1

OVERVIEW AND BACKGROUND

COPYRIGHTED MATERIAL

1.1 INTRODUCTION

In this book, we develop and examine several *mathematical models* consisting of one or more equations that are used in engineering to represent various *physical systems*. Usually, the goal is to solve these equations for the unknown dependent variables, and if that is not possible, the equations can be used to *simulate* the behavior of a system using computer software such as MATLAB.[1] In most engineering courses, the equations are usually *linear* or can be linearized as an approximation, but sometimes they are nonlinear and may be difficult to solve. From such models, it is possible to design and analyze components of a proposed system in order to achieve required performance specifications before developing a prototype and actually implementing the physical system.

Definition: System A *system* is a collection of interacting elements or devices that together result in a more complicated structure than the individual components alone, for the purpose of generating a specific type of signal or realizing a particular process.

The term system, as used in this book, also describes several interrelated equations called a *system of equations*, which are usually linear and can be represented by a

[1] MATLAB® is a registered trademark of The Mathworks, Inc., 3 Apple Hill Drive, Natick, MA.

Mathematical Foundations for Linear Circuits and Systems in Engineering, First Edition. John J. Shynk.
© 2016 John Wiley & Sons, Inc. Published 2016 by John Wiley & Sons, Inc.
Companion Website: http://www.wiley.com/go/linearcircuitsandsystems

matrix equation. The distinction between a physical system and a system of linear equations will be evident from the specific application.

Definition: Mathematical Model A *mathematical model* is an equation or set of equations used to represent a physical system, from which it is possible to predict the properties of the system and its output response to an input, given known parameters, certain variables, and initial conditions.

Generally, we are interested in the *dynamic behavior* of a system over time as it responds to one or more time-varying input signals. A block diagram of a system with single input $x(t)$ and single output $y(t)$ (single-input single-output (SISO)) is shown in Figure 1.1(a), where t is continuous time. The time variable can be defined for the entire real line $\mathcal{R}$: $-\infty < t < \infty$, but often we assume nonnegative $\mathcal{R}^+$: $0 \leq t < \infty$. In this scenario, a mathematical model provides the means to observe how $y(t)$ varies with $x(t)$ over t, assuming known initial conditions (usually at $t = 0$), so that we can predict the future behavior of the system. For the electric circuits described in Chapter 2, the inputs and outputs are *currents* through or *voltages* across the circuit components. For convenience, Table 1.1 summarizes the notation for different sets of numbers used in this book (though quaternions are only briefly discussed in Chapter 4).

Figure 1.1(b) shows a linear SISO system with sinusoidal input $\cos(2\pi f_o t)$ where f_o is ordinary frequency in hertz (Hz). As discussed in Chapter 7, a sinusoidal signal is an *eigenfunction* of a linear system, which means that the output is also sinusoidal with the same frequency f_o. For such a signal, the output differs from the input by having a different magnitude, which is A in the figure, and possibly a phase shift ϕ. This is an important characteristic of linear systems that allows us to investigate them in the so-called *frequency domain*, which provides information about their properties beyond those observed in the time domain.

In order to more easily solve for the unknown variables of a mathematical model, the techniques usually require knowledge of *matrices* and *complex numbers*. The matrices covered in Chapter 3 are useful for describing a system of linear equations

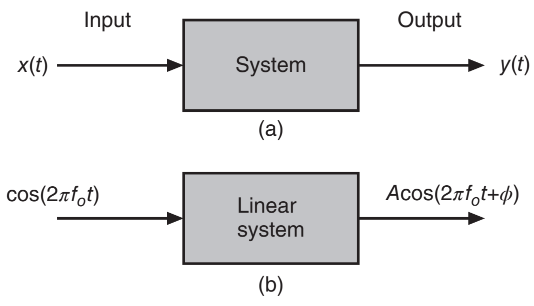


Figure 1.1 Systems with a single input and a single output (SISO). (a) General system with input $x(t)$ and output $y(t)$. (b) Linear system with sinusoidal input and output.

TABLE 1.1 Symbols for Sets of Numbers

Symbol	Domain x	Set
$\mathcal{R}$	$x \in (-\infty, \infty)$	Real numbers
$\mathcal{R}^+$	$x \in [0, \infty)$	Nonnegative real numbers
$\mathcal{Z}$	$x \in \{\dots, -2, -1, 0, 1, 2, \dots\}$	Integers
$\mathcal{Z}^+$	$x \in \{0, 1, 2, \dots\}$	Nonnegative integers
$\mathcal{N}$	$x \in \{1, 2, \dots\}$	Natural numbers
$\mathcal{Q}$	$x = a/b$ with $a, b \in \mathcal{Z}$ and $b \neq 0$	Rational numbers
$\mathcal{I}$	$x = jb$ with $j = \sqrt{-1}$ and $b \in \mathcal{R}$	Imaginary numbers
$\mathcal{C}$	$x = a + jb$ with $j = \sqrt{-1}$ and $a, b \in \mathcal{R}$	Complex numbers
$\mathcal{H}$	$x = a + ib_1 + jb_2 + kb_3$ with $i = j = k = \sqrt{-1}$ and $a, b_1, b_2, b_3 \in \mathcal{R}$	Quaternions

with constant coefficients. Chapter 4 provides the motivation for complex numbers and summarizes many of their properties. Chapter 5 introduces several different waveforms that are used to represent the signals of a system: inputs, outputs, as well as internal waveforms. These include the well-known sinusoidal and exponential signals, as well as the *unit step function* and the *Dirac delta function*. The theory of *generalized functions* and some of their properties are briefly introduced. Systems represented by linear ordinary differential equation*s* (ODEs) are then covered in Chapter 6, where they are solved using conventional *time-domain* techniques. The reader will find that such techniques are straightforward for first- and second-order ODEs, especially for the linear circuits covered in this book, but are more difficult to use for higher order systems.

Chapter 7 describes methods based on the *Laplace transform* that are widely used in engineering to solve linear ODEs with constant coefficients. The Laplace transform converts an ODE into an *algebraic equation* that is more easily solved using matrix techniques. Finally, Chapter 8 introduces methods for analyzing a system in the frequency domain, which provides a characterization of its frequency response to different input waveforms. In particular, we can view linear circuits and systems as *filters* that modify the frequency content of their input signals.

We focus on *continuous-time* systems, which means $\{x(t), y(t)\}$ are defined with *support* $t \in \mathcal{R}$ or $t \in \mathcal{R}^+$ where the functions are nonzero. *Discrete-time* systems and signals are defined for a countable set of time instants such as $\mathcal{Z}$, $\mathcal{Z}^+$, or $\mathcal{N}$. Different but related techniques are used to examine discrete-time systems, though these are beyond the scope of this book.

1.2 MATHEMATICAL MODELS

Consider again the system in Figure 1.1(a) and assume that we have access only to its input $x(t)$ and output $y(t)$ as implied by the block diagram. There is no direct information about the internal structure of the system, and the only way we can learn about its properties is by providing input signals and observing the output signals. Such an unknown system is called a "black box" (because we cannot see inside), and the procedure of examining its input/output characteristics is a type of *reverse engineering*. We mention this because the mathematical models used to represent physical devices and systems are typically verified and even derived from experiments with various types of input/output signals. Such an approach yields the *transfer characteristic* of the system, and for linear and time-invariant (LTI) systems, we can write a specific *transfer function* as described in Chapter 7.

Example 1.1 Suppose input x of an unknown system is varied over $\mathcal{R}$ and we observe the output y shown in Figure 1.2. This characteristic does not change with time, and so we have suppressed the time argument for the input and output. The plot of y is flat for three intervals: $-\infty < x \le -2$, $-1 < x \le 2$, and $3 < x < \infty$, and it is linearly increasing for two intervals: $-2 < x \le -1$ and $2 < x \le 3$. For this *piecewise linear function*, the equation for each interval has the form $y = ax + b$ where $a = \Delta y/\Delta x$ is the *slope* and b is the *ordinate*, which is the point where the line crosses the y-axis if it were extended to $x = 0$. For the first linearly increasing region, the slope is obviously $a = (1 - 0)/[-1 - (-2)] = 1$. When $x = 0$, the extended line crosses the y-axis at $y = 2$, which gives $b = 2$. Similarly, for the second linearly increasing region, $a = (3 - 1)/(3 - 2) = 2$ and $b = -3$. The remaining three regions have zero slope but different ordinates (these equations are of the form $y = b$), and so the overall transfer characteristic for this system is

$$y = \begin{cases} 0, & x \le -2 \\ x + 2, & -2 < x \le -1 \\ 1, & -1 < x \le 2 \\ 2x - 3, & 2 < x \le 3 \\ 3, & x > 3. \end{cases} \tag{1.1}$$

The values of y match at the boundaries for each interval of x as shown in the figure. The *mapping* in (1.1) is a mathematical model for a particular system that can be used to study its behavior even if it is included as part of a larger system. Note that this

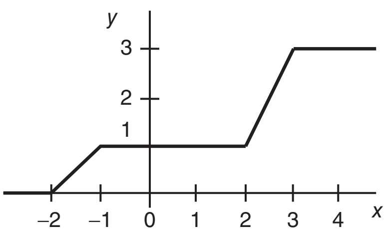


Figure 1.2 Input/output characteristic for the nonlinear system in Example 1.1.

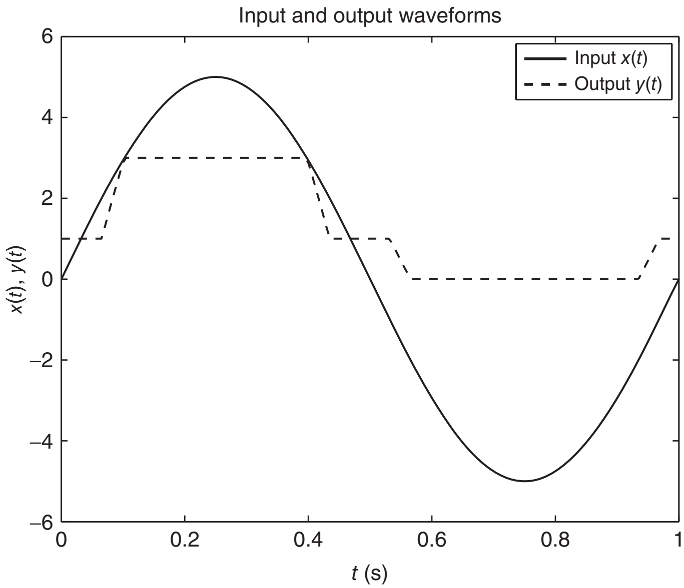


Figure 1.3 Output $y(t)$ for the transfer characteristic in (1.1) in Example 1.1 with input $x(t) = 5\sin(2\pi t)$ for $t \in [0, 1]$.

input/output characteristic does not provide any direct information about the individual components or the internal dynamics of the system. When the input $x(t)$ is a function of time, the output $y(t)$ is also time varying. For example, suppose that $x(t) = 5\sin(2\pi t)$ as illustrated in Figure 1.3 for one period of the sine function with frequency $f_o = 1$ Hz. The output $y(t)$ is computed using (1.1) at each time instant on the closed interval $t \in [0, 1]$ in seconds (s). Observe that $y(t)$ is *truncated* relative to the input waveform due to this particular input/output mapping. Similar results for $y(t)$ can be derived for any input function $x(t)$ by using the model in (1.1).

The output $y(t)$ is *not* sinusoidal because the function in Figure 1.2 is piecewise linear, and so, overall it is *nonlinear*. Sinusoidal signals are not eigenfunctions for nonlinear systems as demonstrated in this example. Eigenfunctions and their defining properties are covered later in Chapter 7. The *fundamental frequency* of the output in Figure 1.3 is $f_o = 1$ Hz because the waveform for all $t \in \mathcal{R}$ consists of repetitions of the 1 s segment dashed curve. The waveform within this segment also has variations, which result in *harmonics* of f_o. This means that sinusoidal components with integer multiples of f_o are also present in $y(t)$. It is possible to determine these harmonics using a *Fourier series* representation of $y(t)$ as discussed in Chapter 5.

Example 1.2 Consider the following mapping:

$$y = 2x - 3, \quad x \in \mathcal{R}, \tag{1.2}$$

which is one component of (1.1) with support extended to the entire real line, and so, the input is not truncated. For $x(t) = 5\sin(2\pi t)$, the output of this system is

$$y(t) = 10\sin(2\pi t) - 3, \tag{1.3}$$

which has the same frequency $f_o = 1$ Hz as the input; there are no harmonics of f_o. However, this system is not linear because it introduces a DC ("direct current") component at $f = 0$ Hz, which causes the output to be shifted downward, as illustrated in Figure 1.4 (the dashed line). The function in (1.2) is actually *affine* because of the nonzero ordinate $b = -3$. A linear function is obtained by dropping the ordinate:

$$y = 2x, \qquad x \in \mathcal{R}, \tag{1.4}$$

which has the output in Figure 1.4 (the dotted line). This is a trivial system because the peak amplitude 10 of the output is unchanged for any input frequency f_o, and the phase shift ϕ is always zero.

A linear system that is modeled by an ODE has a more complicated representation than the simple scaling in (1.4), and the amplitude and phase of its output generally change with frequency f_o. By varying the frequency of the input and observing the output of a linear system, we can derive its frequency response. This representation of a system indicates which frequency components of a signal are attenuated or

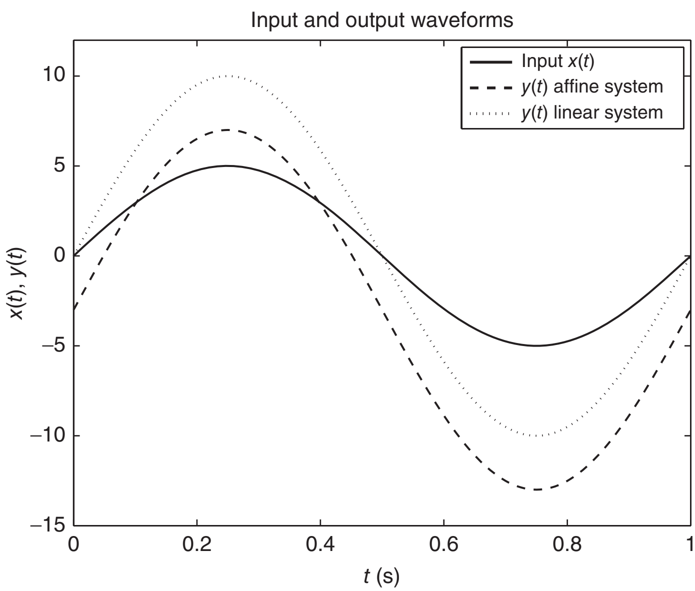


Figure 1.4 Output $y(t)$ for the transfer characteristics in (1.2) and (1.4) in Example 1.2 with input $x(t) = 5\sin(2\pi t)$ for $t \in [0, 1]$.

amplified and whether they are shifted in time. Using this approach, the system can be viewed as a type of filter that modifies the frequency characteristics of the input signal. For example, a low-pass filter retains only low-frequency components while attenuating or blocking high frequencies. It is useful in many applications such as noise reduction in communication systems. The frequency response of a system is investigated further in Chapter 8 where we cover the *Fourier transform*.

Example 1.3 An example of a system of linear equations is

$$a_{11}y_1(t) + a_{12}y_2(t) = x_1(t), \tag{1.5}$$

$$a_{21}y_1(t) + a_{22}y_2(t) = x_2(t), \tag{1.6}$$

where $\{y_1(t), y_2(t)\}$ are unknown outputs, $\{x_1(t), x_2(t)\}$ are known inputs, and $\{a_{mn}\}$ are constant coefficients. (Many books on linear algebra have x and y interchanged. We use the form in (1.5) and (1.6) for notational consistency throughout the book, where known x is the input and unknown y is the output.) These equations can be viewed as a multiple-input multiple-output (MIMO) system as depicted in Figure 1.5. It is straightforward to solve for the unknown variables $\{y_1(t), y_2(t)\}$ by first rearranging (1.6) as

$$y_2(t) = x_2(t)/a_{22} - a_{21}y_1(t)/a_{22}, \tag{1.7}$$

and then substituting (1.7) into (1.5):

$$a_{11}y_1(t) + a_{12}x_2(t)/a_{22} - a_{12}a_{21}y_1(t)/a_{22} = x_1(t), \tag{1.8}$$

which gives

$$y_1(t) = \frac{x_1(t) - a_{12}x_2(t)/a_{22}}{a_{11} - a_{12}a_{21}/a_{22}} = \frac{a_{22}x_1(t) - a_{12}x_2(t)}{a_{11}a_{22} - a_{12}a_{21}}, \tag{1.9}$$

and likewise for the other output:

$$\begin{aligned} y_2(t) &= x_2(t)/a_{22} - (a_{21}/a_{22})\frac{a_{22}x_1(t) - a_{12}x_2(t)}{a_{11}a_{22} - a_{12}a_{21}} \\ &= \frac{a_{11}x_2(t) - a_{21}x_1(t)}{a_{11}a_{22} - a_{12}a_{21}}. \end{aligned} \tag{1.10}$$

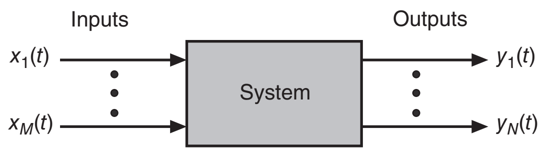


Figure 1.5 Multiple-input and multiple-output (MIMO) system.

The reader may recognize that if (1.5) and (1.6) are written in matrix form as described in Chapter 3, then the denominator in (1.9) and (1.10) is the *determinant* $\det(\mathbf{A}) = a_{11}a_{22} - a_{12}a_{21}$ of the matrix

$$\mathbf{A} \triangleq \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}. \tag{1.11}$$

It is usually convenient to write such systems of equations in matrix form, because it is then straightforward to examine their properties based on the structure and elements of $\mathbf{A}$. Moreover, we can write the solution of the linear equations $\mathbf{Ay}(t) = \mathbf{x}(t)$ via the matrix inverse as $\mathbf{y}(t) = \mathbf{A}^{-1}\mathbf{x}(t)$, where for this two-dimensional matrix, the column vectors are

$$\mathbf{x}(t) \triangleq \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}, \quad \mathbf{y}(t) \triangleq \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix}. \tag{1.12}$$

For a numerical example, let the matrix elements be $a_{11} = a_{21} = a_{22} = 1$ and $a_{12} = -0.1$, and assume the inputs are constant: $x_1(t) = 0$ and $x_2(t) = 1$. Then from (1.9) and (1.10), we have the explicit solution $y_1(t) = 1/11 \approx 0.0909$ and $y_2(t) = 10/11 \approx 0.9091$.

Example 1.4 In this example, we examine a *nonlinear* system to illustrate the difficulty of solving for the output variables of such models. A MIMO system is described by two equations, the first of which is nonlinear:

$$a_{11}y_1(t) + a_{12}\exp(\alpha y_2(t)) = x_1(t), \tag{1.13}$$

$$a_{21}y_1(t) + a_{22}y_2(t) = x_2(t), \tag{1.14}$$

where α and the coefficients $\{a_{mn}\}$ are constant parameters. This system is similar to the one in Example 1.3, except that a_{12} multiplies the *exponential function*

$$\exp(\alpha y_2(t)) \triangleq e^{\alpha y_2(t)}, \tag{1.15}$$

where e is Napier's constant which is reviewed later in this chapter. The inputs are again $\{x_1(t), x_2(t)\}$, and we would like to find a solution for $\{y_1(t), y_2(t)\}$. Unlike the linear system of equations in the previous example, eliminating one variable by substituting one equation into the other does not yield a closed-form solution because of the exponential function. Figure 1.6(a) shows examples of these two equations, obtained by plotting y_1 versus y_2 for the parameters used at the end of Example 1.3 and with $\alpha = 4$. Since $\{y_n\}$ must simultaneously satisfy both equations, it is clear that the solution for this system of equations occurs where the two curves (solid and dashed) in the figure intersect. One approach to finding the solution is *iterative*, where an initial estimate is chosen for y_2, from which it is possible to solve for y_1 using (1.14). This value for y_1 is substituted into (1.13), which is rewritten as follows:

$$y_2 = (1/\alpha)\ln((x_1 - a_{11}y_1)/a_{12}), \tag{1.16}$$

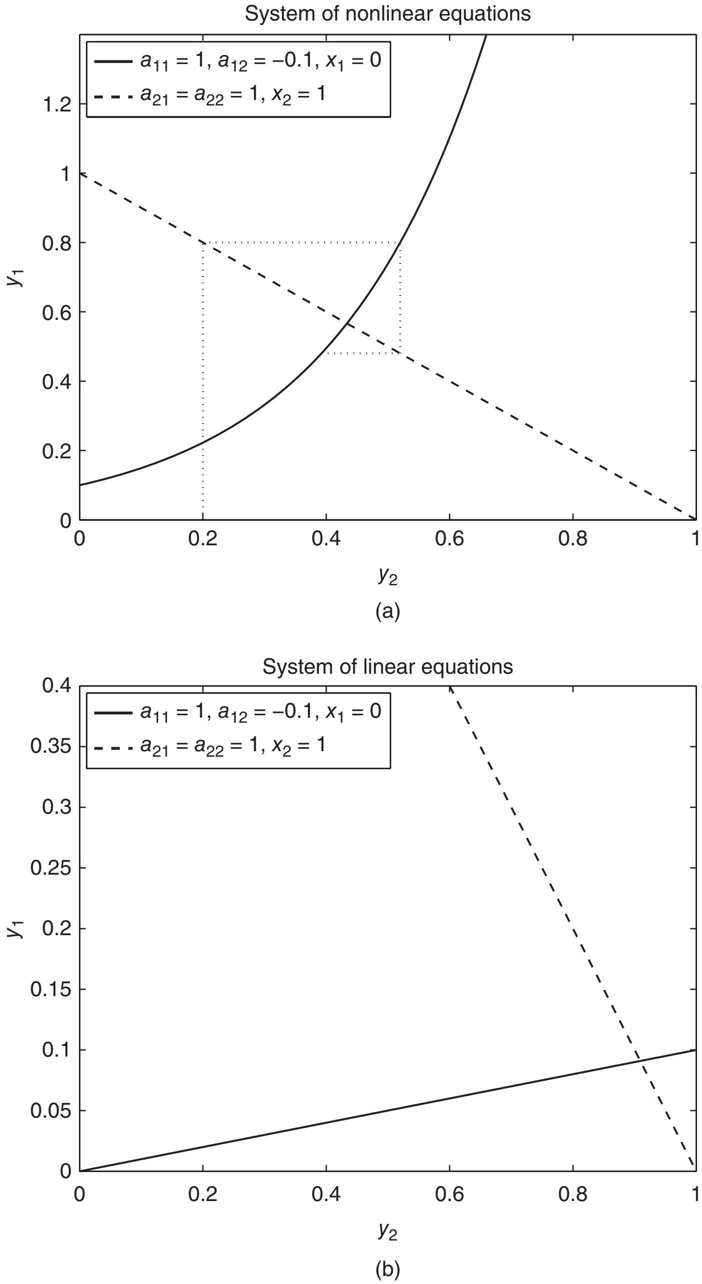


Figure 1.6 Systems of equations. (a) Nonlinear system in (1.13) and (1.14) in Example 1.4 with $\alpha = 4$. (b) Linear system in (1.5) and (1.6) in Example 1.3.

where ln $(\cdot)$ is the *natural logarithm*. This equation yields a new value for y_2, which is used in (1.14) to compute y_1, and the procedure is repeated several times until $\{y_n\}$ no longer change (up to some desired numerical precision), and so they have *converged* to a solution. For the aforementioned parameters and initial value $y_2 = 0.2$, we find using MATLAB that the solution is $y_1 \approx 0.5664$ and $y_2 \approx 0.4336$, which is verified by the intersecting curves in Figure 1.6(a). The first four iterations are denoted by the dotted lines in the figure, which we see approach the solution. For comparison purposes, Figure 1.6(b) shows the two lines for the linear system in Example 1.3 using the same coefficient values. The solution is located where the two lines intersect: $y_1 \approx 0.0909$ and $y_2 \approx 0.9091$. Since this system of equations is linear, we can solve for y_1 and y_2 explicitly as was done in (1.9) and (1.10) (there is no need to perform iterations).

We mention that it is possible to find a type of explicit solution for the system of nonlinear equations in the previous example by using the Lambert W-function described in Appendix F, which includes some examples. Nonlinear circuit equations for the diode are briefly discussed in Chapter 2, and an explicit solution using the Lambert W-function for a simple diode circuit is derived in Appendix F. Although an explicit solution is obtained, it turns out that the Lambert W-function cannot be written in terms of ordinary functions, and so, it must be solved numerically.

The transfer characteristic in Example 1.1 is *static* because it describes the output $y(t)$ for a given input $x(t)$ independently of the time variable t. For many physical systems, the transfer characteristic also depends on other factors, such as the *rate* at which $x(t)$ changes over time. This type of system is modeled by an ODE. In subsequent chapters, we describe techniques used to evaluate and solve linear ODEs for systems in general as in Figure 1.1 and for linear circuits in particular.

Example 1.5 An example of a linear ODE is

$$\frac{d^2}{dt^2}y(t) + a_1\frac{d}{dt}y(t) + a_0y(t) = x(t), \quad t \in \mathcal{R}, \tag{1.17}$$

where time t is the *independent variable*, $y(t)$ is the *unknown* dependent variable, and $x(t)$ is the *known* dependent variable. For the system in Figure 1.1(a), $x(t)$ is the input and $y(t)$ is the output. The coefficients $\{a_0, a_1\}$ are fixed, and the goal is to find a solution for $y(t)$ given these parameters as well as the initial conditions $y(0)$ and $y'(0)$. The superscript denotes the ordinary derivative of $y(t)$ with respect to t, which is then evaluated at $t = 0$:

$$y'(0) \triangleq \left.\frac{d}{dt}y(t)\right|_{t=0.} \tag{1.18}$$

Equation 1.17 is a *second-order* ODE because it contains the second derivative of $y(t)$; higher order derivatives are considered in Chapter 7. An implementation based on *integrators* is illustrated in Figure 1.7. This configuration is preferred in practice because differentiators enhance additive noise in a system (Kailath, 1980), which can overwhelm the signals of interest. Integrators, on the other hand, average out

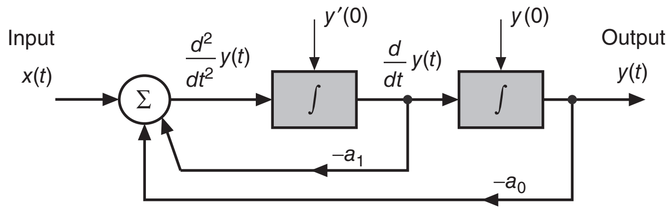


Figure 1.7 Integrator implementation of a second-order linear ODE.

additive noise, which often has a zero average value. This implementation is obtained by bringing the $\{a_0, a_1\}$ terms of (1.17) to the right-hand side of the equation such that the output of the summing element in the figure is

$$\frac{d^2}{dt^2}y(t) = x(t) - a_1\frac{d}{dt}y(t) - a_0y(t). \tag{1.19}$$

The cascaded integrators sequentially yield $dy(t)/dt$ and $y(t)$. The solution to (1.17) when $x(t) = 0$ and $a_0 = a_1 = 2$ is

$$y(t) = \exp(-t)[2\sin(t) + \cos(t)], \quad t \in \mathcal{R}^+, \tag{1.20}$$

where the nonzero initial conditions $y(0) = y'(0) = 1$ have been assumed. This waveform is plotted in Figure 1.8 (the solid line) from which we can easily verify the initial conditions. It is straightforward to show that (1.20) is the solution of (1.19) by differentiating $y(t)$:

$$\begin{aligned}\frac{d}{dt}y(t) &= -\exp(-t)[2\sin(t) + \cos(t)] + \exp(-t)[2\cos(t) - \sin(t)] \\ &= \exp(-t)[\cos(t) - 3\sin(t)],\end{aligned} \tag{1.21}$$

$$\begin{aligned}\frac{d^2}{dt^2}y(t) &= -\exp(-t)[\cos(t) - 3\sin(t)] + \exp(-t)[-\sin(t) - 3\cos(t)] \\ &= \exp(-t)[2\sin(t) - 4\cos(t)].\end{aligned} \tag{1.22}$$

Substituting these expressions into (1.17) with $a_0 = a_1 = 2$, we find that all terms cancel to give 0. By changing the coefficients, a different output response is obtained. For example, when $a_0 = 2$ and $a_1 = 3$, the solution is purely exponential:

$$y(t) = 3\exp(-t) - 2\exp(-2t), \quad t \in \mathcal{R}^+. \tag{1.23}$$

This is also plotted in Figure 1.8 for the same initial conditions and input $x(t) = 0$ (the dashed line). The solutions in (1.20) and (1.23) are known as *underdamped* and *overdamped*, respectively. It turns out that there is a third type of solution for a second-order ODE called *critically damped*, which is obtained by changing the coefficient values. All three solutions are discussed in greater detail in Chapters 6 and 7.

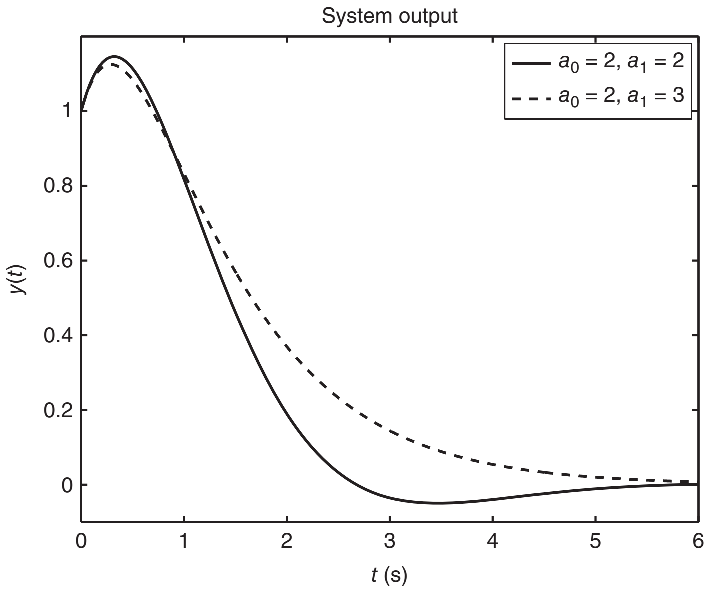


Figure 1.8 Solutions for the second-order ODE in Example 1.5 with constant coefficients. The input is $x(t) = 0$ and the initial conditions are nonzero: $y(0) = y'(0) = 1$.

1.3 FREQUENCY CONTENT

As mentioned earlier, the main goal of this book is to develop mathematical models for circuits and systems, and to describe techniques for finding expressions (solutions) for the dependent variables of interest. In addition, we are interested in the *frequency content* of signals and the *frequency response* of different types of systems. This frequency information illustrates various properties of signals and systems beyond that observed from their time-domain representations.

The most basic signal is sinusoidal with angular frequency $\omega_o = 2\pi f_o$ in radians/second (rad/s) and ordinary frequency f_o in Hz. It turns out that all *periodic* signals can be represented by a sum of sinusoidal signals with fundamental frequency f_o and integer multiples $n f_o$ for $n \in \mathcal{Z}$ called harmonics. For example, the periodic rectangular waveform in Figure 1.9(a) has the frequency *spectrum* shown in Figure 1.9(b), with the magnitude of each frequency component indicated on the vertical axis. Lower harmonics have greater magnitudes, demonstrating that this waveform is dominated by low frequencies. This frequency representation for periodic signals is known as the Fourier series and is covered in Chapter 5. Aperiodic signals, which do not repeat, have a frequency representation known as the Fourier transform. Whereas the Fourier series consists of integer multiples of a fundamental frequency, the Fourier transform is a *continuum of frequencies* as illustrated in Figure 1.10 for triangular and rectangular waveforms. Both of these signals are dominated by low-frequency content.

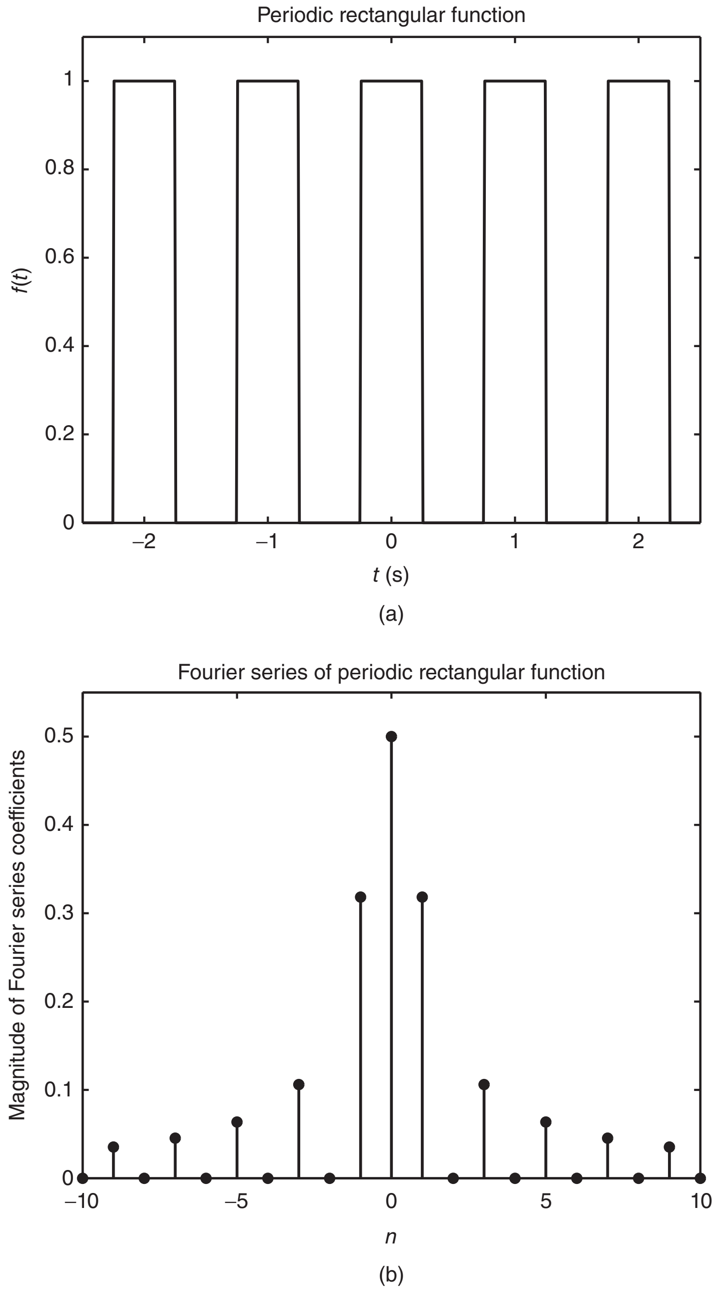


Figure 1.9 Periodic rectangular waveform. (a) Time-domain representation. (b) Magnitude of frequency spectrum: Fourier series with harmonics $n f_o$ and $f_o = 1$ Hz.

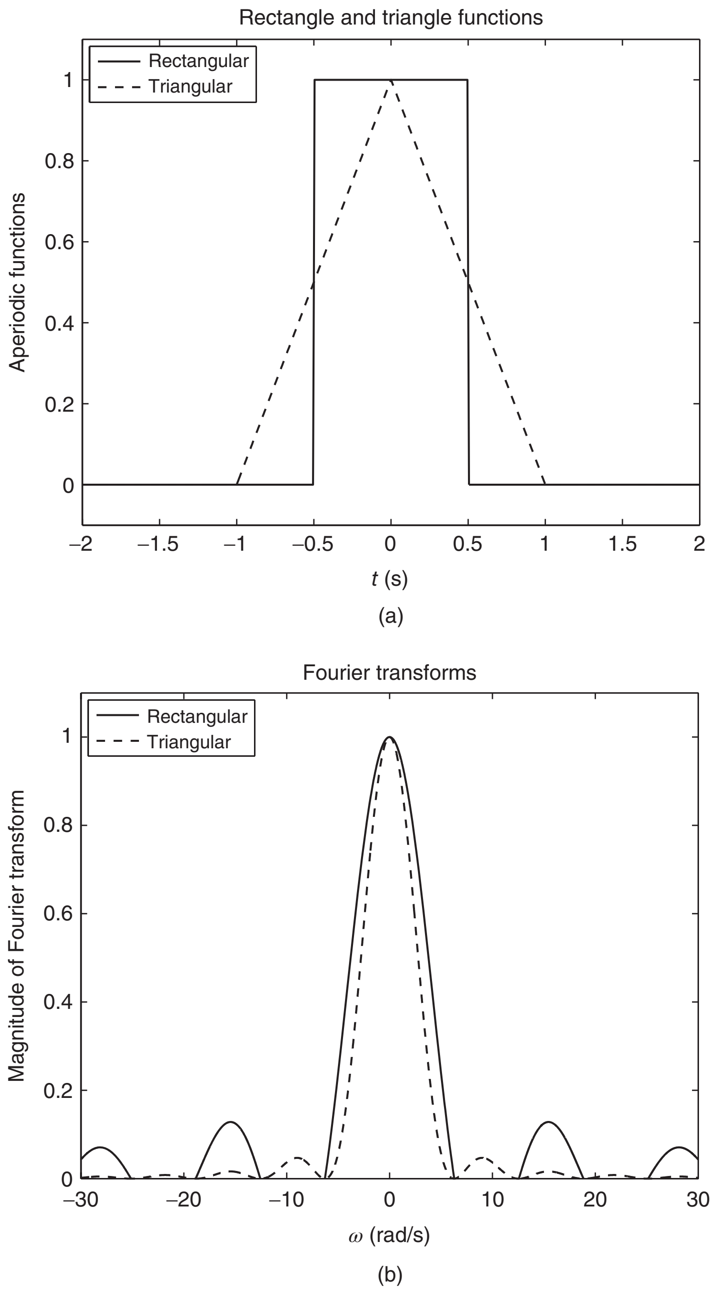
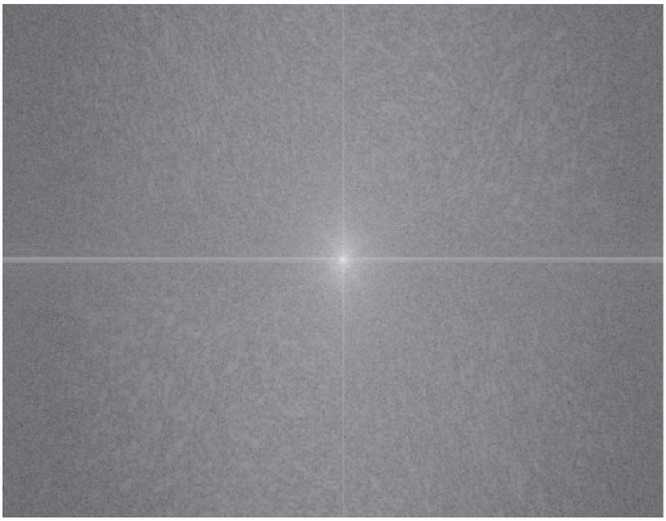


Figure 1.10 Aperiodic waveforms. (a) Time-domain representation. (b) Magnitude of frequency spectrum: Fourier transform.

(a)



(b)

Figure 1.11 Two-dimensional image and spectrum. (a) Spatial representation. (b) Magnitude of frequency spectrum in two dimensions. White denotes a greater magnitude. (The vertical and horizontal white lines are the frequency axes where $\omega_1 = 0$ and $\omega_2 = 0$. A log scale is used to better visualize variations in the spectrum.)

Although we focus on one-dimensional signals in this book, which are generally a function of the independent variable time t, Figure 1.11 shows a two-dimensional *image* and its frequency representation. The two independent variables in Figure 1.11(a) are given by the horizontal (width) and vertical (height) axes, and the information contained in the image is indicated by a gray scale from white to black. Similarly, the magnitude of the spectrum in Figure 1.11(b) is represented by a gray scale, with white denoting a greater magnitude for particular frequencies. Two frequency variables $\{\omega_1, \omega_2\}$ are used in the Fourier transform of a two-dimensional image (Bracewell, 1978); these are the horizontal and vertical axes in Figure 1.11(b). Low frequencies are located around the center of the plot, and high frequencies (positive and negative) extend outward to the edges of the spectrum plot. Once again, we have a signal with mostly low-frequency content; in fact, the spectrum is dominated by the white "star" located about the center where $\omega_1 = \omega_2 = 0$. This occurs because there is not much spatial variation across the image in Figure 1.11(a). In general, greater variations in the time/spatial domain correspond to higher frequencies with greater magnitudes in the Fourier/frequency domain.

Systems are often designed to have a particular frequency response where some frequencies are emphasized and others are attenuated. For example, a system that passes low frequencies and attenuates high frequencies is called a *low-pass filter*. Likewise, systems can be designed to have a high-pass, band-pass, or band-reject frequency response. Conventional amplitude modulated (AM) radio is an example of a system that incorporates band-pass filters to select a transmitted signal located in a specific radio frequency *channel*. Such a channel is defined by a *center frequency* and a *bandwidth* over which the signal can be transmitted without interfering with other signals in nearby channels. The Fourier transform and different types of filters are covered in Chapter 8.

In the rest of this chapter, we provide a review of some basic topics that the reader has probably studied to some extent, and which form the basis of the material covered throughout this book.

1.4 FUNCTIONS AND PROPERTIES

We begin with a summary of basic definitions for functions of a single independent variable.

Definition: Function The *function* $y = f(x)$ is a unique mapping from input x to output y.

Although x yields a single value y, more than one value of x could map to the same y. (Note, however, that it is possible to define multiple-output functions; an example of this is the Lambert W-function discussed in Appendix F.)

Definition: Domain and Range The *domain* of function $f(x)$ consists of those values of x for which the function is defined. The *range* of a function is the set of values $y = f(x)$ generated when x is varied over the domain.

Example 1.6 For $f(x) = x^2$, the *natural* domain is $\mathcal{R}$ (although it is possible to restrict the domain to some finite interval), and the corresponding range is $\mathcal{R}^+$. The domain for $f(x) = \log(x)$ is $\mathcal{R}^+$ and its range is $\mathcal{R}$.

Definition: Support The *support* of a function is the set of x values for which $f(x)$ is nonzero.

Example 1.7 The domain of the *unit step* function is $\mathcal{R}$:

$$u(x) \triangleq \begin{cases} 1, & x \geq 0 \\ 0, & x < 0, \end{cases} \tag{1.24}$$

but its support is $\mathcal{R}^+$. Similarly, the domain of the truncated sine function $\sin(\omega_o t)u(t)$ is $\mathcal{R}$ and its support is $\mathcal{R}^+$. Even though sine is 0 for integer multiples of π, the support is still $\mathcal{R}^+$ because sine is a continuous function and those points (which form a *countable* set) are not excluded from the support.

Definition: Inverse Image and Inverse Function The *inverse image* $x = f^{-1}(y)$ is the set of all values x that map to y. The inverse image of a function may not yield a unique x. If a single $x = f^{-1}(y)$ is generated for each y, then $f(x)$ is *one-to-one* and the inverse image is equivalent to the *inverse function* $x = f^{-1}(y) \triangleq g(y)$.

Example 1.8 For the quadratic function $y = x^2$, it is obvious that each $x \in \mathcal{R}$ gives a single y. Solving for x yields $x = \pm\sqrt{y}$. Since x is not unique for each y, the square root is *not* the inverse function. An inverse function does not exist for $y = x^2$. However, $x = f^{-1}(y) = \pm\sqrt{y}$ describes the inverse image; for example, the inverse image of $y = 9$ is the *set* of values $x = \{-3, 3\}$. The one-to-one function $y = 2x + 1$ has inverse function $x = g(y) = (y - 1)/2$. The natural logarithm $y = \ln(x)$ is also one-to-one with inverse function $x = g(y) = \exp(y)$.

Definition: Linear Function A *linear function* $f(x)$ has the following two properties:

$$f(x_1 + x_2) = f(x_1) + f(x_2), \quad f(\alpha x) = \alpha f(x), \tag{1.25}$$

where $\alpha \in \mathcal{R}$ is any constant.

The line representing a linear function necessarily passes through the origin: $y(x) = 0$ when $x = 0$.

Example 1.9 The circuit model shown in Figure 1.12(a) for a resistor with resistance R has the form $v = Ri$ known as *Ohm's law*. It is a linear function:

$$v_1 = Ri_1, \quad v_2 = Ri_2 \;\Rightarrow\; v_1 + v_2 = R(i_1 + i_2) = Ri_1 + Ri_2, \tag{1.26}$$

$$v = Ri \;\Rightarrow\; \alpha v = R(\alpha i) = \alpha Ri, \tag{1.27}$$

where v is a voltage and i is a current (both are defined in Chapter 2). An example of a nonlinear function is the piecewise linear circuit model for a diode that is in series with resistor R:

$$i = \begin{cases} (v - v_c)/R, & v \geq v_c \\ 0, & v < v_c, \end{cases} \tag{1.28}$$

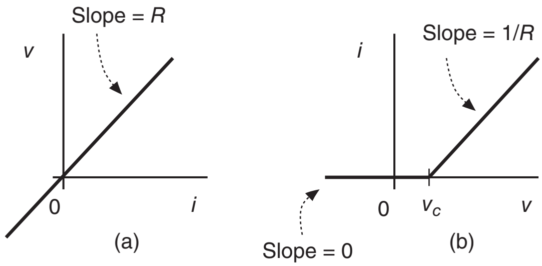


Figure 1.12 Device models used in Example 1.9. (a) Linear model for resistor R. (b) Nonlinear model for diode D with resistance R.

where v_c is a cutoff voltage; typically $v_c \approx 0.7$ V (volt). Although this equation has straight-line components (it is piecewise linear), overall it is nonlinear as depicted in Figure 1.12(b) because it does not satisfy (1.25). Suppose $v_1 = -2$ V such that $i_1 = 0$ A (ampere), and $v_2 = 1.7$ V such that $i_2 = (1/R)$ A. Then $v_1 + v_2 = -0.3$ V $\Rightarrow\ i = 0$ A, which is not equal to $i_1 + i_2 = (1/R)$ A.

The general equation $y = ax + b$ for a line is not linear even though it is straight and is used to describe the different parts of a piecewise linear function (as in Example 1.1). A linear function based on the properties in (1.25) must pass through the origin.

Definition: Affine Function *Affine function* $g(x)$ is a linear function $f(x)$ with additive scalar b such that the ordinate is nonzero:

$$g(x) = f(x) + b. \tag{1.29}$$

An affine function does not satisfy either requirement in (1.25) for a linear function:

$$g(x_1 + x_2) = f(x_1 + x_2) + b \neq g(x_1) + g(x_2) = f(x_1) + b + f(x_2) + b, \tag{1.30}$$

$$g(\alpha x) = f(\alpha x) + b \neq \alpha g(x) = \alpha f(x) + \alpha b, \tag{1.31}$$

where $\alpha \in \mathcal{R}$ is any nonzero constant.

Definition: Continuous Function Function $f(x)$ is *continuous* at x_o if there exists $\epsilon > 0$ for every $\delta > 0$ such that

$$|x - x_o| < \epsilon \ \Rightarrow\ |f(x) - f(x_o)| < \delta. \tag{1.32}$$

More simply we can write

$$\lim_{\epsilon \to 0} |f(x_o + \epsilon) - f(x_o)| = 0, \tag{1.33}$$

where ϵ is either positive or negative such that $f(x_o + \epsilon)$ approaches x_o from the right or the left, respectively.

All the functions shown in Figures 1.8 and 1.12 are continuous. An example of a function that is continuous only *from the right* is shown in Figure 1.13. Approaching x_o from the left, the function jumps to the higher value b. A solid circle indicates that the function is continuous approaching from the right, meaning that the function is b at x_o. A function that is continuous at x_o from the left is similarly defined with the solid and open circles in Figure 1.13 interchanged. If a function is left- and right-continuous at x_o, then it is strictly continuous at that point as defined in (1.32) and (1.33).

Functions of a real variable can have different types of discontinuities. The plot in Figure 1.13 shows a function with a *jump discontinuity*. Another example is the unit step function $u(t)$ in (1.24), which is used extensively throughout this book. Similar to the example in Figure 1.13, $u(t)$ is continuous from the right but not from the left. A function that is *nowhere continuous* is the Dirichlet function, given by

$$f(x) = \begin{cases} 1, & x \in \mathcal{Q} \\ 0, & x \in \mathcal{R} - \mathcal{Q}, \end{cases} \tag{1.34}$$

where $\mathcal{Q}$ is the set of rational numbers. It is not possible to accurately plot this function using MATLAB (or any other mathematics software). Another type of discontinuity is an *infinite discontinuity*, also called an *asymptotic discontinuity*. Examples include

$$f(x) = 1/x, \quad f(x) = 1/(x-1)(x-2), \tag{1.35}$$

where, in the first case, the discontinuity is at $x = 0$, and in the second case, there are discontinuities at $x = \{1, 2\}$. The second function is plotted in Figure 1.14(a), which we see is continuous except at the two points indicated by the vertical dotted lines. With the terminology of functions of complex variables considered later in this book (see Chapter 5 and Appendix E), these *singularities* are called *poles*.

Consider the function

$$f(x) = \frac{\sin(x)}{x}, \tag{1.36}$$

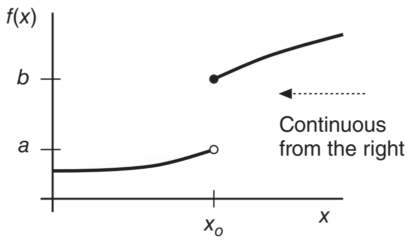


Figure 1.13 Example of a function with a discontinuity at x_o.

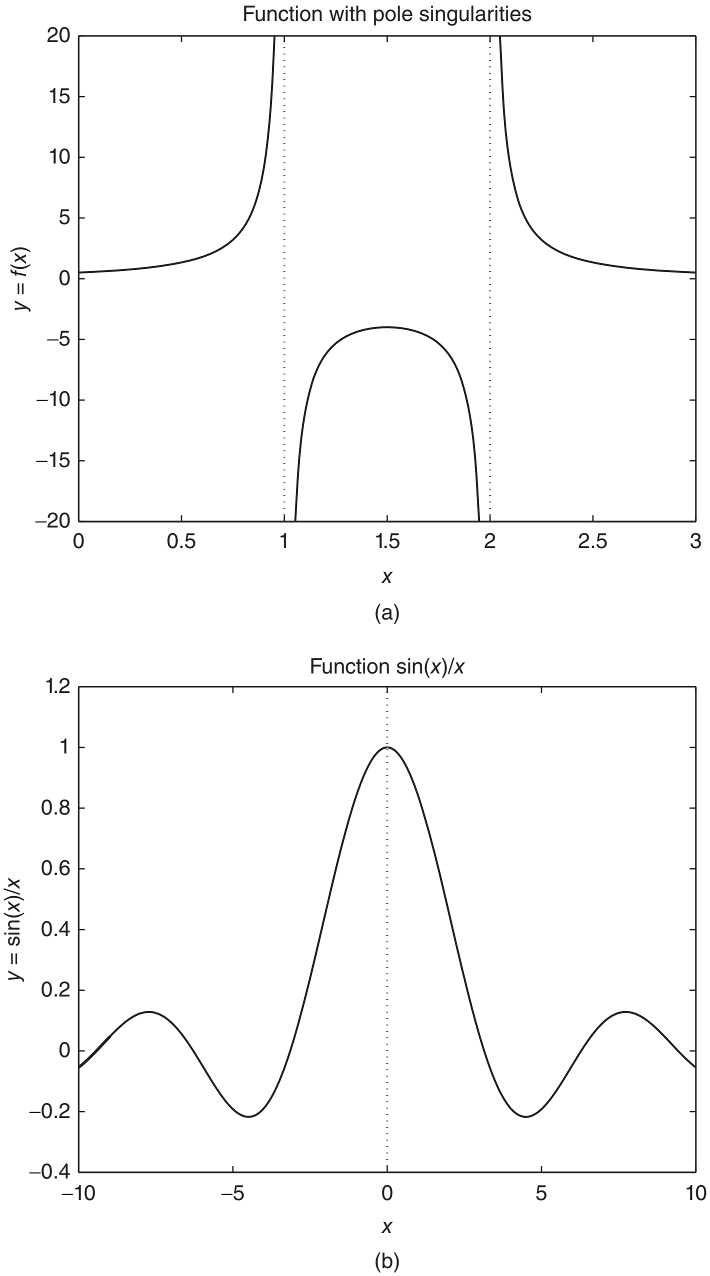


Figure 1.14 (a) Function with two pole singularities at $x = \{1, 2\}$. (b) Function with a removable pole singularity at $x = 0$.

which appears to have a pole at $x = 0$. It turns out, however, that this pole is cancelled by the numerator such that $f(0) = 1$. This can be seen using L'Hôpital's rule

$$\frac{d\sin(x)/dx|_{x=0}}{dx/dx|_{x=0}} = \cos(0) = 1. \tag{1.37}$$

Such singularities are called *removable*. This function, which is plotted in Figure 1.14(b), is known as the unnormalized sinc function, and should not be confused with the usual sinc function $\text{sinc}(x) \triangleq \sin(\pi x)/\pi x$ discussed in subsequent chapters. Another example of a removable singularity is the following rational function, where a factor in the numerator cancels the denominator:

$$f(x) = \frac{x^2 - 1}{x + 1} = x - 1, \tag{1.38}$$

and so $f(-1) = -2$. A function with a singularity for which there is no limit is called an *essential singularity*. The classic example is

$$f(x) = \sin(1/x), \tag{1.39}$$

which is plotted in Figure 1.15. Observe that as x approaches 0, there is no single finite value for the function.

Finally, ordinary functions can be divided into two basic types.

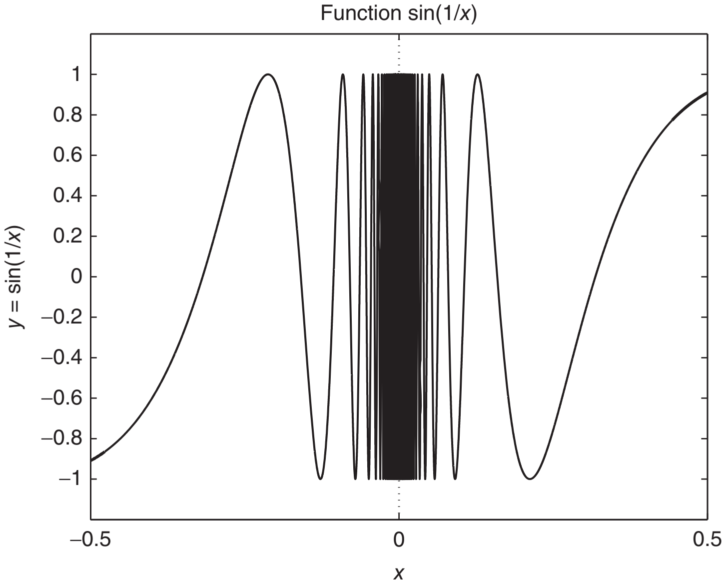


Figure 1.15 Function with an essential singularity at $x = 0$.

Definition: Algebraic Functions and Transcendental Equations An *algebraic function* $f(x)$ satisfies the following polynomial equation:

$$p_n(x)f^n(x) + \ldots + p_1(x)f(x) + p_0(x) = 0, \tag{1.40}$$

where $\{p_m(x)\}$ are polynomials in x and $n \in \mathcal{N}$ (the natural numbers $\{1, 2, \ldots\}$). All other equations are *transcendental equations*, such as those containing exponential, logarithmic, and trigonometric functions.

Example 1.10 Examples of algebraic functions are

$$f(x) = x^4 + x_2 - x + 1, \quad f(x) = \sqrt{x}, \quad f(x) = 1/x^2, \tag{1.41}$$

and examples of transcendental functions are

$$f(x) = \log(x), \quad f(x) = \tan^{-1}(x), \quad f(x) = \cos(x)\tan(x). \tag{1.42}$$

Both types of functions/equations are considered in this book. In Chapter 4, we find that the solutions to some algebraic equations require *complex numbers*. The class of ordinary functions is extended in Chapter 5 to *generalized functions*, which include the Dirac delta function $\delta(x)$ and its derivatives $\delta^{(n)}(x)$.

1.5 DERIVATIVES AND INTEGRALS

In this section, definitions for the ordinary derivative of a function of one independent variable and its Riemann integral are reviewed.

Definition: Derivative The *derivative* of function $f(x)$ is another function that gives the rate of change of $y = f(x)$ as x is varied.

The following notations are used to represent the derivative of $y = f(x)$:

$$\frac{dy}{dx}, \quad \frac{d}{dx}f(x), \quad f'(x), \quad \dot{y}, \tag{1.43}$$

though the last form is usually reserved for the derivative of $y(t)$ with respect to time t: $\dot{y} = dy/dt$. The derivative of a continuous function is generated from the following limit:

$$\frac{d}{dx}f(x) = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}, \tag{1.44}$$

where a *secant line* connects the points $\{x, f(x)\}$ and $\{x + \Delta x, f(x + \Delta x)\}$. As $\Delta x \to 0$, the *family* of secant lines approach the *tangent line* at x as shown for the function in Figure 1.16. The next example demonstrates how to use this definition of the derivative for two of the functions in Example 1.8.

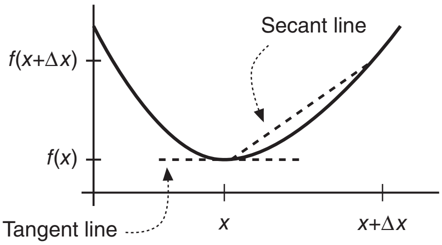


Figure 1.16 Finite approximation of the derivative of $f(x)$ at x.

Example 1.11 For $y = x^2$:

$$\frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{(x+\Delta x)^2 - x^2}{\Delta x} = \lim_{\Delta x \to 0} \frac{2x\Delta x + \Delta x^2}{\Delta x} = 2x, \tag{1.45}$$

and for $y = 2x + 1$:

$$\frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{[2(x+\Delta x)+1] - (2x+1)}{\Delta x} = \lim_{\Delta x \to 0} \frac{2\Delta x}{\Delta x} = 2. \tag{1.46}$$

For the latter affine function, the derivative is a constant equal to the slope. In general, the derivative varies with x, as it does for the quadratic function $f(x) = x^2$.

Example 1.12 Consider the derivative of the absolute value function $y = |x|$:

$$\begin{aligned} \frac{dy}{dx} &= \lim_{\Delta x \to 0} \frac{|x+\Delta x| - |x|}{\Delta x} \\ &= \lim_{\Delta x \to 0} \begin{cases} (x+\Delta x - x)/\Delta x, & x > 0 \\ (-x - \Delta x + x)/\Delta x, & x < 0. \end{cases} \end{aligned} \tag{1.47}$$

Thus, as $\Delta x \to 0$:

$$\frac{dy}{dx} = \begin{cases} 1, & x > 0 \\ -1, & x < 0. \end{cases} \tag{1.48}$$

Although the absolute value function is continuous at all points, its derivative does not exist at $x = 0$ because the ratio in (1.47) is not defined there in the limit. However, since this is usually not an issue in practice, $d|x|/dx = \text{sgn}(x)$ is often used where sgn(x) is the *signum function*:

$$\text{sgn}(x) \triangleq \begin{cases} 1, & x > 0 \\ 0, & x = 0 \\ -1, & x < 0. \end{cases} \tag{1.49}$$

The derivative of function $y(x)$ can be extended to include points where $dy(x)/dx$ is not defined by using the theory of generalized functions, as discussed in Chapter 5. This is even more evident for the derivative of the signum function:

$$\frac{d}{dx}\text{sgn}(x) = 2\delta(x), \tag{1.50}$$

where $\delta(x)$ is the Dirac delta function. This result cannot be derived using the difference approach in (1.44)

It is not necessary to use the limit in (1.44) to find derivatives because many results have already been established for a wide range of functions. For convenience, Appendix C summarizes the derivatives of several ordinary functions. The following important properties of the derivative are provided without proof, which can be used to derive results for more complicated functions.

- *Addition and scalar multiplication*:

$$\frac{d}{dx}[\alpha f(x) + \beta g(x)] = \alpha\frac{d}{dx}f(x) + \beta\frac{d}{dx}g(x), \tag{1.51}$$

  with $\alpha, \beta \in \mathcal{R}$.
- *Product rule*:

$$\frac{d}{dx}[f(x)g(x)] = g(x)\frac{d}{dx}f(x) + f(x)\frac{d}{dx}g(x). \tag{1.52}$$

- *Quotient rule*:

$$\frac{d}{dx}\frac{f(x)}{g(x)} = [1/g^2(x)]\left[g(x)\frac{d}{dx}f(x) - f(x)\frac{d}{dx}g(x)\right]. \tag{1.53}$$

- *Chain rule*:

$$\frac{d}{dx}f(g(x)) = \frac{d}{dg(x)}f(g(x))\frac{d}{dx}g(x). \tag{1.54}$$

As shown earlier, the *independent variable* of the function in a derivative is often suppressed for notational convenience. For example, we usually just write dy/dx, which is the same as $dy(x)/dx$; we also use $y'(x)$ as was done for the initial condition $y'(0)$ in Figure 1.7. For the nth-order derivative, a superscript is used: $y^{(n)}(x)$, or multiple primes $y''(x)$, or multiple dots for time derivatives $\ddot{y}(t)$.

Example 1.13 The chain rule is useful for finding the derivative of a *composite function* where the variable of one equation depends on another variable. Let the two functions be

$$f(y) = 4y^2 - y + 3, \quad y = g(x) = x^2 + 1. \tag{1.55}$$

The derivatives are

$$\frac{d}{dy}f(y) = 8y - 1 = 8g(x) - 1, \quad \frac{d}{dx}g(x) = 2x, \tag{1.56}$$

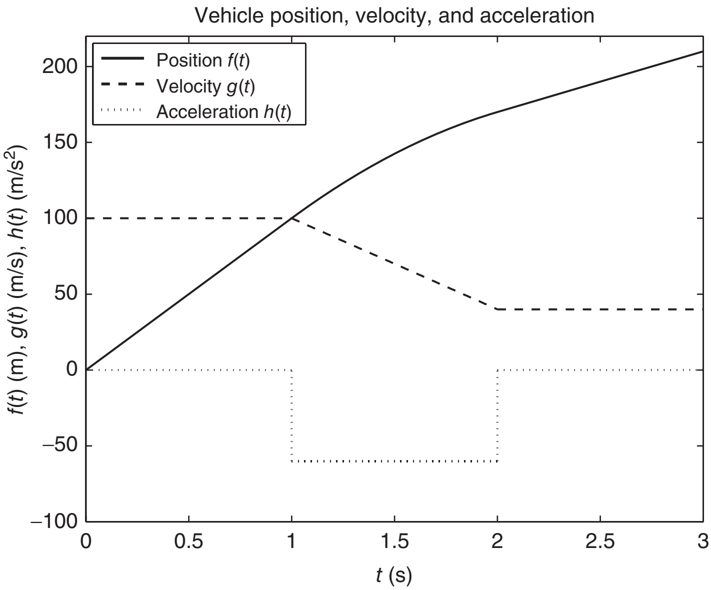


Figure 1.17 Vehicle position, velocity, and acceleration waveforms used in Example 1.14.

and the chain rule yields

$$\frac{d}{dx}f(g(x)) = [8g(x) - 1]2x = 16x^3 + 14x. \tag{1.57}$$

This is verified by substituting $g(x)$ into $f(y)$ and differentiating once with respect to x:

$$f(g(x)) = 4x^4 + 7x^2 + 6 \;\Rightarrow\; \frac{d}{dx}f(g(x)) = 16x^3 + 14x. \tag{1.58}$$

Substituting one equation into the other is usually a tedious process, which is a step the chain rule eliminates. The product and quotient formulas also simplify finding derivatives because it is not necessary to multiply or divide the functions, respectively, before computing derivatives.

Example 1.14 Suppose the position of a vehicle in meters (m) along one Cartesian coordinate over time t is described by the following piecewise linear function:

$$f(t) = \begin{cases} 100t, & 0 \leq t \leq 1 \\ -30t^2 + 160t - 30, & 1 < t \leq 2 \\ 40t + 90, & 2 < t \leq 3, \end{cases} \tag{1.59}$$

where the units of t are seconds (s). The distance traveled versus time is illustrated in Figure 1.17 (the solid line). The velocity is the time derivative of this function, denoted by $g(t) = \dot{f}(t)$ with units m/s:

$$g(t) = \begin{cases} 100, & 0 \leq t \leq 1 \\ -60t + 160, & 1 < t \leq 2 \\ 40, & 2 < t \leq 3. \end{cases} \tag{1.60}$$

In Figure 1.17, we see that the velocity is initially 100 m/s and then it decreases linearly to 40 m/s (the dashed line). The acceleration is the time derivative of the velocity $h(t) = \dot{g}(t) = \ddot{f}(t)$, which has units m/s^2:

$$h(t) = \begin{cases} 0, & 0 \leq t < 1 \\ -60, & 1 \leq t < 2 \\ 0, & 2 \leq t \leq 3. \end{cases} \tag{1.61}$$

The vehicle has nonzero acceleration only when its velocity is decreasing from 100 to 40 m/s (it is actually a *deceleration* because $h(t)$ is negative). Unlike the first two functions, $h(t)$ is not continuous.

Example 1.15 The derivative can be used to find *saddle points*, the *minimum*, or the *maximum* of a function (if they exist). Consider the cubic function $f(x) = (x - 1)^3$ plotted in Figure 1.18. The first derivative is $f'(x) = 3(x - 1)^2$ and the second derivative is $f''(x) = 6(x - 1)$, both of which are also plotted in Figure 1.18. Observe that $f(x)$ has a saddle point at $x = 1$: the derivative $f'(x)$ (the dashed line) is 0 there, but $x = 1$ is neither a maximum nor a minimum of $f(x)$. The second derivative $f''(x)$ is also 0 at $x = 1$, which means that the quadratic function $f'(x)$ has a minimum there. It is a minimum (and not a maximum) because the second derivative of $f'(x)$, given by $f^{(3)}(x) = 6$, is positive.

Definition: Indefinite Integral The *indefinite integral* of $f(x)$ is another function $g(x)$ such that $dg(x)/dx = f(x)$.

The indefinite integral $g(x)$ is also called the *antiderivative*, and so, integration is the inverse operation of differentiation. It is represented by

$$g(x) = \int f(x)dx = F(x) + c, \tag{1.62}$$

where c is a constant independent of x, and $F(x)$ is the antiderivative when $c = 0$. Thus, the antiderivative is not unique; instead, we say it is unique *up to a constant*. The value of c is determined by *boundary conditions*.

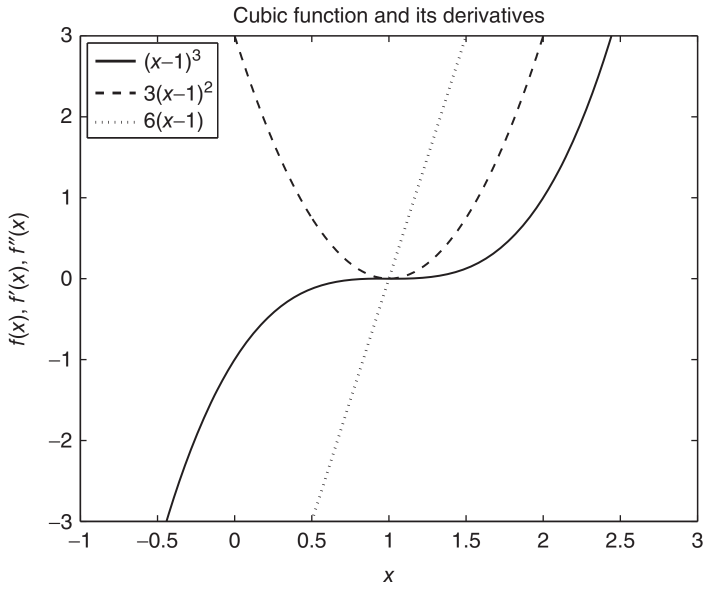


Figure 1.18 Cubic function $f(x) = (x - 1)^3$ in Example 1.15 and its derivatives.

Example 1.16 For the scenario in Example 1.14, the velocity is the indefinite integral of acceleration:

$$g(t) = \int h(t)dt = \begin{cases} c_1, & 0 \leq t \leq 1 \\ -60t + c_2, & 1 < t \leq 2 \\ c_3, & 2 < t \leq 3, \end{cases} \tag{1.63}$$

where $\{c_n\}$ are constants that are determined by the boundary conditions for the subintervals $[0, 1]$, $(1, 2]$, and $(2, 3]$. In order to continue, we need the initial velocity, which in this case is 100 m/s, yielding $c_1 = 100$. Similarly, the final velocity gives $c_3 = 40$ m/s. The middle coefficient is derived by assuming that the velocity does not change instantaneously. Thus, at $t = 1$:

$$-60 \times 1 + c_2 = 100 \implies c_2 = 160 \text{ m/s}, \tag{1.64}$$

which can also be derived at $t = 2$:

$$-60 \times 2 + c_2 = 40 \implies c_2 = 160 \text{ m/s}. \tag{1.65}$$

Combining these terms gives the expression for $g(t)$ in (1.60).

Definition: Definite Integral The *definite integral* of function $f(x)$ is a real number derived from the indefinite integral with specific limits of integration:

$$g(b) - g(a) = \int_a^b f(x)dx. \tag{1.66}$$

It gives the *area* under $f(x)$ on the interval $[a, b]$.

For a definite integral, the constant c appearing in (1.62) is of no concern because it cancels when evaluated at the limits:

$$g(b) - g(a) = [F(b) + c] - [F(a) + c] = F(b) - F(a). \tag{1.67}$$

Note that seemingly simple integrals require special attention. For example, it is not clear how to evaluate

$$\int_a^b f(x)dx = \int_{-1}^1 (1/x)dx, \tag{1.68}$$

because $f(x) = 1/x$ has a singularity at $x = 0$. Such functions are sometimes called *pseudofunctions* and the integral is improper.

Definition: Improper Integral The following integral is *improper* if $f(x)$ is infinite for some x in $[a, b]$:

$$\int_a^b f(x)dx, \tag{1.69}$$

or if $a = -\infty$, $b = \infty$, or both.

In both situations, we must carefully evaluate the integral as demonstrated in the next example.

Example 1.17 The following integral is improper because the function is unbounded at $x = 1$:

$$\int_1^2 \frac{dx}{x-1}. \tag{1.70}$$

This expression is examined by changing the lower limit to ϵ and letting $\epsilon \to 1$:

$$\lim_{\epsilon \to 1} \int_\epsilon^2 \frac{dx}{x-1} = \lim_{\epsilon \to 1} [\ln(|2-1|) - \ln(|\epsilon - 1|)] = \infty. \tag{1.71}$$

Similarly, for

$$\int_2^\infty \frac{dx}{x-1}, \tag{1.72}$$

we have

$$\lim_{\epsilon \to \infty} \int_2^\epsilon \frac{dx}{x-1} = \lim_{\epsilon \to \infty} [\ln(|\epsilon - 1|) - \ln(|2-1|)] = \infty. \tag{1.73}$$

Both of these integrals are *divergent*. Suppose the denominator in (1.70) is squared:

$$\int_1^2 \frac{dx}{(x-1)^2}. \tag{1.74}$$

Then

$$\lim_{\epsilon\to 1}\int_\epsilon^2 \frac{dx}{(x-1)^2} = \lim_{\epsilon\to 1} \left.\frac{-1}{x-1}\right|_\epsilon^2 = \infty, \tag{1.75}$$

which is also divergent. However, the integral

$$\int_2^\infty \frac{dx}{(x-1)^2} \tag{1.76}$$

is *convergent*:

$$\lim_{\epsilon\to \infty}\int_2^\epsilon \frac{dx}{(x-1)^2} = \lim_{\epsilon\to \infty} \left.\frac{-1}{x-1}\right|_2^\epsilon = 1. \tag{1.77}$$

Although $f(x) = 1/\sqrt{x-1}$ is undefined at $x = 1$, the following integral is convergent:

$$\int_1^2 \frac{dx}{\sqrt{x-1}} = 2\sqrt{x-1}\Big|_1^2 = 2. \tag{1.78}$$

The three functions in this example all have a singularity at $x = 1$ as shown in Figure 1.19. Since $1/\sqrt{x-1}$ is *imaginary* for $x < 1$, it is plotted only for $x > 1$. Imaginary and complex numbers are covered in Chapter 4.

The definite integral in (1.66) is known as a *Riemann integral* in order to distinguish it from other types of integrals (such as the Lebesgue integral, which is beyond the scope of this book). It can be defined in terms of the following *Riemann sum*:

$$\int_a^b f(x)dx = \lim_{\Delta x_n\to 0} \sum_{n=0}^{N-1} f(x_n)\Delta x_n, \tag{1.79}$$

such that $N \to \infty$ with $\Delta x_n \triangleq x_{n+1} - x_n$, $x_0 = a$, and $x_N = b$. In a Riemann sum, the interval $[a, b]$ on the x-axis is divided into nonoverlapping subintervals, which together cover the entire interval. This collection of subintervals is called a *partition* of $[a, b]$. Observe that we have used the smaller value x_n of Δ_n for the argument of $f(x_n)$, in which case the sum is known as a *lower Riemann sum*. If instead x_{n+1} is used, then it is called an *upper Riemann sum*. In the limit as $\Delta x_n \to 0$, both sums converge to the same quantity for a continuous function, giving the definite integral of $f(x)$ on $[a, b]$. Examples of the lower and upper Riemann sums are indicated by the shaded regions in Figure 1.20.

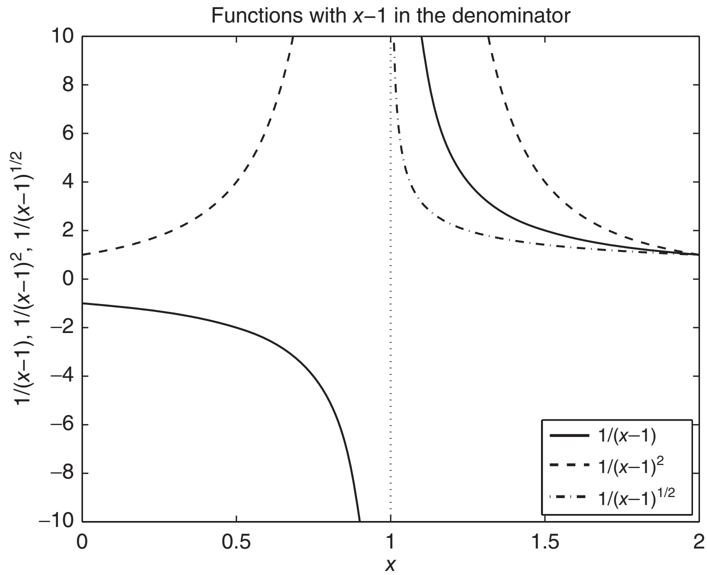


Figure 1.19 Three functions in Example 1.17 with singularities at $x = 1$.

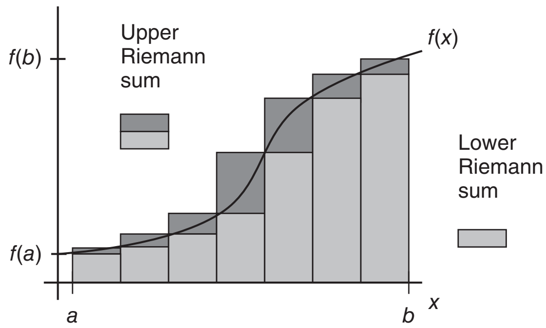


Figure 1.20 Lower and upper Riemann sums approximating the integral of $f(x)$ on $[a, b]$.

Although it is not necessary for the subintervals to have the same width, it is usually convenient to do so with $\Delta x_n = (b - a)/N \triangleq \Delta$ for all n such that $x_n = a + n\Delta$ and (1.79) becomes

$$\int_a^b f(x)dx = \frac{b-a}{N} \lim_{N\to\infty} \sum_{n=0}^{N-1} f(a + n\Delta). \tag{1.80}$$

Example 1.18 Consider again the functions in Example 1.8. The area under $f(x) = x^2$ on $[0, 2]$ is

$$\int_0^2 x^2 dx = \frac{2}{N} \lim_{N\to\infty} \sum_{n=0}^{N-1} (n\Delta)^2 = \frac{8}{N^3} \sum_{n=0}^{N-1} n^2, \tag{1.81}$$

where we have assumed equal-length subintervals and substituted $\Delta = 2/N$. A closed-form expression in Appendix C for the last sum in (1.81) yields

$$\int_0^2 x^2 dx = \lim_{N\to\infty} (8/N^3)(1/6)[(N-1)N][2(N-1)+1]$$
$$= (8/6) \lim_{N\to\infty} \frac{2N^3 - 3N^2 + N}{N^3} = 8/3. \tag{1.82}$$

Since the indefinite integral of $f(x) = x^2$ is $g(x) = x^3/3 + c$, we confirm that the area of $f(x)$ on $[0,2]$ is $8/3$. For $f(x) = 2x+1$ on $[-1,2]$:

$$\int_{-1}^2 (2x+1)dx = \lim_{N\to\infty} \frac{3}{N} \sum_{n=0}^{N-1} [2(-1+n\Delta)+1]$$
$$= \lim_{N\to\infty} \left[(18/N^2) \sum_{n=0}^{N-1} n - (3/N) \sum_{n=0}^{N-1} 1 \right], \tag{1.83}$$

where $\Delta = 3/N$. The last sum is N, and using another closed-form expression from Appendix C for the first sum in (1.83), the area is

$$\int_{-1}^2 (2x+1)dx = \lim_{N\to\infty} (18/N^2)[(N-1)N/2] - 3 = 6. \tag{1.84}$$

The indefinite integral of $f(x) = 2x+1$ is $g(x) = x^2 + x + c$, and from this we verify that the definite integral on $[-1,2]$ is 6.

It is not necessary that the sum in (1.79) be used to derive integrals because many results have already been established for a wide range of functions. Appendix C includes some indefinite integrals as well as a few definite integrals. The following important properties of integration are provided without proof.

- *Integration by parts*:

$$\int w(x)\frac{dv(x)}{dx}dx = w(x)v(x) - \int \frac{dw(x)}{dx}v(x)dx. \tag{1.85}$$

- *Leibniz's integral rule*:

$$\frac{d}{dx}\int_{a(x)}^{b(x)} f(v)dv = f(b(x))\frac{d}{dx}b(x) - f(a(x))\frac{d}{dx}a(x). \tag{1.86}$$

The following expressions are special cases that are used often in engineering problems:

$$\frac{d}{dx}\int_a^x f(v)dv = f(x), \quad \frac{d}{dx}\int_x^b f(v)dv = -f(x). \tag{1.87}$$

Example 1.19 Consider the indefinite integral

$$f(x) = \int x \exp(\alpha x) dx, \tag{1.88}$$

where α is a constant. In order to use integration by parts, we equate the following: $w(x) = x$ and $dv(x)/dx = \exp(\alpha x)$ which yield $dw(x)/dx = 1$ and $v(x) = (1/\alpha)\exp(\alpha x)$. The expression in (1.85) gives

$$f(x) = (x/\alpha)\exp(\alpha x) - (1/\alpha)\int \exp(\alpha x) dx, \tag{1.89}$$

whose integral is straightforward to evaluate:

$$\begin{aligned} f(x) &= (x/\alpha)\exp(\alpha x) - (1/\alpha^2)\exp(\alpha x) + c \\ &= [(\alpha x - 1)/\alpha^2]\exp(\alpha x) + c, \end{aligned} \tag{1.90}$$

where c is the constant of integration. For an example of Leibniz's integral rule, consider

$$g(x) = \int_x^{x^2} \exp(\alpha u) du, \tag{1.91}$$

which has derivative

$$\begin{aligned} \frac{d}{dx} g(x) &= \exp(\alpha x^2)\frac{d}{dx}x^2 - \exp(\alpha x)\frac{d}{dx}x \\ &= 2x\exp(\alpha x^2) - \exp(\alpha x). \end{aligned} \tag{1.92}$$

This is verified by performing the integration:

$$g(x) = (1/\alpha)[\exp(\alpha x^2) - \exp(\alpha x)], \tag{1.93}$$

and then differentiating (using the chain rule):

$$\frac{d}{dx} g(x) = (1/\alpha)[2\alpha x \exp(\alpha x^2) - \alpha \exp(\alpha x)], \tag{1.94}$$

which simplifies to (1.92). Leibniz's integral rule allows us to find the derivative in (1.92) without first computing the integral in (1.91).

Derivatives and integrals appear in the linear ODEs and *integro-differential* equations discussed in Chapters 6 and 7.

1.6 SINE, COSINE, AND π

Next, we discuss some properties of sinusoidal functions and indicate how they arise in practice. Consider the circle shown in Figure 1.21, which has unit radius $r = 1$ and is called the *unit circle*, and so, its circumference is 2π. The famous constant $\pi = 3.141592653589\ldots$ is the ratio of the circumference of any circle and its diameter. Since it cannot be expressed as the ratio of two integers, π is an *irrational number* (of course, this also means that if the circumference of a circle is an integer, then its diameter is not). The circumference in the figure can be divided into 360 equal lengths (arcs), and each "pie slice" projected back to the origin is defined to have an angle of $1°$. The distance along the unit circle yields the corresponding angle in radians. The example in Figure 1.21 illustrates that an angle of $\pi/2$ in the first quadrant relative to the positive horizontal axis is actually one-quarter distance along the circle circumference of 2π: $2\pi/4 = \pi/2$.

It is well known from trigonometry that sine of an angle formed by a right triangle is defined as the ratio of the lengths of the distant side y and the hypotenuse r: $\sin(\theta) \triangleq y/r$. Similarly, cosine of θ is defined as the ratio of the lengths of the adjacent side of a right triangle and its hypotenuse: $\cos(\theta) \triangleq x/r$. Since $x^2 + y^2 = r^2$, we immediately have that for any angle θ:

$$\sin^2(\theta) + \cos^2(\theta) = 1, \tag{1.95}$$

where either $\theta \in [0, 360°]$ or $\theta \in [0, 2\pi]$ radians. It is also clear from Figure 1.21 and the connection between sine and cosine that

$$\sin(\theta \pm \pi/2) = \pm\cos(\theta), \tag{1.96}$$

$$\cos(\theta \pm \pi/2) = \mp\sin(\theta). \tag{1.97}$$

Plotting sine and cosine as functions of θ, we find that sine lags cosine by $\pi/2$ radians (90°).

Suppose now that the angle is written as $\theta(t) = \omega_o t$ so that it varies with time, where ω_o is *angular frequency* with units of rad/s. Thus, with a fixed ω_o, any point on the radial line from the origin to the circle with radius r has the same *constant*

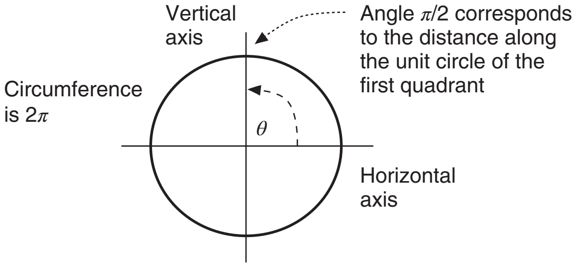


Figure 1.21 Unit circle with radius $r = 1$ and circumference 2π.

angular velocity as it sweeps counterclockwise with increasing t. This result follows because the derivative is a constant $d\theta(t)/dt = \omega_o$. From Figure 1.22, observe how the *functions* $\sin(\omega_o t)$ and $\cos(\omega_o t)$ are generated. As time increases, $\cos(\omega_o t)$ is the *projection* of the end of the radial line onto the horizontal axis; the cosine function is the length of this projection as it varies over $[-1, 1]$ (for $r = 1$). Likewise, $\sin(\omega_o t)$ is the projection of the end of the radial line onto the vertical axis. A projection is defined to be negative for cosine when it is located to the left of the origin on the horizontal axis, and it is negative for sine when it is below the origin on the vertical axis.

Summarizing, the time-varying functions $\sin(\omega_o t)$ and $\cos(\omega_o t)$ follow from the usual definitions of the sine and cosine of an angle, except that the angle varies as $\theta(t) = \omega_o t$. By convention, the angle is defined with respect to the positive horizontal axis as depicted in Figure 1.22 for four different time instants (snapshots). These plots illustrate why the sine and cosine functions are 90° out of phase with respect to each other: as $\sin(\omega_o t)$ increases, $\cos(\omega_o t)$ decreases and vice versa. They are *orthogonal* functions:

$$\int_a^b \sin(\omega_o t)\cos(\omega_o t)dt = 0, \tag{1.98}$$

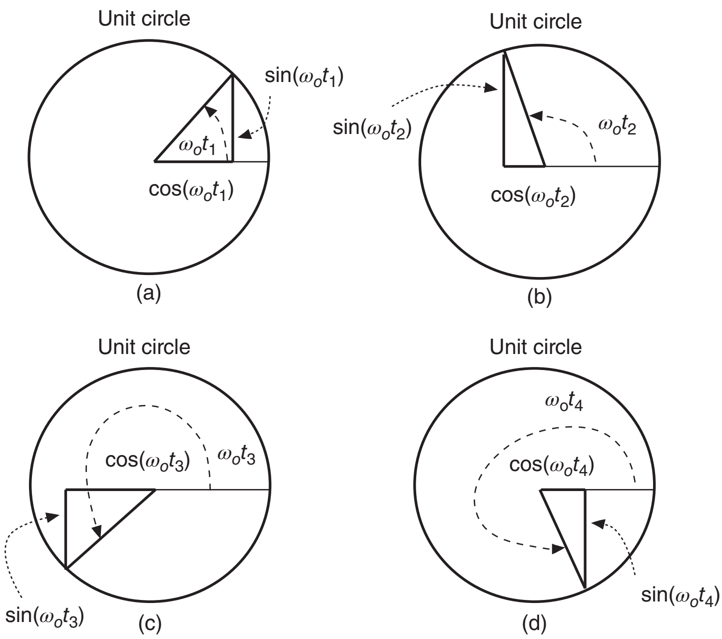


Figure 1.22 Four snapshots of sine and cosine for time-varying angle $\theta(t) = \omega_o t$ with constant angular velocity and $t_1 < t_2 < t_3 < t_4$.

when $(b-a)\omega_o$ is an integer multiple of π. This result is verified by using a trigonometric identity from Appendix C:

$$\begin{aligned}\int_a^b \sin(\omega_o t)\cos(\omega_o t)dt &= (1/2)\int_a^b [\sin(2\omega_o t) + \sin(0)]dt \\ &= (-1/4\omega_o)\cos(2\omega_o t)|_a^b \\ &= [\cos(2\omega_o a) - \cos(2\omega_o b)]/4\omega_o, \end{aligned} \tag{1.99}$$

which is 0 when $\cos(2\omega_o b) = \cos(2\omega_o a)$. Since cosine is periodic with period 2π, we require $2\omega_o b = 2\omega_o a + n2\pi$ for $n \in \mathcal{Z}$, which means $(b-a)\omega_o = n\pi$. Figure 1.23 shows a plot of (1.99) for $a = 0$ and $\omega_o = 1$ rad/s as b is varied from 0 to 5π. The integral is 0 for $b = \{0, \pi, 2\pi, 3\pi, 4\pi, 5\pi\}$, and the maximum area is $1/2$ for this value of ω_o. The orthogonality property is also evident from a geometric viewpoint because the vertical and horizontal dashed lines in Figure 1.22 are orthogonal: they form the previously mentioned right triangle. The fact that the radial line sweeps along a *circle* gives rise to the specific *smooth* shapes of the sine and cosine waveforms, derived as projections on the two axes.

Figure 1.24(a) shows the sine waveform in Figure 1.22 with $\omega_o = 1$ rad/s. The function approaches its maximum with a decreasing derivative, which is the cosine

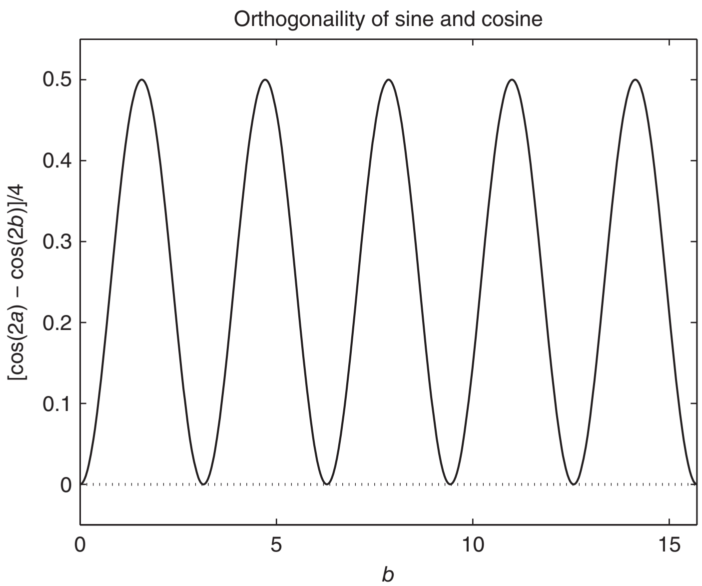


Figure 1.23 Orthogonality of sine and cosine for $a = 0$ and $\omega_o = 1$ rad/s.

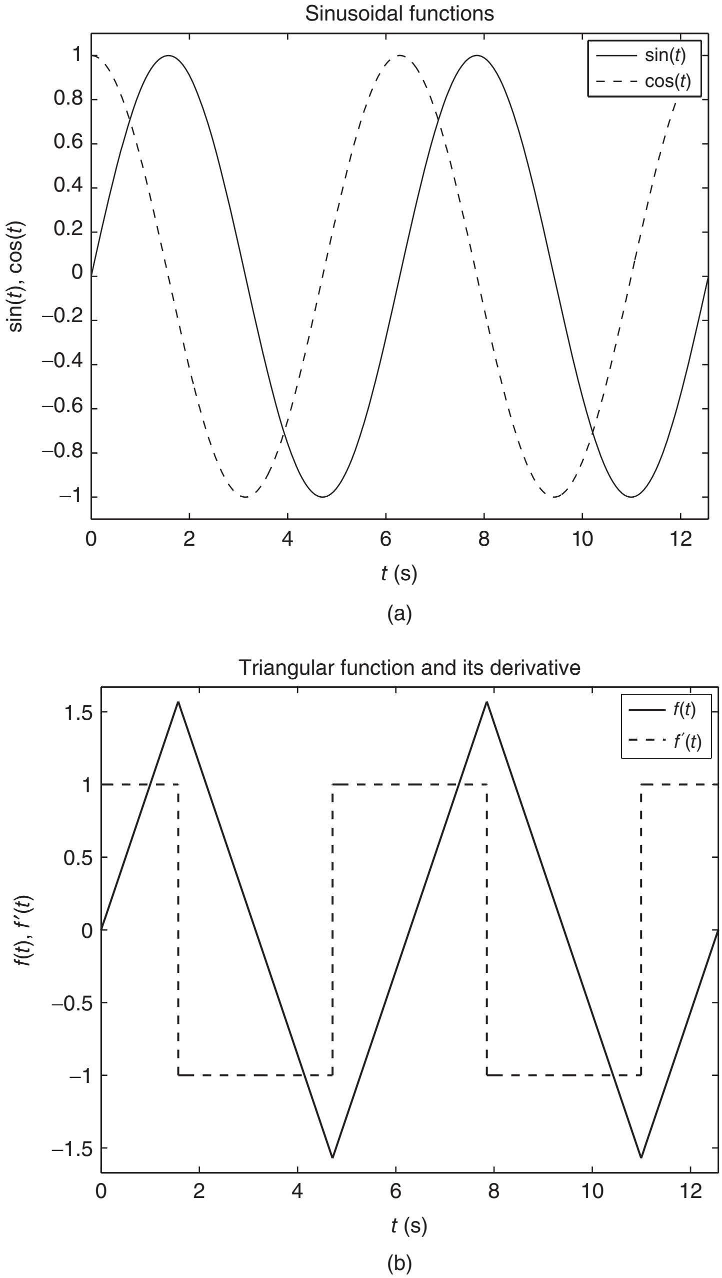


Figure 1.24 Periodic waveforms. (a) Sine waveform sin(t) and its derivative cos(t). (b) Triangular waveform and its rectangular derivative.

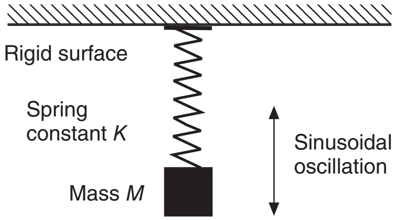


Figure 1.25 Mass on a spring influenced by gravity.

waveform also shown in the figure. (The orthogonality of these two waveforms is also apparent from this figure.) This smooth behavior of its derivative is unlike that of the triangular waveform in Figure 1.24(b) whose derivative is a constant until the function reaches its maximum, at which point the derivative abruptly changes sign. It turns out that many physical phenomena are modeled accurately using sinusoidal functions. Apparently, many physical systems behave in a sinusoidal manner because the underlying physics yield gradual variations rather than abrupt changes. This also means that the physical mechanisms of many systems have the dynamic of constant angular velocity along a circle on the plane as in Figure 1.22.

An example of a mechanical process is an object (mass) attached to a spring as depicted in Figure 1.25. If the object is extended downward and released, its up-and-down trajectory is sinusoidal. As the spring is stretched, its linear velocity gradually decreases and it becomes exactly 0 at its maximum distance, just like a sinusoidal waveform. This behavior is due to the physical properties of the spring and the force of gravity. The object does not have constant *linear* velocity, and it does not abruptly change direction at its minimum and maximum distance from the rigid surface. The amplitude and frequency of the waveform depend on the mass M of the object, the *spring constant* K, and the initial position of the object, which are discussed further in Chapter 2.

We demonstrate in Chapter 4 that the sine and cosine axes as depicted in Figure 1.21 can be represented on the *complex plane*, where the horizontal axis (associated with cosine) is the *real axis* and the vertical axis (associated with sine) is the *imaginary axis*. It turns out that both sine and cosine can be written together using complex notation as follows:

$$\exp(j\omega_o t) = \cos(\omega_o t) + j\sin(\omega_o t), \tag{1.100}$$

where $j \triangleq \sqrt{-1}$ and $\exp(1) = e$ is Napier's constant. This two-dimensional formulation called *Euler's formula* is widely used in engineering to represent signals and waveforms, and $\exp(j\omega_o t)$ is an *eigenfunction* of a linear system as discussed in Chapter 7.

1.7 NAPIER'S CONSTANT *e* AND LOGARITHMS

Napier's constant e is another important irrational number used in mathematics and engineering. It is motivated by the following *compound interest* problem. Suppose one has an initial monetary amount x_o called the *principal*, which accumulates interest at an annual percentage rate of $100r\%$. At the end of 1 year when a single interest payment is made, the new principal is $x_o(1 + r)$, where for now we assume $0 < r \leq 1$. Suppose instead that an interest payment is made after 6 months, and the total amount available then accumulates interest until the end of the year. The amount after one-half year is $x_o(1 + r/2)$. Since this is the principal for the second half of the year, we have a total amount of $x_o(1 + r/2)(1 + r/2) = x_o(1 + r/2)^2$ at the end of the year. Similarly, by dividing the year into thirds, the amount at the end of the year is $x_o(1 + r/3)^3$, and in general, for n interest payments, the principal is $x_o(1 + r/n)^n$ at the end of 1 year.

It can be shown that for $x_o = 1$ and $r = 1$ (corresponding to a 100% interest rate), the limit is Napier's constant:

$$\lim_{n\to\infty} (1 + 1/n)^n = e = 2.718281828459\ldots \tag{1.101}$$

This convergence to e is demonstrated in Figure 1.26. It is an interesting result that the total monetary amount after 1 year of essentially *continuous* interest payments

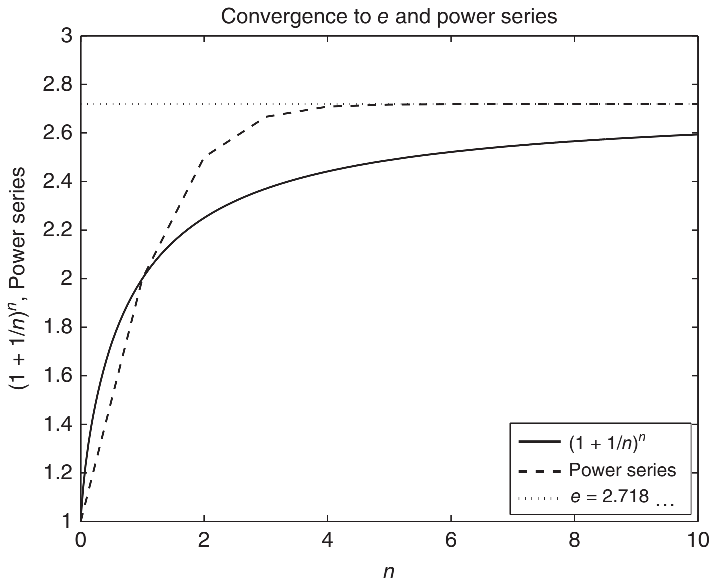


Figure 1.26 Convergence of $(1 + 1/n)^n$ to e and its power series approximation, where n is the upper limit of the sum in (1.104). (The individual points at integer n for the power series have been connected by lines for ease of viewing.)

(because $n \to \infty$) is finite and given exactly by e. For general x_o and r, the limit is

$$\lim_{n\to\infty} x_o(1+r/n)^n = x_o e^r, \tag{1.102}$$

such that $r > 0$ results in a gain on the original principal x_o, and $r < 0$ yields a loss. These correspond to *exponential growth* and *exponential decay*, respectively.

The constant e has the following alternative representations.

- *Limits*:
$$e = \lim_{n\to 0}(1+n)^{1/n}, \quad e^r = \lim_{n\to 0}(1+n/r)^{r/n}. \tag{1.103}$$

- *Power series*:
$$e = \sum_{m=0}^{\infty} \frac{1}{m!}. \tag{1.104}$$

- *Hyperbolic functions*:
$$e = \sinh(1) + \cosh(1). \tag{1.105}$$

Convergence of the power series sum in (1.104) with upper limit n instead of infinity is shown in Figure 1.26. As n is varied over the 11 integers $\{0, \ldots, 10\}$, we find that the sum quickly approaches e; the first six values are 1, 2, 2.5, 2.6667, 2.7083, and 2.7167.

The exponential function based on Napier's constant is defined next, which is discussed further in Chapter 5.

Definition: Exponential Function The *exponential function* is

$$\exp(x) \triangleq e^x. \tag{1.106}$$

It has domain $\mathcal{R}$ and range $\mathcal{R}^+$.

The exponential function has the following properties.

- *Product*:
$$\exp(x)\exp(y) = \exp(xy). \tag{1.107}$$

- *Ratio*:
$$\frac{\exp(x)}{\exp(y)} = \exp(x/y). \tag{1.108}$$

- *Derivative*:
$$\frac{d}{dx}\exp(x) = \exp(x). \tag{1.109}$$

- *Integrals*:
$$\int \exp(x)dx = \exp(x) + c, \quad \int_{-\infty}^{x} \exp(v)dv = \exp(x), \quad \int_0^x \exp(v)dv = \exp(x) - 1. \tag{1.110}$$

- *Power series*:

$$\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}. \tag{1.111}$$

- *Hyperbolic functions*:

$$\exp(x) = \cosh(x) + \sinh(x), \quad \exp(-x) = \cosh(x) - \sinh(x). \tag{1.112}$$

The last property gives $\cosh(x) = (1/2)[\exp(x) + \exp(-x)]$ and $\sinh(x) = (1/2)[\exp(x) - \exp(-x)]$, which is similar to Euler's formula for complex numbers discussed in Chapter 4. The exponential functions in (1.112) and their hyperbolic components are plotted in Figure 1.27.

The exponential function arises *naturally* in many engineering problems because of its unique derivative and integral properties. This is demonstrated by the following example in probability.

Example 1.20 The exponential probability density function (pdf) is

$$f_X(x) = \begin{cases} \alpha \exp(-\alpha x), & x \geq 0 \\ 0, & x < 0, \end{cases} \tag{1.113}$$

where the uppercase notation X denotes a *random variable* with outcomes x, and the parameter $\alpha > 0$ determines the mean and variance of X. This pdf has domain $\mathcal{R}$,

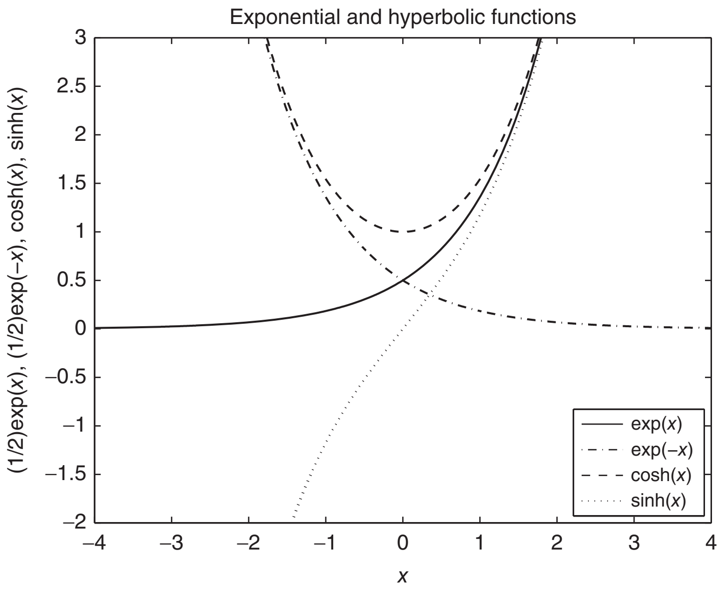


Figure 1.27 Exponential and hyperbolic functions.

support $\mathcal{R}^+$, and range $\mathcal{R}^+$. A valid pdf satisfies the following two conditions:

$$f_X(x) \geq 0, \quad \int_{-\infty}^{\infty} f_X(x)dx = 1. \tag{1.114}$$

These are obviously true for the exponential pdf:

$$\alpha \exp(-\alpha x) \geq 0, \quad \int_0^{\infty} \alpha \exp(-\alpha x)dx = -\exp(-\alpha x)|_0^{\infty} = 1. \tag{1.115}$$

Suppose instead that we are interested in another decaying function such as $f_X(x) = ba^{-x} \geq 0$ for $a, b \geq 0$ and $x \in \mathcal{R}^+$. The integral of this function is

$$b\int_0^{\infty} a^{-x}dx = -\frac{ba^{-x}}{\ln(a)}\bigg|_0^{\infty} = \frac{b}{\ln(a)}, \tag{1.116}$$

where $\ln(\cdot)$ is the natural logarithm defined next. In order for the integral to be 1, it is necessary that $b = \ln(a)$, and so, we must have $a > 1$, yielding the following valid pdf:

$$f_X(x) = \ln(a)a^{-x}, \quad x \in \mathcal{R}^+, \tag{1.117}$$

which has a maximum value of $\ln(a)$ at $x = 0$. Thus, other exponential-like decaying functions are possible, but they require a leading coefficient, and so, they are not the "natural" choice as is $a = e$ with $\ln(a) = 1$. The derivative and integral properties of $\exp(x)$ eliminate such multiplicative scaling of the function. The same reasoning can be used to justify e in the *Gaussian* pdf:

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma}}\exp(-(x-\mu)^2/2\sigma^2), \tag{1.118}$$

where μ and σ are its mean and standard deviation, respectively. Likewise, the pdf of the *Laplace* random variable is

$$f_X(x) = \frac{1}{2\alpha}\exp(-|x|/\alpha), \tag{1.119}$$

with parameter $\alpha > 0$, which determines the variance $2\alpha^2$. The support for these last two pdfs is the entire real line $\mathcal{R}$.

Finally, we consider logarithms and their connection to e.

Definition: Logarithm The *logarithm* of x is the exponent y with base b such that $b^y = x$. It is written as $\log_b(x) = y$ with domain $\mathcal{R}^+$ and range $\mathcal{R}$.

Perhaps the most familiar base is $b = 10$, which yields *common logarithms*. *Binary logarithms* with $b = 2$ are used in the analysis of digital systems. Note that

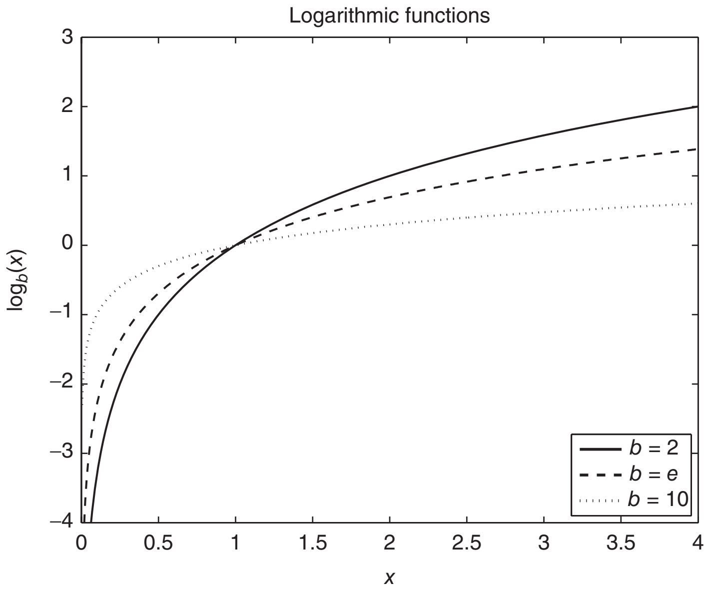


Figure 1.28 Logarithmic functions with different base b.

$\log_b(1) = 0$ for any b as depicted in Figure 1.28 where the base is varied from 2 to 10. The conversion formula of a logarithm from base b_1 to base b_2 is

$$\log_{b_2}(x) = \log_{b_1}(x)/\log_{b_1}(b_2). \tag{1.120}$$

Example 1.21 For $b = 10$, the subscript is often omitted: $\log(x)$ (though in MATLAB log has base e and log10 has base 10). Examples include $\log(1000) = 3$ and $\log(0.1) = -1$. Integer powers of 2 are important numbers in digital systems because their logic is based on the binary number system, usually represented by $\{0, 1\}$. Thus, $b = 2$ such that $\log_2(8) = 3$, $\log_2(64) = 6$, $\log_2(1/2) = -1$, and so on.

The following logarithm appears frequently in engineering applications.

Definition: Natural Logarithm The *natural logarithm* is

$$\ln(x) \triangleq \log_e(x), \tag{1.121}$$

which has domain $\mathcal{R}^+$ and range $\mathcal{R}$. It is also defined by the definite integral:

$$\ln(x) \triangleq \int_1^x (1/v)dv. \tag{1.122}$$

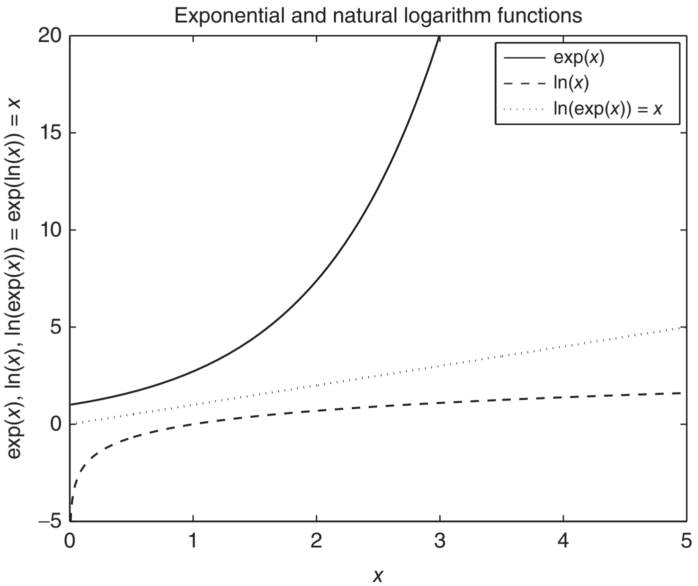


Figure 1.29 Exponential and natural logarithm functions.

This is not an improper integral of the pseudofunction $1/v$ because the limits of integration do not include the origin. From (1.121), we have

$$\ln(\exp(x)) = x, \quad \exp(\ln(x)) = x, \tag{1.123}$$

where it is assumed that $x > 0$ in the second equation. The exponential and natural logarithm functions are plotted in Figure 1.29, where the vertical axis has been limited to 20 because the exponential function increases rapidly (e.g., $\exp(5) \approx 148.41$). Observe the following properties: (i) $\ln(x)$ increases much more slowly than $\exp(x)$ and (ii) $\ln(x) \to -\infty$ as $x \to 0$. We have also included the straight-line plot for $\ln(\exp(x)) = \exp(\ln(x)) = x$, demonstrating that they are in fact inverse functions of each other.

Logarithms have the following properties.

- *Integrals*:

$$\int \log_b(x)dx = x[\log_b(x) - 1/\ln(b)] + c, \quad \int \ln(x)dx = x\ln(x) - x + c. \tag{1.124}$$

- *Sum*:

$$\log_b(x) + \log_b(y) = \log_b(xy). \tag{1.125}$$

- *Difference*:

$$\log_b(x) - \log_b(y) = \log_b(x/y). \tag{1.126}$$

- *Exponent*:

$$\log_b(x^n) = n\log_b(x). \tag{1.127}$$

- *Derivatives*:

$$\frac{d}{dx}\log_b(x) = \frac{1}{x\ln(b)}, \quad \frac{d}{dx}\ln(x) = 1/x. \tag{1.128}$$

- *Limit*:

$$\ln(x) = \lim_{n\to\infty} n(x^{1/n} - 1). \tag{1.129}$$

- *Power series*:

$$\ln(x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}(x-1)^n, \quad \ln(x+1) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}x^n. \tag{1.130}$$

Example 1.22 From the identity $\alpha = \exp(\ln(\alpha))$, we can write

$$\alpha^v = \exp(v\ln(\alpha)). \tag{1.131}$$

Suppose v is a function of x such that

$$\alpha^{v(x)} = \exp(v(x)\ln(\alpha)). \tag{1.132}$$

The right-hand side and the chain rule can be used to find the derivative of functions of this form with x in the exponent:

$$\begin{aligned}\frac{d}{dx}\alpha^{v(x)} &= \frac{d}{dx}\exp(v(x)\ln(\alpha)) \\ &= \exp(v(x)\ln(\alpha))\ln(\alpha)\frac{d}{dx}v(x) \\ &= \ln(\alpha)\alpha^{v(x)}\frac{d}{dx}v(x),\end{aligned} \tag{1.133}$$

where (1.131) has been substituted in the final expression. This result is not the same as the more commonly used derivative

$$\frac{d}{dx}v^n(x) = nv^{n-1}(x)\frac{d}{dx}v(x), \tag{1.134}$$

where n in the exponent is a constant.

We conclude this section with proofs of the derivatives in (1.109) and (1.128) using the limit definition of the derivative in (1.44). For the natural logarithm:

$$\frac{d}{dx}\ln(x) = \lim_{\Delta x\to 0}\frac{\ln(x+\Delta x)-\ln(x)}{\Delta x} = \lim_{\Delta x\to 0}\frac{\ln((x+\Delta x)/x)}{\Delta x}. \tag{1.135}$$

Multiplying and dividing by x and then using the exponent property yield

$$\frac{d}{dx}\ln(x) = \lim_{\Delta x\to 0}\frac{(x/\Delta x)\ln((x+\Delta x)/x)}{x} = (1/x)\lim_{\Delta x\to 0}\ln((1+\Delta x/x)^{x/\Delta x}), \tag{1.136}$$

where $1/x$ has been brought outside the limit. The second form of the limit for e in (1.103) (with x in place of n) gives the final result:

$$\frac{d}{dx}\ln(x) = (1/x)\ln(e) = 1/x. \tag{1.137}$$

The derivative of $\exp(x)$ is obtained from the derivative of the natural logarithm and the chain rule:

$$\frac{d}{dx}\ln(\exp(x)) = \frac{1}{\exp(x)}\frac{d}{dx}\exp(x) \implies \frac{d}{dx}\exp(x) = \exp(x), \tag{1.138}$$

where we have used the fact that the left-hand side equals 1.

PROBLEMS

MATHEMATICAL MODELS

1.1 Sketch the following transfer characteristic:

$$y = \begin{cases} 0, & x < 0 \\ x^2, & 0 \le x < 3 \\ 2x+3, & 3 \le x < 5 \\ 0, & x \ge 5, \end{cases} \tag{1.139}$$

and sketch its output $y(t)$ when the input is the exponential function $x(t) = \exp(t)u(t)$.

1.2 Repeat the previous problem for

$$y = \begin{cases} 0, & x < 0 \\ 2x, & 0 \le x < 2 \\ 4, & 2 \le x < 4 \\ 4\exp(-2(x/2-4)), & x \ge 4, \end{cases} \tag{1.140}$$

and $x(t) = 2tu(t)$.

1.3 (a) For the linear system of equations in (1.5) and (1.6) with

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \tag{1.141}$$

solve for $\{y_1, y_2\}$ given $x_1 = x_2 = 1$. (b) Repeat part (a) for

$$\mathbf{A} = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}. \tag{1.142}$$

1.4 (a) For the nonlinear system in Example 1.4, iteratively solve for y_1 and y_2 by using the same parameter values and inputs, except let the exponential parameter be $\alpha = 5$. (b) Describe the behavior of the iterations for $\alpha = 2$.

1.5 A diode circuit using an exponential model with a series resistor is represented by the following system of equations:

$$a_{11} \exp(\alpha y) + a_{12} = x, \quad a_{21} + a_{22} y = x, \tag{1.143}$$

where the input x is a current (A) and the output y is a voltage (V). The coefficients $\{a_{mn}\}$ depend on the series resistor, the voltage source, and the diode parameters. Iteratively solve for $\{x, y\}$ using the following parameter values: $a_{11} = 10^{-15}$, $a_{12} = -a_{11}$, $a_{21} = 10^{-3}$, $a_{22} = -a_{21}$, and $\alpha = 40$. Let the initial value be $y = 0.6$ V.

1.6 A transistor circuit with a series resistor is represented by the following system of equations:

$$a_{11} y^2 = x, \quad a_{21} y + a_{22} = x, \tag{1.144}$$

where the input x is a current (A) and y is an output voltage (V). The coefficients $\{a_{mn}\}$ depend on the series resistor, the voltage source, and the transistor parameters. Let the parameter values be $a_{11} = 0.5 \times 10^{-3}$, $a_{21} = -10^{-3}$, and $a_{22} = 5 \times 10^{-3}$. Iteratively solve for $\{x, y\}$ assuming the initial value $y = 1$ V.

FUNCTIONS AND PROPERTIES

1.7 Specify the domain, range, and support for the following functions, assuming that x and y are real-valued. (a) $y = \sqrt{x^2 - 1}$. (b) $y = u(x - 2)$ (shifted unit step function). (c) $y = 1/|x - 1|$.

1.8 For real-valued $f(x) = 1/\sqrt{x + 2}$ and $g(x) = |x|$, give the domain, range, and support for the following functions. (a) $y_1 = f(x)/g(x)$. (b) $y_2 = g(x)/f(x)$. (c) $y_3 = f(x)g(x)$.

1.9 Specify the inverse image for each function. (a) $y_1 = x^2 - 5$. (b) $y_2 = |x - 1|/\sqrt{x}$. (c) $y_3 = \mathrm{sgn}(x)u(x + 2)$.

1.10 Find the range of values for x. (a) $|(x-1)/x| < 2$. (b) $|x+2| > 3x$. (c) $x^2 + |x| - 1 > 0$.

1.11 Determine the values of x for which the following functions are continuous. (a) $y_1 = 2x^3 - x^2 + x$. (b) $y_2 = x/(x^2-1)$. (c) $y_3 = \mathrm{sgn}(x-2)\mathrm{sgn}(x+2)$.

1.12 Let $\{x_1, x_2\}$ be the roots of the quadratic equation $ax^2 + bx + c = 0$. Prove that $x_1 + x_2 = -b/a$ and $x_1 x_2 = c/a$.

1.13 It can be shown that if $f(x)$ is a polynomial with real coefficients such that $f(x_1) < 0$ and $f(x_2) > 0$ for real $\{x_1, x_2\}$, then $f(x) = 0$ for some x between x_1 and x_2. Determine if this is the case for the following functions. (a) $f_1(x) = x^3 - 5x^2 + 2x + 8$ with $x_1 = 1$ and $x_2 = 3$. (b) $f_2(x) = x^3 + 2x^2 - 5x - 6$ with $x_1 = -2$ and $x_2 = 0$.

1.14 If function $f(x)$ has a derivative at x_o, then show using the following expression that it is also continuous at x_o:

$$\lim_{x \to x_o} |f(x) - f(x_o)| = \lim_{x \to x_o} \left| (x - x_o) \frac{f(x) - f(x_o)}{x - x_o} \right|. \tag{1.145}$$

DERIVATIVES AND INTEGRALS

1.15 Find the derivative of $y = x^3 + 2x$ using the limit definition.

1.16 Repeat the previous problem for $y = \sqrt{x+2}$.

1.17 Repeat Problem 1.15 for $y = x^2 + 1/x$ assuming $x \neq 0$.

1.18 (a) Use the product and chain rules to write an expression for

$$y = \frac{d}{dx} g^m(x) h^n(f(x)), \tag{1.146}$$

where $\{m, n\}$ are constants. (b) Find the derivative of $x^{2m} \exp(\alpha n \sin(x))$.

1.19 (a) Extend the chain rule to find an expression for

$$y = \frac{d^2}{dx^2} f(g(x)). \tag{1.147}$$

(b) Find the second derivative of $\exp(\alpha \sin(x))$.

1.20 Determine which of the following improper integrals converge.

$$\text{(a)} \int_0^\infty \exp(-\alpha x) \sin(x) dx. \qquad \text{(b)} \int_2^4 \frac{dx}{(x-2)^3}. \qquad \text{(c)} \int_0^\infty \frac{dx}{x^2+4}. \tag{1.148}$$

1.21 Consider the integral transform

$$X(\sigma) \triangleq \int_0^\infty x(t) \exp(-\sigma t) dt, \tag{1.149}$$

where σ is real-valued. Find $X(\sigma)$ for $x(t) = \exp(-t)u(t) + u(t)$ and specify the range of values for σ such that the integral is convergent.

1.22 The current $i(t)$ in a series circuit with resistance R and inductance L is modeled by the following first-order ODE:

$$L\frac{d}{dt}i(t) + Ri(t) = 0. \tag{1.150}$$

(a) Verify that the solution of this equation has the form $i(t) = i(0) \exp(-\alpha t)u(t)$ where $i(0)$ is the initial current, and specify the constant α. (b) Find the value of t such that the current is $1/2e$ its initial value $i(0)$.

1.23 Repeat the previous problem for the integral equation

$$(1/L) \int v(t)dt + v(t)/R = 0, \tag{1.151}$$

where $v(t)$ is a voltage with initial value $v(0)$.

SINE, COSINE, AND π

1.24 Prove the identity $\cos(\theta_2 - \theta_1) = \cos(\theta_1)\cos(\theta_2) + \sin(\theta_1)\sin(\theta_2)$ using an illustration on the unit circle.

1.25 (a) Repeat the previous problem for the double angle formula $\cos(2\theta) = \cos^2(\theta) - \sin^2(\theta)$, and (b) verify this result algebraically using the identity in that problem.

1.26 Find the minimum and maximum of $y = 3\cos(x) + 2\sin(x/2)$ on the interval $x \in [0, \pi]$.

1.27 Solve $\sin^2(\theta) + 2\sin(\theta) - 1 = 0$ for $\theta \in [-\pi/2, \pi/2]$.

1.28 For a general triangle whose sides have lengths x, y, and r, prove the law of cosines:

$$x^2 + y^2 - 2xy\ \cos(\theta) = r^2, \tag{1.152}$$

where θ is the angle formed by the x and y sides.

1.29 Show that

$$\frac{d}{dx}\sin^{-1}(x) = \frac{1}{\sqrt{1-x^2}}. \tag{1.153}$$

NAPIER'S CONSTANT e AND LOGARITHMS

1.30 (a) Use the natural logarithm to find an expression for the derivative of $y = f^{g(x)}(x)$ with respect to x. (b) Find the derivative of $y = x^{\ln(x)}$.

1.31 From the logarithm sum property $\ln(y_1 y_2) = \ln(y_1) + \ln(y_2)$, prove the product property $\exp(x_1 + x_2) = \exp(x_1)\exp(x_2)$ where $y = \exp(x)$.

1.32 Use the fact that $d\ln(y)/dx = (1/y)dy/dx$ to find the derivative of the following functions. (a) $y_1 = x^2\sqrt{x-1}$. (b) $y_2 = x^{2\cos(x)}$.

1.33 Prove that the minimum of $y = x^x$ is located at $x = e^{-1}$.

1.34 Solve $\ln(x-1) - 2\ln(x) = \ln(2)$ for x.

1.35 The time constant of the exponential function $y = 2\exp(-t/\tau)u(t)$ is $\tau > 0$. It is the value of t such that y is $1/e$ times its initial value of 2. (a) Give the number of time constants such that $y = 1/5$. (b) Repeat part (a) by approximating the exponential function using the first two terms of the Maclaurin series expansion in Appendix E.

COMPUTER PROBLEMS

1.36 For the model in Problem 1.1, use MATLAB to plot the input and output for $x(t) = 6\sin(2\pi t)$ on the interval $t \in [0, 1]$.

1.37 A transistor has the following input/output voltage transfer characteristic:

$$y = \begin{cases} A, & x < \alpha \\ A - \beta(x-\alpha)^2, & \alpha \le x < y + \alpha \\ \text{complicated}, & x > y + \alpha. \end{cases} \tag{1.154}$$

Find the upper bound for x in the second region of the transfer characteristic, and approximate the third region using the exponential function $y = y_b\exp(-\beta(x - x_b))u(x - x_b)$, where y_b is the output when the input is $x = x_b$. Repeat the previous problem using this model with input $x(t) = 2\sin(2\pi t) + 2$ on the interval $t \in [0, 1]$. Let the parameters be $A = 5$, $\alpha = 1$, and $\beta = 2$.

1.38 For the model in Problem 1.5, use MATLAB to plot the two functions and show the first few results of the iterative approach for finding the solution for $\{x, y\}$.