Chapter 4

Spectral Analysis and Filtering

4.1 Introduction

In this chapter, we discuss how a given time series can be decomposed into its harmonic parts. This so-called spectral representation can be viewed as the analogue of the Fourier transform for deterministic functions. While time series analysis based on the ACVF is referred to as the *time domain* approach, the use of spectral representations is called the *frequency domain* approach. That both approaches are in fact equivalent will be shown in Section 4.2 below. Section 4.3 contains the *spectral representation theorem* which establishes that every weakly stationary process can be written as a stochastic integral with respect to an orthogonal increment process. In Sections 4.4 and 4.5 we discuss aspects of the estimation in the frequency domain.

In the context of frequency domain time series analysis, it is more convenient to work with stochastic processes possessing values in the complex plain $\mathbb{C}$. This requires an extension of the notion of weak stationarity. Suppose that $(X_t\colon t\in\mathbb{Z})$ is a complex-valued stochastic process. Then, $(X_t\colon t\in\mathbb{Z})$ is called *weakly stationary* if

- $E[X_t]=m\in\mathbb{C}$ for all $t\in\mathbb{Z}$;
- $E[|X_t|^2]=E[X_t\overline{X}_t]<\infty$ for all $t\in\mathbb{Z}$;
- $\gamma(h)=\mathrm{Cov}(X_{t+h},X_t)=E[(X_{t+h}-m)(\overline{X_t-m})]$ is a function of h only.

The ACF of $(X_t\colon t\in\mathbb{Z})$ is defined in the usual way as $\rho(h)=\gamma(h)/\gamma(0)$. We collect some of the basic properties of the complex-valued ACVF in the following remark.

Remark 4.1.1 *If γ is the ACVF of a weakly stationary complex-valued stochastic process, then it satisfies*

(i) γ is nonnegative definite, that is

$$\sum_{j,k=1}^{n} a_j\gamma(j-k)\bar{a}_k\geq 0,$$

for all $n\in\mathbb{N}$ and $a_1,\ldots,a_n\in\mathbb{C}$;

(ii) $\gamma(0)\geq 0$;

(iii) $|\gamma(h)|\leq\gamma(0)$;

(iv) γ is Hermitian, that is, $\gamma(-h) = \overline{\gamma(h)}$.
Similar statements hold for the ACF ρ.

4.2 Herglotz's Theorem and Spectral Densities

In this section, we establish the correspondence between the ACVF γ of a weakly stationary process $(X_t\colon t \in \mathbb{Z})$ and the so-called *spectral distribution function* and *spectral density function* to be defined below. At first, we give a motivating example.

Example 4.2.1 Let the process $(X_t\colon t \in \mathbb{Z})$ be given by

$$X_t = \sum_{j=1}^{n} A(\lambda_j)e^{it\lambda_j}, \qquad t \in \mathbb{Z},$$

where $-\pi < \lambda_1 < \ldots < \lambda_n = \pi$ are frequency parameters and $A(\lambda_1), \ldots, A(\lambda_n)$ are pairwise uncorrelated random variables with $E[A(\lambda_j)] = 0$ and $E[|A(\lambda_j)|^2] = \sigma_j^2$. If the process is to be real-valued, we need that $A(\lambda_n)$ is real-valued (because $\lambda_n = \pi$), and that $\lambda_j = \lambda_{n-j}$ and $A(\lambda_j) = \overline{A(\lambda_{n-j})}$ for $j = 1, \ldots, n-1$. Then, $(X_t\colon t \in \mathbb{Z})$ has zero mean and it is weakly stationary. Its ACVF is given by

$$\gamma(h) = E[X_{t+h}\overline{X}_t] = \sum_{j=1}^{n}\sum_{j=1}^{n} e^{i(t+h)\lambda_j} E\left[A(\lambda_j)\overline{A(\lambda_k)}\right] e^{-it\lambda_k} = \sum_{j=1}^{n} \sigma_j^2 e^{ih\lambda_j}.$$

If we let

$$F(\lambda) = \sum_{j:\lambda_j \leq \lambda} \sigma_j^2, \qquad \lambda \in \mathbb{R},$$

then F is a monotonically increasing, right-continuous (hence measure-defining) function satisfying $F(\lambda) = 0$ $(\lambda \leq -\pi)$ and $F(\lambda) = \sigma^2 := \sigma_1^2 + \ldots + \sigma_n^2$ $(\lambda \geq \pi)$. The function F is referred to as the *spectral distribution function* of the process $(X_t\colon t \in \mathbb{Z})$. With the use of F we can write the ACVF as the Stieltjes integral

$$\gamma(h) = \int_{(-\pi,\pi]} e^{ih\lambda} dF(\lambda), \qquad h \in \mathbb{Z}.$$

The example presented here is no exception. In fact, every weakly stationary centered stochastic process has a unique spectral distribution function which can be used to express its ACVF. This is stated in the next theorems.

Theorem 4.2.1 *If $(X_t\colon t \in \mathbb{Z})$ is a complex-valued stochastic process with ACVF γ, then there is a function F such that*

$$\gamma(h) = \int_{(-\pi,\pi]} e^{ih\lambda} dF(\lambda), \qquad h \in \mathbb{Z}. \tag{4.2.1}$$

Note that in Example 4.2.1, the frequencies λ_j are uniquely determined by the piecewise constant F. The statement in Theorem 4.2.1 is implied by Herglotz's result.

Theorem 4.2.2 (Herglotz) *We have that $\gamma\colon \mathbb{Z} \to \mathbb{C}$ is a nonnegative definite function on the integers if and only if there is a unique increasing, right-continuous function F concentrated on $[-\pi, \pi]$ with $F(-\pi) = 0$ such that (4.2.1) holds.*

Proof. If γ satisfies (4.2.1), then it is clearly Hermitian. Let $a_1, \dots, a_n \in \mathbb{C}$. Then,

$$\begin{aligned}
\sum_{j,k=1}^{n} a_j \gamma(j-k)\bar{a}_k &= \sum_{j,k=1}^{n} a_j \int_{(-\pi,\pi]} e^{i(j-k)\lambda} dF(\lambda)\bar{a}_k \\
&= \int_{(-\pi,\pi]} \sum_{j=1}^{n} a_j e^{ij\lambda} \sum_{k=1}^{n} \bar{a}_k e^{-ik\lambda} dF(\lambda) \\
&= \int_{(-\pi,\pi]} \left| \sum_{j=1}^{n} a_j e^{ij\lambda} \right|^2 dF(\lambda) \\
&\geq 0.
\end{aligned}$$

Therefore γ is nonnegative definite and consequently an ACVF.

To prove the converse, assume that γ is a nonnegative function on the integers. Let

$$f_N(\lambda) = \frac{1}{2\pi N} \sum_{j,k=1}^{N} e^{-ij\lambda}\gamma(j-k)e^{ik\lambda} = \frac{1}{2\pi N} \sum_{|\ell|<N} (N - |\ell|)e^{-i\ell\lambda}\gamma(\ell).$$

Observe that $f_N \geq 0$ because γ is nonnegative definite. Let F_N be the distribution function of the density $f_N \mathbb{I}_{(-\pi,\pi]}$. Then, $F_N(\lambda) = 0$ for $\lambda \leq -\pi$, $F_N(\lambda) = F_N(\pi)$ for $\lambda \geq \pi$ and $F_N(\lambda) = \int_{-\pi}^{\lambda} f_N(\nu)d\nu$ for $\lambda \in (-\pi, \pi]$. Fix $h \in \mathbb{Z}$. Then

$$\int_{-\pi}^{\pi} e^{ih\lambda} dF_N(\lambda) = \sum_{|\ell|<N} \Big(1 - \frac{|\ell|}{N}\Big)\gamma(\ell)\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(h-\ell)\lambda} d\nu = \begin{cases} \Big(1 - \frac{|h|}{N}\Big) & |\ell| < N. \\ 0 & |\ell| \geq N. \end{cases} \tag{4.2.2}$$

Note that $F_N(\pi) < \infty$ for all $N \in \mathbb{N}$. Helly's theorem now implies that there are a distribution function F and a subsequence $(F_{N_j}\colon j \in \mathbb{N})$ of $(F_N\colon N \in \mathbb{N})$ such that

$$\int_{-\pi}^{\pi} g(\lambda)dF_{N_j}(\lambda) \to \int_{-\pi}^{\pi} g(\lambda)dF(\lambda)$$

for all $g \in \mathcal{C}^b[-\pi, \pi]$ such that $g(\pi) = g(-\pi)$. Using (4.2.2) with N_j in place of N, we obtain for $|h| < N_j$ that

$$\int_{-\pi}^{\pi} e^{ih\lambda}dF(\lambda) \leftarrow \int_{-\pi}^{\pi} e^{ih\lambda}dF_{N_j}(\lambda) = \Big(1 - \frac{|h|}{N_j}\Big)\gamma(h) \to \gamma(h)$$

and the limits on the left and right must coincide.

It remains to verify the uniqueness of F. To this end, assume that G is a second such function. Then

$$\gamma(h) = \int_{-\pi}^{\pi} e^{ih\lambda}dF(\lambda) = \int_{-\pi}^{\pi} e^{ih\lambda}dG(\lambda), \qquad h \in \mathbb{Z}.$$

This implies first that $\int_{-\pi}^{\pi} g(\lambda)dF(\lambda) = \int_{-\pi}^{\pi} g(\lambda)dG(\lambda)$ for all $g \in \mathcal{C}^b[-\pi,\pi]$ such that $g(-\pi) = g(\pi)$ and therefore also that $F = G$. $\square$

We are interested in the case when F has a derivative f. The following theorem presents a sufficient condition on the ACVF that ensures that f exists and provides a way of calculating it.

Theorem 4.2.3 *If $\gamma : \mathbb{Z} \to \mathbb{C}$ is nonnegative definite such that $\sum_{h\in\mathbb{Z}} |\gamma(h)| < \infty$, then*

$$f(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \gamma(h) e^{-ih\lambda}, \qquad \lambda \in (-\pi, \pi], \tag{4.2.3}$$

and

$$\gamma(h) = \int_{-\pi}^{\pi} e^{ih\lambda} f(\lambda) d\lambda, \qquad h \in \mathbb{Z}.$$

The function f is called the spectral density function of γ. Note that $dF(\lambda) = f(\lambda)d\lambda$.

Proof. It holds that

$$\int_{-\pi}^{\pi} e^{ih\lambda} f(\lambda) d\lambda = \sum_{j\in\mathbb{Z}} \gamma(j) \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(h-j)\lambda} d\lambda = \gamma(h),$$

where exchanging summation and integration is allowed since

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{j\in\mathbb{Z}} \left| e^{i(h-j)\lambda} \gamma(j) \right| d\lambda < \infty$$

and therefore Fubini's theorem applies. $\square$

Corollary 4.2.1 *If $\gamma \colon \mathbb{Z} \to \mathbb{C}$ is nonnegative definite such that $\sum_{h\in\mathbb{Z}} |\gamma(h)| < \infty$, then γ is the ACVF of a weakly stationary stochastic process $(X_t \colon t \in \mathbb{Z})$ if and only if $f(\lambda) \geq 0$ for all $\lambda \in [-\pi, \pi]$.*

Proof. If γ is an ACVF, then we obtain as in the proof of Herglotz's theorem that

$$0 \leq f_N(\lambda) = \frac{1}{2\pi N} \sum_{|\ell|<N} \left(1 - \frac{|\ell|}{N}\right) e^{-i\ell\lambda} \gamma(\ell) \to f(\lambda) \qquad (N \to \infty),$$

and thus $f(\lambda) \geq 0$ as well as $\gamma(h) = \int_{-\pi}^{\pi} e^{ih\lambda} f(\lambda) d\lambda$, which follows from Theorem 4.2.3. The converse follows also from Theorem 4.2.3. $\square$

Example 4.2.2 Let $\gamma(0) = 1$, $\gamma(1) = \gamma(-1) = \alpha$ and $\gamma(h) = 0$ if $|h| \geq 2$. Then, γ is absolutely summable and Theorem 4.2.3 implies that its spectral density is

$$f(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \gamma(h) e^{-ih\lambda} = \frac{1}{2\pi} \left[1 + \alpha(e^{i\lambda} + e^{-i\lambda})\right] = \frac{1}{2\pi} \left[1 + 2\alpha \cos(\lambda)\right].$$

The last expression is nonnegative if and only if the absolute value of α is less than or equal to 1/2. Therefore, we conclude from Corollary 4.2.1 that γ is an ACVF if and only if $|\alpha| \leq 1/2$. In fact, γ is the ACVF of an MA(1) process.

In the case that γ is real-valued, we obtain

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

which may be easier for computations. In the following, we derive the spectral densities of white noise and AR(1) processes.

Example 4.2.3 (White Noise) If $(Z_t \colon t \in \mathbb{Z})$ is $\mathrm{WN}(0, \sigma^2)$, then $\gamma(0) = \sigma^2$ and $\gamma(h) = 0$ for $h \neq 0$. Therefore, we conclude that

$$f(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \gamma(h) e^{-ih\lambda} = \frac{\sigma^2}{2\pi}, \qquad \lambda \in [-\pi, \pi].$$

Since f is constant, all frequencies contribute in the same way to the spectrum of the white noise process, which in turn explains the name.

Example 4.2.4 (AR(1) Processes) Let $(X_t \colon t \in \mathbb{Z})$ be an autoregressive process of order 1 with autoregressive parameter $|\phi| < 1$. Then, its ACVF is given by (see Example 3.2.1)

$$\gamma(h) = \frac{\sigma^2 \phi^{|h|}}{1 - \phi^2}, \qquad h \in \mathbb{Z}.$$

Writing $z = e^{i\lambda}$, we obtain from Theorem 4.2.3 that the spectral density is

$$\begin{aligned} f(\lambda) &= \frac{\sigma^2}{2\pi(1-\phi^2)} \sum_{h=-\infty}^{\infty} \phi^{|h|} z^{-h} \\ &= \frac{\sigma^2}{2\pi(1-\phi^2)} \left[\frac{1}{1-\phi z} + \frac{\phi/z}{1-\phi/z} \right] \\ &= \frac{\sigma^2}{2\pi(1-\phi^2)} \frac{1 - \phi\overline{z} + \phi\overline{z} - \phi^2}{(1-\phi z)(1-\phi\overline{z})} \\ &= \frac{\sigma^2}{2\pi} \frac{1}{1 - 2\phi\cos(\lambda) + \phi^2}. \end{aligned}$$

We have used that $\overline{z} = 1/z$ and $|z|^2 = 1$. Plotting the spectral density for positive ϕ shows that small frequencies have a higher impact than larger ones. This means that the time series plot will look smoother. For negative ϕ, however, the high frequencies dominate the smaller ones and consequently the time series plot will display more rapid fluctuations around the zero mean (sign changes occur often).

The previous example can be viewed as a special case of the next theorem. To compute spectral densities for general ARMA processes, we state a result for two-sided linear processes.

Theorem 4.2.4 *Let $(Y_t\colon t \in \mathbb{Z})$ be a weakly stationary complex-valued stochastic process with zero mean and spectral distribution function F_Y. Let $(\psi_j\colon j \in \mathbb{Z})$ be an absolutely summable sequence and define the process $(X_t\colon t \in \mathbb{Z})$ by letting*

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

Then, it holds that

(i) $(X_t\colon t \in \mathbb{Z})$ is a weakly stationary process with zero mean;
(ii) the spectral distribution function of $(X_t\colon t \in \mathbb{Z})$ is computed as

$$F_X(\lambda) = \int_{(-\pi,\lambda]} \left| \sum_{j=-\infty}^{\infty} \psi_j e^{-ij\nu} \right|^2 dF_Y(\nu), \qquad \lambda \in [-\pi, \pi].$$

(iii) If $(Y_t\colon t \in \mathbb{Z})$ has a spectral density, then so does $(X_t\colon t \in \mathbb{Z})$. It is given by

$$f_X(\lambda) = |\psi(e^{-i\lambda})|^2 f_Y(\lambda),$$

where $\psi(e^{-i\lambda}) = \sum_{j=-\infty}^{\infty} \psi_j e^{-ij\lambda}$.

Proof. We only show (ii), since it implies (iii) by taking the derivative. First note that, for any $h \in \mathbb{Z}$,

$$\begin{aligned}
E[X_{t+h}\overline{X}_t] &= \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \psi_j \overline{\psi}_k \gamma_Y(h-j+k) \\
&= \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \psi_j \overline{\psi}_k \int_{(-\pi,\pi]} e^{i(h-j+k)\nu} dF_Y(\nu) \\
&= \int_{(-\pi,\pi]} \left(\sum_{j=-\infty}^{\infty} \psi_j e^{-ij\nu} \right) \left(\sum_{k=-\infty}^{\infty} \overline{\psi}_k e^{ik\nu} \right) e^{ih\nu} dF_Y(\nu) \\
&= \int_{(-\pi,\pi]} e^{ih\nu} \left| \sum_{j=-\infty}^{\infty} \psi_j e^{-ij\nu} \right|^2 dF_Y(\nu).
\end{aligned}$$

Thus, $dF_X(\nu) = |\psi(e^{-i\nu})|^2 dF_Y(\nu)$ and an integration of both sides (with integration limits $-\pi$ and λ) yields (ii). $\square$

Spectral densities for ARMA(p,q) processes can now be computed from the foregoing theorem. Note that $\psi(B)$ (with ψ defined in part (iii) of Theorem 4.2.4) is called a *time-invariant filter*, $\psi(e^{-i\,\cdot})$ a *transfer function* and $|\psi(e^{-i\,\cdot})|^2$ a *power transfer function*. More on filters will be covered in Section 3.4 below.

Theorem 4.2.5 *Let $(X_t\colon t \in \mathbb{Z})$ be an ARMA(p,q) process such that $\phi(z)$ and $\theta(z)$ have no common zeroes and such that $\phi(z)$ has no zero on the unit circle. Then, $(X_t\colon t \in \mathbb{Z})$ has spectral density*

$$f_X(\lambda) = \frac{\sigma^2}{2\pi} \frac{|\theta(e^{-i\lambda})|^2}{|\phi(e^{-i\lambda})|^2}, \qquad \lambda \in [-\pi, \pi].$$

Proof. By assumption, the ARMA(p,q) process $(X_t\colon t\in\mathbb{Z})$ can be written in stationary solution form (possibly noncausal) as

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

where $(\psi_j\colon j\in\mathbb{Z})$ is absolutely summable. From Theorem 4.2.4 we know that $(X_t\colon t\in\mathbb{Z})$ admits a spectral density, which can be obtained by first defining $U_t=\phi(B)X_t=\theta(B)Z_t$ and then deducing that

$$f_U(\lambda)=|\phi(e^{-i\lambda})|^2f_X(\lambda)=|\theta(e^{-i\lambda})|^2f_Z(\lambda).$$

Consequently,

$$f_X(\lambda)=\frac{|\theta(e^{-i\lambda})|^2}{|\phi(e^{-i\lambda})|^2}f_Z(\lambda),$$

since $\phi(z)$ does not have zeroes on the unit circle. The proof is complete after an application of Example 4.2.3. $\square$

Example 4.2.5 (a) Let $X_t=Z_t+\theta Z_{t-1}$ be an MA(1) process. Then,

$$f_X(\lambda)=\frac{\sigma^2}{2\pi}|1+\theta e^{-i\lambda}|^2=\frac{\sigma^2}{2\pi}(1+2\theta\cos\lambda+\theta^2).$$

(b) Let $X_t-\phi X_{t-1}=Z_t$ be an AR(1) process for which $|\phi|\neq 1$. Then,

$$f_X(\lambda)=\frac{\sigma^2}{2\pi}\frac{1}{|1-\phi e^{-i\lambda}|^2}=\frac{\sigma^2}{2\pi}\frac{1}{1-2\phi\cos\lambda+\phi^2},$$

which is the same result as in Example 4.2.4.

Theorem 4.2.5 immediately suggests parametric estimators for the spectral density: Obtain estimators for $\phi_1,\ldots,\phi_p$ and $\theta_1,\ldots,\theta_q$ which determine estimated versions of the polynomials $\phi(z)$ and $\theta(z)$. Plug the estimated polynomials back into the spectral density f_X in the theorem to arrive at

$$\hat f_X(\lambda)=\frac{\hat\sigma^2}{2\pi}\frac{|\hat\theta(e^{-i\lambda})|^2}{|\hat\phi(e^{-i\lambda})|^2},\qquad \lambda\in(-\pi,\pi],$$

where $\hat\sigma^2$ denotes an estimator for σ^2. A more popular method, however, is given by the periodogram which is to be introduced in the Section 4.4.

4.3 The Spectral Representation Theorem

In this section we provide the fundamental result that each mean-zero weakly stationary process can be written as the stochastic integral with respect to a certain orthogonal increment process. A process $(Z(\lambda)\colon \lambda \in [-\pi, \pi])$ is called *orthogonal increment process* if $E[Z(\lambda) = 0$, $E[|Z(\lambda)|^2] < \infty$ and

$$E\big[\{Z(\lambda_4) - Z(\lambda_3)\}\{\overline{Z(\lambda_2) - Z(\lambda_1)}\}\big] = 0$$

if $(\lambda_1, \lambda_2) \cap (\lambda_3, \lambda_4) = \emptyset$. It is always assumed that $(Z(\lambda)\colon \lambda \in [-\pi, \pi])$ is *right-continuous*, that is,

$$\big\|Z(\lambda + \delta) - Z(\lambda)\big\|^2 \to 0 \qquad (\delta \searrow 0).$$

Proposition 4.3.1 *If $(Z(\lambda)\colon \lambda \in [-\pi, \pi])$ is an orthogonal increment process, then there is a unique monotonically non-decreasing, right-continuous function F such that (i) $F(\lambda) = 0$ for $\lambda \leq -\pi$, $F(\lambda) = F(\pi)$ for $\lambda \geq \pi$ and*

$$F(\mu) - F(\lambda) = \|Z(\mu) - Z(\lambda)\|^2$$

for all $-\pi \leq \lambda \leq \mu \leq \pi$.

Proof. If $\lambda = -\pi$, then F must satisfy $F(\mu) = \|Z(\mu) - Z(-\pi)\|^2$ for $-\pi \leq \mu \leq \pi$. Let $-\pi \leq \lambda \leq \mu \leq \pi$. Then

$$\begin{aligned} F(\mu) &= \|Z(\mu) - Z(-\pi)\|^2 \\ &= \|Z(\mu) - Z(\lambda) + Z(\lambda) - Z(-\pi)\|^2 \\ &= \|Z(\mu) - Z(\lambda)\|^2 + \|Z(\lambda) - Z(-\pi)\|^2 \\ &\geq F(\lambda), \end{aligned}$$

so that F is non-decreasing. Since

$$F(\mu + \delta) - F(\mu) = \|Z(\mu + \delta) - Z(\mu)\|^2 \to 0 \qquad (\delta \searrow 0),$$

by the assumed right-continuity of $(Z(\lambda)\colon \lambda \in [-\pi, \pi])$, F is also right-continuous. □

In the next step, we shall define the stochastic integration of functions $f \in L^2(F) = L^2([-\pi, \pi], \mathcal{B}, F)$ with respect to an orthogonal increment process $(Z(\lambda)\colon \lambda \in [-\pi, \pi])$. This is done by algebraic induction.

Step 1: For piecewise constant functions

$$f(\lambda) = \sum_{i=0}^{n} f_i I_{(\lambda_{i+1}, \lambda_i]}(\lambda)$$

where $-\pi = \lambda_0 < \lambda_1 < \ldots < \lambda_{n+1} = \pi$, define the stochastic integral to be

$$I(f) = \sum_{i=0}^{n} f_i [Z(\lambda_{i+1} - Z(\lambda_i)].$$

If $\mathcal{D}$ denotes the set of all such piecewise constant functions, then $I\colon \mathcal{D} \to L^2$ defines a linear, inner product preserving mapping. To see this, let $f, g \in \lceil$. Then

$$\begin{aligned}
\langle I(f), I(g)\rangle_{L^2} &= \Big\langle \sum_{j=0}^{n} f_j[Z(\lambda_{j+1}) - Z(\lambda_j)], \sum_{k=0}^{n} g_k[Z(\lambda_{k+1}) - Z(\lambda_k)]\Big\rangle \\
&= \sum_{j=0}^{n} f_j \bar{g}_j [F(\lambda_{j+1}) - F(\lambda_j)] \\
&= \int_{(-\pi,\pi]} f(\lambda) g(\lambda) dF(\lambda) \\
&= \langle f, g\rangle_{L^2(F)}.
\end{aligned}$$

Step 2: Extend the definition to $\bar{\mathcal{D}}$, the set of limit points of sequences in $\mathcal{D}$. By definition, for each $f \in \bar{\mathcal{D}}$ there is $(f_n\colon n \in \mathbb{N}) \subset \mathcal{D}$ such that $\|f_n - f\|_{L^2(F)} \to 0$. It is natural to define

$$I(f) = \lim_{n\to\infty} I(f_n),$$

where the limit is taken in the mean square sense (the limit exists and is unique). It can be shown that the resulting mapping $I\colon \bar{\mathcal{D}} \to L^2$ is linear and inner product preserving. For the latter, note that by the continuity of the inner product

$$\langle I(f), I(g)\rangle_{L^2} = \lim_{n\to\infty} \langle I(f_n), I(g_n)\rangle_{L^2} = \lim_{n\to\infty} \langle f_n, g_n\rangle_{L^2(F)} = \langle f, g\rangle_{L^2(F)}.$$

Step 3: Show that $\bar{\mathcal{D}} = L^2(F)$. Then $I\colon L(F) \to I(\bar{\mathcal{D}})$ is an isometric isomorphism. The following definition makes therefore sense.

Definition 4.3.1 *If $(Z(\lambda)\colon \lambda \in [-\pi, \pi])$ is an orthogonal increment process with distribution function F, then we define*

$$\int_{(-\pi,\pi]} f(\lambda) dZ(\lambda) := I(f)$$

for any $f \in L^2(F)$.

The following properties are immediate: (i) $E[I(f)] = 0$, (ii) $I(af + bg) = aI(f) + bI(g)$ for all $a, b \in \mathbb{C}$ and (iii) $\langle I(f), I(g)\rangle_{L^2} = \langle f, g\rangle_{L^2(F)}$. The foregoing shows that if we select $f = e^{it\cdot}$, then

$$X_t = I(e^{it\cdot}) = \int_{(-\pi,\pi]} e^{it\lambda} dZ(\lambda), \qquad t \in \mathbb{Z},$$

defines a stationary process with mean zero and ACVF

$$\gamma(h) = \int_{(-\pi,\pi]} e^{ih\lambda} dF(\lambda), \qquad h \in \mathbb{Z}.$$

We shall show the converse in what follows. To this end, we need an isomorphism between the "time domain" and the "frequency domain", that is mapping

$$L^2 \supset \bar{\mathcal{H}} = \overline{\mathrm{sp}}(X_t\colon t \in \mathbb{Z}) \to \bar{\mathcal{K}} = \overline{\mathrm{sp}}(e^{it\cdot}\colon t \in \mathbb{Z}) \subset L^2(F),$$

where in fact $\bar{\mathcal{K}} = L^2(F)$. Let

$$T\colon \mathcal{H} \to \mathcal{K}, \quad \sum_{j=1}^{n} a_j X_{t_j} \mapsto \sum_{j=1}^{n} a_j e^{it_j\cdot}.$$

Then, T is well defined and an isometric isomorphism between $\mathcal{H}$ and $\mathcal{K}$. This mapping can in fact be extended. This is stated as a theorem.

Theorem 4.3.1 *(a) If $(X_t\colon t \in \mathbb{Z})$ is a mean zero weakly stationary process with spectral distribution F, then there is a unique mapping $T\colon \bar{\mathcal{H}} \to L^2(F)$ such that $TX_t = e^{it\cdot}$.*

(b) If T is as in (a), then the orthogonal increment process $(Z(\lambda)\colon \lambda \in [-\pi, \pi])$ is given by $Z(\lambda) = T^{-1}(I_{(-\pi,\lambda]}(\cdot))$, $\lambda \in [-\pi, \pi]$.

Proof. (b) The orthogonal increment process $(Z(\lambda)\colon \lambda \in [-\pi, \pi])$ is well-defined by (a). Since $Z(\lambda) \in \overline{\mathrm{sp}}(X_t\colon t \in \mathbb{Z})$, it follows that $E[|Z(\lambda)|^2] < \infty$. Also, $E[Z(\lambda)] = \langle Z(\lambda), 1\rangle = 0$. Let $-\pi \le \lambda_1 \le \lambda_2 \le \lambda_3 \le \lambda_4 \le \pi$. Then

$$\begin{aligned}
\langle Z(\lambda_4) - Z(\lambda_3), &Z(\lambda_2) - Z(\lambda_1)\rangle \\
&= \langle TZ(\lambda_4) - TZ(\lambda_3), TZ(\lambda_2) - TZ(\lambda_1)\rangle \\
&= \langle I_{(\lambda_3,\lambda_4]}(\cdot), I_{(\lambda_1,\lambda_2]}(\cdot)\rangle \\
&= \int_{(-\pi,\pi]} I_{(\lambda_3,\lambda_4]}(\lambda) I_{(\lambda_1,\lambda_2](}\lambda) dF(\lambda) \\
= 0.
\end{aligned}$$

This proves part (b). □

One of the major applications of the spectral representation theorem lies in the field of filtering and signal processing. In this context, a stationary process $(X_t\colon t \in \mathbb{Z})$ is considered an *input* that is filtered through the application of a *filter* $\Psi = (\psi_{t,j}\colon t, j \in \mathbb{Z})$ to yield an *output* process $(Y_t\colon t \in \mathbb{Z})$ given by

$$Y_t = \sum_{j\in\mathbb{Z}} \psi_{t,j} X_j, \qquad t \in \mathbb{Z}.$$

The filter Ψ is said to be *time-invariant* if $\psi_t, j \equiv \psi_{t-j}$ depends on t and j only through the difference $t - j$. Then

$$Y_t = \sum_{j\in\mathbb{Z}} \psi_j X_{t-j}, \qquad t \in \mathbb{Z},$$

is a two-sided MA(∞) process. The filter Ψ is said to be causal if $\psi_j = 0$ for $j < 0$. In the following we focus on time-invariant filters such that $\sum_{j\in\mathbb{Z}} |\psi_j| < \infty$. We know that the spectral distribution of the output $(Y_t\colon t \in \mathbb{Z})$ is given by

$$F_Y(\lambda) = \int_{(-\pi,\pi]} |\psi(e^{i\lambda})|^2 dF_X(\lambda),$$

where $\psi(e^{i\lambda}) = \sum_{j\in\mathbb{Z}} \psi_j e^{-ij\lambda}$. Moreover,

$$Y_t = \int_{(-\pi,\pi]} e^{it\lambda}\psi(e^{-i\lambda} dZ_X(\lambda), \qquad t \in \mathbb{Z},$$

and, if $\psi(e^{-i\lambda}) \neq 0$, then

$$X_t = \int_{(-\pi,\pi]} \pi(e^{-it\lambda} e^{it\lambda} dZ_Y(\lambda), \qquad t \in \mathbb{Z},$$

where $\pi(e^{-i\lambda}) = 1/\psi(e^{-i\lambda})$ and $dZ_Y(\lambda) = \psi(e^{-i\lambda})dZ_X(\lambda)$. This means roughly that $(X_t\colon t \in \mathbb{Z})$ can be decomposed into sinusoids $e^{it\lambda}dZ_X(\lambda)$, $\lambda \in [-\pi,\pi]$, while $(Y_t\colon t \in \mathbb{Z})$ can be decomposed into sinusoids $e^{it\lambda}dZ_Y(\lambda) = \psi(e^{-i\lambda})e^{it\lambda}dZ_X(\lambda)$. This allows an easy computation of the output as it is just the multiplication of the X-sinusoids with the transfer function $\psi(e^{-i\lambda})$.

4.4 The Periodogram

Time series analysis in the frequency domain means analyzing spectral density functions (if they exist) and spectral distribution functions. The fundamental tool for the statistician is the periodogram. To introduce this notion, assume that we have collected a set of n possibly complex-valued observations $x_1, \ldots, x_n$ which we combine in a vector

$$\boldsymbol{x} = (x_1, \ldots, x_n)' \in \mathbb{C}^n.$$

For two elements $\boldsymbol{u}$ and $\boldsymbol{v}$ of $\mathbb{C}^n$, let their inner product be defined by $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = u_1\overline{v}_1 + \ldots + u_n\overline{v}_n$. Our goal here is to express x_t as a linear combination of harmonics, that is,

$$x_t = \frac{1}{\sqrt{n}} \sum_{-\pi<\omega_j\leq\pi} a_j e^{it\omega_j}, \qquad t = 1, \ldots, n,$$

where

$$-\pi < \omega_j = \frac{2\pi j}{n} \leq \pi$$

are called *Fourier frequencies*. Observe that there are always n Fourier frequencies determined by the n integers $-\lfloor \frac{n-1}{2} \rfloor, \ldots, \lfloor \frac{n}{2} \rfloor$ which we combine in the set F_n. Utilizing the notation

$$\boldsymbol{e}_j = \frac{1}{\sqrt{n}} \left(e^{i\omega_j}, e^{i2\omega_j}, \ldots, e^{in\omega_j}\right)'$$

we can write

$$\boldsymbol{x} = \sum_{j \in F_n} a_j \boldsymbol{e}_j \qquad \text{and} \qquad a_j = \langle \boldsymbol{x}, \boldsymbol{e}_j \rangle = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} x_t e^{-it\omega_j}$$

so that the coefficients a_j are determined by taking the inner product of $\boldsymbol{x}$ and $\boldsymbol{e}_j$. (It can be proved that the vectors $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ form an orthonormal basis (ONB) of $\mathbb{C}^n$.) The sequence $(a_j \colon j \in F_n)$ is called *discrete Fourier transform*. It gives rise to the following definition.

Definition 4.4.1 (The Periodogram) *The mapping $I(\omega_j) = I_n(\omega_j) = |a_j|^2$ is called the periodogram of $\boldsymbol{x}$ at frequency ω_j, where $j \in F_n$.*

The quadratic variation in $\boldsymbol{x}$ can be nicely attributed to the periodogram. To see this, write

$$\|\boldsymbol{x}\|^2 = \langle \boldsymbol{x}, \boldsymbol{x} \rangle = \sum_{j \in F_n} \sum_{k \in F_n} a_j \bar{a}_k \langle \boldsymbol{e}_j, \boldsymbol{e}_k \rangle = \sum_{j \in F_n} I_n(\omega_j).$$

Therein, we have used that the $\boldsymbol{e}_j$ are an ONB. The connection between the periodogram and spectral densities will become clear from the following theorem.

Theorem 4.4.1 *If $\omega_j \neq 0$ is a Fourier frequency, then*

$$I_n(\omega_j) = \sum_{|h| < n} \hat{\gamma}(h) e^{-ih\omega_j},$$

where,

$$\hat{\gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-h} (x_{t+h} - m)(\overline{x_t - m}), \qquad 0 \leq h < n,$$

$$\hat{\gamma}(h) = \overline{\hat{\gamma}(-h)}, \qquad -n < h < 0,$$

with $m = n^{-1}(x_1 + \ldots + x_n)$.

Proof. Note first that

$$0 = \frac{1}{n} \sum_{t=1}^{n} e^{-it\omega_j} = \frac{1}{n} \sum_{s=1}^{n} e^{is\omega_j}$$

as long as $\omega_j \neq 0$. From the definition of the periodogram it follows now

$$\begin{aligned}
I_n(\omega_j) &= \frac{1}{n}\left|\sum_{t=1}^{n} x_t e^{-it\omega_j}\right|^2 \\
&= \frac{1}{n}\sum_{t=1}^{n} x_t e^{-it\omega_j}\sum_{s=1}^{n}\overline{x}_s e^{is\omega_j} \\
&= \frac{1}{n}\sum_{t=1}^{n}(x_t - m)e^{-it\omega_j}\sum_{s=1}^{n}(\overline{x_s - m})e^{is\omega_j} \\
&= \frac{1}{n}\sum_{t=1}^{n}\sum_{s=1}^{n}(x_t - m)(\overline{x_s - m})e^{-i(t-s)\omega_j} \\
&= \sum_{|h|<n}\hat{\gamma}(h)e^{-ih\omega_j}.
\end{aligned}$$

The proof is complete. □

If we compare the representation of $I_n(\omega_j)$ given in Theorem 4.4.1 with the form of the spectral density in Theorem 4.2.3, it is clear how the periodogram can be utilized as an estimator for the unknown f. Note, however, the exceptional frequency $\omega_0 = 0$. In this case, it holds that $I_n(0) = n|\bar{x}|^2$. The limiting behavior of $I_n(\omega_j)$ is discussed in the following two theorems. Before, we extend the definition of the periodogram to non-Fourier frequencies. Let

$$\begin{aligned}
I_n(\lambda) &= \begin{cases} I_n(\omega_j), & \lambda \in [0,\pi]\colon \omega_j - \frac{\pi}{n} < \lambda \leq \omega_j + \frac{\pi}{n}, \\ I_n(-\lambda), & \lambda \in [-\pi, 0), \end{cases} \\
&= I_n(g(n,\lambda)),
\end{aligned}$$

where $g(n,\lambda)$ denotes the closest Fourier frequency to λ.

Theorem 4.4.2 *Let $(X_t\colon t \in \mathbb{Z})$ be a weakly stationary stochastic process with spectral density function f. Denote by $I_n(\lambda)$ the (random) periodogram obtained from the random variables $X_1, \ldots, X_n$. Then, it holds that*
(i) $E[I_n(0)] - n\mu^2 \to 2\pi f(0)$ as $n \to \infty$, where $\mu = E[X_t]$;
(ii) $E[I_n(\lambda)] \to 2\pi f(\lambda)$ as $n \to \infty$ for all $\lambda \in (-\pi,\pi]\backslash\{0\}$.
(iii) If $\mu = 0$ the convergence $E[I_n(\lambda)] \to 2\pi f(\lambda)$ is uniform in

Proof. (i) It holds that

$$E[I_n(0)] - n\mu^2 = nE\left[\Big(\frac{1}{n}\sum_{t=1}^{n} X_t\Big)^2\right] - n\mu^2 = n\mathrm{Var}(\bar{X}_n).$$

The right-hand side of this equation converges to the sum of the covariances which is equal to $2\pi f(0)$ according to Theorem 4.2.3.

(ii) Let $\lambda \in (0,\pi]$. Then $g(n,\lambda) \neq 0$ if n is large enough. Therefore

$$\begin{aligned} E[I_n(\lambda)] &= \sum_{|h|<n} \frac{1}{n} \sum_{t=1}^{n-|h|} E\big[(X_{t+h}-\mu)(X_t-\mu)\big] e^{-ihg(n,\lambda)} \\ &= \sum_{|h|<n} \Big(1-\frac{|h|}{n}\Big)\gamma(h) e^{-ihg(n,\lambda)} \\ &\to 2\pi f(\lambda) \end{aligned}$$

because $\sum_{|h|<n}(1-\frac{|h|}{n})\gamma(h)e^{-ih\lambda} \to 2\pi f(\lambda)$ and $g(n,\lambda)\to\lambda$ for all λ .

(iii) Both of the convergences in the previous line are uniform. $\square$

Theorem 4.4.2 shows that the periodogram is asymptotically unbiased if $j \neq 0$. If the mean component $n\mu^2$ is subtracted from the periodogram at frequency zero, it is also unbiased in the limit. The following results indicate, however, that the periodogram is inconsistent.

Assume for the moment that $(X_t\colon t\in\mathbb{Z})$ is a Gaussian white noise with variance σ^2. Then

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

$$\beta(\omega_j) = \sqrt{\frac{2}{n}} \sum_{t=1}^{n} X_t \sin(\omega_j t),$$

$j = 1,\ldots,\lfloor\frac{n-1}{2}\rfloor$, are independent $N(0,\sigma^2)$. In particular, we have that $I_n(\omega_j) = \frac{1}{2}[\alpha^2(\omega_j)+\beta^2(\omega_j)]$ are independent exponentially distributed with parameter $\sigma^2 = 2\pi f_X(\omega_j)$. The result holds more generally.

Theorem 4.4.3 *Let $(Z_t\colon t\in\mathbb{Z}) \sim \mathrm{IID}(0,\sigma^2)$ with periodogram $I_n(\lambda)$.*

(i) If $0<\lambda_1<\ldots<\lambda_m<\pi$, then

$$\big(I_n(\lambda_1),\ldots,I_n(\lambda_m)\big)' \xrightarrow{\mathcal{D}} \mathcal{E} \qquad (n\to\infty),$$

where $\mathcal{E} = (\mathcal{E}_1,\ldots,\mathcal{E}_m)'$ and $\mathcal{E}_1,\ldots,\mathcal{E}_m$ are independent $\mathrm{Exp}(\sigma^2)$ random variables.

(ii) If $E[Z_t^4] = \eta\sigma^4 < \infty$, then

$$\mathrm{Var}\big(I_n(\omega_j)\big) = \begin{cases} \frac{1}{n}(\eta-3)\sigma^4 + 2\sigma^4, & \omega_j = 0,\pi, \\ \frac{1}{n}(\eta-3)\sigma^4 + \sigma^4, & \omega_j \in (0,\pi), \end{cases}$$

and

$$\mathrm{Cov}\big(I_n(\omega_j, I_n(\omega_k)\big) = \tfrac{1}{n}(\eta-3)\sigma^4, \qquad \omega_j \neq \omega_k.$$

Proof. (i) For $\lambda\in(0,\pi)$ let $\alpha(\lambda) = \alpha(g(n,\lambda))$ and $\beta(\lambda) = \beta(g(n,\lambda))$ with $\alpha(\cdot)$ and $\beta(\cdot)$ from above. Then $I_n(\lambda) = \frac{1}{2}[\alpha^2(\lambda)+\beta^2(\lambda)]$ and it suffices to verify that

$$\big(\alpha(\lambda_1),\beta(\lambda_1),\ldots,\alpha(\lambda_m),\beta(\lambda_m)\big)' \xrightarrow{\mathcal{D}} N(\mathbf{0},\sigma^2 I_{2m}),$$

where I_{2m} is the $2m \times 2m$-dimensional identity matrix. Now $g(n,\lambda) \in (0,\pi)$ if n is large enough because $\lambda \in (0,\pi)$. Then,

$$\mathrm{Var}(\alpha(\lambda)) = \mathrm{Var}(\alpha(g(n,\lambda))) = \frac{2\sigma^2}{n}\sum_{t=1}^{n}\cos^2(g(n,\lambda)t) = \sigma^2.$$

Let $\epsilon > 0$. Then

$$\begin{aligned}
&\frac{1}{n}\sum_{t=1}^{n} E\left[\cos^2(g(n,\lambda)t) Z_t^2 \mathbb{I}_{\{|\cos(g(n,\lambda)t)Z_t| > \epsilon\sqrt{n}\sigma\}}\right] \\
&\qquad \le \frac{1}{n}\sum_{t=1}^{n} E\left[Z_t^2 \mathbb{I}_{\{|Z_t| > \epsilon\sqrt{n}\sigma\}}\right] \\
&\qquad = E\left[Z_1^2 \mathbb{I}_{\{|Z_1| > \epsilon\sqrt{n}\sigma\}}\right] \\
&\qquad \to 0.
\end{aligned}$$

Now Lindeberg's condition is satisfied and implies that each $\alpha(\lambda)$ is normally distributed with mean 0 and variance σ^2. A similar reasoning applies to $\beta(\lambda)$. The variance-covariance structure follows from the fact that the independence of the various $\alpha(\lambda_i)$ and $\beta(\lambda_j)$.

(ii) See Brockwell and Davis (1991), Chapter 10. □

Next we consider linear processes

$$X_t = \sum_{j\in\mathbb{Z}} \psi_j Z_{t-j}, \qquad t \in \mathbb{Z}, \tag{4.4.4}$$

where $(Z_t\colon t \in \mathbb{Z}) \sim \mathrm{IID}(0,\sigma^2)$ and $\sum_{j\in\mathbb{Z}} |\psi_j| < \infty$. We know that $f_X(\lambda) = |\psi(e^{-i\lambda})|^2 f_Z(\lambda)$ for all $\lambda \in [-\pi,\pi]$. A similar results holds for periodograms. More precisely,

$$I_{n,X}(\lambda) = |\psi(e^{-i\lambda})|^2 I_{n,Z}(\lambda) + R_n(g(n,\lambda)),$$

where the remainder term $R_n(g(n,\lambda))$ is uniformly $o_P(1)$. This result is the basis for the following theorem which is stated without proof.

Theorem 4.4.4 *Let $(X_t\colon t \in \mathbb{Z})$ be the linear process* (4.4.4) *with spectral density f periodogram I_n.*

(i) If $f(\lambda) > 0$ for all $\lambda \in [-\pi,\pi]$ and if $0 < \lambda_1 < \ldots < \lambda_m < \pi$, then

$$\big(I_n(\lambda_1),\ldots,I_n(\lambda_m)\big)' \xrightarrow{\mathcal{D}} \mathcal{E} \qquad (n \to \infty),$$

where $\mathcal{E} = (\mathcal{E}_1,\ldots,\mathcal{E}_m)'$ and $\mathcal{E}_\ell$ are independent $\mathrm{Exp}(2\pi f(\lambda_\ell))$ *random variables.*

(ii) If $\sum_{j\in\mathbb{Z}} \sqrt{|j|}|\psi_j| < \infty$ and if $E[Z_1^4] = \eta\sigma^4 < \infty$, then

$$\mathrm{Cov}\big(I_n(\omega_j), I_n(\omega_k)\big) \to \begin{cases} 2(2\pi)^2 f^2(\omega_j), & \omega_j = \omega_k = 0,\pi. \\ (2\pi)^2 f^2(\omega_j), & \omega_j = \omega_k \in (0,\pi). \\ 0, & \omega_j \ne \omega_k. \end{cases}$$

From this result, we have that, for any $\epsilon > 0$,

$$P\left(|I(\omega_j) - 2\pi f(\omega_j)| > \epsilon\right) = 1 - \int_{2\pi f(\omega_j)-\epsilon}^{2\pi f(\omega_j)+\epsilon} e^{-x/(2\pi f(\omega_j))}dx > 0.$$

Consequently, the periodogram cannot be a consistent estimator for the spectral density. There are ways to obtain improved periodogram-based estimators. We will return to these in Section 4.5. Here, we continue instead with introducing a couple of tests for hidden periodicities.

Hidden Periodicities. Suppose that $(X_t)_{t\in\mathbb{Z}}$ is a (real-valued) stochastic process which has been observed for $t = 1, \dots, n$. In this paragraph we are interested in testing the hypotheses

$$\begin{array}{lll} H_0: & X_1, \dots, X_n & \text{are generated by Gaussian white noise,} \\ H_A: & X_1, \dots, X_n & \text{are generated by Gaussian white noise} \\ & & \text{and a deterministic periodic component.} \end{array}$$

Under H_0, the vector $\boldsymbol{X} = (X_1, \dots, X_n)^T$ is multivariate normal with independent components and it can be shown that

$$V_k = 2I(\omega_k) \sim \tfrac{1}{2}\chi^2(2), \qquad k = 1, \dots, \left\lfloor \frac{n-1}{2} \right\rfloor =: q$$

and these variables are independent. The key result that will be used to define the test statistic is the following. If we let

$$Y_\ell = \left(\sum_{j=1}^{q} V_j\right)^{-1} \sum_{k=1}^{\ell} V_k = \left(\sum_{j=1}^{q} I(\omega_j)\right)^{-1} \sum_{k=1}^{\ell} I(\omega_k), \qquad \ell = 1, \dots, q-1, \tag{4.4.5}$$

then these variables are distributed as the order statistics of a sample of $q-1$ independent uniform random variables on the interval $(0, 1)$.

Remark 4.4.1 (Fisher's test) *Let $Y_0 = 0$ and $Y_q = 1$. Then, we define*

$$M_q = \max_{1\le\ell\le q}(Y_\ell - Y_{\ell-1}) = \left(\sum_{j=1}^{q} I(\omega_j)\right)^{-1} \max_{1\le\ell\le q} I(\omega_\ell).$$

Note that this quantity is distributed as the length of the largest subinterval of $(0, 1)$ when the interval is randomly partitioned by $q-1$ points drawn from independent, identically distributed random variables uniform on $(0, 1)$. Feller (1971) computed the distribution function of M_q:

$$P(M_q \le a) = \sum_{j=0}^{q} (-1)^j \binom{q}{j} (1 - ja)_+^{q-1},$$

where $x_+ = \max\{x, 0\}$. Fisher's test is now based on the random variable $\xi_q = qM_q$. Given data $x_1, \dots, x_n$ we compute the corresponding observed value a_n of ξ_q and reject the null hypothesis H_0 if $P(\xi_q \ge a_n) < \alpha$, where α is a prespecified significance level.

Remark 4.4.2 (Kolmogorov-Smirnov test) *Recalling the definition of Y_ℓ in (4.4.5), one can develop another test for H_0 against H_A as follows. The empirical distribution function induced by the sequence $Y_1, \ldots, Y_q$ is piecewise constant with jumps of size q^{-1} at Y_ℓ. Under H_0, it is the empirical distribution function of a sample of size $q-1$ from the uniform distribution on $(0,1)$. This suggests testing the hypotheses via the Kolmogorov-Smirnov test by comparing the empirical distribution function with the distribution function of a uniform random variable. For practical purposes, one rejects at level α if the empirical distribution function takes at least one value outside the bounds determined by*

$$x \pm \frac{c_\alpha}{\sqrt{q-1}}, \qquad x \in (0,1),$$

where $c_{.05} = 1.36$ and $c_{.01} = 1.63$.

4.5 Periodogram smoothing

In this section, we show how smoothing can be used to improve the performance of the periodogram. This is similar to the corresponding topics covered in Section 1.3. There are two options for smoothing. It can be performed directly in the frequency domain or in the time domain before taking the Fourier transform. In this section, we discuss both approaches and their connection.

It is throughout this section assumed that the underlying time series $(X_t \colon t \in \mathbb{Z})$ can be written as the two-sided MA(∞) process (4.4.4) for which the weights $(\psi_j \colon j \in \mathbb{Z})$ satisfy the summability condition $\sum_{j\in\mathbb{Z}} \sqrt{|j|}|\psi_j| < \infty$. Following the theory developed in the preceding section, we can approximately write that

$$\frac{I(\omega_j)}{2\pi} \overset{\mathcal{D}}{\approx} f(\omega_j) + f(\omega_j)U_j, \qquad j = 1, \ldots, \left\lfloor \frac{n-1}{2} \right\rfloor,$$

where $(U_j) \sim \mathrm{WN}(0,1)$. Consequently, $E[I(\omega_j)] \approx 2\pi f(\omega_j)$ and $\mathrm{Var}(I(\omega_j)) = 4\pi^2 f^2(\omega_j)$, which suggests that one can use

$$\hat{f}(\omega_j) = \frac{1}{2\pi} \sum_{k=-m}^{m} \frac{1}{2m+1} I(\omega_{j+k}) \tag{4.5.1}$$

as an estimator for $f(\omega_j)$. Heuristically, it follows that

$$E\left[\hat{f}(\omega_j)\right] = \sum_{k=-m}^{m} \frac{1}{2m+1} f(\omega_{j+k}) \approx f(\omega_j)$$

and

$$\mathrm{Var}(\hat{f}(\omega_j)) = \frac{1}{(2m+1)^2} \sum_{k=-m}^{m} f^2(\omega_{j+k}) \approx \frac{1}{2m+1} f^2(\omega_j)$$

provided that the bandwidth m is chosen such that f is approximately constant over the range of Fourier frequencies $[\omega_{j-k}, \omega_{j+k}]$. The estimator just described is called *Daniell filter*.

Note that the variance of $\hat{f}(\omega_j)$ decreases with increasing m. Possibly, however, there may be larger bias with increasing m, too. A compromise is needed which can often be tricky to achieve, since the bias depends on the unknown spectral density function f. (As extreme cases consider a white noise sequence, for which m should be taken large, and a sequence resulting in a spectral density with a single sharp peak, for which m needs to be small.) In practice, one should experiment with a number of choices for the bandwidth.

Instead of working with the Daniell filter, that assigns equal weights to the frequencies inside the bandwidth, one can pick more general weight functions W_n satisfying

$$W_n(k) = W_n(-k), \qquad W_n(k) \geq 0, \qquad \sum_{k=-m}^{m} W_n(k) = 1, \qquad \sum_{k=-m}^{m} W_n^2(k)) \to 0$$

for $m = m(n) \to \infty$ such that $m/n \to 0$ as $n \to \infty$. With these definitions at hand, the estimator in (4.5.1) can be modified to

$$\hat{f}(\omega_j) = \frac{1}{2\pi} \sum_{k=-m}^{m} W_n(k) I(\omega_{j+k}). \tag{4.5.2}$$

Similar (heuristic) arguments as the ones applied above imply that, for $j = 1, \ldots, \lfloor \frac{n-1}{2} \rfloor$,

$$E\left[\hat{f}(\omega_j)\right] = \sum_{k=-m}^{m} W_n(k) f(\omega_{j+k}) \approx f(\omega_j) \sum_{k=-m}^{m} W_n(k) = f(\omega_j),$$

$$\mathrm{Var}(\hat{f}(\omega_j)) = \sum_{k=-m}^{m} W_n^2(k) f^2(\omega_{j+k}) \approx f^2(\omega_j) \sum_{k=-m}^{m} W_n^2(k),$$

where the latter converges to zero by assumption on the weights. The periodogram ordinates at different Fourier frequencies are approximately uncorrelated. Similar results hold true also for non-Fourier frequencies. In practice, one often applies various Daniell filters successively to the original periodogram estimate.

Example 4.5.1 Let $(X_t)_{t\in\mathbb{Z}}$ be the MA(1) process

$$X_t = Z_t - .5Z_{t-1}, \qquad t \in \mathbb{Z},$$

where $(Z_t)_{t\in\mathbb{Z}} \sim \mathrm{WN}(0,1)$. From Theorem 4.2.4, we obtain the spectral density of the MA(1) time series as

$$f_X(\lambda) = \frac{1}{2\pi}|\theta(e^{-i\lambda})|^2 = \frac{1}{2\pi}\left[\frac{5}{4} - \cos(\lambda)\right], \qquad \lambda \in (-\pi, \pi].$$

Note that $\cos(\lambda)$ is monotonically increasing in $(-\pi, 0]$ and monotonically decreasing in $[0, \pi]$. Next, we generate $n = 160$ observations of $(X_t)_{t\in\mathbb{Z}}$, which in turn supplies us with 80 Fourier frequencies in $(0, \pi]$. Apply now successively the Daniell filters with $m^{(1)} = 1$, $m^{(2)} = 3$ and $m^{(3)} = 5$. This leads to the ψ-weights

$$\frac{1}{231}(1, 3, 6, 9, 12, 15, 18, 20, 21, 21, 21, 20, 18, 15, 12, 9, 6, 3, 1)$$

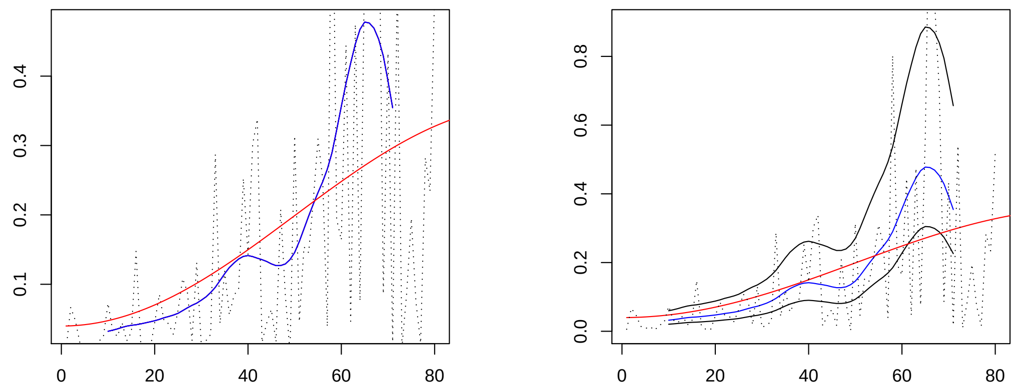

Figure 4.1: The periodogram (dotted), the smoothed periodogram (solid, blue) and the true spectrum (solid, red) of the MA(1) process of Example 4.5.1 (left panel). The corresponding confidence intervals given in Example 4.5.4 (right panel).

and a bandwidth of $m = 9$ so that $2m + 1 = 19$. Figure 4.1 shows in its left panel the original periodogram ordinates (dotted), the true spectrum (red) and the smoothed periodogram (blue) after the application of the three Daniell filters.

A second way to construct estimators for the spectral densities is provided in the following. Using a kernel function w satisfying $w(0) = 1$, $|w(x)| \leq 1$ for all $x \in \mathbb{R}$ and $w(x) = 0$ for all $|x| > 1$, one can define the *lag window estimator*

$$\hat{f}_L(\omega) = \frac{1}{2\pi} \sum_{|h| \leq r} w\Big(\frac{h}{r}\Big) \hat{\gamma}(h) e^{-ih\omega}. \tag{4.5.3}$$

The connection between the earlier estimators and the lag window estimators is as follows. Define the spectral window

$$W(\omega) = \frac{1}{2\pi} \sum_{|h| \leq r} w\Big(\frac{h}{r}\Big) e^{-ih\omega}.$$

Define moreover the extension of the periodogram $\tilde{I}_n(\omega) = \sum_{|h|<n} \hat{\gamma}(h) e^{-ih\omega}$. Then, $\tilde{I}_n(\omega_j) = I_n(\omega_j)$ for all Fourier frequencies and also $\hat{\gamma}(h) = \frac{1}{2\pi} \int_{(-\pi,\pi]} e^{ih\lambda} \tilde{I}_n(\lambda) d\lambda$. There-

fore, we can write

$$\begin{aligned}
\hat{f}_L(\omega) &= \frac{1}{(2\pi)^2}\sum_{|h|\le r} w\Big(\frac{h}{r}\Big)\int_{-\pi}^{\pi} e^{-ih(\omega-\lambda)}\tilde{I}_n(\lambda)d\lambda \\
&= \frac{1}{(2\pi)^2}\int_{-\pi}^{\pi}\Bigg[\sum_{|h|\le r} w\Big(\frac{h}{r}\Big)e^{-ih(\omega-\lambda)}\Bigg]\tilde{I}_n(\lambda)d\lambda \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} W(\omega-\lambda)\tilde{I}_n(\lambda)d\lambda \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} W(\lambda)\tilde{I}_n(\omega+\lambda)d\lambda \\
&\approx \frac{1}{2\pi}\sum_{|k|\le n/2}\frac{2\pi}{n}W(\omega_k)\tilde{I}(\omega+\omega_j) \\
&= \frac{1}{2\pi}\sum_{|k|\le n/2} W_n(k)\tilde{I}(\omega+\omega_j) \\
&\approx \frac{1}{2\pi}\sum_{|k|\le n/2} W_n(k)I(g(n,\omega)+\omega_j),
\end{aligned}$$

where $W_n(k) = \frac{2\pi}{n}W(\omega)$, which is similar to (4.5.2). One can now show that, under regularity conditions,

$$\frac{n}{r}\mathrm{Var}(\hat{f}_L(\omega)) \to \begin{cases} 2f^2(\omega)\int_{-1}^{1} w^2(x)dx, & \omega = 0, \pi. \\ f^2(\omega)\int_{-1}^{1} w^2(x)dx, & \omega \in (0,1). \end{cases}$$

In the following, we discuss a number of popular choices for the kernel function w and the spectral window W. Additional properties related to Fourier approximations of certain functions are briefly discussed in the section below.

Example 4.5.2 (Rectangular Window) Let $w(x) = 1$ for $|x| \le 1$ and $w(x) = 0$ for $|x| > 1$. The spectral window is then

$$W(\omega) = \frac{1}{2\pi}\sum_{|h|\le r} w\Big(\frac{h}{r}\Big)e^{-ih\omega} = \frac{1}{2\pi}\sum_{|h|\le r} e^{-ih\omega} = \frac{1}{2\pi}\frac{\sin((r+1/2)\omega}{\sin(\omega/2)},$$

which is the Dirichlet kernel. In this case, $\mathrm{Var}(\hat{f}_L(\omega)) \sim \frac{2r}{n}f^2(\omega)$. Note that using the rectangular window can lead to negative values for the spectral density estimator because the Dirichlet kernel can be negative.

Example 4.5.3 (Bartlett Window) Let $w(x) = 1 - |x|$ for $|x| \le 1$ and $w(x) = 0$ for $|x| > 1$. The spectral window is then

$$W(\omega) = \frac{1}{2\pi r}\frac{\sin^2(r\omega/2)}{sin^2(\omega/2)}$$

and therefore always positive. The large sample variance satisfies $\mathrm{Var}(\hat{f}_L(\omega)) \sim \frac{2r}{3n}f^2(\omega)$.

Other choices for the spectral window include Blackman and Tukey's cosine bell taper and the Parzen window. Their definitions and properties are discussed in Brockwell and Davis (1991).

To further assess the quality of the spectral density estimator, we construct asymptotic confidence intervals. Recall that our first class of estimators can be given in the form

$$\hat{f}(\omega_j) = \frac{1}{2\pi} \sum_{k=-m}^{m} W_n(k) I(\omega_j + \omega_k). \tag{4.5.4}$$

From the previous section, we know that the collection of periodogram ordinates

$$\left\{ \frac{I(\omega_j + \omega_k)}{\pi f(\omega_j + \omega_k)} : 0 < \omega_j + \omega_k < \pi \right\}$$

is (asymptotically) independent with each variable having a χ^2 distribution with two degrees of freedom. This implies that $\hat{f}(\omega_j)$ is approximately given as a linear combination of $\chi^2(2)$ random variables. This linear combination can in turn be approximated by a single scaled $\chi^2(\nu)$ random variable. To this end, denote by Y a $\chi^2(\nu)$ random variable. We will match mean and variance of $\hat{f}(\omega_j)$ with those of cY. Then, we obtain

$$c\nu = f(\omega_j),$$

$$2c^2\nu = f^2(\omega_j) \sum_{k=-m}^{m} W_n^2(k).$$

Solving for c and ν gives the parameters

$$c = \frac{f(\omega_j)}{2} \sum_{k=-m}^{m} W_n^2(k) \qquad \text{and} \qquad \nu = 2 \left(\sum_{k=-m}^{m} W_n^2(k) \right)^{-1}.$$

To obtain a confidence interval, observe first that the random variable $\nu\hat{f}(\omega_j)/f(\omega_j)$ is distributed as Y. Therefore,

$$P\left(\chi^2_{\alpha/2}(\nu) < \frac{\nu\hat{f}(\omega_j)}{f(\omega_j)} < \chi^2_{1-\alpha/2}(\nu) \right) = 1 - \alpha$$

which leads to the approximate $1-\alpha$ level confidence interval

$$I_{1-\alpha} = \left(\frac{\nu\hat{f}(\omega_j)}{\chi^2_{1-\alpha/2}(\nu)}, \frac{\nu\hat{f}(\omega_j)}{\chi^2_{\alpha/2}(\nu)} \right), \tag{4.5.5}$$

where $\chi^2_{\alpha/2}(\nu)$ and $\chi^2_{1-\alpha/2}(\nu)$ denote the corresponding quantiles.

Example 4.5.4 (Continuation of Example 4.5.1) For the realization of the MA(1) process of Example 4.5.1, we obtain the following confidence interval. First, we compute that $\sum_{k=-9}^{9} W_{160}^2(k) = .07052$. This gives $\nu = 28.36$ along with quantiles

$$\chi^2_{.025}(28.36) = 15.3079 \qquad \text{and} \qquad \chi^2_{.975}(28.36) = 44.4608$$

in the case $\alpha = .05$. The resulting 95% confidence interval is displayed in the right panel of Figure 4.1. It is apparent that the width grows with the values of $\hat{f}(\omega_j)$.

To compute confidence intervals with constant width over all Fourier frequencies, we can take logarithms in (4.5.5). This will then yield the approximate $1-\alpha$ level confidence interval

$$I_{1-\alpha} = \left(\ln \hat{f}(\omega_j) + \ln \nu - \ln \chi^2_{1-\alpha/2}(\nu), \ln \hat{f}(\omega_j) + \ln \nu - \ln \chi^2_{\alpha/2}(\nu)\right). \qquad (4.5.6)$$

for $\ln f(\omega_j)$.

On the other hand, one could approximate $\hat{f}(\omega_j)$ for large n by a sum of $2m+1$ independent random variables. The central limit theorem implies then via Lindeberg's condition that

$$\hat{f}(\omega_j) = AN\left(f(\omega_j), f^2(\omega_j)\sum_{k=-m}^{m} W_n^2(k)\right).$$

This immediately gives the asymptotic $1-\alpha$ level confidence interval

$$\hat{f}(\omega_j) \pm z_{\alpha/2}\hat{f}(\omega_j)\sqrt{\sum_{k=-m}^{m} W_n^2(k)}.$$

Since the width of the interval depends on the values of $\hat{f}(\omega_j)$ it is standard to compute the confidence intervals for $\ln f(\omega_j)$ instead. To do so, we apply that

$$\ln \hat{f}(\omega_j) = AN\left(\ln f(\omega_j), \sum_{k=-m}^{m} W_n^2(k)\right),$$

to obtain

$$\ln \hat{f}(\omega_j) \pm z_{\alpha/2}\sqrt{\sum_{k=-m}^{m} W_n^2(k)}. \qquad (4.5.7)$$

Example 4.5.5 (Continuation of Example 4.5.1) For the realization of the MA(1) process of Example 4.5.4, we obtain the following 95% confidence intervals

$$\left(\ln \hat{f}(\omega_j) - .450,\ \ln \hat{f}(\omega_j) + .617\right)$$

and

$$\ln \hat{f}(\omega_j) \pm .520$$

from (4.5.6) and (4.5.7), respectively.

4.6 Fourier series approximations

In this section we study the approximation of functions $f \in L^2[-\pi,\pi]$ by the Fourier series

$$S_n f = \sum_{j=-n}^{n} \langle f, e_j\rangle e_j,$$

where $e_j(x) = e^{ijx}$ for $j \in \mathbb{Z}$ is an orthonormal system in $L^2[-\pi, \pi]$ and $\langle f, q \rangle = E[f\bar{g}]$ is the usual inner product. It can be shown that $S_n f$ converges in mean square to the infinite series $Sf = \sum_{j \in \mathbb{Z}} \langle f, e_j \rangle e_j$ and that $Sf = f$. Moreover, we can write that

$$\begin{aligned} S_n f(x) &= \sum_{j=-n}^{n} \langle f, e_j \rangle e_j(x) \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x - y) D_n(y) dy, \end{aligned}$$

where the convolution on the right-hand side is with the *Dirichlet kernel*

$$D_n(y) = \sum_{j=-n}^{n} e^{ijy} = \frac{\sin((n + 1/2)y)}{\sin(y/2)}, \qquad y \neq 0,$$

and $D_n(0) = 2n + 1$. The kernel $D_n(y)$ can produce negative values. This is undesirable if the function to be approximated is a nonnegative spectral density. As an alternative one can use the uniform approximation

$$\tilde{S}_n f = \frac{1}{n}(S_0 f + S_1 f + \ldots + S_{n-1} f) \to f$$

which is valid for continuous functions on $[-\pi, \pi]$ such that $f(-\pi) = f(\pi)$. The latter may be rewritten as

$$\tilde{S}_n f(x) = \int_{-\pi}^{\pi} f(x - y) K_n(y) dy,$$

where the convolution is with the *Fejer kernel*

$$K_n(y) = \frac{1}{2\pi n} \sum_{j=0}^{n-1} D_j(y) = \frac{1}{2\pi n} \frac{\sin^2(ny/2)}{\sin^2(y/2)}, \qquad y \neq 0,$$

and $K_n(0) = n/(2\pi)$. The Fejer kernel is positive, even, has period 2π, integrates to one and is more and more concentrated around zero as n increases. The relation of Dirichlet and Fejer kernel to spectral windows was given in the previous section.

4.7 Summary

In this chapter, we have shown how one can perform time series analysis in the frequency domain. At first, we have identified the spectral distribution and density function as the natural frequency domain counterparts of the ACVF, which is the main statistical tool in the time domain. As an estimator for the spectral density, we have introduced the periodogram. The periodogram, however, turns out to be an inconsistent estimator, so that improvements are needed for a meaningful statistical analysis. These can be achieved by smoothing the periodogram with the help of kernel functions which can be applied both in the time domain or directly in the frequency domain.