C H A P T E R 10

Power Spectral Density

INTRODUCTION

Understanding how the strength of a signal is distributed in the frequency domain, relative to the strengths of other ambient signals, is central to the design of any LTI filter intended to extract or suppress the signal. We know this well in the case of deterministic signals, and it turns out to be just as true in the case of random signals. For instance, if a measured waveform is an audio signal (modeled as a random process since the specific audio signal isn't known) with additive disturbance signals, you might want to build a lowpass LTI filter to extract the audio and suppress the disturbance signals. We would need to decide where to place the cutoff frequency of the filter.

There are two immediate challenges we confront in trying to find an appropriate frequency-domain description for a WSS random process. First, individual sample functions typically don't have transforms that are ordinary, well-behaved functions of frequency; rather, their transforms are only defined in the sense of generalized functions. Second, since the particular sample function is determined as the outcome of a probabilistic experiment, its features will actually be random, so we have to search for features of the transforms that are representative of the whole class of sample functions, i.e., of the random process as a whole.

It turns out that the key is to focus on the *expected power* in the signal. This is a measure of signal strength that meshes nicely with the second-moment characterizations we have for WSS processes, as we show in this chapter. For a process that is second-order ergodic, this will also correspond to the time average power in any realization. We introduce the discussion using the case of CT WSS processes, but the DT case follows very similarly.

10.1 EXPECTED INSTANTANEOUS POWER AND POWER SPECTRAL DENSITY

Motivated by situations in which $x(t)$ is the voltage across (or current through) a unit resistor, we refer to $x^2(t)$ as the *instantaneous power* in the signal $x(t)$. When $x(t)$ is WSS, the *expected* instantaneous power is given by

$$E[x^2(t)] = R_{xx}(0) = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_{xx}(j\omega)\, d\omega \ , \tag{10.1}$$

where $S_{xx}(j\omega)$ is the CTFT of the autocorrelation function $R_{xx}(\tau)$. Furthermore, when $x(t)$ is ergodic in correlation, so that time averages and ensemble averages are equal in correlation computations, then (10.1) also represents the time-average power in any ensemble member. Note that since $R_{xx}(\tau) = R_{xx}(-\tau)$, we know $S_{xx}(j\omega)$ is always *real* and *even* in ω; a simpler notation such as $P_{xx}(\omega)$ might therefore have been more appropriate for it, but we shall stick to $S_{xx}(j\omega)$ to avoid a proliferation of notational conventions, and to keep apparent the fact that this quantity is the Fourier transform of $R_{xx}(\tau)$.

The integral above suggests that we might be able to consider the expected (instantaneous) power (or, assuming the process is ergodic, the time-average power) in a frequency band of width $d\omega$ to be given by $(1/2\pi)S_{xx}(j\omega)\,d\omega$. To examine this thought further, consider extracting a band of frequency components of $x(t)$ by passing $x(t)$ through an ideal bandpass filter, shown in Figure 10.1.

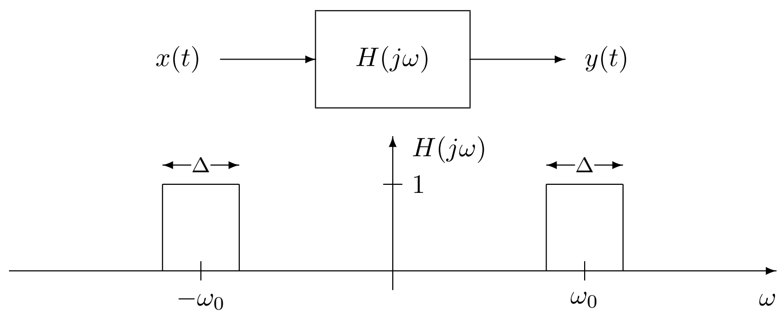

FIGURE 10.1 Ideal bandpass filter to extract a band of frequencies from input, $x(t)$.

Because of the way we are obtaining $y(t)$ from $x(t)$, the expected power in the output $y(t)$ can be interpreted as the expected power that $x(t)$ has in the selected passband. Using the fact that

$$S_{yy}(j\omega) = |H(j\omega)|^2 S_{xx}(j\omega)\ , \tag{10.2}$$

we see that this expected power can be computed as

$$E\{y^2(t)\} = R_{yy}(0) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} S_{yy}(j\omega)\,d\omega = \frac{1}{2\pi}\int_{\text{passband}} S_{xx}(j\omega)\,d\omega\ . \tag{10.3}$$

Thus

$$\frac{1}{2\pi}\int_{\text{passband}} S_{xx}(j\omega)\,d\omega \tag{10.4}$$

is indeed the expected power of $x(t)$ in the passband. It is therefore reasonable to call $S_{xx}(j\omega)$ the **power spectral density (PSD)** of $x(t)$. Note that the instantaneous power of $y(t)$, and hence the expected instantaneous power $E[y^2(t)]$, is always nonnegative, no matter how narrow the passband, It follows that, in addition to being real and even in ω, the PSD is always nonnegative, $S_{xx}(j\omega) \geq 0$ for all ω. While the PSD $S_{xx}(j\omega)$ is the Fourier transform of the autocorrelation function, it is useful to have a name for the *Laplace* transform of the autocorrelation function; we shall refer to $S_{xx}(s)$ as the *complex* PSD.

Exactly parallel results apply for the DT case, leading to the conclusion that $S_{xx}(e^{j\Omega})$ is the power spectral density of $x[n]$.

10.2 EINSTEIN-WIENER-KHINCHIN THEOREM ON EXPECTED TIME-AVERAGED POWER

The previous section defined the PSD as the transform of the autocorrelation function, and provided an interpretation of this transform. We now develop an alternative route to the PSD. Consider a random realization $x(t)$ of a WSS process. We have already mentioned the difficulties with trying to take the CTFT of $x(t)$ directly, so we proceed indirectly. Let $x_T(t)$ be the signal obtained by windowing $x(t)$, so it equals $x(t)$ in the interval $(-T\,,\,T)$ but is 0 outside this interval. Thus

$$x_T(t) = w_T(t)\,x(t)\ , \tag{10.5}$$

where we define the *window* function $w_T(t)$ to be 1 for $|t| < T$ and 0 otherwise. Let $X_T(j\omega)$ denote the Fourier transform of $x_T(t)$; note that because the signal $x_T(t)$ is nonzero only over the *finite* interval $(-T, T)$, its Fourier transform is typically well defined. We know that the **energy spectral density (ESD)** $\overline{S}_{xx}(j\omega)$ of $x_T(t)$ is given by

$$\overline{S}_{xx}(j\omega) = |X_T(j\omega)|^2 \tag{10.6}$$

and that this ESD is actually the Fourier transform of $x_T(\tau) * x_T^{\leftarrow}(\tau)$, where $x_T^{\leftarrow}(t) = x_T(-t)$. We thus have the CTFT pair

$$x_T(\tau) * x_T^{\leftarrow}(\tau) = \int_{-\infty}^{\infty} w_T(\alpha)w_T(\alpha-\tau)x(\alpha)x(\alpha-\tau)\,d\alpha \Leftrightarrow |X_T(j\omega)|^2\ , \tag{10.7}$$

or, dividing both sides by $2T$ (which is valid, since scaling a signal by a constant scales its Fourier transform by the same amount),

$$\frac{1}{2T}\int_{-\infty}^{\infty} w_T(\alpha)w_T(\alpha-\tau)x(\alpha)x(\alpha-\tau)\,d\alpha \Leftrightarrow \frac{1}{2T}\,|X_T(j\omega)|^2\ . \tag{10.8}$$

The quantity on the right is what we defined (for the DT case) as the **periodogram** of the finite-length signal $x_T(t)$.

Because the Fourier transform operation is linear, the Fourier transform of the expected value of a signal is the expected value of the Fourier transform. We may therefore take expectations of both sides in the preceding equation. Since $E[x(\alpha)x(\alpha-\tau)] = R_{xx}(\tau)$, we conclude that

$$R_{xx}(\tau)\Lambda(\tau) \Leftrightarrow \frac{1}{2T}\,E[|X_T(j\omega)|^2]\ , \tag{10.9}$$

where $\Lambda(\tau)$ is a triangular pulse of height 1 at the origin and decaying to 0 at $|\tau| = 2T$, the result of carrying out the convolution $w_T * w_T^{\leftarrow}(\tau)$ and dividing by

$2T$. Now taking the limit as T goes to ∞, we arrive at the result

$$R_{xx}(\tau) \Leftrightarrow S_{xx}(j\omega) = \lim_{T\to\infty} \frac{1}{2T} E[|X_T(j\omega)|^2] \, . \tag{10.10}$$

This is the **Einstein-Wiener-Khinchin** theorem (proved by Wiener, and independently by Khinchin, in the early 1930's, but — as only recently recognized — stated by Einstein in 1914).

The result is important to us because it underlies a basic method for estimating $S_{xx}(j\omega)$: with a given T, compute the periodogram for several realizations of the random process (i.e., in several independent experiments), and average the results. Increasing the number of realizations over which the averaging is done will reduce the noise in the estimate, while repeating the entire procedure for larger T will improve the frequency resolution of the estimate.

10.2.1 System Identification Using Random Processes as Input

Consider the problem of determining or "identifying" the impulse response $h[n]$ of a stable LTI system from measurements of the input $x[n]$ and output $y[n]$, as indicated in Figure 10.2.

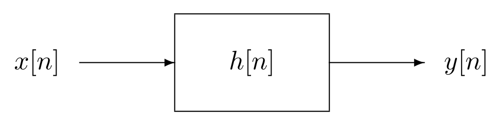

FIGURE 10.2 System with impulse response $h[n]$ to be determined.

The most straightforward approach is to choose the input to be a unit impulse $x[n] = \delta[n]$, and to measure the corresponding output $y[n]$, which by definition is the impulse response. It is often the case in practice, however, that we do not wish to — or are unable to — pick this simple input.

For instance, to obtain a reliable estimate of the impulse response in the presence of measurement errors, we may wish to use a more "energetic" input, one that excites the system more strongly. There are generally limits to the amplitude we can use on the input signal, so to get more energy we have to cause the input to act over a longer time. We could then compute $h[n]$ by evaluating the inverse transform of $H(e^{j\Omega})$, which in turn could be determined as the ratio $Y(e^{j\Omega})/X(e^{j\Omega})$. Care has to be taken, however, to ensure that $X(e^{j\Omega}) \neq 0$ for any Ω; in other words, the input has to be sufficiently "rich". In particular, the input cannot be just a finite linear combination of sinusoids (unless the LTI system is such that knowledge of its frequency response at a finite number of frequencies serves to determine the frequency response at all frequencies — which would be the case with a lumped system, i.e., a finite-order system, except that one would need to know an upper bound on the order of the system so as to have a sufficient number of sinusoids combined in the input).

The above constraints might suggest using a *randomly generated input* signal. For instance, suppose we let the input be a *Bernoulli process*, with $x[n]$ for each n taking the value $+1$ or -1 with equal probability, independently of the values taken at other times. This process is (strict- and) wide-sense stationary, with mean value 0 and autocorrelation function $R_{xx}[m] = \delta[m]$. The corresponding power spectral density $S_{xx}(e^{j\Omega})$ is flat at the value 1 over the entire frequency range $\Omega \in [-\pi, \pi]$; evidently the expected power of $x[n]$ is distributed evenly over all frequencies. A process with flat power spectrum is referred to as a **white process** (a term that is motivated by the rough notion that white light contains all visible frequencies in equal amounts); a process that is not white is termed **colored**.

Now consider what the DTFT $X(e^{j\Omega})$ might look like for a typical sample function of a Bernoulli process. A typical sample function is not absolutely summable or square summable, and so does not fall into either of the categories for which we know that there are nicely behaved DTFTs. We might expect that the DTFT exists in some generalized-function sense (since the sample functions are bounded, and therefore do not grow faster than polynomially with n for large $|n|$), and this is indeed the case, but it is not a simple generalized function; not even as "nice" as the impulses or impulse trains or doublets that we are familiar with.

When the input $x[n]$ is a Bernoulli process, the output $y[n]$ will also be a WSS random process, and $Y(e^{j\Omega})$ will again not be a pleasant transform to deal with. However, recall that

$$R_{yx}[m] = h[m] * R_{xx}[m] \; , \tag{10.11}$$

so if we can estimate the cross-correlation of the input and output, we can determine the impulse response (for this case where $R_{xx}[m] = \delta[m]$) as $h[m] = R_{yx}[m]$. For a more general random process at the input, with a more general $R_{xx}[m]$, we can solve for $H(e^{j\Omega})$ by taking the Fourier transform of (10.11), obtaining

$$H(e^{j\Omega}) = \frac{S_{yx}(e^{j\Omega})}{S_{xx}(e^{j\Omega})} \; . \tag{10.12}$$

If the input is not accessible, and only its autocorrelation (or equivalently its PSD) is known, then we can still determine the *magnitude* of the frequency response, as long as we can estimate the autocorrelation (or PSD) of the output. In this case, we have

$$|H(e^{j\Omega})|^2 = \frac{S_{yy}(e^{j\Omega})}{S_{xx}(e^{j\Omega})} \; . \tag{10.13}$$

Given additional constraints or knowledge about the system, one can often determine a lot more (or even everything) about $H(e^{j\omega})$ from knowledge of its magnitude.

10.2.2 Invoking Ergodicity

How does one estimate $R_{yx}[m]$ and/or $R_{xx}[m]$ in an example such as the one above? The usual procedure is to assume (or prove) that the signals x and y are **ergodic**. What ergodicity permits — as we have noted earlier — is the replacement of an expectation or *ensemble average* by a *time average*, when computing the expected value of various functions of random variables associated with a stationary random process. Thus a WSS process $x[n]$ would be called *mean-ergodic* if

$$E\{x[n]\} = \lim_{N\to\infty} \frac{1}{2N+1} \sum_{k=-N}^{N} x[k] \; . \tag{10.14}$$

(The convergence on the right hand side involves a sequence of random variables, so there are subtleties involved in defining it precisely, but we bypass these issues in 6.011.) Similarly, for a pair of *jointly-correlation-ergodic* processes, we could replace the cross-correlation $E\{y[n+m]x[n]\}$ by the time average of $y[n+m]x[n]$.

What ergodicity generally requires is that values taken by a typical sample function over time be representative of the values taken across the ensemble. Intuitively, what this requires is that the correlation between samples taken at different times falls off fast enough. For instance, a sufficient condition for a WSS process $x[n]$ with finite variance to be mean-ergodic turns out to be that its autocovariance function $C_{xx}[m]$ tends to 0 as $|m|$ tends to ∞, which is the case with most of the examples we deal with in these notes. A more precise (necessary and sufficient) condition for mean-ergodicity is that the time-averaged autocovariance function $C_{xx}[m]$, weighted by a triangular window, be 0:

$$\lim_{L\to\infty} \frac{1}{2L+1} \sum_{m=-L}^{L} \left(1 - \frac{|m|}{L+1}\right) C_{xx}[m] = 0 \; . \tag{10.15}$$

A similar statement holds in the CT case. More stringent conditions (involving fourth moments rather than just second moments) are needed to ensure that a process is second-order ergodic; however, these conditions are typically satisfied for the processes we consider, where the correlations decay exponentially with lag.

10.2.3 Modeling Filters and Whitening Filters

There are various detection and estimation problems that are relatively easy to formulate, solve, and analyze when some random process that is involved in the problem — for instance, the set of measurements — is white, i.e., has a flat spectral density. When the process is colored rather than white, the easier results from the white case can still often be invoked in some appropriate way if:

(a) the colored process is the result of passing a white process through some LTI *modeling* or *shaping* filter, which shapes the white process at the input into one that has the spectral characteristics of the given colored process at the output; or

(b) the colored process is transformable into a white process by passing it through an LTI *whitening* filter, which flattens out the spectral characteristics of the colored process presented at the input into those of the white noise obtained at the output.

Thus, a modeling or shaping filter is one that converts a white process to some colored process, while a whitening filter converts a colored process to a white process.

An important result that follows from thinking in terms of modeling filters is the following (stated and justified rather informally here — a more careful treatment is beyond our scope):

Key Fact: A real function $R_{xx}[m]$ is the autocorrelation function of a real-valued WSS random process if and only if its transform $S_{xx}(e^{j\Omega})$ is real, even and non-negative. The transform in this case is the PSD of the process.

The *necessity* of these conditions on the transform of the candidate autocorrelation function follows from properties we have already established for autocorrelation functions and PSDs.

To argue that these conditions are also *sufficient*, suppose $S_{xx}(e^{j\Omega})$ has these properties, and assume for simplicity that it has no impulsive part. Then it has a real and even square root, which we may denote by $\sqrt{S_{xx}(e^{j\Omega})}$. Now construct a (possibly noncausal) *modeling filter* whose frequency response $H(e^{j\Omega})$ equals this square root; the unit-sample reponse of this filter is found by inverse-transforming $H(e^{j\Omega}) = \sqrt{S_{xx}(e^{j\Omega})}$. If we then apply to the input of this filter a (zero-mean) unit-variance white noise process, e.g., a Bernoulli process that has equal probabilities of taking $+1$ and -1 at each DT instant independently of every other instant, then the output will be a WSS process with PSD given by $|H(e^{j\Omega})|^2 = S_{xx}(e^{j\Omega})$, and hence with the specified autocorrelation function.

If the transform $S_{xx}(e^{j\Omega})$ had an impulse at the origin, we could capture this by adding an appropriate constant (determined by the impulse strength) to the output of a modeling filter, constructed as above by using only the non-impulsive part of the transform. For a pair of impulses at frequencies $\Omega = \pm\Omega_o \neq 0$ in the transform, we could similarly add a term of the form $A\cos(\Omega_o n + \Theta)$, where A is deterministic (and determined by the impulse strength) and Θ is independent of all other other variables, and uniform in $[0, 2\pi]$.

Similar statements can be made in the CT case.

We illustrate below the logic involved in designing a whitening filter for a particular example; the logic for a modeling filter is similar (actually, inverse) to this.

Consider the following discrete-time system shown in Figure 10.3.

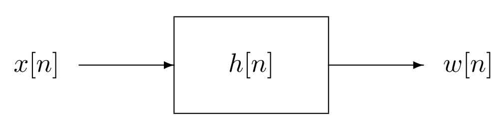

FIGURE 10.3 A discrete-time whitening filter.

Suppose that $x[n]$ is a process with autocorrelation function $R_{xx}[m]$ and PSD $S_{xx}(e^{j\Omega})$, i.e., $S_{xx}(e^{j\Omega}) = \mathcal{F}\{R_{xx}[m]\}$. We would like $w[n]$ to be a white noise output with variance σ_w^2.

We know that

$$S_{ww}(e^{j\Omega}) = |H(e^{j\Omega})|^2 \, S_{xx}(e^{j\Omega}) \tag{10.16}$$

or,

$$|H(e^{j\Omega})|^2 = \frac{\sigma_w^2}{S_{xx}(e^{j\Omega})}. \tag{10.17}$$

This then tells us what the squared magnitude of the frequency response of the LTI system must be to obtain a white noise output with variance σ_w^2. If we have $S_{xx}(e^{j\Omega})$ available as a rational function of $e^{j\Omega}$ (or can model it that way), then we can obtain $H(e^{j\Omega})$ by appropriate factorization of $|H(e^{j\Omega})|^2$.

EXAMPLE 10.1 Whitening filter

Suppose that

$$S_{xx}(e^{j\Omega}) = \frac{5}{4} - \cos(\Omega). \tag{10.18}$$

Then, to whiten $x(t)$, we require a stable LTI filter for which

$$|H(e^{j\Omega})|^2 = \frac{1}{(1 - \frac{1}{2}e^{j\Omega})(1 - \frac{1}{2}e^{-j\Omega})}, \tag{10.19}$$

or equivalently,

$$H(z)H(1/z) = \frac{1}{(1 - \frac{1}{2}z)(1 - \frac{1}{2}z^{-1})} \ . \tag{10.20}$$

The filter is constrained to be stable in order to produce a WSS output. One choice of $H(z)$ that results in a causal filter is

$$H(z) = \frac{1}{1 - \frac{1}{2}z^{-1}} \ , \tag{10.21}$$

with region of convergence (ROC) given by $|z| > \frac{1}{2}$. This system function could be multiplied by the system function $A(z)$ of any *allpass* system, i.e., a system function satisfying $A(z)A(z^{-1}) = 1$, and still produce the same whitening action, because $|A(e^{j\Omega})|^2 = 1$.

10.3 SAMPLING OF BANDLIMITED RANDOM PROCESSES

A WSS random process is termed *bandlimited* if its PSD is bandlimited, i.e., is zero for frequencies outside some finite band. For deterministic signals that are bandlimited, we can sample at or above the Nyquist rate and recover the signal exactly. We examine here whether we can do the same with bandlimited random processes.

In the discussion of sampling and DT processing of CT signals in your prior courses, the derivations and discussion rely heavily on picturing the effect in the frequency domain, i.e., tracking the Fourier transform of the continuous-time signal through the C/D (sampling) and D/C (reconstruction) process. While the arguments can alternatively be carried out directly in the time domain, for deterministic finite-energy signals the frequency domain development seems more conceptually clear.

As you might expect, results similar to the deterministic case hold for the reconstruction of bandlimited random processes from samples. However, since these stochastic signals do not possess Fourier transforms except in the generalized sense, we carry out the development for random processes directly in the time domain. An essentially parallel argument could have been used in the time domain for deterministic signals (by examining the total energy in the reconstruction error rather than the expected instantaneous power in the reconstruction error, which is what we focus on below).

The basic sampling and bandlimited reconstruction process should be familiar from your prior studies in signals and systems, and is depicted in Figure 10.4 below. In this figure we have explicitly used bold upper-case symbols for the signals to underscore that they are random processes.

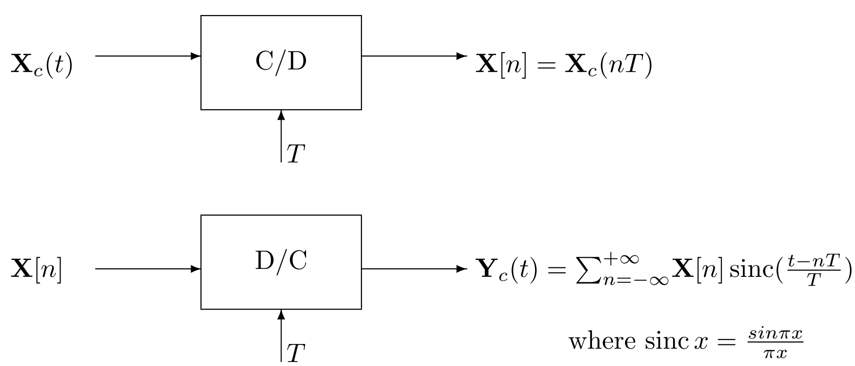

FIGURE 10.4 C/D and D/C for random processes.

For the deterministic case, we know that if $x_c(t)$ is bandlimited to less than $\frac{\pi}{T}$, then with the D/C reconstruction defined as

$$y_c(t) = \sum_n x[n] \, \mathrm{sinc}(\frac{t - nT}{T}) \tag{10.22}$$

it follows that $y_c(t) = x_c(t)$. In the case of random processes, what we show below is that, under the condition that $S_{x_c x_c}(j\omega)$, the power spectral density of $\mathbf{X}_c(t)$, is bandlimited to less that $\frac{\pi}{T}$, the mean square value of the error between $\mathbf{X}_c(t)$ and $\mathbf{Y}_c(t)$ is zero; i.e., if

$$S_{x_c x_c}(j\omega) = 0 \qquad |w| \geq \frac{\pi}{T} \, , \tag{10.23}$$

then

$$\mathcal{E} \triangleq E\{[\mathbf{X}_c(t) - \mathbf{Y}_c(t)]^2\} = 0 \,. \tag{10.24}$$

This, in effect, says that there is "zero power" in the error. (An alternative proof to the one below is outlined in Problem 13 at the end of this chapter.)

To develop the above result, we expand the error and use the definitions of the C/D (or sampling) and D/C (or ideal bandlimited interpolation) operations in Figure 10.4 to obtain

$$\mathcal{E} = E\{\mathbf{X}_c^2(t)\} + E\{\mathbf{Y}_c^2(t)\} - 2E\{\mathbf{Y}_c(t)\mathbf{X}_c(t)\} \,. \tag{10.25}$$

We first consider the last term, $E\{\mathbf{Y}_c(t)\mathbf{X}_c(t)\}$:

$$\begin{aligned} E\{\mathbf{Y}_c(t)\mathbf{X}_c(t)\} &= E\{\sum_{n=-\infty}^{+\infty} \mathbf{X}_c(nT)\,\mathrm{sinc}(\frac{t-nT}{T})\,\mathbf{X}_c(t)\} \\ &= \sum_{n=-\infty}^{+\infty} R_{x_cx_c}(nT-t)\,\mathrm{sinc}(\frac{nT-t}{T}) \qquad (10.26) \\ & \qquad (10.27) \end{aligned}$$

where, in the last expression, we have invoked the symmetry of sinc(.) to change the sign of its argument from the expression that precedes it.

Equation (10.26) can be evaluated using Parseval's relation in discrete time, which states that

$$\sum_{n=\infty}^{+\infty} v[n]w[n] = \frac{1}{2\pi}\int_{-\pi}^{\pi} V(e^{j\Omega})W^*(e^{j\Omega})d\Omega \tag{10.28}$$

To apply Parseval's relation, note that $R_{x_cx_c}(nT-t)$ can be viewed as the result of the C/D or sampling process depicted in Figure 10.5, in which the input is considered to be a function of the variable τ:

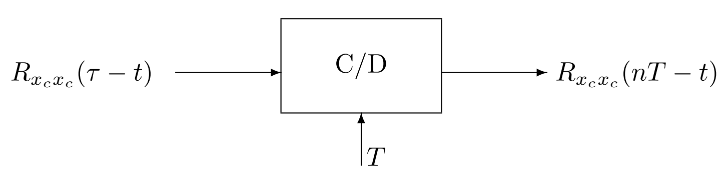

FIGURE 10.5 C/D applied to $R_{x_cx_c}(\tau - t)$.

The CTFT (in the variable τ) of $R_{x_cx_c}(\tau-t)$ is $e^{-j\omega t}S_{x_cx_c}(j\omega)$, and since this is bandlimited to $|\omega| < \frac{\pi}{T}$, the DTFT of its sampled version $R_{x_cx_c}(nT-t)$ is

$$\frac{1}{T}e^{\frac{-j\Omega t}{T}}S_{x_cx_c}(j\frac{\Omega}{T}) \tag{10.29}$$

in the interval $|\Omega| < \pi$. Similarly, the DTFT of $\text{sinc}(\frac{nT-t}{T})$ is $e^{\frac{-j\Omega t}{T}}$. Consequently, under the condition that $S_{x_cx_c}(j\omega)$ is bandlimited to $|\omega| < \frac{\pi}{T}$,

$$\begin{aligned} E\{\mathbf{Y}_c(t)\mathbf{X}_c(t)\} &= \frac{1}{2\pi T}\int_{-\pi}^{\pi} S_{x_cx_c}(\frac{j\Omega}{T})d\Omega \\ &= \frac{1}{2\pi}\int_{-(\pi/T)}^{(\pi/T)} S_{x_cx_c}(j\omega)d\omega \\ &= R_{x_cx_c}(0) = E\{\mathbf{X}_c^2(t)\} \end{aligned} \tag{10.30}$$

Next, we expand the middle term in equation (10.25):

$$\begin{aligned} E\{\mathbf{Y}_c^2(t)\} &= E\{\sum_n\sum_m \mathbf{X}_c(nT)\mathbf{X}_c(mT)\,\text{sinc}(\frac{t-nT}{T})\,\text{sinc}(\frac{t-mT}{T})\} \\ &= \sum_n\sum_m R_{x_cx_c}(nT-mT)\,\text{sinc}(\frac{t-mT}{T})\,\text{sinc}(\frac{t-mT}{T})\ . \end{aligned} \tag{10.31}$$

With the substitution $n - m = r$, we can express 10.31 as

$$E\{\mathbf{Y}_c^2(t)\} = \sum_r R_{x_cx_c}(rT)\sum_m \text{sinc}(\frac{t-mT}{T})\,\text{sinc}(\frac{t-mT-rT}{T})\ . \tag{10.32}$$

Using the identity

$$\sum_n \text{sinc}(n-\theta_1)\text{sinc}(n-\theta_2) = \text{sinc}(\theta_2-\theta_1)\ , \tag{10.33}$$

which again comes from Parseval's theorem (see Problem 12 at the end of this chapter), we have

$$\begin{aligned} E\{\mathbf{Y}_c^2(t)\} &= \sum_r R_{x_cx_c}(rT)\,\text{sinc}(r) \\ &= R_{x_cx_c}(0) = E\{\mathbf{X}_c^2\} \end{aligned} \tag{10.34}$$

since $\text{sinc}(r) = 1$ if $r = 0$ and zero otherwise. Substituting 10.31 and 10.34 into 10.25, we obtain the result that $\mathcal{E} = 0$, as desired.