Chapter 4

Spectral Analysis

4.1 Introduction

Many of the time series discussed in the previous chapters displayed strong periodic components: The sunspot numbers of Example 1.1.1, the number of trapped lynx of Example 1.1.2 and the Australian wine sales data of Example 1.4.1. Often, there is an obvious choice for the period d of this cyclical part such as an annual pattern in the wine sales. Given d, one could then proceed by removing the seasonal effects as in Section 1.4. In the first two examples it is, however, somewhat harder to determine the precise value of d. In this chapter, a general method is therefore discussed to deal with the periodic components of a time series. To complicate matters, it is usually the case that several cyclical patterns are simultaneously present in a time series. As an example recall the southern oscillation index (SOI) data which exhibits both an annual pattern and a so-called El Niño pattern.

The sine and cosine functions are the prototypes of periodic functions. They are going to be utilized here to describe cyclical behavior in time series. Before doing so, a *cycle* is defined to be one complete period of a sine or cosine function over a time interval of length 2π. Define also the *frequency*

$$\omega = \frac{1}{d}$$

as the number of cycles per observation, where d denotes the period of a time series (that is, the number of observations in a cycle). For monthly observations with an annual period, $d = 12$ and hence $\omega = 1/12 = .083$ cycles per observation. Now reconsider the process

$$X_t = R\sin(2\pi\omega t + \varphi)$$

as introduced in Example 1.2.2, using the convention $\lambda = 2\pi\omega$. To include randomness in this process, choose the amplitude R and the phase φ to be random variables. An equivalent representation of this process is given by

$$X_t = A\cos(2\pi\omega t) + B\sin(2\pi\omega t),$$

with $A = R\sin(\varphi)$ and $B = R\cos(\varphi)$ usually being independent standard normal variates. Then, $R^2 = A^2 + B^2$ is a χ-squared random variable with 2 degrees of freedom and $\varphi = \tan^{-1}(B/A)$ is

uniformly distributed on $(-\pi, \pi]$. Moreover, R and φ are independent. Choosing now the value of ω one particular periodicity can be described. To accommodate more than one, it seems natural to consider mixtures of these periodic series with multiple frequencies and amplitudes:

$$X_t = \sum_{j=1}^{m} \big[A_j \cos(2\pi\omega_j t) + B_j \sin(2\pi\omega_j t)\big], \qquad t \in \mathbb{Z},$$

where $A_1, \ldots, A_m$ and $B_1, \ldots, B_m$ are independent random variables with zero mean and variances $\sigma_1^2, \ldots, \sigma_m^2$, and $\omega_1, \ldots, \omega_m$ are *distinct* frequencies. It can be shown that $(X_t \colon t \in \mathbb{Z})$ is a weakly stationary process with lag-h ACVF

$$\gamma(h) = \sum_{j=1}^{m} \sigma_j^2 \cos(2\pi\omega_j h), \qquad h \in \mathbb{Z}.$$

The latter result yields in particular that $\gamma(0) = \sigma_1^2 + \ldots + \sigma_m^2$. The variance of X_t is consequently the sum of the component variances.

Example 4.1.1. Let $m = 2$ and choose $A_1 = B_1 = 1$, $A_2 = B_2 = 4$ to be constant as well as $\omega_1 = 1/12$ and $\omega_2 = 1/6$. This means that

$$X_t = X_t^{(1)} + X_t^{(2)} = \big[\cos(2\pi t/12) + \sin(2\pi t/12)\big] + \big[4\cos(2\pi t/6) + 4\sin(2\pi t/6)\big]$$

is the sum of two periodic components of which one exhibits an annual cycle and the other a cycle of six months. For all processes involved, realizations of $n = 48$ observations (4 years of data) are displayed in Figure 4.1. Also shown is a fourth time series plot which contains the X_t distorted by standard normal independent noise, $\tilde{X}_t$. The corresponding R code is

```
> t = 1:48
> x1 = cos(2*pi*t/12)+sin(2*pi*t/12)
> x2 = 4*cos(2*pi*t/6)+4*sin(2*pi*t/6)
> x = x1+x2
> tildex = x+rnorm(48)
```

Note that the squared amplitude of $X_t^{(1)}$ is $1^2 + 1^2 = 2$. The maximum and minimum values of $X_t^{(1)}$ are therefore $\pm\sqrt{2}$. Similarly, we obtain $\pm\sqrt{32}$ for the second component.

For a statistician it is now important to develop tools to recover the periodicities from the data. The branch of statistics concerned with this problem is called spectral analyis. The standard method in this area is based on the *periodogram* which is introduced now. Suppose for the moment that the frequency parameter $\omega_1 = 1/12$ in Example 4.1.1 is known. To obtain estimates of A_1 and B_1, one could try to run a regression using the explanatory variables $Y_{t,1} = \cos(2\pi t/12)$ or $Y_{t,2} = \sin(2\pi t/12)$ to compute the least squares estimators

$$\hat{A}_1 = \frac{\sum_{t=1}^n X_t Y_{t,1}}{\sum_{t=1}^n Y_{t,1}^2} = \frac{2}{n}\sum_{t=1}^{n} X_t \cos(2\pi t/12),$$

$$\hat{B}_1 = \frac{\sum_{t=1}^n X_t Y_{t,2}}{\sum_{t=1}^n Y_{t,2}^2} = \frac{2}{n}\sum_{t=1}^{n} X_t \sin(2\pi t/12).$$

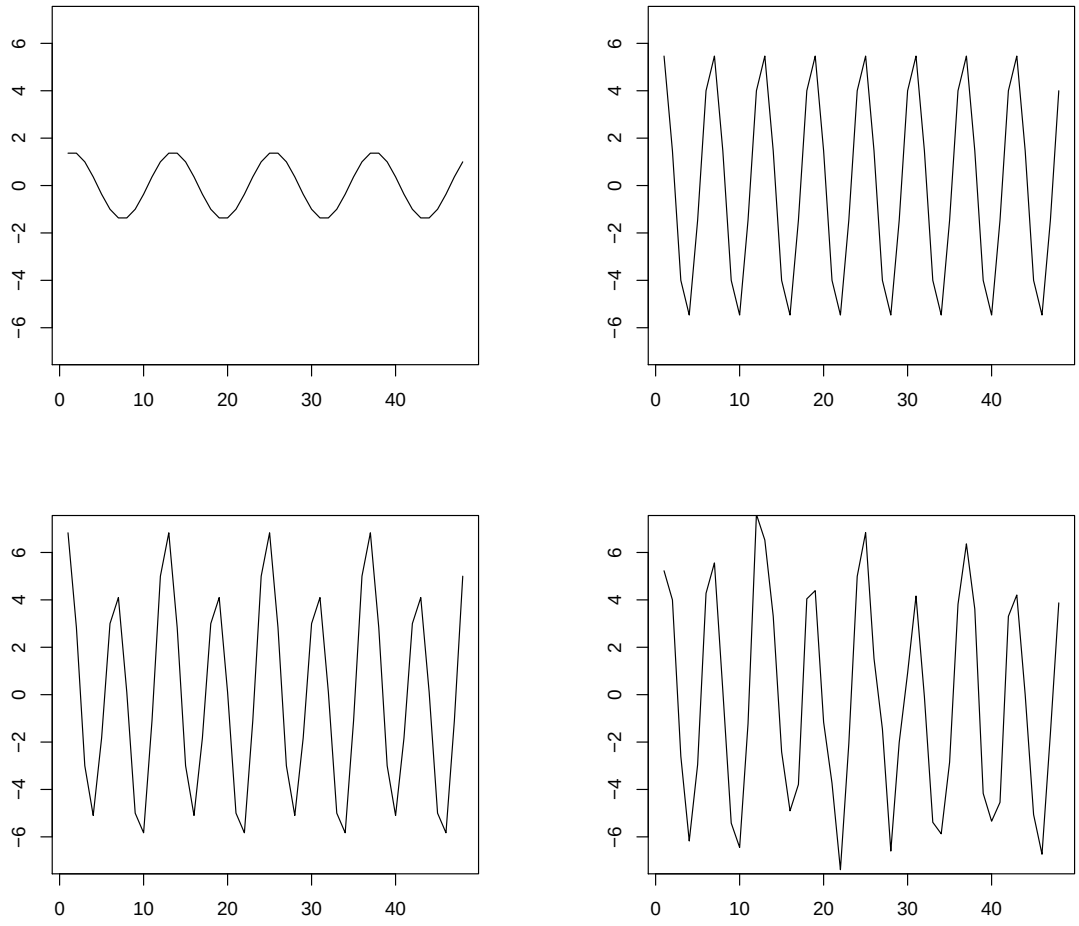

Figure 4.1: Time series plots of $(X_t^{(1)})$, $(X_t^{(2)})$, (X_t) and $(\tilde{X}_t)$.

Since, in general, the frequencies involved will not be known to the statistician prior to the data analysis, the foregoing suggests to pick a number of potential ω's, say j/n for $j = 1, \ldots, n/2$ and to run a long regression of the form

$$X_t = \sum_{j=0}^{n/2} \big[A_j \cos(2\pi jt/n) + B_j \sin(2\pi jt/n)\big]. \tag{4.1.1}$$

This leads to least squares estimates $\hat{A}_j$ and $\hat{B}_j$ of which the "significant" ones should be selected. Note that the regression in (4.1.1) is a perfect one because there are as many unknowns as variables! Note also that

$$P(j/n) = \hat{A}_j^2 + \hat{B}_j^2$$

is essentially (up to a normalization) an estimator for the correlation between the time series X_t and the corresponding sum of the periodic cosine and sine functions at frequency j/n. The collection of all $P(j/n)$, $j = 1, \ldots, n/2$, is called the *scaled periodogram*. It can be computed quickly via an algorithm known as the fast Fourier transform (FFT) which in turn is based on the discrete Fourier transform (DFT)

$$d(j/n) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} X_t \exp(-2\pi ijt/n).$$

The frequencies j/n are called the Fourier or fundamental frequencies. Since $\exp(-ix) = \cos(x) - i\sin(x)$ and $|z|^2 = z\bar{z} = (a+ib)(a-ib) = a^2 + b^2$ for any complex number $z = a + ib$, it follows that

$$I(j/n) = |d(j/n)|^2 = \frac{1}{n}\left(\sum_{t=1}^{n} X_t \cos(2\pi jt/n)\right)^2 + \frac{1}{n}\left(\sum_{t=1}^{n} X_t \sin(2\pi jt/n)\right)^2.$$

The quantity $I(j/n)$ is referred to as the *periodogram*. It follows immediately that the periodogram and the scaled periodogram are related via the identity $4I(j/n) = nP(j/n)$.

Example 4.1.2. Using the expressions and notations of Example 4.1.1, the periodogram and the scaled periodogram are computed in R as follows:

```
> t = 1:48
> I = abs(fft(x)/sqrt(48))^2
> P = 4*I/48
> f = 0:24/48
> plot(f, P[1:25], type="l")
> abline(v=1/12)
> abline(v=1/6)
```

The corresponding (scaled) periodogram for $(\tilde{X}_t)$ can be obtained in a similar fashion. The scaled periodograms are shown in the left and middle panel of Figure 4.2. The right panel displays the scaled periodogram of another version of $(\tilde{X}_t)$ in which the standard normal noise has been replaced with normal noise with variance 9. From these plots it can be seen that the six months periodicity is clearly visible in the graphs (see the dashed vertical lines at $x = 1/6$). The less pronounced annual cycle

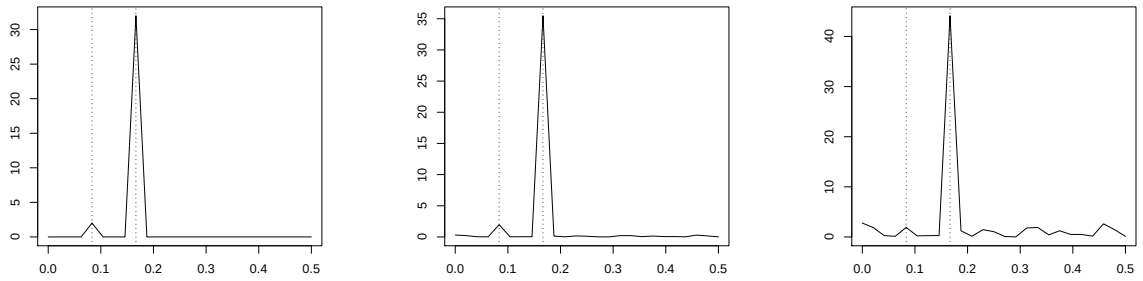

Figure 4.2: The scaled periodograms of (X_t), $(\tilde{X}_t^{(1)})$ and $(\tilde{X}_t^{(2)})$.

(vertical line at $x = 1/12$) is still visible in the first two scaled periodograms but is lost if the noise variance is increased as in the right plot. Note, however, that the y-scale is different for all three plots.

In the ideal situation that we observe the periodic component without additional contamination by noise, we can furthermore see why the periodogram may be useful in uncovering the variance decomposition from above. We have shown in the lines preceding Example 4.1.1 that the squared amplitudes of $X_t^{(1)}$ and $X_t^{(2)}$ are 2 and 32, respectively. These values are readily read from the scaled periodogram in the left panel of Figure 4.2. The contamination with noise alters these values.

In the next section, it is established that the time domain approach (based on properties of the ACVF, that is, regression on past values of the time series) and the frequency domain approach (using a periodic function approach via fundamental frequencies, that is, regression on sine and cosine functions) are equivalent. Some details are given on the spectral density (the population counterpart of the periodogram) and on properties of the periodogram itself.

4.2 The spectral density and the periodogram

The fundamental technical result which is at the core of spectral analysis states that any (weakly) stationary time series can be viewed (approximately) as a random superposition of sine and cosine functions varying at various frequencies. In other words, the regression in (4.1.1) is approximately true for all weakly stationary time series. In Chapters 1–3, it is shown how the characteristics of a stationary stochastic process can be described in terms of its ACVF $\gamma(h)$. The first goal in this section is to introduce the quantity corresponding to $\gamma(h)$ in the frequency domain.

Definition 4.2.1 (Spectral Density)**.** *If the ACVF $\gamma(h)$ of a stationary time series $(X_t)_{t\in\mathbb{Z}}$ satisfies the condition*

$$\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty,$$

then there exists a function f defined on $(-1/2, 1/2]$ such that

$$\gamma(h) = \int_{-1/2}^{1/2} \exp(2\pi i\omega h) f(\omega) d\omega, \qquad h \in \mathbb{Z},$$

and

$$f(\omega) = \sum_{h=-\infty}^{\infty} \gamma(h) \exp(-2\pi i \omega h), \qquad \omega \in (-1/2, 1/2].$$

The function f is called the spectral density of the process $(X_t \colon t \in \mathbb{Z})$.

Definition 4.2.1 (which contains a theorem part as well) establishes that each weakly stationary process can be equivalently described in terms of its ACVF or its spectral density. It also provides the formulas to compute one from the other. Time series analysis can consequently be performed either in the time domain (using $\gamma(h)$) or in the frequency domain (using $f(\omega)$). Which approach is the more suitable one cannot be decided in a general fashion but has to be reevaluated for every application of interest.

In the following, several basic properties of the spectral density are collected and evaluated f for several important examples. That the spectral density is analogous to a probability density function is established in the next proposition.

Proposition 4.2.1. *If $f(\omega)$ is the spectral density of a weakly stationary process $(X_t \colon t \in \mathbb{Z})$, then the following statements hold:*

(a) $f(\omega) \geq 0$ for all ω. This follows from the positive definiteness of $\gamma(h)$;

(b) $f(\omega) = f(-\omega)$ and $f(\omega + 1) = f(\omega)$;

(c) The variance of $(X_t \colon t \in \mathbb{Z})$ is given by

$$\gamma(0) = \int_{-1/2}^{1/2} f(\omega) d\omega.$$

Part (c) of the proposition states that the variance of a weakly stationary process is equal to the integrated spectral density over all frequencies. This property is revisited below, when a spectral analysis of variance (spectral ANOVA) will be discussed. In the following three examples are presented.

Example 4.2.1 (White Noise)**.** If $(Z_t \colon t \in \mathbb{Z}) \sim \text{WN}(0, \sigma^2)$, then its ACVF is nonzero only for $h = 0$, in which case $\gamma_Z(h) = \sigma^2$. Plugging this result into the defining equation in Definition 4.2.1 yields that

$$f_Z(\omega) = \gamma_Z(0) \exp(-2\pi i \omega 0) = \sigma^2.$$

The spectral density of a white noise sequence is therefore constant for all $\omega \in (-1/2, 1/2]$, which means that every frequency ω contributes equally to the overall spectrum. This explains the term "white" noise (in analogy to "white" light).

Example 4.2.2 (Moving Average)**.** Let $(Z_t \colon t \in \mathbb{Z}) \sim \text{WN}(0, \sigma^2)$ and define the time series $(X_t \colon t \in \mathbb{Z})$ by

$$X_t = \tfrac{1}{2} \left(Z_t + Z_{t-1} \right), \qquad t \in \mathbb{Z}.$$

It can be shown that

$$\gamma_X(h) = \frac{\sigma^2}{4} \left(2 - |h| \right), \qquad h = 0, \pm 1,$$

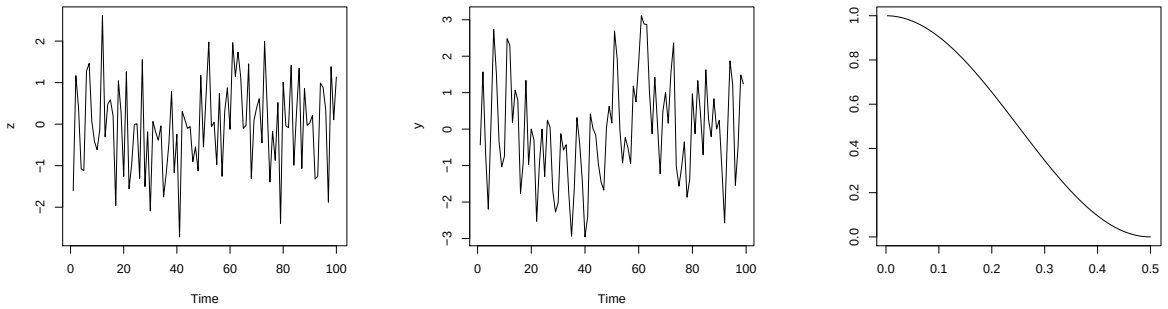

Figure 4.3: Time series plots of white noise $(Z_t\colon t \in \mathbb{Z})$ (left), two-point moving average $(X_t\colon t \in \mathbb{Z})$ (middle) and spectral density of $(X_t\colon t \in \mathbb{Z})$ (right).

and that $\gamma_X = 0$ otherwise. Therefore,

$$\begin{aligned} f_X(\omega) &= \sum_{h=-1}^{1} \gamma_X(h) \exp(2\pi i \omega h) \\ &= \frac{\sigma^2}{4} \left[\exp(-2\pi i\omega(-1)) + 2\exp(-2\pi i\omega 0) + \exp(-2\pi i\omega 1)\right] \\ &= \frac{\sigma^2}{2} \left[1 + \cos(2\pi\omega)\right] \end{aligned}$$

using that $\exp(ix) = \cos(x) + i\sin(x)$, $\cos(x) = \cos(-x)$ and $\sin(x) = -\sin(-x)$. It can be seen from the two time series plots in Figure 4.3 that the application of the two-point moving average to the white noise sequence smoothes the sample path. This is due to an attenuation of the higher frequencies which is visible in the form of the spectral density in the right panel of Figure 4.3. All plots have been obtained using Gaussian white noise with $\sigma^2 = 1$.

Example 4.2.3 (AR(2) Process)**.** Let $(X_t\colon t \in \mathbb{Z})$ be an AR(2) process which can be written in the form

$$Z_t = X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2}, \qquad t \in \mathbb{Z}.$$

In this representation, it can be seen that the ACVF γ_Z of the white noise sequence can be obtained as

$$\begin{aligned} \gamma_Z(h) &= E\left[\left(X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2}\right)\left(X_{t+h} - \phi_1 X_{t+h-1} - \phi_2 X_{t+h-2}\right)\right] \\ &= (1 + \phi_1^2 + \phi_2^2)\gamma_X(h) + (\phi_1\phi_2 - \phi_1)[\gamma_X(h+1) + \gamma_X(h-1)] \\ &\quad - \phi_2[\gamma_X(h+2) + \gamma_X(h-2)] \end{aligned}$$

Now it is known from Definition 4.2.1 that

$$\gamma_X(h) = \int_{-1/2}^{1/2} \exp(2\pi i\omega h) f_X(\omega) d\omega \qquad \text{and} \qquad \gamma_Z(h) = \int_{-1/2}^{1/2} \exp(2\pi i\omega h) f_Z(\omega) d\omega,$$

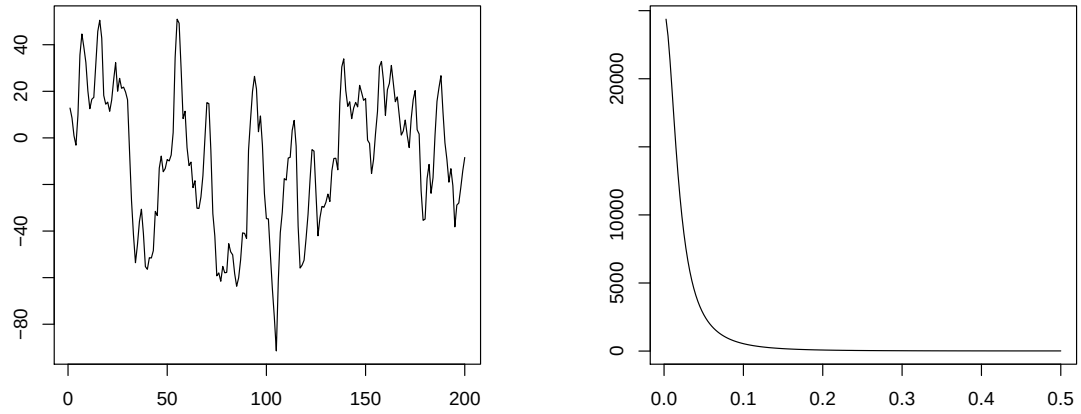

Figure 4.4: Time series plot and spectral density of the AR(2) process in Example 4.2.3.

where $f_X(\omega)$ and $f_Z(\omega)$ denote the respective spectral densities. Consequently,

$$\begin{aligned}
\gamma_Z(h) &= \int_{-1/2}^{1/2} \exp(2\pi i\omega h) f_Z(\omega) d\omega \\
&= (1+\phi_1^2+\phi_2^2)\gamma_X(h) + (\phi_1\phi_2-\phi_1)[\gamma_X(h+1)+\gamma_X(h-1)] - \phi_2[\gamma_X(h+2)+\gamma_X(h-2)] \\
&= \int_{-1/2}^{1/2} \big[(1+\phi_1^2+\phi_2^2) + (\phi_1\phi_2-\phi_1)(\exp(2\pi i\omega)+\exp(-2\pi i\omega)) \\
&\qquad -\phi_2(\exp(4\pi i\omega)+\exp(-4\pi i\omega))\big] \exp(2\pi i\omega h) f_X(\omega) d\omega \\
&= \int_{-1/2}^{1/2} \big[(1+\phi_1^2+\phi_2^2) + 2(\phi_1\phi_2-\phi_1)\cos(2\pi\omega) - 2\phi_2\cos(4\pi\omega)\big] \exp(2\pi i\omega h) f_X(\omega) d\omega.
\end{aligned}$$

The foregoing implies together with $f_Z(\omega) = \sigma^2$ that

$$\sigma^2 = \big[(1+\phi_1^2+\phi_2^2) + 2(\phi_1\phi_2-\phi_1)\cos(2\pi\omega) - 2\phi_2\cos(4\pi\omega)\big] f_X(\omega).$$

Hence, the spectral density of an AR(2) process has the form

$$f_X(\omega) = \sigma^2 \left[(1+\phi_1^2+\phi_2^2) + 2(\phi_1\phi_2-\phi_1)\cos(2\pi\omega) - 2\phi_2\cos(4\pi\omega)\right]^{-1}.$$

Figure 4.4 displays the time series plot of an AR(2) process with parameters $\phi_1 = 1.35$, $\phi_2 = -.41$ and $\sigma^2 = 89.34$. These values are very similar to the ones obtained for the recruitment series in Section 3.5. The same figure also shows the corresponding spectral density using the formula just derived.

With the contents of this Section, it has so far been established that the spectral density $f(\omega)$ is a population quantity describing the impact of the various periodic components. Next, it is verified that the periodogram $I(\omega_j)$ introduced in Section 4.1 is the sample counterpart of the spectral density.

Proposition 4.2.2. *Let $\omega_j = j/n$ denote the Fourier frequencies. If $I(\omega_j) = |d(\omega_j)|^2$ is the periodogram based on observations $X_1, \ldots, X_n$ of a weakly stationary process $(X_t \colon t \in \mathbb{Z})$, then*

$$I(\omega_j) = \sum_{h=-n+1}^{n-1} \hat{\gamma}_n(h) \exp(-2\pi i \omega_j h), \qquad j \neq 0.$$

If $j = 0$, then $I(\omega_0) = I(0) = n\bar{X}_n^2$.

Proof. Let first $j \neq 0$. Using that $\sum_{t=1}^{n} \exp(-2\pi i \omega_j t) = 0$, it follows that

$$\begin{aligned} I(\omega_j) &= \frac{1}{n} \sum_{t=1}^{n} \sum_{s=1}^{n} (X_t - \bar{X}_n)(X_s - \bar{X}_n) \exp(-2\pi i \omega_j (t-s)) \\ &= \frac{1}{n} \sum_{h=-n+1}^{n-1} \sum_{t=1}^{n-|h|} (X_{t+|h|} - \bar{X}_n)(X_t - \bar{X}_n) \exp(-2\pi i \omega_j h) \\ &= \sum_{h=-n+1}^{n-1} \hat{\gamma}_n(h) \exp(-2\pi i \omega_j h), \end{aligned}$$

which proves the first claim of the proposition. If $j = 0$, the relations $\cos(0) = 1$ and $\sin(0) = 0$ imply that $I(0) = n\bar{X}_n^2$. This completes the proof. □

More can be said about the periodogram. In fact, one can interpret spectral analysis as a spectral analysis of variance (ANOVA). To see this, let first

$$d_c(\omega_j) = \mathrm{Re}(d(\omega_j)) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} X_t \cos(2\pi\omega_j t),$$

$$d_s(\omega_j) = \mathrm{Im}(d(\omega_j)) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} X_t \sin(2\pi\omega_j t).$$

Then, $I(\omega_j) = d_c^2(\omega_j) + d_s^2(\omega_j)$. Let us now go back to the introductory example and study the process

$$X_t = A_0 + \sum_{j=1}^{m} \big[A_j \cos(2\pi\omega_j t) + B_j \sin(2\pi\omega_j t)\big],$$

where $m = (n-1)/2$ and n odd. Suppose $X_1, \ldots, X_n$ have been observed. Then, using regression techniques as before, it can be seen that $A_0 = \bar{X}_n$ and

$$A_j = \frac{2}{n} \sum_{t=1}^{n} X_t \cos(2\pi\omega_j t) = \frac{2}{\sqrt{n}} d_c(\omega_j),$$

$$B_j = \frac{2}{n} \sum_{t=1}^{n} X_t \sin(2\pi\omega_j t) = \frac{2}{\sqrt{n}} d_s(\omega_j).$$

Therefore,

$$\sum_{t=1}^{n} (X_t - \bar{X}_n)^2 = 2 \sum_{j=1}^{m} \big[d_c^2(\omega_j) + d_s^2(\omega_j)\big] = 2 \sum_{j=1}^{m} I(\omega_j)$$

and the following ANOVA table is obtained. If the underlying stochastic process exhibits a strong periodic pattern at a certain frequency, then the periodogram will most likely pick these up.

Source	df	SS	MS
ω_1	2	$2I(\omega_1)$	$I(\omega_1)$
ω_2	2	$2I(\omega_2)$	$I(\omega_2)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
ω_m	2	$2I(\omega_m)$	$I(\omega_m)$
Total	$n-1$	$\sum_{t=1}^{n}(X_t - \bar{X}_n)^2$	

Example 4.2.4. Consider the $n = 5$ data points $X_1 = 2$, $X_2 = 4$, $X_3 = 6$, $X_4 = 4$ and $X_5 = 2$, which display a cyclical but nonsinusoidal pattern. This suggests that $\omega = 1/5$ is significant and $\omega = 2/5$ is not. In R, the spectral ANOVA can be produced as follows.

```
> x = c(2,4,6,4,2), t=1:5
> cos1 = cos(2*pi*t*1/5)
> sin1 = sin(2*pi*t*1/5)
> cos2 = cos(2*pi*t*2/5)
> sin2 = sin(2*pi*t*2/5)
```

This generates the data and the independent cosine and sine variables. Now run a regression and check the ANOVA output.

```
> reg = lm(x~cos1+sin1+cos2+sin2)
> anova(reg)
```

This leads to the following output. According to previous reasoning (check the last table!), the pe-

```
Response:   x
           Df  Sum Sq  Mean Sq  F  value Pr(>F)
cos1        1  7.1777   7.1777
cos2        1  0.0223   0.0223
sin1        1  3.7889   3.7889
sin2        1  0.2111   0.2111
Residuals   0  0.0000
```

riodogram at frequency $\omega_1 = 1/5$ is given as the sum of the `cos1` and `sin1` coefficients, that is, $I(1/5) = (d_c(1/5) + d_s(1/5))/2 = (7.1777 + 3.7889)/2 = 5.4833$. Similarly, $I(2/5) = (d_c(2/5) + d_s(2/5))/2 = (0.0223 + 0.2111)/2 = 0.1167$. Note, however, that the mean squared error is computed differently in R. We can compare these values with the periodogram:

```
> abs(fft(x))^ 2/5
  [1]  64.8000000  5.4832816  0.1167184  0.1167184  5.4832816
```

The first value here is $I(0) = n\bar{X}_n^2 = 5 * (18/5)^2 = 64.8$. The second and third value are $I(1/5)$ and $I(2/5)$, respectively, while $I(3/5) = I(2/5)$ and $I(4/5) = I(1/5)$ complete the list.

In the next section, some large sample properties of the periodogram are discussed to get a better understanding of spectral analysis.

4.3 Large sample properties

Let $(X_t\colon t\in\mathbb{Z})$ be a weakly stationary time series with mean μ, absolutely summable ACVF $\gamma(h)$ and spectral density $f(\omega)$. Proceeding as in the proof of Proposition 4.2.2, one obtains

$$I(\omega_j)=\frac{1}{n}\sum_{h=-n+1}^{n-1}\sum_{t=1}^{n-|h|}(X_{t+|h|}-\mu)(X_t-\mu)\exp(-2\pi i\omega_j h),$$

provided $\omega_j\neq 0$. Using this representation, the limiting behavior of the periodogram can be established.

Proposition 4.3.1. *Let $I(\cdot)$ be the periodogram based on observations $X_1,\dots,X_n$ of a weakly stationary process $(X_t\colon t\in\mathbb{Z})$, then, for any $\omega\neq 0$,*

$$E[I(\omega_{j:n})]\to f(\omega)\qquad(n\to\infty),$$

where $\omega_{j:n}=j_n/n$ with $(j_n)_{n\in\mathbb{N}}$ chosen such that $\omega_{j:n}\to\omega$ as $n\to\infty$. If $\omega=0$, then

$$E[I(0)]-n\mu^2\to f(0)\qquad(n\to\infty).$$

Proof. There are two limits involved in the computations of the periodogram mean. First, take the limit as $n\to\infty$. This, however, requires secondly that for each n we have to work with a different set of Fourier frequencies. To adjust for this, we have introduced the notation $\omega_{j:n}$. If $\omega_j\neq 0$ is a Fourier frequency (n fixed!), then

$$E[I(\omega_j)]=\sum_{h=-n+1}^{n-1}\left(\frac{n-|h|}{n}\right)\gamma(h)\exp(-2\pi i\omega_j h).$$

Therefore ($n\to\infty$!),

$$E[I(\omega_{j:n})]\to\sum_{h=-\infty}^{\infty}\gamma(h)\exp(-2\pi i\omega h)=f(\omega),$$

thus proving the first claim. The second follows from $I(0)=n\bar{X}_n^2$ (see Proposition 4.2.2), so that $E[I(0)]-n\mu^2=n(E[\bar{X}_n^2]-\mu^2)=n\mathrm{Var}(\bar{X}_n)\to f(0)$ as $n\to\infty$ as in Chapter 2. The proof is complete. □

Proposition 4.3.1 shows that the periodogram $I(\omega)$ is asymptotically unbiased for $f(\omega)$. It is, however, inconsistent. This is implied by the following proposition which is given without proof. It is not surprising considering that each value $I(\omega_j)$ is the sum of squares of only two random variables irrespective of the sample size.

Proposition 4.3.2. *If $(X_t\colon t\in\mathbb{Z})$ is a (causal or noncausal) weakly stationary time series such that*

$$X_t=\sum_{j=-\infty}^{\infty}\psi_j Z_{t-j},\qquad t\in\mathbb{Z},$$

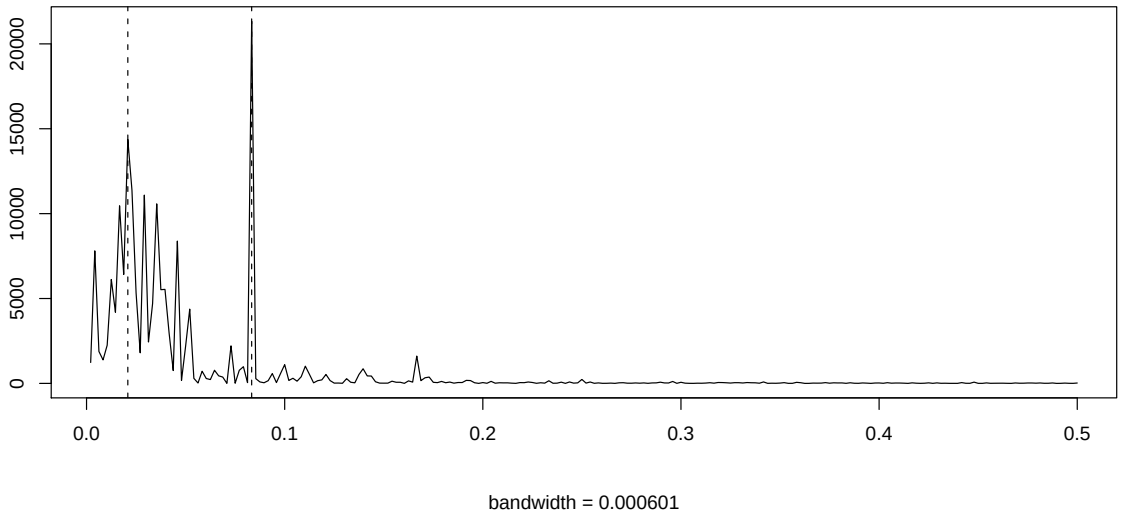

Figure 4.5: Periodogram of the recruitment data discussed in Example 4.3.1.

with $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$ *and* $(Z_t)_{t\in\mathbb{Z}} \sim$ *WN*$(0, \sigma^2)$*, then*

$$\left(\frac{2I(\omega_{1:n})}{f(\omega_1)}, \ldots, \frac{2I(\omega_{m:n})}{f(\omega_m)}\right) \xrightarrow{\mathcal{D}} (\xi_1, \ldots, \xi_m),$$

where $\omega_1, \ldots, \omega_m$ *are* m *distinct frequencies with* $\omega_{j:n} \to \omega_j$ *and* $f(\omega_j) > 0$*. The variables* $\xi_1, \ldots, \xi_m$ *are independent, identical chi-squared distributed with two degrees of freedom.*

The result of this proposition can be used to construct confidence intervals for the value of the spectral density at frequency ω. To this end, denote by $\chi_2^2(\alpha)$ the lower tail probability of the chi-squared variable ξ_j, that is,

$$P(\xi_j \leq \chi_2^2(\alpha)) = \alpha.$$

Then, Proposition 4.3.2 implies that an approximate confidence interval with level $1 - \alpha$ is given by

$$\frac{2I(\omega_{j:n})}{\chi_2^2(1-\alpha/2)} \leq f(\omega) \leq \frac{2I(\omega_{j:n})}{\chi_2^2(\alpha/2)}.$$

Proposition 4.3.2 also suggests that confidence intervals can be derived simultaneously for several frequency components. Before confidence intervals are computed for the dominant frequency of the recruitment data return for a moment to the computation of the FFT which is the basis for the periodogram usage. To ensure a quick computation time, highly composite integers n' have to be used. To achieve this in general, the length of time series is adjusted by padding the original but detrended data by adding zeroes. In R, spectral analysis is performed with the function `spec.pgram`. To find out which n' is used for your particular data, type `nextn(length(x))`, assuming that your series is in `x`.

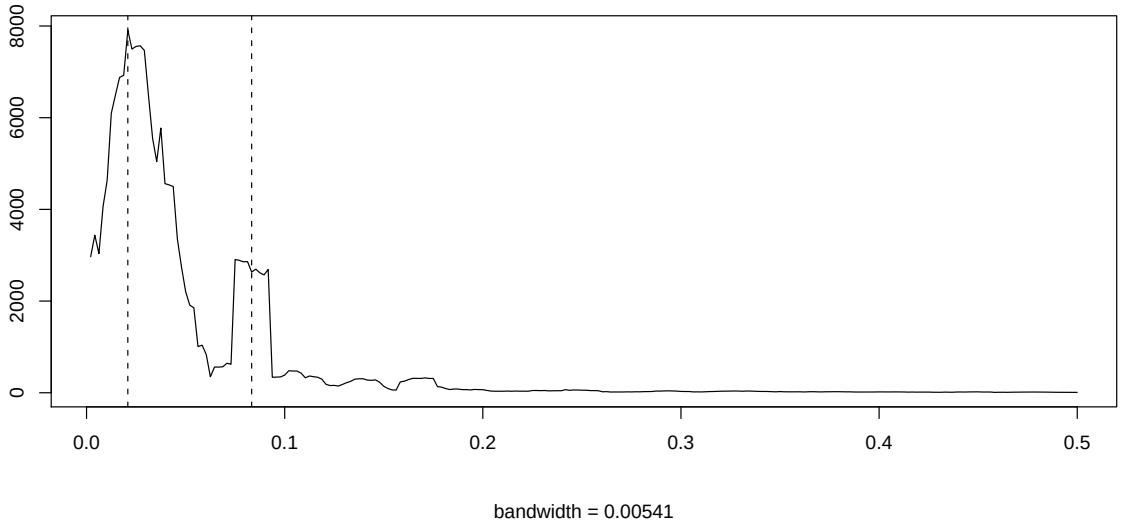

Figure 4.6: Averaged periodogram of the recruitment data discussed in Example 4.3.1.

Example 4.3.1. Figure 4.5 displays the periodogram of the recruitment data which has been discussed in Example 3.3.5. It shows a strong annual frequency component at $\omega = 1/12$ as well as several spikes in the neighborhood of the El Niño frequency $\omega = 1/48$. Higher frequency components with $\omega > .3$ are virtually absent. Even though an AR(2) model was fitted to this data in Chapter 3 to produce future values based on this fit, it is seen that the periodogram here does not validate this fit as the spectral density of an AR(2) process (as computed in Example 4.2.3) is qualitatively different. In R, the following commands can be used (`nextn(length(rec))` gives $n' = 480$ here if the recruitment data is stored in `rec` as before).

```
> rec.pgram = spec.pgram(rec, taper=0, log="no")
> abline(v=1/12, lty=2)
> abline(v=1/48, lty=2)
```

The function `spec.pgram` allows you to fine-tune the spectral analysis. For our purposes, we always use the specifications given above for the raw periodogram (`taper` allows you, for example, to exclusively look at a particular frequency band, `log` allows you to plot the log-periodogram and is the R standard).

To compute the confidence intervals for the two dominating frequencies $1/12$ and $1/48$, you can use the following R code, noting that $1/12 = 40/480$ and $1/48 = 10/480$.

```
> rec.pgram$spec[40]
  [1]  21332.94
> rec.pgram$spec[10]
  [1]  14368.42
```

```
> u = qchisq(.025, 2); l = qchisq(.975, 2)
> 2*rec.pgram$spec[40]/l
> 2*rec.pgram$spec[40]/u
> 2*rec.pgram$spec[10]/l
> 2*rec.pgram$spec[10]/u
```

Using the numerical values of this analysis, the following confidence intervals are obtained at the level $\alpha = .1$:

$$f(1/12) \in (5783.041, 842606.2) \qquad \text{and} \qquad f(1/48) \in (3895.065, 567522.5).$$

These are much too wide and alternatives to the raw periodogram are needed. These are provided, for example, by a smoothing approach which uses an averaging procedure over a band of neighboring frequencies. This can be done as follows.

```
> k = kernel("daniell",4)
> rec.ave = spec.pgram(rec, k, taper=0, log="no")
> abline(v=1/12, lty=2)
> abline(v=1/48, lty=2)
> rec.ave$bandwidth
  [1]  0.005412659
```

The resulting smoothed periodogram is shown in Figure 4.6. It is less noisy, as is expected from taking averages. More precisely, a two-sided Daniell filter with $m = 4$ was used here with $L = 2m + 1$ neighboring frequencies

$$\omega_k = \omega_j + \frac{k}{n}, \qquad k = -m, \ldots, m,$$

to compute the periodogram at $\omega_j = j/n$. The resulting plot in Figure 4.6 shows, on the other hand, that the sharp annual peak has been flattened considerably. The bandwidth reported in R can be computed as $b = L/(\sqrt{12}n)$. To compute confidence intervals one has to adjust the previously derived formula. This is done by taking changing the degrees of freedom from 2 to $df = 2Ln/n'$ (if the zeroes where appended) and leads to

$$\frac{df}{\chi^2_{df}(1-\alpha/2)} \sum_{k=-m}^{m} f\Big(\omega_j + \frac{k}{n}\Big) \le f(\omega) \le \frac{df}{\chi^2_{df}(\alpha/2)} \sum_{k=-m}^{m} f\Big(\omega_j + \frac{k}{n}\Big)$$

for $\omega \approx \omega_j$. For the recruitment data the following R code can be used:

```
> df = ceiling(rec.ave$df)
> u=qchisq(.025,df), l = qchisq(.975,df)
> df*rec.ave$spec[40]/l
> df*rec.ave$spec[40]/u
> df*rec.ave$spec[10]/l
> df*rec.ave$spec[10]/u
```

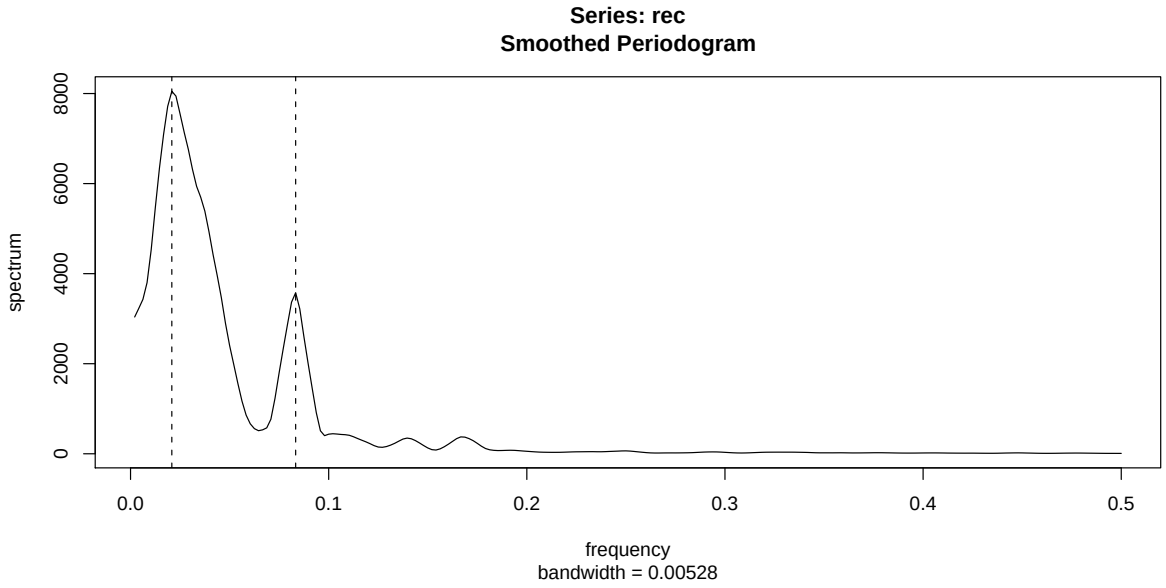

Figure 4.7: The modified Daniell periodogram of the recruitment data discussed in Example 4.3.1.

to get the confidence intervals

$$f(1/12) \in (1482.427, 5916.823) \qquad \text{and} \qquad f(1/48) \in (4452.583, 17771.64).$$

The compromise between the noisy raw periodogram and further smoothing as described here (with $L = 9$) reverses the magnitude of the $1/12$ annual frequency and the $1/48$ El Niño component. This is due to the fact that the annual peak is a very sharp one, with neighboring frequencies being basically zero. For the $1/48$ component, there are is a whole band of neighboring frequency which also contribute (the El Niño phenomenon is irregular and does only on average appear every four years). Moreover, the annual cycle is now distributed over a whole range. One way around this issue is provided by the use of other kernels such as the modified Daniell kernel given in R as `kernel("modified.daniell", c(3,3))`. This leads to the spectral density in Figure 4.7.

4.4 Linear filtering

A linear filter uses specific coefficients $(\psi_s \colon s \in \mathbb{Z})$, called the impulse response function, to transform a weakly stationary input series $(X_t \colon t \in \mathbb{Z})$ into an output series $(Y_t \colon t \in \mathbb{Z})$ via

$$Y_t = \sum_{s=-\infty}^{\infty} \psi_s X_{t-s}, \qquad t \in \mathbb{Z},$$

where $\sum_{s=-\infty}^{\infty} |\psi_s| < \infty$. Then, the frequency response function

$$\Psi(\omega) = \sum_{s=-\infty}^{\infty} \psi_s \exp(-2\pi i \omega s)$$

is well defined. Note that the two-point moving average of Example 4.2.2 and the differenced sequence ∇X_t are examples of linear filters. On the other hand, *any* causal ARMA process can be identified as a linear filter applied to a white noise sequence. Implicitly this concept was already used to compute the spectral densities in Exampels 4.2.2 and 4.2.3. To investigate this in further detail, let $\gamma_X(h)$ and $\gamma_Y(h)$ denote the ACVF of the input process $(X_t\colon t\in\mathbb{Z})$ and the output process $(Y_t\colon t\in\mathbb{Z})$, respectively, and denote by $f_X(\omega)$ and $f_Y(\omega)$ the corresponding spectral densities. The following is the main result in this section.

Theorem 4.4.1. *Under the assumptions made in this section, it holds that* $f_Y(\omega)=|\Psi(\omega)|^2 f_X(\omega)$.

Proof. First note that

$$\begin{aligned}\gamma_Y(h) &= E\big[(Y_{t+h}-\mu_Y)(Y_t-\mu_Y)\big]\\ &= \sum_{r=-\infty}^{\infty}\sum_{s=-\infty}^{\infty}\psi_r\psi_s\gamma(h-r+s)\\ &= \sum_{r=-\infty}^{\infty}\sum_{s=-\infty}^{\infty}\psi_r\psi_s\int_{-1/2}^{1/2}\exp(2\pi i\omega(h-r+s))f_X(\omega)d\omega\\ &= \int_{-1/2}^{1/2}\Big(\sum_{r=-\infty}^{\infty}\psi_r\exp(-2\pi i\omega r)\Big)\Big(\sum_{s=-\infty}^{\infty}\psi_s\exp(2\pi i\omega s)\Big)\exp(2\pi i\omega h)f_X(\omega)d\omega\\ &= \int_{-1/2}^{1/2}\exp(2\pi i\omega h)|\Psi(\omega)|^2 f_X(\omega)d\omega.\end{aligned}$$

Now identify $f_Y(\omega)=|\Psi(\omega)|^2 f_X(\omega)$, which is the assertion of the theorem. □

Theorem 4.4.1 suggests a way to compute the spectral density of a causal ARMA process. To this end, let $(Y_t\colon t\in\mathbb{Z})$ be such a causal ARMA(p,q) process satisfying $Y_t=\psi(B)Z_t$, where $(Z_t\colon t\in\mathbb{Z})\sim \mathrm{WN}(0,\sigma^2)$ and

$$\psi(z)=\frac{\theta(z)}{\phi(z)}=\sum_{s=0}^{\infty}\psi_s z^s,\qquad |z|\le 1.$$

with $\theta(z)$ and $\phi(z)$ being the moving average and autoregressive polynomial, respectively. Note that the $(\psi_s\colon s\in\mathbb{N}_0)$ can be viewed as a special impulse response function.

Corollary 4.4.1. *If* $(Y_t\colon t\in\mathbb{Z})$ *be a causal ARMA(*p,q*) process. Then, its spectral density is given by*

$$f_Y(\omega)=\sigma^2\frac{|\theta(e^{-2\pi i\omega})|^2}{|\phi(e^{-2\pi i\omega})|^2}.$$

Proof. Apply Theorem 4.4.1 with input sequence $(Z_t\colon t\in\mathbb{Z})$. Then $f_Z(\omega)=\sigma^2$, and moreover the frequency response function is

$$\Psi(\omega)=\sum_{s=0}^{\infty}\psi_s\exp(-2\pi i\omega s)=\psi(e^{-2\pi i\omega})=\frac{\theta(e^{-2\pi i\omega})}{\phi(e^{-2\pi i\omega})}.$$

Since $f_Y(\omega) = |\Psi(\omega)|^2 f_X(\omega)$, the proof is complete. □

Corollary 4.4.1 gives an easy approach to define parametric spectral density estimates for causal ARMA(p, q) processes by simply replacing the population quantities by appropriate sample counterparts. This gives the spectral density estimator

$$\hat{f}(\omega) = \hat{\sigma}_n^2 \frac{|\hat{\theta}(e^{-2\pi i\omega})|^2}{|\hat{\phi}(e^{-2\pi i\omega})|^2}.$$

Now any of the estimation techniques discussed in Section 3.5 may be applied when computing $\hat{f}(\omega)$.

4.5 Summary

In this chapter, the basic methods for frequency domain time series analysis were introduced. These are based on a regression of the given data on cosine and sine functions varying at the Fourier frequencies. On the population side, spectral densities were identified as the frequency domain counterparts of absolutely summable autocovariance functions. These are obtained from one another by the application of (inverse) Fourier transforms. On the sample side, the periodogram has been shown to be an estimator for the unknown spectral density. Since it is an incosistent estimator, various techniques have been discussed to overcome this fact. Finally, linear filters were introduced which can, for example, be used to compute spectral densities of causal ARMA processes and to derive parametric spectral density estimators other than the periodogram.