4

Spectral Analysis and Filtering

4.1 Introduction

The notion that a time series exhibits repetitive or regular behavior over time is of fundamental importance because it distinguishes time series analysis from classical statistics, which assumes complete independence over time. We have seen how dependence over time can be introduced through models that describe in detail the way certain empirical data behaves, even to the extent of producing forecasts based on the models. It is natural that models based on predicting the present as a regression on the past, such as are provided by the celebrated ARIMA or state-space forms, will be attractive to statisticians, who are trained to view nature in terms of linear models. In fact, the difference equations used to represent these kinds of models are simply the discrete versions of linear differential equations that may, in some instances, provide the ideal physical model for a certain phenomenon. An alternate version of the way nature behaves exists, however, and is based on a decomposition of an empirical series into its regular components.

In this chapter, we argue that the concept of regularity of a series can best be expressed in terms of periodic variations of the underlying phenomenon that produced the series, expressed as Fourier frequencies being driven by sines and cosines. Such a possibility was discussed in Chapters 1 and 2. From a regression point of view, we may imagine a system responding to various driving frequencies by producing linear combinations of sine and cosine functions. Expressed in these terms, the time domain approach may be thought of as regression of the present on the past, whereas the frequency domain approach may be considered as regression of the present on periodic sines and cosines.

Frequency domain approaches are the focus of this chapter. To illustrate the two methods for generating series with a single primary periodic component, consider Figure 1.9, which was generated from a simple second-order autoregressive model, and the middle and bottom panels of Figure 1.11, which were generated by adding a cosine wave with a period of 50 points to white noise. Both series exhibit strong periodic fluctuations, illustrating that both models can generate time series with regular behavior. As discussed in Example 2.8, a fundamental objective of spectral analysis is to identify the dominant frequencies in a series and to find an explanation of the system from which the measurements were derived.

Of course, the primary justification for any alternate model must lie in its potential for explaining the behavior of some empirical phenomenon. In this sense, an explanation involving only a few kinds of primary oscillations becomes simpler and more physically meaningful than a collection of parameters estimated for some selected difference equation. It is the tendency of observed data to show periodic kinds of fluctuations that justifies the use of frequency domain methods. Many of the examples in §1.2 are time series representing real phenomena that are driven by periodic components. The speech recording of the syllable *aa...hh* in Figure 1.3 contains a complicated mixture of frequencies related to the opening and closing of the glottis. Figure 1.5 shows the monthly SOI, which we later explain as a combination of two kinds of periodicities, a seasonal periodic component of 12 months and an El Niño component of about three to five years. Of fundamental interest is the return period of the El Niño phenomenon, which can have profound effects on local climate. Also of interest is whether the different periodic components of the new fish population depend on corresponding seasonal and El Niño-type oscillations. We introduce the coherence as a tool for relating the common periodic behavior of two series. Seasonal periodic components are often pervasive in economic time series; this phenomenon can be seen in the quarterly earnings series shown in Figure 1.1. In Figure 1.6, we see the extent to which various parts of the brain will respond to a periodic stimulus generated by having the subject do alternate left and right finger tapping. Figure 1.7 shows series from an earthquake and a nuclear explosion. The relative amounts of energy at various frequencies for the two phases can produce statistics, useful for discriminating between earthquakes and explosions.

In this chapter, we summarize an approach to handling correlation generated in stationary time series that begins by transforming the series to the frequency domain. This simple linear transformation essentially matches sines and cosines of various frequencies against the underlying data and serves two purposes as discussed in Example 2.8 and Example 2.9. The periodogram that was introduced in Example 2.9 has its population counterpart called the power spectrum, and its estimation is a main goal of spectral analysis. Another purpose of exploring this topic is statistical convenience resulting from the periodic components being nearly uncorrelated. This property facilitates writing likelihoods based on classical statistical methods.

An important part of analyzing data in the frequency domain, as well as the time domain, is the investigation and exploitation of the properties of the time-invariant linear filter. This special linear transformation is used similarly to linear regression in conventional statistics, and we use many of the same terms in the time series context. We have previously mentioned the coherence as a measure of the relation between two series at a given frequency, and we show later that this coherence also measures the performance of the best linear filter relating the two series. Linear filtering can also be an important step in isolating a signal embedded in noise. For example, the lower panels of Figure 1.11 contain a signal contaminated with an additive noise, whereas the upper panel contains the pure signal. It might also be appropriate to ask whether a linear filter transformation exists that could be applied to the lower panel to produce a series closer to the signal in the upper panel. The use of filtering for reducing noise will also be a part of the presentation in this chapter. We emphasize, throughout, the analogy between filtering techniques and conventional linear regression.

Many frequency scales will often coexist, depending on the nature of the problem. For example, in the Johnson & Johnson data set in Figure 1.1, the predominant frequency of oscillation is one cycle per year (4 quarters), or .25 cycles per observation. The predominant frequency in the SOI and fish populations series in Figure 1.5 is also one cycle per year, but this corresponds to 1 cycle every 12 months, or .083 cycles per observation. For simplicity, we measure frequency, ω, at cycles per time point and discuss the implications of certain frequencies in terms of the problem context. Of descriptive interest is the *period* of a time series, defined as the number of points in a cycle, i.e., $1/\omega$. Hence, the predominant period of the Johnson & Johnson series is $1/.25$ or 4 quarters per cycle, whereas the predominant period of the SOI series is 12 months per cycle.

4.2 Cyclical Behavior and Periodicity

As previously mentioned, we have already encountered the notion of periodicity in numerous examples in Chapters 1, 2 and 3. The general notion of periodicity can be made more precise by introducing some terminology. In order to define the rate at which a series oscillates, we first define a cycle as one complete period of a sine or cosine function defined over a unit time interval. As in (1.5), we consider the periodic process

$$x_t = A\cos(2\pi\omega t + \phi) \tag{4.1}$$

for $t = 0, \pm 1, \pm 2, \ldots$, where ω is a frequency index, defined in cycles per unit time with A determining the height or *amplitude* of the function and ϕ, called the *phase*, determining the start point of the cosine function. We can introduce random variation in this time series by allowing the amplitude and phase to vary randomly.

As discussed in Example 2.8, for purposes of data analysis, it is easier to use a trigonometric identity[1] and write (4.1) as

[1] $\cos(\alpha \pm \beta) = \cos(\alpha)\cos(\beta) \mp \sin(\alpha)\sin(\beta)$.

$$x_t = U_1 \cos(2\pi\omega t) + U_2 \sin(2\pi\omega t), \tag{4.2}$$

where $U_1 = A\cos\phi$ and $U_2 = -A\sin\phi$ are often taken to be normally distributed random variables. In this case, the amplitude is $A = \sqrt{U_1^2 + U_2^2}$ and the phase is $\phi = \tan^{-1}(-U_2/U_1)$. From these facts we can show that if, and only if, in (4.1), A and ϕ are independent random variables, where A^2 is chi-squared with 2 degrees of freedom, and ϕ is uniformly distributed on $(-\pi, \pi)$, then U_1 and U_2 are independent, standard normal random variables (see Problem 4.2).

The above random process is also a function of its frequency, defined by the parameter ω. The frequency is measured in cycles per unit time, or in cycles per point in the above illustration. For $\omega = 1$, the series makes one cycle per time unit; for $\omega = .50$, the series makes a cycle every two time units; for $\omega = .25$, every four units, and so on. In general, for data that occur at discrete time points will need at least two points to determine a cycle, so the highest frequency of interest is .5 cycles per point. This frequency is called the *folding frequency* and defines the highest frequency that can be seen in discrete sampling. Higher frequencies sampled this way will appear at lower frequencies, called *aliases*; an example is the way a camera samples a rotating wheel on a moving automobile in a movie, in which the wheel appears to be rotating at a different rate. For example, movies are recorded at 24 frames per second. If the camera is filming a wheel that is rotating at the rate of 24 cycles per second (or 24 Hertz), the wheel will appear to stand still (that's about 110 miles per hour in case you were wondering).

Consider a generalization of (4.2) that allows mixtures of periodic series with multiple frequencies and amplitudes,

$$x_t = \sum_{k=1}^{q} \left[U_{k1} \cos(2\pi\omega_k t) + U_{k2} \sin(2\pi\omega_k t)\right], \tag{4.3}$$

where U_{k1}, U_{k2}, for $k = 1, 2, \ldots, q$, are independent zero-mean random variables with variances σ_k^2, and the ω_k are distinct frequencies. Notice that (4.3) exhibits the process as a sum of independent components, with variance σ_k^2 for frequency ω_k. Using the independence of the Us and the trig identity in footnote 1, it is easy to show[2] (Problem 4.3) that the autocovariance function of the process is

$$\gamma(h) = \sum_{k=1}^{q} \sigma_k^2 \cos(2\pi\omega_k h), \tag{4.4}$$

and we note the autocovariance function is the sum of periodic components with weights proportional to the variances σ_k^2. Hence, x_t is a mean-zero stationary processes with variance

[2] For example, for x_t in (4.2) we have $\mathrm{cov}(x_{t+h}, x_t) = \sigma^2\{\cos(2\pi\omega[t+h])\cos(2\pi\omega t) + \sin(2\pi\omega[t+h])\sin(2\pi\omega t)\} = \sigma^2\cos(2\pi\omega h)$, noting that $\mathrm{cov}(U_1, U_2) = 0$.

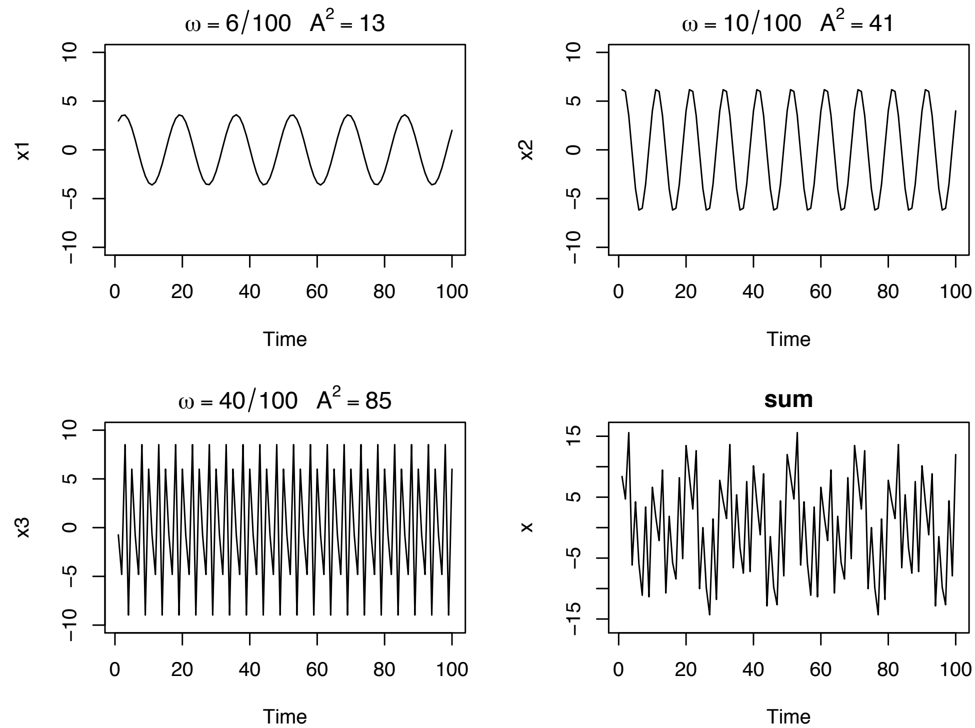

Fig. 4.1. Periodic components and their sum as described in Example 4.1.

$$\gamma(0) = E(x_t^2) = \sum_{k=1}^{q} \sigma_k^2, \tag{4.5}$$

which exhibits the overall variance as a sum of variances of each of the component parts.

Example 4.1 A Periodic Series

Figure 4.1 shows an example of the mixture (4.3) with $q = 3$ constructed in the following way. First, for $t = 1, \ldots, 100$, we generated three series

$x_{t1} = 2\cos(2\pi t\, 6/100) + 3\sin(2\pi t\, 6/100)$

$x_{t2} = 4\cos(2\pi t\, 10/100) + 5\sin(2\pi t\, 10/100)$

$x_{t3} = 6\cos(2\pi t\, 40/100) + 7\sin(2\pi t\, 40/100)$

These three series are displayed in Figure 4.1 along with the corresponding frequencies and squared amplitudes. For example, the squared amplitude of x_{t1} is $A^2 = 2^2 + 3^2 = 13$. Hence, the maximum and minimum values that x_{t1} will attain are $\pm\sqrt{13} = \pm 3.61$.

Finally, we constructed

$$x_t = x_{t1} + x_{t2} + x_{t3}$$

and this series is also displayed in Figure 4.1. We note that x_t appears to behave as some of the periodic series we saw in Chapters 1 and 2. The systematic sorting out of the essential frequency components in a time series, including their relative contributions, constitutes one of the main objectives of spectral analysis.

The R code to reproduce Figure 4.1 is

```
x1 = 2*cos(2*pi*1:100*6/100) + 3*sin(2*pi*1:100*6/100)
x2 = 4*cos(2*pi*1:100*10/100) + 5*sin(2*pi*1:100*10/100)
x3 = 6*cos(2*pi*1:100*40/100) + 7*sin(2*pi*1:100*40/100)
x = x1 + x2 + x3
par(mfrow=c(2,2))
plot.ts(x1, ylim=c(-10,10), main=expression(omega==6/100~~~A^2==13))
plot.ts(x2, ylim=c(-10,10), main=expression(omega==10/100~~~A^2==41))
plot.ts(x3, ylim=c(-10,10), main=expression(omega==40/100~~~A^2==85))
plot.ts(x,  ylim=c(-16,16), main="sum")
```

Example 4.2 The Scaled Periodogram for Example 4.1

In §2.3, Example 2.9, we introduced the periodogram as a way to discover the periodic components of a time series. Recall that the scaled periodogram is given by

$$P(j/n) = \left(\frac{2}{n}\sum_{t=1}^{n} x_t \cos(2\pi t j/n)\right)^2 + \left(\frac{2}{n}\sum_{t=1}^{n} x_t \sin(2\pi t j/n)\right)^2, \qquad (4.6)$$

and it may be regarded as a measure of the squared correlation of the data with sinusoids oscillating at a frequency of $\omega_j = j/n$, or j cycles in n time points. Recall that we are basically computing the regression of the data on the sinusoids varying at the fundamental frequencies, j/n. As discussed in Example 2.9, the periodogram may be computed quickly using the fast Fourier transform (FFT), and there is no need to run repeated regressions.

The scaled periodogram of the data, x_t, simulated in Example 4.1 is shown in Figure 4.2, and it clearly identifies the three components x_{t1}, x_{t2}, and x_{t3} of x_t. Note that

$$P(j/n) = P(1 - j/n), \quad j = 0, 1, \ldots, n-1,$$

so there is a mirroring effect at the folding frequency of $1/2$; consequently, the periodogram is typically not plotted for frequencies higher than the folding frequency. In addition, note that the heights of the scaled periodogram shown in the figure are

$$P(6/100) = 13, \quad P(10/100) = 41, \quad P(40/100) = 85,$$

$P(j/n) = P(1-j/n)$ and $P(j/n) = 0$ otherwise. These are exactly the values of the squared amplitudes of the components generated in Example 4.1. This outcome suggests that the periodogram may provide some insight into the variance components, (4.5), of a real set of data.

Assuming the simulated data, `x`, were retained from the previous example, the R code to reproduce Figure 4.2 is

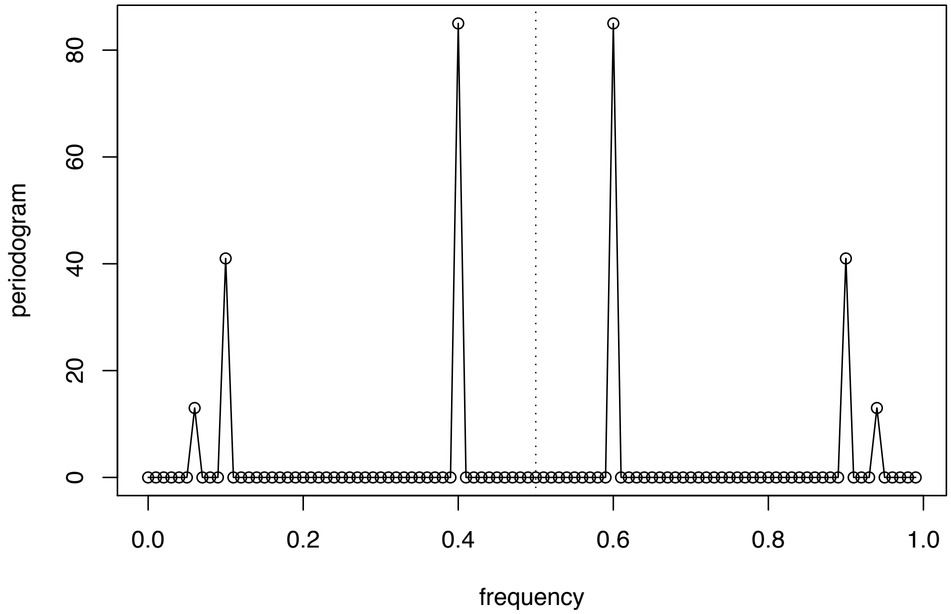

Fig. 4.2. Periodogram of the data generated in Example 4.1.

```
P = abs(2*fft(x)/100)^2;  Fr = 0:99/100
plot(Fr, P, type="o", xlab="frequency", ylab="periodogram")
```

If we consider the data x_t in Example 4.1 as a color (waveform) made up of primary colors x_{t1}, x_{t2}, x_{t3} at various strengths (amplitudes), then we might consider the periodogram as a prism that decomposes the color x_t into its primary colors (spectrum). Hence the term *spectral analysis*.

Another fact that may be of use in understanding the periodogram is that for any time series sample $x_1, \dots, x_n$, where n is odd, we may write, *exactly*

$$x_t = a_0 + \sum_{j=1}^{(n-1)/2} \left[a_j \cos(2\pi t\, j/n) + b_j \sin(2\pi t\, j/n) \right], \tag{4.7}$$

for $t = 1, \dots, n$ and suitably chosen coefficients. If n is even, the representation (4.7) can be modified by summing to $(n/2 - 1)$ and adding an additional component given by $a_{n/2} \cos(2\pi t\, 1/2) = a_{n/2}(-1)^t$. The crucial point here is that (4.7) is exact for any sample. Hence (4.3) may be thought of as an approximation to (4.7), the idea being that many of the coefficients in (4.7) may be close to zero. Recall from Example 2.9 that

$$P(j/n) = a_j^2 + b_j^2, \tag{4.8}$$

so the scaled periodogram indicates which components in (4.7) are large in magnitude and which components are small. We also saw (4.8) in Example 4.2.

The periodogram, which was introduced in Schuster (1898) and used in Schuster (1906) for studying the periodicities in the sunspot series (shown in

Figure 4.31 in the Problems section) is a sample based statistic. In Example 4.2, we discussed the fact that the periodogram may be giving us an idea of the variance components associated with each frequency, as presented in (4.5), of a time series. These variance components, however, are population parameters. The concepts of population parameters and sample statistics, as they relate to spectral analysis of time series can be generalized to cover stationary time series and that is the topic of the next section.

4.3 The Spectral Density

The idea that a time series is composed of periodic components, appearing in proportion to their underlying variances, is fundamental in the spectral representation. The result is quite technical because it involves stochastic integration; that is, integration with respect to a stochastic process. The essence of the result is that (4.3) is approximately true for any stationary time series. In other words, we have the following.

Property 4.1 Spectral Representation of a Stationary Process

In nontechnical terms, any stationary time series may be thought of, approximately, as the random superposition of sines and cosines oscillating at various frequencies.

Given that (4.3) is approximately true for all stationary time series, the next question is whether a meaningful representation for its autocovariance function, like the one displayed in (4.4), also exists. The answer is yes. The following example will help explain the result.

Example 4.3 A Periodic Stationary Process

Consider a periodic stationary random process given by (4.2), with a fixed frequency ω_0, say,

$$x_t = U_1 \cos(2\pi\omega_0 t) + U_2 \sin(2\pi\omega_0 t),$$

where U_1 and U_2 are independent zero-mean random variables with equal variance σ^2. The number of time periods needed for the above series to complete one cycle is exactly $1/\omega_0$, and the process makes exactly ω_0 cycles per point for $t = 0, \pm1, \pm2, \ldots$. It is easily shown that[3]

$$\begin{aligned}\gamma(h) = \sigma^2 \cos(2\pi\omega_0 h) &= \frac{\sigma^2}{2} e^{-2\pi i\omega_0 h} + \frac{\sigma^2}{2} e^{2\pi i\omega_0 h}\\ &= \int_{-1/2}^{1/2} e^{2\pi i\omega h} dF(\omega)\end{aligned}$$

[3] Some identities may be helpful here: $e^{i\alpha} = \cos(\alpha) + i\sin(\alpha)$ and consequently, $\cos(\alpha) = (e^{i\alpha} + e^{-i\alpha})/2$ and $\sin(\alpha) = (e^{i\alpha} - e^{-i\alpha})/2i$.

using a Riemann–Stieltjes integration, where $F(\omega)$ is the function defined by

$$F(\omega) = \begin{cases} 0 & \omega < -\omega_0, \\ \sigma^2/2 & -\omega_0 \le \omega < \omega_0, \\ \sigma^2 & \omega \ge \omega_0. \end{cases}$$

The function $F(\omega)$ behaves like a cumulative distribution function for a discrete random variable, except that $F(\infty) = \sigma^2 = \text{var}(x_t)$ instead of one. In fact, $F(\omega)$ is a cumulative distribution function, not of probabilities, but rather of variances associated with the frequency ω_0 in an analysis of variance, with $F(\infty)$ being the total variance of the process x_t. Hence, we term $F(\omega)$ the *spectral distribution function*.

A representation such as the one given in Example 4.3 always exists for a stationary process. In particular, if x_t is stationary with autocovariance $\gamma(h) = E[(x_{t+h} - \mu)(x_t - \mu)]$, then there exists a unique monotonically increasing function $F(\omega)$, called the spectral distribution function, that is bounded, with $F(-\infty) = F(-1/2) = 0$, and $F(\infty) = F(1/2) = \gamma(0)$ such that

$$\gamma(h) = \int_{-1/2}^{1/2} e^{2\pi i\omega h} \, dF(\omega). \tag{4.9}$$

A more important situation we use repeatedly is the case when the autocovariance function is absolutely summable, in which case the spectral distribution function is absolutely continuous with $dF(\omega) = f(\omega)\, d\omega$, and the representation (4.9) becomes the motivation for the property given below.

Property 4.2 The Spectral Density

If the autocovariance function, $\gamma(h)$, of a stationary process satisfies

$$\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty, \tag{4.10}$$

then it has the representation

$$\gamma(h) = \int_{-1/2}^{1/2} e^{2\pi i\omega h} f(\omega)\, d\omega \quad h = 0, \pm 1, \pm 2, \ldots \tag{4.11}$$

as the inverse transform of the spectral density, which has the representation

$$f(\omega) = \sum_{h=-\infty}^{\infty} \gamma(h) e^{-2\pi i\omega h} \quad -1/2 \le \omega \le 1/2. \tag{4.12}$$

This spectral density is the analogue of the probability density function; the fact that $\gamma(h)$ is non-negative definite ensures

$$f(\omega) \geq 0$$

for all ω. It follows immediately from (4.12) that

$$f(\omega) = f(-\omega)$$

verifying the spectral density is an even function. Because of the evenness, we will typically only plot $f(\omega)$ for $\omega \geq 0$. In addition, putting $h = 0$ in (4.11) yields

$$\gamma(0) = \mathrm{var}(x_t) = \int_{-1/2}^{1/2} f(\omega)\, d\omega,$$

which expresses the total variance as the integrated spectral density over all of the frequencies. We show later on, that a linear filter can isolate the variance in certain frequency intervals or bands.

Analogous to probability theory, $\gamma(h)$ in (4.11) is the characteristic function[4] of the spectral density $f(\omega)$ in (4.12). These facts should make it clear that, when the conditions of Property 4.2 are satisfied, *the autocovariance function, $\gamma(h)$, and the spectral density function, $f(\omega)$, contain the same information.* That information, however, is expressed in different ways. The autocovariance function expresses information in terms of lags, whereas the spectral density expresses the same information in terms of cycles. Some problems are easier to work with when considering lagged information and we would tend to handle those problems in the time domain. Nevertheless, other problems are easier to work with when considering periodic information and we would tend to handle those problems in the spectral domain.

We note that the autocovariance function, $\gamma(h)$, in (4.11) and the spectral density, $f(\omega)$, in (4.12) are Fourier transform pairs. In particular, this means that if $f(\omega)$ and $g(\omega)$ are two spectral densities for which

$$\gamma_f(h) = \int_{-1/2}^{1/2} f(\omega) e^{2\pi i\omega h}\, d\omega = \int_{-1/2}^{1/2} g(\omega) e^{2\pi i\omega h}\, d\omega = \gamma_g(h) \tag{4.13}$$

for all $h = 0, \pm1, \pm2, \ldots,$ then

$$f(\omega) = g(\omega). \tag{4.14}$$

We also mention, at this point, that we have been focusing on the frequency ω, expressed in cycles per point rather than the more common (in statistics) alternative $\lambda = 2\pi\omega$ that would give radians per point. Finally, the absolute summability condition, (4.10), is not satisfied by (4.4), the example that we have used to introduce the idea of a spectral representation. The condition, however, is satisfied for ARMA models.

It is illuminating to examine the spectral density for the series that we have looked at in earlier discussions.

[4] If $M_X(\lambda) = E(e^{\lambda X})$ for $\lambda \in \mathbb{R}$ is the moment generating function of random variable X, then $\varphi_X(\lambda) = M_X(i\lambda)$ is the characteristic function.

Example 4.4 White Noise Series

As a simple example, consider the theoretical power spectrum of a sequence of uncorrelated random variables, w_t, with variance σ_w^2. A simulated set of data is displayed in the top of Figure 1.8. Because the autocovariance function was computed in Example 1.16 as $\gamma_w(h) = \sigma_w^2$ for $h = 0$, and zero, otherwise, it follows from (4.12), that

$$f_w(\omega) = \sigma_w^2$$

for $-1/2 \le \omega \le 1/2$. Hence the process contains equal power at all frequencies. This property is seen in the realization, which seems to contain all different frequencies in a roughly equal mix. In fact, the name white noise comes from the analogy to white light, which contains all frequencies in the color spectrum at the same level of intensity. Figure 4.3 shows a plot of the white noise spectrum for $\sigma_w^2 = 1$.

If x_t is ARMA, its spectral density can be obtained explicitly using the fact that it is a linear process, i.e., $x_t = \sum_{j=0}^{\infty} \psi_j w_{t-j}$, where $\sum_{j=0}^{\infty} |\psi_j| < \infty$. In the following property, we exhibit the form of the spectral density of an ARMA model. The proof of the property follows directly from the proof of a more general result, Property 4.7 given on page 221, by using the additional fact that $\psi(z) = \theta(z)/\phi(z)$; recall Property 3.1.

Property 4.3 The Spectral Density of ARMA

If x_t is ARMA(p, q), $\phi(B)x_t = \theta(B)w_t$, its spectral density is given by

$$f_x(\omega) = \sigma_w^2 \frac{|\theta(e^{-2\pi i\omega})|^2}{|\phi(e^{-2\pi i\omega})|^2} \tag{4.15}$$

where $\phi(z) = 1 - \sum_{k=1}^{p} \phi_k z^k$ and $\theta(z) = 1 + \sum_{k=1}^{q} \theta_k z^k$.

Example 4.5 Moving Average

As an example of a series that does not have an equal mix of frequencies, we consider a moving average model. Specifically, consider the MA(1) model given by

$$x_t = w_t + .5w_{t-1}.$$

A sample realization is shown in the top of Figure 3.2 and we note that the series has less of the higher or faster frequencies. The spectral density will verify this observation.

The autocovariance function is displayed in Example 3.4 on page 90, and for this particular example, we have

$$\gamma(0) = (1 + .5^2)\sigma_w^2 = 1.25\sigma_w^2; \quad \gamma(\pm 1) = .5\sigma_w^2; \quad \gamma(\pm h) = 0 \text{ for } h > 1.$$

Substituting this directly into the definition given in (4.12), we have

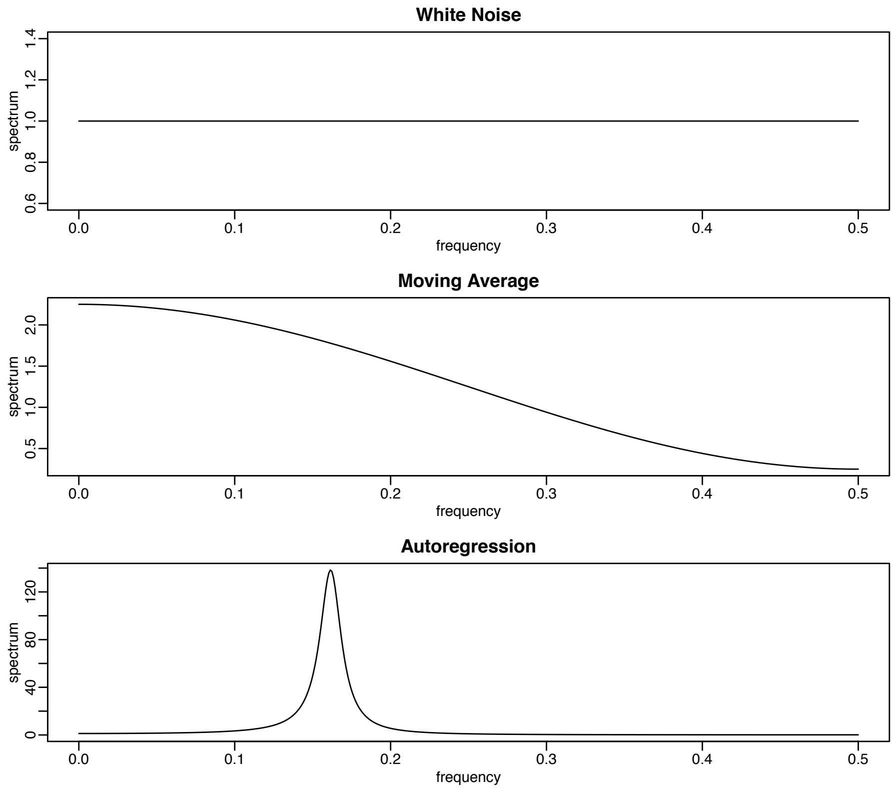

Fig. 4.3. Theoretical spectra of white noise (top), a first-order moving average (middle), and a second-order autoregressive process (bottom).

$$
\begin{aligned}
f(\omega) = \sum_{h=-\infty}^{\infty} \gamma(h)\, e^{-2\pi i\omega h} &= \sigma_w^2 \left[1.25 + .5\left(e^{-2\pi i\omega} + e^{2\pi\omega}\right)\right] \\
&= \sigma_w^2 \left[1.25 + \cos(2\pi\omega)\right].
\end{aligned}
\tag{4.16}
$$

We can also compute the spectral density using Property 4.3, which states that for an MA, $f(\omega) = \sigma_w^2 |\theta(e^{-2\pi i\omega})|^2$. Because $\theta(z) = 1 + .5z$, we have

$$
\begin{aligned}
|\theta(e^{-2\pi i\omega})|^2 = |1 + .5e^{-2\pi i\omega}|^2 &= (1 + .5e^{-2\pi i\omega})(1 + .5e^{2\pi i\omega}) \\
&= 1.25 + .5\left(e^{-2\pi i\omega} + e^{2\pi\omega}\right)
\end{aligned}
$$

which leads to agreement with (4.16).

Plotting the spectrum for $\sigma_w^2 = 1$, as in the middle of Figure 4.3, shows the lower or slower frequencies have greater power than the higher or faster frequencies.

Example 4.6 A Second-Order Autoregressive Series

We now consider the spectrum of an AR(2) series of the form

$$x_t - \phi_1 x_{t-1} - \phi_2 x_{t-2} = w_t,$$

for the special case $\phi_1 = 1$ and $\phi_2 = -.9$. Figure 1.9 on page 14 shows a sample realization of such a process for $\sigma_w = 1$. We note the data exhibit a strong periodic component that makes a cycle about every six points.

To use Property 4.3, note that $\theta(z) = 1$, $\phi(z) = 1 - z + .9z^2$ and

$$\begin{aligned}|\phi(e^{-2\pi i\omega})|^2 &= (1 - \mathrm{e}^{-2\pi i\omega} + .9\mathrm{e}^{-4\pi i\omega})(1 - \mathrm{e}^{2\pi i\omega} + .9\mathrm{e}^{4\pi i\omega})\\ &= 2.81 - 1.9(\mathrm{e}^{2\pi i\omega} + \mathrm{e}^{-2\pi i\omega}) + .9(\mathrm{e}^{4\pi i\omega} + \mathrm{e}^{-4\pi i\omega})\\ &= 2.81 - 3.8\cos(2\pi\omega) + 1.8\cos(4\pi\omega).\end{aligned}$$

Using this result in (4.15), we have that the spectral density of x_t is

$$f_x(\omega) = \frac{\sigma_w^2}{2.81 - 3.8\cos(2\pi\omega) + 1.8\cos(4\pi\omega)}.$$

Setting $\sigma_w = 1$, the bottom of Figure 4.3 displays $f_x(\omega)$ and shows a strong power component at about $\omega = .16$ cycles per point or a period between six and seven cycles per point and very little power at other frequencies. In this case, modifying the white noise series by applying the second-order AR operator has concentrated the power or variance of the resulting series in a very narrow frequency band.

The spectral density can also be obtained from first principles, without having to use Property 4.3. Because $w_t = x_t - x_{t-1} + .9x_{t-2}$ in this example, we have

$$\begin{aligned}\gamma_w(h) &= \mathrm{cov}(w_{t+h}, w_t)\\ &= \mathrm{cov}(x_{t+h} - x_{t+h-1} + .9x_{t+h-2},\ x_t - x_{t-1} + .9x_{t-2})\\ &= 2.81\gamma_x(h) - 1.9[\gamma_x(h+1) + \gamma_x(h-1)] + .9[\gamma_x(h+2) + \gamma_x(h-2)]\end{aligned}$$

Now, substituting the spectral representation (4.11) for $\gamma_x(h)$ in the above equation yields

$$\begin{aligned}\gamma_w(h) &= \int_{-1/2}^{1/2}\big[2.81 - 1.9(\mathrm{e}^{2\pi i\omega} + \mathrm{e}^{-2\pi i\omega}) + .9(\mathrm{e}^{4\pi i\omega} + \mathrm{e}^{-4\pi i\omega})\big]\mathrm{e}^{2\pi i\omega h} f_x(\omega)d\omega\\ &= \int_{-1/2}^{1/2}\big[2.81 - 3.8\cos(2\pi\omega) + 1.8\cos(4\pi\omega)\big]\mathrm{e}^{2\pi i\omega h} f_x(\omega)d\omega.\end{aligned}$$

If the spectrum of the white noise process, w_t, is $g_w(\omega)$, the uniqueness of the Fourier transform allows us to identify

$$g_w(\omega) = [2.81 - 3.8\cos(2\pi\omega) + 1.8\cos(4\pi\omega)]\, f_x(\omega).$$

But, as we have already seen, $g_w(\omega) = \sigma_w^2$, from which we deduce that

$$f_x(\omega) = \frac{\sigma_w^2}{2.81 - 3.8\cos(2\pi\omega) + 1.8\cos(4\pi\omega)}$$

is the spectrum of the autoregressive series.

To reproduce Figure 4.3, use the `spec.arma` script (see §R.1):

```
par(mfrow=c(3,1))
spec.arma(log="no", main="White Noise")
spec.arma(ma=.5, log="no", main="Moving Average")
spec.arma(ar=c(1,-.9), log="no", main="Autoregression")
```

The above examples motivate the use of the power spectrum for describing the theoretical variance fluctuations of a stationary time series. Indeed, the interpretation of the spectral density function as the variance of the time series over a given frequency band gives us the intuitive explanation for its physical meaning. The plot of the function $f(\omega)$ over the frequency argument ω can even be thought of as an analysis of variance, in which the columns or block effects are the frequencies, indexed by ω.

Example 4.7 Every Explosion has a Cause (cont)

In Example 3.3, we discussed the fact that explosive models have causal counterparts. In that example, we also indicated that it was easier to show this result in general in the spectral domain. In this example, we give the details for an AR(1) model, but the techniques used here will indicate how to generalize the result.

As in Example 3.3, we suppose that $x_t = 2x_{t-1} + w_t$, where $w_t \sim$ iid $\mathrm{N}(0, \sigma_w^2)$. Then, the spectral density of x_t is

$$f_x(\omega) = \sigma_w^2\, |1 - 2e^{-2\pi i\omega}|^{-2}. \tag{4.17}$$

But,

$$|1 - 2e^{-2\pi i\omega}| = |1 - 2e^{2\pi i\omega}| = |(2e^{2\pi i\omega})\, (\tfrac{1}{2}e^{-2\pi i\omega} - 1)| = 2\, |1 - \tfrac{1}{2}e^{-2\pi i\omega}|.$$

Thus, (4.17) can be written as

$$f_x(\omega) = \tfrac{1}{4}\sigma_w^2\, |1 - \tfrac{1}{2}e^{-2\pi i\omega}|^{-2},$$

which implies that $x_t = \frac{1}{2}x_{t-1} + v_t$, with $v_t \sim$ iid $\mathrm{N}(0, \frac{1}{4}\sigma_w^2)$ is an equivalent form of the model.

4.4 Periodogram and Discrete Fourier Transform

We are now ready to tie together the periodogram, which is the sample-based concept presented in §4.2, with the spectral density, which is the population-based concept of §4.3.

Definition 4.1 *Given data* $x_1, \dots, x_n$, *we define the* **discrete Fourier transform (DFT)** *to be*

$$d(\omega_j) = n^{-1/2} \sum_{t=1}^{n} x_t e^{-2\pi i\omega_j t} \tag{4.18}$$

for $j = 0, 1, \dots, n-1$, *where the frequencies* $\omega_j = j/n$ *are called the* **Fourier** *or* **fundamental frequencies**.