:h expresses in quantitative form the bandwidth and noise limitations on infor-ion transmission. In the light of this theory it is possible to draw some conclusion ut the relative merits of conventional systems and also get some hints as to how er systems can be designed.

The text concludes, in Chap. 10, with digital data transmission, the most dly expanding area in communication engineering. This topic not only involves he previous material but also gives added meaning to the mathematical theory of munication.

Each chapter contains several exercises designed to clarify and reinforce the :epts and analytic techniques as they are introduced. Students are strongly ouraged to test their grasp of the material by working these exercises. Answers e been provided where appropriate.

Certain optional or more advanced topics are interspersed through the text. ntified by the symbol ★, these sections can be omitted without serious loss of tinuity. Other optional material of a supplementary nature has been collected in three appendices at the back of the book.

Also at the back the reader will find several tables and a list of selected supple-ntary reading. The former contain most of the mathematical relations and numer- data needed to work the problems at the end of the chapters. The latter serves n annotated bibliography of books and papers for the benefit of those who wish ursue a topic in greater depth.

2

SIGNALS, SPECTRA, AND FILTERS

Electrical communication signals are time-varying quantities, such as voltage or current. The usual description of a signal $v(t)$ is in the *time domain*, where the independent variable is t. But for communications work, it is often more convenient to describe signals in the *frequency domain*, where the independent variable is f. Roughly speaking, we think of the time function as being composed of a number of frequency components, each with appropriate amplitude and phase. Thus, while the signal physically exists in the time domain, we can say that it consists of those components in its frequency-domain description, called the *spectrum*.

Spectral analysis, based on the Fourier series and transform, is a powerful tool in communication engineering. Accordingly, we will concentrate primarily on Fourier theory rather than on other techniques such as Laplace transforms and time-domain analysis. There are several reasons for this emphasis.

First, the frequency domain is essentially a steady-state viewpoint; and for many purposes it is reasonable to restrict attention to the steady-state behavior of a communication system. Indeed, considering the multitude of possible signals that a system may handle, detailed transient solutions for each would be an impossible task. Second, the spectral approach allows us to treat entire classes of signals that have similar properties in the frequency domain. This not only gives insight to analysis but is invaluable for design. It is quite unlikely, for example, that such a significant technique as single-sideband modulation could have been developed without the aid of spectral concepts. Third, many components of a communication system can be classified as linear and time-invariant devices; when this is so, we can describe them by their *frequency-response characteristics* which, in turn, further expedites analysis and design work.

This chapter therefore is devoted to a review† and elaboration of Fourier analysis of signals and frequency-response characteristics of system components, particularly those frequency-selective components known as filters. However, the spectral approach should not be thought of as the only method used by communication engineers. There are some problems where it cannot be applied directly, and some problems where other techniques are more convenient. Thus, each new problem must be approached with an open mind and a good set of analytic tools.

As the first step in much of our work we will write equations representing signals or components. But one must bear in mind that such equations are only mathematical *models* of physical entities, usually imperfect models. In fact, a completely faithful description of the simplest signal or component would be prohibitively complex in mathematical form and consequently useless for engineering purposes. Hence the models we seek are ones that represent, with minimum complexity, those properties that are pertinent to the problem at hand. This sometimes leads to constructing several different models for the same thing, according to need. Then, given a particular problem, the choice of which model to use is based on understanding the physical phenomena involved and the limitations of the mathematics; in short, it is engineering.

2.1 AC SIGNALS AND NETWORKS

Figure 2.1 represents a two-port electrical network or system being driven by an input signal $x(t)$ which produces the output signal $y(t)$. We say that $x(t)$ is a sinusoidal or AC signal if

$$x(t) = A_x \cos(\omega_0 t + \theta_x) \qquad -\infty < t < \infty \tag{1}$$

where A_x is the *amplitude* (in volts or amperes), ω_0 is the *angular frequency* (in radians per second), and θ_x is the *phase* (in radians or degrees). It is also convenient to introduce the *cyclical frequency* $f_0 = \omega_0/2\pi$ (in cycles per second or hertz) so the *period* (in seconds) is $T_0 = 1/f_0 = 2\pi/\omega_0$. The significance of T_0 is that $x(t)$ repeats itself every T_0 seconds for all time. Obviously, no real signal lasts forever, but Eq. (1) is a convenient and useful representation or model for a sinusoidal signal of finite duration if the duration is much longer than the period.

† It is assumed that the reader has some background in transform theory and linear systems analysis. Accordingly, certain proofs and derivations have been omitted.

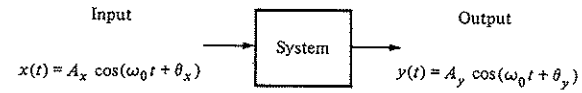


FIGURE 2.1
System in AC steady-state condition.

Now assume that the network is *linear*, *time-invariant*, and *asymptotically stable*. In essence, this means that superposition applies, there are no time-varying parameters, and the natural behavior decays with time. Under these conditions, the output signal will also be sinusoidal at the same frequency as the input, differing only in amplitude and phase, i.e.,

$$y(t) = A_y \cos(\omega_0 t + \theta_y) \qquad -\infty < t < \infty \tag{2}$$

Therefore, given the parameters of the input signal and the network's characteristics, we need only solve for A_y and θ_y to completely describe the resulting output signal. This is, of course, the familiar AC steady-state problem.

It is well known that such problems are most easily solved using exponential time functions of the form $e^{j\omega t}$ rather than sinusoidal functions. Take the case, for instance, of a circuit having complex impedance $Z(j\omega)$; if the current through the circuit is $e^{j\omega t}$ then, by the definition of impedance, the resulting voltage is simply $Z(j\omega)e^{j\omega t}$. Similarly, for the two-port network case of Fig. 2.1, we define the network's *transfer function* $H(j\omega)$ such that $y(t) = H(j\omega)e^{j\omega t}$ when $x(t) = e^{j\omega t}$; that is,

$$H(j\omega) \triangleq \frac{y(t)}{x(t)} \qquad \text{when } x(t) = e^{j\omega t} \tag{3}$$

a definition† we will generalize subsequently. Combining Eq. (3) with the superposition principle, it follows that if $x(t)$ is a linear combination of exponentials, say

$$x(t) = \alpha_1 e^{j\omega_1 t} + \alpha_2 e^{j\omega_2 t} + \cdots \tag{4a}$$

then

$$y(t) = H(j\omega_1)\alpha_1 e^{j\omega_1 t} + H(j\omega_2)\alpha_2 e^{j\omega_2 t} + \cdots \tag{4b}$$

where α_1 and α_2 are constants, $H(j\omega_1)$ represents $H(j\omega)$ evaluated at $\omega = \omega_1$, etc.

However, we have not yet solved the AC problem stated at the outset, which now entails converting sinusoids to exponentials. At this point most circuit theory students would probably write $x(t) = \text{Re}\,[A_x e^{j\theta_x} e^{j\omega_0 t}]$ but, preparing the way for future developments, we want an expression more like Eq. (4*a*). For that purpose we invoke a corollary of Euler's theorem,‡ to wit:

$$\cos\phi = \tfrac{1}{2}(e^{j\phi} + e^{-j\phi}) \tag{5}$$

† The symbol $\triangleq$ stands for "equals by definition."
‡ Euler's theorem is the most versatile of the trigonometric identities. A short table of these useful relations is given in Table B at the end of the book.

Thus, the sinusoidal input $x(t)$ in Eq. (1) can be rewritten as

$$x(t) = A_x \tfrac{1}{2}[e^{j(\omega_0 t + \theta_x)} + e^{-j(\omega_0 t + \theta_x)}]$$
$$= \frac{A_x}{2} e^{j\theta_x} e^{j\omega_0 t} + \frac{A_x}{2} e^{-j\theta_x} e^{-j\omega_0 t} \qquad (6a)$$

and since this has the same form as Eq. (4*a*), appropriate substitution in Eq. (4*b*) gives

$$y(t) = H(j\omega_0) \frac{A_x}{2} e^{j\theta_x} e^{j\omega_0 t} + H(-j\omega_0) \frac{A_x}{2} e^{-j\theta_x} e^{-j\omega_0 t} \qquad (6b)$$

While this is a correct result, it can be tidied up and made more understandable via the following two steps.

First, since the transfer function is in general a complex quantity, we will express it in the polar form

$$H(j\omega) = |H(j\omega)| e^{j \arg[H(j\omega)]} \qquad (7)$$

where $|H(j\omega)|$ is the *magnitude* and $\arg[H(j\omega)]$ is the *angle*.† Second, despite the fact that $H(j\omega)$ is complex, $y(t)$ ought to be a real function of time since $x(t)$ is real; and this will be true if and only if

$$H(-j\omega) = H^*(j\omega)$$
$$= |H(j\omega)| e^{-j \arg[H(j\omega)]} \qquad (8)$$

where $H^*(j\omega)$ is the complex conjugate of $H(j\omega)$. The complex-conjugate relationship does, in fact, hold for any real network.

Applying Eqs. (7) and (8) to Eq. (6*b*) and simplifying yields

$$y(t) = |H(j\omega_0)| \frac{A_x}{2} [e^{j(\omega_0 t + \theta_x + \arg[H(j\omega_0)])} + e^{-j(\omega_0 t + \theta_x + \arg[H(j\omega_0)])}]$$
$$= \underbrace{|H(j\omega_0)| A_x}_{A_y} \cos(\omega_0 t + \underbrace{\theta_x + \arg[H(j\omega_0)]}_{\theta_y}) \qquad (9)$$

in which we have identified

$$A_y = |H(j\omega_0)| A_x \qquad \theta_y = \theta_x + \arg[H(j\omega_0)] \qquad (10)$$

† That is, if H_r and H_i are the real and imaginary parts of $H(j\omega)$, respectively, then

$$|H(j\omega)| = \sqrt{H_r^2 + H_i^2} \qquad \arg[H(j\omega)] = \arctan \frac{H_i}{H_r}$$

The reader who has difficulties with such manipulations is advised to brush up on the subject of complex numbers.

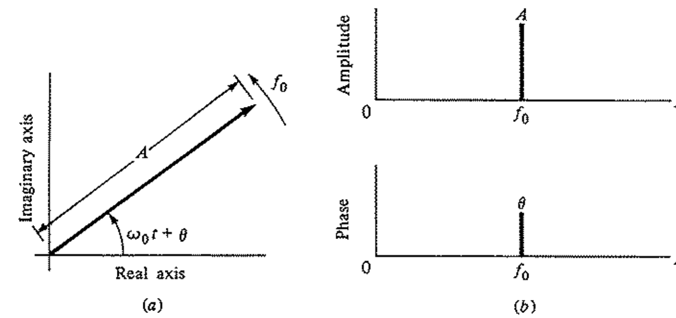


FIGURE 2.2
Representations of $A \cos(\omega_0 t + \theta)$. (*a*) Phasor diagram; (*b*) line spectrum.

Thus, we have these very simple equations for the output amplitude and phase of the AC steady-state response, providing we know the network's transfer function. More will be said about the latter after we have looked at the frequency-domain interpretation.

EXERCISE 2.1 Write $H(j\omega_0)$ and $H(-j\omega_0)$ in polar form and show that the imaginary part of Eq. (6*b*) is not zero when $H(-j\omega_0) \neq H^*(j\omega_0)$.

Phasors and Line Spectra

Besides facilitating network analysis, converting sinusoids to exponentials also underlies the notion of the frequency domain by way of phasor diagrams. To introduce this idea, consider the arbitrary sinusoid

$$v(t) = A \cos(\omega_0 t + \theta) \qquad \omega_0 = 2\pi f_0$$

which can be written as

$$A \cos(\omega_0 t + \theta) = \text{Re}[Ae^{j(\omega_0 t + \theta)}]$$
$$= \text{Re}[Ae^{j\theta} e^{j\omega_0 t}] \qquad (11)$$

This is called a *phasor representation* because the term inside the brackets may be viewed as a rotating vector in a complex plane whose axes are the real and imaginary parts, as Fig. 2.2*a* illustrates. The phasor has length A, rotates counterclockwise at a rate f_0 revolutions per second, and at time $t = 0$ makes an angle θ with respect to the positive real axis. At any time t the projection of the phasor on the real axis—i.e., its real part—equals the sinusoid $v(t)$.

Note carefully that only three parameters are needed to specify a phasor: amplitude, relative phase, and rotational frequency. To describe the same phasor in the *frequency domain*, we see that it is defined only for the particular frequency f_0. With this frequency we must associate the corresponding amplitude and phase. Hence, a suitable frequency-domain description would be the *line spectrum* of Fig. 2.2*b*, which consists of two plots, amplitude versus frequency and phase versus frequency. While Fig. 2.2*b* appears simple to the point of being trivial, it does have great conceptual value, especially when applied to more complicated signals. But before taking that step, four standard conventions used in constructing line spectra should be stated.

1 In all our spectral drawings the independent variable will be *cyclical frequency* f in hertz, rather than radian frequency ω, and any specific frequency such as f_0, being a constant, will be identified by a subscript. We will, however, use ω with or without subscripts as a shorthand notation for $2\pi f$ in various equations since that combination occurs so often.

2 Phase angles will be measured with respect to *cosine* waves or, equivalently, with respect to the positive real axis of the phasor diagram. Hence, sine waves need to be converted to cosines via the identity

$$\sin \omega t = \cos (\omega t - 90°) \tag{12}$$

3 We regard amplitude as always being a *positive* quantity; when negative signs appear, they must be absorbed in the phase, e.g.,

$$-A \cos \omega t = A \cos (\omega t \pm 180°) \tag{13}$$

and it does not matter whether one takes $+180°$ or $-180°$ since the phasor ends up in the same place either way.

4 Phase angles usually are expressed in degrees even though other angles are inherently in radians—for instance, ωt in Eqs. (12) and (13) is a radian measure while $-90°$ and $\pm 180°$ clearly are in degrees. No confusion should result from this mixed notation since angles expressed in degrees will always carry the appropriate symbol.

Illustrating these conventions and carrying the idea of line spectrum further, suppose a signal consists of a sum of sinusoids, such as

$$w(t) = 2 + 6 \cos (2\pi 10t + 30°) + 3 \sin 2\pi 30t - 4 \cos 2\pi 35t$$

Converting the constant (DC) term to a zero-frequency sinusoid and applying Eqs. (12) and (13) gives

$$w(t) = 2 \cos 2\pi 0t + 6 \cos (2\pi 10t + 30°) + 3 \cos (2\pi 30t - 90°) + 4 \cos (2\pi 35t - 180°)$$

so the spectrum, Fig. 2.3, has amplitude and phase lines at 0, 10, 30, and 35 Hz.

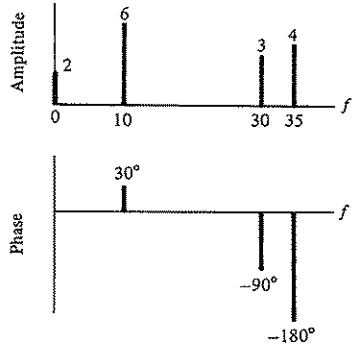


FIGURE 2.3
The line spectrum of $2 + 6 \cos (2\pi 10t + 30°) + 3 \sin 2\pi 30t - 4 \cos 2\pi 35t$.

Figures 2.2*b* and 2.3, called *one-sided* or *positive-frequency line spectra*, can be generated for any linear combination of sinusoids. But there is another spectral representation which is only slightly more complicated and turns out to be much more useful. It is based on writing a sinusoid as a sum of two exponentials, similar to Eq. (6*a*), i.e.,

$$A \cos (\omega_0 t + \theta) = \frac{A}{2} e^{j\theta} e^{j\omega_0 t} + \frac{A}{2} e^{-j\theta} e^{-j\omega_0 t} \tag{14}$$

which we will call the *conjugate-phasor* representation since the two terms are complex conjugates of each other. The corresponding diagram is shown in Fig. 2.4*a*, where there are now two phasors having equal lengths but opposite angles and directions of rotation. Hence, it is the vector sum at any time that equals $v(t)$. Note that the sum always falls on the real axis—as it should since $v(t)$ is real.

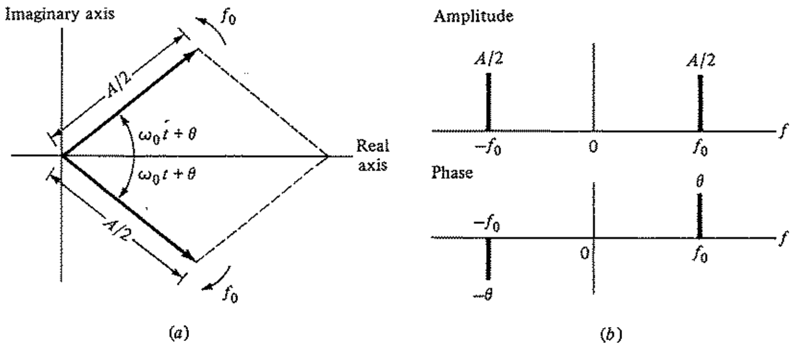


FIGURE 2.4
(*a*) Conjugate-phasor representation of $A \cos (\omega_0 t + \theta)$; (*b*) two-sided line spectrum.

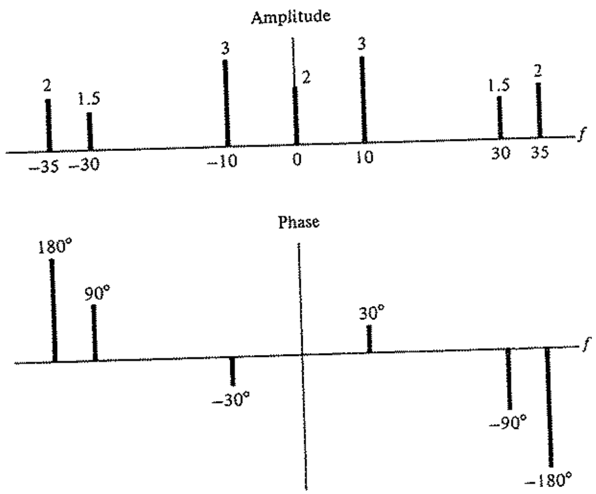


FIGURE 2.5
The two-sided version of Fig. 2.3.

A line spectrum taken from the conjugate-phasor expression must include negative frequencies to allow for the two rotational directions. Thus in the *two-sided line spectrum* of Fig. 2.4*b*, half of the original amplitude is associated with each of the two frequencies $\pm f_0$. The rules for constructing such spectra are quite simple; a little thought will show that the amplitude lines have even symmetry in f, while the phase lines have odd symmetry. The symmetry is a direct consequence of $v(t)$ being a real (noncomplex) function of time. So, for example, Fig. 2.5 is the two-sided version of Fig. 2.3.

There is one significant difference in interpreting the two-sided line spectrum compared to the positive-frequency spectrum. A single line in the latter represents a cosine wave, via Re $[e^{j\omega t}]$. But in the two-sided case, one line by itself represents a single phasor, and the conjugate term is required to get a real function of time. Thus, whenever we speak of some frequency interval in a two-sided spectrum, such as f_1 to f_2, we must include the corresponding negative-frequency interval, $-f_1$ to $-f_2$. A simple notation specifying both intervals is $f_1 \leq |f| \leq f_2$.

It should be emphasized that these line spectra are just pictorial ways of representing certain signals in terms of the form $e^{j\omega t}$. The effect on a signal by a network then reduces to a steady-state AC problem. We also point out that, in some ways, the amplitude spectrum is more important than the phase spectrum. Both parts are required, of course, to unambiguously define a signal in the time domain, but the

amplitude spectrum by itself shows what frequencies are present and in what proportions—that is, it tells us the signal's *frequency content*. The specific advantage of the two-sided version will become apparent as we go along.

EXERCISE 2.2 Explain why the zero-frequency term results in just one amplitude line in Fig. 2.5, and state how the figure would be changed if that term were negative.

Transfer Functions and Frequency Response

Just as a signal can be described in the frequency domain by way of its spectrum, a network or system can be described in terms of its frequency-response characteristics as determined from the transfer function $H(j\omega)$. Emphasizing the frequency-domain viewpoint, it is now convenient to introduce a new notation $H(f)$, defined by

$$H(f) \triangleq H(j\omega) \qquad \text{with } \omega = 2\pi f \tag{15}$$

which, henceforth, will be called either the frequency-response function or the transfer function. Although Eq. (15) modestly violates formal mathematical notation, it simply means that $H(f)$ is identical to $H(j\omega)$ with ω replaced by $2\pi f$. It then follows from Eq. (8) that, for a real network,

$$H(-f) = H^*(f) \tag{16a}$$

or in polar form,

$$|H(-f)| = |H(f)| \qquad \arg[H(-f)] = -\arg[H(f)] \tag{16b}$$

a property known as *hermitian symmetry*.

For the interpretation of $H(f)$, suppose the input signal is a single phasor

$$x(t) = A_x e^{j\theta_x} e^{j\omega_0 t} \qquad \omega_0 = 2\pi f_0$$

so, from Eq. (4*b*), the output is

$$\begin{aligned} y(t) &= H(f_0) A_x e^{j\theta_x} e^{j\omega_0 t} \\ &= \underbrace{|H(f_0)| A_x}_{A_y} \exp j\underbrace{(\theta_x + \arg[H(f_0)])}_{\theta_y} e^{j\omega_0 t} \end{aligned}$$

Hence, just as in the AC case, the input and output signal parameters are related by

$$\frac{A_y}{A_x} = |H(f_0)| \qquad \theta_y - \theta_x = \arg[H(f_0)] \tag{17}$$

so $|H(f_0)|$ is the ratio of amplitudes and arg $[H(f_0)]$ is the phase difference, both at the specific frequency f_0. Generalizing, we conclude that $|H(f)|$ gives the system's *amplitude ratio* (sometimes called *amplitude response* or *gain*) and arg $[H(f)]$ gives the *phase shift*, both as continuous functions of frequency. Plots of these two versus f give a frequency-domain representation of the system, analogous to the amplitude and phase spectrum of a signal. Moreover, the hermitian symmetry of Eq. (16) means that the amplitude ratio will be an even function of frequency while the phase shift will be an odd function. An illustrative example is presented below after brief consideration of the question of determining $H(f)$ for a particular system.

Given the circuit diagram of a network, finding $H(f)$ is nothing more than a phasor analysis problem, i.e., we assume the input is $e^{j2\pi ft}$ and calculate the output, which will be $H(f)e^{j2\pi ft}$. All the standard electrical engineering tools—Ohm's law for complex impedance, Kirchhoff's laws, etc.—can be brought to bear. Alternatively, if a system is described by a linear differential equation with constant coefficients, of the general form

$$a_n \frac{d^n y}{dt^n} + \cdots + a_1 \frac{dy}{dt} + a_0 y(t) = b_m \frac{d^m x}{dt^m} + \cdots + b_1 \frac{dx}{dt} + b_0 x(t) \tag{18a}$$

then

$$H(f) = \frac{b_m(j2\pi f)^m + \cdots + b_1(j2\pi f) + b_0}{a_n(j2\pi f)^n + \cdots + a_1(j2\pi f) + a_0} \tag{18b}$$

whose derivation is left as an exercise. A third method, involving the system's impulse response, is presented in Sect. 2.5.

EXERCISE 2.3 Derive Eq. (18*b*) by substituting $x(t) = e^{j2\pi ft}$ and $y(t) = H(f)e^{j2\pi ft}$ in Eq. (18*a*) and solving for $H(f)$.

Example 2.1 RC Lowpass Filter

The network of Fig. 2.6*a*, virtually a classic in communications, is called an RC lowpass filter; $x(t)$ is the input voltage and $y(t)$ is the output voltage under open-circuit (unloaded) conditions. Finding the transfer function is a simple matter since, with $x(t) = e^{j2\pi ft}$, application of the voltage-divider relation yields

$$y(t) = \frac{Z_C}{R + Z_C} e^{j2\pi ft}$$

where the capacitor's impedance is $Z_C = 1/j2\pi fC$. Thus,

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

$$= \frac{1}{1 + j(f/B)} \tag{19}$$

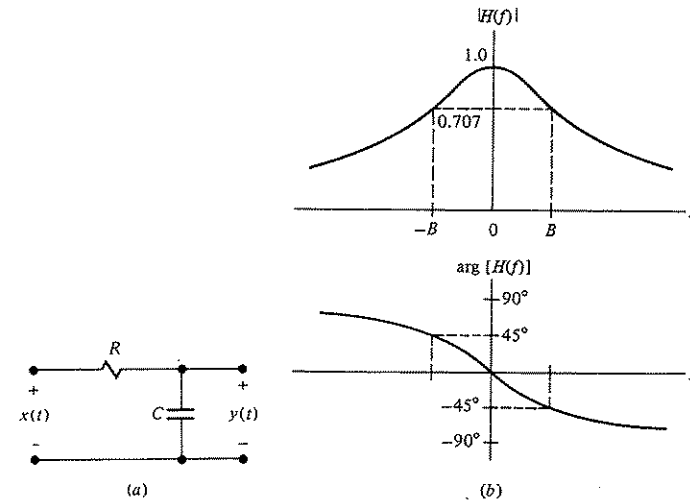


FIGURE 2.6
RC lowpass filter. (*a*) Circuit; (*b*) transfer function

in which we have defined the system parameter

$$B \triangleq \frac{1}{2\pi RC}$$

Conversion to polar form gives the amplitude ratio and phase shift as

$$|H(f)| = \frac{1}{\sqrt{1 + (f/B)^2}} \tag{20a}$$

$$\arg [H(f)] = -\arctan \frac{f}{B} \tag{20b}$$

which are plotted in Fig. 2.6*b*. ////

EXERCISE 2.4 Suppose the resistor and capacitor are interchanged in Fig. 2.6*a*. Find the new $H(f)$ and, by sketching $|H(f)|$, justify calling this circuit an RC *highpass* filter.

2.2 PERIODIC SIGNALS AND FOURIER SERIES

A signal $v(t)$ is said to be periodic with repetition period T_0 if, for any integer m,

$$v(t \pm mT_0) = v(t) \qquad -\infty < t < \infty \tag{1}$$

Since this implies a signal that lasts forever, our earlier remarks about the difference between a signal and its mathematical model are pertinent here. Assuming that Eq. (1) is a reasonable model, then Fourier series expansion can be invoked to decompose $v(t)$ into a linear combination of sinusoids or phasors—which, in turn, leads to the signal's line spectrum. The essential proviso for Fourier expansion is that $v(t)$ have well-defined average power, and because average power and time averages in general are commonly used terms in communications, we take a brief digression to formalize the concepts.

The *average* of an arbitrary time function $v(t)$ will be denoted by $\langle v(t)\rangle$ and defined in general as

$$\langle v(t)\rangle \triangleq \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} v(t)\,dt \tag{2}$$

If $v(t)$ happens to be periodic, Eq. (2) reduces to the average over any interval exactly T_0 seconds long, i.e.,

$$\langle v(t)\rangle = \frac{1}{T_0}\int_{T_0} v(t)\,dt \tag{3}$$

where $\int_{T_0}$ stands for integration from t_1 to $t_1 + T_0$ with t_1 an arbitrary constant. Average power is, of course, the time average of instantaneous power; e.g., if $v(t)$ is the voltage across a 1-ohm (Ω) resistance, the instantaneous power is $v^2(t)$ and the average power is $\langle v^2(t)\rangle$. For our purposes it is convenient to assume all resistances are normalized to unity so that, whether $v(t)$ is a periodic voltage or current, its *average power P* will be defined as

$$P \triangleq \langle |v(t)|^2\rangle = \frac{1}{T_0}\int_{T_0} |v(t)|^2\,dt \tag{4}$$

The signal $v(t)$ is then said to have well-defined average power if the integral (4) exists and yields a finite quantity. Note also that we have allowed for the possibility of complex signals by writing $|v(t)|^2 = v(t)v^*(t)$ instead of $v^2(t)$.

EXERCISE 2.5 Use Eqs. (3) and (4) to show that if

$$v(t) = A\cos(\omega_0 t + \theta) \qquad \omega_0 = \frac{2\pi}{T_0} \tag{5a}$$

then

$$\langle v(t)\rangle = 0 \qquad \langle |v(t)|^2\rangle = \frac{A^2}{2} \tag{5b}$$

Thus the average power of a sinusoid depends only on its amplitude.

Fourier Series and Line Spectra

The *exponential Fourier series* expansion of a periodic signal $v(t)$ is

$$v(t) = \sum_{n=-\infty}^{\infty} c(nf_0)e^{j2\pi nf_0 t} \qquad f_0 = \frac{1}{T_0} \tag{6}$$

where $c(nf_0)$ is the nth Fourier coefficient

$$c(nf_0) \triangleq \frac{1}{T_0}\int_{T_0} v(t)e^{-j2\pi nf_0 t}\,dt \tag{7}$$

Equation (6) states that $v(t)$ can be expressed as a linear combination or weighted sum of phasors at the frequencies $f = nf_0 = 0, \pm f_0, \pm 2f_0, \ldots$, the weighting factors being given by Eq. (7). If $v(t)$ has well-defined average power, then the summation on the right of Eq. (6) converges to $v(t)$ everywhere it is finite and continuous, which would be true of any physical signal. Insofar as engineering purposes are concerned, we may view the series as being identical to $v(t)$.

Actually, one seldom carries out the summation of Eq. (6) to find $v(t)$; instead, given a periodic signal, we use Eq. (7) to find its Fourier coefficients and, from that, the line spectrum. Before addressing that aspect, a few points regarding $c(nf_0)$ are in order. In general, the coefficients are complex quantities, even when the signal is real; therefore, we can write

$$c(nf_0) = |c(nf_0)|\,e^{j\arg[c(nf_0)]}$$

If $v(t)$ is a *real* function, then replacing n by $-n$ in Eq. (7) shows that

$$c(-nf_0) = c^*(nf_0) \tag{8}$$

so again we have conjugate phasors and hermitian symmetry. Finally, note that $c(nf_0)$ does not depend on time since Eq. (7) is a definite integral with t being the variable of integration. It is helpful, however, to regard $c(nf_0)$ as a function of frequency f defined only for the discrete frequencies $f = nf_0$.

Turning to the spectral interpretation, we see from Eq. (6) that a periodic signal contains only those frequency components that are *integer multiples of the fundamental frequency* $f_0 = 1/T_0$ or, in other words, all the frequencies are *harmonics* of the fundamental. Since the coefficient of the nth harmonic is $c(nf_0)$, its amplitude and phase are $|c(nf_0)|$ and $\arg[c(nf_0)]$, respectively. Therefore, we have a two-sided line spectrum with $|c(nf_0)|$ giving the amplitude and $\arg[c(nf_0)]$ the phase. Some of the important properties of such spectra are listed below.

1 All spectral lines are equally spaced by f_0 since all the frequencies are harmonically related to the fundamental.

2 The DC component equals the *average value* of the signal since setting $n = 0$ in Eq. (7) yields

$$c(0) = \frac{1}{T_0}\int_{T_0} v(t)\,dt = \langle v(t)\rangle \qquad (9)$$

Therefore, calculated values of $c(0)$ may be checked by inspection of $v(t)$—which is a wise practice since the integration frequently yields an indeterminate form for $c(0)$.

3 If $v(t)$ is *real*, the amplitude spectrum has even symmetry while the phase spectrum has odd symmetry, i.e.,

$$|c(-nf_0)| = |c(nf_0)| \qquad \arg[c(-nf_0)] = -\arg[c(nf_0)] \qquad (10)$$

which follows from Eq. (8).

4 If a real signal has *even symmetry* in time, such that

$$v(-t) = v(t) \qquad (11a)$$

then $c(nf_0)$ is entirely *real* and

$$\arg[c(nf_0)] = 0 \quad \text{or} \quad \pm 180^\circ \qquad (11b)$$

where $\pm 180^\circ$ corresponds to $c(nf_0)$ being negative. Conversely, if a real signal has *odd time symmetry*, i.e.,

$$v(-t) = -v(t) \qquad (12a)$$

then $c(nf_0)$ is entirely *imaginary* and

$$\arg[c(nf_0)] = \pm 90^\circ \qquad (12b)$$

which stems from the fact that $\pm j = e^{\pm j\pi/2} = e^{\pm j90^\circ}$. Proving these symmetry relations is left for the reader.

One final point before taking up an example: When $v(t)$ is real, we can draw upon Eq. (10) and regroup the exponential series in conjugate-phasor pairs of the form

$$c(nf_0)e^{j2\pi nf_0t} + c^*(nf_0)e^{-j2\pi nf_0t} = 2|c(nf_0)|\cos(2\pi nf_0 t + \arg[c(nf_0)])$$

so that Eq. (6) becomes

$$v(t) = c(0) + \sum_{n=1}^{\infty} |2c(nf_0)|\cos(2\pi nf_0 t + \arg[c(nf_0)]) \qquad (13)$$

By this process we have arrived at a *trigonometric* Fourier series, and in a sense have come full circle, for $v(t)$ is now described as a sum of sinusoids rather than conjugate phasors. Indeed, Eq. (13) is less versatile than Eq. (6) and certainly lacks the attractive symmetry of the exponential series. We shall, however, have some occasions for its use and for the corresponding positive-frequency line spectrum.

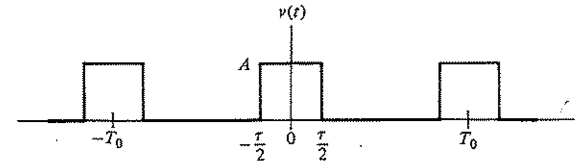


FIGURE 2.7
Rectangular pulse train.

Example 2.2 Rectangular Pulse Train

As an important example of the ideas we have discussed, let us find the line spectrum of the periodic waveform in Fig. 2.7, called a rectangular pulse train. To calculate $c(nf_0)$, we take the range of integration in Eq. (7) as† $[-T_0/2, T_0/2]$ and observe that in this interval

$$v(t) = \begin{cases} A & |t| < \dfrac{\tau}{2} \\ 0 & |t| > \dfrac{\tau}{2} \end{cases}$$

where τ is the pulse duration and A is the amplitude. Note, incidentally, that this signal model has stepwise *discontinuities* at $t = \pm\tau/2$, etc., and values of $v(t)$ are *undefined* wherever it is discontinuous. This illustrates one of the possible differences between a physical signal and its mathematical model, since a physical signal never has an abrupt stepwise transition. However, this model is useful if the actual transition time is very small compared to the pulse duration, and the undefined values at the discontinuity points have no effect on the calculation of $c(nf_0)$.

Proceeding with that calculation, we have

$$\begin{aligned} c(nf_0) &= \frac{1}{T_0}\int_{-T_0/2}^{T_0/2} v(t)e^{-j2\pi nf_0t}\,dt \\ &= \frac{1}{T_0}\int_{-\tau/2}^{\tau/2} Ae^{-j2\pi nf_0t}\,dt \\ &= \frac{A}{-j2\pi nf_0T_0}(e^{-j\pi nf_0\tau} - e^{+j\pi nf_0\tau}) \\ &= \frac{A}{\pi n}\sin \pi nf_0\tau \end{aligned}$$

where we used the fact that $f_0T_0 = 1$ and $e^{j\phi} - e^{-j\phi} = 2j\sin\phi$.

† The notation $[-T_0/2, T_0/2]$ stands for $-T_0/2 \le t \le T_0/2$.

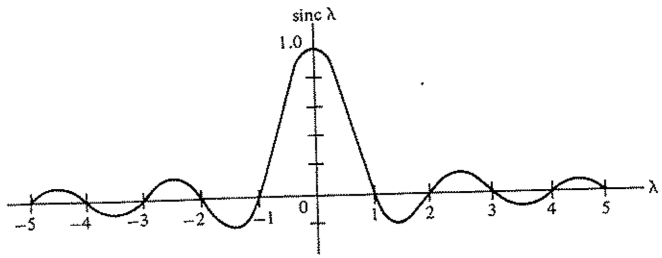


FIGURE 2.8
The function sinc $\lambda = (\sin \pi\lambda)/\pi\lambda$.

To somewhat simplify notation in the above result, we introduce a new function called the *sinc* function† and defined by

$$\operatorname{sinc} \lambda \triangleq \frac{\sin \pi\lambda}{\pi\lambda} \tag{14}$$

where λ is the independent variable. This function will be quite important owing to its relation to averages of exponentials and sinusoids; in particular, as is easily proved,

$$\frac{1}{T}\int_{-T/2}^{T/2} e^{\pm j2\pi ft}\,dt = \frac{1}{T}\int_{-T/2}^{T/2} \cos 2\pi ft\,dt = \operatorname{sinc} fT \tag{15}$$

in which T is an arbitrary constant not necessarily related to f. Figure 2.8 shows that sinc λ is an even function having its peak at $\lambda = 0$ and zero crossings at all other integer values of λ, i.e.,

$$\operatorname{sinc} \lambda = \begin{cases} 1 & \lambda = 0 \\ 0 & \lambda = \pm 1, \pm 2, \ldots \end{cases}$$

Numerical values of sinc λ and $\operatorname{sinc}^2 \lambda$ are given in Table C.

Using the sinc function, the series coefficients for the rectangular pulse train become

$$c(nf_0) = Af_0\tau \frac{\sin \pi n f_0 \tau}{\pi n f_0 \tau} = Af_0\tau \operatorname{sinc} nf_0\tau \tag{16}$$

which is independent of time and strictly real—the latter because $v(t)$ happens to be real and even. Therefore, the amplitude spectrum is $|c(nf_0)| = Af_0\tau|\operatorname{sinc} nf_0\tau|$, as shown in Fig. 2.9*a* for the case where $f_0\tau = 1/4$. Such plots are facilitated by regarding the continuous function $Af_0\tau|\operatorname{sinc} f\tau|$ as the *envelope* of the lines, indicated by the

† Some authors use the so-called *sampling function*, Sa $(\lambda) \triangleq (\sin \lambda)/\lambda$; note that sinc λ = Sa $(\pi\lambda)$.

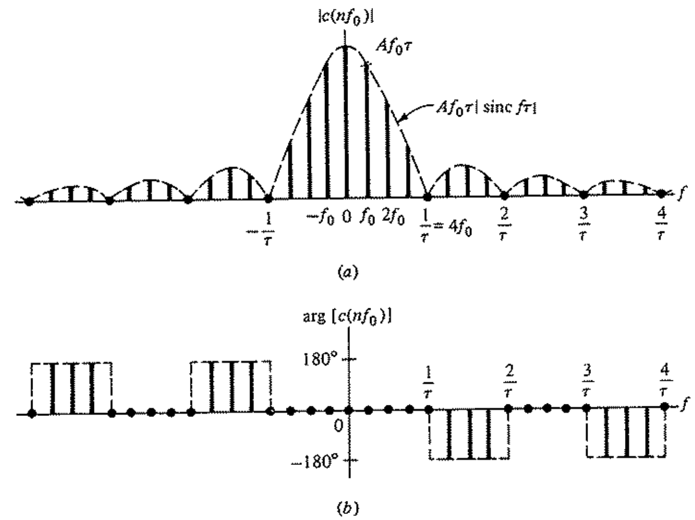


FIGURE 2.9
Spectrum of rectangular pulse train with $f_0\tau = \frac{1}{4}$. (*a*) Amplitude; (*b*) phase.

dashed curve. The spectral lines at $\pm 4f_0$, $\pm 8f_0$, etc., are "missing" since they fall precisely at multiples of $1/\tau$ where the envelope equals zero. The DC component has amplitude $c(0) = Af_0\tau = A\tau/T_0$ which should be recognized as the average value of $v(t)$ by inspection of Fig. 2.7. Note, incidentally, that τ/T_0 is the ratio of "on" time to period, frequently designated as the *duty cycle* in pulse electronics work.

Figure 2.9*b*, the phase spectrum, is constructed by observing that $c(nf_0)$ is always real but sometimes negative. Hence, arg $[c(nf_0)]$ takes on the values 0° and $\pm 180°$, according to the polarity of sinc $nf_0\tau$. Both $+180°$ and $-180°$ have been used to preserve the odd symmetry, although this is more or less arbitrary in such cases. ////

EXERCISE 2.6 Sketch the amplitude spectrum of a rectangular pulse train for each of the following cases: $\tau = T_0/5$, $\tau = T_0/2$, $\tau = T_0$. In the last case the pulse train degenerates into a constant for all time; how does this show up in the spectrum?

EXERCISE 2.7 Show that the *square wave* in Fig. 2.10 has

$$c(nf_0) = \begin{cases} A \operatorname{sinc} \dfrac{n}{2} & n = \pm 1, \pm 3, \ldots \\ 0 & n = 0, \pm 2, \pm 4, \ldots \end{cases} \tag{17}$$

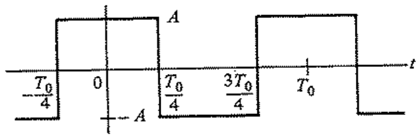


FIGURE 2.10
Square wave.

(*Hint:* The manipulation below is useful for simplifying the answer in this and similar problems, particularly when ϕ_1 or ϕ_2 is zero.)

$$e^{j\phi_1} \pm e^{j\phi_2} = [e^{j(\phi_1-\phi_2)/2} \pm e^{-j(\phi_1-\phi_2)/2}]e^{j(\phi_1+\phi_2)/2}$$

$$= \begin{cases} 2\cos\dfrac{\phi_1-\phi_2}{2}\,e^{j(\phi_1+\phi_2)/2} \\ 2j\sin\dfrac{\phi_1-\phi_2}{2}\,e^{j(\phi_1+\phi_2)/2} \end{cases} \tag{18}$$

Parseval's Power Theorem

This famous theorem relates the average power P of a periodic signal to its Fourier series coefficients. To derive the relationship, we start with the definition of P, Eq. (4), write $|v(t)|^2 = v(t)v^*(t)$, replace $v^*(t)$ by its Fourier series, and interchange the order of summation and integration, as follows:

$$P = \frac{1}{T_0}\int_{T_0} v(t)\overbrace{\left[\sum_{n=-\infty}^{\infty} c^*(nf_0)e^{-j2\pi nf_0 t}\right]}^{v^*(t)} dt$$

$$= \sum_{n=-\infty}^{\infty} \underbrace{\left[\frac{1}{T_0}\int_{T_0} v(t)e^{-j2\pi nf_0 t}\,dt\right]}_{c(nf_0)} c_v^*(nf_0)$$

Thus

$$P = \sum_{n=-\infty}^{\infty} c(nf_0)c^*(nf_0) = \sum_{n=-\infty}^{\infty} |c(nf_0)|^2 \tag{19}$$

The spectral interpretation of this result is extraordinarily simple; i.e., average power can be found by squaring and adding the heights of the amplitude lines. Observe that Eq. (19) does not involve the phase spectrum arg $[c(nf_0)]$, underscoring our prior comment about the dominant role of the amplitude spectrum relative to a signal's frequency content.

For further interpretation of Eq. (19), recall that the exponential Fourier series expands $v(t)$ as a sum of phasors each of the form $c(nf_0)e^{j2\pi nf_0 t}$. Now it is easily shown that the average power of each of these is

$$\langle |c(nf_0)e^{j2\pi nf_0 t}|^2\rangle = |c(nf_0)|^2$$

Therefore, Parseval's theorem implies *superposition of average power* in that the total average power of $v(t)$ is the sum of the average powers of its phasor components.

Periodic Steady-State Response

In Sect. 2.1 we found the AC steady-state response of a network by expressing the sinusoidal input as a sum of phasors. Those results are readily extended to the case of an arbitrary periodic input signal by the method presented here.

First, expand the input $x(t)$ as

$$x(t) = \sum_{n=-\infty}^{\infty} c_x(nf_0)e^{j2\pi nf_0 t} \tag{20a}$$

where $c_x(nf_0)$ is found from Eq. (7). Then, since each of the above phasors produces an output of the form $H(nf_0)c_x(nf_0)\exp(j2\pi nf_0 t)$, and since superposition is assumed, the total output is

$$y(t) = \sum_{n=-\infty}^{\infty} c_y(nf_0)e^{j2\pi nf_0 t} \tag{20b}$$

where

$$c_y(nf_0) = H(nf_0)c_x(nf_0) \tag{21}$$

Therefore, the periodic steady-state response is a periodic signal having the same fundamental frequency as the input, whose Fourier series coefficients $c_y(nf_0)$ equal the respective coefficients of the input signal multiplied by the network's transfer function $H(f)$ evaluated at $f = nf_0$. Moreover, the average power in the output signal is

$$P_y = \sum_{n=-\infty}^{\infty} |c_y(nf_0)|^2 = \sum_{n=-\infty}^{\infty} |H(nf_0)|^2|c_x(nf_0)|^2 \tag{22}$$

by application of Parseval's theorem.

The frequency-domain interpretation of these results is best seen by converting Eq. (21) to polar form, thus:

$$\begin{aligned} |c_y(nf_0)| &= |H(nf_0)|\,|c_x(nf_0)| \\ \arg[c_y(nf_0)] &= \arg[c_x(nf_0)] + \arg[H(nf_0)] \end{aligned} \tag{23}$$

Putting this into words, the output amplitude spectrum equals the input amplitude spectrum *times* the network's amplitude ratio, whereas the output phase spectrum equals the input phase spectrum *plus* the network's phase shift at the frequencies in question. The phase relationship is additive instead of multiplicative simply because arguments add when exponentials are multiplied.

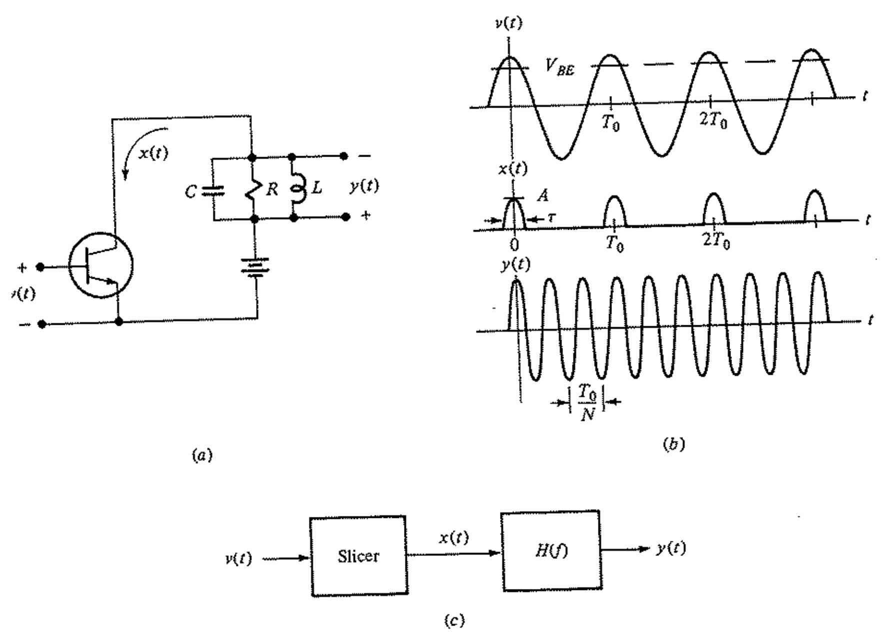


FIGURE 2.11
Frequency multiplier. (*a*) Circuit; (*b*) waveforms; (*c*) block diagram.

Equations (20) to (22) provide a theoretical solution to the periodic steady-state analysis problem. Practically speaking, however, determining the actual *shape* of the output waveform $y(t)$ from Eq. (20*b*) is a tedious process—unless there are only a few significant terms in the sum. The following example falls in this category and demonstrates how the spectral interpretation helps estimate the number of significant terms.

Example 2.3 Frequency Multiplier

Figure 2.11*a* is the circuit diagram of a frequency multiplier, a device often used in communication systems. The input voltage $v(t)$ is a sinusoid at a specified frequency f_0 and the output $y(t)$ is supposed to be a sinusoid at a multiple of the input frequency, say Nf_0. Usually, $N = 2$ or 3, i.e., a *frequency doubler* or *tripler*, and greater multiplication factors are obtained by connecting doublers and triplers in tandem.

The waveforms shown in Fig. 2.11*b* illustrate the essential operations, as follows. The transistor is a silicon *npn* type without base-emitter bias so the current $x(t)$ is negligible unless $v(t)$ exceeds the base-emitter voltage drop V_{BE}. Thus, if the amplitude

of $v(t)$ is just slightly greater than V_{BE}, as indicated, $x(t)$ consists of short pulses having period $T_0 = 1/f_0$. From our previous work we know that such a waveform in general will contain *all harmonics* of f_0, and the role of the parallel RLC circuit† is to pick out the Nth harmonic and reject the rest. This implies that the circuit is *tuned* or *resonant* at Nf_0. For analysis purposes these functions can be represented in block-diagram form, per Fig. 2.11*c*, which we will take as our model of the device.

While conceivably one could exactly calculate the Fourier series coefficients of $x(t)$, there is really no need for such precision. Rather, we will approximate $x(t)$ as a *rectangular* pulse train of amplitude A and duration $\tau \ll T_0$; hence, for small values of n,

$$c_x(nf_0) \approx Af_0\tau \qquad |n| \ll \frac{T_0}{\tau} \tag{24}$$

where we have used the results of Example 2.2 together with sinc $\lambda \approx 1$ for $|\lambda| \ll 1$. It is likewise assumed that $N \ll T_0/\tau$.

Next we need the transfer function $H(f)$ of the tuned circuit. Noting that its input $x(t)$ is a current and the output $y(t)$ is a voltage, $H(f)$ is identical to the circuit's impedance $Z(j\omega)$ with $\omega = j2\pi f$. Routine analysis gives‡

$$H(f) = \frac{R}{1 + jQ\left(\dfrac{f^2 - f_r^2}{ff_r}\right)} \tag{25}$$

where

$$f_r \triangleq \frac{1}{2\pi\sqrt{LC}} \qquad Q \triangleq \frac{R}{2\pi f_r L} = R\sqrt{\frac{C}{L}}$$

which are the resonant frequency and quality factor, respectively. It has been assumed in Eq. (25) that $Q > \frac{1}{2}$ so the circuit actually is resonant; as a matter of fact, the present application requires $Q \gg 1$. The corresponding amplitude ratio $|H(f)|$ is plotted in Fig. 2.12*a* directly above the input amplitude spectrum $|c_x(nf_0)|$, Fig. 2.12*b*. Negative frequencies have been omitted for convenience, since we know that both functions have even symmetry.

If we recall that $|c_y(nf_0)| = |H(nf_0)|\,|c_x(nf_0)|$, Fig. 2.12 suggests that all the harmonics in $y(t)$ except Nf_0 can be made to have negligible amplitude if

$$f_r = Nf_0 \qquad \text{and} \qquad Q \gg N/2$$

† As a rule the resistance R represents losses in L and any coupled load rather than being a distinct circuit element.
‡ See, for instance, Close (1966, chap. 6).

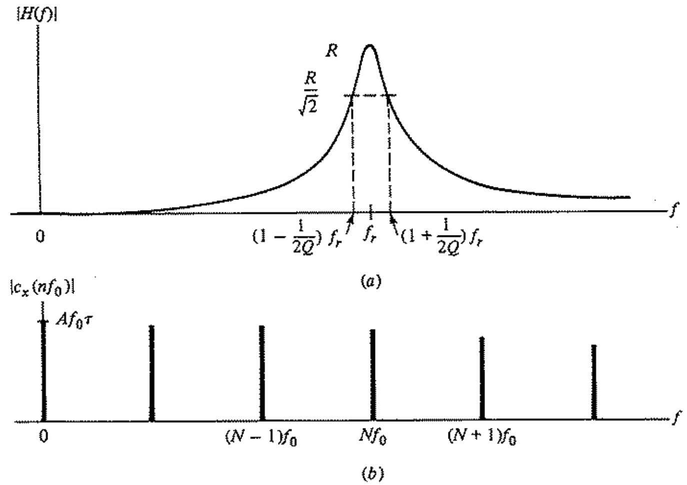


FIGURE 2.12
(a) Amplitude ratio of tuned circuit, $Q = 10$; (b) amplitude spectrum of input signal.

the condition on Q ensuring that $[1 + (1/2Q)]f_r < (N + 1)f_0$ and $[1 - (1/2Q)f_r] > (N - 1)f_0$ so the tuned circuit "passes" only Nf_0. Therefore, the output amplitude spectrum consists essentially of two terms, $c_y(\pm Nf_0)$, and

$$\begin{aligned} y(t) &\approx H(Nf_0)c_x(Nf_0)e^{j2\pi Nf_0t} + H(-Nf_0)c_x(-Nf_0)e^{-j2\pi Nf_0t} \\ &= 2RAf_0\,\tau \cos 2\pi Nf_0\,t \end{aligned}$$

where we have used Eqs. (24) and (25) and converted the result to sinusoidal form.

////

2.3 NONPERIODIC SIGNALS AND FOURIER TRANSFORMS

We have described a periodic signal as one with a repeating characteristic applied for a long time interval, theoretically infinite. Now we consider nonperiodic signals whose effects are concentrated over a brief period of time. Such signals may be *strictly timelimited*, so $v(t)$ is identically zero outside of a specified interval, or *asymptotically timelimited*, so $v(t) \to 0$ as $t \to \pm\infty$. In either case it is assumed that the signal's total *energy* is well-defined, energy being measured in the same normalized sense as was power in the previous section, namely,

$$E \triangleq \int_{-\infty}^{\infty} |v(t)|^2\,dt \tag{1}$$

Implied by this definition is the fact that if E is finite, then both the average value and average power equal zero, an observation pursued more fully in Sect. 2.6. Here we are concerned with the frequency-domain description of nonperiodic energy signals via the Fourier transform.

Fourier Transforms and Continuous Spectra

A periodic signal can be represented by its exponential Fourier series

$$v(t) = \sum_{n=-\infty}^{\infty} \overbrace{\left[\frac{1}{T_0}\int_{T_0} v(t)e^{-j2\pi nf_0t}\,dt\right]}^{c_v(nf_0)} e^{j2\pi nf_0t} \tag{2}$$

where the integral expression for $c_v(nf_0)$ has been written out in full. According to the *Fourier integral theorem* there is a similar representation for a *nonperiodic* signal, namely,

$$v(t) = \int_{-\infty}^{\infty} \underbrace{\left[\int_{-\infty}^{\infty} v(t)e^{-j2\pi ft}\,dt\right]}_{V(f)} e^{j2\pi ft}\,df \tag{3}$$

The bracketed term is the *Fourier transform* of $v(t)$, symbolized by $V(f)$ or $\mathscr{F}[v(t)]$ and defined as

$$V(f) = \mathscr{F}[v(t)] \triangleq \int_{-\infty}^{\infty} v(t)e^{-j2\pi ft}\,dt \tag{4}$$

which is an integration over all time. The theorem (3) states that $v(t)$ can be found by the *inverse Fourier transform* of $V(f)$,

$$v(t) = \mathscr{F}^{-1}[V(f)] \triangleq \int_{-\infty}^{\infty} V(f)e^{j2\pi ft}\,df \tag{5}$$

which is an integration over all frequency.

Equations (4) and (5) are often referred to as the Fourier integrals or the Fourier transform pair† and, at first glance, they seem to be a closed circle of operations. In a given problem, however, one usually knows either $V(f)$ or $v(t)$ but not both. If $V(f)$ is known, $v(t)$ can be found by carrying out the inverse transform (5), and vice versa when finding $V(f)$ from $v(t)$.

Turning to the frequency-domain picture, a comparison of Eqs. (2) and (3) indicates that $V(f)$ plays the same role for nonperiodic signals that $c_v(nf_0)$ plays for

† Somewhat different definitions apply when $\omega = 2\pi f$ is used as the independent variable of the frequency domain.

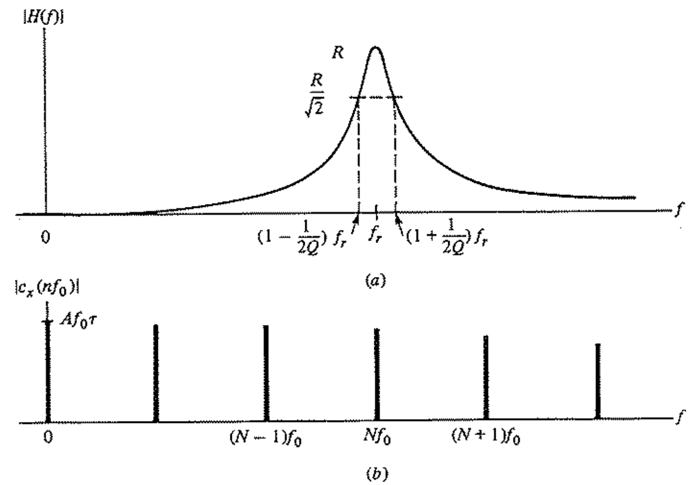


FIGURE 2.12
(a) Amplitude ratio of tuned circuit, $Q = 10$; (b) amplitude spectrum of input signal.

the condition on Q ensuring that $[1 + (1/2Q)]f_r < (N+1)f_0$ and $[1 - (1/2Q)f_r] > (N-1)f_0$ so the tuned circuit "passes" only Nf_0. Therefore, the output amplitude spectrum consists essentially of two terms, $c_y(\pm Nf_0)$, and

$$y(t) \approx H(Nf_0)c_x(Nf_0)e^{j2\pi Nf_0 t} + H(-Nf_0)c_x(-Nf_0)e^{-j2\pi Nf_0 t}$$
$$= 2RAf_0 \tau \cos 2\pi N f_0 t$$

where we have used Eqs. (24) and (25) and converted the result to sinusoidal form. ////

2.3 NONPERIODIC SIGNALS AND FOURIER TRANSFORMS

We have described a periodic signal as one with a repeating characteristic applied for a long time interval, theoretically infinite. Now we consider nonperiodic signals whose effects are concentrated over a brief period of time. Such signals may be *strictly timelimited*, so $v(t)$ is identically zero outside of a specified interval, or *asymptotically timelimited*, so $v(t) \to 0$ as $t \to \pm\infty$. In either case it is assumed that the signal's total energy is well-defined, energy being measured in the same normalized sense as was power in the previous section, namely,

$$E \triangleq \int_{-\infty}^{\infty} |v(t)|^2 \, dt \qquad (1)$$

Implied by this definition is the fact that if E is finite, then both the average value and average power equal zero, an observation pursued more fully in Sect. 2.6. Here we are concerned with the frequency-domain description of nonperiodic energy signals via the Fourier transform.

Fourier Transforms and Continuous Spectra

A periodic signal can be represented by its exponential Fourier series

$$v(t) = \sum_{n=-\infty}^{\infty} \overbrace{\left[\frac{1}{T_0}\int_{T_0} v(t)e^{-j2\pi n f_0 t}\,dt\right]}^{c_v(nf_0)} e^{j2\pi n f_0 t} \qquad (2)$$

where the integral expression for $c_v(nf_0)$ has been written out in full. According to the *Fourier integral theorem* there is a similar representation for a *nonperiodic* signal, namely,

$$v(t) = \int_{-\infty}^{\infty} \underbrace{\left[\int_{-\infty}^{\infty} v(t)e^{-j2\pi f t}\,dt\right]}_{V(f)} e^{j2\pi f t}\,df \qquad (3)$$

The bracketed term is the *Fourier transform* of $v(t)$, symbolized by $V(f)$ or $\mathscr{F}[v(t)]$ and defined as

$$V(f) = \mathscr{F}[v(t)] \triangleq \int_{-\infty}^{\infty} v(t)e^{-j2\pi f t}\,dt \qquad (4)$$

which is an integration over all time. The theorem (3) states that $v(t)$ can be found by the *inverse Fourier transform* of $V(f)$,

$$v(t) = \mathscr{F}^{-1}[V(f)] \triangleq \int_{-\infty}^{\infty} V(f)e^{j2\pi f t}\,df \qquad (5)$$

which is an integration over all frequency.

Equations (4) and (5) are often referred to as the Fourier integrals or the Fourier transform pair† and, at first glance, they seem to be a closed circle of operations. In a given problem, however, one usually knows either $V(f)$ or $v(t)$ but not both. If $V(f)$ is known, $v(t)$ can be found by carrying out the inverse transform (5), and vice versa when finding $V(f)$ from $v(t)$.

Turning to the frequency-domain picture, a comparison of Eqs. (2) and (3) indicates that $V(f)$ plays the same role for nonperiodic signals that $c_v(nf_0)$ plays for

† Somewhat different definitions apply when $\omega = 2\pi f$ is used as the independent variable of the frequency domain.

periodic signals. Thus, $V(f)$ is the *spectrum* of the nonperiodic signal $v(t)$. But $V(f)$ is a continuous function defined for all values of f whereas $c_v(nf_0)$ is defined only for discrete frequencies. Therefore, a nonperiodic signal will have a *continuous spectrum* rather than a line spectrum. Again, comparing Eqs. (2) and (3) helps explain this difference: in the periodic case we return to the time domain by *summing* discrete-frequency phasors while in the nonperiodic case we *integrate* a continuous frequency function.

Like $c_v(nf_0)$, $V(f)$ generally is a complex function so that $|V(f)|$ is the amplitude spectrum and $\arg[V(f)]$ is the phase spectrum. Other important properties of $V(f)$, paralleling properties of $c_v(nf_0)$, are listed below.

1 If $v(t)$ is *real*, then $V(-f) = V^*(f)$ and

$$|V(-f)| = |V(f)| \qquad \arg[V(-f)] = -\arg[V(f)] \tag{6}$$

Hence, the spectrum has hermitian symmetry.

2 If $v(t)$ has either even or odd *time symmetry*, then Eq. (4) simplifies to

$$V(f) = \begin{cases} 2\int_0^\infty v(t)\cos\omega t\,dt & v(t)\text{ even} \\ -j2\int_0^\infty v(t)\sin\omega t\,dt & v(t)\text{ odd} \end{cases} \tag{7}$$

where we have written ω in place of $2\pi f$ for notational convenience, a practice frequently used hereafter. It follows from Eq. (7) that if $v(t)$ is also real, then $V(f)$ is purely real or imaginary, respectively.

3 The value of $V(f)$ at $f = 0$ equals the *net area* of $v(t)$, i.e.,

$$V(0) = \int_{-\infty}^{\infty} v(t)\,dt \tag{8}$$

which compares with the periodic case where $c_v(0)$ equals the average value of $v(t)$.

EXERCISE 2.8 Integrals of the general form $\int_{-T}^{T} w(t)\,dt$ simplify when the integrand is symmetrical; specifically, for any constant T,

$$\int_{-T}^{T} w(t)\,dt = \begin{cases} 2\int_0^T w(t)\,dt & \text{if } w(-t) = w(t) \quad (9a) \\ 0 & \text{if } w(-t) = -w(t) \quad (9b) \end{cases}$$

Use Eq. (9) to derive Eq. (7) from Eq. (4).

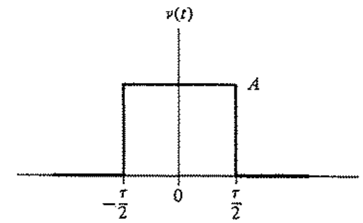


FIGURE 2.13
The rectangular pulse $v(t) = A\Pi(t/\tau)$.

Example 2.4 Rectangular Pulse

In Example 2.2, Sect. 2.2, we found the line spectrum of a rectangular pulse train. Now we will consider the continuous spectrum of the *single* rectangular pulse shown in Fig. 2.13. This is so common a signal model that it deserves a symbol of its own. Let us therefore adopt the notation

$$\Pi\left(\frac{t}{\tau}\right) \triangleq \begin{cases} 1 & |t| < \dfrac{\tau}{2} \\ 0 & |t| > \dfrac{\tau}{2} \end{cases} \tag{10}$$

which stands for a *rectangular function* with unit height or amplitude having width or duration τ centered at $t = 0$. Thus, $v(t) = A\Pi(t/\tau)$ in Fig. 2.13.

Since $v(t)$ has even symmetry, its Fourier transform is

$$\begin{aligned} V(f) &= 2\int_0^\infty v(t)\cos\omega t\,dt \\ &= 2\int_0^{\tau/2} A\cos\omega t\,dt = \frac{2A}{\omega}\sin\frac{\omega\tau}{2} \\ &= A\tau\,\mathrm{sinc}\,f\tau \end{aligned} \tag{11}$$

and $V(0) = A\tau$ which clearly equals the pulse's area. The corresponding spectrum is plotted in Fig. 2.14. This figure should be compared with Fig. 2.7 to illustrate our previous discussion of line spectra and continuous spectra.

It is apparent from $|V(f)|$ that the significant portion of the spectrum is in the range $|f| < 1/\tau$ since $|V(f)| \ll V(0)$ for $|f| > 1/\tau$. We therefore may take $1/\tau$ as a measure of the spectral "width." Now if the pulse duration is reduced (small τ), the frequency width is increased, whereas increasing the duration reduces the spectral width. Thus, short pulses have broad spectra, long pulses have narrow spectra. This phenomenon is called *reciprocal spreading* and is a general property of all signals, pulses or not, because high-frequency components are demanded by rapid time

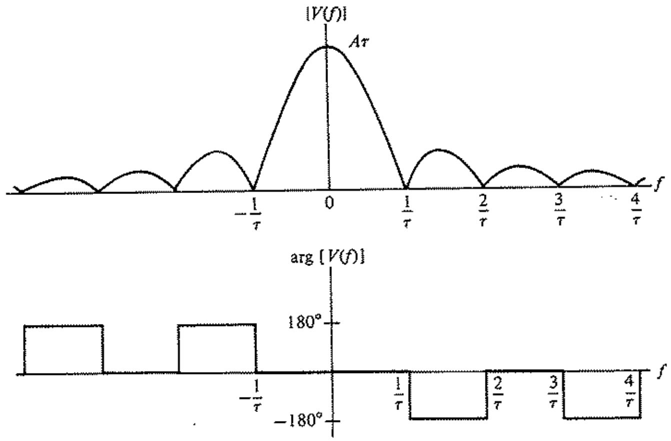


FIGURE 2.14
Spectrum of a rectangular pulse, $V(f) = A\tau \operatorname{sinc} f\tau$.

variations whereas smoother, slower time variations require relatively little high-frequency content. ////

Example 2.5 Exponential Pulse

Consider a decaying exponential function of the form $v(t) = Ae^{-t/T}$, $t > 0$. To ensure that the energy is finite, we further specify that $v(t) = 0$ for $t < 0$. Introducing the *unit step function* notation

$$u(t) \triangleq \begin{cases} 1 & t > 0 \\ 0 & t < 0 \end{cases} \tag{12}$$

we can write

$$v(t) = Ae^{-t/T}u(t)$$

which will be called an exponential pulse.

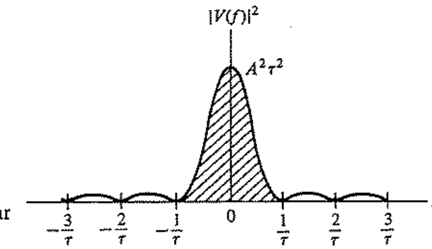
The integration for $V(f)$ is a simple problem, resulting in

$$V(f) = \int_0^{\infty} Ae^{-t/T}e^{-j\omega t}\,dt = \frac{AT}{1 + j2\pi fT}$$

so the amplitude and phase spectra are

$$|V(f)| = \frac{AT}{\sqrt{1 + (2\pi fT)^2}} \qquad \arg[V(f)] = -\arctan 2\pi fT$$

FIGURE 2.15
Energy spectral density of a rectangular pulse, $|V(f)|^2 = A^2\tau^2 \operatorname{sinc}^2 f\tau$.



Thus, unlike the previous example, the phase is a smooth curve between $+90°$ ($f = -\infty$) and $-90°$ ($f = +\infty$). Sharp-eyed readers will probably spot similarities between $V(f)$ and the transfer function of an RC lowpass filter, Example 2.1. The similarity is not accidental, and we will explain why in Sect. 2.5. ////

Rayleigh's Energy Theorem

Analogous to Parseval's power theorem, Rayleigh's energy theorem relates the total energy E of a signal to its amplitude spectrum. Actually, the theorem is a special case of an interesting integral relationship

$$\int_{-\infty}^{\infty} v(t)w^*(t)\,dt = \int_{-\infty}^{\infty} V(f)W^*(f)\,d_f \tag{13}$$

where $V(f) = \mathscr{F}[v(t)]$ and $W(f) = \mathscr{F}[w(t)]$. Proof of Eq. (13) follows the same method used to prove Parseval's theorem, Eq. (19), Sect. 2.2. We obtain Rayleigh's theorem by taking $w(t) = v(t)$ so the left-hand side of Eq. (13) is the energy of $v(t)$—by the definition (1)—and hence

$$E = \int_{-\infty}^{\infty} V(f)V^*(f)\,df = \int_{-\infty}^{\infty} |V(f)|^2\,df \tag{14}$$

Therefore, integrating the square of the amplitude spectrum $|V(f)|^2$ over all frequency yields the total energy.

The value of Eq. (14) lies not so much in computing E, since the time-domain integration of $|v(t)|^2$ often is easier. Rather, it implies that $|V(f)|^2$ gives the distribution of energy in the frequency domain, and therefore may be termed the *energy spectral density*. By this we mean that the energy in any differential frequency band $f \pm df/2$ equals $|V(f)|^2\,df$. That interpretation, in turn, lends quantitative support to the notion of spectral width in the sense that most of the energy of a given signal should be contained in the range of frequencies taken to be the spectral width.

By way of illustration, Fig. 2.15 is the energy spectral density of a rectangular pulse whose spectral width was previously claimed to be $|f| < 1/\tau$. The energy in that band is the shaded area in the figure, i.e.,

$$\int_{-1/\tau}^{1/\tau} |V(f)|^2\, df = \int_{-1/\tau}^{1/\tau} (A\tau)^2 \operatorname{sinc}^2 f\tau\, df = 0.92A^2\tau$$

whose evaluation entails numerical methods. But the total signal energy is $E = \int_{-\infty}^{\infty} |v(t)|^2\, dt = A^2\tau$ (found by inspection!) so the asserted spectral width encompasses more than 90 percent of the total energy.

Transform Theorems

Given below are some of the many other theorems associated with Fourier transforms. They are included not just as manipulation exercises but for two very practical reasons. First, the theorems are invaluable when interpreting spectra, for they express relationships between time-domain and frequency-domain operations. Second, one can build up an extensive catalog of transform pairs by applying the theorems to known pairs—and such a catalog will be useful as we seek new signal models.

In stating the theorems, we indicate a signal and its transform (or spectrum) by lowercase and uppercase letters, e.g., $V(f) = \mathscr{F}[v(t)]$ and $v(t) = \mathscr{F}^{-1}[V(f)]$. This is also denoted more compactly by $v(t) \leftrightarrow V(f)$. Table A at the back of the book lists the theorems and transform pairs covered here, plus a few others.

Linearity (Superposition)

For the constants α and β

$$\alpha v(t) + \beta w(t) \leftrightarrow \alpha V(f) + \beta W(f) \qquad (15)$$

This theorem simply states that linear combinations in the time domain become linear combinations in the frequency domain. Although proof of the theorem is trivial, its importance cannot be overemphasized. From a practical viewpoint Eq. (15) greatly facilitates spectral analysis when the signal in question is a linear combination of functions whose individual spectra are known. From a theoretical viewpoint it underscores the applicability of the Fourier transform for the study of linear systems.

Time Delay

If a signal $v(t)$ is delayed in time by t_d seconds, producing the new signal $v(t - t_d)$ the spectrum is modified by a linear phase shift of slope $-2\pi t_d$, that is,

$$v(t - t_d) \leftrightarrow V(f)e^{-j\omega t_d} \qquad (16)$$

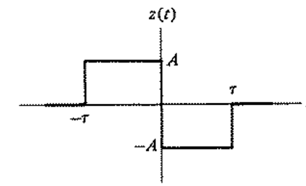


FIGURE 2.16

Translation of a signal in time thus changes the spectral phase but not the amplitude. Note that if t_d is a negative number, the signal is *advanced* in time and the phase shift has *positive* slope. Since time advancement is a physical impossibility, we conclude that in actual signal processing the spectral phase will have negative slope, though not necessarily linear.

Proof of this theorem is accomplished by change of variable $\lambda = t - t_d$ in the transform integral. (Observe that time is indeed a dummy variable in the direct transform, just as frequency is a dummy variable in the inverse transform.) The change-of-variable technique is basic to the proof of most transform theorems, and is demonstrated here:

$$\begin{aligned}\mathscr{F}[v(t - t_d)] &= \int_{-\infty}^{\infty} v(t - t_d)e^{-j\omega t}\, dt \\ &= \int_{-\infty}^{\infty} v(\lambda)e^{-j\omega(\lambda + t_d)}\, d\lambda \\ &= \left[\int_{-\infty}^{\infty} v(\lambda)e^{-j\omega\lambda}\, d\lambda\right] e^{-j\omega t_d}\end{aligned}$$

The integral in brackets is just $V(f)$, so $\mathscr{F}[v(t - t_d)] = V(f)e^{-j\omega t_d}$.

EXERCISE 2.9 The signal in Fig. 2.16 can be written as

$$z(t) = A\Pi\left(\frac{t + t_d}{\tau}\right) - A\Pi\left(\frac{t - t_d}{\tau}\right) \qquad t_d = \frac{\tau}{2} \qquad (17a)$$

Apply the linearity and time-delay theorems to the results of Example 2.4 to obtain

$$Z(f) = j2A\tau \operatorname{sinc} f\tau \sin \pi f\tau \qquad (17b)$$

Then sketch the amplitude spectrum and compare with Fig. 2.14.

Scale Change

Time delay is equivalent to translation of the time origin. Another geometric operation is scale change, in which the time axis is expanded, compressed, or reversed. Thus, $v(at)$ is a compressed version of $v(t)$ when a is positive and greater than 1.

Similarly, if a is negative and less than 1, $v(at)$ is the expanded image of $v(t)$ reversed in time. Such operations may occur in playback of recorded signals, for example.

The scale-change theorem says that

$$v(at) \leftrightarrow \frac{1}{|a|} V\left(\frac{f}{a}\right) \tag{18}$$

which formally expresses the property of reciprocal spreading encountered in Example 2.4; for if the signal is compressed in time by the factor a, its spectrum is expanded in frequency by $1/a$, and conversely. The theorem is proved by change of variables, considering positive and negative values of a separately.

Duality

The powerful concept of duality is well known in circuit analysis. In spectral analysis there is also duality, a duality between the time and frequency domains that stems from the similarity of the Fourier transform integrals. The duality theorem says that if

$$v(t) \leftrightarrow V(f)$$

then the transform of the *time* function $V(t)$ is

$$\mathscr{F}[V(t)] = v(-f) \tag{19}$$

as proved by interchanging t and f in the Fourier transform integrals.

In the form of Eq. (19) this theorem is rather abstract, and it is difficult to visualize its use in generating new transform pairs. The following example should help to clarify the procedure.

Example 2.6 Sinc Pulse

Consider the signal $z(t) = A \operatorname{sinc} 2Wt$, a sinc function in time. (Although the idea of a sinc pulse may seem strange at first, it plays a major role in the study of digital data transmission.) Recalling the transform pair of Example 2.4, $A\Pi(t/\tau) \leftrightarrow A\tau \operatorname{sinc} f\tau$, we apply duality by writing $z(t)$ in the form

$$z(t) = V(t) = \frac{A}{2W} 2W \operatorname{sinc} t2W$$

so

$$Z(f) = v(-f) = \frac{A}{2W} \Pi\left(\frac{-f}{2W}\right)$$

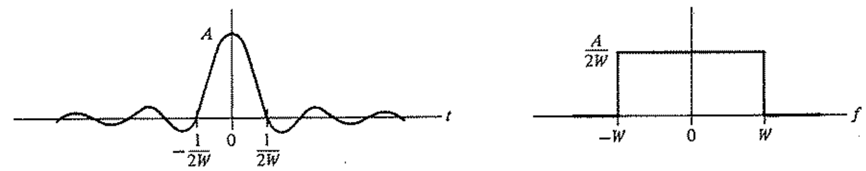


FIGURE 2.17
A sinc pulse and its bandlimited spectrum.

Because the rectangular function has even symmetry, $\Pi(-f/2W) = \Pi(f/2W)$, and we have derived the new transform pair

$$A \operatorname{sinc} 2Wt \leftrightarrow \frac{A}{2W} \Pi\left(\frac{f}{2W}\right) \tag{20}$$

which is shown in Fig. 2.17. As can be seen, the spectrum of a sinc pulse has clearly defined spectral width W. In fact, the spectrum is zero for $|f| > W$, and the signal is said to be *bandlimited* in W. This is our first encounter with a signal that is strictly bandlimited in frequency. Note that the signal itself is only asymptotically limited in time. ////

Frequency Translation (Modulation)

Duality can be used to generate transform theorems as well as transform pairs. For example, a dual of the time-delay theorem (16) is

$$v(t)e^{j\omega_c t} \leftrightarrow V(f - f_c) \qquad \omega_c = 2\pi f_c \tag{21}$$

We designate this as *frequency translation* or *complex modulation*, since multiplying a time function by $e^{j\omega_c t}$ causes its spectrum to be translated in frequency by $+f_c$.

To see the effects of frequency translation, let $v(t)$ have the bandlimited spectrum of Fig. 2.18*a*, where the amplitude and phase are plotted on the same axes using solid

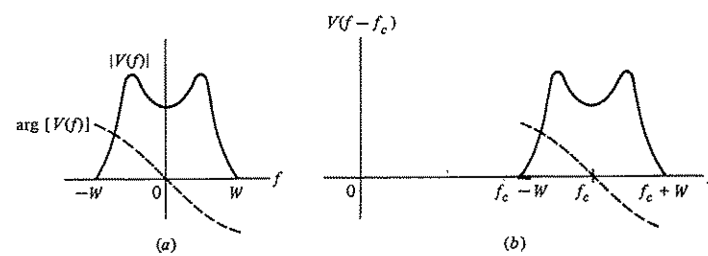


FIGURE 2.18

and broken lines, respectively. Inspection of the translated spectrum $V(f-f_c)$ in Fig. 2.18*b* reveals the following:

1 The significant components are concentrated around the frequency f_c.
2 Though $V(f)$ was bandlimited in W, $V(f-f_c)$ has a spectral width of $2W$. Translation has therefore doubled spectral width. Stated another way, the negative-frequency portion of $V(f)$ now appears at positive frequencies.
3 $V(f-f_c)$ is not hermitian but does have symmetry with respect to translated origin at $f=f_c$.

These considerations may appear somewhat academic in view of the fact that $v(t)e^{j\omega_c t}$ is not a real time function and cannot occur as a communication signal. However, signals of the form $v(t)\cos(\omega_c t+\theta)$ are common—in fact, they are the basis of carrier modulation—and by direct extension of Eq. (21) we have the following *modulation theorem:*

$$v(t)\cos(\omega_c t+\theta)\leftrightarrow\frac{e^{j\theta}}{2}V(f-f_c)+\frac{e^{-j\theta}}{2}V(f+f_c) \qquad (22)$$

In words, multiplying a signal by a sinusoid translates its spectrum *up and down* in frequency by f_c. All the comments about complex modulation also apply here. In addition, the resulting spectrum is hermitian, which it must be if $v(t)\cos(\omega_c t+\theta)$ is a real function of time. The theorem is easily proved with the aid of Euler's theorem and Eq. (21).

Example 2.7 RF Pulse

Consider the finite-duration sinusoid of Fig. 2.19*a*, sometimes referred to as an RF pulse when f_c falls in the radio-frequency band. Since

$$z(t)=A\Pi\left(\frac{t}{\tau}\right)\cos\omega_c t \qquad (23a)$$

we have immediately

$$Z(f)=\frac{A\tau}{2}\operatorname{sinc}(f-f_c)\tau+\frac{A\tau}{2}\operatorname{sinc}(f+f_c)\tau \qquad (23b)$$

by setting $v(t)=A\Pi(t/\tau)$ and $V(f)=A\tau\operatorname{sinc} f\tau$ in Eq. (22). The resulting amplitude spectrum is sketched in Fig. 2.19*b* for the case of $f_c\gg 1/\tau$ so the two translated sinc functions have negligible overlap.

Because this is a sinusoid of finite duration, its spectrum is continuous and contains more than just the frequencies $f=\pm f_c$. Those other frequencies stem from the fact that $z(t)=0$ for $|t|>\tau/2$, and the smaller τ is, the larger the spectral spread around $\pm f_c$—reciprocal spreading, again. On the other hand, had we been dealing

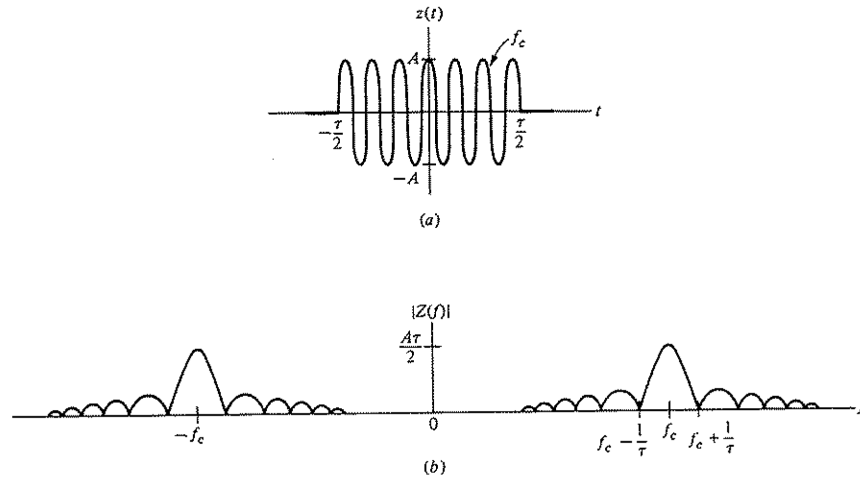


FIGURE 2.19
An RF pulse and its amplitude spectrum, $f_c\gg 1/\tau$.

with a sinusoid of *infinite* duration, the frequency-domain representation would be a two-sided *line* spectrum containing only the discrete frequencies $\pm f_c$. ////

Differentiation and Integration

Certain processing techniques involve differentiating or integrating a signal. The frequency-domain effects of these operations are indicated in the theorems below. A word of caution, however: The theorems should not be applied before checking to make sure that the differentiated or integrated signal is Fourier-transformable, i.e., has well-defined energy. And the fact that $v(t)$ has finite energy is not a guarantee that the same is true for its derivative or integral.

To derive the differentiation theorem, we replace $v(t)$ by the inverse transform integral and interchange the order of operations, as follows:

$$\frac{d}{dt}v(t)=\frac{d}{dt}\left[\int_{-\infty}^{\infty}V(f)e^{j2\pi ft}\,df\right]$$
$$=\int_{-\infty}^{\infty}V(f)\left(\frac{d}{dt}e^{j2\pi ft}\right)df$$
$$=\int_{-\infty}^{\infty}\underbrace{[j2\pi fV(f)]}_{\mathscr{F}\left[\frac{dv(t)}{dt}\right]}e^{j2\pi ft}\,df$$

Referring back to the Fourier integral theorem (3) reveals that the bracketed term must be $\mathscr{F}[dv(t)/dt]$, so

$$\frac{d}{dt}v(t) \leftrightarrow j2\pi f V(f)$$

and by iteration

$$\frac{d^n}{dt^n}v(t) \leftrightarrow (j2\pi f)^n V(f) \tag{24}$$

which is the *differentiation theorem*.

Now suppose we generate another time function from $v(t)$ by carrying out the operation $\int_{-\infty}^{t} v(\lambda)\,d\lambda$, where the dummy variable λ is required since the independent variable t is the upper limit of integration. The *integration theorem* says that

$$\int_{-\infty}^{t} v(\lambda)\,d\lambda \leftrightarrow \frac{1}{j2\pi f}V(f) \tag{25}$$

whose proof involves the same method used above. One can also generalize Eq. (25) to multiple integration but the notation is cumbersome.

Inspecting these theorems, we can say that differentiation enhances the high-frequency components of a signal while integration suppresses high-frequency components. Spectral interpretation thus agrees with the time-domain viewpoint that differentiation accentuates time variations while integration smoothes them out.

Example 2.8 Triangular Pulse

To illustrate the integration theorem—and obtain yet another useful transform pair—let us integrate the signal $z(t)$ in Fig. 2.16 and divide it by the constant τ. This produces

$$w(t) = \frac{1}{\tau}\int_{-\infty}^{t} z(\lambda)\,d\lambda = \begin{cases} A\left(1 - \dfrac{|t|}{\tau}\right) & |t| < \tau \\ 0 & |t| > \tau \end{cases}$$

a *triangular* pulse shape shown in Fig. 2.20*a*. (The reader is strongly encouraged to check this result using the graphical interpretation of integration.) Then, taking $Z(f)$ from Eq. (17*b*), we have

$$W(f) = \frac{1}{\tau}\frac{1}{j2\pi f}Z(f) = \frac{j2A\tau \operatorname{sinc} f\tau \sin \pi f\tau}{j2\pi f\tau} = A\tau \operatorname{sinc}^2 f\tau$$

as sketched in Fig. 2.20*b*. Comparing this spectrum with Fig. 2.14 shows that the triangular pulse has less high-frequency content than a rectangular pulse with amplitude A and duration τ, even though they both have the same area. The difference is traced to the fact that the triangular pulse is spread over 2τ seconds and does not have the sharp, stepwise time variations of the rectangular shape.

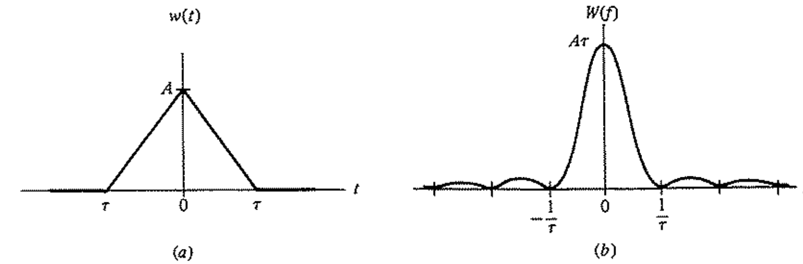


FIGURE 2.20
A triangular pulse and its spectrum.

This transform pair can be written more compactly by defining the *triangular function*

$$\Lambda\left(\frac{t}{\tau}\right) \triangleq \begin{cases} 1 - \dfrac{|t|}{\tau} & |t| < \tau \\ 0 & |t| > \tau \end{cases} \tag{26}$$

Then $w(t) = A\Lambda(t/\tau)$ and

$$A\Lambda\left(\frac{t}{\tau}\right) \leftrightarrow A\tau \operatorname{sinc}^2 f\tau \tag{27}$$

It so happens that triangular functions can be generated from rectangular functions by another mathematical operation, namely, convolution. And convolution happens to be the next major subject on the agenda, so we will take another look at this example soon. ////

2.4 CONVOLUTION AND IMPULSES

The mathematical operation known as convolution ranks high among the analytic tools used by communication engineers. For one reason, it is a good model of the physical processes that go on in a linear system; for another, it helps to further our understanding of the relationships between the time domain and the frequency domain. In both cases, convolution goes hand in hand with that curious engineering fiction called the impulse or delta function. This section deals with these concepts as they relate to signals; they are applied to linear systems in the next section.

Convolution Integral

The convolution of two functions of the same variable, say $v(t)$ and $w(t)$, is defined as

$$v * w(t) \triangleq \int_{-\infty}^{\infty} v(\lambda)w(t-\lambda)\,d\lambda \qquad (1)$$

where $v * w(t)$ merely stands for the operation on the right-hand side of Eq. (1) and the asterisk (*) has nothing to do with complex conjugation. Equation (1) is the *convolution integral*, often denoted by $v * w$ when the independent variable is unambiguous. At other times the notation $[v(t)] * [w(t)]$ is necessary for clarity. Note carefully that the independent variable here is t, the same as the independent variable of the functions being convolved; the integration is always performed with respect to a dummy variable (such as λ) and t is a constant insofar as the integration is concerned.

Calculating $v * w(t)$ is no more difficult than ordinary integration when the two functions are continuous for all t. Often, however, one or both of the functions is defined in a piecewise fashion, and the graphical interpretation of convolution illustrated in Fig. 2.21 becomes especially helpful. Figures 2.21*a* and *b* are the functions involved here, but the integrand in Eq. (1) is $v(\lambda)w(t-\lambda)$. Of course, $v(\lambda)$ is nothing more than $v(t)$ with t replaced by λ, Fig. 2.21*c*. But $w(t-\lambda)$ as a function of λ must be obtained by two steps: first, $w(-\lambda)$ is $w(t)$ reversed in time with t replaced by λ; then, for a given value of t, sliding $w(-\lambda)$ to the right t units yields $w(t-\lambda)$. Figure 2.21*d* shows $w(t-\lambda)$ for the case of $t = t_1 > 0$, illustrating that the value of t always equals the distance from the origin of $v(\lambda)$ to the shifted origin of $w(-\lambda)$. Finally, $v(\lambda)$ and $w(t-\lambda)$ are multiplied and the area of the product equals $v * w(t)$ for that particular value of t, Fig. 2.21*e*.

As $v * w(t)$ is evaluated for $-\infty < t < \infty$, the plot of $w(t-\lambda)$ moves from the left to right with respect to $v(\lambda)$, and the actual form of the convolution integration may change depending on the value of t. In Fig. 2.21, for instance, it follows that

$$v * w(t) = 0 \qquad t < 0$$

since $w(t-\lambda)$ does not overlap $v(\lambda)$ and the area of the product is zero. Similarly,

$$v * w(t) = \begin{cases} \displaystyle\int_0^t v(\lambda)w(t-\lambda)\,d\lambda & 0 < t < T \\[2ex] \displaystyle\int_{t-T}^t v(\lambda)w(t-\lambda)\,d\lambda & t > T \end{cases}$$

since $v(\lambda) = 0$ for $\lambda < 0$ and $w(t-\lambda) = 0$ for $\lambda < t - T$ and $\lambda > t$. Note in these cases that t appears as a limit of integration. Simple sketches of the functions involved help one discover these different cases.

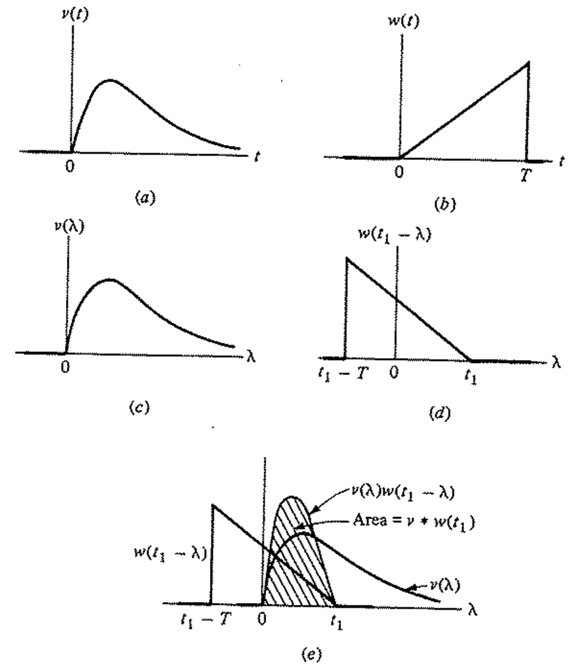


FIGURE 2.21
The graphical interpretation of convolution.

Further study of Fig. 2.21 should reveal that $v * w(t) = w * v(t)$, i.e., we get the same result by reversing v and sliding it past w. This property and several other convolution properties are listed below for reference.

$$v * w = w * v \qquad (2a)$$

$$v * (w * z) = (v * w) * z \qquad (2b)$$

$$(\alpha v + \beta w) * z = \alpha(v * z) + \beta(w * z) \qquad (2c)$$

$$\frac{d}{dt}(v * w) = v * \frac{dw}{dt} = \frac{dv}{dt} * w \qquad (3)$$

Example 2.9 Convolution of Rectangular Pulses

The convolution of two rectangular pulses, Fig. 2.22*a*, is relatively simple using the graphical interpretation, and the problem breaks up into three cases: $|t| > (\tau_1 + \tau_2)/2$, $(\tau_1 - \tau_2)/2 < |t| < (\tau_1 + \tau_2)/2$, and $|t| < (\tau_1 - \tau_2)/2$, assuming $\tau_1 \geq \tau_2$. The result is

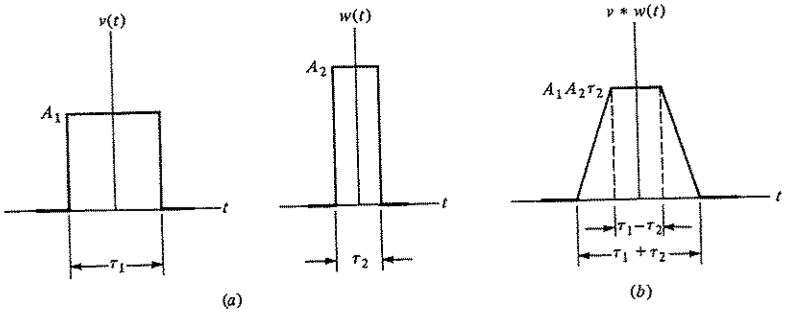


FIGURE 2.22
Convolution of rectangular pulses.

a *trapezoidal* function, Fig. 2.22*b*, which degenerates into a triangular function in the special case of $\tau_1 = \tau_2$. ////

EXERCISE 2.10 By carrying out all the details, confirm the result asserted in Example 2.9.

Convolution Theorems

Having defined convolution, we now give the theorems pertaining to convolution and transforms, which are two in number, namely,

$$v * w(t) \leftrightarrow V(f)W(f) \tag{4}$$

$$v(t)w(t) \leftrightarrow V * W(f) \tag{5}$$

These theorems state that convolution in the time domain becomes multiplication in the frequency domain, while multiplication in the time domain becomes convolution in the frequency domain. Both of these relationships are important for future work.

The proof of Eq. (4) uses the time-delay theorem, as follows:

$$\begin{aligned}\mathscr{F}[v * w(t)] &= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} v(\lambda)w(t-\lambda)\,d\lambda\right]e^{-j\omega t}\,dt \\ &= \int_{-\infty}^{\infty} v(\lambda)\left[\int_{-\infty}^{\infty} w(t-\lambda)e^{-j\omega t}\,dt\right]d\lambda \\ &= \int_{-\infty}^{\infty} v(\lambda)[W(f)e^{-j\omega\lambda}]\,d\lambda \\ &= \left[\int_{-\infty}^{\infty} v(\lambda)e^{-j\omega\lambda}\,d\lambda\right]W(f) = V(f)W(f)\end{aligned}$$

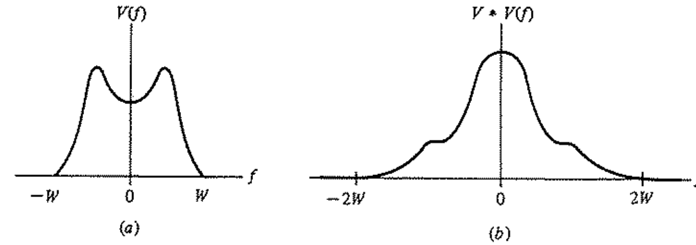


FIGURE 2.23

Equation (5) can be proved by writing out the transform of $v(t)w(t)$ and replacing $w(t)$ by the inversion integral $\mathscr{F}^{-1}[W(f)]$.

EXERCISE 2.11 Use Eq. (4) together with the results of Example 2.9 to obtain the transform of a triangular pulse, Eq. (27), Sect. 2.3.

Example 2.10 The Spectrum of $v^2(t)$

Suppose $v(t)$ is bandlimited in W, with a spectrum as shown in Fig. 2.23*a*. What then is the spectrum of $v^2(t)$? From Eq. (5) we must convolve $V(f)$ with itself, the result being something like Fig. 2.23*b*. Without any further specific knowledge of $v(t)$ we reach this important conclusion: when $v(t)$ is bandlimited in W, $v^2(t)$ is bandlimited in $2W$. (Note the difference between this operation and the modulation theorem, both of which double spectral width.) The process may be iterated for $v^3(t)$, etc., with predictable conclusions. ////

Unit Impulse

Previous examples have demonstrated that convolution is a *smoothing* operation; i.e., the result is "smoother" and "longer" or "wider" than either of the functions involved. But there is one notable exception to this rule, namely, when one of the functions is a *unit impulse* or *Dirac delta function* $\delta(t)$; in that case, providing $v(t)$ is continuous,

$$[v(t)] * [\delta(t)] = v(t) \tag{6}$$

so convolving with an impulse merely reproduces the other function in its entirety.

Actually, $\delta(t)$ is not a function in the strict mathematical sense; rather, it is a member of that special class known as *generalized functions* or *distributions*. And

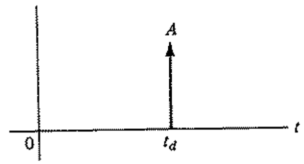


FIGURE 2.24
The graphical representation of $A\delta(t - t_d)$.

because it is not a function, the unit impulse is defined by an assignment rule or process instead of a conventional equation. Specifically, given any ordinary function $v(t)$ that is continuous at $t = 0$, $\delta(t)$ is defined by

$$\int_{t_1}^{t_2} v(t)\,\delta(t)\,dt = \begin{cases} v(0) & t_1 < 0 < t_2 \\ 0 & \text{otherwise} \end{cases} \qquad (7)$$

a rule that assigns a number—either $v(0)$ or 0—to the process on the left-hand side.

Taking $v(t) = 1$ for all t, it follows from Eq. (7) that

$$\int_{-\infty}^{\infty} \delta(t)\,dt = \int_{-\epsilon}^{\epsilon} \delta(t)\,dt = 1 \qquad (8a)$$

which may be interpreted by saying that $\delta(t)$ has *unit area* concentrated at the discrete point $t = 0$ and no net area elsewhere. Carrying this argument further suggests that

$$\delta(t) = 0 \qquad t \neq 0 \qquad (8b)$$

Equations (8*a*) and (8*b*) are the more familiar "definitions" of the impulse, and lead to the common graphical representation. For instance, the picture of $A\delta(t - t_d)$ is shown in Fig. 2.24, where the letter A next to the arrowhead means that $A\delta(t - t_d)$ has area or weight A located at $t = t_d$. It should be noted that Eq. (8*b*) is not an assignment rule, and distribution theory, strictly interpreted, does not specify values for the impulse other than in the integral sense. It is, however, consistent with Eq. (7) and helps us visualize impulse properties under the operation of integration.

Two of the most important integration properties are

$$\int_{-\infty}^{\infty} v(t)\,\delta(t - t_d)\,dt = v(t_d) \qquad (9)$$

$$[v(t)] * [\delta(t - t_d)] = v(t - t_d) \qquad (10)$$

both of which can be derived from Eq. (7). Equation (9) is called the *sampling* property since the indicated operation picks out or samples the values of $v(t)$ at $t = t_d$ where $\delta(t - t_d)$ is "located." On the other hand, convolving $v(t)$ with $\delta(t - t_d)$ is a *replication* property since, according to Eq. (10), it reproduces the entire function $v(t)$ displaced by t_d units. The difference between Eqs. (9) and (10) should be clearly

understood: sampling picks out a particular value, i.e., a number, while convolving repeats the function completely.

By definition, the impulse has no mathematical or physical meaning unless it appears under integration. Even so, it is convenient to state three nonintegral relations as simplifications that can be made before integration since they are consistent with what would happen after integration. Specifically, in view of the sampling property (9), we can just as well replace $v(t)$ by $v(t_d)$, so that

$$v(t)\,\delta(t - t_d) = v(t_d)\,\delta(t - t_d) \qquad (11)$$

whose justification stems from integrating both sides over $-\infty < t < \infty$. Similarly one can justify the scale-change relationship

$$\delta(at) = \frac{1}{|a|}\,\delta(t) \qquad a \neq 0 \qquad (12)$$

which says that, relative to the independent variable t, $\delta(at)$ is an impulse having weight $1/|a|$. The special case of $a = -1$ indicates the even-symmetry property $\delta(t) = \delta(-t)$. Finally, relating the unit impulse to the unit step $u(t)$ defined in Eq. (13), Sect. 2.3, we see from Eq. (7) that

$$\int_{-\infty}^{t} \delta(\lambda)\,d\lambda = \begin{cases} 1 & t > 0 \\ 0 & t < 0 \end{cases}$$
$$= u(t)$$

Differentiating both sides then yields

$$\delta(t) = \frac{du(t)}{dt} \qquad (13)$$

which is not a definition of $\delta(t)$ but a consequence of the assignment rule (7).

Although an impulse does not exist physically, there are numerous conventional functions that have all the properties of $\delta(t)$ in the limit as some parameter ϵ goes to zero. In particular, if the function $\delta_\epsilon(t)$ is such that

$$\lim_{\epsilon \to 0} \int_{-\infty}^{\infty} v(t)\,\delta_\epsilon(t)\,dt = v(0) \qquad (14a)$$

then we say that

$$\lim_{\epsilon \to 0} \delta_\epsilon(t) = \delta(t) \qquad (14b)$$

Two functions satisfying Eq. (14*a*) are

$$\delta_\epsilon(t) = \frac{1}{\epsilon}\,\Pi\left(\frac{t}{\epsilon}\right) \qquad (15)$$

$$\delta_\epsilon(t) = \frac{1}{\epsilon}\,\text{sinc}\,\frac{t}{\epsilon} \qquad (16)$$

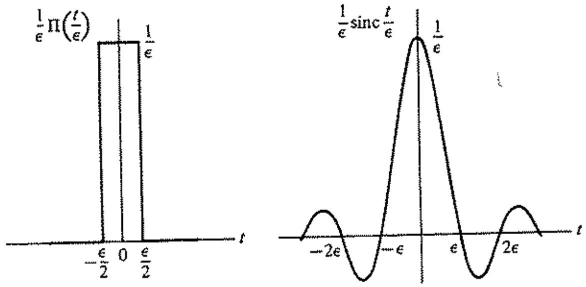


FIGURE 2.25
Two functions that become impulses as $\epsilon \to 0$.

which are plotted in Fig. 2.25. One can easily show that Eq. (15) satisfies Eq. (14*a*) by expanding $v(t)$ in a Maclaurin series prior to integrating. An argument for Eq. (16) will be given shortly when we consider impulses and transforms.

EXERCISE 2.12 Use Eq. (16) to prove that

$$\lim_{F\to\infty} \int_{-F}^{F} e^{\pm j2\pi ft}\, df = \delta(t) \qquad (17)$$

Impulses in Frequency

Sections 2.2 and 2.3 drew the distinction between periodic power signals and nonperiodic energy signals, one class being described by line spectra, the other by continuous spectra. This means that we have something of a quandary if a signal consists of periodic and nonperiodic parts since different frequency-domain representations would be required. This quandary is solved by allowing impulses in the frequency domain as the representation of discrete frequency components. Such impulses are derived by limiting operations on conventional Fourier transform pairs, and may be dubbed *transforms in the limit*. Note, however, that the corresponding time functions are power signals having infinite or undefined energy, and the concept of energy spectral density no longer applies.

As a starting point of this discussion, consider the signal $v(t) = A$, a constant for all time. Referring to Fig. 2.17*a*, we let $v(t)$ be a *sinc pulse* with $W \to 0$, i.e.,

$$v(t) = \lim_{W\to 0} A \text{ sinc } 2Wt = A$$

But we already have the transform pair $A \text{ sinc } 2Wt \leftrightarrow (A/2W)\Pi(f/2W)$, so

$$\mathscr{F}[v(t)] = \lim_{W\to 0} \frac{A}{2W}\Pi\left(\frac{f}{2W}\right) = A\,\delta(f)$$

where we have invoked Eq. (15) with t replaced by f and ϵ replaced by $2W$. Therefore,

$$A \leftrightarrow A\delta(f) \qquad (18)$$

and the spectrum of a constant in the time domain is an impulse in the frequency domain at $f = 0$. This result agrees with intuition in that a constant signal has no time variation and its spectral content ought to be confined to $f = 0$. The impulsive form results simply because we use integration to return to the time domain, via the inverse transform, and an impulse is required to concentrate nonzero area at a discrete point in frequency. Checking this argument mathematically gives

$$\mathscr{F}^{-1}[A\,\delta(f)] = \int_{-\infty}^{\infty} A\,\delta(f)e^{j2\pi ft}\,dt = Ae^{j2\pi ft}\Big|_{f=0} = A$$

which justifies Eq. (18) for our purposes. Note that the impulse has been integrated to obtain a physical quantity, namely, the signal $v(t) = A$.

As an alternate to the above procedure for deriving Eq. (18), we could have begun with a rectangular pulse, $A\Pi(t/\tau)$, and let $\tau \to \infty$ to get a constant for all time. Then, since $\mathscr{F}[A\Pi(t/\tau)] = A\tau \text{ sinc } f\tau$, agreement with Eq. (18) requires that

$$\lim_{\tau\to\infty} A\tau \text{ sinc } f\tau = A\delta(f)$$

And this supports the earlier assertion Eq. (16) that a sinc function becomes an impulse under appropriate limiting conditions.

To generalize Eq. (18), direct application of the frequency-translation and modulation theorems yields

$$Ae^{j\omega_c t} \leftrightarrow A\delta(f - f_c) \qquad (19)$$

$$A\cos(\omega_c t + \theta) \leftrightarrow \frac{Ae^{j\theta}}{2}\,\delta(f - f_c) + \frac{Ae^{-j\theta}}{2}\,\delta(f + f_c) \qquad (20)$$

Thus, the spectrum of a single phasor is an impulse at $f = f_c$ while the spectrum of a sinusoid has two impulses, Fig. 2.26. Going even further in this direction, if $v(t)$ is an arbitrary periodic signal whose exponential Fourier series is

$$v(t) = \sum_{n=-\infty}^{\infty} c_v(nf_0)e^{j2\pi nf_0 t} \qquad (21a)$$

then its Fourier transform is

$$V(f) = \sum_{n=-\infty}^{\infty} c_v(nf_0)\,\delta(f - nf_0) \qquad (21b)$$

where superposition allows one to transform the sum term by term.

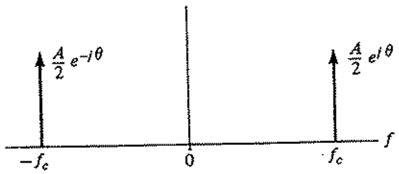


FIGURE 2.26
The spectrum of $A \cos (\omega_c t + \theta)$.

By now it should be obvious from Eqs. (18) to (21) that any two-sided line spectrum can be converted to a "continuous" spectrum using this rule: convert the spectral lines to impulses whose weights equal the line heights. The phase portion of the line spectrum is absorbed by letting the impulse weights be complex numbers, e.g., the weights in Eq. (21*b*) are $c_v(nf_0) = |c_v(nf_0)|e^{j \arg [c_v(nf_0)]}$. Hence, with the aid of transforms in the limit, we can represent both periodic and nonperiodic signals by continuous spectra. In addition, the transform theorems developed in Sect. 2.3 can now be applied to periodic signals. That strange beast the impulse function thereby emerges as a key to unifying spectral analysis.

But one may well ask: What is the difference between the line spectrum and the "continuous" spectrum of a periodic signal? Obviously there can be no physical difference; the difference lies in the mathematical conventions. To return to the time domain from the line spectrum, we sum the phasors which the lines represent. To return to the time domain from the continuous spectrum, we integrate the impulses to get phasors.

EXERCISE 2.13 Prove Eq. (20) by carrying out the inverse transform of the right-hand side.

Impulses in Time

The time-domain impulse may seem a trifle farfetched as a signal model, but the next section will show conditions where it is both reasonable and highly useful. Here we are concerned with the transform

$$A\delta(t) \leftrightarrow A \qquad (22)$$

which is derived by Fourier transformation using Eq. (17) to evaluate the integral. Since the transform of the time impulse has *constant amplitude*, its spectrum contains all frequencies in equal proportion.

The reader may have observed that Eq. (22) is the dual of $A \leftrightarrow A\delta(f)$. This dual relationship has its roots in reciprocal spreading, Eqs. (18) and (22) being the two extremes; i.e., a constant signal of infinite duration has "zero" spectral width, whereas an impulse in time has "zero" duration and infinite spectral width.

Applying the time-delay theorem to Eq. (22) yields the more general pair

$$A\delta(t - t_d) \leftrightarrow Ae^{-j\omega t_d} \qquad (23)$$

Problem 2.34 outlines how this pair, together with the differentiation theorem, provides a shortcut method for finding certain other transforms.

2.5 SYSTEM RESPONSE AND FILTERS

Let us return to the input-output or system response problem as posed at the beginning of the chapter, save that now we allow the input $x(t)$ to be more or less arbitrary. We will still assume that the system is linear, time-invariant, and asymptotically stable, and we will add the constraint that there is no stored energy in the system when the input is applied. The system's response $y(t)$ will be formulated using both time-domain and frequency-domain analysis.

Impulse Response and Time-Domain Analysis

In linear system theory, the *impulse response* $h(t)$ of a system is defined as the output that results when the input is a unit impulse, i.e.,

$$h(t) \triangleq y(t) \qquad \text{when } x(t) = \delta(t) \qquad (1)$$

The response to an arbitrary input $x(t)$ is then found by convolving $h(t)$ with $x(t)$, so

$$y(t) = h * x(t) = \int_{-\infty}^{\infty} h(\lambda)x(t - \lambda)\, d\lambda \qquad (2)$$

Often called the *superposition integral*, Eq. (2) is the basis of time-domain system analysis. This method therefore requires knowing the impulse response as well as the ability to carry out the convolution.

Several techniques are available for finding the impulse response of a system, given its mathematical model. If the model is a simple block diagram without feedback loops, $h(t)$ usually can be found from inspection by invoking the definition (1). Other times it is easier to calculate the *step response* $y_u(t)$, defined by

$$y_u(t) \triangleq y(t) \qquad \text{when } x(t) = u(t) \qquad (3a)$$

in which case

$$h(t) = \frac{dy_u(t)}{dt} \qquad (3b)$$

This follows since $y_u(t) = [h(t)] * [u(t)]$ so $dy_u(t)/dt = [h(t)] * [du(t)/dt] = [h(t)] * [\delta(t)] = h(t)$, the pertinent relations being Eqs. (3), (6), and (13), Sect. 2.4. When these methods do not work, the problem is probably a candidate for frequency-domain analysis.

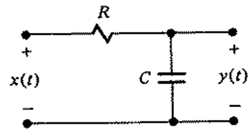


FIGURE 2.27
RC lowpass filter.

Example 2.11 Step, Impulse, and Pulse Response of an RC Lowpass Filter

Again consider the RC lowpass filter, Fig. 2.27. If $x(t)$ is a unit step, Fig. 2.28*a*, the step response is well known to be

$$y_u(t) = (1 - e^{-t/RC})u(t) \tag{4a}$$

as plotted in Fig. 2.28*b*. Differentiation then yields Fig. 2.28*c*, namely,

$$h(t) = \frac{dy_u(t)}{dt} = \frac{1}{RC}e^{-t/RC}u(t) \tag{4b}$$

Now consider the response to a rectangular pulse of duration τ starting at $t = 0$, i.e., $x(t) = A\Pi[(t - \tau/2)/\tau]$. Putting $x(t)$ and $h(t)$ in Eq. (2) gives, with the help of the graphical interpretation of convolution,

$$y(t) = \begin{cases} 0 & t < 0 \\ A(1 - e^{-t/RC}) & 0 < t < \tau \\ A(1 - e^{-\tau/RC})e^{-(t-\tau)/RC} & t > \tau \end{cases} \tag{5}$$

which is sketched in Fig. 2.29 for two values of τ/RC. ////

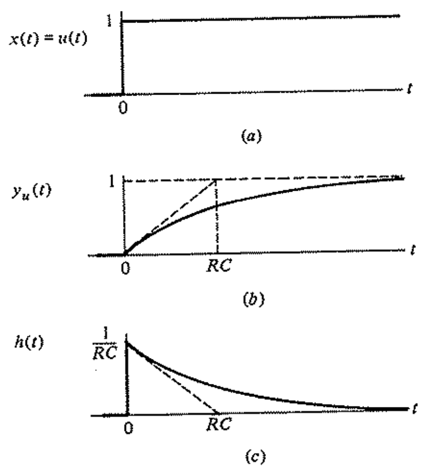


FIGURE 2.28
Waveforms for RC lowpass filter. (*a*) Unit step input; (*b*) step response $y_u(t)$; impulse response $h(t) = dy_u(t)/dt$.

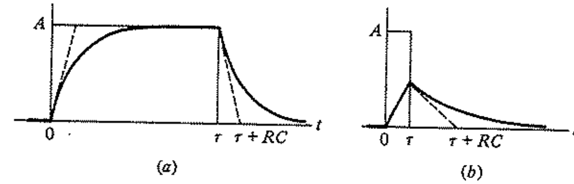


FIGURE 2.29
Rectangular pulse response of an RC lowpass filter. (*a*) $\tau/RC \gg 1$; (*b*) $\tau/RC \ll 1$.

Transfer Function and Frequency-Domain Analysis

Linking time-domain analysis to the frequency domain, let the input in the superposition integral (2) be $x(t) = e^{j2\pi ft}$ for $-\infty < t < \infty$. Then,

$$\begin{aligned} y(t) &= \int_{-\infty}^{\infty} h(\lambda)e^{j2\pi f(t-\lambda)}\,d\lambda \\ &= \left[\int_{-\infty}^{\infty} h(\lambda)e^{-j2\pi f\lambda}\,d\lambda\right]e^{j2\pi ft} \end{aligned} \tag{6}$$

and the expression in brackets is recognized as a Fourier transform integral. To identify that integral, recall that the transfer function $H(f)$ of a system was previously defined such that, when $x(t) = e^{j2\pi ft}$, $y(t) = H(f)e^{j2\pi ft}$—but Eq. (6) has precisely this form if we take

$$H(f) = \mathscr{F}[h(t)] = \int_{-\infty}^{\infty} h(t)e^{-j2\pi ft}\,dt \tag{7}$$

Therefore, the impulse response and transfer function of a given system constitute a Fourier transform pair, and all the properties of $H(f)$ stated earlier can be derived from Eq. (7).

Equation (7) also explains why the Fourier transform of the exponential pulse, Example 2.5, Sect. 2.3, turned out to have the same frequency dependence as the transfer function of an RC lowpass filter, Example 2.1, Sect. 2.1. In hindsight, this demonstrates the $h(t) \leftrightarrow H(f)$ pair, since Example 2.11 showed that the impulse response of an RC lowpass filter is an exponential pulse.

As for frequency-domain analysis per se, we take the Fourier transform of the superposition integral and apply the convolution theorem, Eq. (4), Sect. 2.4. Thus

$$\mathscr{F}[y(t)] = \mathscr{F}[h * x(t)] = \mathscr{F}[h(t)]\mathscr{F}[x(t)]$$

or

$$Y(f) = H(f)X(f) \tag{8}$$

as illustrated schematically in Fig. 2.30 along with the time-domain relation. This elegantly simple equation, the basis of frequency-domain analysis, says that the

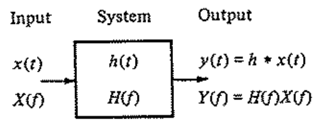


FIGURE 2.30
Input-output relations for a linear time-invariant system.

output spectrum $Y(f)$ equals the input spectrum $X(f)$ multiplied by the transfer function $H(f)$. The corresponding amplitude and phase spectra are

$$|Y(f)| = |H(f)|\,|X(f)| \qquad (9)$$
$$\arg[Y(f)] = \arg[X(f)] + \arg[H(f)]$$

which should be compared to the periodic steady-state result, Eq. (23), Sect. 2.2. Additionally, if the output $y(t)$ is an energy signal, its energy spectral density and total energy are given by

$$|Y(f)|^2 = |H(f)|^2\,|X(f)|^2 \qquad (10)$$

$$E_y = \int_{-\infty}^{\infty} |H(f)|^2\,|X(f)|^2\,df \qquad (11)$$

from Rayleigh's energy theorem.

Further interpretation of Eq. (8) is afforded if we take $x(t)$ as a unit impulse; then, since $X(f) = \mathscr{F}[\delta(t)] = 1$,

$$Y(f) = H(f) \qquad \text{when } x(t) = \delta(t) \qquad (12)$$

in agreement with the fact that the transfer function is the transform of the impulse response. Viewed from the frequency domain, the spectrum of the input signal has all frequency components in equal proportion in this case, so the output spectrum is shaped entirely by the transfer function $H(f)$.

Finally, one can return to the time domain by taking the inverse transform of Eq. (8), i.e.,

$$y(t) = \mathscr{F}^{-1}[H(f)X(f)] = \int_{-\infty}^{\infty} H(f)X(f)e^{j2\pi ft}\,df \qquad (13)$$

Contrasting Eqs. (13) and (8), it appears that the output spectrum is more easily obtained than the output time function, which indeed is the case if $H(f)$ and $X(f)$ are known. The power of frequency-domain analysis rests on the simple relationship of input and output spectra. Furthermore, an experienced communication engineer can often infer all he needs to know from the spectrum. (Much of this chapter has pointed toward making such inferences.) On the other hand, if specific details of the time function are to be investigated, the superposition integral may be easier than Eq. (13) for finding $y(t)$.

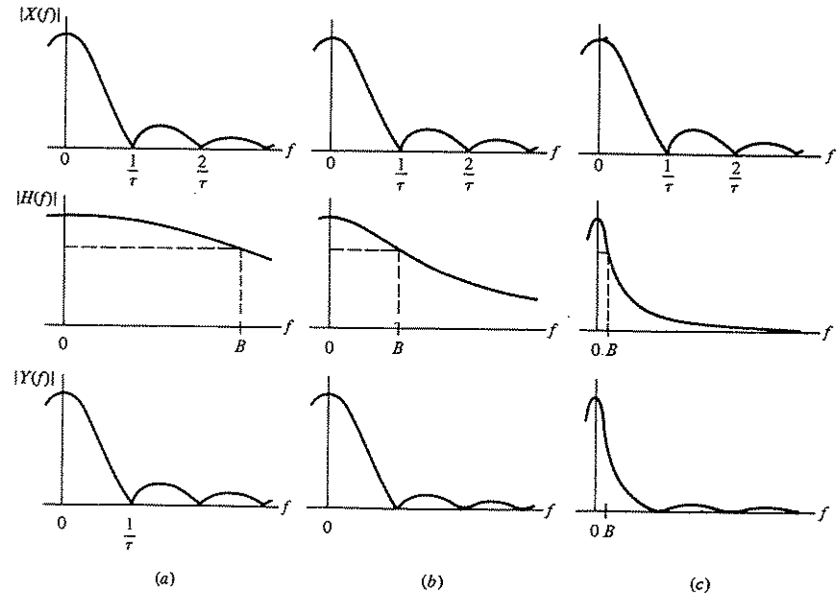


FIGURE 2.31
Frequency-domain analysis of the rectangular pulse response of an RC lowpass filter. (a) $B \gg 1/\tau$; (b) $B \approx 1/\tau$; (c) $B \ll 1/\tau$.

Example 2.12

To illustrate how far one can go just in terms of the frequency domain, suppose a rectangular pulse of duration τ is applied to an RC lowpass filter with $B = 1/2\pi RC$. Our prior studies have shown that most of the pulse's spectral content is in $|f| < 1/\tau$, while the filter responds primarily to frequencies in the range $|f| < B$. Clearly, the shape of the output spectrum—and, hence, the output signal—depends on the relative values of $1/\tau$ and B. Figure 2.31 gives plots of $|X(f)|$, $|H(f)|$, and $|Y(f)| = |H(f)|\,|X(f)|$ for the three cases $B \gg 1/\tau$, $B \approx 1/\tau$, and $B \ll 1/\tau$.

In the first case, Fig. 2.31*a*, the filter passes all the significant frequency components since $H(f) \approx 1$ for $|f| < 1/\tau$. Therefore, $Y(f) \approx X(f)$ and $y(t) \approx x(t)$, so the output time function should look very much like the input. As a direct check on that conclusion, Fig. 2.29*a* shows the actual waveforms under this condition. In the second case, Fig. 2.31*b*, the shape of the output spectrum depends on both $X(f)$ and $H(f)$ so $y(t)$ will differ substantially from $x(t)$. We then say that the output is *distorted* in the sense that it does not resemble the input, but a more precise statement entails actually going through the time-domain calculations. In the third case, Fig. 2.31*c*, the input spectrum is constant or "flat" over $|f| < B$ so $Y(f) \approx A\tau H(f)$. Since the output spectrum now depends primarily on $H(f)$, the output signal will look like the filter's *impulse response*, i.e., $y(t) \approx A\tau h(t)$. Again, this is confirmed by the time-domain result plotted in Fig. 2.29*b*.

Extrapolating this last case to other filters and other pulse shapes, we state the following rule of thumb: If the input spectrum is essentially constant over the frequency band where the filter has significant response, then the output signal is essentially the impulse response of the filter. Under such conditions, it is quite reasonable to model the input signal as being an impulse. ////

Parallel, Cascade, and Feedback Connections

More often than not, a communication system comprises many interconnected units or subsystems. When the subsystems in question are described by individual transfer functions, it is possible and desirable to lump them together and speak of the overall system transfer function. The corresponding relations are given below for two subsystems connected in parallel, cascade, and feedback. More complicated configurations can be analyzed by successive application of these basic rules. One essential assumption must be made, however, namely, that any interaction or *loading* effects have been accounted for in the individual transfer functions so that they represent the actual response of the subsystems in the context of the overall system.

Figure 2.32*a* diagrams two subsystems in *parallel*; both units have the same input and their outputs are summed to get the system's output. From superposition it follows that $Y(f) = [H_1(f) + H_2(f)]X(f)$ so the overall transfer function is

$$H(f) = H_1(f) + H_2(f) \qquad \text{Parallel connection} \qquad (14)$$

In the *cascade* connection, Fig. 2.32*b*, the output of the first unit is the input to the second, so $Y(f) = H_2(f)[H_1(f)X(f)]$ and

$$H(f) = H_1(f)H_2(f) \qquad \text{Cascade connection} \qquad (15)$$

The *feedback* connection, Fig. 2.32*c*, differs from the other two in that the output is sent back through $H_2(f)$ and subtracted from the input. Thus, $Y(f) = H_1(f)[X(f) - H_2(f)Y(f)]$ and rearranging yields $Y(f) = \{H_1(f)/[1 + H_1(f)H_2(f)]\}X(f)$ so

$$H(f) = \frac{H_1(f)}{1 + H_1(f)H_2(f)} \qquad \text{Feedback connection} \qquad (16)$$

This case is more properly termed the *negative* feedback connection as distinguished from positive feedback, where the returned signal is added to the input instead of subtracted.

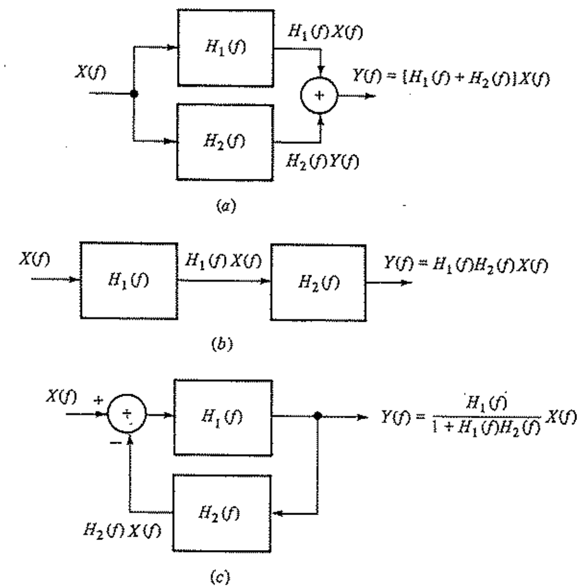


FIGURE 2.32
(*a*) Parallel connection; (*b*) cascade connection; (*c*) feedback connection.

Example 2.13 Zero-Order Hold

The zero-order hold system, Fig. 2.33*a*, has several applications in electrical communication. Here we take it as an instructive exercise of the parallel and cascade relations. But first we need the individual transfer functions, determined as follows: the upper branch of the parallel section is a straight-through path so, trivially, $H_1(f) = 1$; the lower branch produces pure time delay of T seconds and sign inversion, and lumping them together gives $H_2(f) = -e^{-j2\pi fT}$ by application of the time-delay theorem; using the integration theorem, the integrator in the final block has $H_3(f) = 1/j2\pi f$. Figure 2.33*b* is the equivalent block diagram in terms of these transfer functions.

Having gotten this far, the rest of the work is easy. We combine the parallel branches in $H_{12}(f) = H_1(f) + H_2(f)$ and use the cascade rule to obtain

$$\begin{aligned} H(f) &= H_{12}(f)H_3(f) = [H_1(f) + H_2(f)]H_3(f) \\ &= [1 - e^{-j2\pi fT}]\frac{1}{j2\pi f} \\ &= \frac{e^{j\pi fT} - e^{-j\pi fT}}{j2\pi f} e^{-j\pi fT} = \frac{\sin \pi fT}{\pi f} e^{-j\pi fT} \\ &= T \operatorname{sinc} fT e^{-j\pi fT} \end{aligned} \qquad (17)$$

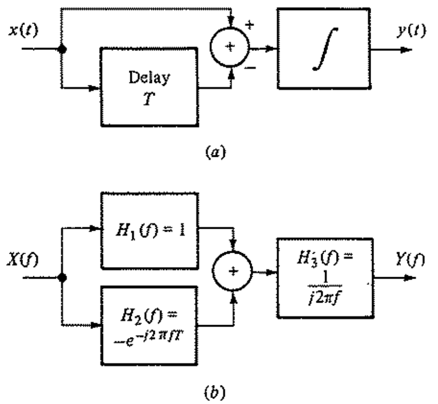


FIGURE 2.33
Block diagrams of zero-order hold. (*a*) Time domain; (*b*) frequency domain.

Hence we have the unusual result that the amplitude ratio of this system is a *sinc function* in frequency!

To confirm this result by another route, let us calculate the impulse response $h(t)$ drawing upon the definition that $y(t) = h(t)$ when $x(t) = \delta(t)$. Inspection of Fig. 2.33*a* shows that the input to the integrator then is $x(t) - x(t - T) = \delta(t) - \delta(t - T)$, so

$$h(t) = \int_{-\infty}^{t} [\delta(\lambda) - \delta(\lambda - T)]\, d\lambda = \begin{cases} 1 & 0 < t < T \\ 0 & \text{otherwise} \end{cases}$$

$$= \Pi\left(\frac{t - T/2}{T}\right) \tag{18}$$

and the impulse response is a rectangular pulse, Fig. 2.34. The reader should have little trouble identifying Eq. (17) as the transform of Eq. (18). ////

EXERCISE 2.14 Derive Eq. (17) by letting $x(t) = e^{j2\pi ft}$ in Fig. 2.33*a* and finding $y(t)$.

Real and Ideal Filters

Systems or networks that exhibit frequency-selective characteristics are called *filters*. A lowpass filter (LPF), for instance, passes only "low" frequencies—in the sense that

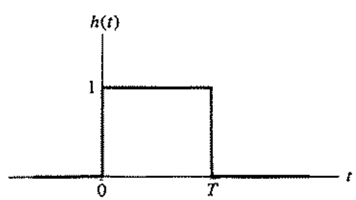


FIGURE 2.34
Impulse response of zero-order hold.

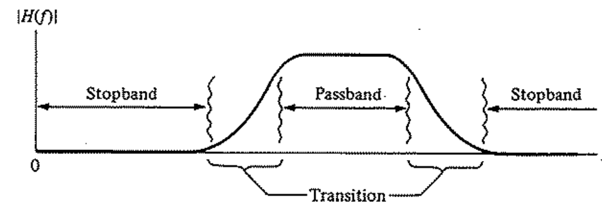


FIGURE 2.35
Passband and stopbands of a typical bandpass filter.

its amplitude ratio $|H(f)|$ is much greater at low frequencies than at high frequencies. The RC lowpass filter and the zero-order hold just discussed illustrate this type of frequency response. Similarly, there are *highpass* filters (HPF), *bandpass* filters (BPF), and *band-rejection* filters. The tuned circuit in the frequency multiplier acts as a BPF.

Ideally, a filter should have a sharp boundary between its *passband* and *stopband* or rejection band. Then the bandwidth of the filter could be unambiguously measured as the width of the passband. However, actual amplitude ratio curves do not have pronounced demarcation points, so the ends of the passband and stopband are ambiguous and there is a *transition* band between them. Figure 2.35 illustrates this characteristic for a typical bandpass filter; note that the negative-frequency portion has been omitted and the various bands are indicated only in terms of positive frequencies, in recognition of the even symmetry of $|H(f)|$ plus the fact that negative frequencies are something of a fiction introduced for analytic convenience.

When dealing with real filters, the bandwidth usually is taken to be the range of *positive frequencies* over which $|H(f)|$ falls no lower than $1/\sqrt{2}$ times the maximum value of $|H(f)|$ in the passband. This particular bandwidth convention—and there are others—is called the *half-power* or *3-decibel (dB) bandwidth*. The name stems from the fact that a sinusoidal input at the band-edge frequency would come out with its average power reduced by $(1/\sqrt{2})^2 = \frac{1}{2}$ compared to a sinusoid at the center of the passband. Converting this ratio to decibels† gives $10 \log_{10} \frac{1}{2} \approx -3$ dB so the power ratio at the edge of the passband is 3 dB below the center.

As an example of this convention, the half-power bandwidth of the RC LPF (Fig. 2.6) is $B = 1/2\pi RC$ since $|H(f)|_{max} = |H(0)| = 1$ and $|H(B)| = 1/\sqrt{2}$; the passband is $0 \le f \le B$. By the same token, the bandwidth of the tuned-circuit BPF in Fig. 2.12 is $B = f_r/Q$ since $|H(f)|_{max} = |H(f_r)| = R$ and $|H(f_r \pm f_r/2Q| = R/\sqrt{2}$. (At this point the reader should take another look at Figs. 2.6 and 2.12 in the light of these comments, especially noting that B is measured in terms of the positive-frequency range.)

† See Table E.

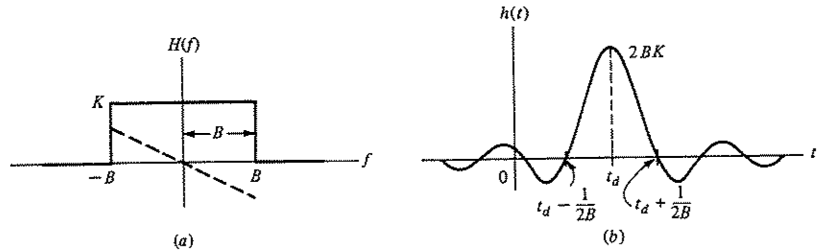


FIGURE 2.36
Ideal lowpass filter. (*a*) Transfer function; (*b*) impulse response.

More sophisticated filter designs have much more selective characteristics than our simple examples, to the point approaching stepwise frequency transitions. In the limit, we define an *ideal* LPF as having the rectangular characteristic

$$H(f) = Ke^{-j\omega t_d}\Pi\left(\frac{f}{2B}\right) \tag{19}$$

where K is the amplification, t_d is the time delay, and B is the bandwidth. Figure 2.36*a* plots this transfer function. Note that the bandwidth is unambiguous and the response outside the passband is identically zero. Ideal BPFs and HPFs are similarly defined, and BPFs will come up for further discussion in Chap. 5.

In advanced network theory it is shown that ideal filters cannot be physically realized. We skip the general proof here and give instead an argument based on impulse response. Consider, for example, the impulse response of an ideal LPF. By definition and using Eq. (19),

$$h(t) = \mathscr{F}^{-1}[H(f)] = \mathscr{F}^{-1}\left[Ke^{-j\omega t_d}\Pi\left(\frac{f}{2B}\right)\right]$$

$$= 2BK \operatorname{sinc} 2B(t - t_d) \tag{20}$$

which is plotted in Fig. 2.36*b*. Since $h(t)$ is the response to $\delta(t)$ and $h(t)$ has nonzero values for $t < 0$, *the output appears before the input is applied.* Such a filter is said to be *anticipatory*, and the portion of the output appearing before the input is called a *precursor*. Without doubt, such behavior is physically impossible, and hence the filter must be nonrealizable. Similar results are found for the bandpass and highpass case.

Fictitious though they may be, ideal filters are still conceptually useful in the study of communication systems, and practical filters can be designed that come quite close to being ideal, at least for engineering purposes. In fact, as the number of reactive elements increases without limit, the transfer function can be made arbitrarily close to that of an ideal filter. But at the same time, the filter time delay increases without limit. As a side point we observe that the infinite time delay means the precursors will always appear after the input is applied, which must be true of a real filter.

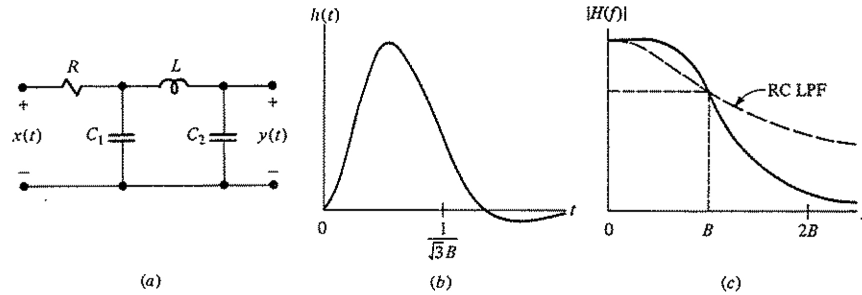


FIGURE 2.37
Third-order Butterworth filter. (*a*) Circuit ($R = 100$, $C_1 = 1/400\pi B$, $C_2 = 3C_1$, $L = 200/3\pi B$); (*b*) impulse response; (*c*) amplitude ratio.

Example 2.14 Butterworth Lowpass Filter

Figure 2.37*a* is the circuit diagram of a third-order Butterworth filter with 3-dB bandwidth B. Its transfer function and impulse response are

$$H(f) = \left\{1 - 2\left(\frac{f}{B}\right)^2 + j\left[2\left(\frac{f}{B}\right) - \left(\frac{f}{B}\right)^3\right]\right\}^{-1}$$

$$h(t) = 2\pi B\left[e^{-2\pi Bt} - \frac{2}{\sqrt{3}}e^{-\pi Bt}\cos(\pi\sqrt{3}Bt + 30°)\right]u(t)$$

as plotted in Fig. 2.37*b* and *c*. The similarity to an ideal LPF is apparent. Figure 2.37*c* also shows that this filter has a narrower transition region than the simple RC LPF.

For the case of M reactive elements, the amplitude ratio of an Mth-order Butterworth is

$$|H(f)| = \left[1 + \left(\frac{f}{B}\right)^{2M}\right]^{-1/2} \tag{21}$$

which is said to be *maximally flat* since the first M derivatives of $|H(f)|$ are zero at $f = 0$. (Incidentally, the RC LPF is a first-order Butterworth.) Other types of filter designs are given in the literature. ////

Bandlimiting and Timelimiting

Earlier we said that a signal $v(t)$ is *bandlimited* if there is a constant W such that

$$\mathscr{F}[v(t)] = 0 \qquad |f| > W \tag{22}$$

i.e., the spectrum has no content outside $|f| < W$. Similarly, a *timelimited* signal has the property that, for the constants $t_1 < t_2$,

$$v(t) = 0 \qquad t < t_1 \text{ and } t > t_2 \tag{23}$$

so the signal "starts" at time t_1 and "ends" at time t_2. Here we consider the meaning of these two definitions.

Ideal filters and bandlimited signals are concepts that go hand in hand. Indeed, passing an arbitrary signal through an ideal LPF produces a bandlimited signal at the output. But it has been seen that the impulse response of an ideal LPF is a sinc pulse existing for all time. We now assert that any signal emerging from an ideal LPF will exist for all time; stated another way, a bandlimited signal cannot be strictly timelimited. Conversely, a strictly timelimited signal cannot be bandlimited. In short, *bandlimiting and timelimiting are mutually incompatible*. A general proof of the assertion is difficult and will not be attempted here.† However, every transform pair encountered in this chapter is consistent therewith; e.g., see Figs. 2.13, 2.14, 2.17, 2.20.

This observation has implications for the signal and system models used in the study of communication systems. Since a signal cannot be simultaneously bandlimited and timelimited, we should either abandon bandlimited signals (and ideal filters) or accept signal models which exist for all time. But a physically real signal *is* strictly timelimited; it starts and it stops, or is turned on and off. On the other hand, the concept of bandlimited spectra is too powerful and appealing for engineering purposes to be dismissed entirely.

Fortunately, resolution of the dilemma is really not so difficult, requiring but a small compromise. Although a strictly timelimited signal is not strictly bandlimited, its spectrum can be *essentially* zero outside a certain frequency range, in the sense that the neglected frequency components contain an inconsequential portion of the total energy; e.g., consider $|f| \gg 1/\tau$ in the spectrum of a rectangular pulse. Similarly, a strictly bandlimited signal can be virtually zero outside a certain time interval; e.g., $\operatorname{sinc} 2Wt \approx 0$ for $|t| \gg 1/2W$. Therefore, it is not inappropriate to speak of signals that are both bandlimited and timelimited for most practical purposes.

† See Wozencraft and Jacobs (1965, app. 5B).

2.6 CORRELATION AND SPECTRAL DENSITY

In Sects. 2.3 and 2.4, two particular cases were observed. Nonperiodic energy signals were represented in the frequency domain by Fourier transforms that are continuous functions of frequency, free of impulses; the signal energy is

$$E = \int_{-\infty}^{\infty} |v(t)|^2\, dt = \int_{-\infty}^{\infty} v(t)v^*(t)\, dt \tag{1}$$

and the *energy spectral density* $|V(f)|^2$ gives its distribution in frequency; the power averaged over all time is zero since E is finite. On the other hand, periodic power signals were represented in the frequency domain by impulsive spectra resulting from transforms in the limit; the average power is

$$P = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} |v(t)|^2\, dt = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} v(t)v^*(t)\, dt \tag{2}$$

and the concept of energy spectral density does not apply since the total energy must be infinite when $P \neq 0$. However, it may be meaningful to speak of the distribution of power in the frequency domain, as described by a *power spectral density*.

The purpose of this section is to develop more fully the spectral density concept in a form that applies to both of the above cases as well as to the case of *random signals* — the latter anticipating the needs of Chap. 3. This generality is obtained at the price of using the somewhat abstract viewpoint of *signal space*, but the price is a worthwhile long-range investment that pays off now and in our future work. In the interest of getting to useful results as quickly as possible, many of the signal space derivations have been relegated to Appendix A, which the reader can consult for details omitted here.

Scalar Product, Norm, and Orthogonality

Let $v(t)$ and $w(t)$ be two signals of the same *class*, i.e., either energy type or power type. Their *scalar product* is a quantity — possibly complex — denoted† as $\langle v(t),w(t)\rangle$ and defined by

$$\langle v(t),w(t)\rangle \triangleq \begin{cases} \displaystyle\int_{-\infty}^{\infty} v(t)w^*(t)\, dt & \text{Energy signals} \quad (3a) \\ \displaystyle\lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} v(t)w^*(t)\, dt & \text{Power signals} \quad (3b) \end{cases}$$

† This notation is distinguished from $\langle v(t)\rangle$, the *time average* of $v(t)$; the scalar product symbol always has *two* functions separated by a comma.

If $v(t)$ and $w(t)$ happen to be periodic with period T_0, Eq. (3*b*) simplifies to

$$\langle v(t),w(t)\rangle = \frac{1}{T_0}\int_{T_0} v(t)w^*(t)\,dt \qquad \text{Periodic signals} \qquad (3c)$$

Another scalar product definition will be given in Chap. 3 covering the case of random signals. The value of the scalar product concept is that $\langle v(t),w(t)\rangle$ has certain invariant properties regardless of the specific definition.

Before stating some of those properties, note that the scalar product of $v(t)$ with itself equals either the energy E or power P. Generalizing, the *norm* of $v(t)$ is defined by

$$\|v\| \triangleq \langle v(t),v(t)\rangle^{1/2} \qquad (4)$$

a real nonnegative quantity, not to be confused with the function $|v(t)|$. Therefore

$$\|v\|^2 = \langle v(t), v(t)\rangle = \begin{cases} E & \text{Energy signals} \\ P & \text{Power signals} \end{cases} \qquad (5)$$

as follows from Eqs. (1) to (4).

Schwarz's inequality links Eqs. (3) and (4) in the form

$$|\langle v(t),w(t)\rangle| \le \|v\|\,\|w\| \qquad (6a)$$

which establishes an upper bound on the magnitude of the scalar product. The upper bound is achieved only when the two signals are directly proportional, i.e.,

$$|\langle v(t),w(t)\rangle| = \|v\|\,\|w\| \qquad \text{if } w(t) = \alpha v(t) \qquad (6b)$$

where α is an arbitrary constant. On the other hand, suppose that $v(t)$ and $w(t)$ are *orthogonal*, meaning

$$\langle v(t),w(t)\rangle = 0 \qquad (7)$$

Under this condition

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2 \qquad (8)$$

proof of which is given in Appendix A along with the proof of Eq. (6).

In signal space theory, orthogonal signals are viewed as perpendicular vectors and Eq. (8) becomes equivalent to the pythagorean theorem. Here we interpret $\langle v(t),w(t)\rangle = 0$ as the condition for *superposition of energy or power*. For instance, suppose $v(t)$ and $w(t)$ are orthogonal energy signals with $E_v = \|v\|^2$ and $E_w = \|w\|^2$; then the signal $z(t) = v(t) + w(t)$ has energy $E_z = \|v + w\|^2$ which, from Eq. (8), equals $E_v + E_w$.

It is also useful to enumerate some specific relationships between the functions

or waveforms $v(t)$ and $w(t)$ that result in their being orthogonal. Three different conditions under which $\langle v(t),w(t)\rangle = 0$ are:

1. When $v(t)$ and $w(t)$ have *opposite symmetry*, i.e., one is an even function and the other is odd.
2. When $v(t)$ and $w(t)$ are nonoverlapping or *disjoint in time*, i.e., one equals zero when the other is nonzero, and vice versa.
3. When $v(t)$ and $w(t)$ are *disjoint in frequency*, i.e., their spectra do not overlap.

The first two conditions are derived directly from Eq. (3); Eq. (13), Sect. 2.3, leads to the third. Bear in mind, however, that these are not the only relationships giving rise to orthogonality.

As a final point, note that the scalar product measures the degree of *similarity* between two signals. If they are proportional (i.e., similar), then $|\langle v(t),w(t)\rangle|$ is maximum; if they are orthogonal (i.e., dissimilar), then $\langle v(t),w(t)\rangle = 0$.

Correlation Functions

For any two signals of the same type, the *crosscorrelation* of $v(t)$ with $w(t)$ is defined† as

$$R_{vw}(\tau) \triangleq \langle v(t), w(t-\tau)\rangle \qquad (9)$$

a scalar product in which the second signal is displaced or shifted in time τ seconds. The displacement τ is arbitrary and, in fact, the crosscorrelation has τ as its independent variable, the time variable t having been washed out by the scalar product operation. By extension of the arguments just given, $R_{vw}(\tau)$ measures the similarity between $v(t)$ and $w(t-\tau)$ as a function of the displacement τ of $w(t)$ with respect to $v(t)$. Therefore, the crosscorrelation is a more comprehensive measure than the regular scalar product, for it detects any time-shifted similarities that would be ignored by $\langle v(t),w(t)\rangle$.

Now suppose we form the correlation of $v(t)$ with itself, i.e.,

$$R_v(\tau) \triangleq R_{vv}(\tau) = \langle v(t), v(t-\tau)\rangle \qquad (10)$$

which is called the *autocorrelation function*. Physically, autocorrelation has the same interpretation as crosscorrelation save that it compares a signal with itself displaced in time. But this means that $R_v(\tau)$ will tell us something about the *time variation* of $v(t)$, at least in an integrated or averaged sense. For instance, if $|R_v(\tau)|$ is large, then we infer that $v(t-\tau)$ is very similar (proportional) to $\pm v(t)$ for that particular value of τ; conversely, if $R_v(\tau) = 0$ for some value of τ, then we know that $v(t)$ and $v(t-\tau)$ are orthogonal. This interpretation presently is invoked in conjunction with the

† The definition $\langle v(t), w(t+\tau)\rangle$ is also used.

spectral density function. First, however, we need to list two mathematical properties of autocorrelation.

Applying Eqs. (5) and (6) to Eq. (10) yields

$$R_v(0) = \langle v(t),v(t) \rangle = \|v\|^2 \tag{11a}$$

$$|R_v(\tau)| \le R_v(0) \tag{11b}$$

so that $R_v(\tau)$ has a maximum at $\tau = 0$ where it equals the signal energy or power $\|v\|^2$. It can also be shown from Eqs. (3) and (10) that

$$R_v(-\tau) = R_v^*(\tau) \tag{12}$$

Thus, if $v(t)$ is *real*, $R_v(\tau)$ is *real* and has *even symmetry*.

Consider now the sum of two signals, say $z(t) = v(t) + w(t)$, whose correlation is $R_z(\tau) = R_v(\tau) + R_{vw}(\tau) + R_{wv}(\tau) + R_w(\tau)$. If the component signals are *orthogonal for all* τ, i.e., if

$$R_{vw}(\tau) = R_{wv}(\tau) = 0 \tag{13a}$$

then

$$R_z(\tau) = R_v(\tau) + R_w(\tau) \tag{13b}$$

and, setting $\tau = 0$,

$$\|z\|^2 = \|v\|^2 + \|w\|^2 \tag{13c}$$

In this case, the signals are said to be *incoherent* and we have superposition of correlation functions as well as superposition of energy or power.

Example 2.15 Autocorrelation of a Sinusoid

Let us calculate the autocorrelation of the periodic power signal

$$z(t) = A\cos(\omega_0 t + \theta) \qquad \omega_0 = 2\pi/T_0$$

While this may be done directly using Eqs. (3c) and (10), it is easier and more instructive to write

$$z(t) = \underbrace{\frac{A}{2}e^{j\theta}e^{j\omega_0 t}}_{v(t)} + \underbrace{\frac{A}{2}e^{-j\theta}e^{-j\omega_0 t}}_{w(t)}$$

This happens to be a better method for the case at hand because $v(t)$ and $w(t-\tau)$ are orthogonal for all τ—since they are disjoint in frequency—and, moreover, $w(t) = v^*(t)$. Thus, $R_{vw}(\tau) = R_{wv}(\tau) = 0$ and $R_w(\tau) = R_v^*(\tau)$, so Eq. (13b) becomes $R_z(\tau) = R_v(\tau) + R_v^*(\tau) = 2\,\mathrm{Re}\,[R_v(\tau)]$. Incidentally, although not all problems can be simplified this much, the student should be alert to such possibilities.

Proceeding with the calculation, we have

$$R_v(\tau) = \frac{1}{T_0}\int_{T_0}\left[\frac{A}{2}e^{j\theta}e^{j\omega_0 t}\right]\left[\frac{A}{2}e^{-j\theta}e^{-j\omega_0(t-\tau)}\right]dt = \frac{A^2}{4}e^{j\omega_0\tau} \tag{14}$$

and hence

$$R_z(\tau) = 2\,\mathrm{Re}\,[R_v(\tau)] = \frac{A^2}{2}\cos\omega_0\tau \tag{15}$$

so that $R_z(0) = A^2/2 = \|z\|^2$ as predicted. Note that the autocorrelation of a sinusoid is another sinusoid at the same frequency but in the "τ domain" rather than the time domain. The phase parameter θ has dropped out owing to the averaging effect of correlation. Because of this fact we conclude that the autocorrelation does not uniquely define a signal, e.g., all $z(t) = A\cos(\omega_0 t + \theta)$ have the same $R_z(\tau)$ regardless of θ. ////

EXERCISE 2.15 The autocorrelation of an energy signal is a type of *convolution* since replacing t with λ in Eq. (3a) gives

$$R_v(\tau) = \int_{-\infty}^{\infty} v(\lambda)v^*(\lambda-\tau)\,d\lambda = [v(\tau)] * [v^*(-\tau)] \tag{16}$$

Use this to show that a rectangular pulse

$$v(t) = A\Pi\left(\frac{t}{T}\right) \tag{17a}$$

has a triangular autocorrelation

$$R_v(\tau) = A^2 T\Lambda\left(\frac{\tau}{T}\right) \tag{17b}$$

Underscoring the previously observed nonuniqueness, Eq. (17b) also holds when $v(t) = A\Pi[(t-t_d)/T]$ for any value of t_d.

Spectral Density Functions

In view of our observation that $R_v(\tau)$ gives information about the time-domain behavior of $v(t)$, it seems plausible to investigate the Fourier transform of $R_v(\tau)$ as a possible basis for frequency-domain analysis. Consider, therefore,

$$G_v(f) \triangleq \mathscr{F}[R_v(\tau)] = \int_{-\infty}^{\infty} R_v(\tau)e^{-j2\pi f\tau}\,d\tau \tag{18a}$$

which is called the *spectral density function* for reasons soon explained. Having made this definition, it follows that $\mathscr{F}^{-1}[G_v(f)]$ equals $R_v(\tau)$, i.e.,

$$R_v(\tau) = \int_{-\infty}^{\infty} G_v(f)e^{+j2\pi f\tau}\,df \qquad (18b)$$

so we have the Fourier transform pair

$$R_v(\tau) \leftrightarrow G_v(f) \qquad (19)$$

where τ takes the place of t. Equation (19) bears the name of the *Wiener-Kinchine theorem*. In the special but important case where $v(t)$ is *real*, $G_v(f)$ is *real* and *even* since $R_v(\tau)$ is real and even.

The fundamental property of $G_v(f)$ is that integrating it over all frequency yields $\|v\|^2$, as is easily proved by setting $\tau = 0$ in Eq. (18*b*), i.e.,

$$\int_{-\infty}^{\infty} G_v(f)\,df = R_v(0) = \|v\|^2 \qquad (20)$$

To interpret this, and thereby justify the name of $G_v(f)$, we recall that $\|v\|^2$ is the energy or power associated with $v(t)$. Therefore, one can argue that $G_v(f)$ tells how the energy or power is distributed in the frequency domain and deserves being called the spectral density.

Supporting that view, let $v(t)$ be an energy signal and $V(f)$ its spectrum. From Eq. (16), $R_v(\tau) = [v(\tau)] * [v^*(-\tau)]$; so, invoking the convolution theorem,

$$G_v(f) = \mathscr{F}[R_v(\tau)] = \mathscr{F}[v(\tau)]\mathscr{F}[v^*(-\tau)]$$

Clearly, $\mathscr{F}[v(\tau)] = \mathscr{F}[v(t)] = V(f)$, and it is a routine exercise to show that $\mathscr{F}[v^*(-\tau)] = \{\mathscr{F}[v(t)]\}^* = V^*(f)$. Therefore,

$$G_v(f) = V(f)V^*(f) = |V(f)|^2 \qquad (21a)$$

which was identified in Sect. 2.3 as the *energy spectral density*. Inversion of Eq. (21*a*) gives

$$R_v(\tau) = \mathscr{F}^{-1}[|V(f)|^2] = \int_{-\infty}^{\infty} |V(f)|^2 e^{j2\pi f\tau}\,df \qquad (21b)$$

an alternate expression for the autocorrelation function of an energy signal. Furthermore, with $\tau = 0$,

$$\|v\|^2 = \int_{-\infty}^{\infty} |V(f)|^2\,df \qquad (21c)$$

which, in retrospect, is just *Rayleigh's energy theorem*. The reader may wish to check the consistency of Eq. (21*b*) with Eq. (17).

Now consider a periodic power signal expressed in Fourier series form

$$v(t) = \sum_{n=-\infty}^{\infty} c_v(nf_0)e^{j2\pi nf_0t} \qquad f_0 = \frac{1}{T_0}$$

Using Eq. (14) and the fact that all terms in the series are mutually incoherent,

$$R_v(\tau) = \sum_{n=-\infty}^{\infty} |c_v(nf_0)|^2 e^{j2\pi nf_0\tau} \qquad (22a)$$

and Fourier transformation gives

$$G_v(f) = \sum_{n=-\infty}^{\infty} |c_v(nf_0)|^2\,\delta(f - nf_0) \qquad (22b)$$

But is this the *power spectral density* of a periodic signal? The answer is clearly affirmative since a periodic signal can be decomposed into terms of the form $c_v(nf_0)e^{j2\pi nf_0t}$ each of which represents an amount of power equal to $|c_v(nf_0)|^2$ located exactly at the frequency $f = nf_0$. The spectral density is impulsive in this case simply because impulses are required to indicate that nonzero units of power are concentrated at discrete frequencies. Finally, integrating Eq. (22*b*) over all f—or simply setting $\tau = 0$ in Eq. (22*a*)—gives *Parseval's power theorem*.

To summarize, while we have not explicitly proved that $G_v(f)$ represents the spectral density for other types of signals, we have presented strong evidence in favor of that conclusion based on two cases where there is a firm intuitive notion of what the spectral density should be. Concluding this chapter, spectral density functions are used in conjunction with transfer functions to determine input-output relations. But before doing so, it must be pointed out that $G_v(f)$, like $R_v(\tau)$, does not uniquely represent $v(t)$. True, a given signal has only one spectral density function; however, that spectral density function may apply to other signals. Phase-shifted sinusoids and time-delayed rectangular pulses are simple examples. And in the case of random signals, two or more drastically different waveforms can have the same spectral density—meaning that their averages are the same even though the waveforms are different.

Input-Output Relations

As diagramed in Fig. 2.38, let $x(t)$ be the input to a linear time-invariant system having inpulse response $h(t)$ and transfer function $H(f)$. If it happens that $x(t)$ is an energy signal with spectrum $X(f)$, the output energy spectral density is

$$|Y(f)|^2 = |H(f)|^2\,|X(f)|^2$$

or

$$G_y(f) = |H(f)|^2\,G_x(f) \qquad (23)$$

since $G_x(f) = |X(f)|^2$ is the energy spectral density of the input, etc.

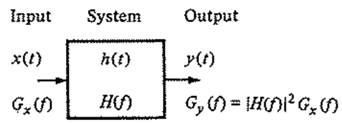


FIGURE 2.38

Although we derived Eq. (23) for energy signals, its form suggests that $|H(f)|^2$ always relates the input and output spectral density functions, irrespective of signal type. This, indeed, is true; Eq. (23) applies for any type of input whose spectral density function exists. The general proof is relatively straightforward but tedious, and will only be outlined here.

First, the autocorrelation $R_y(\tau)$ of the output is found by inserting

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

into $R_y(\tau) = \langle y(t), y(t-\tau)\rangle$. Upon manipulation, this yields the awesome expression

$$R_y(\tau) = \iint_{-\infty}^{\infty} h(\lambda)h^*(\mu)R_x(\tau+\mu-\lambda)\,d\mu\,d\lambda \qquad (24)$$

where μ is another dummy variable. Then, taking the Fourier transform, one finally gets to

$$G_y(f) = \left[\int_{-\infty}^{\infty} h(\lambda)e^{-j2\pi f\lambda}\,d\lambda\right]\left[\int_{-\infty}^{\infty} h^*(\mu)e^{+j2\pi f\mu}\,d\mu\right]G_x(f)$$

so $G_y(f) = H(f)H^*(f)G_x(f)$, as asserted.

To reiterate the significance of this result, given an input signal of almost any type, the corresponding spectral density function at the output of a linear system is found by multiplying the input spectral density by $|H(f)|^2$. Having thus obtained $G_y(f)$, application of Eq. (20) gives the output signal's *energy* or *power* as

$$\|y\|^2 = \int_{-\infty}^{\infty} G_y(f)\,df = \int_{-\infty}^{\infty} |H(f)|^2 G_x(f)\,df \qquad (25)$$

Often, especially in those problems involving random signals, $\|y\|^2$ is precisely the information being sought—and Eq. (25) offers the most direct route to that information. Other information about $y(t)$ may be gleaned from

$$R_y(\tau) = \mathscr{F}^{-1}[G_y(f)] = \int_{-\infty}^{\infty} |H(f)|^2 G_x(f)e^{j2\pi f\tau}\,df \qquad (26)$$

Note, by the way, that this inverse transform is probably easier to deal with than Eq. (24).

Equation (23) also expedites spectral density calculations, whether or not filtering actually is involved. Suppose, for instance, that $w(t) = dv(t)/dt$, $G_v(f)$ is known, and one desires to find $G_w(f)$. Conceptually, $w(t)$ could be generated by passing $v(t)$ through an ideal differentiator, for which $H(f) = j2\pi f$—see Eq. (24), Sect. 2.3. Therefore, using Eq. (23), if

$$w(t) = \frac{dv(t)}{dt} \qquad (27a)$$

then

$$G_w(f) = (2\pi f)^2 G_v(f) \qquad (27b)$$

Likewise, if

$$w(t) = \int_{-\infty}^{t} v(\lambda)\,d\lambda \qquad (28a)$$

then

$$G_w(f) = (2\pi f)^{-2} G_v(f) \qquad (28b)$$

Further illustrations of the use of Eqs. (23), (25), and (26) are presented in the next chapter, as applied to random signals.

2.7 PROBLEMS

2.1 (Sect. 2.1) Use Eq. (9) to find $y(t)$ when $x(t) = 4\cos 2\pi 10t$ and $H(j\omega) = 15 + j(\omega/\pi)$. *Ans.*: $100\cos(2\pi 10t + 53°)$.

2.2 (Sect. 2.1) Find and sketch $|H(f)|$ and $\arg[H(f)]$ for each of the following transfer functions:
(*a*) $(10 + jf)/(1 + jf)$
(*b*) $(1 + jf)/(10 + jf)$
(*c*) $(1 - jf)/(1 + jf)$

2.3 (Sect. 2.1) If the capacitor in Fig. 2.6*a* is replaced by an inductor, show that $H(f) = j(f/B)/[1 + j(f/B)]$ where $B = R/2\pi L$. Sketch the amplitude ratio and phase shift.

2.4 (Sect. 2.1) Referring to Fig. 2.6*a*, suppose $x(t) = 10\cos 2\pi f_0 t$ and the filter has $B = 3$ kHz. If $A_y = 2$, what is the value of f_0?

2.5 (Sect. 2.2) Find $\langle v(t)\rangle$ and P for $v(t) = Ae^{j(\omega_0 t+\theta)}$. *Ans.*: 0, A^2.

2.6 (Sect. 2.2) Find $\langle v(t)\rangle$ and P for the full-rectified sinusoid $v(t) = A|\sin 2\pi t/T_0|$.

2.7 (Sect. 2.2) When $v(-t) = v(t)$, show that

$$c(nf_0) = \frac{2}{T_0}\int_0^{T_0/2} v(t)\cos 2\pi n f_0 t\,dt$$

and use this to prove Eq. (11). Carry out a similar analysis for $v(-t) = -v(t)$.