```
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 $x_1, \ldots, x_n$, 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 $j = 0, 1, \ldots, n-1$, where the frequencies $\omega_j = j/n$ are called the **Fourier** *or* **fundamental frequencies***.*

If n is a highly composite integer (i.e., it has many factors), the DFT can be computed by the fast Fourier transform (FFT) introduced in Cooley and Tukey (1965). Also, different packages scale the FFT differently, so it is a good idea to consult the documentation. R computes the DFT defined in (4.18) without the factor $n^{-1/2}$, but with an additional factor of $e^{2\pi i\omega_j}$ that can be ignored because we will be interested in the squared modulus of the DFT. Sometimes it is helpful to exploit the inversion result for DFTs which shows the linear transformation is one-to-one. For the inverse DFT we have,

$$x_t = n^{-1/2}\sum_{j=0}^{n-1} d(\omega_j)\mathrm{e}^{2\pi i\omega_j t} \tag{4.19}$$

for $t = 1, \dots, n$. The following example shows how to calculate the DFT and its inverse in R for the data set $\{1, 2, 3, 4\}$; note that R writes a complex number $z = a + ib$ as `a+bi`.

```
(dft = fft(1:4)/sqrt(4))
  [1]  5+0i  -1+1i  -1+0i  -1-1i
(idft = fft(dft, inverse=TRUE)/sqrt(4))
  [1] 1+0i  2+0i  3+0i  4+0i
(Re(idft))  # keep it real
  [1] 1  2  3  4
```

We now define the periodogram as the squared modulus[5] of the DFT.

Definition 4.2 *Given data $x_1, \dots, x_n$, we define the* **periodogram** *to be*

$$I(\omega_j) = |d(\omega_j)|^2 \tag{4.20}$$

for $j = 0, 1, 2, \dots, n-1$.

Note that $I(0) = n\bar{x}^2$, where $\bar{x}$ is the sample mean. In addition, because $\sum_{t=1}^{n}\exp(-2\pi it\frac{j}{n}) = 0$ for $j \neq 0$,[6] we can write the DFT as

$$d(\omega_j) = n^{-1/2}\sum_{t=1}^{n}(x_t - \bar{x})e^{-2\pi i\omega_j t} \tag{4.21}$$

for $j \neq 0$. Thus, for $j \neq 0$,

$$\begin{aligned} I(\omega_j) = |d(\omega_j)|^2 &= n^{-1}\sum_{t=1}^{n}\sum_{s=1}^{n}(x_t - \bar{x})(x_s - \bar{x})e^{-2\pi i\omega_j(t-s)} \\ &= n^{-1}\sum_{h=-(n-1)}^{n-1}\sum_{t=1}^{n-|h|}(x_{t+|h|} - \bar{x})(x_t - \bar{x})e^{-2\pi i\omega_j h} \\ &= \sum_{h=-(n-1)}^{n-1}\widehat{\gamma}(h)e^{-2\pi i\omega_j h} \end{aligned} \tag{4.22}$$

---

[5] Recall that if $z = a + ib$, then $\bar{z} = a - ib$, and $|z|^2 = z\bar{z} = a^2 + b^2$.

[6] $\sum_{t=1}^{n} z^t = z\frac{1-z^n}{1-z}$ for $z \neq 1$.

where we have put $h = t - s$, with $\widehat{\gamma}(h)$ as given in (1.34).[7]

Recall, $P(\omega_j) = (4/n)I(\omega_j)$ where $P(\omega_j)$ is the scaled periodogram defined in (4.6). Henceforth we will work with $I(\omega_j)$ instead of $P(\omega_j)$. In view of (4.22), the periodogram, $I(\omega_j)$, is the sample version of $f(\omega_j)$ given in (4.12). That is, we may think of the periodogram as the "sample spectral density" of x_t.

It is sometimes useful to work with the real and imaginary parts of the DFT individually. To this end, we define the following transforms.

Definition 4.3 *Given data $x_1, \ldots, x_n$, we define the* **cosine transform**

$$d_c(\omega_j) = n^{-1/2} \sum_{t=1}^{n} x_t \cos(2\pi\omega_j t) \tag{4.23}$$

and the **sine transform**

$$d_s(\omega_j) = n^{-1/2} \sum_{t=1}^{n} x_t \sin(2\pi\omega_j t) \tag{4.24}$$

where $\omega_j = j/n$ for $j = 0, 1, \ldots, n-1$.

We note that $d(\omega_j) = d_c(\omega_j) - i\, d_s(\omega_j)$ and hence

$$I(\omega_j) = d_c^2(\omega_j) + d_s^2(\omega_j). \tag{4.25}$$

We have also discussed the fact that spectral analysis can be thought of as an analysis of variance. The next example examines this notion.

Example 4.8 Spectral ANOVA

Let $x_1, \ldots, x_n$ be a sample of size n, where for ease, n is odd. Then, recalling Example 2.9 on page 67 and the discussion around (4.7) and (4.8),

$$x_t = a_0 + \sum_{j=1}^{m} \left[a_j \cos(2\pi\omega_j t) + b_j \sin(2\pi\omega_j t)\right], \tag{4.26}$$

where $m = (n-1)/2$, is exact for $t = 1, \ldots, n$. In particular, using multiple regression formulas, we have $a_0 = \bar{x}$,

$$a_j = \frac{2}{n} \sum_{t=1}^{n} x_t \cos(2\pi\omega_j t) = \frac{2}{\sqrt{n}} d_c(\omega_j)$$

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

Hence, we may write

[7] Note that (4.22) can be used to obtain $\widehat{\gamma}(h)$ by taking the inverse DFT of $I(\omega_j)$. This approach was used in Example 1.27 to obtain a two-dimensional ACF.

$$(x_t - \bar{x}) = \frac{2}{\sqrt{n}} \sum_{j=1}^{m} [d_c(\omega_j)\cos(2\pi\omega_j t) + d_s(\omega_j)\sin(2\pi\omega_j t)]$$

for $t = 1, \dots, n$. Squaring both sides and summing we obtain

$$\sum_{t=1}^{n}(x_t - \bar{x})^2 = 2\sum_{j=1}^{m}\left[d_c^2(\omega_j) + d_s^2(\omega_j)\right] = 2\sum_{j=1}^{m} I(\omega_j)$$

using the results of Problem 2.10(d) on page 81. Thus, we have partitioned the sum of squares into harmonic components represented by frequency ω_j with the periodogram, $I(\omega_j)$, being the mean square regression. This leads to the ANOVA table for n odd:

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	$n-1$	$\sum_{t=1}^{n}(x_t - \bar{x})^2$	

This decomposition means that if the data contain some strong periodic components, the periodogram values corresponding to those frequencies (or near those frequencies) will be large. On the other hand, the corresponding values of the periodogram will be small for periodic components not present in the data.

The following is an R example to help explain this concept. We consider $n = 5$ observations given by $x_1 = 1, x_2 = 2, x_3 = 3, x_4 = 2, x_5 = 1$. Note that the data complete one cycle, but not in a sinusoidal way. Thus, we should expect the $\omega_1 = 1/5$ component to be relatively large but not exhaustive, and the $\omega_2 = 2/5$ component to be small.

```
x = c(1, 2, 3, 2, 1)
c1 = cos(2*pi*1:5*1/5);  s1 = sin(2*pi*1:5*1/5)
c2 = cos(2*pi*1:5*2/5);  s2 = sin(2*pi*1:5*2/5)
omega1 = cbind(c1, s1);  omega2 = cbind(c2, s2)
anova(lm(x~omega1+omega2))    # ANOVA Table
```

```
              Df    Sum Sq   Mean Sq
  omega1       2   2.74164   1.37082
  omega2       2    .05836    .02918
  Residuals    0    .00000
```

```
abs(fft(x))^2/5         # the periodogram (as a check)
```

```
  [1] 16.2  1.37082  .029179  .029179  1.37082
   #   I(0)   I(1/5)   I(2/5)   I(3/5)   I(4/5)
```

Note that $\bar{x} = 1.8$, and $I(0) = 16.2 = 5 \times 1.8^2 (= n\bar{x}^2)$. Also, note that

$$I(1/5) = 1.37082 = \texttt{Mean Sq}(\omega_1) \quad \text{and} \quad I(2/5) = .02918 = \texttt{Mean Sq}(\omega_2)$$

and $I(j/5) = I(1-j/5)$, for $j = 3, 4$. Finally, we note that the sum of squares associated with the residuals (SSE) is zero, indicating an exact fit.

We are now ready to present some large sample properties of the periodogram. First, let μ be the mean of a stationary process x_t with absolutely summable autocovariance function $\gamma(h)$ and spectral density $f(\omega)$. We can use the same argument as in (4.22), replacing $\bar{x}$ by μ in (4.21), to write

$$I(\omega_j) = n^{-1} \sum_{h=-(n-1)}^{n-1} \sum_{t=1}^{n-|h|} (x_{t+|h|} - \mu)(x_t - \mu) e^{-2\pi i \omega_j h} \tag{4.27}$$

where ω_j is a non-zero fundamental frequency. Taking expectation in (4.27) we obtain

$$E\left[I(\omega_j)\right] = \sum_{h=-(n-1)}^{n-1} \left(\frac{n-|h|}{n}\right) \gamma(h) e^{-2\pi i \omega_j h}. \tag{4.28}$$

For any given $\omega \neq 0$, choose a sequence of fundamental frequencies $\omega_{j:n} \to \omega$[8] from which it follows by (4.28) that, as $n \to \infty$[9]

$$E\left[I(\omega_{j:n})\right] \to f(\omega) = \sum_{h=-\infty}^{\infty} \gamma(h) e^{-2\pi i h \omega}. \tag{4.29}$$

In other words, under absolute summability of $\gamma(h)$, the spectral density is the long-term average of the periodogram.

To examine the asymptotic distribution of the periodogram, we note that if x_t is a normal time series, the sine and cosine transforms will also be jointly normal, because they are linear combinations of the jointly normal random variables $x_1, x_2, \ldots, x_n$. In that case, the assumption that the covariance function satisfies the condition

$$\theta = \sum_{h=-\infty}^{\infty} |h||\gamma(h)| < \infty \tag{4.30}$$

is enough to obtain simple large sample approximations for the variances and covariances. Using the same argument used to develop (4.28) we have

$$\operatorname{cov}[d_c(\omega_j), d_c(\omega_k)] = n^{-1} \sum_{s=1}^{n} \sum_{t=1}^{n} \gamma(s-t) \cos(2\pi\omega_j s) \cos(2\pi\omega_k t), \tag{4.31}$$

$$\operatorname{cov}[d_c(\omega_j), d_s(\omega_k)] = n^{-1} \sum_{s=1}^{n} \sum_{t=1}^{n} \gamma(s-t) \cos(2\pi\omega_j s) \sin(2\pi\omega_k t), \tag{4.32}$$

---

8 By this we mean $\omega_{j:n} = j_n/n$, where $\{j_n\}$ is a sequence of integers chosen so that j_n/n is the closest Fourier frequency to ω; consequently, $|j_n/n - \omega| \leq \frac{1}{2n}$.

9 From Definition 4.2 we have $I(0) = n\bar{x}^2$, so the analogous result of (4.29) for the case $\omega = 0$ is $E[I(0)] - n\mu^2 = n \operatorname{var}(\bar{x}) \to f(0)$ as $n \to \infty$.

and

$$\operatorname{cov}[d_s(\omega_j), d_s(\omega_k)] = n^{-1} \sum_{s=1}^{n} \sum_{t=1}^{n} \gamma(s-t) \sin(2\pi\omega_j s) \sin(2\pi\omega_k t), \tag{4.33}$$

where the variance terms are obtained by setting $\omega_j = \omega_k$ in (4.31) and (4.33). In Appendix C, §C.2, we show the terms in (4.31)-(4.33) have interesting properties under assumption (4.30), namely, for $\omega_j, \omega_k \neq 0$ or $1/2$,

$$\operatorname{cov}[d_c(\omega_j), d_c(\omega_k)] = \begin{cases} f(\omega_j)/2 + \varepsilon_n & \omega_j = \omega_k, \\ \varepsilon_n & \omega_j \neq \omega_k, \end{cases} \tag{4.34}$$

$$\operatorname{cov}[d_s(\omega_j), d_s(\omega_k)] = \begin{cases} f(\omega_j)/2 + \varepsilon_n & \omega_j = \omega_k, \\ \varepsilon_n & \omega_j \neq \omega_k, \end{cases} \tag{4.35}$$

and

$$\operatorname{cov}[d_c(\omega_j), d_s(\omega_k)] = \varepsilon_n, \tag{4.36}$$

where the error term ε_n in the approximations can be bounded,

$$|\varepsilon_n| \leq \theta/n, \tag{4.37}$$

and θ is given by (4.30). If $\omega_j = \omega_k = 0$ or $1/2$ in (4.34), the multiplier $1/2$ disappears; note that $d_s(0) = d_s(1/2) = 0$, so (4.35) does not apply.

Example 4.9 Covariance of Sine and Cosine Transforms

For the three-point moving average series of Example 1.9 and $n = 256$ observations, the theoretical covariance matrix of the vector $\boldsymbol{d} = (d_c(\omega_{26}), d_s(\omega_{26}), d_c(\omega_{27}), d_s(\omega_{27}))'$ is

$$\operatorname{cov}(\boldsymbol{d}) = \begin{pmatrix} .3752 & -.0009 & -.0022 & -.0010 \\ -.0009 & .3777 & -.0009 & .0003 \\ -.0022 & -.0009 & .3667 & -.0010 \\ -.0010 & .0003 & -.0010 & .3692 \end{pmatrix}.$$

The diagonal elements can be compared with half the theoretical spectral values of $\frac{1}{2}f(\omega_{26}) = .3774$ for the spectrum at frequency $\omega_{26} = 26/256$, and of $\frac{1}{2}f(\omega_{27}) = .3689$ for the spectrum at $\omega_{27} = 27/256$. Hence, the cosine and sine transforms produce nearly uncorrelated variables with variances approximately equal to one half of the theoretical spectrum. For this particular case, the uniform bound is determined from $\theta = 8/9$, yielding $|\varepsilon_{256}| \leq .0035$ for the bound on the approximation error.

If $x_t \sim \text{iid}(0, \sigma^2)$, then it follows from (4.30)-(4.36), Problem 2.10(d), and a central limit theorem[10] that

[10] If $Y_j \sim \text{iid}(0, \sigma^2)$ and $\{a_j\}$ are constants for which $\sum_{j=1}^{n} a_j^2 / \max_{1 \leq j \leq n} a_j^2 \to \infty$ as $n \to \infty$, then $\sum_{j=1}^{n} a_j Y_j \sim \text{AN}\Big(0, \sigma^2 \sum_{j=1}^{n} a_j^2\Big)$. AN is read *asymptotically normal*.

$$d_c(\omega_{j:n}) \sim \text{AN}(0, \sigma^2/2) \quad \text{and} \quad d_s(\omega_{j:n}) \sim \text{AN}(0, \sigma^2/2) \tag{4.38}$$

jointly and independently, and independent of $d_c(\omega_{k:n})$ and $d_s(\omega_{k:n})$ provided $\omega_{j:n} \to \omega_1$ and $\omega_{k:n} \to \omega_2$ where $0 < \omega_1 \neq \omega_2 < 1/2$. We note that in this case, $f_x(\omega) = \sigma^2$. In view of (4.38), it follows immediately that as $n \to \infty$,

$$\frac{2I(\omega_{j:n})}{\sigma^2} \xrightarrow{d} \chi_2^2 \quad \text{and} \quad \frac{2I(\omega_{k:n})}{\sigma^2} \xrightarrow{d} \chi_2^2 \tag{4.39}$$

with $I(\omega_{j:n})$ and $I(\omega_{k:n})$ being asymptotically independent, where χ_ν^2 denotes a chi-squared random variable with ν degrees of freedom.

Using the central limit theory of §C.2, it is fairly easy to extend the results of the iid case to the case of a linear process.

Property 4.4 Distribution of the Periodogram Ordinates

If

$$x_t = \sum_{j=-\infty}^{\infty} \psi_j w_{t-j}, \qquad \sum_{j=-\infty}^{\infty} |\psi_j| < \infty \tag{4.40}$$

where $w_t \sim iid(0, \sigma_w^2)$, *and* (4.30) *holds, then for any collection of* m *distinct frequencies* $\omega_j \in (0, 1/2)$ *with* $\omega_{j:n} \to \omega_j$

$$\frac{2I(\omega_{j:n})}{f(\omega_j)} \xrightarrow{d} \text{iid } \chi_2^2 \tag{4.41}$$

provided $f(\omega_j) > 0$, *for* $j = 1, \ldots, m$.

This result is stated more precisely in Theorem C.7 of §C.3. Other approaches to large sample normality of the periodogram ordinates are in terms of cumulants, as in Brillinger (1981), or in terms of mixing conditions, such as in Rosenblatt (1956a). Here, we adopt the approach used by Hannan (1970), Fuller (1996), and Brockwell and Davis (1991).

The distributional result (4.41) can be used to derive an approximate confidence interval for the spectrum in the usual way. Let $\chi_\nu^2(\alpha)$ denote the lower α probability tail for the chi-squared distribution with ν degrees of freedom; that is,

$$\Pr\{\chi_\nu^2 \le \chi_\nu^2(\alpha)\} = \alpha. \tag{4.42}$$

Then, an approximate $100(1-\alpha)\%$ confidence interval for the spectral density function would be of the form

$$\frac{2\, I(\omega_{j:n})}{\chi_2^2(1-\alpha/2)} \le f(\omega) \le \frac{2\, I(\omega_{j:n})}{\chi_2^2(\alpha/2)}. \tag{4.43}$$

Often, nonstationary trends are present that should be eliminated before computing the periodogram. Trends introduce extremely low frequency components in the periodogram that tend to obscure the appearance at higher frequencies. For this reason, it is usually conventional to center the data prior to a spectral analysis using either mean-adjusted data of the form $x_t - \bar{x}$ to eliminate the zero or d-c component or to use detrended data of the form $x_t - \widehat{\beta}_1 - \widehat{\beta}_2 t$ to eliminate the term that will be considered a half cycle by the spectral analysis. Note that higher order polynomial regressions in t or nonparametric smoothing (linear filtering) could be used in cases where the trend is nonlinear.

As previously indicated, it is often convenient to calculate the DFTs, and hence the periodogram, using the fast Fourier transform algorithm. The FFT utilizes a number of redundancies in the calculation of the DFT when n is highly composite; that is, an integer with many factors of $2, 3$, or 5, the best case being when $n = 2^p$ is a factor of 2. Details may be found in Cooley and Tukey (1965). To accommodate this property, we can pad the centered (or detrended) data of length n to the next highly composite integer n' by adding zeros, i.e., setting $x^c_{n+1} = x^c_{n+2} = \cdots = x^c_{n'} = 0$, where x^c_t denotes the centered data. This means that the fundamental frequency ordinates will be $\omega_j = j/n'$ instead of j/n. We illustrate by considering the periodogram of the SOI and Recruitment series, as has been given in Figure 1.5 of Chapter 1. Recall that they are monthly series and $n = 453$ months. To find n' in R, use the command `nextn(453)` to see that $n' = 480$ will be used in the spectral analyses by default [use `help(spec.pgram)` to see how to override this default].

Example 4.10 Periodogram of SOI and Recruitment Series

Figure 4.4 shows the periodograms of each series, where the frequency axis is labeled in multiples of $\Delta = 1/12$. As previously indicated, the centered data have been padded to a series of length 480. We notice a narrow-band peak at the obvious yearly (12 month) cycle, $\omega = 1\Delta = 1/12$. In addition, there is considerable power in a wide band at the lower frequencies that is centered around the four-year (48 month) cycle $\omega = \frac{1}{4}\Delta = 1/48$ representing a possible El Niño effect. This wide band activity suggests that the possible El Niño cycle is irregular, but tends to be around four years on average. We will continue to address this problem as we move to more sophisticated analyses.

Noting $\chi^2_2(.025) = .05$ and $\chi^2_2(.975) = 7.38$, we can obtain approximate 95% confidence intervals for the frequencies of interest. For example, the periodogram of the SOI series is $I_S(1/12) = .97$ at the yearly cycle. An approximate 95% confidence interval for the spectrum $f_S(1/12)$ is then

$$[2(.97)/7.38,\ 2(.97)/.05] = [.26,\ 38.4],$$

which is too wide to be of much use. We do notice, however, that the lower value of .26 is higher than any other periodogram ordinate, so it is safe to say that this value is significant. On the other hand, an approximate 95% confidence interval for the spectrum at the four-year cycle, $f_S(1/48)$, is

$$[2(.05)/7.38,\ 2(.05)/.05] = [.01,\ 2.12],$$

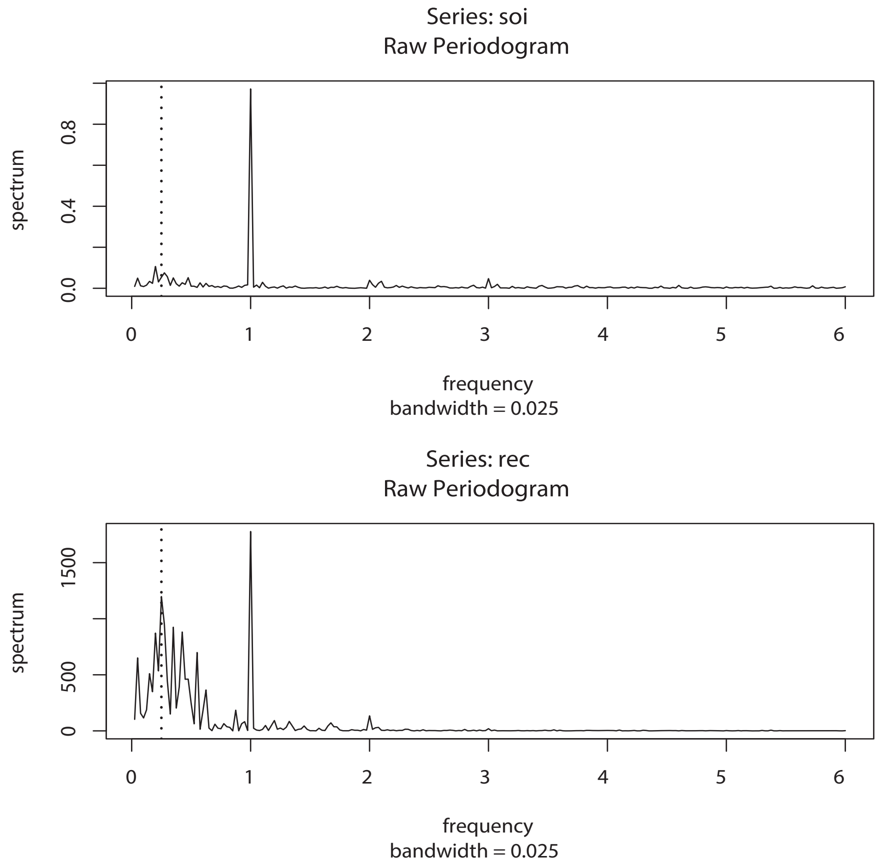

Fig. 4.4. Periodogram of SOI and Recruitment, $n = 453$ ($n' = 480$), where the frequency axis is labeled in multiples of $\Delta = 1/12$. Note the common peaks at $\omega = 1\Delta = 1/12$, or one cycle per year (12 months), and $\omega = \frac{1}{4}\Delta = 1/48$, or one cycle every four years (48 months).

which again is extremely wide, and with which we are unable to establish significance of the peak.

We now give the R commands that can be used to reproduce Figure 4.4. To calculate and graph the periodogram, we used the `spec.pgram` command in R. We note that the value of Δ is the reciprocal of the value of `frequency` used in `ts()` when making the data a time series object. If the data are not time series objects, `frequency` is set to 1. Also, we set `log="no"` because R will plot the periodogram on a $\log_{10}$ scale by default. Figure 4.4 displays a `bandwidth`. We will discuss bandwidth and tapering in the next section, so ignore these concepts for the time being.

```
require(astsa)  # needed for mvspec() - otherwise use spec.pgram()
par(mfrow=c(2,1))
soi.per = mvspec(soi, log="no")
abline(v=1/4, lty="dotted")
rec.per = mvspec(rec, log="no")
abline(v=1/4, lty="dotted")
```

The confidence intervals for the SOI series at the yearly cycle, $\omega = 1/12 = 40/480$, and the possible El Niño cycle of four years $\omega = 1/48 = 10/480$ can be computed in R as follows:

```
soi.per$spec[40]   # 0.97223; soi pgram at freq 1/12 = 40/480
soi.per$spec[10]   # 0.05372; soi pgram at freq 1/48 = 10/480
#  conf intervals  -  returned value:
U = qchisq(.025,2)     #  0.05063
L = qchisq(.975,2)     #  7.37775
2*soi.per$spec[10]/L   #  0.01456
2*soi.per$spec[10]/U   #  2.12220
2*soi.per$spec[40]/L   #  0.26355
2*soi.per$spec[40]/U   # 38.40108
```

The example above makes it clear that the periodogram as an estimator is susceptible to large uncertainties, and we need to find a way to reduce the variance. Not surprisingly, this result follows if we think about the periodogram, $I(\omega_j)$ as an estimator of the spectral density $f(\omega)$ and realize that it is the sum of squares of only two random variables for any sample size. The solution to this dilemma is suggested by the analogy with classical statistics where we look for independent random variables with the same variance and average the squares of these common variance observations. Independence and equality of variance do not hold in the time series case, but the covariance structure of the two adjacent estimators given in Example 4.9 suggests that for neighboring frequencies, these assumptions are approximately true.

4.5 Nonparametric Spectral Estimation

To continue the discussion that ended the previous section, we introduce a frequency band, $\mathcal{B}$, of $L << n$ contiguous fundamental frequencies, centered around frequency $\omega_j = j/n$, which is chosen close to a frequency of interest, ω. For frequencies of the form $\omega^* = \omega_j + k/n$, let

$$\mathcal{B} = \left\{ \omega^* \colon\ \omega_j - \frac{m}{n} \le \omega^* \le \omega_j + \frac{m}{n} \right\}, \tag{4.44}$$

where

$$L = 2m + 1 \tag{4.45}$$

is an odd number, chosen such that the spectral values in the interval $\mathcal{B}$,

$$f(\omega_j + k/n), \quad k = -m, \dots, 0, \dots, m$$

are approximately equal to $f(\omega)$.

We now define an averaged (or smoothed) periodogram as the average of the periodogram values, say,