# STATISTICAL FOURIER ANALYSIS:
# CLARIFICATIONS AND INTERPRETATIONS

by D.S.G. Pollock (University of Leicester)
*Email:* stephen_pollock@sigmapi.u-net.com

This paper expounds some of the results of Fourier theory that are essential to the statistical analysis of time series. It employs the algebra of circulant matrices to expose the structure of the discrete Fourier transform and to elucidate the filtering operations that may be applied to finite data sequences.

An ideal filter with a gain of unity throughout the pass band and a gain of zero throughout the stop band is commonly regarded as incapable of being realised in finite samples. It is shown here that, to the contrary, such a filter can be realised both in the time domain and in the frequency domain.

The algebra of circulant matrices is also helpful in revealing the nature of statistical processes that are band limited in the frequency domain. In order to apply the conventional techniques of autoregressive moving-average modelling, the data generated by such processes must be subjected to anti-aliasing filtering and sub sampling. These techniques are also described.

It is argued that band-limited processes are more prevalent in statistical and econometric time series than is commonly recognised.

## 1. Introduction

Statistical Fourier analysis is an important part of modern time-series analysis, yet it frequently poses an impediment that prevents a full understanding of temporal stochastic processes and of the manipulations to which their data are amenable. This paper provides a survey of the theory that is not overburdened by inessential complications, and it addresses some enduring misapprehensions.

Amongst these is a misunderstanding of the effects in the frequency domain of linear filtering operations. It is commonly maintained, for example, that an ideal frequency-selective filter that preserves some of the elements of a time series whilst nullifying all others is incapable of being realised in finite samples.

This paper shows that such finite-sample filters are readily available and that they possess time-domain and frequency-domain representations that are both tractable. Their representations are directly related to the classical Wiener–Kolmogorov theory of filtering, which presupposes that the data can be treated as if they form a doubly-infinite or a semi-infinite sequence.

A related issue, which the paper aims to clarify, concerns the nature of continuous stochastic processes that are band-limited in frequency. The paper provides a model of such processes that challenges the common supposition that the natural *primum mobile* of any continuous-time stochastic process is Wiener process, which consists of a steam of infinitesimal impulses.

The sampling of a Wiener process is inevitably accompanied by an aliasing effect, whereby elements of high frequencies are confounded with elements of frequencies that fall within the range that is observable in sampling. Here, it is shown, to the contrary, that there exist simple models of continuous band-limited processes for which no aliasing need occur in sampling.

The spectral theory of time series is a case of a non-canonical Fourier theory. It has some peculiarities that originally caused considerable analytic difficulties, which were first overcome in a satisfactory manner by Norbert Wiener (1930) via his theory of a generalised harmonic analysis. The statistical filtering theory that is appropriate to such time series originated with Wiener (1941) and Kolmogorov (1941a).

Many of the difficulties of the spectral representation of a time series can be avoided by concentrating exclusively on its correlation structure. Then, one can exploit a simple canonical Fourier relationship that exists between the autocovariance function of a discrete stationary process and its spectral density function.

This relationship is by virtue of an inverted form of the classical Fourier series relationship that exists between a continuous, or piecewise continuous, periodic function and its transform, which is a sequence of Fourier coefficients. (In the inverted form of the relationship, which is described as the discrete-time Fourier transform, it is the sequence that is the primary function and the continuous periodic function that is its transform.) However, it is important to have a mathematical model of the process itself, and this is where some of the complications arise.

Figure 1 depicts what may be described as the canonical Fourier transforms. For conciseness, they have been expressed using complex exponentials,

whereas they might be more intelligible when expressed in terms of ordinary trigonometrical functions. When dealing with the latter, only positive frequencies are considered, whereas complex exponentials are defined for both positive and negative frequencies.

In the diagrams, the real-valued time-domain functions are symmetric, with the consequence that their frequency-domain transforms are also real-valued and symmetric. Under other conditions, the transforms would be complex-valued, giving rise to a pair of functions.

## 2. Canonical Fourier Analysis

The classical Fourier series, illustrated in Figure 1.(ii), expresses a continuous, or piecewise continuous, period function $x(t) = x(t+T)$ in terms of a sum of sine and cosine functions of harmonically increasing frequencies $\{\omega_j = 2\pi j/T; j = 1, 2, 3, \ldots\}$:

$$
\begin{aligned}
x(t) &= \alpha_0 + \sum_{j=1}^{\infty} \alpha_j \cos(\omega_j t) + \sum_{j=1}^{\infty} \beta_j \sin(\omega_j t) \\
&= \alpha_0 + \sum_{j=1}^{\infty} \rho_j \cos(\omega_j t - \theta_j).
\end{aligned}
\tag{1}
$$

Here, $\omega_1 = 2\pi/T$ is the fundamental frequency or angular velocity, which corresponds to a trigonometric function that completes a single cycle in the interval $[0, T]$, or in any other interval of length $T$ such as $[-T/2, T/2]$. The second expression depends upon the definitions

$$
\rho_j^2 = \alpha_j^2 + \beta_j^2 \quad \text{and} \quad \theta_j = \tan^{-1}(\beta_j/\alpha_j).
\tag{2}
$$

The equality follows from the trigonometrical identity

$$
\cos(A - B) = \cos(A)\cos(B) + \sin(A)\sin(B).
\tag{3}
$$

In representing the periodic nature of $x(t)$, it is helpful to consider mapping the interval $[0, T]$, over which the function is defined, onto the circumference of a circle, such that the end points of the interval coincide. Then, successive laps or circuits of the circle will generate successive cycles of the function.

According to Euler's equations, there are

$$
\cos(\omega_j t) = \frac{1}{2}(e^{i\omega_j t} + e^{-i\omega_j t}) \quad \text{and} \quad \sin(\omega_j t) = \frac{-i}{2}(e^{i\omega_j t} - e^{-i\omega j t}).
\tag{4}
$$

Therefore, equation (1) can be expressed as

$$
x(t) = \alpha_0 + \sum_{j=1}^{\infty} \frac{\alpha_j + i\beta_j}{2} e^{-i\omega_j t} + \sum_{j=1}^{\infty} \frac{\alpha_j - i\beta_j}{2} e^{i\omega_j t},
\tag{5}
$$

which can be written concisely as

$$
x(t) = \sum_{j=-\infty}^{\infty} \xi_j e^{i\omega_j t},
\tag{6}
$$

**(i) The Fourier integral:**

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \xi(\omega)e^{i\omega t}d\omega \quad \longleftrightarrow \quad \xi(\omega) = \int_{-\infty}^{\infty} x(t)e^{-i\omega t}dt$$

**(ii) The classical Fourier series:**

$$x(t) = \sum_{j=-\infty}^{\infty} \xi_j e^{i\omega_j t} \quad \longleftrightarrow \quad \xi_j = \frac{1}{T} \int_{0}^{T} x(t)e^{-i\omega_j t}dt$$

**(iii) The discrete-time Fourier transform:**

$$x_t = \frac{1}{2\pi} \int_{-\pi}^{\pi} \xi(\omega)e^{i\omega t}d\omega \quad \longleftrightarrow \quad \xi(\omega) = \sum_{t=-\infty}^{\infty} x_t e^{-i\omega t}$$

**(iv) The discrete Fourier transform:**

$$x_t = \sum_{j=0}^{T-1} \xi_j e^{i\omega_j t} \quad \longleftrightarrow \quad \xi_j = \frac{1}{T} \sum_{t=0}^{T-1} x_t e^{-i\omega_j t}$$

**Figure 1.** The classes of the Fourier transforms.

4

where

$$\xi_0 = \alpha_0, \quad \xi_j = \frac{\alpha_j - \mathrm{i}\beta_j}{2} \quad \text{and} \quad \xi_{-j} = \xi_j^* = \frac{\alpha_j + \mathrm{i}\beta_j}{2}. \tag{7}$$

The inverse of the classical transform of (6) is demonstrated by writing

$$\xi_j = \frac{1}{T} \int_0^T x(t) e^{-\mathrm{i}\omega_j t} dt = \frac{1}{T} \int_0^T \left\{ \sum_{k=-\infty}^{\infty} \xi_k e^{\mathrm{i}\omega_k t} \right\} e^{-\mathrm{i}\omega_j t} dt$$

$$= \frac{1}{T} \sum_{k=-\infty}^{\infty} \xi_k \left\{ \int_0^T e^{\mathrm{i}(\omega_k - \omega_j)t} dt \right\} = \xi_j, \tag{8}$$

where the final equality follows from an orthogonality condition in the form of

$$\int_0^T e^{\mathrm{i}(\omega_k - \omega_j)t} dt = \begin{cases} 0, & \text{if } j \neq k; \\ T, & \text{if } j = k. \end{cases} \tag{9}$$

The relationship between the continuous periodic function and its Fourier transform can be summarised by writing

$$x(t) = \sum_{j=-\infty}^{\infty} \xi_j e^{\mathrm{i}\omega_j t} \quad \longleftrightarrow \quad \xi_j = \frac{1}{T} \int_0^T x(t) e^{-\mathrm{i}\omega_j t} dt. \tag{10}$$

The question of the conditions that are sufficient for the existence of such a Fourier relationship is an essential one; and there are a variety of Fourier theories. The theory of Fourier series is concerned with establishing the conditions under which the partial sums converge to the function, in some specified sense, as the number of the terms increases.

At the simplest level, it is sufficient for convergence that $x(t)$ should be continuous and bounded in the interval $[0, T]$. However, the classical theory of Fourier series is concerned with the existence of the relationship in the case where $x(t)$ is bounded but is also permitted to have a finite number of maxima and minima and a finite number of jump discontinuities. It can be shown that, in that case, as successive terms are added, the Fourier series of (1) converges to

$$\frac{1}{2}\{x(t+0) + x(t-0)\}, \tag{11}$$

where $x(t+0)$ is the value as $t$ is approached from the right and $x(t-0)$ is the value as it is approached from the left. If $x(t)$ is continuous at the point in question, then the Fourier series converges to $x(t)$.

Here, the relevant criterion is that of overall convergence in mean-square rather than of pointwise convergence. The question of the mode of converge was the issue at stake in the paradox known as Gibbs' phenomenon, which concerns the absence of pointwise convergence in the presence jump discontinuities.

This phenomenon is illustrated in Figure 2, where it is apparent that not all of the oscillations in the partial sums that approximate a periodic square wave (i.e. a time-domain rectangle) are decreasing at a uniform rate as the
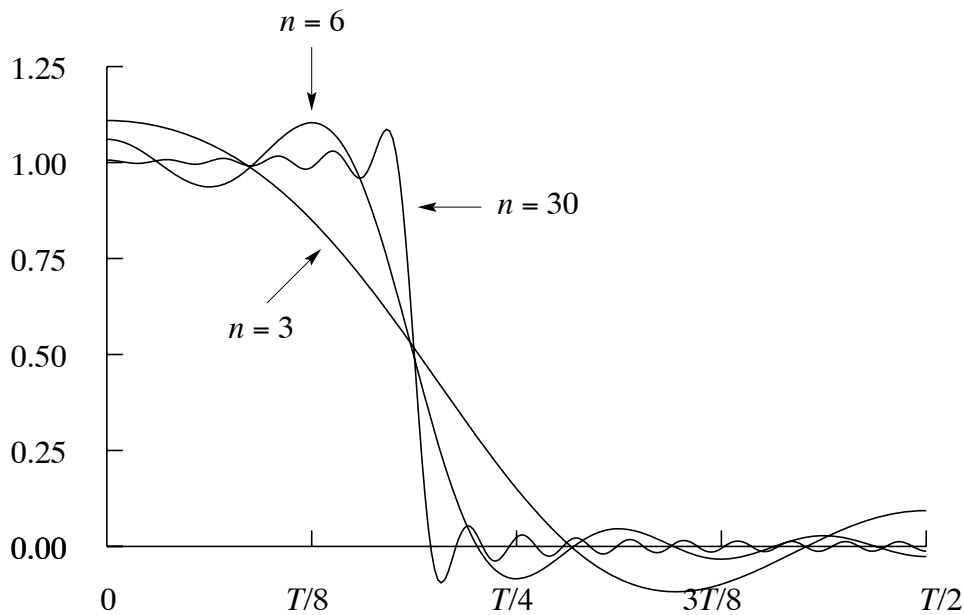
**Figure 2.** The Fourier approximation of a square wave. The diagram shows approximations to the positive half of a rectangle defined on the interval $[-T/2, T/2]$ and symmetric with respect to the vertical axis through zero.

number of terms $n$ increases. Instead, the oscillations that are adjacent to the point of discontinuity are tending to a limiting amplitude, which is about 9% of the jump. However, as $n$ increases, the width of these end-oscillations becomes vanishingly small; and thus the mean-square convergence of the Fourier series is assured. Gibbs' phenomenon is analysed in detail by Carslaw (1930); and Carslaw (1925) has also recounted the history of its discovery.

Parseval's theorem asserts that, under the stated conditions, which guarantee mean-square convergence, there is

$$\frac{1}{T} \int_0^T |x(t)|^2 dt = \alpha_0^2 + \frac{1}{4} \sum_{j=1}^{\infty} (\alpha_j^2 + \beta_j^2) = \sum_{j=-\infty}^{\infty} |\xi_j|^2, \tag{12}$$

where $|\xi_j|^2 = \xi_j \xi_j^* = \rho_j^2/4$. This indicates that the energy of the function $x(t)$, which becomes its average power if we divide by $T$, is equal to the sum of the energies of the sinusoidal components. Here, it is essential that the square integral converges, which is to say that the function possesses a finite energy. The fulfilment of such an energy or power condition characterises what we have chosen to describe as the canonical Fourier transforms.

There are some powerful symmetries between the two domains of the Fourier transforms. The discrete-time Fourier transform, which is fundamental to time-series analysis, is obtained by interchanging the two domains of the classical Fourier series transform. It effects the transformation of a sequence $x(t) = \{x_t; t = 0, \pm 1, \pm 2, \ldots\}$ within the time domain, that is square summable,

6

into a continuous periodic function in the frequency domain via an integral expression that is mean-square convergent. It is illustrated in Figure 1.(iii). We may express the relationship in question by writing

$$x_t = \frac{1}{2\pi} \int_{-\pi}^{\pi} \xi(\omega) e^{i\omega t} d\omega \quad \longleftrightarrow \quad \xi(\omega) = \sum_{t=-\infty}^{\infty} x_t e^{-i\omega t}. \qquad (13)$$

The periodicity is now in the frequency domain such that $\xi(\omega) = \xi(\omega + 2\pi)$ for all $\omega$. The periodic function completes a single cycle in any interval of length $2\pi$; and, sometimes, it may be appropriate to define the function over the interval $[0, 2\pi]$ instead of the interval $[-\pi, \pi]$. In that case, the relationship of (13) will be unaffected; and its correspondence with (10), which defines the classical Fourier series, is clarified.

There remain two other Fourier transforms, which embody a complete symmetry between the two domains. The first of these is the Fourier integral transform of Figure 1.(i) which is denoted by

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \xi(\omega) e^{i\omega t} d\omega \quad \longleftrightarrow \quad \xi(\omega) = \int_{-\infty}^{\infty} x(t) e^{-i\omega t} dt. \qquad (14)$$

Here, apart from the scaling, which can be revised to achieve symmetry, there is symmetry between the transform and its inverse. (The factor $1/2\pi$ can be eliminated from the first integral by changing the variable from angular velocity $\omega$, measured in radians per period, to frequency $f = \omega/2\pi$, measured in cycles per period.)

One of the surprises on first encountering the Fourier integral is the aperiodic nature of the transform, which is in spite of the cyclical nature of the constituent sinusoidal elements. However, this is readily explained. Consider the sum of two elements within a Fourier synthesis that are at the frequencies $\omega_m$ and $\omega_n$ respectively. Then, their sum will constitute a periodic function of period $\tau = 2\pi/\omega_*$ if and only if $\omega_m = m\omega_*$ and $\omega_n = n\omega_*$ are integer multiples of a common frequency $\omega_*$.

For a sum of many sinusoidal elements to constitute a periodic function, it is necessary and sufficient that all of the corresponding frequencies should be integer multiples of a fundamental frequency. Whereas this condition is fulfilled by the classical Fourier series of (10), it cannot be fulfilled by a Fourier integral comprising frequencies that are arbitrary irrational numbers.

A sequence that has been sampled at the integer time points from a continuous aperiodic function that is square-integrable will have a transform that is a periodic function. This outcome is manifest in the discrete-time Fourier transform, where it can be seen that the period in question has the length of $2\pi$ radians. Let $\xi(\omega)$ be the transform of the continuous aperiodic function, and let $\xi_S(\omega)$ be the transform of the sampled sequence $\{x_t; t = 0, \pm 1, \pm 2, \ldots\}$. Then

$$x_t = \frac{1}{2\pi} \int_{-\infty}^{\infty} \xi(\omega) e^{i\omega t} d\omega = \frac{1}{2\pi} \int_{-\pi}^{\pi} \xi_S(\omega) e^{i\omega t} d\omega. \qquad (15)$$
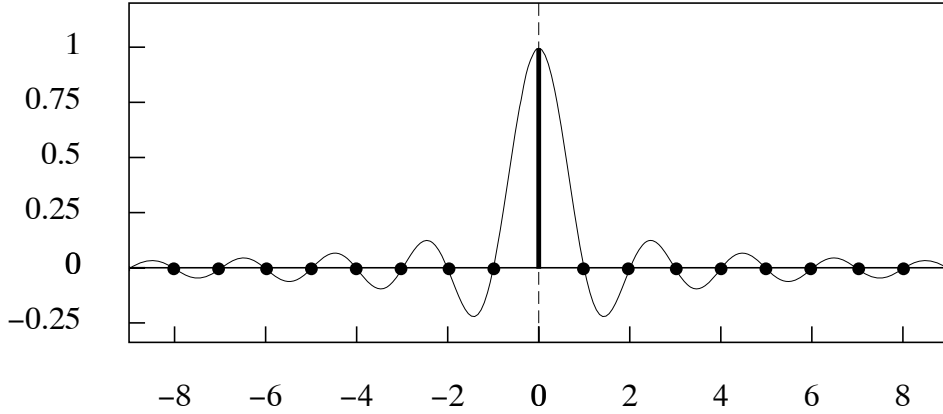
**Figure 3.** The sinc function wave-packet $\psi_0(t) = \sin(\pi t)/\pi t$ comprising frequencies in the interval $[0, \pi]$.

The equality of the two integrals implies that

$$\xi_S(\omega) = \sum_{j=-\infty}^{\infty} \xi(\omega + 2j\pi). \tag{16}$$

Thus, the periodic function $\xi_S(\omega)$ is obtained by wrapping $\xi(\omega)$ around a circle of circumference of $2\pi$ and adding the coincident ordinates.

The relationship between the Fourier series and the discrete-time Fourier transform enables us to infer that a similar effect will arise from the regular sampling of a continuous function of frequency. Thus, the corresponding time-domain function will be wrapped around a circle of circumferences $T = 2\pi/\omega_1$, where $\omega_1$ is both the fundamental frequency and the sampling interval in the frequency domain.

The Fourier integral is employed extensively in mathematical physics. In quantum mechanics, for example, it is used in one domain to describe a localised wave train and, in the other domain, to describe its frequency composition, which is a resolution of its energy amongst a set of frequencies. There is an inverse relationship between the dispersion of the wave train in space and the dispersion of its frequency components. The product of the two dispersions is bounded below according to the famous Heisenberg uncertainty principle.

A simple and important example of the Fourier integral is afforded by the sinc function wave packet and its transform, which is a frequency-domain rectangle. Figure 3 depicts a continuous sinc function of which the Fourier transform is a rectangle on the frequency interval $[-\pi, \pi]$:

$$\psi_0(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{\mathrm{i}\omega t} d\omega(t) = \left[ \frac{e^{\mathrm{i}t\omega}}{\mathrm{i}2\pi t} \right]_{-\pi}^{\pi} = \frac{\sin \pi t}{\pi t}. \tag{17}$$

Here, the final equality is by virtue of (4), which expresses a sine function as a combination of complex exponential functions.

The function $\psi_0(t)$ can also be construed as a wave packet centred on time $t = 0$. The figure also represents a sampled version the sinc function. This would be obtained, in the manner of a classical Fourier series, from the rectangle on the interval $[-\pi, \pi]$, if this were regarded as a single cycle of a periodic function.

The function $\psi_0(t)$ with $t \in \mathcal{I} = \{0, \pm 1, \pm 2, \ldots\}$ is nothing but the unit impulse sequence. Therefore, the set of all sequences $\{\psi_0(t - k); t, k \in \mathcal{I}\}$, obtained by integer displacements $k$ of $\psi_0(t)$, constitutes the ordinary orthogonal Cartesian basis in the time domain for the set of all real-valued time series.

When $t \in \mathcal{R}$ is a real-valued index of continuous time, the set of displaced sinc functions $\{\psi_0(t - k); t \in \mathcal{R}, k \in \mathcal{I}\}$ constitute a basis for the set of continuous functions of which the frequency content is bounded by the Nyquist value of $\pi$ radians per unit time interval. In common with their discretely sampled counterparts, the sequence of continuous sinc functions at integer displacements constitutes an orthogonal basis.

To demonstrate the orthogonality, consider the fact that the corresponding frequency-domain rectangle is an idempotent function. When multiplied by itself it does not change. The time-domain operation corresponding to this frequency-domain multiplication is an autoconvolution. The time-domain functions are real and symmetric with $\psi_0(k-t) = \psi_0(t-k)$, so their autoconvolution is the same as their autocorrelation:

$$\gamma(k) = \int \psi_0(t)\psi_0(k - t)dt = \int \psi_0(t)\psi_0(t - k)dt. \qquad (18)$$

Therefore, the sinc function is its own autocovariance function: $\gamma(k) = \psi_0(k)$. The zeros of the sinc function that are found at integer displacements from the centre correspond the orthogonality of sinc functions separated from each other by these distances.

A sinc function wave packet that is limited to the frequency band $[\alpha, \beta] \in [0, \pi]$, if we are talking only of positive frequencies, has the functional form of

$$\begin{aligned}
\psi(t) &= \frac{1}{\pi t}\{\sin(\beta t) - \sin(\alpha t)\} \\
&= \frac{2}{\pi t}\cos\{(\alpha + \beta)t/2\}\sin\{(\beta - \alpha)t/2\} \qquad (19) \\
&= \frac{2}{\pi t}\cos(\gamma t)\sin(\delta t),
\end{aligned}$$

where $\gamma = (\alpha+\beta)/2$ is the centre of the band and $\delta = (\beta-\alpha)/2$ is half its width. The equality follows from the identity $\sin(A + B) - \sin(A - B) = 2\cos A\sin B$.

Finally, we must consider the Fourier transform that is the most important one from the point of view of this paper. This is the discrete Fourier transform of Figure 1.(iv) that maps from a sequence of $T$ data $\{x_t; t = 0, 1, \ldots, T - 1\}$ in the time domain to a set of $T$ ordinates $\{\xi_j; j = 0, 1, \ldots, T - 1\}$ in the frequency domain at the points $\{\omega_j = 2\pi j/T; j = 0, 1, \ldots, T - 1\}$, which are equally spaced within the interval $[0, 2\pi]$, and vice versa. The relationship can be expressed in terms of sines and cosines; but it is more conveniently expressed

**Table 1. The classes of Fourier transforms\***

|  | **Periodic** | **Aperiodic** |
|---|---|---|
| **Continuous** | Discrete aperiodic *Fourier series* | Continuous aperiodic *Fourier integral* |
| **Discrete** | Discrete periodic *Discrete FT* | Continuous periodic *Discrete-time FT* |

\* The class of the Fourier transform depends upon the nature of the function which is transformed. This function may be discrete or continuous and it may be periodic or aperiodic. The names and the natures of corresponding transforms are shown in the cells of the table.

in terms of complex exponentials, even in the case where the sequence $\{x_t\}$ is a set of real data values:

$$x_t = \sum_{j=0}^{T-1} \xi_j e^{\mathrm{i}\omega_j t} \quad \longleftrightarrow \quad \xi_j = \frac{1}{T} \sum_{t=0}^{T-1} x_t e^{-\mathrm{i}\omega_j t}. \qquad (20)$$

Since all sequences in data analysis are finite, the discrete Fourier transform is used, in practice, to represent all manner of Fourier transforms.

Although the sequences in both domains are finite and of an equal number of elements, it is often convenient to regard both of them as representing single cycles of periodic functions defined over the entire set of positive and negative integers. These infinite sequences are described and the periodic extensions of their finite counterparts.

The various Fourier transforms may be summarised in a table that records the essential characteristics of the function that is to be transformed, namely whether it is discrete or continuous and whether it is periodic or aperiodic. This is table 1.

There are numerous accounts of Fourier analysis that can be accessed in pursuit of a more thorough treatment. Much of the text of Carslaw (1930) is devoted to the issues in real analysis that were raised in the course of the development of Fourier analysis. The book was first published in 1906, which is close to the dates of some of the essential discoveries that clarified such matters as Gibbs' phenomenon. It continues to be a serviceable and authoritative text. Titchmarsh (1937), from the same era, provides a classical account of the Fourier integral.

The more recent text of Kreider, Kuller, Ostberg and Perkins (1966) deals with some of the analytic complications of Fourier analysis in an accessible manner, as does the brief text of Churchill and Brown (1978). The recent text of Stade (2005) is an extensive resource. That of Nahin (2006) is both engaging and insightful.

Many of the texts that are intended primarily for electrical engineers are also useful to statisticians and econometricians. Baher (1990) gives much of the relevant analytic details, whereas Brigham (1988) provides a practical text that avoids analytic complications.

Following the discoveries of Wiener (1930) and Kolmogorov (1941b), the mathematical exposition of the spectral analysis of time series developed rapidly. Detailed treatments of the analytic issues are to be found in some of the classical texts of time series analysis, amongst which are those of Wold (1938), Doob (1953), Grenander and Rosenblat (1957), Hannan (1960), Yaglom (1962) and Rozanov (1967). Yaglom and Rozanov were heirs to a Russian tradition that began with Kolmogorov (1941b)

It was not until the 1960's that spectral analysis began to find widespread application; and the book of Jenkins and Watts (1968) was symptomatic of the increasing practicality of the subject. Econometricians also began to take note of spectral analysis in the 1960's; and two influential books that brought it to their attention were those of Granger and Hatanaka (1964) and Fishman (1969). Nerlove, Grether and Carvalho (1979) continued the pursuit of the spectral analysis of economic data.

Latterly, Brillinger (1975), Priestly (1981), Rosenblatt (1985) and Brockwell and Davis (1987) have treated Fourier analysis with a view to the statistical analysis of time series, as has Pollock (1999).

## 3. Representations of the Discrete Fourier Transform

It is often helpful to express the discrete Fourier transform in terms of the notation of the $z$-transform. The $z$-transforms of the sequences $\{x_t; t = 0, \dots, T-1\}$ and $\{\xi_j; j = 0, \dots, T-1\}$ are the polynomials

$$\xi(z) = \sum_{j=0}^{T-1} \xi_j z^j \quad \text{and} \quad x(z) = \sum_{t=0}^{T-1} x_t z^t, \tag{21}$$

wherein $z$ is an indeterminate algebraic symbol that is commonly taken to be a complex number, in accordance with the fact that the equations $\xi(z) = 0$ and $x(z) = 0$ have solutions within the complex plane.

The polynomial equation $z^T = 1$ is important in establishing the connection between the $z$-transform of a sequence of $T$ elements and the corresponding discrete Fourier transform. The solutions of the equation are the complex numbers

$$W_T^j = \exp\left(\frac{-\mathrm{i}2\pi j}{T}\right) = \cos\left(\frac{2\pi j}{T}\right) - \mathrm{i}\sin\left(\frac{2\pi j}{T}\right); j = 0, 1, \dots, T-1. \tag{22}$$

These constitute a set of points, described as the $T$ roots of unity, that are equally spaced around the circumference of the unit circle at angles of $\omega_j = 2\pi j/T$ radians. Multiplying any of these angles by $T$ will generate a multiple of $2\pi$ radians, which coincides with an angle of zero radians, which is the angle, or argument, attributable to unity within the complex plane.

Using this conventional notation for the roots of unity, we can write the discrete Fourier transform of (20) and its inverse as

$$\xi_j = \frac{1}{T} \sum_{t=0}^{T-1} x_t W_T^{jt} \quad \text{and} \quad x_t = \sum_{j=0}^{T-1} \xi_j W_T^{-jt}. \tag{23}$$

The advantage of the $z$-transforms is that they enable us to deploy the ordinary algebra of polynomials and power series to manipulate the objects of Fourier analysis. In econometrics, it is common to replace $z$ by the so-called lag operator $L$, which operates on doubly-infinite sequences and which has the effect that $Lx(t) = x(t-1)$.

It is also useful to replace $z$ by a variety of matrix operators. The matrix lag operator $L_T = [e_1, e_2, \ldots, e_{T-1}, 0]$ is derived from the identity matrix $I_T = [e_0, e_1, \ldots, e_{T-1}]$ of order $T$ by deleting its leading column and appending a column of zeros to the end of the array.

Setting $z = L_T$ within the polynomial $x(z) = x_0 + x_1 z + \cdots + x_{T-1} z^{T-1}$ gives rise to a banded lower-triangular Toeplitz matrix of order $T$. The matrices $I_T = L_T^0, L_T, L_T^2, \ldots, L_T^{T-1}$ form a basis for the vector space comprising all such matrices. The ordinary polynomial algebra can be applied to these matrices with the provision that the argument $L_T$ is nilpotent of degree $T$, which is to say that $L_T^q = 0$ for all $q \geq T$.

## Circulant Matrices

A matrix argument that has properties that more closely resemble those of the complex exponentials is the circulant matrix $K_T = [e_1, \ldots, e_{T-1}, e_0]$, which is formed by carrying the leading column of the identity matrix $I_T$ to the back of the array. This is an orthonormal matrix of which the transpose is the inverse, such that $K_T' K_T = K_T K_T' = I_T$.

The powers of the matrix form a $T$-periodic sequence such that $K_T^{T+q} = K_T^q$ for all $q$. The periodicity of these powers is analogous to the periodicity of the powers of the argument $z = \exp\{-i2\pi/T\}$, which is to be found in the Fourier transform of a sequence of $T$ elements.

The matrices $K_T^0 = I_T, K_T, \ldots, K_T^{T-1}$ form a basis for the set of all circulant matrices of order $T$—a circulant matrix $X = [x_{ij}]$ of order $T$ being defined as a matrix in which the value of the generic element $x_{ij}$ is determined by the index $\{(i-j) \bmod T\}$. This implies that each column of $X$ is equal to the previous column rotated downwards by one element. The generic circulant matrix has a form that may be illustrated follows:

$$X = \begin{bmatrix} x_0 & x_{T-1} & x_{T-2} & \ldots & x_1 \\ x_1 & x_0 & x_{T-1} & \ldots & x_2 \\ x_2 & x_1 & x_0 & \ldots & x_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{T-1} & x_{T-2} & x_{T-3} & \ldots & x_0 \end{bmatrix}. \tag{24}$$

There exists a one-to-one correspondence between the set of all polynomials of degree less than $T$ and the set of all circulant matrices of order $T$. Thus, if $x(z)$ is a polynomial of degree less that $T$, then there exits a corresponding circulant matrix

$$X = x(K_T) = x_0 I_T + x_1 K_T + \cdots + x_{T-1} K_T^{T-1}. \tag{25}$$

A convergent sequence of an indefinite length can also be mapped into a circulant matrix. If $\{c_i\}$ is an absolutely summable sequence obeying the

condition that $\sum |c_i| < \infty$, then the $z$-transform of the sequence, which is defined by $c(z) = \sum c_j z^j$, is an analytic function on the unit circle. In that case, replacing $z$ by $K_T$ gives rise to a circulant matrix $C^\circ = c(K_T)$ with finite-valued elements. In consequence of the periodicity of the powers of $K_T$, it follows that

$$C^\circ = \left\{ \sum_{j=0}^{\infty} c_{jT} \right\} I_T + \left\{ \sum_{j=0}^{\infty} c_{(jT+1)} \right\} K_T + \cdots + \left\{ \sum_{j=0}^{\infty} c_{(jT+T-1)} \right\} K_T^{T-1} \quad (26)$$

$$= c_0^\circ I_T + c_1^\circ K_T + \cdots + c_{T-1}^\circ K_T^{T-1}.$$

Given that $\{c_i\}$ is a convergent sequence, it follows that the sequence of the matrix coefficients $\{c_0^\circ, c_1^\circ, \ldots, c_{T-1}^\circ\}$ converges to $\{c_0, c_1, \ldots, c_{T-1}\}$ as $T$ increases. Notice that the matrix $c^\circ(K_T) = c_0^\circ I_T + c_1^\circ K_T + \cdots + c_{T-1}^\circ K_T^{T-1}$, which is derived from a polynomial $c^\circ(z)$ of degree $T-1$, is a synonym for the matrix $c(K_T)$, which is derived from the $z$-transform of an infinite convergent sequence.

The polynomial representation is enough to establish that circulant matrices commute in multiplication and that their product is also a polynomial in $K$. That is to say

$$\begin{array}{l} \text{If } X = x(K_T) \text{ and } Y = y(K_T) \text{ are circulant matrices,} \\ \text{then } XY = YX \text{ is also a circulant matrix.} \end{array} \quad (27)$$

The matrix operator $K_T$ has a spectral factorisation that is particularly useful in analysing the properties of the discrete Fourier transform. To demonstrate this factorisation, we must first define the so-called Fourier matrix. This is a symmetric matrix $U = T^{-1/2}[W^{jt}; t, j = 0, \ldots, T-1]$, of which the generic element in the $j$th row and $t$th column is $W^{jt}$, where $W = \exp\{-i2\pi/T\}$ is the first root of unity, from which we are now omitting the subscripted $T$ for ease of notation. On taking account of the $T$-periodicity of $W^q$, the matrix can be written explicitly as

$$U = \frac{1}{\sqrt{T}} \begin{bmatrix} 1 & 1 & 1 & \ldots & 1 \\ 1 & W & W^2 & \ldots & W^{T-1} \\ 1 & W^2 & W^4 & \ldots & W^{T-2} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & W^{T-1} & W^{T-2} & \ldots & W \end{bmatrix}. \quad (28)$$

The second row and the second column of this matrix contain the $T$ roots of unity. The conjugate matrix is defined as $\bar{U} = T^{-1/2}[W^{-jt}; t, j = 0, \ldots, T-1]$; and, by using $W^{-q} = W^{T-q}$, this can be written explicitly as

$$\bar{U} = \frac{1}{\sqrt{T}} \begin{bmatrix} 1 & 1 & 1 & \ldots & 1 \\ 1 & W^{T-1} & W^{T-2} & \ldots & W \\ 1 & W^{T-2} & W^{T-4} & \ldots & W^2 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & W & W^2 & \ldots & W^{T-1} \end{bmatrix}. \quad (29)$$

The matrix $U$ is a unitary, which is to say that it fulfils the condition

$$\bar{U}U = U\bar{U} = I. \tag{30}$$

The operator $K$, from which we now omit the subscript, can be factorised as

$$K = \bar{U}DU = U\bar{D}\bar{U}, \tag{31}$$

where

$$D = \text{diag}\{1, W, W^2, \ldots, W^{T-1}\} \tag{32}$$

is a diagonal matrix whose elements are the $T$ roots of unity, which are found on the circumference of the unit circle in the complex plane. Observe also that $D$ is $T$-periodic, such that $D^{q+T} = D^q$, and that $K^q = \bar{U}D^qU = U\bar{D}^q\bar{U}$ for any integer $q$. Since the powers of $K$ form the basis for the set of circulant matrices, it follows that all circulant matrices are amenable to a spectral factorisation based on (31).

Circulant matrices have represented a mathematical curiosity ever since their first appearance in the literature in a paper by Catalan (1846). The literature on circulant matrices, from their introduction until 1920, was summarised in four papers by Muir (1911)–(1923). A recent treatise on the subject, which contains a useful bibliography, has been provided by Davis (1979); but his book does not deal with problems in time-series analysis. An up-to-date account, orientated towards statistical signal processing, has been provided by Gray (2002).

**The Matrix Discrete Fourier Transform**

The matrices $U$ and $\bar{U}$ are entailed in the discrete Fourier transform. Thus, the equations of (20) or (23) can be written as

$$\xi = T^{-1/2}Ux = T^{-1}Wx \qquad \text{and} \qquad x = T^{1/2}\bar{U}\xi = \bar{W}\xi, \tag{33}$$

where $x = [x_0, x_1, \ldots x_{T-1}]'$ and $\xi = [\xi_0, \xi_1, \ldots, \xi_{T-1}]'$, and where $W = [W^{jt}]$ and $\bar{W} = [W^{-jt}]$ are the matrices of (28) and (29) respectively, freed from their scaling factors. We may note, in particular, that

$$\iota = T^{1/2}Ue_0 = T^{1/2}\bar{U}e_0 \qquad \text{and} \qquad e_0 = T^{-1/2}U\iota = T^{-1/2}\bar{U}\iota, \tag{34}$$

where $e_0 = [1, 0, \ldots 0]'$ is the leading column of the identity matrix and $\iota = [1, 1, \ldots, 1]'$ is the summation vector. Thus, a spike or an impulse located at $t = 0$ in the time domain or at $j = 0$ in the frequency domain is transformed into a constant sequence in the other domain and vice versa. These are extreme examples of the inverse relationship between the dispersion of a sequence and that of its transform.

Given the circulant the matrix $X = x(K)$ defined in (25), the discrete Fourier transform and its inverse may be represented by the following spectral factorisations:

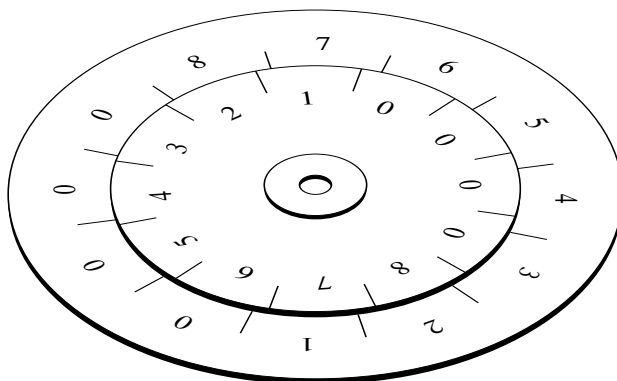$$X = \bar{U}x(D)U \qquad \text{and} \qquad x(D) = UX\bar{U}. \tag{35}$$

**Figure 4.** A device for finding the circular correlation of two sequences. The upper disc is rotated clockwise through successive angles of 30 degrees. Adjacent numbers on the two discs are multiplied and the products are summed and divided by the sample size $T$ to obtain the circular covariances.

Here, $x(D) = T\mathrm{diag}\{\xi_0, \xi_1, \ldots, \xi_{T-1}\}$ is a diagonal matrix containing the spectral ordinates of the data. The equations of (33) are recovered by post multiplying the $X$ by $e_0$ and $x(D)$ by $\iota$ and using the results of (34).

Assuming that the data are mean adjusted, i.e. that $\iota'X = 0$, the cross product of the matrix $X$ gives rise to the matrix $C^\circ$ of the circular autocovariances of the data. Thus

$$
\begin{aligned}
T^{-1}X'X = C^\circ &= \bar{U}c^\circ(D)U \\
&= T^{-1}\bar{U}x(\bar{D})x(D)U,
\end{aligned}
\tag{36}
$$

where we have used $X' = Ux(D)\bar{U} = \bar{U}x(\bar{D})U$.

The values of the elements of $[c_0^\circ, c_1^\circ, \ldots, c_{T-1}^\circ]' = c^\circ = C^\circ e_0$ are given by the formula

$$
c_\tau^\circ = \frac{1}{T}\sum_{t=0}^{T-1} x_t x_{t+\tau}; \quad \text{where} \quad x_t = x_{(t \bmod T)}
$$

or, equivalently,

$$
\begin{aligned}
c_\tau^\circ &= \frac{1}{T}\sum_{t=0}^{T-1-\tau} x_t x_{t+\tau} + \frac{1}{T}\sum_{t=0}^{\tau-1} x_t x_{t+T-\tau} \\
&= c_\tau + c_{T-\tau},
\end{aligned}
\tag{37}
$$

where $c_\tau$ is the ordinary empirical autocovariance of lag $\tau$.

The vector $c^\circ$ of circular autocovariances is obtained via a circular correlation of the data vector $x$. One can envisage the data disposed at $T$ equally-spaced points around the circumferences of two discs on a common axis. The same numbers on the two discs are aligned and products are formed and added. This generates the sum of squares. Then, the upper disc is rotated by an angle of $2\pi/T$ radians and the displaced numbers are multiplied together and added. The process is repeated. The $T$ values formed via a complete rotation are each

15

divided by $T$, and the result is the sequence of circular autocovariances. This device is illustrated in Figure 4, where $T = 12$ and where identical sequences on the two discs end with four zeros.

The core of the factorisation of the matrix of $C^\circ$ of (36) is the real-valued diagonal matrix

$$c^\circ(D) = T^{-1}x(\bar{D})x(D) = T\text{Diag}\{|\xi_0|^2, |\xi_1|^2, \ldots, |\xi_{T-1}|^2\}, \qquad (38)$$

of which the graph of the elements is described as the periodogram. (The frequencies indexed by $j = 0, \ldots, T-1$ extend from 0 to $2\pi(T-1)/T$. However, for real-valued data, there is $\xi_j = \xi_{T-j}$, and so it is customary to plot the periodogram ordinates only over a frequency range from 0 to $\pi$, as in Figure 14.)

Since $e_0' C^\circ e_0 = c_0^\circ = c_0$ and $Ue_0 = \bar{U}e_0 = T^{-1/2}\iota$, we get the following expression for the variance:

$$c_0 = T^{-1}\sum_{j=0}^{T-1} x_t^2$$

$$= T^{-2}\iota' x(\bar{D})x(D)\iota = \sum_{j=0}^{T-1} |\xi_j|^2. \qquad (39)$$

This is the discrete version of Parseval's theorem. In statistical terms, it represents a frequency-specific analysis of variance.

## 4. Fourier Analysis of Temporal Sequences

It is clear that the discrete-time Fourier transform is inappropriate to a doubly-infinite stationary stochastic sequence $y(t) = \{y_t; t = 0, \pm 1, \pm 2, \ldots\}$ defined over the set of positive and negative integers. Such a sequence, which has infinite energy, is not summable.

In the appendix, it is demonstrated there there exists a spectral representation of $y(t)$ in the form of

$$y(t) = \int_{-\pi}^{\pi} e^{\mathrm{i}\omega t}dZ(\omega), \qquad (40)$$

where $Z(\omega)$ is continuous non-differentiable complex-valued stochastic process defined on the interval $[-\pi, \pi]$. The increments $dZ(\omega)$, $dZ(\lambda)$ of the process are uncorrelated for all $\lambda \neq \omega$. However, the statistical expectation

$$E\{dZ(\omega)dZ^*(\omega)\} = dF(\omega) = f(\omega)d\omega \qquad (41)$$

comprises a function $F(\omega)$, known as the spectral distribution function, which is continuous and differentiable almost everywhere in the interval, provided that $y(t)$ contains no perfectly regular components of finite power. If there are regular components within $y(t)$, then $F(\omega)$ will manifest jump discontinuities at the corresponding frequencies.

The discrete-time transform comes into effect when it is applied not to the data sequence itself but, instead, to its autocovariance function

$$\gamma(\tau) = \{\gamma_t; \tau = 0, \pm 1, \pm 2, \ldots\} \quad \text{where} \quad \gamma_\tau = \gamma_{-\tau} = E(y_t y_{t-\tau}). \quad (42)$$

To ensure that its transform is mean-square convergent, the autocovariance sequence must the square summable such that $\sum \gamma_\tau^2 < \infty$. This is a necessary and sufficient condition for the stationarity of the process $y(t)$.

The Fourier transform of the autocovariance function is the spectral density function or power spectrum $f(\omega)$, which, in view of the symmetry of $\gamma(\tau)$, can be expressed as

$$
\begin{aligned}
f(\omega) &= \frac{1}{2\pi} \Big\{ \gamma_0 + \sum_{\tau=1}^{\infty} \gamma_\tau (e^{i\omega\tau} + e^{-i\omega\tau}) \Big\} \\
&= \frac{1}{2\pi} \Big\{ \gamma_0 + 2 \sum_{\tau=1}^{\infty} \gamma_\tau \cos(\omega\tau) \Big\}.
\end{aligned}
\quad (43)
$$

Notice, however, that compared with equation (14), the scaling factor $1/2\pi$ has migrated from the Fourier integral to the series. Now, the inverse transform takes the form of

$$\gamma_\tau = \int_{-\pi}^{\pi} e^{i\omega\tau} f(\omega) d\omega. \quad (44)$$

Setting $\tau = 0$ gives $\gamma_0 = \int f(\omega) d\omega$, which indicates that the power of the process, which is expressed by its variance, is equal to the integral of the spectral density function over the interval $[-\pi, \pi]$. This a form of Parseval's theorem. The sequence $y(t)$ is not square summable and, therefore, its energy cannot be expressed in terms of a sum of squares. Instead, its power is expressed, in the time domain, via the statistical moment $\gamma_0 = E(y_t^2)$.

In the frequency domain, this power is attributed to an infinite set of sinusoidal components, of which the measure of power is provided by the continuous spectral density function $f(\omega)$. One can see, by a direct analogy with a continuous probability density function, that the power associated with a component at a particular frequency is vanishingly small, i.e. of measure zero.

Often, it is assumed that, in addition to square summability, the autocovariances obey the stronger condition that $\sum |\gamma_\tau| < \infty$, which is the condition of absolute summability. This condition is satisfied by processes generated by finite-order autoregressive moving-average (ARMA) models, which can be represented by setting $d = 0$ in the equation

$$y(z) = \frac{\theta(z)}{(1-z)^d \phi(z)} \varepsilon(z), \quad (45)$$

which describes an autoregressive integrated moving-average (ARIMA) model. Here $\varepsilon(z)$ and $y(z)$ are, respectively, the $z$-transform of a white-noise forcing function and of the output sequence and $\theta(z)$ and $\phi(z)$ are polynomials of finite degree. Provided that $d = 0$ and that the roots of the equation

$\phi(z) = \phi_0 + \phi_1 z + \cdots + \phi_p z^p = 0$ lie outside the unit circle, the spectral density function of the process will be bounded and everywhere continuous.

The absolute summability of the ARMA autocovariances is a consequence of the summability of the infinite sequence of moving-average coefficients generated by the series expansion of the rational transfer function $\theta(z)/\phi(z)$ that maps from the white-noise forcing function $\varepsilon(t)$ to the output $y(t)$. These coefficients constitute the impulse response of the transfer function; and the condition of their absolute summability is the bounded input–bounded output (BIBO) stability condition of linear systems analysis. (see Pollock 1999, for example.)

The autocovariance generating function of an ARMA process, which is the $z$-transform of the two-sided sequence $\{\gamma_\tau; \tau = 0, \pm 1, \pm 2, \ldots\}$, is given by

$$\gamma(z) = \sigma^2 \frac{\theta(z)\theta(z^{-1})}{\phi(z)\phi(z^{-1})}, \tag{46}$$

where $\sigma^2 = V\{\varepsilon(t)\}$ is the variance of the white-noise forcing function. The spectral density function $f(\omega) = \gamma(\exp\{i\omega\})$ is produced when the locus of $z$ is the unit circle in the complex plane.

The conditions of absolute and square summability are violated by the ARIMA models that can be used to describe trended processes. Unless there are finite starting values at a finite distance from the current value, the variance of such a process will be unbounded. (If equation (45) is to remain viable when $d = 1$, then it is necessary to define $y(z)$ and $\varepsilon(z)$ to be the $z$-transforms of finite sequences that incorporate the appropriate initial conditions—see Pollock (2008), for example.)

The spectral density function of an ARIMA process, which is defined in respect of the doubly infinite sequence $y(t)$ will not have a finite-valued integral. Therefore, it is liable to be described as a pseudo spectral density function. See, for example, Maravall and Piece (1987). The unit roots within the factor $(1 - z)^d$ of equation (45) contribute infinite power to the ARIMA process from within the neighbourhood of the zero frequency, where the pseudo spectrum is unbounded.

In a seasonal ARIMA model, which is used to model persistent seasonal fluctuations, the denominator of the transfer function incorporates the factor

$$S(z) = 1 + z + z^2 + \cdots + z^{s-1} = \prod_{j=1}^{s-1}(1 - e^{2\pi j/s}z). \tag{47}$$

where $s$ is the number of seasons or months of the year. Here, the generic quadratic factor is

$$1 - 2\cos(2\pi j/s) + z^2 = (1 - e^{2\pi j/s}z)(1 - e^{2\pi(s-j)/s}z), \tag{48}$$

where $j < s/2$. This factor, which will contribute infinite power to the pseudo spectrum, corresponds to an unbounded spike at the frequency of $\omega_j = 2\pi j/s$ which is the $j$th harmonic of the fundamental seasonal frequency. In the
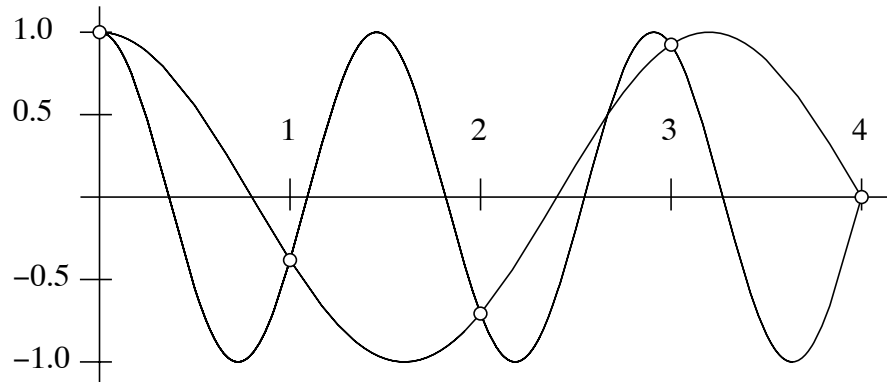
**Figure 5.** The values of the function $\cos\{(11/8)\pi t\}$ coincide with those of its alias $\cos\{(5/8)\pi t\}$ at the integer points $\{t = 0, \pm 1, \pm 2, \ldots\}$.

numerator of the transfer function, there should be a polynomial $R(z) = 1 + \rho z + \rho^2 z^2 + \cdots + \rho^{s-1} z^{s-1}$ with a value of $\rho < 1$ close to unity, which will serve to counteract the effects of $S(z)$ at the non-seasonal frequencies.

Latterly, econometricians and others have been focussing their attention on so-called long-memory processes that are characterised by autocovariance sequences that are square summable but not absolutely summable. Such processes are commonly modelled by replacing the difference operator of an ARIMA model by a corresponding fractional difference operator $(1-z)^d$, where $d \in [0, 0.5)$. Such models have spectra that are unbounded at the zero frequency. When $|d| \geq 0.5$, the condition of the square summability of the autocovariances is violated and the process is no longer stationary.

Models incorporating the factional difference operator were proposed by Granger and Joyeux (1980) and Hoskin (1981). They have been described in detail by Beran (1994) and by Palma (2007). Granger and Ding (1996) have considered the provenance of long-memory processes; and they have provided some alternative explanations.

Persistent seasonal fluctuations can also be described by long-memory models. Hoskin (1984) noted that, when $|d| < 0.5$ and $|\phi| < 1$, the process defined by $y(z) = \varepsilon(z)/(1 - 2\phi z + z^2)^d$ would generate a sequence displaying long-term persistence and quasi periodic behaviour. The transfer function of this model, which incorporates a fractional power of the seasonal polynomial of (48), is a so-called Gegenbauer polynomial. The lead of Hoskin has been followed by other authors, including Gray, Zhang and Woodward, (1989) Hidalgo (1997), Artech and Robinson (2000) and McCoy and Stephens (2004), who have defined models with spectra that are unbounded at multiple frequencies.

**The Problem of Aliasing**

Notice that the (positive) frequencies in the spectral analysis of a discrete-time process are limited to the interval $[0, \pi]$, bounded by the so-call Nyquist frequency $\pi$. The process of sampling imposes a limit of the observability of
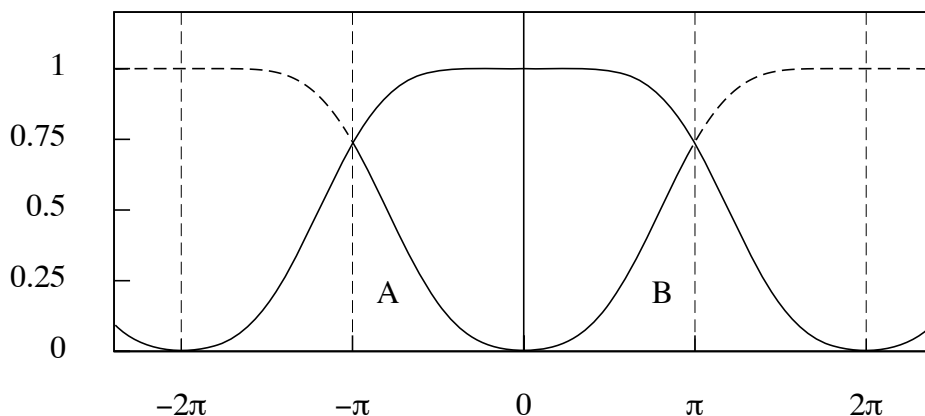
**Figure 6.** The function $f(\lambda/2 + \pi)$, represented by the semi-continuous line, superimposed upon the function $f(\lambda/2)$. The sum of these functions is a $2\pi$-periodic function $f_H(\lambda)$, which represents the spectral density of a subsampled process. The segments $A$ and $B$ of $f(\lambda/2 + \pi)$ that fall within the interval $[-\pi, \pi]$ may be construed as the effects of wrapping $f(\lambda/2)$, defined over the interval $[-2\pi, 2\pi]$, around a circle of circumference $2\pi$.

high-frequency components. To be observable in sampled data, a sinusoidal motion must take no less than two sample periods to complete its cycle, which limits its frequency to no more than $\pi$ radians per period.

A motion of higher angular velocity will be mistaken for one of a velocity that lies within the observable interval, and this is described as the problem of aliasing. The problem is illustrated by Figure 5 which shows that, by sampling it at the integer points $\{t = 0, \pm 1, \pm 2, \ldots\}$, a sinusoid with frequency of $(11/8)\pi$ radians per period, which is in excess of the limiting Nyquist value of $\pi$, will be mistaken for one with a frequency of $(5/8)\pi$.

The extent to which aliasing is a problem depends upon the structure of the particular time series under analysis. It will be suggested later that, for many econometric time series, the problem does not arise, for the reason that their (positive) frequencies are band-limited to a subinterval of the range $[0, \pi]$. However, this fact, which has its own problems, is not commonly recognised in econometric time-series analysis.

To understand the statistical aspects of aliasing, we may consider a stationary stochastic process $y(t) = \{y_t; t = 0, \pm 1, \pm 2, \ldots\}$ with an autocovariance function $\gamma(\tau) = \{\gamma_\tau; \tau = 0, \pm 1, \pm 2, \ldots\}$. The relationship between the spectral density function $f(\omega)$ and the autocovariance function is a follows:

$$\gamma(\tau) = \int_{-\pi}^{\pi} f(\omega) e^{i\omega\tau}\, d\omega \quad \longleftrightarrow \quad f(\omega) = \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} \gamma_\tau e^{-i\omega\tau}. \qquad (49)$$

When alternate values are selected from the data sequence, the autocovariance function is likewise subsampled, and there is

$$\gamma(2\tau) = \int_{-\pi}^{\pi} f(\omega)e^{i\omega(2\tau)}d\omega$$

$$= \frac{1}{2}\int_{-2\pi}^{2\pi} f(\lambda/2)e^{i\lambda\tau}d\lambda \tag{50}$$

$$= \frac{1}{2}\left\{\int_{-2\pi}^{-\pi} f(\lambda/2)e^{i\lambda\tau}d\lambda + \int_{-\pi}^{\pi} f(\lambda/2)e^{i\lambda\tau}d\lambda + \int_{\pi}^{2\pi} f(\lambda/2)e^{i\lambda t}d\lambda\right\}.$$

Here, we have defined $\lambda = 2\omega$, and we have used the change of variable technique to obtain the second equality.

Within the integrand of (50), there is $\exp\{i\lambda\tau\} = \exp\{i(\lambda\tau \pm 2\pi)\}$. Also, $f(\{\lambda/2\} - \pi) = f(\{\lambda/2\} + \pi)$, by virtue of the $2\pi$-periodicity of the function. Therefore,

$$\int_{-2\pi}^{-\pi} f(\lambda/2)e^{i\lambda\tau}d\lambda = \int_{0}^{\pi} f(\{\lambda/2\} - \pi)e^{i(\lambda\tau-2\pi)}d\lambda$$

$$= \int_{0}^{\pi} f(\{\lambda/2\} + \pi)e^{i\lambda\tau}, \tag{51}$$

where the first equality is an identity and the second exploits the results above. Likewise,

$$\int_{\pi}^{2\pi} f(\lambda/2)e^{i\lambda\tau}d\lambda = \int_{-\pi}^{0} f(\{\lambda/2\} + \pi)e^{i(\lambda\tau+2\pi)}d\lambda$$

$$= \int_{-\pi}^{0} f(\{\lambda/2\} + \pi)e^{i\lambda\tau}. \tag{52}$$

It follows that within the expression of (50), the first integral may be translated to the interval $[0, \pi]$, whereas the third integral may be translated to the interval $[-\pi, 0]$. After their translations, the first and the third integrands can be combined to form the segment of the function $f(\pi + \lambda/2)$ that falls in the interval $[-\pi, \pi]$. The consequence is that

$$\gamma(2\tau) = \frac{1}{2}\int_{-\pi}^{\pi} \{f(\lambda/2) + f(\pi + \lambda/2)\}e^{i\lambda\tau}d\lambda. \tag{53}$$

It follows that

$$\gamma(2\tau) \longleftrightarrow f_H(\lambda) = \frac{1}{2}\{f(\lambda/2) + f(\pi + \lambda/2)\}, \tag{54}$$

where $f_H(\lambda) = f_H(\lambda + 2\pi)$ is the spectral density function of the subsampled process. Figure 6 shows the relationship between $f(\lambda/2)$ and $f(\pi + \lambda/2)$.

The superimposition of the shifted frequency function $f(\pi + \lambda/2)$ upon the function $f(\lambda/2)$ corresponds to a process of aliasing. To visualise the process, one can imagine a single cycle of the original function $f(\omega)$ over the interval

$[-\pi, \pi]$. The dilated function $f(\lambda/2)$ manifests a single cycle over the interval $[-2\pi, 2\pi]$. By wrapping this segment twice around the circle defined by $\exp\{-i\lambda\}$ with $\lambda \in [-\pi, \pi]$, we obtain the aliased function $\{f(\lambda/2)+f(\pi+\lambda/2)\}$.

Observe that, if $f(\omega) = 0$ for all $|\omega| > \pi/2$, then the support of the dilated function would be the interval $[-\pi, \pi]$, and the first and third integrals would be missing from the final expression of (50). In that case, there would be no aliasing, and the function $f_H(\lambda)$ would represent the spectrum of the original process accurately.

In the case of a finite sample, the frequency-domain representation and the time-domain representation are connected via the discrete Fourier transform. We may assume that the size $T$ of the original sample is an even number. Then, the elements of the sequence $\{x_0, x_2, x_4, \ldots, x_{T-1}\}$, obtained by selecting alternate points, are described by

$$
\begin{aligned}
x_{2t} &= \sum_{j=0}^{T-1} \xi_j \exp\{i\omega_1 2tj\}, \quad \text{where} \quad \omega_1 = \frac{2\pi}{T} \\
&= \sum_{j=0}^{(T/2)-1} \{\xi_j + \xi_{j+(T/2)}\} \exp\{i\omega_2 tj\}, \quad \text{where} \quad \omega_2 = 2\omega_1 = \frac{4\pi}{T}.
\end{aligned}
\tag{55}
$$

Here, the second equality follow from the fact that

$$
\exp\{i\omega_2 j\} = \exp\{i\omega_2[j \bmod (T/2)]\}.
$$

To envisage the effect of equation (55), we can imagine wrapping the sequence $\xi_j; j = 0, 1, \ldots, T-1$ twice around a circle of circumference $T/2$ so that its elements fall on $T/2$ points at angles of $\omega_{2j} = 4\pi j/T; j = 1, 2, \ldots, T/2$ from the horizontal.

Undoubtedly, the easiest way to recognise the effect of aliasing in the case of a finite sample is to examine the effect upon the matrix representation of the discrete Fourier transform. For conciseness, we adopt the expressions under (33), and we may take the case were $T = 8$. Subsampling by a factor of 2 is a matter of removing from the matrix $\bar{U}$ alternate rows corresponding to the odd-valued indices. This gives

$$
\begin{aligned}
\begin{bmatrix} x_0 \\ x_2 \\ x_4 \\ x_6 \end{bmatrix} &= \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & W_8^6 & W_8^4 & W_8^2 & 1 & W_8^6 & W_8^4 & W_8^2 \\ 1 & W_8^4 & 1 & W_8^4 & 1 & W_8^4 & 1 & W_8^4 \\ 1 & W_8^2 & W_8^4 & W_8^6 & 1 & W_8^2 & W_8^4 & W_8^6 \end{bmatrix} \begin{bmatrix} \xi_0 \\ \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \\ \xi_5 \\ \xi_6 \\ \xi_7 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & W_4^3 & W_4^2 & W_4 \\ 1 & W_4^2 & 1 & W_4^2 \\ 1 & W_4 & W_4^2 & W_4^3 \end{bmatrix} \begin{bmatrix} \xi_0 + \xi_4 \\ \xi_1 + \xi_5 \\ \xi_2 + \xi_6 \\ \xi_3 + \xi_7 \end{bmatrix}.
\end{aligned}
\tag{56}
$$

We see that, in the final vector on the RHS, the elements $\xi_0$, $\xi_1$, $\xi_2$, $\xi_3$ are conjoined with the elements $\xi_4$, $\xi_5$, $\xi_6$, $\xi_7$, which are associated with frequencies that exceed the Nyquist rate.

## 5. Linear Filters in Time and Frequency

In order to avoid the effects of aliasing in the process of subsampling the data, it is appropriate to apply an anti-aliasing filter to remove from the data those frequency components that are in excess of the Nyquist frequency of the sub-sampled date. In the case of downsampling by a factor of two, we should need to remove, via a preliminary operation, those components of frequencies in excess of $\pi/2$.

A linear filter is defined by a set of coefficients $\{\psi_k; k = -p, \ldots, q\}$, which are applied to the data sequence $y(t)$ via a process of linear convolution to give a processed sequence

$$x(t) = \sum_{k=-p}^{q} \psi_k y(t - k). \tag{57}$$

On defining the $z$-transforms $\psi(z) = \sum_k \psi_k z^k$, $y(z) = \sum_t y_t z^t$ and $x(z) = \sum_t x_t z^t$, we may represent the filtering process by the following polynomial or series equation:

$$x(z) = \psi(z)y(z). \tag{58}$$

Here, $\psi(z)$ is described as the transfer function of the filter.

To ensure that the sequence $x(t)$ is bounded if $y(t)$ is bounded, it is necessary and sufficient for the coefficient sequence to be absolutely summable, such that $\sum_k |\psi_k| < \infty$. This is known by engineers as the bounded input–bounded output (BIBO) condition. The condition guarantees the existence of the discrete-time Fourier transform of the sequence. However, a meaning must be sought for this transform.

Consider, therefore, the mapping of a (doubly-infinite) cosine sequence $y(t) = \cos(\omega t)$ through a filter defined by the coefficients $\{\psi_k\}$. This produces the output

$$
\begin{aligned}
x(t) &= \sum_k \psi_k \cos(\omega[t - k]) \\
&= \sum_k \psi_k \cos(\omega k)\cos(\omega t) + \sum_k \psi_k \sin(\omega k)\sin(\omega t) \\
&= \alpha \cos(\omega t) + \beta \sin(\omega t) = \rho \cos(\omega t - \theta),
\end{aligned}
\tag{59}
$$

where $\alpha = \sum_k \psi_k \cos(\omega k)$, $\beta = \sum_k \psi_k \sin(\omega k)$, $\rho = \sqrt{(\alpha^2 + \beta^2)}$ and $\theta = \tan^{-1}(\beta/\alpha)$. These results follow in view of the trigonometrical identity of (3).

The equation indicates that the filter serves to alter the amplitude of the cosine signal by a factor of $\rho$ and to displace it in time by a delay of $\tau = \theta/\omega$ periods, corresponding to an angular displacement of $\theta$ radians. A (positive) phase delay is inevitable if the filter is operating in real time, with $j \geq 0$, which is to say that only the present and past values of $y(t)$ are comprised by the filter. However, if the filter is symmetric, with $\psi_{-k} = \psi_k$, which necessitates working off-line, then $\beta = 0$ and, therefore, $\theta = 0$.

Observe that the values of $\rho$ and $\theta$ are specific to the frequency $\omega$ of the cosine signal. The discrete-time Fourier transform of the filter coefficients

allows the gain and the phase effects to be assessed over the entire range of frequencies in the interval $[0, \pi]$. The transform may be obtained by setting $z = \exp\{-i\omega\} = \cos(\omega) - i\sin(\omega)$ within $\psi(z)$, which constrains this argument to lie on the unit circle in the complex plane. The resulting function

$$\psi(\exp\{-i\omega\}) = \sum_k \psi_k \cos(\omega k) - i \sum_k \psi_k \sin(\omega k)$$
$$= \alpha(\omega) - i\beta(\omega) \tag{60}$$

is the frequency response function, which is, in general, a periodic complex-valued function of $\omega$ with a period of $2\pi$. In the case of a symmetric filter, it becomes a real-valued and even function, which is symmetric about $\omega = 0$.

For a more concise notation, we may write $\psi(\omega)$ in place of $\psi(\exp\{-i\omega\})$. In that case, the relationship between the coefficients and their transform can be conveyed by writing $\{\psi_k\} \longleftrightarrow \psi(\omega)$.

The effect of a linear filter on a stationary stochastic process, which is not summable, can be represented in terms of the autocovariance generating functions $f_y(z)$ and $f_x(z)$ of the sequence and its transform, respectively. Thus

$$f_x(z) = |\psi(z)|^2 f_y(z), \tag{61}$$

where $|\psi(z)|^2 = \psi(z)\psi(z^{-1})$. Setting $z = \exp\{-i\omega\}$ gives the relationship between the spectral density functions, which entails the squared gain $|\psi(\omega)|^2 = |\psi(e^{-i\omega})|^2$ of the filter.

**The Ideal Filter**

Imagine that it is required to remove from a data sequence the components of frequencies in excess of $\omega_d$ within the interval $[0, \pi]$. This requirement may arise from an anti-aliasing operation, preliminary to sub sampling, or it may arise from a desire to isolate a component that lies in the frequency range $[0, \omega_d]$. We shall discuss the motivation in more detail later.

The ideal filter for the purpose would be one that preserves all components with (positive) frequencies in the subinterval $[0, \omega_d)$ and nullifies all those with frequencies in the subinterval $(\omega_d, \pi]$. The specification of the frequency response function of such a filter over the interval $[-\pi, \pi]$ is

$$\psi(\omega) = \begin{cases} 1, & \text{if } |\omega| \in (0, \omega_d), \\ 1/2, & \text{if } \omega = \pm\omega_d, \\ 0, & \text{otherwise.} \end{cases} \tag{62}$$

The coefficients of the filter are given by the sinc function:

$$\psi_k = \frac{1}{2\pi} \int_{-\omega_d}^{\omega_d} e^{i\omega k} d\omega = \begin{cases} \dfrac{\omega_d}{\pi}, & \text{if } k = 0; \\ \dfrac{\sin(\omega_d k)}{\pi k}, & \text{if } k \neq 0 \,. \end{cases} \tag{63}$$
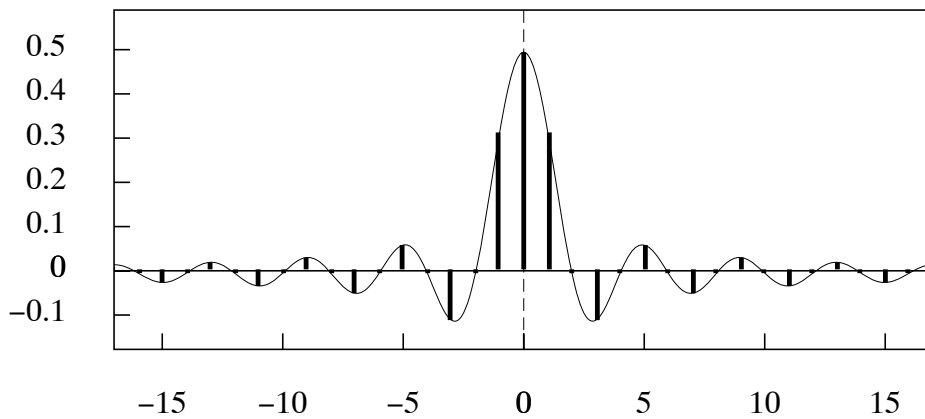
**Figure 7.** The central coefficients of the Fourier transform of the frequency response of an ideal low pass filter with a cut-off point at $\omega = \pi/2$. The sequence of coefficients extends indefinitely in both directions.
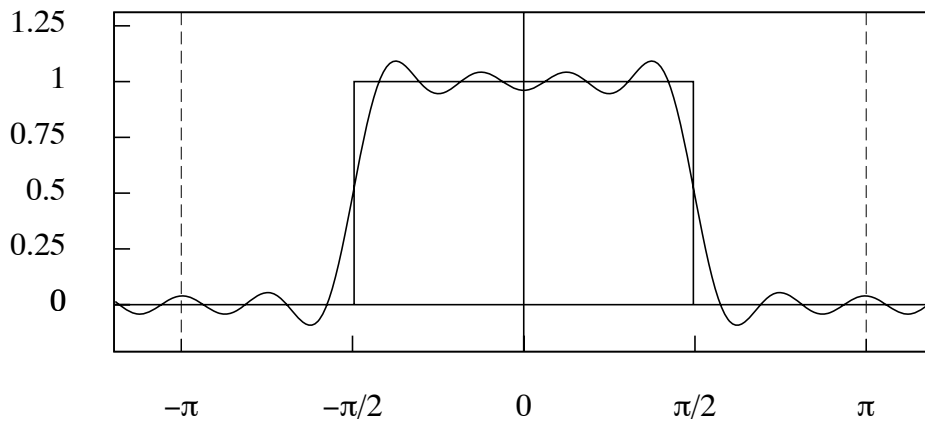


**Figure 8.** The result of applying a 17-point rectangular window to the coefficients of an ideal lowpass filter with a cut-off point at $\omega = \pi/2$.
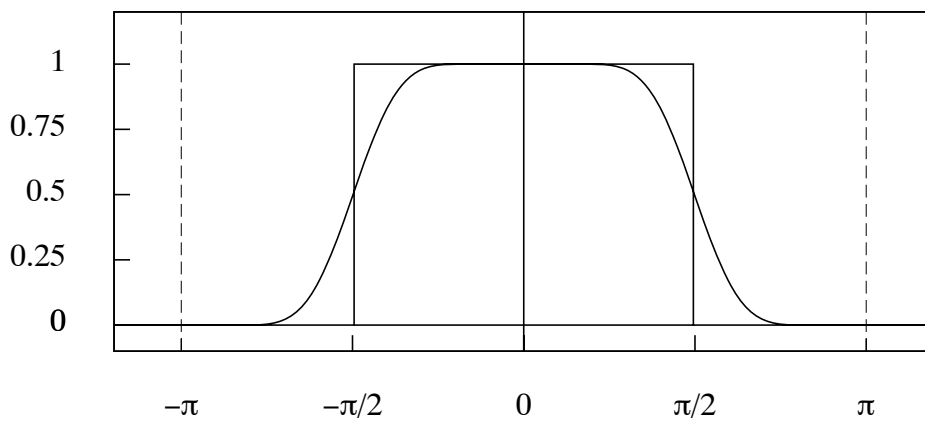


**Figure 9.** The result of applying a 17-point Blackman window to the coefficients of an ideal lowpass filter with a cut-off point at $\omega = \pi/2$.

These coefficients form a doubly-infinite sequence and, in order to apply such a filter to a data sample of size $T$, it is customary to truncate the sequence, retaining only a limited number of its central elements (Figure 7).

The truncation gives rise to a partial-sum approximation of the ideal frequency response that has undesirable characteristics. In particular, there is a ripple effect whereby the gain of the filter fluctuates within the pass band, where it should be constant with a unit value, and within the stop band, where it should be zero-valued. Within the stop band, there is a corresponding problem of leakage whereby the truncated filter transmits elements that ought to be blocked (Figure 8). This is nothing but Gibbs' effect, which has been discussed already in section 3 and illustrated in Figure 2.

The classical approach to these problems, which has been pursued by electrical engineers, has been to modulate the truncated filter sequence with a so-called window sequence, which applies a gradual taper to the higher-order filter coefficients. (A full account of this has been given by Pollock, 1999.)

The effect is to suppress the leakage that would otherwise occur in regions of the stop band that are remote from the regions where the transitions occur between stop band and pass band. The detriment of this approach is that it exacerbates the extent of the leakage within the transition regions (Figure 9).

The implication here seems to be that, if we wish to make sharp cuts in the frequency domain to separate closely adjacent spectral structures, then we require a filter of many coefficients and a very long data sequence to match it. However, Figures 8 and 9 are based upon the discrete-time Fourier transform, which entails an infinite sequence in the time domain and a continuous function in the frequency domain; and their appearances are liable be misleading.

In practice, one must work with a finite number of data points that transform into an equal number of frequency ordinates. If one is prepared to work in the frequency domain, then one can impose an ideal frequency selection upon these ordinates by setting to zero those that lie within the stop band of the ideal filter and conserving those that lie within the pass band. One can ignore whatever values a continuous frequency response function might take in the interstices between these frequency ordinates, because they have no implications for the finite data sequence.

The same effects can be achieved by operations performed within the time domain. These entail a circular filter, which is applied to the data by a process of circular convolution that is akin to the process of circular correlation described and illustrated in section 3. In the case of circular convolution, the data sequence and the sequence of filter coefficients are disposed around the circle in opposite directions—clockwise and anti clockwise. The set of $T$ coefficients of a circular filter are derived by applying the (inverse) discrete Fourier transform to a sample of the ordinates of the ideal frequency response, taken at the $T$ Fourier frequencies $\omega_j = 2\pi j/T; j = 0, 1, \ldots, T-1$.

To understand the nature of the filter, consider evaluating the frequency response of the ideal filter at one of the points $\omega_j$. The response, expressed in

terms of the doubly-infinite coefficient sequence of the ideal filter, is

$$\psi(\omega_j) = \sum_{k=-\infty}^{\infty} \psi_k W^{jk}, \tag{64}$$

where $W^j = \exp\{-i\omega_j\} = \cos(\omega_j) - i\sin(\omega_j)$. However, since $W^q$ is a $T$-periodic function, there is $W \uparrow q = W \uparrow (q \bmod T)$, where the upward arrow signifies exponentiation. Therefore,

$$\psi(\omega_j) = \left\{ \sum_{q=-\infty}^{\infty} \psi_{qT} \right\} + \left\{ \sum_{q=-\infty}^{\infty} \psi_{qT+1} \right\} W^j + \cdots$$

$$+ \left\{ \sum_{q=-\infty}^{\infty} \psi_{qT+T-1} \right\} W^{j(T-1)} \tag{65}$$

$$= \psi_0^\circ + \psi_1^\circ W^j + \cdots + \psi_{T-1}^\circ W^{j(T-1)}, \quad \text{for} \quad j = 0, 1, 2, \ldots, T-1.$$

Thus, the coefficients $\psi_0^\circ, \psi_1^\circ, \ldots, \psi_{T-1}^\circ$ of the circular filter would be obtained by wrapping the infinite sequence of sinc function coefficients of the ideal filter around a circle of circumference $T$ and adding the overlying coefficients. By applying a sampling process to the continuous frequency response function, we have created an aliasing effect in the time domain. However, summing the sinc function coefficients is not a practical way of obtaining the coefficients of the finite-sample filter. They should be obtained, instead, by transforming the frequency-response sequence or else from one of the available analytic formulae.

In the case where $\omega_d = d\omega_1 = 2\pi d/T$, the coefficients of the filter that fulfils the specification of (61) within a finite sample of $T$ points are given by

$$\phi_d^\circ(k) = \begin{cases} \dfrac{2d}{T}, & \text{if } k = 0 \\[2ex] \dfrac{\cos(\omega_1 k/2) \sin(d\omega_1 k)}{T \sin(\omega_1 k/2)}, & \text{for } k = 1, \ldots, [T/2], \end{cases} \tag{66}$$

where $\omega_1 = 2\pi/T$ and where $[T/2]$ is the integral part of $T/2$.

A more general specification for a bandpass filter is as follows:

$$\phi(\omega) = \begin{cases} 1, & \text{if } |\omega| \in (-\omega_a, \omega_b), \\[1ex] 1/2, & \text{if } \omega = \pm\omega_a, \pm\omega_b, \\[1ex] 0, & \text{for } \omega \text{ elsewhere in } [-\pi, \pi), \end{cases} \tag{67}$$

where $\omega_a = a\omega_1$ and $\omega_b = b\omega_1$. This can be fulfilled by subtracting one lowpass filter from another to create

$$\begin{aligned} \phi_{[a,b]}^\circ(k) &= \phi_b^\circ(k) - \phi_a^\circ(k) \\[1ex] &= \frac{\cos(\omega_1 k/2)\{\sin(b\omega_1 k) - \sin(a\omega_1 k)\}}{T \sin(\omega_1 k/2)} \\[1ex] &= 2\cos(g\omega_1 k) \frac{\cos(\omega_1 k/2) \sin(d\omega_1 k)}{T \sin(\omega_1 k/2)}. \end{aligned} \tag{68}$$

Here, $2d = b - a$ is the width of the pass band (measured in terms of a number of sampled points) and $g = (a + b)/2$ is the index of its centre. These results have been demonstrated by Pollock (2009).

The relationship between the time-domain and the frequency-domain filtering can be elucidated by considering the matrix representation of the filter, which is obtained by replacing the generic argument $W^j = \exp\{-i\omega_j\}$ in (64) and (65) by the matrix operator $K$. This gives

$$\psi^\circ(K) = \Psi^\circ = \bar{U}\psi^\circ(D)U. \tag{69}$$

Replacing $z$ in the $z$-transform $y(z)$ of the data sequence by $K$ gives the circulant matrix $Y = \bar{U}y(D)U$. Multiplying the data matrix by the filter matrix gives

$$\begin{aligned} x(K) = X = \Psi^\circ Y &= \{\bar{U}\psi^\circ(D)U\}\{\bar{U}y(D)U\} \\ &= \bar{U}\{\psi^\circ(D)y(D)\}U = \bar{U}\psi^\circ(D)UY, \end{aligned} \tag{70}$$

where the final equality follows in view of the identity $UY = y(D)U$.

Here, there is a convolution in the time domain, represented by the product $X = \Psi^\circ Y$ of two circulant matrices, and a modulation in the frequency domain, represented by the product $x(D) = \psi^\circ(D)y(D)$ of two diagonal matrices. Of these matrices, $\psi^\circ(D)$ comprises the zeros and units of the ideal frequency response, whereas $y(D)$ comprises the spectral ordinates of the data.

It should be emphasised that $\psi(\omega)$ and $\psi^\circ(\omega_j)$ are distinct functions of which, in general, the values coincide only at the roots of unity. Strictly speaking, these are the only points for which the latter function is defined. Nevertheless, there may be some interest in discovering the values that $\psi^\circ(\omega_j)$ would deliver if its argument were free to travel around the unit circle.

Figure 10 is designed to satisfy such curiosity. Here, it can be seen that, within the stop band, the function $\psi^\circ(\omega)$ is zero-valued only at the Fourier frequencies. Since the data are compounded from sinusoidal functions with Fourier frequencies, this is enough to eliminate from the sample all elements that fall within the stop band and to preserve all that fall within the pass band.

In Figure 10, the transition between the pass band and the stop band occurs at a Fourier frequency. This feature is also appropriate to other filters that we might design, that have a more gradual transition with a mid point that also falls on a Fourier frequency. Figure 11 shows that it is possible, nevertheless, to make an abrupt transition within the space between two adjacent frequencies.

An alternative way of generating the filtered sequence within the time domain is to create a periodic extension of the data sequence via successive replications of the data. Then, the filtered values can be generated via an ordinary linear convolution of the wrapped filter coefficients of $\psi_0(z)$ with the data. It can be seen that these products are synonymous with the product of the infinite sequence of sinc function coefficients with an infinite periodic extension of the data.

A modulation in the frequency domain, which requires only $T$ multiplications, is a far less laborious operation than the corresponding circular convolution in the time domain, which requires $T^2$ multiplications and $T(T-1)$
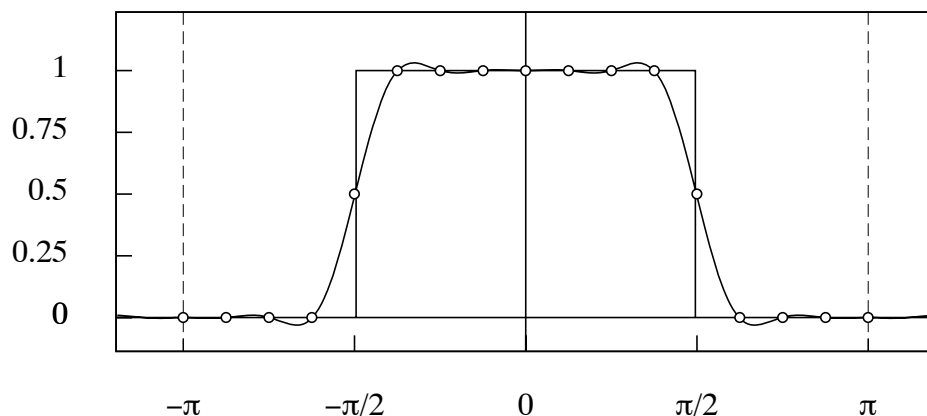
**Figure 10.** The frequency response of the 16-point wrapped filter defined over the interval $[-\pi, \pi)$. The values at the Fourier frequencies are marked by circles and dots.
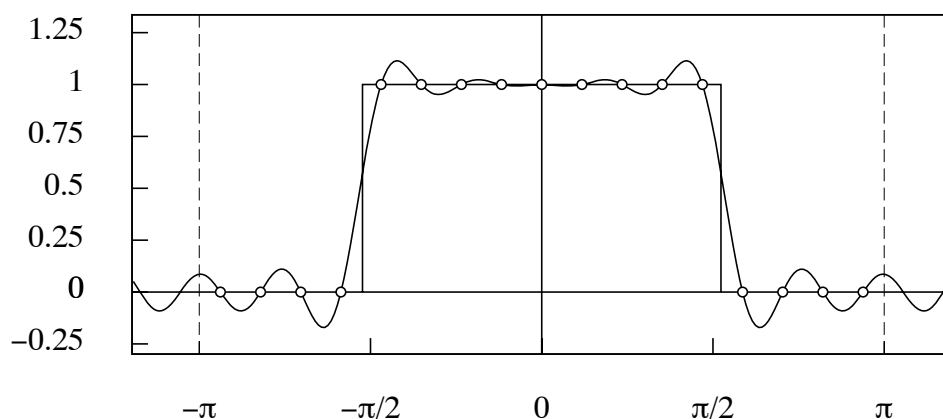


**Figure 11.** The frequency response of the 17-point wrapped filter defined over the interval $[-\pi, \pi)$. The values at the Fourier frequencies are marked by circles.

additions. To be set against the relative efficiency of the frequency-domain operations is the need to carry the data to that domain via a Fourier transform and to carry the products of the modulation back to the time domain via an inverse transformation. However, given the efficiency of the Fast Fourier transform, the advantage remains with the frequency-domain operations. Also, the time-domain filter may have to be formed by applying an inverse Fourier transform to a set of ordinates sampled from the frequency response of the filter.

The ideal lowpass filter has been implemented in a program that can be downloaded from the following web address:

<div align="center">

`http://www.le.ac.uk/users/dsgp1/`

</div>

Various devices are available within the program for dealing with trended data sequences. A trend function can be interpolated through the data to deliver a sequence of residual deviations to which the filter can be applied. Thereafter, the filtered residuals and the trend can be recombined to produce

<div align="center">

29

</div>

a trend–cycle trajectory.

Alternatively, the trend can be eliminated from the data by a twofold application of the difference operator. The filter can be applied to the differenced data and the result can be reinflated via the summation operator, which is the inverse of the difference operator.

The program also uses the method of frequency-domain filtering to create an ideal bandstop filter that is able to eliminate the seasonal component of a trended econometric data sequence.

## Wiener–Kolmogorov Theory

The classical Wiener–Kolmogorov filtering theory concerns a doubly infinite data sequence $y(t) = \{y_t; t = 0, \pm 1, \pm 2, \ldots\}$, which is the sum of a signal sequence $\xi(t)$ and a noise sequence $\eta(t)$:

$$y(t) = \xi(t) + \eta(t). \tag{71}$$

The signal and noise components are assumed to be generated by mutually independent stationary stochastic processes of zero mean. Therefore, the autocovariance generating function of the data, which may be represented by

$$\gamma_y(z) = \gamma_\xi(z) + \gamma_\eta(z), \tag{72}$$

is the sum of the autocovariance generating functions of the components.

On this basis, a filter may be formulated that produces a minimum mean-square-error estimate of the signal. Its transfer function is given by

$$\psi(z) = \frac{\gamma_\xi(z)}{\gamma_\xi(z) + \gamma_\eta(z)}. \tag{73}$$

The infinite-sample filter is a theoretical construct from which a practical implementation must be derived.

A straightforward way of deriving a filter that is applicable to the finite data sequence contained in a vector $y = [y_0, y_1, \ldots, y_{T-1}]'$ is to replace the argument $z$ within the various autocovariance generating functions by the matrix lag operator $L_T = [e_1, \ldots, e_{T-1}, 0]$ and to replace $z^{-1}$ by $F_T = L_T'$. Then, the generating function $\gamma_y(z) = \gamma_0 + \sum_{\tau=1}^{\infty} \gamma_\tau(z^{-\tau} + z^\tau)$ gives rise to matrix

$$\Omega_y = \gamma_0 I_T + \sum_{\tau=1}^{T-1} \gamma_\tau(F_T^\tau + L_T^\tau), \tag{74}$$

where the limit of the summation is when $\tau = T-1$, since $L_T^q = 0$ for all $q \geq T$. This is an ordinary autocovariance matrix. The autocovariance matrices of the components $\Omega_\xi$ and $\Omega_\eta$ are derived likewise, and then there is $\Omega_y = \Omega_\xi + \Omega_\eta$.

Using the matrices in place of the autocovariance generating functions in (72) gives rise to the following estimate $x$ of the signal component $\xi$:

$$x = \Omega_\xi(\Omega_\xi + \Omega_\eta)^{-1}y \tag{75}$$

It is necessary to preserve the order of the factors of this product since, unlike the corresponding $z$-transforms, $\Omega_\xi$ and $\Omega_y^{-1} = (\Omega_\xi + \Omega_\eta)^{-1}$ do not commute in multiplication. Such filters have been discussed in greater detail by Pollock (2000), (2001), where the matter of trended data sequences is also dealt with.

For an alternative finite-sample representation of the filter, we may replace $z$ in (72) and (73) by $K_T = \bar{U}DU$ and $z^{-1}$ by $K_T' = K_T^{T-1} = UD\bar{U}$. Then, the generating functions are replaced by the following matrices of circular autocovariances:

$$\Omega_\xi^\circ = \gamma_\xi(K) = \bar{U}\gamma_\xi(D)U, \quad \Omega_\eta^\circ = \gamma_\eta(K) = \bar{U}\gamma_\eta(D)U$$

$$\text{and} \quad \Omega_y^\circ = \Omega_\xi^\circ + \Omega_\eta^\circ = \gamma_y(K) = \bar{U}\gamma_y(D)U. \tag{76}$$

Putting these matrices in place of the autocovariance matrices of (75) gives rise to the following estimate of the signal:

$$x = \bar{U}\gamma_\xi(D)\{\gamma_\xi(D) + \gamma_\eta(D)\}^{-1}Uy = \bar{U}J_\xi Uy. \tag{77}$$

The estimate may be obtained in the following manner. First, a Fourier transform is applied to the data vector $y$ to give $Uy$. Then, the elements of the transformed vector are multiplied by those of the diagonal weighting matrix $J_\xi = \gamma_\xi(D)\{\gamma_\xi(D) + \gamma_\eta(D)\}^{-1}$. Finally, the products are carried back to the time domain by the inverse Fourier transform.

The same method is indicated by equation (70), which represent the application of the ideal filter. When it is post multiplied by $e_0$, the latter delivers the expression $x = \bar{U}\psi^\circ(D)Uy$, which can also be given the foregoing interpretation.

## 6. The Processes Underlying the Data

In this section, we shall discuss the relationship between the sampled data and the processes that generate it, which are presumed to operate in continuous time.

If the continuous process contains no sinusoidal elements with frequencies in excess of the Nyquist value of $\pi$ radians, then a perfect reconstruction of the underlying function can be obtained from its sampled values by using the sinc function $\sin(\pi t)/\pi t$ of Figure 3 as an interpolator. To effect the interpolation, each of the data points is associated with a sinc function scaled by the sampled value. These functions, which are set at unit displacements along the real line, are added to form a continuous envelope.

The effect of a sinc function at any integer point $t$ is to disperse the mass of the sampled value over the adjacent points of the real line. The function has a value of unity at $t$ and a value of zero on all other integer points. Therefore, in the process of interpolation, it adds nothing to the values at the adjacent integer points. However, it does attribute values to the non-integer points that lie in the interstices, thereby producing a continuous function from discrete data. Moreover, the data can be recovered precisely by sampling the continuous function at the integers.

Only if the sampled data is free of aliasing, will this process of interpolation succeed in recovering the original continuous function. The question of whether aliasing has occurred is an empirical one that relates to the specific nature of the process underlying the data. However, the models of continuous-time processes that are commonly adopted by econometricians and other statisticians suggest that severe aliasing is bound to occur.

The discrete-time white-noise process $\varepsilon(t)$ that is the *primum mobile* of linear stochastic processes of the ARMA and ARIMA varieties is commonly thought to originate from sampling a Wiener process $W(t)$ defined by the following conditions:

(a) $W(0) = 0$,

(b) $E\{W(t)\} = 0$, for all $t$,

(c) $W(t)$ is normal,

(d) $dW(s), dW(t)$ for all $t \neq s$ are independent stationary increments,

(e) $\varepsilon(t) = \dfrac{1}{h}\{W(t+h) - W(t)\}$ for some $h > 0$.

The increments $dW(s), dW(t)$ are impulses that have a uniform power spectrum distributed over the entire real line, with infinitesimal power in any finite interval. Sampling $W(t)$ at regular intervals entails a process of aliasing whereby the spectral power of the cumulated increments gives rise to a uniform spectrum of finite power over the interval $[-\pi, \pi]$.

An alternative approach is to imagine that the discrete-time white noise is derived by sampling a continuous-time process that is inherently limited in frequency to the interval $[-\pi, \pi]$. The sinc function of Figure 3, which is limited in frequency to this interval, can also be construed as a wave packet centered on time $t = 0$. A continuous stochastic function that is band-limited by the Nyquist frequency of $\pi$ will be generated by the superposition of such functions arriving at regular or irregular intervals of time and having random amplitudes. No aliasing would occur in the sampling of such a process.

(The wave packet is a concept from quantum mechanics—see, for example Dirac (1958) and Schiff (1981)—that has become familiar to statisticians via wavelet analysis—see, for example, Daubechies (1992) and Percival and Walden (2000).)

**Band-Limited Processes**

It is implausible to assume that a Wiener process could be the *primum mobile* of a stochastic process that is band limited within a sub interval of the (positive) frequency range $[0, \pi]$. Such processes appear to be quite common amongst the components of econometric time-series.

Figure 12 displays a sequence of the logarithms of the quarterly series of U.K. Gross Domestic Product (GDP) over the period from 1970 to 2005. Interpolated through this sequence is a quadratic trend, which represents the growth path of the economy.
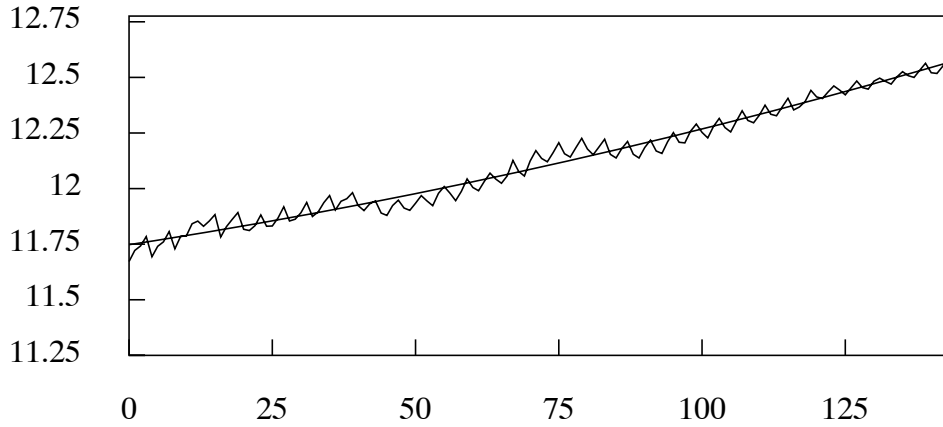
**Figure 12.** The quarterly sequence of the logarithms of the GDP in the U.K. for the years 1970 to 2005, inclusive, together with a quadratic trend interpolated by least squares regression.
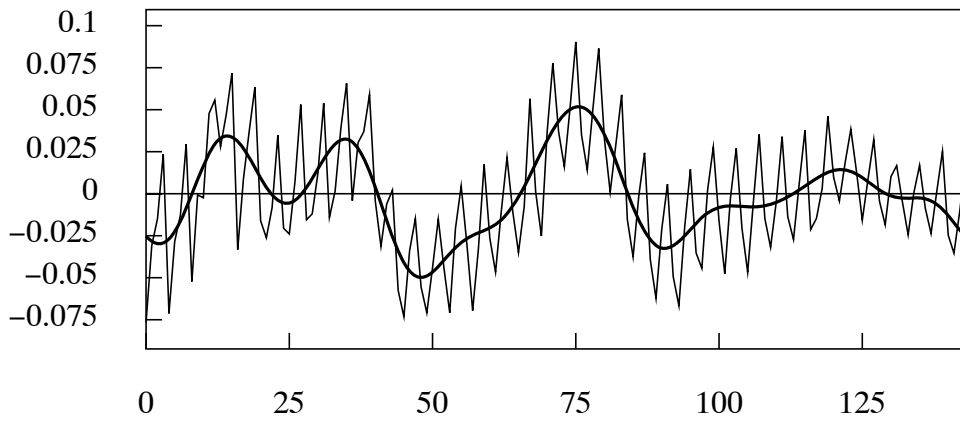


**Figure 13.** The residual sequence from fitting a quadratic trend to the income data of Figure 12. The interpolated line represents the business cycle.
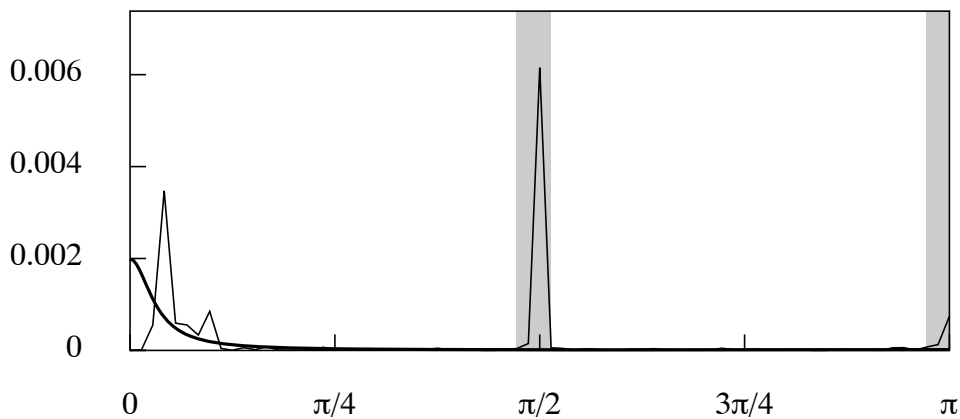


**Figure 14.** The periodogram of the residuals obtained by fitting a quadratic trend through the logarithmic sequence of U.K. GDP. The parametric spectrum of a fitted AR(2) model is superimposed upon the periodogram.
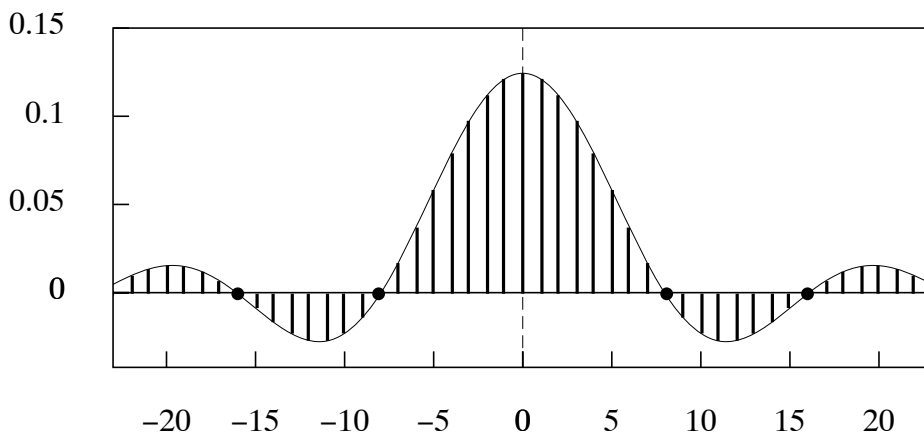
**Figure 15.** The sinc function wave-packet $\psi_3(t) = \sin(\pi t/8)/\pi t$ comprising frequencies in the interval $[0, \pi/8]$.

The deviations from this growth path are a combination of the low-frequency business cycle with the high-frequency fluctuations that are due to the seasonal nature of economic activity. These deviations are represented in Figure 13, which also shows an interpolated continuous function that is designed to represent the business cycle.

The periodogram of the deviations is shown in Figure 14. This gives a clear indication of the separability of the business cycle and the seasonal fluctuations. The spectral structure extending from zero frequency up to $\pi/8$ belongs to the business cycle. The prominent spikes located at the frequency $\pi/2$ and at the limiting Nyquist frequency of $\pi$ are the property of the seasonal fluctuations. Elsewhere in the periodogram, there are wide dead spaces, which are punctuated by the spectral traces of minor elements of noise.

The slowly varying continuous function interpolated through the deviations of Figure 13 has been created by combining a set of sine and cosine functions of increasing frequencies in the manner of equation (1), but with the summation extending no further than the limiting frequency of the business cycle, which is $\pi/8$.

A process that is band limited to the frequency interval $[0, \pi/8]$ would be derived from the random arrival of wave packets of random amplitudes, likewise band limited in frequency. The sinc function is but one example such a packet; and it has the disadvantage of being widely dispersed in time. Its advantage is that it provides an orthogonal basis for the set of band-limited functions. The basis set is

$$\{\psi_3(t - 8k); k = 0, \pm 1, \pm 2, \ldots\} \quad \text{with} \quad \psi_3(t) = \frac{\sin(\pi t/8)}{\pi t}, \qquad (78)$$

wherein the functions are displaced, from one to the next, by 8 sample points. (Here the subscript on $\psi_3$ is in reference to the fact that $8 = 2^3$ and that the interval $[0, \pi/8]$ arises from the 3rd dyadic subdivision of $[0, \pi]$.) The function at $t = 0$ is represented in Figure 15. Whereas a continuous band-limited stochastic function would be constituted by adding the envelopes of the succession of

34

continuous sinc functions, its sampled version would be obtained by adding the sampled ordinates that are also represented in Figure 15.

A further point concerns the autocovariances of a band-limited process. The autocovariance function of an ordinary linear stochastic process of the ARMA variety is positive definite, which corresponds to the nonzero property of the spectral density function. Likewise, the matrix of autocovariances is positive definite, which implies that all of its eignenvalues are positive. Therefore, it is not possible to represent a band-limited process by such means.

However, it is straightforward to represent the autocovariances of a band-limited process via a circulant matrix of the form

$$\gamma(K) = \Gamma = \bar{U}\gamma(D)U. \tag{79}$$

Here, the diagonal matrix $\gamma(D)$ represents the spectral density of the process, sampled at $T$ frequency points. The band-limited nature of the process will be reflected by the zero-valued elements of the diagonal of $\gamma(D)$ and by the singularity of $\Gamma$.

## Fitting ARMA Models to the Business Cycle

One might wish to fit a parametric model to the detrended data of Figure 13. For this purpose, the multiplicative seasonal ARIMA model of Box and Jenkins (1976) seems, at first, to be a natural choice. Such a model embodies a seasonal difference operator of the form $(1 - z^4) = (1 - z)(1 + z + z^2 + z^3)$, of which the ordinary difference operator $(1 - z)$ is a factor. However, the data of Figure 13, have already been detrended. Therefore, the presence or the redundant difference operator renders the model inappropriate. A more appropriate recourse is to remove the seasonal component from the data before fitting an autoregressive model aimed at representing the business cycle component.

The seasonal component can be removed from the data of Figure 13 by nullifying the spectral ordinates at the seasonal frequencies and at some adjacent frequencies. This operation, which can be performed by the program mentioned in Section 5, is represented in Figure 14 by the highlighted bands that correspond to the stop bands of the deseasonalising filter. Also represented in Figure 14 is the parametric spectrum of a second-order autoregressive AR(2) model that has been fitted to the seasonally adjusted data,

One would expect to be able to characterise the low-frequency dynamics of the business cycle component via a pair of conjugate complex autoregressive roots to be found within the estimated model. The arguments of such roots should characterise the angular velocity of the cycle; and their moduli should represent the degree of its damping. However, it is found that the roots in question are real-valued. It has been reported by Pagan (1997), amongst others, that this is the almost invariable outcome of such exercises.

We might wish, nevertheless, to use an ARMA model to characterise the business cycle component that is supported on the interval $[0, \pi/8]$ and which is represented by the smooth continuous function that has been interpolated through the data of Figure 13. In general, when we wish to fit an ARMA model to a component of the data that is supported on a frequency interval $[0, \omega_d]$,

where $\omega_d < \pi$, it is necessary to map the component onto the full frequency interval $[0, \pi]$. For this purpose, it may be necessary to synthesise a continuous function from the Fourier ordinates within the subinterval $[0, \omega_d]$. The function can then be resampled at the lesser rate of one observation in every $\pi/\omega_d$ periods to produce data that are appropriate to the purpose.

In the case of the business cycle component of Figure 13, which has a spectral signature that falls within the interval $[0, \pi/8]$, it is sufficient to take one in eight of the data points of the anti-aliased sequence that has been obtained by subjecting the original data to an ideal lowpass filter with a cut-off at $\pi/8$. The periodogram of the resulting subsampled sequence, which has the frequency range of $[0, \pi]$, will have a profile identical to that of the spectral of structure that occupies the interval $[0, \pi/8]$ in Figure 14.

The AR(2) model that fits these data has a pair of conjugate complex roots with an argument of $\theta = 1.35$ radians ($77°$), which corresponds to a cyclical period of 37 quarters, and a modulus of $\rho = 0.695$. However, the evidence of the interpolated function of Figure 13 and of the periodogram of Figure 14 suggests that the estimated period is an average of the durations longer major cycle and a shorter minor cycle. However, with only 18 points available in the subsampled data, it seems that there is insufficient information to justify the presentation of the AR(4) model that reflects these putative cycles in two pairs of conjugate complex roots.

## The Recognition of the Frequency Limitation

Some reasons should be given for why the prevalence of band limited processes has been overlooked by econometricians. One evident reason is that, in the past, econometricians have rarely examined the periodograms of their data sequences.

However, the band limited nature of a process is liable to be obscured whenever there are significant disjunctions in the periodic extension of the data at the points where the end of one replication of the sample joins the beginning of the next. Such disjunctions will give rise to a slew of spectral power that will serve to conceal the underlying structure of a band-limited process.

The presence of a barely perceptible trend in the data will be enough to cause such a disjunction. The periodic extension of a finite linearly trending sequence will have the appearance of the serrated edge of a carpenter's saw. The spectrum of a saw tooth function has the *one-over-f* profile of a rectangular hyperbola, which will mask the finer spectral details. (See Hamming 1989, for example.) Granger (1966) has described such spectra as typical of econometric time series. His characterisation has greatly influenced the perceptions of econometricians. It has suggested to many of them that the spectral analysis of economic data is liable to be difficult if not fruitless.

Several authors, including Granger and Hyung (2004), have observed that structural breaks within the processes generating the data can give rise to slowly decaying autocorrelations and to other properties of long-memory processes, including *one-over-f* periodograms that decay hyperbolically as the frequency rises. Such breaks are not dissimilar to the end-of-sample disjunctions which are

considered here. However, the latter seem to provide a more fruitful explanation for the occurrence of misleading long-memory properties.

A third reason for the failure to perceive the band-limited nature of some of these process may be that econometricians have been unduly influenced by the prevalence of the Wiener process as a model of the forcing functions of continuous-time stochastic processes. A Wiener process, which has a frequency band of infinite width, is the antithesis of a band-limited process.

Within the computer program that is an adjunct of this paper, there are several provisions for ensuring that the data can be wrapped seamlessly around the circumference of a circle in a manner that avoids disjunctions in the periodic extension.

If the data are to be detrended by the interpolation of a polynomial function, then extra weight can be given to the points at the beginning and the end of the sample to ensure that the polynomial passes through their midst. This may have the desired effect upon the residual sequence, which will be used thereafter in the analysis.

If the data have already been detrended, then it may be desirable to add tapered extrapolations at both ends which converge on a common asymptote. The extrapolations are liable to be discarded after the data have been filtered. Alternatively, if the detrended data display prominent seasonal fluctuations of a pattern that is changing over the course of the sample, then a segment may be interpolated into the circular data in which the pattern of the final year of the sample gradually morphs into the pattern of the first year.

## References

Artech, J., and P.M. Robinson, (2000), Semiparametric Inference in Seasonal and Cyclical Long Memory Processes, *Journal of Time Series Analysis,* 21, 1–25.

Baher, H., (1990), *Analog and Digital Signal Processing,* John Wiley and Sons, Chichester.

Beran, J., (1994), *Statistics for Long-memory Processes,* Chapman and Hall, London.

Box, G.E.P., and G.M. Jenkins, (1976), *Time Series Analysis: Forecasting and Control: Revised Edition,* Holden-Day, San Francisco.

Brigham, E.O., (1988), *The Fast Fourier Transform and its Applications,* Prentice-Hall, Englewood Cliffs, New Jersey.

Brillinger, D.R., (1975), *Time Series: Data Analysis and Theory,* Holt, Rinehart and Winston, New York.

Brockwell, P.J., and R.A. Davis, (1987), *Time Series: Theory and Methods,* Springer-Verlag, New York.

Carslaw, H.S., (1925), A Historical Note on Gibbs' Phenomenon in Fourier Series and Integrals, *Bulletin of the American Mathematical Society,* 31, 420–424.

Carslaw, H.S., (1930), *Introduction to the Theory of Fourier's Series and Integrals,* Third Edition, Macmillan, London.

Catalan, E., (1846), Récherches sur les Déterminants, *Bulletin de l'Academie Royale de Belgique,* 13, 534–555.

Churchill, R.V., and J.W. Brown, (1978), *Fourier Series and Boundary Value Problems,* Third edition, McGraw-Hill, New York.

Daubechies, I., (1992), *Ten Lectures on Wavelets,* Society for Industrial and Applied Mathematics, Philadelphia.

Davis, P.J., (1979), *Circulant Matrices,* John Wiley and Sons, New York.

Dirac, P.A.M., (1958), *The Principles of Quantum Mechanics, Fourth Edition,* Oxford University Press, Oxford.

Fishman, G.S., (1969), *Spectral Methods in Econometrics,* Harvard University Press, Cambridge, Massachusetts.

Fuller, W.A., (1976), *Introduction to Statistical Time Series,* John Wiley and Sons, New York.

Granger, C.W.J, (1966), The Typical Spectral Shape of an Economic Variable, *Econometrica,* 34, 150–161.

Granger, C.W.J., and Z. Ding, (1996), Varieties of Long Memory Models, *Journal of Econometrics,* 73, 61–77.

Granger, C.W.J., and M. Hatanaka, (1964), *Spectral Analysis of Economic Time Series,* Princeton University Press, Princeton, New Jersey.

Granger, C.W.J., and N. Hyung, (2004), Occasional Structural Breaks and Long Memory with an Application to the S&P 500 Absolute Stock Returns, *Journal of Empirical Finance,* 3, 399–421.

Granger, C.W.J., and R. Joyeux, (1980), An Introduction to Long-memory Time Series and Fractional Differencing, *Journal of Time Series Analysis,* 1, 15–30.

Gray, R.M., (2002), *Toeplitz and Circulant Matrices: A Review,* Information Systems Laboratory, Department of Electrical Engineering, Stanford University, Stanford, California, `rmgray@stanford.edu`.

Gray, H.L., N.F. Zhang, and W.A. Woodward, (1989), On Generalized Fractional Processes, *Journal of Time Series Analysis,* 10, 233–57.

Grenander, U., and M. Rosenblatt, (1957), *Statistical Analysis of Stationary Time Series,* John Wiley and Sons, New York.

Hamilton, J.D., (1994), *Time Series Analysis,* Princeton University Press, Princeton, New Jersey.

Hamming, R.W., (1989), *Digital Filters: Third Edition,* Prentice-Hall, Englewood Cliffs, New Jersey.

Hannan, E.J., (1960), *Time-series Analysis,* Methuen, London.

Hidalgo, J., (1997), Spectral Analysis for Bivariate Time Series with Long Memory, *Econometric Theory,* 12, 773–792.

Hosking J.R.M., (1981), Fractional Differencing, *Biometrika,* 68, 165–176.

Jenkins, G.M., and D.G. Watts, (1968), *Spectral Analysis and its Applications,* Holden-Day, Oakland, California.

Kolmogorov, A.N., (1941a), Interpolation und Extrapolation von stationären zuffaligen Folgen, *Bulletin de l'academie des sciences de U.S.S.R.,* Ser. Math., 5, 3–14.

Kolmogorov, A.N., (1941b) Stationary Sequences in a Hilbert Space, (in Russian), *Bulletin of the Moscow State University,* 2, 1–40.

Kreider, D.L., R.G. Kuller, D.R. Osterberg and F.W. Perkins, (1966), *An Introduction to Linear Analysis,* Addison Wesley Publishing Co., Reading, Massachusetts.

Maravall A., and D.A. Pierce, (1987), A Prototypical Seasonal Adjustment Model, *Journal of Time Series Analysis,* 8, 177–193.

McCoy, E.J., and D.A. Stephens, (2004), Bayesian Time Series Analysis of Periodic Behaviour and Spectral Structure, *International Journal of Forecasting,* 20, 713–730.

Muir, T., (1906), The Theory of Circulants in the Historical Order of Development up to 1860, *Proceedings of the Royal Society of Edinburgh,* 26, 390–398.

Muir, T., (1911), The Theory of Circulants from 1861 to 1880, *Proceedings of the Royal Society of Edinburgh,* 32, 136–149.

Muir, T., (1915), The Theory of Circulants from 1880 to 1900, *Proceedings of the Royal Society of Edinburgh,* 36, 151–179.

Muir, T., (1923), The Theory of Circulants from 1900 to 1920, *Proceedings of the Royal Society of Edinburgh,* 44, 218–241.

Nerlove, M., D.M. Grether and J.L. Carvalho, (1979), *Analysis of Economic Time Series: a Synthesis,* Academic Press, New York.

Pagan, A., (1997), Towards an Understanding of Some Business Cycle Characteristics, *The Australian Economic Review,* 30, 1–15.

Palma, W., (2007), *Long-Memory Time Series: Theory and Methods,* John Wiley and Sons, Hoboken, New Jersey.

Percival, D.B., and A.T. Walden, (2000), *Wavelet Methods for Time Series Analysis,* Cambridge University Press, Cambridge.

Pollock, D.S.G., (1999), *A Handbook of Time-Series Analysis, Signal Processing and Dynamics,* Academic Press, London.

Pollock, D.S.G., (2000), Trend Estimation and De-Trending via Rational Square-Wave Filters, *Journal of Econometrics,* 99, 317–334.

Pollock, D.S.G., (2001), Filters for Short Non-stationary Sequences, *Journal of Forecasting,* 20, 341–355.

Pollock, D.S.G., (2008), Investigating Economic Trends and Cycles, Chapter 3.1 in T.C. Mills and K. Patterson (eds.), *Palgrave Handbook of Econometrics: Vol. 2 Applied Econometrics,* Palgrave Macmillan, Basingstoke.

Pollock, D.S.G., (2009), The Realisation of Finite-Sample Frequency-Selective Filters, *Journal of Statistical Planning and Inference,* 139, 1541–1558.

Priestley, M.B., (1981), *Spectral Analysis and Time Series,* Academic Press, London.

Rosenblatt, M., (1985), *Stationary Sequences and Random Fields,* Birkhàuser, Basel.

Rozanov, Y.A., (1967), *Stationary Random Processes,* Holden-Day, San Francisco.

Schiff, L.I., (1981), *Quantum Mechanics,* McGraw-Hill Book Co., New York.

Stade, E., (2005), *Fourier Analysis,* John Wiley and Sons, Hoboken, New Jersey.

Titchmarsh, E.C., (1937), *Introduction to the Theory of Fourier Integrals,* Clarendon Press, Oxford.

Wiener, N., (1930), Generalised Harmonic Analysis, *Acta Mathematica,* 55, 117–258.

Wiener, N., (1941), *Extrapolation, Interpolation and Smoothing of Stationary Time Series.* Report on the Services Research Project DIC-6037. Published in book form in 1949 by MIT Technology Press and John Wiley and Sons, New York.

Yaglom, A.M., (1962), *An Introduction to the Theory of Stationary Random Functions,* Prentice-Hall, Englewood Cliffs, New Jersey.

### Appendix: The Spectral Representation of a Stationary Process

Consider a sample $\{y_0, y_1, \ldots, y_{T-1}\}$ of $T$ observations on a stationary stochastic process. These are amenable to a discrete Fourier transform, whereby they can be expressed as

$$y_t = \sum_{j=0}^{n} \left\{ \alpha_j \cos(\omega_j t) + \beta_j \sin(\omega_j t) \right\}, \qquad (A.1)$$

where $\omega_j = 2\pi j/T$ and $n$ is the integral part of $T/2$.

By allowing $n$ to tend to infinity, it is possible to express a sequence of indefinite length in terms of a sum of sine and cosine functions. However, in the limit as $n \to \infty$, the coefficients $\alpha_j, \beta_j$ tend to vanish; and, therefore, an alternative representation in terms of differentials is called for.

By writing $\alpha_j = dA(\omega_j)$, $\beta_j = dB(\omega_j)$, where $A(\omega)$, $B(\omega)$ are step functions with discontinuities at the points $\{\omega_j; j = 0, \ldots, n\}$, the expression (A.1) can be rendered as

$$y_t = \sum_j \left\{ \cos(\omega_j t) dA(\omega_j) + \sin(\omega_j t) dB(\omega_j) \right\}. \tag{A.2}$$

In the limit, as $n \to \infty$, the summation is replaced by an integral to give the expression

$$y(t) = \int_0^\pi \left\{ \cos(\omega t) dA(\omega) + \sin(\omega t) dB(\omega) \right\}. \tag{A.3}$$

Here, $\cos(\omega t)$ and $\sin(\omega t)$, and therefore $y(t)$, may be regarded as infinite sequences defined over the entire set of positive and negative integers.

Since $A(\omega)$ and $B(\omega)$ are discontinuous functions for which no derivatives exist, one must avoid using $\alpha(\omega) d\omega$ and $\beta(\omega) d\omega$ in place of $dA(\omega)$ and $dB(\omega)$. Moreover, the integral in (A.3) is a Fourier–Stieltjes integral.

In order to derive a statistical theory for the process that generates $y(t)$, one must make some assumptions concerning the functions $A(\omega)$ and $B(\omega)$. So far, the sequence $y(t)$ has been interpreted as a realisation of a stochastic process. If $y(t)$ is regarded as the stochastic process itself, then the functions $A(\omega)$, $B(\omega)$ must, likewise, be regarded as stochastic processes defined over the interval $[0, \pi]$. A single realisation of these processes now corresponds to a single realisation of the process $y(t)$.

The first assumption to be made is that the functions $dA(\omega)$ and $dB(\omega)$ represent a pair of stochastic processes of zero mean, which are indexed on the continuous parameter $\omega$. Thus

$$E\{dA(\omega)\} = E\{dB(\omega)\} = 0. \tag{A.4}$$

The second and third assumptions are that the two processes are mutually uncorrelated and that non-overlapping increments within each process are uncorrelated. Thus

$$\begin{aligned} E\{dA(\omega)dB(\lambda)\} &= 0 \quad \text{for all} \quad \omega, \lambda, \\ E\{dA(\omega)dA(\lambda)\} &= 0 \quad \text{if} \quad \omega \neq \lambda, \\ E\{dB(\omega)dB(\lambda)\} &= 0 \quad \text{if} \quad \omega \neq \lambda. \end{aligned} \tag{A.5}$$

The final assumption is that the variance of the increments is given by

$$\begin{aligned} V\{dA(\omega)\} = V\{dB(\omega)\} &= 2dF(\omega) \\ &= 2f(\omega)d\omega. \end{aligned} \tag{A.6}$$

It can be seen that, unlike $A(\omega)$ and $B(\omega)$, $F(\omega)$ is a continuous function that is differentiable except, at most, on a set of measure zero. The function $F(\omega)$ and its derivative $f(\omega)$ are the spectral distribution function and the spectral density function, respectively.

In order to express (A.3) in terms of complex exponentials, we may define a pair of conjugate complex stochastic processes:

$$
\begin{aligned}
dZ(\omega) &= \frac{1}{2}\big\{dA(\omega) - idB(\omega)\big\}, \\
dZ^*(\omega) &= \frac{1}{2}\big\{dA(\omega) + idB(\omega)\big\}.
\end{aligned}
\tag{A.7}
$$

Also, the domain of the functions $A(\omega)$, $B(\omega)$ may be extended from $[0, \pi]$ to $[-\pi, \pi]$ by regarding $A(\omega)$ as an even function such that $A(-\omega) = A(\omega)$ and by regarding $B(\omega)$ as an odd function such that $B(-\omega) = -B(\omega)$. Then there is

$$
dZ^*(\omega) = dZ(-\omega).
\tag{A.8}
$$

From the conditions of (A.5), it follows that

$$
\begin{aligned}
E\big\{dZ(\omega)dZ^*(\lambda)\big\} &= 0 \quad \text{if} \quad \omega \neq \lambda, \\
E\big\{dZ(\omega)dZ^*(\omega)\big\} &= f(\omega)d\omega.
\end{aligned}
\tag{A.9}
$$

These results may be used to reexpress (A.3) as

$$
\begin{aligned}
y(t) &= \int_0^\pi \left\{ \frac{(e^{i\omega t} + e^{-i\omega t})}{2}dA(\omega) - i\frac{(e^{i\omega t} - e^{-i\omega t})}{2}dB(\omega) \right\} \\
&= \int_0^\pi \left\{ e^{i\omega t}\frac{\{dA(\omega) - idB(\omega)\}}{2} + e^{-i\omega t}\frac{\{dA(\omega) + idB(\omega)\}}{2} \right\} \\
&= \int_0^\pi \left\{ e^{i\omega t}dZ(\omega) + e^{-i\omega t}dZ^*(\omega) \right\}.
\end{aligned}
\tag{A.10}
$$

When the integral is extended over the range $[-\pi, \pi]$, this becomes

$$
y(t) = \int_{-\pi}^\pi e^{i\omega t}dZ(\omega),
\tag{A.11}
$$

which is commonly described as the spectral representation of the process $y(t)$.