6 Spectral Analysis and Linear Filters

Up to now we have viewed a time series as a time indexed sequence of random variables. The class of ARMA process was seen as an adequate class of models for the analysis of stationary time series. This approach is usually termed as time series analysis in the *time domain*. There is, however, an equivalent perspective which views a time series as overlayed waves of different frequencies. This view point is termed in time series analysis as the analysis in the *frequency domain*. The decomposition of a time series into sinusoids of different frequencies is called the *spectral representation*. The estimation of the importance of the waves at particular frequencies is referred to as *spectral or spectrum estimation*. Priestley (1981) provides an excellent account of these methods. The use of frequency domain methods, in particular *spectrum estimation*, which originated in the natural sciences was introduced to economics by Granger (1964).[1] Notably, he showed that most of the fluctuations in economic time series can be attributed low frequencies cycles (Granger 1966).

Although both approaches are equivalent, the analysis in the frequency domain is more convenient when it comes to the analysis and construction of *linear filters*. The application of a filter to a time series amounts to take some moving-average of the time series. These moving-average may extend, at least in theory, into the infinite past, but also into the infinite future. A causal ARMA process $\{X_t\}$ may be regarded as filtered white-noise process with filter weights given by ψ_j, $j = 1, 2, \ldots$ In economics, filters are usually applied to remove cycles of a particular frequency, like seasonal cycles (for example Christmas sales in a store), or to highlight particular cycles, like business cycles.

[1] The use of spectral methods in the natural sciences can be traced many centuries back. The modern statistical approach builds on to the work of N. Wiener, G. U. Yule, J. W. Tukey, and many others. See the interesting survey by Robinson (1982).

K. Neusser, *Time Series Econometrics*, Springer Texts in Business and Economics,
DOI 10.1007/978-3-319-32862-1_6

From a mathematical point of view, the equivalence between time and frequency domain analysis rest on the theory of *Fourier series*. An adequate representation of this theory is beyond the scope of this book. The interested reader may consult Brockwell and Davis (1991, chapter 4). An introduction to the underlying mathematical theory can be found in standard textbooks like Rudin (1987).

6.1 The Spectral Density

In the following, we assume that $\{X_t\}$ is a mean-zero (centered) stationary stochastic process with autocovariance function $\gamma(h)$, $h = 0, \pm 1, \pm 2, \ldots$ Mathematically, $\gamma(h)$ represents an double-infinite sequence which can be mapped into a real valued function $f(\lambda)$, $\lambda \in \mathbb{R}$, by the Fourier transform. This function is called the *spectral density function* or *spectral density*. Conversely, we retrieve from the spectral density each covariance. Thus, we have a one-to-one relation between autocovariance functions and spectral densities: both objects summarize the same properties of the time series, but represent them differently.

Definition 6.1 (Spectral Density)**.** *Let $\{X_t\}$ be a mean-zero stationary stochastic process absolutely summable autocovariance function γ then the function*

$$f(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \gamma(h) e^{-\imath h \lambda}, \qquad -\infty < \lambda < \infty, \tag{6.1}$$

is called the spectral density function *or* spectral density *of $\{X_t\}$. Thereby $\imath$ denotes the imaginary unit (see Appendix A).*

The sine is an odd function whereas the cosine and the autocovariance function are even functions.[2] This implies that the spectral density can be rewritten as:

$$\begin{aligned} f(\lambda) &= \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \gamma(h)(\cos(h\lambda) - \imath \sin(h\lambda)) \\ &= \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \gamma(h) \cos(h\lambda) + 0 = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \gamma(h) \cos(-h\lambda) \\ &= \frac{\gamma(0)}{2\pi} + \frac{1}{\pi} \sum_{h=1}^{\infty} \gamma(h) \cos(h\lambda). \end{aligned} \tag{6.2}$$

[2] A function f is called even if $f(-x) = f(x)$; the function is called odd if $f(-x) = -f(x)$. Thus, we have $\sin(-\theta) = -\sin(\theta)$ and $\cos(-\theta) = \cos(\theta)$.

Because of the periodicity of the cosine function, i.e. because

$$f(\lambda + 2k\pi) = f(\lambda), \qquad \text{for all } k \in \mathbb{Z},$$

it is sufficient to consider the spectral density only in the interval $(-\pi, \pi]$. As the cosine is an even function so is f. Thus, we restrict the analysis of the spectral density $f(\lambda)$ further to the domain $\lambda \in [0, \pi]$.

In practice, we often use the *period* or *oscillation length* instead of the radiant λ. They are related by the formula:

$$\text{period length} = \frac{2\pi}{\lambda}. \tag{6.3}$$

If, for example, the data are quarterly observations, a value of 0.3 for λ corresponds to a period length of approximately 21 quarters.

Remark 6.1. We gather some properties of the spectral density function f:

- Because $f(0) = \frac{1}{2\pi}\sum_{h=-\infty}^{\infty}\gamma(h)$, the long-run variance of $\{X_t\}$ J (see Sect. 4.4) equals $2\pi f(0)$, i.e. $2\pi f(0) = J$.
- f is an even function so that $f(\lambda) = f(-\lambda)$.
- $f(\lambda) \geq 0$ for all $\lambda \in (-\pi, \pi]$. The proof of this proposition can be found in Brockwell and Davis (1996, chapter 4). This property corresponds to the non-negative definiteness of the autocovariance function (see property 4 in Theorem 1.1 of Sect. 1.3).
- The single autocovariances are the Fourier-coefficients of the spectral density f:

$$\gamma(h) = \int_{-\pi}^{\pi} e^{\imath h\lambda} f(\lambda) d\lambda = \int_{-\pi}^{\pi} \cos(h\lambda) f(\lambda) d\lambda.$$

  For $h = 0$, we therefore get $\gamma(0) = \int_{-\pi}^{\pi} f(\lambda) d\lambda$.

The last property allows us to compute the autocovariances from a given spectral density. It shows how time and frequency domain analysis are related to each other and how a property in one domain is reflected as a property in the other.

These properties of a non-negative definite function can be used to characterize the spectral density of a stationary process $\{X_t\}$ with autocovariance function γ.

Theorem 6.1 (Properties of a Spectral Density). *A function f defined on $(-\pi, \pi]$ is the spectral density of a stationary process if and only if the following properties hold:*

- $f(\lambda) = f(-\lambda)$;
- $f(\lambda) \geq 0$;
- $\int_{-\pi}^{\pi} f(\lambda) d\lambda < \infty$.

Corollary 6.1. *An absolutely summable function γ is the autocovariance function of a stationary process if and only if*

$$f(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \gamma(h) e^{-\imath h\lambda} \geq 0, \qquad \text{for all } \lambda \in (-\pi, \pi].$$

In this case f is called the spectral density of γ.

The function $f(\lambda)/\gamma(0)$ can be considered as a density function of some probability distribution defined on $[-\pi, \pi]$ because $\frac{f(\lambda)}{\gamma(0)} \geq 0$ and

$$\int_{-\pi}^{\pi} \frac{f(\lambda)}{\gamma(0)} d\lambda = 1.$$

The corresponding cumulative distribution function G is then defined as:

$$G(\lambda) = \int_{-\pi}^{\lambda} \frac{f(\omega)}{\gamma(0)} d\omega, \qquad -\pi \leq \lambda \leq \pi.$$

It satisfies: $G(-\pi) = 0$, $G(\pi) = 1$, $1 - G(\lambda) = G(\lambda)$, und $G(0) = 1/2$. The autocorrelation function ρ is then given by

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \int_{-\pi}^{\pi} e^{\imath h\lambda} dG(\lambda).$$

Some Examples

Some relevant examples illustrating the above are:

white noise: Let $\{X_t\}$ be a white noise process with $X_t \sim \mathrm{WN}(0, \sigma^2)$. For this process all autocovariances, except $\gamma(0)$, are equal to zero. The spectral density therefore is equal to

$$f(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \gamma(h) e^{-\imath h\lambda} = \frac{\gamma(0)}{2\pi} = \frac{\sigma^2}{2\pi}.$$

Thus, the spectral density is equal to a constant which is proportional to the variance. This means that no particular frequency dominates the spectral density. This is the reason why such a process is called white noise.

MA(1): Let $\{X_t\}$ be a MA(1) process with autocovariance function

$$\gamma(h) = \begin{cases} 1, \ h = 0; \\ \rho, \ h = \pm 1; \\ 0, \ \text{otherwise}. \end{cases}$$

The spectral density therefore is:

$$f(\lambda) = \frac{1}{2\pi}\sum_{h=-\infty}^{\infty}\gamma(h)e^{-\imath h\lambda} = \frac{\rho e^{\imath\lambda} + 1 + \rho e^{-\imath\lambda}}{2\pi} = \frac{1 + 2\rho\cos\lambda}{2\pi}.$$

Thus, $f(\lambda) \geq 0$ if and only $|\rho| \leq 1/2$. According to Corollary 6.1 above, γ is the autocovariance function of a stationary stochastic process if and only if $|\rho| \leq 1/2$. This condition corresponds exactly to the condition derived in the time domain (see Sect. 1.3). The spectral density for $\rho = 0.4$ or equivalently $\theta = 0.5$, respectively for $\rho = -0.4$ or equivalently $\theta = -0.5$, and $\sigma^2 = 1$ is plotted in Fig. 6.1a. As the process is rather smooth when the first order autocorrelation is positive, the spectral density is large in the neighborhood of zero and small in the neighborhood of π. For a negative autocorrelation the picture is just reversed.

AR(1): The spectral density of an AR(1) process $X_t = \phi X_{t-1} + Z_t$ with $Z_t \sim \mathrm{WN}(0, \sigma^2)$ is:

$$\begin{aligned} f(\lambda) &= \frac{\gamma(0)}{2\pi}\left(1 + \sum_{h=1}^{\infty}\phi^h\left(e^{-\imath h\lambda} + e^{\imath h\lambda}\right)\right) \\ &= \frac{\sigma^2}{2\pi(1-\phi^2)}\left(1 + \frac{\phi e^{\imath\lambda}}{1-\phi e^{\imath\lambda}} + \frac{\phi e^{-\imath\lambda}}{1-\phi e^{-\imath\lambda}}\right) = \frac{\sigma^2}{2\pi}\frac{1}{1 - 2\phi\cos\lambda + \phi^2} \end{aligned}$$

The spectral density for $\phi = 0.6$ and $\phi = -0.6$ and $\sigma^2 = 1$ are plotted in Fig. 6.1b. As the process with $\phi = 0.6$ exhibits a relatively large positive autocorrelation so that it is rather smooth, the spectral density takes large values for low frequencies. In contrast, the process with $\phi = -0.6$ is rather volatile due to the negative first order autocorrelation. Thus, high frequencies are more important than low frequencies as reflected in the corresponding figure.

Note that, as ϕ approaches one, the spectral density evaluated at zero tends to infinity, i.e. $\lim_{\lambda\downarrow 0} f(\lambda) = \infty$. This can be interpreted in the following way. As the process gets closer to a random walk more and more weight is given to long-run fluctuations (cycles with very low frequency or very high periodicity) (Granger 1966).

6.2 Spectral Decomposition of a Time Series

Consider the simple harmonic process $\{X_t\}$ which just consists of a cosine and a sine wave:

$$X_t = A\cos(\omega t) + B\sin(\omega t). \tag{6.4}$$

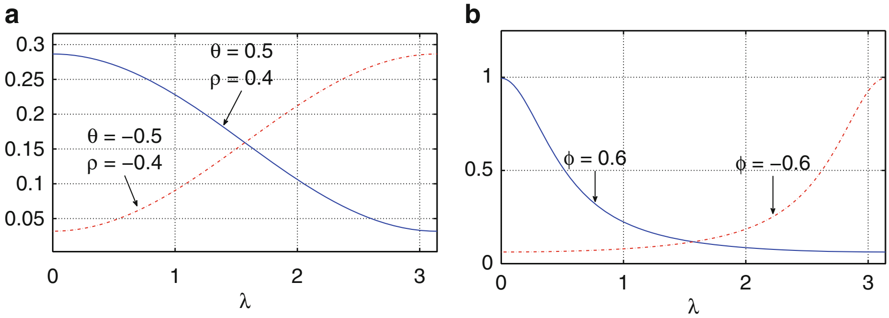

Fig. 6.1 Examples of spectral densities with $Z_t \sim \text{WN}(0, 1)$. (**a**) MA(1) process. (**b**) AR(1) process

Thereby A and B are two uncorrelated random variables with $\mathbb{E}A = \mathbb{E}B = 0$ and $\mathbb{V}A = \mathbb{V}B = 1$. The autocovariance function of this process is $\gamma(h) = \cos(\omega h)$. This autocovariance function cannot be represented as $\int_{-\pi}^{\pi} e^{\imath h\lambda} f(\lambda) d\lambda$. However, it can be regarded as the Fourier transform of a discrete distribution function F:

$$\gamma(h) = \cos(\omega h) = \int_{(-\pi,\pi]} e^{\imath h\lambda} dF(\lambda),$$

where

$$F(\lambda) = \begin{cases} 0, & \text{for } \lambda < -\omega; \\ 1/2, & \text{for } -\omega \leq \lambda < \omega; \\ 1, & \text{for } \lambda \geq \omega. \end{cases} \tag{6.5}$$

The integral with respect to the discrete distribution function is a so-called Riemann-Stieltjes integral.[3] F is a step function with jumps at $-\omega$ and ω and step size of $1/2$ so that the above integral equals $\frac{1}{2}e^{-\imath h\omega} + \frac{1}{2}e^{-\imath h\omega} = \cos(h\omega)$.

These considerations lead to a representation, called the *spectral representation*, of the autocovariance function as the Fourier transform a distribution function over $[-\pi, \pi]$.

[3]The Riemann-Stieltjes integral is a generalization of the Riemann integral. Let f and g be two bounded functions defined on the interval $[a, b]$ then the Riemann-Stieltjes integral $\int_a^b f(x) dg(x)$ is defined as $\lim_{n\to\infty} \sum_{i=1}^n f(\xi_i)[g(x_i) - g(x_{i-1})]$ where $a = x_1 < x_2 < \cdots < x_{n-1} < x_n = b$. For $g(x) = x$ we obtain the standard Riemann integral. If g is a step function with a countable number of steps x_i of height h_i then $\int_a^b f(x) dg(x) = \sum_i f(x_i) h_i$.

Theorem 6.2 (Spectral Representation). *γ is the autocovariance function of a stationary process $\{X_t\}$ if and only if there exists a right-continuous, nondecreasing, bounded function F on $(-\pi, \pi]$ with the properties $F(-\pi) = 0$ and*

$$\gamma(h) = \int_{(-\pi,\pi]} e^{\imath h\lambda} dF(\lambda). \tag{6.6}$$

F is called the spectral distribution function of γ.

Remark 6.2. If the spectral distribution function F has a density f such that $F(\lambda) = \int_{-\pi}^{\lambda} f(\omega)d\omega$ then f is called the *spectral density* and the time series is said to have a continuous spectrum.

Remark 6.3. According to the Lebesgue-Radon-Nikodym Theorem (see, for example, Rudin (1987)), the spectral distribution function F can be represented uniquely as the sum of a distribution function F_Z which is absolutely continuous with respect to the Lebesgue measure and a discrete distribution function F_V. The distribution function F_Z corresponds to the regular part of the Wold Decomposition (see Theorem 3.1 in Sect. 3.2) and has spectral density

$$f_Z(\lambda) = \frac{\sigma^2}{2\pi} \left|\Psi(e^{-\imath\lambda})\right|^2 = \frac{\sigma^2}{2\pi} \left|\sum_{j=0}^{\infty} \psi_j e^{-\imath j\lambda}\right|^2.$$

The discrete distribution F_V corresponds to the deterministic part $\{V_t\}$.

The process (6.4) considers just a single frequency ω. We may, however, generalize this process by superposing several sinusoids. This leads to the class of *harmonic processes*:

$$X_t = \sum_{j=1}^{k} A_j \cos(\omega_j t) + B_j \sin(\omega_j t), \qquad 0 < \omega_1 < \cdots < \omega_k < \pi \tag{6.7}$$

where $A_1, B_1, \ldots, A_k, B_k$ are random variables which are uncorrelated with each other and which have means $\mathbb{E}A_j = \mathbb{E}B_j = 0$ and variances $\mathbb{V}A_j = \mathbb{V}B_j = \sigma_j^2$, $j = 1, \ldots, k$. The autocovariance function of such a process is given by $\gamma(h) = \sum_{j=1}^{k} \sigma_j^2 \cos(\omega_j h)$. According to the spectral representation theorem the corresponding distribution function F can be represented as a weighted sum of distribution functions like those in Eq. (6.5):

$$F(\lambda) = \sum_{j=1}^{k} \sigma_j^2 F_j(\lambda)$$

with

$$F_j(\lambda) = \begin{cases} 0, & \text{for } \lambda < -\omega_j; \\ 1/2, & \text{for } -\omega_j \leq \lambda < \omega_j; \\ 1, & \text{for } \lambda \geq \omega_j. \end{cases}$$

This generalization points to the following properties:

- Each of the k components $A_j \cos(\omega_j t) + B_j \sin(\omega_j t)$, $j = 1, \ldots, k$, is completely associated to a specific frequency ω_j.
- The k components are uncorrelated with each other.
- The variance of each component is σ_j^2. The contribution of each component to the variance of X_t given by $\sum_{j=1}^k \sigma_j^2$ therefore is σ_j^2.
- F is a nondecreasing step-function with jumps at frequencies $\omega = \pm\omega_j$ and step sizes $\frac{1}{2}\sigma_j^2$.
- The corresponding probability distribution is discrete with values $\frac{1}{2}\sigma_j^2$ at the frequencies $\omega = \pm\omega_j$ and zero otherwise.

The interesting feature of harmonic processes as represented in Eq. (6.7) is that *every* stationary process can be represented as the superposition of uncorrelated sinusoids. However, in general infinitely many (even uncountably many) of these processes have to be superimposed. The generalization of (6.7) then leads to the spectral representation of a stationary stochastic process:

$$X_t = \int_{(-\pi,\pi]} e^{\imath t\lambda} \mathrm{d}Z(\lambda). \tag{6.8}$$

Thereby $\{Z(\lambda)\}$ is a complex-valued stochastic process with uncorrelated increments defined on the interval $(-\pi, \pi]$. The above representation is known as the *spectral representation* of the process $\{X_t\}$.[4] Note the analogy to the spectral representation of the autocovariance function in Eq. (6.6).

For the harmonic processes in Eq. (6.7), we have:

$$\mathrm{d}Z(\lambda) = \begin{cases} \frac{A_j + \imath B_j}{2}, & \text{for } \lambda = -\omega_j \text{ and } j = 1, 2, \ldots, k; \\ \frac{A_j - \imath B_j}{2}, & \text{for } \lambda = \omega_j \text{ and } j = 1, 2, \ldots, k; \\ 0, & \text{otherwise.} \end{cases}$$

In this case the variance of $\mathrm{d}Z$ is given by:

$$\mathbb{E}\mathrm{d}Z(\lambda)\overline{\mathrm{d}Z(\lambda)} = \begin{cases} \frac{\sigma_j^2}{2}, & \text{if } \lambda = \pm\omega_j; \\ 0, & \text{otherwise.} \end{cases}$$

[4] A mathematically precise statement is given in Brockwell and Davis (1991, chapter 4) where also the notion of stochastic integration is explained.

In general, we have:

$$= \begin{cases} F(\lambda) - F(\lambda^-), & \text{discrete spectrum;} \\ f(\lambda)\mathrm{d}\lambda, & \text{continuous spectrum.} \end{cases}$$

Thus, a large jump of the spectrum at frequency λ is associated with a large sinusoidal component with frequency λ.[5]

6.3 The Periodogram and the Estimation of Spectral Densities

Although the spectral distribution function is uniquely determined, its estimation from a finite sample with realizations $\{x_1, x_2, \ldots, x_T\}$ is not easy. This has to do with the problem of estimating a function from a finite number of points. We will present two-approaches: a non-parametric and a parametric one.

6.3.1 Non-Parametric Estimation

A simple estimator of the spectral density, $\hat{f}_T(\lambda)$, can be obtained by replacing in the defining equation (6.1) the theoretical autocovariances γ by their estimates $\hat{\gamma}$. However, instead of a simple sum, we consider a weighted sum:

$$\hat{f}_T(\lambda) = \frac{1}{2\pi} \sum_{|h| \leq \ell_T} k\left(\frac{h}{\ell_T}\right) \hat{\gamma}(h) \mathrm{e}^{-\imath h \lambda}. \tag{6.9}$$

The weighting function k, also known as the *lag window*, is assumed to have exactly the same properties as the kernel function introduced in Sect. 4.4. This correspondence is not accidental, indeed the long-run variance defined in Eq. (4.1) is just 2π times the spectral density evaluated at $\lambda = 0$. Thus, one might choose a weighting, kernel or lag window from Table 4.1, like the Bartlett-window, and use it to estimate the spectral density. The lag truncation parameter is chosen in such a way that $\ell_T \to \infty$ as $T \to \infty$. The rate of divergence should, however, be smaller than T so that $\frac{\ell_T}{T}$ approaches zero as T goes to infinity. As an estimator of the autocovariances one uses the estimator given in Eq. (4.2) of Sect. 4.2.

The above estimator is called an *indirect spectral estimator* because it requires the estimation of the autocovariances in the first step. The *periodogram* provides an alternative *direct spectral estimator*. For this purpose, we represent the observations as linear combinations of sinusoids of specific frequencies. These so-called *Fourier frequencies* are defined as $\omega_k = \frac{2\pi k}{T}$, $k = -\lfloor\frac{T-1}{2}\rfloor, \ldots, \lfloor\frac{T}{2}\rfloor$. Thereby $\lfloor x \rfloor$ denotes

[5]Thereby $F(\lambda^-)$ denotes the left-sided limit, i.e. $F(\lambda^-) = \lim_{\omega \uparrow \lambda} F(\omega)$.

the largest integer smaller or equal to x. With this notation, the observations x_t, $t = 1, \ldots, T$, can be represented as a sum of sinusoids:

$$x_t = \sum_{k=-\lfloor \frac{T-1}{2} \rfloor}^{\lfloor \frac{T}{2} \rfloor} a_k e^{\imath \omega_k t} = \sum_{k=-\lfloor \frac{T-1}{2} \rfloor}^{\lfloor \frac{T}{2} \rfloor} a_k (\cos(\omega_k t) + \imath \sin(\omega_k t)).$$

The coefficients $\{a_k\}$ are the *discrete Fourier-transform* of the observations $\{x_1, x_2, \ldots, x_T\}$. The periodogram I_T is then defined as follows.

Definition 6.2 (Periodogram)**.** *Given observations* $\{x_1, x_2, \ldots, x_T\}$*, the periodogram is defined as the function*

$$I_T(\lambda) = \frac{1}{T} \left| \sum_{t=1}^{T} x_t e^{-\imath t \lambda} \right|^2 .$$

For each Fourier-frequency ω_k, the periodogram $I_T(\omega_k)$ equals $|a_k|^2$. This implies that

$$\sum_{t=1}^{T} |x_t|^2 = \sum_{k=-\lfloor \frac{T-1}{2} \rfloor}^{\lfloor \frac{T}{2} \rfloor} |a_k|^2 = \sum_{k=-\lfloor \frac{T-1}{2} \rfloor}^{\lfloor \frac{T}{2} \rfloor} I_T(\omega_k).$$

The value of the periodogram evaluated at the Fourier-frequency ω_k is therefore nothing but the contribution of the sinusoid with frequency ω_k to the variation of $\{x_t\}$ as measured by sum of squares. In particular, for any Fourier-frequency different from zero we have that

$$I_T(\omega_k) = \sum_{h=-T+1}^{T-1} \hat{\gamma}(h) e^{-\imath h \omega_k}.$$

Thus the periodogram represents, disregarding the proportionality factor 2π, the sample analogue of the spectral density and therefore carries the same information.

Unfortunately, it turns out that the periodogram is not a consistent estimator of the spectral density. In particular, the covariance between $I_T(\lambda_1)$ and $I_T(\lambda_2)$, $\lambda_1 \neq \lambda_2$, goes to zero for T going to infinity. The periodogram thus has a tendency to get very jagged for large T leading to the detection of spurious sinusoids. A way out of this problem is to average the periodogram over neighboring frequencies, thereby reducing its variance. This makes sense because the variance is relatively constant within a small frequency band. The averaging (smoothing) of the periodogram over neighboring frequencies leads to the class of *discrete spectral average estimators* which turn out to be consistent:

$$\hat{f}_T(\lambda) = \frac{1}{2\pi} \sum_{|h| \leq \ell_T} K_T(h) I_T \left(\tilde{\omega}_{T,\lambda} + \frac{2\pi h}{T} \right) \tag{6.10}$$

where $\tilde{\omega}_{T,\lambda}$ denotes the multiple of $\frac{2\pi}{T}$ which is closest to λ. ℓ_T is the bandwidth of the estimator, i.e. the number of ordinates over which the average is taken. ℓ_T satisfies the same properties as in the case of the indirect spectral estimator (6.9): $\ell_T \to \infty$ and $\ell_T/T \to 0$ for $T \to \infty$. Thus, as T goes to infinity, on the one hand the average is taken over more and more values, but on the other hand the frequency band over which the average is taken is getting smaller and smaller. The *spectral weighting function* or *spectral window* K_T is a positive even function satisfying $\sum_{|h|\leq \ell_T} K_T(h) = 1$ and $\sum_{|h|\leq \ell_T} K_T^2(h) \to 0$ for $T \to \infty$. It can be shown that under these conditions the discrete spectral average estimator is mean-square consistent. Moreover, the estimator in Eq. (6.9) can be approximated by a corresponding discrete spectral average estimator by defining the spectral window as

$$K_T(\omega) = \frac{1}{2\pi} \sum_{|h|\leq \ell_T} k\left(\frac{h}{\ell_T}\right) \mathrm{e}^{-\imath h\omega}$$

or vice versa

$$k(h) = \int_{-\pi}^{\pi} K_T(\omega) \mathrm{e}^{-\imath h\omega} \mathrm{d}\omega.$$

Thus, the lag and the spectral window are related via the Fourier transform. For details and the asymptotic distribution the interested reader is referred to Brockwell and Davis (1991, Chapter10). Although the indirect and the direct estimator give approximately the same result when the kernels used are related as in the equation above, the direct estimator (6.10) is usually preferred in practice because it is, especially for long times series, computationally more efficient, in particular in connection with the fast fourier transformation (FFT).[6]

A simple spectral weighting function, known as the Daniell spectral window, is given by $K_T(h) = (2\ell_T + 1)^{-1}$ when $|h| \leq \ell_T$ and 0 otherwise and where $\ell_T = \sqrt{T}$. It averages over $2\ell_T + 1$ values within a frequency band of approximate width $\frac{4\pi}{\sqrt{T}}$. This function corresponds to the Daniell kernel function or Daniell lag window $k(x) = \sin(\pi x)/(\pi x)$ for $|x| \leq 1$ and zero otherwise (see Sect. 4.4). In practice, the sample size is fixed and the researcher is faced with a trade-off between variance and bias. On the one hand, a weighting function which averages over a wide frequency band produces a smooth spectral density, but has probably a large bias because the estimate of $f(\lambda)$ depends on frequencies which are rather far away from λ. On the other hand, a weighting function which averages only over a small frequency band produces a small bias, but probably a large variance. It is thus advisable in practice

[6] The FFT is seen as one of the most important numerical algorithms ever as it allows a rapid computation of Fourier transformations and its inverse. The FFT is widely in digital signal processing.

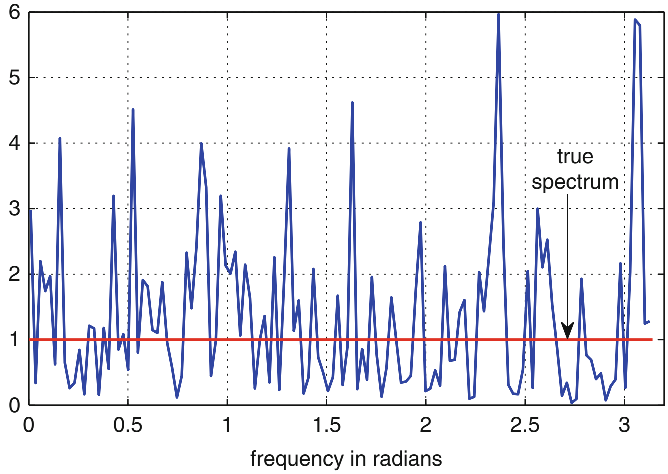

Fig. 6.2 Raw periodogram of a white noise time series ($X_t \sim \text{WN}(0,1)$, $T = 200$)

to work with alternative weighting functions and to choose the one which delivers a satisfying balance between bias and variance.

The following two examples demonstrate the large variance of the periodogram. The first example consists of 200 observations from a simulated white noise time series with variance equal. Whereas the true spectrum is constant equal to one, the raw periodogram, i.e. the periodogram without smoothing, plotted in Fig. 6.2 is quite erratic. However, it is obvious that by taking averages of adjacent frequencies the periodogram becomes smoother and more in line with the theoretical spectrum. The second example consists of 200 observations of a simulated AR(2) process. Figure 6.3 demonstrates again the jaggedness of the raw periodogram. However, these erratic movements are distributed around the true spectrum. Thus, by smoothing one can hope to get closer to the true spectrum and even detect the dominant cycle with radian equal to one. It is also clear that by smoothing over a too large range, in the extreme over all frequencies, no cycle could be detected.

Figure 6.4 illustrates these considerations with real life data by estimating the spectral density of quarterly growth rates of real investment in constructions for the Swiss economy using alternative weighting functions. To obtain a better graphical resolution we have plotted the estimates on a logarithmic scale. All three estimates show a peak (local maximum) at the frequency $\lambda = \frac{\pi}{2}$. This corresponds to a wave with a period of one year. The estimator with a comparably wide frequency band (dotted line) smoothes the minimum $\lambda = 1$ away. The estimator with a comparable small frequency band (dashed line), on the contrary, reveals additional waves with frequencies $\lambda = 0.75$ and 0.3 which correspond to periods of approximately two, respectively five years. Whether these waves are just artifacts of the weighting function or whether there really exist cycles of that periodicity remains open.

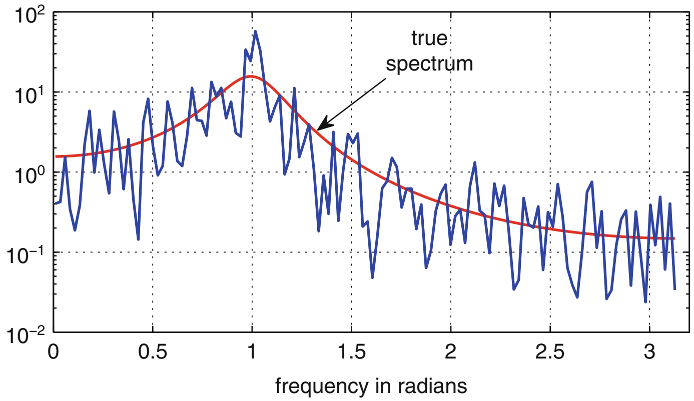

Fig. 6.3 Raw periodogram of an AR(2) process ($X_t = 0.9X_{t-1} - 0.7X_{t-2} + Z_t$ with $Z_t \sim \text{WN}(0, 1)$, $T = 200$)

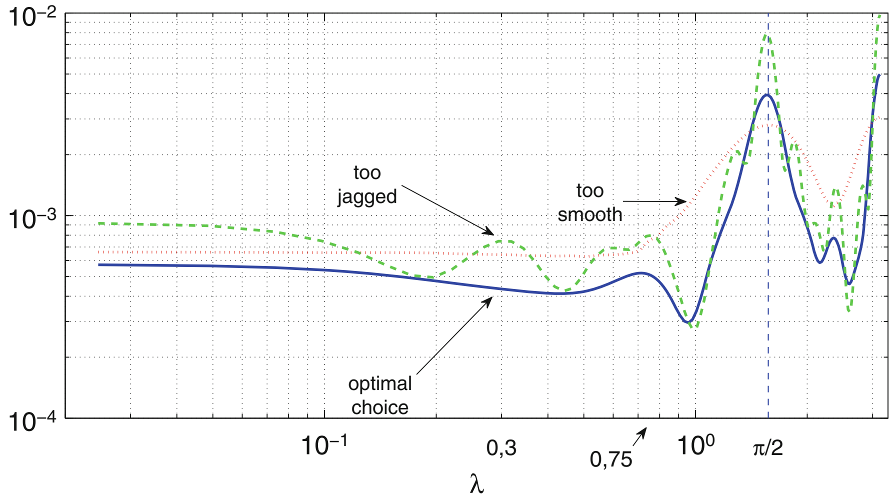

Fig. 6.4 Non-parametric direct estimates of a spectral density with alternative weighting functions

6.3.2 Parametric Estimation

An alternative to the nonparametric approaches just outlined consists in the estimation of an ARMA model and followed by deducing the spectral density from it. This approach was essentially first proposed by Yule (1927).

Theorem 6.3 (Spectral Density of ARMA Processes)**.** *Let $\{X_t\}$ be a causal ARMA(p,q) process given by $\Phi(L)X_t = \Theta(L)Z_t$ and $Z_t \sim WN(0, \sigma^2)$. Then the spectral density f_X is given by*

$$f_X(\lambda) = \frac{\sigma^2}{2\pi} \frac{\left|\Theta(e^{-\imath\lambda})\right|^2}{\left|\Phi(e^{-\imath\lambda})\right|^2}, \qquad -\pi \le \lambda \le \pi. \tag{6.11}$$

Proof. $\{X_t\}$ is generated by applying the linear filter $\Psi(L)$ with transfer function $\Psi(\mathrm{e}^{-\imath\lambda}) = \frac{\Theta(\mathrm{e}^{-\imath\lambda})}{\Phi(\mathrm{e}^{-\imath\lambda})}$ to $\{Z_t\}$ (see Sect. 6.4). Formula (6.11) is then an immediate consequence of Theorem 6.5 because the spectral density of $\{Z_t\}$ is equal to $\frac{\sigma^2}{2\pi}$. □

Remark 6.4. As the spectral density of an ARMA process $\{X_t\}$ is given by a quotient of trigonometric functions, the process is said to have a *rational spectral density*.

The spectral density of the AR(2) process $X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + Z_t$ with $Z_t \sim \mathrm{WN}(0, \sigma^2)$, for example, is then given by

$$f_X(\lambda) = \frac{\sigma^2}{2\pi(1 + \phi_1^2 + 2\phi_2 + \phi_2^2 + 2(\phi_1\phi_2 - \phi_1)\cos\lambda - 4\phi_2\cos^2\lambda)}.$$

The spectral density of an ARMA(1,1) process $X_t = \phi X_{t-1} + Z_t + \theta Z_{t-1}$ with $Z_t \sim \mathrm{WN}(0, \sigma^2)$ is

$$f_X(\lambda) = \frac{\sigma^2(1 + \theta^2 + 2\theta\cos\lambda)}{2\pi(1 + \phi^2 + 2\phi\cos\lambda)}.$$

An estimate of the spectral density is then obtained by replacing the unknown coefficients of the ARMA model by their corresponding estimates and by applying Formula (6.11) to the estimated model. Figure 6.5 compares the nonparametric to the parametric method based on an AR(4) model using the same data as in Fig. 6.4. Both methods produce similar estimates. They clearly show waves of periodicity of half a year and a year, corresponding to frequencies $\frac{\pi}{2}$ and π. The nonparametric estimate is, however, more volatile in the frequency band [0.6, 1] and around 2.5.

6.4 Linear Time-Invariant Filters

Time-invariant linear filters are an indispensable tool in time series analysis. Their objective is to eliminate or amplify waves of a particular periodicity. For example, they may be used to purge a series from seasonal movements. The seasonally adjusted time series should then reflect more strongly the business cyclical movements which are viewed to have period length between two and eight years. The spectral analysis provides just the right tools to construct and analyze such filters.

Definition 6.3. *$\{Y_t\}$ is the output of the* linear time-invariant filter *(LTF)* $\Psi = \{\psi_j, j = 0, \pm 1, \pm 2, \ldots\}$ *applied to the input $\{X_t\}$ if*

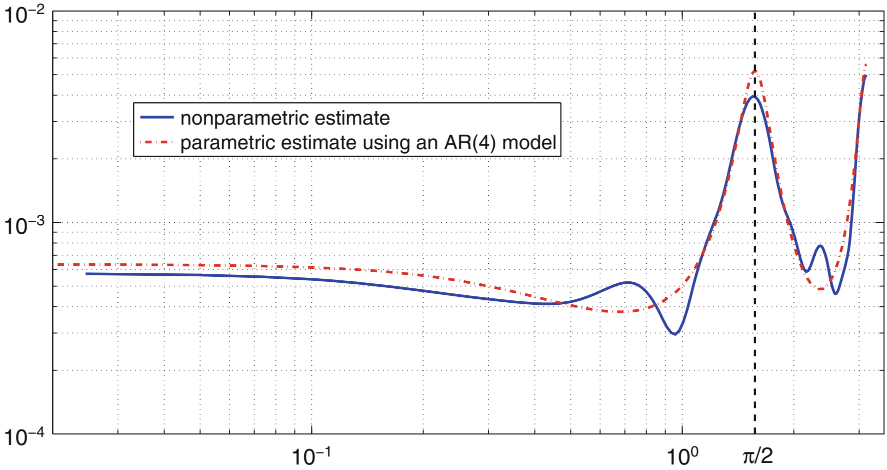

Fig. 6.5 Comparison of nonparametric and parametric estimates of the spectral density of the growth rate of investment in the construction sector

$$Y_t = \Psi(L)X_t = \sum_{j=-\infty}^{\infty} \psi_j X_{t-j} \qquad \textit{with} \sum_{j=-\infty}^{\infty} |\psi_j| < \infty.$$

The filter is called causal *or* one-sided *if* $\psi_j = 0$ *for* $j < 0$*; otherwise it is called two-sided.*

Remark 6.5. Time-invariance in this context means that the lagged process $\{Y_{t-s}\}$ is obtained for all $s \in \mathbb{Z}$ from $\{X_{t-s}\}$ by applying the same filter Ψ.

Remark 6.6. MA processes, causal AR processes and causal ARMA processes can be viewed as filtered white noise processes.

It is important to recognize that the application of a filter systematically changes the autocorrelation properties of the original time series. This may be warranted in some cases, but may lead to the "discovery" of spurious regularities which just reflect the properties of the filter. See the example of the Kuznets filter below.

Theorem 6.4 (Autocovariance Function of Filtered Process)**.** *Let* $\{X_t\}$ *be a mean-zero stationary process with autocovariance function* γ_X*. Then the filtered process* $\{Y_t\}$ *defined as*

$$Y_t = \sum_{j=-\infty}^{\infty} \psi_j X_{t-j} = \Psi(\mathrm{L})X_t$$

with $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$ *is also a mean-zero stationary process with autocovariance function* γ_Y*. Thereby the two autocovariance functions are related as follows:*

$$\gamma_Y(h) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \psi_j \psi_k \gamma_X(h+k-j), \qquad h = 0, \pm 1, \pm 2, \ldots$$

Proof. We first show the existence of the output process $\{Y_t\}$. For this end, consider the sequence of random variables $\{Y_t^{(m)}\}_{m=1,2,\ldots}$ defined as

$$Y_t^{(m)} = \sum_{j=-m}^{m} \psi_j X_{t-j}.$$

To show that the limit for $m \to \infty$ exists in the mean square sense, it is, according to Theorem C.6, enough to verify the Cauchy criterion

$$\mathbb{E}\left|Y_t^{(m)} - Y_t^{(n)}\right|^2 \longrightarrow 0, \qquad \text{for } m, n \to \infty.$$

Taking without loss of generality $m > n$, Minkowski's inequality (see Theorem C.2 or triangular inequality) leads to

$$\left(\mathbb{E}\left|\sum_{j=-m}^{m} \psi_j X_{t-j} - \sum_{j=-n}^{n} \psi_j X_{t-j}\right|^2\right)^{1/2} \le \left(\mathbb{E}\left|\sum_{j=n+1}^{m} \psi_j X_{t-j}\right|^2\right)^{1/2} + \left(\mathbb{E}\left|\sum_{j=-n-1}^{-m} \psi_j X_{t-j}\right|^2\right)^{1/2}.$$

Using the Cauchy-Bunyakovskii-Schwarz inequality and the stationarity of $\{X_t\}$, the first term on the right hand side is bounded by

$$\begin{aligned}\left(\mathbb{E}\sum_{j,k=n+1}^{m} |\psi_j X_{t-j} \psi_k X_{t-k}|\right)^{1/2} &\le \left(\sum_{j,k=n+1}^{m} |\psi_j||\psi_k| \mathbb{E}(|X_{t-j}||X_{t-k}|)\right)^{1/2} \\ &\le \left(\sum_{j,k=n+1}^{m} |\psi_j||\psi_k| (\mathbb{E}X_{t-j}^2)^{1/2} (\mathbb{E}X_{t-k}^2)^{1/2}\right)^{1/2} \\ &= \gamma_X(0)^{1/2} \sum_{j=n+1}^{m} |\psi_j|.\end{aligned}$$

As $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$ by assumption, the last term converges to zero. Thus, the limit of $\{Y_t^{(m)}\}$, $m \to \infty$, denoted by S_t, exists in the mean square sense with $\mathbb{E}S_t^2 < \infty$.

In remains to show that S_t and $\sum_{j=-\infty}^{\infty} \psi_j X_{t-j}$ are actually equal with probability one. This is established by noting that

$$\mathbb{E}\,|S_t - \Psi(\mathrm{L})X_t|^2 = \mathbb{E}\liminf_{m\to\infty}\left|S_t - \sum_{j=-m}^{m} X_{t-j}\right|^2$$
$$\leq \liminf_{m\to\infty}\mathbb{E}\left|S_t - \sum_{j=-m}^{m} X_{t-j}\right|^2 = 0$$

where use has been of Fatou's lemma.

The stationarity of $\{Y_t\}$ can be checked as follows:

$$\mathbb{E}Y_t = \lim_{m\to\infty}\sum_{j=-m}^{m}\psi_j\mathbb{E}X_{t-j} = 0,$$

$$\mathbb{E}Y_tY_{t-h} = \lim_{m\to\infty}\mathbb{E}\left[\left(\sum_{j=-m}^{m}\psi_jX_{t-j}\right)\left(\sum_{k=-m}^{m}\psi_kX_{t-h-k}\right)\right]$$
$$= \sum_{j=-\infty}^{\infty}\sum_{k=-\infty}^{\infty}\psi_j\psi_k\gamma_X(h+k-j).$$

Thus, $\mathbb{E}Y_t$ and $\mathbb{E}Y_tY_{t-h}$ are finite and independent of t. $\{Y_t\}$ is therefore stationary. □

Corollary 6.2. *If $X_t \sim \mathrm{WN}(0,\sigma^2)$ and $Y_t = \sum_{j=0}^{\infty}\psi_jX_{t-j}$ with $\sum_{j=0}^{\infty}|\psi_j| < \infty$ then the above expression for $\gamma_Y(h)$ simplifies to*

$$\gamma_Y(h) = \sigma^2\sum_{j=0}^{\infty}\psi_j\psi_{j+|h|}.$$

Remark 6.1. In the proof of the existence of $\{Y_t\}$, the assumption of the stationarity of $\{X_t\}$ can be weakened by assuming only $\sup_t \mathbb{E}X_t^2 < \infty$.

Theorem 6.5. *Under the conditions of Theorem 6.4, the spectral densities of $\{X_t\}$ and $\{Y_t\}$ are related as*

$$f_Y(\lambda) = \left|\Psi(\mathrm{e}^{-\imath\lambda})\right|^2 f_X(\lambda) = \Psi(\mathrm{e}^{\imath\lambda})\Psi(\mathrm{e}^{-\imath\lambda})f_X(\lambda)$$

where $\Psi(\mathrm{e}^{-\imath\lambda}) = \sum_{j=-\infty}^{\infty}\psi_j\mathrm{e}^{-\imath j\lambda}$. $\Psi(\mathrm{e}^{-\imath\lambda})$ is called the transfer function *of the filter.*

To understand the effect of the filter Ψ, consider the simple harmonic process $X_t = 2\cos(\lambda t) = e^{\imath\lambda t} + e^{-\imath\lambda t}$. Passing $\{X_t\}$ through the filter Ψ leads to a transformed time series $\{Y_t\}$ defined as

$$Y_t = 2\left|\Psi(e^{-\imath\lambda})\right| \cos\left(\lambda\left(t - \frac{\theta(\lambda)}{\lambda}\right)\right)$$

where $\theta(\lambda) = \arg\Psi(e^{-\imath\lambda})$. The filter therefore amplifies some frequencies by the factor $g(\lambda) = \left|\Psi(e^{-\imath\lambda})\right|$ and delays X_t by $\frac{\theta(\lambda)}{\lambda}$ periods. Thus, we have a change in amplitude given by the *amplitude gain* function $g(\lambda)$ and a phase shift given by the *phase gain* function $\theta(\lambda)$. If the gain function is bigger than one the corresponding frequency is amplified. On the other hand, if the value is smaller than one the corresponding frequency is dampened.

Examples of Filters

- First differences (changes with respect to previous period):

$$\Psi(L) = \Delta = 1 - L.$$

  The transfer function of this filter is $(1 - e^{-\imath\lambda})$ and the gain function is $2(1 - \cos\lambda)$. These functions take the value zero for $\lambda = 0$. Thus, the filter eliminates the trend which can be considered as a wave with an infinite period length.
- Change with respect to same quarter last year, assuming that the data are quarterly observations:

$$\Psi(L) = 1 - L^4.$$

  The transfer function and the gain function are $1 - e^{-4\imath\lambda}$ and $2(1 - \cos(4\lambda))$, respectively. Thus, the filter eliminates all frequencies which are multiples of $\frac{\pi}{2}$ including the zero frequency. In particular, it eliminates the trend and waves with periodicity of four quarters.
- A famous example of a filter which led to wrong conclusions is the Kuznets filter (see Sargent 1987, 273–276). Assuming yearly data, this filter is obtained from two transformations carried out in a row. The first transformation which should eliminate cyclical movements takes centered five year moving averages. The second one take centered non-overlapping first differences. Thus, the filter can be written as:

$$\Psi(L) = \underbrace{\frac{1}{5}\left(L^{-2} + L^{-1} + 1 + L + L^2\right)}_{\text{first transformation}} \underbrace{\left(L^{-5} - L^5\right)}_{\text{second transformation}} .$$

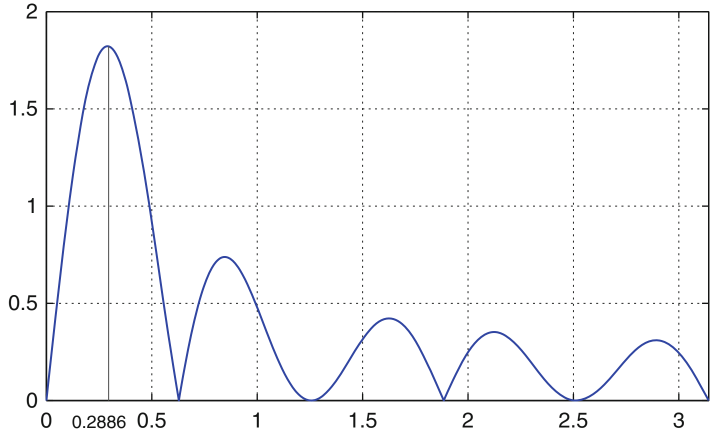

Fig. 6.6 Transfer function of the Kuznets filters

Figure 6.6 gives a plot of the transfer function of the Kuznets filter. Thereby it can be seen that all frequencies are dampened, except those around $\lambda = 0.2886$. The value $\lambda = 0.2886$ corresponds to a wave with periodicity of approximately $2\pi/0.2886 = 21.77 \approx 22$ years. Thus, as first claimed by Howrey (1968), even a filtered white noise time series would exhibit a 22 year cycle. This demonstrates that cycles of this length, related by Kuznets (1930) to demographic processes and infrastructure investment swings, may just be an artefact produced by the filter and are therefore not endorsed by the data.

6.5 Some Important Filters

6.5.1 Construction of Low- and High-Pass Filters

For some purposes it is desirable to eliminate specific frequencies. Suppose, we want to purge a time series from all movements with frequencies above λ_c, but leave those below this value unchanged. The transfer function of such an ideal low-pass filter would be:

$$\Psi(\mathrm{e}^{-\imath\lambda}) = \begin{cases} 1, \text{ for } \lambda \leq \lambda_c; \\ 0, \text{ for } \lambda > \lambda_c. \end{cases}$$

By expanding $\Psi(\mathrm{e}^{-\imath\lambda})$ into a Fourier-series $\Psi(\mathrm{e}^{-\imath\lambda}) = \sum_{j=-\infty}^{\infty} \psi_j \mathrm{e}^{-\imath j\lambda}$, it is possible to determine the appropriate filter coefficients $\{\psi_j\}$. In the case of a low-pass filter they are given by:

$$\psi_j = \frac{1}{2\pi}\int_{-\lambda_c}^{\lambda_c} e^{-\imath j\omega}\mathrm{d}\omega = \begin{cases} \frac{\lambda_c}{\pi}, & j = 0; \\ \frac{\sin(j\lambda_c)}{j\pi}, & j \neq 0. \end{cases}$$

The implementation of the filter in practice is not straightforward because only a finite number of coefficients can be used. Depending on the number of observations, the filter must be truncated such that only those ψ_j with $|j| \leq q$ are actually employed. The problem becomes more severe as one gets to the more recent observations because less future observations are available. For the most recent period even no future observation is available. This problem is usually overcome by replacing the missing future values by their corresponding forecast. Despite this remedy, the filter works best in the middle of the sample and is more and more distorted as one approaches the beginning or the end of the sample.

Analogously for low-pass filters, it is possible to construct high-pass filters. Figure 6.7 compares the transfer function of an ideal high-pass filter with two filters truncated at $q = 8$ and $q = 32$, respectively. Obviously, the transfer function with the higher q approximates the ideal filter better. In the neighborhood of the critical frequency, in our case $\pi/16$, however, the approximation remains inaccurate. This is known as the Gibbs phenomenon.

6.5.2 The Hodrick-Prescott Filter

The Hodrick-Prescott filter (HP-Filter) has gained great popularity in the macroeconomic literature, particularly in the context of the real business cycles theory. This high-pass filter is designed to eliminate the trend and cycles of high periodicity and to emphasize movements at business cycles frequencies (see Hodrick and Prescott 1980; King and Rebelo 1993; Brandner and Neusser 1992).

One way to introduce the HP-filter is to examine the problem of decomposing a time series $\{X_t\}$ additively into a growth component $\{G_t\}$ and a cyclical component $\{C_t\}$:

$$X_t = G_t + C_t.$$

This decomposition is, without further information, not unique. Following the suggestion of Whittaker (1923), the growth component should be approximated by a smooth curve. Based on this recommendation Hodrick and Prescott suggest to solve the following *restricted least-squares problem* given a sample $\{X_t\}_{t=1,\ldots,T}$:

$$\sum_{t=1}^{T} (X_t - G_t)^2 + \lambda \sum_{t=2}^{T-1} [(G_{t+1} - G_t) - (G_t - G_{t-1})]^2 \quad \longrightarrow \min_{\{G_t\}}.$$

The above objective function has two terms. The first one measures the fit of $\{G_t\}$ to the data. The closer $\{G_t\}$ is to $\{X_t\}$ the smaller this term becomes. In the limit when $G_t = X_t$ for all t, the term is minimized and equal to zero. The second term measures the smoothness of the growth component by looking at the discrete analogue to the second derivative. This term is minimized if the changes of the growth component from one period to the next are constant. This, however, implies that G_t is a linear

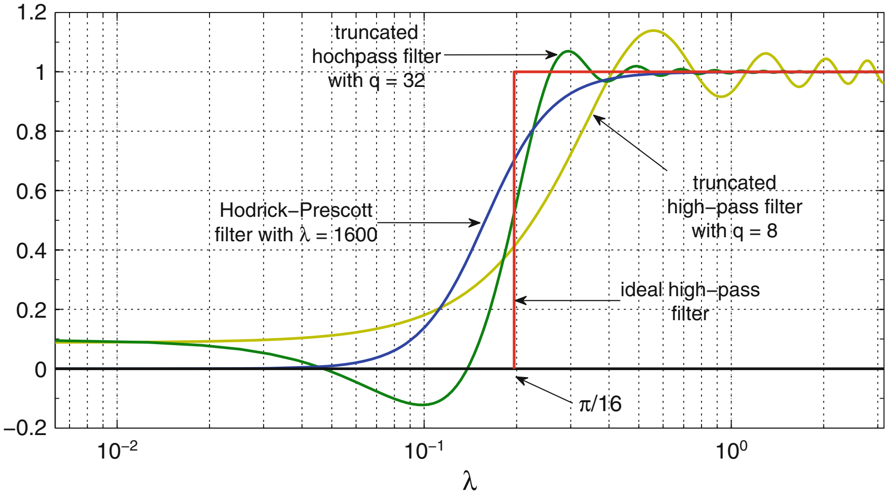

Fig. 6.7 Transfer function of HP-filter in comparison to high-pass filters

function. Thus the above objective function represents a trade-off between fitting the data and smoothness of the approximating function. This trade-off is governed by the meta-parameter λ which must be fixed a priori.

The value of λ depends on the critical frequency and on the periodicity of the data (see Uhlig and Ravn 2002, for the latter). Following the proposal by Hodrick and Prescott (1980) the following values for λ are common in the literature:

$$\lambda = \begin{cases} 6.25, & \text{yearly observations;} \\ 1600, & \text{quarterly observations;} \\ 14400, & \text{monthly observations.} \end{cases}$$

It can be shown that these choices for λ practically eliminate waves of periodicity longer than eight years. The cyclical or business cycle component is therefore composed of waves with periodicity less than eight years. Thus, the choice of λ implicitly defines the business cycle. Figure 6.7 compares the transfer function of the HP-filter to the ideal high-pass filter and two approximate high-pass filters.[7]

As an example, Fig. 6.8 displays the HP-filtered US logged GDP together with the original series in the upper panel and the implied business cycle component in the lower panel.

[7]As all filters, the HP-filter systematically distorts the properties of the time series. Harvey and Jaeger (1993) show how the blind application of the HP-filter can lead to the detection of spurious cyclical behavior.

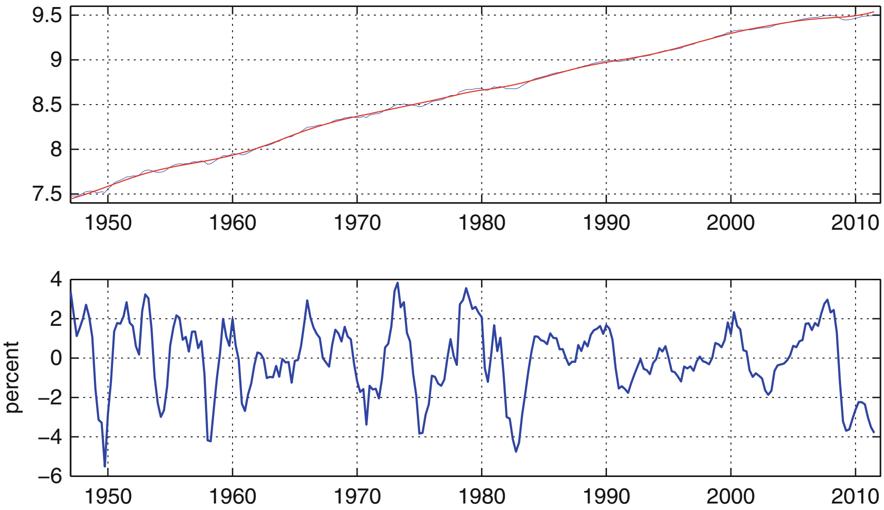

Fig. 6.8 HP-filtered US logged GDP (*upper panel*) and cyclical component (*lower panel*)

6.5.3 Seasonal Filters

Besides the elimination of trends, the removal of seasonal movements represents another important application in practice. Seasonal movements typically arise when a time series is observed several times within a year giving rise to the possibility of waves with periodicity less than twelve month in the case of monthly observations, respectively of four quarters in the case of quarterly observations. These cycles are usually considered to be of minor economic interest because they are due to systematic seasonal variations in weather conditions or holidays (Easter, for example).[8] Such variations can, for example, reduce construction activity during the winter season or production in general during holiday time. These cycles have usually quite large amplitude so that they obstruct the view to the economically and politically more important business cycles. In practice, one may therefore prefer to work with seasonally adjusted series. This means that one must remove the seasonal components from the time series in a preliminary stage of the analysis. In section, we will confine ourself to only few remarks. Comprehensive treatment of seasonality can be found in Hylleberg (1986) and Ghysels and Osborn (2001).

Two simple filters for the elimination of seasonal cycles in the case of quarterly data are given by the one-sided filter

$$\Psi(L) = (1 + L + L^2 + L^3)/4$$

[8] An exception to this view is provided by Miron (1996).

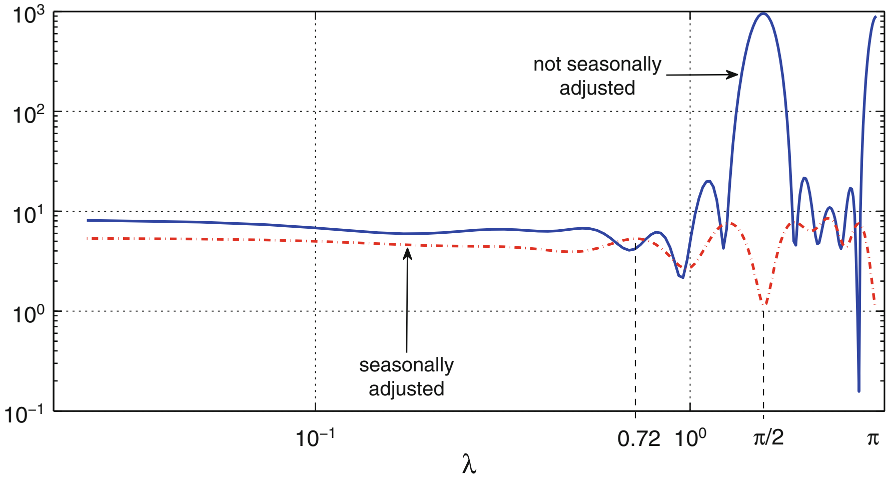

Fig. 6.9 Transfer function of growth rate of investment in the construction sector with and without seasonal adjustment

or the two-sided filter

$$\Psi(L) = \frac{1}{8}L^2 + \frac{1}{4}L + \frac{1}{4} + \frac{1}{4}L^{-1} + \frac{1}{8}L^{-2}.$$

In practice, the so-called X–11-Filter or its enhanced versions X–12 and X–13 filter developed by the United States Census Bureau are often applied. This filter is a two-sided filter which makes, in contrast to two examples above, use of all sample observations. As this filter not only adjusts for seasonality, but also corrects for outliers, a blind mechanical use is not recommended. Gómez and Maravall (1996) developed an alternative method known under the name TRAMO-SEATS. More details on the implementation of both methods can be found in Eurostat (2009).

Figure 6.9 shows the effect of seasonal adjustment using TRAMO-SEATS by looking at the corresponding transfer functions of the growth rate of construction investment. One can clearly discern how the yearly and the half-yearly waves corresponding to the frequencies $\pi/2$ and π are dampened. On the other hand, the seasonal filter weakly amplifies a cycle of frequency 0.72 corresponding to a cycle of periodicity of two years.

6.5.4 Using Filtered Data

Whether or not to use filtered, especially seasonally adjusted, data is still an ongoing debate. Although the use of unadjusted data together with a correctly specified model is clearly the best choice, there is a nonnegligible uncertainty in modeling economic time series. Thus, in practice one faces several trade-offs which must be taken into account and which may depend on the particular context (Sims 1974, 1993; Hansen and Sargent 1993). One the one hand, the use of adjusted data may disregard important information on the dynamics of the time series and introduce some biases. On the other hand, the use of unadjusted data encounters the risk of misspecification, especially because usual measures of fit may put too large emphasis on fitting the seasonal frequencies thereby neglecting other frequencies.

6.6 Exercises

Exercise 6.6.1.

(i) *Show that the process defined in Eq.* (6.4) *has an autocovariance function equal to* $\gamma(h) = \cos(\omega h)$.

(ii) *Show that the process defined in Eq.* (6.7) *has autocovariance function*

$$\gamma(h) = \sum_{j=1}^{k} \sigma_j^2 \cos(\omega_j h)$$

Exercise 6.6.2. *Compute the transfer and the gain function for the following filters:*

(i) $\Psi(\mathrm{L}) = 1 - \mathrm{L}$

(ii) $\Psi(\mathrm{L}) = 1 - \mathrm{L}^4$