BAND-LIMITED PROCESSES IN DISCRETE AND CONTINUOUS TIME

by D.S.G. POLLOCK
University of Leicester

In the theory of stochastic differential equations, it is commonly assumed that the forcing function is a Wiener process. Such a process has an infinite bandwidth in the frequency domain. In practice, however, all stochastic processes have a limited bandwidth.

A theory of band-limited linear stochastic processes is described that reflects this reality, and it is shown how the corresponding ARMA models can be estimated. By ignoring the limitation on the frequencies of the forcing function, in the process of fitting a conventional ARMA model, one is liable to derive estimates that are severely biased.

If the maximum frequency in the sampled data is less than the Nyquist value, then the underlying continuous function can be reconstituted by sinc function or Fourier interpolation. The estimation biases can be avoided by sampling the continuous process at a rate corresponding to the maximum frequency of the forcing function. Then, there is a direct correspondence between the parameters of the band-limited ARMA model and those of an equivalent continuous-time process.

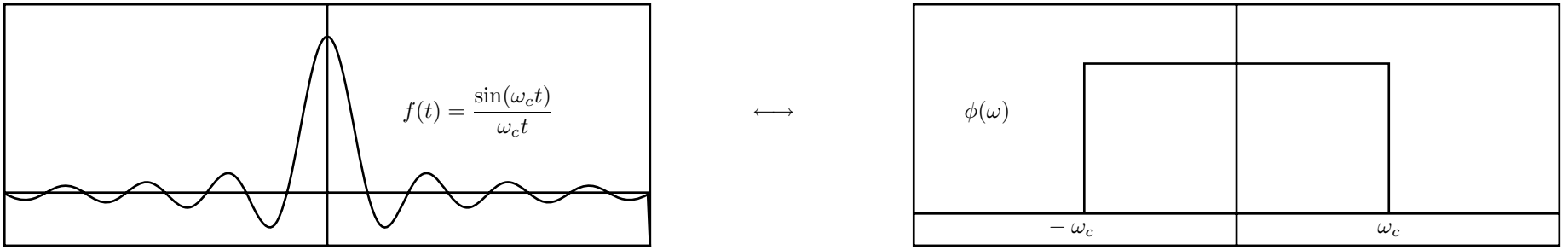
1. Time-Limited versus Band-Limited Processes

Stochastic processes in continuous time are usually modelled by filtered versions of Wiener processes which have infinite bandwidth. This seems inappropriate for modelling the slowly evolving trajectories of macroeconomic data. Therefore, we shall model these as processes that are limited in frequency.

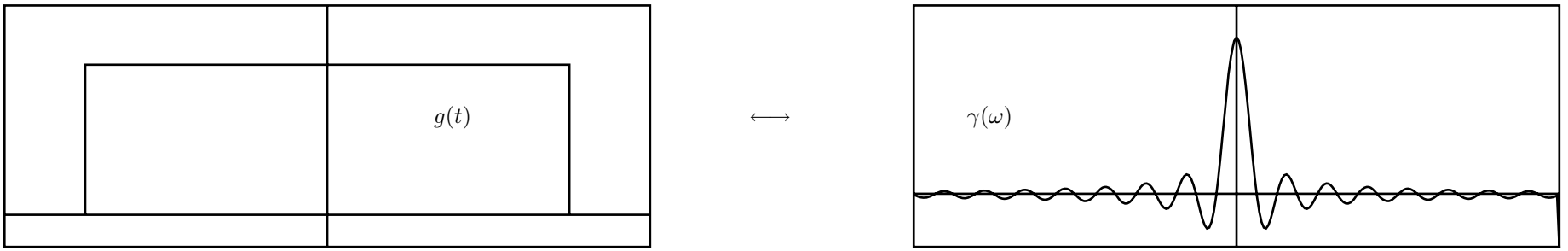
A function cannot be simultaneously limited in frequency and limited in time. One must choose either a band-limited function, which extends infinitely in time, or a time-limited function, which extents over an infinite range of frequencies.

A band-limited function is analytic. It possesses derivatives of all orders. Knowing its derivatives, allows one to find the turning points. It should also be possible to extrapolate the function indefinitely with perfect accuracy.

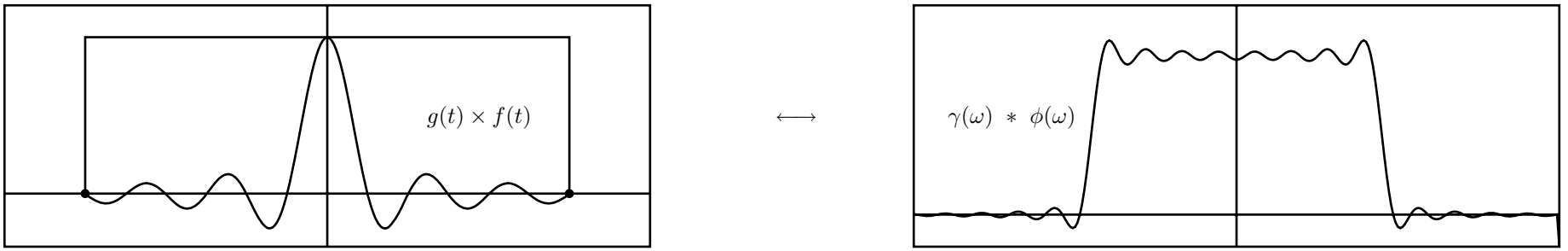
In a statistical context, the perfect predictability of band-limited functions may lead one to question their relevance. Eventually, I shall dispel the awkward implication of perfect predictability.



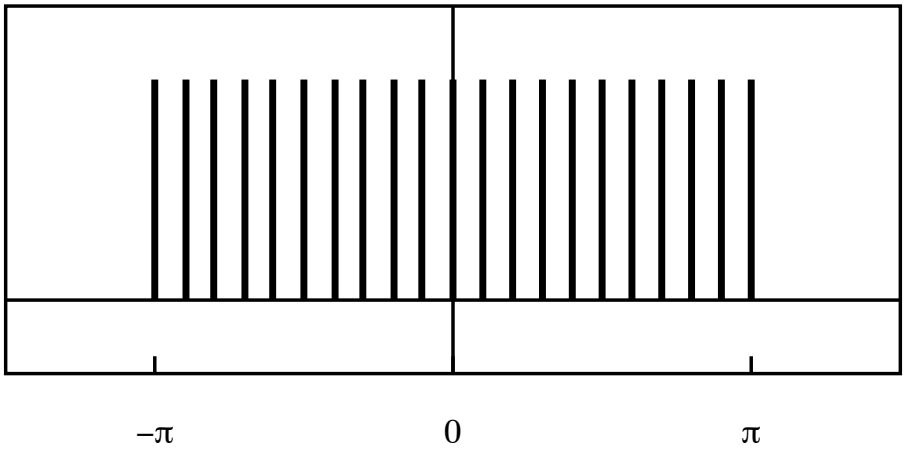
The Fourier Transform of a time-domain sinc function is a rectangle in the frequency domain.



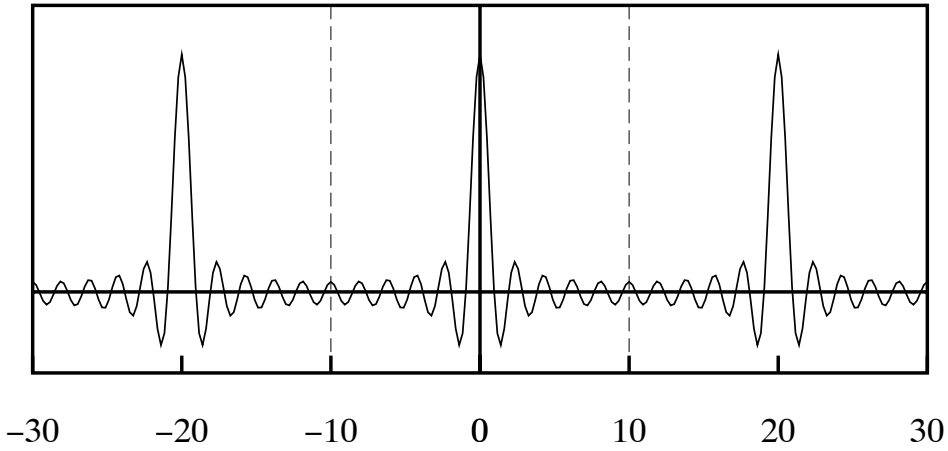
The Fourier Transform of a time-domain rectangle is a sinc function in the frequency domain.



Truncating the time domain sinc function causes leakage in the frequency domain. The time-domain truncation corresponds to a convolution in the frequency domain.



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



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

2. The Definition of a Wiener Process

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

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

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

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

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

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

The increments $dW(s), dW(t)$ are impulses that have a uniform power spectrum distributed over the entire real line.

Sampling $W(t)$ at regular intervals entails a process of aliasing whereby the spectral power of the cumulated increments gives rise to a uniform spectrum of finite power over the interval $[-\pi, \pi]$

An inspection of the periodograms of macroeconomic data sequences reveals that their components tend to reside in strictly limited frequency bands; and it is improbable that these should have arisen from the filtering of Wiener processes.

3. Polynomial Regression and the Difference Operator

Let $I_T = [e_0, e_1, \ldots, e_{T-1}]$ be the identity matrix and let $L_T = [e_1, \ldots, e_{T-1}, 0]$ be the matrix lag operator. Then $\nabla_T^p = (I_T - L_T)^p$ takes p-fold differences of the data.

When $T = 6$, the matrix second-order difference operator is

$$\nabla_6^2 = \left[\begin{array}{cccccc} 1 & 0 & 0 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 & 0 & 0 \\ \hline 1 & -2 & 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 1 & -2 & 1 \end{array}\right] = \begin{bmatrix} Q_*' \\ Q' \end{bmatrix}.$$

Here, Q_*' is liable to be discarded.

If y is the data vector, then the deviations from an interpolated linear trend are given by

$$e = Q(Q'Q)^{-1}Q'y.$$

This vector contains the same information as the vector $Q'y$ of the differences.

The periodogram of the polynomial residuals allows us to discern the spectral structure across the entire frequency range.

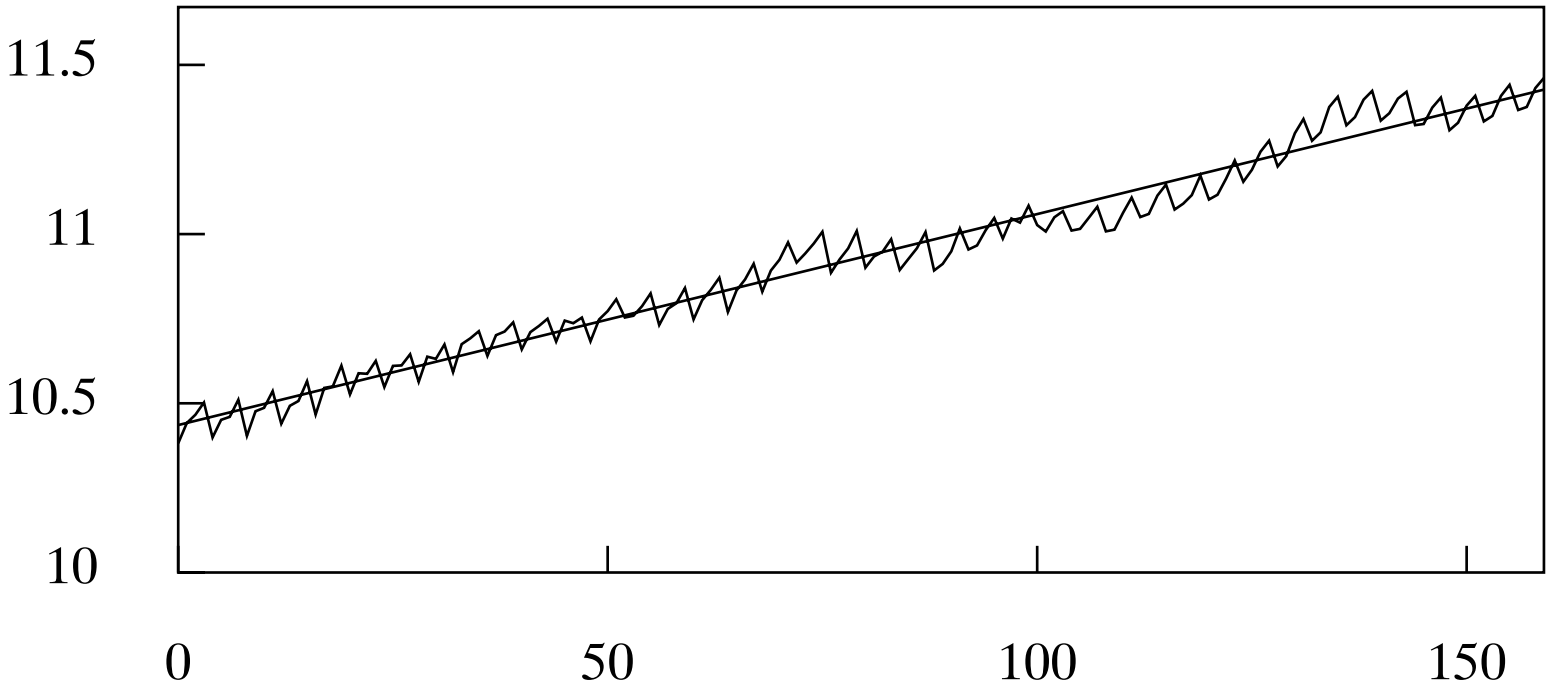


Figure 2. The quarterly series of the logarithms of consumption in the U.K., for the years 1955 to 1994, together with a linear trend interpolated by least squares regression.

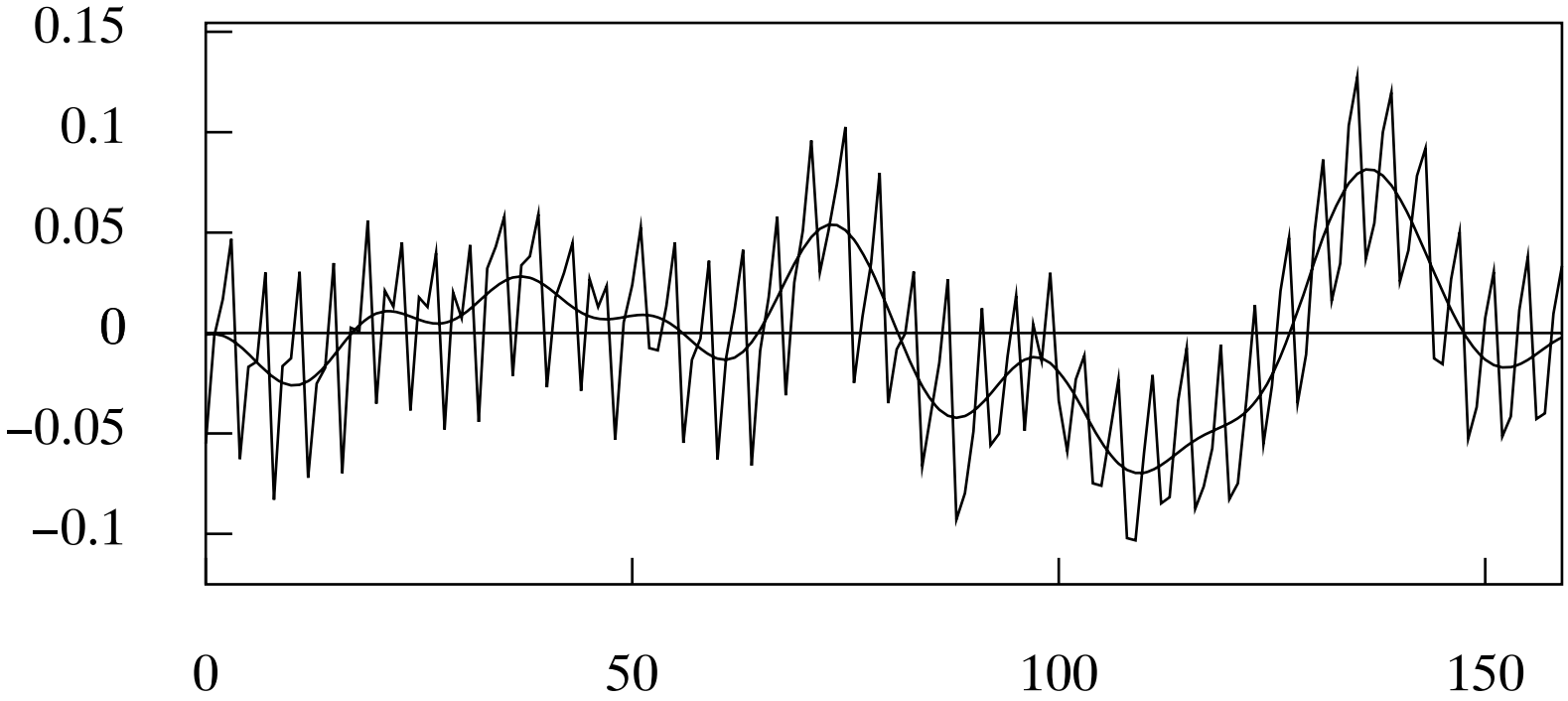


Figure 3. The residual sequence from fitting a linear trend to the logarithmic consumption data. The interpolated line, which represents the business cycle, has been synthesised from the Fourier ordinates in the frequency interval $[0, \pi/8]$.

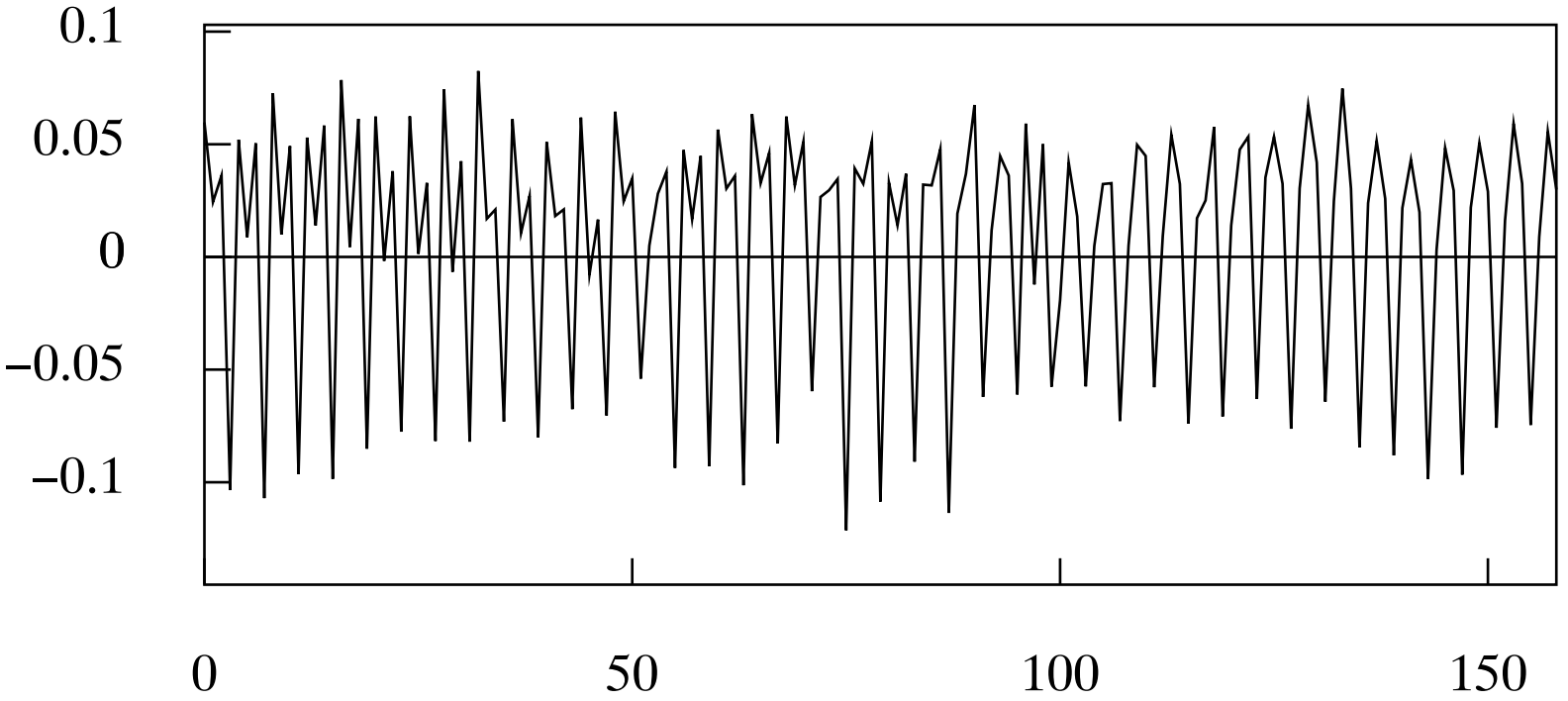


Figure 4. The differences of the logarithmic consumption data.

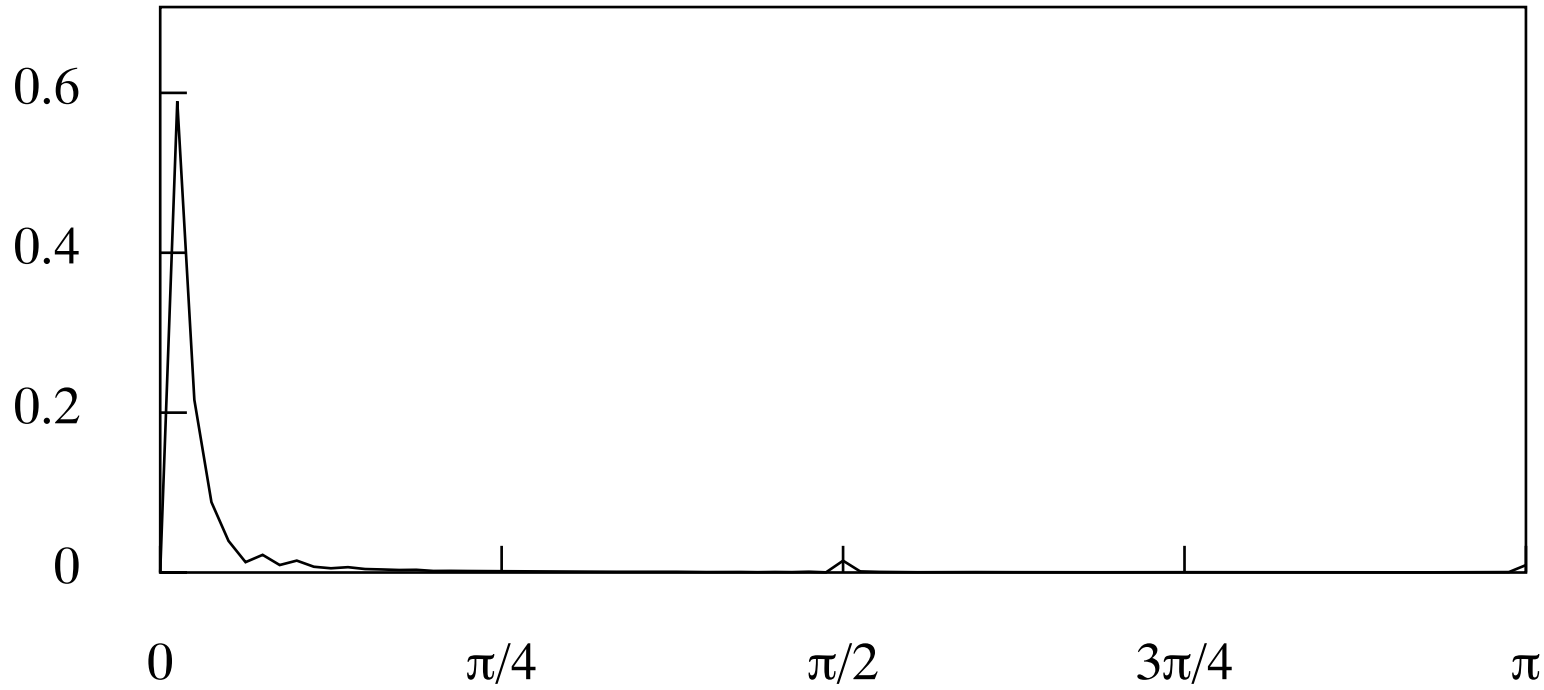


Figure 5. The periodogram of the logarithms of consumption in the U.K., for the years 1955 to 1994.

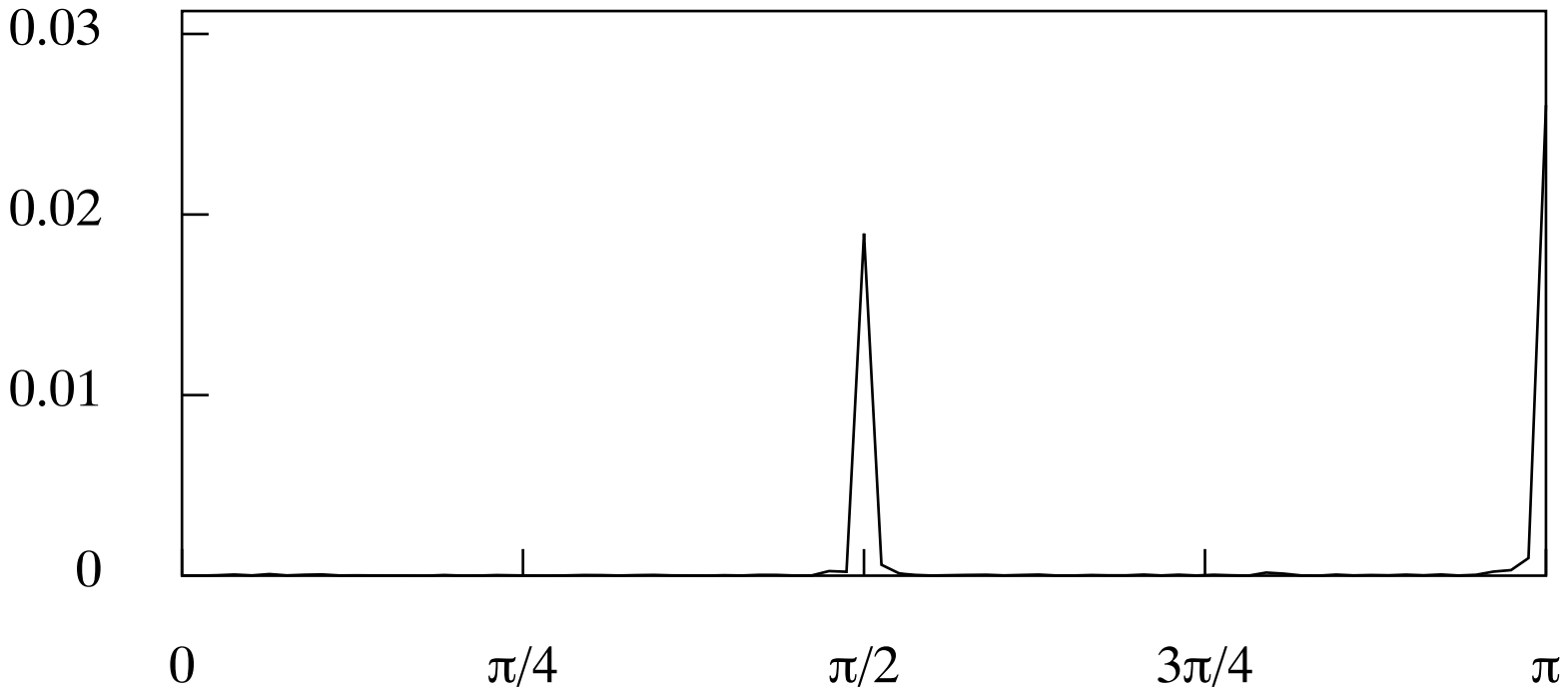


Figure 6. The periodogram of the first differences of the the logarithmic consumption data.

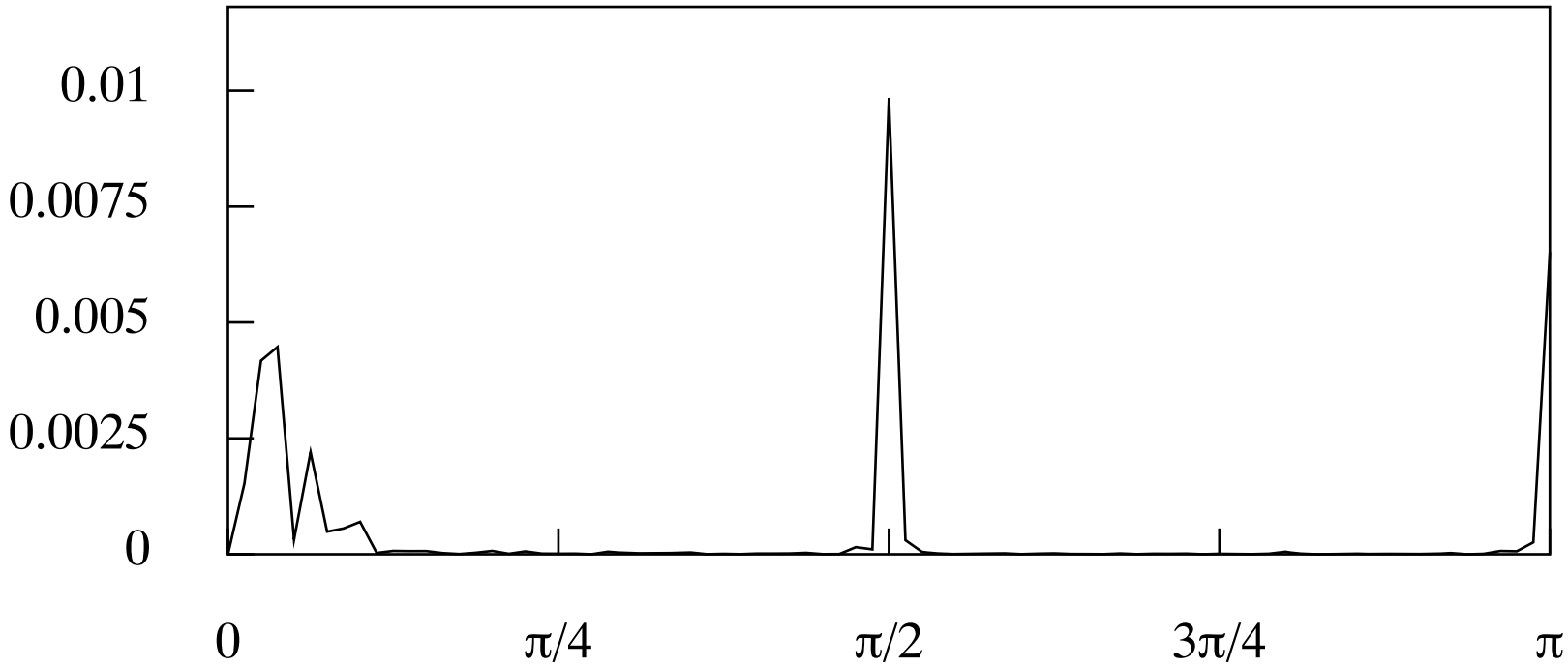


Figure 7. The periodogram of the residual sequence obtained from the linear detrending of the logarithmic consumption data.

4. The Fourier Synthesis of the Business Cycle

The slowly varying continuous function interpolated through the deviations of Figure 3 has been created by combining a set of sine and cosine functions of regularly spaced frequencies extending to the limiting frequency of the business cycle, which is $\omega_d = \pi/8$.

For a sample of size T, the Fourier frequencies are $\omega_j = 2\pi j/T; j = 0, 1, \ldots, [T/2]$, where $[T/2]$ is the integral part of $T/2$. Thus,

$$\begin{aligned} z(t) &= \sum_{j=0}^{d} \{\alpha_j \cos(\omega_j t) + \beta_j \sin(\omega_j t)\} \\ &= \sum_{j=-d}^{d} \xi_j e^{\mathrm{i}\omega_j t} \quad \text{with} \quad \omega_d = 2\pi d/T, \end{aligned}$$

where $\xi_j = (\alpha_j - \mathrm{i}\beta_j)/2$ and $\xi_{-j} = \xi_j^* = (\alpha_i + \mathrm{i}\beta_j)/2$, and where we have used

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

With $\omega_d = 2\pi d/T = \pi/8$ and $T = 160$, there is $d = 10$.

5. Sampling and the Shannon–Nyquist Theorem

A sequence sampled from a square-integrable continuous aperiodic function will have a periodic transform, with a period of 2π radians. Consider

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

For a sample element of $\{x_t; t = 0, \pm 1, \pm 2, \ldots\}$, there is

$$x_t = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{\mathrm{i}\omega t}\xi_S(\omega)d\omega \longleftrightarrow \xi_S(\omega) = \sum_{k=-\infty}^{\infty} x_k e^{-ik\omega}.$$

Therefore, at $x_t = x(t)$, there is

$$x_t = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{\mathrm{i}\omega t}\xi(\omega)d\omega = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{\mathrm{i}\omega t}\xi_S(\omega)d\omega.$$

The equality of the two integrals implies that

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

If $\xi(\omega)$ is not band limited to $[-\pi, \pi]$, then $\xi_S(\omega) \neq \xi(\omega)$ and aliasing will occur.

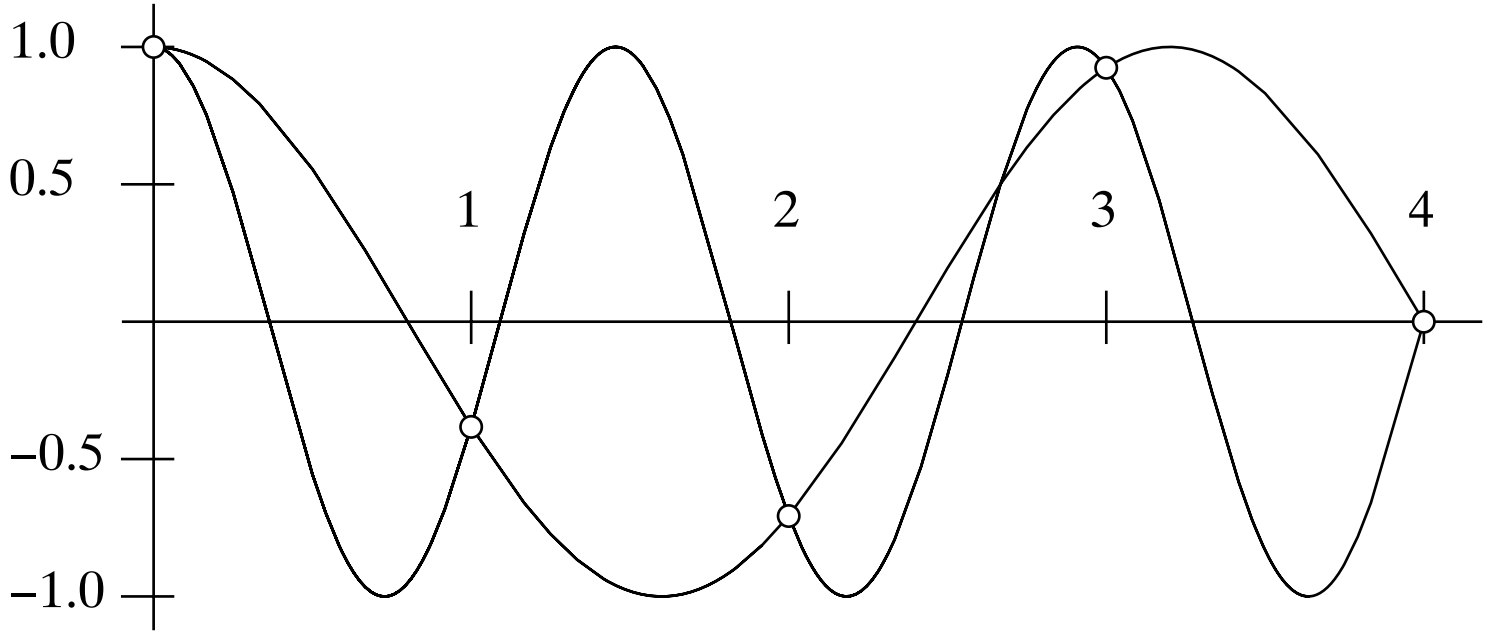


Figure 8. The values of the function $\cos\{(11/8)\pi t\}$ coincide with those of its alias $\cos\{(5/8)\pi t\}$ at the integer points $\{t = 0, \pm 1, \pm 2, \ldots\}$.

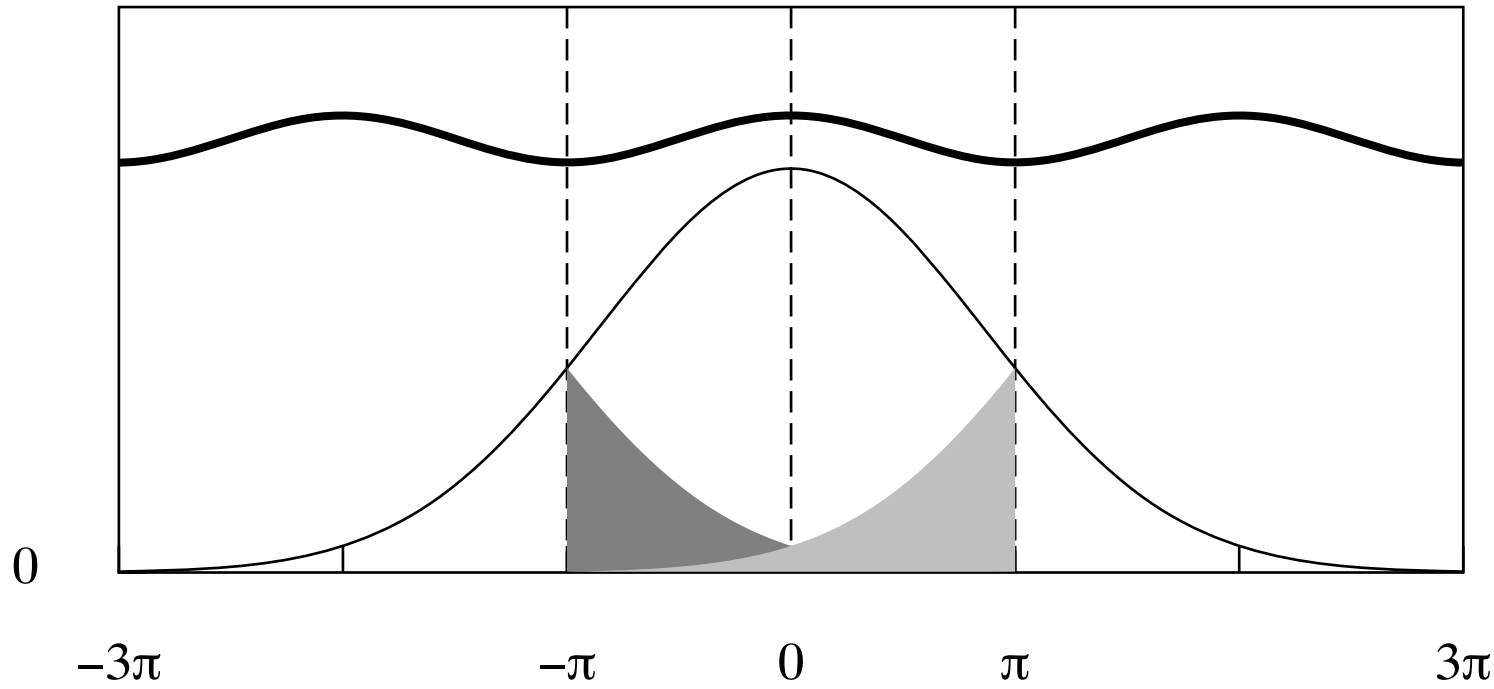


Figure 9. The figure illustrates the aliasing effect of regular sampling. The bell-shaped function supported on the interval $[-3\pi, 3\pi]$ is the spectrum of a continous-time process. The spectrum of the sampled process, represented by the heavy line, is a periodic function of period 2π.

The effect of sampling is to wrap the spectrum around a circle of radius 2π and to add the overlying parts. The same effect is obtained by folding the branches of the function, supported on $[-3\pi, -\pi]$ and $[\pi, 3\pi]$, onto the interval $[-\pi, \pi]$.

6. The Problem of Aliasing

Consider the case of a pure cosine wave of unit amplitude and zero phase whose frequency ω lies in the interval $\pi < \omega < 2\pi$. Let $\omega^* = 2\pi - \omega$. Then

$$\begin{aligned}\cos(\omega t) &= \cos\left\{(2\pi - \omega^*)t\right\} \\ &= \cos(2\pi)\cos(\omega^* t) + \sin(2\pi)\sin(\omega^* t) = \cos(\omega^* t);\end{aligned}$$

which indicates that ω and ω^* are observationally indistinguishable. Here, $\omega^* \in [0, \pi]$ is the alias of $\omega > \pi$.

Imagine that, on his daily walk, a person observes the sea level at 6am and 6pm. He should notice a very gradual recession and advance of the water level; the frequency of the cycle being $f = 1/28$, which amounts to one tide in 14 days. Therefore, he might fail to recognise the daily cycle of the tides.

In fact, the true frequency is $f = 1 - 1/28$, which gives 27 tides in 14 days. Observing the sea level every six hours should enable him to infer the correct frequency.

The spectrum of a sampled process is obtained by wrapping the spectrum of the underlying continuous-time process around a circle of a circumference twice the value of the Nyquist frequency (which is π radians per sample period) and adding the overlying ordinates.

7. Sampling and Sinc-Function Interpolation

If $\xi(\omega) = \xi_S(\omega)$ is a continuous function band-limited to the interval $[-\pi, \pi]$, then it may be regarded as a periodic function of a period of 2π. Putting

$$\xi_S(\omega) = \sum_{k=-\infty}^{\infty} x_k e^{-\mathrm{i}k\omega} \quad \text{into} \quad x(t) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{\mathrm{i}\omega t}\xi_S(\omega)d\omega$$

gives

$$x(t) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left\{\sum_{k=-\infty}^{\infty} x_k e^{-\mathrm{i}k\omega}\right\}e^{\mathrm{i}\omega t}d\omega = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty} x_k\left\{\int_{-\pi}^{\pi} e^{\mathrm{i}\omega(t-k)}\right\}d\omega.$$

The integral on the RHS is evaluated as

$$\int_{-\pi}^{\pi} e^{i\omega(t-k)}d\omega = 2\frac{\sin\{\pi(t-k)\}}{t-k}.$$

Putting this into the RHS gives a weighted sum of sinc functions:

$$x(t) = \sum_{k=-\infty}^{\infty} x_k \frac{\sin\{\pi(t-k)\}}{\pi(t-k)} = \sum_{k=-\infty}^{\infty} x_k\phi_0(t-k), \quad \text{where} \quad \phi_0(t-k) = \frac{\sin\{\pi(t-k)\}}{\pi(t-k)}.$$

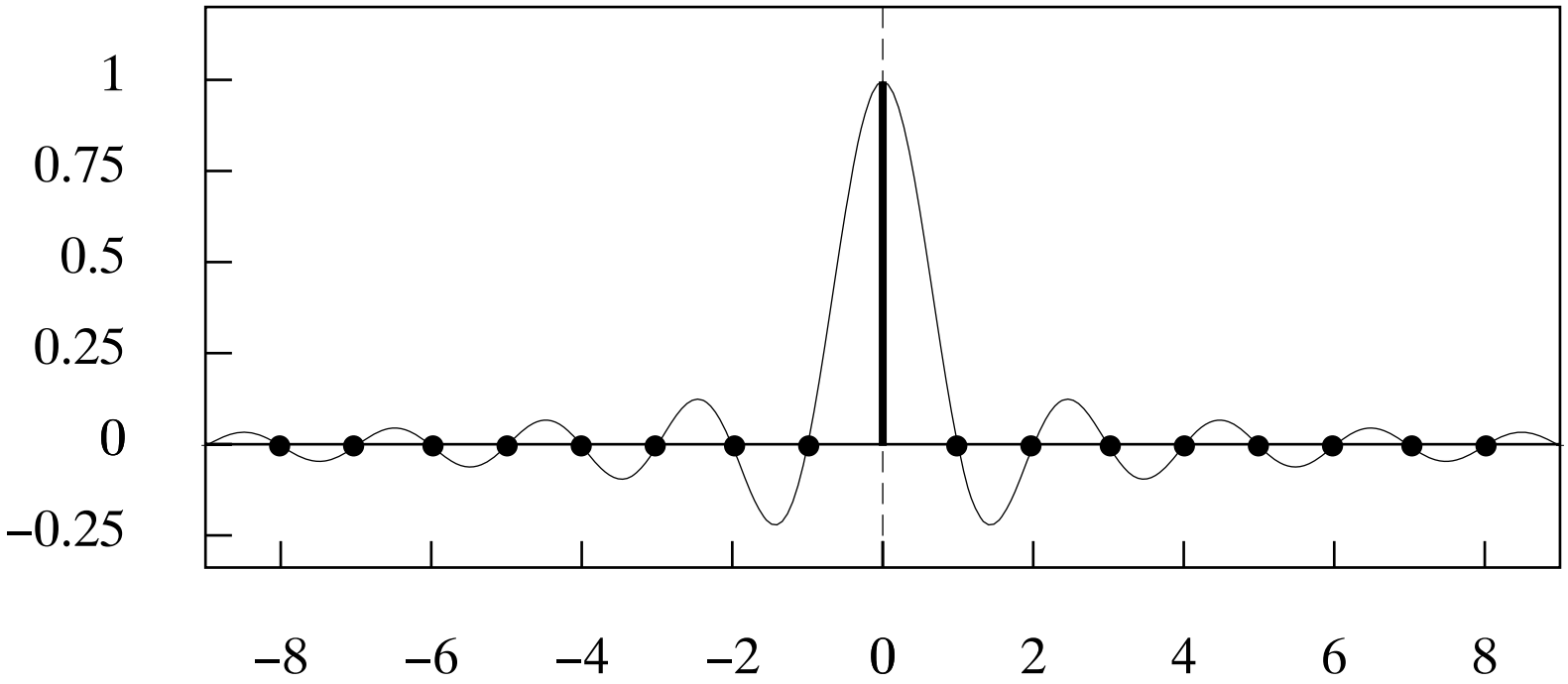


Figure 10. The sinc function wave-packet $\phi_0(t) = \sin(\pi t)/\pi t$ comprising frequencies in the interval $[0, \pi]$.

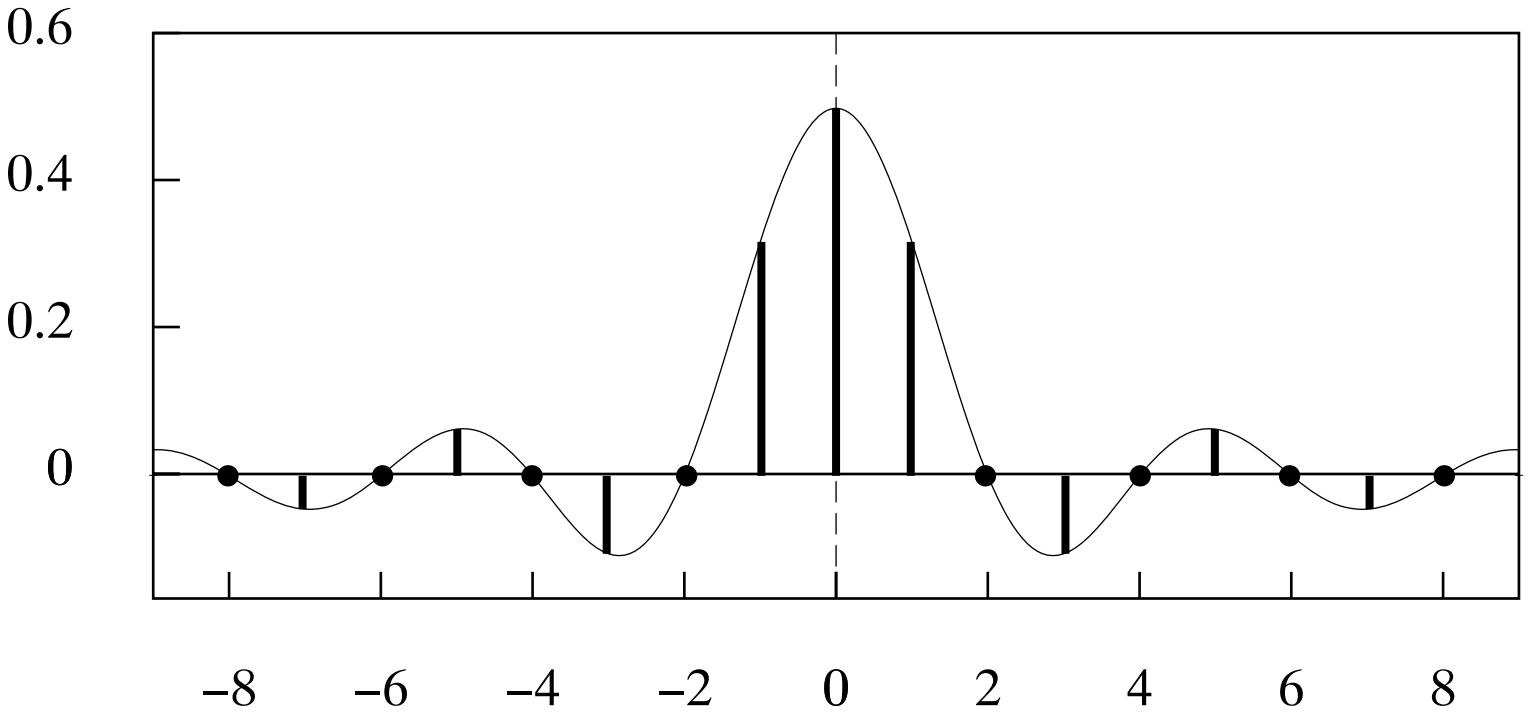


Figure 11. The sinc function wave-packet $\phi_1(t) = \sin(\pi t/2)/\pi t$ comprising frequencies in the interval $[0, \pi/2]$.

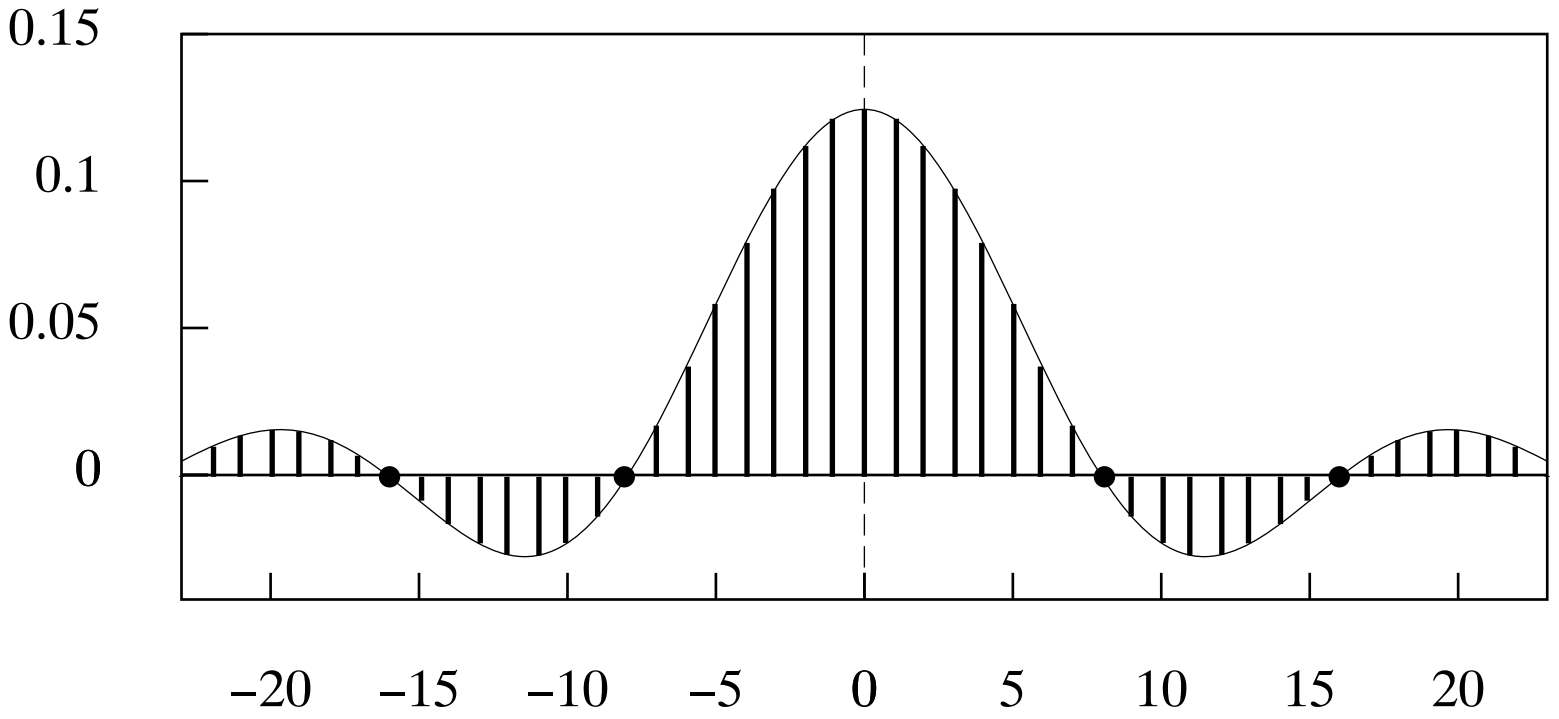


Figure 12. The sinc function wave-packet $\phi_3(t) = \sin(\pi t/8)/\pi t$ comprising frequencies in the interval $[0, \pi/8]$.

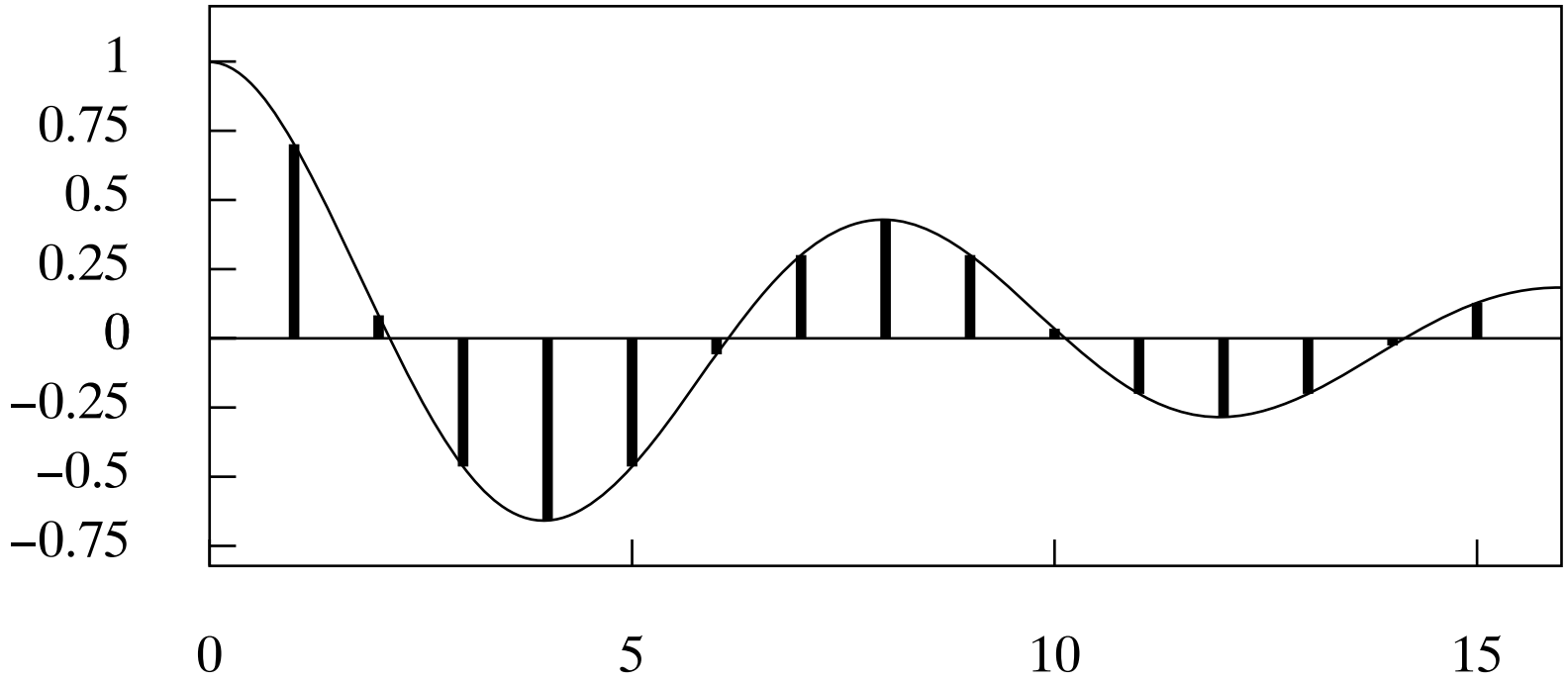


Figure. A continuous autocovariance function and its sampled ordinates

8. The Dirichlet Kernel

A time-domain function on a finite support can be regarded as a single cycle of a periodic function. Let ξ_j^S be the jth ordinate from the discrete Fourier transform of $T = 2n$ points sampled from the function. If the function is band-limited to π radians, then there is

$$x(t) = \sum_{j=0}^{T-1} \xi_j^S e^{\mathrm{i}\omega_j t} \longleftrightarrow \xi_j^S = \frac{1}{T} \sum_{t=0}^{T-1} x_t e^{-\mathrm{i}\omega_j t}, \qquad \omega_j = \frac{2\pi j}{T}.$$

Putting the expression for the Fourier ordinates into the series expansion of the time-domain function and commuting the summation signs gives

$$x(t) = \sum_{j=0}^{T-1} \left\{ \frac{1}{T} \sum_{k=0}^{T-1} x_k e^{\mathrm{i}\omega_j k} \right\} e^{\mathrm{i}\omega_j t} = \frac{1}{T} \sum_{k=0}^{T-1} x_k \left\{ \sum_{j=0}^{T-1} e^{\mathrm{i}\omega_j (t-k)} \right\}.$$

The inner summation gives rise to the Dirichlet kernel:

$$\phi_n^\circ(t) = \sum_{t=0}^{T-1} e^{\mathrm{i}\omega_j t} = \frac{\sin([n-1/2]\omega_1 t)}{\sin(\omega_1 t/2)}, \qquad \omega_1 = \frac{2\pi}{T} = \frac{\pi}{n}.$$

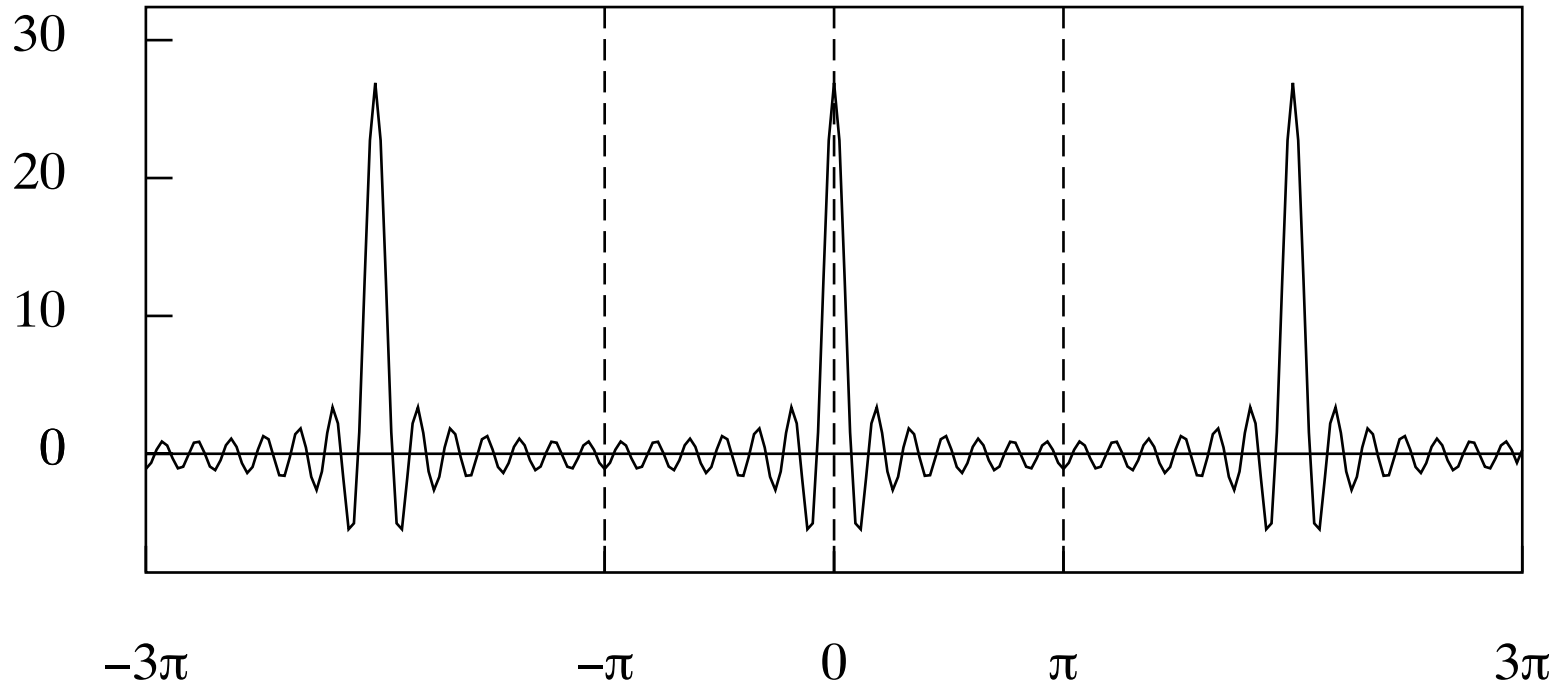


Figure 13. Three Cycles of a Dirichlet Function $\delta(t) = \sin(n\omega t)/\sin(\omega t)$.

9. Band-limited Wave Packets

We can model the process underlying the sampled data as a sequence of band-limited wave packets of random amplitudes, which may also have varying profiles. These may arrive at random intervals in the manner of a Poisson process.

Since the wave packets are band limited, it will always be possible to resolve their sum into a sequence of sinc functions at regular displacements along the time axis, having amplitudes that can be described by a discrete stochastic process.

A white-noise process can be derived by sampling a process composed of a stream of sinc function wave packets of independent and identically distributed amplitudes, arriving at regular intervals.

The generic sinc function is

$$\delta_n(t-k) = \frac{\sin\{n(t-k)\}}{\pi(t-k)},$$

where $\Delta = 1/n$, is the interval between successive functions. In the limit, as $\Delta \to 0$, this becomes Dirac's delta function $\delta(t-k)$, also described as the unit impulse.

In the limit, the stochastic sequence of sinc functions scaled by $1\sqrt{n}$ becomes a Wiener process.

10. ARMA Modelling of Band-Limited Processes

A value sampled at an arbitrary point from a band-limited process will be expressed as a sum of values sampled from an infinite number of wave packets.

The sum will include values sampled from the tails of receding wave packets as well as values that represent the onset of the waves that lie ahead. A discrete ARMA model that looks in one direction only, appears, at first sight, to be incapable of modelling such a process.

This problem in ARMA modelling, which arises from the interference of the waves, in common with the problem of aliasing, can be overcome by sampling the data at the critical rate that corresponds to the highest frequency present in the underlying process.

In that case, there will be no interference, either from the waves that lie ahead of the current sample point or from the waves that have passed by, and the relationship of the current value to the previous sampled values can be described by a discrete ARMA model.

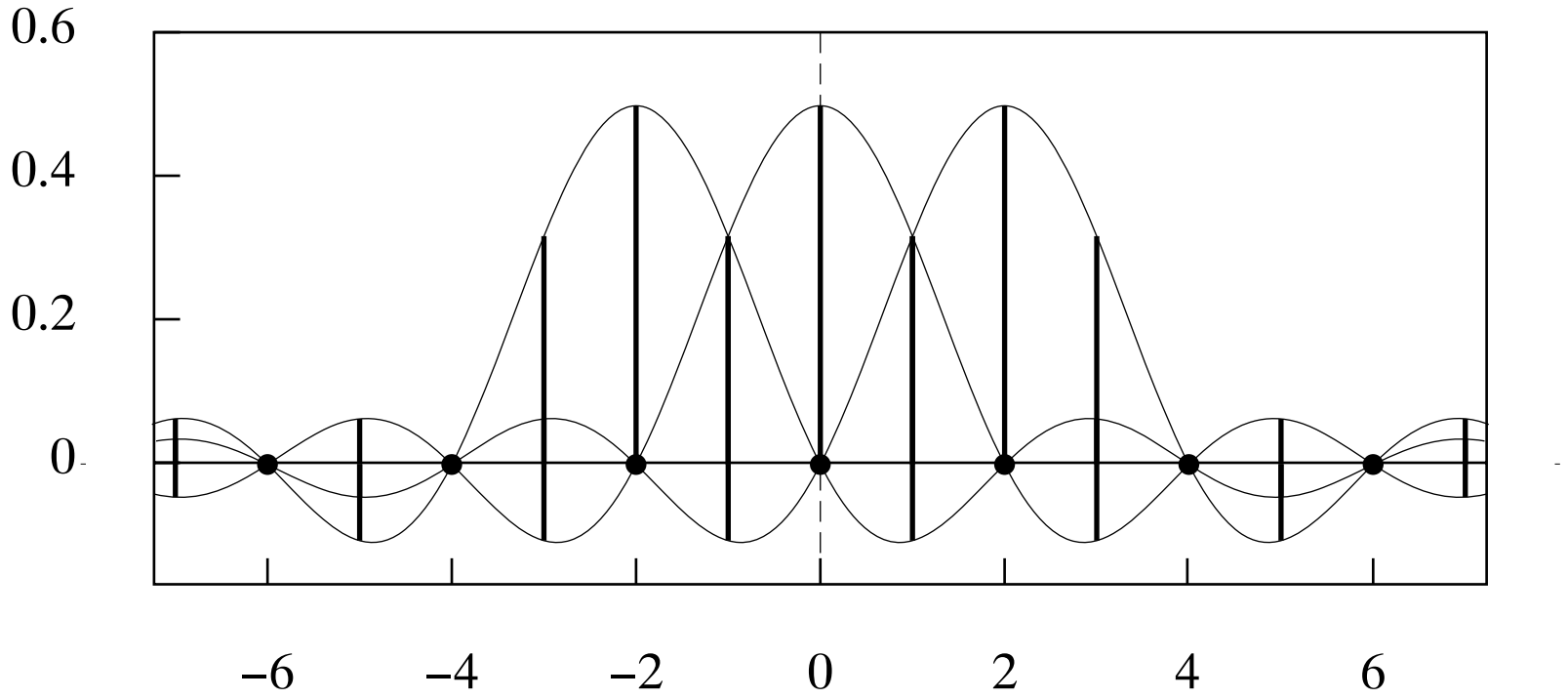


Figure 14. The wave packets $\phi_2(t)$ and $\phi_2(t-k)$ suffer no interference when $k \in \{\pm 2, \pm 4, \pm 6, \ldots\}$.

11. The Seasonal Adjustment of the Data

In order to conduct the initial experiments in autoregressive estimation, we need first to seasonally adjust the data in a manner that is typical of the governmental statistical agencies.

We use a filter that is derived by applying the Wiener–Kolmogorov principle to a model that attributes the seasonal fluctuations to complex autoregressive unit roots, or poles, in which the arguments, i.e. the angles within their polar exponential expressions, correspond to the seasonal frequency and its harmonic frequencies.

In order to confine their effects to the vicinities of the seasonal frequencies, the poles are balanced by zeros, i.e. roots of the moving average operator, with the same arguments and with moduli close to the perimeter of the unit circle.

The effect of the filter is to remove from the data the elements at the seasonal frequencies and to leave the rest largely intact. Because it is synthesised from very few elements, the estimated seasonal component has a regular appearance. The seasonally adjusted data, which contains elements of noise, is rough and irregular by contrast.

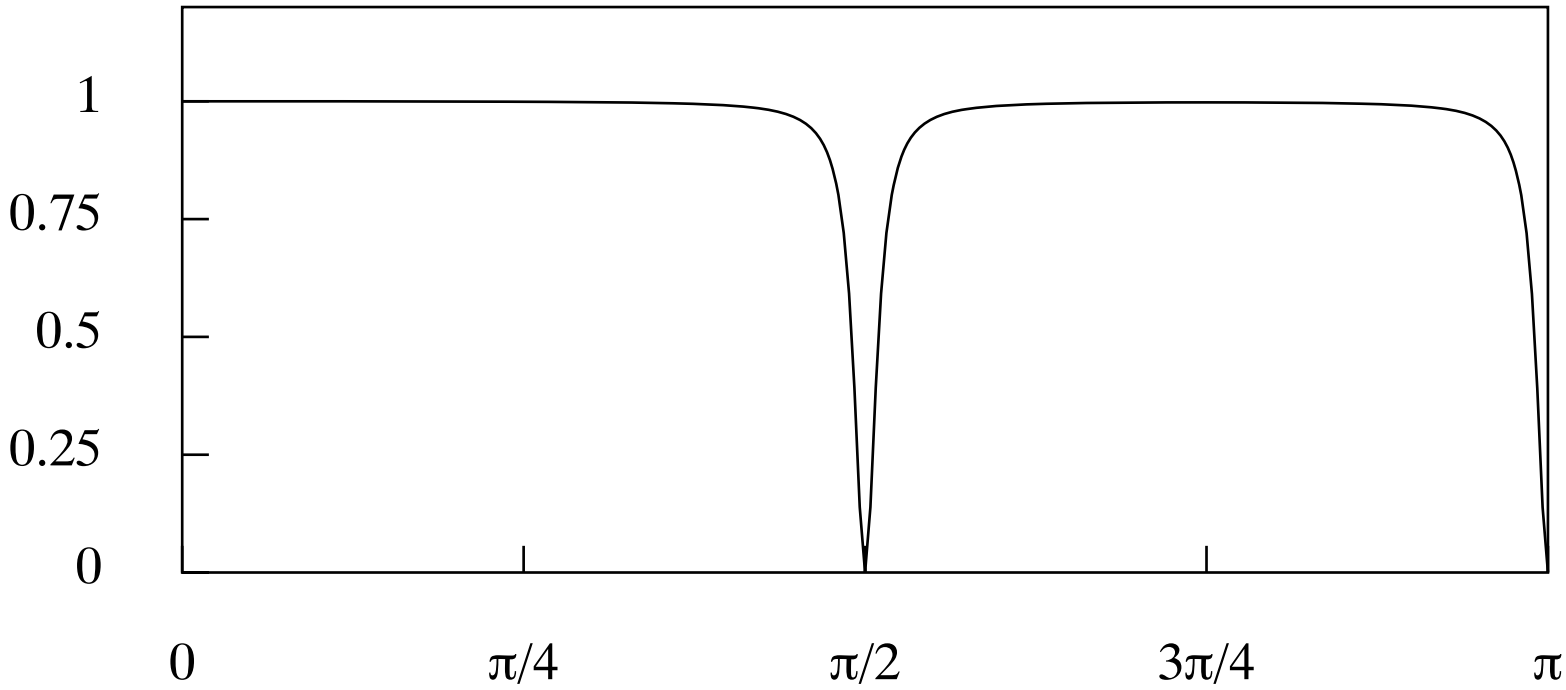


Figure 15. The squared gain of a seasonal adjustment filter to be applied to the quarterly detrended logarithmic consumption data.

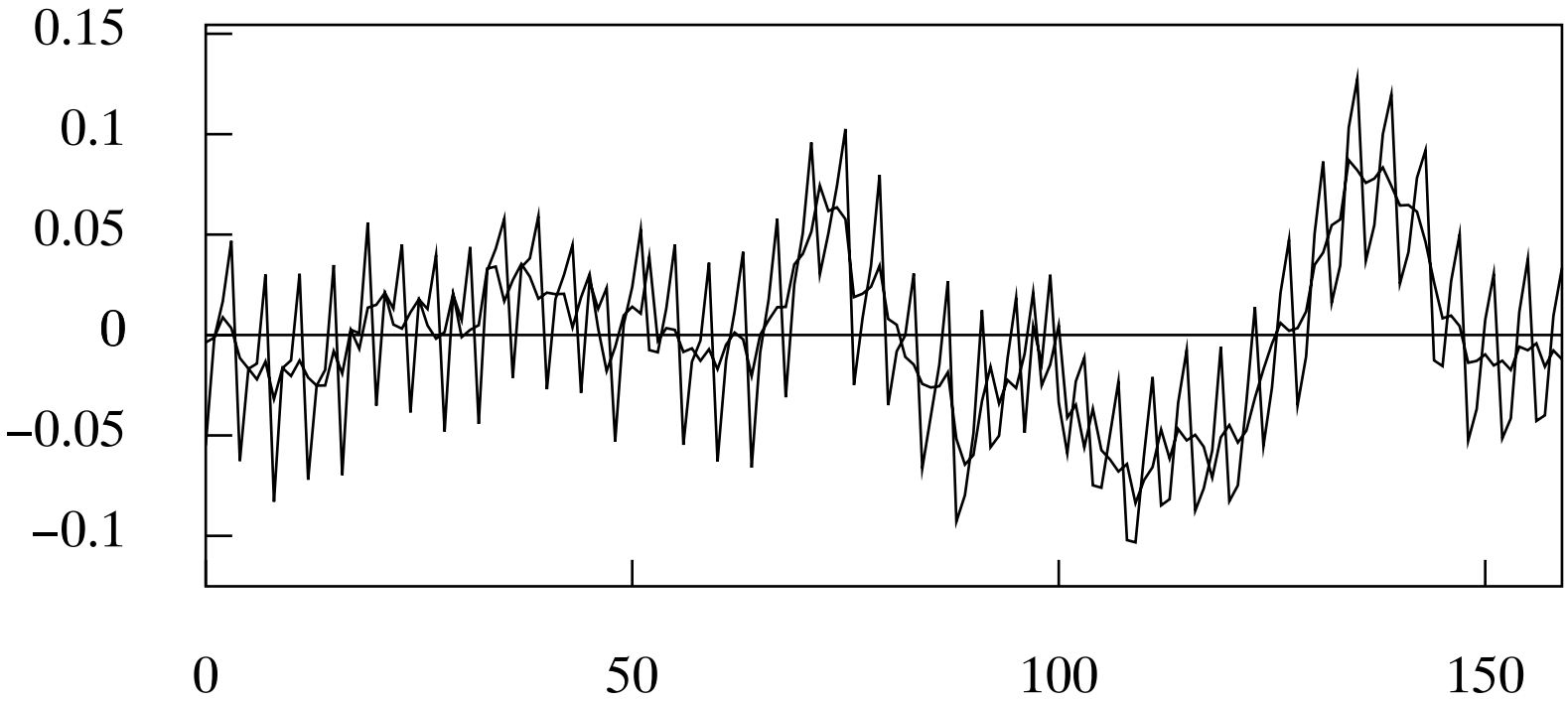


Figure 16a. The plot of 160 quarterly observations on an unidentified data series with a superimposed seasonally-adjusted sequence.

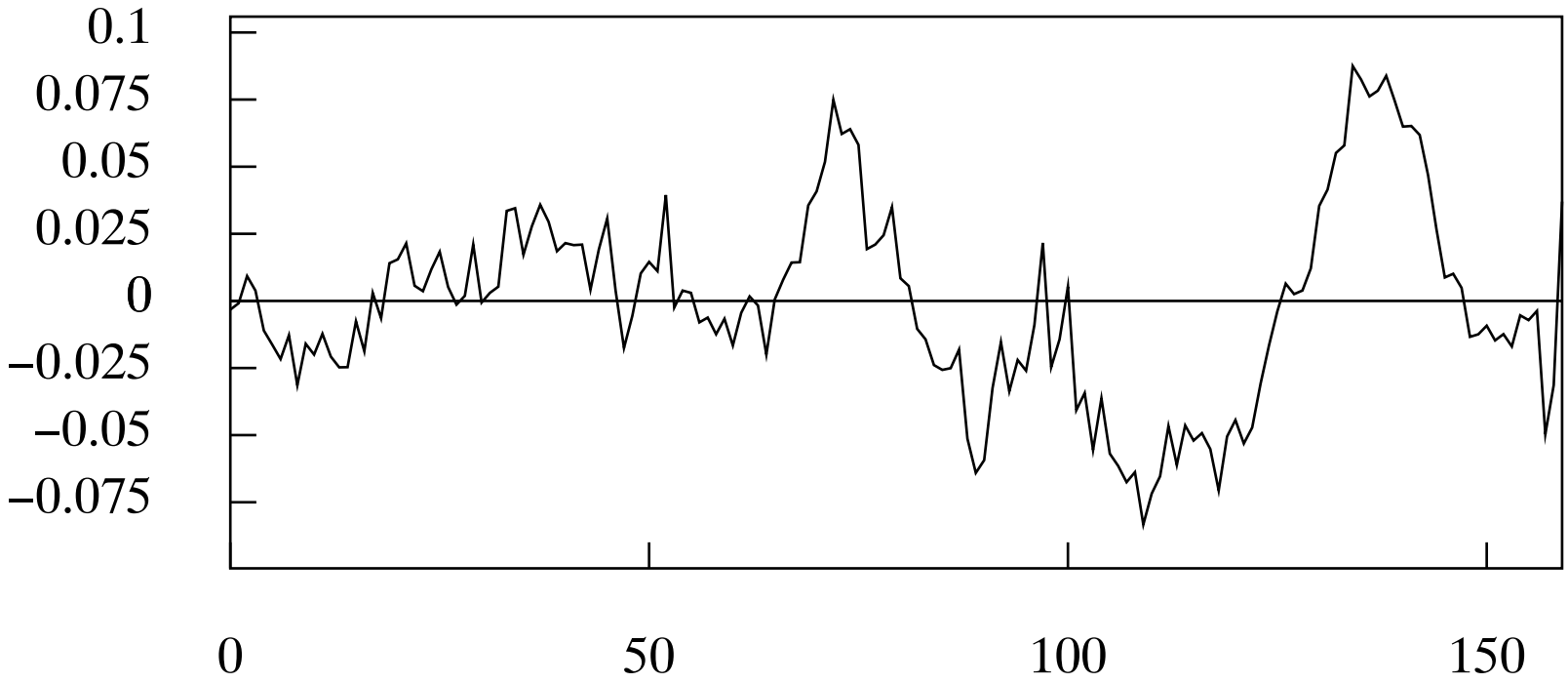


Figure 16. The seasonally-adjusted detrended logarithmic consumption data.

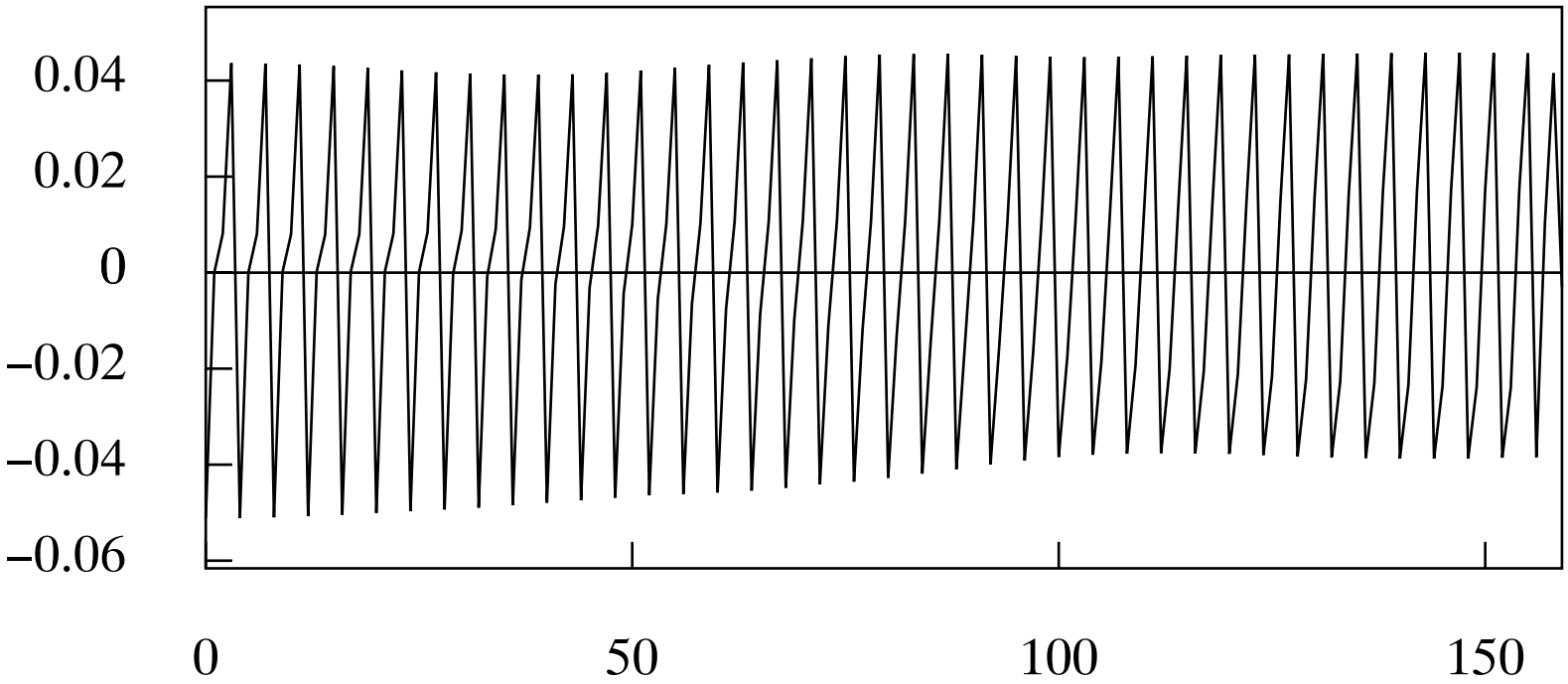


Figure 17. The seasonal component extracted from the detrended logarithmic consumption data in the process of seasonal adjustment.

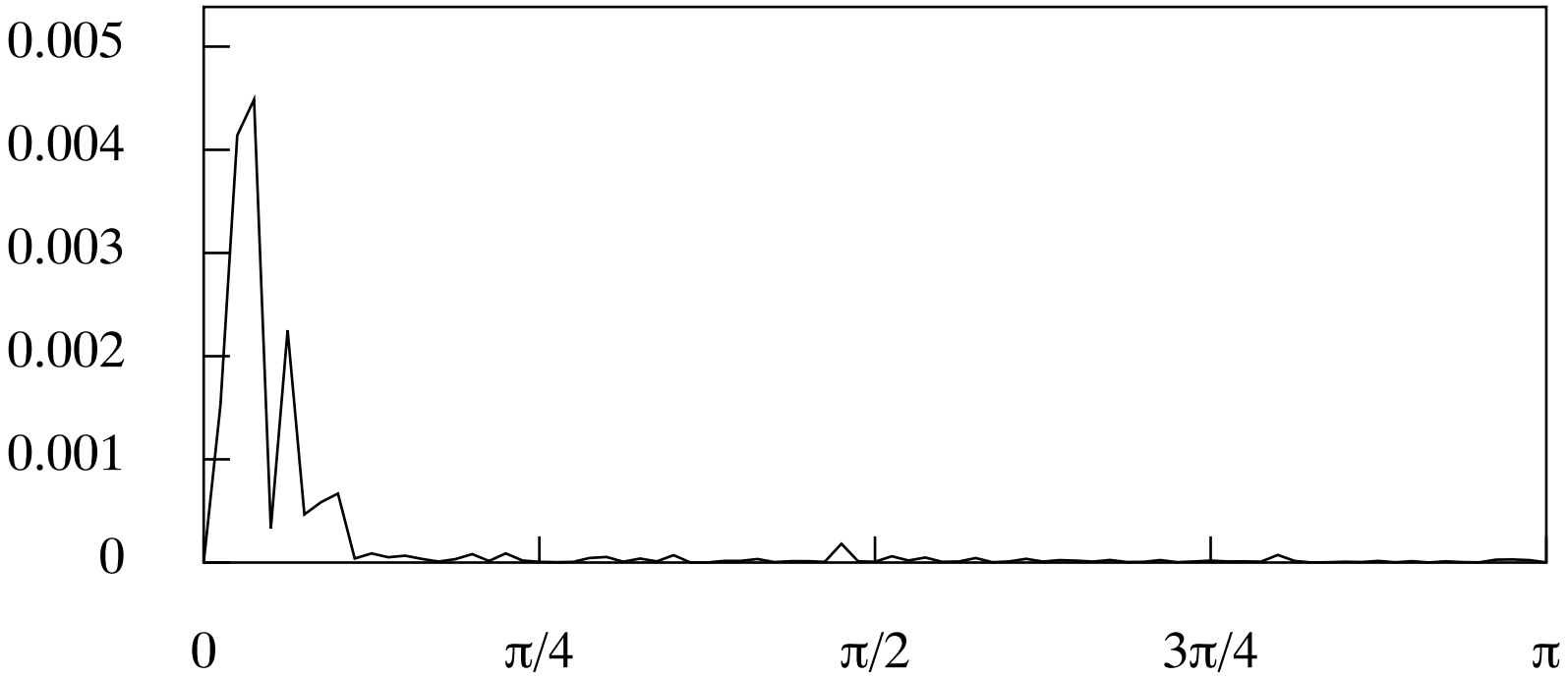


Figure 18. The periodogram of the seasonally-adjusted data.

12. Experiments with Empirical Data

In the first experiment, a second-order autoregressive model is estimated by applying the standard procedures to the seasonally adjusted data. Whereas one would hope to characterise the low-frequency fluctations by a pair of conjugate complex roots, the procedure delivers only real-valued roots.

In the second experiment, a weighted Whittle estimator is applied to the same data. Weights of unity are associated to the spectral ordinates within a low-frequency band and zero weights are associated to the ordinates outside the band. The results are quite the opposite of the previous ones. Now, the estimator delivers complex roots that are very close to the perimeter of the unit circle and which give rise to an excessively prominent spectral spike.

In a third experiment, an ordinary autoregressive estimator is applied to data, representing the low-frequency component, that have been obtained by resampling the continuous-time band-limited component at such a rate that their periodogram is supported on the full Nyquist interval. This procedure delivers the only reasonable estimates.

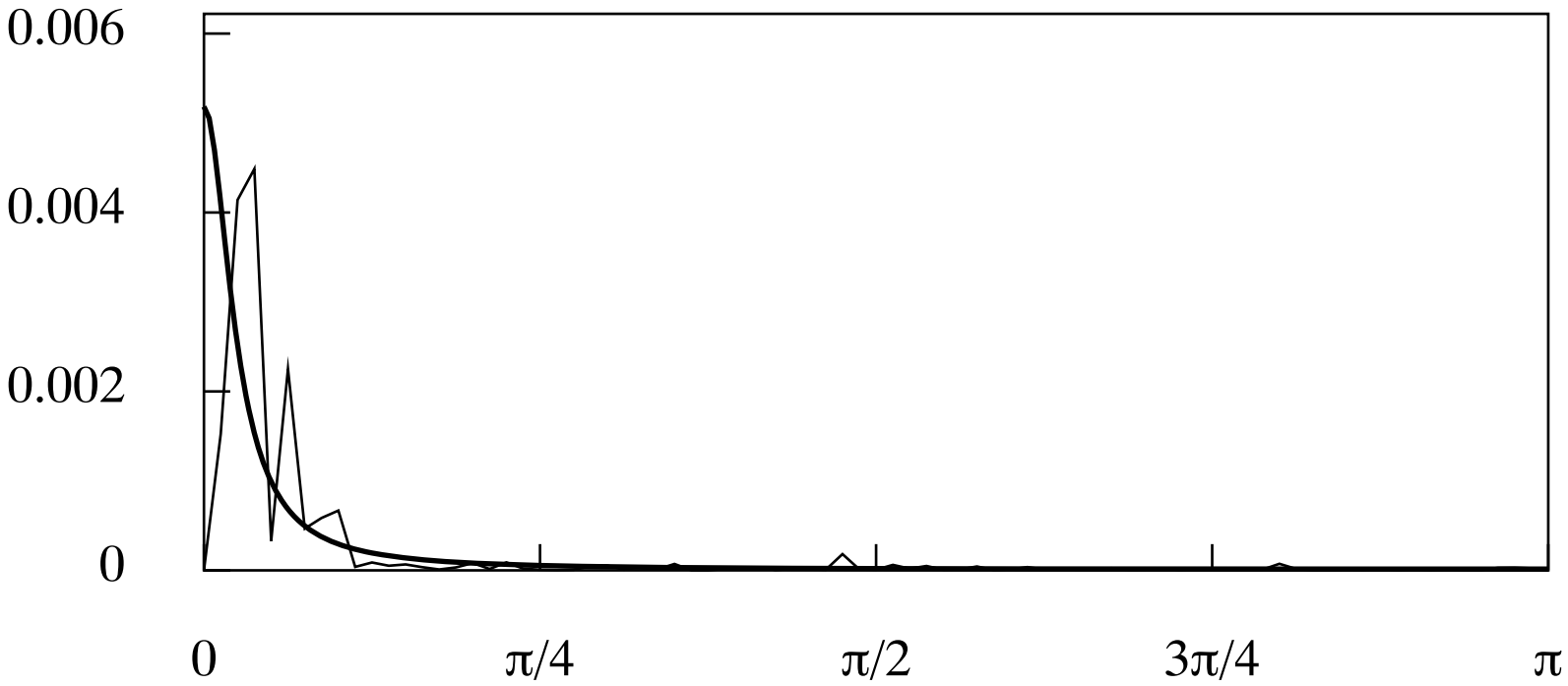


Figure 19. The spectrum of an AR(2) model fitted to the detrended, seasonally-adjusted logarithmic consumption data, superimposed on the periodogram.

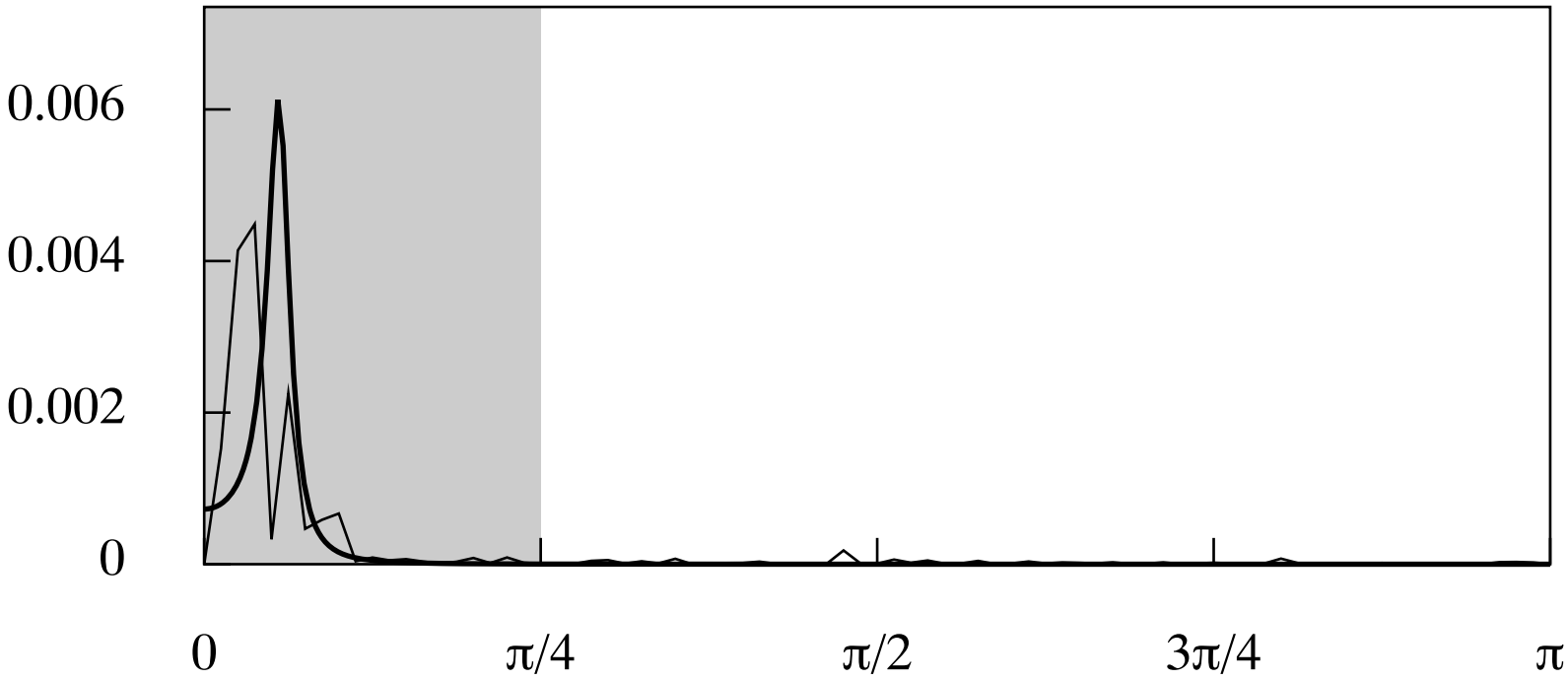


Figure 20. The spectrum of an AR(2) model fitted to the detrended, seasonally-adjusted logarithmic consumption data via a band-limited autoregressive estimation, superimposed on the periodogram.

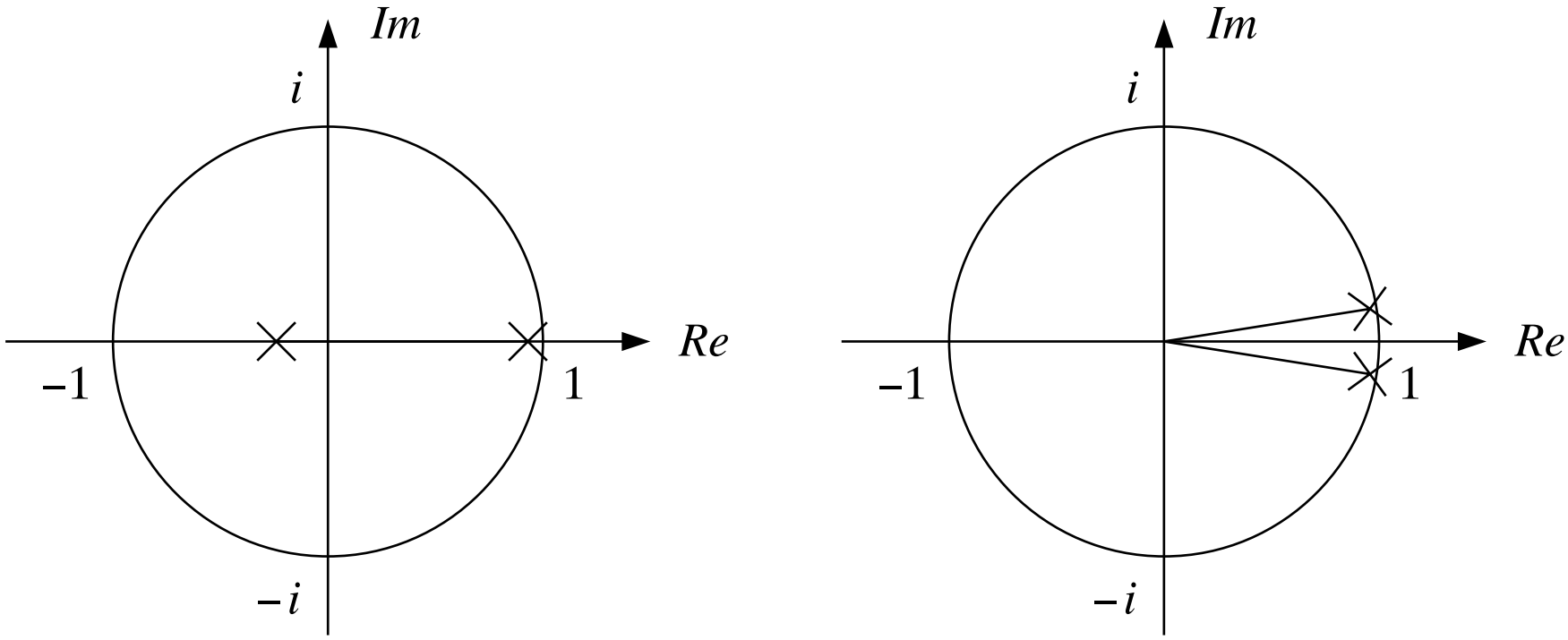


Figure 21. The poles of the AR(2) models fitted to the detrended logarithmic consumption data. (a) is from unrestricted estimator and (b) is from the band-limited estimator.

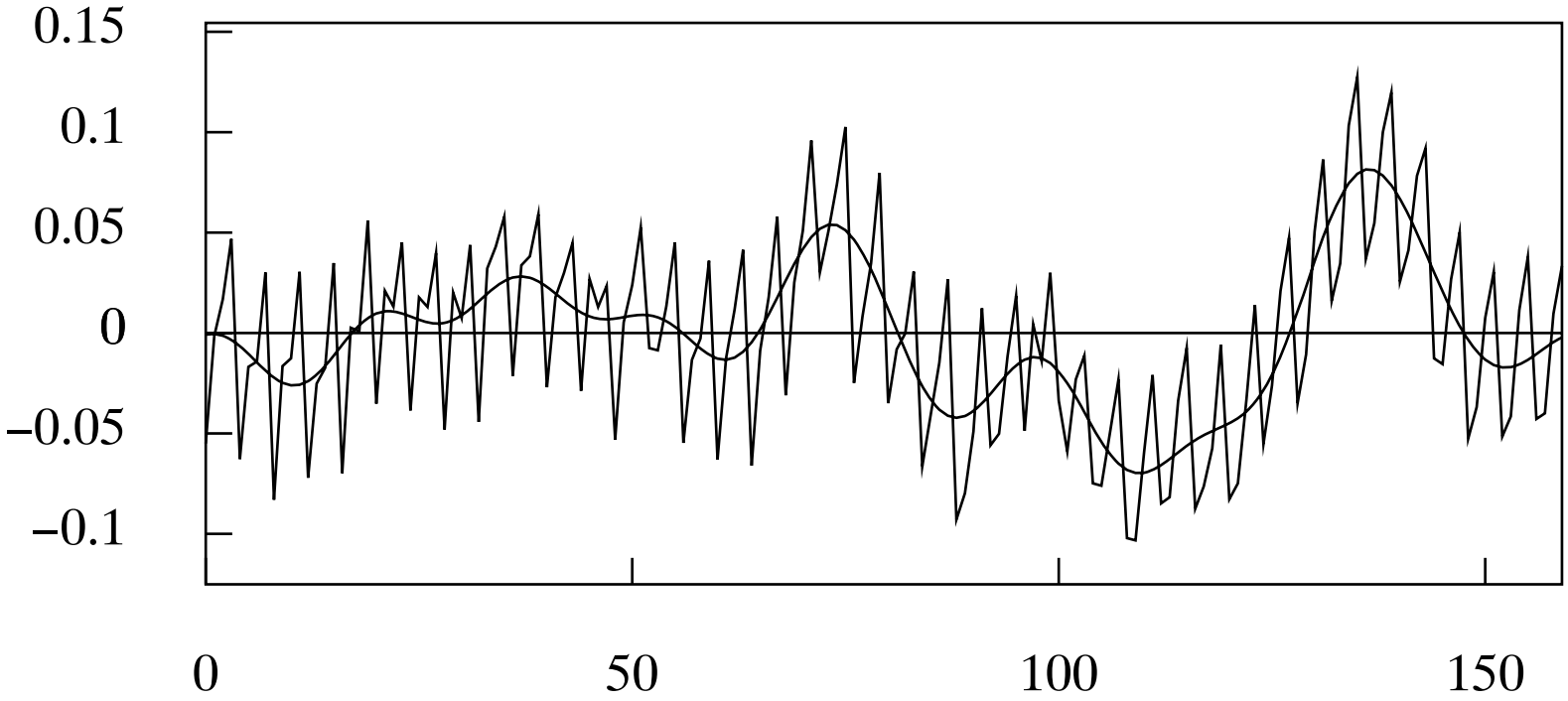


Figure 22. The residual sequence from fitting a linear trend to the logarithmic consumption data. The interpolated line, which represents the business cycle, has been synthesised from the Fourier ordinates in the frequency interval $[0, \pi/8]$.

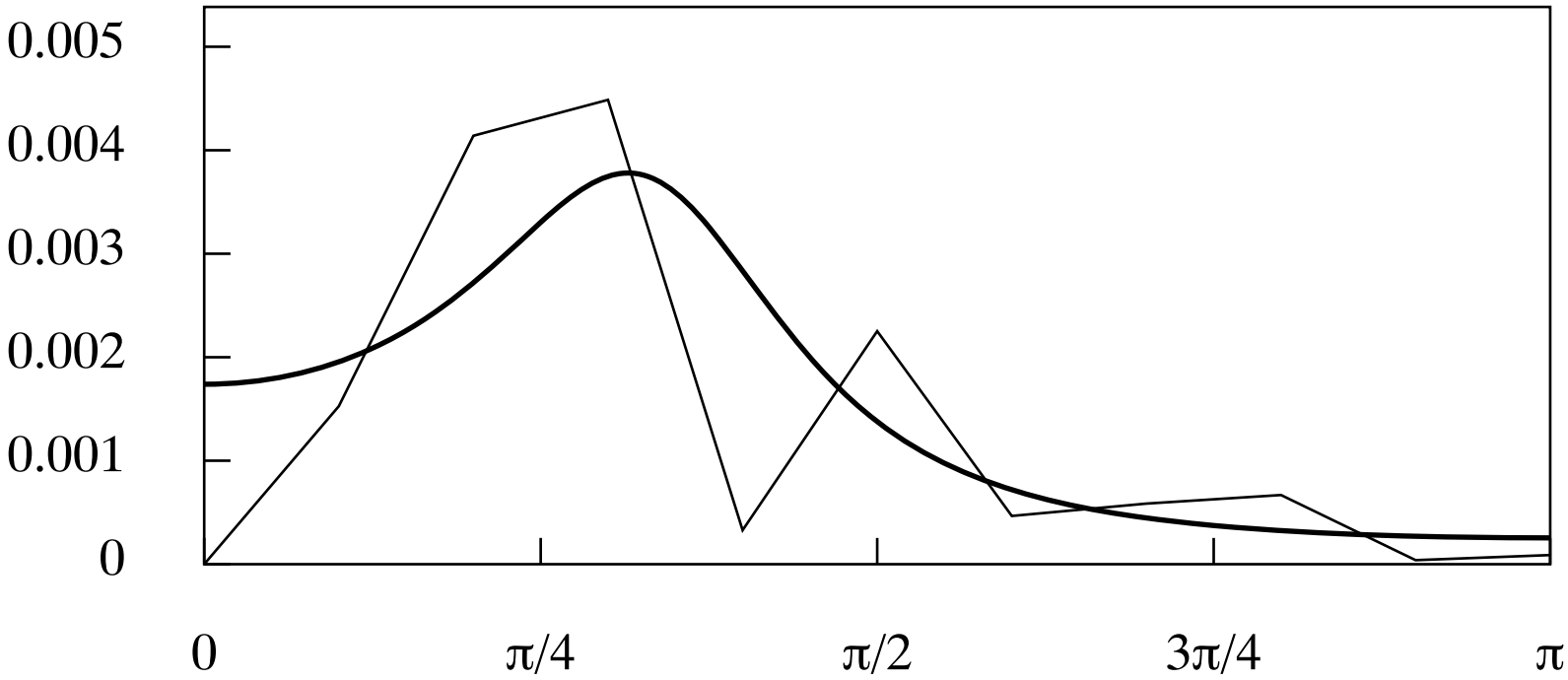


Figure 23. The periodogram of the subsampled anti-aliased data with the parametric spectrum of an estimated AR(2) model superimposed.

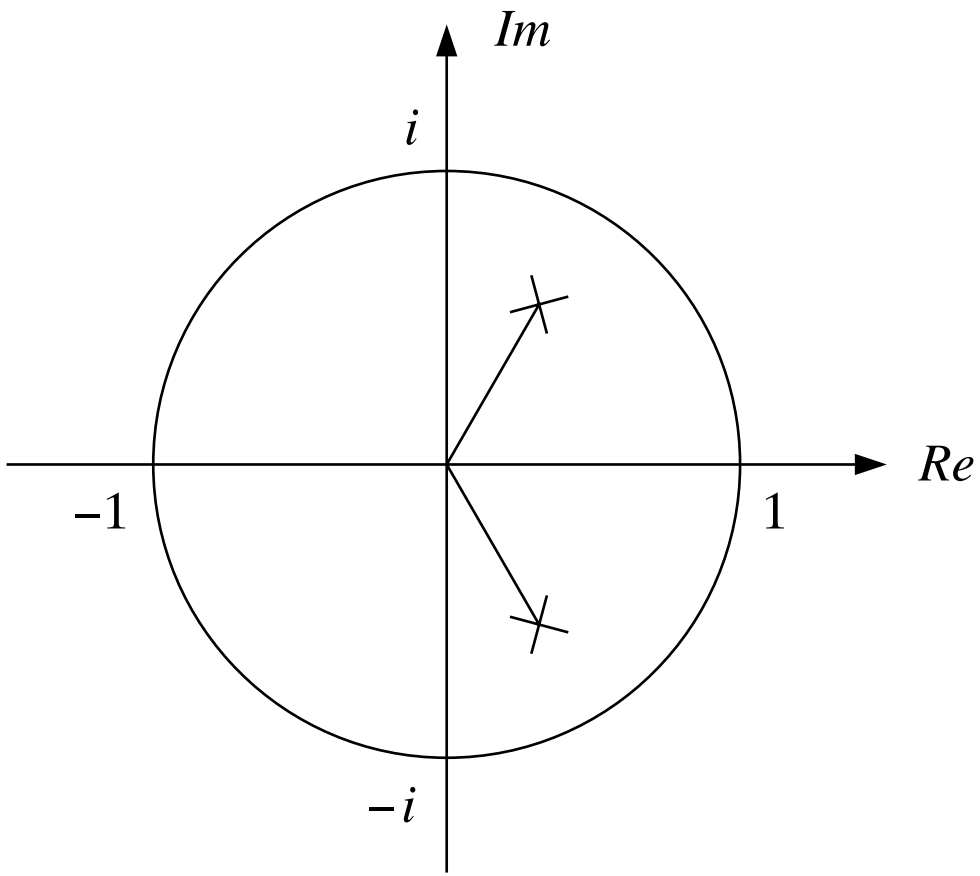


Figure 24. The poles of the AR(2) model fitted to 20 points subsampled at the rate of 1 in 8 from data that has been subjected to an anti-aliasing lowpass filter with cut off at $\pi/8$ radians.

Estimation of Low-Frequency Components via the ZAR Method

Another means of estimating an autoregressive model with improved resolution of low-frequency cycles has been proposed by Morton and Tunincliffe–Wilson (2004).

It is proposed to apply a lowpass filter of the form $(1 - \theta z)^p$ to the white-noise forcing function, where p is the order of the autoregressive operator:

$$(1 + \phi_1 z + \cdots + \phi_p z^p)y(z) = (1 - \theta z)^p \varepsilon(z).$$

The model can also be expressed in the form of

$$(\xi_0 + \xi_1 w + \cdots + \xi_p w^p)y(z) = \varepsilon(z),$$

where

$$w = \frac{z - \theta}{1 - \theta z}.$$

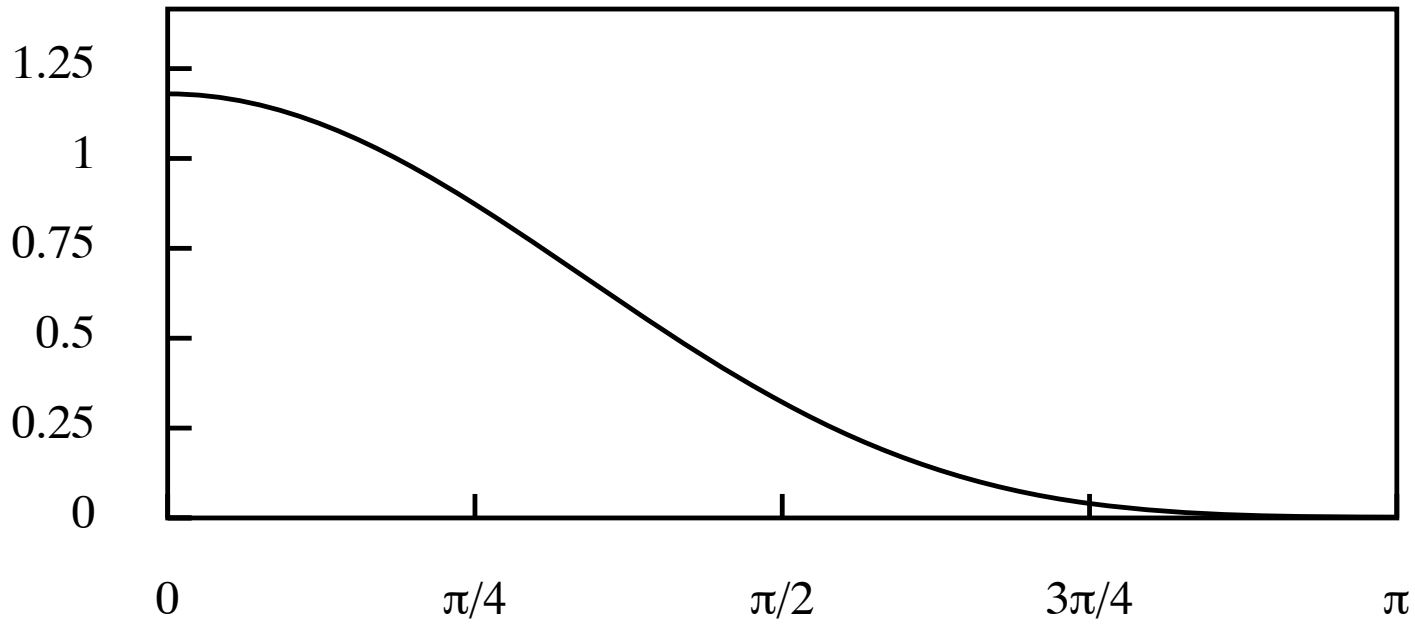
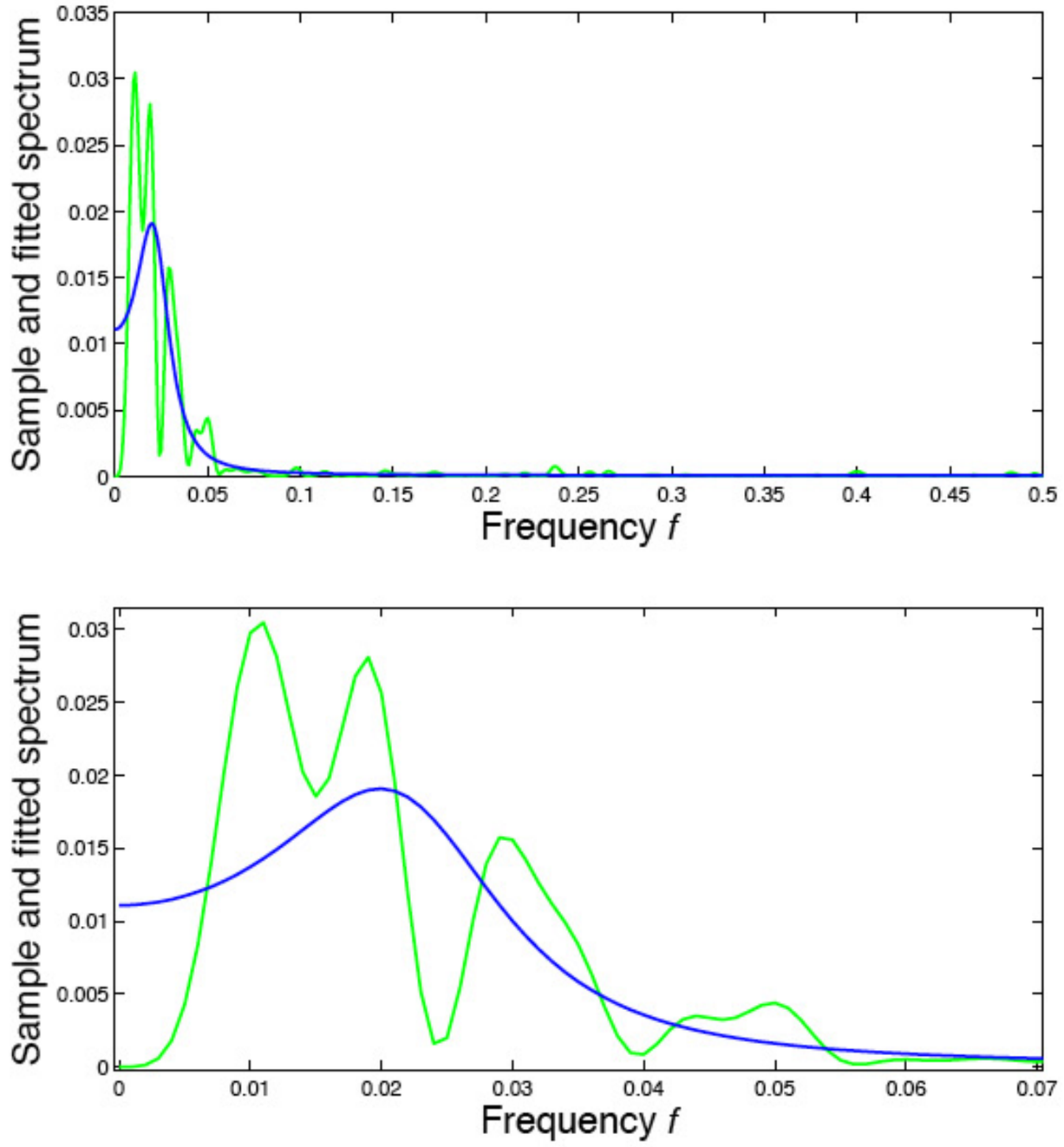


Figure. The squared gain of the lowpass filter $(1 + 0.65z)^2$.


Sample and fitted spectrum
Frequency f
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
0.01
0.02
0.03
0.04
0.05
0.06
0.07

13. Experiments with Pseudo-Random Data

Some sampling experiments have been conducted that have been designed to reaffirm what has been revealed, in the first instance, via the empirical data. An extensive interactive computer program has been written for the purpose. This is available on my Leicester website. I invite others to pursue their own explorations via this program.

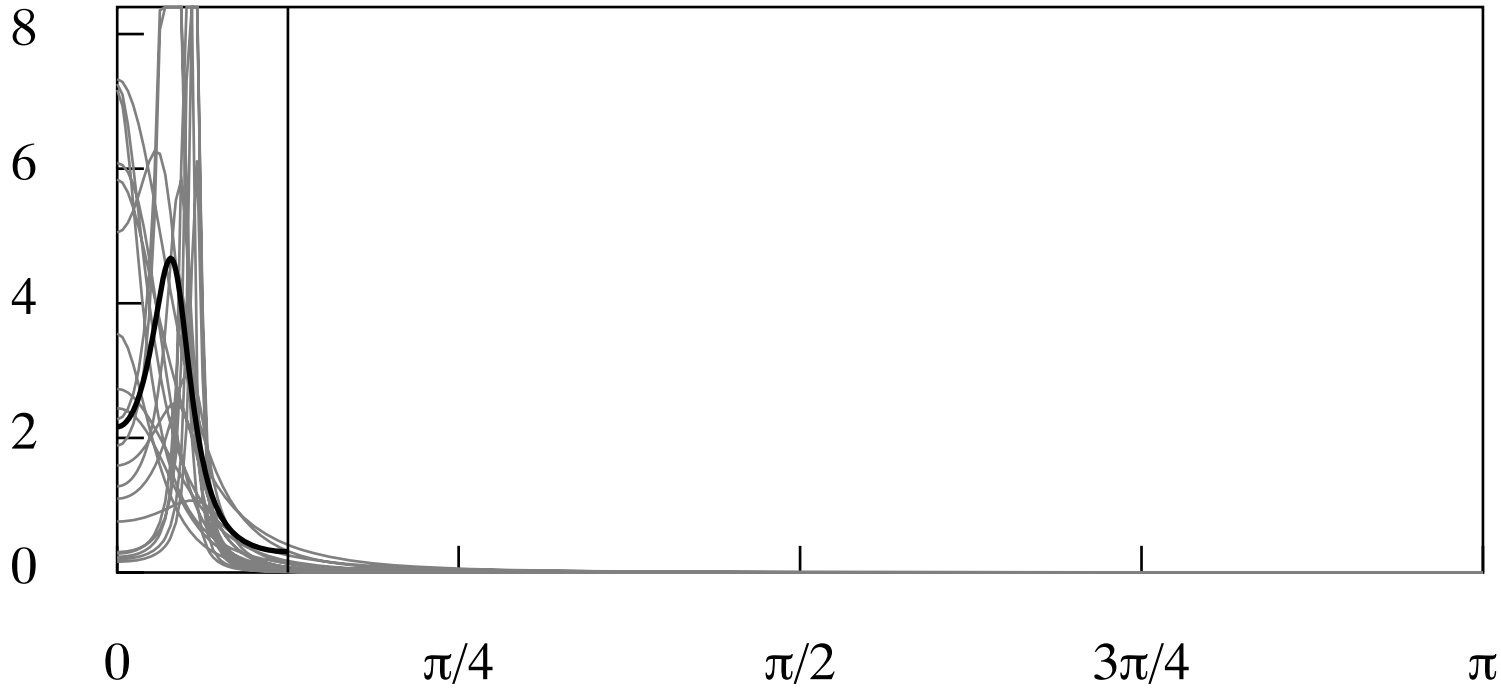


Figure 25. The spectra of 20 AR(2) models estimated from band-limited data supported on the interval $[0, \pi/8]$. The spectrum of the AR(2) model used in generating the data is described by the heavy black line.

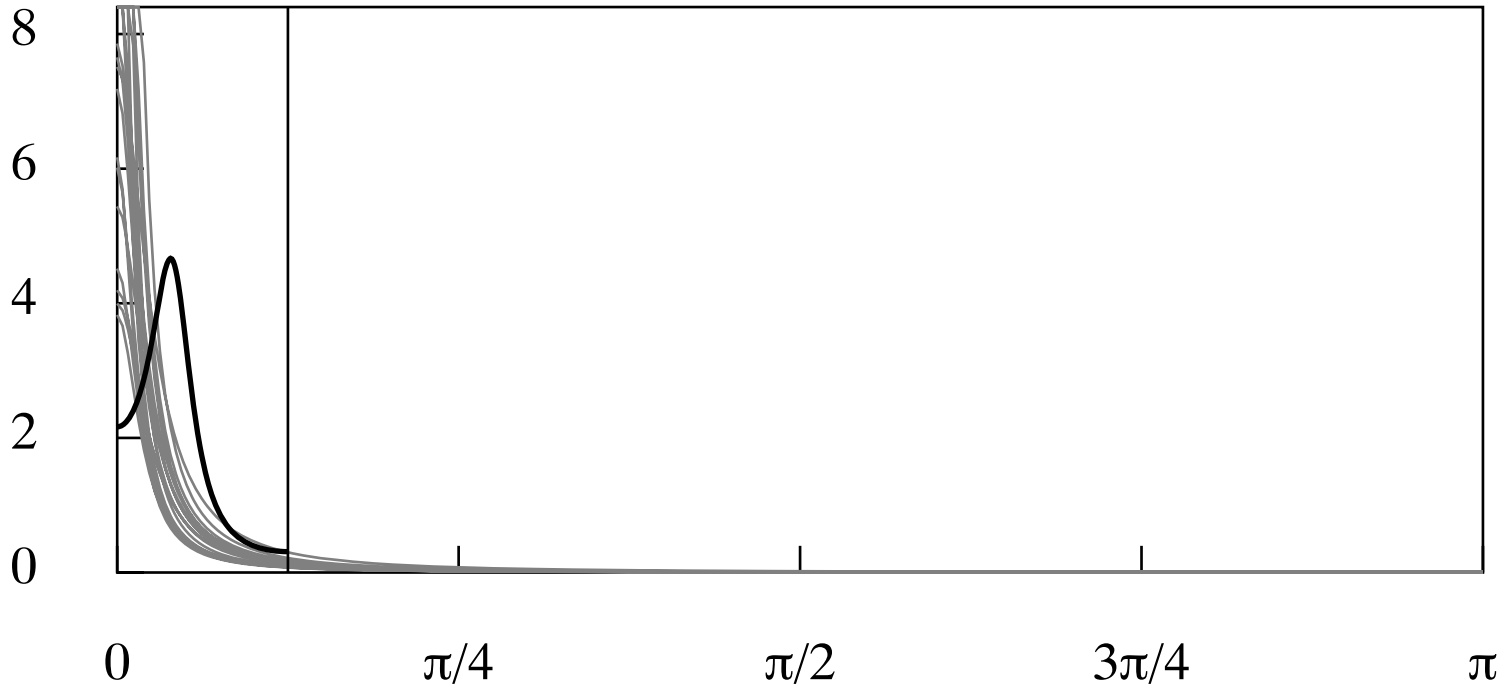


Figure 26. The spectra of 20 AR(2) models estimated from band-limited data contaminated by white noise. The spectrum of the AR(2) model used in generating the data is described by the heavy black line.

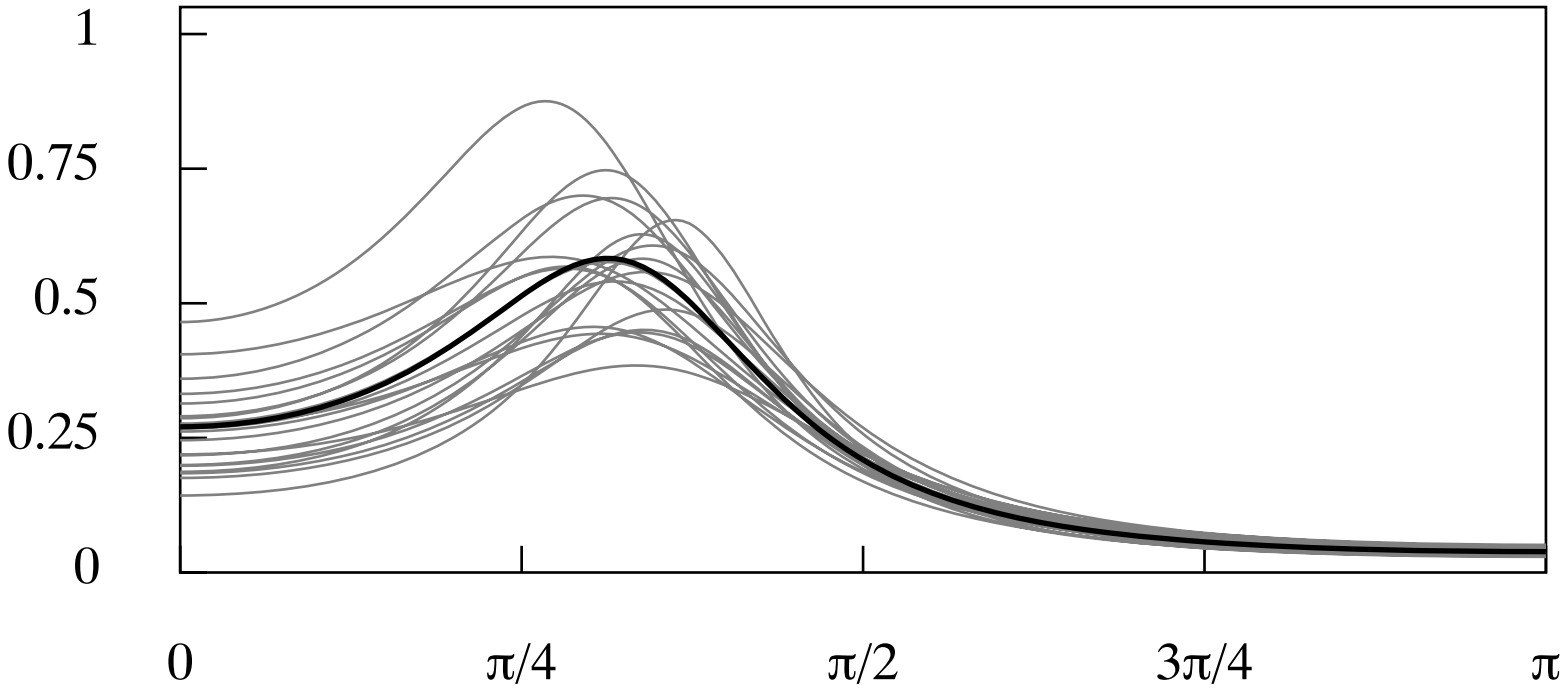


Figure 27. The spectra of 20 AR(2) models estimated from band-limited data subsampled at the critical rate. The spectrum of the AR(2) model used in generating the data is described by the heavy black line.

15. Perfect Predictability of Band-Limited Processes

A band-limited function is analytic. It possesses a series expansion. If we knew all of its derivatives, then we could extrapolate such a function perfectly.

For a stochastic process defined over the the real line, one would need to sample the values of a denumerable infinity of points or derivatives from a finite interval of the past in order to make perfect predictions of the future. This is not practical.

What comes to mind, when one thinks of the notion of perfect predictability, is the concept of Laplacian causality, which is encapsulated in the following familiar quotation:

> "We may regard the present state of the universe as the effect of its past and the cause of its future. An intellect which, at a certain moment, would know all forces that set nature in motion, and all positions of all items of which nature is composed, if it were also vast enough to submit these data to analysis, would embrace, in a single formula, the movements of the greatest bodies of the universe and those of the tiniest atom. For such an intellect, nothing would be uncertain and the future, as much as the past, would be present before its eyes."
>
> Pierre-Simon Laplace, (1814), *Essai Philosophique sur les Probabilités.*

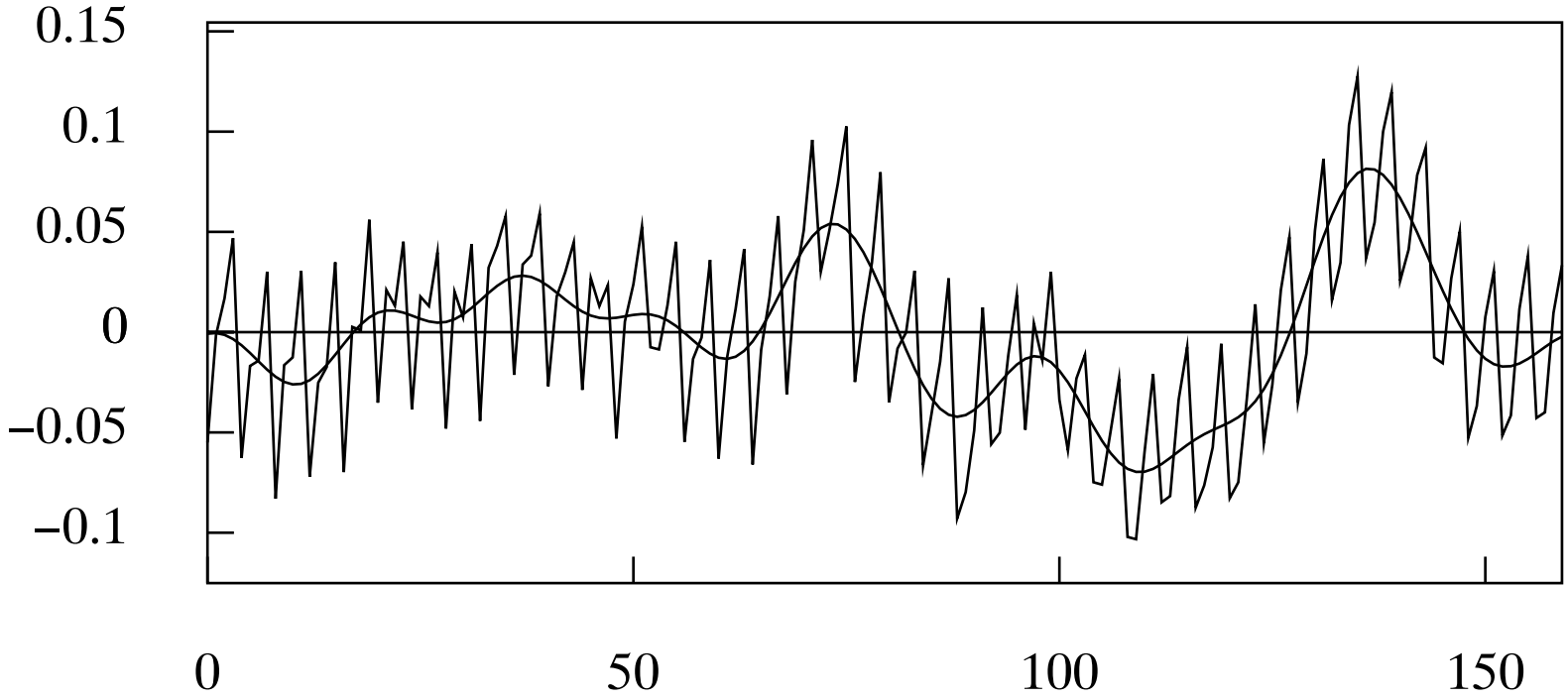


Figure 16. The residual sequence from fitting a linear trend to the logarithmic consumption data with an interpolated function representing the business cycle.

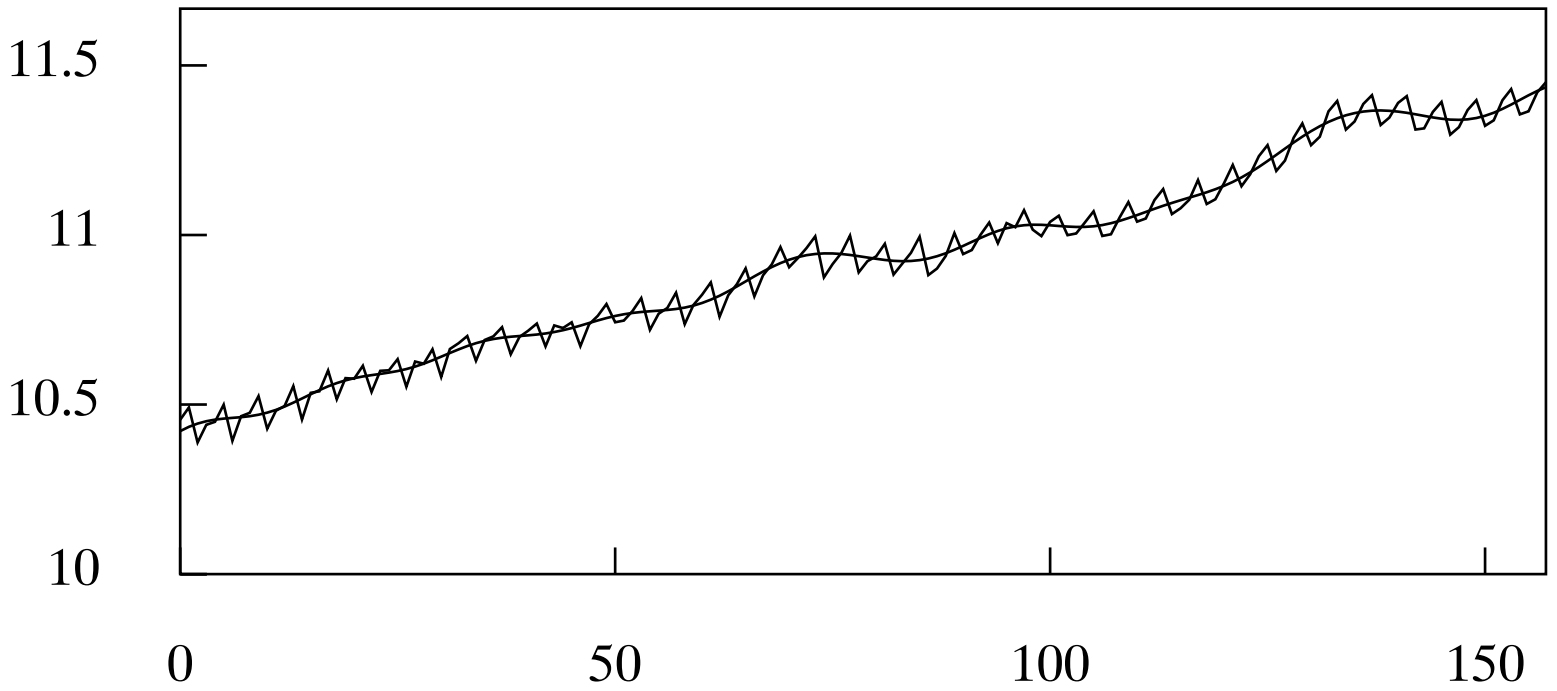


Figure 17. The trend-cycle component of U.K. consumption determined by the frequency-domain method, superimposed on the logarithmic data.

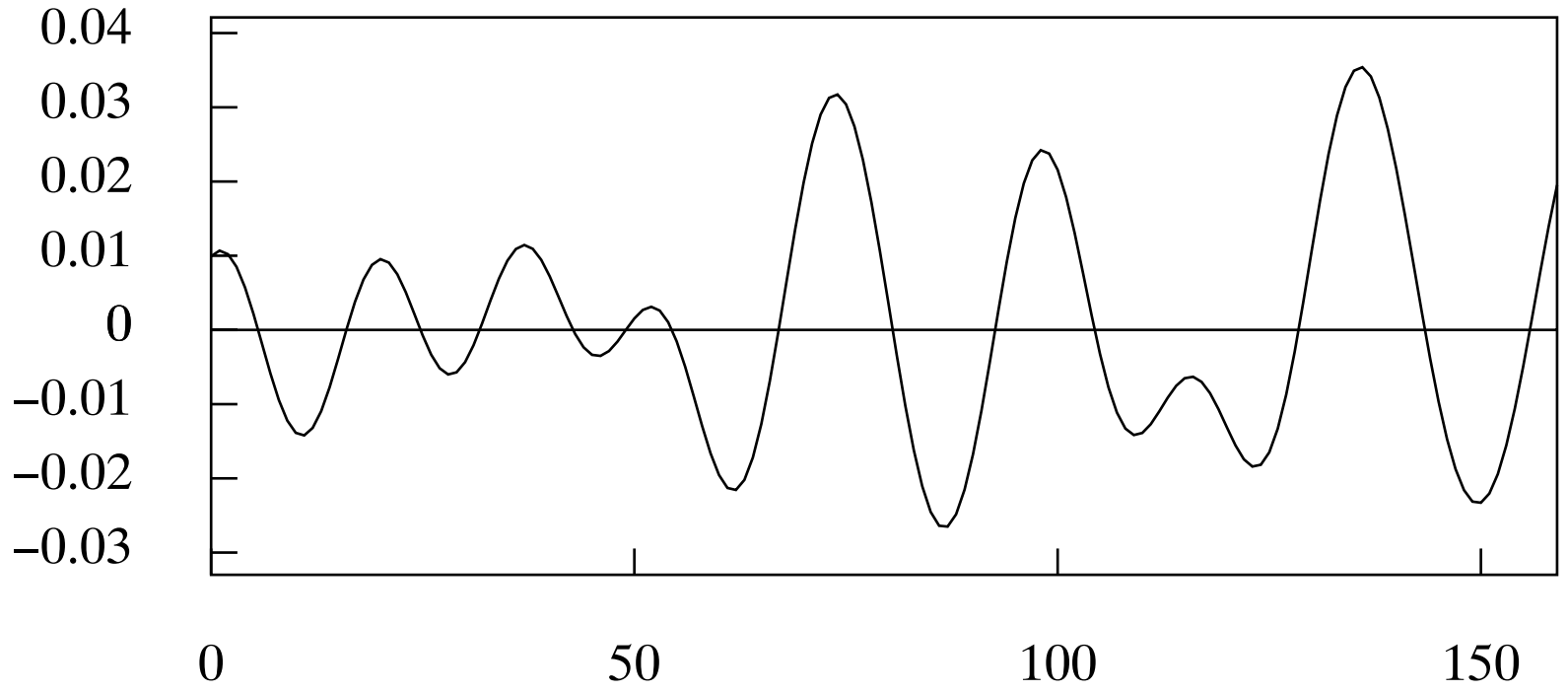


Figure 18. The business cycle determined by the Hodrick–Prescott Filter.

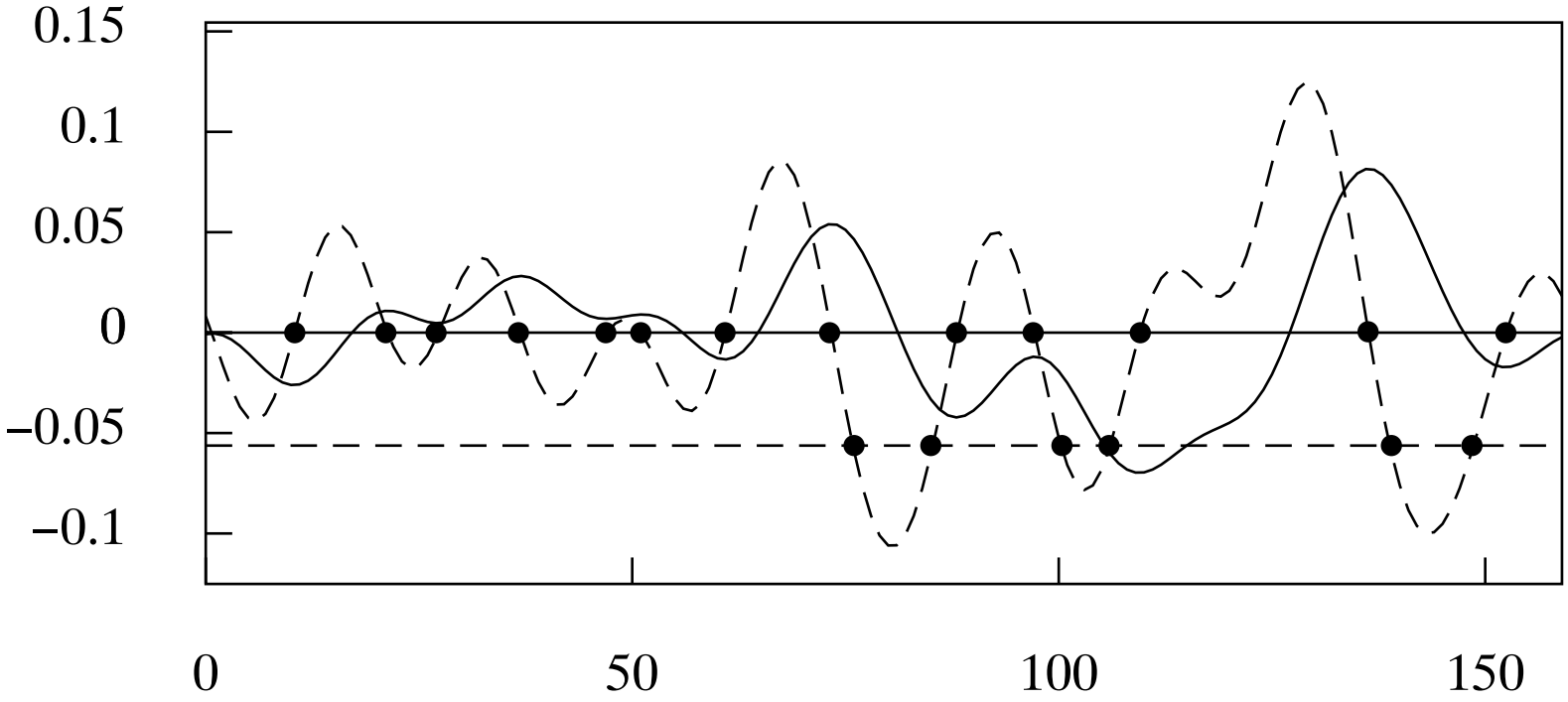


Figure 19. The turning points of the business cycle marked on the horizontal axis by black dots. The solid line is the business cycle of Figure 16. The broken line is the derivative function.

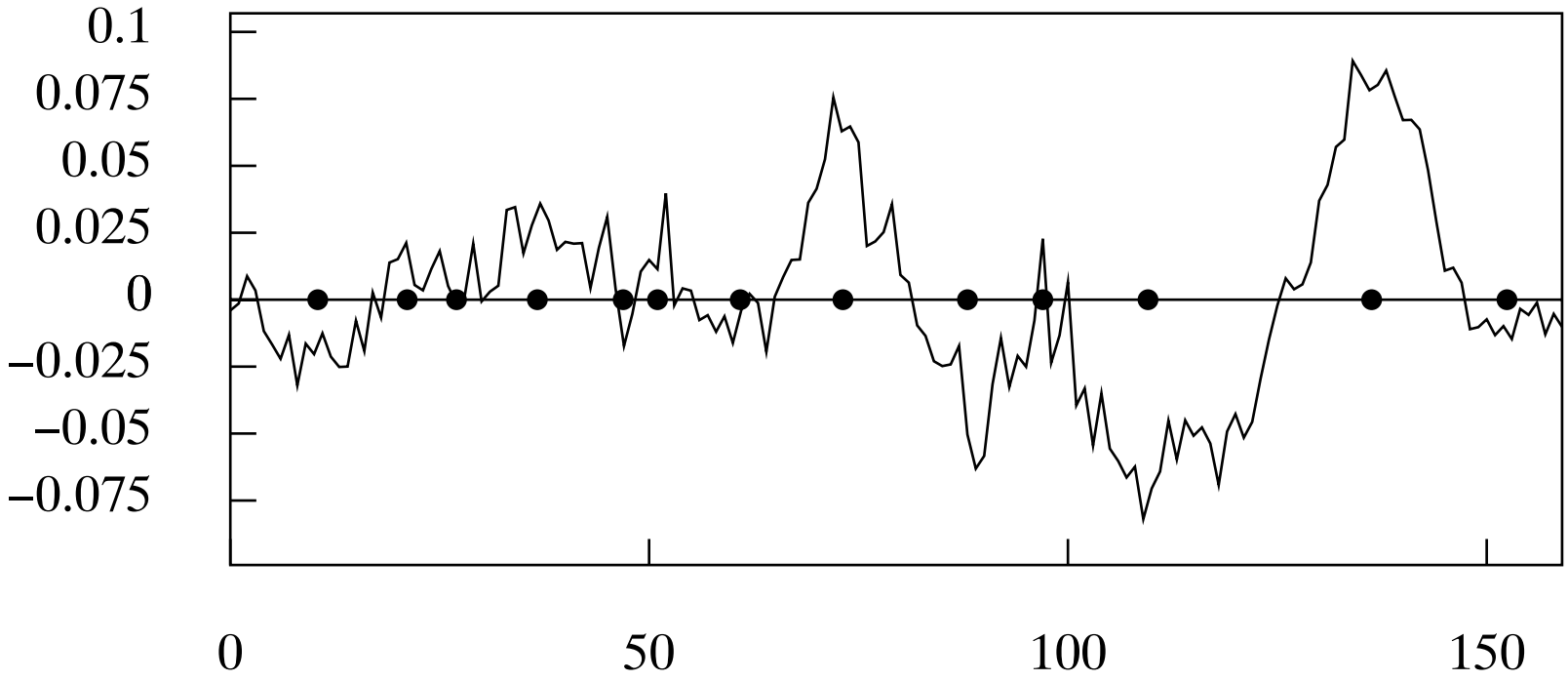


Figure 20. A sequence derived from the detrended data of Figure 16 via the time-domain method of seasonal adjustment.