Chapter 17

Wide Sense Stationary Random Processes

17.1 Introduction

Having introduced the concept of a random process in the previous chapter, we now wish to explore an important subclass of stationary random processes. This is motivated by the very restrictive nature of the stationarity condition, which although mathematically expedient, is almost never satisfied in practice. A somewhat weaker type of stationarity is based on requiring the mean to be a constant in time and the covariance sequence to depend only on the separation in time between the two samples. We have already encountered these types of random processes in Examples 16.9–16.11. Such a random process is said to be stationary in the wide sense or *wide sense stationary* (WSS). It is also termed a *weakly stationary* random process to distinguish it from a stationary process, which is said to be *strictly stationary.* We will use the former terminology to refer to such a process as a WSS random process. In addition, as we will see in Chapter 19, if the random process is Gaussian, then wide sense stationarity implies stationarity. For this reason alone, it makes sense to explore WSS random processes since the use of Gaussian random processes for modeling is ubiquitous.

Once we have discussed the concept of a WSS random process, we will be able to define an extremely important measure of the WSS random process—the *power spectral density* (PSD). This function extends the idea of analyzing the behavior of a deterministic signal by decomposing it into a sum of sinusoids of different frequencies to that of a random process. The difference now is that the amplitudes and phases of the sinusoids will be random variables and so it will be convenient to quantify the *average power* of the various sinusoids. This description of a random phenomenon is important in nearly every scientific field that is concerned with the analysis of time series data such as systems control [Box and Jenkins 1970], signal processing [Schwartz and Shaw 1975], economics [Harvey 1989], geophysics [Robinson 1967], vibration testing [McConnell 1995], financial analysis [Taylor 1986], and others. As an example, in Figure 17.1 the Wolfer sunspot data [Tong 1990] is shown, with the data points connected by straight lines for easier viewing. It measures the average number of sunspots visually observed through a telescope each year. The importance of the sunspot number is that as it increases, an increase in solar flares occurs. This has the effect of disrupting all radio communications as the solar flare particles reach the earth. Clearly from the data we see a periodic type property. The estimated PSD of this data set is shown in Figure 17.2. We see that the distribution of power versus frequency is highest at a frequency of about 0.09 cycles per year. This means that the random process exhibits a large periodic component with a period of about $1/0.09 \approx 11$ years per cycle, as is also evident from Figure 17.1. This is a powerful prediction tool and therefore is of great interest. How the PSD is actually estimated will be discussed in this chapter, but before doing so, we will need to lay some groundwork.

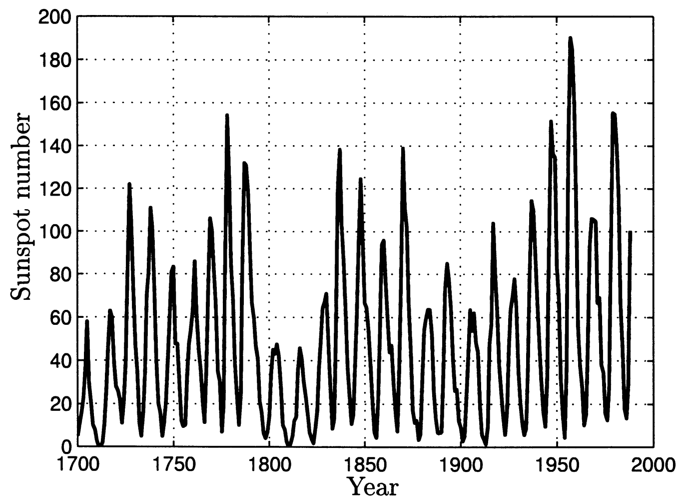


Figure 17.1: Annual number of sunspots – Wolfer sunspot data.

17.2 Summary

A less restrictive form of stationarity, termed wide sense stationarity, is defined by (17.4) and (17.5). The conditions require the mean to be the same for all n and the covariance sequence to depend only on the time difference between the samples. A random process that is stationary is also wide sense stationary as shown in Section 17.3. The autocorrelation sequence is defined by (17.9) with n being arbitrary. It is the covariance between two samples separated by k units for a zero mean WSS random process. Some of its properties are summarized by Properties 17.1–17.4. Under certain conditions the mean of a WSS random process can be found by using

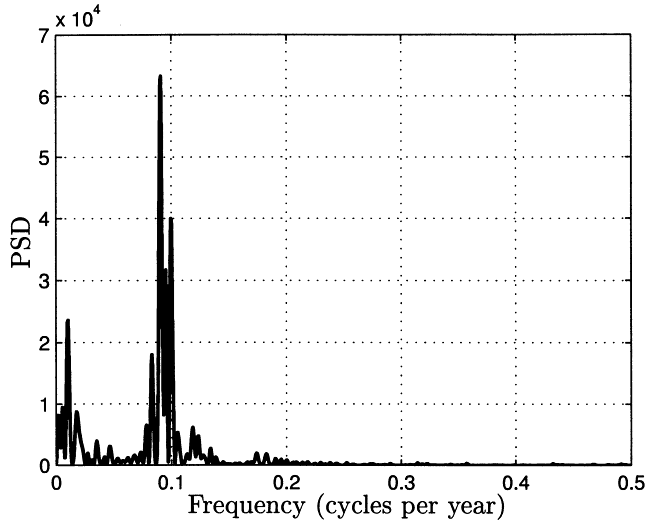


Figure 17.2: Estimated power spectral density for Wolfer sunspot data of Figure 17.1. The sample mean has been computed and removed from the data prior to estimation of the PSD.

the temporal average of (17.25). Such a process is said to be ergodic in the mean. For this to be true the variance of the temporal average given by (17.28) must converge to zero as the number of samples averaged becomes large. The power spectral density (PSD) of a WSS random process is defined by (17.30) and can be evaluated more simply using (17.34). The latter relationship says that the PSD is the Fourier transform of the autocorrelation sequence. It measures the amount of average power per unit frequency or the distribution of average power with frequency. Some of its properties are summarized in Properties 17.7–17.12. From a finite segment of a realization of the random process the autocorrelation sequence can be estimated using (17.43) and the PSD can be estimated by using the averaged periodogram estimate of (17.44) and (17.45). The analogous definitions for a continuous-time WSS random process are given in Section 17.8. Also, an important example is described that relates sampled continuous-time white Gaussian noise to discrete-time white Gaussian noise. Finally, an application of the use of PSDs to random vibration testing is given in Section 17.9.

17.3 Definition of WSS Random Process

Consider a discrete-time random process $X[n]$, which is defined for $-\infty < n < \infty$ with n an integer. Previously, we defined the mean and covariance sequences of

$X[n]$ to be

$$\mu_X[n] = E[X[n]] \quad -\infty < n < \infty \tag{17.1}$$

$$c_X[n_1, n_2] = E[(X[n_1] - \mu_X[n_1])(X[n_2] - \mu_X[n_2])] \quad \begin{array}{c} -\infty < n_1 < \infty \\ -\infty < n_2 < \infty \end{array} \tag{17.2}$$

where n_1, n_2 are integers. Having knowledge of these sequences allows us to assess important characteristics of the random process such as the mean level and the correlation between samples. In fact, based on only this information we are able to predict $X[n_2]$ based on observing $X[n_1] = x[n_1]$ as

$$\hat{X}[n_2] = \mu_X[n_2] + \frac{c_X[n_1, n_2]}{c_X[n_1, n_1]}(x[n_1] - \mu_X[n_1]) \tag{17.3}$$

which is just the usual linear prediction formula of (7.41) with x replaced by $x[n_1]$ and Y replaced by $X[n_2]$, and which makes use of the mean and covariance sequences defined in (17.1) and (17.2), respectively. However, since in general the mean and covariance change with time, i.e., they are nonstationary, it would be exceedingly difficult to estimate them in practice. To extend the practical utility we would like the mean not to depend on time and the covariance only to depend on the separation between samples or on $|n_2 - n_1|$. This will allow us to estimate these quantities as described later. Thus, we are led to a weaker form of stationarity known as *wide sense stationarity*. A random process is defined to be WSS if

$$\mu_X[n] = \mu \text{ (a constant)} \quad -\infty < n < \infty \tag{17.4}$$

$$c_X[n_1, n_2] = g(|n_2 - n_1|) \quad -\infty < n_1 < \infty, -\infty < n_2 < \infty \tag{17.5}$$

for some function g. Note that since

$$c_X[n_1, n_2] = E[X[n_1]X[n_2]] - E[X[n_1]]E[X[n_2]]$$

these conditions are equivalent to requiring that $X[n]$ satisfy

$$\begin{aligned} E[X[n]] &= \mu & -\infty < n < \infty \\ E[X[n_1]X[n_2]] &= h(|n_2 - n_1|) & -\infty < n_1 < \infty, -\infty < n_2 < \infty \end{aligned}$$

for some function h. *The mean should not depend on time and the average value of the product of two samples should depend only upon the time interval between the samples.* Some examples of WSS random processes have already been given in Examples 16.9–16.11. For the MA process of Example 16.10 we showed that

$$\begin{aligned} \mu_X[n] &= 0 \quad -\infty < n < \infty \\ c_X[n_1, n_2] &= \begin{cases} \frac{1}{2}\sigma_U^2 & |n_2 - n_1| = 0 \\ \frac{1}{4}\sigma_U^2 & |n_2 - n_1| = 1 \\ 0 & |n_2 - n_1| > 1. \end{cases} \end{aligned}$$

It is seen that every random variable $X[n]$ for $-\infty < n < \infty$ has a mean of zero and the covariance for two samples depends only on the time interval between the samples, which is $|n_2 - n_1|$. Also, this implies that the variance does not depend on time since $\text{var}(X[n]) = c_X[n,n] = \sigma_U^2/2$ for $-\infty < n < \infty$. In contrast to this behavior consider the random processes for which typical realizations are shown in Figure 16.7. In Figure 16.7a the mean changes with time (with the variance being constant) and in Figure 16.7b the variance changes with time (with the mean being constant). Clearly, these random processes are not WSS.

A WSS random process is a special case of a stationary random process. To see this recall that if $X[n]$ is stationary, then from (16.3) with $N = 1$ and $n_1 = n$, we have

$$p_{X[n+n_0]} = p_{X[n]} \qquad \text{for all } n \text{ and for all } n_0.$$

As a consequence, if we let $n = 0$, then

$$p_{X[n_0]} = p_{X[0]} \qquad \text{for all } n_0$$

and since the PDF does not depend on the particular time n_0, the mean must not depend on time. Thus,

$$\mu_X[n] = \mu \qquad -\infty < n < \infty. \tag{17.6}$$

Next, using (16.3) with $N = 2$, we have

$$p_{X[n_1+n_0],X[n_2+n_0]} = p_{X[n_1],X[n_2]} \qquad \text{for all } n_1, n_2 \text{ and } n_0. \tag{17.7}$$

Now if $n_0 = -n_1$ we have from (17.7)

$$p_{X[0],X[n_2-n_1]} = p_{X[n_1],X[n_2]}$$

and if $n_0 = -n_2$, we have

$$p_{X[n_1-n_2],X[0]} = p_{X[n_1],X[n_2]}.$$

This results in

$$\begin{aligned} p_{X[n_1],X[n_2]} &= p_{X[0],X[n_2-n_1]} \\ p_{X[n_1],X[n_2]} &= p_{X[n_1-n_2],X[0]} \end{aligned}$$

which leads to

$$\begin{aligned} E[X[n_1]X[n_2]] &= E[X[0]X[n_2-n_1]] \\ E[X[n_1]X[n_2]] &= E[X[n_1-n_2]X[0]] = E[X[0]X[n_1-n_2]]. \end{aligned}$$

Finally, these two conditions combine to give

$$E[X[n_1]X[n_2]] = E[X[0]X[|n_2-n_1|]] \tag{17.8}$$

which along with the mean being constant with time yields the second condition for wide sense stationarity of (17.5) that

$$c_X[n_1, n_2] = E[X[n_1]X[n_2]] - E[X[n_1]]E[X[n_2]] = E[X[0]X[|n_2 - n_1|]] - \mu^2.$$

This proves the assertion that a stationary random process is WSS but the converse is not generally true (see Problem 17.5).

17.4 Autocorrelation Sequence

If $X[n]$ is WSS, then as we have seen $E[X[n_1]X[n_2]]$ depends only on the separation in time between the samples. We can therefore define a new joint moment by letting $n_1 = n$ and $n_2 = n + k$ to yield

$$r_X[k] = E[X[n]X[n+k]] \tag{17.9}$$

which is called the *autocorrelation sequence* (ACS). It depends only on the time difference between samples which is $|n_2 - n_1| = |(n+k) - n| = |k|$ so that *the value of n used in the definition is arbitrary.* It is termed the *auto*correlation sequence (ACS) since it measures the correlation between two samples of the *same* random process. Later we will have occasion to define correlation between two different random processes (see Section 19.3). Note that the time interval between samples is also called the *lag*. An example of the computation of the ACS is given next.

Example 17.1 – A Differencer

Define a random process as $X[n] = U[n] - U[n-1]$, where $U[n]$ is an IID random process with mean μ and variance σ_U^2. A realization of this random process for which $U[n]$ is a Gaussian random variable for all n is shown in Figure 17.3. Although $U[n]$ was chosen here to be a sequence of Gaussian random variables for the sake of displaying the realization in Figure 17.3, the ACS to be found will be the same regardless of the PDF of $U[n]$. This is because it relies on *only the first two moments of $U[n]$ and not its PDF.* The ACS is found as

$$\begin{aligned}
r_X[k] &= E[X[n]X[n+k]] \\
&= E[(U[n] - U[n-1])(U[n+k] - U[n+k-1])] \\
&= E[U[n]U[n+k]] - E[U[n]U[n+k-1]] \\
&\quad - E[U[n-1]U[n+k]] + E[U[n-1]U[n+k-1]].
\end{aligned}$$

But for $n_1 \neq n_2$

$$\begin{aligned}
E[U[n_1]U[n_2]] &= E[U[n_1]]E[U[n_2]] \qquad \text{(independence)} \\
&= \mu^2
\end{aligned}$$

and for $n_1 = n_2 = n$

$$E[U[n_1]U[n_2]] = E[U^2[n]] = E[U^2[0]] = \sigma_U^2 + \mu^2 \quad \text{(identically distributed)}.$$

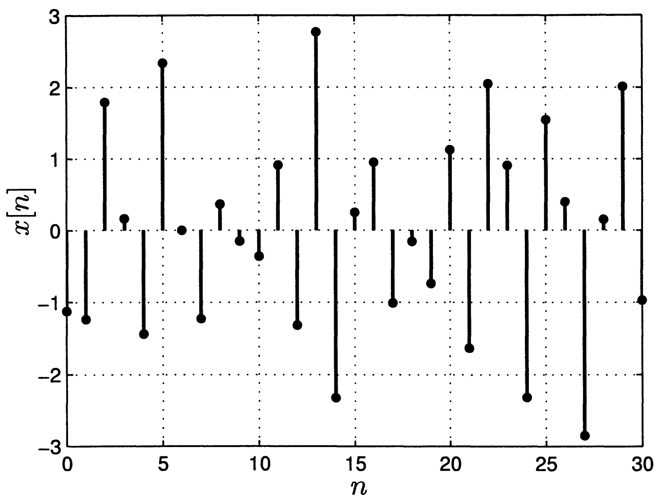


Figure 17.3: Typical realization of a differenced IID Gaussian random process with $U[n] \sim \mathcal{N}(1, 1)$.

Combining these results we have that

$$E[U[n_1]U[n_2]] = \mu^2 + \sigma_U^2 \delta[n_2 - n_1]$$

and therefore the ACS becomes

$$r_X[k] = 2\sigma_U^2 \delta[k] - \sigma_U^2 \delta[k-1] - \sigma_U^2 \delta[k+1]. \tag{17.10}$$

This is shown in Figure 17.4. Several observations can be made. The only nonzero correlation is between adjacent samples and this correlation is negative. This accounts for the observation that the realization shown in Figure 17.3 exhibits many adjacent samples that are opposite in sign. Some other observations are that $r_X[0] > 0$, $|r_X[k]| \leq r_X[0]$ for all k, and finally $r_X[-k] = r_X[k]$. In words, the ACS has a maximum at $k = 0$, which is positive, and is a symmetric sequence about $k = 0$ (also called an *even sequence*). These properties hold in general as we now prove.

◇

Property 17.1 – ACS is positive for the zero lag or $r_X[0] > 0$.

Proof:

$$r_X[k] = E[X[n]X[n+k]] \qquad \text{(definition)}$$

so that with $k = 0$ we have $r_X[0] = E[X^2[n]] > 0$.

□

Note that $r_X[0]$ is the *average power* of the random process at all sample times n. One can view $X[n]$ as the voltage across a 1 ohm resistor and hence $x^2[n]/1$

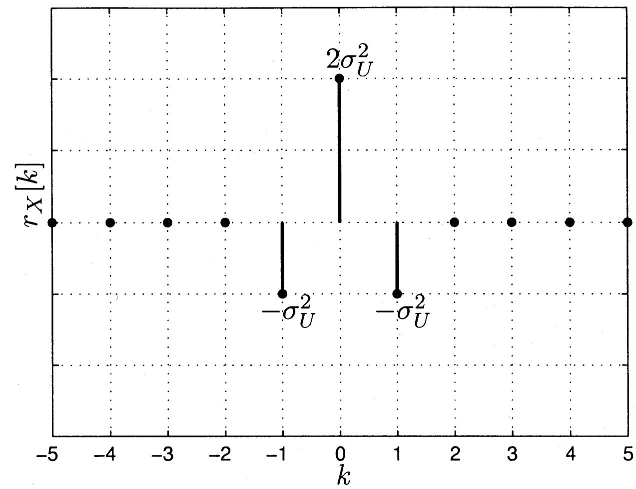


Figure 17.4: Autocorrelation sequence for differenced random process.

is the power for any particular realization of $X[n]$ at time n. The *average* power $E[X^2[n]] = r_X[0]$ does not change with time.

Property 17.2 – ACS is an even sequence or $r_X[-k] = r_X[k]$.

Proof:

$$\begin{aligned} r_X[k] &= E[X[n]X[n+k]] \qquad \text{(definition)} \\ r_X[-k] &= E[X[n]X[n-k]] \end{aligned}$$

and letting $n = m + k$ since the choice of n in the definition of the ACS is arbitrary, we have

$$\begin{aligned} r_X[-k] &= E[X[m+k]X[m]] \\ &= E[X[m]X[m+k]] \\ &= E[X[n]X[n+k]] \qquad \text{(ACS not dependent on } n) \\ &= r_X[k]. \end{aligned}$$

□

Property 17.3 – Maximum absolute value of ACS is at $k = 0$ or $|r_X[k]| \leq r_X[0]$.

Note that it is possible for some values of $r_X[k]$ for $k \neq 0$ to also equal $r_X[0]$. As an example, for the randomly phased sinusoid of Example 16.11 we had $c_X[n_1, n_2] = \frac{1}{2}\cos[2\pi(0.1)(n_2 - n_1)]$ with a mean of zero. Thus, $r_X[k] = \frac{1}{2}\cos[2\pi(0.1)k]$ and therefore $r_X[10] = r_X[0]$. Hence, the property says that no value of the ACS can *exceed* $r_X[0]$, although there may be multiple values of the ACS that are equal to $r_X[0]$.

Proof: The proof is based on the Cauchy-Schwarz inequality, which from Appendix 7A is

$$|E_{V,W}[VW]| \leq \sqrt{E_V[V^2]}\sqrt{E_W[W^2]}$$

with equality holding if and only if $W = cV$ for c a constant. Letting $V = X[n]$ and $W = X[n+k]$, we have

$$|E[X[n]X[n+k]]| \leq \sqrt{E[X^2[n]]}\sqrt{E[X^2[n+k]]}$$

from which it follows that

$$|r_X[k]| \leq \sqrt{r_X[0]}\sqrt{r_X[0]} = |r_X[0]| = r_X[0] \qquad \text{(since } r_X[0] > 0\text{)}.$$

Note that equality holds if and only if $X[n+k] = cX[n]$ for all n. This implies perfect predictability of a sample based on the realization of another sample spaced k units ahead or behind in time (see Problem 17.10 for an example involving periodic random processes).

□

Property 17.4 – ACS measures the predictability of a random process. The correlation coefficient for two samples of a zero mean WSS random process is

$$\rho_{X[n],X[n+k]} = \frac{r_X[k]}{r_X[0]} \tag{17.11}$$

For a nonzero mean the expression is easily modified (see Problem 17.11).
Proof: Recall that the correlation coefficient for two random variables V and W is defined as

$$\rho_{V,W} = \frac{\operatorname{cov}(V,W)}{\sqrt{\operatorname{var}(V)\operatorname{var}(W)}}.$$

Assuming that V and W are zero mean, this becomes

$$\rho_{V,W} = \frac{E_{V,W}[VW]}{\sqrt{E_V[V^2]E_W[W^2]}}$$

and letting $V = X[n]$ and $W = X[n+k]$, we have

$$\begin{aligned}
\rho_{X[n],X[n+k]} &= \frac{E[X[n]X[n+k]]}{\sqrt{E[X^2[n]]E[X^2[n+k]]}} \\
&= \frac{r_X[k]}{\sqrt{r_X[0]r_X[0]}} \\
&= \frac{r_X[k]}{|r_X[0]|} \\
&= \frac{r_X[k]}{r_X[0]} \qquad \text{(from Property 17.1).}
\end{aligned}$$

□

As an example, for the differencer of Example 17.1 we have from Figure 17.4

$$\rho_{X[n],X[n+k]} = \begin{cases} 1 & k=0 \\ -\frac{1}{2} & k=\pm 1 \\ 0 & \text{otherwise.} \end{cases}$$

As mentioned previously, the adjacent samples are negatively correlated and the magnitude of the correlation coefficient is now seen to be 1/2.

We next give some more examples of the computation of the ACS.

Example 17.2 – White noise

White noise is defined as *a WSS random process with zero mean, identical variance* σ^2*, and uncorrelated samples.* It is a more general case of the white noise random process first described in Example 16.9. There we assumed the stronger condition of zero mean *IID* samples (hence they must have the same variance due to the identically distributed assumption and also be uncorrelated due to the independence assumption). In addition, it was assumed there that each sample had a *Gaussian PDF.* Note, however, that the definition given above for white noise does not specify a particular PDF. To find the ACS we note that from the definition of the white noise random process

$$\begin{aligned} r_X[k] &= E[X[n]X[n+k]] \\ &= E[X[n]]E[X[n+k]] = 0 \quad k \neq 0 \qquad \text{(uncorrelated and zero mean samples)} \\ &= E[X^2[n]] = \sigma^2 \quad k = 0 \qquad \text{(equal variance samples).} \end{aligned}$$

Therefore, we have that

$$r_X[k] = \sigma^2 \delta[k]. \tag{17.12}$$

Could you predict $X[1]$ from a realization of $X[0]$?

◇

As an aside, for WSS random processes, we can find the covariance sequence from the ACS and the mean since

$$\begin{aligned} c_X[n_1, n_2] &= E[X[n_1]X[n_2]] - \mu_X[n_1]\mu_X[n_2] \\ &= r_X[n_2 - n_1] - \mu^2. \end{aligned} \tag{17.13}$$

Another property of the ACS that is evident from (17.13) concerns the behavior of the ACS as $k \to \infty$. Letting $n_1 = n$ and $n_2 = n + k$, we have that

$$r_X[k] = c_X[n, n+k] + \mu^2. \tag{17.14}$$

If two samples becomes uncorrelated or $c_X[n, n+k] \to 0$ as $k \to \infty$, then we see that $r_X[k] \to \mu^2$ as $k \to \infty$. Thus, as another property of the ACS we have the following.

Property 17.5 – ACS approaches μ^2 as $k \to \infty$

This assumes that the samples become uncorrelated for large lags, which is usually the case.

□

If the mean is zero, then from (17.14)

$$r_X[k] = c_X[n, n+k] \tag{17.15}$$

and the ACS approaches zero as the lag increases. We continue with some more examples.

◇

Example 17.3 – MA random process

This random process was shown in Example 16.10 to have a zero mean and a covariance sequence

$$c_X[n_1, n_2] = \begin{cases} \frac{\sigma_U^2}{2} & n_1 = n_2 \\ \frac{\sigma_U^2}{4} & |n_2 - n_1| = 1 \\ 0 & \text{otherwise.} \end{cases} \tag{17.16}$$

Since the covariance sequence depends only on $|n_2 - n_1|$, $X[n]$ is WSS from (17.15). Specifically, the ACS follows from (17.15) and (17.16) with $k = n_2 - n_1$ as

$$r_X[k] = \begin{cases} \frac{\sigma_U^2}{2} & k = 0 \\ \frac{\sigma_U^2}{4} & k = \pm 1 \\ 0 & \text{otherwise.} \end{cases}$$

See Figure 16.13 for a plot of the ACS (replace Δn with k.) Could you predict $X[1]$ from a realization of $X[0]$?

◇

Example 17.4 – Randomly phased sinusoid

This random process was shown in Example 16.11 to have a zero mean and a covariance sequence $c_X[n_1, n_2] = \frac{1}{2}\cos[2\pi(0.1)(n_2 - n_1)]$. Since the covariance sequence depends only on $|n_2 - n_1|$, $X[n]$ is WSS. Hence, from (17.15) we have that

$$r_X[k] = \frac{1}{2}\cos[2\pi(0.1)k].$$

See Figure 16.14 for a plot of the ACS (replace Δn with k.) Could you predict $X[1]$ from a realization of $X[0]$?

◇

In determining predictability of a WSS random process, it is convenient to consider the linear predictor, which depends only on the first two moments. Then, the MMSE linear prediction of $X[n_0+k]$ given $x[n_0]$ is from (17.3) and (17.13) with $n_1 = n_0$ and $n_2 = n_0 + k$

$$\hat{X}[n_0+k] = \mu + \frac{r_X[k]-\mu^2}{r_X[0]-\mu^2}(x[n_0]-\mu) \quad \text{for all } k \text{ and } n_0.$$

For a zero mean random process this becomes

$$\begin{aligned} \hat{X}[n_0+k] &= \frac{r_X[k]}{r_X[0]}x[n_0] \\ &= \rho_{X[n_0],X[n_0+k]}x[n_0] \quad \text{for all } k \text{ and } n_0. \end{aligned}$$

One last example is the autoregressive random process which we will use to illustrate several new concepts for WSS random processes.

Example 17.5 – Autoregressive random process

An autoregressive (AR) random process $X[n]$ is defined to be a WSS random process with a zero mean that evolves according to the recursive difference equation

$$X[n] = aX[n-1] + U[n] \qquad -\infty < n < \infty \tag{17.17}$$

where $|a| < 1$ and $U[n]$ is WGN. The WGN random process $U[n]$ (see Example 16.6), has a zero mean and variance σ_U^2 for all n and its samples are all independent with a Gaussian PDF. The name *autoregressive* is due to the *regression* of $X[n]$ upon $X[n-1]$, which is another sample of the same random process, hence, the prefix *auto*. The evolution of $X[n]$ proceeds, for example, as

$$\begin{aligned} &\vdots \\ X[0] &= aX[-1] + U[0] \\ X[1] &= aX[0] + U[1] \\ X[2] &= aX[1] + U[2] \\ &\vdots \end{aligned}$$

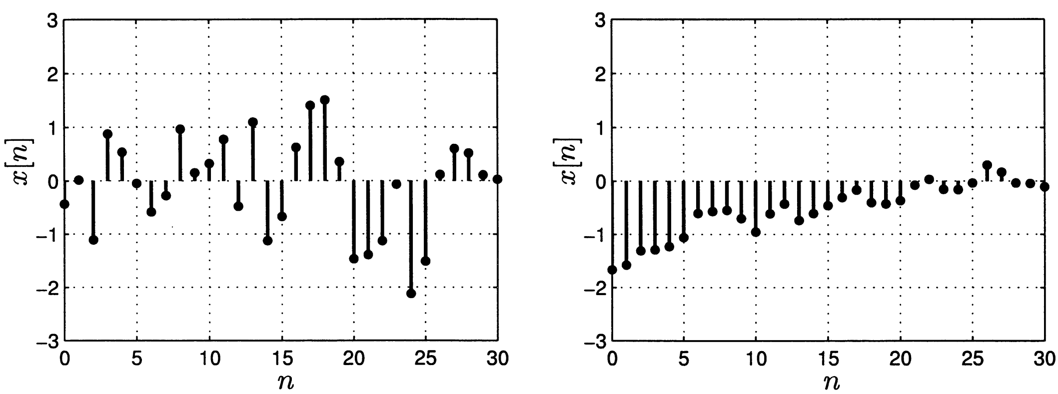
Note that $X[n]$ depends only upon the present and past values of $U[n]$ since for example

$$\begin{aligned} X[2] &= aX[1] + U[2] = a(aX[0]+U[1]) + U[2] = a^2X[0] + aU[1] + U[2] \\ &= a^2(aX[-1]+U[0]) + aU[1] + U[2] = a^3X[-1] + a^2U[0] + aU[1] + U[2] \\ &\vdots \\ &= \sum_{k=0}^{\infty} a^k U[2-k] \end{aligned} \tag{17.18}$$

where the term involving $a^k U[2-k]$ decays to zero as $k \to \infty$ since $|a| < 1$. We see that $X[2]$ depends only on $\{U[2], U[1], \ldots\}$ and it is therefore uncorrelated with $\{U[3], U[4], \ldots\}$. More generally, it can be shown that (see also Problem 19.6)

$$E[X[n]U[n+k]] = 0 \qquad k \geq 1. \tag{17.19}$$

It is seen from (17.18) that in order for the recursion to be stable and hence $X[n]$ to be WSS it is required that $|a| < 1$. The AR random process can be used to model a wide variety of physical random processes with various ACSs, depending upon the choice of the parameters a and σ_U^2. Some typical realizations of the AR random process for different values of a are shown in Figure 17.5. The WGN random process $U[n]$ has been chosen to have a variance $\sigma_U^2 = 1 - a^2$. We will soon see that this choice of variance results in $r_X[0] = 1$ for both AR processes shown in Figure 17.5. The MATLAB code used to generate the realizations shown is given below.



(a) $a = 0.25$, $\sigma_U^2 = 1 - a^2$ (b) $a = 0.98$, $\sigma_U^2 = 1 - a^2$

Figure 17.5: Typical realizations of autoregressive random process with different parameters.

```
clear all
randn('state',0)
a1=0.25;a2=0.98;
varu1=1-a1^2;varu2=1-a2^2;
varx1=varu1/(1-a1^2);varx2=varu2/(1-a2^2); % this is r_X[0]
x1(1,1)=sqrt(varx1)*randn(1,1); % set initial condition X[-1]
                                % see Problems 17.17, 17.18
x2(1,1)=sqrt(varx2)*randn(1,1);
for n=2:31
   x1(n,1)=a1*x1(n-1)+sqrt(varu1)*randn(1,1);
   x2(n,1)=a2*x2(n-1)+sqrt(varu2)*randn(1,1);
end
```

We next derive the ACS. In Chapter 18 we will see how to alternatively obtain the ACS using results from linear systems theory. Using (17.17) we have for $k \geq 1$

$$\begin{aligned} r_X[k] &= E[X[n]X[n+k]] \\ &= E[X[n](aX[n+k-1]+U[n+k])] \\ &= aE[X[n]X[n+k-1]] \qquad \text{(using (17.19))} \\ &= ar_X[k-1]. \end{aligned} \tag{17.20}$$

The solution of this recursive linear difference equation is readily seen to be $r_X[k] = ca^k$, for c any constant and for $k \geq 1$. For $k = 1$ we have that $r_X[1] = ca$ and so from (17.20) $r_X[1] = ar_X[0]$, which implies $c = r_X[0]$. In Problem 17.15 it is shown that

$$r_X[0] = \frac{\sigma_U^2}{1-a^2}$$

so that for all $k \geq 0$, $r_X[k] = r_X[0]a^k$ becomes

$$r_X[k] = \frac{\sigma_U^2}{1-a^2}a^k.$$

Finally, noting that $r_X[-k] = r_X[k]$ from Property 17.2, we obtain the ACS as

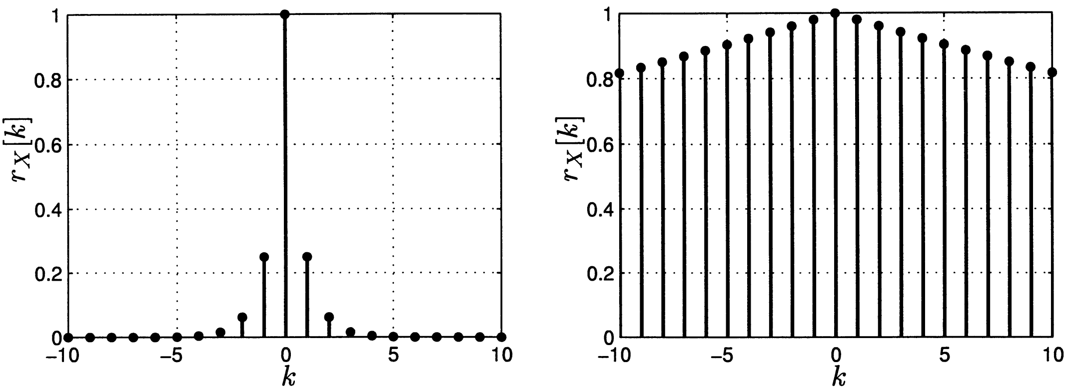
$$r_X[k] = \frac{\sigma_U^2}{1-a^2}a^{|k|} \qquad -\infty < k < \infty. \tag{17.21}$$

(See also Problem 17.16 for an alternative derivation of the ACS.) The ACS is plotted in Figure 17.6 for $a = 0.25$ and $a = 0.98$ and $\sigma_U^2 = 1-a^2$. For both values of a the value of σ_U^2 has been chosen to ensure that $r_X[0] = 1$. Note that for $a = 0.25$ the ACS dies off very rapidly which means that the random process samples quickly become uncorrelated as the separation between them increases. This is consistent with the typical realization shown in Figure 17.5a. For $a = 0.98$ the ACS decays very slowly, indicating a strong positive correlation between samples, and again being consistent with the typical realization shown in Figure 17.5b. In either case the samples become uncorrelated as $k \to \infty$ since $|a| < 1$ and therefore, $r_X[k] \to 0$ as $k \to \infty$ in accordance with Property 17.5. However, the random process with the slower decaying ACS is more predictable.

◊

One last property that is necessary for a sequence to be a valid ACS is the property of positive definiteness. As its name implies, it is related to the positive definite property of the covariance matrix. As an example, consider the random vector $\mathbf{X} = [X[0]\, X[1]]^T$. Then we know from the proof of Property 9.2 (covariance matrix is positive semidefinite) that if $Y = a_0X[0] + a_1X[1]$ cannot be made equal to a constant by any choice of a_0 and a_1, then

$$\text{var}(Y) = \underbrace{\begin{bmatrix} a_0 & a_1 \end{bmatrix}}_{\mathbf{a}^T} \underbrace{\begin{bmatrix} \text{cov}(X[0],X[0]) & \text{cov}(X[0],X[1]) \\ \text{cov}(X[1],X[0]) & \text{cov}(X[1],X[1]) \end{bmatrix}}_{\mathbf{C}_X} \underbrace{\begin{bmatrix} a_0 \\ a_1 \end{bmatrix}}_{\mathbf{a}} > 0.$$



(a) $a = 0.25$, $\sigma_U^2 = 1 - a^2$ (b) $a = 0.98$, $\sigma_U^2 = 1 - a^2$

Figure 17.6: The autocorrelation sequence for autoregressive random processes with different parameters.

Since this holds for all $\mathbf{a} \neq \mathbf{0}$, the covariance matrix $\mathbf{C}_X$ is by definition positive definite (see Appendix C). (If it were possible to choose a_0 and a_1 so that $Y = c$, for c a constant, then $X[1]$ would be perfectly predictable from $X[0]$ as $X[1] = -(a_0/a_1)X[0] + (c/a_1)$. Therefore, we could have $\text{var}(Y) = \mathbf{a}^T\mathbf{C}_X\mathbf{a} = 0$, and $\mathbf{C}_X$ would only be positive *semi*definite.) Now if $X[n]$ is a zero mean WSS random process

$$\text{cov}(X[n_1], X[n_2]) = E(X[n_1]X[n_2]) = r_X[n_2 - n_1]$$

and the covariance matrix becomes

$$\mathbf{C}_X = \begin{bmatrix} r_X[0] & r_X[1] \\ r_X[-1] & r_X[0] \end{bmatrix} = \underbrace{\begin{bmatrix} r_X[0] & r_X[1] \\ r_X[1] & r_X[0] \end{bmatrix}}_{\mathbf{R}_X}.$$

Therefore, the covariance matrix, which we now denote by $\mathbf{R}_X$ and which is called the *autocorrelation matrix*, must be positive definite. This implies that all the *principal minors* (see Appendix C) are positive. For the 2×2 case this means that

$$\begin{aligned} r_X[0] &> 0 \\ r_X^2[0] - r_X^2[1] &> 0 \end{aligned} \qquad (17.22)$$

with the first condition being consistent with Property 17.1 and the second condition producing $r_X[0] > |r_X[1]|$. The latter condition is nearly consistent with Property 17.3 with the slight difference, that $|r_X[1]|$ may equal $r_X[0]$ being excluded. This is because we assumed that $X[1]$ was not perfectly predictable from knowledge of $X[0]$. If we allow perfect predictability, then the autocorrelation matrix is only positive semidefinite and the $>$ sign in the second equation of (17.22) would be replaced with $\geq$. In general the $N \times N$ autocorrelation matrix $\mathbf{R}_X$ is given as the covariance matrix of the *zero mean* random vector $\mathbf{X} = [X[0]\, X[1] \dots X[N-1]]^T$ as

$$\mathbf{R}_X = \begin{bmatrix} r_X[0] & r_X[1] & r_X[2] & \dots & r_X[N-1] \\ r_X[1] & r_X[0] & r_X[1] & \dots & r_X[N-2] \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_X[N-1] & r_X[N-2] & r_X[N-3] & \dots & r_X[0] \end{bmatrix}. \tag{17.23}$$

For a sequence to be a valid ACS the $N \times N$ autocorrelation matrix must be positive semidefinite for all $N = 1, 2, \dots$ and positive definite if we exclude the possibility of perfect predictability [Brockwell and Davis 1987]. This imposes a large number of constraints on $r_X[k]$ and hence not all sequences satisfying Properties 17.1–17.3 are valid ACSs (see also Problem 17.19). In summary, for our last property of the ACS we have the following.

Property 17.6 – ACS is a positive semidefinite sequence.

Mathematically, this means that $r_X[k]$ must satisfy

$$\mathbf{a}^T \mathbf{R}_X \mathbf{a} \geq 0$$

for all $\mathbf{a} = [a_0\, a_1 \dots a_{N-1}]^T$ and where $\mathbf{R}_X$ is the $N \times N$ autocorrelation matrix given by (17.23). This must hold for all $N \geq 1$.

□

17.5 Ergodicity and Temporal Averages

When a random process is WSS, its mean does not depend on time. Hence, the random variables $\dots, X[-1], X[0], X[1], \dots$ all have the same mean. Then, at least as far as the mean is concerned, when we observe a realization of a random process, it is as if we are observing multiple realizations of the same random variable. This suggests that we may be able to determine the value of the mean from a single infinite length realization. To pursue this idea further we plot three realizations of an IID random process whose marginal PDF is Gaussian with mean $\mu_X[n] = \mu = 1$ and a variance $\sigma_X^2[n] = \sigma^2 = 1$ in Figure 17.7. If we let $x_i[18]$ denote the ith realization at time $n = 18$, then by definition of $E[X[18]]$

$$\lim_{M \to \infty} \frac{1}{M} \sum_{m=1}^{M} x_m[18] = E[X[18]] = \mu_X[18] = \mu = 1. \tag{17.24}$$

This is because as we observe all realizations of the random variable $X[18]$ they will conform to the Gaussian PDF (recall that $X[n] \sim \mathcal{N}(1,1)$). In fact, the original definition of expected value was based on the relationship given in (17.24). This

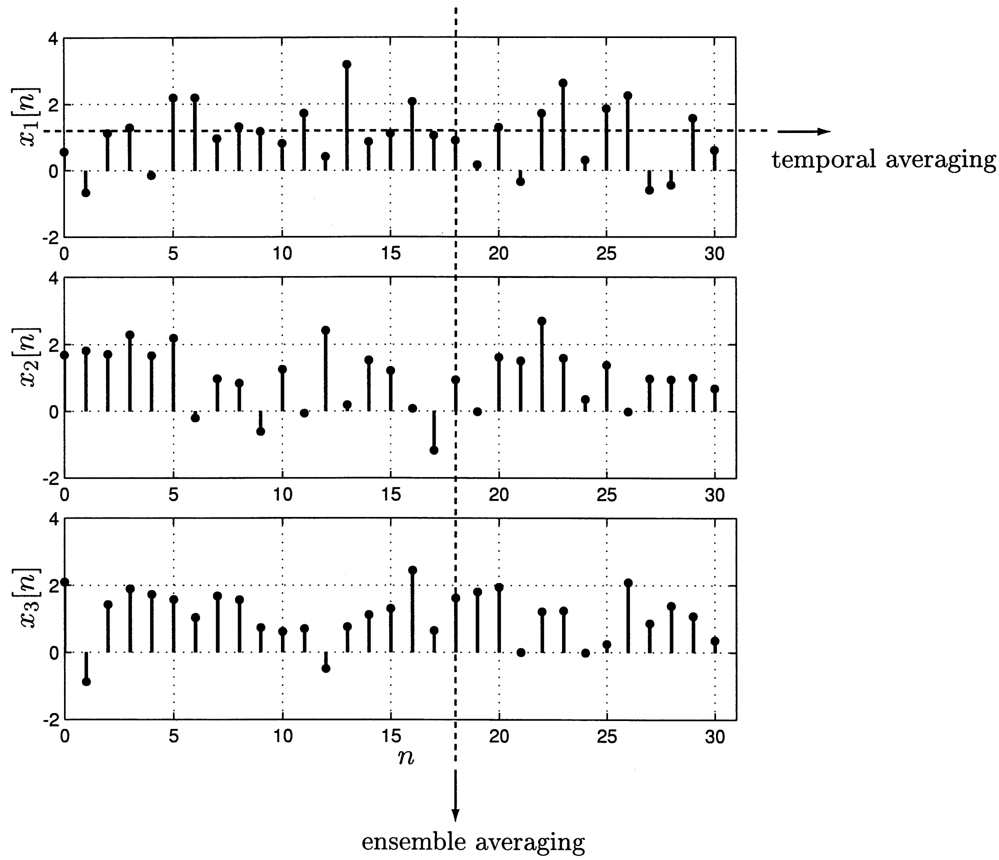


Figure 17.7: Several realizations of WSS random process with $\mu_X[n] = \mu = 1$. Vertical dashed line indicates "ensemble averaging" while horizontal dashed line indicates "temporal averaging."

type of averaging is called "averaging down the ensemble" and consequently is just a restatement of our usual notion of the expected value of a random variable. However, if we are given only a single realization such as $x_1[n]$, then it seems reasonable that

$$\hat{\mu}_N = \frac{1}{N} \sum_{n=0}^{N-1} x_1[n]$$

should also converge to μ as $N \to \infty$. This type of averaging is called "temporal averaging" since we are averaging the samples in time. If it is true that the temporal average converges to μ, then we can state that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} x_1[n] = \mu = E[X[18]] = \lim_{M\to\infty} \frac{1}{M} \sum_{m=1}^{M} x_m[18]$$

and it is said that *temporal averaging is equivalent to ensemble averaging* or that the random process is *ergodic in the mean.* This property is of great practical importance since it assures us that by averaging enough samples of the realization, we can determine the mean of the random process. For the case of an IID random process ergodicity holds due to the law of large numbers (see Chapter 15). Recall that if $X_1, X_2, \ldots, X_N$ are IID random variables with mean μ and variance σ^2, then the sample mean random variable has the property that

$$\frac{1}{N}\sum_{i=1}^{N} X_i \to E[X] = \mu \qquad \text{as } N \to \infty.$$

Hence, if $X[n]$ is an IID random process, the conditions required for the law of large numbers to hold are satisfied, and we can immediately conclude that

$$\hat{\mu}_N = \frac{1}{N}\sum_{n=0}^{N-1} X[n] \to \mu. \tag{17.25}$$

Now the assumptions required for a random process to be IID are overly restrictive for (17.25) to hold. More generally, if $X[n]$ is a WSS random process, then since $E[X[n]] = \mu$, it follows that $E[\hat{\mu}_N] = (1/N)\sum_{n=0}^{N-1} E[X[n]] = \mu$. Therefore, the only further condition required for ergodicity in the mean is that

$$\lim_{N\to\infty} \operatorname{var}(\hat{\mu}_N) = 0.$$

In the case of the IID random process it is easily shown that $\operatorname{var}(\hat{\mu}_N) = \sigma^2/N \to 0$ as $N \to \infty$ and the condition is satisfied. More generally, however, the random process samples are correlated so that evaluation of this variance is slightly more complicated. We illustrate this computation next.

Example 17.6 – General MA random process

Consider the general MA random process given as $X[n] = (U[n] + U[n-1])/2$, where $E[U[n]] = \mu$ and $\operatorname{var}(U[n]) = \sigma_U^2$ for $-\infty < n < \infty$ and the $U[n]$'s are all uncorrelated. This is similar to the MA process of Example 16.10 but is more general in that the mean of $U[n]$ is not necessarily zero, the samples of $U[n]$ are only uncorrelated, and hence, not necessarily independent, and the PDF of each sample need not be Gaussian. The general MA process $X[n]$ is easily shown to be WSS and to have a mean sequence $\mu_X[n] = \mu$ (see Problem 17.20). To determine if it is ergodic in the mean we must compute the $\operatorname{var}(\hat{\mu}_N)$ and show that it converges to zero as $N \to \infty$. Now

$$\operatorname{var}(\hat{\mu}_N) = \operatorname{var}\left(\frac{1}{N}\sum_{n=0}^{N-1} X[n]\right).$$

Since the $X[n]$'s are now correlated, we use (9.26), where $\mathbf{a} = [a_0\, a_1 \ldots a_{N-1}]^T$ with $a_n = 1/N$, to yield

$$\operatorname{var}(\hat{\mu}_N) = \operatorname{var}\left(\sum_{n=0}^{N-1} a_n X[n]\right) = \mathbf{a}^T \mathbf{C}_X \mathbf{a}. \tag{17.26}$$

The covariance matrix has (i, j) element

$$[\mathbf{C}_X]_{ij} = E[(X[i] - E[X[i]])(X[j] - E[X[j]])] \qquad i = 0, 1, \ldots, N-1; j = 0, 1, \ldots, N-1.$$

But

$$\begin{aligned} X[n] - E[X[n]] &= \frac{1}{2}(U[n] + U[n-1]) - \frac{1}{2}(\mu + \mu) \\ &= \frac{1}{2}[(U[n] - \mu) + (U[n-1] - \mu)] \\ &= \frac{1}{2}[\bar{U}[n] + \bar{U}[n-1]] \end{aligned}$$

where $\bar{U}[n]$ is a *zero mean* random variable for each value of n. Thus,

$$\begin{aligned} [\mathbf{C}_X]_{ij} &= \frac{1}{4} E[(\bar{U}[i] + \bar{U}[i-1])(\bar{U}[j] + \bar{U}[j-1])] \\ &= \frac{1}{4}\left(E[\bar{U}[i]\bar{U}[j]] + E[\bar{U}[i]\bar{U}[j-1]] + E[\bar{U}[i-1]\bar{U}[j]] + E[\bar{U}[i-1]\bar{U}[j-1]]\right) \end{aligned}$$

and since $E[\bar{U}[n_1]\bar{U}[n_2]] = \operatorname{cov}(U[n_1], U[n_2]) = \sigma_U^2 \delta[n_2 - n_1]$ (all the $U[n]$'s are uncorrelated), we have

$$[\mathbf{C}_X]_{ij} = \frac{1}{4}\left(\sigma_U^2 \delta[j-i] + \sigma_U^2 \delta[j-1-i] + \sigma_U^2 \delta[j-i+1] + \sigma_U^2 \delta[j-i]\right).$$

Finally, we have the required covariance matrix

$$[\mathbf{C}_X]_{ij} = \begin{cases} \frac{1}{2}\sigma_U^2 & i = j \\ \frac{1}{4}\sigma_U^2 & |i-j| = 1 \\ 0 & \text{otherwise.} \end{cases} \tag{17.27}$$

Using this in (17.26) produces

$$\operatorname{var}(\hat{\mu}_N)$$

$$
\begin{aligned}
&= \mathbf{a}^T \mathbf{C}_X \mathbf{a} \\
&= \begin{bmatrix} \frac{1}{N} & \frac{1}{N} & \cdots & \frac{1}{N} \end{bmatrix}
\begin{bmatrix}
\frac{\sigma_U^2}{2} & \frac{\sigma_U^2}{4} & 0 & 0 & \cdots & 0 & 0 & 0 \\
\frac{\sigma_U^2}{4} & \frac{\sigma_U^2}{2} & \frac{\sigma_U^2}{4} & 0 & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \cdots & \frac{\sigma_U^2}{4} & \frac{\sigma_U^2}{2} & \frac{\sigma_U^2}{4} \\
0 & 0 & 0 & 0 & \cdots & 0 & \frac{\sigma_U^2}{4} & \frac{\sigma_U^2}{2}
\end{bmatrix}
\begin{bmatrix} \frac{1}{N} \\ \frac{1}{N} \\ \vdots \\ \frac{1}{N} \end{bmatrix} \\
&= \frac{1}{N^2}\sum_{i=0}^{N-1}\frac{\sigma_U^2}{2} + \frac{1}{N^2}\sum_{i=0}^{N-2}\frac{\sigma_U^2}{4} + \frac{1}{N^2}\sum_{i=1}^{N-1}\frac{\sigma_U^2}{4} \\
&= \frac{\sigma_U^2}{2N} + \frac{\sigma_U^2}{4}\frac{N-1}{N^2} + \frac{\sigma_U^2}{4}\frac{N-1}{N^2} \to 0 \qquad \text{as } N \to \infty.
\end{aligned}
$$

Finally, we see that the general MA random process is ergodic in the mean.

◇

In general, it can be shown that for a WSS random process to be ergodic in the mean, the variance of the sample mean

$$\operatorname{var}(\hat{\mu}_N) = \frac{1}{N}\sum_{k=-(N-1)}^{N-1}\left(1 - \frac{|k|}{N}\right)(r_X[k] - \mu^2) \tag{17.28}$$

must converge to zero as $N \to \infty$ (see Problem 17.23 for the derivation of (17.28)). For this to occur, the covariance sequence $r_X[k] - \mu^2$ must decay to zero at a fast enough rate as $k \to \infty$, which is to say that as the samples are spaced further and further apart, they must eventually become uncorrelated. A little reflection on the part of the reader will reveal that ergodicity requires a single realization of the random process to display the behavior of the entire ensemble of realizations. If not, ergodicity will not hold. Consider the following simple nonergodic random process.

Example 17.7 – Random DC level

Define a random process as $X[n] = A$ for $-\infty < n < \infty$, where $A \sim \mathcal{N}(0,1)$. Some realizations are shown in Figure 17.8. This random process is WSS since

$$
\begin{aligned}
\mu_X[n] &= E[X[n]] = E[A] = 0 = \mu & -\infty < n < \infty \quad & \text{(not dependent on } n\text{)} \\
r_X[k] &= E[X[n]X[n+k]] = E[A^2] = 1 & & \text{(not dependent on } n\text{)}.
\end{aligned}
$$

However, it should be clear that $\hat{\mu}_N$ will not converge to $\mu = 0$. Referring to the realization $x_1[n]$ in Figure 17.8, the sample mean will produce -0.43 no matter how large N becomes. In addition, it can be shown that $\operatorname{var}(\hat{\mu}_N) = 1$ (see Problem 17.24). Each realization is *not* representative of the *ensemble of realizations.*

◇

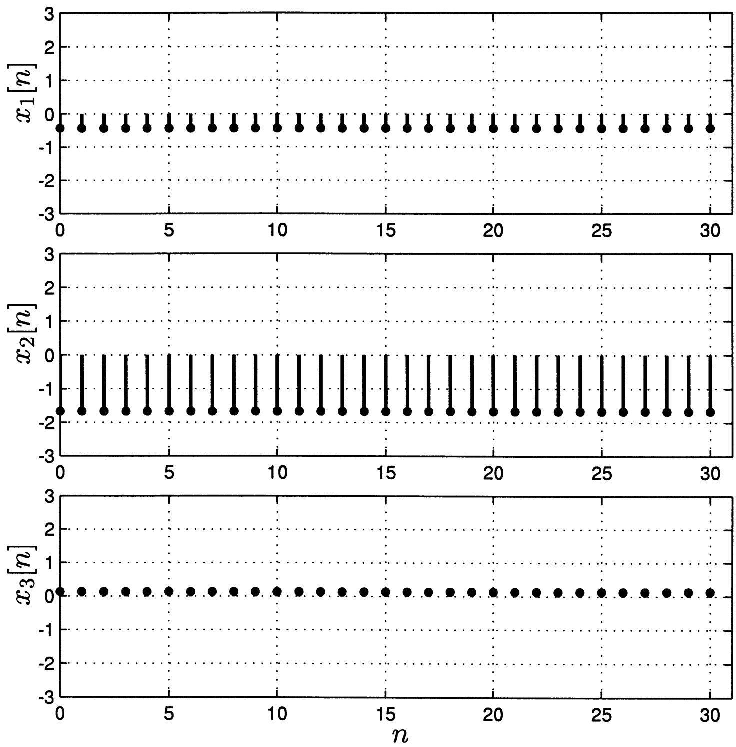


Figure 17.8: Several realizations of the random DC level process.

17.6 The Power Spectral Density

The ACS measures the correlation between samples of a WSS random process. For example, the AR random process was shown to have the ACS

$$r_X[k] = \frac{\sigma_U^2}{1-a^2} a^{|k|}$$

which for $a = 0.25$ and $a = 0.98$ is shown in Figure 17.6, along with some typical realizations in Figure 17.5. Note that when the ACS dies out rapidly (see Figure 17.6a), the realization is more rapidly varying in time (see Figure 17.5a). In contrast, when the ACS decays slowly (see Figure 17.6b), the realization varies slowly (see Figure 17.5b). It would seem that the ACS is related to the rate of change of the random process. For deterministic signals the rate of change is usually measured by examining a discrete-time Fourier transform [Jackson 1991]. Signals with high frequency content exhibit rapid fluctutations in time while signals with only low frequency content exhibit slow variations in time. For WSS random processes we will be interested in the *power* at the various frequencies. In particular, we will introduce the measure known as the *power spectral density* (PSD) and show that it quantifies the distribution of power with frequency. Before doing so, however, we consider the following deterministically motivated measure of power with frequency based on the discrete-time Fourier transform

$$\hat{P}_X(f) = \frac{1}{N}\left|\sum_{n=0}^{N-1} X[n]\exp(-j2\pi f n)\right|^2. \tag{17.29}$$

This is a normalized version of the magnitude-squared discrete-time Fourier transform of the random process over the time interval $0 \le n \le N-1$. It is called the *periodogram* since its original purpose was to find periodicities in random data sets [Schuster 1898]. In (17.29) f denotes the discrete-time frequency, which is assumed to be in the range $-1/2 \le f \le 1/2$ for reasons that will be elucidated later. The $1/N$ factor is required to normalize $\hat{P}_X(f)$ to be interpretable as a power spectral *density* or power per unit frequency. The use of a "hat" is meant to convey the notion that this quantity is an estimator. As we now show, the periodogram is not a suitable measure of the distribution of power with frequency, although it would be for some deterministic signals (such as periodic discrete-time signals with period N). As an example, we plot $\hat{P}_X(f)$ in Figure 17.9 for the realizations given in Figure 17.5. We

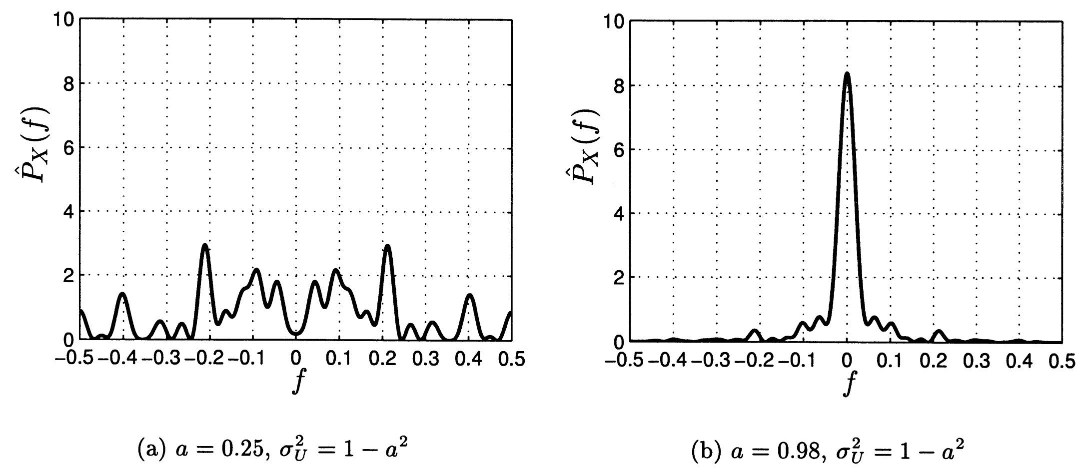


Figure 17.9: Periodogram for autoregressive random process with different parameters. The realizations shown in Figure 17.5 were used to generate these estimates.

see that the periodogram in Figure 17.9a exhibits many random fluctuations. Other realizations will also produce similar seemingly random curves. However, it does seem to produce a reasonable result—for the periodogram in Figure 17.9a there is more high frequency power than for the periodogram in Figure 17.9b. The reason for the random nature of the plot is that (17.29) is a function of N random variables and hence is a random variable itself for each frequency. As such, it exhibits the variability of a random process for which the usual dependence on time is replaced by frequency. What we would actually like is an *average* measure of the power distribution with frequency, suggesting the need for an expected value. Also, to ensure that we capture the entire random process behavior, an infinite length realization is required. We are therefore led to the following more suitable definition of the PSD

$$P_X(f) = \lim_{M\to\infty} \frac{1}{2M+1} E\left[\left|\sum_{n=-M}^{M} X[n]\exp(-j2\pi fn)\right|^2\right]. \tag{17.30}$$

The function $P_X(f)$ is called the *power spectral density* (PSD) and when integrated provides a measure of the average power within a band of frequencies. It is completely analogous to the PDF in that to find the *average power* of the random process in the frequency band $f_1 \leq f \leq f_2$ we should find the area under the PSD curve.

Fourier analysis of a random process yields no phase information.

In our definition of the PSD we are using the magnitude-squared of the Fourier transform. It is obvious then, that the PSD does not tell us anything about the phases of the Fourier transform of the random process. This is in contrast to a Fourier transform of a deterministic signal. There the inverse Fourier transform can be viewed as a decomposition of the signal into sinusoids of different frequencies with deterministic amplitudes and phases. For a random process a similar decomposition called the *spectral representation theorem* [Brockwell and Davis 1987] yields sinusoids of different frequencies with *random amplitudes and random phases.* The PSD is essentially the *expected value of the power of the random sinusoidal amplitudes* per unit of frequency. No phase information is retained and therefore no phase information can be extracted from knowledge of the PSD.

We next give an example of the computation of a PSD.

Example 17.8 – White noise

Assume that $X[n]$ is white noise (see Example 17.2) and therefore, has a zero mean and ACS $r_X[k] = \sigma^2\delta[k]$. Then,

$$\begin{aligned}
P_X(f) &= \lim_{M\to\infty} \frac{1}{2M+1} E\left[\sum_{n=-M}^{M} X[n]\exp(j2\pi fn) \sum_{m=-M}^{M} X[m]\exp(-j2\pi fm)\right] \\
&= \lim_{M\to\infty} \frac{1}{2M+1} \sum_{n=-M}^{M}\sum_{m=-M}^{M} \underbrace{E[X[n]X[m]]}_{r_X[m-n]} \exp[-j2\pi f(m-n)] \qquad (17.31) \\
&= \lim_{M\to\infty} \frac{1}{2M+1} \sum_{n=-M}^{M}\sum_{m=-M}^{M} \sigma^2\delta[m-n]\exp[-j2\pi f(m-n)] \\
&= \lim_{M\to\infty} \frac{1}{2M+1} \sum_{n=-M}^{M} \sigma^2 \\
&= \lim_{M\to\infty} \sigma^2 = \sigma^2. \qquad (17.32)
\end{aligned}$$

Hence, for white noise the PSD is

$$P_X(f) = \sigma^2 \qquad -1/2 \leq f \leq 1/2.$$

As first mentioned in Chapter 16 white noise contains equal contributions of average power at all frequencies.

◇

A more straightforward approach to obtaining the PSD is based on knowledge of the ACS. From (17.31) we see that

$$P_X(f) = \lim_{M\to\infty} \frac{1}{2M+1} \sum_{n=-M}^{M}\sum_{m=-M}^{M} r_X[m-n]\exp[-j2\pi f(m-n)]. \qquad (17.33)$$

This can be simplified using the formula (see Problem 17.26)

$$\sum_{n=-M}^{M}\sum_{m=-M}^{M} g[m-n] = \sum_{k=-2M}^{2M} (2M+1-|k|)g[k]$$

which results from considering $g[m-n]$ as an element of the $(2M+1)\times(2M+1)$ matrix $\mathbf{G}$ with elements $[\mathbf{G}]_{mn} = g[m-n]$ for $m = -M, \ldots, M$ and $n = -M, \ldots, M$ and then summing all the elements. Using this relationship in (17.33) produces

$$\begin{aligned}
P_X(f) &= \lim_{M\to\infty} \frac{1}{2M+1} \sum_{k=-2M}^{2M} (2M+1-|k|) r_X[k]\exp(-j2\pi fk) \\
&= \lim_{M\to\infty} \sum_{k=-2M}^{2M} \left(1 - \frac{|k|}{2M+1}\right) r_X[k]\exp(-j2\pi fk).
\end{aligned}$$

Assuming that $\sum_{k=-\infty}^{\infty} |r_X[k]| < \infty$, the limit can be shown to produce the final result (see Problem 17.27)

$$P_X(f) = \sum_{k=-\infty}^{\infty} r_X[k] \exp(-j2\pi f k) \tag{17.34}$$

which says that *the power spectral density is the discrete-time Fourier transform of the ACS.* This relationship is known as the *Wiener-Khinchine* theorem. Some examples follow.

Example 17.9 – White noise

From Example 17.2 $r_X[k] = \sigma^2 \delta[k]$ and so

$$\begin{aligned} P_X(f) &= \sum_{k=-\infty}^{\infty} r_X[k] \exp(-j2\pi f k) \\ &= \sum_{k=-\infty}^{\infty} \sigma^2 \delta[k] \exp(-j2\pi f k) \\ &= \sigma^2. \end{aligned}$$

This is shown in Figure 17.10. Note that the total average power in $X[n]$, which is $r_X[0] = \sigma^2$, is given by the area under the PSD curve.

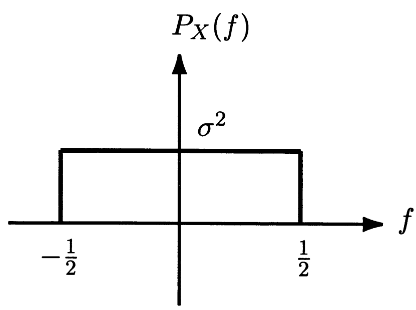


Figure 17.10: PSD of white noise.

◇

Example 17.10 – AR random process

From (17.21) we have that

$$r_X[k] = \frac{\sigma_U^2}{1-a^2} a^{|k|} \qquad -\infty < k < \infty$$

and from (17.34)

$$\begin{aligned} P_X(f) &= \sum_{k=-\infty}^{\infty} r_X[k]\exp(-j2\pi fk) \\ &= \frac{\sigma_U^2}{1-a^2}\sum_{k=-\infty}^{\infty} a^{|k|}\exp(-j2\pi fk) \\ &= \frac{\sigma_U^2}{1-a^2}\left[\sum_{k=-\infty}^{-1} a^{-k}\exp(-j2\pi fk) + \sum_{k=0}^{\infty} a^k \exp(-j2\pi fk)\right] \\ &= \frac{\sigma_U^2}{1-a^2}\left[\sum_{k=1}^{\infty}[a\exp(j2\pi f)]^k + \sum_{k=0}^{\infty}[a\exp(-j2\pi f)]^k\right]. \end{aligned}$$

Since $|a\exp(\pm j2\pi f)| = |a| < 1$, we can use the formula $\sum_{k=k_0}^{\infty} z^k = z^{k_0}/(1-z)$ for z a complex number with $|z| < 1$ to evaluate the sums. This produces

$$\begin{aligned} P_X(f) &= \frac{\sigma_U^2}{1-a^2}\left(\frac{a\exp(j2\pi f)}{1-a\exp(j2\pi f)} + \frac{1}{1-a\exp(-j2\pi f)}\right) \\ &= \frac{\sigma_U^2}{1-a^2}\frac{a\exp(j2\pi f)(1-a\exp(-j2\pi f)) + (1-a\exp(j2\pi f))}{(1-a\exp(j2\pi f))(1-a\exp(-j2\pi f))} \\ &= \frac{\sigma_U^2}{1-a^2}\frac{1-a^2}{|1-a\exp(-j2\pi f)|^2} \\ &= \frac{\sigma_U^2}{|1-a\exp(-j2\pi f)|^2}. \end{aligned} \tag{17.35}$$

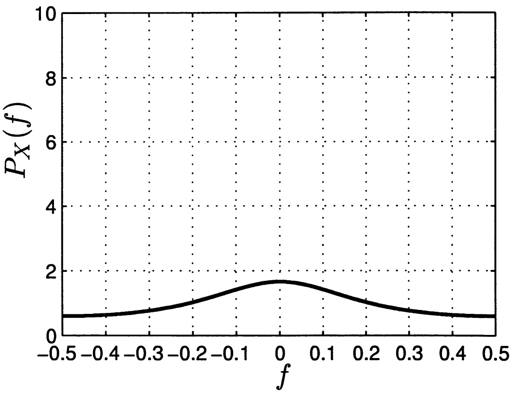
This can also be written in real form as

$$P_X(f) = \frac{\sigma_U^2}{1+a^2-2a\cos(2\pi f)} \qquad -1/2 \le f \le 1/2. \tag{17.36}$$

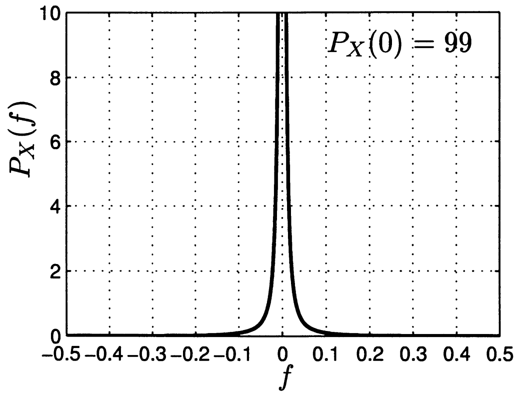
For $a = 0.25$ and $a = 0.98$ and $\sigma_U^2 = 1 - a^2$, the PSDs are plotted in Figure 17.11. Note that the total average power in each PSD is the same, being $r_X[0] = \sigma_U^2/(1-a^2) = 1$. As expected the more noise-like random process has a PSD (see Figure 17.11a) with more high frequency average power than the slowly varying random process (see Figure 17.11b) which has all its average power near $f = 0$ (or at DC).

◇

From the previous example, we observe that the PSD exhibits the properties of being a real nonnegative function of frequency, consistent with our notion of power as a nonnegative physical quantity, of being symmetric about $f = 0$, and of being periodic with period one (see (17.36)). We next prove that these properties are true in general.



(a) $a = 0.25$, $\sigma_U^2 = 1 - a^2$



(b) $a = 0.98$, $\sigma_U^2 = 1 - a^2$

Figure 17.11: Power spectral densities for autoregressive random process with different parameters. The periodograms, which are estimated PSDs, were given in Figure 17.9.

Property 17.7 – PSD is a real function.
The PSD is also given by the real function

$$P_X(f) = \sum_{k=-\infty}^{\infty} r_X[k]\cos(2\pi f k). \tag{17.37}$$

Proof:

$$\begin{aligned}
P_X(f) &= \sum_{k=-\infty}^{\infty} r_X[k]\exp(-j2\pi f k) \\
&= \sum_{k=-\infty}^{\infty} r_X[k](\cos(2\pi f k) - j\sin(2\pi f k)) \\
&= \sum_{k=-\infty}^{\infty} r_X[k]\cos(2\pi f k) - j\sum_{k=-\infty}^{\infty} r_X[k]\sin(2\pi f k).
\end{aligned}$$

But

$$\sum_{k=-\infty}^{\infty} r_X[k]\sin(2\pi f k) = \sum_{k=-\infty}^{-1} r_X[k]\sin(2\pi f k) + \sum_{k=1}^{\infty} r_X[k]\sin(2\pi f k)$$

since the $k = 0$ term is zero, and letting $l = -k$ in the first sum we have

$$\begin{aligned}\sum_{k=-\infty}^{\infty} r_X[k]\sin(2\pi f k) &= \sum_{l=1}^{\infty} r_X[-l]\sin(2\pi f(-l)) + \sum_{k=1}^{\infty} r_X[k]\sin(2\pi f k) \\ &= \sum_{k=1}^{\infty} r_X[k](-\sin(2\pi f k) + \sin(2\pi f k)) = 0 \quad (r_X[-l] = r_X[l])\end{aligned}$$

from which (17.37) follows.

□

Property 17.8 – PSD is nonnegative.

$$P_X(f) \geq 0$$

Proof: Follows from (17.30) but can also be shown to follow from the positive semidefinite property of the ACS [Brockwell and Davis 1987]. (See also Problem 17.19.)

□

Property 17.9 – PSD is symmetric about $f = 0$.

$$P_X(-f) = P_X(f)$$

Proof: Follows from (17.37).

□

Property 17.10 – PSD is periodic with period one.

$$P_X(f+1) = P_X(f)$$

Proof: From (17.37) we have

$$\begin{aligned}P_X(f+1) &= \sum_{k=-\infty}^{\infty} r_X[k]\cos(2\pi(f+1)k) \\ &= \sum_{k=-\infty}^{\infty} r_X[k]\cos(2\pi f k + 2\pi k) \\ &= \sum_{k=-\infty}^{\infty} r_X[k]\cos(2\pi f k) \qquad (\cos(2\pi k) = 1,\ \sin(2\pi k) = 0) \\ &= P_X(f)\end{aligned}$$

□

Property 17.11 – ACS recovered from PSD using inverse Fourier transform

$$r_X[k] = \int_{-\frac{1}{2}}^{\frac{1}{2}} P_X(f)\exp(j2\pi fk)df \qquad -\infty < k < \infty \tag{17.38}$$

$$= \int_{-\frac{1}{2}}^{\frac{1}{2}} P_X(f)\cos(2\pi fk)df \qquad -\infty < k < \infty \tag{17.39}$$

Proof: (17.38) follows from properties of discrete-time Fourier transform [Jackson 1991]. (17.39) follows from Property 17.9 (see Appendix B.5 and also Problem 17.49).

□

Property 17.12 – PSD yields average power over band of frequencies.
To obtain the average power in the frequency band $f_1 \leq f \leq f_2$ we need only find the area under the PSD curve for this band. The average *physical* power is obtained as twice this area since the negative frequencies account for half of the average power (recall Property 17.9). Hence,

$$\text{Average physical power in } [f_1, f_2] = 2\int_{f_1}^{f_2} P_X(f)df. \tag{17.40}$$

The proof of this property requires some concepts to be described in the next chapter, and thus, we defer the proof until Section 18.4. Note, however, that if $f_1 = 0$ and $f_2 = 1/2$, then the average power in this band is

$$\begin{aligned}
\text{Average physical power in } [0, 1/2] &= 2\int_0^{1/2} P_X(f)df \\
&= \int_{-\frac{1}{2}}^{\frac{1}{2}} P_X(f)df \quad \text{(due to symmetry of PSD)} \\
&= \int_{-\frac{1}{2}}^{\frac{1}{2}} P_X(f)\exp(j2\pi f(0))df \\
&= r_X[0] \qquad \text{(from (17.38))}
\end{aligned}$$

which we have already seen yields the total average power since $r_X[0] = E[X^2[n]]$. Hence, we see that the *total average power* is obtained by integrating the PSD over all frequencies to yield

$$r_X[0] = \int_{-\frac{1}{2}}^{\frac{1}{2}} P_X(f)df. \tag{17.41}$$

□

⚠ **Definitions of PSD are not consistent.**

In some texts, especially ones describing the use of the PSD for physical measurements, the definition of the PSD is slightly different. The alternative definition relies on the relationship of (17.40) to define the PSD as $G_X(f) = 2P_X(f)$. It is called the *one-sided PSD* and its advantage is that it yields directly the average power over a band when integrated over the band. As can be seen from (17.40)

$$\text{Average physical power in } [f_1, f_2] = \int_{f_1}^{f_2} G_X(f)df.$$

⚠

A final comment concerns the periodicity of the PSD. We have chosen the frequency interval $[-1/2, 1/2]$ over which to display the PSD. The rationale for this choice arises from the practical situation in which a *continuous-time* WSS random process (see Section 17.8) is sampled to produce a discrete-time WSS random process. Then, if the continuous-time random process $X(t)$ has a PSD that is bandlimited to W Hz and is sampled at F_s samples/sec, the discrete-time PSD $P_X(f)$ will have discrete-time frequency units of W/F_s. For Nyquist rate sampling of $F_s = 2W$, the maximum discrete-time frequency will be $f = W/F_s = 1/2$. Hence, our choice of the frequency interval $[-1/2, 1/2]$ corresponds to the continuous-time frequency interval of $[-W, W]$ Hz. The discrete-time frequency is also referred to as the *normalized frequency*, the normalizing factor being F_s.

17.7 Estimation of the ACS and PSD

Recall from our discussion of ergodicity that in the problem of mean estimation for a WSS random process, we were restricted to observing only a finite number of samples of *one realization* of the random process. If the random process is ergodic in the mean, then we saw that as the number of samples increases to infinity, the temporal average $\hat{\mu}_N$ will converge to the ensemble average μ. To apply this result to estimation of the ACS consider the problem of estimating the ACS for lag $k = k_0$ which is

$$r_X[k_0] = E[X[n]X[n + k_0]].$$

Then by defining the product random process $Y[n] = X[n]X[n + k_0]$ we see that

$$r_X[k_0] = E[Y[n]] \qquad -\infty < n < \infty$$

or the desired quantity to be estimated is just the mean of the random process $Y[n]$. The mean of $Y[n]$ does not depend on n. This suggests that we replace the observed values of $X[n]$ with those of $Y[n]$ by using $y[n] = x[n]x[n+k_0]$, and then use a temporal average to estimate the ensemble average. Hence, we have the temporal average estimate

$$\begin{aligned} \hat{r}_X[k_0] &= \frac{1}{N}\sum_{n=0}^{N-1} y[n] \\ &= \frac{1}{N}\sum_{n=0}^{N-1} x[n]x[n+k_0]. \end{aligned} \tag{17.42}$$

Also, since $r_X[-k] = r_X[k]$, we need only estimate the ACS for $k \geq 0$. There is one slight modification that we need to make to the estimate. Assuming that $\{x[0], x[1], \ldots, x[N-1]\}$ are observed, we must choose the upper limit on the summation in (17.42) to satisfy the constraint $n + k_0 \leq N-1$. This is because $x[n+k_0]$ is unobserved for $n + k_0 > N-1$. With this modification we have as our estimate of the ACS (and now replacing the specific lag of k_0 by the more general lag k)

$$\hat{r}_X[k] = \frac{1}{N-k}\sum_{n=0}^{N-1-k} x[n]x[n+k] \qquad k = 0, 1, \ldots, N-1. \tag{17.43}$$

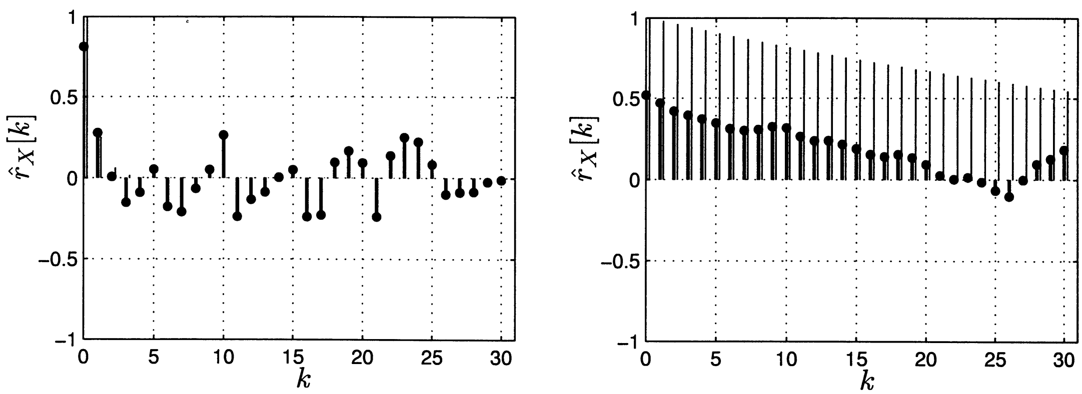
We have also changed the $1/N$ averaging factor to $1/(N-k)$. This is because the number of terms in the sum is only $N-k$. For example, if $N = 4$ so that we observe $\{x[0], x[1], x[2], x[3]\}$, then (17.43) yields the estimates

$$\begin{aligned} \hat{r}_X[0] &= \frac{1}{4}(x^2[0] + x^2[1] + x^2[2] + x^2[3]) \\ \hat{r}_X[1] &= \frac{1}{3}(x[0]x[1] + x[1]x[2] + x[2]x[3]) \\ \hat{r}_X[2] &= \frac{1}{2}(x[0]x[2] + x[1]x[3]) \\ \hat{r}_X[3] &= x[0]x[3]. \end{aligned}$$

As k increases, the distance between the samples increases and so there are less products available for averaging. In fact, for $k > N-1$, we cannot estimate the value of the ACS at all. With the estimate given in (17.43) we see that $E[\hat{r}_X[k]] = r_X[k]$ for $k = 0, 1, \ldots, N-1$. In order for the estimate to converge to the true value as $N \to \infty$, i.e, for the random process to be *ergodic in the autocorrelation* or

$$\lim_{N\to\infty} \hat{r}_X[k] = \lim_{N\to\infty} \frac{1}{N-k}\sum_{n=0}^{N-1-k} x[n]x[n+k] = r_X[k] \qquad k = 0, 1, \ldots$$

we require that $\text{var}(\hat{r}_X[k]) \to 0$ as $N \to \infty$. This will generally be true if $r_X[k] \to 0$ as $k \to \infty$ for a zero mean random process but see Problem 17.25 for a case where this is not required. To illustrate the estimation performance consider the AR random process described in Example 17.5. The true ACS and the estimated one using (17.43) and based on the realizations shown in Figure 17.5 are shown in Figure 17.12. The estimated ACS is shown as the dark lines while the true ACS as given by (17.21) is shown as light lines, which are slightly displaced to the right for easier viewing. Note that in Figure 17.12 the estimated values for k large exhibit



(a) $a = 0.25$, $\sigma_U^2 = 1 - a^2$

(b) $a = 0.98$, $\sigma_U^2 = 1 - a^2$

Figure 17.12: Estimated ACSs (dark lines) and the true ACSs given in Figure 17.6 (light lines) for the AR random process realizations shown in Figure 17.5.

a large error. This is due to the fewer number of products, i.e., $N - k = 31 - k$, that are available for averaging in (17.43). In the case of $k = 30$ the estimate is $\hat{r}_X[30] = x[0]x[30]$, which as you might expect is very poor since there is no averaging at all! Clearly, for accurate estimates of the ACS we require that $k_{\max} \ll N$. The MATLAB code used to estimate the ACS for Figure 17.12 is given below.

```
n=[0:30]';N=length(n);
a1=0.25;a2=0.98;
varu1=1-a1^2;varu2=1-a2^2;
r1true=(varu1/(1-a1^2))*a1.^n; % see (17.21)
r2true=(varu2/(1-a2^2))*a2.^n;
for k=0:N-1
   r1est(k+1,1)=(1/(N-k))*sum(x1(1:N-k).*x1(1+k:N));
   r2est(k+1,1)=(1/(N-k))*sum(x2(1:N-k).*x2(1+k:N));
end
```

To estimate the PSD requires somewhat more care than the ACS. We have already seen that the periodogram estimate of (17.29) is not suitable. There are many ways to estimate the PSD based on either (17.30) or (17.34). We illustrate one approach based on (17.30). Others may be found in [Jenkins and Watts 1968, Kay 1988]. Since we only have a segment of a single realization of the random process, we cannot implement the expectation operation required in (17.30). Note that the operation of $E[\cdot]$ represents an average down the ensemble or equivalently an average over multiple realizations. To obtain some averaging, however, we can break up the data $\{x[0], x[1], \ldots, x[N-1]\}$ into I nonoverlapping blocks, with each block having a total of L samples. We assume for simplicity that there is an integer number of blocks so that $N = IL$. The implicit assumption in doing so is that each block exhibits the statistical characteristics of a single realization and so we can mimic the averaging down the ensemble by averaging temporally across successive blocks of data. Once again, the assumption of ergodicity is being employed. Thus, we first break up the data set into the I nonoverlapping data blocks

$$y_i[n] = x[n + iL] \qquad n = 0, 1, \ldots, L-1; i = 0, 1, \ldots, I-1$$

where each data block has a length of L samples. Then, for each data block we compute a periodogram as

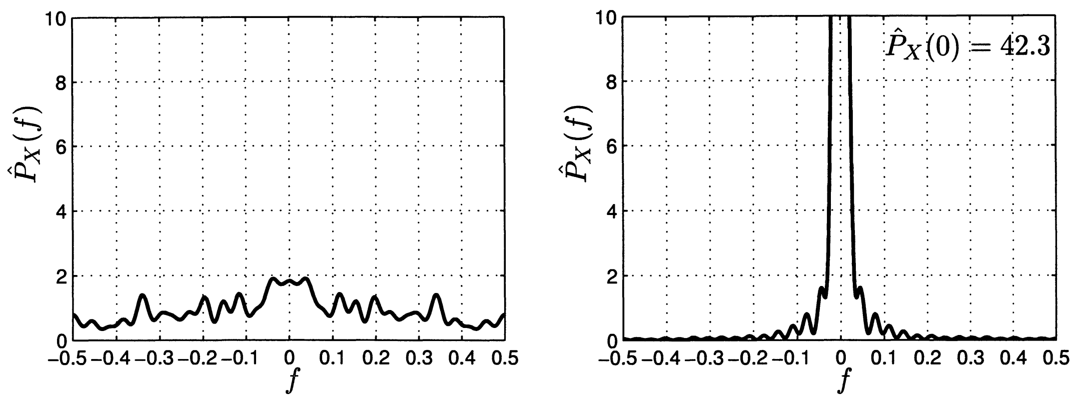
$$\hat{P}_X^{(i)}(f) = \frac{1}{L}\left|\sum_{n=0}^{L-1} y_i[n] \exp(-j2\pi f n)\right|^2 \tag{17.44}$$

and then average all the periodograms together to yield the final PSD estimate as

$$\hat{P}_{\text{av}}(f) = \frac{1}{I}\sum_{i=0}^{I-1} \hat{P}_X^{(i)}(f). \tag{17.45}$$

This estimate is called the *averaged periodogram*. It can be shown that under some conditions, $\lim_{N\to\infty} \hat{P}_{\text{av}}(f) = P_X(f)$. Once again we are calling upon an ergodicity type of property in that we are averaging the periodograms obtained in time instead of the theoretical ensemble averaging. Of course, for convergence to hold as $N \to \infty$, we must have $L \to \infty$ and $I \to \infty$ as well.

As an example, we examine the averaged periodogram estimates for the two AR processes whose PSDs are shown in Figure 17.11. The number of data samples was $N = 310$, which was broken up into $I = 10$ nonoverlapping blocks of data with $L = 31$ samples in each one. By comparing the spectral estimates in Figure 17.13 with those of Figure 17.9, it is seen that the averaging has yielded a better estimate. Of course, the price paid is that the data set needs to be $I = 10$ times as long! The MATLAB code used to implement the averaged periodogram estimate is given next. A fast Fourier transform (FFT) is used to compute the Fourier transform of the $y_i[n]$ sequences at the frequencies $f = -0.5 + k\Delta_f$, where $k = 0, 1, \ldots, 1023$ and $\Delta_f = 1/1024$ (see [Kay 1988] for a more detailed description).



(a) $a = 0.25$, $\sigma_U^2 = 1 - a^2$ (b) $a = 0.98$, $\sigma_U^2 = 1 - a^2$

Figure 17.13: Power spectral density estimates using the averaged periodogram method for autoregressive processes with different parameters. The true PSDs are shown in Figure 17.11.

```
Nfft=1024; % set FFT size
Pav1=zeros(Nfft,1);Pav2=Pav1; % set up arrays with desired dimension
f=[0:Nfft-1]'/Nfft-0.5; % set frequencies for later plotting
                        % of PSD estimate
for i=0:I-1
   nstart=1+i*L;nend=L+i*L; % set up beginning and end points
                            % of ith block of data
   y1=x1(nstart:nend);
   y2=x2(nstart:nend);
% take FFT of block, since FFT outputs samples of Fourier
% transform over frequency range [0,1), must shift FFT outputs
% for [1/2,1) to [-1/2, 0), then take complex magnitude-squared,
% normalize by L and average
Pav1=Pav1+(1/(I*L))*abs(fftshift(fft(y1,Nfft))).^2;
Pav2=Pav2+(1/(I*L))*abs(fftshift(fft(y2,Nfft))).^2;
end
```

17.8 Continuous-Time WSS Random Processes

In this section we give the corresponding definitions and formulas for continuous-time WSS random processes. A more detailed description can be found in [Papoulis 1965]. Also, an important example is described to illustrate the use of these formulas.

A continuous-time random process $X(t)$ for $-\infty < t < \infty$ is defined to be WSS if the *mean function* $\mu_X(t)$ satisfies

$$\mu_X(t) = E[X(t)] = \mu \qquad -\infty < t < \infty \tag{17.46}$$

which is to say it is constant in time and an *autocorrelation function* (ACF) can be defined as

$$r_X(\tau) = E[X(t)X(t+\tau)] \qquad -\infty < \tau < \infty \tag{17.47}$$

which is not dependent on the value of t. Thus, $E[X(t_1)X(t_2)]$ depends only on $|t_2 - t_1|$. Note the use of the "parentheses" indicates that the argument of the ACF is continuous and serves to distinguish $r_X[k]$ from $r_X(\tau)$. The ACF has the following properties.

Property 17.13 – ACF is positive for the zero lag or $r_X(0) > 0$.
The total average power is $r_X(0) = E[X^2(t)]$.

□

Property 17.14 – ACF is an even function or $r_X(-\tau) = r_X(\tau)$.

□

Property 17.15 – Maximum value of ACF is at $\tau = 0$ or $|r_X(\tau)| \leq r_X(0)$.

□

Property 17.16 – ACF measures the predictability of a random process.
The correlation coefficient for two samples of a zero mean WSS random process is

$$\rho_{X(t),X(t+\tau)} = \frac{r_X(\tau)}{r_X(0)}.$$

□

Property 17.17 – ACF approaches μ^2 as $\tau \to \infty$.
This assumes that the samples become uncorrelated for large lags, which is usually the case.

□

Property 17.18 – $r_X(\tau)$ is a positive semidefinite function.
See [Papoulis 1965] for the definition of a positive semidefinite function. This property assumes that the some samples of $X(t)$ may be perfectly predictable. If it is not, then the ACF is positive definite.

□

The PSD is defined as

$$P_X(F) = \lim_{T\to\infty} \frac{1}{T} E\left[\left|\int_{-T/2}^{T/2} X(t)\exp(-j2\pi Ft)dt\right|^2\right] \quad -\infty < F < \infty \tag{17.48}$$

where F is the frequency in Hz. We use a capital F to denote continuous-time or analog frequency. By the Wiener-Khinchine theorem this is equivalent to the continuous-time Fourier transform of the ACF

$$\begin{aligned} P_X(F) &= \int_{-\infty}^{\infty} r_X(\tau)\exp(-j2\pi F\tau)d\tau & (17.49) \\ &= \int_{-\infty}^{\infty} r_X(\tau)\cos(2\pi F\tau)d\tau. & (17.50) \end{aligned}$$

(See also Problem 17.49.) The PSD has the usual interpretation as the average power distribution with frequency. In particular, it is the average power per Hz. The average physical power in a frequency band $[F_1, F_2]$ is given by

$$\text{Average physical power in } [F_1, F_2] = 2\int_{F_1}^{F_2} P_X(F)dF$$

where again the 2 factor reflects the additional contribution of the negative frequencies. The properties of the PSD are as follows:

Property 17.19 – PSD is a real function.

The PSD is given by the real function

$$P_X(F) = \int_{-\infty}^{\infty} r_X(\tau)\cos(2\pi F\tau)d\tau$$

□

Property 17.20 – PSD is nonnegative.

$$P_X(F) \geq 0$$

□

Property 17.21 – PSD is symmetric about $F = 0$.

$$P_X(-F) = P_X(F)$$

□

Property 17.22 – ACF recovered from PSD using inverse Fourier transform

$$r_X(\tau) = \int_{-\infty}^{\infty} P_X(F)\exp(j2\pi F\tau)dF \qquad -\infty < \tau < \infty \tag{17.51}$$

$$= \int_{-\infty}^{\infty} P_X(F)\cos(2\pi F\tau)dF \qquad -\infty < \tau < \infty. \tag{17.52}$$

(See also Problem 17.49.)

□

Unlike the PSD for a discrete-time WSS random process, the PSD for a continuous-time WSS random process is *not* periodic. We next illustrate these definitions and formulas with an example of practical importance.

Example 17.11 – Obtaining discrete-time WGN from continuous-time WGN

A common model for a continuous-time noise random process $X(t)$ in a physical system is a WSS random process with a zero mean. In addition, due to the origin of noise as microscopic fluctuations of a large number of electrons, or molecules, etc., a central limit theorem can be employed to assert that $X(t)$ is a Gaussian random variable for all t. The average power of the noise in a band of frequencies is observed to be the same for all bands up to some upper frequency limit, at which the average power begins to decrease. For instance, consider thermal noise in a conductor due to random fluctuations of the electrons about some mean velocity. The average power versus frequency is predicted by physics to be constant until a cutoff frequency of about $F_c = 1000$ GHz at room temperature [Bell Telephone Labs 1970]. Hence, we can assume that the PSD of the noise has a PSD shown in Figure 17.14 as the true PSD. To further simplify the mathematically modeling without sacrificing the realism of the model, we can observe that all physical systems will only pass frequency components that are much lower than F_c—typically the bandwidth of the system is W Hz as shown in Figure 17.14. Any frequencies above W Hz will be cut off by the system. Therefore, the noise output of the system will be the same whether we use the true PSD or the modeled one shown in Figure 17.10. The modeled PSD is given by

$$P_X(F) = \frac{N_0}{2} \qquad -\infty < F < \infty.$$

This is clearly a physically impossible PSD in that the total average power is $r_X(0) = \int_{-\infty}^{\infty} P_X(F)df = \infty$. However, its use simplifies much systems analysis (see Problem 17.50). The corresponding ACF is from (17.51) the inverse Fourier transform, which is

$$r_X(\tau) = \frac{N_0}{2}\delta(\tau) \tag{17.53}$$

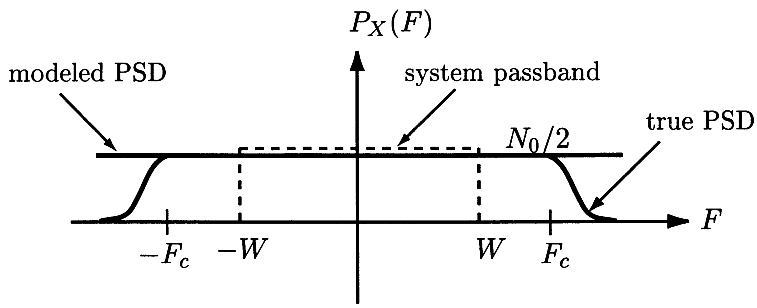


Figure 17.14: True and modeled PSDs for continuous-time white Gaussian noise.

and is seen to be an impulse at $\tau = 0$. Again the nonphysical nature of this model is manifest by the value $r_X(0) = \infty$. A continuous-time WSS Gaussian random process with zero mean and the ACF given by (17.53) is called continuous-time *white Gaussian noise* (WGN) (see also Example 20.6). It is a standard model in many disciplines.

Now as was previously mentioned, all physical systems are bandlimited to W Hz, which is typically chosen to ensure that a desired signal with a bandwidth of W Hz is not distorted. Modern signal processing hardware first bandlimits the continuous-time waveform to a maximum of W Hz using a lowpass filter and then samples the output of the filter at the Nyquist rate of $F_s = 2W$ samples/sec. The samples are then input into a digital computer. An important question to answer is: What are the statistical characteristics of the noise samples that are input to the computer? To answer this question we let Δ_t be the time interval between successive samples. Also, let $X(t)$ be the noise at the output of an ideal lowpass filter ($H(F) = 1$ for $|F| \leq W$ and $H(F) = 0$ for $|F| > W$) over the system passband shown in Figure 17.14. Then, the noise samples can be represented as

$$X(t)|_{t=n\Delta_t} = X[n] \qquad \text{for } -\infty < n < \infty.$$

Since $X(t)$ is bandlimited to W Hz and prior to filtering had the modeled PSD shown in Figure 17.14, its PSD is

$$P_X(F) = \begin{cases} \frac{N_0}{2} & |F| \leq W \\ 0 & |F| > W. \end{cases}$$

The noise samples $X[n]$ comprise a discrete-time random process. Its characteristics follow those of $X(t)$. Since $X(t)$ is Gaussian, then so is $X[n]$ (being just a sample). Also, since $X(t)$ is zero mean, so is $X[n]$ for all n. Finally, we inquire as to whether $X[n]$ is WSS, i.e., can we define an ACS? To answer this we first note that $X[n] = X(n\Delta_t)$ and recall that $X(t)$ is WSS. Then from the definition of the ACS

$$\begin{aligned} E[X[n]X[n+k]] &= E[X(n\Delta_t)X((n+k)\Delta_t)] \\ &= r_X(k\Delta_t) \quad \text{(definition of continuous-time ACF)} \end{aligned}$$

which does not depend on n, and so $X[n]$ is a zero mean discrete-time WSS random process with ACS

$$r_X[k] = r_X(k\Delta_t). \tag{17.54}$$

It is seen to be a sampled version of the continuous-time ACF. To explicitly evaluate the ACS we have from (17.51)

$$\begin{aligned} r_X(\tau) &= \int_{-\infty}^{\infty} P_X(F)\exp(j2\pi F\tau)dF \\ &= \int_{-W}^{W} \frac{N_0}{2}\exp(j2\pi F\tau)dF \\ &= \frac{N_0}{2}\int_{-W}^{W}\cos(2\pi F\tau)dF \qquad \text{(sine component is odd function)} \\ &= \frac{N_0}{2}\frac{\sin(2\pi F\tau)}{2\pi\tau}\Bigg|_{-W}^{W} \\ &= N_0 W\frac{\sin(2\pi W\tau)}{2\pi W\tau} \end{aligned} \tag{17.55}$$

which is shown in Figure 17.15. Now since $r_X[k] = r_X(k\Delta_t) = r_X(k/(2W))$, we

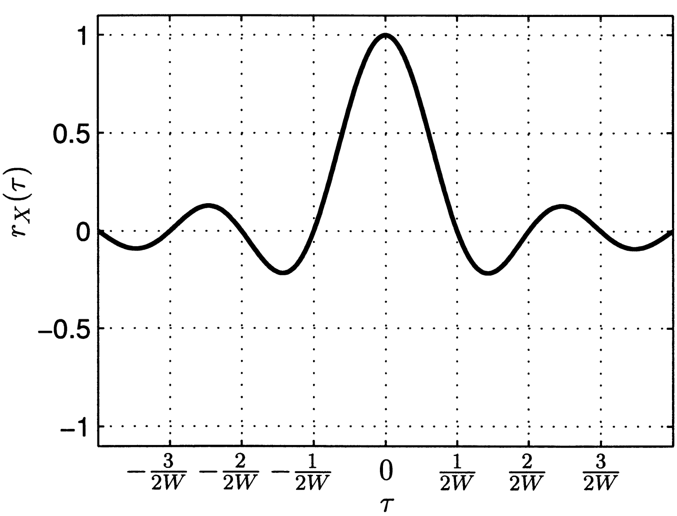


Figure 17.15: ACF for bandlimited continuous-time WGN with $N_0W = 1$.

see from Figure 17.15 that for $k = \pm1, \pm2, \ldots$ the ACS is zero, being the result of sampling the continuous-time ACF at its zeros. The only nonzero value is for $k = 0$, which is $r_X[0] = r_X(0) = N_0W$ from (17.55). Therefore, we finally observe that the ACS of the noise samples is

$$r_X[k] = N_0W\delta[k]. \tag{17.56}$$

The discrete-time noise random process is indeed WSS and has the ACS of (17.56). The PSD corresponding to this ACS has already been found and is shown in Figure 17.10, where $\sigma^2 = N_0 W$. Therefore, $X[n]$ *is a discrete-time white Gaussian noise random process.* This example justifies the use of the WGN model for discrete-time systems analysis.

◇

Sampling faster gives only marginally better performance.

It is sometimes argued that by sampling the output of a system lowpass filter whose cutoff frequency is W Hz at a rate greater than $2W$, we can improve the performance of a signal processing system. For example, consider the estimation of the mean μ based on samples $Y[n] = \mu + X[n]$ for $n = 0, 1, \ldots, N-1$ where $E[X[n]] = 0$, $\mathrm{var}(X[n]) = \sigma^2$, and the $X[n]$ samples are uncorrelated. The obvious estimate is the sample mean or $(1/N)\sum_{n=0}^{N-1} Y[n]$, whose expectation is μ and whose variance is σ^2/N. Clearly, if we could increase N, then the variance could be reduced and a better estimate would result. This suggests sampling the continuous-time random process at a rate faster than $2W$ samples/sec. The fallacy, however, is that as the sampling rate increases, *the noise samples become correlated* as can be seen by considering a sampling rate of $4W$ for which the time interval between samples becomes $\tau = \Delta_t/2 = 1/(4W)$. Then, as observed from Figure 17.15, the correlation between successive samples is $r_X(1/(4W)) = 0.6$. In effect, by sampling faster we are not obtaining any new realizations of the noise samples but nearly repetitions of the same noise samples. As a result, the variance will *not* decrease as $1/N$ but at a slower rate (see also Problem 17.51).

17.9 Real-World Example – Random Vibration Testing

Anyone who has ever traveled in a jet knows that upon landing, the cabin can vibrate greatly. This is due to the air currents outside the cabin which interact with the metallic aircraft surface. These pressure variations give rise to vibrations which are referred to as *turbulent boundary layer noise.* A manufacturer that intends to attach an antenna or other device to an aircraft must be cognizant of this vibration and plan for it. It is customary then to subject the antenna to a random vibration test in the lab to make sure it is not adversely affected in flight [McConnell 1995]. To do so the antenna would be mounted on a shaker table and the table shaken in a manner to simulate the turbulent boundary layer (TBL) noise. The problem the manufacturer faces is how to provide the proper vibration signal to the table, which presumably will then be transmitted to the antenna. We now outline a possible solution to this problem.

The National Aeronautics and Space Administration (NASA) has determined PSD models for the TBL noise through physical modeling and experimentation. A reasonable model for the *one-sided* PSD of TBL noise upon reentry of a space vehicle, such as the space shuttle, into the earth's atmosphere is given by [NASA 2001]

$$G_X(F) = \begin{cases} G_X(500) & 0 \le F < 500 \text{ Hz} \\ \frac{9\times 10^{14} r^2}{F+11364} & 500 \le F \le 50000 \text{ Hz} \end{cases}$$

where r represents a reference value which is 20μPa. A μPa is a unit of pressure equal to 10^{-6} nt/m^2. This PSD is shown in Figure 17.16 referenced to the standard unit so that $r = 1$. Note that it has a lowpass type of characteristic. In order

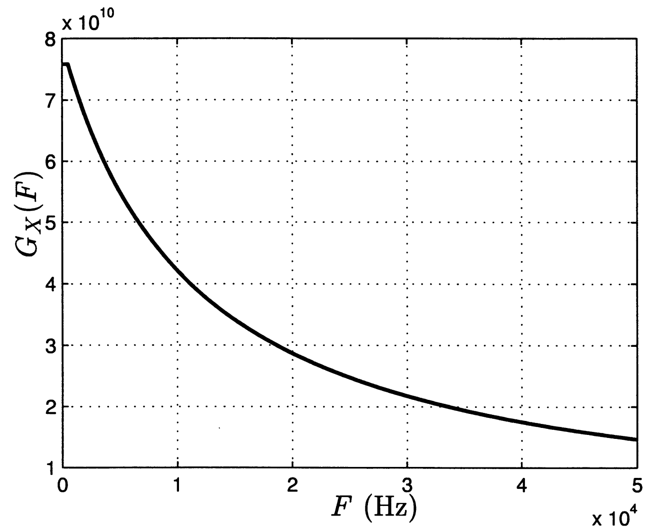


Figure 17.16: Continuous-time one-sided PSD for TBL noise.

to provide a signal to the shaker table that is random and has the PSD shown in Figure 17.16, we will assume that the signal is produced in a digital computer and then converted via a digital-to-analog convertor to a continuous-time signal. Hence, we need to produce a discrete-time WSS random process within the computer that has the proper PSD. Recalling our discussion in Section 17.8 we know that $r_X[k] = r_X(k\Delta_t)$ and since the highest frequency in the PSD is $W = 50,000$ Hz, we choose $\Delta_t = 1/(2W) = 1/100,000$. This produces the discrete-time PSD shown in Figure 17.17 and is given by $P_X(f) = (1/(2\Delta_t))G_X(f/\Delta_t)$. (We have divided by two to obtain the usual *two-sided* PSD. Also, the sampling operation introduces a factor of $1/\Delta_t$ [Jackson 1991].) To generate a realization of a discrete-time WSS random process with PSD given in Figure 17.17 we will use the AR model introduced in

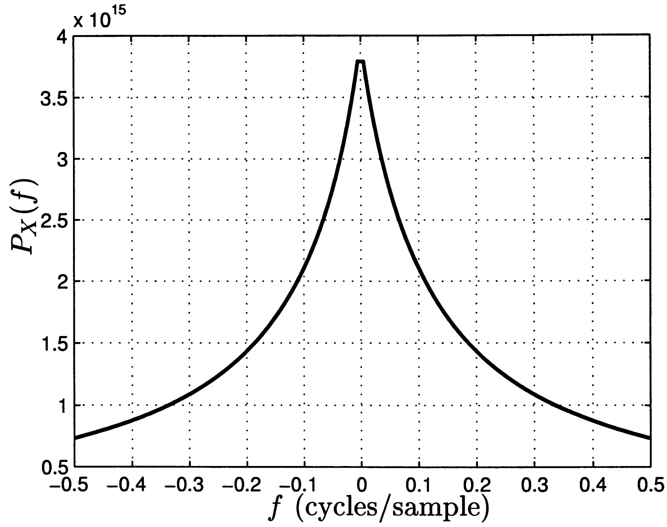


Figure 17.17: Discrete-time PSD for TBL noise.

Example 17.5. From the ACS we can determine values of a and σ_U^2 if we know $r_X[0]$ and $r_X[1]$ since

$$a = \frac{r_X[1]}{r_X[0]} \tag{17.57}$$

$$\sigma_U^2 = r_X[0](1-a^2) = r_X[0]\left[1-\left(\frac{r_X[1]}{r_X[0]}\right)^2\right]. \tag{17.58}$$

Knowing a and σ_U^2 will allow us to use the defining recursive difference equation, $X[n] = aX[n-1] + U[n]$, of an AR random process to generate the realization. To obtain the first two lags of the ACS we use (17.39)

$$r_X[0] = \int_{-\frac{1}{2}}^{\frac{1}{2}} P_X(f)df$$

$$r_X[1] = \int_{-\frac{1}{2}}^{\frac{1}{2}} P_X(f)\cos(2\pi f)df$$

where $P_X(f)$ is given in Figure 17.17. These can be evaluated numerically by replacing the integrals with approximating sums to yield $r_X[0] = 1.5169 \times 10^{15}$ and $r_X[1] = 4.8483 \times 10^{14}$. Then, using (17.57) and (17.58), we have the AR parameters $a = 0.3196$ and $\sigma_U^2 = 1.362 \times 10^{15}$. With these parameters the AR PSD (see (17.36)) and the true PSD (shown in Figure 17.17) are plotted in Figure 17.18. The agreement between them is fairly good except near $f = 0$. Hence, with these values of the parameters a random process realization could be synthesized within a digital computer and then converted to analog form to drive the shaker table.

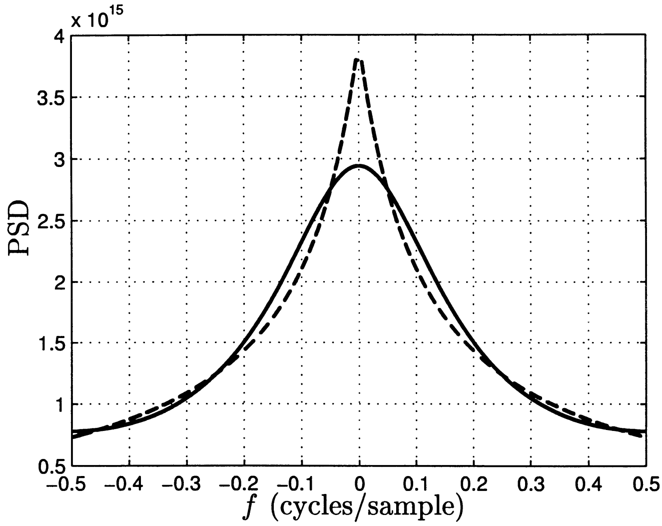


Figure 17.18: Discrete-time PSD and its AR PSD model for TBL noise. The true PSD is shown as the dashed line and the AR PSD model as the solid line.

References

Bell Telephone Laboratories, *Transmission Systems for Communications*, Western Electric Co, Winston-Salem, NC, 1970.

Box, E.P.G., G.M. Jenkins, *Time Series Analysis, Forecasting and Control*, Holden-Day, San Francisco, 1970.

Brockwell, P.J., R.A. Davis, *Time Series: Theory and Methods*, Springer-Verlag, New York, 1987.

Harvey, A.C., *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge University Press, New York, 1989.

Jackson, L.B., *Signals, Systems, and Transforms*, Addison-Wesley, Reading, MA 1991.

Jenkins, G.M., D.G. Watts, *Spectral Analysis and Its Applications*, Holden-Day, San Francisco, 1968.

Kay, S., *Modern Spectral Estimation, Theory and Applications*, Prentice-Hall, Englewood Cliffs, NJ, 1988.

McConnell, K.G., *Vibration Testing*, John Wiley & Sons, New York, 1995.

NASA, *Dynamic Environmental Criteria, NASA Technical Handbook*, pp. 50–51, NASA-HDBK-7005, March 13, 2001.

Papoulis, A., *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, New York, 1965.

Robinson, E.A., *Multichannel Time Series Analysis with Digital Computer Programs*, Holden-Day, San Francisco, 1967.

Schuster, A., "On the Investigation of Hidden Periodicities with Applications to a Supposed 26 Day Period of Meterological Phenomena," *Terr. Magn.*, Vol. 3, pp. 13–41, March 1898.

Schwartz, M., L. Shaw, *Signal Processing, Discrete Spectral Analysis, Detection, and Estimation*, McGraw-Hill, New York, 1975.

Taylor, S., *Modelling Financial Time Series*, John Wiley & Sons, New York, 1986.

Tong, H., *Non-linear Time Series*, Oxford University Press, New York, 1990.

Problems

17.1 (⌣) **(w)** A Bernoulli random process $X[n]$ for $-\infty < n < \infty$ consists of independent random variables with each random variable taking on the values $+1$ and -1 with probabilities p and $1-p$, respectively. Is this random process WSS? If it is WSS, find its mean sequence and autocorrelation sequence.

17.2 (w) Consider the random process defined as $X[n] = a_0U[n] + a_1U[n-1]$ for $-\infty < n < \infty$, where a_0 and a_1 are constants, and $U[n]$ is an IID random process with each $U[n]$ having a mean of zero and a variance of one. Is this random process WSS? If it is WSS, find its mean sequence and autocorrelation sequence.

17.3 (w) A sinusoidal random process is defined as $X[n] = A\cos(2\pi f_0 n)$ for $-\infty < n < \infty$, where $0 < f_0 < 0.5$ is a discrete-time frequency, and $A \sim \mathcal{N}(0,1)$. Is this random process WSS? If it is WSS, find its mean sequence and autocorrelation sequence.

17.4 (f) A WSS random process has $E[X[0]] = 1$ and a covariance sequence $c_X[n_1, n_2] = 2\delta[n_2 - n_1]$. Find the ACS and plot it.

17.5 (⌣) **(w)** A random process $X[n]$ for $-\infty < n < \infty$ consists of independent random variables with

$$X[n] \sim \begin{cases} \mathcal{N}(0,1) & \text{for } n \text{ even} \\ \mathcal{U}(-\sqrt{3}, \sqrt{3}) & \text{for } n \text{ odd.} \end{cases}$$

Is this random process WSS? Is it stationary?

17.6 (w) The random processes $X[n]$ and $Y[n]$ are both WSS. Every sample of $X[n]$ is independent of every sample of $Y[n]$. Is $Z[n] = X[n] + Y[n]$ WSS? If it is WSS, find its mean sequence and autocorrelation sequence.

17.7 (w) The random processes $X[n]$ and $Y[n]$ are both WSS. Every sample of $X[n]$ is independent of every sample of $Y[n]$. Is $Z[n] = X[n]Y[n]$ WSS? If it is WSS, find its mean sequence and autocorrelation sequence.

17.8 (f) For the ACS $r_X[k] = (1/2)^k$ for $k \geq 0$ and $r_X[k] = (1/2)^{-k}$ for $k < 0$, verify that Properties 17.1–17.3 are satisfied.

17.9 (⌣̈) (w) For the sequence $r_X[k] = ab^{|k|}$ for $-\infty < k < \infty$, determine the values of a and b that will result in a valid ACS.

17.10 (w) A periodic WSS random process with period P is defined to be a random process $X[n]$ whose ACS satisfies $r_X[k + P] = r_X[k]$ for all k. An example is the randomly phased sinusoid of Example 17.10 for which $P = 10$. Show that the correlation coefficient for two samples of a *zero mean* periodic random process that are separated by P samples is one. Comment on the predictability of $X[n + P]$ based on $X[n] = x[n]$.

17.11 (w) A WSS random process has an ACS $r_X[k]$ and mean μ. Find the correlation coefficient for two samples of the random process that are separated by k samples.

17.12 (⌣̈) (w) Which of the sequences in Figure 17.19 cannot be valid ACSs? If the sequence cannot be an ACS, explain why not.

17.13 (w) For the randomly phased sinusoid described in Example 17.4 find the optimal linear prediction of $X[1]$ based on observing $X[0] = x[0]$, and also of $X[10]$ based on observing $X[0] = x[0]$. Can either of these samples be perfectly predicted? Explain why or why not.

17.14 (w) For the AR random process described in Example 17.10 find the optimal linear prediction of $X[n_0+k_0]$ based on observing $X[n_0] = x[n_0]$. How accurate is your prediction in terms of MSE as k_0 increases?

17.15 (t) In this problem we derive $r_X[0]$ for the AR random process described in Example 17.5. To do so assume that $X[n]$ can be written as

$$X[n] = \sum_{k=0}^{\infty} a^k U[n-k]. \tag{17.59}$$

This was shown to be true in Example 17.5. Then verify that $r_X[0]$ can be written as

$$r_X[0] = \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} a^k a^l E[U[n-k]U[n-l]]$$

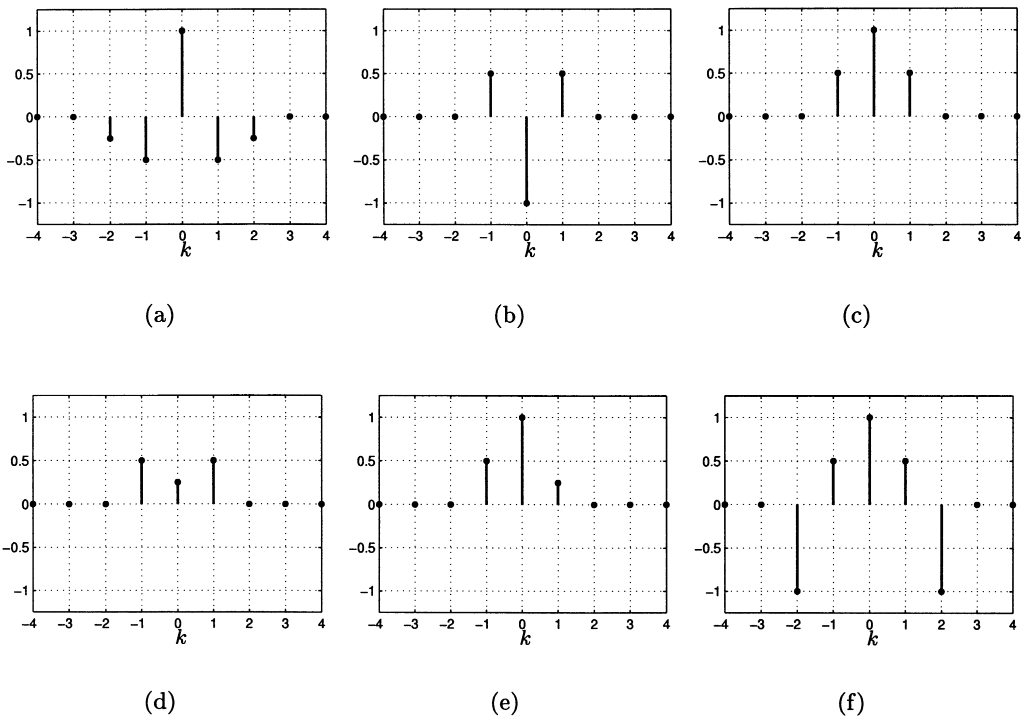


Figure 17.19: Possible ACSs for Problem 17.12.

and use the properties of the $U[n]$ random process to finish the derivation.

17.16 (t) Using a similar approach to the one used in Problem 17.15 derive the ACS for the AR random process described in Example 17.5. Hint: Start with the definition of the ACS and use (17.59).

17.17 (☺) (w) To generate a realization of an AR process on the computer we can use the recursive difference equation $X[n] = aX[n-1] + U[n]$ for $n \geq 0$. However, in doing so, we soon realize that the initial condition $X[-1]$ is required. Assume that we set $X[-1] = 0$ and use the recursion $X[0] = U[0], X[1] = aX[0] + U[1], \ldots$. Determine the mean and variance of $X[n]$ for $n \geq 0$, where as usual $U[n]$ consists of uncorrelated random variables with zero mean and variance σ_U^2. Does the mean depend on n? Does the variance depend on n? What happens as $n \to \infty$? Hint: First show that $X[n]$ can be written as $X[n] = \sum_{k=0}^{n} a^k U[n-k]$ for $n \geq 0$.

17.18 (w) This problem continues Problem 17.17. Instead of letting $X[-1] = 0$, set $X[-1]$ equal to a random variable with mean 0 and a variance of $\sigma_U^2/(1-a^2)$ and that is uncorrelated with $U[n]$ for $n \geq 0$. Find the mean and variance of

$X[0]$. Explain your results and why this makes sense.

17.19 (⌣) (w) An example of a sequence that is not positive semidefinite is $r[0] = 1$, $r[-1] = r[1] = -7/8$ and equals zero otherwise. Compute the determinant of the 1×1 principal minor, the 2×2 principal minor, and the 3×3 principal minor of the 3×3 autocorrelation matrix $\mathbf{R}_X$ using these values. Also, plot the discrete-time Fourier transform of $r[k]$. Why do you think the positive semidefinite property is important?

17.20 (⌣) (w) For the general MA random process of Example 17.6 show that the process is WSS.

17.21 (f) Use (17.28) to show that the MA random process defined in Example 17.6 is ergodic in the mean.

17.22 (t,f) Show that a WSS random process whose ACS satisfies $r_X[k] = \mu^2$ for $k > k_0 \geq 0$ must be ergodic in the mean.

17.23 (t) Prove (17.28) by using the relationship

$$\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} g[i-j] = \sum_{k=-(N-1)}^{N-1} (N-|k|)g[k].$$

Try verifying this relationship for $N = 3$.

17.24 (f) For the random DC level defined in Example 17.7 prove that $\operatorname{var}(\hat{\mu}_N) = 1$.

17.25 (f) Explain why the randomly phased sinusoid defined in Example 17.4 is ergodic in the mean. Next show that it is ergodic in the ACS in that

$$\lim_{N\to\infty} \hat{r}_X[k] = \lim_{N\to\infty} \frac{1}{N-k} \sum_{n=0}^{N-1-k} X[n]X[n+k] = \frac{1}{2}\cos(2\pi(0.1)k) = r_X[k] \quad k \geq 0$$

by computing $\hat{r}_X[k]$ directly. Hint: Use the fact that $\lim_{N\to\infty}(1/(N-k))\sum_{n=0}^{N-1-k}\cos(2\pi f n + \phi) = 0$ for any $0 < f < 1$ and any phase angle ϕ. This is because the temporal average of an infinite duration sinusoid is zero.

17.26 (t) Show that the formula

$$\sum_{m=-M}^{M}\sum_{n=-M}^{M} g[m-n] = \sum_{k=-2M}^{2M} (2M+1-|k|)g[k]$$

is true for $M = 1$.

17.27 (t) Argue that

$$\lim_{M\to\infty}\sum_{k=-2M}^{2M}\underbrace{\left(1-\frac{|k|}{2M+1}\right)}_{w[k]} r_X[k]\exp(-j2\pi fk)=\sum_{k=-\infty}^{\infty} r_X[k]\exp(-j2\pi fk)$$

by drawing pictures of $r_X[k]$, which decays to zero, and overlay it with $w[k]$ as M increases.

17.28 (⌣̈) (w) For the differenced random process defined in Example 17.1 determine the PSD. Explain your results.

17.29 (f) Determine the PSD for the randomly phased sinusoid described in Example 17.4. Is this result reasonable? Hint: The discrete-time Fourier transform of $\exp(j2\pi f_0 n)$ for $-1/2 < f_0 < 1/2$ is $\delta(f - f_0)$ over the frequency interval $-1/2 \le f \le 1/2$.

17.30 (⌣̈) (w) A random process is defined as $X[n] = AU[n]$, where $A \sim \mathcal{N}(0, \sigma_A^2)$ and $U[n]$ is white noise with variance σ_U^2. The random variable A is independent of all the samples of $U[n]$. Determine the PSD of $X[n]$.

17.31 (w) Find the PSD of the random process $X[n] = (1/2)^{|n|}U[n]$ for $-\infty < n < \infty$, where $U[n]$ is white noise with variance σ_U^2.

17.32 (w) Find the PSD of the random process $X[n] = a_0U[n] + a_1U[n-1]$, where a_0, a_1 are constants and $U[n]$ is white noise with variance $\sigma_U^2 = 1$.

17.33 (w) A Bernoulli random process consists of IID Bernoulli random variables taking on values $+1$ and -1 with equal probabilities. Determine the PSD and explain your results.

17.34 (⌣̈) (w) A random process is defined as $X[n] = U[n] + \mu$ for $-\infty < n < \infty$, where $U[n]$ is white noise with variance σ_U^2. Find the ACS and PSD and plot your results.

17.35 (w,c) Consider the AR random process defined in Example 17.5 and described further in Example 17.10 with $-1 < a < 0$ and for some $\sigma_U^2 > 0$. Plot the PSD for several values of a and explain your results.

17.36 (f,c) Plot the corresponding PSD for the ACS

$$r_X[k] = \begin{cases} 1 & k = 0 \\ 1/2 & k = \pm 1 \\ 1/4 & k = \pm 2 \\ 0 & \text{otherwise.} \end{cases}$$

17.37 (w) If a random process has the PSD $P_X(f) = 1 + \cos(2\pi f)$, are the samples of the random process uncorrelated?

17.38 (⌣) (f) If a random process has the PSD $P_X(f) = |1 + \exp(-j2\pi f) + (1/2)\exp(-j4\pi f)|^2$, determine the ACS.

17.39 (c) For the AR random processes whose ACSs are shown in Figure 17.6 generate a realization of $N = 2000$ samples for each process. Use the MATLAB code segment given in Section 17.4 to do this. Then, estimate the ACS for $k = 0, 1, \ldots, 30$ and plot the results. Compare your results to those shown in Figure 17.12 and explain.

17.40 (⌣) (w) A PSD is given as $P_X(f) = a + b\cos(2\pi f)$ for some constants a and b. What values of a and b will result in a valid PSD?

17.41 (f) A PSD is given as

$$P_X(f) = \begin{cases} 2 - 8f & 0 \le f \le 1/4 \\ 0 & 1/4 < f \le 1/2. \end{cases}$$

Plot the PSD and find the total average power in the random process.

17.42 (⌣) (c) Plot 50 realizations of the randomly phased sinusoid described in Example 17.4 with $N = 50$, and overlay the samples in a scatter diagram plot such as shown in Figure 16.15. Explain the results by referring to the PDF of Figure 16.12. . Next estimate the following quantities: $E[X[10]], E[X[12]], E[X[10]X[12]], E[X[12]X[14]]$ by averaging down the ensemble, and compare your simulated results to the theoretical values.

17.43 (c) In this problem we support the results of Problem 17.18 by using a computer simulation. Specifically, generate $M = 10,000$ realizations of the AR random process $X[n] = 0.95X[n-1] + U[n]$ for $n = 0, 1, \ldots, 49$, where $U[n]$ is WGN with $\sigma_U^2 = 1$. Do so two ways: for the first set of realizations let $X[-1] = 0$ and for the second set of realizations let $X[-1] \sim \mathcal{N}(0, \sigma_U^2/(1-a^2))$, using a different random variable for each realization. Now estimate the variance for each sample time n, which is $r_X[0]$, by averaging $X^2[n]$ down the ensemble of realizations. Do you obtain the theoretical result of $r_X[0] = \sigma_U^2/(1-a^2)$?

17.44 (⌣) (c) Generate a realization of discrete-time white Gaussian noise with variance $\sigma_X^2 = 1$. For $N = 64$, $N = 128$, and $N = 256$, plot the periodogram. What is the true PSD? Does your estimated PSD get closer to the true PSD as N increases? If not, how could you improve your estimate?

17.45 (c) Generate a realization of an AR random process of length $N = 31,000$ with $a = 0.25$ and $\sigma_U^2 = 1 - a^2$. Break up the data set into 1000 nonoverlapping blocks of data and compute the periodogram for each block. Finally, average the periodograms together for each point in frequency to determine the final averaged periodogram estimate. Compare your results to the theoretical PSD shown in Figure 17.11a.

17.46 (f) A continuous-time randomly phased sinusoid is defined by $X(t) = \cos(2\pi F_0 t + \Theta)$, where $\Theta \sim \mathcal{U}(0, 2\pi)$. Determine the mean function and ACF for this random process.

17.47 (⌣) (f) For the PSD $P_X(F) = \exp(-|F|)$, determine the average power in the band [10, 100] Hz.

17.48 (w) If a PSD is given as $P_X(F) = \exp(-|F/F_0|)$, what happens to the ACF as F_0 increases and also as $F_0 \to \infty$?

17.49 (t) Based on (17.49) derive (17.50), and also based on (17.51) derive (17.52).

17.50 (⌣) (w) A continuous-time white noise random process $U(t)$ whose PSD is given as $P_U(F) = N_0/2$ is integrated to form the continuous-time MA random process

$$X(t) = \frac{1}{T} \int_{t-T}^{t} U(\xi) d\xi.$$

Determine the mean function and the variance of $X(t)$. Does $X(t)$ have infinite total average power?

17.51 (⌣) (w,c) Consider a continuous-time random process $X(t) = \mu + U(t)$, where $U(t)$ is zero mean and has the ACF given in Figure 17.15. If $X(t)$ is sampled at twice the Nyquist rate, which is $F_s = 4W$, determine the ACS of $X[n]$. Next using (17.28) find the variance of the sample mean estimator $\hat{\mu}_N$ for $N = 20$. Is it half of the variance of the sample mean estimator if we had sampled at the Nyquist rate and used $N = 10$ samples in our estimate? Note that in either case the total length of the data interval in seconds is the same, which is $20/(4W) = 10/(2W)$.

17.52 (f) A PSD is given as

$$P_X(f) = \left|1 + \frac{1}{2}\exp(-j2\pi f)\right|^2.$$

Model this PSD by using an AR PSD as was done in Section 17.9. Plot the true PSD and the AR model PSD.