A general **cubic** equation

$$y^3 + py^2 + qy + r = 0$$

may be reduced to the **depressed cubic** form

$$x^3 + ax + b = 0$$

by substituting

$$y = x - \tfrac{p}{3}$$

This yields

$$a = \tfrac{1}{3}(3q - p^2) \qquad b = \tfrac{1}{27}(2p^3 - 9pq + 27r)$$

Now let

$$A = \sqrt[3]{-\tfrac{b}{2} + \sqrt{\tfrac{b^2}{4} + \tfrac{a^3}{27}}}, \qquad B = \sqrt[3]{-\tfrac{b}{2} - \sqrt{\tfrac{b^2}{4} + \tfrac{a^3}{27}}}$$

The solution of the depressed cubic is

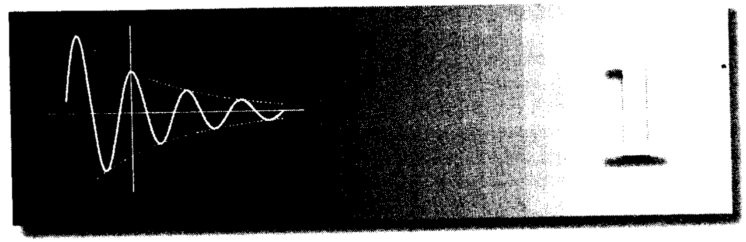
$$x = A + B, \qquad x = -\tfrac{A+B}{2} + \tfrac{A-B}{2}\sqrt{-3}, \qquad x = -\tfrac{A+B}{2} - \tfrac{A-B}{2}\sqrt{-3}$$

and

$$y = x - \tfrac{p}{3}$$

References

1. Asimov, Isaac, *Asimov on Numbers,* Bell Publishing Co., N.Y., 1982.
2. Calinger, R., Ed., *Classics of Mathematics,* Moore Publishing Co.,Oak Park, IL., 1982.
3. Hogben, Lancelot, *Mathematics in the Making,* Doubleday & Co. Inc., New York, 1960.
4. Cajori, Florian, *A History of Mathematics*, 4th ed., Chelsea, New York, 1985.
5. Encyclopaedia Britannica, 15th ed., *Micropaedia*, vol. 11, p. 1043, 1982.
6. Singh, Jagjit, *Great Ideas of Modern Mathematics,* Dover, New York, 1959.
7. Dunham, William, *Journey through Genius,* Wiley, New York, 1990.



Introduction to Signals and Systems

In this chapter we shall discuss certain basic aspects of signals. We shall also introduce important basic concepts and qualitative explanations of the how's and why's of systems theory, thus building a solid foundation for understanding the quantitative analysis in the remainder of the book.

Signals

A **signal**, as the term implies, is a set of information or data. Examples include a telephone or a television signal, monthly sales of a corporation, or the daily closing prices of a stock market (e.g., the Dow Jones averages). In all these examples, the signals are functions of the independent variable *time*. This is not always the case, however. When an electrical charge is distributed over a body, for instance, the signal is the charge density, a function of *space* rather than time. In this book we deal almost exclusively with signals that are functions of time. The discussion, however, applies equally well to other independent variables.

Systems

Signals may be processed further by **systems**, which may modify them or extract additional information from them. For example, an antiaircraft gun operator may want to know the future location of a hostile moving target that is being tracked by his radar. Knowing the radar signal he knows the past location and velocity of the target. By properly processing the radar signal (the input) he can approximately estimate the future location of the target. Thus, a system is an entity that *processes* a set of signals (**inputs**) to yield another set of signals (**outputs**). A system may be made up of physical components, as in electrical, mechanical, or hydraulic systems (hardware realization), or it may be an algorithm that computes an output from an input signal (software realization).

1.1 Size of a Signal

The size of any entity is a number that indicates the largeness or strength of that entity. Generally speaking, the signal amplitude varies with time. How can a signal that exists over a certain time interval with varying amplitude be measured by one number that will indicate the signal size or signal strength? Such a measure must consider not only the signal amplitude, but also its duration. For instance, if we are to devise a single number V as a measure of the size of a human being, we must consider not only his or her width (girth), but also the height. If we make a simplifying assumption that the shape of a person is a cylinder of variable radius r (which varies with the height h) then a reasonable measure of the size of a person of height H is the person's volume V, given by

$$V = \pi \int_0^H r^2(h)\, dh$$

Signal Energy

Arguing in this manner, we may consider the area under a signal $f(t)$ as a possible measure of its size, because it takes account of not only the amplitude, but also the duration. However, this will be a defective measure because $f(t)$ could be a large signal, yet its positive and negative areas could cancel each other, indicating a signal of small size. This difficulty can be corrected by defining the signal size as the area under $f^2(t)$, which is always positive. We call this measure the **signal energy** E_f, defined (for a real signal) as

$$E_f = \int_{-\infty}^{\infty} f^2(t)\, dt \tag{1.1}$$

This definition can be generalized to a complex valued signal $f(t)$ as

$$E_f = \int_{-\infty}^{\infty} |f(t)|^2\, dt \tag{1.2}$$

There are also other possible measures of signal size, such as the area under $|f(t)|$. The energy measure, however, is not only more tractable mathematically, but is also more meaningful (as shown later) in the sense that it is indicative of the energy that can be extracted from the signal.

Signal Power

The signal energy must be finite for it to be a meaningful measure of the signal size. A necessary condition for the energy to be finite is that the signal amplitude $\to 0$ as $|t| \to \infty$ (Fig. 1.1a). Otherwise the integral in Eq. (1.1) will not converge.

In some cases, for instance, when the amplitude of $f(t)$ does not $\to 0$ as $|t| \to \infty$ (Fig. 1.1b), then, the signal energy is infinite. A more meaningful measure of the signal size in such a case would be the time average of the energy, if it exists. This measure is called the **power** of the signal. For a signal $f(t)$, we define its power P_f as

$$P_f = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} f^2(t)\, dt \tag{1.3}$$

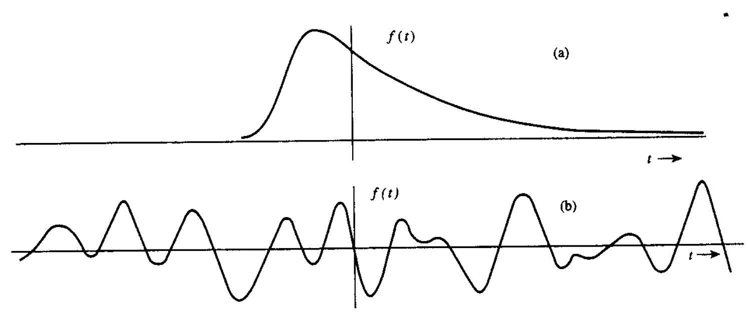


Fig. 1.1 Examples of Signals: (a) a signal with finite energy (b) a signal with finite power.

We can generalize this definition for a complex signal $f(t)$ as

$$P_f = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} |f(t)|^2\, dt \tag{1.4}$$

Observe that the signal power P_f is the time average (mean) of the signal amplitude squared, that is, the *mean-squared* value of $f(t)$. Indeed, the square root of P_f is the familiar **rms** (root mean square) value of $f(t)$.

The mean of an entity averaged over a large time interval approaching infinity exists if the entity is either periodic or has a statistical regularity. If such a condition is not satisfied, the average may not exist. For instance, a ramp signal $f(t) = t$ increases indefinitely as $|t| \to \infty$, and neither the energy nor the power exists for this signal.

Comments

The signal energy as defined in Eq. (1.1) or Eq. (1.2) does not indicate the actual energy of the signal because the signal energy depends not only on the signal, but also on the load. It can, however, be interpreted as the energy dissipated in a normalized load of a 1-ohm resistor if a voltage $f(t)$ were to be applied across the 1-ohm resistor (or if a current $f(t)$ were to be passed through the 1-ohm resistor). The measure of "energy" is, therefore indicative of the energy capability of the signal and not the actual energy. For this reason the concepts of conservation of energy should not be applied to this "signal energy". Parallel observation applies to "signal power" defined in Eq. (1.3) or (1.4). These measures are but convenient indicators of the signal size, which prove useful in many applications. For instance, if we approximate a signal $f(t)$ by another signal $g(t)$, the error in the approximation is $e(t) = f(t) - g(t)$. The energy (or power) of $e(t)$ is a convenient indicator of the goodness of the approximation. It provides us with a quantitative measure of determining the closeness of the approximation. In communication systems, during transmission over a channel, message signals are corrupted by unwanted signals (noise). The quality of the received signal is judged by the relative sizes of the

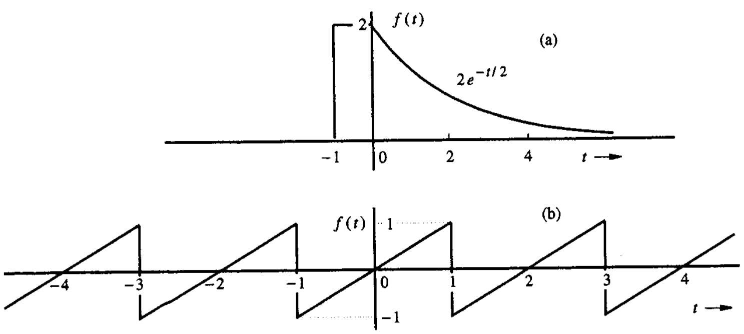


Fig. 1.2 Signals for Example 1.1.

desired signal and the unwanted signal (noise). In this case the ratio of the message signal and noise signal powers (signal to noise power ratio) is a good indication of the received signal quality.

Units of Energy and Power: Equations (1.1) and (1.2) are not correct dimensionally. This is because here we are using the term *energy* not in its conventional sense, but to indicate the signal size. The same observation applies to Eqs. (1.3) and (1.4) for power. The units of energy and power, as defined here, depend on the nature of the signal $f(t)$. If $f(t)$ is a voltage signal, its energy E_f has units of V^2s (volts squared-seconds) and its power P_f has units of V^2 (volts squared). If $f(t)$ is a current signal, these units will be A^2s (amperes squared-seconds) and A^2 (amperes squared), respectively.

■ **Example 1.1**

Determine the suitable measures of the signals in Fig 1.2.

In Fig. 1.2a, the signal amplitude $\to 0$ as $|t| \to \infty$. Therefore the suitable measure for this signal is its energy E_f given by

$$E_f = \int_{-\infty}^{\infty} f^2(t)\, dt = \int_{-1}^{0} (2)^2\, dt + \int_{0}^{\infty} 4e^{-t}\, dt = 4 + 4 = 8$$

In Fig. 1.2b, the signal amplitude does not $\to 0$ as $|t| \to \infty$. However, it is periodic, and therefore its power exists. We can use Eq. (1.3) to determine its power. We can simplify the procedure for periodic signals by observing that a periodic signal repeats regularly each period (2 seconds in this case). Therefore, averaging $f^2(t)$ over an infinitely large interval is identical to averaging this quantity over one period (2 seconds in this case). Thus

$$P_f = \frac{1}{2}\int_{-1}^{1} f^2(t)\, dt = \frac{1}{2}\int_{-1}^{1} t^2\, dt = \frac{1}{3}$$

Recall that the signal power is the square of its rms value. Therefore, the rms value of this signal is $1/\sqrt{3}$. ■

■ **Example 1.2**

Determine the power and the rms value of

(a) $f(t) = C\cos(\omega_0 t + \theta)$ **(b)** $f(t) = C_1\cos(\omega_1 t + \theta_1) + C_2\cos(\omega_2 t + \theta_2)$ $(\omega_1 \neq \omega_2)$.
(c) $f(t) = De^{j\omega_0 t}$.

(a) This is a periodic signal with period $T_0 = 2\pi/\omega_0$. The suitable measure of this signal is its power. Because it is a periodic signal, we may compute its power by averaging its energy over one period $T_0 = 2\pi/\omega_0$. However, for the sake of demonstration, we shall solve this problem by averaging over an infinitely large time interval using Eq (1.3).

$$P_f = \lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2} C^2\cos^2(\omega_0 t + \theta)\, dt = \lim_{T\to\infty}\frac{C^2}{2T}\int_{-T/2}^{T/2}[1 + \cos(2\omega_0 t + 2\theta)]\, dt$$

$$= \lim_{T\to\infty}\frac{C^2}{2T}\int_{-T/2}^{T/2} dt + \lim_{T\to\infty}\frac{C^2}{2T}\int_{-T/2}^{T/2}\cos(2\omega_0 t + 2\theta)\, dt$$

The first term on the right-hand side is equal to $C^2/2$. Moreover, the second term is zero because the integral appearing in this term represents the area under a sinusoid over a very large time interval T with $T \to \infty$. This area is at most equal to the area of half the cycle because of cancellations of the positive and negative areas of a sinusoid. The second term is this area multiplied by $C^2/2T$ with $T \to \infty$. Clearly this term is zero, and

$$P_f = \frac{C^2}{2} \tag{1.5a}$$

This shows that a sinusoid of amplitude C has a power $C^2/2$ regardless of the value of its frequency ω_0 $(\omega_0 \neq 0)$ and phase θ. The rms value is $C/\sqrt{2}$. If the signal frequency is zero (dc or a constant signal of amplitude C), the reader can show that the power is C^2.

(b) In Chapter 4, we show that a sum of two sinusoids may or may not be periodic, depending on whether the ratio ω_1/ω_2 is a rational number or not. Therefore, the period of this signal is not known. Hence, its power will be determined by averaging its energy over T seconds with $T \to \infty$. Thus,

$$P_f = \lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2}[C_1\cos(\omega_1 t + \theta_1) + C_2\cos(\omega_2 t + \theta_2)]^2\, dt$$

$$= \lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2} {C_1}^2\cos^2(\omega_1 t + \theta_1)\, dt + \lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2} {C_2}^2\cos^2(\omega_2 t + \theta_2)\, dt$$

$$+ \lim_{T\to\infty}\frac{2C_1C_2}{T}\int_{-T/2}^{T/2}\cos(\omega_1 t + \theta_1)\cos(\omega_2 t + \theta_2)\, dt$$

The first and second integrals on the right-hand side are the powers of the two sinusoids, which are ${C_1}^2/2$ and ${C_2}^2/2$ as found in part **(a)**. Arguing as in part **(a)**, we see that the third term is zero, and we have†

$$P_f = \frac{{C_1}^2}{2} + \frac{{C_2}^2}{2} \tag{1.5b}$$

and the rms value is $\sqrt{({C_1}^2 + {C_2}^2)/2}$.

We can readily extend this result to a sum of any number of sinusoids with distinct frequencies. Thus, if

†This is true only if $\omega_1 \neq \omega_2$. If $\omega_1 = \omega_2$, the integrand of the third term contains a constant $\cos(\theta_1 - \theta_2)$, and the third term $\to 2C_1C_2\cos(\theta_1 - \theta_2)$ as $T \to \infty$.

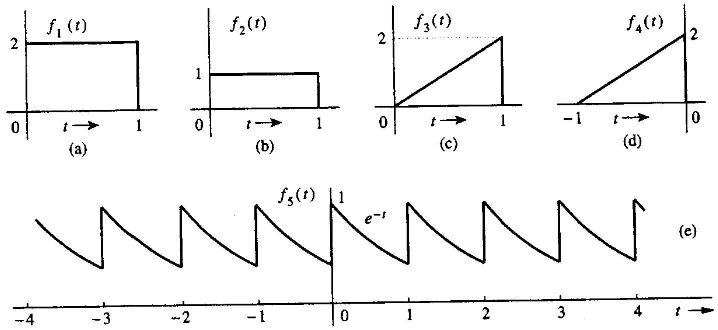


Fig. 1.3 Signals for Exercise E1.1.

$$f(t) = \sum_{n=1}^{\infty} C_n \cos(\omega_n t + \theta_n)$$

where none of the two sinusoids have identical frequencies, then

$$P_f = \frac{1}{2}\sum_{n=1}^{\infty} {C_n}^2 \tag{1.5c}$$

(c) In this case the signal is complex, and we use Eq. (1.4) to compute the power.

$$P_f = \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} |De^{j\omega_0 t}|^2\, dt$$

Recall that $|e^{j\omega_0 t}| = 1$ so that $|De^{j\omega_0 t}|^2 = |D|^2$, and

$$P_f = |D|^2 \tag{1.5d}$$

The rms value is $|D|$. ■

Comment: In part **(b)** we have shown that the power of the sum of two sinusoids is equal to the sum of the powers of the sinusoids. It appears that the power of $f_1(t) + f_2(t)$ is $P_{f_1} + P_{f_2}$. Unfortunately, this conclusion is not true in general. It is true only under a certain condition (orthogonality) discussed later in Sec. 3.1-3.

△ **Exercise E1.1**

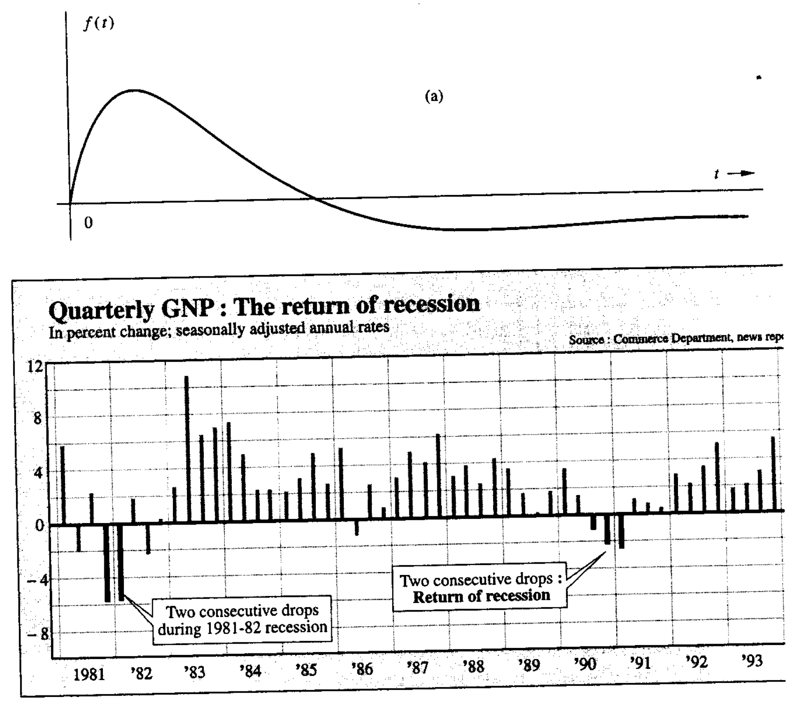
Show that the energies of the signals in Figs. 1.3a,b,c and d are 4, 1, 4/3, and 4/3, respectively. Observe that doubling a signal quadruples the energy, and time-shifting a signal has no effect on the energy. Show also that the power of the signal in Fig. 1.3e is 0.4323. What is the rms value of signal in Fig. 1.3e? ▽

△ **Exercise E1.2**

Redo Example 1.2a to find the power of a sinusoid $C\cos(\omega_0 t + \theta)$ by averaging the signal energy over one period $T_0 = 2\pi/\omega_0$ (rather than averaging over the infinitely large interval). Show also that the power of a constant signal $f(t) = C_0$ is C_0^2, and its rms value is C_0. ▽

△ **Exercise E1.3**

Show that if $\omega_1 = \omega_2$, the power of $f(t) = C_1\cos(\omega_1 t + \theta_1) + C_2\cos(\omega_2 t + \theta_2)$ is $[{C_1}^2 + {C_2}^2 + 2C_1C_2\cos(\theta_1 - \theta_2)]/2$, which is not equal to $({C_1}^2 + {C_2}^2)/2$. ▽



(b)

Fig. 1.4 Continuous-time and Discrete-time Signals.

1.2 Classification of Signals

There are several classes of signals. Here we shall consider only the following classes, which are suitable for the scope of this book:

1. Continuous-time and discrete-time signals
2. Analog and digital signals
3. Periodic and aperiodic signals
4. Energy and power signals
5. Deterministic and probabilistic signals

1.2-1 Continuous-Time and Discrete-Time Signals

A signal that is specified for every value of time t (Fig. 1.4a) is a **continuous-time signal**, and a signal that is specified only at discrete values of t (Fig. 1.4b) is **a discrete-time signal**. Telephone and video camera outputs are continuous-time

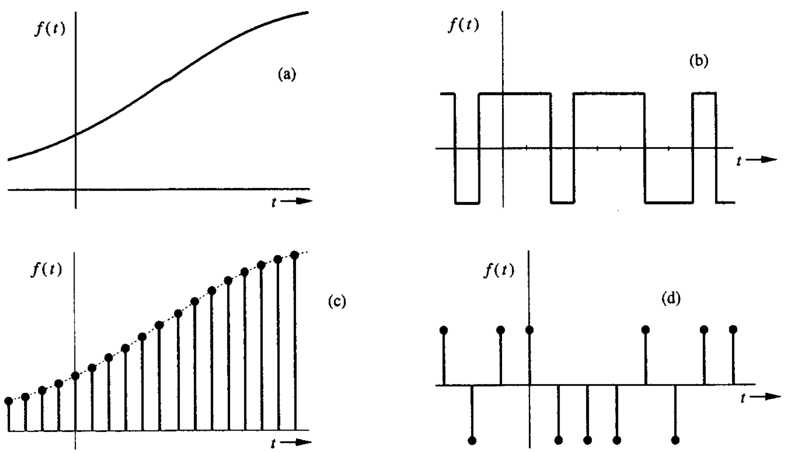


Fig. 1.5 Examples of Signals: (a) analog, continuous-time (b) digital, continuous-time (c) analog, discrete-time (d) digital, discrete-time.

signals, whereas the quarterly gross national product (GNP), monthly sales of a corporation, and stock market daily averages are discrete-time signals.

1.2-2 Analog and Digital Signals

The concept of continuous-time is often confused with that of analog. The two are not the same. The same is true of the concepts of discrete-time and digital. A signal whose amplitude can take on any value in a continuous range is an **analog signal**. This means that an analog signal amplitude can take on an infinite number of values. A **digital signal**, on the other hand, is one whose amplitude can take on only a finite number of values. Signals associated with a digital computer are digital because they take on only two values (binary signals). A digital signal whose amplitudes can take on M values is an **M-ary** signal of which binary ($M = 2$) is a special case. The terms *continuous-time* and *discrete-time* qualify the nature of a signal along the time (horizontal) axis. The terms *analog* and *digital*, on the other hand, qualify the nature of the signal amplitude (vertical axis). Figure 1.5 shows examples of various types of signals. It is clear that analog is not necessarily continuous-time and digital need not be discrete-time. Figure 1.5c shows an example of an analog discrete-time signal. An analog signal can be converted into a digital signal [analog-to-digital (A/D) conversion] through quantization (rounding off), as explained in Sec. 5.1-3.

1.2-3 Periodic and Aperiodic Signals

A signal $f(t)$ is said to be **periodic** if for some positive constant T_0

$$f(t) = f(t + T_0) \qquad \text{for all } t \tag{1.6}$$

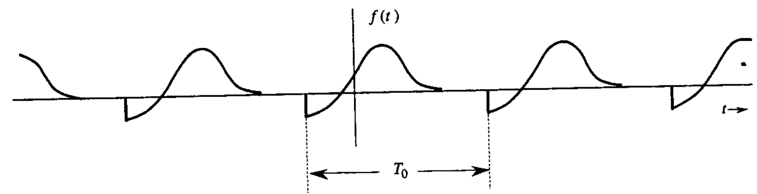


Fig. 1.6 A periodic signal of period T_0.

The *smallest* value of T_0 that satisfies the periodicity condition (1.6) is the **period** of $f(t)$. The signals in Figs. 1.2b and 1.3e are periodic signals with periods 2 and 1, respectively. A signal is **aperiodic** if it is not periodic. Signals in Figs. 1.2a, 1.3a, 1.3b, 1.3c, and 1.3d are all aperiodic.

By definition, a periodic signal $f(t)$ remains unchanged when time-shifted by one period. For this reason a periodic signal must start at $t = -\infty$ because if it starts at some finite instant, say $t = 0$, the time-shifted signal $f(t + T_0)$ will start at $t = -T_0$ and $f(t + T_0)$ would not be the same as $f(t)$. Therefore a *periodic signal, by definition, must start at* $t = -\infty$ *and continuing forever, as illustrated in Fig. 1.6.*

Another important property of a periodic signal $f(t)$ is that $f(t)$ can be generated by **periodic extension** of any segment of $f(t)$ of duration T_0 (the period). As a result we can generate $f(t)$ from any segment of $f(t)$ with a duration of one period by placing this segment and the reproduction thereof end to end ad infinitum on either side. Figure 1.7 shows a periodic signal $f(t)$ of period $T_0 = 6$. The shaded portion of Fig. 1.7a shows a segment of $f(t)$ starting at $t = -1$ and having a duration of one period (6 seconds). This segment, when repeated forever in either direction, results in the periodic signal $f(t)$. Figure 1.7b shows another shaded segment of $f(t)$ of duration T_0 starting at $t = 0$. Again we see that this segment, when repeated forever on either side, results in $f(t)$. The reader can verify that this construction is possible with any segment of $f(t)$ starting at any instant as long as the segment duration is one period.

It is helpful to label signals that start at $t = -\infty$ and continue for ever as everlasting signals. Thus, an everlasting signal exists over the entire interval $-\infty < t < \infty$. The signals in Figs. 1.1b and 1.2b are examples of everlasting signals. Clearly, a periodic signal, by definition, is an everlasting signal.

A signal that does not start before $t = 0$ is a **causal** signal. In other words, $f(t)$ is a causal signal if

$$f(t) = 0 \qquad t < 0 \tag{1.7}$$

Signals in Figs. 1.3a, b, c, as well as in Figs. 1.9a and 1.9b are causal signals. A signal that starts before $t = 0$ is a **noncausal** signal. All the signals in Figs. 1.1 and 1.2 are noncausal. Observe that an everlasting signal is always noncausal but a noncausal signal is not necessarily everlasting. The everlasting signal in Fig. 1.2b is noncausal; however, the noncausal signal in Fig. 1.2a is not everlasting. A signal that is zero for all $t \geq 0$ is called an **anticausal** signal.

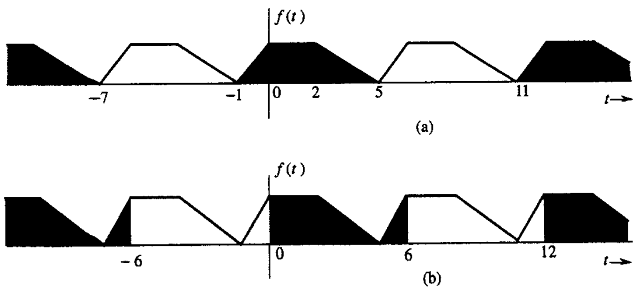


Fig. 1.7 Generation of a periodic signal by periodic extension of its segment of one-period duration.

Comment:

A true everlasting signal cannot be generated in practice for obvious reasons. Why should we bother to postulate such a signal? In later chapters we shall see that certain signals (including an everlasting sinusoid) which cannot be generated in practice *do* serve a very useful purpose in the study of signals and systems.

1.2-4 Energy and Power Signals

A signal with finite energy is an **energy signal**, and a signal with finite and nonzero power is a **power signal**. Signals in Fig. 1.2a and 1.2b are examples of energy and power signals, respectively. Observe that power is the time average of energy. Since the averaging is over an infinitely large interval, a signal with finite energy has zero power, and a signal with finite power has infinite energy. Therefore, a signal cannot both be an energy and a power signal. If it is one, it cannot be the other. On the other hand, there are signals that are neither energy nor power signals. The ramp signal is such an example.

Comments

All practical signals have finite energies and are therefore energy signals. A power signal must necessarily have infinite duration; otherwise its power, which is its energy averaged over an infinitely large interval, will not approach a (nonzero) limit. Clearly, it is impossible to generate a true power signal in practice because such a signal has infinite duration and infinite energy.

Also, because of periodic repetition, periodic signals for which the area under $|f(t)|^2$ over one period is finite are power signals; however, not all power signals are periodic.

△ **Exercise E1.4**

Show that an everlasting exponential e^{-at} is neither an energy nor a power signal for any real value of a. However, if a is imaginary, it is a power signal with power $P_f = 1$ regardless of the value of a. ▽

1.2-5 Deterministic and Random Signals

A signal whose physical description is known completely, either in a mathematical form or a graphical form, is a **deterministic signal**. A signal whose values cannot be predicted precisely but are known only in terms of probabilistic description, such as mean value, mean squared value, and so on is a **random signal**. In this book we shall exclusively deal with deterministic signals. Random signals are beyond the scope of this study.

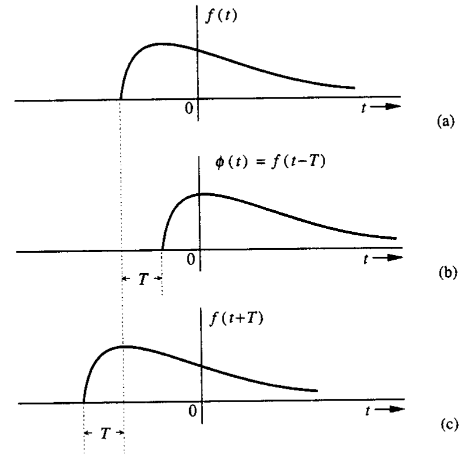


Fig. 1.8 Time shifting a signal.

1.3 Some Useful Signal Operations

We discuss here three useful signal operations: shifting, scaling, and inversion. Since the independent variable in our signal description is time, these operations are discussed as time shifting, time scaling, and time inversion (or folding). However, this discussion is valid for functions having independent variables other than time (e.g., frequency or distance).

1.3-1 Time Shifting

Consider a signal $f(t)$ (Fig. 1.8a) and the same signal delayed by T seconds (Fig. 1.8b), which we shall denote by $\phi(t)$. Whatever happens in $f(t)$ (Fig. 1.8a) at some instant t also happens in $\phi(t)$ (Fig. 1.8b) T seconds later at the instant $t+T$. Therefore

$$\phi(t+T) = f(t) \tag{1.8}$$

and

$$\phi(t) = f(t-T) \tag{1.9}$$

Therefore, to time-shift a signal by T, we replace t with $t-T$. Thus $f(t-T)$ represents $f(t)$ time-shifted by T seconds. If T is positive, the shift is to the right (delay). If T is negative, the shift is to the left (advance). Thus, $f(t-2)$ is $f(t)$ delayed (right-shifted) by 2 seconds, and $f(t+2)$ is $f(t)$ advanced (left-shifted) by 2 seconds.

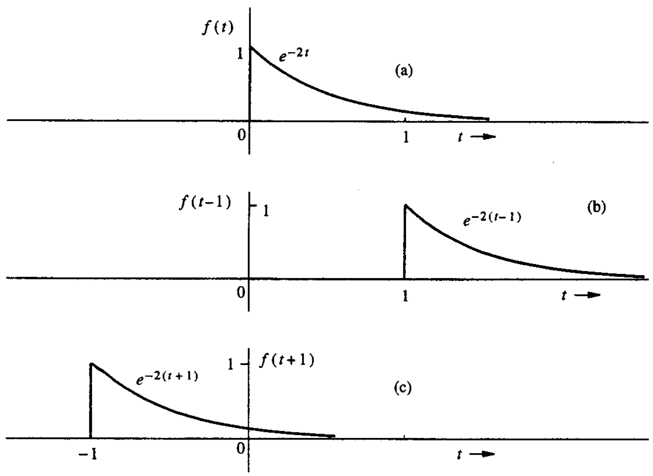


Fig. 1.9 (a) signal $f(t)$ (b) $f(t)$ delayed by 1 second (c) $f(t)$ advanced by 1 second.

■ **Example 1.3**

An exponential function $f(t) = e^{-2t}$ shown in Fig. 1.9a is delayed by 1 second. Sketch and mathematically describe the delayed function. Repeat the problem if $f(t)$ is advanced by 1 second.

The function $f(t)$ can be described mathematically as

$$f(t) = \begin{cases} e^{-2t} & t \geq 0 \\ 0 & t < 0 \end{cases} \tag{1.10}$$

Let $f_d(t)$ represent the function $f(t)$ delayed (right-shifted) by 1 second as illustrated in Fig. 1.9b. This function is $f(t-1)$; its mathematical description can be obtained from $f(t)$ by replacing t with $t-1$ in Eq. (1.10). Thus

$$f_d(t) = f(t-1) = \begin{cases} e^{-2(t-1)} & t-1 \geq 0 \quad \text{or} \quad t \geq 1 \\ 0 & t-1 < 0 \quad \text{or} \quad t < 1 \end{cases} \tag{1.11}$$

Let $f_a(t)$ represent the function $f(t)$ advanced (left-shifted) by 1 second as depicted in Fig. 1.9c. This function is $f(t+1)$; its mathematical description can be obtained from $f(t)$ by replacing t with $t+1$ in Eq. (1.10). Thus

$$f_a(t) = f(t+1) = \begin{cases} e^{-2(t+1)} & t+1 \geq 0 \quad \text{or} \quad t \geq -1 \\ 0 & t+1 < 0 \quad \text{or} \quad t < -1 \end{cases} \tag{1.12}$$

■

△ **Exercise E1.5**

Write a mathematical description of the signal $f_3(t)$ in Fig. 1.3c. This signal is delayed by 2 seconds. Sketch the delayed signal. Show that this delayed signal $f_d(t)$ can be described mathematically as $f_d(t) = 2(t-2)$ for $2 \leq t \leq 3$, and equal to 0 otherwise. Now repeat the procedure if the signal is advanced (left-shifted) by 1 second. Show that this advanced signal $f_a(t)$ can be described as $f_a(t) = 2(t+1)$ for $-1 \leq t \leq 0$, and equal to 0 otherwise. ▽

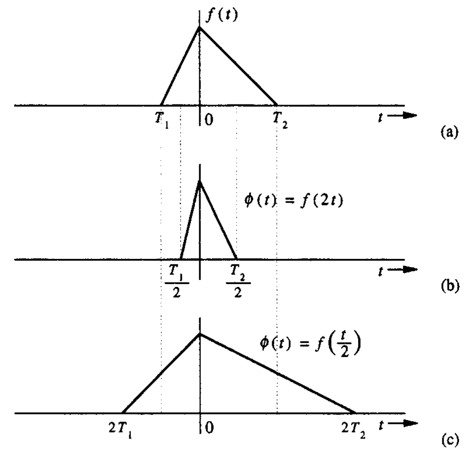


Fig. 1.10 Time scaling a signal.

1.3-2 Time Scaling

The compression or expansion of a signal in time is known as **time scaling**. Consider the signal $f(t)$ of Fig. 1.10a. The signal $\phi(t)$ in Fig. 1.10b is $f(t)$ compressed in time by a factor of 2. Therefore, whatever happens in $f(t)$ at some instant t also happens to $\phi(t)$ at the instant $t/2$, so that

$$\phi\left(\tfrac{t}{2}\right) = f(t) \tag{1.13}$$

and

$$\phi(t) = f(2t) \tag{1.14}$$

Observe that because $f(t) = 0$ at $t = T_1$ and T_2, we must have $\phi(t) = 0$ at $t = T_1/2$ and $T_2/2$, as shown in Fig. 1.10b. If $f(t)$ were recorded on a tape and played back at twice the normal recording speed, we would obtain $f(2t)$. In general, if $f(t)$ is compressed in time by a factor a $(a > 1)$, the resulting signal $\phi(t)$ is given by

$$\phi(t) = f(at) \tag{1.15}$$

Using a similar argument, we can show that $f(t)$ expanded (slowed down) in time by a factor a $(a > 1)$ is given by

$$\phi(t) = f\left(\tfrac{t}{a}\right) \tag{1.16}$$

Figure 1.10c shows $f(\frac{t}{2})$, which is $f(t)$ expanded in time by a factor of 2. Observe that in time scaling operation, the origin $t = 0$ is the anchor point, which remains unchanged under scaling operation because at $t = 0$, $f(t) = f(at) = f(0)$.

In summary, to time-scale a signal by a factor a, we replace t with at. If $a > 1$, the scaling results in compression, and if $a < 1$, the scaling results in expansion.

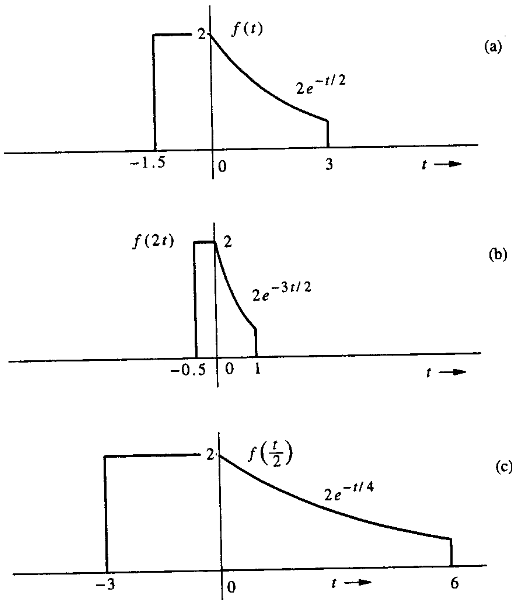


Fig. 1.11 (a) signal $f(t)$ (b) signal $f(3t)$ (c) signal $f(\frac{t}{2})$.

■ **Example 1.4**

Figure 1.11a shows a signal $f(t)$. Sketch and describe mathematically this signal time-compressed by factor 3. Repeat the problem for the same signal time-expanded by factor 2.

The signal $f(t)$ can be described as

$$f(t) = \begin{cases} 2 & -1.5 \le t < 0 \\ 2\,e^{-t/2} & 0 \le t < 3 \\ 0 & \text{otherwise} \end{cases} \tag{1.17}$$

Figure 1.11b shows $f_c(t)$, which is $f(t)$ time-compressed by factor 3; consequently, it can be described mathematically as $f(3t)$, which is obtained by replacing t with $3t$ in the right-hand side of Eq. 1.17. Thus

$$f_c(t) = f(3t) = \begin{cases} 2 & -1.5 \le 3t < 0 \quad \text{or} \quad -0.5 \le t < 0 \\ 2\,e^{-3t/2} & 0 \le 3t < 3 \quad \text{or} \quad 0 \le t < 1 \\ 0 & \text{otherwise} \end{cases} \tag{1.18a}$$

Observe that the instants $t = -1.5$ and 3 in $f(t)$ correspond to the instants $t = -0.5$, and 1 in the compressed signal $f(3t)$.

Figure 1.11c shows $f_e(t)$, which is $f(t)$ time-expanded by factor 2; consequently, it can be described mathematically as $f(t/2)$, which is obtained by replacing t with $t/2$ in $f(t)$. Thus

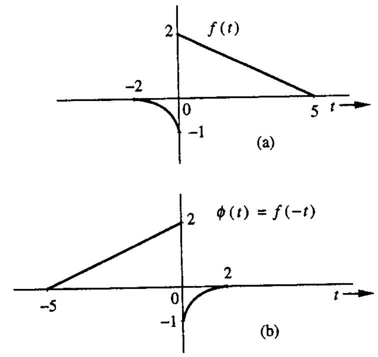


Fig. 1.12 Time inversion (reflection) of a signal.

$$f_e(t) = f\left(\frac{t}{2}\right) = \begin{cases} 2 & -1.5 \le \frac{t}{2} < 0 \quad \text{or} \quad -3 \le t < 0 \\ 2\,e^{-t/4} & 0 \le \frac{t}{2} < 3 \quad \text{or} \quad 0 \le t < 6 \\ 0 & \text{otherwise} \end{cases} \tag{1.18b}$$

Observe that the instants $t = -1.5$ and 3 in $f(t)$ correspond to the instants $t = -3$ and 6 in the expanded signal $f(\frac{t}{2})$. ■

△ **Exercise E1.6**

Show that the time-compression by a factor n ($n > 1$) of a sinusoid results in a sinusoid of the same amplitude and phase, but with the frequency increased n-fold. Similarly the time expansion by a factor n ($n > 1$) of a sinusoid results in a sinusoid of the same amplitude and phase, but with the frequency reduced by a factor n. Verify your conclusion by sketching a sinusoid sin $2t$ and the same sinusoid compressed by a factor 3 and expanded by a factor 2. ▽

1.3-3 Time Inversion (Time Reversal)

Consider the signal $f(t)$ in Fig. 1.12a. We can view $f(t)$ as a rigid wire frame hinged at the vertical axis. To time-invert $f(t)$, we rotate this frame 180° about the vertical axis. This time inversion or folding [the reflection of $f(t)$ about the vertical axis] gives us the signal $\phi(t)$ (Fig. 1.12b). Observe that whatever happens in Fig. 1.12a at some instant t also happens in Fig. 1.12b at the instant $-t$. Therefore

$$\phi(-t) = f(t)$$

and

$$\phi(t) = f(-t) \tag{1.19}$$

Therefore, to time-invert a signal we replace t with $-t$. Thus, the time inversion of signal $f(t)$ yields $f(-t)$. Consequently, the mirror image of $f(t)$ about the vertical axis is $f(-t)$. Recall also that the mirror image of $f(t)$ about the horizontal axis is $-f(t)$.

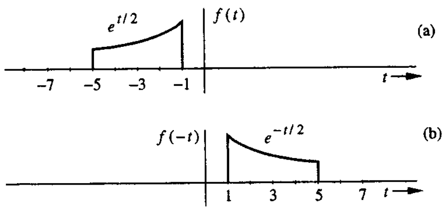


Fig. 1.13 An example of time inversion.

■ **Example 1.5**

For the signal $f(t)$ illustrated in Fig. 1.13a, sketch $f(-t)$, which is time inverted $f(t)$.

The instants -1 and -5 in $f(t)$ are mapped into instants 1 and 5 in $f(-t)$. Because $f(t) = e^{t/2}$, we have $f(-t) = e^{-t/2}$. The signal $f(-t)$ is depicted in Fig. 1.13b. We can describe $f(t)$ and $f(-t)$ as

$$f(t) = \begin{cases} e^{t/2} & -1 \geq t > -5 \\ 0 & \text{otherwise} \end{cases}$$

and its time inverted version $f(-t)$ is obtained by replacing t with $-t$ in $f(t)$ as

$$f(-t) = \begin{cases} e^{-t/2} & -1 \geq -t > -5 \quad \text{or} \quad 1 \leq t < 5 \\ 0 & \text{otherwise} \end{cases}$$ ■

1.3-4 Combined Operations

Certain complex operations require simultaneous use of more than one of the above operations. The most general operation involving all the three operations is $f(at - b)$, which is realized in two possible sequences of operation:

1. Time-shift $f(t)$ by b to obtain $f(t-b)$. Now time-scale the shifted signal $f(t-b)$ by a (that is, replace t with at) to obtain $f(at - b)$.
2. Time-scale $f(t)$ by a to obtain $f(at)$. Now time-shift $f(at)$ by $\frac{b}{a}$ (that is, replace t with $(t - \frac{b}{a})$) to obtain $f[a(t - \frac{b}{a})] = f(at - b)$. In either case, if a is negative, time scaling involves time inversion.

For instance, the signal $f(2t - 6)$ can be obtained in two ways: first, delay $f(t)$ by 6 to obtain $f(t - 6)$ and then time-compress this signal by factor 2 (replace t with $2t$) to obtain $f(2t - 6)$. Alternately, we first time-compress $f(t)$ by factor 2 to obtain $f(2t)$, then delay this signal by 3 (replace t with $t - 3$) to obtain $f(2t - 6)$.

1.4 Some Useful Signal Models

In the area of signals and systems, the step, the impulse, and the exponential functions are very useful. They not only serve as a basis for representing other signals, but their use can simplify many aspects of the signals and systems.

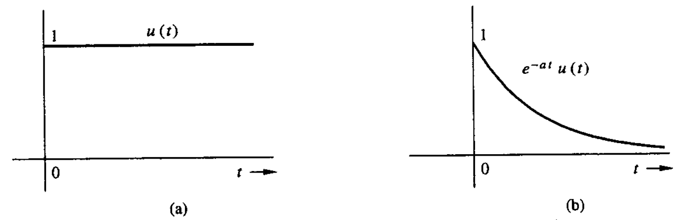


Fig. 1.14 (a) Unit step function $u(t)$ (b) exponential $e^{-at}u(t)$.

1. Unit Step Function u(t)

In much of our discussion, the signals begin at $t = 0$ (causal signals). Such signals can be conveniently described in terms of unit step function $u(t)$ shown in Fig. 1.14a. This function is defined by

$$u(t) = \begin{cases} 1 & t \geq 0 \\ 0 & t < 0 \end{cases} \tag{1.20}$$

If we want a signal to start at $t = 0$ (so that it has a value of zero for $t < 0$), we only need to multiply the signal with $u(t)$. For instance, the signal e^{-at} represents an everlasting exponential that starts at $t = -\infty$. The causal form of this exponential illustrated in Fig. 1.14b can be described as $e^{-at}u(t)$.

The unit step function also proves very useful in specifying a function with different mathematical descriptions over different intervals. Examples of such functions appear in Fig. 1.11. These functions have different mathematical descriptions over different segments of time as seen from Eqs. (1.17), (1.18a), and (1.18b). Such a description often proves clumsy and inconvenient in mathematical treatment. Using the unit step function, we can describe such functions by a single expression that is valid for all t.

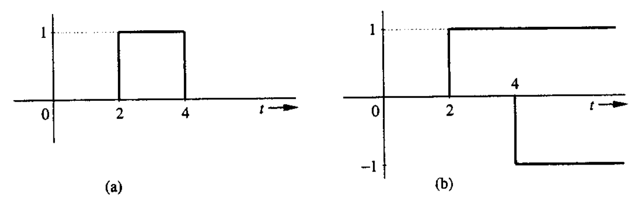


Fig. 1.15 Representation of a rectangular pulse by step functions.

Consider, for example, the rectangular pulse depicted in Fig. 1.15a. We can express such a pulse in terms of familiar step functions by observing that the pulse $f(t)$ can be expressed as the sum of the two delayed unit step functions as shown in Fig. 1.15b. The unit step function $u(t)$ delayed by T seconds is $u(t - T)$. From Fig. 1.15b, it is clear that

$$f(t) = u(t - 2) - u(t - 4)$$

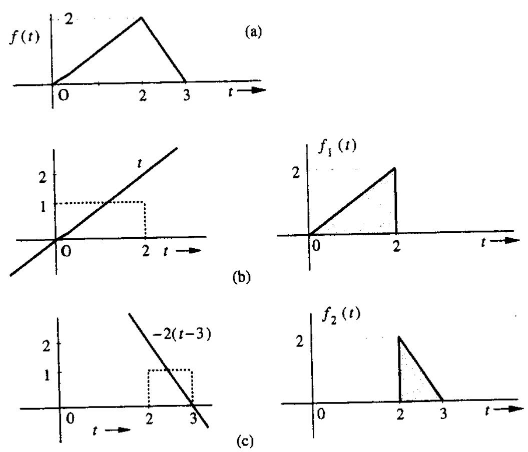


Fig. 1.16 Representation of a signal defined interval by interval.

■ **Example 1.6**

Describe the signal in Fig. 1.16a.

The signal illustrated in Fig. 1.16a can be conveniently handled by breaking it up into the two components $f_1(t)$ and $f_2(t)$, depicted in Figs. 1.16b and 1.16c respectively. Here, $f_1(t)$ can be obtained by multiplying the ramp t by the gate pulse $u(t) - u(t-2)$, as shown in Fig. 1.16b. Therefore

$$f_1(t) = t\,[u(t) - u(t-2)]$$

The signal $f_2(t)$ can be obtained by multiplying another ramp by the gate pulse illustrated in Fig. 1.16c. This ramp has a slope -2; hence it can be described by $-2t+c$. Now, because the ramp has a zero value at $t = 3$, the constant $c = 6$, and the ramp can be described by $-2(t-3)$. Also, the gate pulse in Fig. 1.16c is $u(t-2) - u(t-3)$. Therefore

$$f_2(t) = -2(t-3)\,[u(t-2) - u(t-3)]$$

and

$$\begin{aligned} f(t) &= f_1(t) + f_2(t) \\ &= t\,[u(t) - u(t-2)] - 2(t-3)\,[u(t-2) - u(t-3)] \\ &= tu(t) - 3(t-2)u(t-2) + 2(t-3)u(t-3) \end{aligned}$$ ■

■ **Example 1.7**

Describe the signal in Fig. 1.11a by a single expression valid for all t.

Over the interval from -1.5 to 0, the signal can be described by a constant 2, and over the interval from 0 to 3, it can be described by $2e^{-t/2}$. Therefore

$$\begin{aligned} f(t) &= \underbrace{2[u(t+1.5) - u(t)]}_{f_1(t)} + \underbrace{2e^{-t/2}[u(t) - u(t-3)]}_{f_2(t)} \\ &= 2u(t+1.5) - 2(1 - e^{-t/2})u(t) - 2e^{-t/2}u(t-3) \end{aligned}$$

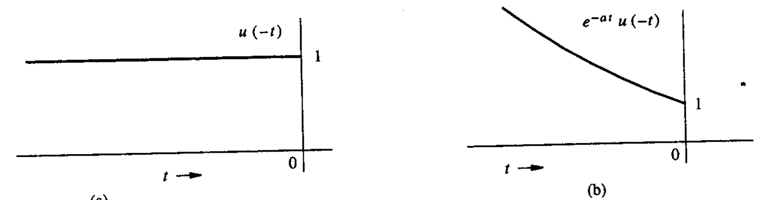


Fig. 1.17 The Signal for Exercise E1.7.

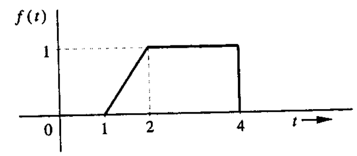


Fig. 1.18 The signal for Exercise E1.8.

Compare this expression with the expression for the same function found in Eq. 1.17. ■

△ **Exercise E1.7**

Show that the signals depicted in Figs. 1.17a and 1.17b can be described as $u(-t)$, and $e^{-at}u(-t)$, respectively. ▽

△ **Exercise E1.8**

Show that the signal shown in Fig. 1.18 can be described as

$$f(t) = (t-1)u(t-1) - (t-2)u(t-2) - u(t-4)$$ ▽

2. The Unit Impulse Function $\delta(t)$

The unit impulse function $\delta(t)$ is one of the most important functions in the study of signals and systems. This function was first defined by P. A. M Dirac as

$$\begin{aligned} &\delta(t) = 0 \qquad t \neq 0 \\ &\int_{-\infty}^{\infty} \delta(t)\,dt = 1 \end{aligned} \tag{1.21}$$

We can visualize an impulse as a tall, narrow rectangular pulse of unit area, as illustrated in Fig. 1.19b. The width of this rectangular pulse is a very small value $\epsilon \to 0$. Consequently, its height is a very large value $1/\epsilon$. The unit impulse therefore can be regarded as a rectangular pulse with a width that has become infinitesimally small, a height that has become infinitely large, and an overall area that has been maintained at unity. Thus $\delta(t) = 0$ everywhere except at $t = 0$, where it is undefined. For this reason a unit impulse is represented by the spear-like symbol in Fig. 1.19a.

Other pulses, such as exponential pulse, triangular pulse, or Gaussian pulse may also be used in impulse approximation. The important feature of the unit impulse function is not its shape but the fact that its effective duration (pulse width)

Fig. 1.19 A unit impulse and its approximation.

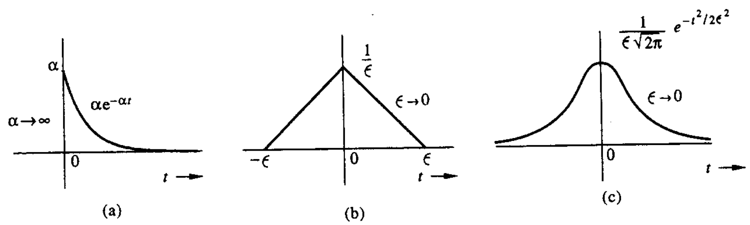


Fig. 1.20 Other possible approximations to a unit impulse.

approaches zero while its area remains at unity. For example, the exponential pulse $\alpha e^{-\alpha t}u(t)$ in Fig. 1.20a becomes taller and narrower as α increases. In the limit as $\alpha \to \infty$, the pulse height $\to \infty$, and its width or duration $\to 0$. Yet, the area under the pulse is unity regardless of the value of α because

$$\int_0^\infty \alpha e^{-\alpha t}\, dt = 1 \tag{1.22}$$

The pulses in Figs. 1.20b and 1.20c behave in a similar fashion.

From Eq. (1.21), it follows that the function $k\delta(t) = 0$ for all $t \neq 0$, and its area is k. Thus, $k\delta(t)$ is an impulse function whose area is k (in contrast to the unit impulse function, whose area is 1).

Multiplication of a Function by an Impulse

Let us now consider what happens when we multiply the unit impulse $\delta(t)$ by a function $\phi(t)$ that is known to be continuous at $t = 0$. Since the impulse exists only at $t = 0$, and the value of $\phi(t)$ at $t = 0$ is $\phi(0)$, we obtain

$$\phi(t)\delta(t) = \phi(0)\delta(t) \tag{1.23a}$$

Similarly, if $\phi(t)$ is multiplied by an impulse $\delta(t - T)$ (impulse located at $t = T$), then

$$\phi(t)\delta(t - T) = \phi(T)\delta(t - T) \tag{1.23b}$$

provided $\phi(t)$ is continuous at $t = T$.

Sampling Property of the Unit Impulse Function

From Eq. (1.23a) it follows that

$$\begin{aligned}\int_{-\infty}^{\infty} \phi(t)\delta(t)\, dt &= \phi(0) \int_{-\infty}^{\infty} \delta(t)\, dt \\ &= \phi(0)\end{aligned} \tag{1.24a}$$

provided $\phi(t)$ is continuous at $t = 0$. This result means that *the area under the product of a function with an impulse $\delta(t)$ is equal to the value of that function at the instant where the unit impulse is located.* This property is very important and useful, and is known as the **sampling** or **sifting property** of the unit impulse.

From Eq. (1.23b) it follows that

$$\int_{-\infty}^{\infty} \phi(t)\delta(t - T)\, dt = \phi(T) \tag{1.24b}$$

Equation (1.24b) is just another form of sampling or sifting property. In the case of Eq. (1.24b), the impulse $\delta(t - T)$ is located at $t = T$. Therefore, the area under $\phi(t)\delta(t - T)$ is $\phi(T)$, the value of $\phi(t)$ at the instant where the impulse is located (at $t = T$). In these derivations we have assumed that the function is continuous at the instant where the impulse is located.

Unit Impulse as a Generalized Function

The definition of the unit impulse function given in Eq. (1.21) is not mathematically rigorous, which leads to serious difficulties. First, the impulse function does not define a unique function: for example, it can be shown that $\delta(t) + \dot{\delta}(t)$ also satisfies Eq. (1.21).[1] Moreover, $\delta(t)$ is not even a true function in the ordinary sense. An ordinary function is specified by its values for all time t. The impulse function is zero everywhere except at $t = 0$, and at this only interesting part of its range it is undefined. These difficulties are resolved by defining the impulse as a generalized function rather than an ordinary function. A **generalized function** is defined by its effect on other functions instead of by its value at every instant of time.

In this approach the impulse function is defined by the sampling property [Eq. (1.24)]. We say nothing about what the impulse function is or what it looks like. Instead, the impulse function is defined in terms of its effect on a test function $\phi(t)$. We define a unit impulse as a function for which the area under its product with a function $\phi(t)$ is equal to the value of the function $\phi(t)$ at the instant where the impulse is located. It is assumed that $\phi(t)$ is continuous at the location of the impulse. Therefore, either Eq. (1.24a) or (1.24b) can serve as a definition of the impulse function in this approach. Recall that the sampling property [Eq. (1.24)] is the consequence of the classical (Dirac) definition of impulse in Eq. (1.21). In contrast, *the sampling property [Eq. (1.24)] defines the impulse function in the generalized function approach.*

We now present an interesting application of the generalized function definition of an impulse. Because the unit step function $u(t)$ is discontinuous at $t = 0$, its derivative du/dt does not exist at $t = 0$ in the ordinary sense. We now show that this derivative *does* exist in the generalized sense, and it is, in fact, $\delta(t)$. As a proof, let us evaluate the integral of $(du/dt)\phi(t)$, using integration by parts:

$$\int_{-\infty}^{\infty} \frac{du}{dt}\phi(t)\,dt = u(t)\phi(t)\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} u(t)\dot{\phi}(t)\,dt \tag{1.25}$$

$$= \phi(\infty) - 0 - \int_{0}^{\infty} \dot{\phi}(t)\,dt$$

$$= \phi(\infty) - \phi(t)|_0^\infty$$

$$= \phi(0) \tag{1.26}$$

This result shows that du/dt satisfies the sampling property of $\delta(t)$. Therefore it is an impulse $\delta(t)$ in the generalized sense—that is,

$$\frac{du}{dt} = \delta(t) \tag{1.27}$$

Consequently

$$\int_{-\infty}^{t} \delta(\tau)\,d\tau = u(t) \tag{1.28}$$

These results can also be obtained graphically from Fig. 1.19b. We observe that the area from $-\infty$ to t under the limiting form of $\delta(t)$ in Fig. 1.19b is zero if $t < 0$ and unity if $t \geq 0$. Consequently

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

$$= u(t) \tag{1.29}$$

Derivatives of impulse function can also be defined as generalized functions (see Prob. 1.4-10).

△ **Exercise E1.9**
Show that

(a) $(t^3 + 3)\delta(t) = 3\delta(t)$ (b) $\left[\sin\left(t^2 - \frac{\pi}{2}\right)\right]\delta(t) = -\delta(t)$

(c) $e^{-2t}\delta(t) = \delta(t)$ (d) $\frac{\omega^2+1}{\omega^2+9}\delta(\omega - 1) = \frac{1}{5}\delta(\omega - 1)$

Hint: Use Eqs. (1.23). ▽

△ **Exercise E1.10**
Show that

(a) $\int_{-\infty}^{\infty} \delta(t)e^{-j\omega t}\,dt = 1$ (b) $\int_{-\infty}^{\infty} \delta(t-2)\cos\left(\frac{\pi t}{4}\right)dt = 0$

(c) $\int_{-\infty}^{\infty} e^{-2(x-t)}\delta(2-t)\,dt = e^{-2(x-2)}$

Hint: In part c recall that $\delta(x)$ is located at $x = 0$. Therefore $\delta(2-t)$ is located at $2 - t = 0$; that is at $t = 2$. ▽

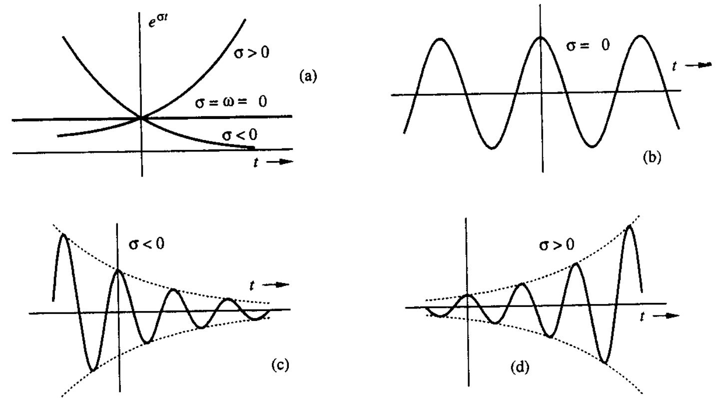


Fig. 1.21 Sinusoids of complex frequency $\sigma + j\omega$.

3. The Exponential Function e^{st}

One of the most important functions in the area of signals and systems is the exponential signal e^{st}, where s is complex in general, given by

$$s = \sigma + j\omega$$

Therefore

$$e^{st} = e^{(\sigma+j\omega)t} = e^{\sigma t}e^{j\omega t} = e^{\sigma t}(\cos\,\omega t + j\sin\,\omega t) \tag{1.30a}$$

If $s^* = \sigma - j\omega$ (the conjugate of s), then

$$e^{s^*t} = e^{\sigma - j\omega} = e^{\sigma t}e^{-j\omega t} = e^{\sigma t}(\cos\,\omega t - j\sin\,\omega t) \tag{1.30b}$$

and

$$e^{\sigma t}\cos\,\omega t = \frac{1}{2}(e^{st} + e^{s^*t}) \tag{1.30c}$$

Comparison of this equation with Euler's formula shows that e^{st} is a generalization of the function $e^{j\omega t}$, where the frequency variable $j\omega$ is generalized to a complex variable $s = \sigma + j\omega$. For this reason we designate the variable s as the **complex frequency**. From Eqs. (1.30) it follows that the function e^{st} encompasses a large class of functions. The following functions are special cases of e^{st}:

1 A constant $k = ke^{0t}$ $(s = 0)$
2 A monotonic exponential $e^{\sigma t}$ $(\omega = 0,\ s = \sigma)$
3 A sinusoid $\cos\,\omega t$ $(\sigma = 0,\ s = \pm j\omega)$
4 An exponentially varying sinusoid $e^{\sigma t}\cos\,\omega t$ $(s = \sigma \pm j\omega)$

These functions are illustrated in Fig. 1.21.

The complex frequency s can be conveniently represented on a **complex frequency plane** (s plane) as depicted in Fig. 1.22. The horizontal axis is the real axis (σ axis), and the vertical axis is the imaginary axis ($j\omega$ axis). The absolute value of the imaginary part of s is $|\omega|$ (the *radian* frequency), which indicates the frequency

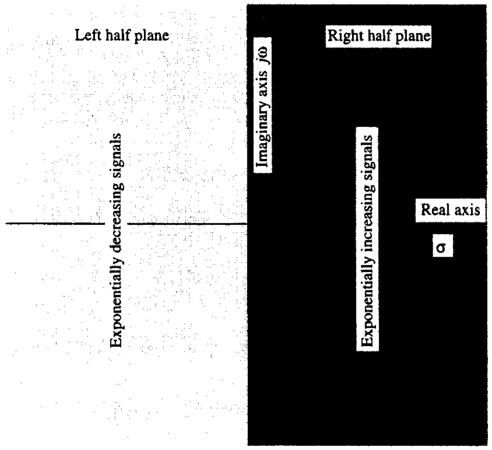


Fig. 1.22 Complex frequency plane.

of oscillation of e^{st}; the real part σ (the *neper* frequency) gives information about the rate of increase or decrease of the amplitude of e^{st}. For signals whose complex frequencies lie on the real axis (σ-axis, where $\omega = 0$), the frequency of oscillation is zero. Consequently these signals are monotonically increasing or decreasing exponentials (Fig. 1.21a). For signals whose frequencies lie on the imaginary axis ($j\omega$ axis where $\sigma = 0$), $e^{\sigma t} = 1$. Therefore, these signals are conventional sinusoids with constant amplitude (Fig. 1.21b). The case $s = 0$ ($\sigma = \omega = 0$) corresponds to a constant (dc) signal because $e^{0t} = 1$. For the signals illustrated in Figs. 1.21c and 1.21d, both σ and ω are nonzero; the frequency s is complex and does not lie on either axis. The signal in Fig. 1.21c decays exponentially. Therefore, σ is negative, and s lies to the left of the imaginary axis. In contrast, the signal in Fig. 1.21d *grows* exponentially. Therefore, σ is positive, and s lies to the right of the imaginary axis. Thus the s-plane (Fig. 1.21) can be differentiated into two parts: the **left half-plane** (LHP) corresponding to exponentially decaying signals and the **right half-plane** (RHP) corresponding to exponentially growing signals. The imaginary axis separates the two regions and corresponds to signals of constant amplitude.

An exponentially growing sinusoid $e^{2t}\cos(5t+\theta)$, for example, can be expressed as a sum of exponentials $e^{(2+j5)t}$ and $e^{(2-j5)t}$ with complex frequencies $2+j5$ and $2-j5$, respectively, which lie in the RHP. An exponentially decaying sinusoid $e^{-2t}\cos(5t+\theta)$ can be expressed as a sum of exponentials $e^{(-2+j5)t}$ and $e^{(-2-j5)t}$ with complex frequencies $-2+j5$ and $-2-j5$, respectively, which lie in the LHP. A constant amplitude sinusoid $\cos(5t+\theta)$ can be expressed as a sum of exponentials e^{j5t} and e^{-j5t} with complex frequencies $\pm j5$, which lie on the imaginary axis. Observe that the monotonic exponentials $e^{\pm 2t}$ are also generalized sinusoids with complex frequencies ± 2.

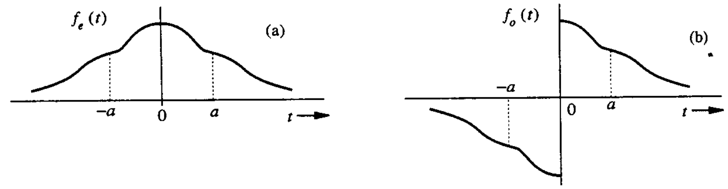


Fig. 1.23 An even and an odd function of t.

1.5 Even and Odd Functions

A function $f_e(t)$ is said to be an **even function** of t if

$$f_e(t) = f_e(-t) \tag{1.31}$$

and a function $f_o(t)$ is said to be an **odd function** of t if

$$f_o(t) = -f_o(-t) \tag{1.32}$$

An even function has the same value at the instants t and $-t$ for all values of t. Clearly, $f_e(t)$ is symmetrical about the vertical axis, as shown in Fig. 1.23a. On the other hand, the value of an odd function at the instant t is the negative of its value at the instant $-t$. Therefore, $f_o(t)$ is anti-symmetrical about the vertical axis, as depicted in Fig. 1.23b.

1.5-1 Some Properties of Even and Odd Functions

Even and odd functions have the following property:

even function × odd function = odd function
odd function × odd function = even function
even function × even function = even function

The proofs of these facts are trivial and follow directly from the definition of odd and even functions [Eqs. (1.31) and (1.32)].

Area

Because $f_e(t)$ is symmetrical about the vertical axis, it follows from Fig. 1.23a that

$$\int_{-a}^{a} f_e(t)\,dt = 2\int_{0}^{a} f_e(t)\,dt \tag{1.33a}$$

It is also clear from Fig. 1.23b that

$$\int_{-a}^{a} f_o(t)\,dt = 0 \tag{1.33b}$$

These results can also be proved formally by using the definitions in Eqs. (1.31) and (1.32). We leave them as an exercise for the reader.

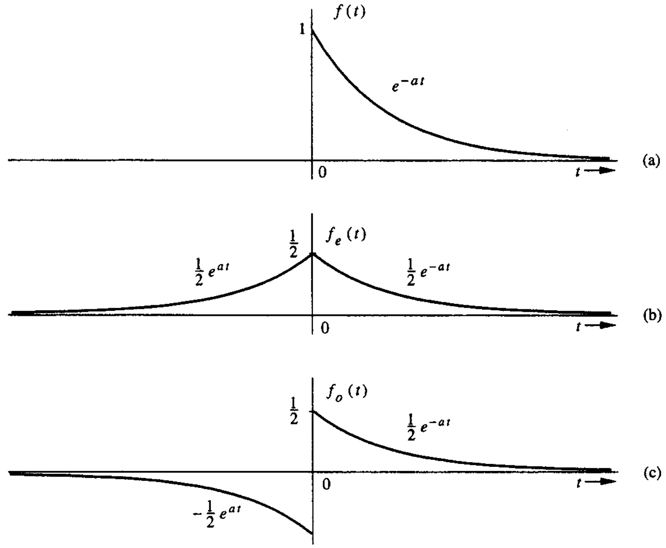


Fig. 1.24 Finding an even and odd components of a signal.

1.5-2 Even and Odd Components of a Signal

Every signal $f(t)$ can be expressed as a sum of even and odd components because

$$f(t) = \underbrace{\tfrac{1}{2}\left[f(t) + f(-t)\right]}_{\text{even}} + \underbrace{\tfrac{1}{2}\left[f(t) - f(-t)\right]}_{\text{odd}} \tag{1.34}$$

From the definitions in Eqs. (1.31) and (1.32), we can clearly see that the first component on the right-hand side is an even function, while the second component is odd. This is apparent from the fact that replacing t by $-t$ in the first component yields the same function. The same maneuver in the second component yields the negative of that component.

Consider the function

$$f(t) = e^{-at}u(t)$$

Expressing this function as a sum of the even and odd components $f_e(t)$ and $f_o(t)$, we obtain

$$f(t) = f_e(t) + f_o(t)$$

where [from Eq. (1.34)]

$$f_e(t) = \tfrac{1}{2}\left[e^{-at}u(t) + e^{at}u(-t)\right] \tag{1.35a}$$

and

$$f_o(t) = \tfrac{1}{2}\left[e^{-at}u(t) - e^{at}u(-t)\right] \tag{1.35b}$$

The function $e^{-at}u(t)$ and its even and odd components are illustrated in Fig. 1.24.

■ **Example 1.8**

Find the even and odd components of e^{jt}.

From Eq. (1.34)

$$e^{jt} = f_e(t) + f_o(t)$$

where

$$f_e(t) = \tfrac{1}{2}\left[e^{jt} + e^{-jt}\right] = \cos t$$

and

$$f_o(t) = \tfrac{1}{2}\left[e^{jt} - e^{-jt}\right] = j \sin t$$ ■

1.6 Systems

As mentioned in Sec. 1.1, systems are used to process signals in order to modify or to extract additional information from the signals. A system may consist of physical components (hardware realization) or may consist of an algorithm that computes the output signal from the input signal (software realization).

A system is characterized by its **inputs**, its **outputs** (or **responses**), and the **rules of operation** (or **laws**) adequate to describe its behavior. For example, in electrical systems, the laws of operation are the familiar voltage-current relationships for the resistors, capacitors, inductors, transformers, transistors, and so on, as well as the laws of interconnection (i.e., Kirchhoff's laws). Using these laws, we derive mathematical equations relating the outputs to the inputs. These equations then represent a **mathematical model** of the system. Thus a system is characterized by its inputs, its outputs, and its mathematical model.

A system can be conveniently illustrated by a "black box" with one set of accessible terminals where the input variables $f_1(t), f_2(t), \ldots, f_j(t)$ are applied and another set of accessible terminals where the output variables $y_1(t), y_2(t), \ldots, y_k(t)$ are observed. Note that the direction of the arrows for the variables in Fig. 1.25 is always from cause to effect.

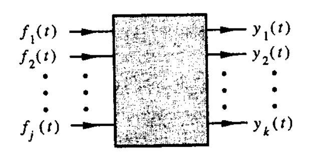


Fig. 1.25 Representation of a system.

The study of systems consists of three major areas: mathematical modeling, analysis, and design. Although we shall be dealing with mathematical modeling, our main concern is with analysis and design. The major portion of this book is devoted to the analysis problem—how to determine the system outputs for the given inputs and a given mathematical model of the system (or rules governing the system). To a lesser extent, we will also consider the problem of design or synthesis—how to construct a system which will produce a desired set of outputs for the given inputs.

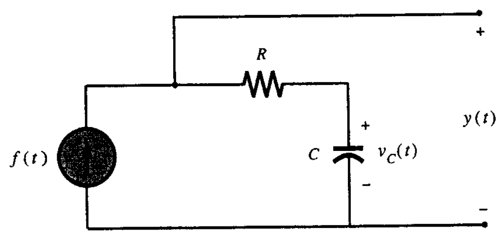


Fig. 1.26 An example of a simple electrical system.

Data Needed to Compute System Response

In order to understand what data we need to compute a system response, consider a simple RC circuit with a current source $f(t)$ as its input (Fig. 1.26). The output voltage $y(t)$ is given by

$$y(t) = Rf(t) + \frac{1}{C}\int_{-\infty}^{t} f(\tau)\,d\tau \tag{1.36a}$$

The limits of the integral on the right-hand side are from $-\infty$ to t because this integral represents the capacitor charge due to the current $f(t)$ flowing in the capacitor, and this charge is the result of the current flowing in the capacitor from $-\infty$. Now, Eq. (1.36a) can be expressed as

$$y(t) = Rf(t) + \frac{1}{C}\int_{-\infty}^{0} f(\tau)\,d\tau + \frac{1}{C}\int_{0}^{t} f(\tau)\,d\tau \tag{1.36b}$$

The middle term on the right-hand side is $v_C(0)$, the capacitor voltage at $t = 0$. Therefore

$$y(t) = v_C(0) + Rf(t) + \frac{1}{C}\int_{0}^{t} f(\tau)\,d\tau \tag{1.36c}$$

This equation can be readily generalized as

$$y(t) = v_C(t_0) + Rf(t) + \frac{1}{C}\int_{t_0}^{t} f(\tau)\,d\tau \tag{1.36d}$$

From Eq. (1.36a), the output voltage $y(t)$ at an instant t can be computed if we know the input current flowing in the capacitor throughout its entire past ($-\infty$ to t). Alternatively, if we know the input current $f(t)$ from some moment t_0 onward, then, using Eq. (1.36d), we can still calculate $y(t)$ for $t \geq t_0$ from a knowledge of the input current, provided we know $v_C(t_0)$, the initial capacitor voltage (voltage at t_0). Thus $v_C(t_0)$ contains all the relevant information about the circuit's entire past ($-\infty$ to t_0) that we need to compute $y(t)$ for $t \geq t_0$. Therefore, the response of a system at $t > t_0$ can be determined from its input(s) during the interval t_0 to t and from certain **initial conditions** at $t = t_0$.

In the preceding example, we needed only one initial condition. However, in more complex systems, several initial conditions may be necessary. We know, for example, that in passive RLC networks, the initial values of all inductor currents and all capacitor voltages† are needed to determine the outputs at any instant $t \geq 0$ if the inputs are given over the interval $[0, t]$.

1.7 Classification of Systems

Systems may be classified broadly in the following categories:‡

1. Linear and nonlinear systems;
2. Constant-parameter and time-varying-parameter systems;
3. Instantaneous (memoryless) and dynamic (with memory) systems;
4. Causal and noncausal systems;
5. Lumped-parameter and distributed-parameter systems;
6. Continuous-time and discrete-time systems;
7. Analog and Digital systems;

1.7-1 Linear and Nonlinear Systems

The Concept of Linearity

A system whose output is proportional to its input is an *example* of a linear system. But linearity implies more than this; it also implies **additivity property**, implying that if several causes are acting on a system, then the total effect on the system due to all these causes can be determined by considering each cause separately while assuming all the other causes to be zero. The total effect is then the sum of all the component effects. This property may be expressed as follows: for a linear system, if a cause c_1 acting alone has an effect e_1, and if another cause c_2, also acting alone, has an effect e_2, then, with both causes acting on the system, the total effect will be $e_1 + e_2$. Thus, if

$$c_1 \longrightarrow e_1 \quad \text{and} \quad c_2 \longrightarrow e_2 \tag{1.37}$$

then for all c_1 and c_2

$$c_1 + c_2 \longrightarrow e_1 + e_2 \tag{1.38}$$

In addition, a linear system must satisfy the **homogeneity** or scaling property, which states that for arbitrary real or imaginary number k, if a cause is increased k-fold, the effect also increases k-fold. Thus, if

$$c \longrightarrow e$$

then for all real or imaginary k

$$kc \longrightarrow ke \tag{1.39}$$

Thus, linearity implies two properties: homogeneity (scaling) and additivity$. Both these properties can be combined into one property (**superposition**), which is expressed as follows: If

$$c_1 \longrightarrow e_1 \quad \text{and} \quad c_2 \longrightarrow e_2$$

then for all values of constants k_1 and k_2,

† Strictly speaking, independent inductor currents and capacitor voltages.

‡ Other classifications, such as deterministic and probabilistic systems, are beyond the scope of this text and are not considered.

$ A linear system must also satisfy the additional condition of **smoothness**, where small changes in the system's inputs must result in small changes in its outputs.[2]

$$k_1c_1 + k_2c_2 \longrightarrow k_1e_1 + k_2e_2 \tag{1.40}$$

This is true for all c_1 and c_2.

It may appear that additivity implies homogeneity. Unfortunately, there are cases where homogeneity does not follow from additivity. See the case in Exercise E1.11 below.

△ **Exercise E1.11**

Show that a system with the input (cause) $c(t)$ and the output (effect) $e(t)$ related by $e(t) = \text{Re}\{c(t)\}$ satisfies the additivity property but violates the homogeneity property. Hence, such a system is not linear.

Hint: show that Eq. (1.39) is not satisfied when k is complex. ▽

Response of a Linear System

For the sake of simplicity, we discuss below only **single-input, single-output (SISO)** systems. But the discussion can be readily extended to **multiple-input, multiple-output (MIMO)** systems.

A system's output for $t \geq 0$ is the result of two independent causes: the initial conditions of the system (or the system state) at $t = 0$ and the input $f(t)$ for $t \geq 0$. If a system is to be linear, the output must be the sum of the two components resulting from these two causes: first, the **zero-input response** component that results only from the initial conditions at $t = 0$ with the input $f(t) = 0$ for $t \geq 0$, and then the **zero-state response** component that results only from the input $f(t)$ for $t \geq 0$ when the initial conditions (at $t = 0$) are assumed to be zero. When all the appropriate initial conditions are zero, the system is said to be in **zero state.** The system output is zero when the input is zero only if the system is in zero state.

In summary, a linear system response can be expressed as the sum of a zero-input and a zero-state component:

$$\textbf{Total response = zero-input response + zero-state response} \tag{1.41}$$

This property of linear systems which permits the separation of an output into components resulting from the initial conditions and from the input is called the **decomposition property.**

For the RC circuit of Fig. 1.26, the response $y(t)$ was found to be [see Eq. (1.36c)]

$$y(t) = \underbrace{v_C(0)}_{z-i \text{ component}} + \underbrace{Rf(t) + \frac{1}{C}\int_0^t f(\tau)\,d\tau}_{z-s \text{ component}} \tag{1.42}$$

From Eq. (1.42), it is clear that if the input $f(t) = 0$ for $t \geq 0$, the output $y(t) = v_C(0)$. Hence $v_C(0)$ is the zero-input component of the response $y(t)$. Similarly, if the system state (the voltage v_C in this case) is zero at $t = 0$, the output is given by the second component on the right-hand side of Eq. (1.42). Clearly this is the zero-state component of the response $y(t)$.

In addition to the decomposition property, linearity implies that both the zero-input and zero-state components must obey the principle of superposition with respect to each of their respective causes. For example, if we increase the initial condition k-fold, the zero-input component must also increase k-fold. Similarly, if we increase the input k-fold, the zero-state component must also increase k-fold.

These facts can be readily verified from Eq. (1.42) for the RC circuit in Fig. 1.26. For instance, if we double the initial condition $v_C(0)$, the zero-input component doubles; if we double the input $f(t)$, the zero-state component doubles. ▪

■ **Example 1.9**

Show that the system described by the equation

$$\frac{dy}{dt} + 3y(t) = f(t) \tag{1.43}$$

is linear.

Let the system response to the inputs $f_1(t)$ and $f_2(t)$ be $y_1(t)$ and $y_2(t)$, respectively. Then

$$\frac{dy_1}{dt} + 3y_1(t) = f_1(t)$$

and

$$\frac{dy_2}{dt} + 3y_2(t) = f_2(t)$$

Multiplying the first equation by k_1, the second with k_2, and adding them yields

$$\frac{d}{dt}\left[k_1y_1(t) + k_2y_2(t)\right] + 3\left[k_1y_1(t) + k_2y_2(t)\right] = k_1f_1(t) + k_2f_2(t)$$

But this equation is the system equation [Eq. (1.43)] with

$$f(t) = k_1f_1(t) + k_2f_2(t)$$

and

$$y(t) = k_1y_1(t) + k_2y_2(t)$$

Therefore, when the input is $k_1f_1(t) + k_2f_2(t)$, the system response is $k_1y_1(t) + k_2y_2(t)$. Consequently, the system is linear. Using this argument, we can readily generalize the result to show that a system described by a differential equation of the form

$$\frac{d^ny}{dt^n} + a_{n-1}\frac{d^{n-1}y}{dt^{n-1}} + \cdots + a_0y = b_m\frac{d^mf}{dt^m} + \cdots + b_1\frac{df}{dt} + b_0f \tag{1.44}$$

is a linear system. The coefficients a_i and b_i in this equation can be constants or functions of time. ■

△ **Exercise E1.12**

Show that the system described by the following equation is linear:

$$\frac{dy}{dt} + t^2y(t) = (2t+3)f(t) \quad ▽$$

△ **Exercise E1.13**

Show that a system described by the following equation is nonlinear:

$$y(t)\frac{dy}{dt} + 3y(t) = f(t) \quad ▽$$

More Comments on Linear Systems

Almost all systems observed in practice become nonlinear when large enough signals are applied to them. However, many systems show linear behavior for small signals. The analysis of nonlinear systems is generally difficult. Nonlinearities can arise in so many ways that describing them with a common mathematical form is impossible. Not only is each system a category in itself, but even for a given system, changes in initial conditions or input amplitudes may change the nature of the problem. On the other hand, the superposition property of linear systems is a powerful unifying principle which allows for a general solution. The superposition property (linearity) greatly simplifies the analysis of linear systems. Because of the decomposition property, we can evaluate separately the two components of the output. The zero-input component can be computed by assuming the input to be zero, and the zero-state component can be computed by assuming zero initial conditions. Moreover, if we express an input $f(t)$ as a sum of simpler functions,

$$f(t) = a_1 f_1(t) + a_2 f_2(t) + \cdots + a_m f_m(t)$$

then, by virtue of linearity, the response $y(t)$ is given by

$$y(t) = a_1 y_1(t) + a_2 y_2(t) + \cdots + a_m y_m(t) \tag{1.45}$$

where $y_k(t)$ is the zero-state response to an input $f_k(t)$. This apparently trivial observation has profound implications. As we shall see repeatedly in later chapters, it proves extremely useful and opens new avenues for analyzing linear systems.

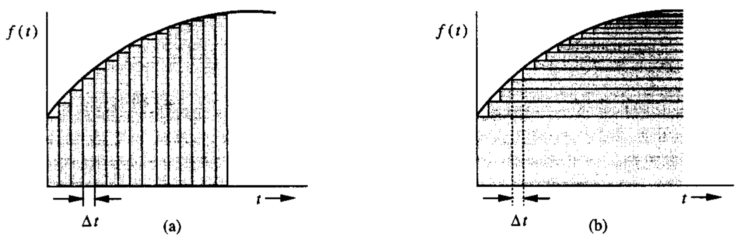


Fig. 1.27 Signal representation in terms of impulse and step components.

As an example, consider an arbitrary input $f(t)$ such as the one shown in Fig. 1.27a. We can approximate $f(t)$ with a sum of rectangular pulses of width Δt and of varying heights. The approximation improves as $\Delta t \to 0$, when the rectangular pulses become impulses spaced Δt seconds apart (with $\Delta t \to 0$). Thus, an arbitrary input can be replaced by a weighted sum of impulses spaced Δt ($\Delta t \to 0$) seconds apart. Therefore, if we know the system response to a unit impulse, we can immediately determine the system response to an arbitrary input $f(t)$ by adding the system response to each impulse component of $f(t)$. A similar situation is depicted in Fig. 1.27b, where $f(t)$ is approximated by a sum of step functions of varying magnitude and spaced Δt seconds apart. The approximation improves as Δt becomes smaller. Therefore, if we know the system response to a unit step input, we can compute the system response to any arbitrary input $f(t)$ with relative ease. Time-domain analysis of linear systems (discussed in Chapter 2) uses this approach.

In Chapters 4,6,10, and 11 we employ the same approach but instead use sinusoids or exponentials as our basic signal components. There, we show that any arbitrary input signal can be expressed as a weighted sum of sinusoids (or exponentials) having various frequencies. Thus a knowledge of the system response to a sinusoid enables us to determine the system response to an arbitrary input $f(t)$.

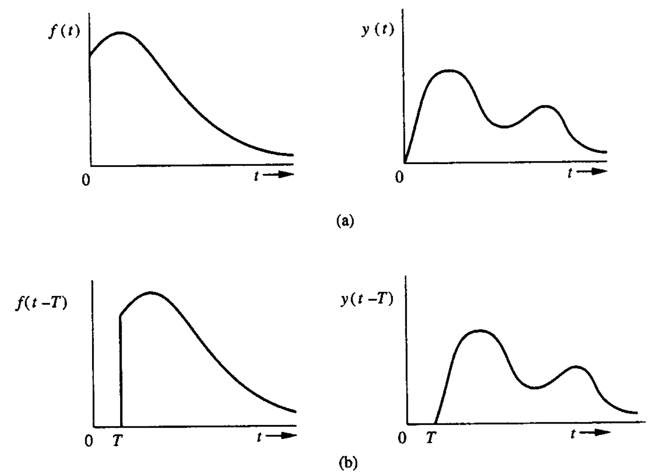


Fig. 1.28 Time-invariance property.

1.7-2 Time-Invariant and Time-Varying Parameter Systems

Systems whose parameters do not change with time are **time-invariant** (also **constant-parameter**) systems. For such a system, if the input is delayed by T seconds, the output is the same as before but delayed by T (assuming identical initial conditions). This property is expressed graphically in Fig. 1.28.

It is possible to verify that the system in Fig. 1.26 is a time-invariant system. Networks composed of RLC elements and other commonly used active elements such as transistors are time-invariant systems. A system with an input-output relationship described by a linear differential equation of the form (1.44) is a linear time-invariant (LTI) system when the coefficients a_i and b_i of such equation are constants. If these coefficients are functions of time, then the system is a linear **time-varying** system. The system described in exercise E1.12 is an example of a linear time-varying system. Another familiar example of a time-varying system is the carbon microphone, in which the resistance R is a function of the mechanical pressure generated by sound waves on the carbon granules of the microphone. An equivalent circuit for the microphone appears in Fig. 1.29. The response is the current $i(t)$, and the equation describing the circuit is

$$L\frac{di(t)}{dt} + R(t)i(t) = f(t)$$

One of the coefficients in this equation, $R(t)$, is time-varying.

△ **Exercise E1.14**
Show that a system described by the following equation is time-varying parameter system:

$$y(t) = (\sin t)\, f(t-2)$$

Hint: Show that the system fails to satisfy the time-invariance property. ▽

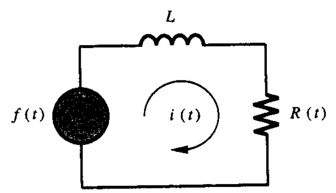


Fig. 1.29 An example of a linear time-varying system.

1.7-3 Instantaneous and Dynamic Systems

As observed earlier, a system's output at any instant t generally depends upon the entire past input. However, in a special class of systems, the output at any instant t depends only on its input at that instant. In resistive networks, for example, any output of the network at some instant t depends only on the input at the instant t. In these systems, past history is irrelevant in determining the response. Such systems are said to be **instantaneous** or **memoryless** systems. More precisely, a system is said to be instantaneous (or memoryless) if its output at any instant t depends, at most, on the strength of its input(s) at the same instant but not on any past or future values of the input(s). Otherwise, the system is said to be **dynamic** (or a system with memory). A system whose response at t is completely determined by the input signals over the past T seconds [interval from $(t-T)$ to t] is a **finite-memory system** with a memory of T seconds. Networks containing inductive and capacitive elements generally have infinite memory because the response of such networks at any instant t is determined by their inputs over the entire past $(-\infty, t)$. This is true for the RC circuit of Fig. 1.26.

In this book we will generally examine dynamic systems. Instantaneous systems are a special case of dynamic systems.

1.7-4 Causal and Noncausal Systems

A **causal** (also known as a **physical** or **non-anticipative**) system is one for which the output at any instant t_0 depends only on the value of the input $f(t)$ for $t \le t_0$. In other words, the value of the output at the present instant depends only on the past and present values of the input $f(t)$, not on its future values. To put it simply, in a causal system the output cannot start before the input is applied. If the response starts before the input, it means that the system knows the input in the future and acts on this knowledge before the input is applied. A system that violates the condition of causality is called a **noncausal** (or **anticipative**) system.

Any practical system that operates in real time† must necessarily be causal. We do not yet know how to build a system that can respond to future inputs (inputs not yet applied). A noncausal system is a prophetic system that knows the future input and acts on it in the present. Thus, if we apply an input starting at $t = 0$ to a noncausal system, the output would begin even before $t = 0$. As an example, consider the system specified by

$$y(t) = f(t-2) + f(t+2) \tag{1.46}$$

†In real-time operations, the response to an input is essentially simultaneous (contemporaneous) with the input itself.

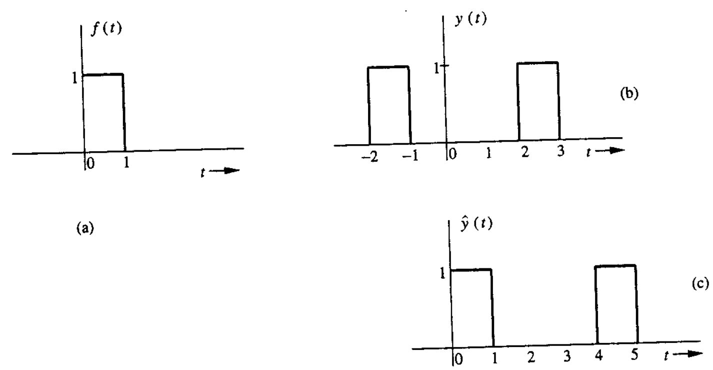


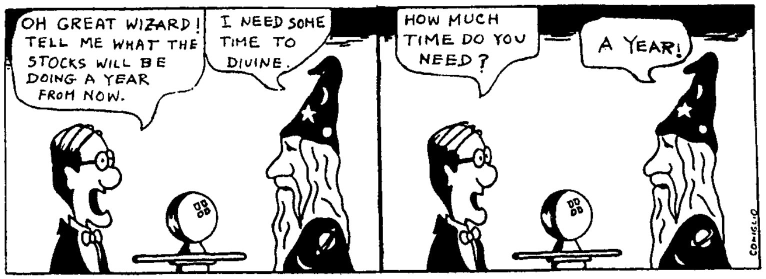
Fig. 1.30 A noncausal system and its realization by a delayed causal system.

For the input $f(t)$ illustrated in Fig. 1.30a, the output $y(t)$, as computed from Eq. (1.46) (shown in Fig. 1.30b), starts even before the input is applied. Equation (1.46) shows that $y(t)$, the output at t, is given by the sum of the input values two seconds before and two seconds after t (at $t-2$ and $t+2$ respectively). But if we are operating the system in real time at t, we do not know what the value of the input will be two seconds later. Thus it is impossible to implement this system in real time. For this reason, noncausal systems are unrealizable in **real time**.

Why Study Noncausal Systems?

From the above discussion it may seem that noncausal systems have no practical purpose. This is not the case; they are valuable in the study of systems for several reasons. First, noncausal systems *are* realizable when the independent variable is other than "time" (e.g., **space**). Consider, for example, an electric charge of density $q(x)$ placed along the x-axis for $x \ge 0$. This charge density produces an electric field $E(x)$ that is present at every point on the x-axis from $x = -\infty$ to ∞. In this case the input [i.e., the charge density $q(x)$] starts at $x = 0$, but its output [the electric field $E(x)$] begins before $x = 0$. Clearly, this space charge system is noncausal. This discussion shows that only temporal systems (systems with time as independent variable) must be causal in order to be realizable. The terms "before" and "after" have a special connection to causality only when the independent variable is time. This connection is lost for variables other than time. Nontemporal systems, such as those occurring in optics, can be noncausal and still realizable.

Moreover, even for temporal systems, such as those used for signal processing, the study of noncausal systems is important. In such systems we may have all input data prerecorded. (This often happens with speech, geophysical, and meteorological signals, and with space probes.) In such cases, the input's future values are available to us. For example, suppose we had a set of input signal records available for the system described by Eq. (1.46). We can then compute $y(t)$ since, for any t, we need only refer to the records to find the input's value two seconds before and two seconds after t. Thus, noncausal systems can be realized, although not in real time. We



Noncausal systems are realizable with time delay!

may therefore be able to realize a noncausal system, provided that we are willing to accept a time delay in the output. Consider a system whose output $\hat{y}(t)$ is the same as $y(t)$ in Eq. (1.46) delayed by two seconds (Fig 1.30c), so that

$$\begin{aligned}\hat{y}(t) &= y(t-2)\\ &= f(t-4) + f(t)\end{aligned}$$

Here the value of the output $\hat{y}$ at any instant t is the sum of the values of the input f at t and at the instant four seconds earlier [at $(t-4)$]. In this case, the output at any instant t does not depend on future values of the input, and the system is causal. The output of this system, which is $\hat{y}(t)$, is identical to that in Eq. (1.46) or Fig. 1.30b except for a delay of two seconds. Thus, a noncausal system may be realized or satisfactorily approximated in real time by using a causal system with a delay.

A third reason for studying noncausal systems is that they provide an upper bound on the performance of causal systems. For example, if we wish to design a filter for separating a signal from noise, then the optimum filter is invariably a noncausal system. Although unrealizable, this noncausal system's performance acts as the upper limit on what can be achieved and gives us a standard for evaluating the performance of causal filters.

At first glance, noncausal systems may seem inscrutable. Actually, there is nothing mysterious about these systems and their approximate realization through using physical systems with delay. If we want to know what will happen one year from now, we have two choices: go to a prophet (an unrealizable person) who can give the answers immediately, or go to a wise man and allow him a delay of one year to give us the answer! If the wise man is truly wise, he may even be able to shrewdly guess the future very closely with a delay of less than a year by studying trends. Such is the case with noncausal systems—nothing more and nothing less.

△ **Exercise E1.15**
Show that a system described by the equation below is noncausal:

$$y(t) = \int_{t-5}^{t+5} f(\tau)\, d\tau$$

Show that this system can be realized physically if we accept a delay of 5 seconds in the output.
▽

1.7-5 Lumped-Parameter and Distributed-Parameter Systems

In the study of electrical systems, we make use of voltage-current relationships for various components (Ohm's law, for example). In doing so, we implicitly assume that the current in any system component (resistor, inductor, etc.) is the same at every point throughout that component. Thus, we assume that electrical signals are propagated instantaneously throughout the system. In reality, however, electrical signals are electromagnetic space waves requiring some finite propagation time. An electric current, for example, propagates through a component with a finite velocity and therefore may exhibit different values at different locations in the same component. Thus, an electric current is a function not only of time but also of space. However, if the physical dimensions of a component are small compared to the wavelength of the signal propagated, we may assume that the current is constant throughout the component. This is the assumption made in **lumped-parameter systems**, where each component is regarded as being lumped at one point in space. Such an assumption is justified at lower frequencies (higher wavelength). Therefore, in lumped-parameter models, signals can be assumed to be functions of time alone. For such systems, the system equations require only one independent variable (time) and therefore are ordinary differential equations.

In contrast, for **distributed-parameter systems** such as transmission lines, waveguides, antennas, and microwave tubes, the system dimensions cannot be assumed to be small compared to the wavelengths of the signals; thus the lumped-parameter assumption breaks down. The signals here are functions of space as well as of time, leading to mathematical models consisting of partial differential equations.[3] The discussion in this book will be restricted to lumped-parameter systems only.

1.7-6 Continuous-Time and Discrete-Time Systems

Distinction between discrete-time and continuous-time signals is discussed in Sec. 1.2-1. Systems whose inputs and outputs are continuous-time signals are **continuous-time systems**. On the other hand, systems whose inputs and outputs are discrete-time signals are **discrete-time systems**. If a continuous-time signal is sampled, the resulting signal is a discrete-time signal. We can process a continuous-time signal by processing its samples with a discrete-time system.

1.7-7 Analog and Digital Systems

Analog and digital signals are discussed in Sec. 1.2-2. A system whose input and output signals are analog is an **analog system**; a system whose input and output signals are digital is a **digital system**. A digital computer is an example of a digital (binary) system. Observe that a digital computer is an example of a system that is digital as well as discrete-time.

Additional Classification of Systems

There are additional classes of systems, such as **invertible** and **noninvertible** systems. A system $\mathcal{S}$ performs certain operation(s) on input signal(s). If we can obtain the input $f(t)$ back from the output $y(t)$ by some operation, the system

$\mathcal{S}$ is said to be **invertible**. For a noninvertible system, different inputs can result in the same output (as in a rectifier), and it is impossible to determine the input for a given output. Therefore, for an invertible system, it is essential that distinct inputs result in distinct outputs so that there is one-to-one mapping between an input and the corresponding output. This ensures that every output has a unique input. Consequently, the system is invertible. The system that achieves this inverse operation [of obtaining $f(t)$ from $y(t)$] is the **inverse system** of $\mathcal{S}$. For instance, A system whose input and output are related by equation $y(t) = a\,f(t) + b$ is an invertible system. But a rectifier, specified by the equation $y(t) = |f(t)|$ is noninvertible because the rectification operation cannot be undone.

An ideal differentiator is noninvertible because integration of its output cannot restore the original signal unless we know one piece of information about the signal. For instance, if $f(t) = 3t + 5$, the output of the differentiator is $y(t) = 3$. If this output is applied to an integrator, the output is $3t + c$, where c is an arbitrary constant. If we know one piece of information about $f(t)$, such as $f(0) = 5$, we can determine the input to be $f(t) = 3t + 5$. Thus, a differentiator along with one piece of information (known as auxiliary condition) is an invertible system.† Similarly, a system consisting of a cascade of two differentiators is invertible, if we know two independent pieces of information (auxiliary conditions) about the input signal.

In addition, systems can also be classified as **stable** or **unstable** systems. The concept of stability is discussed in more depth in later chapters.

△ **Exercise E1.16**

Show that a system described by the equation $y(t) = f^2(t)$ is noninvertible. ▽

1.8 System Model: Input-output Description

As mentioned earlier, systems theory encompasses a variety of systems, such as electrical, mechanical, hydraulic, acoustic, electromechanical, and chemical, as well as social, political, economic, and biological. The first step in analyzing any system is the construction of a system model, which is a mathematical expression or a rule that satisfactorily approximates the dynamical behavior of the system. In this chapter we shall consider only the continuous-time systems. (Modeling of discrete-time systems is discussed in Chapter 8.)

To construct a system model, we must study the relationships between different variables in the system. In electrical systems, for example, we must determine a satisfactory model for the voltage-current relationship of each element, such as Ohm's law for a resistor. In addition, we must determine the various constraints on voltages and currents when several electrical elements are interconnected. These are the laws of interconnection—the well-known Kirchhoff's voltage and current laws (KVL and KCL). From all these equations, we eliminate unwanted variables to obtain equation(s) relating the desired output variable(s) to the input(s). The following examples demonstrate the procedure of deriving input-output relationships for some LTI electrical systems.

†The additional piece of information cannot be just any information. For instance, in the above example, if we are given $\dot{f}(0) = 0$, it will not help in determining c, and the system is noninvertible.

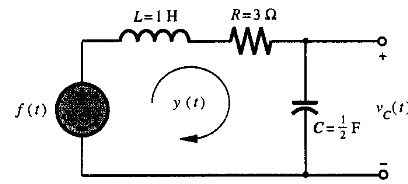


Fig. 1.31 Circuit for Example 1.10.

■ **Example 1.10**

For the series RLC circuit of Fig. 1.31, find the input-output equation relating the input voltage $f(t)$ to the output current (loop current) $y(t)$.

Application of the Kirchhoff's voltage law around the loop yields

$$v_L(t) + v_R(t) + v_C(t) = f(t) \tag{1.47}$$

By using the voltage-current laws of each element (inductor, resistor, and capacitor), we can express this equation as

$$\frac{dy}{dt} + 3y(t) + 2\int_{-\infty}^{t} y(\tau)\,d\tau = f(t) \tag{1.48}$$

Differentiating both sides of this equation obtains

$$\frac{d^2y}{dt^2} + 3\frac{dy}{dt} + 2y(t) = \frac{df}{dt} \tag{1.49}$$

This differential equation is the input-output relationship between the output $y(t)$ and the input $f(t)$. ■

It proves convenient to use a compact notation D for the differential operator $\frac{d}{dt}$. Thus

$$\frac{dy}{dt} \equiv Dy(t) \tag{1.50}$$

$$\frac{d^2y}{dt^2} \equiv D^2y(t) \tag{1.51}$$

and so on. With this notation, Eq. (1.49) can be expressed as

$$\left(D^2 + 3D + 2\right) y(t) = Df(t) \tag{1.52}$$

The differential operator is the inverse of the integral operator, so we can use the operator $1/D$ to represent integration†.

$$\int_{-\infty}^{t} y(\tau)\,d\tau \equiv \frac{1}{D}y(t) \tag{1.53}$$

† Use of operator $1/D$ for integration generates some subtle mathematical difficulties because the operators D and $1/D$ do not commute. For instance, we know that $D(1/D) = 1$ because $\frac{d}{dt}[\int_{-\infty}^{t} y(\tau)\,d\tau] = y(t)$. However, $\frac{1}{D}D$ is not necessarily unity. Use of Cramer's rule in solving simultaneous integro-differential equations will always result in cancellation of operators $1/D$ and D. This procedure may yield erroneous results in those cases where the factor D occurs in the

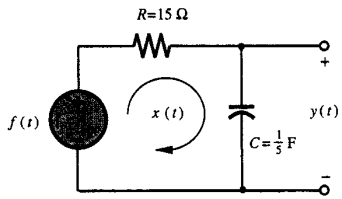


Fig. 1.32 Circuit for Example 1.11.

Consequently, the loop equation (1.48) can be expressed as

$$\left(D + 3 + \frac{2}{D}\right) y(t) = f(t) \tag{1.54}$$

Multiplying both sides by D, that is, differentiating Eq. (1.54), we obtain

$$\left(D^2 + 3D + 2\right) y(t) = Df(t) \tag{1.55}$$

which is identical to Eq. (1.52).

Recall that Eq. (1.55) is not an algebraic equation, and $D^2 + 3D + 2$ is not an algebraic term that multiplies $y(t)$; it is an operator that operates on $y(t)$. It means that we must perform the following operations on $y(t)$: take the second derivative of $y(t)$ and add to it 3 times the first derivative of $y(t)$ and 2 times $y(t)$. Clearly, a polynomial in D multiplied by $y(t)$ represents a certain differential operation on $y(t)$.

■ Example 1.11

Find the equation relating the input to output for the series RC circuit of Fig. 1.32 if the input is the voltage $f(t)$ and output is **(a)** the loop current $x(t)$ **(b)** the capacitor voltage $y(t)$.

The loop equation for the circuit is

$$Rx(t) + \frac{1}{C}\int_{-\infty}^{t} x(\tau)\,d\tau = f(t) \tag{1.56}$$

or

$$15x(t) + 5\int_{-\infty}^{t} x(\tau)\,d\tau = f(t) \tag{1.57}$$

With operational notation, this equation can be expressed as

---

numerator as well as in the denominator. This happens, for instance, in circuits with all-inductor loops or all-capacitor cutsets. To eliminate this problem, avoid the integral operation in system equations so that the resulting equations are differential rather than integro-differential. In electrical circuits, this can be done by using charge (instead of current) variables in loops containing capacitors and using current variables for loops without capacitors. In the literature this problem of commutativity of D and $1/D$ is largely ignored. As mentioned earlier, such procedure gives erroneous results only in special systems, such as the circuits with all-inductor loops or all-capacitor cutsets. Fortunately such systems constitute a very small fraction of the systems we deal with. For further discussion of this topic and a correct method of handling problems involving integrals, see Ref. 4

$$15x(t) + \frac{5}{D}x(t) = f(t) \tag{1.58}$$

Multiplying both sides of the above equation by D (that is, differentiating the above equation), we obtain

$$(15D + 5)x(t) = Df(t) \tag{1.59a}$$

or

$$15\frac{dx}{dt} + 5x(t) = \frac{df}{dt} \tag{1.59b}$$

Moreover,

$$x(t) = C\frac{dy}{dt} = \frac{1}{5}Dy(t)$$

Substitution of this result in Eq. (1.59a) yields

$$(3D + 1)y(t) = f(t) \tag{1.60}$$

or

$$3\frac{dy}{dt} + y(t) = f(t) \tag{1.61}$$

■

△ **Exercise E1.17**

For the RLC circuit in Fig. 1.31, find the input-output relationship if the output is the inductor voltage $v_L(t)$.

Hint: $v_L(t) = LDy(t) = Dy(t)$. Answer: $\left(D^2 + 3D + 2\right) v_L(t) = D^2 f(t)$ ▽

△ **Exercise E1.18**

For the RLC circuit in Fig. 1.31, find the input-output relationship if the output is the capacitor voltage $v_C(t)$.

Hint: $v_C(t) = \frac{1}{CD}y(t) = \frac{2}{D}y(t)$. Answer: $\left(D^2 + 3D + 2\right) v_C(t) = 2f(t)$ ▽

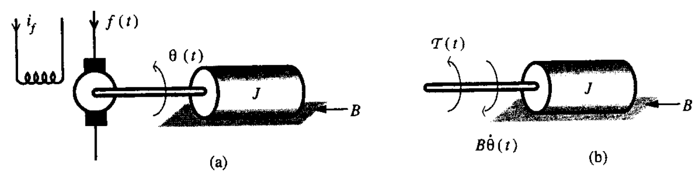


Fig. 1.33 Armature controlled dc motor.

■ Example 1.12

In rotational systems, the equations of motion are similar to those in translational systems. In place of force F, we have torque $\mathcal{T}$. In place of mass M, we have **moment of inertia** J (the rotational mass), and in place of linear acceleration $\ddot{x}$, we have angular acceleration $\ddot{\theta}$. The equation of motion for rotational systems is $\mathcal{T} = J\ddot{\theta}$ (in place of $F = M\ddot{x}$).

A wide variety of electromechanical systems convert electrical signals into mechanical motion (mechanical energy) and vice versa. Here we consider a rather simple example of an armature-controlled dc motor (with a constant field current i_f) driven by a current source $f(t)$, as depicted in Fig. 1.33a. Let $\theta(t)$ be the angular position of the rotor. The torque $\mathcal{T}(t)$ generated in the rotor is proportional to the armature current $f(t)$. Therefore

$$\mathcal{T}(t) = K_T f(t) \tag{1.62}$$

where K_T is a constant of the motor. This torque drives a mechanical load whose free-body diagram is illustrated in Fig. 1.33b. The viscous damping (with coefficient B), which is proportional to the angular velocity $\dot{\theta}$, dissipates a torque $B\dot{\theta}(t)$. If J is the moment of inertia of the load (including the rotor of the motor), then the net torque $\mathcal{T}(t) - B\dot{\theta}(t)$ available must equal to $J\ddot{\theta}(t)$;

$$J\ddot{\theta}(t) = \mathcal{T}(t) - B\dot{\theta}(t) \tag{1.63}$$

Thus

$$\left(JD^2 + BD\right)\theta(t) = \mathcal{T}(t)$$
$$= K_T f(t) \tag{1.64}$$

which can be expressed as

$$D(D+a)\theta(t) = K_1 f(t) \tag{1.65}$$

where $a = B/J$ and $K_1 = K_T/J$. ■

1.8-1 Internal and External Descriptions of a System

With a knowledge of the internal structure of a system, we can write system equations yielding an **internal description** of the system. In contrast, the system description seen from the system's input and output terminals is the system's **external description**. To understand an external description, suppose that a system is enclosed in a "black box" with only its input(s) and output(s) terminals accessible. In order to describe or characterize such a system, we must perform some measurements at these terminals. For example, we might apply a known input, such as a unit impulse or a unit step, and then measure the system's output. The description provided by such a measurement is an external description of the system.

Suppose the circuit in Fig. 1.34a with the input $f(t)$ and the output $y(t)$ is enclosed inside a "black box" with only the input and output terminals accessible. Under these conditions the only way to describe or specify the system is with external measurements. We can, for example, apply a known voltage $f(t)$ at the input terminals and measure the resulting output voltage $y(t)$. From this information we can describe or characterize the system. This is the *external description.*

Assuming zero initial capacitor voltage, the input voltage $f(t)$ produces a current i (Fig. 1.34a), which divides equally between the two branches because of the balanced nature of the circuit. Thus, the voltage across the capacitor continues to remain zero. Therefore, for the purpose of computing the current i, the capacitor may be removed or replaced by a short. The resulting circuit is equivalent to that shown in Fig. 1.34b. It is clear from Fig. 1.34b that $f(t)$ sees a net resistance of $5\,\Omega$, and

$$i(t) = \frac{1}{5}f(t)$$

Also, because $y(t) = 2 \times (i/2) = i$,

$$y(t) = \frac{1}{5}f(t) \tag{1.66}$$

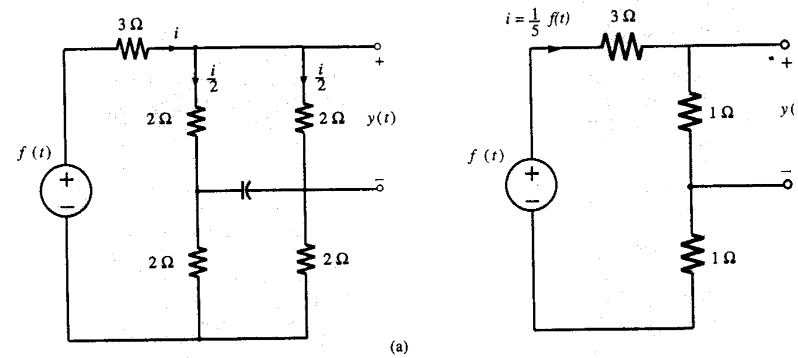


Fig. 1.34 A system that cannot be described by external measurements.

The equivalent system as seen from the system's external terminals is depicted in Fig. 1.34b. Clearly, for the external description, the capacitor does not exist. For most systems, the external and internal descriptions are identical, but there are a few exceptions, as in the present case, where the external description gives an inadequate picture of the systems. This happens when the system is **uncontrollable** and/or **unobservable**. Figures 1.35a and 1.35b show a structural representation of simple uncontrollable and unobservable systems respectively. In Fig. 1.35a we note that part of the system (subsystem S_2) inside the box cannot be controlled by the input $f(t)$. In Fig. 1.35b some of the system outputs (those in subsystem S_2) cannot be observed from the output terminals. If we try to describe either of these systems by applying an external input $f(t)$ and then measuring the output $y(t)$, the measurement will not characterize the complete system but only the part of the system (here S_1) that is both controllable and observable (linked to both the input and output). Such systems are undesirable in practice and should be avoided in any system design. The system in Fig. 1.35a can be shown to be neither controllable nor observable. It can be represented structurally as a combination of the systems in Figs. 1.35a and 1.35b.

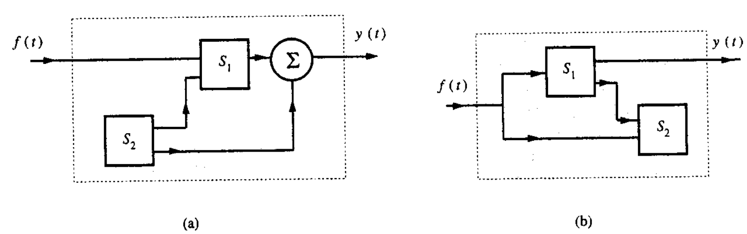


Fig. 1.35 Structures of uncontrollable and unobservable systems.

1.9 Summary

A *signal* is a set of information or data. A *system* processes input signals to modify them or extract additional information from them to produce output signals (response). A system may be made up of physical components (hardware realization) or may be an algorithm that computes an output signal from an input signal (software realization).

A convenient measure of the size of a signal is its energy if it is finite. If the signal energy is infinite, the appropriate measure is its power, if it exists. The signal power is the time average of its energy (averaged over the entire time interval from $-\infty$ to ∞). For periodic signals the time averaging need be performed only over one period in view of the periodic repetition of the signal. Signal power is also equal to the mean squared value of the signal (averaged over the entire time interval from $t = -\infty$ to ∞).

Signals can be classified in several ways as follows:

1. A *continuous-time signal* is specified for a continuum of values of the independent variable (such as time t). A *discrete-time signal* is specified only at a finite or a countable set of time instants.
2. An *analog signal* is a signal whose amplitude can take on any value over a continuum. On the other hand, a signal whose amplitudes can take on only a finite number of values is a *digital signal*. The terms *discrete-time* and *continuous-time* qualify the nature of a signal along the time axis (horizontal axis). The terms *analog* and *digital*, on the other hand, qualify the nature of the signal amplitude (vertical axis).
3. A *periodic signal* $f(t)$ is defined by the fact that $f(t) = f(t + T_0)$ for some T_0. The smallest value of T_0 for which this relationship is satisfied is called the *period*. A periodic signal remains unchanged when shifted by an integral multiple of its period. A periodic signal can be generated by a periodic extension of any segment of $f(t)$ of duration T_0. Finally, a periodic signal, by definition, must exist over the entire time interval $-\infty < t < \infty$. A signal is *aperiodic* if it is not periodic.

   An *everlasting signal* starts at $t = -\infty$ and continues forever to $t = \infty$. A *causal signal* is a signal that is zero for $t < 0$. Hence, periodic signals are everlasting signals.
4. A signal with finite energy is an *energy signal*. Similarly a signal with a finite and nonzero power (mean square value) is a *power signal*. A signal can either be an energy signal or a power signal, but not both. However, there are signals that are neither energy nor power signals.
5. A signal whose physical description is known completely in a mathematical or graphical form is a *deterministic signal*. A *random signal* is known only in terms of its probabilistic description such as mean value, mean square value, and so on, rather than its mathematical or graphical form.

A signal $f(t)$ delayed by T seconds (right-shifted) is given by $f(t - T)$; on the other hand, $f(t)$ advanced by T (left-shifted) is given by $f(t + T)$. A signal $f(t)$ time-compressed by a factor a ($a > 1$) is given by $f(at)$; on the other hand, the same

signal time-expanded by factor a is given by $f(\frac{t}{a})$. The same signal time-inverted is given by $f(-t)$.

The unit step function $u(t)$ is very useful in representing causal signals and signals with different mathematical descriptions over different intervals.

In the classical definition, the unit impulse function $\delta(t)$ is characterized by unit area, and the fact that it is concentrated at a single instant $t = 0$. The impulse function has a sampling (or sifting) property, which states that the area under the product of a function with a unit impulse is equal to the value of that function at the instant where the impulse is located (assuming the function to be continuous at the impulse location). In the modern approach, the impulse function is viewed as a generalized function and is defined by the sampling property.

The exponential function e^{st}, where s is complex, encompasses a large class of signals that includes a constant, a monotonic exponential, a sinusoid, and an exponentially varying sinusoid.

A signal that is symmetrical about the vertical axis ($t = 0$) is an *even* function of time, and a signal that is antisymmetrical about the vertical axis is an *odd* function of time. The product of an even function with an odd function results in an odd function. However, the product of an even function with an even function or an odd function with an odd function results in an even function. The area under an odd function from $t = -a$ to a is always zero regardless of the value of a. On the other hand, the area under an even function from $t = -a$ to a is two times the area under the same function from $t = 0$ to a (or from $t = -a$ to 0). Every signal can be expressed as a sum of odd and even function of time.

A system processes input signals to produce output signals (response). The input is the cause and the output is its effect. In general, the output is affected by two causes: the internal conditions of the system (such as the initial conditions) and the external input.

Systems can be classified in several ways:

1. Linear systems are characterized by the linearity property, which implies superposition; if several causes (such as various inputs and initial conditions) are acting on a linear system, the total effect (response) is the sum of the responses from each cause, assuming that all the remaining causes are absent. A system is nonlinear if it is not linear.
2. Time-invariant systems are characterized by the fact that system parameters do not change with time. The parameters of time-varying parameter systems change with time.
3. For memoryless (or instantaneous) systems, the system response at any instant t depends only on the present value of the input (value at t). For systems with memory (also known as dynamic systems), the system response at any instant t depends not only on the present value of the input, but also on the past values of the input (values before t).
4. In contrast, if a system response at t also depends on the future values of the input (values of input beyond t), the system is noncausal. In causal systems, the response does not depend on the future values of the input. Because of the dependence of the response on the future values of input, the effect (response)

of noncausal systems occurs before cause. When the independent variable is time (temporal systems), the noncausal systems are prophetic systems, and therefore, unrealizable, although close approximation is possible with some time delay in the response. Noncausal systems with independent variables other than time (e.g., space) are realizable.

5. If the dimensions of system elements are small compared to the wavelengths of the signals, we may assume that each element is lumped at a single point in space, and the system may be considered as a lumped-parameter system. The signals under this assumption are functions of time only. If this assumption does not hold, the signals are functions of space and time; such a system is a distributed-parameter system.
6. Systems whose inputs and outputs are continuous-time signals are continuous-time systems; systems whose inputs and outputs are discrete-time signals are discrete-time systems. If a continuous-time signal is sampled, the resulting signal is a discrete-time signal. We can process a continuous-time signal by processing the samples of this signal with a discrete-time system.
7. Systems whose inputs and outputs are analog signals are analog systems; those whose inputs and outputs are digital signals are digital systems.
8. If we can obtain the input $f(t)$ back from the output $y(t)$ of a system S by some operation, the system S is said to be invertible. Otherwise the system is noninvertible.

The system model derived from a knowledge of the internal structure of the system is its internal description. In contrast, an external description of a system is its description as seen from the system's input and output terminals; it can be obtained by applying a known input and measuring the resulting output. In the majority of practical systems, an external description of a system so obtained is equivalent to its internal description. In some cases, however, the external description fails to give adequate information about the system. Such is the case with the so-called uncontrollable or unobservable systems.

References

1. Papoulis, A., *The Fourier Integral and Its Applications*, McGraw-Hill, New York, 1962.
2. Kailath, T., *Linear Systems*, Prentice-Hall, Englewood Cliffs, New Jersey, 1980.
3. Lathi, B.P., *Signals, Systems, and Communication,* Wiley, New York, 1965.
4. Lathi, B.P., *Signals and Systems*, Berkeley-Cambridge Press, Carmichael, California, 1987.

Problems

1.1-1 Find the energies of the signals illustrated in Fig. P1.1-1. Comment on the effect on energy of sign change, time shifting, or doubling of the signal. What is the effect on the energy if the signal is multiplied by k?

1.1-2 Repeat Prob. 1.1-1 for the signals in Fig. P1.1-2.

1.1-3 (a) Find the energies of the pair of signals $x(t)$ and $y(t)$ depicted in Figs. P1.1-3a and b. Sketch and find the energies of signals $x(t) + y(t)$ and $x(t) - y(t)$. Can you make any observation from these results?

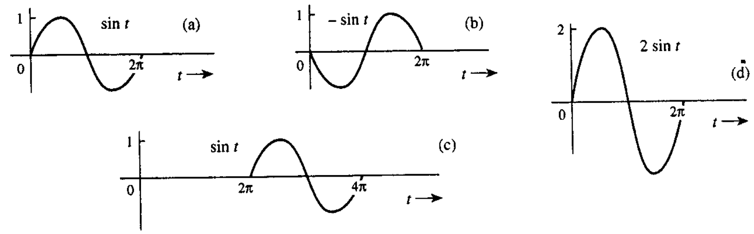


Fig. P1.1-1

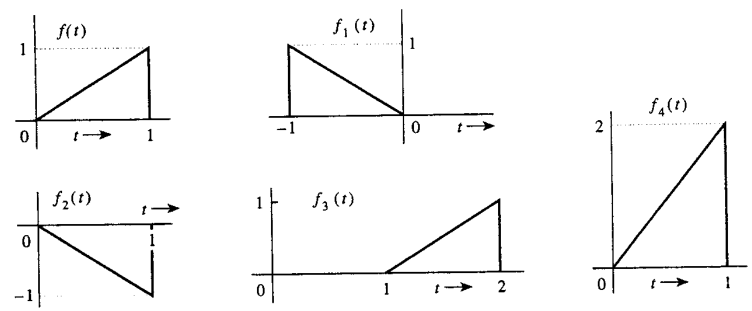


Fig. P1.1-2

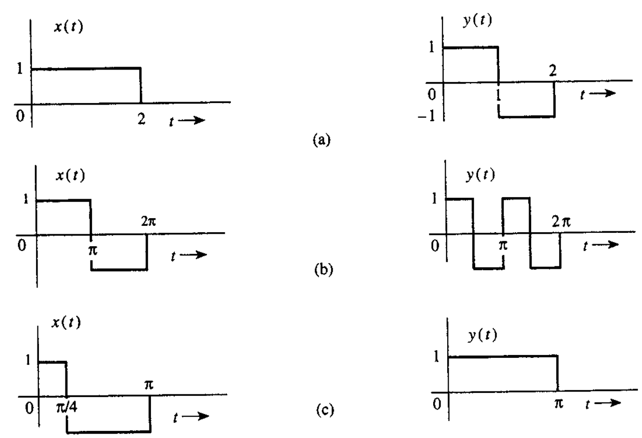


Fig. P1.1-3

(b) Repeat part (a) for the signal pair illustrated in Fig. P1.1-3c. Is your observation in part (a) still valid?

1.1-4 Find the power of the periodic signal $f(t)$ shown in Fig. P1.1-4. Find also the powers and the rms values of: **(a)** $-f(t)$ **(b)** $2f(t)$ **(c)** $cf(t)$. Comment.

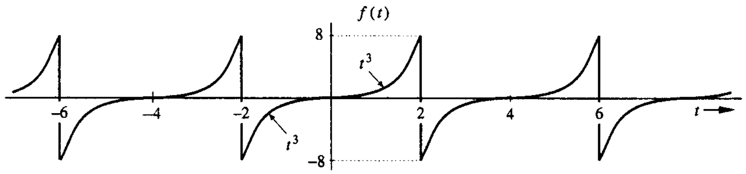


Fig. P1.1-4

1.1-5 Show that the power of a signal

$$f(t) = \sum_{k=m}^{n} D_k e^{j\omega_k t} \quad \text{is} \quad P_f = \sum_{k=m}^{n} |D_k|^2$$

assuming all frequencies to be distinct, that is, $\omega_i \neq \omega_k$ for all $i \neq k$

1.1-6 Determine the power and the rms value for each of the following signals:

(a) $10 \cos\left(100t + \frac{\pi}{3}\right)$ **(b)** $10 \cos\left(100t + \frac{\pi}{3}\right) + 16 \sin\left(150t + \frac{\pi}{5}\right)$

(c) $(10 + 2 \sin 3t) \cos 10t$ **(d)** $10 \cos 5t \cos 10t$

(e) $10 \sin 5t \cos 10t$ **(f)** $e^{j\alpha t} \cos \omega_0 t$

1.3-1 In Fig. P1.3-1, the signal $f_1(t) = f(-t)$. Express signals $f_2(t)$, $f_3(t)$, $f_4(t)$, and $f_5(t)$ in terms of signals $f(t)$, $f_1(t)$, and their time-shifted, time-scaled or time-inverted versions. For instance $f_2(t) = f(t - T) + f_1(t - T)$.

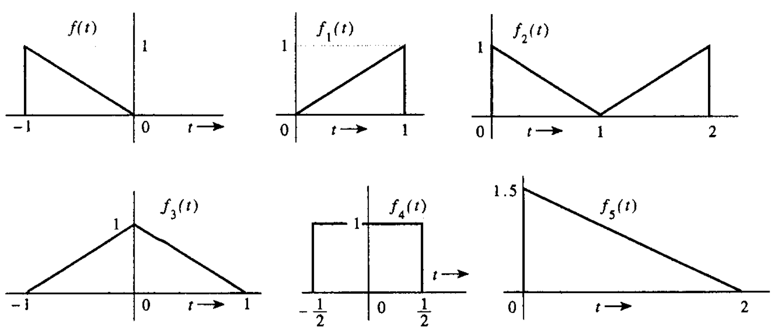


Fig. P1.3-1

1.3-2 For the signal $f(t)$ depicted in Fig. P1.3-2, sketch the signals: **(a)** $f(-t)$ **(b)** $f(t+6)$ **(c)** $f(3t)$ **(d)** $f(\frac{t}{2})$.

1.3-3 For the signal $f(t)$ illustrated in Fig. P1.3-3, sketch **(a)** $f(t-4)$ **(b)** $f(\frac{t}{1.5})$ **(c)** $f(-t)$ **(d)** $f(2t-4)$ **(e)** $f(2-t)$.

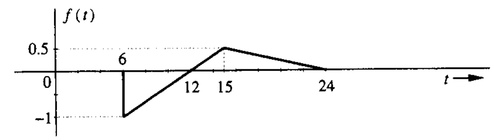


Fig. P1.3-2

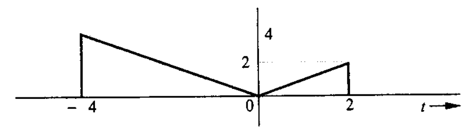


Fig. P1.3-3

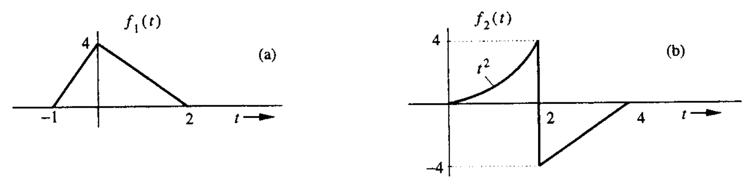


Fig. P1.4-2

1.4-1 Sketch the signals **(a)** $u(t-5) - u(t-7)$ **(b)** $u(t-5) + u(t-7)$ **(c)** $t^2[u(t-1) - u(t-2)]$ **(d)** $(t-4)[u(t-2) - u(t-4)]$

1.4-2 Express each of the signals in Fig. P1.4-2 by a single expression valid for all t.

1.4-3 For an energy signal $f(t)$ with energy E_f, show that the energy of any one of the signals $-f(t)$, $f(-t)$ and $f(t-T)$ is E_f. Show also that the energy of $f(at)$ as well as $f(at-b)$ is E_f/a. This shows that time-inversion and time-shifting does not affect signal energy. On the other hand, time compression of a signal ($a > 1$) reduces the energy, and time expansion of a signal ($a < 1$) increases the energy. What is the effect on signal energy if the signal is multiplied by a constant a?

1.4-4 Simplify the following expressions:

(a) $\left(\frac{\sin t}{t^2+2}\right)\delta(t)$ **(b)** $\left(\frac{j\omega+2}{\omega^2+9}\right)\delta(\omega)$

(c) $\left[e^{-t}\cos(3t - 60°)\right]\delta(t)$ **(d)** $\left(\frac{\sin\left[\frac{\pi}{2}(t-2)\right]}{t^2+4}\right)\delta(t-1)$

(e) $\left(\frac{1}{j\omega+2}\right)\delta(\omega+3)$ **(f)** $\left(\frac{\sin k\omega}{\omega}\right)\delta(\omega)$

Hint: Use Eq. (1.23). For part **(f)** use L'Hôpital's rule.

1.4-5 Evaluate the following integrals:

(a) $\int_{-\infty}^{\infty} \delta(\tau) f(t-\tau)\, d\tau$

(b) $\int_{-\infty}^{\infty} f(\tau)\delta(t-\tau)\, d\tau$

(c) $\int_{-\infty}^{\infty} \delta(t)e^{-j\omega t}\, dt$

(d) $\int_{-\infty}^{\infty} \delta(t-2)\sin \pi t\, dt$

(e) $\int_{-\infty}^{\infty} \delta(t+3)e^{-t}\, dt$

(f) $\int_{-\infty}^{\infty} (t^3+4)\delta(1-t)\, dt$

(g) $\int_{-\infty}^{\infty} f(2-t)\delta(3-t)\, dt$

(h) $\int_{-\infty}^{\infty} e^{(x-1)}\cos\left[\frac{\pi}{2}(x-5)\right]\delta(x-3)\, dx$

Hint: $\delta(x)$ is located at $x=0$. For example, $\delta(1-t)$ is located at $1-t=0$, and so on.

1.4-6 (a) Find and sketch df/dt for the signal $f(t)$ shown in Fig. P1.3-3.
(b) Find and sketch d^2f/dt^2 for the signal $f(t)$ depicted in Fig. P1.4-2a.

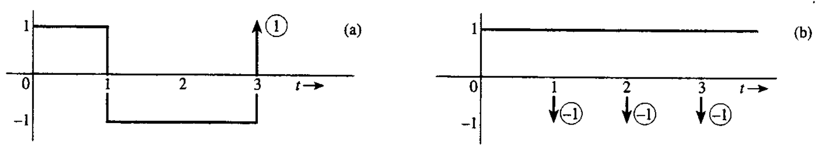


Fig. P1.4-7

1.4-7 Find and sketch $\int_{-\infty}^{t} f(x)\, dx$ for the signal $f(t)$ illustrated in Fig. P1.4-7.

1.4-8 Using the generalized function definition, show that $\delta(t)$ is an even function of t.
Hint: Start with Eq. (1.24a) as the definition of $\delta(t)$. Now change variable $t=-x$ to show that

$$\int_{-\infty}^{\infty} \phi(t)\delta(-t)\, dt = \phi(0)$$

1.4-9 Prove that

$$\delta(at) = \frac{1}{|a|}\delta(t)$$

Hint: Show that

$$\int_{-\infty}^{\infty} \phi(t)\delta(at)\, dt = \frac{1}{|a|}\phi(0)$$

..4-10 Show that

$$\int_{-\infty}^{\infty} \dot{\delta}(t)\phi(t)\, dt = -\dot{\phi}(0)$$

where $\phi(t)$ and $\dot{\phi}(t)$ are continuous at $t=0$, and $\phi(t) \to 0$ as $t \to \pm\infty$. This integral defines $\dot{\delta}(t)$ as a generalized function. Hint: Use integration by parts.

..4-11 A sinusoid $e^{\sigma t}\cos \omega t$ can be expressed as a sum of exponentials e^{st} and e^{-st} (Eq. (1.30c) with complex frequencies $s=\sigma+j\omega$ and $s=\sigma-j\omega$. Locate in the complex plane the frequencies of the following sinusoids: (a) $\cos 3t$ (b) $e^{-3t}\cos 3t$ (c) $e^{2t}\cos 3t$ (d) e^{-2t} (e) e^{2t} (f) 5.

1.5-1 Find and sketch the odd and the even components of (a) $u(t)$ (b) $tu(t)$ (c) $\sin \omega_0 t\, u(t)$ (d) $\cos \omega_0 t\, u(t)$ (e) $\sin \omega_0 t$ (f) $\cos \omega_0 t$.

1.6-1 Write the input-output relationship for an ideal integrator. Determine the zero-input and zero-state components of the response.

1.7-1 For the systems described by the equations below, with the input $f(t)$ and output $y(t)$, determine which of the systems are linear and which are nonlinear.

(a) $\frac{dy}{dt} + 2y(t) = f^2(t)$

(b) $\frac{dy}{dt} + 3ty(t) = t^2 f(t)$

(e) $\left(\frac{dy}{dt}\right)^2 + 2y(t) = f(t)$

(d) $\frac{dy}{dt} + y^2(t) = f(t)$

(c) $3y(t) + 2 = f(t)$

(f) $\frac{dy}{dt} + (\sin t)y(t) = \frac{df}{dt} + 2f(t)$

(g) $\frac{dy}{dt} + 2y(t) = f(t)\frac{df}{dt}$

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

1.7-2 For the systems described by the equations below, with the input $f(t)$ and output $y(t)$, determine which of the systems are time-invariant parameter systems and which are time-varying parameter systems.

(a) $y(t) = f(t-2)$

(b) $y(t) = f(-t)$

(c) $y(t) = f(at)$

(d) $y(t) = t\, f(t-2)$

(e) $y(t) = \int_{-5}^{5} f(\tau)\, d\tau$

(f) $y(t) = \left(\frac{df}{dt}\right)^2$

1.7-3 For a certain LTI system with the input $f(t)$, the output $y(t)$ and the two initial conditions $x_1(0)$ and $x_2(0)$, following observations were made:

$f(t)$	$x_1(0)$	$x_2(0)$	$y(t)$
0	1	-1	$e^{-t}u(t)$
0	2	1	$e^{-t}(3t+2)u(t)$
$u(t)$	-1	-1	$2u(t)$

Determine $y(t)$ when both the initial conditions are zero and the input $f(t)$ is as shown in Fig. P1.7-3.
Hint: There are three causes: the input and each of the two initial conditions. Because of linearity property, if a cause is increased by a factor k, the response to that cause also increases by the same factor k. Moreover, if causes are added, the corresponding responses add.

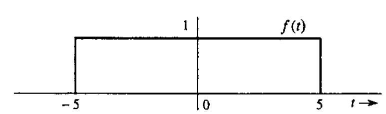


Fig. P1.7-3

1.7-4 A system is specified by its input-output relationship as

$$y(t) = f^2(t) \Big/ \left(\frac{df}{dt}\right)$$

Show that the system satisfies the homogeneity property but not the additivity property.

1.7-5 Show that the circuit in Fig. P1.7-5 is zero-state linear but is not zero-input linear. Assume all diodes to have identical (matched) characteristics.
Hint: In zero state (when the initial capacitor voltage $v_c(0) = 0$), the circuit is linear. If the input $f(t) = 0$, and $v_c(0)$ is nonzero, the current $y(t)$ does not exhibit linearity with respect to its cause $v_c(0)$.

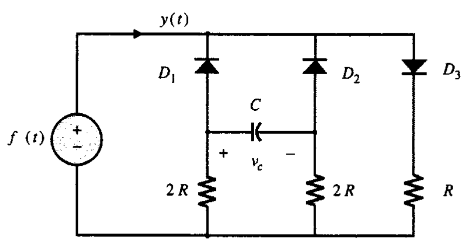


Fig. P1.7-5

1.7-6 The inductor L and the capacitor C in Fig. P1.7-6 are nonlinear, which makes the circuit nonlinear. The remaining 3 elements are linear. Show that the output $y(t)$ of this nonlinear circuit satisfies the linearity conditions with respect to the input $f(t)$ and the initial conditions (all the initial inductor currents and capacitor voltages). Recognize that a current source is an open circuit when the current is zero.

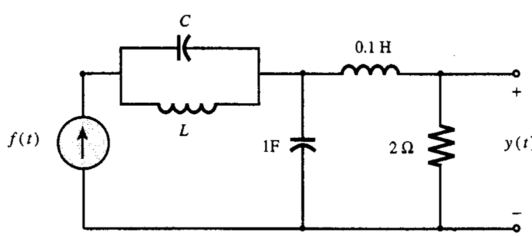


Fig. P1.7-6

1.7-7 For the systems described by the equations below, with the input $f(t)$ and output $y(t)$, determine which of the systems are causal and which are noncausal.

(a) $y(t) = f(t-2)$ **(c)** $y(t) = f(at) \quad a > 1$
(b) $y(t) = f(-t)$ **(d)** $y(t) = f(at) \quad a < 1$

1.7-8 For the systems described by the equations below, with the input $f(t)$ and output $y(t)$, determine which of the systems are invertible and which are noninvertible. For the invertible systems, find the input-output relationship of the inverse system.

(a) $y(t) = \int_{-\infty}^{t} f(\tau)\,d\tau$ **(c)** $y(t) = f^n(t) \quad n$, integer
(b) $y(t) = f(3t-6)$ **(d)** $y(t) = \cos[f(t)]$

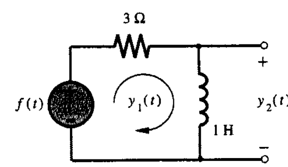


Fig. P1.8-1

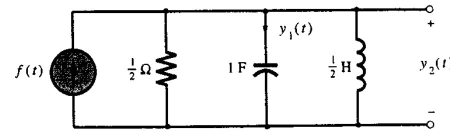


Fig. P1.8-2

1.8-1 For the circuit depicted in Fig. P1.8-1, find the differential equations relating outputs $y_1(t)$ and $y_2(t)$ to the input $f(t)$.

1.8-2 Repeat Prob. 1.8-1 for the circuit in Fig. P1.8-2.

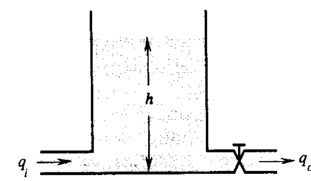


Fig. P1.8-3

1.8-3 Water flows into a tank at a rate of q_i units/s and flows out through the outflow valve at a rate of q_0 units/s (Fig. P1.8-3). Determine the equation relating the outflow q_0 to the input q_i. The outflow rate is proportional to the head h. Thus $q_0 = Rh$ where R is the valve resistance. Determine also the differential equation relating the head h to the input q_i.
(Hint: The net inflow of water in time Δt is $(q_i - q_0)\Delta t$. This inflow is also $A\Delta h$ where A is the cross section of the tank.)