# ECONOMETRIC FILTERS

by D.S.G. Pollock

University of Leicester

A variety of filters that are commonly employed by econometricians are analysed with a view to determining their effectiveness in extracting well-defined components of economic data sequences.

These components can be defined in terms of their spectral structures—i.e. their frequency content—and it is argued that the process of econometric signal extraction should be guided by a careful appraisal of the periodogram of the detrended data sequence.

Whereas it is true that many annual and quarterly economic data sequences are amenable to relatively unsophisticated filtering techniques, it is often the case that monthly data that exhibit strong seasonal fluctuations require a far more delicate approach.

In such cases, it may be appropriate to use filters that work directly in the frequency domain by selecting or modifying the spectral ordinates of a Fourier decomposition of data that have been subject to a preliminary detrending.

## The Frequency Domain

A sequence $y(t) = \{y_t; t = 0, 1, \ldots, T-1\}$ can be projected onto a trigonometrical basis:

$$y_t = \sum_{j=0}^{[T/2]} \rho_j \cos(\omega_j t + \theta_j) = \sum_{j=0}^{[T/2]} \{\alpha_j \cos(\omega_j t) + \beta_j \sin(\omega_j t)\}.$$

Here, the Fourier frequencies $\omega_j = 2\pi j/T; j = 0, 1, \ldots, [T/2]$ are evenly distributed in the interval $[0, \pi]$ and $[T/2]$ is the integer quotient of the division of $T$ by 2.

Euler's equations enable us to express the trigonometrical functions in terms of complex exponentials:

$$\cos\theta = \frac{e^{\mathrm{i}\theta} + e^{-\mathrm{i}\theta}}{2} \quad \text{and} \quad \sin\theta = \frac{-\mathrm{i}}{2}(e^{\mathrm{i}\theta} - e^{-\mathrm{i}\theta}).$$

Define $\zeta_j = (\alpha_j - \mathrm{i}\beta_j)/2$ and $\zeta_{-j} = (\alpha_j + \mathrm{i}\beta_j)/2 = \zeta_{T-j}$. Then, we have

$$y_t = \sum_{j=1-[T/1]}^{[T/2]} \zeta_j e^{\mathrm{i}\omega_j t} = \sum_{j=0}^{T-1} \zeta_j e^{\mathrm{i}\omega_j t}.$$

## The Periodogram and the Spectrum

The periodogram is the sequence of the squared amplitude coefficients $\rho_j^2 = \alpha_j^2 + \beta_j^2$ scaled by $T$. It is also the Fourier transform of the autocovariance function $c(\tau)$:

$$T\rho_j^2 = \sum_{\tau=1-T}^{T-1} c_\tau \cos(\omega_j \tau).$$

A stationary stochastic process also has an expression in terms of trigonometrical functions, described as its spectral representation:

$$y(t) = \int_0^\pi \Big\{ \cos(\omega t) dA(\omega) + \sin(\omega t) dB(\omega) \Big\} = \int_{-\infty}^{\infty} e^{\mathrm{i}\omega t} dZ(\omega).$$

Here, $dZ(\omega) = \{A(\omega) - \mathrm{i}B(\omega)\}/2$, where $A(\omega)$, $B(\omega)$ are mutually uncorrelated stochastic processes with infinitesimal increments $dA(\omega)$, $dB(\omega)$ such that

$$E\{dA(\omega)\} = E\{dB(\omega)\} = 0 \quad \text{and} \quad V\{dA(\omega)\} = V\{dB(\omega)\} = 2dF(\omega) = 2f(\omega)d\omega.$$

The analytic function $f(\omega)$ is the spectrum of the process, which is also the Fourier transform of the autocovariance sequence $\gamma(\tau)$:

$$f(\omega) = \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} \gamma_\tau \cos(\omega_\tau) = \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} \gamma_\tau e^{\mathrm{i}\omega_\tau}.$$

## A Requirement for Detrending

In a Fourier analysis, the data are treated as a single cycle of a periodic function defined on the real line or on the circumference of a circle. A trended sequence will given rise to a sawtooth function, which has a one-over-$f$ periodogram, with a dominant low-frequency component.

To assess the cyclical elements of the data, one must detrend the data. The ordinates of a linear trend function are given by

$$\begin{aligned} x &= y - Q(Q'Q)^{-1}Q'y \\ &= y - e, \end{aligned}$$

where $e$ is the vector of the residual sequence, and where

$$Q' = \begin{bmatrix} 1 & -2 & 1 & \ldots & 0 & 0 \\ 0 & 1 & -2 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & 0 \\ 0 & 0 & 0 & \ldots & -2 & 1 \end{bmatrix}$$

is the matrix version of the twofold difference operator.

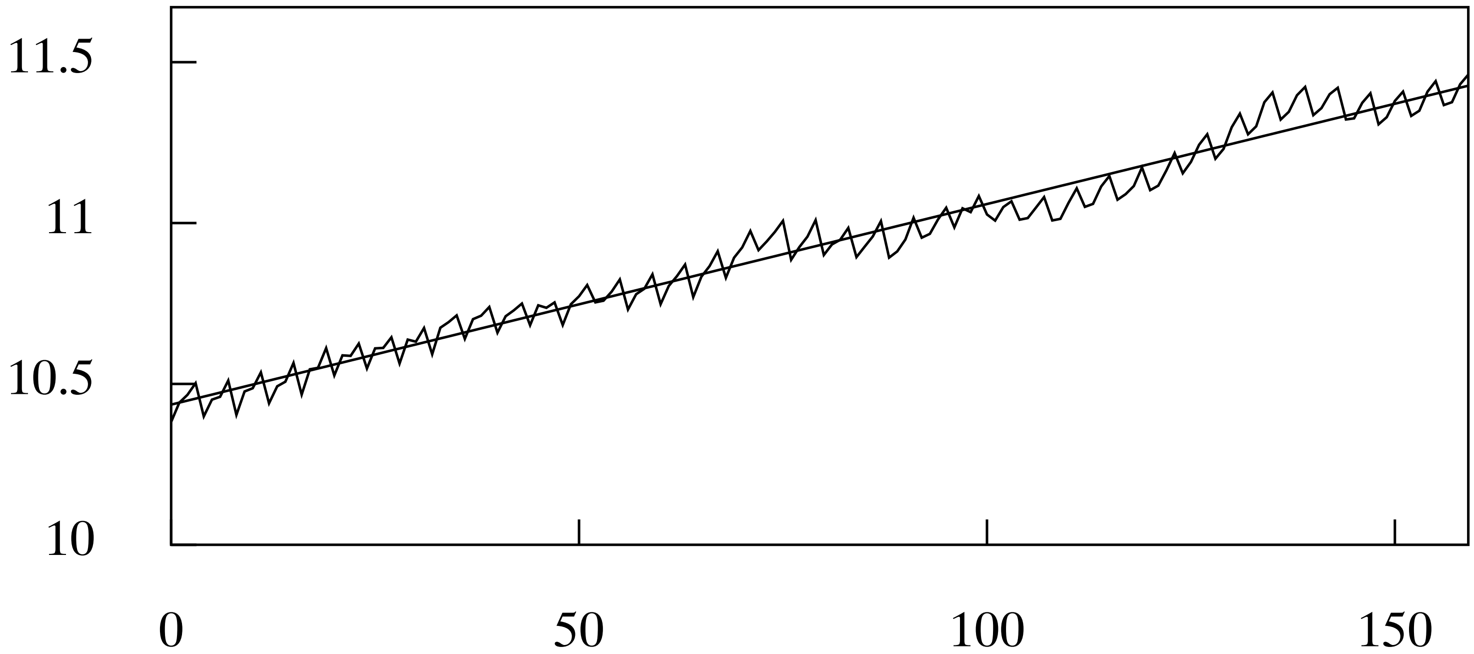


**Figure 1.** The quarterly sequence of the logarithms of household consumption expenditure in the U.K. for the years 1955 to 1994 with an interpolated linear trend.

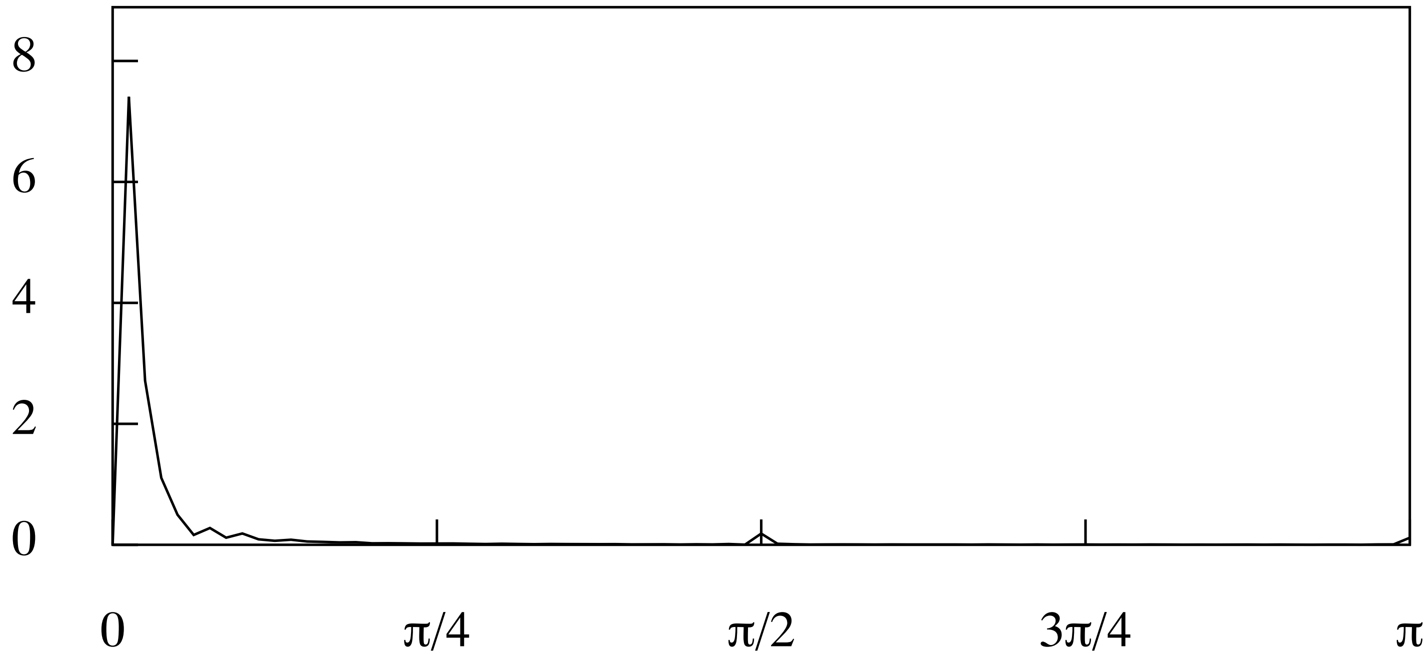


**Figure 2.** The periodogram of the logarithmic consumption data.

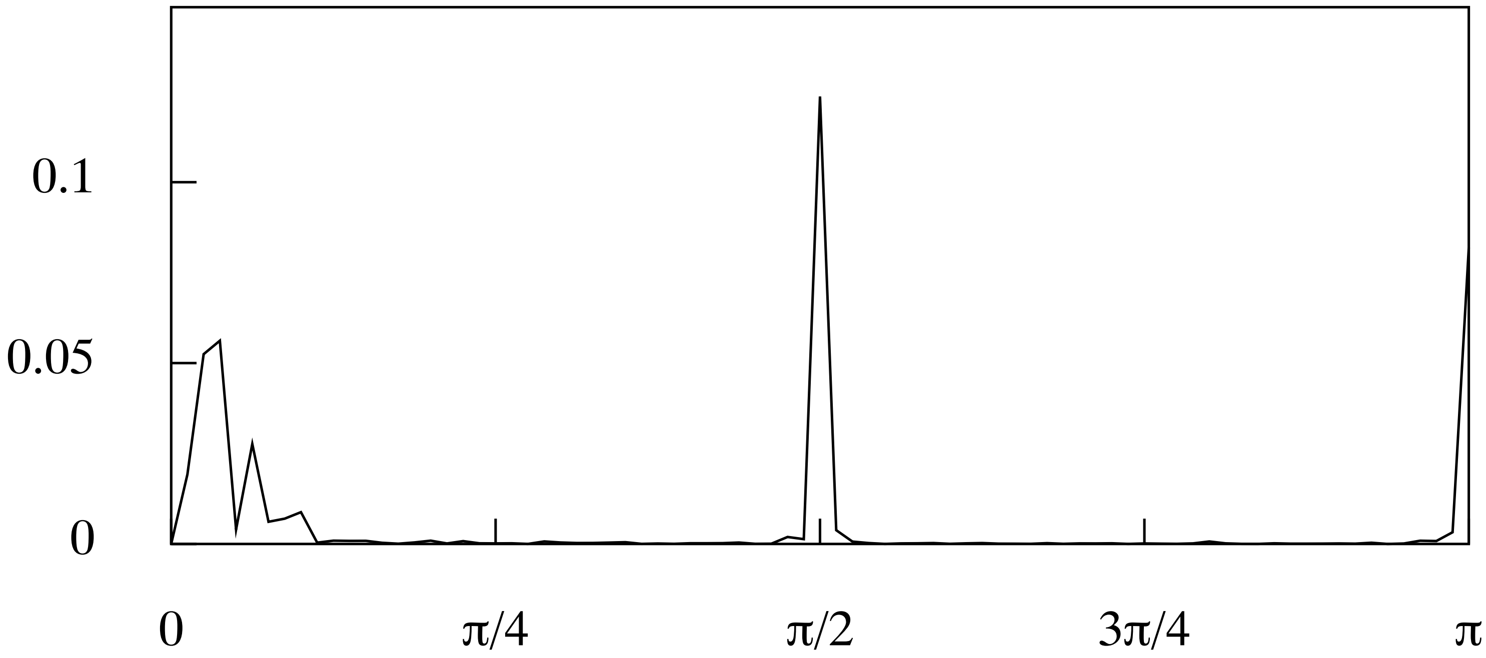


**Figure 3.** The periodogram of the residual sequence from the linear detrending of the logarithmic consumption data.

## Local Polynomials: The Henderson Filters

The Henderson filters are derived fitting a polynomial to the points that fall within a moving window, spanning $2m+1$ data points. At each step, the central ordinate of the fitted polynomial replaces the corresponding data value.

The outcome is a set of moving-average coefficients $\psi_j; j = 0, \pm 1, \ldots, \pm m$ disposed symmetrically around the central value $\psi_0$, with $\psi_{-j} = \psi_j$. These are applied throughout the sample except at the beginning and the end.

In the filters derived by Robert Henderson (1916, 1924), a cubic polynomial is fitted to the windowed data points. They are used within the X-11 family of seasonal adjustment programs.

Figure 4 displays the coefficients of the 23-point Henderson filter and Figure 5 shows the effect of applying the Henderson filter directly to the logarithmic consumption data. The filter appears to do a reasonable job of estimating the trend-cycle function.

Notice that the filter runs to the ends of the sample, whereas one might expect it to fall short, leaving $m = 11$ points unprocessed at either end. This is achieved by adapting the coefficients of the filter as it nears the ends.

## The Effects of a Linear Filter

A linear filter combines the values of an input sequence $x(t)$ to create an output sequence

$$y(t) = \sum_j \psi_j x(t-j).$$

The effects of the filter can be shown by considering a complex exponential input sequence of the form $x(t) = \cos(\omega t) + \mathrm{i}\sin(\omega t) = \exp\{\mathrm{i}\omega \mathrm{t}\}$. The corresponding output is

$$y(t) = \sum_j \psi_j e^{\mathrm{i}\omega(t-j)} = \left\{\sum_j \psi_j e^{-\mathrm{i}\omega j}\right\} e^{\mathrm{i}\omega t} = \psi(\omega)e^{\mathrm{i}\omega t}.$$

The effects are summarised by the complex function

$$\psi(\omega) = |\psi(\omega)|e^{-\theta(\omega)}.$$

The modulus $|\psi(\omega)|$ alters the amplitudes of the cyclical elements of the data, which is the *gain effect.* The argument $\theta(\omega)$ displaces the elements in time, which is the *phase effect.*

A phase effect can be avoided if the filter coefficients are disposed symmetrically about a central point, such that the filter reaches equally forward and backwards in time.

## The Gain of the Henderson Filter

In the case of a symmetric Henderson filter, where $\psi_j = \psi_{-j}$, the associated complex exponential functions combine to form $\cos(\omega_j) = \{\exp(-\mathrm{i}\omega j) + \exp(\mathrm{i}\omega j)\}/2$. Therefore, the frequency response function is real-valued and there is no phase effect.

The frequency response of the filter allows the business-cycle component of the consumption data to be transmitted in full. There are no other significant elements of the data that fall within the pass band of the filter. Therefore, it serves the purpose of extracting the trend-cyle function well enough.

It will be observed that the frequency response function of the Henderson filter, in Figure 6, shows a very gradual transition from the pass band, where the elements of the Fourier decomposition are fully preserved, to the stop band, where they should be wholly nullified. There are circumstances where one would wish to have a more rapid transition.

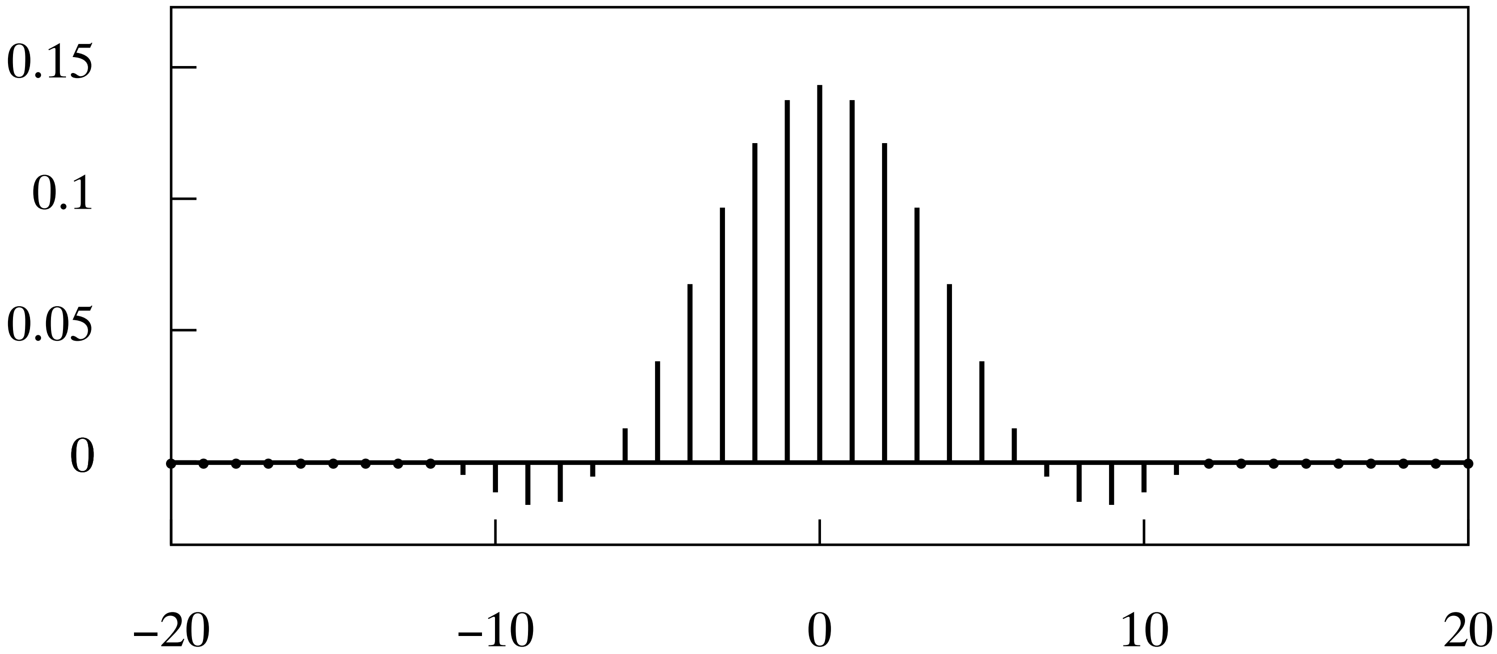


**Figure 4.** The coefficients of the symmetric Henderson filter of 23 points.

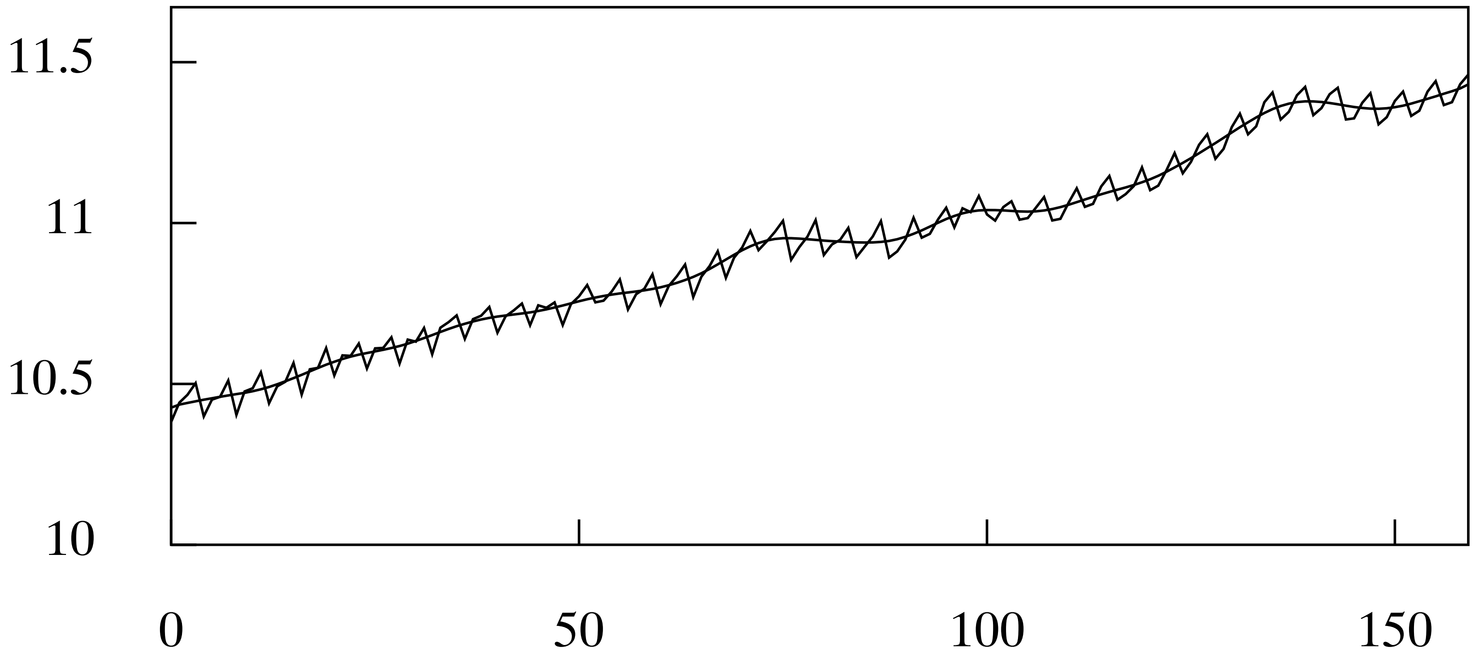


**Figure 5.** A trend, determined by a Henderson filter with 23 coefficients, interpolated through the 160 points of the logarithmic consumption data.

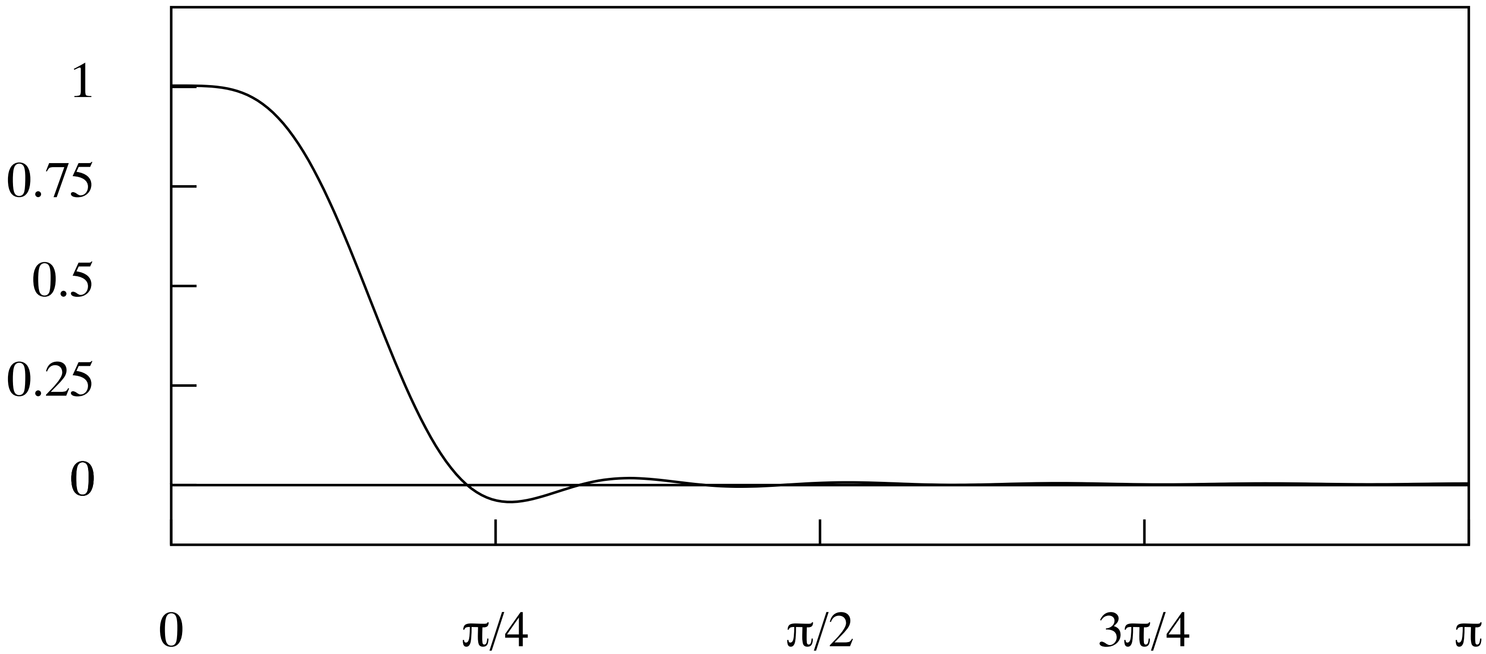


**Figure 6.** The frequency response function of the Henderson moving-average filter of 23 terms.

## Approximate Bandpass Filters

According to the definition of Burns and Mitchell (1946), the business cycle comprises all the elements of the data that have cyclical durations of no less than one-and-a-half years and of no more than eight years.

Baxter and King (1999) have formed a filter from the inverse Fourier transform of the rectangle, defined on $[\alpha, \beta] \in [0, \pi]$, that constitutes the ideal frequency response. For quarterly data, the values in radians are $\alpha = \pi/16$ ($11.25^\circ$) and $\beta = \pi/3$ ($60^\circ$).

The Fourier transform of a frequency-domain rectangle is a doubly-infinite sequence of filter coefficients, of which the central values are displayed in Figure 7. The coefficients are the sampled ordinates of the function

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

described as a displaced sinc function, where $k \in \{0, \pm 1, \pm 2, \ldots\}$. Here, $\gamma$, which is the displacement parameter, represents the centre of the pass band, whereas $\delta$ is half its width.

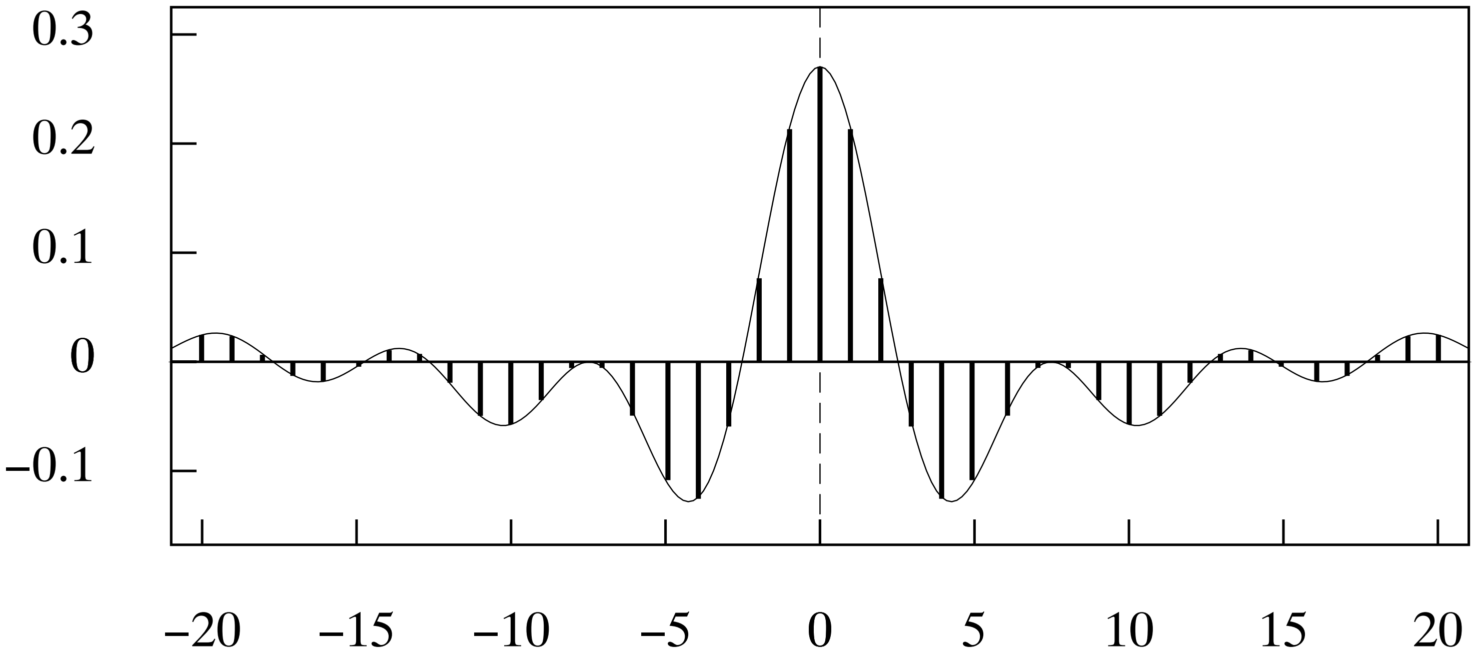


**Figure 7.** The central coefficients of the ideal bandpass filter defined on the frequency interval $[\pi/16, \pi/3]$.

## Truncations of the Ideal Filter

The sequence of coefficients must be drastically truncated. Figure 8 shows the frequency response of a truncated band pass filter of 25 coefficients, compared with the rectangle of the ideal frequency response.

The truncated filter allows elements within the stop bands to be transmitted to a significant extent. This so-called problem of leakage greatly subverts the original intentions.

Christiano and Fitzgerald (2001) assumed that a random-walk process has generated the data. Then, the optimal forecasts and backcasts are obtained by horizontal extrapolations of the values at the ends of the sample.

Thus, the filtered value at time $t$ may be denoted by

$$\begin{aligned} x_t = Ay_0 + \psi_t y_0 + \cdots + \psi_1 y_{t-1} + \psi_0 y_t \\ + \psi_1 y_{t+1} + \cdots + \psi_{T-1-t} y_{T-1} + By_{T-1}, \end{aligned}$$

where $A$ and $B$ are the sums of the extra-sample coefficients at either end.

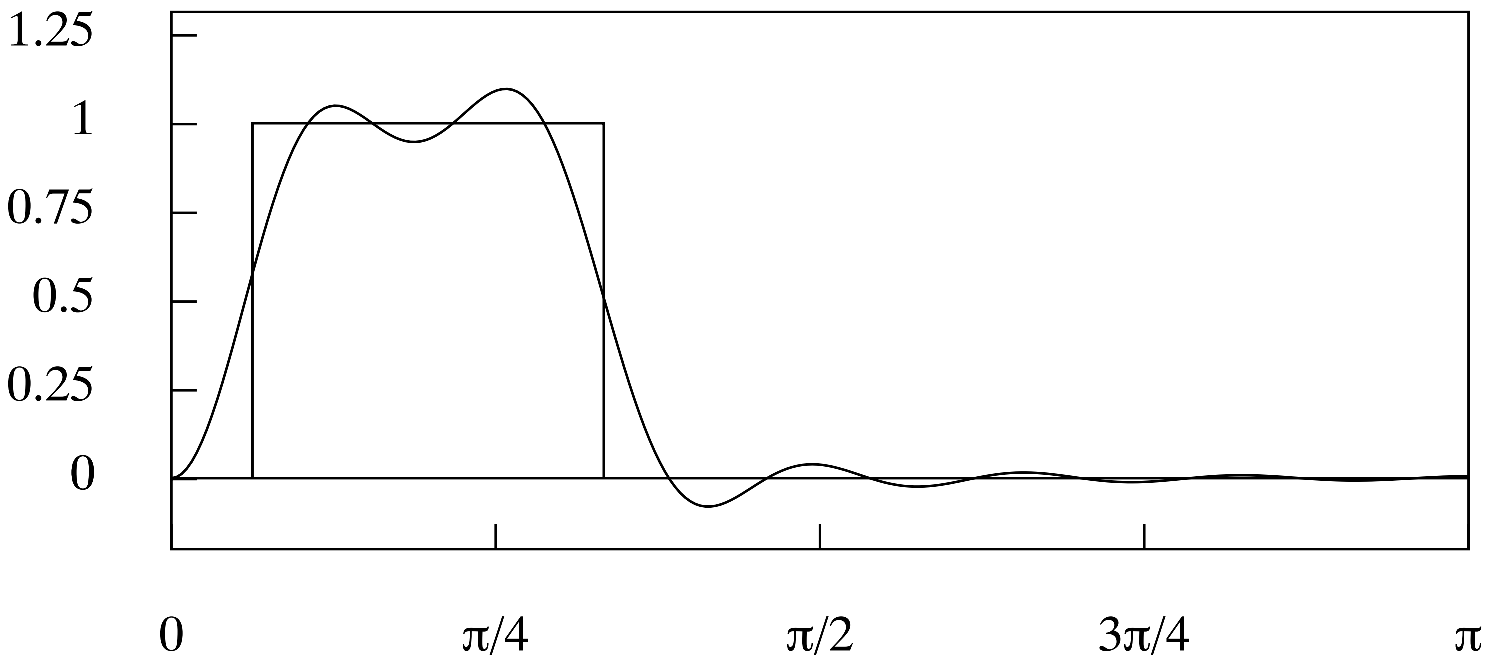


**Figure 8.** The rectangular frequency response of the ideal bandpass filter defined on the interval $[\pi/16, \pi/3]$, together with the frequency response of the truncated filter of 25 coefficients.

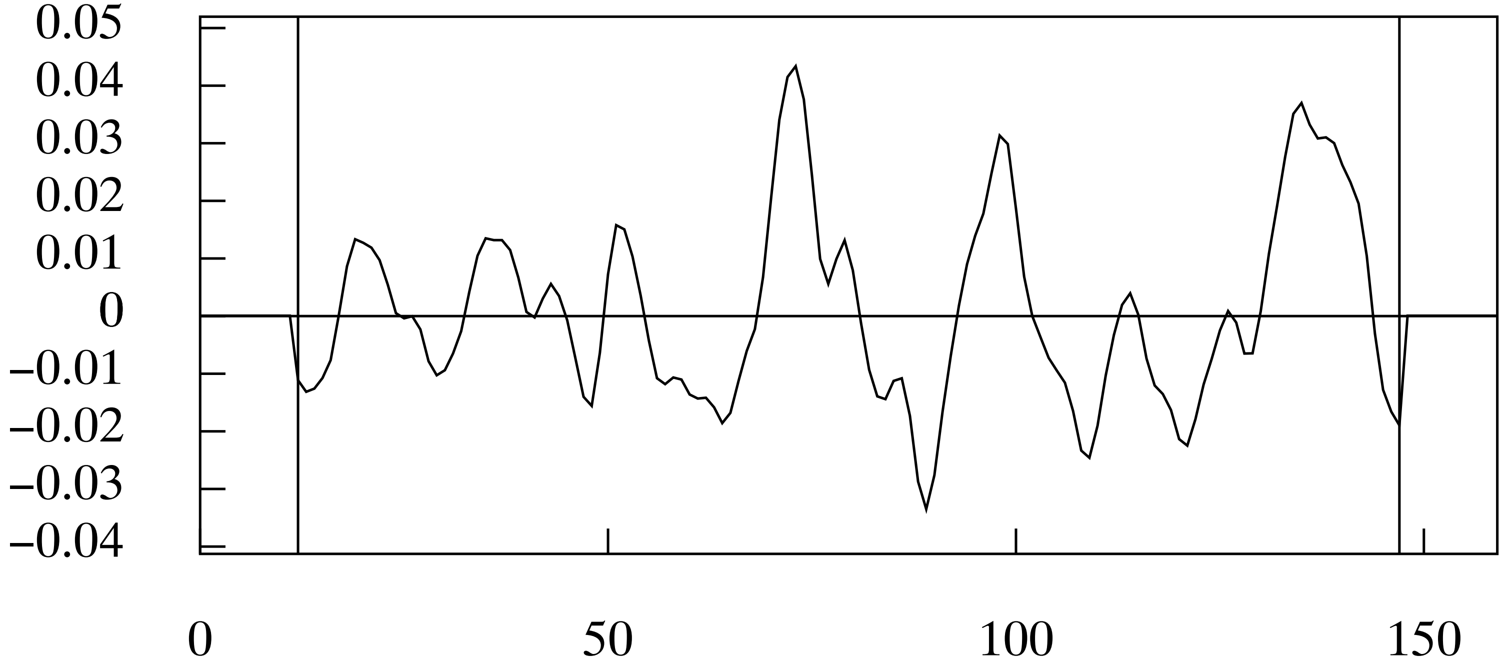


**Figure 9.** The effect of applying the truncated bandpass filter of 25 coefficients to the quarterly logarithmic data on U.K. consumption.

## Dealing with Trended Data

Trended data should be reduced to a mean-reverting residual sequence with a zero mean. Then, the extra-sample values of the detrended sequence may be represented by zeros, which might stand for their unconditional expectations.

Instead of fitting a truncated filter within the confines of a finite data sequence, one can run the data sequence along the central part of the infinite sequence of filter coefficients. Then, the data are treated as the moving average and the filter coefficients are treated as the data.

This is equivalent to setting the required extra-sample values to zero. An alternative interpretation of this procedure is derived by considering a banded Toeplitz matrix $\Psi$ of the same order as the sample.

In the case of $T = 4$, there is

$$\Psi = \begin{bmatrix} \psi_0 & \psi_1 & \psi_2 & \psi_3 \\ \psi_1 & \psi_0 & \psi_1 & \psi_2 \\ \psi_2 & \psi_1 & \psi_0 & \psi_1 \\ \psi_3 & \psi_2 & \psi_1 & \psi_0 \end{bmatrix}.$$

The vector $x = \Psi d$ of the filtered values is obtained by premultiplying the vector $d$ of the detrended data by this matrix.

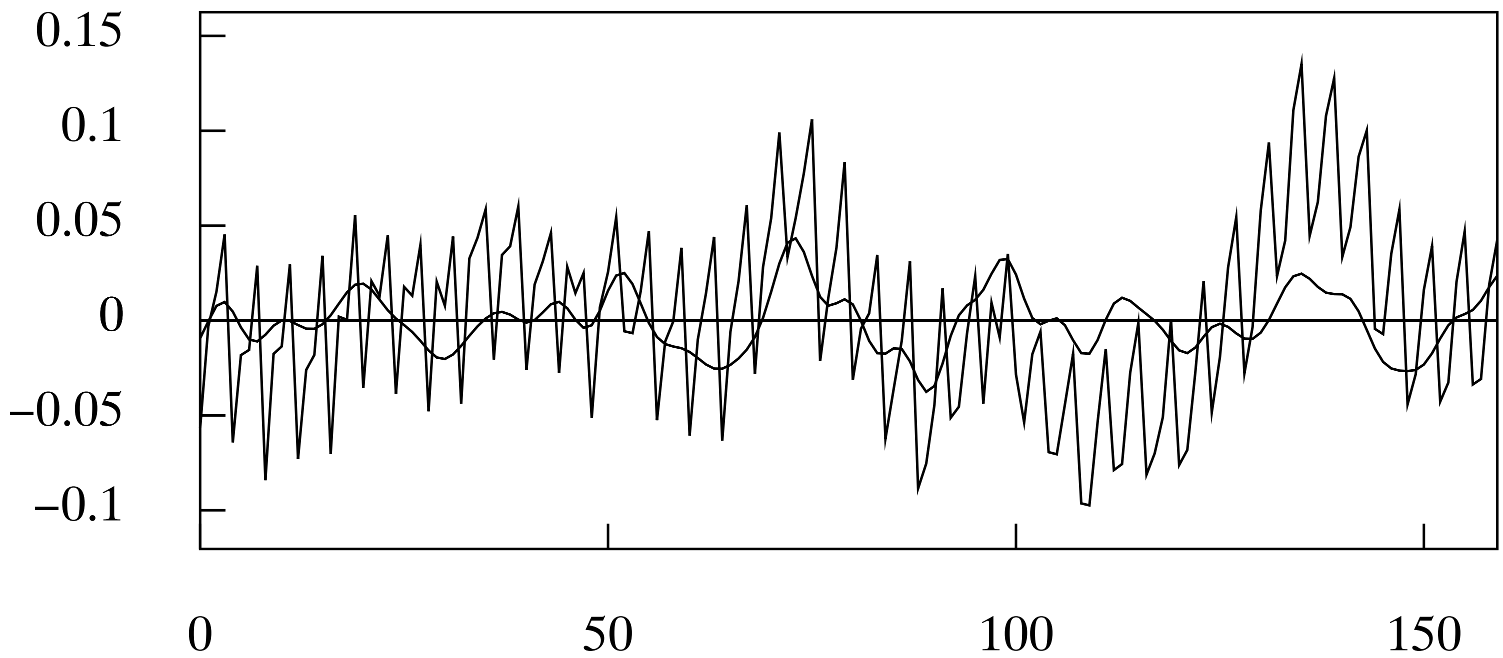


**Figure 10.** The effect of applying the the filter of Christiano and FItzgerald to the quarterly logarithmic data on U.K. consumption.

## Wiener–Kolmogorov Filters

A time-invariant IIR filter can be expressed as the ratio of $z$-transform polynomials:

$$\psi(z) = \frac{\theta(z^{-1})\theta(z)}{\phi(z^{-1})\phi(z)}.$$

In place of $z$, one might put the lag operator, familiar to econometricians. Such a filter must, of necessity, be applied to $y(t)$ to produce $x(t)$ in two separate passes, running forwards and backwards in time and described, respectively, by the equations

$$\text{(i)} \quad \phi(z)q(z) = \theta(z)y(z) \qquad \text{and} \qquad \text{(ii)} \quad \phi(z^{-1})x(z) = \theta(z^{-1})q(z).$$

A lowpass Wiener Kolmogorov filter is aimed at extracting the signal component $\xi(t)$ from a data sequence $y(t) = \xi(t) + \eta(t)$, where $\eta(t)$ is noise. It takes the form of

$$\psi(z) = \frac{\gamma^{\xi\xi}(z)}{\gamma^{\xi\xi(z)} + \gamma^{\eta\eta(z)}},$$

where $\gamma^{\xi\xi}(z)$ and $\gamma^{\eta\eta}(z)$ are the autocovariance generating functions of the signal and the noise respectively. The complementary highpass filter is

$$\psi^c(z) = 1 - \psi(z) = \frac{\gamma^{\eta\eta}(z)}{\gamma^{\xi\xi}(z) + \gamma^{\eta\eta}(z)}.$$

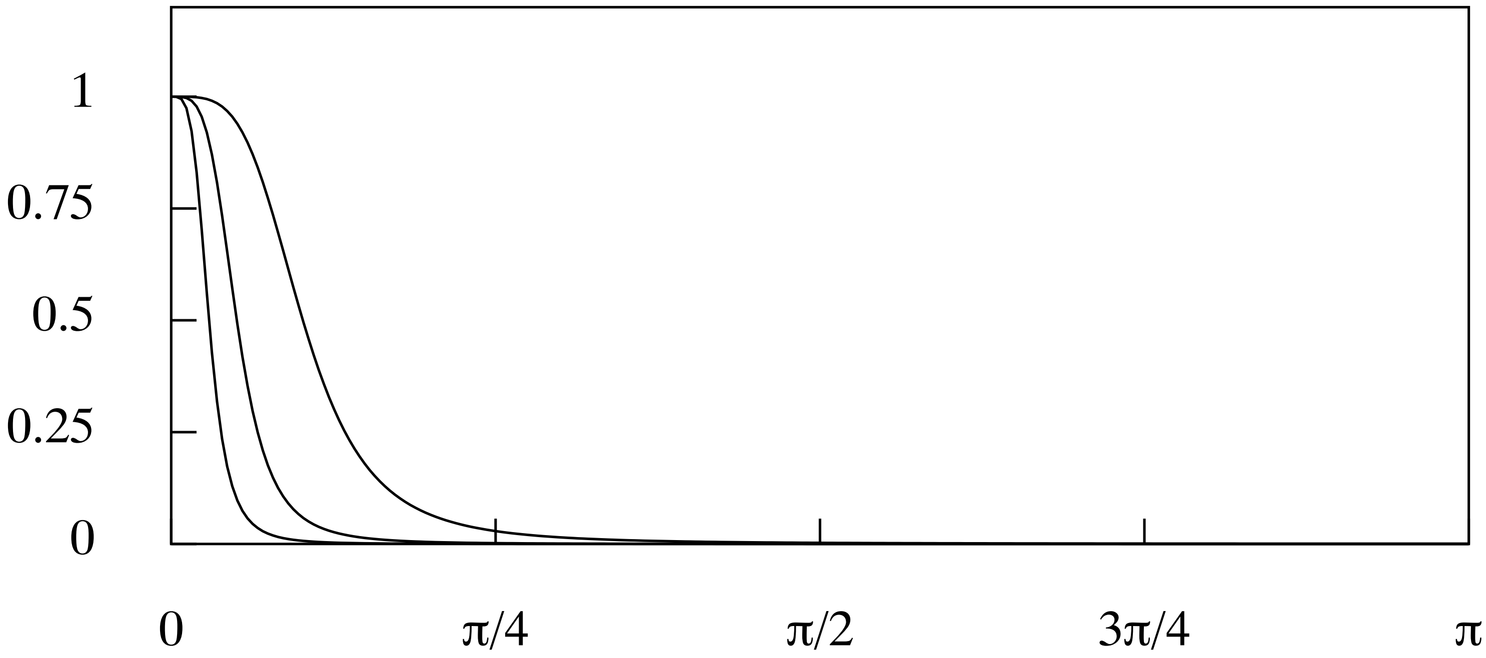


**Figure 11.** The frequency response function of the Hodrick–Prescott lowpass smoothing filter—or Leser filter—for various values of the smoothing parameter.

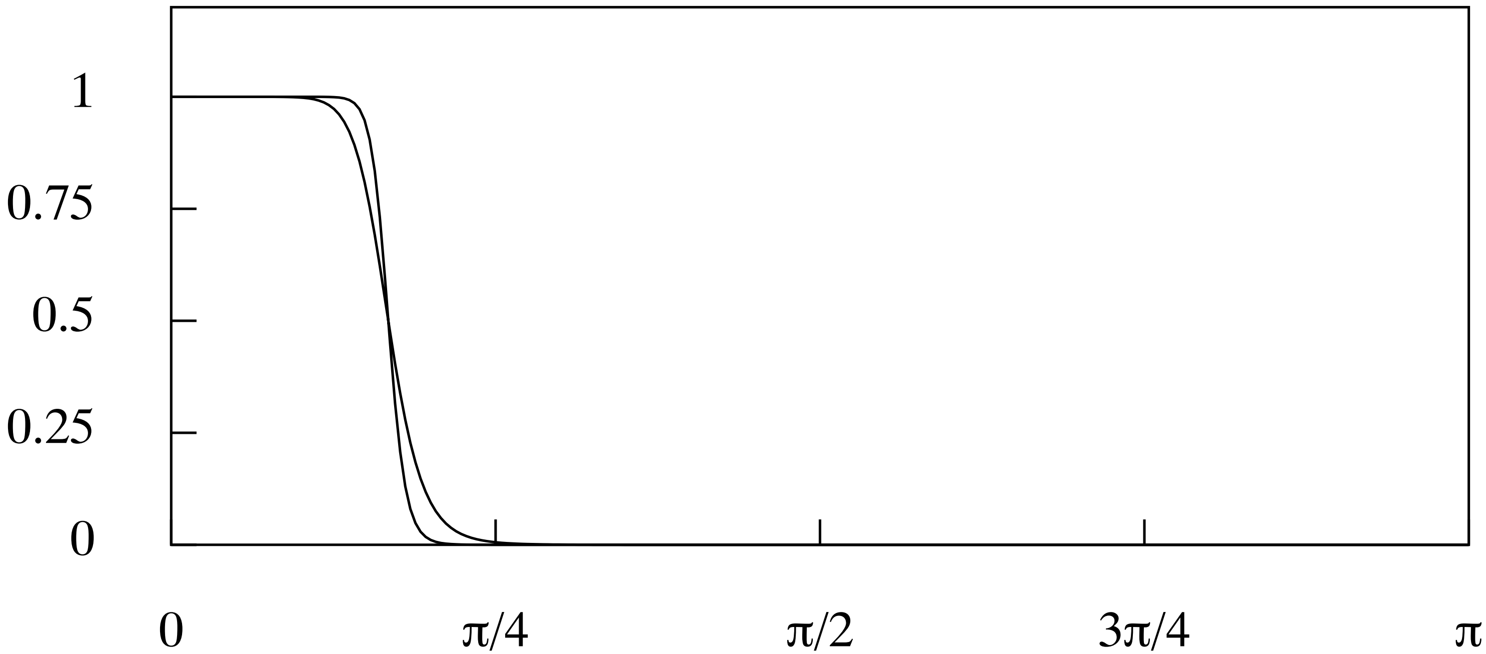


**Figure 12.** The frequency response function of the Butterworth filters of orders $n = 6$ and $n = 12$ with a nominal cut-off point of $\pi/6$ radians ($30°$).

## Ways of Coping with Trended Data

There are several ways of catering to trended data sequences. One way is to remove a linear trend or a polynomial trend of higher degree from the data, and, thereafter, to filter the residual sequence. The low-frequency filtered sequence can be added back to the trend, if it is required to represent the trend-cycle.

An alternative way is to apply a twofold difference operator to the data. The differenced data can be filtered and, thereafter, they can be reinflated by a double summation, which represents the inverse of the differencing operation.

The requirement for an explicit estimation of the initial conditions can be avoided if attention is concentrated on the highpass filter. The initial conditions that are required for the inflation of a differenced sequence that has been subjected to a highpass filter are nothing but zero values.

The complementary lowpass sequence can be obtained by subtracting the inflated product of the highpass filter from the data.

It is remarkable that, given the appropriate conditions, all three methods of dealing with the problems of a trended sequence are algebraically equivalent.

## Finite Sample W–K filters

Finite sample W–K filters overcome the end-of-sample problem by adapting their coefficients as they move through the data.

The so-called Hodrick–Prescott filter, properly attributable to Conrad Leser (1961), comprises a single adjustable smoothing parameter $\lambda$. The equation of the lowpass time-varying filter is

$$\begin{aligned} x &= y - Q(\lambda^{-1}I + Q'Q)^{-1}Q'y \\ &= y - h, \end{aligned}$$

where $h$ represents the highpass component. As $\lambda \to \infty$, the equation converges on that of the linear detrending regression.

The equation of the Butterworth lowpass filter is

$$x = y - \Sigma Q(\lambda^{-1}M + Q'\Sigma Q)^{-1}Q'y.$$

Here, the matrices are

$$\Sigma = \{2I_T - (L_T + L'_T)\}^{n-2} \quad \text{and} \quad M = \{2I_T + (L_T + L'_T)\}^{n},$$

where $L_T$ is the lag-operator matrix of the order of the sample size $T$, which has units on the first subdiagonal and zeros elsewhere. It can be verified that

$$Q'\Sigma Q = \{2I_T - (L_T + L'_T)\}^{n}.$$

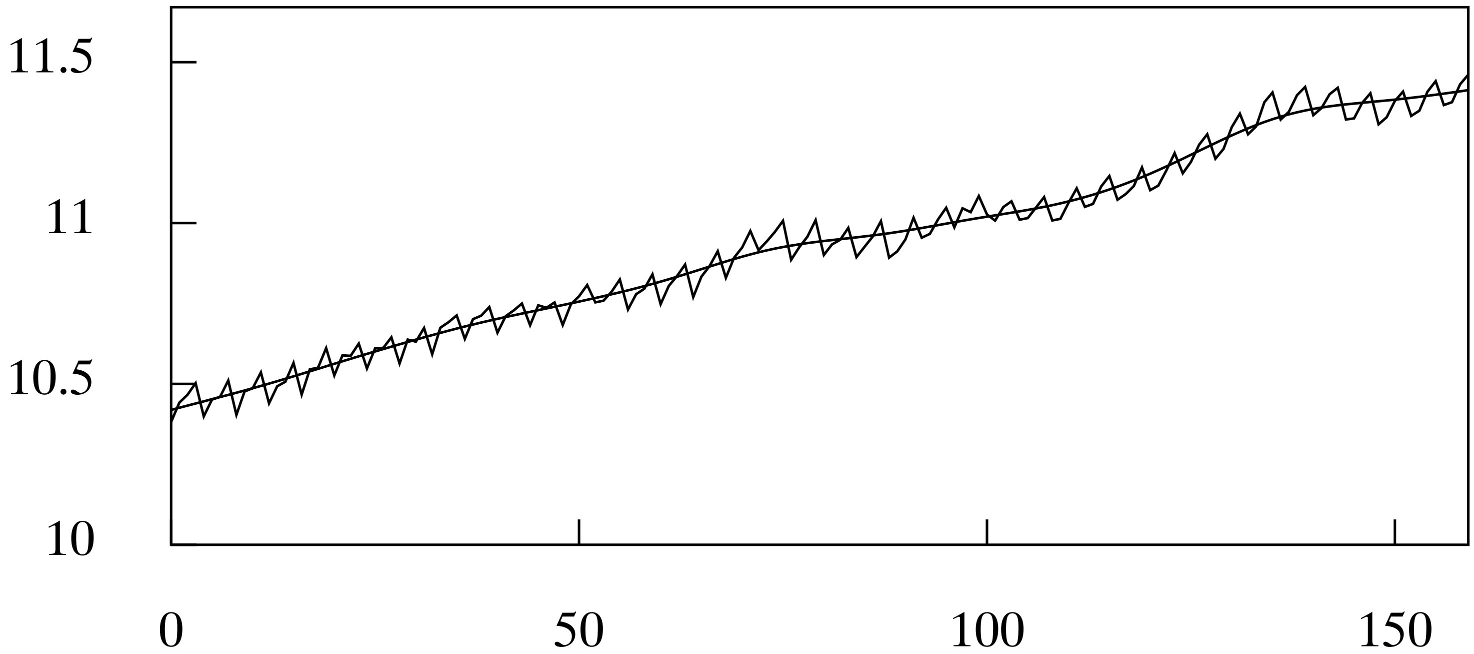


**Figure 13.** The trend-cycle function obtained bu applying a lowpass Hodrick–Prescott filter with a smoothing parameter of $\lambda = 1,600$ to 160 points of the logarithmic consumption data.

## Frequency-Domain Filters

Components of the data with well-defined spectral structures can be synthesised from their spectral ordinates. The Fourier transform of the ideal filter with a rectangular frequency response is the sinc function, which has infinitely many coefficients.

The finite-sample version of the filter, the Dirichelet kernel, is obtained by wrapping the coefficients around a circle of circumference $T$ and by adding the overlying coefficients.

The filtered values would be obtained by the circular convolution of the data with the coefficients of the filter—which is equivalent to applying the sinc function to an indefinite periodic extension of the data sequence by a linear convolution.

The circular convolution can be represented by the matrix equation $x = \Psi^\circ y$, where, for $T = 4$, there is

$$\Psi^\circ = \begin{bmatrix} \psi_0^\circ & \psi_1^\circ & \psi_2^\circ & \psi_1^\circ \\ \psi_1^\circ & \psi_0^\circ & \psi_1^\circ & \psi_2^\circ \\ \psi_2^\circ & \psi_1^\circ & \psi_0^\circ & \psi_1^\circ \\ \psi_1^\circ & \psi_2^\circ & \psi_1^\circ & \psi_0^\circ \end{bmatrix}.$$

The data are treated as a circular sequence, with the effect that the filtered values towards the end of the sequence are formed partly from data values at the beginning of the sequence—and vice versa for the filtered values at the beginning the sequence.

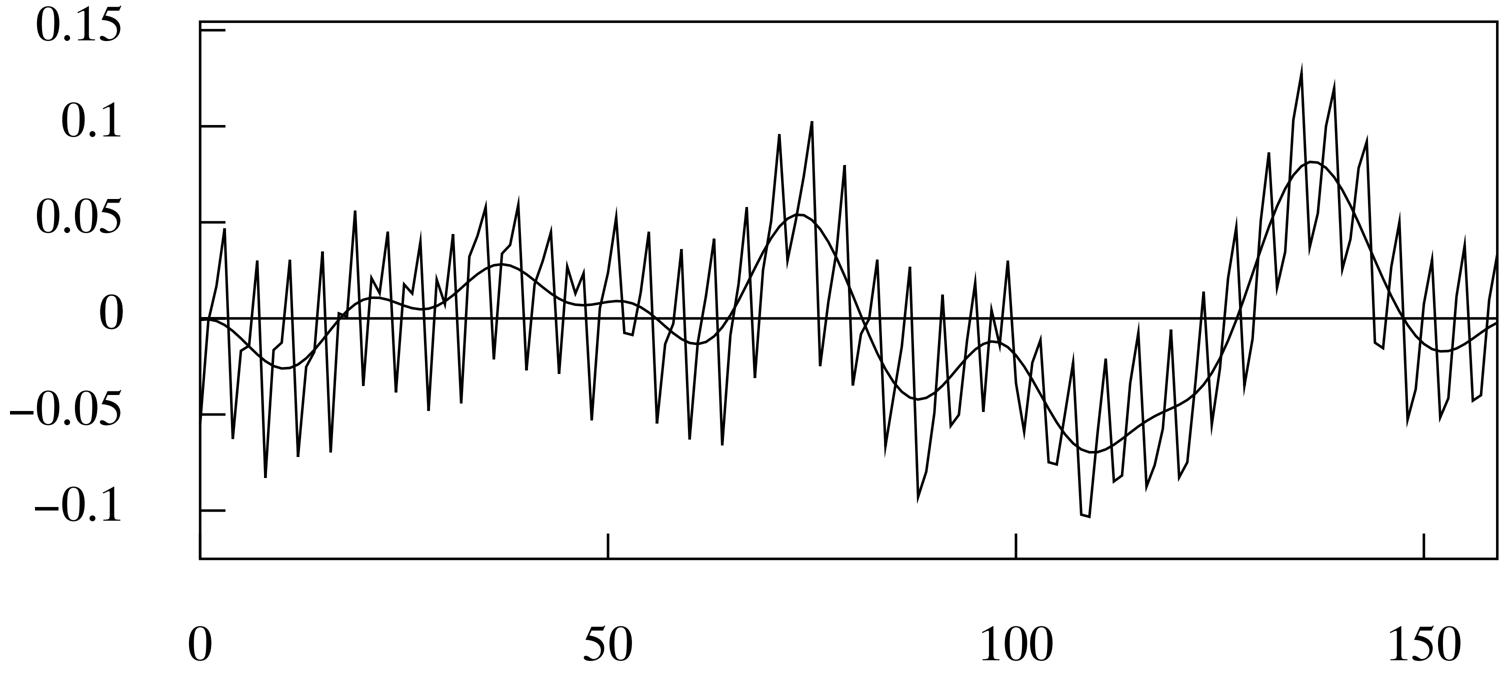


**Figure 14.** The residual sequence from fitting a linear trend to the logarithmic consumption data with an interpolated line representing the business cycle, obtained by the frequency-domain method.

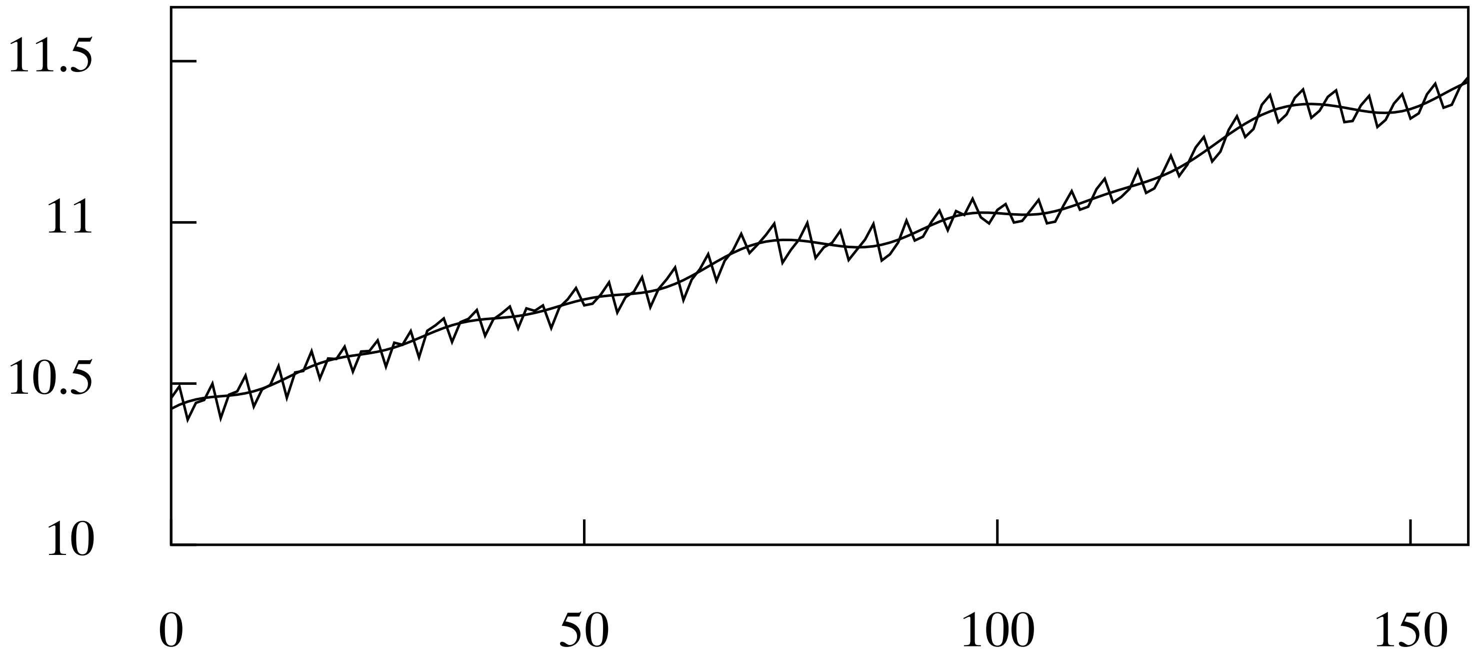


**Figure 15.** The trend–cycle component of U.K. consumption determined by the frequency-domain method, superimposed on the logarithmic data.

## Turning Points

As a sum of trigonometrical functions, the business-cycle trajectory is an analytic function of which the derivatives exist of all orders.

This implies that it is straightforward to find the maxima and minima of the function, and hence the turning points of the business cycle, by identifying the points where the first derivative is zero-valued.

The simplicity of this procedure contrasts markedly with the complexity of some other well-known procedures for locating the turning points of the business cycle, such as that of Bry and Boschan (1971).

A distinction can be made between the turning points of the business cycle function and the turning points of the trend–cycle function. Whenever there is a trend, the latter are less numerous and, whenever there is an upward trend, the downturns occur later in time and the upturns occur sooner.

In Figure 16, the turning points of the business cycle are marked on the horizontal line at zero. The turning points of the trend-cycle function are marked on the horizontal line at -0.05, which represents a baseline of 5 percent growth.

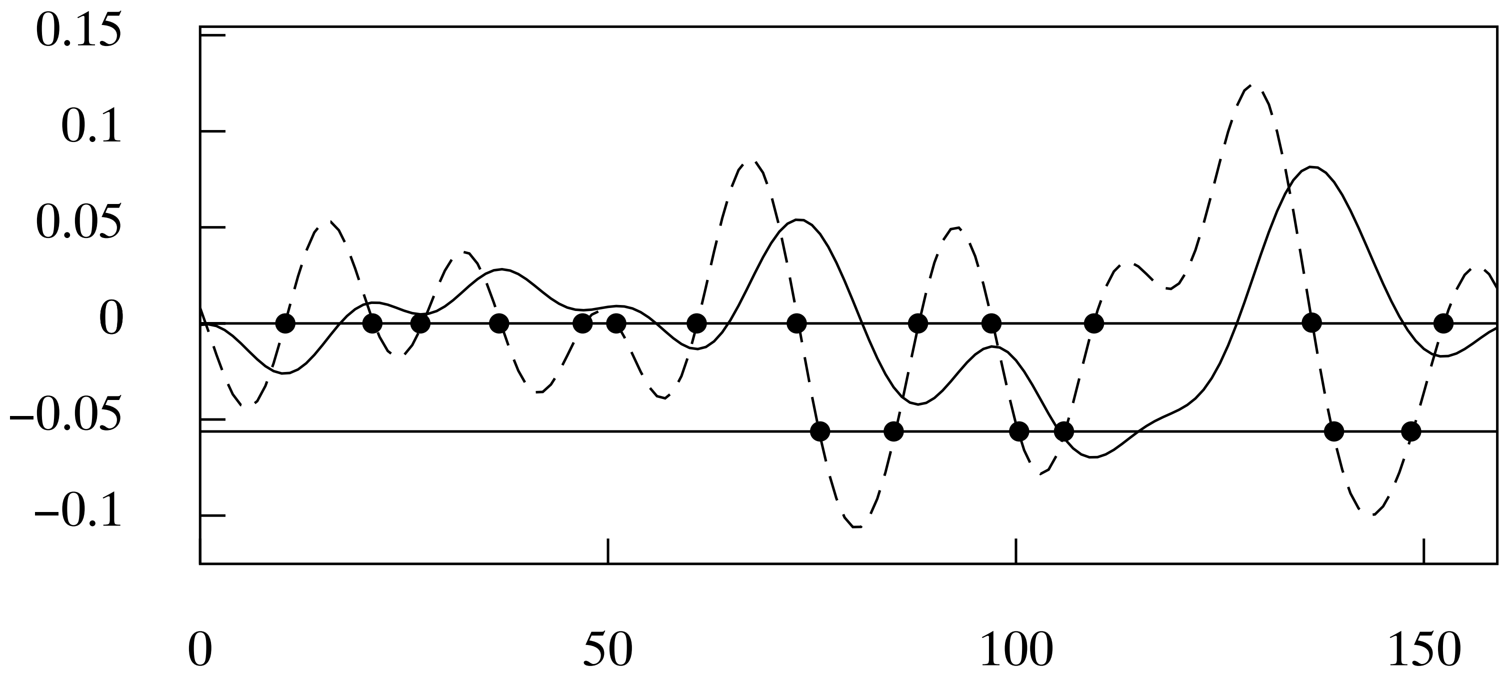


**Figure 16.** The turning points of the business cycle marked on the horizontal axis by black dots. The solid line is the business cycle of Figure 14. The broken line is the derivative function.

## Filtering Monthly Seasonal Data

With monthly data, a filter with a rapid transition from pass band to stop band may be required to isolate the rest of the data from the low-frequency trend-cycle component. The transition is liable to occur before $\pi/6$ radians ($30^\circ$ degrees), which is the fundamental seasonal frequency. A frequency domain filter will serve the purpose.

The data will be subject to polynomial detrending followed by a lowpass filtering to remove the residual low-frequency component that is part of the trend-cycle function.

After polynomial detrending, there may be a disjunction in the circular data sequence, where the end joins the beginning, or, equally, in the periodic extension of the data, where one replication of the data sequence is succeeded by the next.

To eliminate this disjunction, a run of synthetic data may be inserted between the end and the beginning, which makes a gradual transition from one to the other.

When the data show a strong seasonal variation, which has evolved over the course of the sample, a segment of pseudo data may be constructed by morphing the pattern of seasonal variation so that it changes from the pattern at the end of the data sequence to the pattern at the beginning.

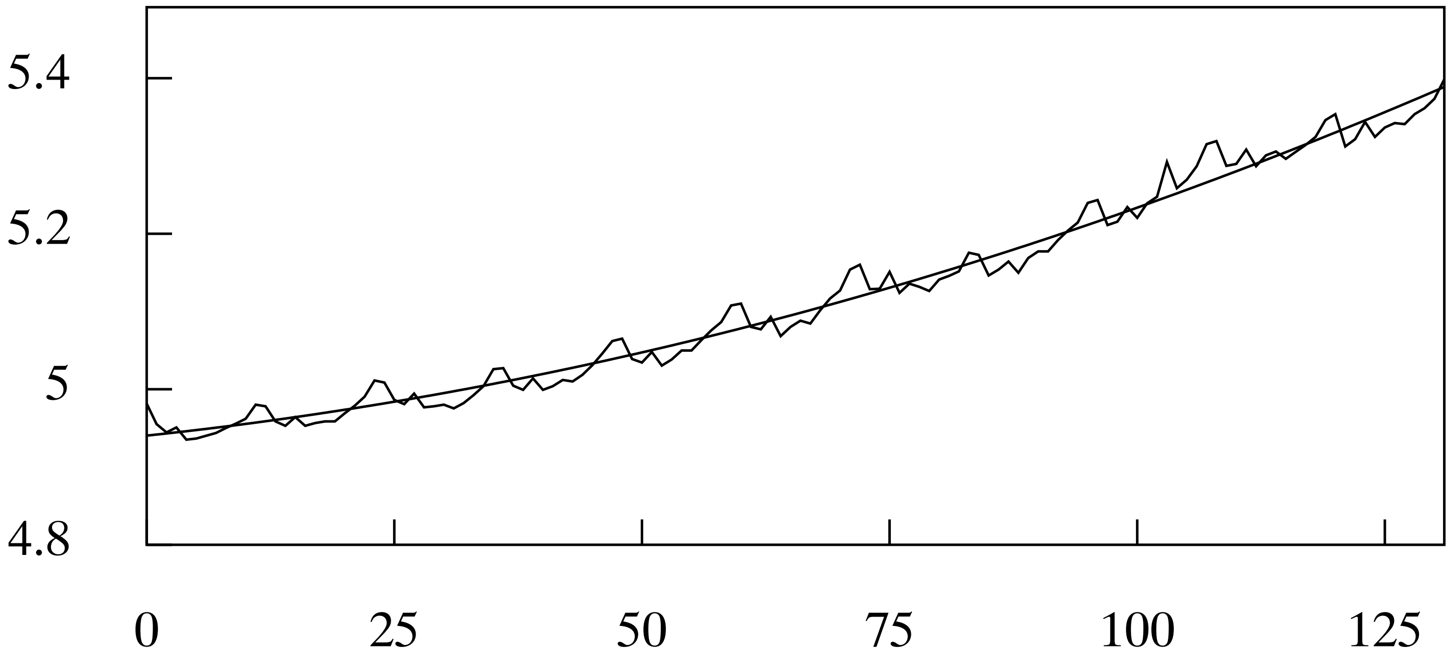


**Figure 17.** The plot of 132 monthly observations on the logarithms of the U.S. money supply, beginning in January 1960. A quadratic function has been interpolated through the data.

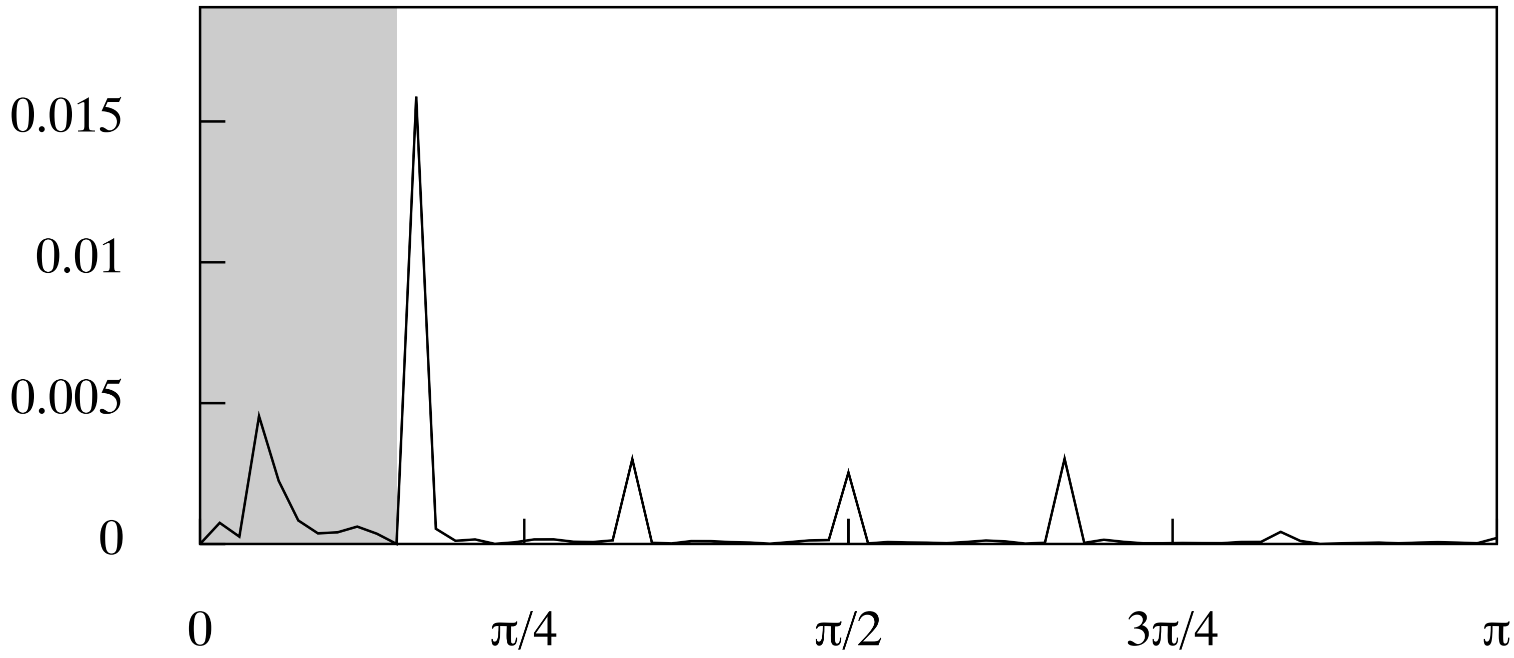


**Figure 18.** The periodogram of the residuals from the quadratic detrending of the logarithmic money-supply data.

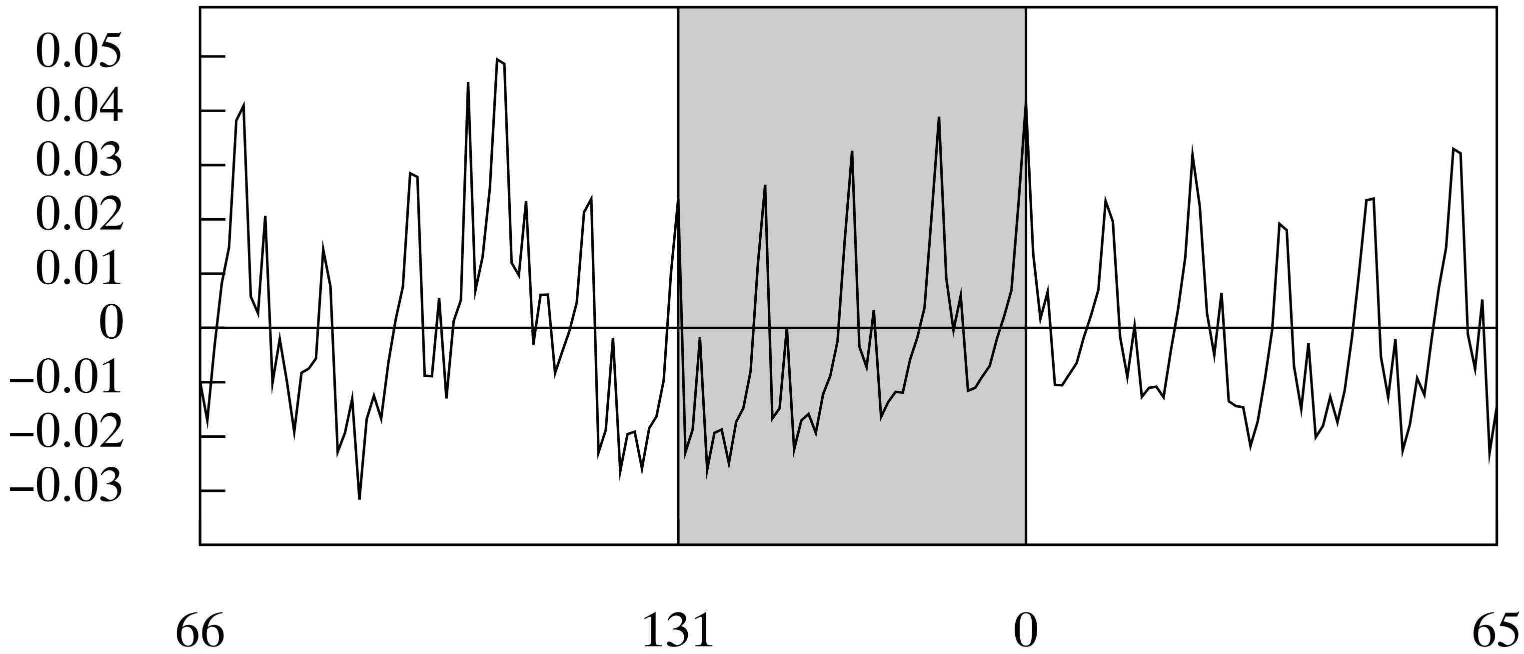


**Figure 19.** The residuals from a linear detrending of the sales data, with an interpolation of four years length inserted between the end and the beginning of the circularised sequence, marked by the shaded band.

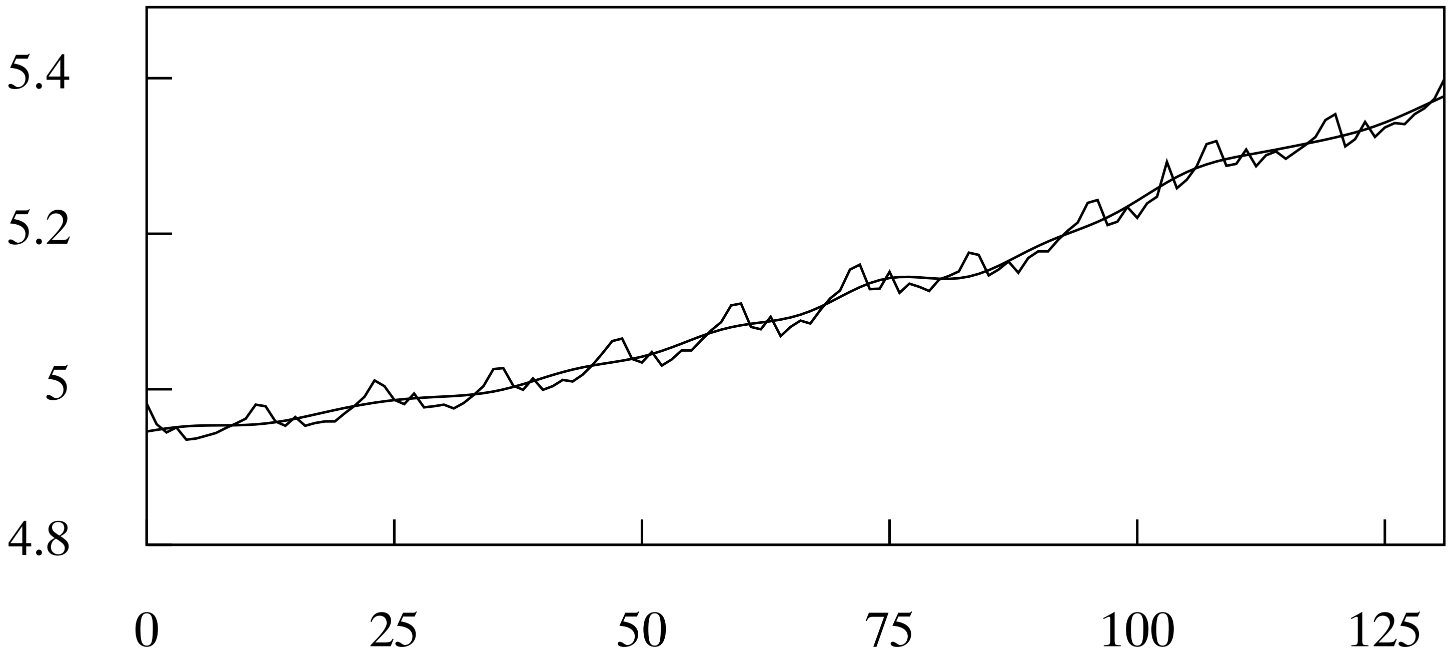


**Figure 20.** The plot of the logarithms of 132 monthly observations on the U.S. money supply, beginning in January 1960. A trend-cycle, estimated by the Fourier method, has been interpolated through the data.

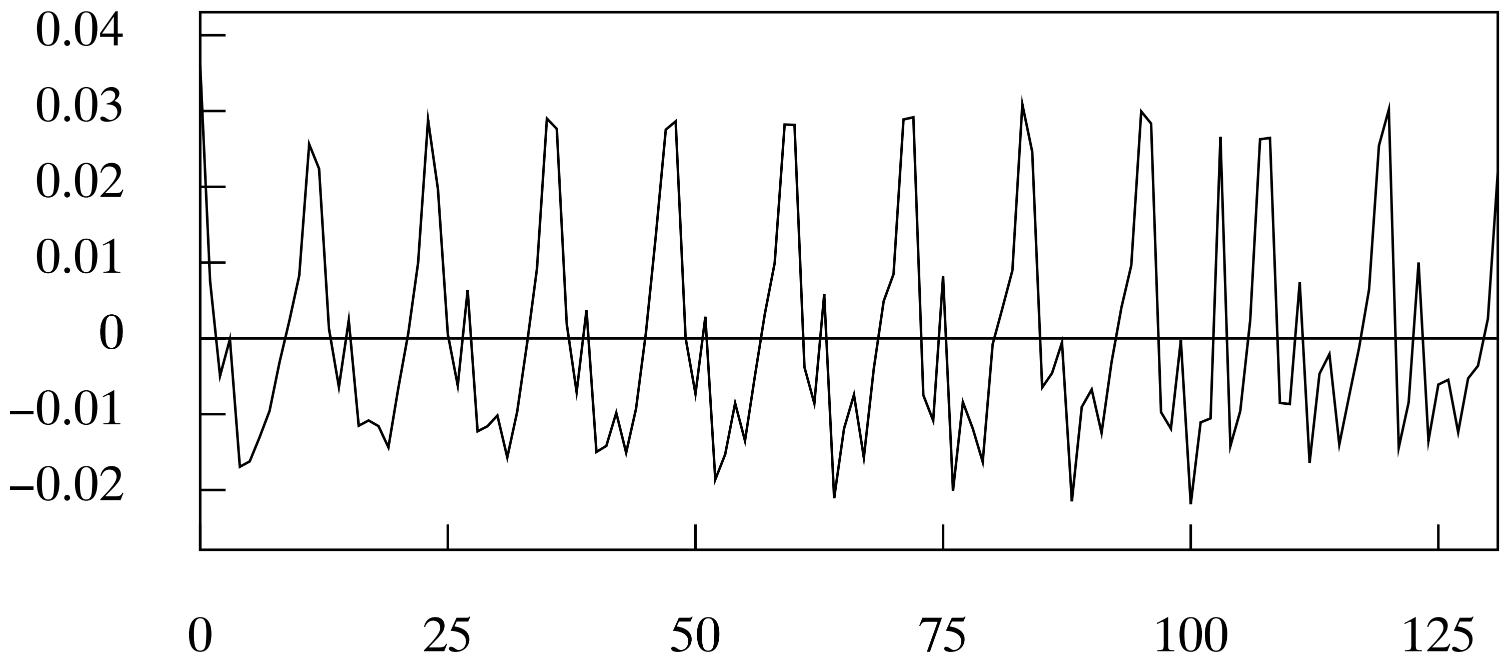


**Figure 21.** The sequence of residual deviations of the logarithmic money supply data from the estimated trend-cycle function.

## Interrupted Trends

Sometimes, there are major interruptions that halt the steady progress of the economy and which can give rise to wide deviations from an interpolated polynomial trend.

If such interruptions are deemed to have an enduring effect on the underlying trajectory of the economy, then it may be appropriate to describe them as structural breaks and to absorb them into the trend.

A device that will serve this purpose is a form of the Hodrick–Prescott filter in which the smoothing parameter can take different values in different localities.

In the vicinity of the break, the smoothing parameter can be set to a sufficiently low value to allow the function to absorb the break. Elsewhere, it should be set to a high value to make it sufficiently stiff to prevent it from absorbing the cyclical fluctuations of the data.

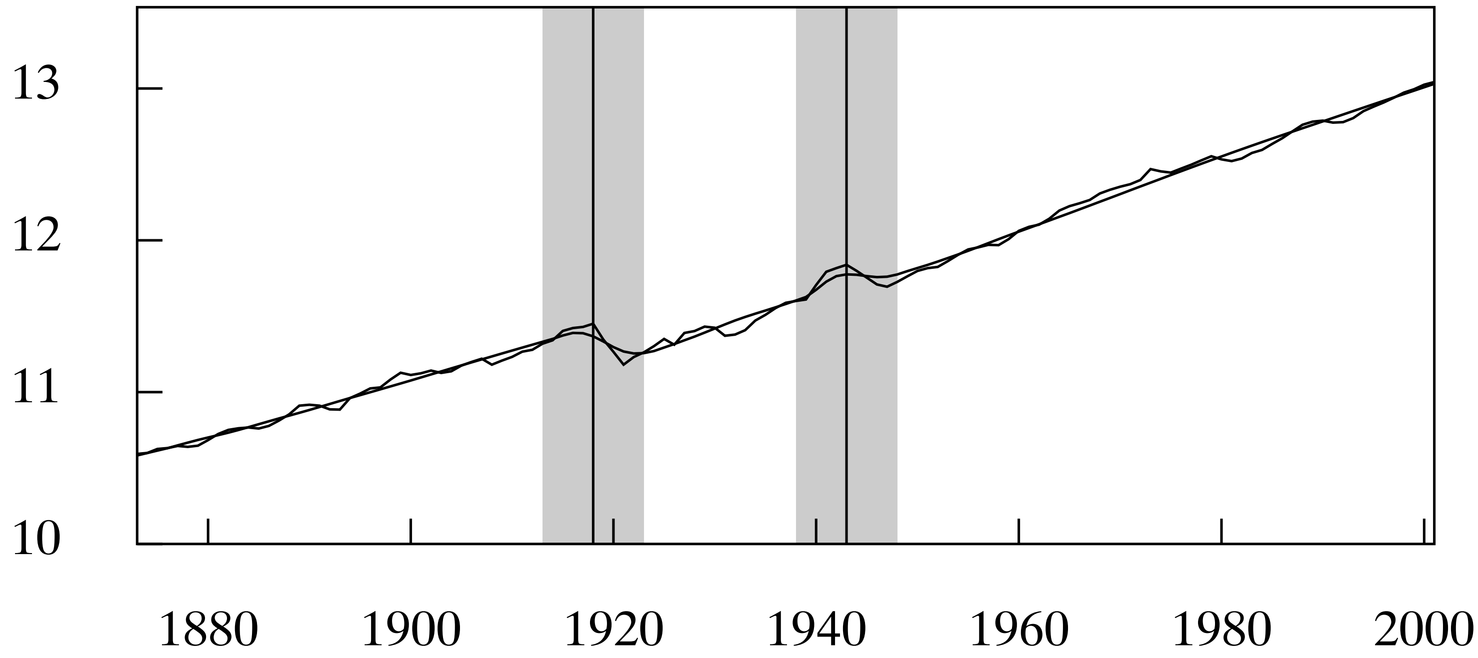


**Figure 22.** The logarithms of annual U.K. real GDP from 1873 to 2001 with an interpolated trend. The trend is estimated via a filter with a variable smoothing parameter.