show clearly that the bias and variance of the estimator are dependent on both the number of segments N as well as on the true MSC value. Computed values are listed in Table 1.

These results have been obtained for idealised data averaged using a simple overlapping protocol. In many real situations the methods used to average the constituent spectra are more complicated than those given in the above equations and those used by Carter *et al.* It is not always possible to arrive analytically at measures of bias and variance and frequently empirical methods are employed to obtain such estimates. In the context of this introductory paper it is more appropriate to consider the basic features of a few averaging processes. Let the nth periodogram be

$$X_n(\omega) = a_n(\omega) + jb_n(\omega)$$

The autospectrum is built up using the following average:

$$\begin{aligned} N\hat{G}_{xx}(\omega) &= \sum_{n=1}^{N} |X_n(\omega)|^2 \\ &= a_1^2 + b_1^2 \\ &+ a_2^2 + b_2^2 \\ &\vdots \\ &a_N^2 + b_N^2 \\ &\overline{\qquad\qquad} \\ &A_N^2 + B_N^2 \end{aligned} \quad (42)$$

This method of addition is equivalent to the time-domain operation of averaging the *acvfs* of the N segments. The reader is warned that an average taken on the basis of the separate real and imaginary components of each segmental spectrum will have a totally different interpretation. The operation consists of the following summation:

$$\begin{aligned} &a_1 + jb_1 \\ &+a_2 + jb_2 \\ &\vdots \quad \vdots \\ &a_N + jb_N \\ &\overline{\qquad\qquad} \\ &A'_N + jB'_N \end{aligned} \quad (43)$$

Then $G'_{xx}(\omega) = (A'_N)^2 + (B'_N)^2$

The operation preserves phase in the individual segments and is equivalent to the computation of the periodogram of the time-domain average of the segment recordings. It causes much of the signal energy to be lost as most segments will be incoherent when lined up one below the other.

A not unrelated argument applies to the estimation of the cross-spectrum averaged over all of the response epochs. The estimate could be made in two ways. If the individual FFTs of a pair of segments are $a_n(\omega) + jb_n(\omega)$ and $c_n(\omega) + jd_n(\omega)$ then the cross-spectrum for that segment pair is $G_{xyn}(\omega) = (a_n c_n + b_n d_n) + j(b_n c_n - a_n d_n) = g_n + jh_n$ say. This spectrum is the frequency-domain equivalent of the circular cross-covariance function for the two segments. It carries in its phase an imprint of the relative lag between the two segments just as in the time domain the *ccvf* preserves such lag information. If we form the averaged cross-spectrum thus:

$$\begin{aligned} N\hat{b}_{xyN} &= g_1 + jh_1 \\ &+g_2 + jh_2 \\ &\vdots \\ &g_N + jh_N \\ &\overline{\qquad\qquad} \\ &G_N + jH_N \end{aligned} \quad (44)$$

Table 1 Computed values of (a) bias and (b) variance for the magnitude-squared coherence function for a range of coherence values and three different ensemble sizes $N = 32$, 64 and 128. Computations are based on the approximate expressions given in the text

(a) Bias

$\|\gamma\|^2$	32	64	128
0·0	0·0313	0·0156	0·0078
0·1	0·0253	0·0127	0·0063
0·2	0·0200	0·0100	0·0050
0·3	0·0153	0·0077	0·0038
0·4	0·0112	0·0056	0·0028
0·5	0·0078	0·0039	0·0020
0·6	0·0050	0·0025	0·0012
0·7	0·0028	0·0014	0·0007
0·8	0·0012	0·0006	0·0003
0·9	0·0003	0·0002	0·0001
1·0	0·0000	0·0000	0·0000

(b) Variance

$\|\gamma\|^2$	32	64	128
0·0	0·0010	0·0002	0·0001
0·1	0·0051	0·0025	0·0013
0·2	0·0080	0·0040	0·0020
0·3	0·0092	0·0046	0·0023
0·4	0·0090	0·0045	0·0022
0·5	0·0078	0·0039	0·0020
0·6	0·0060	0·0030	0·0015
0·7	0·0039	0·0020	0·0010
0·8	0·0020	0·0010	0·0005
0·9	0·0006	0·0003	0·0001
1·0	0·0000	0·0000	0·0000

we have the frequency-domain equivalent of adding together all of the individual time domain *ccvfs*. Clearly if the phase relationships between the two component signals in successive pairs of recording segments are not consistent from one segment to the next then addition in this way will lead to a reduction in the total value of the sum and hence a reduced magnitude for the coherence function. This makes good sense as inconsistent time delays in successive segments will imply dissimilarity between the two signals. It is of course still possible to adopt a restrictive approach and to reduce the effect of variable delay as a contributor to coherence and this would imply addition of the following form:

$$\begin{array}{r} g_1^2 + h_1^2 \\ +g_2^2 + h_2^2 \\ \vdots \\ g_N^2 + h_N^2 \\ \hline G_N^2 + H_N^2 \end{array}$$

and

$$N\hat{G}_{xyn} = G_N^2 + H_N^2 \qquad (45)$$

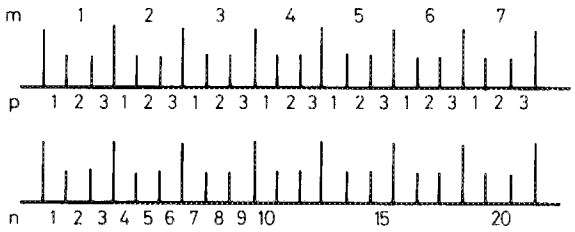

Fig. 10 *An averaging scheme for the MSC function. The data are broken up into 21 segments and the autospectra (see eqn. 46) are formed on the basis of a phase-free average of all 21 subrecordings. The cross-spectra are obtained from a two-stage process. Groups of three segmental cross-spectra are averaged taking phase into account. The seven-part averages are then themselves averaged with phase discarded*

$M = 7 \quad P = 3 \quad N = 21$

The operation would yield a coherence estimate which is sensitive to differing spectral content due to additive noise or nonlinear effects but which is comparatively insensitive to relative phase fluctuations. In some instances relative phase between segments and between spectral components may change rapidly in a manner which masks much similarity between the two signals under examination. In this case it may be sensible to perform averaging according to eqn. 45 over a few segments and then to proceed with further averaging over groups of segments according to eqn. 44. Indeed considerable flexibility is available to the analyst in the manner in which the averages are chosen (TICK, 1967) and very frequently the benefits of experience are required to design an adequate match between the averaging protocol and the statistical properties of the data on hand.

As an example of the averaging complexity which is possible Fig. 10 shows diagramatically a pair of recordings broken up into 21 segment pairs. These have been further subdivided into groups of three pairs. The cross-power spectrum would for example be estimated on the basis of a two-stage averaging process, the first of which included phase changes over the restricted range of the three segments. The second stage would then form an average of the seven results of the first stage averaging and would not include phase over the longer intervals. If $a_n + jb_n$ and $c_n + jd_n$ represent the FFTs of the nth segments in each of the two signals the averaging scheme becomes

$$C_{xy}(\omega) = \frac{1/M \sum_{m=1}^{M}([1/P \sum_{n=P_m-M+1}^{P_m} (a_n c_n + b_n d_n)]^2 + [1/P \sum_{n=P_m-M+1}^{P_m} (a_n d_n - b_n c_n)]^2)}{[1/N \sum_{n=1}^{N} (a_n^2 + b_n^2)][1/N \sum_{n=1}^{N} (c_n^2 + d_n^2)]} \qquad (46)$$

where $M = 7$, $N = 21$ and $P = 3$. The process could under certain requirements be further complicated by forming the averaged MSC of a set of averages formed in the manner shown above. There is a broad range of choice and the ideas we have presented here really must only touch the surface of this complex subject. We would suggest that if the reader intends to make serious use of coherence calculations a comprehensive study of relevant literature is undertaken as a preliminary measure. We

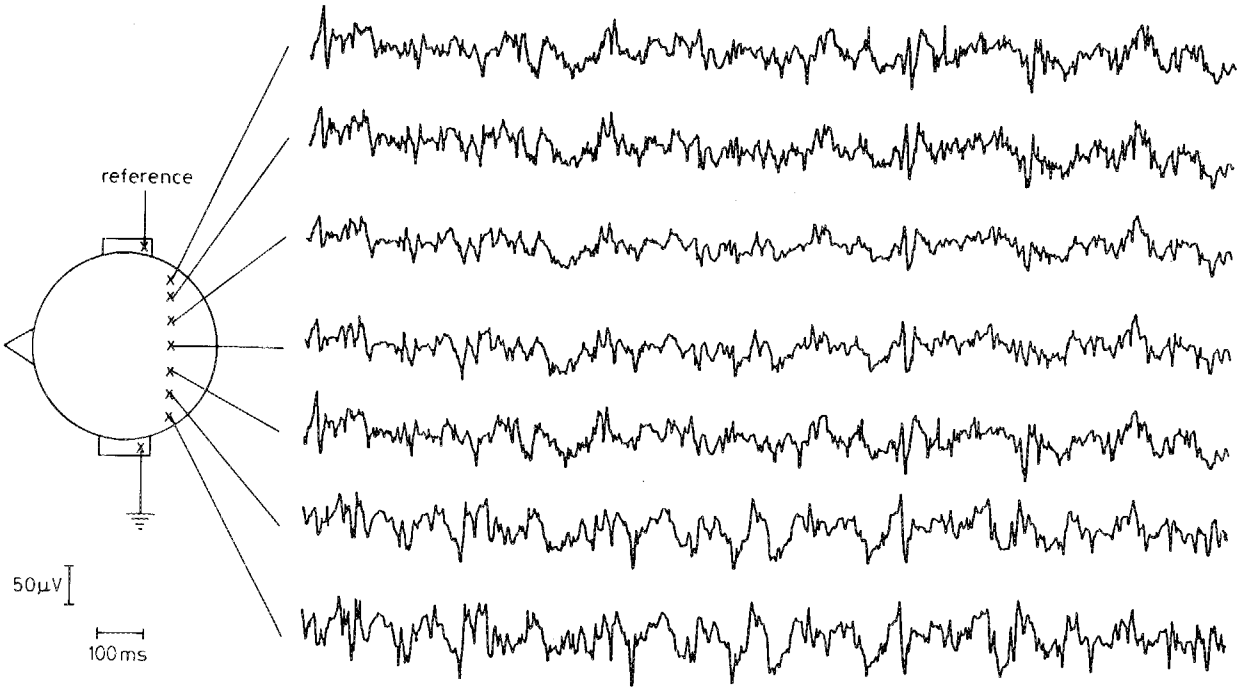

Fig. 11 *Diagram of an array of electrodes placed on the human scalp for recording of visually evoked potentials, together with typical responses*

include here a Bibliography which will serve as useful background material; they include the original work of WIENER (1930) on the subject in 1930 and an excellent review by TICK (1967).

No discussion of a technique is complete without an example of its use. Fig. 11 shows a diagram of an array of electrodes place on the human scalp in an experiment designed to record visually evoked potentials (VEPs). The problem was to establish the extent to which background (i.e. non-evoked) activity was consistently similar between the several electrode channels. 128 segments of EEG sampled at 128 Hz and each of length 1024 points were

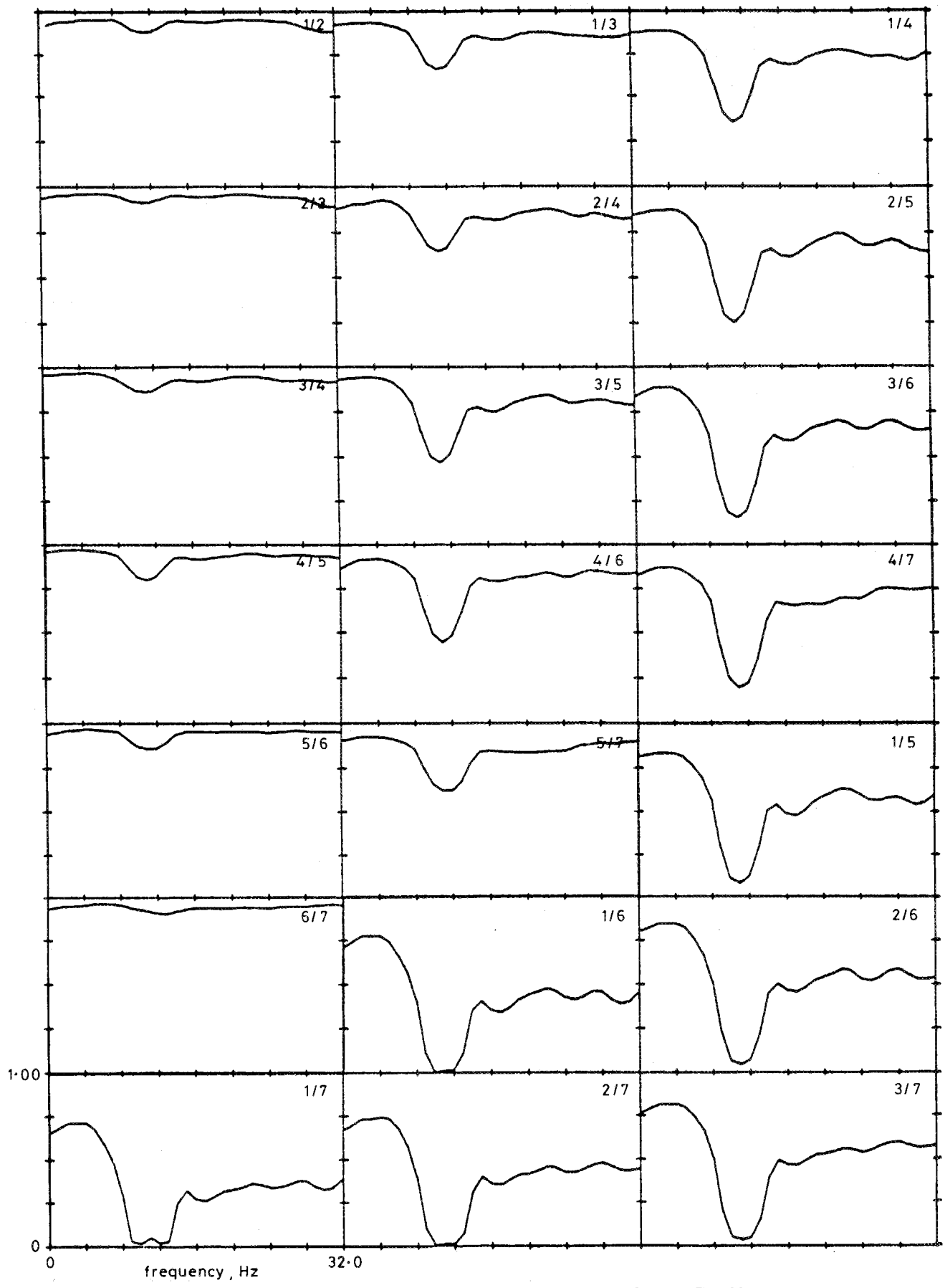

Fig. 12 *Magnitude squared coherence functions between pairs of electrodes in the montage shown in Fig. 11*

used to estimate the MSC between pairs of electrode channels. The results are shown in Fig. 12, and we see that there is considerable variability in MSC across the available frequency band and between different electrode pairs. A similar exercise was carried out for the evoked potentials themselves and from the results of both studies it was possible to choose pairs of electrodes in which *either* the EPs were coherent and the background was not *or* the background activity was coherent in both channels and the EPs were not. Signals from appropriate channel pairs were then combined in an adaptive noise-cancelling algorithm designed to enhance recognition of VEPs without resort to the coherent average (WINSKI, 1986).

7 Conclusion

This paper concludes our series of three tutorials on what are termed the classical methods of signal analysis. We have seen how signal recordings can be regarded as samples of the activity of some underlying process and how they can be considered as combinations of random and deterministic components. Their properties can be estimated and they can be compared to each other on the basis of similarity of their individual properties or on the basis of joint properties such as cross-covariance or coherence functions. Description and comparison of signals can be carried out in the time domain, or the frequency domain, or both domains. Many operations in one domain (time or frequency) are equivalent to other operations in the other domain (i.e. frequency or time). All measures of signal properties are regarded as estimates of some underlying process; these estimates must be based on compromises between potential errors and uncertainties. This idea has been considered in detail for the case of the estimation of the power spectrum of a signal from a number of sample recordings of the signal. We have shown that a number of techniques are available which can be used to control the compromise between frequency-domain resolution and statistical bias on individual spectrum components on the one hand, and increased uncertainty due to estimator variance on the other. Both the properties of the source data *and* the estimation technique employed affect the eventual compromise that will be made. These ideas were developed to include frequency-domain signal comparison by means of coherence function analysis.

It has been found that in some cases it is not possible to obtain an acceptable compromise between, for instance, frequency resolution, bias and variance, and new techniques are required. These generally fall under the title 'modern spectrum analysis' and will form the subject matter for our next and final paper in this series.

References

CARTER, G. C., KNAPP, C. H. and NUTTALL, A. H. (1973) Estimation of the magnitude-squared coherence function via overlapped fast Fourier transform processing. *IEEE Trans.*, **AU-21,** 337–344.

CHALLIS, R. E. and KITNEY, R. I. (1982) The design of digital filters for biomedical signal processing (in three parts). Part 1. *J. Biomed. Eng.*, **4,** 267–278.

CHALLIS, R. E. and KITNEY, R. I. (1983*a*) The design of digital filters for biomedical signal processing (in three parts). Part 2. *Ibid.*, **5,** 19–30.

CHALLIS, R. E. and KITNEY, R. I. (1983*b*) The design of digital filters for biomedical signal processing (in three parts). Part 3. *Ibid.*, **5,** 91–102.

CHALLIS, R. E. and KITNEY, R. I. (1990) Biomedical signal processing (in four parts). Part 1 Time-domain methods. *Med. & Biol. Eng. & Comput.*, **28,** 509–524.

CHALLIS, R. E. and KITNEY, R. I. (1991) Biomedical signal processing (in four parts). Part 2. The frequency transforms and their inter-relationships. *Ibid.*, **29,** 1–17.

CHATFIELD, C. (1984) *The analysis of time series: an introduction,* 3rd edn. Chapman & Hall, London.

ELLIOTT, D. F. and RAO, K. R. (1982) *Fast transforms, algorithms, analyses, applications.* Academic Press, New York.

JENKINS, G. M. and WATTS, D. G. (1968) *Spectral analysis and its applications.* Holden-Day Inc., San Francisco.

NUTTALL, A. H. and CARTER, G. C. (1981) An approximation to the cumulative distribution function of the magnitude-squared coherence estimate. *IEEE Trans.*, **ASSP-29,** 932–934.

OPPENHEIM, A. V. and SCHAFER, R. W. (1975) *Digital signal processing.* Prentice Hall, New Jersey.

TICK, L. J. (1967) Estimation of coherency. In *Spectral analysis of time series.* HARRIS, B. (Ed.), John Wiley, New York.

WELCH, P. D. (1967) The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. *IEEE Trans.*, **AU-15,** 70–73.

WINSKI, R. W. (1986) Adaptive techniques for signal enhancement in the human electroencephalogram. Ph.D. Thesis, University of Keele, UK.

Bibliography

BENIGNUS, V. A. (1969*a*) Estimation of the coherence spectrum and its confidence interval using the fast Fourier transform. *IEEE Trans.*, **AU-17,** 145–150.

BENIGNUS, V. A. (1969*b*) Estimation of coherence spectrum of non-Gaussian time series populations. *Ibid.*, **AU-17,** 198–201.

CARTER, G. C., KNAPP, C. H. and NUTTALL, A. H. (1973) Estimation of the magnitude-squared coherence function via overlapped fast Fourier transform processing. *Ibid.*, **AU-21,** 337–344.

CARTER, G. C. (1980) Bias in magnitude-squared coherence estimation (correspondence). *Ibid.*, **ASSP-28,** 97–98.

JONES, A. G. (1981) Transformed coherence functions of multivariate studies. *Ibid.*, **ASSP-29,** 317–319.

NUTTALL, A. H. and CARTER, G. C. (1976) Bias of the estimate of magnitude-squared coherence. *Ibid.*, **ASSP-24,** 582–583.

NUTTALL, A. H. and CARTER, G. C. (1981) An approximation to the cumulative distribution function of the magnitude-squared coherence estimate. *Ibid.*, **ASSP-29,** 932–934.

PISANENKO, V. F. (1973) The retrieval of harmonics from a covariance function. *Geophys. J.R. Ast. Soc.*, **33,** 347–366.

STEARNS, S. D. (1981) Tests of coherence unbiasing methods. *IEEE Trans.*, **ASSP-29,** 321–323.

TICK, L. J. (1967) Estimation of coherency. In *Spectral analysis of time series.* HARRIS, B. (Ed.), John Wiley, New York.

WIENER, N. (1930) Generalised harmonic analysis. *Acta Math.*, **55,** 117–258.

WINSKI, R. W. (1986) Adaptive techniques for signal enhancement in the human electroencephalogram. Ph.D. Thesis, University of Keele, UK.

Authors' biographies

Richard E. Challis is Professor of Industrial Applications of Signal Processing in the Electronic Engineering Group at the University of Keele, UK. He is a graduate of Imperial College, London, and carried out research for his Ph.D. there. His career as an engineer includes posts in engineering industries, the medical service and universities in the UK and overseas. At present he directs the Ultrasonics & Digital Signal Processing Laboratory at Keele and is involved in a number of research programmes in the biomedical area, and in industrial instrumentation.

Richard I. Kitney is currently Professor of Biomedical Systems in the Electrical Engineering Department of Imperial College, London, UK. After obtaining his Ph.D. at Imperial College he held a number of academic posts at Chelsea College and at Imperial College.