TRENDS CYCLES AND SEASONS: ECONOMETRIC METHODS OF SIGNAL EXTRACTION

D.S.G. Pollock

University of Leiccester

stephen_pollock@sigmapi.u-net.com

A variety of models and methods are available for decomposing an econometric data sequence into a trend (or trend-cycle) component, a seasonal component and an irregular component. If logarithms of the data are taken, then this will be an additative decomposition.

This paper presents a general time-domain procedure based on finite-sample Wiener-Kolmogorov (W-K) filtering, which can implement the available model-based methods for achieving the triple decomposition. It also presents a method that operates in the frequency domain that has certain advantages over the time-domain methods.

The matrix algebra of the W-K method is unavoidably dense. Therefore, only a summary outline of the procedue will be presented.

Classical Wiener–Kolmogorov Filters

The archetypal Wiener–Kolmogorove filter is a linear time-invariant device that is aimed at decomposing a stationary data sequence $y(t) = \xi(t) + \eta(t)$ into its signal component $\xi(t)$ and its noise component $\eta(t)$.

The filter presupposes that the data sequence and its components are stationary and of indefinite duration. If

$$y(t) = \xi(t) + \eta(t) = \phi(L)\varepsilon(t) + \theta(L)\nu(t),$$

then the signal extraction filter is

$$\psi(z) = \frac{\gamma^{\xi\xi}(z)}{\gamma^{\xi\xi}(z) + \gamma^{\eta\eta}(z)} = \frac{\sigma_\varepsilon^2\phi(z^{-1})\phi(z)}{\sigma_\varepsilon^2\phi(z^{-1})\phi(z) + \sigma_\eta^2\theta(z^{-1})\theta(z)}.$$

There is no recognition of the start-up or initial-conditions problem.

The finite-sample version of the filter is a time-varying device in which the autocovariance generating functions are replaced by the corresponding dispersion matrices:

$$\Psi = (\Omega_\xi^{-1} + \Omega_\eta^{-1})^{-1}\Omega_\eta^{-1} = \Omega_\xi(\Omega_\xi + \Omega_\eta)^{-1}.$$

Filters for Non-stationary Sequences

The time-invariant filters can be applied to trended data, provided that the start-up problem is handled appropriately. A trended data sequence may be extrapolated at both ends to provide support for the filter.

It is appropriate to extract the stationary noise component $\eta(t)$ from the data and then to estimate the trended signal component $\xi(t)$ by subtraction:

$$\xi(t) = y(t) - \eta(t).$$

Let the matrix difference operator Q' be effective in reducing the finite data vector $y = \xi + \eta$ to stationarity to give

$$g = Q'y = Q'\xi + Q'\eta \quad \text{with} \quad D(Q'\xi) = \Omega_\varepsilon \quad \text{and} \quad D(Q'\eta) = Q'\Omega_\eta Q.$$

Then, the matrix transformation that extracts the signal from the data vector y is

$$\Psi y = (Q\Omega_\varepsilon^{-1}Q' + \Omega_\eta^{-1})\Omega_\eta^{-1}y = \{I - \Omega_\eta Q(\Omega_\varepsilon + Q'\Omega_\eta Q)^{-1}Q'\}y.$$

There is no explicit treatment of the start-up problem, since the appropriate initial values for the noise sequence are their zero-valued expectations.

A Model with Two Non-Stationary Components

We may also consider the case of a data sequence composed of two nonstationary components:

$$y(t) = \xi(t) + \kappa(t)$$

In the case of a finite data sequence of which the vector is $y = \xi + \kappa$, there are the following reductions to stationarity:

$$Q'_\xi \xi = \varepsilon \quad \text{with} \quad D(\varepsilon) = \Omega_\varepsilon \quad \text{and} \quad Q'_\kappa \kappa = \zeta \quad \text{with} \quad D(\zeta) = \Omega_\zeta.$$

Then, the estimates of the two components are

$$x = (Q_\xi \Omega_\varepsilon^{-1} Q'_\xi + Q_\kappa \Omega_\zeta^{-1} Q'_\kappa)^{-1} Q'_\xi y,$$

$$k = (Q_\kappa \Omega_\zeta^{-1} Q'_\xi + Q_\xi \Omega_\zeta^{-1} Q'_\xi)^{-1} Q'_\kappa y.$$

Since neither $Q_\xi \Omega_\varepsilon^{-1} Q'_\xi$ nor $Q_\kappa \Omega_\zeta^{-1} Q'_\kappa$ are invertible, the matrix inversion lemma is not applicable. We are faced with inverting a full matrix of order T, which may tax the computer's memory. Therefore, we must look for a better way.

The Model with a Trend and with Seasonal Fluctuations

$$
\begin{aligned}
y(t) &= \xi(t) + \eta(t) \\
&= \{\rho(t) + \kappa(t)\} + \eta(t),
\end{aligned}
$$

with

$\rho(t)$ non-stationary trend,

$\kappa(t)$ non-stationary seasonal component,

$\xi(t) = \rho(t) + \kappa(t)$ combined non-stationary component,

$\eta(t)$ stationary irregular component.

Reductions to Stationarity

$$
\begin{aligned}
\varepsilon(t) &= \nabla^2(L)\rho(t), \quad \text{reducing the trend} \\
\zeta(t) &= \Sigma(L)\kappa(t), \quad \text{reducing the seasonal component} \\
g(t) &= \Delta(L)y(t) = \nabla^2(L)\Sigma(L)y(t). \quad \text{reducing the data}
\end{aligned}
$$

Estimates

From the stationary sequence $g(t)$, we can find an estimates of the stationary elements of the model via Wiener–Kolmogorov filtering:

$\eta(t)$	the stationary irregular component
$\varepsilon(t)$	the forcing function of the trend
$\zeta(t)$	the forcing function of the seasonal fluctuations

Then, the combined non-stationary component can be estimated via the equation

$$\xi(t) = y(t) - \eta(t).$$

Estimates of the trend component and the seasonal component can be found via the equations

$$\rho(t) = \nabla^{-2}(L)\varepsilon(t) \quad \text{and} \quad \kappa(t) = \Sigma^{-1}(L)\zeta(t).$$

These equations entail integration (the inverse of differencing). The necessary initial conditions are available via the relationship

$$\xi(t) = \nabla^{-2}(L)\varepsilon(t) + \Sigma^{-1}(L)\zeta(t)$$

The Airline Pasenger Model and TRAMO-SEATS

The airline passenger model of Box and Jenkins is the basis of the TRAMO-SEATS procedures. The model is represented by the equation

$$y(z) = \frac{N(z)}{\Delta(z)}\varepsilon(z) = \left\{ \frac{(1-\phi z)(1-\Theta z^s)}{(1-z)(1-\ z^s)} \right\} \varepsilon(z),$$

where $\Delta(z) = (1-z)(1-z^s) = \nabla^2(z)\Sigma(z)$. The model is subject to a canonical decomposition, which defines the three components and which maximises the variance attributed to the irregular noise component.

The filters can be derived from a *meta model* of the form

$$\begin{aligned} y(z) &= \frac{U(z)}{\nabla^2(z)}\varepsilon(z) + \frac{V(z)}{\Sigma(z)}\zeta(z) + \eta(z) \\ &= \rho(z) + \kappa(z) + \eta(z). \end{aligned}$$

Whereas no explicit expressions are available for $U(z)$ and $V(z)$, the expressions for $\Omega_\varepsilon(z) = U(z^{-1})U(z)$ and $\Omega_\zeta(z) = V(z^{-1})V(z)$ have been provided by Hillmer and Tao (1982). Thus, the necessary elements for the finite-sample implementation of the filters are readily available.

The Frequency Response Functions of the Canonical Filters

Spectral theory presupposes stationary stochastic processes of infinite duration. Therefore, in analysing the effects of the canonical decomposition, we have to consider the linear time-invariant versions of the filters.

The frequency responses of the filters indicate the extent to which the trigonometric elements, into which such data sequences are decomposed, are liable to have their amplitudes amplified or attenuated.

Since the time-invariant versions of the filters have symmetric sequences of coefficients, there are no associated phase effects; and the filters are characterised by their gain or their squared gain functions, which are defined over the the Nyquist interval of $[0, \pi)$.

The value of π radians per sample interval represents the maximum frequency, or angular velocity, that is observable in sampled data.

According to the principal of canonical decomposition, neither the trend component nor the seasonal component should contain any extraneous white noise. The principal is reflected in the gain function of the trend-extraction filter, which attains a value of zero at the limiting Nyquist frequency.

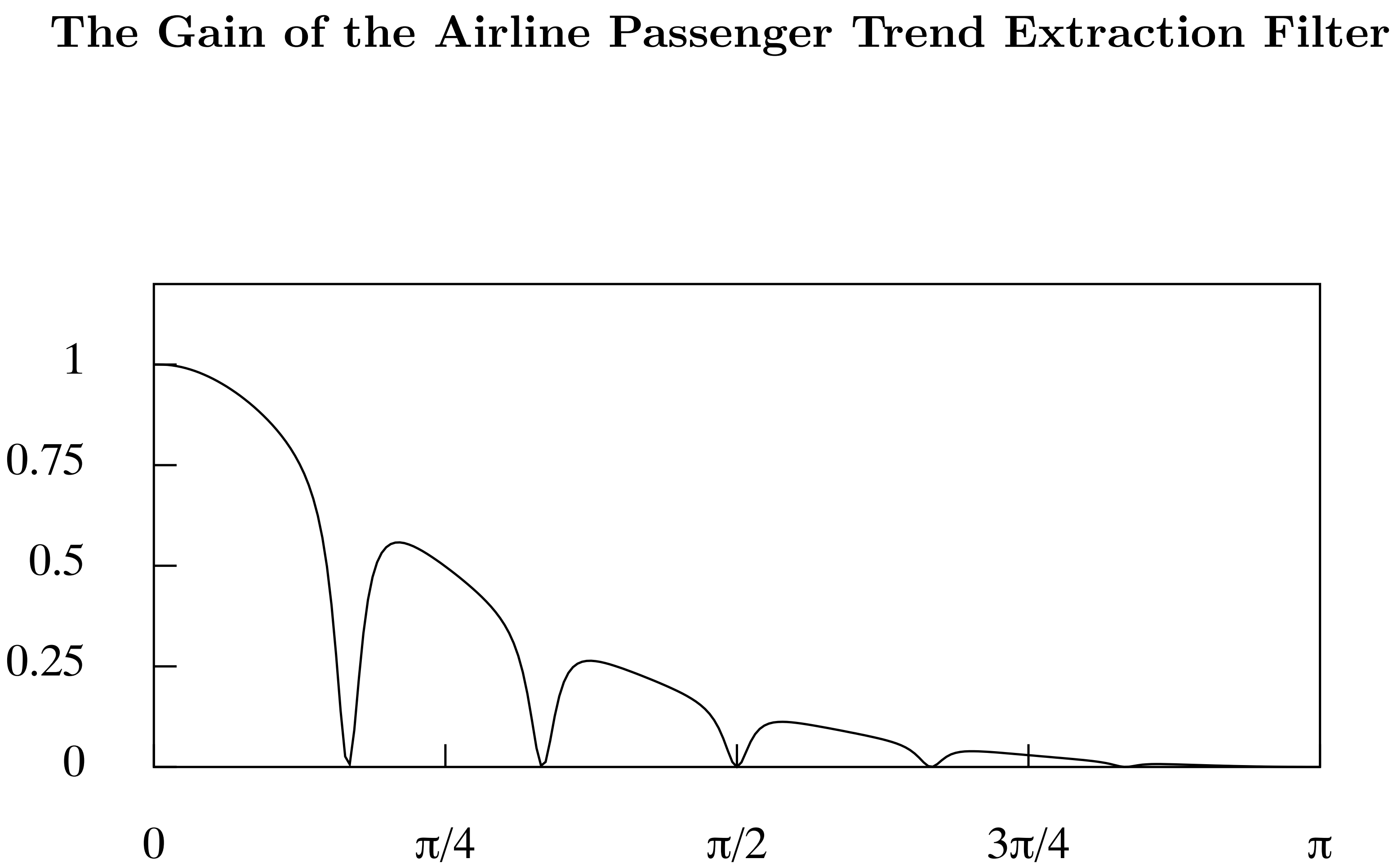

Figure 1. The gain of the trend extraction filter associated with the monthly airline passenger model.

The Gain of the Airline Passenger Seasonal-adjustment Filter

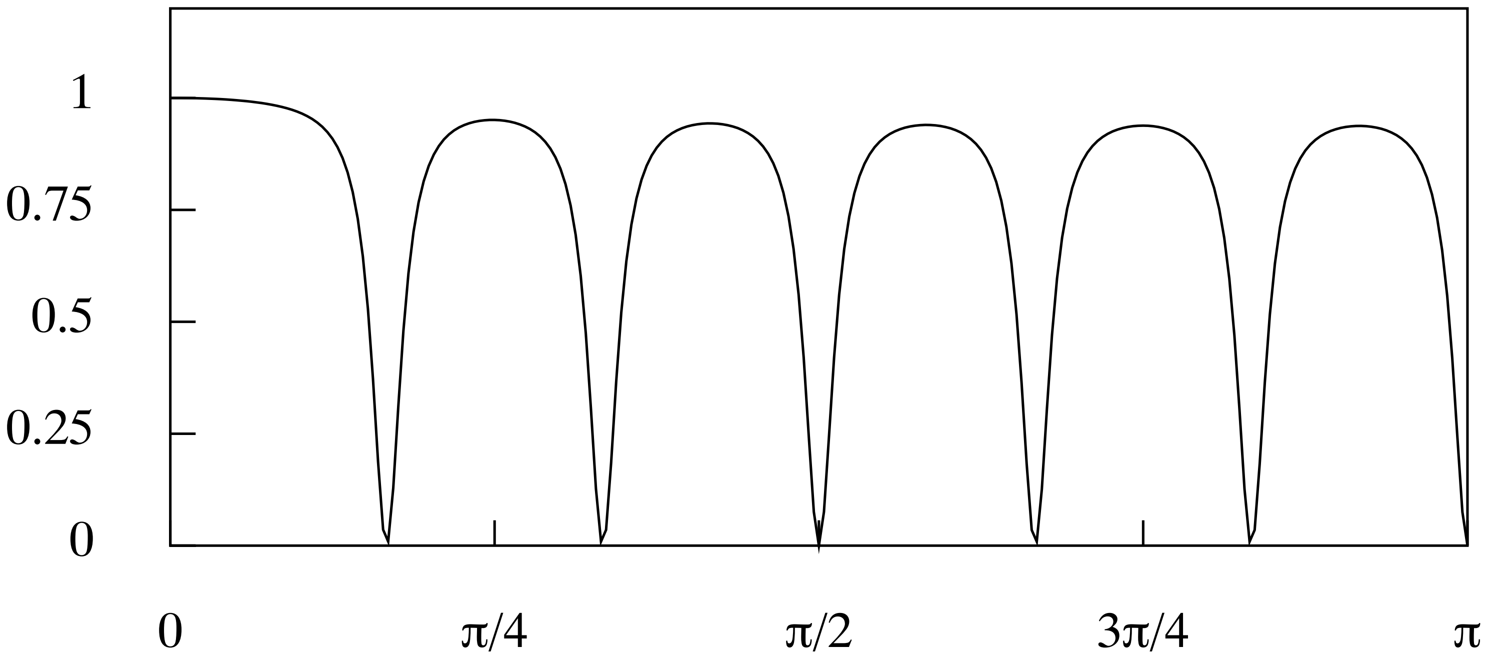

Figure 2. The gain of the seasonal-adjustment filter associated with the monthly airline passenger model.

The Structural Times Series Model of STAMP

The equation of a structural time series model, which is the basis of the STAMP program, is

$$\begin{aligned} y(z) &= \left\{ \frac{z\nu(z)}{\nabla^2(z)} + \frac{\xi(z)}{\nabla(z)} \right\} + \frac{\zeta(z)}{\Sigma(z)} + \eta(z) \\ &= \rho(z) + \kappa(z) + \eta(z), \end{aligned}$$

wherein $\zeta(z), \xi(z), \nu(z)$ and $\eta(z)$ are the z-transforms of white-noise sequences. The estimation of the model requires only the determination of the variances of the four white-noise processes.

The first two terms on the RHS of can be combined to give

$$\rho(z) = \frac{z\nu(z)}{\nabla^2(z)} + \frac{\xi(z)}{\nabla(z)} = \frac{1 - \mu z}{\nabla^2(z)} \varepsilon(z),$$

which stands for a doubly integrated first-order moving-average process wherein $\varepsilon(t)$ is also a white-noise process.

The procedures of STAMP do not adhere to the principal of canonical decomposition, which may be imposed nevertheless.

The Gain of the STAMP Trend Extraction Filter

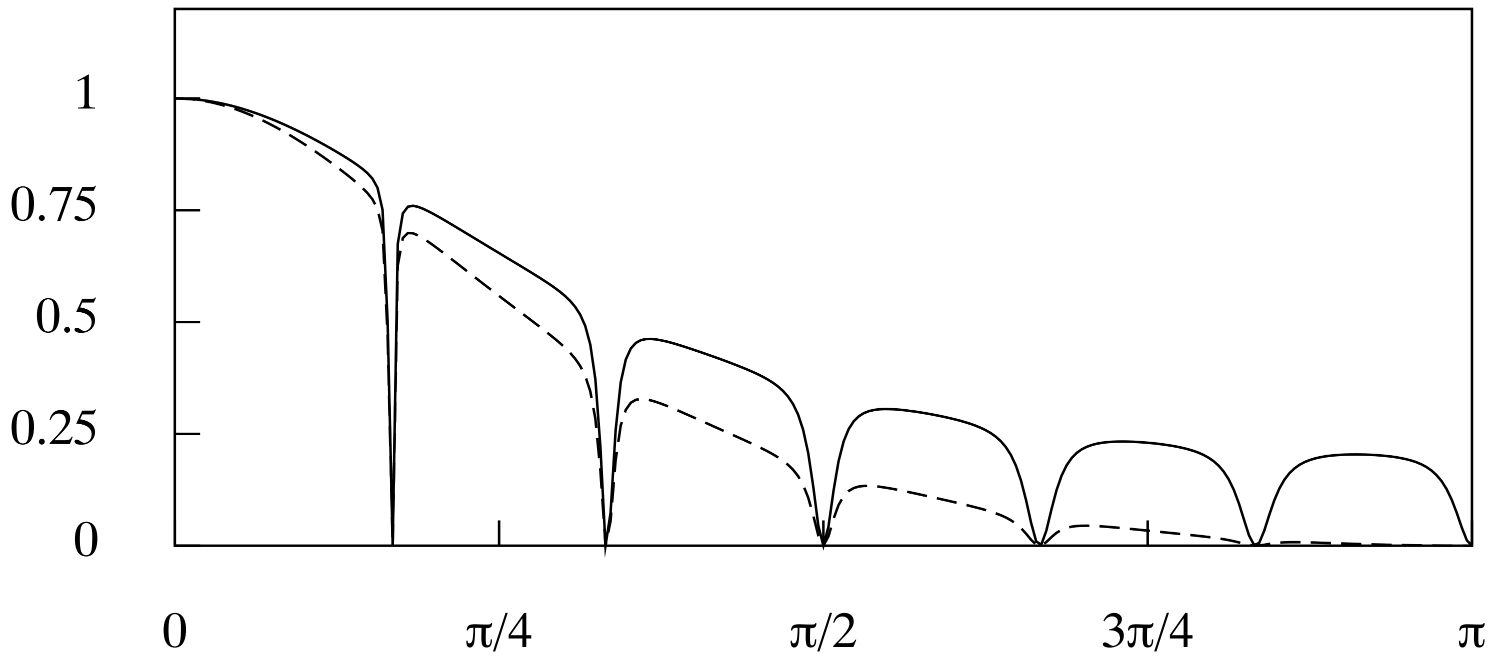

Figure 3. The gain of the trend extraction filter associated with the structural time series model (the solid line) together with that of the canonical version of the filter (the broken line).

The Model-based Procedure of IDEOLOG

The model-based procedure of IDEOLOG is for investigting and illustrating the possibilities for seasonal adjustment via a Wiener–Kolmogorov filter.

The parameters of the filter are not estimated. Instead, they are specified by the user. The equation of the model is

$$\begin{aligned} y(z) &= \rho(z) + \frac{\Sigma(\theta z)}{\Sigma(z)}\zeta(z) + \eta(z) \\ &= \rho(z) + \kappa(z) + \eta(z), \end{aligned}$$

Here, $\rho(z)$ is a trend fuction—a polynomial or a Hodrick–Prescott trend, for example.

The model of the seasonal component embodies the numerator polynomial

$$\Sigma(\theta z) = 1 + \theta z + \theta^2 z^2 + \cdots + \theta^{s-1} z^{s-1} = \frac{1 - \theta^s z^s}{1 - \theta z},$$

where $\theta < 1$ is a positive number close to unity. The second adjustable parameter is the variance ratio $\lambda = \sigma_\eta^2/\sigma_\zeta^2$.

The Gain of Seasonal-adjustment Filter of IDEOLOG

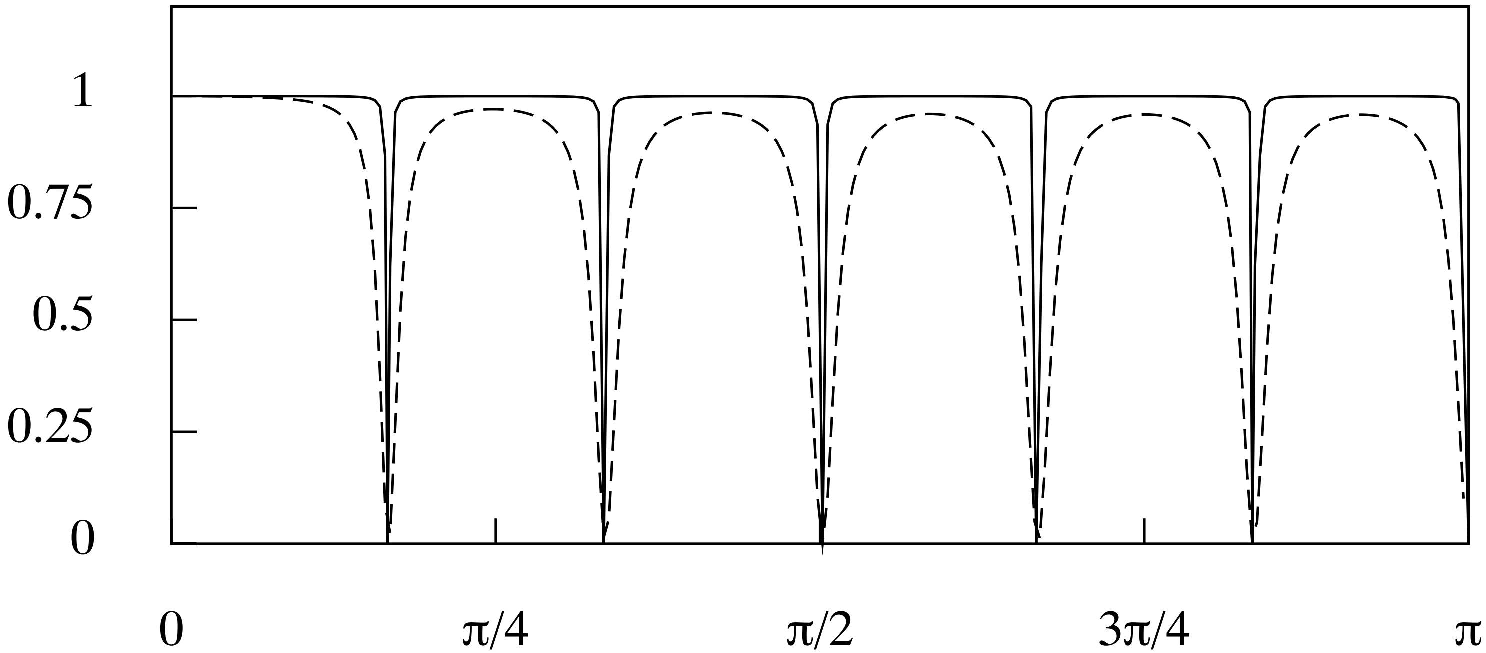

Figure 4. The gain of the seasonal adjustment filter associated with the monthly version of the heuristic model. The solid line corresponds to the parameters $\theta = 0.99$ and $\lambda = \sigma_\eta^2/\sigma_\zeta^2 = 0.125$ and the broken line corresponds to the parameters $\theta = 0.6$ and $\lambda = 0.125$.

A Data Sequence Sesasonally Adjusted by IDEOLOG

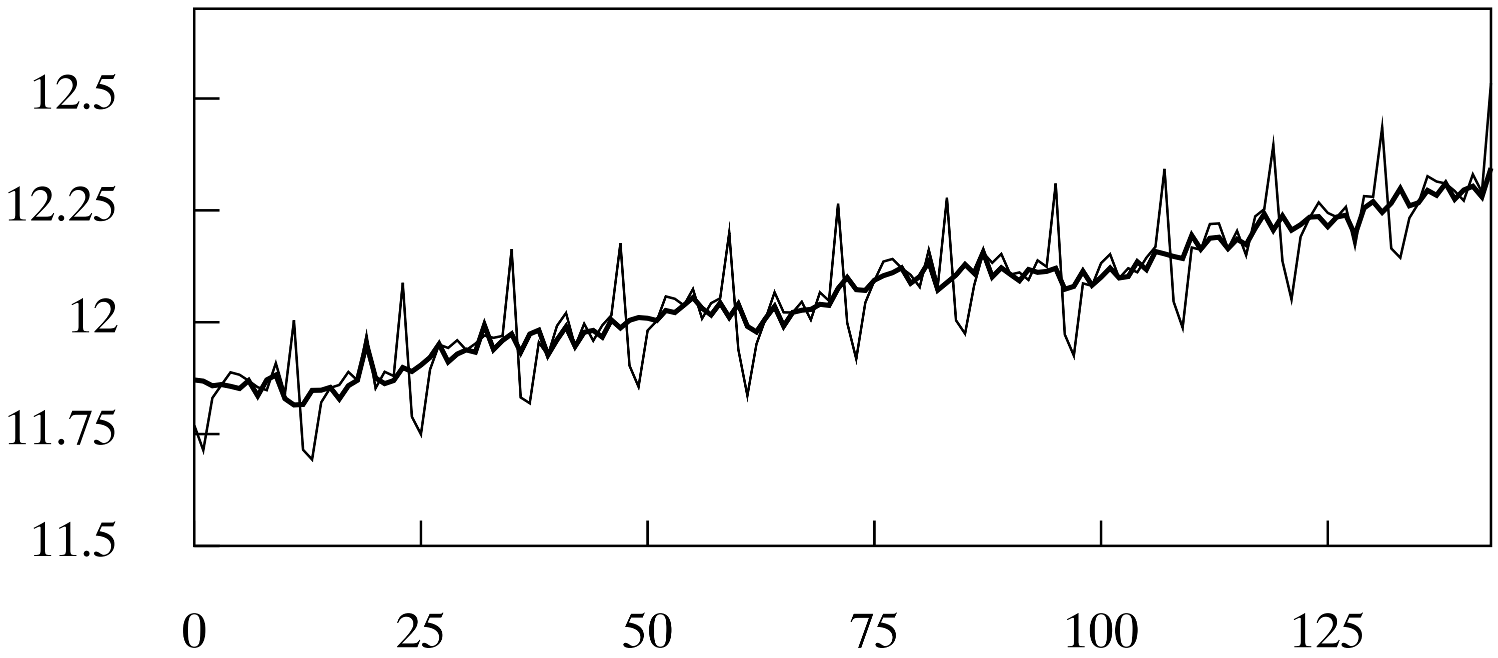

Figure 5. The effect of applying seasonal adjustment filter of the heuristic model to the logarithms of the monthly index of U.S. total sales from January 1953 to December 1964. The parameters of the model are $\theta = 0.99$ and $\lambda = \sigma_\eta^2/\sigma_\zeta^2 = 0.125$.

Filtering in the Frequency Domain

A so-called ideal filter aims to isolate a well-defined subset of the trigonometric or spectral elements into which a data sequence can be resolved. Such a filter has a rectangular frequency response with precisely defined cut-off points.

The Fourier transform of a frequency-domain rectangle on the inteval $[a, b] \in [0, \pi)$ is a sinc function, which gives rise to a doubly-infinite sequence of filter coefficients:

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

However, if the frequency response is sampled at T points equally spaced in $[0, \pi)$, which are thereafter tranformed to the time domain, then the result is a wrapped sinc function, or a Dirichlet kernel, which can be applied to the T data points by a circular convolution.

In practice, it is more efficient to operate in the frequency domain by selecting the Fourier coefficients of the data that are to be preserved and by nullifying the remainder. Then, the inverse Fourier transform into the time domain will generate the filtered data.

The Periodic Dirichlet Kernel

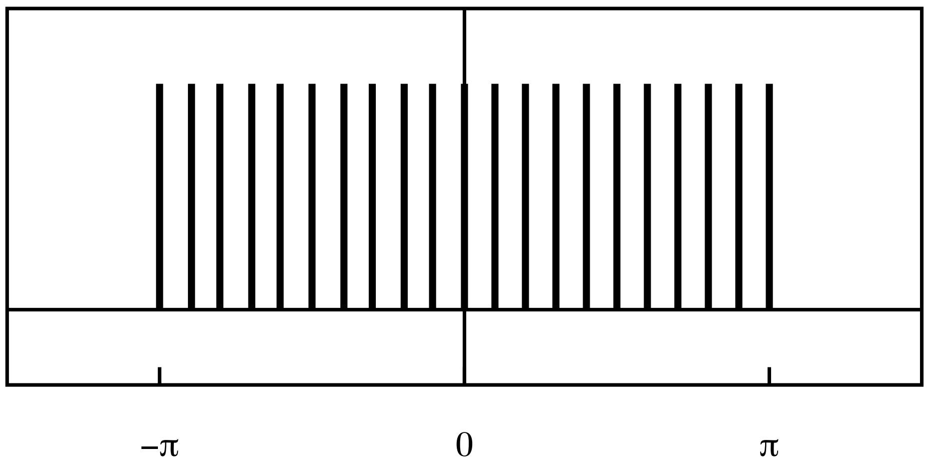

Figure 6. The frequency-domain rectangle sampled at $M = 21$ points.

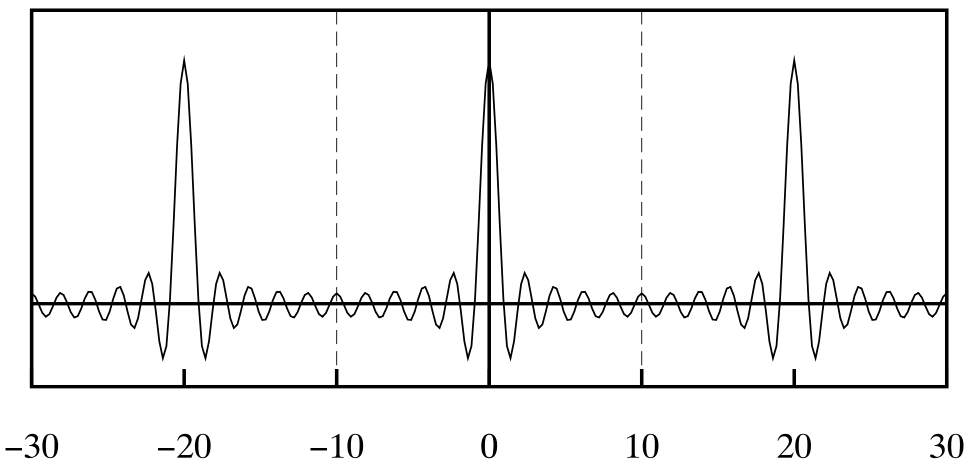

Figure 7. The Dirichlet function $\sin(\pi t)/\sin(\pi t/M)$ obtained from inverse Fourier transform of a frequency-domain rectangle sampled at $M = 21$ points

The Requirement to Detrend a Circular Data Sequence

If the data are trended, then their periodic extension generates a saw-tooth profile that has a *on-over-f* spectrum. To show the finer details of the spectral structure, the data must be detrended. A polynomial function can often serve as a preliminary representation of the trend.

Sometimes, there remains a significant disjunction in the periodic extension of the detrended data, where the end of one replication is joined to the beginning of the next replication.

One way of of eliminating such a break is to interpolate a sequence of pseudo data into the circular sequence to achieve an orderly transition between the end and the beginning. At an appropriate stage, the pseudo data can be discarded.

The pseudo data can be created from a convex combination of the final segment of the data and the initial segment. The weighting of the final segment can vary from unity to zero over successive interations of the two segments. The trajectory of the weights can be governed by a half cycle of a raised cosine function:

$$\frac{\cos(\omega)+1}{2} \qquad \text{with} \qquad \omega \in [0, 90].$$

Logarithmic Data with an Interpolated Polynomial

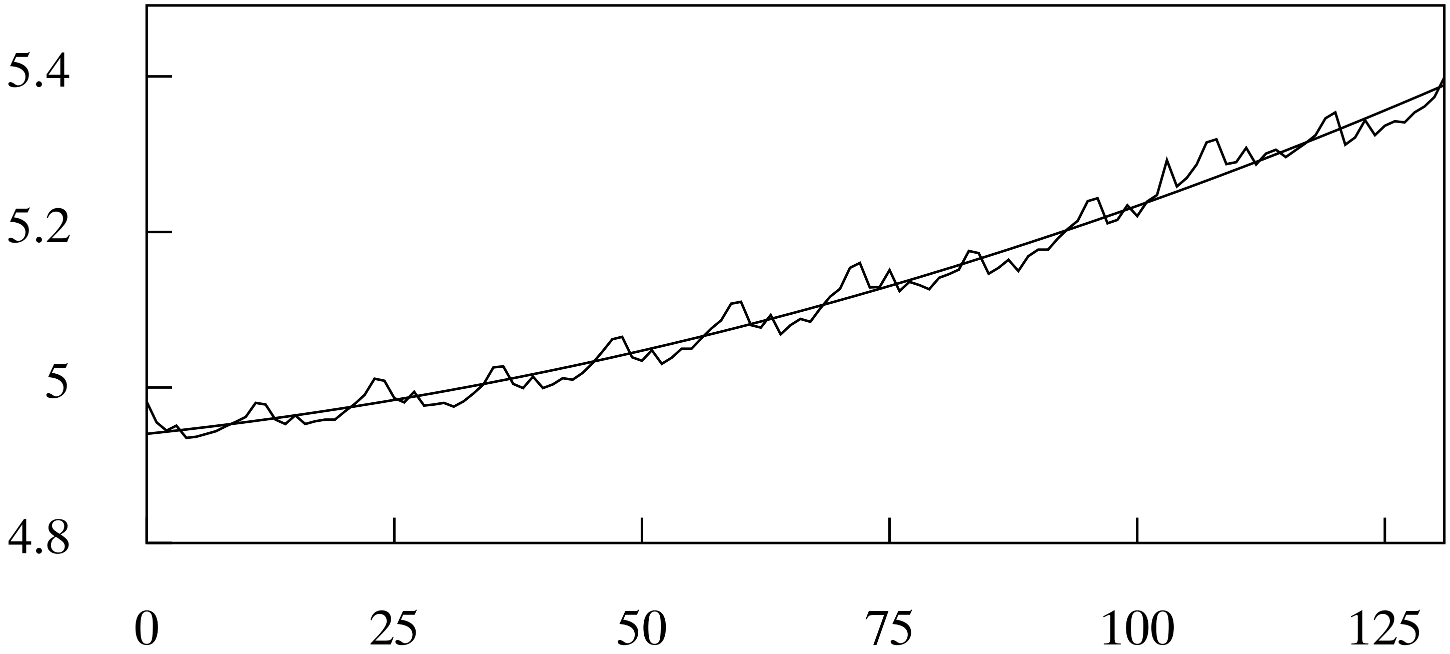

Figure 8. The plot of 132 monthly observations on the logarithms of the U.S. money supply, beginning in January 1960. A quadratic function has been interpolated through the data.

The Periodogram of the Detrended Data

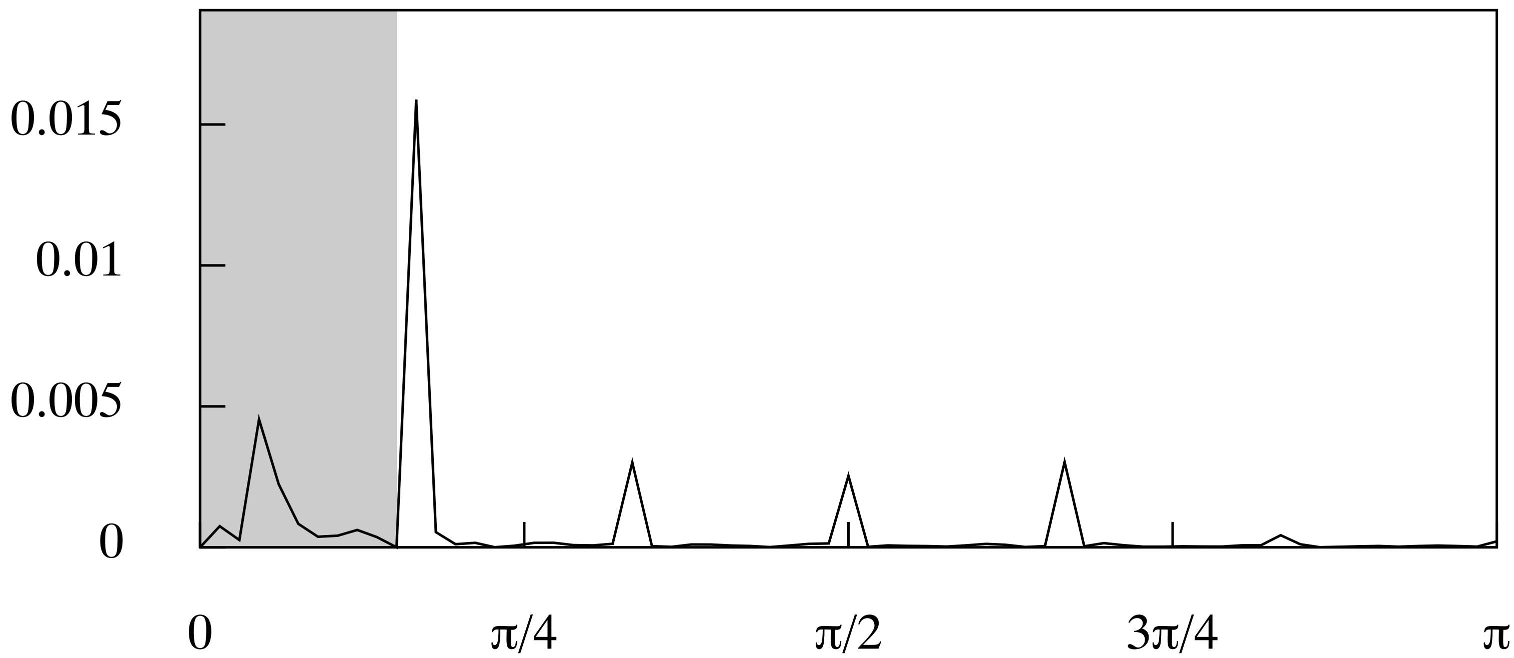

Figure 9. The periodogram of the residuals from the quadratic detrending of the logarithmic money-supply data.

The Detended Data with Interpolated Pseudo Data

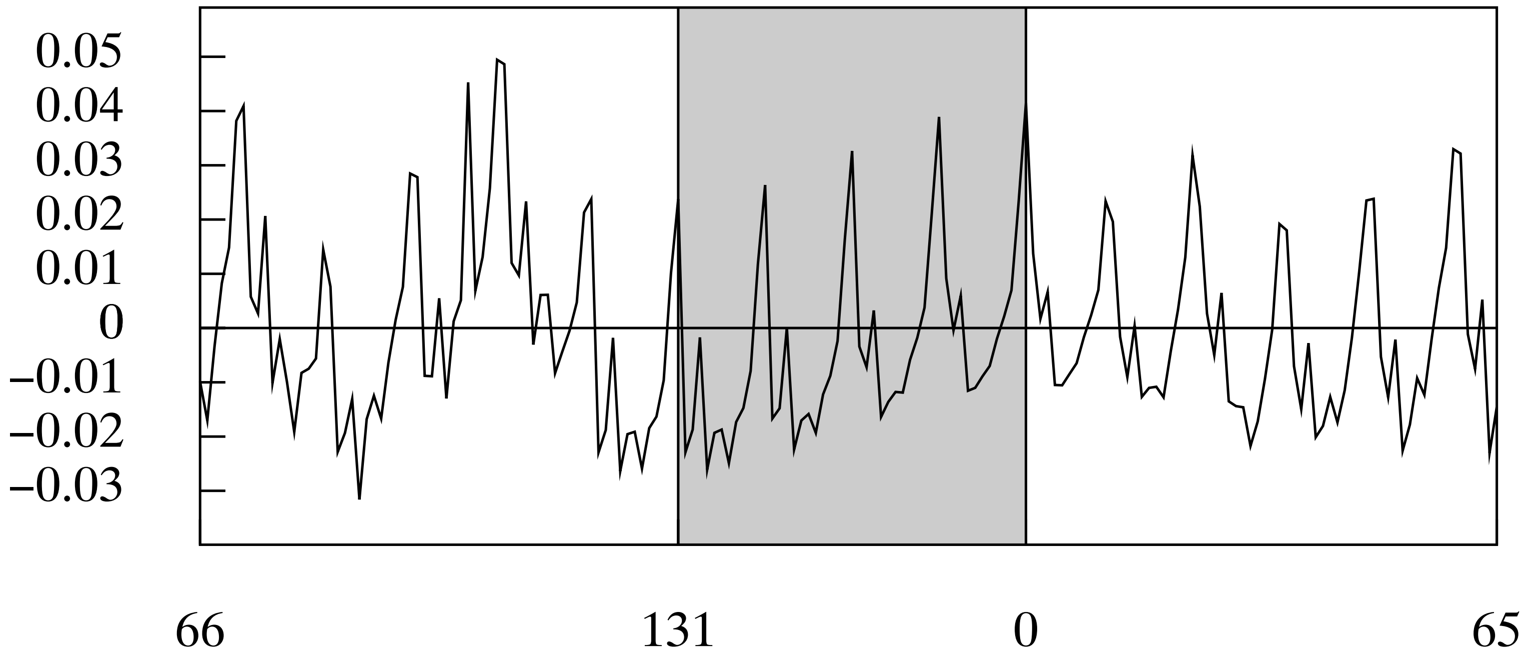

Figure 10. The residuals from a linear detrending of the logarithmic money-supply data, with an interpolation of four years length inserted between the end and the beginning of the circularised sequence, marked by the shaded band.

Logarithmic Data with a Trend-Cycle

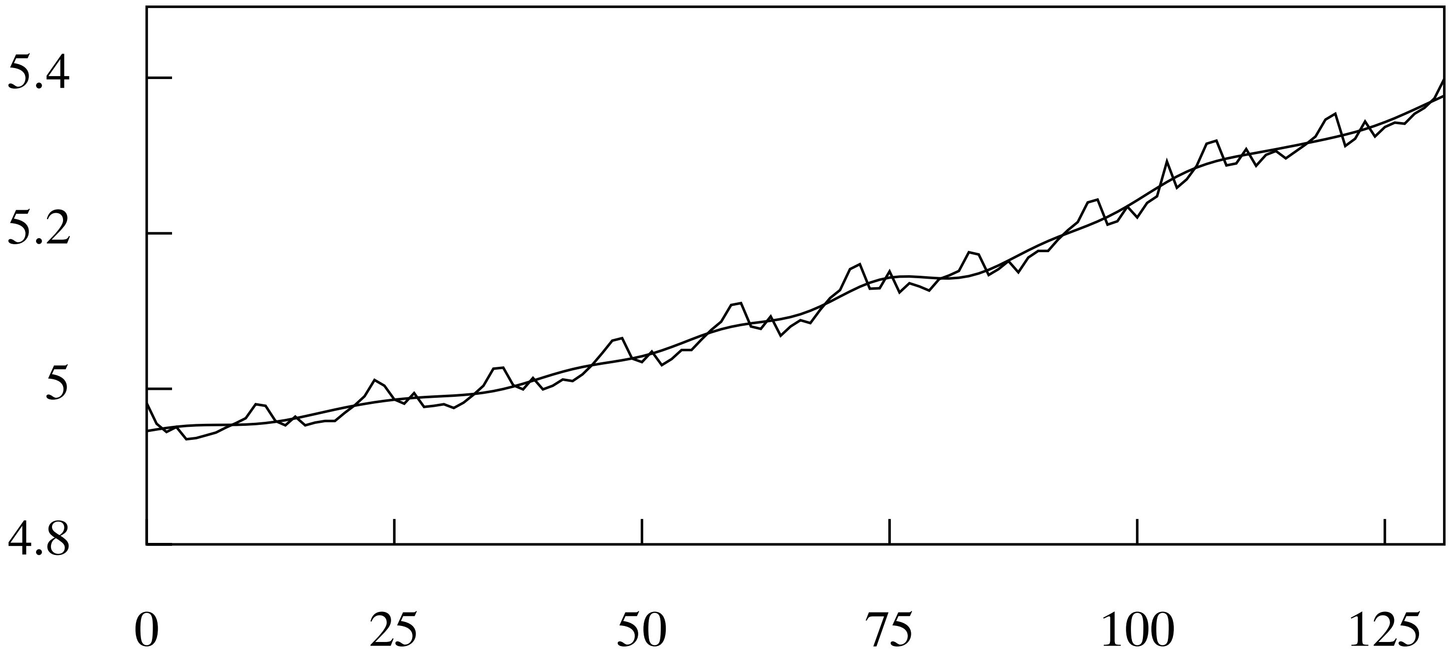

Figure 11. The plot of the logarithms of 132 monthly observations on the U.S. money supply, beginning in January 1960. A trend-cycle, estimated by the Fourier method, has been interpolated through the data.

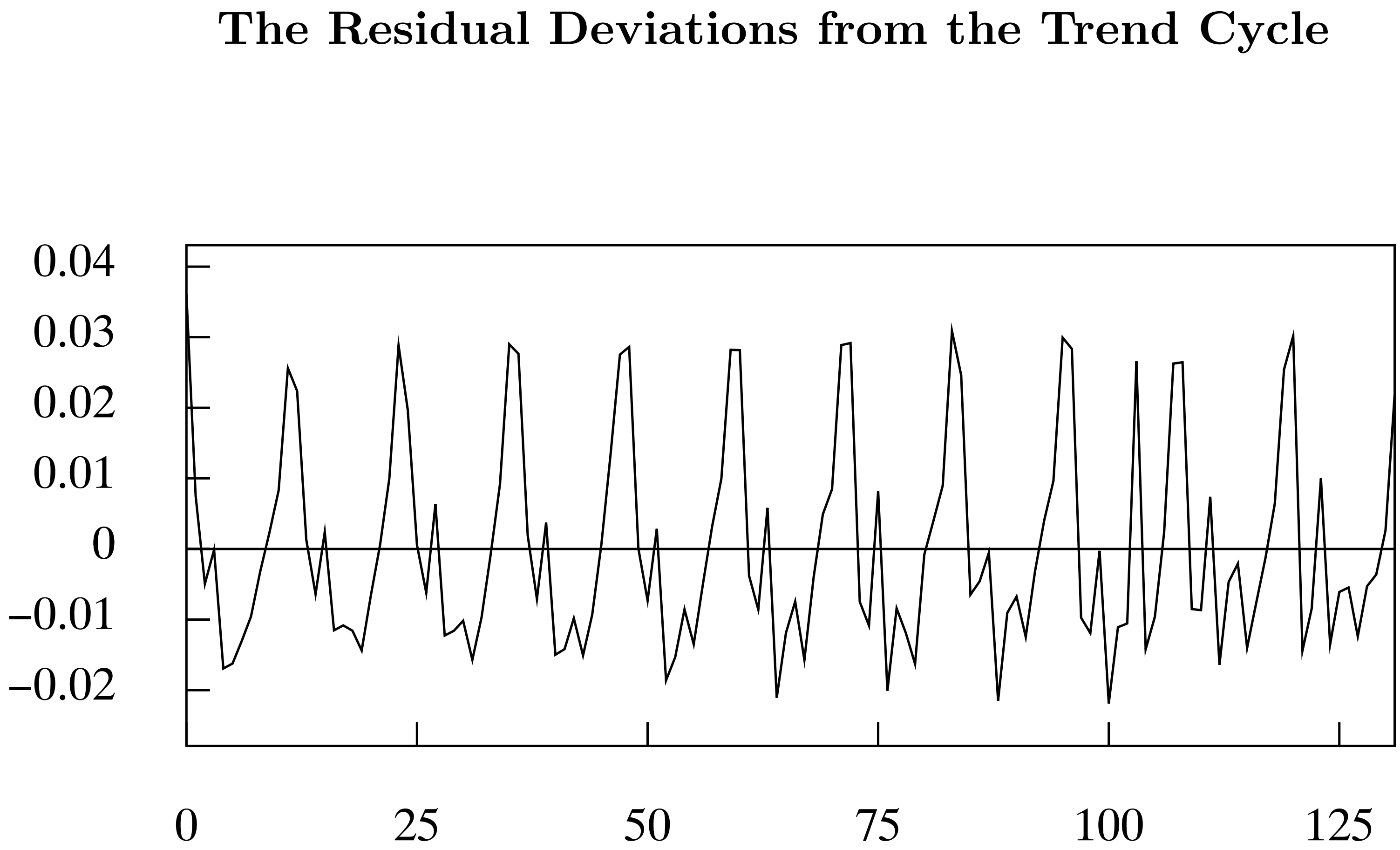

Figure 12. The sequence of residual deviations of the logarithmic money supply data from the estimated trend-cycle function.