University of Leicester

DEPARTMENT OF ECONOMICS

REALISATIONS OF FINITE-SAMPLE FREQUENCY-SELECTIVE FILTERS

D.S.G. Pollock, University of Leicester, UK

Working Paper No. 08/32
September 2008

Realisations of Finite-Sample Frequency-Selective Filters

D.S.G. Pollock

Department of Economics, University of Leicester, LE1 7RH, UK

Abstract

A filtered data sequence can be obtained by multiplying the Fourier ordinates of the data by the ordinates of the frequency response of the filter and by applying the inverse Fourier transform to carry the product back to the time domain. Using this technique, it is possible, within the constraints of a finite sample, to design an ideal frequency-selective filter that will preserve all elements within a specified range of frequencies and that will remove all elements outside it.

Approximations to ideal filters that are implemented in the time domain are commonly based on truncated versions of the infinite sequences of coefficients derived from the Fourier transforms of rectangular frequency response functions. An alternative to truncating an infinite sequence of coefficients is to wrap it around a circle of a circumference equal in length to the data sequence and to add the overlying coefficients.

The coefficients of the wrapped filter can also be obtained by applying a discrete Fourier transform to a set of ordinates sampled from the frequency response function. Applying the coefficients to the data via circular convolution produces results that are identical to those obtained by a multiplication in the frequency domain, which constitutes a more efficient approach.

Key words: Signal extraction, Linear filtering, Frequency-domain analysis

1 Introduction: The Problem of the Ideal Filter

Recently, business cycle analysts have become interested in extracting, from macroeconomic indices, data components that fall within specified intervals of the frequency spectrum. Examples are to be found in the papers of Baxter and King (1999), Christiano and Fitzgerald (2003) and Iacobucci and Noullez (2005). In particular, Baxter and King have proposed that, according

Email address: stephen_pollock@sigampi.u-net.com (D.S.G. Pollock).

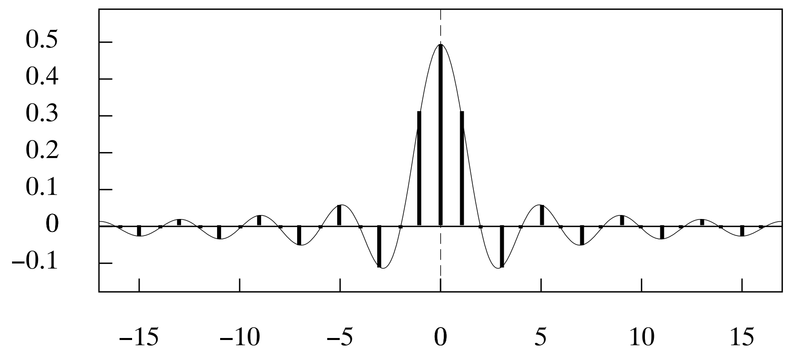

Fig. 1. The central coefficients of the Fourier transform of the frequency response of an ideal lowpass filter with a cut-off point at $\omega = \pi/2$. The sequence of coefficients extends indefinitely in both directions.

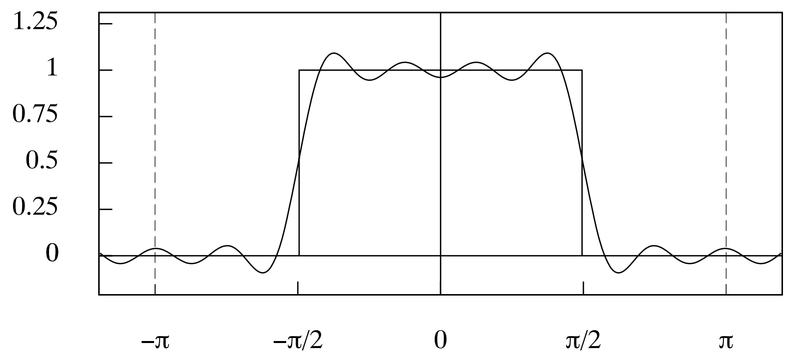

Fig. 2. The frequency response of a filter obtained by applying a 17-point rectangular window to the coefficients of an ideal lowpass filter with a cut-off point at $\omega = \pi/2$, superimposed upon the frequency response of the ideal filter.

to the definition of Burns and Mitchell (1946), the business cycles should comprise cyclical elements with durations of no less than 18 months and of no more than 8 years.

It is commonly believed that, in the case of finite-length samples, it is impossible to design a filter that will preserve completely all elements within a specified range of frequencies and that will remove all elements outside it. A filter that would achieve such an objective is described as an ideal filter.

This belief is based on the fact that, when a (classical) Fourier transform is applied to a periodic square wave or boxcar function, representing the ideal frequency response of the filter, the result is a symmetric doubly-infinite sequence of filter coefficients. To obtain a practical filter, it seems that one must truncate the sequence, retaining only a limited number of its central elements (Figure 1).

The truncation gives rise to a filter of which the frequency response has certain 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

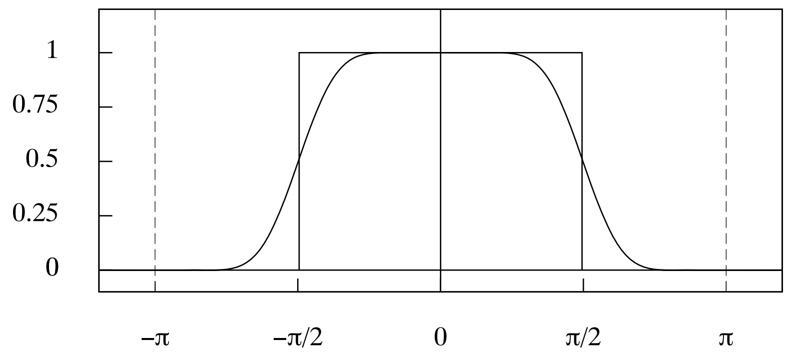

Fig. 3. The frequency response of a filter obtained by applying a 17-point Blackman window to the coefficients of an ideal lowpass filter with a cut-off point at $\omega = \pi/2$.

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 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 3).

The purpose of this paper is to show that none of the above-mentioned problems need afflict the filtering of finite data sequences. It shows that an alternative to truncating the filter is to wrap the infinite sequence of coefficients around a circle of a circumference T equal to the length of the data sequence. The overlying coefficients are added to give the coefficients of the wrapped filter.

It is impractical to perform the operation of filter wrapping in the time domain by summing the infinite sequences of the overlying coefficients directly. Instead, one may resort to the equivalent operation of sampling the frequency response function of the filter at T equally-spaced points. The coefficients of the wrapped filter may be obtained by applying a discrete Fourier transform to this frequency-domain sample to carry its effects into the time domain.

The wrapped filter can be applied, via an ordinary convolution, to a periodic extension of the data sequence. Alternatively, it can be applied, via circular convolution, to the ordinary data sequence, with the same results. However, such a convolution can be realised most effectively via an equivalent modulation in the frequency domain of the Fourier transform of the data, followed by an inverse Fourier transform to carry the results back to the time domain.

This implementation of an ideal filter is but one instance of a general approach to the problem of finite-sample filter design, which we shall expound in this paper, that recognises the finite nature of the data sample at the outset. Other approaches begin, in effect, with the assumption of a doubly-infinite sample; and then they makes amends for the fact that the sample is finite by resorting to a variety of ingenious adaptations.

In order to pursue the circular approach successfully, it is necessary to detrend the data and to ensure that there are no radical disjunctions in the periodic extension of the detrended data where the end of one replication of the sequence meets the beginning of the next replication.

The data can be detrended by differencing. Once the relevant components have been extracted from the differenced data, the corresponding components of the trended data can be recovered by a process of anti-differencing, or cumulation, that requires some initial conditions. These are readily available. In the case of highpass or bandpass filtering, the cumulation process can be avoided. If the cumulation operator is cancelled with the differencing operator that is embodied by the filter, then a reduced filter is derived that will deliver the required product directly.

This paper has a frequency-domain orientation. Reference to the frequency domain is becoming increasingly common amongst statisticians and econometricians. Thus, for example, Haywood and Tunnicliffe Wilson (1997) and Proietti (2005) have recently devised modified lowpass filters in reference to their effects in the frequency domain.

There is clear evidence that the central statistical agencies, which are responsible for producing seasonally adjusted data series and for estimating the trends in official statistics, are placing increasing emphasis in the frequency domain. Examples from the U.S. Census Bureau are provided by the recent papers of Findley and Martin (2003) and of Bell and Martin (2004).

Amongst Europeans, the SEATS–TRAMO program for the canonical analysis of unobserved components in time series has been influential in fostering a growing awareness of the frequency domain (see Caporello and Maravall 2004).

The procedures that are described in this paper have been incorporated in a new computer program, which implements various filters that can be used for extracting the components of an economic data sequence and for producing smoothed and seasonally-adjusted data from monthly and quarterly sequences. The program can be downloaded from the following web address:

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

It is accompanied by a collection of data and by various log files, which record steps that can be taken in processing some typical economic data.

2 Approximations to the Ideal Filter

The theory underlying the spectral analysis of statistical time series deals preponderantly with sequences that are defined over the entire set of positive and negative integers. Such a sequence, which may be denoted by $y(t) = \{y_t; t = 0, \pm 1, \pm 2, \ldots\}$, can be described as a linear combination of trigonometrical functions of which the frequencies, denominated in radians per sampling interval, range for zero to the limiting Nyquist value of π.

An infinite sequence generated by a stationary stochastic process is liable to be expressed as a weighted integral of a non denumerable set of sines and cosines indexed by a frequency value ω that varies continuously within the interval $[0, \pi]$. Since

$$\cos(\omega t) = \frac{1}{2}\left(e^{\mathrm{i}\omega t} + e^{-\mathrm{i}\omega t}\right) \quad \text{and} \quad \sin(\omega t) = \frac{-\mathrm{i}}{2}\left(e^{\mathrm{i}\omega t} - e^{-\mathrm{i}\omega t}\right), \tag{1}$$

the value generated at time t can also be expressed as a weighted integral over the interval $[-\pi, \pi]$ of a complex exponential function $\exp\{\mathrm{i}\omega t\}$:

$$y_t = \int_{-\pi}^{\pi} e^{\mathrm{i}\omega t} dZ(\omega). \tag{2}$$

Here, the complex element $dZ(\omega)$, which constitutes the stochastic weighting function, represents the infinitesimal increments of a cumulative function $Z(\omega)$ that is everywhere continuous but nowhere differentiable. The expectation of the squared modulus of $dZ(\omega)$ constitutes an increment of the cumulative spectrum: $dF(\omega) = E\{dZ(\omega)dZ^*(\omega)\}$. In the case of a purely stochastic process, the cumulative spectrum $F(\omega)$ is an analytic function of which the derivative $f(\omega)$ is described as the spectral density function or the "spectum".

A time-invariant linear filter forms a weighted combination of adjacent elements of the sequence $y(t)$. The filter is defined by the sequence of these weights or filter coefficients, which is the impulse response of the filter. Its effect can also be represented by the manner in which it alters the sinusoidal elements of which $y(t)$ is composed.

Mapping a (doubly-infinite) cosine sequence $y(t) = \cos(\omega t)$, of a given frequency ω, through a filter defined by the coefficients $\{\phi_k\}$ produces the output

$$\begin{aligned} x(t) &= \sum_k \phi_k \cos(\omega[t-k]) \\ &= \sum_k \phi_k \cos(\omega k)\cos(\omega t) + \sum_k \phi_k \sin(\omega k)\sin(\omega t) \\ &= \alpha\cos(\omega t) + \beta\sin(\omega t) = \lambda\cos(\omega t - \theta), \end{aligned} \tag{3}$$

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

The effect of the filter is to alter the amplitude of the cosine via the gain factor λ and to induce a delay that corresponds to the phase angle θ. It is apparent that, if the filter is symmetric about the central coefficient ϕ_0, with

$\phi_{-k} = \phi_k$, then $\beta = \sum_k \phi_k \sin(\omega k) = 0$ and, therefore, $\theta = 0$. That is to say, a symmetric filter that looks equally forward and backwards in time has no phase effect.

The z-transform of the sequence of filter coefficients is the polynomial

$$\phi(z) = \sum_k \phi_k z^k, \tag{4}$$

wherein z stands for a complex number. Setting $z = \exp\{-\mathrm{i}\omega\} = \cos(\omega) - \mathrm{i}\sin(\omega)$ constrains this number to lie on the unit circle in the complex plane. The resulting function

$$\begin{aligned}\phi(\exp\{-\mathrm{i}\omega\}) &= \sum_k \phi_k \cos(\omega k) - \mathrm{i}\sum_k \phi_k \sin(\omega k)\\ &= \alpha(\omega) - \mathrm{i}\beta(\omega)\end{aligned} \tag{5}$$

is the frequency response function, which is, in general, a periodic complex-valued function of ω with a period of 2π. In the case of a symmetric filter, it becomes a real-valued and even function, which is symmetric about $\omega = 0$. When the frequency response function is defined over the interval $[-\pi, \pi)$, or equally over the interval $[0, 2\pi)$, it conveys all of the information concerning the gain and the phase effects of the filter.

For a more concise notation, we may write $\phi(\omega)$ in place of $\phi(\exp\{-\mathrm{i}\omega\})$. This allows us to denote the frequency response by

$$\phi(\omega) = |\phi(\omega)|e^{-\mathrm{i}\theta(\omega)}, \quad \text{where} \quad |\phi(\omega)| = \sqrt{\alpha^2(\omega) + \beta^2(\omega)}. \tag{6}$$

Here, $|\phi(\omega)|$ denotes the gain of the filter at the frequency ω, whereas $\theta(\omega)$ indicates the phase effect. In the case of a symmetric filter, there is $\theta(\omega) = 0$ and $\phi(\omega) = \sum_k \phi_k \cos(\omega k)$; and $\exp\{-\mathrm{i}\theta(\omega)\}$ is evaluated as either $+1$ or -1, according to the sign of $\phi(\omega)$.

An ideal frequency-selective filter has the effect of nullifying all trigonometric sequences of which the frequencies fall within the stop band and of preserving, without alteration, all those of which the frequencies fall within the pass band.

The ideal phase-neutral lowpass filter with a cut-off at frequency $\omega = \alpha$ has the following frequency response over the interval $[-\pi, \pi]$:

$$\phi(\omega) = \begin{cases} 1, & \text{if } |\omega| \in (0, \alpha), \\ 1/2, & \text{if } \omega = \pm\alpha, \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

Here, the gain of the filter coincides with its frequency response. The coefficients of a filter may be obtained via the (inverse) Fourier transform of $\phi(\omega)$. In the case of the ideal filter, they are given by the sampled ordinates of a sinc function:

$$\phi_k = \frac{1}{2\pi}\int_{-\alpha}^{\alpha} e^{i\omega k} d\omega = \begin{cases} \alpha, & \text{if } k = 0; \\ \dfrac{\sin(\alpha k)}{\pi k}, & \text{if } k \neq 0. \end{cases} \tag{8}$$

The coefficients constitute a doubly-infinite sequence which sums to unity. Figure 1 shows the central coefficients of the ideal lowpass filter with a cut-off frequency of $\alpha = \pi/2$.

The coefficients of a bandpass filter with a gain of unity within the interval $[\alpha, \beta]$ and a gain of zero outside the interval are given by $\phi_k = \{\sin(\beta k) - \sin(\alpha k)\}/\pi k$, when $k \neq 0$, together with $\phi_0 = \beta - \alpha$. This is just the difference of two lowpass filters. The sum of the coefficients of a bandpass filter is zero.

In practice, all data sequences are finite, and it is impossible to apply a filter that has an infinite number of coefficients. However, a practical filter may be obtained by selecting a limited number of the central coefficients of an ideal infinite-sample filter. In the case of a truncated filter based on $2q+1$ central coefficients, the elements of the filtered sequence are given by

$$\begin{aligned} x_t = {} & \phi_q y_{t-q} + \phi_{q-1} y_{t-q+1} + \cdots + \phi_1 y_{t-1} + \phi_0 y_t \\ & + \phi_1 y_{t+1} + \cdots + \phi_{q-1} y_{t+q-1} + \phi_q y_{t+q}. \end{aligned} \tag{9}$$

Given a sample $y_0, y_1, \ldots, y_{T-1}$ of T data points, only $T - 2q$ processed values $x_q, x_{q+1}, \ldots, x_{T-q-1}$ are available, since the filter cannot reach the ends of the sample, unless the sample is extrapolated.

The usual effect of the truncating the filter will be to cause a considerable spectral leakage. Thus, if the filter is applied to trended data, then it is liable to transmit some powerful low-frequency elements that will give rise to cycles of high amplitudes within the filtered output. Another effect of the truncation will be a violation of the condition regarding the sum of the filter coefficients.

If the coefficients of the truncated bandpass or highpass filter are adjusted so that they sum to zero, then the z-transform polynomial $\phi(z)$ of the coefficient sequence will contain two roots of unit value. The adjustments may be made by subtracting $\sum_k \phi_k/(2q+1)$ from each coefficient. The sum of the adjusted coefficients is $\phi(1) = 0$, from which it follows that $1 - z$ is a factor of $\phi(z)$. The condition of symmetry, which is that $\phi(z) = \phi(z^{-1})$, implies that $1 - z^{-1}$ is also a factor. Therefore, the polynomial contains the factor

$$(1 - z)(1 - z^{-1}) = -z^{-1}(1 - z)^2, \tag{10}$$

within which $\nabla^2(z) = (1 - z)^2$ corresponds to the square of the difference operator.

Since it incorporates the factor $\nabla^2(z)$, the effect of applying the filter to a data sequence with a linear trend will be to produce an untrended sequence with a zero mean. The effect of applying it to a sequence with a quadratic trend will be to produce an untrended sequence with a nonzero mean. Such filters have been used by Baxter and King (1999) in extracting the business cycle from strongly trended aggregate economic indices.

It is possible to remove $\nabla^2(z)$ from $\phi(z) = \psi(z)\nabla^2(z)$. Then, the corresponding differencing operator can be applied to the data with the aim of reducing it to stationarity before applying a reduced filter, of which $\psi(z)$ is the z-transform. However, when the computations are wholly within the time domain, such an approach has no practical advantage over the approach that applies the symmetric filter $\phi(z)$ directly to the undifferenced data.

An alternative filter that is designed to reach the ends of the sample has been proposed by Christiano and Fitzgerald (2003). The filter is described by the equation

$$\begin{aligned} x_t = &Ay_0 + \phi_t y_0 + \cdots + \phi_1 y_{t-1} + \phi_0 y_t \\ &+ \phi_1 y_{t+1} + \cdots + \phi_{T-1-t} y_{T-1} + By_{T-1}. \end{aligned} \tag{11}$$

This equation comprises the entire data sequence $y_0, \ldots, y_{T-1}$; and the value of t determines which of the coefficients of the infinite-sample filter are involved in producing the current output. The value of x_0 is generated by looking forwards to the end of the sample, whereas the value of x_{T-1} is generated by looking backwards to the beginning of the sample.

If the process generating the data is stationary, then it is appropriate to set $A = B = 0$, which is tantamount to approximating the extra-sample elements by zeros. In the case of a data sequence that appears to follow a first-order random walk, it has been proposed to set A and B to the values of the sums of the coefficients that lie beyond the span of the data on either side. Since the filter coefficients must sum to zero, it follows that

$$A = -(\frac{1}{2}\phi_0 + \phi_1 + \cdots + \phi_t) \quad \text{and} \quad B = -(\frac{1}{2}\phi_0 + \phi_1 + \cdots + \phi_{T-t-1}). \tag{12}$$

The effect is tantamount to extending the sample at either end by constant sequences comprising the first and the last sample values respectively. For data that have the appearance of having been generated by a first-order random walk with a constant drift, it is appropriate to extract a linear trend before filtering the residual sequence. In fact, this has proved to be the usual practice in most circumstances.

Christiano and Fitzgerald (2003) have also proposed another time-varying filter that is intended to be a superior approximation to the ideal bandpass filter. The coefficients of this filter, which may be denoted by $\phi_j^{(t)}$ and which are determined for each value of t, are comprised by the equation

$$\begin{aligned} x_t = &\phi_t^{(t)} y_0 + \cdots + \phi_1^{(t)} y_{t-1} + \phi_0^{(t)} y_t \\ &+ \phi_1^{(t)} y_{t+1} + \cdots + \phi_{T-1-t}^{(t)} y_{T-1}. \end{aligned} \tag{13}$$

Let $\phi(z)$ be the z-transform of the coefficients of the ideal infinite-sample filter and let $\phi^{(t)}(z)$ be the z-transform of the finite-sample filter for time t.

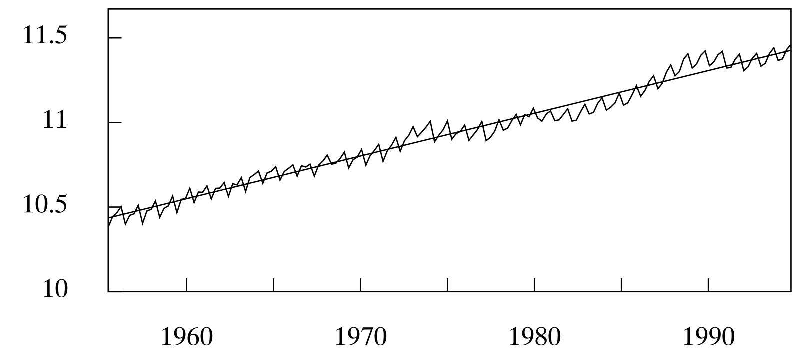

Fig. 4. The quarterly sequence of the logarithms of consumption in the U.K., for the years 1955 to 1994, together with a linear trend interpolated by least-squares regression.

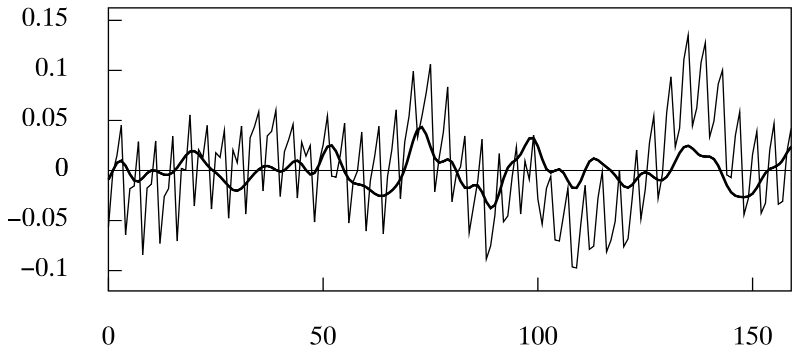

Fig. 5. The sequence derived by applying the bandpass filter of Christiano and Fitzgerald to the quarterly logarithmic data on U.K. consumption.

Setting $z = \exp\{-i\omega\}$ gives the frequency response functions of the filters. It is proposed that the coefficients of the finite-sample filter should be the values that jointly minimise the function

$$\int_{-\pi}^{\pi} |\phi(e^{-i\omega}) - \phi^{(t)}(e^{-i\omega})|^2 f(\omega) d\omega, \tag{14}$$

where $f(\omega)$ is the spectral density function of the process generating the data.

The intention of this criterion is to minimise the discrepancy between the finite-sample filter and the ideal filter in those regions of the frequency domain, indicated by the values of $f(\omega)$, where it matters most. However, given that a data sequence of T elements is represented in the frequency domain by a set of complex exponential functions defined on T frequency values, described as the Fourier frequencies, there is no cause for assessing the discrepancy at every frequency in the interval $[\pi, \pi]$.

Moreover, it is possible to devise an ideal finite-sample filter that eliminates the discrepancy completely at the T Fourier frequencies. To demonstrate this point, it is necessary to consider the discrete Fourier transform of the finite data sequence.

Example. Figure 4 represents the logarithms of the data on quarterly real household expenditure in the U.K. for the period 1955–1994, through which a linear trend has been interpolated that passes through the midst of the data points of the first and the final years. The residual deviations of the data from the trend are shown in Figure 5. Superimposed upon these residuals is a sequence that has been generated by subjecting them to the filter of Christiano and Fitzgerald (2003) with a nominal pass band over the interval $[\pi/16, \pi/3]$.

This range of frequencies corresponds to cycles of durations of no less than one-and-a-half years and not exceeding 8 years. According to Baxter and King (1999), the range accords with a definition of the business cycle that was proposed by Burns and Mitchell (1946). The filtered sequence fails to capture some of the salient low-frequency fluctuations of the data; and it contains some high-frequency fluctuations that we would not normally be regarded as part of the business cycle.

3 The Discrete Fourier Transform

The discrete Fourier transform is a one-to-one mapping from a set of T data points to a set of T coefficients associated with a set of harmonically related trigonometric functions. The vectors of the ordinates sampled from the trigonometric functions constitute an orthogonal basis of the T-dimensional space that contains the data vector.

The inverse Fourier transform, which is a mapping from the coefficients to the data, gives rise to the following equation, which describes the Fourier synthesis of the data:

$$y_t = \sum_{t=0}^{[T/2]} \{\alpha_j \cos(\omega_j t) + \beta_j \sin(\omega_j t)\}; \quad t = 0, 1, \ldots, T-1. \tag{15}$$

Here, $[T/2]$ denotes the integer quotient of the division of T by 2. The harmonically related Fourier frequencies $\omega_j = 2\pi j/T; j = 0, \ldots, [T/2]$, which are equally spaced in the interval $[0, \pi]$, are integer multiples of the fundamental frequency $\omega_1 = 2\pi/T$, which relates to a sinusoidal function that completes a single cycle in the time spanned by the sample. A stochastic nature is imparted to y_t by the coefficients α_j, β_j, which are to be regarded as random variables

The temporal index t of the above equation ranges from 0 to $T-1$. However, strictly for analytic purposes, we may regard the data sequence as a single cycle of a periodic function defined over the entire set of positive and negative integers, which is described as the periodic extension the data. Considering the periodic extension does not entail making any assumption that the data have been generated by an underlying periodic process.

For mathematical convenience, we may express the trigonometric functions of (15) in terms of complex exponential functions:

$$\cos(\omega_j t) = \frac{1}{2}\left(e^{\mathrm{i}\omega_j t} + e^{-\mathrm{i}\omega_j t}\right), \qquad \sin(\omega_j t) = \frac{-\mathrm{i}}{2}\left(e^{\mathrm{i}\omega_j t} - e^{-\mathrm{i}\omega_j t}\right). \tag{16}$$

Then, on defining

$$\zeta_j = \frac{\alpha_j - \mathrm{i}\beta_j}{2} \qquad \text{and} \quad \zeta_{T-j} = \frac{\alpha_j + \mathrm{i}\beta_j}{2}, \tag{17}$$

equation (15) can be written as

$$y_t = \sum_{j=0}^{T-1} \zeta_j e^{\mathrm{i}\omega_j t} = \sum_{j=0}^{T-1} \zeta_j W^{jt}; \quad t = 0, 1, \ldots, T-1, \tag{18}$$

where $W^{jt} = \exp\{2\pi jt/T\}$. Here, W^q is a T-period function of q such that $W \uparrow q = W \uparrow (q \bmod T)$, where the upward arrow signifies exponentiation. The complex values $W^0 = 1, W, W^2, \ldots, W^{T-1}$, which describe one cycle of the function, are known as the T roots of unity.

Using the exponential notation, the Fourier transform and its inverse can be denoted by

$$\zeta_j = \frac{1}{T}\sum_{t=0}^{T-1} y_t e^{-\mathrm{i}\omega_j t} dt \quad \longleftrightarrow y_t = \sum_{j=0}^{T-1} \zeta_j e^{\mathrm{i}\omega_j t}. \tag{19}$$

For a matrix representation of these transforms, one may define

$$\begin{aligned} U &= T^{-1/2}[\exp\{-\mathrm{i}2\pi tj/T\}; t, j = 0, \ldots, T-1], \\ \bar{U} &= T^{-1/2}[\exp\{\mathrm{i}2\pi tj/T\}; t, j = 0, \ldots, T-1], \end{aligned} \tag{20}$$

which are unitary complex matrices such that $U\bar{U} = \bar{U}U = I_T$. Then,

$$\zeta = T^{-1/2}Uy \quad \longleftrightarrow \quad y = T^{1/2}\bar{U}\zeta. \tag{21}$$

where $y = [y_0, y_1, \ldots y_{T-1}]'$ and $\zeta = [\zeta_0, \zeta_1, \ldots \zeta_{T-1}]'$ are the vectors of the data and of their spectral ordinates, respectively.

Observe that, under the assumption that the data are generated by a stationary stochastic process, the limiting form of equation (18), as $T \to \infty$, is equation (2).

4 The Ideal Finite-Sample Filter

In terms of the frequency domain, the process of filtering a finite data sequence consists of altering the values of the spectral ordinates within the vector $\zeta = [\zeta_0, \zeta_1, \ldots, \zeta_{T-1}]'$. These ordinates, which correspond to the Fourier frequencies

$$\omega_j = 2\pi j/T; \quad j = 0, 1, 2, \ldots, T-1, \tag{22}$$

may be envisaged as a set of spikes erected on the circumference of the unit circle in the complex plane at locations that are indicated by the T roots of unity

$$W^j = \exp\{-\mathrm{i}\omega_j\} = \cos(\omega_j) - \mathrm{i}\sin(\omega_j); \quad j = 0, 1, 2, \ldots, T-1. \tag{23}$$

The frequency response at ω_j, determined in accordance with the specification (7) of the ideal filter, is

$$\lambda_j = \phi(\omega_j) = \sum_{k=-\infty}^{\infty} \phi_k W^{jk}. \tag{24}$$

Since W^q is a T-periodic function, it follows that

$$\begin{aligned}\lambda_j &= \Big\{\sum_{q=-\infty}^{\infty} \phi_{qT}\Big\} + \Big\{\sum_{q=-\infty}^{\infty} \phi_{qT+1}\Big\}W^j + \cdots + \Big\{\sum_{q=-\infty}^{\infty} \phi_{qT+T-1}\Big\}W^{j(T-1)} \\ &= \phi_0^\circ + \phi_1^\circ W^j + \cdots + \phi_{T-1}^\circ W^{j(T-1)}, \quad \text{for} \quad j = 0, 1, 2, \ldots, T-1. \end{aligned} \tag{25}$$

These equations serve to determine the circular filter coefficients $\phi_0^\circ, \phi_1^\circ, \ldots, \phi_{T-1}^\circ$. Since the filter coefficients constitute a symmetric sequence, with $\phi_j = \phi_{-j}$, it follows that $\phi_j^\circ = \phi_{T-j}^\circ$.

Let $\phi^\circ = [\phi_0^\circ, \phi_1^\circ, \ldots, \phi_{T-1}^\circ]'$ be the vector of the coefficients of a circular filter and let $\lambda = [\lambda_0, \lambda_1, \ldots, \lambda_{T-1}]'$ be the vector of the values of the frequency response at the Fourier frequencies. Then, in terms of the matrices of (20), the mapping from ϕ° to λ and the corresponding inverse mapping can be represented by

$$\lambda = T^{-1/2}U\phi^\circ \quad \longleftrightarrow \quad \phi^\circ = T^{1/2}\bar{U}\lambda. \tag{26}$$

The filtering operation can be performed by multiplying the spectral ordinates within the vector ζ by the weights within λ. The resulting vector may be transformed from the frequency domain to the time domain to produce the filtered output.

Let $\Lambda = \mathrm{diag}\{\lambda_0, \lambda_1, \ldots, \lambda_{T-1}\}$ be the diagonal matrix of the weights. Then, with reference to (21), it can be see that the weighted values of the spectral ordinates are given by the vector

$$\Lambda\zeta = T^{-1/2}\Lambda U y. \tag{27}$$

Subjecting this vector to the inverse Fourier transform gives the filtered output

$$x = T^{1/2}\bar{U}\Lambda\zeta = \{\bar{U}\Lambda U\}y = \Phi^\circ y, \tag{28}$$

where $\Phi^\circ = \bar{U}\Lambda U$ is the matrix of the filtering operation in the time domain. The next section is devoted to revealing the nature of this matrix and of the associated time-domain filtering operation.

5 Filtering via Circular Convolution

Consider the following matrix equation:

$$
\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & W & 0 & 0 \\ 0 & 0 & W^2 & 0 \\ 0 & 0 & 0 & W^3 \end{bmatrix}
\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & W & W^2 & W^3 \\ 1 & W^2 & W^4 & W^6 \\ 1 & W^3 & W^6 & W^9 \end{bmatrix}
\tag{29}
$$
$$
= \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & W & W^2 & W^3 \\ 1 & W^2 & W^4 & W^6 \\ 1 & W^3 & W^6 & W^9 \end{bmatrix}
\begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}
= \begin{bmatrix} 1 & 1 & 1 & 1 \\ W & W^2 & W^3 & 1 \\ W^2 & W^4 & W^6 & 1 \\ W^3 & W^6 & W^9 & 1 \end{bmatrix}
$$

The first equality can be represented in summary notation by

$$DU = UK. \tag{30}$$

This example is readily generalised to encompass matrices of any order.

In general, $K = [e_1, e_2, \ldots, e_{T-1}, e_0]$ is a circulant matrix operator that is formed from the identity matrix $I = [e_0, e_1, e_2, \ldots, e_{T-1}]$ by moving the leading vector to the back of the array, whereas $D = \mathrm{diag}\{1, W, W^2, \ldots, W^{T-1}\}$ is a diagonal matrix containing the roots of unity. The remaining matrix U is in accordance with the definitions of (20). More extensive accounts of the algebra of circulant matrices have been given by Pollock (1999, 2002).

From (30), it follows that

$$K = \bar{U}DU \quad \text{and} \quad K^q = \bar{U}D^qU, \tag{31}$$

where the second identity depends upon the result that $U\bar{U} = I$. We may also observe that $\Lambda = \mathrm{diag}\{\lambda_0, \lambda_1, \ldots, \lambda_{T-1}\}$, which contains the ordinates sampled at the Fourier frequencies from the frequency response function of the filter, can be written as

$$\Lambda = \phi^\circ(D), \tag{32}$$

where $\phi^\circ(D)$ is obtained by replacing z by D in the z-transform $\phi^\circ(z) = \phi^\circ_0 + \phi^\circ_1 z + \phi^\circ_2 z^2 + \cdots + \phi^\circ_{T-1} z^{T-1}$. It follows from (31) and (32) that the filter matrix is

$$\Phi^\circ = \bar{U}\Lambda U = \bar{U}\phi^\circ(D)U = \phi^\circ(K). \tag{33}$$

where $\Phi^\circ = \phi^\circ(K)$ is obtained by replacing z by K in the z-transform $\phi^\circ(z)$. Thus, it can be seen that

$$\begin{aligned} \Phi^\circ &= \phi_0 I + \phi_1 K + \cdots + \phi_{T-1} K^{T-1} \\ &= [\phi^\circ, K\phi^\circ, \ldots, K^{T-1}\phi^\circ], \end{aligned} \tag{34}$$

where $\phi^\circ = [\phi^\circ_0, \phi^\circ_1, \ldots, \phi^\circ_{T-1}]'$ is the vector of the coefficients of the circular filter. This is manifestly a circulant matrix. In consequence of the condition that $\phi^\circ_j = \phi^\circ_{T-j}$, it is also a symmetric matrix. The form of the matrix is adequately represented by the case where $T = 4$:

$$\Phi^\circ = \begin{bmatrix} \phi^\circ_0 & \phi^\circ_1 & \phi^\circ_2 & \phi^\circ_1 \\ \phi^\circ_1 & \phi^\circ_0 & \phi^\circ_1 & \phi^\circ_2 \\ \phi^\circ_2 & \phi^\circ_1 & \phi^\circ_0 & \phi^\circ_1 \\ \phi^\circ_1 & \phi^\circ_2 & \phi^\circ_1 & \phi^\circ_0 \end{bmatrix}. \tag{35}$$

Applying the filter matrix Φ° to the data vector $y = [y_0, y_1, \ldots, y_{T-1}]'$ gives

$$x = \Phi^\circ y = \bar{U}\Lambda U y. \tag{36}$$

This equation indicates that there are two ways of forming the filtered vector x. The first way is via the circular convolution in the time domain of the vector ϕ° of the filter coefficients and the vector y of the data. To elucidate this operation, define the circulant data matrix $Y = [y, Ky, K^2y, \ldots, K^{T-1}y]$ and observe that circulant matrices commute in multiplication. It follows that

$$\Phi^\circ y = \Phi^\circ Y e_0 = Y\Phi^\circ e_0 = Y\phi^\circ. \tag{37}$$

This expression manifests a symmetry that puts the data vector y and the filter vector ϕ° on an equal footing.

The second way of obtaining the filtered output is via the Fourier transform and its inverse. First, the discrete Fourier transform is applied to the data vector to carry it into the frequency domain. Then a differential weighting is applied to the elements of the resulting vector $\zeta = Uy$ to give $\Lambda\zeta = \Lambda U y$. Finally, the filtered vector $x = \bar{U}\Lambda U y$ is obtained by applying the inverse Fourier transform.

6 The Finite-Sample Frequency Response

It should be emphasised that $\phi(z)$ and $\phi^\circ(z)$ 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 $\phi^\circ(z)$ would deliver if its argument were free to travel around the unit circle.

Figure 6 is designed to satisfy such curiosity. Here, it can be seen that, within the stop band, the function $\phi^\circ(\omega)$ 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.

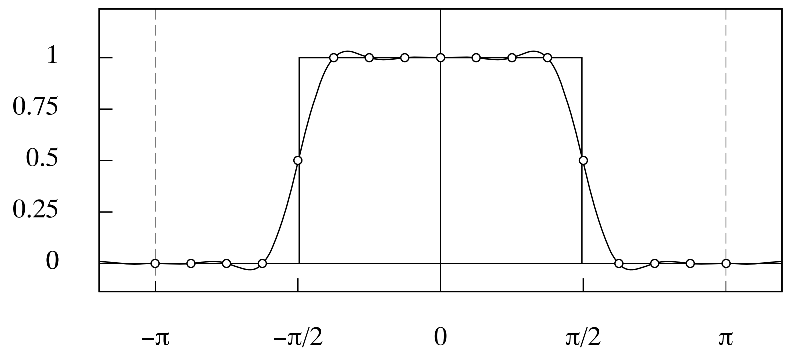

Fig. 6. The frequency response of the 16-point wrapped filter defined over the interval $[-\pi, \pi)$. The values at the Fourier frequencies are marked by circles. (Note that, when the horizontal axis is wrapped around the circumference of the unit circle, the points at π and $-\pi$ coincide. Therefore, only one of them is included in the interval.)

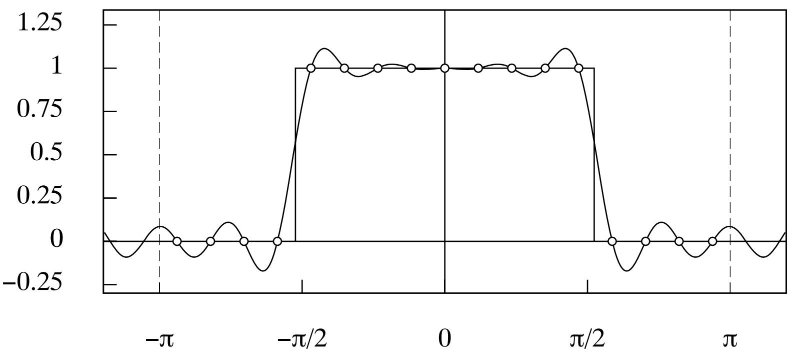

Fig. 7. The frequency response of the 17-point wrapped filter defined over the interval $[-\pi, \pi)$. The values at the Fourier frequencies are marked by circles.

This explains the seeming paradox whereby we are able to achieve a perfect frequency selection via a finite filter. In theory, we do have a doubly-infinite sequence at our disposal in the form of the periodic extension of the data. Applying a wrapped filter to a finite data sequence by circular convolution is equivalent to applying an infinite, unwrapped filter to its periodic extension by ordinary linear convolution.

In Figure 6, the transition between the pass band and the stop band occurs at a Fourier frequency. This feature is also appropriate to other filters that one can design, which have a more gradual transition with a mid point that also falls on a Fourier frequency. Figure 7 shows that it is possible, nevertheless, to make an abrupt transition within the space between two adjacent frequencies. The formulae for the filter coefficients in both cases are given in the appendix.

The method of filter design that we are pursuing in this paper allows considerable flexibility in specifying the form of the frequency response function. Often, there is an advantage in departing from the ideal specification so as to allow a more gradual transition between the pass band and the stop band.

However, for freely specified responses, it may be difficult to find analytic expressions for the corresponding filter coefficients.

In the case of the ideal function, the coefficients of the wrapped filter are readily available. In the appendix, the coefficients are found of the lowpass filter that is obtained by sampling the following periodic frequency response function at T Fourier points $\omega_j = 2\pi j/T$ that lie in the interval $[-\pi, \pi)$:

$$\phi(\omega) = \begin{cases} 1, & \text{if } \omega \in (-\omega_d, \omega_d), \\ 1/2, & \text{if } \omega = \pm\omega_d, \\ 0, & \text{for } \omega \text{ elsewhere in } [-\pi, \pi). \end{cases} \tag{38}$$

Here, $\pm\omega_d = \pm d\omega_1 = \pm 2\pi d/T$ are the points of discontinuity. The filter coefficients are given by

$$\phi_d^\circ(k) = \begin{cases} \dfrac{2d}{T}, & \text{if } k = 0 \\ \dfrac{\cos(\omega_1 k/2)\sin(d\omega_1 k)}{T\sin(\omega_1 k/2)}, & \text{for } k = 1, \ldots, [T/2], \end{cases} \tag{39}$$

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

One might wish to construct a wrapped filter according to the more general bandpass specification:

$$\phi(\omega) = \begin{cases} 1, & \text{if } |\omega| \in (-\omega_a, \omega_b), \\ 1/2, & \text{if } \omega = \pm\omega_a, \pm\omega_b, \\ 0, & \text{for } \omega \text{ elsewhere in } [-\pi, \pi), \end{cases} \tag{40}$$

where $\omega_a = a\omega_1$ and $\omega_b = b\omega_1$. For the ideal filter, this can be achieved by subtracting one filter from another to create

$$\begin{aligned} \phi_{[a,b]}^\circ(t) &= \phi_b^\circ(t) - \phi_a^\circ(t) \\ &= \frac{\cos(\omega_1 t/2)\{\sin(b\omega_1 t) - \sin(a\omega_1 t)\}}{T\sin(\omega_1 t/2)} \\ &= 2\cos(g\omega_1 t)\frac{\cos(\omega_1 t/2)\sin(d\omega_1 t)}{T\sin(\omega_1 t/2)}. \end{aligned} \tag{41}$$

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. The final expression follows from the identity $\sin(A + B) - \sin(A - B) = 2\cos A \sin B$. The expression can be interpreted as the result of shifting a lowpass filter with a cut-off frequency at the point d so that its centre is moved from 0 to the point g. The technique of frequency shifting is not confined to the ideal frequency response. It can be applied to any frequency response function.

Example. In applying the methods of frequency-domain filtering, it is necessary to ensure that the data have no trend. It is also important to ensure that

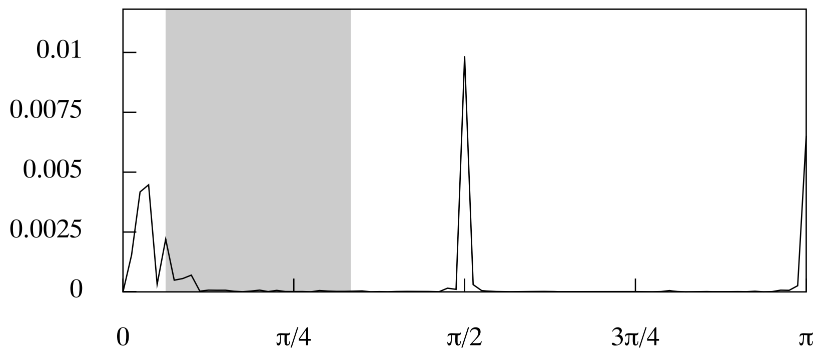

Fig. 8. The periodogram of the residual sequence obtained from the linear detrending of the logarithmic consumption data. A band, with a lower bound of $\pi/16$ radians and an upper bound of $\pi/3$ radians, is masking the periodogram.

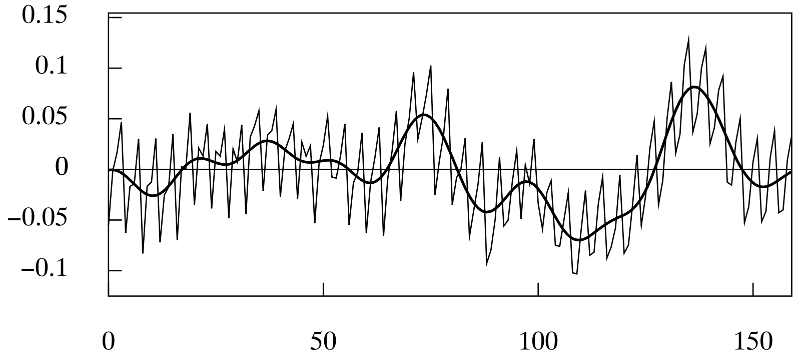

Fig. 9. The residual sequence from fitting a linear trend to the logarithmic consumption data, together with an interpolated line representing the business cycle.

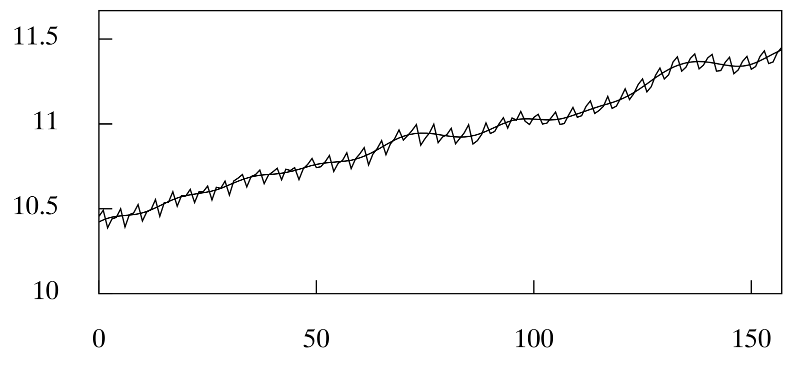

Fig. 10. The trend-cycle component of U.K. consumption determined by the frequency-domain method, superimposed on the logarithmic data.

there are no 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 replication. Equivalently, there must be a smooth transition between the finishing and the starting points when the sequence of sample points is wrapped around a circle of circumference T.

Both objectives can be achieved by using a weighted least-squares regression procedure to interpolate a polynomial trend function through the data in a manner that ensures that it passes through the midst of the scatters of end points. This usually requires extra weight to be given to the end points. For the logarithmic consumption data of Figure 4, an ordinary least-squares regression generates a linear trend that passes fortuitously through the end points of the sample; and, therefore, the residual sequence can be wrapped seamlessly around the circle.

The periodogram of the regression residuals is shown in figure 8. From this, it is apparent that the low-frequency structure of the data extends no further in frequency than $\pi/8$ radians per sample interval. Superimposed upon this periodogram is a highlighted band that covers the interval $[\pi/16, \pi/3]$, which is the nominal pass band of the filter that has been used to generate the sequence in Figure 5, which purports to represent the business cycle.

The Fourier ordinates that fall in the interval $[0, \pi/8]$ have been used in synthesising the continuous function that has been superimposed upon the residual sequence in Figure 9. This function provides a better representation of the business cycle than does the function of Figure 5. To obtain a representation of the trend-cycle component, the business cycle function of Figure 9 can be added to the linear trend of Figure 4. The result is shown in Figure 10.

7 Filtering Trended Sequences

In the U.K., the period 1955–1994 was one of steady economic growth. Therefore, the line that has been interpolated through the logarithmic consumption data in Figure 4 serves to represent the underlying trajectory of the economy; and it provides a benchmark for measuring the fluctuations of the business cycle.

The program that has been mentioned in the introduction to this paper has a variety of facilities that can be used in estimating trajectories of a more variable nature. Polynomial functions can often serve the purpose. Flexible smoothing devices are also available that will allow the trend to absorb structural breaks, such as those that occurred at the ends of the two world wars and at the start of the inter-war recession.

The sequence of the residual deviations of the data from the estimated trend are liable to be subjected to a process of lowpass filtering of the sort that has been described in the previous section. When the filtered sequence is added to the trend, the result will constitute an estimate of the trend-cycle component of the data.

The trend-cycle component will hardly be affected by varying the way in which the trend is depicted, since the estimate of the cycle will contain the compensating variations. This is bound to be the case whenever, after the extraction of the linear component, the Fourier elements of the trend function fall within the band of frequencies spanned by the cyclical component. For, if some of the motions that would otherwise belong to the cycle are assigned the trend, then there will be no alteration in the composite trend-cycle component.

The essential requirement is to fit the trend in such a way as to avoid any end-of-sample disjunctions in the residual sequence. Such disjunctions will give rise to accidental frequency-domain components that tend to obscure other components that may be of genuine interest.

For estimating the trend-cycle component, some alternative procedures are available that depend on the difference operator to eliminate the trend from the data. After the differenced data have been filtered, they may need to be reinflated to produce estimates of the corresponding components of the original data sequence. The relevant anti-differencing operations can be performed either in the time domain or in the frequency domain

The matrix that takes the p-th (backward) difference of a vector of order T is given by

$$\nabla_T^p = (I - L_T)^p, \tag{42}$$

where $L_T = [e_1, e_2, \ldots, e_{T-1}, 0]$ is the matrix lag operator that is formed from the identity matrix $I_T = [e_0, e_1, e_2, \ldots, e_{T-1}]$ by deleting the leading vector and appending a zero vector to the end of the array.

The differencing matrix may be partitioned such that $\nabla_T^p = [Q_*, Q]'$, where Q_*' has p rows. The inverse matrix is partitioned conformably to give $\nabla_T^{-p} = [S_*, S]$. It follows that

$$\begin{bmatrix} S_* & S \end{bmatrix} \begin{bmatrix} Q_*' \\ Q' \end{bmatrix} = S_*Q_*' + SQ' = I_T, \tag{43}$$

and that

$$\begin{bmatrix} Q_*' \\ Q' \end{bmatrix} \begin{bmatrix} S_* & S \end{bmatrix} = \begin{bmatrix} Q_*'S_* & Q_*'S \\ Q'S_* & Q'S \end{bmatrix} = \begin{bmatrix} I_d & 0 \\ 0 & I_{T-d} \end{bmatrix}. \tag{44}$$

When the difference operator is applied to the data vector y, the first p elements of the product, which are in g_*, are not true differences and they are liable to be discarded:

$$\nabla_T^p y = \begin{bmatrix} Q_*' \\ Q' \end{bmatrix} y = \begin{bmatrix} g_* \\ g \end{bmatrix}. \tag{45}$$

However, if the elements of g_* are available, then the vector y can be recovered from $g = Q'y$ via the equation

$$y = S_*g_* + Sg. \tag{46}$$

The columns of the matrix S_* provide a basis for the set of polynomials of degree $p-1$ defined over the integer values $t = 0, 1, \ldots, T-1$. Therefore, S_*g_*

is a vector of polynomial ordinates, whilst g_* can be regarded as a vector of p polynomial parameters.

For an example of the differencing operator, we may consider the case where the degree of differencing is $p = 2$ and the length of the data sequence is $T = 5$. Then, there is

$$\nabla_5^2 = \begin{bmatrix} Q'_* \\ Q' \end{bmatrix} = \left[\begin{array}{ccccc} 1 & 0 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 & 0 \\ \hline 1 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -2 & 1 \end{array}\right]. \tag{47}$$

A trended data sequence may be filtered as follows. First, the data is reduced to stationarity by differencing it an appropriate number of times. (We rarely need to difference the data more than twice.) Next, the relevant filters are applied to the differenced data to isolate its components. Finally, the components of the differenced data may be integrated, with an appropriate choice of initial conditions, to provide estimates of the components of the original trended sequence.

The initial conditions will be determined according to a criterion that assumes different forms depending on whether the component to be extracted is trended or non-trended. A trended component will contain the zero-frequency Fourier element, which is a constant vector, together with other elements of adjacent frequencies in the vicinity of the zero. We presume that there can be only one trended component. The remaining components will comprise fluctuations around mean values of zero.

In accordance with this categorisation, we can represent the generic decomposition of the data vector as

$$y = x + h, \tag{48}$$

where x is the trend or trend-cycle component and h is the complementary detrended component, which might be subject to further decompositions. The differenced data would be

$$\begin{aligned} Q'y &= Q'x + Q'h \\ &= d + k = g. \end{aligned} \tag{49}$$

The vectors d and k require to be cumulated to form

$$x = S_* d_* + Sd \qquad \text{and} \qquad h = S_* k_* + Sk. \tag{50}$$

However, given the adding-up constraint that is posed by (48), the initial conditions within d_* and k_* must be equivalent.

The initial conditions should be chosen so as to ensure that the trend is aligned with the data as closely as possible or, equivalently, that the deviations of the trend from the data are minimised. This entails minimising the quadratic norm $h'h = (y - x)'(y - x)$. The criterion for finding k_* is, therefore,

$$\text{Minimise} \quad (S_*k_* + Sk)'(S_*k_* + Sk) \quad \text{with respect to} \quad k_*. \tag{51}$$

The solution for the starting values is

$$k_* = -(S_*'S_*)^{-1}S_*'Sk. \tag{52}$$

The equivalent criterion for finding d_* is

$$\text{Minimise} \quad (y - S_*d_* - Sd)'(y - S_*d_* - Sd) \quad \text{with respect to} \quad d_*. \tag{53}$$

The solution for the starting values is

$$d_* = (S_*'S_*)^{-1}S_*'(y - Sd). \tag{54}$$

In terms of the notation

$$P_* = S_*(S_*'S_*)^{-1}S_*', \tag{55}$$

the equations of (50) can be written as

$$x = P_*y + (I - P_*)Sd, \qquad \text{and} \qquad h = (I - P_*)Sk. \tag{56}$$

Since $(I - P_*)\iota = 0$, where $\iota = [1, 1, \ldots, 1]'$ is the summation vector, it follows that $\iota'h = 0$, which is to say that the detrended data have a mean of zero. Then, given that $S(d+k) = Sg$ and that $(I-P_*)Sg = (I-P_*)(Sg+S_*g_*) = (I-P_*)y$, since $(I - P_*)S_* = 0$, it follows that

$$\begin{aligned} x + h &= (I - P_*)S(d + k) + P_*y \\ &= (I - P_*)y + P_*y = y, \end{aligned} \tag{57}$$

which is to say that the sum of the estimated components is the original data vector, in accordance with (48). In practice, it is redundant to compute both x and h. Only one of them is required, since the other can be found by subtracting from y.

In extracting the stationary component h contained within a trended data sequence y, the business of reinflating the filtered sequence by summation can be avoided, thereby dispensing with the initial conditions. The highpass filter that would serve to extract $k = Q'h$ from $g = Q'y$ will contain an implicit differencing operator, which serves to nullify the low-frequency elements of the data. If the filter is symmetric, then it will embody at least a twofold differencing operator. The need for reinflation can be avoided by cancelling the inflating summation operator with the differencing factors within the filter.

We may begin by considering the symmetric version of the twofold differencing operator, which is to be applied to the data at the outset. This is

$$\begin{aligned}N(z) &= z^{-1} - 2 + z = z^{-1}(1-z)^2\\ &= z^{-1}\nabla^2(z).\end{aligned} \tag{58}$$

The matrix version of the operator is obtained by setting $z = L_T$ and $z^{-1} = L'_T$, which gives

$$N(L_T) = N_T = L_T - 2I_T + L'_T. \tag{59}$$

The first and the final rows of this matrix do not deliver true differences. Therefore, they are liable to be deleted, with the effect that the two end points are lost from the twice-differenced data. Deleting the rows $e'_0 N_T$ and $e'_{T-1}N_T$ from N_T gives the matrix Q', which can also be obtained by from $\nabla^2_T = (I_T - L_T)^2$ by deleting the matrix Q'_*, which comprises the first two rows $e'_0\nabla^2_T$ and $e'_1\nabla^2_T$. In the case of $T = 5$, there is

$$N_5 = \begin{bmatrix} Q'_{-1} \\ Q' \\ Q_{+1} \end{bmatrix} = \left[\begin{array}{ccccc} -2 & 1 & 0 & 0 & 0 \\ \hline 1 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -2 & 1 \\ \hline 0 & 0 & 0 & 1 & -2 \end{array}\right]. \tag{60}$$

On deleting the first and last elements of the vector $N_T y$, which are $Q'_{-1}y = e'_1\nabla^2_T y$ and $Q_{+1}y$, respectively, we get $Q'y = [q_1, \ldots, q_{T-2}]'$.

The vector $Q'y$ of differenced data is to be used both in the procedure that reinflates the filtered sequence by summation and in the present procedure that avoids doing so. However, in the absence of a summation, the missing elements are not restored to the filtered vector, nor is there any accompanying phase alteration.

The loss of the two elements from either end of the (centrally) twice-differenced data can be overcome by supplementing the original data vector y with two extrapolated end points y_{-1} and y_T. Alternatively, the differenced data may be supplemented by attributing appropriate values to q_0 and q_{T-1}. These could be zeros or some combination of the adjacent values. In either case, we will obtain a vector of order T denoted by $q = [q_0, q_1, \ldots q_{T-1}]'$.

In describing the method for implementing a highpass filter, let Λ be the matrix that selects the appropriate ordinates of the Fourier transform $\gamma = Uq$ of the twice differenced data. These ordinates must be reinflated to compensate for the differencing operation, which has the frequency response

$$f(\omega) = 2\cos(\omega) - 2. \tag{61}$$

The response of the anti-differencing operation is $1/f(\omega)$; and γ is reinflated by pre-multiplying by the diagonal matrix

$$V = \mathrm{diag}\{v_0, v_1, \ldots, v_{T-1}\}, \tag{62}$$

comprising the values $v_j = 1/f(\omega_j); j = 0, \ldots, T-1$, where $\omega_j = 2\pi j/T$.

Let $H = V\Lambda$ be the matrix that is applied to $\gamma = Uq$ to generate the Fourier ordinates of the filtered vector. The resulting vector is transformed to the time domain to give

$$h = \bar{U}H\gamma = \bar{U}HUq. \tag{63}$$

It will be see that $f(\omega)$ is zero-valued when $\omega = 0$ and that $1/f(\omega)$ is unbounded in the neighbourhood of $\omega = 0$. Therefore, a frequency-domain reinflation is available only when there are no nonzero Fourier ordinates in this neighbourhood. That is to say, it can work only in conjunction with highpass or bandpass filtering. However, it is straightforward to construct a lowpass filter that complements the highpass filter. The low-frequency trend component that is complementary to h is

$$x = y - h = y - \bar{U}HUq. \tag{64}$$

Example. Given that Q and S_* are complementary matrices such that $\text{Rank}[Q, S_*] = T$ and $Q'S_* = 0$, it follows that

$$P_* = S_*(S_*'S_*)^{-1}S_*' = I - Q(Q'Q)^{-1}Q'. \tag{65}$$

Therefore, since S_* provides the basis for a polynomial regression, the residual vector can be expressed as $r = Q(Q'Q)^{-1}Q'y$. This contains exactly the same information as the vector $g = Q'y$ of the differenced data.

The difference operator has the effect of nullifying the element of zero frequency and of attenuating radically the adjacent low-frequency elements. Therefore, the low-frequency spectral structures of the data are not perceptible in the periodogram of the differenced sequence. However, the periodogram of the residuals of the polynomial regression can be relied upon to reveal the spectral structures at all frequencies.

Figure 8 reveals that the trend-cycle structure of the consumption data falls within the interval $[0, \pi/8]$. To isolate corresponding data component, an ideal lowpass filter with a cut-off frequency of $\pi/8$ can be applied to the twice-differenced data, wherafter the filtered sequence can be reinflated within the time domain via the twofold summation operator, with the help of some initial conditions calculated according to equation (54).

Alternatively, an ideal highpass filter with a cut-off frequency of $\pi/8$ can be applied to the twice differenced data, which can be reinflated thereafter within the frequency domain by modulating the filtered spectral ordinates by the frequency response of the summation operator. Then, the result can be transformed to the time domain to produce a high-frequency component that can be subtracted from the data to deliver the trend-cycle estimate.

Both of these procedures have been implemented in the program mentioned in the introduction. Their results are indistinguishable. Nor do these differ from the result, represented in Figure 10, that has been obtained by

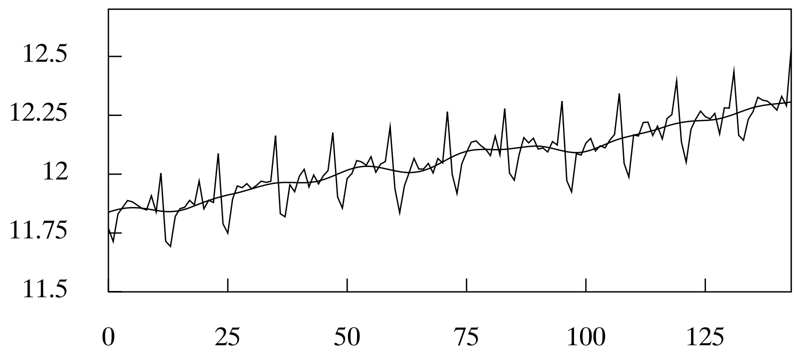

Fig. 11. The logarithms of the monthly data on retail sales in U.S. for the years 1953 to 1964, together with an interpolated trend-cycle.

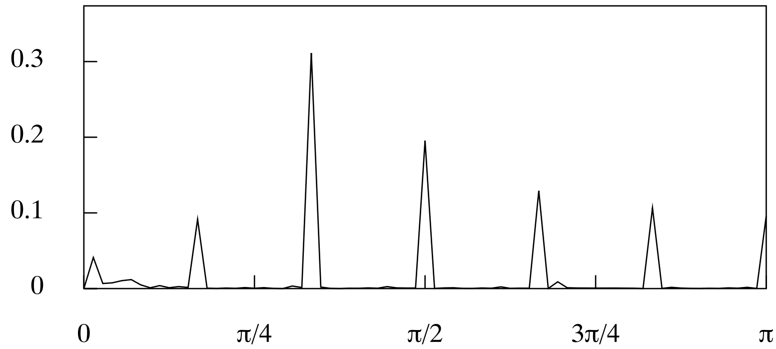

Fig. 12. The periodogram of residuals obtained by fitting a linear trend through the logarithmic sales data of Figure 11.

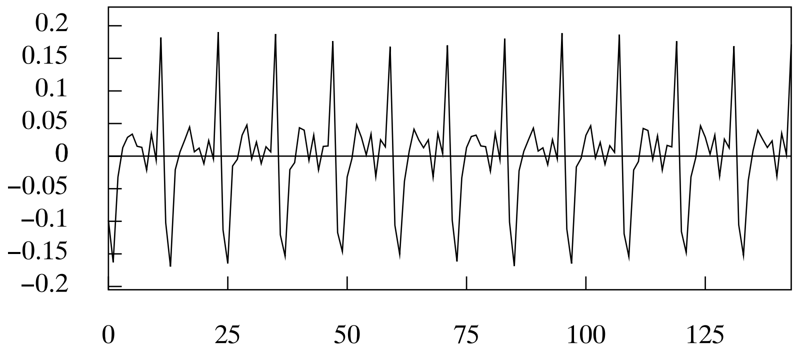

Fig. 13. The seasonal fluctuations of the data of Figure 11, which are virtually identical to the deviations of the data from the interpolated trend.

adding a filtered low-frequency component to the linear function of Figure 4 that served to detrend the data.

8 Extrapolations

A symmetric moving-average filter with $2q + 1$ constant coefficients is unable to process the first q and the final q sample points. Therefore, within the programs that are commonly employed to produce seasonally adjusted data, it has become common to use estimated ARIMA models to extrapolate the data so that linear time-invariant filters can reach the ends of the sample.

The frequency-domain methods that are described in the paper involve a mapping of the T data points onto the circumference of a circle. The equivalent time-domain filter would be a moving average of order T that rotates around the circle.

The circular disposition of the data ensures that there are no end-of-sample effects of the sort described above. However, a careful detrending is required so as to ensure that there are no radical disjunctions in start-finish area where the end of the data sequence is joined to its beginning. This is achieved within the program, where necessary, by ensuring that the trend function passes through the midst of the scatter of points at the ends of the sample.

In addition to such detrending, one might employ the device of tapering to ensure that both ends of the detrended data decline to the horizontal axis. (see Bloomfield 1976, for example.) The disadvantage of applying tapering to the sample is that it falsifies the values at the ends, which is particularly inconvenient if, as is often the case in economics, attention is focussed on the most recent data.

To avoid this difficulty, the program applies tapering only to some appropriate extrapolations of the sample, which can be added to it either before or after it has been detrended. At the end of the process, the extrapolations are discarded. When such extrapolations are used, the circular nature of the theoretical moving-average operator is effectively effaced.

Although this moving average has a span equal to the circumference of the circular data, its coefficients are liable to be highly concentrated around its centre—with values that decline rapidly as the distance from the centre increases. Therefore, the circularity of the convolution process will only be apparent, if at all, where the beginning and the end of the data sequence are joined. If these are the ends of the extrapolations, as opposed to the ends of sample, and if the extrapolations are to be discarded, then there will be no effective circularity.

The detended data sequence can be extrapolated by reflecting it around the endpoints. Then, the additions can be truncated and tapered. However, if the data show strong seasonal fluctuations, then a tapered sequence based on successive repetitions of the ultimate seasonal cycle can be added to the upper end and a similar sequence can be added to the beginning

An alternative extrapolation is available for monthly seasonal data that does not employ any tapering. Instead, it interpolates a sequence into the circle, between the start and the finish of the detrended data, in which, as one travels in an anticlockwise direction (i.e. the direction of time), the seasonal

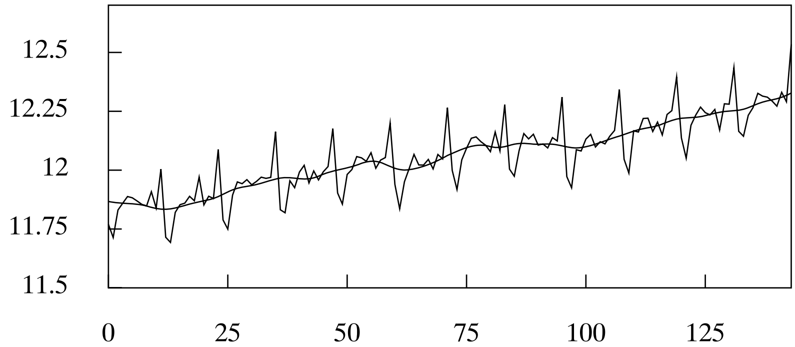

Fig. 14. The logarithms of the data on U.S. retail sales with an interpolated trend-cycle determined by the X-12-ARIMA program.

pattern of the last twelve months gradually morphs into the pattern of the first twelve months.

Example. A data sequence that demands to handled with care is provided by the 144 monthly observations on retail sales in the US for the years 1953 to 1964, which were recorded in the paper of Shiskin, Young and Musgrave (1967) that presented the X-11 program of the U.S. Census Bureau. This program, which was based on the Henderson (1916) filter, is intended for seasonal adjustment and trend estimation.

Figure 11 show the logarithms of the data, and Figure 12 shows the periodogram of the residuals from fitting a linear trend. The periodogram has a prominent spike at the seasonal frequency of $\pi/6$ and at the harmonic frequencies of $\pi/3$, $\pi/2$, $2\pi/3$, $5\pi/6$ and π. With the exception of the interval $(2\pi/3, 5\pi/6)$, which does contain one significant ordinate, the interstices between these seasonal frequencies are virtual dead spaces. The presence of some nonzero ordinates in the interval $[0, \pi/6)$, which covers the trend frequencies, indicates that the log-linear detrending is inadequate.

The trend-cycle component that is portrayed in Figure 11 is based on the Fourier ordinates of the twice-differenced data that lie in the interval $[0, \pi/6)$. To avoid some distortionary end effects, the sample has been extrapolated. A linear function has been interpolated through the data and tapered sequences based on repetitions of the first and the final year have been added to the lower and upper branches of the trend line. The exact details are recorded in a log file that accompanies the program mentioned in the introduction.

Figure 13 shows a sequence of seasonal fluctuations that has been synthesised from a selection of the Fourier ordinates of the twice-differenced data. The ordinates correspond to the seasonal frequency and its harmonics and to the two frequency points immediately above $2\pi/3$. The periodogram of Figure 12 has provided the necessary guidance in selecting these frequencies.

The differenced seasonal vector, synthesised from these elements, may be denoted by k. The vector h of the seasonal fluctuations is obtained by cumulating k via the formula $h = S_*k_* + Sk$ of (50), wherein k_* is calculated according to (52). It transpires that this synthesised seasonal vector, which is represented by Figure 13, it virtually indistinguishable from the vector of the residuals obtained by subtracting the trend-cycle vector x of Figure 11 from the corresponding data vector y.

It is interesting to compare these results with those of Figure 14 that have been obtained via the X-12-ARIMA program of the U.S. Bureau of the Census. This program has been described in detail by Findley *et al* (1998); and it has been implemented in conjunction with the SEATS–TRAMO program of Caporello and Maravall (2004) by the Statistical Office of the European Communities under a common interface named Demetra—see Eurostat Eurostat (2002).

The representation of the trend-cycle in Figure 14 is very similar to that of Figure 11, but there is a perceptible difference in the two at the beginning of the sample, which is due to different ways of handling the end-of-sample problem. In the case of X-12-ARIMA, a seasonal autoregressive moving-average model of order $(p, d, q) \times (P, D, Q) = (0, 1, 1) \times (0, 1, 1)$, determined automatically by the program, has been used to create the extrapolations. These consist of indefinite replications of the seasonal patterns of the first and the final year of the data superimposed upon horizontal linear extrapolations.

In the case of the Figure 11, the asymptote of the extrapolations is an inclined straight line fitted by weighted least-square regression to the logarithmic data. The pre-sample extrapolation of the data declines towards this asymptote, with the effect that the estimated trend-cyle is somewhat depressed at the beginning of the sample.

This is appropriate if it is deemed that the values of first year of the sample were above the trend. The effect can be alleviated by ensuring that straight line passes through the scatter of points at the beginning of the sample, by giving extra weight to these points. In that case, the results will resemble those of Figure 14 more closely.

A Appendix: The Wrapped Coefficients of the Ideal Lowpass Filter

Let the sample size be T, and consider a set of T frequency-domain ordinates sampled, at the Fourier frequencies $\omega_j = 2\pi j/T$ that fall within the interval $[-\pi, \pi)$, from a boxcar function, centred on $\omega_0 = 0$. If the cut-off points are at $\pm\omega_d = \pm d\omega_1 = \pm 2\pi d/T$, then the ordinates of the sample will be

$$\lambda_j = \begin{cases} 1, & \text{if } j \in \{1-d, \ldots, d-1\}, \\ 1/2, & \text{if } j = \pm d, \\ 0, & \text{otherwise.} \end{cases} \tag{A.1}$$

Their (discrete) Fourier transform is the sequence of the coefficients

$$\phi_k^\circ = \frac{1}{T}\sum_j \lambda_j e^{\mathrm{i}\omega_j k}, \tag{A.2}$$

defined for $k = 0, 1, \ldots, T-1$.

The ordinates $\lambda_d = \lambda_{-d} = 1/2$ cause some inconvenience in evaluating this transform. To overcome this, we may begin by evaluating the function

$$S^+(z) = z^{1-d} + \cdots + z^{-1} + 1 + z + \cdots + z^d, \tag{A.3}$$

where $z = e^{\mathrm{i}\omega}$, together with the function $S^-(z) = z^{-1}S^+(z)$. Then we may form the symmetric function $\phi^\circ(z) = \{S^-(z) + S^+(z)\}/(2T)$, wherafter we may set $\omega = \omega_k = 2\pi k/T = k\omega_1$ to obtain the kth coefficient.

First, consider

$$\begin{aligned} S^+(z) &= z^{1-d}(1 + z + \cdots + z^{2d-1}) \\ &= z^{1-d}\frac{(1-z^{2d})}{1-z}. \end{aligned} \tag{A.4}$$

Multiplying top and bottom by $z^{-1/2}$ gives

$$\begin{aligned} S^+(z) &= z^{(1/2)-d}\frac{(1-z^{2d})}{z^{-1/2}-z^{1/2}} \\ &= z^{1/2}\frac{(z^{-d}-z^d)}{z^{-1/2}-z^{1/2}}. \end{aligned} \tag{A.5}$$

Then, by setting $z = e^{i\omega}$, we get

$$S^+(e^{\mathrm{i}\omega}) = e^{\mathrm{i}\omega/2}\frac{(e^{\mathrm{i}\omega d} - e^{-\mathrm{i}\omega d})}{e^{\mathrm{i}\omega/2} - e^{-\mathrm{i}\omega/2}} = e^{\mathrm{i}\omega/2}\frac{\sin(\omega d)}{\sin(\omega/2)}. \tag{A.6}$$

When $\omega = \omega_k = k\omega_1$, this becomes

$$S^+(k) = e^{\mathrm{i}\omega_1 k/2}\frac{\sin(d\omega_1 k)}{\sin(\omega_1 k/2)}, \tag{A.7}$$

which is the Dirichlet function multiplied by a complex exponential that owes its presence to the non-symmetric nature of $S^+(z)$. There is also

$$S^-(k) = e^{-\mathrm{i}\omega_1 k/2}\frac{\sin(d\omega_1 k)}{\sin(\omega_1 k/2)}, \tag{A.8}$$

Therefore, for $k \neq 0$, there is

$$\phi^\circ(k) = \frac{1}{2T}\{S^-(k) + S^+(k)\} = \frac{\cos(\omega_1 k/2)\sin(d\omega_1 k)}{T\sin(\omega_1 k/2)}, \tag{A.9}$$

whereas for $k = 0$ there is $\phi_0^\circ = 2d/T$, which comes from setting $z = e^0 = 1$ in the expression for $S^+(z)$ of (A.3) and in the analogous expression for $S^-(z)$.

Setting $d = T/4$ in (A.9) gives the wrapped version of the lowpass half band filter that is the subject of Section 2 of the paper.

In an alternative specification of the ideal filter, the cut-off points fall between Fourier frequencies. Then, the ordinates sampled from the frequency response function are

$$\lambda_j = \begin{cases} 1, & \text{if } j \in \{1-d, \ldots, d-1\}, \\ 0, & \text{otherwise.} \end{cases} \tag{A.10}$$

In place of $\{S^-(z) + S^+(z)\}/2$, there is

$$\begin{aligned} S(z) &= z^{1-d} + \cdots z^{-1} + 1 + z + \cdots z^{d-1} \\ &= \frac{z^{(1/2)-d} - z^{d-1/2}}{z^{-1/2} - z^{1/2}}. \end{aligned} \tag{A.11}$$

Then, for $k = 0$, there is $\phi_0^\circ = (2d-1)/T$, whereas, for $k \neq 0$, the formula is

$$\phi^\circ(k) = \frac{\sin([d-1/2]\omega_1 k)}{T\sin(\omega_1 k/2)}. \tag{A.12}$$

References

[1] Baxter, M., and King, R.G., 1999. Measuring business cycles: approximate bandpass filters for economic time series. *Review of Economics and Statistics,* 81, 575–593.

[2] Bell, W.M., and Martin, D.E.K., 2004. Computation of asymmetric signal extraction filters and mean squared error for ARIMA components models. *Journal of Time Series Analysis,* 25, 603–623.

[3] Bloomfield, P., 1976. *Fourier Analysis of Time Series: An Introduction.* John Wiley and Sons, New York.

[4] Burns, A.M., and Mitchell, W.C., 1946. *Measuring Business Cycles.* New York: National Bureau of Economic Research.

[5] Caporello, G., and Maravall, A., 2004. *Program TSW: Revised Reference Manua.* Working Paper, Research Department, Banco de España.

[6] Christiano, L.J., and Fitzgerald, T.J., 2003. The bandpass filter. *International Economic Review,* 44, 435–465.

[7] Eurostat, 2002. *Demetra 2.0 User Manual: Seasonal Adjustment Interface for Tramo/Seats and X12-ARIMA.* the Statistical Office of the European Communities.
`http://circa.europa.eu/irc/dsis/eurosam/info/data/demetra.htm`

[8] Findley, D.F., Monsell, B.C, Bell, W.R., Otto, M.C., and Chen, B., 1998. New capabilities and methods of the X12-ARIMA seasonal adjustment program. *Journal of Business and Economic Statistics,*16, 127–152.

[9] Findley, D.F, and Martin, D.E.K., 2003. *Frequency Domain Analyses of SEATS and X–11/12–ARIMA Seasonal Adjustment Filters for Short and Moderate-Length Time Series*. Statistical Research Division U.S. Bureau of the Census Washington D.C. 20233.

[10] Haywood, J., and Tunnicliffe Wilson, G., 1997. Fitting time series models by minimising multistep-ahead errors: a frequency domain approach. *Journal of the Royal Statistical Society, Series B (Methodological),* 59, 237–254.

[11] Henderson, R., 1916. Note on graduation by adjusted averages. *Transactions of the Actuarial Society of America,* 17, 43–48.

[12] Iacobucci, Alessandra, and Noullez, A., 2005. A frequency selective filter for short-length time series. *Computational Economics,* 25, 75–102.

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

[14] Pollock, D.S.G., 2002. Circulant matrices and time-series analysis. *The International Journal of Mathematical Education in Science and Technology,* 33, 213–230.

[15] Proietti, T., 2005. Forecasting and signal extraction with misspecified Models. *Journal of Forecasting,* 24, 539–556.

[16] Shiskin, J., Young A.H., and Musgrave, J.C., 1967. *The X-11 Variant of the Census Method II Seasonal Adjustment Program.* Technical Paper no. 15, U.S. Dept. of Commerce, Bureau of the Census.
`http://www.census.gov/srd/www/historicpapers.html`