SHARP FILTERS FOR BUSINESS-CYCLE ANALYSIS

by D.S.G. Pollock

Queen Mary and Westfield College

University of London

This paper describes a technique for constructing linear filters which are intended as tools for the analysis of business cycles. Such filters are adapted to the task of extracting the trend from an econometric times series when it is capable of being described in terms of a well-defined range of frequencies with a firm upper limit.

1. Introduction

This paper describes a new technique for constructing linear filters which are intended as tools for the analysis of business cycles.

In business-cycle analysis, much of the interest lies in the comparison of the trajectories of economic variables from collections of neighbouring countries or from the same country in different periods. For such purposes, it is usually sufficient to define a trend in broad terms as the low-frequency motion which underlies a series, and any filtering technique which attenuates the high-frequency elements of the series which are noisy or distracting, such as the annual seasonal cycle, will serve the purpose of representing the trend.

However, it is also interesting to compare the relative amplitudes of secular cycles and seasonal cycles as well as the time lags in the impacts of external economic events. For such purposes, it is essential to remove the trend from the data series; and this calls for an exacting definition of the trend which avoids confusing it with the other motions which are the objects of the study.

The filters which have been used hitherto by economists for the purposes of detrending, such as the Hodrick–Prescott filter [7], usually show a gradual transition between the pass band, which selects the low-frequency elements which correspond to the trend, and the stop band, which suppresses the elements of higher frequency which are not part of the trend. Such filters are not suited to the task of extracting a trend when it is described in terms of a narrowly-defined range of frequencies with a firm upper limit. The Hodrick–Prescott filter, which is closely related to the smoothing spline of Reinsch [12], has been analysed by Kydland and Prescott [8], by Cogley and Nason [3] and by Harvey and Jaeger [6], amonst others.

In some areas of science and engineering, the techniques for constructing frequency-selective filters are well established. In audio-acoustic engineering, for example, the data series are liable to lengthy and, in a broad sense, they manifest the property of stationarity. In such cases, moving-average filters, or FIR (finite impulse response) filters as they are also know, can be used to produce a phase-free frequency response of which the gain approximates the ideal square-wave with a high degree of accuracy. Such accuracy is purchased at the cost of a wide filter span; but this causes no difficulty.

In econometric analysis, the data series are often of a strictly limited duration and they are liable to be strongly trended. In such cases, we cannot afford to use a wide-span filter. Moreover, the non-stationarity of the series means that we are bound to pay particular attention to the start-up problem, which entails the question of how we should begin the process of filtering at one end or the other of the data series without sacrificing any of the data.

In this paper, we present a new technique for designing a frequency-selective filter which is based upon a rational-function approximation to the ideal square wave. The resulting filter uses relatively few coefficients and it also displays a rapid transition between the pass band and the stop band. The technique employs some tools which are also used in the process of estimating an ARMA time-series model.

There are some costs which have to be paid for obtaining the favourable characteristics of the rational filter. The first cost is that the filter operates on the border of instability. It tends to accumulate rounding error rapidly unless both the filter coefficients and the filtered sequence are represented with high precision.

The second cost of the sharpness of the frequency response is that the filter violates the principle of complementarity which requires that every high-pass filter should be matched by a corresponding low-pass filter such that the sum of the outputs of the two filters is equal to the original series. In our case, the complementary filter would entail complex-valued coefficients and therefore, for practical purposes, it cannot be defined.

In the final section of the paper, we provide a example of the use of a rational filter in detrending a series of unemployment statistics. The detrended sequence shows a remarkable regularity in its seasonal fluctuations which is not revealed when the Hodrick–Prescott filter is used. The example demonstrates the advantages of a sharp filter with a well-defined cut-off point and with a rapid transition between the pass band and the stop band.

2. Non-Recursive Square-Wave Filters

A non-recursive discrete-time filtering operation entails a linear combination of successive elements of a signal sequence $y(t) = \{y_t; t = 0, \pm 1, \pm 2, \ldots\}$. Such an operation can be represented by the equation

$$x(t) = \psi(L)y(t) = \sum_{j=-p}^{q} \psi_j y(t-j) \tag{1}$$

wherein

$$\psi(L) = \psi_{-p}L^{-p} + \cdots + \psi_{-1}L^{-1} + \psi_0 I + \psi_1 L^1 + \cdots + \psi_q L^q \tag{2}$$

is described as an FIR (finite impulse response) filter. Here L stands for the lag operator whose effect on the sequence $y(t)$ is described by the equation $Ly(t) = y(t-1)$.

The frequency-response function $\psi(\omega)$ of the filter is the Fourier transform of the sequence $\{\psi_j\}$ of the filter's coefficients:

$$\psi(\omega) = \sum_j \psi_j e^{-i\omega j}. \tag{3}$$

This function can also be depicted as the result of setting $z = e^{-i\omega}$ in the z-transform $\psi(z) = \sum \psi_j z^j$ of the sequence. We may describe $\psi(z)$ as the generating function of the filter.

In general, $\psi(\omega)$ is a complex-valued function which can be expressed as

$$\psi(\omega) = |\psi(\omega)|e^{-i\varphi(\omega)}, \tag{4}$$

where $|\psi(\omega)| = \sqrt{\psi(\omega)\psi(-\omega)}$ is the modulus of the function and $-\varphi(\omega)$ is its argument. The modulus of the complex function represents the gain of the filter, which is the factor by which filter alters the amplitudes of the cyclical elements of a signal. The argument of the function represents the phase effect of the filter, which is its propensity to impose delays on the signal elements.

A simple idea for the design of a linear filter is to specify the desired frequency-response function $\psi(\omega)$ in terms of its gain and phase characteristics and then to attempt to find the corresponding filter coefficients by applying an inverse Fourier transform.

An ideal frequency-selective filter is one for which the gain is unity over a certain range of frequencies, described as the passband, and zero over the remaining frequencies, which constitute the stopband. Thus the gain an ideal low-pass filter with a cut-off point at the frequency ω_c is given by the square-wave function

$$|\psi(\omega)| = \begin{cases} 1, & \text{if } |\omega| \leq \omega_c; \\ 0, & \text{if } \omega_c < |\omega| \leq \pi. \end{cases} \tag{5}$$

If a condition of symmetry is imposed such that $\psi(-\omega) = \psi(\omega)$, then $\psi(\omega) = |\psi(\omega)|$ becomes a real-valued function, and the filter has no phase effect. In that case, the function is also idempotent such that $\psi(\omega) = \psi^2(\omega)$.

The square-wave function is also positive semi-definite in the sense that

$$0 \leq \frac{1}{2\pi i}\oint \lambda(z)\psi(z)\lambda(z^{-1})\frac{dz}{z} = \frac{1}{2\pi i}\oint |\lambda(z)\psi(z)|^2 \frac{dz}{z}, \tag{6}$$

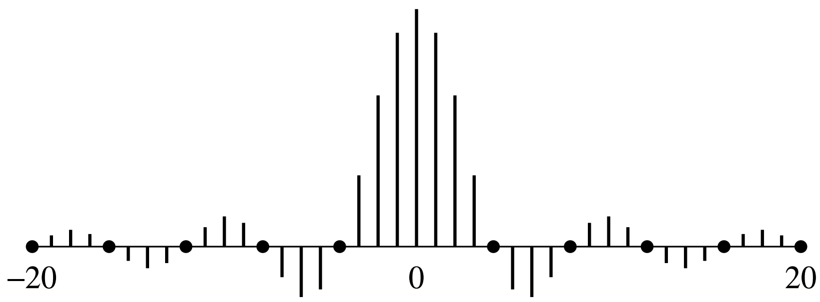

Figure 1. The central coefficients of the Fourier transform of a square wave with a jump at $\omega_c = \pi/4$. The sequence of coefficients, which represents the impulse response of an ideal low-pass filter, extends indefinitely in both directions.

where $\lambda(z)$ is any polynomial or power series in z. When the locus of this contour integral is the unit circle, the expression stands for the integral of the squared gain of the composite filter $\lambda(\omega)\psi(\omega)$. Observe that an equality holds when $\lambda(z) = 1 - \psi(z)$. In that case, $\lambda(\omega)$ represents an ideal high-pass filter which is the complement of $\psi(\omega)$ and which nullifies the output of the latter which is nonzero only over the frequency interval $(-\omega_c, \omega_c)$

Given that the function $\psi(\omega)$ is real-valued and even, i.e. symmetric about point $\omega = 0$, it follows that its transform will give rise to a sequence of filter coefficients $\{\psi_j\}$ which is also real-valued and even with $\psi_{-j} = \psi_j$. The latter is a necessary and sufficient condition for the absence of a phase effect. Thus, for the ideal square-wave filter, we have

$$\psi_j = \frac{1}{2\pi}\int_{-\omega_c}^{\omega_c} e^{i\omega j} d\omega = \begin{cases} \dfrac{\omega_c}{\pi}, & \text{if } j = 0; \\ \dfrac{\sin(\omega_c j)}{\pi j}, & \text{if } j \neq 0\ . \end{cases} \tag{7}$$

However these coefficients constitute a sequence which extends indefinitely in both directions; and, since a practical filter can contain only a limited number of coefficients, we are bound to accept approximations to the ideal filter.

In general, we can represent the practical versions of the square-wave filter as the product of the desired impulse response and a finite-duration window function $\{\kappa_j\}$. The simple process of truncating the sequence $\{\psi_j\}$ entails a

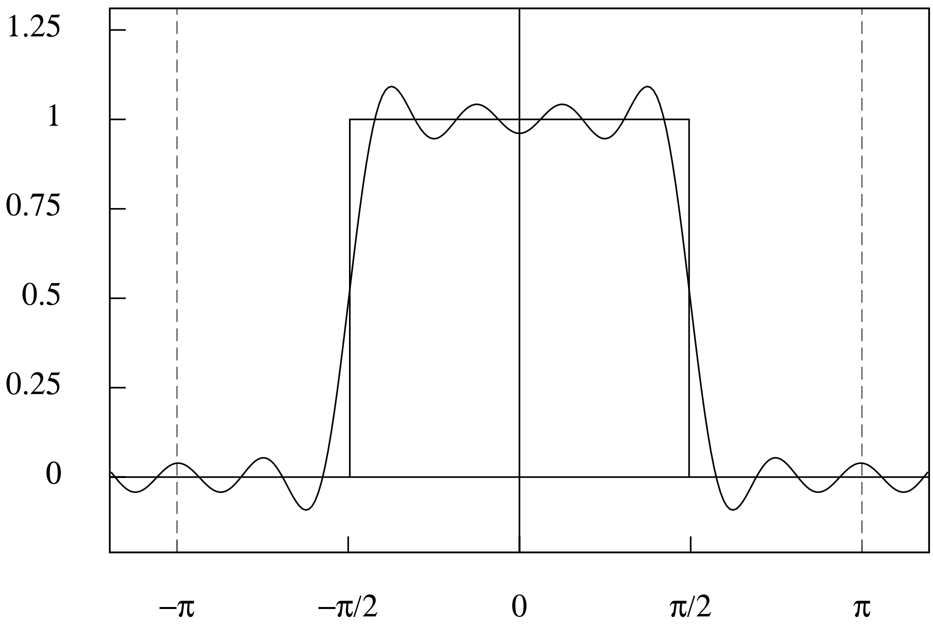

Figure 2. The result of applying a 17-point rectangular window to the coefficients of an ideal low-pass filter with a cut-off point at $\omega = \pi/2$.

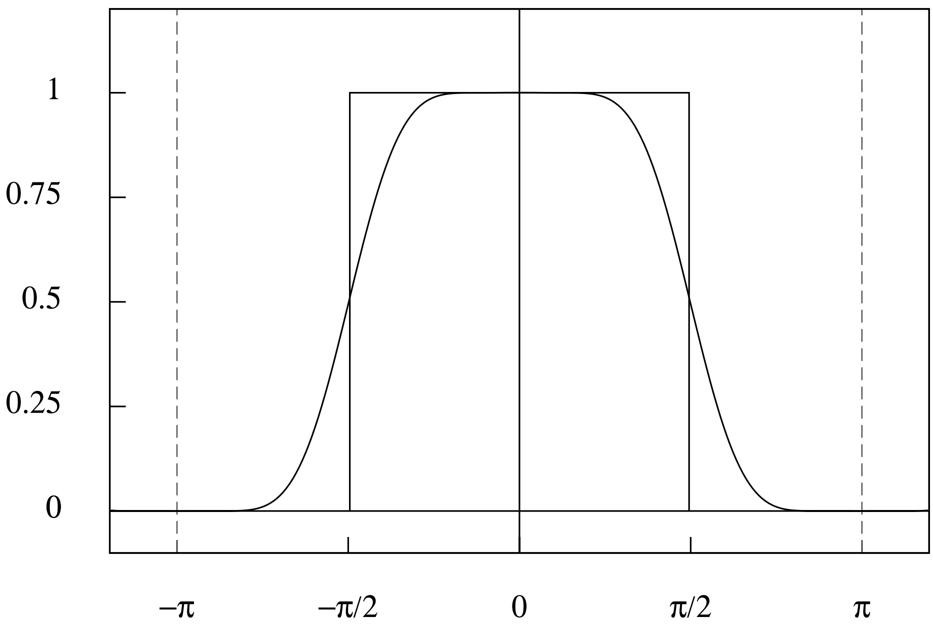

Figure 3. The result of applying a 17-point Blackman window to the coefficients of an ideal low-pass filter with a cut-off point at $\omega = \pi/2$.

rectangular window defined by

$$(8)\qquad \kappa_j = \begin{cases} 1, & \text{if } |j| \leq M; \\ 0, & \text{if } |j| > M. \end{cases}$$

The various effects of such a truncation of the ideal impulse-response function can be seen by examining the transform of the windowed sequence

$$(9)\qquad \begin{aligned} \{\psi_{Mj}\} &= \{\psi_j \kappa_j; j = 0, \pm 1, \ldots, \pm M\} \\ &= \{\psi_j; j = 0, \pm 1, \ldots, \pm M\}. \end{aligned}$$

This is represented in Figure 2 for the case where $M = 8$. First, the sharp transition in the ideal response from unity to zero at the cut-off frequency of $\omega = \omega_c$ has been converted to a gradual transition. Next, in the passband, where $|\omega| < \omega_c$, the constant value of $\psi(\omega) = 1$ has given way to a series of ripples which attain a maximum amplitude at the end of the interval. Finally, in the stopband, where $|\omega| > \omega_c$, the zero value of $\psi(\omega)$ has also given way to a series of ripples. The greatest of these ripples is the one which is adjacent to ω_c; and it is the counterpart of the ripple at the end of the passband. The ripples in the stopband give rise to a problem known as spectral leakage—for their presence implies that the filter will fail to annihilate completely some of the high-frequency elements of a signal.

With one exception, the magnitude of the ripples and the extent of the leakage will be reduced by an increase in the value of M, which corresponds to the span of the filter. The exception concerns the oscillations which are adjacent to the point of discontinuity at $\omega = \omega_c$. Their magnitude will tend to a value which is about 9% of the jump, albeit that the width of the oscillations, and hence the value of their integral, will become vanishingly small. The persistence of these oscillations, which implies the failure of the Fourier transform of the filter coefficients to converge uniformly to the ideal square wave, is know as Gibb's phenomenon [4] [5]. (See also Carslaw [2]).

Some of these problems can be avoided by adopting an alternative window sequence in place of the rectangular sequence of (8). Whereas windows are readily available which suppress the ripples in the passband and the leakage in the stop band, they will do so invariably at the cost of a more gradual transition between the two bands. One window function which is used quite commonly to improve the performance of a practical square-wave filter is the Blackman window (see Blackman and Tukey [1]) which is defined by

$$(10)\qquad \beta_j = 0.42 + 0.5 \cos\left(\frac{\pi j}{M}\right) + 0.08 \cos\left(\frac{2\pi j}{M}\right), \qquad \text{where} \qquad |j| \leq M.$$

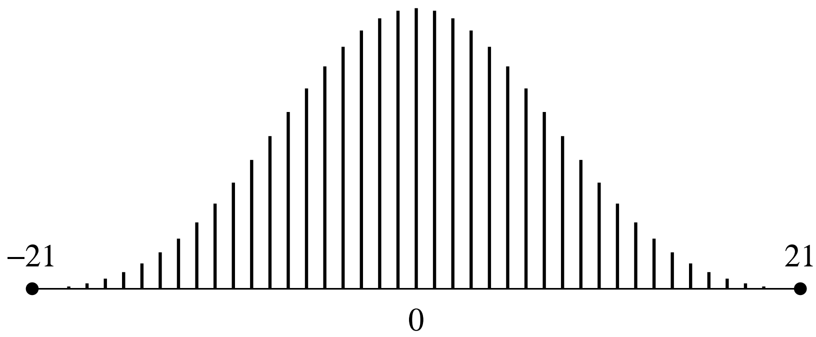

Figure 4. The coefficients of a Blackman window.

The sequence formed by applying the Blackman window to the square-wave coefficients may be denoted by

$$\{\psi_{Bj}\} = \{\psi_j\beta_j; j = 0, \pm 1, \ldots, \pm M\}. \tag{11}$$

Figure 3 shows the effect of applying the Blackman window to the sequence of the coefficients which have given rise to Figure 2. The comparison of the two figures shows clearly that the cost eliminating the passband ripples and the leakage into the stopband is a much-reduced rate of transition between the two bands. It is with the aim of avoiding this cost that we will now pursue alternative techniques of approximating the ideal square wave.

3. Recursive Square-Wave Filters

A recursive discrete-time filtering operation is one which entails a process of feedback whereby the output of the filter becomes part of its input after a minimum delay of one period. Such an operation may be represented by the equation

$$\begin{aligned} x(t) = \theta_0 y(t) + \theta_1 y(t-1) + \cdots + \theta_q y(t-q) \\ - \phi_1 x(t-1) - \cdots - \phi_p x(t-p), \end{aligned} \tag{12}$$

which can be written in terms of lag-operator polynomials as

$$\phi(L)x(t) = \theta(L)y(t). \tag{13}$$

The rational function $\theta(L)/\phi(L)$ is commonly described as an IIR (infinite impulse response) filter. The condition is invariably imposed that the roots of the polynomial $\phi(z) = 0$ must lie outside the unit circle; and this is the stability condition which is necessary and sufficient for the existence of a convergent series expansion of the rational operator. The coefficients of this expansion constitute the impulse response of the filter.

An ordinary recursive filter, which has a one-sided power-series expansion, is bound to impose a phase lag on the processed series $x(t)$; and this lag will take different values at different frequencies. However, unless there is a need to perform the operations in real time, it is possible to eliminate the phase lag by applying the filter a second time in the reverse direction. Such a process of bidirectional filtering may be described by the equations

$$\text{(14)} \qquad \text{(i)} \quad \phi(L)w(t) = \theta(L)y(t) \qquad \text{and} \qquad \text{(ii)} \quad \phi(F)x(t) = \theta(F)w(t),$$

where $F = L^{-1}$ is the forward-shift operator whose effect on the sequence $w(t)$ is described by the equation $Fw(t) = w(t+1)$. The two filtering operations can be represented as a combined operation by defining a symmetric two-sided rational filter in the form of

$$\text{(15)} \qquad \psi_R(L) = \frac{\theta(F)\theta(L)}{\phi(F)\phi(L)}.$$

The function $\psi_R(z)$ is amenable to a Laurent expansion which gives rise to a symmetric sequence of coefficients $\{\psi_{Rj}\}$. Applying the Fourier transform to these coefficients in the manner of equation (3) generates the frequency-response function $\psi_R(\omega)$ of the bidirectional filter. In practice however, given the doubly-infinite nature of the Laurent expansion, we should calculate the response directly from the coefficients of $\theta(z)$ and $\phi(z)$.

The objective in designing a rational square-ware filter is to determine the orders of the operators $\theta(L)$ and $\phi(L)$ and the values of their coefficients so as to ensure that the frequency-response function $\psi_R(\omega)$ is an effective approximation to the ideal square wave. The sequence of square-wave coefficients defined in (7) constitutes a positive-semi-definite function which is analogous to the autocovariance generating function of a stationary stochastic process. The matter of determining the operators $\theta(L)$ and $\phi(L)$ is therefore analogous to the problem of inferring the parameters of an ARMA process from its autocovariances.

To show how the coefficients of $\theta(z)$ and $\phi(z)$ are found, let us imagine that $\psi_R(z)$ is know via the coefficients of the Laurent expansion. Then an equation can be written in the form of

$$\text{(16)} \qquad \begin{aligned} \psi_R(z)\phi(z) &= \frac{\theta(z)\theta(z^{-1})}{\phi(z^{-1})} \\ &= \delta(z) \end{aligned}$$

where, in consequence of the degree q of the polynomial $\theta(z)$, the expansion

$$\delta(z) = \{\cdots + \delta_{-1}z^{-1} + \delta_0 + \delta_1 z^1 + \cdots + \delta_q z^q\} \tag{17}$$

has z^q as the highest power of z. Dividing both sides of equation by z^{q+1} eliminates the nonnegative powers of z. Therefore, using the notation of Whittle [13], we can denote the series in the nonnegative powers by $[\psi_R(z)\phi(z)/z^{q+1}]_+ = 0$; and, with a slight elaboration of the notation, we can also write

$$\left[\frac{\psi_R(z)\phi(z)}{z^{q+1}}\right]_{(0,p)} = 0, \tag{18}$$

where the subscript $(0, p)$ indicates only the terms in $z^0, z^1, \ldots, z^p$ have been taken. Given the normalisation $\phi_0 = 1$, equation (18), which is analogous to the normal equations of a linear regression, will serve to determine the coefficients of $\phi(z)$.

Once $\phi(z)$ has been determined, the equation

$$\phi(z)\psi_R(z)\phi(z^{-1}) = \theta(z)\theta(z^{-1}) \tag{19}$$

can be formed from which the terms associated with the powers $z^0, z^{\pm 1}, \ldots, z^{\pm q}$ may be extracted:

$$\left[\phi(z)\psi_R(z)\phi(z^{-1})\right]_{(-q,q)} = \theta(z)\theta(z^{-1}). \tag{20}$$

Here we have omitted to apply the subscript to the RHS for the reason that z^q and z^{-q} are manifestly the highest powers z in the product $\theta(z)\theta(z^{-1})$.

Solving equation (20) for the coefficients of $\theta(z)$ is a matter of finding the Cramér–Wold decomposition of the Laurent polynomial on the LHS. The algorithm of Wilson [14], which is based on the Newton–Raphson procedure, is an effective way of achieving the factorisation; and versions which are coded in *C* and in *Pascal* have been provided by Pollock [11] (See, also, Laurie [9], [10]).

The remaining issue of this section is the question of how to specify the coefficients of the Laurent expansion of the function $\psi_R(z)$ from which the coefficients of the rational filter are to be determined by the algorithm described above. The simplest prescription is to offer to the algorithm the leading coefficients $\psi_j; j = 0, \ldots, p + q$ of the ideal square-wave $\psi(z)$. However, the ideal function is positive semi-definite, and, for the algorithm to work, $\psi_R(z)$ must be positive definite.

In fact, when they are calculated in finite-precision arithmetic in a manner which reflects the definition of (7), the square-wave coefficients are unlikely to constitute even an positive semi-definite function.

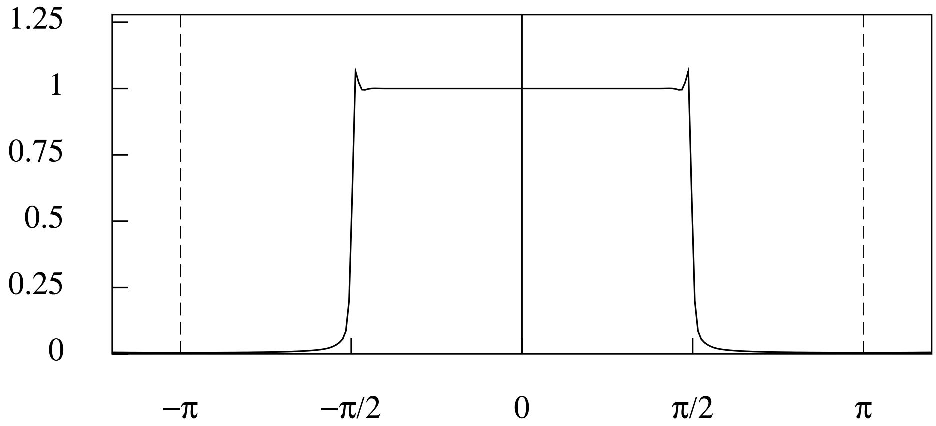

Figure 5. The gain of a bidirectional low-pass rational filter with a numerator order of $q = 8$ and a denominator order of $p = 12$. The cut-off point is at $\pi/2$.

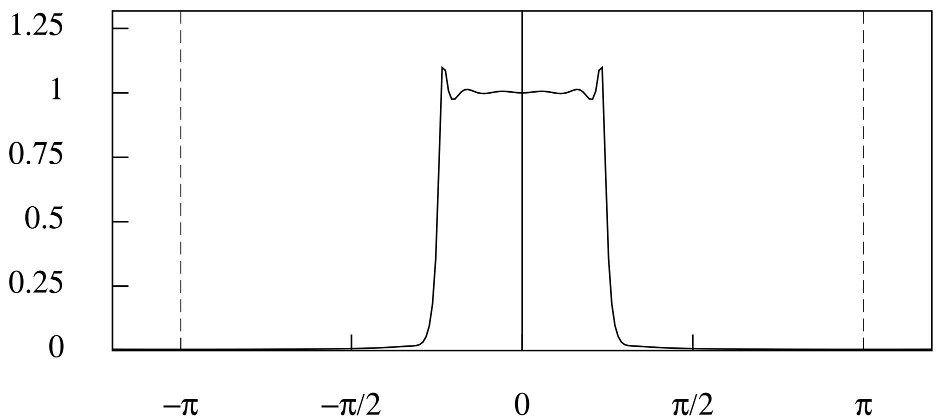

Figure 6. The gain of a bidirectional low-pass rational filter with a numerator order of $q = 6$ and a denominator order of $p = 8$. The cut-off point is at $\pi/4$.

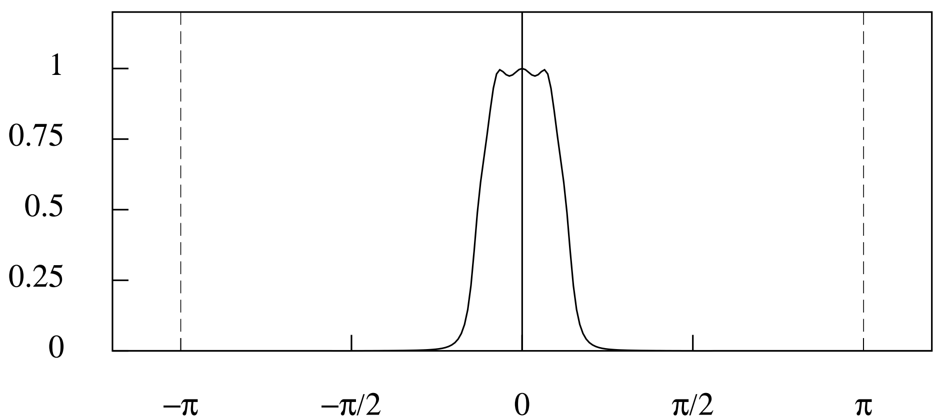

Figure 7. The gain of a bidirectional rational filter with $p = q = 6$ and with a cut-off point $\pi/8$. A Blackman window of 30 points has been applied to the coefficients of the ideal square wave prior to calculating the rational coefficients.

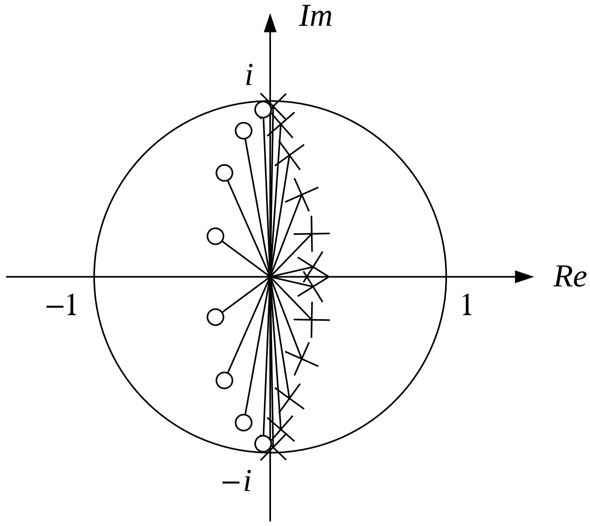

Figure 8. The pole–zero diagram of the rational function $\theta(z^{-1})/\phi(z^{-1})$. The degree of $\theta(z^{-1})$ is $q = 8$ and the degree of $\phi(z^{-1})$ is $p = 12$. The cut-off point of the corresponding filter is at $\pi/2$. The poles are marked with crosses and the zeros with circles.

The objective is therefore to find a positive-definite function which approximates to $\psi(z)$ as closely as possible. For this purpose, it is appropriate to specify that

$$\psi_R(z) = \psi_T(z)\psi_T(z^{-1}), \tag{21}$$

where $\psi_T(z)$ is the z-transform of a sequence $\{\psi_{Tj}\}$ of coefficients formed from the ideal square-wave sequence $\{\psi_j\}$ by processes of truncation and windowing. The autoconvolution of the $\{\psi_{Tj}\}$, which is entailed in the product on the RHS of equation (21), will ensure that, in practice, the function $\psi_R(z)$ is positive definite. The properties of symmetry and idempotency whereby $\psi(z) = \psi(z)\psi(z^{-1})$, which characterise the ideal square-wave filter, will ensure that $\psi_R(z)$ will be close to $\psi(z)$ if $\psi_T(z)$ is close.

Two examples of $\{\psi_{Tj}\}$, which can be used in forming $\psi_R(z)$, are provided by the ordinary truncated sequence of (9) and sequence of (11) which is obtained by applying the Blackman windows to the truncated sequence. These sequences have already been investigated from the point of view of their use in forming FIR filters. The conclusion was that, for realistic values of the parameter M which denotes the order of the filter, there were severe inadequacies in the FIR filters. Now, for the purposes of forming the function $\psi_R(z)$, we are proposing to take much larger values of M. Experience has demonstrated that the appropriate value lies anywhere between 30 and 50.

Figures 5 and 6 represent the gain functions of two rational square-wave filters, each calculated from a truncated sequence of the coefficients of an ideal square wave. The truncation point is given by $M = 50$ in both cases. By any standards, the transition between the pass band and the stop band is extremely rapid. It will be noticed that the profiles of both of these filters have spikes or ears at the ends of the pass band. This feature leads to the violation of the condition that the gain of the filter should never exceed unity, which is entailed in the principle of complementarity which has been enunciated in the introduction. This is one of the costs of securing a rapid transition.

The rapid transition of the gain function is due to the particular placement in the complex plane of the poles and zeros of the filter. Figure 8, which is the pole–zero diagram corresponding to the gain function of Figure 5, reveals two pole–zero pairs which lie close to the intersections of the unit circle with the imaginary axis. The location of these poles, which accounts for the abrupt transition in the vicinity of $\pi/2$, implies that the filter is operating on the border of instability.

Figure 7 represents the gain of a rational square-wave filter calculated from a sequence of coefficients obtained by applying a Blackman window with $M = 30$ to the coefficients of an ideal square wave. The effect of the window has been to reduce the amplitude of the ripples on the profile of the pass band, including the ears at the ends. This has been at the cost of a slower transition between the pass band and the stop band.

4. Filtering Finite Sequences

In the preceding sections, the assumption has been made that the domain of the filter is the set of sequences indexed by $\{t = 0, \pm 1, \pm 2, \ldots\}$ which is the set of all positive and negative integers. However, the data which will be processed in practice has only a limited duration.

The principle problem which affects the filtering of limited samples is how to represent the values which fall outside the sample period. In cases where the data can be modelled by a stationary stochastic process with a mean of zero, it may be appropriate to represent the extra-sample values by zeros; and the filter coefficients can be adapted accordingly when the ends of the sample are reached.

In econometric analysis, it is common to find data sequences which are nonstationary and for which the extra-sample values are unbounded. Then there is no possibility of replacing the values by zeros; and a common recourse is to extend the sample by extrapolation. However, the difficulties of filtering a nonstationary sequence can be circumvented by applying the filter to a transformed sequence which has been reduced to stationarity by successive differencing. Thereafter, the filtered version of the trended sequence can be recovered from its differenced version.

The method of differencing can be summarised in the existing notation which relates to indefinite sequences. Let the differencing operator be denoted by

$$(1-L)^d = \delta(L) = 1 + \delta_1 L + \cdots + \delta_d L^d, \tag{22}$$

and let the differenced sequence and its filtered version be denoted by

$$d(t) = \delta(L)y(t) \qquad \text{and} \qquad z(t) = \psi(L)d(t) \tag{23}$$

respectively. Then the filtered version of the original data sequence is

$$x(t) = \psi(L)y(t) = \delta^{-1}(L)z(t) \tag{24}$$

and, given a set of initial conditions $x_0, x_1, \ldots, x_{d-1}$, the succeeding values of $x(t)$ can be generated by a simple process of accumulation based on the equation

$$x_t = z_t - \delta_1 x_{t-1} - \cdots - \delta_d x_{t-d}. \tag{25}$$

The matter of finding the initial conditions with which to begin the process of recovering the $x(t)$ from $z(t)$ is straightforward.

An alternative approach in generating the filtered version of a trended sequence makes use of the residual sequence obtained by filtering the differenced sequence $d(t)$. Thus, if

$$h(t) = d(t) - z(t) \tag{26}$$

is the residual sequence, then, instead of equation (24), there is

$$x(t) = y(t) - \delta^{-1}(L)h(t), \tag{27}$$

and, therefore, we have the option of cumulating $h(t)$ instead of $z(t)$. Moreover, since $\delta^{-1}(L)h(t)$ is likely to be stationary, the danger is reduced that the process of cumulation will be affected by numerical rounding errors.

In representing the finite-sample versions of the filter and in developing the associated algorithms, a vector notation is called for; and it is helpful to employ the finite-sample version of the lag operator which is in the form of a matrix. Let the data be indexed by $t = 0, \ldots, T-1$ and let the identity matrix be denoted by

$$I_T = [e_0, e_1, \ldots, e_{T-1}], \tag{28}$$

where e_j represents a column vector with a unit in the position j—counting from zero—and with zeros elsewhere. Then the finite-sample lag operator is the matrix

$$L_T = [e_1, \ldots, e_{T-1}, 0] \tag{29}$$

which has units on the first subdiagonal and zeros elsewhere. This matrix may be formed by deleting the leading vector of the identity matrix and by appending a zero vector to the end of the array.

The finite-sample operator L_T is distinguished from the ordinary operator $L = L_\infty$ by the fact that it is nilpotent of degree T such that $(L_T)^T = 0$. There is also a straightforward correspondence between negative powers of L and the powers or $F_T = L_T'$; but it will be observed that the product of F_T and L_T is not the identity operator. The fault lies in the loss of units from either end of the identity matrix I_T.

The lag-operator polynomials which have characterised our analysis can be converted to matrix operators of order T simply by replacing the L by L_T. Two such matrices, which are of primary importance in this account, are the d-fold differencing matrix $\Delta = \delta(L_T)$ and its inverse $\Sigma = \Delta^{-1} = \delta^{-1}(L_T)$ which is the (d-fold) summation matrix.

Taking differences within a vector entails a loss of information. Thus, if $\Delta = [Q_*, Q]'$, where Q_*' has d rows, then the d-th differences of the vector $x = [x_0, \ldots, x_{T-1}]'$ are the elements of the vector $z = [z_d, \ldots, z_{T-1}]'$ which is found in the equation

$$\begin{bmatrix} z_* \\ z \end{bmatrix} = \begin{bmatrix} Q_*' \\ Q' \end{bmatrix} x. \tag{30}$$

The vector $z_* = \Delta_* x_*$ in this equation is a transform of the vector $x_* = [x_0, \ldots, x_{d-1}]'$ of the initial conditions which are associated with the equation (25). The matrix of the transformation is the operator $\Delta_* = \delta(L_d)$.

To recover x from z requires the d initial conditions or constants of integration which are provided by z_*. Thus, if $\Sigma = [S_*, S]$, where S_* has d columns, then

$$x = S_* z_* + Sz. \tag{31}$$

This equation represents a non-recursive version of the cumulation described by equation (25).

The initial conditions can be determined by evaluating the criterion

$$\begin{aligned} &\text{Minimise} \quad (y - x)'(y - x) = (y - S_* z_* - Sz)'(y - S_* z_* - Sz) \\ &\text{With respect to} \quad z_*. \end{aligned} \tag{32}$$

The minimising vector is

$$z_* = (S_*' S_*)^{-1} S_*'(y - Sz). \tag{33}$$

On defining the operator $P_* = S_*(S_*' S_*)^{-1} S_*'$, which is a symmetric idempotent matrix, the estimated trend may be expressed as

$$x = P_* y + (I - P_*) Sz. \tag{34}$$

The disadvantage of this approach to finding the initial conditions is that it requires successive rows of the matrix S to be generated in the process of forming the vector Sz. Also, the elements of S_* must be generated and stored.

We can circumvent the problem of the initial conditions altogether by seeking the solution to the following problem:

$$\text{Minimise} \quad (y - x)'(y - x) \quad \text{Subject to} \quad Q'x = z, \tag{35}$$

The minimisation is accomplished by evaluating the Lagrangean function

$$L(x, \mu) = (y - x)'(y - x) + 2\mu'(Q'x - z). \tag{36}$$

By differentiating the function with respect to x and setting the result to zero, we obtain the condition

$$(y - x) - Q\mu = 0, \tag{37}$$

whence, on premultiplying by Q' and rearranging, we get

$$\mu = (Q'Q)^{-1} Q'(y - x). \tag{38}$$

Putting the final expression for μ into (37) and using the condition $Q'x = z$ gives

$$x = y - Q(Q'Q)^{-1}(Q'y - z). \tag{39}$$

This will be recognised as a disguised form of equation (27).

The advantage of this approach to recovering x from z is that, typically, the matrix Q comprises only a handful of distinct elements. Therefore the computation is relatively undemanding.

It is easy to demonstrate the equivalence of the solutions under (34) and (39). Consider writing (39) as

$$x = y - P_Q(y - x), \tag{40}$$

where $P_Q = Q(Q'Q)^{-1}Q'$. Next, one can recognise that S_* and Q are complementary matrices such that $[S_*, Q]$ has full rank whilst $Q'S_* = 0$. It follows that $P_Q = I - P_*$. This enables us to rewrite equation (40) as

$$x = P_*y + (I - P_*)x. \tag{41}$$

The equivalence of the equations (34) and (41) follows from the identity

$$\begin{aligned}(I - P_*)x &= (I - P_*)(S_*z_* + Sz)\\ &= (I - P_*)Sz.\end{aligned} \tag{42}$$

By a judicious use of the finite-sample lag operator, we can also represent the finite-sample version of the rational filter as a direct adaptation of the version given under (15). Consider the numerator of filter:

$$\theta(z^{-1})\theta(z) = M(z) = m_0 + m_1(z^{-1} + z) + \cdots + m_q(z^{-q} + z^q), \tag{43}$$

Setting $z = L_{T-d}$ and $z^{-1} = F_{T-d}$ gives rise to a Toeplitz matrix M. Likewise the denominator

$$\phi(z^{-1})\phi(z) = W(z) = w_0 + w_1(z^{-1} + z) + \cdots + w_p(z^{-p} + z^p), \tag{44}$$

gives rise to a Toeplitz matrix which we shall denote by W. Therefore a finite-sample version of the equation

$$z(t) = \psi_R(L)d(t) = \frac{\theta(F)\theta(L)}{\phi(F)\phi(L)}d(t) = \frac{M(L)}{W(L)}d(t) \tag{45}$$

is given by

$$z = MW^{-1}d = Mg, \qquad \text{where} \qquad g = W^{-1}d. \tag{46}$$

The vector g can be found via a Cholesky factorisation which sets $W = U'U$ where U is an upper-triangular matrix with zero elements above the pth supra-diagonal band. The equation $Wg = d$ can be cast in the form of $U'f = d$, where $f = Ug$, and it can be solved for f via a simple recursion which finds one element at a time. Then g can be recovered from f by a similar recursion running in the opposite direction. The main use of computer memory is in storing the elements of the matrix U' which has no simple structure, unlike the $\phi(L_T)$ which is a lower-triangular Toeplitz matrix.

The Toeplitz matrix M has zeros everywhere above the qth supra-diagonal band and below the qth subdiagonal band. Therefore the vector $z = Mg$ can be found from g via an operation of matrix multiplication which requires very little computer memory.

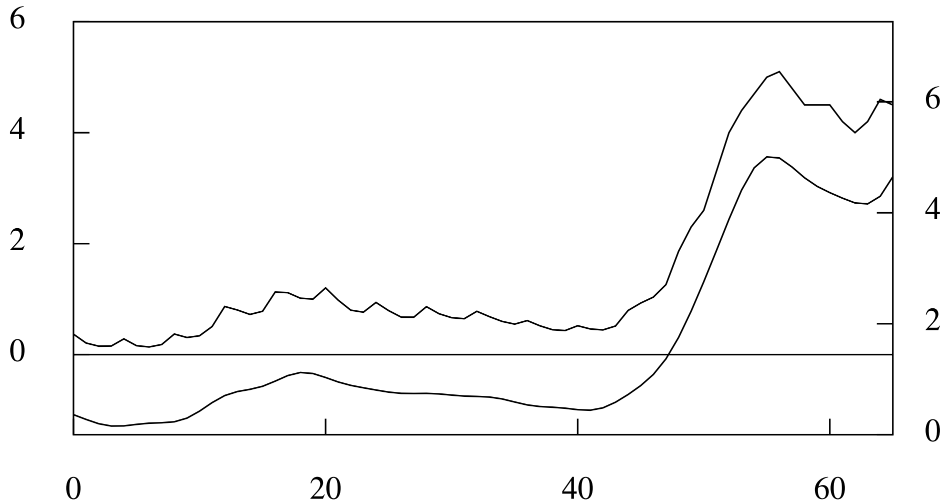

Figure 9. The quarterly figures on Swiss unemployment from 1980.1 to 1996.2 (upper panel) together with their trend obtained by smoothing the series using a rational low-pass filter with a numerator order of 7 and a denominator order of 8.

5. The Uses of the Rational Filter

We shall illustrate the uses of the rational filter by applying it to a series of 66 figures which constitute the quarterly unemployment statistics for Switzerland from the first quarter of 1980 through to the second quarter of 1996. The graph of this series is given in Figure 9. Our objective is to discover the pattern of the seasonal fluctuations which surround the longer-term trend. A casual inspection of the graph would suggest that the seasonal motions have been in abeyance in the period of rapidly increasing unemployment in the third segment of the series, only to be resumed when unemployment is stabilised at a higher level at the end of the series.

The angular velocity of the seasonal fluctuation is $\pi/2$ radians, or 90 degrees, per period; and our objective of removing the trend would be fulfilled by eliminating every component of a lesser frequency. In fact, we shall choose a nominal cut-off point for the filter of 75 degrees. This places the transition between the pass band and the stop band in an area which corresponds to a dead space in the periodogram of the data where there are no elements of any significant power. The existence of this dead space allows us to use a filter of relatively low orders which has a more gradual transition than would be tolerated in more exacting circumstances. The effects of the choices of the cut-off point and the filter orders can be seen in Figure 10.

Figure 11 shows the residuals of the series after the trend has been extracted using the low-pass filter. What is remarkable about this series is its

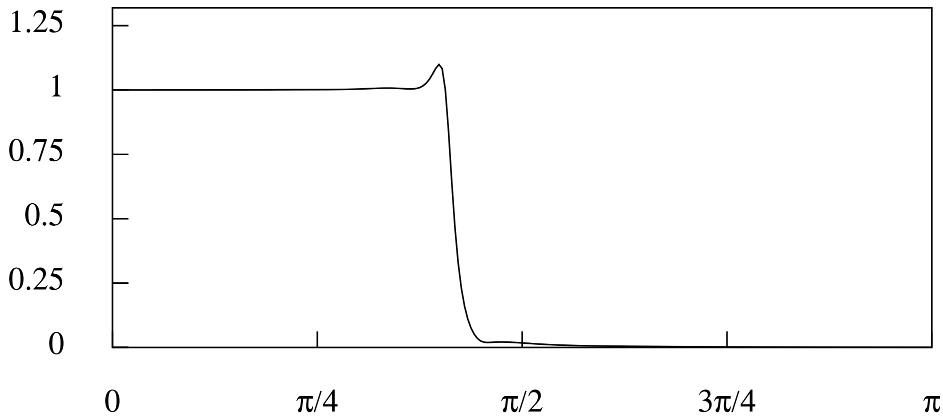

Figure 10. The gain of a low-pass rational square-wave filter with a numerator order of 7 and a denominator order of 8. The cut-off point is at 75.0 degrees.

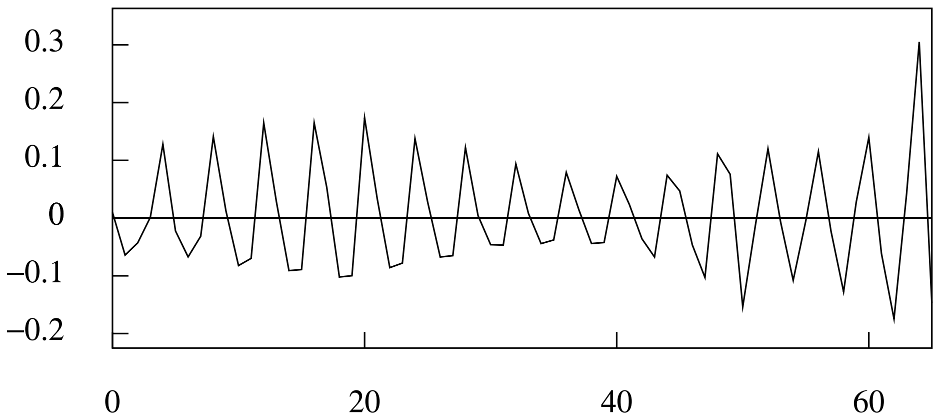

Figure 11. The residual sequence from detrending the Swiss unemployment figures using the rational filter.

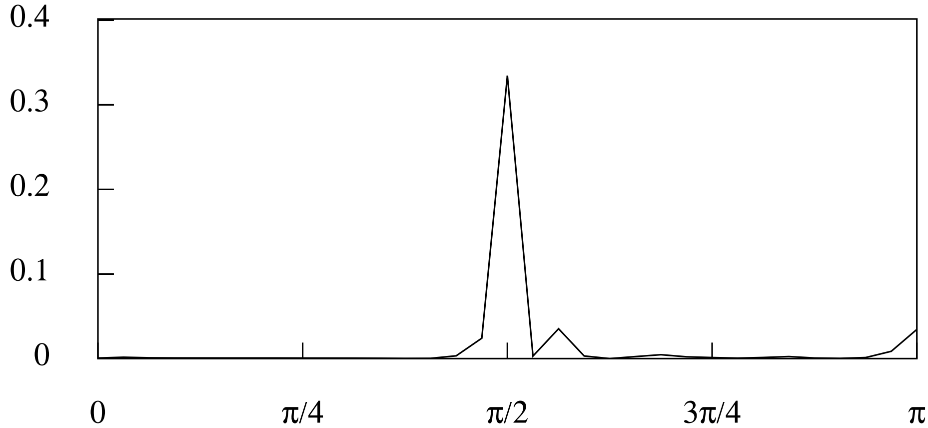

Figure 12. The periodogram of the residual sequence obtained by detrending the Swiss unemployment figures using the rational filter.

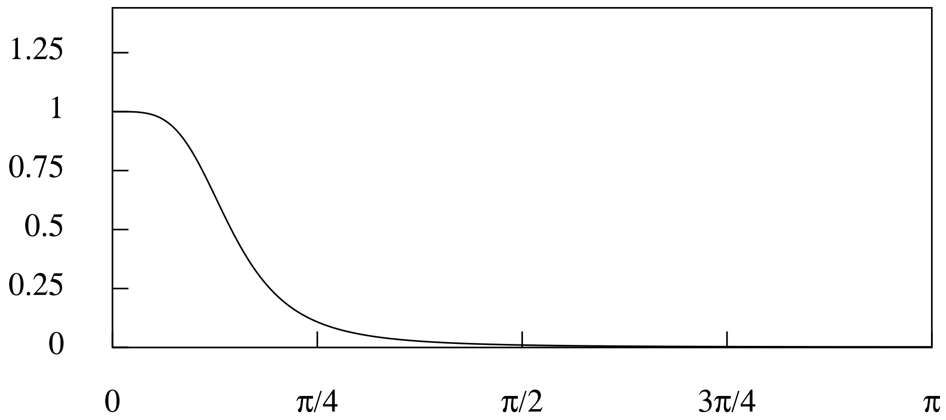

Figure 13. The gain of the Hodrick–Prescott low-pass filter with a smoothing parameter of 24.

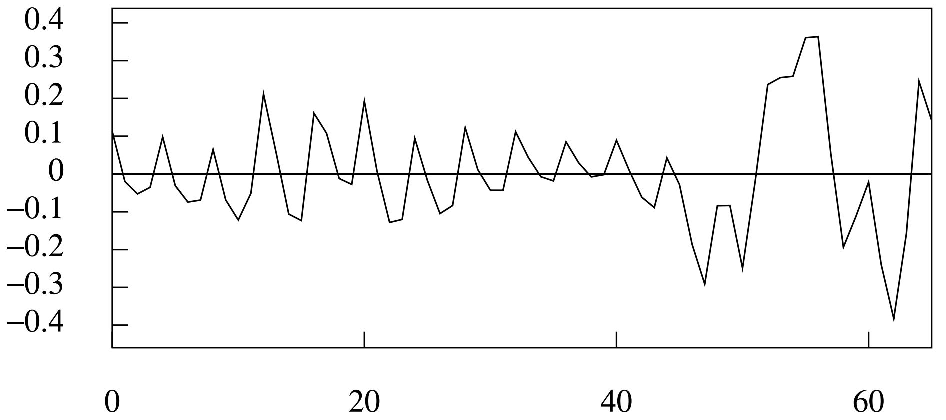

Figure 14. The residual sequence from detrending the Swiss unemployment figures using the Hodrick–Prescott filter.

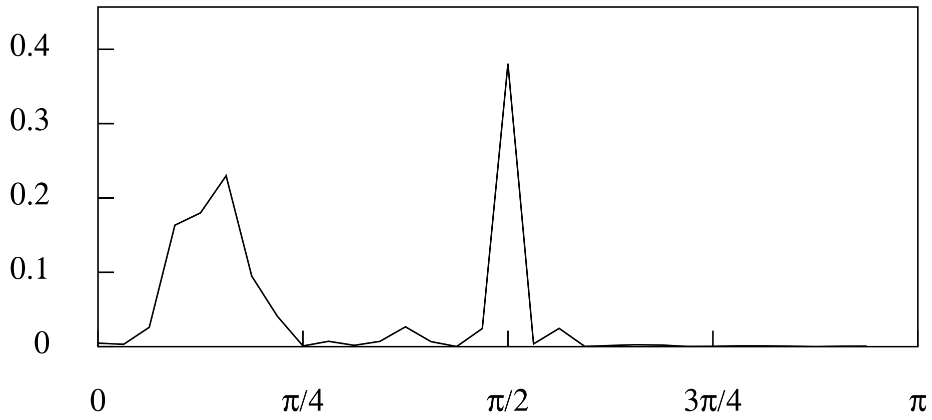

Figure 15. The periodogram of the residual sequence obtained by detrending the Swiss unemployment figures using the Hodrick–Prescott filter.

regularity. The amplitudes of the seasonal fluctuations are clearly related to the level of unemployment. Thus, in times of high employment, there appears to be a widespread hoarding of labour which would be subject, at other times, to seasonal unemployment. This is a feature which one would not have detected by inspecting the original data series. It also transpires, from Figure 11, that, far from being in abeyance during the period of rapidly increasing unemployment, the seasonal fluctuations were present and were of a steadily increasing amplitude.

The regularity of the residual series is reflected in its periodogram which is represented in Figure 12. Here, the complete absence of any elements of a frequency below the cut-off point is a powerful testimony to the efficacy of the rational filter. The tall spike centred at 90 degrees, or $\pi/2$ radians, represents the power of the seasonal fluctuations.

The effects of a parallel analysis of the unemployment figures which has used the Hodrick–Prescott filter are represented in Figures 13 to 15. The filter fails to remove from the residual sequence some of the motions which ought to be attributed to the trend. The consequences is that the regularity of the seasonal effect in not apparent in the residual sequence and the false impression is strengthened that the effect is largely in abeyance during the period of the rapid increase in unemployment. The fault of the filter is evident in Figure 15 which shows that it has allowed some powerful low-frequency elements to pass through into the residual series.

This example emphasises the need to use sharp filters in analysing economic time series.

References

[1] Blackman , R.B., and T.W. Tukey, (1959), *The Measurement of the Power Spectrum from the Point of View of Communications Engineering,* Dover Publications: New York.

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

[3] Cogley, T., and J.M. Nason, (1995), Effects of the Hodrick–Prescott Filter on Trend and Difference Stationary Time Series, Implications for Business Cycle Research, *Journal of Economic Dynamics and Control,* **19,** 253–278.

[4] Gibbs, J.W., (1898), Fourier's Series: A Letter to the Editor, *Nature, December 29 1898,* **58,** 200.

[5] Gibbs, J.W., (1899), Fourier's Series: A Letter to the Editor, *Nature, April 27th 1899,* **59,** 606.

[6] Harvey, A.C., and A. Jaeger, (1993), Detrending, Stylised Facts and the Business Cycle, *Journal of Applied Econometrics,* **8,** 231–247.

[7] Hodrick, R., and E. Prescott, (1980), Post-war U.S. Business Cycles: An Empirical Investigation, Working Paper, Carnegie–Mellon University, Pittsburgh, Pennsylvania,

[8] Kydland, F.E., and C. Prescott, (1990), Business Cycles: Real Facts and a Monetary Myth, *Federal Reserve Bank of Minneapolis Quarterly Review,* **14,** 3–18.

[9] Laurie, D.P., (1980), Efficient Implementation of Wilson's Algorithm for Factorising a Self-Reciprocal Polynomial, *BIT,* **20,** 257–259.

[10] Laurie, D.P., (1982), Cramér-Wold Factorisation, Algorithm AS 175, *Applied Statistics,* **31,** 86–90.

[11] Pollock, D.S.G, (1998), *Time-Series Analysis, Signal Processing and Dynamics,* The Academic Press, London.

[12] Reinsch, C.H., (1976), Smoothing by Spline Functions, *Numerische Mathematik,* **10,** 177–183.

[13] Whittle, P., (1983), *Prediction and Regulation by Linear Least-Square Methods, Second Revised Edition,* Basil Blackwell, Oxford.

[14] Wilson, G.T., (1969), Factorisation of the Covariance Generating Function of a Pure Moving Average Process, *SIAM Journal of Numerical Analysis,* **6,** 1–7.