chapter

3

Signal Transmission and Filtering

CHAPTER OUTLINE

Signal transmission is the process whereby an electrical waveform gets from one location to another, ideally arriving without distortion. In contrast, signal filtering is an operation that purposefully distorts a waveform by altering its spectral content. Nonetheless, most transmission systems and filters have in common the properties of linearity and time invariance. These properties allow us to model both transmission and filtering in the time domain in terms of the impulse response, or in the frequency domain in terms of the frequency response.

This chapter begins with a general consideration of system response in both domains. Then we'll apply our results to the analysis of signal transmission and distortion for a variety of media and systems such as fiber optics and satellites. We'll examine the use of various types of filters and filtering in communication systems. Some related topics—notably transmission loss, Hilbert transforms, and correlation—are also included as starting points for subsequent development.

OBJECTIVES

After studying this chapter and working the exercises, you should be able to do each of the following:

1. State and apply the input–output relations for an LTI system in terms of its impulse response $h(t)$, step response $g(t)$, or transfer function $H(f)$ (Sect. 3.1).
2. Use frequency-domain analysis to obtain an exact or approximate expression for the output of a system (Sect. 3.1).
3. Find $H(f)$ from the block diagram of a simple system (Sect. 3.1).
4. Distinguish between amplitude distortion, delay distortion, linear distortion, and nonlinear distortion (Sect. 3.2).
5. Identify the frequency ranges that yield distortionless transmission for a given channel, and find the equalization needed for distortionless transmission over a specified range (Sect. 3.2).
6. Use dB calculations to find the signal power in a cable transmission system with amplifiers (Sect. 3.3).
7. Discuss the characteristics of and requirements for transmission over fiber optic and satellite systems (Sect. 3.3).
8. Identify the characteristics and sketch $H(f)$ and $h(t)$ for an ideal LPF, BPF, or HPF (Sect. 3.4).
9. Find the 3 dB bandwidth of a real LPF, given $H(f)$ (Sect. 3.4).
10. State and apply the bandwidth requirements for pulse transmission (Sect. 3.4).
11. State and apply the properties of the Hilbert transform (Sect. 3.5).
12. Define the crosscorrelation and auto-correlation functions for power or energy signals, and state their properties (Sect. 3.6).
13. State the Wiener-Kinchine theorem and the properties of spectral density functions (Sect. 3.6).
14. Given $H(f)$ and the input correlation or spectral density function, find the output correlation or spectral density (Sect. 3.6).

3.1 RESPONSE OF LTI SYSTEMS

Figure 3.1–1 depicts a system inside a "black box" with an external **input signal $x(t)$** and an **output signal $y(t)$.** In the context of electrical communication, the system usually would be a two-port network driven by an applied voltage or current at the input port, producing another voltage or current at the output port. Energy storage elements and other internal effects may cause the output waveform to look quite different from the input. But regardless of what's in the box, the system is characterized by an **excitation-and-response** relationship between input and output.

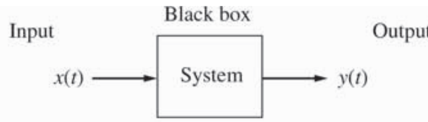

Figure 3.1–1 System showing external input and output.

Here we're concerned with the special but important class of **linear time-invariant** systems—or LTI systems for short. We'll develop the input–output relationship in the time domain using the superposition integral and the system's impulse response. Then we'll turn to frequency-domain analysis expressed in terms of the system's transfer function.

Impulse Response and the Superposition Integral

Let Fig. 3.1–1 be an LTI system having no internal stored energy at the time the input $x(t)$ is applied. The output $y(t)$ is then the **forced response** due entirely to $x(t)$, as represented by

$$y(t) = F[x(t)] \tag{1}$$

where $F[x(t)]$ stands for the functional relationship between input and output. The **linear** property means that Eq. (1) obeys the **principle of superposition.** Thus, if

$$x(t) = \sum_k a_k x_k(t) \tag{2a}$$

where a_k are constants, then

$$y(t) = \sum_k a_k F[x_k(t)] \tag{2b}$$

The **time-invariance** property means that the system's characteristics remain **fixed with time.** Thus, a time-shifted input $x(t - t_d)$ produces

$$F[x(t - t_d)] = y(t - t_d) \tag{3}$$

so the output is time-shifted but otherwise unchanged.

Most LTI systems consist entirely of **lumped-parameter** elements (such as resistors capacitors, and inductors), as distinguished from elements with spatially distributed phenomena (such as transmission lines). Direct analysis of a lumped-parameter system starting with the element equations leads to the input–output relation as a linear differential equation in the form

$$a_n \frac{d^n y(t)}{dt^n} + \cdots + a_1 \frac{dy(t)}{dt} + a_0 y(t) = b_m \frac{d^m x(t)}{dt^m} + \cdots + b_1 \frac{dx(t)}{dt} + b_0 x(t) \tag{4}$$

where the a's and b's are constant coefficients involving the element values. The number of independent energy-storage elements determines the value of n, known as the *order* of the system. Unfortunately, Eq. (4) doesn't provide us with a direct expression for $y(t)$.

To obtain an explicit input–output equation, we must first define the system's **impulse response** function

$$h(t) \triangleq F[\delta(t)] \tag{5}$$

which equals the forced response when $x(t) = \delta(t)$. Now any continuous input signal can be written as the convolution $x(t) = x(t) * \delta(t)$, so

$$y(t) = F\left[\int_{-\infty}^{\infty} x(\lambda)\delta(t-\lambda)\,d\lambda\right]$$
$$= \int_{-\infty}^{\infty} x(\lambda)F[\delta(t-\lambda)]\,d\lambda$$

in which the interchange of operations is allowed by virtue of the system's linearity. Now, from the time-invariance property, $F[\delta(t-\lambda)] = h(t-\lambda)$ and hence

$$y(t) = \int_{-\infty}^{\infty} x(\lambda)h(t-\lambda)\,d\lambda \tag{6a}$$

$$= \int_{-\infty}^{\infty} h(\lambda)x(t-\lambda)\,d\lambda \tag{6b}$$

where we have drawn upon the commutativity of convolution.

Either form of Eq. (6) is called the **superposition integral.** It expresses the forced response as a convolution of the input $x(t)$ with the impulse response $h(t)$. System analysis in the time domain therefore requires knowledge of the impulse response along with the ability to carry out the convolution.

Various techniques exist for determining $h(t)$ from a differential equation or some other system model. However, you may be more comfortable taking $x(t) = u(t)$ and calculating the system's **step response**

$$g(t) \triangleq F[u(t)] \tag{7a}$$

from which

$$h(t) = \frac{dg(t)}{dt} \tag{7b}$$

This derivative relation between the impulse and step response follows from the general convolution property

$$\frac{d}{dt}[v(t) * w(t)] = v(t) * \left[\frac{dw(t)}{dt}\right]$$

Thus, since $g(t) = h(t) * u(t)$ by definition, $dg(t)/dt = h(t) * [du(t)/dt] = h(t) * \delta(t) = h(t)$.

Time Response of a First-Order System

EXAMPLE 3.1–1

The simple RC circuit in Fig. 3.1–2 has been arranged as a two-port network with input voltage $x(t)$ and output voltage $y(t)$. The reference voltage polarities are indicated by the $+/-$ notation where the assumed higher potential is indicated by the + sign. This circuit is a first-order system governed by the differential equation

$$RC\frac{dy(t)}{dt} + y(t) = x(t)$$

Similar expressions describe certain transmission lines and cables, so we're particularly interested in the system response.

From either the differential equation or the circuit diagram, the step response is readily found to be

$$g(t) = (1 - e^{-t/RC})u(t) \tag{8a}$$

Interpreted physically, the capacitor starts at zero initial voltage and charges toward $y(\infty) = 1$ with time constant RC when $x(t) = u(t)$. Figure 3.1–3*a* plots this behavior, while Fig. 3.1–3*b* shows the corresponding impulse response

$$h(t) = \frac{1}{RC}e^{-t/RC}u(t) \tag{8b}$$

obtained by differentiating $g(t)$. Note that $g(t)$ and $h(t)$ are *causal* waveforms since the input equals zero for $t < 0$.

The response to an arbitrary input $x(t)$ can now be found by putting Eq. (8*b*) in the superposition integral. For instance, take the case of a rectangular pulse applied at $t = 0$, so $x(t) = A$ for $0 < t < \tau$. The convolution $y(t) = h(t) * x(t)$ divides into three parts, like the example back in Fig. 2.4–1 with the result that

$$y(t) = \begin{cases} 0 & t < 0 \\ A(1 - e^{-t/RC}) & 0 < t < \tau \\ A(1 - e^{-\tau/RC})e^{-(t-\tau)/RC} & t > \tau \end{cases} \tag{9}$$

as sketched in Fig. 3.1–4 for three values of τ/RC.

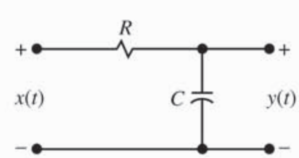

Figure 3.1–2 RC lowpass filter.

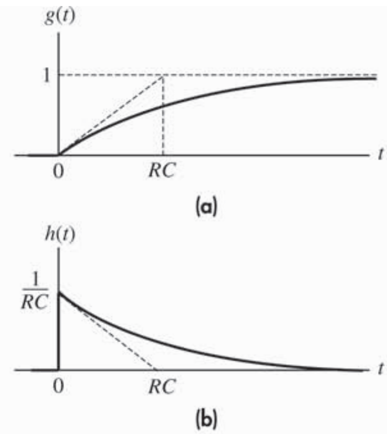

Figure 3.1–3 Output of an RC lowpass filter: (*a*) step response; (*b*) impulse response.

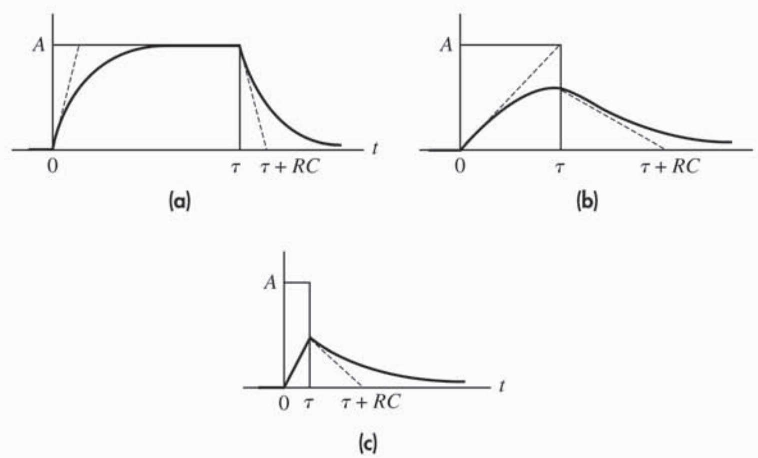

Figure 3.1–4 Rectangular pulse response of an RC lowpass filter: (*a*) $\tau \gg RC$; (*b*) $\tau \approx RC$; (*c*) $\tau \ll RC$.

EXERCISE 3.1–1 Let the resistor and the capacitor be interchanged in Fig. 3.1–2. Find the step and impulse response.

Transfer Functions and Frequency Response

Time-domain analysis becomes increasingly difficult for higher-order systems, and the mathematical complications tend to obscure significant points. We'll gain a different and often clearer view of system response by going to the frequency domain. As a first step in this direction, we define the system **transfer function** to be the Fourier transform of the impulse response, namely,

$$H(f) \triangleq \mathcal{F}[h(t)] = \int_{-\infty}^{\infty} h(t)e^{-j2\pi ft}\,dt \tag{10}$$

This definition requires that $H(f)$ exists, at least in a limiting sense. In the case of an **unstable** system, $h(t)$ **grows with time** and $H(f)$ does not exist.

When $h(t)$ is a *real* time function, $H(f)$ has the hermitian symmetry

$$H(-f) = H^*(f) \tag{11a}$$

so that

$$|H(-f)| = |H(f)| \qquad \arg H(-f) = -\arg H(f) \tag{11b}$$

We'll assume this property holds unless otherwise stated.

The frequency-domain interpretation of the transfer function comes from $y(t) = h * x(t)$ with a *phasor* input, say

$$x(t) = A_x e^{j\phi_x} e^{j2\pi f_0 t} \qquad -\infty < t < \infty \tag{12a}$$

The stipulation that $x(t)$ persists for all time means that we're dealing with *steady-state* conditions, like the familiar case of ac steady-state circuit analysis. The steady-state forced response is

$$\begin{aligned} y(t) &= \int_{-\infty}^{\infty} h(\lambda)A_x e^{j\phi_x} e^{j2\pi f_0(t-\lambda)}\,d\lambda \\ &= \left[\int_{-\infty}^{\infty} h(\lambda)e^{-j2\pi f_0\lambda}d\lambda\right] A_x e^{j\phi_x} e^{j2\pi f_0 t} \\ &= H(f_0)A_x e^{j\phi_x} e^{j2\pi f_0 t} \end{aligned}$$

where, from Eq. (10), $H(f_0)$ equals $H(f)$ with $f = f_0$. Converting $H(f_0)$ to polar form then yields

$$y(t) = A_y e^{j\phi_y} e^{j2\pi f_0 t} \qquad -\infty < t < \infty \tag{12b}$$

in which we have identified the output phasor's amplitude and angle

$$A_y = |H(f_0)|A_x \qquad \phi_y = \arg H(f_0) + \phi_x \tag{13}$$

Using conjugate phasors and superposition, you can similarly show that if

$$x(t) = A_x \cos(2\pi f_0 t + \phi_x)$$

then

$$y(t) = A_y \cos(2\pi f_0 t + \phi_y)$$

with A_y and ϕ_y as in Eq. (13).

Since $A_y/A_x = |H(f_0)|$ at any frequency f_0, we conclude that $|H(f)|$ represents the system's **amplitude ratio** as a function of frequency (sometimes called the **amplitude response** or **gain**). By the same token, arg $H(f)$ represents the **phase shift,** since $\phi_y - \phi_x = \arg H(f_0)$. Plots of $|H(f)|$ and arg $H(f)$ versus frequency give us the frequency-domain representation of the system or, equivalently, the system's **frequency response.** Henceforth, we'll refer to $H(f)$ as either the transfer function or frequency-response function.

Now let $x(t)$ be any signal with spectrum $X(f)$. Calling upon the convolution theorem, we take the transform of $y(t) = h(t) * x(t)$ to obtain

$$Y(f) = H(f)X(f) \tag{14}$$

This elegantly simple result constitutes the basis of frequency-domain system analysis. It says that

> The output spectrum *Y(f)* equals the input spectrum *X(f)* multiplied by the transfer function *H(f)*.

The corresponding amplitude and phase spectra are

$$|Y(f)| = |H(f)||X(f)|$$
$$\arg Y(f) = \arg H(f) + \arg X(f)$$

which compare with the single-frequency expressions in Eq. (13). If $x(t)$ is an energy signal, then $y(t)$ will be an energy signal whose spectral density and total energy are given by

$$|Y(f)|^2 = |H(f)|^2|X(f)|^2 \tag{15a}$$

$$E_y = \int_{-\infty}^{\infty} |H(f)|^2|X(f)|^2\,df \tag{15b}$$

as follows from Rayleigh's energy theorem.

Equation (14) sheds new light on the meaning of the system transfer function and the transform pair $h(t) \leftrightarrow H(f)$. For if we let $x(t)$ be a unit impulse, then $X(f) = 1$ and $Y(f) = H(f)$ — in agreement with the definition $y(t) = h(t)$ when $x(t) = \delta(t)$. From the frequency-domain viewpoint, the "flat" input spectrum $X(f) = 1$ contains all frequencies in equal proportion and, consequently, the output spectrum takes on the shape of the transfer function $H(f)$.

Figure 3.1–5 summarizes our input–output relations in both domains. Clearly, when $H(f)$ and $X(f)$ are given, the output spectrum $Y(f)$ is much easier to find than the output signal $y(t)$. In principle, we could compute $y(t)$ from the inverse transform

$$y(t) = \mathscr{F}^{-1}[H(f)X(f)] = \int_{-\infty}^{\infty} H(f)X(f)e^{j2\pi ft}\,df$$

But this integration does not necessarily offer any advantages over time-domain convolution. Indeed, the power of frequency-domain system analysis largely depends on

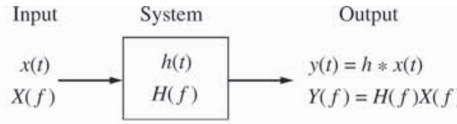

Figure 3.1–5 Input–output relations for an LTI system.

staying in that domain and using our knowledge of spectral properties to draw inferences about the output signal.

Finally, we point out two ways of determining $H(f)$ that don't involve $h(t)$. If you know the *differential equation* for a lumped-parameter system, you can immediately write down its transfer function as the ratio of polynomials

$$H(f) = \frac{b_m(j2\pi f)^m + \cdots + b_1(j2\pi f) + b_0}{a_n(j2\pi f)^n + \cdots + a_1(j2\pi f) + a_0} \tag{16}$$

whose coefficients are the same as those in Eq. (4). Equation (16) follows from Fourier transformation of Eq. (4).

Alternatively, if you can calculate a system's **steady-state phasor response,** Eqs. (12) and (13) show that

$$H(f) = \frac{y(t)}{x(t)} \qquad \text{when} \qquad x(t) = e^{j2\pi ft} \tag{17}$$

This method corresponds to impedance analysis of electrical circuits, but is equally valid for any LTI system. Furthermore, Eq. (17) may be viewed as a special case of the s domain transfer function $H(s)$ used in conjunction with Laplace transforms. Since $s = \sigma + j\omega$ in general, $H(f)$ is obtained from $H(s)$ simply by letting $s = j2\pi f$. These methods assume, of course, that the system is stable.

Frequency Response of a First-Order System

EXAMPLE 3.1–2

The RC circuit from Example 3.1–1 has been redrawn in Fig. 3.1–6a with the impedances $Z_R = R$ and $Z_C = 1/j\omega C$ replacing the elements. Since $y(t)/x(t) = Z_C/(Z_C + Z_R)$ when $x(t) = e^{j\omega t}$, Eq. (17) gives

$$H(f) = \frac{(1/j2\pi fC)}{(1/j2\pi fC) + R} = \frac{1}{1 + j2\pi fRC}$$

$$= \frac{1}{1 + j(f/B)} \tag{18a}$$

where we have introduced the system parameter

$$B \triangleq \frac{1}{2\pi RC} \tag{18b}$$

Identical results would have been obtained from Eq. (16), or from $H(f) = \mathscr{F}[h(t)]$. (In fact, the system's impulse response has the same form as the causal exponential pulse discussed in Example 2.2–2.) The amplitude ratio and phase shift are

$$|H(f)| = \frac{1}{\sqrt{1 + (f/B)^2}} \qquad \arg H(f) = -\arctan\frac{f}{B} \tag{18c}$$

as plotted in Fig. 3.1–6*b* for $f \geq 0$. The hermitian symmetry allows us to omit $f < 0$ without loss of information.

The amplitude ratio $|H(f)|$ has special significance relative to any *frequency-selective* properties of the system. We call this particular system a **lowpass filter** because it has almost no effect on the amplitude of low-frequency components, say $|f| \ll B$, while it drastically reduces the amplitude of high-frequency components, say $|f| \gg B$. The parameter B serves as a measure of the filter's **passband** or **bandwidth.**

To illustrate how far you can go with frequency-domain analysis, let the input $x(t)$ be an arbitrary signal whose spectrum has negligible content for $|f| > W$. There are three possible cases to consider, depending on the relative values of B and W:

1. If $W \ll B$, as shown in Fig. 3.1–7*a*, then $|H(f)| \approx 1$ and $\arg H(f) \approx 0$ over the signal's frequency range $|f| < W$. Thus, $Y(f) = H(f)X(f) \approx X(f)$ and $y(t) \approx x(t)$ so we have **undistorted transmission** through the filter.
2. If $W \approx B$, as shown in Fig. 3.1–7*b*, then $Y(f)$ depends on both $H(f)$ and $X(f)$. We can say that the output is **distorted**, since $y(t)$ will differ significantly from $x(t)$, but time-domain calculations would be required to find the actual waveform.

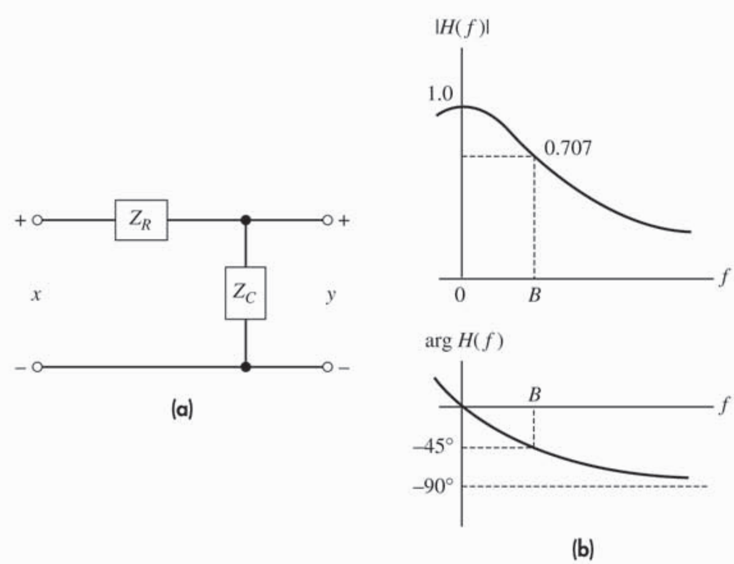

Figure 3.1–6 RC lowpass filter. (*a*) circuit; (*b*) transfer function.