HST.582J / 6.555J / 16.456J Biomedical Signal and Image Processing
Spring 2007

Harvard-MIT Division of Health Sciences and Technology
HST.582J: Biomedical Signal and Image Processing, Spring 2007
Course Director: Dr. Julie Greenberg

Chapter 12 - RANDOM SIGNALS AND LINEAR SYSTEMS

Introduction

In Chapter 2, we saw that the impulse response completely characterizes a linear, time-invariant system because the response to an arbitrary, but known, input can be computed by convolving the input with the impulse response. The impulse response plays a key role for random signals as it does for deterministic signals, with the important difference that it is used for computing *time averages* of the output from averages of the input. Specifically, we will show that knowing the impulse response suffices to derive the mean and autocorrelation function of the output from the mean and autocorrelation function of the input. That autocorrelation functions are involved is to be expected, since we showed in Chapter 11 that these functions naturally arise when processing random signals by linear filters.

In Chapter 3, we introduced Fourier analysis for deterministic signals, and showed that this concept leads to simplifications in the analysis and design of linear, time invariant systems. Frequency-domain techniques are as powerful for stationary random signals as they are for deterministic signals. They lead to the concepts of *power spectrum* and *Wiener filters,* which have numerous applications to system identification and signal detection in noise.

12.1 Response of LTI systems to random signals

Our goal in this section is to derive general formulas for the mean and autocorrelation of the response of a linear system to a stationary random signal, given both the system impulse response and the mean and autocorrelation function of the input.

12.1.1 Mean of $y[n]$

Let $x[n]$ be a random signal used as input to an LTI system with impulse response $h[n]$. The mean of the output $y[n]$ is:

$$< y[n] > = < x[n] * h[n] > = < \sum_{m=-\infty}^{\infty} h[m]x[n-m] >_n = \sum_{m=-\infty}^{\infty} h[m] \; < x[n-m] >_n \quad (12.1)$$

In this expression, the notation $< . >_n$ is used to specify that averaging is over the time variable n rather than the parameter m. Because we average over n, $h[m]$ is constant, and we can write $< h[m]x[n-m] >_n = \; h[m] \; < x[n-m] >_n$. By stationarity, we further obtain:

$$< y[n] > = < x[n] > \sum_{n=-\infty}^{\infty} h[n] \quad (12.2a)$$

As a special case of (12.2), if $x[n]$ has zero mean and the system is stable, the output also has zero mean.

Further simplification can be obtained by introducing the frequency response $H(f)$ and using the initial value theorem:

$$< y[n] > = \ H(0) \ < x[n] > \tag{12.2b}$$

This formula has an intuitive interpretation: The DC component (mean) of the output is the DC component of the input multiplied by the frequency response evaluated at DC.

12.1.2 Crosscorrelation function between $x[n]$ and $y[n]$

To obtain the autocorrelation function of the output, it is easier to first derive the crosscorrelation function between input and output.

$$R_{xy}[k] \ = < x[n]y[n{+}k] >_n = < x[n] \sum_{m=-\infty}^{\infty} h[m]x[n{+}k{-}m] >_n = \sum_{m=-\infty}^{\infty} h[m] \ R_x[k{-}m] \tag{12.3}$$

This gives the simple result:

$$R_{xy}[k] \ = \ h[k] \ * \ R_x[k] \tag{12.4}$$

Thus, the crosscorrelation function between the input and the output is the convolution of the autocorrelation function of the input with the impulse response of the filter.

As an important special case, if the input $w[n]$ is zero-mean, white noise with variance σ_w^2, the crosscorrelation function between input and output is

$$R_{wy}[k] \ = \ \sigma_w^2 \ h[k] \tag{12.5}$$

This result is the basis for a widely-used method of *system identification:* In order to measure the impulse response of an unknown LTI system, a white signal is used as input to the system, and the crosscorrelation function between input and output is computed, giving an estimate of the impulse response. It can be shown that this method will also work if the white noise is not the only input (provided that the other inputs are uncorrelated with the white noise), and if the linear system is followed by a memoryless nonlinearity *(Price's theorem).* The mathematician Norbert Wiener has further shown that, in principle, white noise inputs can be used for the identification of a wider class of nonlinear systems, but such methods are difficult to implement because small measurement errors or computational inaccuracies can greatly affect the results.

12.1.3 Autocorrelation function of $y[n]$

We can now give a general formula for the autocorrelation function of the output of an LTI system:

$$R_y[k] \ = < y[n] \ y[n-k] >_n \ = < y[n] \sum_{m=-\infty}^{\infty} h[m] \ x[n-k-m] >_n = \sum_{m=-\infty}^{\infty} h[m] \ R_{xy}[m+k] \tag{12.6}$$

Again, we obtain a simple result:

$$R_y[k] \;=\; h[-k] \;*\; R_{xy}[k] \tag{12.7}$$

Combining (12.4) and (12.7) yields:

$$R_y[k] \;=\; h[-k] \;*\; (h[k] \;*\; R_x[k]) \;=\; (h[k] \;*\; h[-k]) \;*\; R_x[k] \;=\; \tilde{R}_h[k] \;*\; R_x[k] \tag{12.8a}$$

with

$$\tilde{R}_h[k] \;\stackrel{\triangle}{=}\; h[k] \;*\; h[-k] \;=\; \sum_{n=-\infty}^{\infty} h[n]\; h[n+k] \tag{12.8b}$$

The function $\tilde{R}_h[k]$ is called the *deterministic autocorrelation function* of $h[k]$. It resembles a true autocorrelation function, with the important difference that the autocorrelation function of a random signal is an *average* over time, applicable to signals that have a finite mean power (but an infinite energy), while a deterministic autocorrelation function is a *sum* over time applicable to signals that have a finite energy, but zero mean power. Deterministic autocorrelation function have similar properties to those of true autocorrelation functions: They are even, and have a maximum at the origin.

Because $y[n] - \mu_y$ is the response of the filter to the centered signal $x[n] - \mu_x$, (12.8a) also holds for the autocovariance function, which is the autocorrelation function of the centered signal:

$$C_y[k] \;=\; \tilde{R}_h[k] \;*\; C_x[k] \tag{12.9}$$

12.1.4 Example

For example, consider a first-order FIR filter that approximates a differentiator:

$$y[n] \;=\; x[n] \;-\; x[n-1]$$

The deterministic autocorrelation function can be determined by inspection as:

$$\tilde{R}_h[k] \;=\; 2\;\delta[k] \;-\; \delta[k-1] \;-\; \delta[k+1]$$

From (12.8a), the autocorrelation of the output is given as a function of the input as:

$$R_y[k] \;=\; 2\;R_x[k] \;-\; R_x[k-1] \;+\; R_x[k+1]$$

12.1.5 White noise inputs

The case of white-noise inputs is again of special interest. For these inputs, $C_w[k] \;=\; \sigma_w^2\;\delta[k]$, so that the autocovariance function of the output signal is given by:

$$C_y[k] \;=\; \sigma_w^2\;\tilde{R}_h[k] \tag{12.10}$$

and, in particular, the variance of the output signal is:

$$\sigma_y^2 \;=\; C_y[0] \;=\; \sigma_w^2 \sum_{n=-\infty}^{\infty} h[n]^2 \;=\; \sigma_w^2\; E_h \tag{12.11}$$

Thus, the variance of the output is equal to the variance of the white input multiplied by the energy in the impulse response. This simple relationship only holds for white inputs: In general, the variance of a filtered random signal depends not only on the variance of the input, but also on the values of its covariance for non-zero lags.

12.1.6 Example 1: First-order autoregressive filter

Consider now a first-order lowpass filter with unit sample-response $a^n\ u[n]$ excited by zero-mean, white noise. The deterministic autocorrelation function of the filter is:

$$\tilde{R}_h[k] \;=\; \sum_{n=0}^{\infty} a^n\ a^{n+|k|} \;=\; a^{|k|} \sum_{n=0}^{\infty} a^{2n} \;=\; \frac{a^{|k|}}{1\;-\;a^2} \tag{12.12}$$

If the variance of a zero-mean, white input is σ_w^2, then the autocorrelation function of the output is:

$$R_y[k] \;=\; \sigma_w^2\ \frac{a^{|k|}}{1\;-\;a^2} \tag{12.13}$$

and in particular the variance is

$$\sigma_y^2 \;=\; \frac{\sigma_w^2}{1\;-\;a^2} \tag{12.14}$$

$R_y[k]$ is shown in Fig. 12.1A for several values of the parameter a. As a increases from 0 to 1, the autocorrelation spreads to larger lags, indicating that the memory of the system increases.

12.1.7 Example 2: N-point moving average

We introduce another example that will be useful in Chapter 13 for estimating parameters of random signals from finite data: the rectangular, or "boxcar" filter of length N. The impulse response of the filter is:

$$\Pi_N[n] \;\triangleq\; \begin{cases} 1 & \text{if } 0 \le n \le N-1 \\ 0 & \text{otherwise} \end{cases} \tag{12.15}$$

It can be verified that its deterministic autocorrelation function is a triangular function of width $2N-1$ and height N:

$$\Lambda_N[k] \triangleq \Pi_N[k] * \Pi_N[-k] = \begin{cases} N-|k| & \text{if } |k| < N \\ 0 & \text{otherwise} \end{cases} \tag{12.16}$$

If a random signal $x[n]$ with covariance function $C_x[k]$ is used as input to this rectangular filter, the variance of the output $y[n]$ will be:

$$\sigma_y^2 \;=\; C_y[0] \;=\; (C_x[k]\ *\ \Lambda_N[k])_{/k=0} \;=\; \sum_{n=-(N-1)}^{N-1} (N\;-\;|k|)\ C_x[k] \tag{12.17}$$

Thus, σ_y^2 depends on the values of $C_x[k]$ for $-(N-1)\ \le\ k\ \le\ N-1$. This is a general result for all FIR filters of length N.

In the special case of a white noise input with variance σ_w^2, this expression simplifies to:

$$\sigma_y^2 \;=\; \sigma_w^2\ \Lambda_N[0] \;=\; N\ \sigma_w^2 \tag{12.18}$$

Thus, adding N consecutive samples of a white signal increases the variance by a factor of N. The rectangular filter also has the effect of introducing correlation between $N-1$ successive samples, i.e. $C_y[k] \neq 0$ for $|k| < N$. Figure 12.1C shows $C_y[k]$ for different values of N.

12.1.8 Generalization to two inputs and two outputs

The preceding results can be generalized to the case of two correlated inputs processed through two linear systems. Specifically, assume that the random signals $x_1[n]$ and $x_1[n]$ are processed through the filters $h_1[n]$ and $h_2[n]$, respectively, to give the two outputs $y_1[n]$ and $y_2[n]$. Further assume that the crosscorrelation function $R_{x_1x_2}[k]$ is known. We will derive an expression for $R_{y_1y_2}[k]$:

$$R_{y_1y_2}[k] \;= < y_1[n]\ y_2[n+k] >_n = < \sum_m\ h_1[m]\ x_1[n-m]\ y_2[n+k] >_n$$

$$R_{y_1y_2}[k] \;=\; \sum_m\ h_1[m]\ R_{x_1y_2}[k+m] \;=\; h_1[-k]\ *\ R_{x_1y_2}[k] \tag{12.19}$$

To complete the proof, we need to derive an expression for $R_{x_1y_2}[k]$

$$R_{x_1y_2}[k] \;= < x_1[n]\ y_2[n+k] >_n = < x_1[n]\ \sum_m\ h_2[m]\ x_2[n+k-m] >_n$$

$$R_{x_1y_2}[k] \;=\; \sum_m\ h_2[m]\ R_{x_1x_2}[k-m] \;=\; h_2[k]\ *\ R_{x_1x_2}[k] \tag{12.20a}$$

By symmetry, we also have

$$R_{x_2y_1}[k] \;=\; h_1[k]\ *\ R_{x_2x_1}[k] \;=\; h_1[k]\ *\ R_{x_1x_2}[-k] \tag{12.20b}$$

Combining (12.19) with (12.20a) yields:

$$R_{y_1y_2}[k] \;=\; h_1[-k]\ *\ h_2[k]\ *\ R_{x_1x_2}[k] \;=\; \tilde{R}_{h_1h_2}[k]\ *\ R_{x_1x_2}[k], \tag{12.21a}$$

where

$$\tilde{R}_{h_1h_2}[k] \;\stackrel{\triangle}{=} h_1[-k]\ *\ h_2[k] \;=\; \sum_n\ h_1[n]\ h_2[n+k] \tag{12.21b}$$

is the *deterministic crosscorrelation function* of $h_1[n]$ and $h_2[n]$. Three cases are of special interest:

1. If the two inputs are uncorrelated, i.e. if $R_{x_1x_2}[k] \;=\; 0$ for all k, then the outputs are also uncorrelated.

2. On the other hand, if the two inputs are identical, i.e. if $x_1[n] \;=\; x_2[n] \;=\; x[n]$, then

   $$R_{y_1y_2}[k] \;=\; \tilde{R}_{h_1h_2}[k]\ *\ R_x[k],$$

   In general, $R_{y_1y_2}[k] \neq 0$. Thus, the outputs of two filters excited by the same signal are, in general correlated. The only exception is if $\tilde{R}_{h_1h_2}[k] \;=\; 0$ for all k, a condition best expressed in the frequency domain.

3. If in addition $h_1[n] \;=\; h_2[n] \;=\; h[n]$, so that $y_1[n] \;=\; y_2[n] \;=\; y[n]$, then (12.21) reduces to (12.8), as expected.

12.2 The power spectrum

12.2.1 Definition and properties

We have just derived a general expression for the autocorrelation function of the response $y[n]$ of a linear filter $h[n]$ to a random input $x[n]$:

$$R_y[k] \;=\; R_x[k] \;*\; (h[k] \;*\; h[-k]) \;=\; R_x[k] \;*\; \tilde{R}_h[k] \qquad (12.22)$$

Because this equation includes a double convolution, it can be simplified by introducing the discrete time Fourier transform $S_x(f)$ of the autocorrelation function:

$$S_x(f) \;\triangleq\; \sum_{k=-\infty}^{\infty} R_x[k] \; e^{-j2\pi fk} \qquad (12.23a)$$

$S_x(f)$ is called the *power spectrum* of the random signal $x[n]$. The autocorrelation function can be recovered from the power spectrum by means of an inverse DTFT:

$$R_x[k] \;=\; \int_{-\frac{1}{2}}^{\frac{1}{2}} S_x(f) \; e^{j2\pi fk} \; df \qquad (12.23b)$$

Because $R_x[k]$ is always real and even, its transform $S_x(f)$ is also real and even:

$$R_x[k] \;=\; R_x[-k] \;=\; R_x^*[k] \longleftrightarrow S_x(f) \;=\; S_x(-f) \;=\; S_x^*(f) \qquad (12.24)$$

Reporting (12.23a) into (12.22), and making use of the convolution theorem leads to a simple expression for the power spectrum $S_y(f)$ of the output signal as a function of the power spectrum of the input:

$$S_y(f) \;=\; S_x(f) \; H(f) \; H(-f) \;=\; S_x(f) \; |H(f)|^2 \qquad (12.25)$$

Thus, the power spectrum of the output of an LTI system is the power spectrum of the input multiplied by the magnitude squared of the frequency response. This simple result has many important applications.

12.2.2 Physical interpretation

The power spectrum of a random signal represents the contribution of each frequency component of the signal to the total power in the signal. To see this, we first note that, applying the initial value theorem to (12.23b), the signal power is equal to the area under the power spectrum:

$$P_x \;=\; R_x[0] \;=\; \int_{-\frac{1}{2}}^{\frac{1}{2}} S_x(f) \; df \qquad (12.26)$$

Suppose now that $x[n]$ is input to a narrow bandpass filter $B(f)$ with center frequency f_0 and bandwidth Δf:

$$|B(f)| \triangleq \begin{cases} 1 & \text{for } f_0 - \frac{\Delta f}{2} < f < f_0 + \frac{\Delta f}{2} \text{ modulo } 1 \\ 0 & \text{otherwise} \end{cases} \qquad (12.27)$$

The power in the output signal $y[n]$ is:

$$P_y \;=\; \int_{-\frac{1}{2}}^{\frac{1}{2}} S_y(f)\; df \;=\; \int_{-\frac{1}{2}}^{\frac{1}{2}} S_x(f)\; |B(f)|^2\; df \tag{12.28}$$

Because the filter is very narrow, this integral is approximately the value of the integrand for $f \;=\; f_0$ multiplied by the width of the integration interval:

$$P_y \;\approx\; S_x(f_0)\; |B(f_0)|^2\; \Delta f \;\approx\; S_x(f_0)\; \Delta f \tag{12.29}$$

Thus, the power in the bandpass-filtered signal is equal to the power spectrum of the input at the center of the passband multiplied by the filter bandwidth. This shows that $S_x(f_0)$ represents the contribution of frequency components near f_0 to the power of $x[n]$. Because the power at the output of any bandpass filter is always nonnegative, the power spectrum must be nonnegative for all frequencies:

$$S_x(f) \;\geq\; 0 \qquad \text{for all} \;\; f \tag{12.30}$$

Writing this constraint for $f \;=\; 0$ gives an important property of autocorrelation functions:

$$\sum_{k=-\infty}^{\infty} R_x[k] \;=\; S_x(0) \;\geq\; 0 \tag{12.31}$$

In general, because the positivity condition holds not only for $f \;=\; 0$, but for all frequencies, it strongly constrains the set of possible autocorrelation functions.

12.2.3 Example 1: Sine wave

The autocorrelation function of a sine wave $s[n]$ with amplitude A and frequency f_0 is a cosine wave:

$$R_s[k] \;=\; \frac{A^2}{2}\; \cos\; 2\pi f_0 n \tag{12.32}$$

Therefore, the power spectrum consists of impulses at frequencies $\pm\; f_0$ modulo 1:

$$S_s(f) \;=\; \frac{A^2}{4}\; \left(\; \tilde{\delta}(f \;-\; f_0) \;+\; \tilde{\delta}(f \;+\; f_0)\; \right) \tag{12.33}$$

If the sine wave is input to a filter with frequency response $H(f)$, the power spectrum of the output will be

$$S_y(f) \;=\; |H(f_0)|^2\; \frac{A^2}{4}\; \left(\; \tilde{\delta}(f \;-\; f_0) \;+\; \tilde{\delta}(f \;+\; f_0)\; \right) \tag{12.34}$$

As expected, this is the power spectrum of a sine wave with amplitude $A|H(f_0)|$. In general, the presence of impulses in the power spectrum implies that the signal contains periodic components. A special case is the DC component, which appears as an impulse at the origin in the power spectrum. It is generally desirable to remove these periodic components, including the DC, before estimating the power spectrum of a random signal.

12.2.4 Example 2: White noise

The autocorrelation function of zero-mean, white noise $w[n]$ is an impulse at the origin:

$$R_w[k] \ = \ \sigma_w^2 \ \delta[k] \tag{12.35}$$

Therefore, its power spectrum is a constant equal to the variance:

$$S_w(f) \ = \ \sigma_w^2 \tag{12.36}$$

i.e., white noise has equal power at all frequencies. This result justifies the term "white noise" by analogy with white light which contains all visible frequencies of the electromagnetic spectrum.

From (12.25), if white noise, is used as input to a filter with frequency response $H(f)$, the power spectrum of the output $y[n]$ is equal to the magnitude square of the frequency response within a multiplicative factor:

$$S_y(f) \ = \ \sigma_w^2 \ |H(f)|^2 \tag{12.37}$$

This result has two important interpretations.

1. An arbitrary random signal with power spectrum $S_x(f)$ which does not contain impulses can be considered as the response of a filter with frequency response $|H(f)| \ = \ \sqrt{S_x(f)}$ to zero-mean, white noise with unit variance. This point of view is useful because it replaces the problem of estimating a power spectrum by the simpler one of estimating parameters of a linear filter. For example, in *autoregressive spectral estimation,* an all-pole filter model is fitted to random data. This technique is described in Chapter 8.

2. To generate noise with an arbitrary power spectrum, it suffices to pass white noise through a filter whose frequency response is the square root of the desired spectrum. This is always possible because, as shown above, power spectra are nonnegative.

12.2.5 Example 3: Differentiator

Let $y[n]$ be the response of the first-order differentiator to zero-mean, white noise, i.e. $y[n] \ = \ w[n] \ - \ w[n-1]$. We have previously derived the autocorrelation function:

$$R_y[k] \ = \ \sigma_w^2 \ (2 \ \delta[k] \ - \ \delta[k-1] \ - \ \delta[k+1])$$

Therefore, the power spectrum is

$$S_y(f) \ = \ 2 \ \sigma_w^2 \ (1 \ - \ \cos \ 2\pi f)$$

Appropriately, $S_y(f) \ = \ 0$ for $f \ = \ 0$ and increases monotonically to reach 4 for $f \ = \ \frac{1}{2}$.

12.2.6 Example 4: First-order autoregressive process

A first-order autoregressive process is obtained by passing zero-mean, white noise through a first-order, recursive lowpass filter:

$$y[n] \ = \ a \ y[n-1] \ + \ w[n] \tag{12.38}$$

The filter frequency response is:

$$H(f) \;=\; \frac{1}{1 \;-\; a\; e^{-j2\pi f}} \tag{12.39}$$

Applying (12.37) gives the power spectrum of the output:

$$S_y(f) \;=\; \sigma_w^2\; |H(f)|^2 \;=\; \frac{\sigma_w^2}{1 \;-\; 2\; a \cos\; 2\pi f \;+\; a^2} \tag{12.40}$$

Taking the inverse Fourier transform of $S_y(f)$ gives the autocorrelation function that we derived previously:

$$R_y[k] \;=\; \sigma_w^2\; \frac{a^{|k|}}{1 \;-\; a^2} \tag{12.13}$$

Fig. 12.1B shows $S_y(f)$ for different values of a. As a approaches 1, and the effective duration of the impulse response (and autocorrelation function) increases, the power spectrum becomes increasingly centered around the origin.

12.2.7 Example 5: Rectangular filter

Our last example is the output of a rectangular filter of length N to a zero-mean, white noise input. The autocorrelation function has a triangular shape:

$$R_y[k] \;=\; \sigma_w^2\; \Lambda_N[k], \tag{12.41}$$

where $\Lambda_N[k]$ is defined in (12.16). Therefore, the power spectrum is the square of a periodic sinc function:

$$S_y(f) \;=\; \sigma_w^2\; \frac{\sin^2\; \pi N f}{\sin^2\; \pi f} \tag{12.42}$$

The width of the main lobe is $1/N$, and the height N^2. Fig. 12.1D shows $S_y(f)$ for different values of N. As for the autoregressive process, increasing the filter length makes the output power spectrum increasing lowpass.

12.2.8 Physical units of the power spectrum

For continuous-time signals, the power spectrum is defined by the CTFT of the autocorrelation function $R_x(\tau)$:

$$S_x(F) \;=\; \int_{-\infty}^{\infty} R_x(\tau)\; e^{-j2\pi F\tau}\; d\tau \tag{12.43a}$$

$$R_x(\tau) \;=\; \int_{-\infty}^{\infty} S_x(F)\; e^{j2\pi F\tau}\; dF \tag{12.43b}$$

If the discrete-time signal $x[n]$ is obtained by sampling $x(t)$ at intervals of T_s, the integral in (12.43a) can be approximated by the sum:

$$S_x(F) \;\approx\; T_s \sum_{-\infty}^{\infty} R_x(kT_s)\; e^{-j2\pi F k T_s} \tag{12.44}$$

Thus, a more physical definition of the power spectrum of a discrete random signal $x[n]$ would be

$$S'_x(F) \triangleq T_s \sum_{-\infty}^{\infty} R_x[k]\, e^{-j2\pi F k T_s} \qquad (12.45a)$$

This definition differs from (12.23a) by a multiplicative factor:

$$S'_x(F) = T_s\, S_x(F/F_s) \qquad (12.46)$$

The power spectrum defined in (12.6) has units of Volt^2/Hz if $x[n]$ is expressed in Volts. The autocorrelation function (in Volt^2) can be recovered from $S'_x(F)$ through the formula:

$$R_x[k] = \int_{-F_s/2}^{F_s/2} S'_x(F)\, e^{j2\pi F k T_s}\, dF \qquad (12.45b)$$

Definition (12.45) can be used as an alternative to (12.23) in situations when physical units are important. Because this chapter is primarily concerned with mathematical properties, we use Definition (12.23).

12.2.9 The periodogram - Wiener-Kitchnin theorem - optional

Can the power spectrum be directly derived from a random signal without introducing the autocorrelation function as an intermediate? Specifically, assume that we have N samples of a random signal $\{x[n],\ n = 0, ...N-1\}$. Because the power spectrum represents the contribution of each frequency band to the total power, one might expect that it would be approximately proportional to the magnitude square of the Fourier transform of the data:

$$\hat{S}_x(f) = C\, |X_N(f)|^2 = C \left| \sum_{n=0}^{N-1} x[n]\, e^{-j2\pi f n} \right|^2 \qquad (12.47)$$

The proportionality constant C can be determined from dimensionality considerations. Specifically, we want the integral of the power spectrum over all frequencies to be equal to the mean power:

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} \hat{S}_x(f)\, df = P_x \qquad (12.48)$$

If N is sufficiently large, one has from Parseval's theorem

$$P_x \approx \frac{1}{N} \sum_{n=0}^{N-1} x[n]^2 = \frac{1}{N} \int_{-\frac{1}{2}}^{\frac{1}{2}} |X_N(f)|^2\, df \qquad (12.49)$$

Comparing with (12.47), it is clear that the proportionality constant C must equal $1/N$ for the power to be conserved. We are thus led to define the *periodogram* $\hat{S}_x(f)$ as an estimate of the power spectrum from the data $\{x[n],\ n = 0, ...N-1\}$:

$$\hat{S}_x(f) \triangleq \frac{1}{N} |X_N(f)|^2 = \frac{1}{N} \left| \sum_{n=0}^{N-1} x[n]\, e^{-j2\pi f n} \right|^2 \qquad (12.50)$$

One might hope that, as N goes to infinity, the periodogram would approach the power spectrum. Unfortunately, this is not the case. Because the periodogram is the transform of a signal segment of length N, it can be exactly reconstructed from N samples at frequencies $f_k = k/N$. The periodogram is said to have N *degrees of freedom.* Thus, as N is increased, more and more independent frequency samples must be computed, so that there is no averaging involved in the computation of the periodogram. For this reason, periodograms are very irregular in shape, regardless of the number of data points.

Despite these difficulties, we can derive a relation between the periodogram and the power spectrum by averaging periodograms for multiple data segments. Specifically, suppose that we have $N \times M$ samples of $\{x[n],\ n = 0, ..., NM - 1\}$. We divide the data in M consecutive segments $x_k[n]$, each of length N:

$$x_k[n] = \begin{cases} x[n + kN] & \text{for } 0 \leq n \leq N - 1 \\ 0 & \text{otherwise} \end{cases} \qquad k = 0, \ldots, M - 1 \tag{12.51}$$

and we compute a periodogram for each segment. The power spectrum is then the limit when M and N become large of the average periodogram:

$$S_x(f) = \lim_{M \to \infty, N \to \infty} \frac{1}{M} \sum_{k=0}^{M-1} \frac{1}{N} |X_k(f)|^2, \tag{12.52}$$

where $X_k(f)$ is the DTFT of the k-th segment $x_k[n]$. This identity between the power spectrum and the limit of the average periodogram is known as the *Wiener-Kitchnin theorem.* Averaging periodograms keeps the degrees of freedom (N) in the periodogram at only a fraction of the number of data points ($N \times M$), so that each frequency sample of the periodogram is effectively an average of a large number (M) of data points. In Chapter 13, we introduce techniques for reliably estimating the power spectrum from finite data records.

12.3 The cross spectrum

12.3.1 Definition

We have shown above that the crosscorrelation function between the input and the output of a linear filter is expressed by

$$R_{xy}[k] = h[k] * R_x[k] \tag{12.53}$$

This convolution can be simplified by introducing the DTFT of the cross-correlation function

$$S_{xy}(f) \triangleq \sum_{k=-\infty}^{\infty} R_{xy}[k]\, e^{-j2\pi fk} \tag{12.54}$$

$S_{xy}(f)$ is called the *cross-spectrum* of the signals $x[n]$ and $y[n]$. Unlike the power spectrum, it is not, in general, a positive nor even a real function. The order of the two signals x and y is important because

$$R_{yx}[k] = R_{xy}[-k] \longleftrightarrow S_{yx}(f) = S_{xy}(-f) = S^*_{xy}(f) \tag{12.55}$$

Taking the Fourier transform of (12.53) yields a simple relation between the cross-spectrum and the power spectrum of the input:

$$S_{xy}(f) = H(f)\, S_x(f) \tag{12.56}$$

This relation is particularly simple if the input is zero-mean, white noise $w[n]$ with variance σ_w^2:

$$S_{wy}(f) \;=\; \sigma_w^2 \; H(f) \tag{12.57}$$

Thus, in order to measure the frequency response of an unknown system, it suffices to estimate the cross-spectrum between a white noise input and the filter response. This technique is often used in system identification.

12.3.2 Physical interpretation

As the power spectrum, the cross spectrum evaluated at frequency f_0 has a simple physical interpretation if we introduce the ideal bandpass filter:

$$|B(f)| \triangleq \begin{cases} 1 & \text{for } f_0 - \frac{\Delta f}{2} < f < f_0 + \frac{\Delta f}{2} \text{ modulo } 1 \\ 0 & \text{otherwise} \end{cases} \tag{12.58}$$

Specifically, consider the arrangement of Fig. 12.2: $u[n]$ is the response of $B(f)$ to $x[n]$, and $v[n]$ the response of $B(f)$ to $y[n]$. Note that both $u[n]$ and $v[n]$ are complex because the impulse response $b[n]$ of the one-sided filter $B(f)$ is complex. We will show that

$$< u[n] \; v^*[n] > \approx \; S_{xy}(f_0) \; \Delta f \tag{12.59}$$

where * denotes the complex conjugate. Specializing Equation (12.21) with $x_1[n] \;=\; x[n]$, $x_2[n] \;=\; y[n]$, $h_1[n] \;=\; h_2[n] \;=\; b[n]$, $y_1[n] \;=\; u[n]$, and $y_2[n] \;=\; v[n]$, we obtain1

$$R_{uv}[k] \;\triangleq < u[n] \; v^*[n+k] > = \; b^*[-k] \;*\; b[k] \;*\; R_{xy}[k] \tag{12.60a}$$

[1] In the frequency domain, this becomes:

$$S_{uv}(f) \;=\; |B(f)|^2 \; S_{xy}(f) \tag{12.60b}$$

Applying the initial value theorem to the above expression gives:

$$< u[n] \; v^*[n] > = \; R_{uv}[0] \;=\; \int_{-\frac{1}{2}}^{\frac{1}{2}} \; |B(f)|^2 \; S_{xy}(f) \; df \tag{12.61}$$

For small Δf, the integral on the right side becomes approximately $S_{xy}(f_0) \; \Delta f$, completing the proof of (12.59). Thus, the cross spectrum evaluated at f_0 can be interpreted as the mean product of the frequency components of $x[n]$ and $y[n]$ centered at f_0.

12.3.3 Generalization to two inputs and two outputs

We have derived an expression for the crosscorrelation function between the outputs of two filters excited by two correlated inputs:

$$R_{y_1 y_2}[k] \;=\; h_1[-k] \;*\; h_2[k] \;*\; R_{x_1 x_2}[k] \tag{12.21a}$$

[1]For complex signals fuch as $u[n]$ and $v[n]$, a complex conjugate as in (12.60a) must be introduced in the definitions of the cross- and autocorrelation functions to ensure that the power is positive.

This formula takes a simple form in the frequency domain:

$$S_{y_1y_2}(f) \;=\; H_1^*(f)\; H_2(f)\; S_{x_1x_2}(f) \tag{12.62}$$

This result implies that the filter outputs can be uncorrelated (i.e. $S_{y_1y_2}(f) \;=\; 0$) even though the inputs are correlated so long that $|H_1(f)|\;|H_2(f)| \;=\; 0$ for all f, meaning that the filter frequency responses do not ovelap. This is the case, for example for a lowpass filter and a highpass filter having the same cutoff frequency.

12.4 Wiener filters

12.4.1 Linear least-squares estimation

In the preceding section, we have assumed that $h[n]$ and $x[n]$ were known, and derived an expression for the cross spectrum $S_{xy}(f)$. Many applications lead to a different problem that was first solved by the mathematician Norbert Wiener in the 1940's: Given two random signals $x[n]$ and $y[n]$, what is the filter $h[n]$ that does the best job of producing $y[n]$ from $x[n]$? This problem has important applications in both system modeling and signal conditioning. Specifically, let $\hat{y}[n]$ be the estimate of $y[n]$ obtained by processing $x[n]$ through the filter $h[n]$. By "best" filter, we mean the one that minimizes the mean power in the *error signal* $e[n] \stackrel{\triangle}{=} y[n] \;-\; \hat{y}[n]$:

$$P_e \;= < e[n]^2 > = < y([n] \;-\; \hat{y}[n])^2 > = < (y[n] \;-\; \sum_k h[k]\; x[n-k]\;)^2 > \tag{12.63}$$

Because the relation between the data $x[n]$ and the estimate $y[n]$ is a linear one, this is a *linear, least-squares estimation* problem. Making use of the linearity time averages, the *mean-square estimation error* P_e can be expanded into the expression:

$$P_e \;=\; P_y \;-\; 2\sum_k h[k]\; R_{xy}[k] \;+\; \sum_k \sum_l h[k]\; h[l]\; R_x[k-l] \tag{12.64}$$

The power in the error signal is a quadratic function of the filter coefficients $h[k]$. Therefore it has a single minimum which can be determined by setting to zero the partial derivatives of P_e with respect to the $h[k]$:

$$\frac{\partial P_e}{\partial h[k]} \;=\; 0 \tag{12.65}$$

This yields the system of linear equations:

$$R_{xy}[k] \;=\; \sum_l h[l]\; R_x[k-l] \tag{12.66}$$

It is easily verified that, if (12.66) holds, the prediction error can be written as:

$$P_e \;=\; P_y \;-\; \sum_k h[k]\; R_{xy}[k] \;=\; P_y \;-\; P_{\hat{y}} \tag{12.67a}$$

Thus, the power in the desired signal $y[n]$ is the sum of the power in the estimate $\hat{y}[n]$ and the power in the error signal $e[n]$. In fact it can be shown more generally that:

$$R_e[k] \;=\; R_y[k] \;-\; R_{\hat{y}}[k] \tag{12.67b}$$

For any two random signals $u[n]$ and $v[n]$, one has:

$$R_{u+v}[k] \;=\; R_u[k] \;+\; R_v[k] \;+\; R_{uv}[k] \;+\; R_{vu}[k] \tag{12.68}$$

Taking $u[n] \;=\; \hat{y}[n]$ and $v[n] \;=\; e[n]$, (12.67b) implies that $R_{\hat{y}e}[k] \;=\; 0$ for all k, i.e. that the error signal $e[n]$ is uncorrelated with the estimate $\hat{y}[n]$. Because $\hat{y}[n]$ is a weighted sum of input samples $x[n-k]$, this also means that the error is uncorrelated with the observations $x[n-k]$, i.e. the $R_{xe}[k] \;=\; 0$ equals zero for all k. The result that the error is uncorrelated with the observations is a general property of linear, least-squares estimation which can be used to derive the system of equations (12.66).

12.4.2 Non-causal Wiener filter

In general, solving the system of equations (12.66) requires knowledge of $R_{xy}[k]$ and $R_x[k]$ for all k. The exact solution depends on constraints on the filter $h[n]$. For example, if $h[n]$ is constrained to be a causal, FIR filter of length N, i.e. if it is zero outside of the interval $[0,\; N-1]$, (12.66) reduces to a system of N linear equations with N unknowns that can be solved by standard techniques. This is what we do for the special case $y[n] \;=\; x[n+1]$ when deriving the Yule-Walker equations for linear prediction in Chapter 8. There is another case in which the solution to (12.66) is easy to find: When there are no constraints on the filter, i.e. when (12.66) is to be solved for $-\infty < k < \infty$. In this case, the right side of (12.66) is the convolution $R_x[k] \,*\, h[k]$, so that a solution can be obtained by means of the Fourier transform:

$$H(f) \;=\; \frac{S_{xy}(f)}{S_x(f)} \tag{12.69}$$

$H(f)$ is called the *(non-causal) discrete-time Wiener filter.*

Note that (12.69) is the same as (12.56). This means that, if $y[n]$ were exactly derived from $x[n]$ by a filtering operation, the filter that provides the least-squares estimate of $y[n]$ from $x[n]$ would be the actual one.

12.4.3 Application to filtering of additive noise

A very common problem in signal processing is to filter a noisy signal in order to estimate a desired signal. Specifically, suppose that the noisy signal $x[n]$ is the sum of the desired signal $y[n]$ plus a disturbance $d[n]$ that is uncorrelated with $y[n]$:

$$x[n] \;=\; y[n] \;+\; d[n] \tag{12.70}$$

One has:

$$R_{xy}[k] \;=\; R_y[k] \;+\; R_{yd}[k] \;=\; R_y[k] \tag{12.71}$$

because $y[n]$ and $d[n]$ are uncorrelated. Therefore:

$$S_{xy}(f) \;=\; S_y(f) \tag{12.72}$$

Similarly, one has:

$$R_x[k] \;=\; R_y[k] \;+\; R_{yd}[k] \;+\; R_{dy}[k] \;+\; R_d[k] \;=\; R_y[k] \;+\; R_d[k] \tag{12.73}$$

so that:

$$S_x(f) \;=\; S_y(f) \;+\; S_d(f) \tag{12.74}$$

The optimal filter for estimating $y[n]$ from the noisy signal $x[n]$ is

$$H(f) \;=\; \frac{S_{xy}(f)}{S_x(f)} \;=\; \frac{S_y(f)}{S_y(f) \;+\; S_d(f)} \tag{12.75}$$

As expected, $H(f) \;\approx\; 1$ for frequencies where the signal-to-noise ratio $S_y(f)/S_d(f)$ is large, while $H(f) \;\approx\; 0$ when the signal-to-noise ratio is small. Also note that, because power spectra are real and even, $H(f)$ is also real and even, which means that $h[n]$ is symmetric with respect to the origin, and therefore non-causal. In applications that require causality, a causal filter could be obtained by approximating $h[n]$ by a finite-impulse response filter, then delaying the impulse response of the FIR filter by half its length.

12.4.4 Example

Let the disturbance be zero-mean, white noise with variance σ_w^2, and the signal a first-order autoregressive process with parameter a. The signal spectrum is given by (12.40):

$$S_y(f) \;=\; \frac{1}{1 \;-\; 2\,a \cos\; 2\pi f \;+\; a^2}$$

The Wiener filter is thus:

$$H(f) \;=\; \frac{S_y(f)}{S_y(f) \;+\; \sigma_w^2} \;=\; \frac{1}{1 \;+\; \sigma_w^2\;(1 \;-\; 2\,a \cos\; 2\pi f \;+\; a^{2)}}$$

Fig. 12.3A shows the spectra of the signal and the noise for $s_w^2 \;=\; 10$ and $a \;=\; 0.9$, while Fig. 12.3B shows the Wiener filter $H(f)$.

12.4.5 Applications of Wiener filters

Wiener filters have two main applications, *system identification,* and *signal conditioning.* In system identification, the goal is to model the unknown system that produces a known output $y[n]$ from a known input $x[n]$. This can be achieved in any of two ways. In *direct system identification* (Fig. 12.4A), the unknown system and the Wiener filter are placed in parallel, in the sense that both receive the same input $x[n]$. The goal is to find the filter $H(f)$ such that its response $\hat{y}[n]$ to $x[n]$ best estimates the output $y[n]$ of the unknown filter. On the other hand, in *inverse system identification* (Fig. 12.4B), the unknown filter and the Wiener filter are placed in series: The output $x[n]$ of the unknown system is used as input to the Wiener filter, and the goal is to make the output of the Wiener filter $\hat{y}[n]$ best estimate the input $y[n]$ to the unknown system. Thus, if $y[n]$ and $x[n]$ are related by a filter $G(f)$, the Wiener filter $H(f)$ would ideally be $1/G(f)$. This is the approach taken in linear prediction discussed in Chapter 8.

In conditioning applications of Wiener filters (Fig. 12.4C), the goal is to either cancel out the noise from a noisy signal, or to detect a signal in additive noise. In both cases, the signal to be estimated $y[n]$ is assumed to be the sum of two uncorrelated components $u[n]$ and $v[n]$. The signal to be filtered $x[n]$ is related to $v[n]$ by an unknown system (and therefore $x[n]$ and $v[n]$ are correlated), but $x[n]$ and $u[n]$ are uncorrelated. In *detection applications,* $v[n]$ is the signal and $u[n]$ the noise, so that the output $\hat{y}[n]$ of the Wiener filter is effectively an estimate of the signal $v[n]$ from the observations $x[n]$. The error signal $e[n] = y[n] - \hat{y}[n]$ is then an estimate of the noise $u[n]$. On the other hand, in *cancellation applications,* $u[n]$ is the signal, and $v[n]$ (and therefore $x[n]$) is noise. Thus, the output $\hat{y}[n]$ of the Wiener filter is an estimate of the noise $v[n]$ from $x[n]$, while the error signal $e[n] = y[n] - \hat{y}[n]$ is an estimate of the signal $u[n]$. This technique can be used, for example, to cancel 60-Hz components from recordings of the electrocardiogram.

12.5 Matched filters

The Wiener filter (12.75) gives the optimum linear estimate of a desired random signal $y[n]$ corrupted by additive noise $d[n]$. To implement this filter, the desired signal $y[n]$ does not have to be known exactly, only its power spectrum is needed. A different kind of optimum filter, the *matched filter,* is used in applications when the desired signal is known exactly. This is the case, for example, in radar and sonar applications, where the echo closely resembles the emitted pulse, except for a delay and a scale factor. We will first derive the matched filter for the case of additive white noise, then treat the general case of noise with an arbitrary, but known, power spectrum.

12.5.1 White-noise case

Our goal is to detect a known signal $s[n]$ in additive white noise $w[n]$. We further assume that $s[n]$ has finite energy, so that it is well localized in time. Specifically, let $x[n]$ be the sum of $s[n]$ and $w[n]$. We would like to design a digital filter $h[n]$ that optimizes our chances of detecting the known signal, in the sense that the signal-to-noise ratio at the output of the filter would be maximized for a particular time n_0. n_0 is the time when the signal is detected. The filter output $y[n]$ can be written as:

$$y[n] = h[n] * x[n] = h[n] * s[n] + h[n] * w[n] \tag{12.76}$$

This is the sum of a term $y_s[n] = h[n] * s[n]$ due to the signal and a term $y_w[n] = h[n] * w[n]$ due to the noise. We want to maximize the following signal-to-noise ratio:

$$SNR = \frac{\text{Power in } s[n] * h[n] \text{ at time } n_0}{\text{Mean power in } w[n] * h[n]} = \frac{y_s[n_0]^2}{P_{y_w}} \tag{12.77}$$

From (12.11), the mean power due to the white noise input is:

$$P_{y_w} = \sigma_w^2 \sum_{n=-\infty}^{\infty} h[n]^2 \tag{12.78}$$

The power due to the signal at time n_0 is:

$$y_s[n_0]^2 = \left[\sum_{n=-\infty}^{\infty} h[n]\; s[n_0 - n] \right]^2 \tag{12.79}$$

From the Cauchy-Schwarz inequality, one has:

$$\left[\sum_{n=-\infty}^{\infty} h[n]s[n_0-n]\right]^2 \leq \left[\sum_{n=-\infty}^{\infty} h[n]^2\right]\left[\sum_{n=-\infty}^{\infty} s[n_0-n]^2\right] \tag{12.80}$$

Therefore, the signal-to-noise ratio in (12.77) is bounded by:

$$SNR \;=\; \frac{y_s[n_0]^2}{P_{y_w}} \;\leq\; \frac{\sum_{-\infty}^{\infty} s[n]^2}{\sigma_w^2} \;=\; \frac{E_s}{\sigma_w^2} \tag{12.81}$$

Equality is obtained if and only if the two signals are proportional to each other:

$$h[n] \;=\; C\; s[n_0-n] \tag{12.82}$$

Thus, the impulse response of the filter that optimizes the signal-to-noise ratio is proportional to the signal reversed in time and delayed by n_0. Such a filter is called a *matched filter.* If $s[n]$ is of finite duration, $h[n]$ is an FIR filter, and the delay n_0 must be at least as long as the signal in order to make the filter causal. If causality is not an issue, n_0 can be set to zero.

The matched filter can also be expressed in the frequency domain:

$$H(f) \;=\; C\; S^*(f)\; e^{-j2\pi f n_0} \tag{12.83}$$

The signal-to-noise ratio at the matched filter output is the ratio of signal energy E_s to the noise power σ_w^2. Figure 12.5 illustrates the improvement in signal-to-noise ratio achieved by a matched filter for nerve impulses buried in white noise.

12.5.2 General case

The concept of matched filter can be easily extended to the general case of detecting a known signal in additive noise with arbitrary power spectrum. The basic idea is to introduce a *whitening filter* that transforms the additive noise into white noise, then apply the results of the preceding section for the white noise case.

Specifically, let $x[n] \;=\; s[n] \;+\; d[n]$ be the sum the the known signal $s[n]$ and noise $d[n]$ with power spectrum $S_d(f)$. A whitening filter $H_1(f)$ transforms $d[n]$ into white noise $w[n]$. From (12.25), $H_1(f)$ must be the inverse square root of $S_d(f)$ (within a phase factor):

$$H_1(f) = \frac{e^{j\Phi(f)}}{\sqrt{S_d(f)}}, \tag{12.84}$$

where $\Phi(f)$ is an arbitrary phase function. For example, $\phi(f)$ could be chosen to make $h_1[n]$ causal. The white signal $w[n] \;=\; d[n] \;*\; h_1[n]$ is called the *innovation* of $d[n]$.

The response $x_1[n]$ of $H_1(f)$ to $x[n]$ can be written as

$$x_1[n] \;=\; h_1[n] \;*\; x[n] \;=\; h_1[n] \;*\; s[n] \;+\; h_1[n] \;*\; d[n] \;=\; s_1[n] \;+\; w[n], \tag{12.85}$$

where $s_1[n] \;=\; s[n] \;*\; h_1[n]$. To maximize the SNR with input $x[n]$, it suffices to find a filter $H_2(f)$ that maximizes the SNR with input $x_1[n]$. From (12.83), the optimum (matched) filter $H_2(f)$ for detecting $s_1[n]$ in white noise $w[n]$ is:

$$H_2(f) \;=\; C\; S_1^*(f)\; e^{-j2\pi f n_0}, \tag{12.86a}$$

with

$$S_1(f) \;=\; S(f)\; H_1(f) \qquad (12.86b)$$

The overall matched filter $H(f)$ is the cascade of the whitening filter $H_1(f)$ and the matched filer $H_2(f)$ for white noise:

$$H(f) \;=\; H_1(f)\; H_2(f) \;=\; C\; S^*(f)\; |H_1(f)|^2\; e^{-j2\pi f n_0} \;=\; C\; \frac{S^*(f)}{S_d(f)}\; e^{-j2\pi f n_0} \qquad (12.87)$$

Note that the arbitrary phase factor $\Phi(f)$ cancels out when $H_1(f)$ is multiplied by its complex conjugate.

It is interesting to contrast the matched filter (12.87) with the Wiener filter (12.75). Both filters depend on the ratio of a signal spectrum to the noise power spectrum $S_d(f)$, so that filter attenuation will be large in frequency regions where the signal-to-noise ratio is low. The two filters differ in that, for the Wiener filter, the signal *power spectrum* $S_s(f)$ appears in the numerator, while for the matched filter it is the DTFT conjugate $S^*(f)$, which includes both a magnitude and a phase. Because the matched filter preserves phase information about the signal, it is effective even in situations when the signal and the noise occupy the same frequency region, while the Wiener filter is ineffective (flat) in this case. To render a matched filter ineffective, the noise *waveform* at the time of the signal would have to match the signal waveform, a very unlikely occurrence.

Matched filters provide optimal signal detection under the assumption of an exactly known signal. Fortunately, they are still effective when the signal is only approximately known. This makes matched filters suitable for applications in which the signal is not completely reproducible, such as detection of the QRS complex in the electrocardiogram, or detection of action potentials recorded with microelectrodes from single neurons.

12.6 Summary

The mean and autocorrelation function of the output of a filter with impulse response $h[n]$ can be expressed as a function of the mean and autocorrelation function of the input by means of the formulas:

$$< y[n] > = < x[n] > \sum_{n=-\infty}^{\infty} h[n]$$

$$R_{xy}[k] \;=\; h[k] \;*\; R_x[k]$$

$$R_y[k] \;=\; R_x[k] \;*\; (h[k] \;*\; h[-k])$$

These relations are particularly simple for white noise, for which samples at different times are uncorrelated.

These formulas can be further simplified by introducing the power spectrum $S_x(f)$, which is the Fourier transform of the autocorrelation function:

$$S_x(f) \;=\; \sum_{k=-\infty}^{\infty} R_x[k]\; e^{-j2\pi f k}$$

The power spectrum for a particular frequency f_0 represents the contribution of the frequency band centered at f_0 to the total power in the signal. The power spectrum of the output of a linear filter is equal to the power spectrum of the input multiplied by the magnitude square of the frequency response:

$$S_y(f) \;=\; |H(f)|^2 \; S_x(f)$$

This property is useful for analyzing the responses of linear systems to random signals, and for generating random signals with arbitrary spectral characteristics.

The cross spectrum $S_{xy}(f)$ of two random signals $x[n]$ and $y[n]$ is the Fourier transform of the crosscorrelation function $R_{xy}[k]$. If $y[n]$ is the response of a linear filter to $x[n]$, the cross spectrum is the product of the power spectrum of the input by the filter frequency response. Conversely, given two signals $x[n]$ and $y[n]$, the frequency response of the linear filter that best characterizes the relation between the two signals in a least-squares sense is:

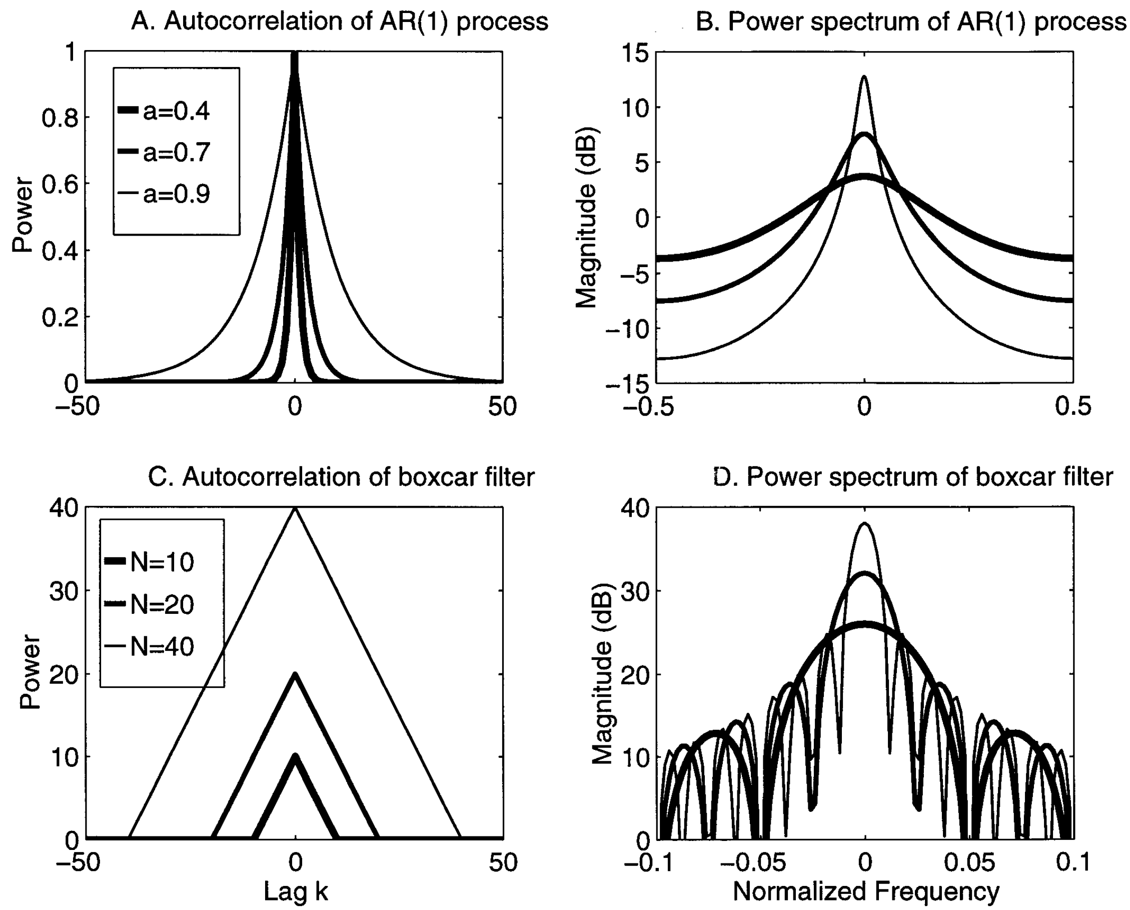
$$H(f) \;=\; \frac{S_{xy}(f)}{S_x(f)}$$

This Wiener filter has many applications to signal conditioning and system identification.

While the Wiener filter is useful for estimating a *random* signal in additive noise, the matched filter is used to detect a *known* signal in noise. The matched filter takes a particularly simple form for white noise, in which case its impulse response is the signal waveform reversed in time.

12.7 Further reading

Papoulis and Pillai: Chapters 9, 13



A. Autocorrelation of AR(1) process
B. Power spectrum of AR(1) process
C. Autocorrelation of boxcar filter
D. Power spectrum of boxcar filter
a=0.4
a=0.7
a=0.9
N=10
N=20
N=40
Power
Magnitude (dB)
Lag k
Normalized Frequency

Figure 12.1:

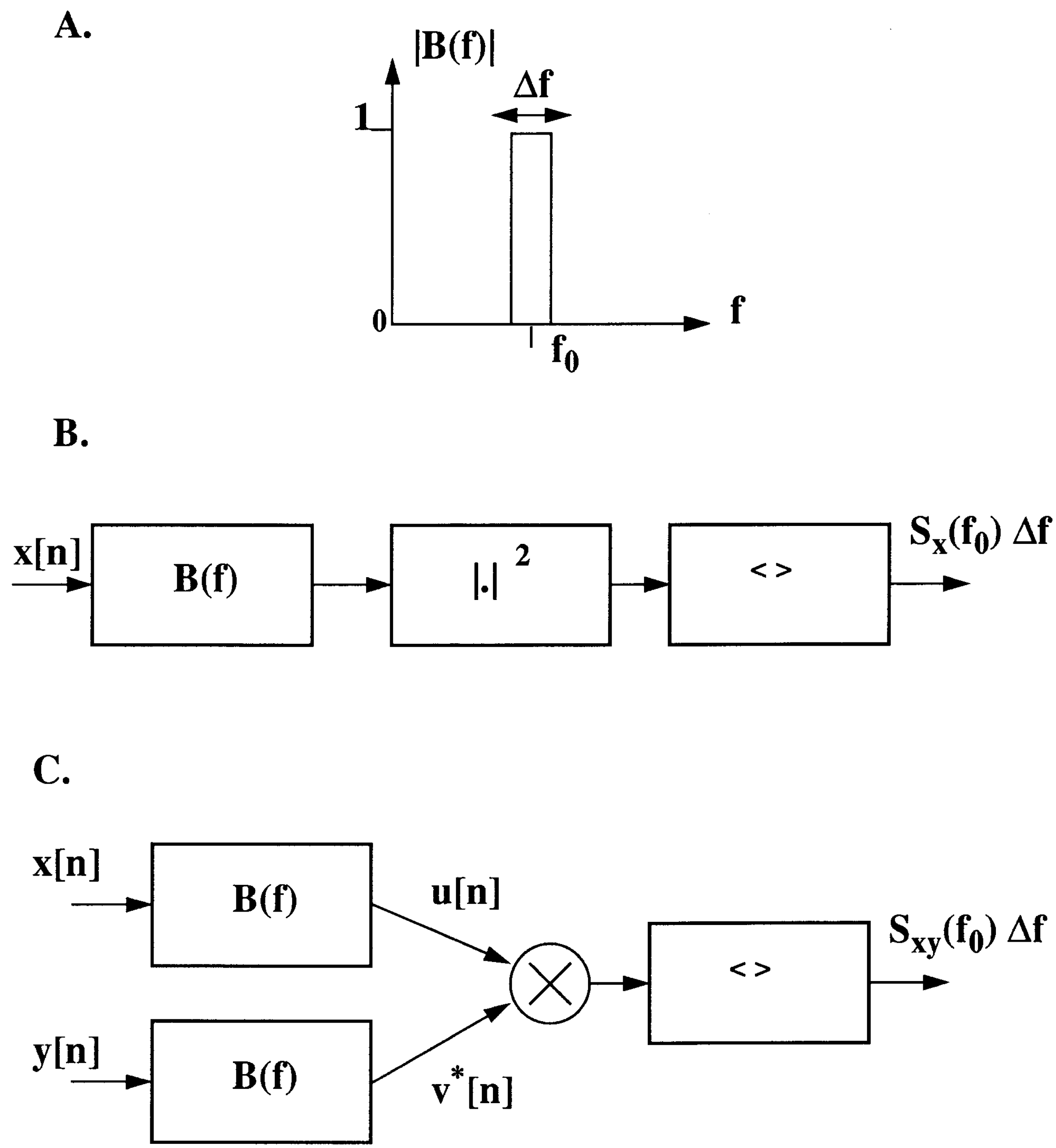


Figure 12.2: A. Magnitude of the frequency response of the ideal bandpass filter B(f). B. Physical interpretation of the power spectrum. C. Physical interpretation of the cross-spectrum.

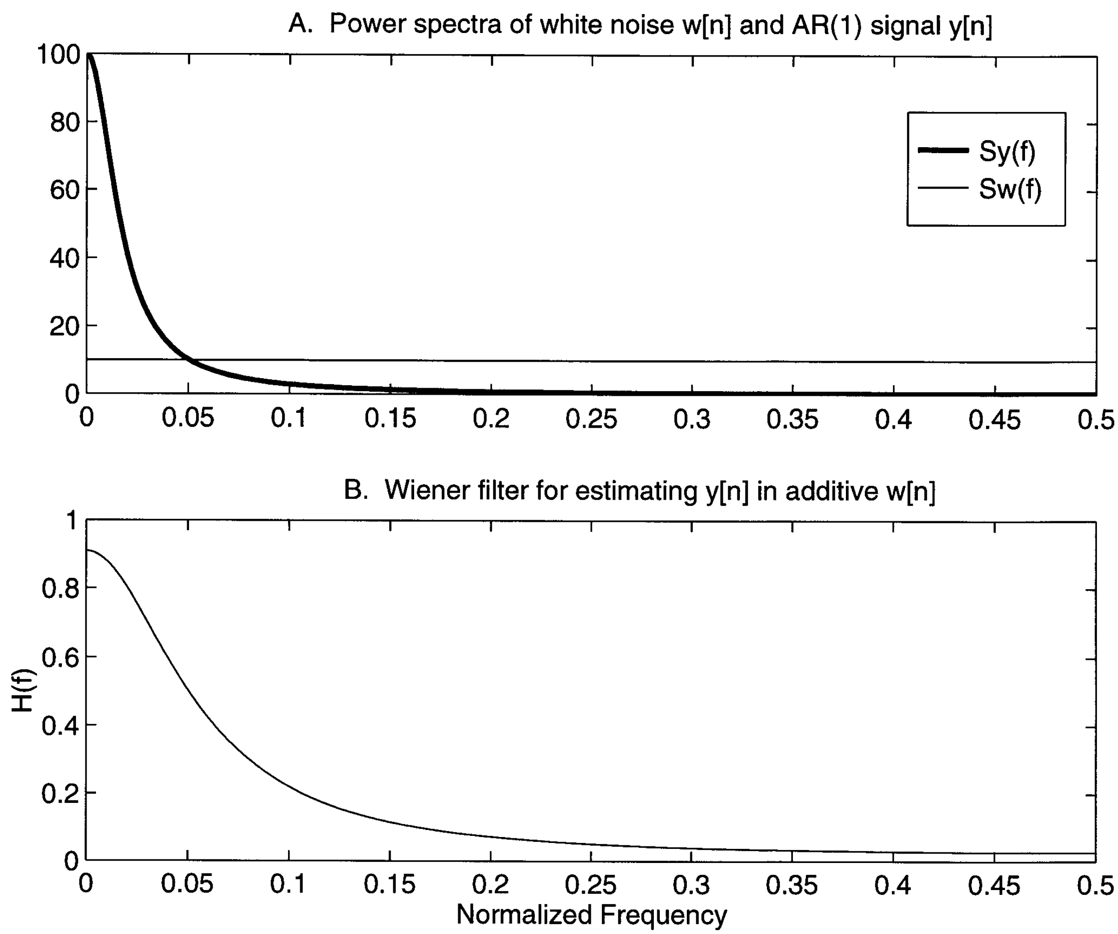

A. Power spectra of white noise w[n] and AR(1) signal y[n]
Sy(f)
Sw(f)
100
80
60
40
20
0
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
B. Wiener filter for estimating y[n] in additive w[n]
H(f)
1
0.8
0.6
0.4
0.2
Normalized Frequency

Figure 12.3:

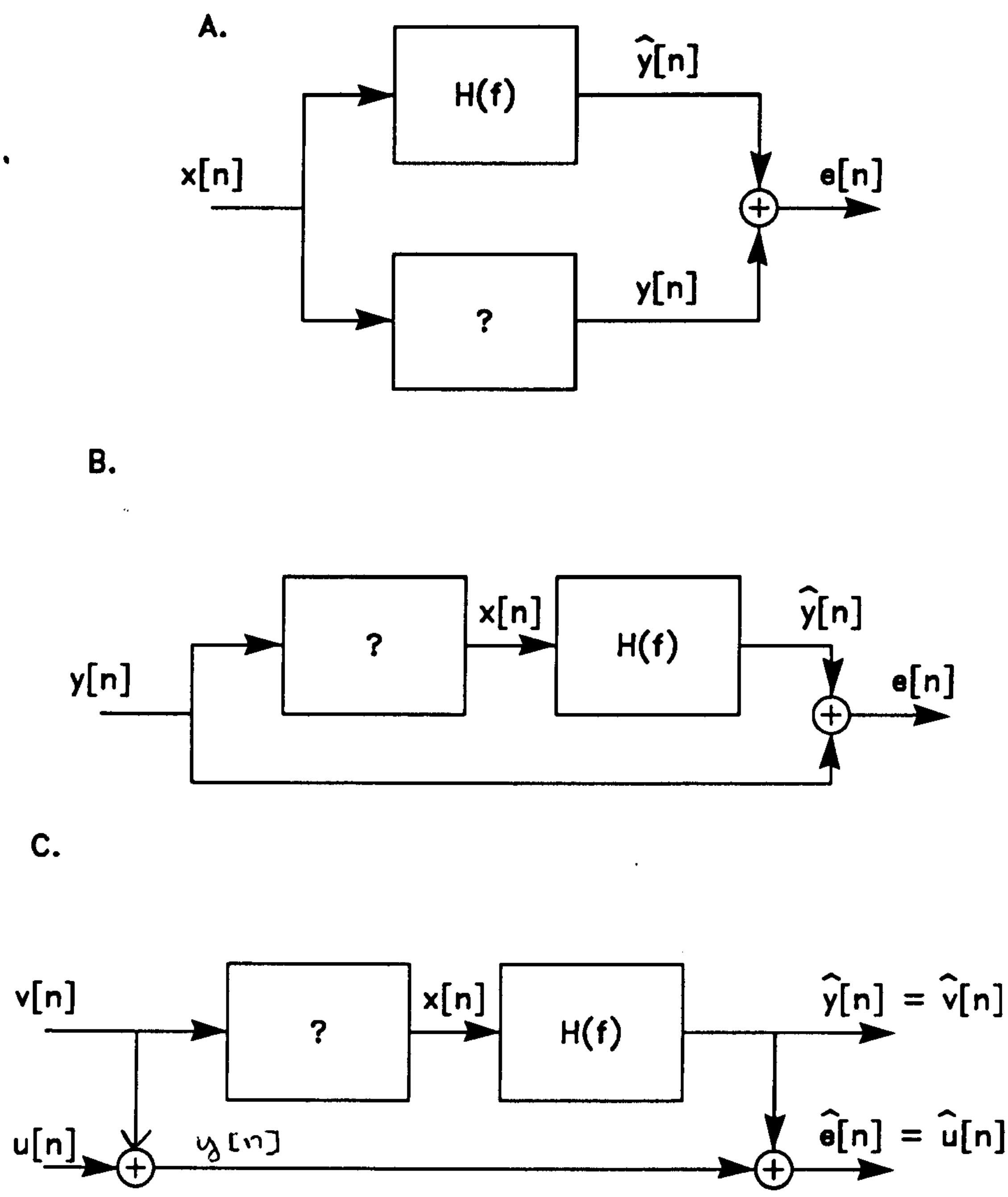


Figure 12.4: Applications of Wiener filters. A. Direct system identification. B. Inverse system identification. C. Noise cancellation and signal detection. In each case, the Wiener filter is indicated by H(f), and the unknown system by a question mark.

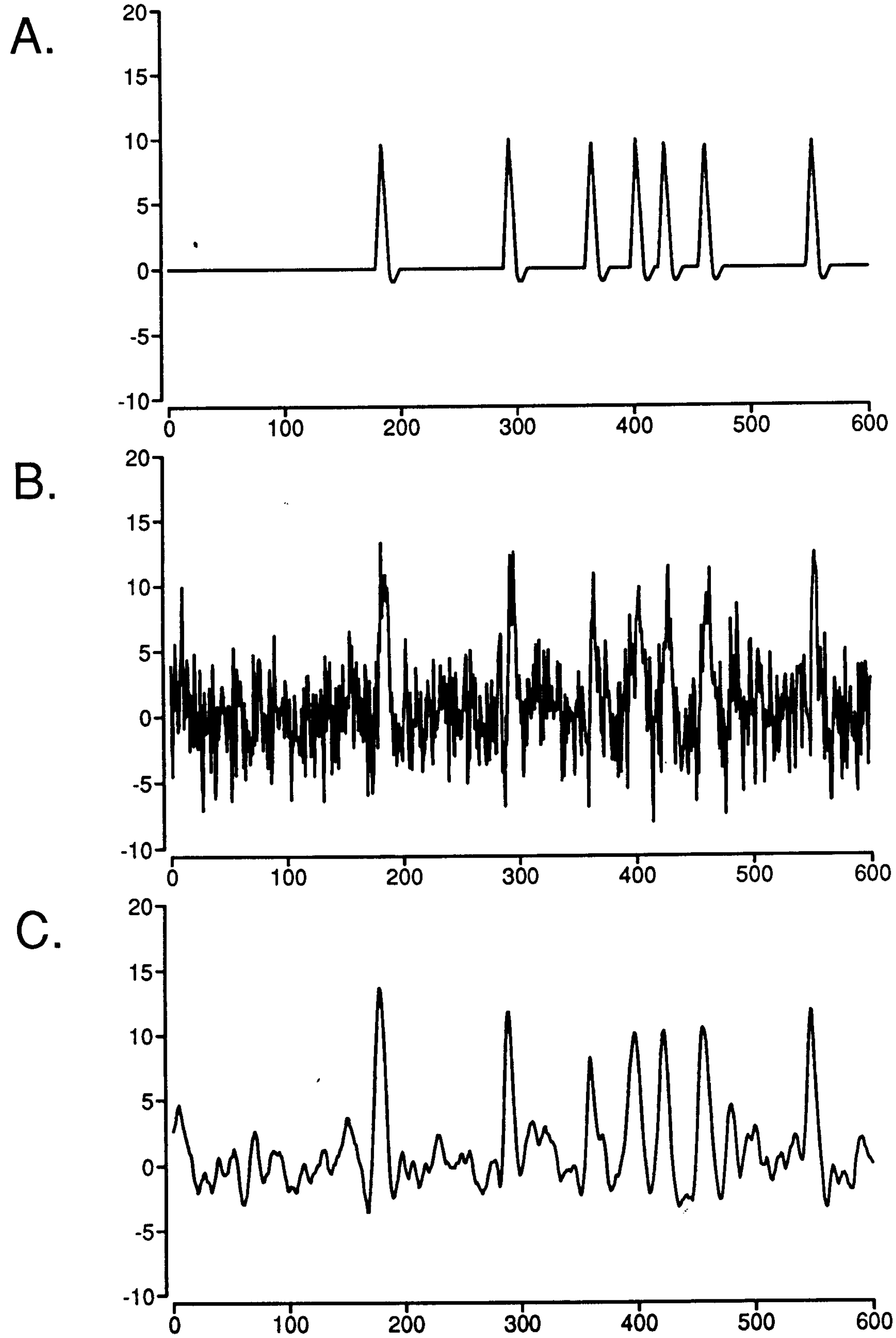


Figure 12.5: A. Nerve Impulses. B. The same impulses corrupted by white noise. C. Result of processing the corrupted impulses by a matched filter.