PART A

Ordinary Differential Equations (ODEs)

Many physical laws and relations can be expressed mathematically in the form of differential equations. Thus it is natural that this book opens with the study of differential equations and their solutions. Indeed, many engineering problems appear as differential equations.

The main objectives of Part A are twofold: the study of ordinary differential equations and their most important methods for solving them and the study of modeling.

Ordinary differential equations **(ODEs)** are differential equations that depend on a single variable. The more difficult study of partial differential equations (PDEs), that is, differential equations that depend on several variables, is covered in Part C.

Modeling is a crucial general process in engineering, physics, computer science, biology, medicine, environmental science, chemistry, economics, and other fields that translates a physical situation or some other observations into a "mathematical model." Numerous examples from engineering (e.g., mixing problem), physics (e.g., Newton's law of cooling), biology (e.g., Gompertz model), chemistry (e.g., radiocarbon dating), environmental science (e.g., population control), etc. shall be given, whereby this process is explained in detail, that is, how to set up the problems correctly in terms of differential equations.

For those interested in solving ODEs numerically on the computer, look at Secs. 21.1–21.3 of Chapter 21 of Part F, that is, ***numeric methods for ODEs***. These sections are kept independent by design of the other sections on numerics. ***This allows for the study of numerics for ODEs directly after Chap. 1 or 2.***

CHAPTER 1

First-Order ODEs

Chapter 1 begins the study of ordinary differential equations (ODEs) by deriving them from physical or other problems (***modeling***), solving them by standard mathematical methods, and interpreting solutions and their graphs in terms of a given problem. The simplest ODEs to be discussed are ODEs ***of the first order*** because they involve only the first derivative of the unknown function and no higher derivatives. These unknown functions will usually be denoted by $y(x)$ or $y(t)$ when the independent variable denotes time t. The chapter ends with a study of the existence and uniqueness of solutions of ODEs in Sec. 1.7.

Understanding the basics of ODEs requires solving problems by hand (paper and pencil, or typing on your computer, but first without the aid of a CAS). In doing so, you will gain an important conceptual understanding and feel for the basic terms, such as ODEs, direction field, and initial value problem. If you wish, you can use your **Computer Algebra System (CAS)** for checking solutions.

COMMENT. ***Numerics for first-order ODEs can be studied immediately after this chapter.*** See Secs. 21.1–21.2, which are independent of other sections on numerics.

Prerequisite: Integral calculus.
Sections that may be omitted in a shorter course: 1.6, 1.7.
References and Answers to Problems: App. 1 Part A, and App. 2.

1.1 Basic Concepts. Modeling

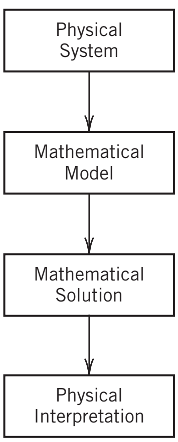


Fig. 1. Modeling, solving, interpreting

If we want to solve an engineering problem (usually of a physical nature), we first have to formulate the problem as a mathematical expression in terms of variables, functions, and equations. Such an expression is known as a mathematical **model** of the given problem. The process of setting up a model, solving it mathematically, and interpreting the result in physical or other terms is called *mathematical modeling* or, briefly, **modeling**.

Modeling needs experience, which we shall gain by discussing various examples and problems. (Your computer may often help you in *solving* but rarely in *setting up* models.)

Now many physical concepts, such as velocity and acceleration, are derivatives. Hence a model is very often an equation containing derivatives of an unknown function. Such a model is called a **differential equation**. Of course, we then want to find a solution (a function that satisfies the equation), explore its properties, graph it, find values of it, and interpret it in physical terms so that we can understand the behavior of the physical system in our given problem. However, before we can turn to methods of solution, we must first define some basic concepts needed throughout this chapter.

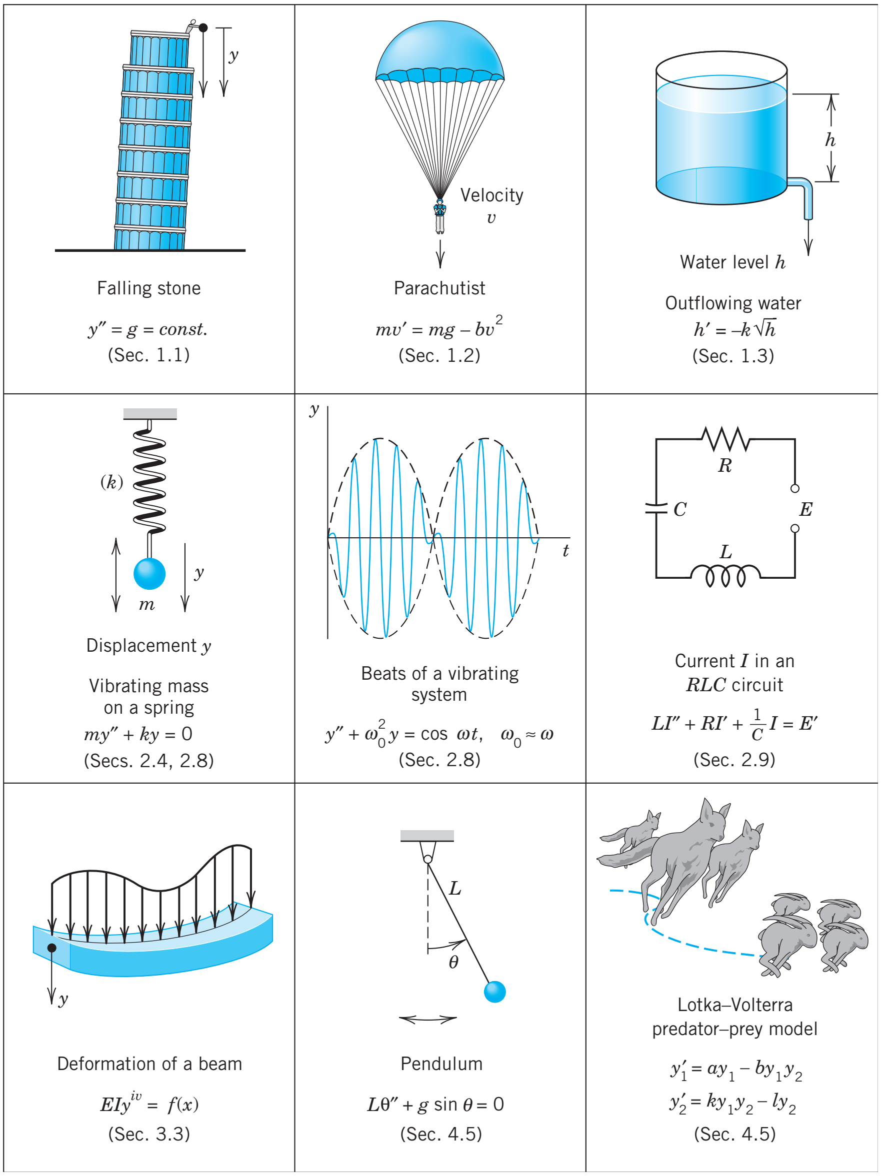


Fig. 2. Some applications of differential equations

An **ordinary differential equation (ODE)** is an equation that contains one or several derivatives of an unknown function, which we usually call $y(x)$ (or sometimes $y(t)$ if the independent variable is time t). The equation may also contain y itself, known functions of x (or t), and constants. For example,

$$y' = \cos x \tag{1}$$

$$y'' + 9y = e^{-2x} \tag{2}$$

$$y'y''' - \tfrac{3}{2}y'^2 = 0 \tag{3}$$

are ordinary differential equations (ODEs). Here, as in calculus, y' denotes dy/dx, $y'' = d^2y/dx^2$, etc. The term *ordinary* distinguishes them from *partial differential equations* (PDEs), which involve partial derivatives of an unknown function of *two or more* variables. For instance, a PDE with unknown function u of two variables x and y is

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0.$$

PDEs have important engineering applications, but they are more complicated than ODEs; they will be considered in Chap. 12.

An ODE is said to be of **order** n if the nth derivative of the unknown function y is the highest derivative of y in the equation. The concept of order gives a useful classification into ODEs of first order, second order, and so on. Thus, (1) is of first order, (2) of second order, and (3) of third order.

In this chapter we shall consider **first-order ODEs**. Such equations contain only the first derivative y' and may contain y and any given functions of x. Hence we can write them as

$$F(x, y, y') = 0 \tag{4}$$

or often in the form

$$y' = f(x, y).$$

This is called the *explicit form*, in contrast to the *implicit form* (4). For instance, the implicit ODE $x^{-3}y' - 4y^2 = 0$ (where $x \neq 0$) can be written explicitly as $y' = 4x^3y^2$.

Concept of Solution

A function

$$y = h(x)$$

is called a **solution** of a given ODE (4) on some open interval $a < x < b$ if $h(x)$ is defined and differentiable throughout the interval and is such that the equation becomes an identity if y and y' are replaced with h and h', respectively. The curve (the graph) of h is called a **solution curve**.

Here, **open interval** $a < x < b$ means that the endpoints a and b are not regarded as points belonging to the interval. Also, $a < x < b$ includes *infinite intervals* $-\infty < x < b$, $a < x < \infty$, $-\infty < x < \infty$ (the real line) as special cases.

EXAMPLE 1 Verification of Solution

Verify that $y = c/x$ (c an arbitrary constant) is a solution of the ODE $xy' = -y$ for all $x \neq 0$. Indeed, differentiate $y = c/x$ to get $y' = -c/x^2$. Multiply this by x, obtaining $xy' = -c/x$; thus, $xy' = -y$, the given ODE. ■

EXAMPLE 2 **Solution by Calculus. Solution Curves**

The ODE $y' = dy/dx = \cos x$ can be solved directly by integration on both sides. Indeed, using calculus, we obtain $y = \int \cos x \, dx = \sin x + c$, where c is an arbitrary constant. This is a ***family of solutions***. Each value of c, for instance, 2.75 or 0 or -8, gives one of these curves. Figure 3 shows some of them, for $c = -3, -2, -1, 0, 1, 2, 3, 4$. ■

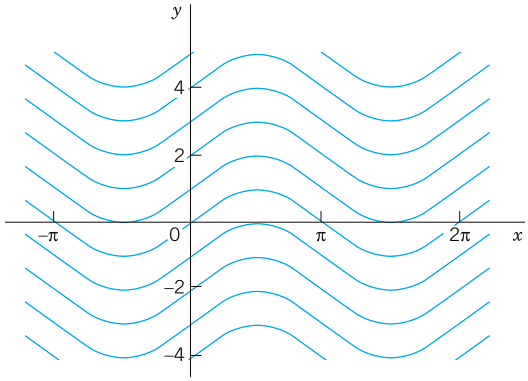


Fig. 3. Solutions $y = \sin x + c$ of the ODE $y' = \cos x$

EXAMPLE 3 **(A) Exponential Growth. (B) Exponential Decay**

From calculus we know that $y = ce^{0.2t}$ has the derivative

$$y' = \frac{dy}{dt} = 0.2e^{0.2t} = 0.2y.$$

Hence y is a solution of $y' = 0.2y$ (Fig. 4A). This ODE is of the form $y' = ky$. With positive-constant k it can model exponential growth, for instance, of colonies of bacteria or populations of animals. It also applies to humans for small populations in a large country (e.g., the United States in early times) and is then known as **Malthus's law**.[1] We shall say more about this topic in Sec. 1.5.

(B) Similarly, $y' = -0.2$ (with a minus on the right) has the solution $y = ce^{-0.2t}$, (Fig. 4B) modeling **exponential decay**, as, for instance, of a radioactive substance (see Example 5). ■

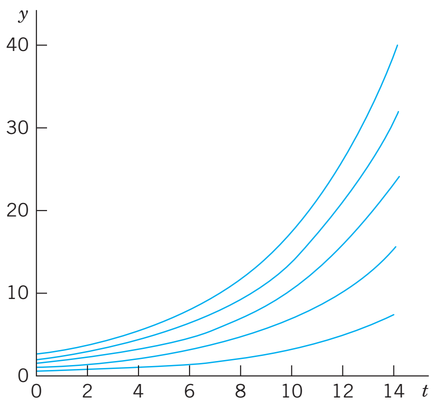


Fig. 4A. Solutions of $y' = 0.2y$ in Example 3 **(exponential growth)**

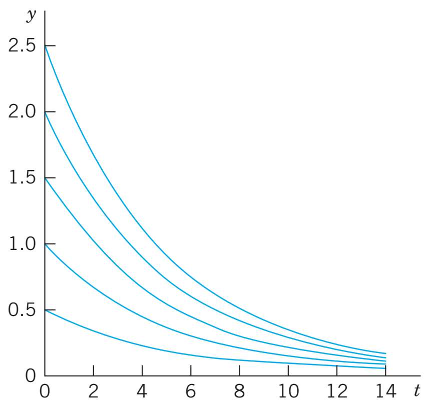


Fig. 4B. Solutions of $y' = -0.2y$ in Example 3 **(exponential decay)**

---

[1]Named after the English pioneer in classic economics, THOMAS ROBERT MALTHUS (1766–1834).

We see that each ODE in these examples has a solution that contains an arbitrary constant c. Such a solution containing an arbitrary constant c is called a **general solution** of the ODE.

(We shall see that c is sometimes not completely arbitrary but must be restricted to some interval to avoid complex expressions in the solution.)

We shall develop methods that will give general solutions *uniquely* (perhaps except for notation). Hence we shall say ***the*** general solution of a given ODE (instead of ***a*** general solution).

Geometrically, the general solution of an ODE is a family of infinitely many solution curves, one for each value of the constant c. If we choose a specific c (e.g., $c = 6.45$ or 0 or -2.01) we obtain what is called a **particular solution** of the ODE. A particular solution does not contain any arbitrary constants.

In most cases, general solutions exist, and every solution not containing an arbitrary constant is obtained as a particular solution by assigning a suitable value to c. Exceptions to these rules occur but are of minor interest in applications; see Prob. 16 in Problem Set 1.1.

Initial Value Problem

In most cases the unique solution of a given problem, hence a particular solution, is obtained from a general solution by an **initial condition** $y(x_0) = y_0$, with given values x_0 and y_0, that is used to determine a value of the arbitrary constant c. Geometrically this condition means that the solution curve should pass through the point (x_0, y_0) in the xy-plane. An ODE, together with an initial condition, is called an **initial value problem**. Thus, if the ODE is explicit, $y' = f(x, y)$, the initial value problem is of the form

$$y' = f(x, y), \qquad y(x_0) = y_0. \tag{5}$$

EXAMPLE 4 Initial Value Problem

Solve the initial value problem

$$y' = \frac{dy}{dx} = 3y, \qquad y(0) = 5.7.$$

Solution. The general solution is $y(x) = ce^{3x}$; see Example 3. From this solution and the initial condition we obtain $y(0) = ce^0 = c = 5.7$. Hence the initial value problem has the solution $y(x) = 5.7e^{3x}$. This is a particular solution. ■

More on Modeling

The general importance of modeling to the engineer and physicist was emphasized at the beginning of this section. We shall now consider a basic physical problem that will show the details of the typical steps of modeling. Step 1: the transition from the physical situation (the physical system) to its mathematical formulation (its mathematical model); Step 2: the solution by a mathematical method; and Step 3: the physical interpretation of the result. This may be the easiest way to obtain a first idea of the nature and purpose of differential equations and their applications. Realize at the outset that your ***computer*** (your ***CAS***) may perhaps give you a hand in Step 2, but Steps 1 and 3 are basically your work.

And Step 2 requires a solid knowledge and good understanding of solution methods available to you—*you* have to choose the method for your work by hand or by the computer. Keep this in mind, and always check computer results for errors (which may arise, for instance, from false inputs).

EXAMPLE 5 Radioactivity. Exponential Decay

Given an amount of a radioactive substance, say, 0.5 g (gram), find the amount present at any later time.

Physical Information. Experiments show that at each instant a radioactive substance decomposes—and is thus decaying in time—proportional to the amount of substance present.

Step 1. Setting up a mathematical model of the physical process. Denote by $y(t)$ the amount of substance still present at any time t. By the physical law, the time rate of change $y'(t) = dy/dt$ is proportional to $y(t)$. This gives the ***first-order ODE***

$$\frac{dy}{dt} = -ky \tag{6}$$

where the constant k is positive, so that, because of the minus, we do get decay (as in [B] of Example 3). The value of k is known from experiments for various radioactive substances (e.g., $k = 1.4 \cdot 10^{-11}\ \mathrm{sec}^{-1}$, approximately, for radium ${}^{226}_{88}\mathrm{Ra}$).

Now the given initial amount is 0.5 g, and we can call the corresponding instant $t = 0$. Then we have the ***initial condition*** $y(0) = 0.5$. This is the instant at which our observation of the process begins. It motivates the term *initial condition* (which, however, is also used when the independent variable is not time or when we choose a t other than $t = 0$). Hence the mathematical model of the physical process is the ***initial value problem***

$$\frac{dy}{dt} = -ky, \qquad y(0) = 0.5. \tag{7}$$

Step 2. Mathematical solution. As in (B) of Example 3 we conclude that the ODE (6) models exponential decay and has the general solution (with arbitrary constant c but definite given k)

$$y(t) = ce^{-kt}. \tag{8}$$

We now determine c by using the initial condition. Since $y(0) = c$ from (8), this gives $y(0) = c = 0.5$. Hence the particular solution governing our process is (cf. Fig. 5)

$$y(t) = 0.5e^{-kt} \qquad (k > 0). \tag{9}$$

Always check your result—it may involve human or computer errors! Verify by differentiation (chain rule!) that your solution (9) satisfies (7) as well as $y(0) = 0.5$:

$$\frac{dy}{dt} = -0.5ke^{-kt} = -k \cdot 0.5e^{-kt} = -ky, \qquad y(0) = 0.5e^{0} = 0.5.$$

Step 3. Interpretation of result. Formula (9) gives the amount of radioactive substance at time t. It starts from the correct initial amount and decreases with time because k is positive. The limit of y as $t \to \infty$ is zero. ■

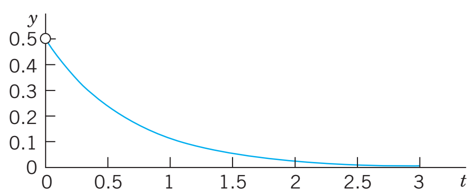


Fig. 5. Radioactivity (Exponential decay, $y = 0.5e^{-kt}$, with $k = 1.5$ as an example)

PROBLEM SET 1.1

1–8 CALCULUS

Solve the ODE by integration or by remembering a differentiation formula.

1. $y' + 2 \sin 2\pi x = 0$
2. $y' + xe^{-x^2/2} = 0$
3. $y' = y$
4. $y' = -1.5y$
5. $y' = 4e^{-x} \cos x$
6. $y'' = -y$
7. $y' = \cosh 5.13x$
8. $y''' = e^{-0.2x}$

9–15 VERIFICATION. INITIAL VALUE PROBLEM (IVP)

(a) Verify that y is a solution of the ODE. **(b)** Determine from y the particular solution of the IVP. **(c)** Graph the solution of the IVP.

9. $y' + 4y = 1.4, \quad y = ce^{-4x} + 0.35, \quad y(0) = 2$
10. $y' + 5xy = 0, \quad y = ce^{-2.5x^2}, \quad y(0) = \pi$
11. $y' = y + e^x, \quad y = (x + c)e^x, \quad y(0) = \frac{1}{2}$
12. $yy' = 4x, \quad y^2 - 4x^2 = c \; (y > 0), \quad y(1) = 4$
13. $y' = y - y^2, \quad y = \dfrac{1}{1 + ce^{-x}}, \quad y(0) = 0.25$
14. $y' \tan x = 2y - 8, \quad y = c \sin^2 x + 4, \quad y(\frac{1}{2}\pi) = 0$
15. Find two constant solutions of the ODE in Prob. 13 by inspection.
16. **Singular solution.** An ODE may sometimes have an additional solution that cannot be obtained from the general solution and is then called a *singular solution.* The ODE $y'^2 - xy' + y = 0$ is of this kind. Show by differentiation and substitution that it has the general solution $y = cx - c^2$ and the singular solution $y = x^2/4$. Explain Fig. 6.

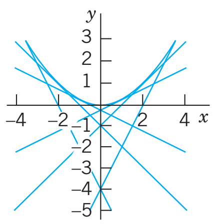


Fig. 6. Particular solutions and singular solution in Problem 16

17–20 MODELING, APPLICATIONS

These problems will give you a first impression of modeling. Many more problems on modeling follow throughout this chapter.

17. **Half-life.** The *half-life* measures exponential decay. It is the time in which half of the given amount of radioactive substance will disappear. What is the half-life of ${}^{226}_{88}\mathrm{Ra}$ (in years) in Example 5?
18. **Half-life.** Radium ${}^{224}_{88}\mathrm{Ra}$ has a half-life of about 3.6 days.
    (a) Given 1 gram, how much will still be present after 1 day?
    (b) After 1 year?
19. **Free fall.** In dropping a stone or an iron ball, air resistance is practically negligible. Experiments show that the acceleration of the motion is constant (equal to $g = 9.80 \text{ m/sec}^2 = 32 \text{ ft/sec}^2$, called the **acceleration of gravity**). Model this as an ODE for $y(t)$, the distance fallen as a function of time t. If the motion starts at time $t = 0$ from rest (i.e., with velocity $v = y' = 0$), show that you obtain the familiar law of free fall
$$y = \tfrac{1}{2}gt^2.$$
20. **Exponential decay. Subsonic flight.** The efficiency of the engines of subsonic airplanes depends on air pressure and is usually maximum near 35,000 ft. Find the air pressure $y(x)$ at this height. *Physical information.* The rate of change $y'(x)$ is proportional to the pressure. At 18,000 ft it is half its value $y_0 = y(0)$ at sea level. *Hint.* Remember from calculus that if $y = e^{kx}$, then $y' = ke^{kx} = ky$. Can you see without calculation that the answer should be close to $y_0/4$?

1.2 Geometric Meaning of $y' = f(x, y)$. Direction Fields, Euler's Method

A first-order ODE

$$y' = f(x, y) \tag{1}$$

has a simple geometric interpretation. From calculus you know that the derivative $y'(x)$ of $y(x)$ is the slope of $y(x)$. Hence a solution curve of (1) that passes through a point (x_0, y_0) must have, at that point, the slope $y'(x_0)$ equal to the value of f at that point; that is,

$$y'(x_0) = f(x_0, y_0).$$

Using this fact, we can develop graphic or numeric methods for obtaining approximate solutions of ODEs (1). This will lead to a better conceptual understanding of an ODE (1). Moreover, such methods are of practical importance since many ODEs have complicated solution formulas or no solution formulas at all, whereby numeric methods are needed.

Graphic Method of Direction Fields. Practical Example Illustrated in Fig. 7. We can show directions of solution curves of a given ODE (1) by drawing short straight-line segments (lineal elements) in the xy-plane. This gives a **direction field** (or *slope field*) into which you can then fit (approximate) solution curves. This may reveal typical properties of the whole family of solutions.

Figure 7 shows a direction field for the ODE

$$y' = y + x \tag{2}$$

obtained by a CAS (Computer Algebra System) and some approximate solution curves fitted in.

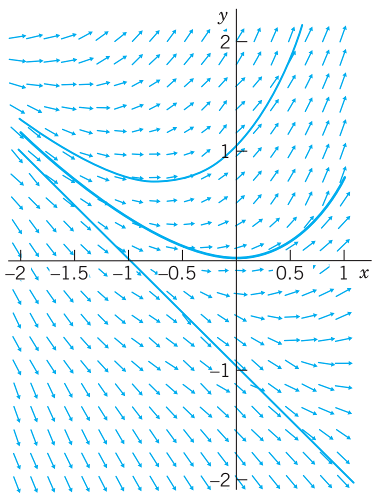


Fig. 7. Direction field of $y' = y + x$, with three approximate solution curves passing through (0, 1), (0, 0), (0, −1), respectively

If you have no CAS, first draw a few *level curves* $f(x, y) = \text{const}$ of $f(x, y)$, then parallel lineal elements along each such curve (which is also called an **isocline**, meaning a curve of equal inclination), and finally draw approximation curves fit to the lineal elements.

We shall now illustrate how numeric methods work by applying the simplest numeric method, that is Euler's method, to an initial value problem involving ODE (2). First we give a brief description of Euler's method.

Numeric Method by Euler

Given an ODE (1) and an initial value $y(x_0) = y_0$, **Euler's method** yields approximate solution values at equidistant x-values $x_0, x_1 = x_0 + h, x_2 = x_0 + 2h, \cdots$, namely,

$$
\begin{aligned}
y_1 &= y_0 + hf(x_0, y_0) \qquad &&\text{(Fig. 8)} \\
y_2 &= y_1 + hf(x_1, y_1), &&\text{etc.}
\end{aligned}
$$

In general,

$$y_n = y_{n-1} + hf(x_{n-1}, y_{n-1})$$

where the step h equals, e.g., 0.1 or 0.2 (as in Table 1.1) or a smaller value for greater accuracy.

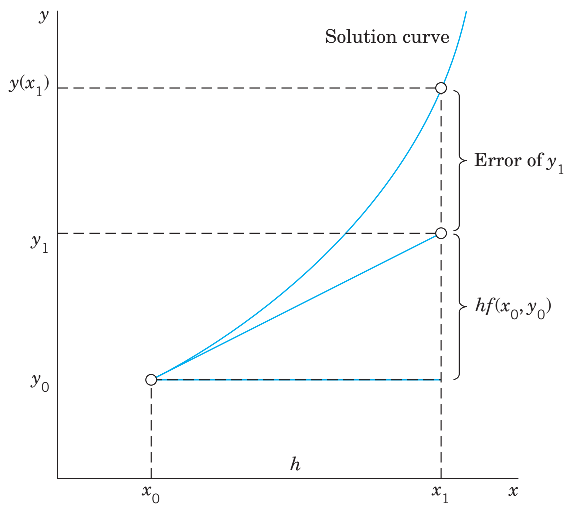


Fig. 8. First Euler step, showing a solution curve, its tangent at (x_0, y_0), step h and increment $hf(x_0, y_0)$ in the formula for y_1

Table 1.1 shows the computation of $n = 5$ steps with step $h = 0.2$ for the ODE (2) and initial condition $y(0) = 0$, corresponding to the middle curve in the direction field. We shall solve the ODE exactly in Sec. 1.5. For the time being, verify that the initial value problem has the solution $y = e^x - x - 1$. The solution curve and the values in Table 1.1 are shown in Fig. 9. These values are rather inaccurate. The errors $y(x_n) - y_n$ are shown in Table 1.1 as well as in Fig. 9. Decreasing h would improve the values, but would soon require an impractical amount of computation. Much better methods of a similar nature will be discussed in Sec. 21.1.

Table 1.1. **Euler method for $y' = y + x, y(0) = 0$ for $x = 0, \cdots, 1.0$ with step $h = 0.2$**

n	x_n	y_n	$y(x_n)$	Error
0	0.0	0.000	0.000	0.000
1	0.2	0.000	0.021	0.021
2	0.4	0.04	0.092	0.052
3	0.6	0.128	0.222	0.094
4	0.8	0.274	0.426	0.152
5	1.0	0.488	0.718	0.230

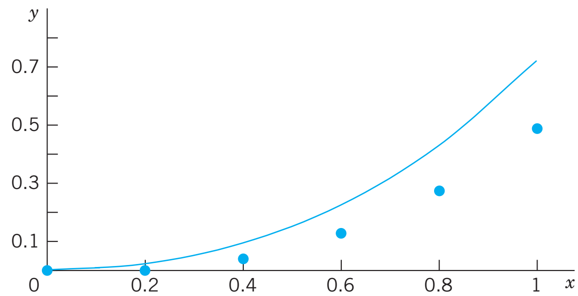


Fig. 9. Euler method: Approximate values in Table 1.1 and solution curve

PROBLEM SET 1.2

1–8 DIRECTION FIELDS, SOLUTION CURVES

Graph a direction field (by a CAS or by hand). In the field graph several solution curves by hand, particularly those passing through the given points (x, y).

1. $y' = 1 + y^2, \quad (\frac{1}{4}\pi, 1)$
2. $yy' + 4x = 0, \quad (1, 1), (0, 2)$
3. $y' = 1 - y^2, \quad (0, 0), (2, \frac{1}{2})$
4. $y' = 2y - y^2, \quad (0, 0), (0, 1), (0, 2), (0, 3)$
5. $y' = x - 1/y, \quad (1, \frac{1}{2})$
6. $y' = \sin^2 y, \quad (0, -0.4), (0, 1)$
7. $y' = e^{y/x}, \quad (2, 2), (3, 3)$
8. $y' = -2xy, \quad (0, \frac{1}{2}), (0, 1), (0, 2)$

9–10 ACCURACY OF DIRECTION FIELDS

Direction fields are very useful because they can give you an impression of all solutions without solving the ODE, which may be difficult or even impossible. To get a feel for the accuracy of the method, graph a field, sketch solution curves in it, and compare them with the exact solutions.

9. $y' = \cos \pi x$
10. $y' = -5y^{1/2}$ (Sol. $\sqrt{y} + \frac{5}{2}x = c$)
11. **Autonomous ODE.** This means an ODE not showing x (the independent variable) *explicitly*. (The ODEs in Probs. 6 and 10 are autonomous.) What will the level curves $f(x, y) = \text{const}$ (also called *isoclines* = curves of equal inclination) of an autonomous ODE look like? Give reason.

12–15 MOTIONS

Model the motion of a body B on a straight line with velocity as given, $y(t)$ being the distance of B from a point $y = 0$ at time t. Graph a direction field of the model (the ODE). In the field sketch the solution curve satisfying the given initial condition.

12. Product of velocity times distance constant, equal to 2, $y(0) = 2$.
13. Distance = Velocity × Time, $\quad y(1) = 1$
14. Square of the distance plus square of the velocity equal to 1, initial distance $1/\sqrt{2}$
15. **Parachutist.** Two forces act on a parachutist, the attraction by the earth mg (m = mass of person plus equipment, $g = 9.8 \text{ m/sec}^2$ the acceleration of gravity) and the air resistance, assumed to be proportional to the square of the velocity $v(t)$. Using **Newton's second law** of motion (mass × acceleration = resultant of the forces), set up a model (an ODE for $v(t)$). Graph a direction field (choosing m and the constant of proportionality equal to 1). Assume that the parachute opens when $v = 10$ m/sec. Graph the corresponding solution in the field. What is the limiting velocity? Would the parachute still be sufficient if the air resistance were only proportional to $v(t)$?

16. CAS PROJECT. Direction Fields. Discuss direction fields as follows.

(a) Graph portions of the direction field of the ODE (2) (see Fig. 7), for instance, $-5 \leqq x \leqq 2$, $-1 \leqq y \leqq 5$. Explain what you have gained by this enlargement of the portion of the field.

(b) Using implicit differentiation, find an ODE with the general solution $x^2 + 9y^2 = c \ (y > 0)$. Graph its direction field. Does the field give the impression that the solution curves may be semi-ellipses? Can you do similar work for circles? Hyperbolas? Parabolas? Other curves?

(c) Make a conjecture about the solutions of $y' = -x/y$ from the direction field.

(d) Graph the direction field of $y' = -\frac{1}{2}y$ and some solutions of your choice. How do they behave? Why do they decrease for $y > 0$?

17–20 EULER'S METHOD

This is the simplest method to explain numerically solving an ODE, more precisely, an initial value problem (IVP). (More accurate methods based on the same principle are explained in Sec. 21.1.) Using the method, to get a feel for numerics as well as for the nature of IVPs, solve the IVP numerically with a PC or a calculator, 10 steps. Graph the computed values and the solution curve on the same coordinate axes.

17. $y' = y, \quad y(0) = 1, \quad h = 0.1$

18. $y' = y, \quad y(0) = 1, \quad h = 0.01$

19. $y' = (y - x)^2, \quad y(0) = 0, \quad h = 0.1$
Sol. $y = x - \tanh x$

20. $y' = -5x^4y^2, \quad y(0) = 1, \quad h = 0.2$
Sol. $y = 1/(1 + x)^5$

1.3 Separable ODEs. Modeling

Many practically useful ODEs can be reduced to the form

$$g(y)\,y' = f(x) \tag{1}$$

by purely algebraic manipulations. Then we can integrate on both sides with respect to x, obtaining

$$\int g(y)\,y'\,dx = \int f(x)\,dx + c. \tag{2}$$

On the left we can switch to y as the variable of integration. By calculus, $y'\,dx = dy$, so that

$$\int g(y)\,dy = \int f(x)\,dx + c. \tag{3}$$

If f and g are continuous functions, the integrals in (3) exist, and by evaluating them we obtain a general solution of (1). This method of solving ODEs is called the **method of separating variables**, and (1) is called a **separable equation**, because in (3) the variables are now separated: x appears only on the right and y only on the left.

EXAMPLE 1 Separable ODE

The ODE $y' = 1 + y^2$ is separable because it can be written

$$\frac{dy}{1 + y^2} = dx. \qquad \text{By integration,} \qquad \arctan y = x + c \qquad \text{or} \qquad y = \tan(x + c).$$

It is very important to introduce the constant of integration immediately when the integration is performed. If we wrote $\arctan y = x$, then $y = \tan x$, and *then* introduced c, we would have obtained $y = \tan x + c$, which is not a solution (when $c \neq 0$). Verify this. ■

EXAMPLE 2 **Separable ODE**

The ODE $y' = (x + 1)e^{-x}y^2$ is separable; we obtain $y^{-2}\,dy = (x + 1)e^{-x}\,dx$.

By integration, $-y^{-1} = -(x + 2)e^{-x} + c, \qquad y = \dfrac{1}{(x + 2)e^{-x} - c}.$ ■

EXAMPLE 3 **Initial Value Problem (IVP). Bell-Shaped Curve**

Solve $y' = -2xy,\ y(0) = 1.8.$

Solution. By separation and integration,

$$\frac{dy}{y} = -2x\,dx, \qquad \ln y = -x^2 + \tilde{c}, \qquad y = ce^{-x^2}.$$

This is the general solution. From it and the initial condition, $y(0) = ce^0 = c = 1.8$. Hence the IVP has the solution $y = 1.8e^{-x^2}$. This is a particular solution, representing a bell-shaped curve (Fig. 10). ■

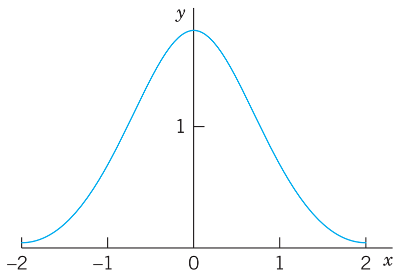


Fig. 10. Solution in Example 3 (bell-shaped curve)

Modeling

The importance of modeling was emphasized in Sec. 1.1, and separable equations yield various useful models. Let us discuss this in terms of some typical examples.

EXAMPLE 4 **Radiocarbon Dating**[2]

In September 1991 the famous Iceman (Oetzi), a mummy from the Neolithic period of the Stone Age found in the ice of the Oetztal Alps (hence the name "Oetzi") in Southern Tyrolia near the Austrian–Italian border, caused a scientific sensation. When did Oetzi approximately live and die if the ratio of carbon ${}^{14}_{6}C$ to carbon ${}^{12}_{6}C$ in this mummy is 52.5% of that of a living organism?

Physical Information. In the atmosphere and in living organisms, the ratio of radioactive carbon ${}^{14}_{6}C$ (made radioactive by cosmic rays) to ordinary carbon ${}^{12}_{6}C$ is constant. When an organism dies, its absorption of ${}^{14}_{6}C$ by breathing and eating terminates. Hence one can estimate the age of a fossil by comparing the radioactive carbon ratio in the fossil with that in the atmosphere. To do this, one needs to know the half-life of ${}^{14}_{6}C$, which is 5715 years (*CRC Handbook of Chemistry and Physics,* 83rd ed., Boca Raton: CRC Press, 2002, page 11–52, line 9).

Solution. ***Modeling.*** Radioactive decay is governed by the ODE $y' = ky$ (see Sec. 1.1, Example 5). By separation and integration (where t is time and y_0 is the initial ratio of ${}^{14}_{6}C$ to ${}^{12}_{6}C$)

$$\frac{dy}{y} = k\,dt, \qquad \ln|y| = kt + c, \qquad y = y_0 e^{kt} \qquad (y_0 = e^c).$$

---

[2]Method by WILLARD FRANK LIBBY (1908–1980), American chemist, who was awarded for this work the 1960 Nobel Prize in chemistry.

Next we use the half-life $H = 5715$ to determine k. When $t = H$, half of the original substance is still present. Thus,

$$y_0 e^{kH} = 0.5y_0, \qquad e^{kH} = 0.5, \qquad k = \frac{\ln 0.5}{H} = -\frac{0.693}{5715} = -0.0001213.$$

Finally, we use the ratio 52.5% for determining the time t when Oetzi died (actually, was killed),

$$e^{kt} = e^{-0.0001213t} = 0.525, \qquad t = \frac{\ln 0.525}{-0.0001213} = 5312. \qquad \textit{Answer:} \qquad \text{About 5300 years ago.}$$

Other methods show that radiocarbon dating values are usually too small. According to recent research, this is due to a variation in that carbon ratio because of industrial pollution and other factors, such as nuclear testing. ■

EXAMPLE 5 Mixing Problem

Mixing problems occur quite frequently in chemical industry. We explain here how to solve the basic model involving a single tank. The tank in Fig. 11 contains 1000 gal of water in which initially 100 lb of salt is dissolved. Brine runs in at a rate of 10 gal/min, and each gallon contains 5 lb of dissoved salt. The mixture in the tank is kept uniform by stirring. Brine runs out at 10 gal/min. Find the amount of salt in the tank at any time t.

Solution. ***Step 1. Setting up a model.*** Let $y(t)$ denote the amount of salt in the tank at time t. Its time rate of change is

$$y' = \text{Salt inflow rate} - \text{Salt outflow rate} \qquad \textbf{Balance law.}$$

5 lb times 10 gal gives an inflow of 50 lb of salt. Now, the outflow is 10 gal of brine. This is $10/1000 = 0.01$ $(= 1\%)$ of the total brine content in the tank, hence 0.01 of the salt content $y(t)$, that is, $0.01\ y(t)$. Thus the model is the ODE

$$y' = 50 - 0.01y = -0.01(y - 5000). \tag{4}$$

Step 2. Solution of the model. The ODE (4) is separable. Separation, integration, and taking exponents on both sides gives

$$\frac{dy}{y - 5000} = -0.01\,dt, \qquad \ln|y - 5000| = -0.01t + c^*, \qquad y - 5000 = ce^{-0.01t}.$$

Initially the tank contains 100 lb of salt. Hence $y(0) = 100$ is the initial condition that will give the unique solution. Substituting $y = 100$ and $t = 0$ in the last equation gives $100 - 5000 = ce^0 = c$. Hence $c = -4900$. Hence the amount of salt in the tank at time t is

$$y(t) = 5000 - 4900e^{-0.01t}. \tag{5}$$

This function shows an exponential approach to the limit 5000 lb; see Fig. 11. Can you explain physically that $y(t)$ should increase with time? That its limit is 5000 lb? Can you see the limit directly from the ODE?

The model discussed becomes more realistic in problems on pollutants in lakes (see Problem Set 1.5, Prob. 35) or drugs in organs. These types of problems are more difficult because the mixing may be imperfect and the flow rates (in and out) may be different and known only very roughly. ■

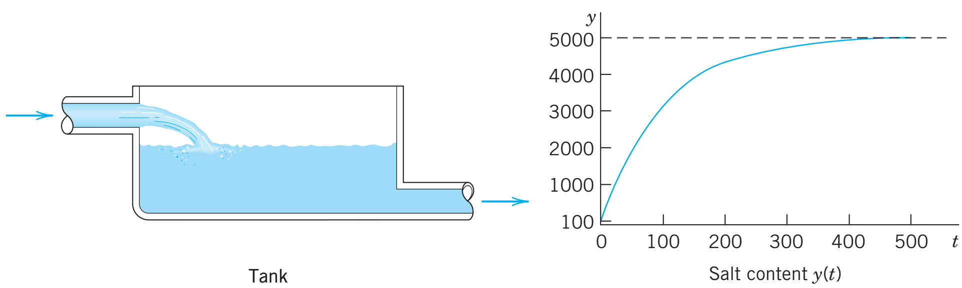


Fig. 11. Mixing problem in Example 5

EXAMPLE 6 Heating an Office Building (Newton's Law of Cooling[3])

Suppose that in winter the daytime temperature in a certain office building is maintained at 70°F. The heating is shut off at 10 P.M. and turned on again at 6 A.M. On a certain day the temperature inside the building at 2 A.M. was found to be 65°F. The outside temperature was 50°F at 10 P.M. and had dropped to 40°F by 6 A.M. What was the temperature inside the building when the heat was turned on at 6 A.M.?

Physical information. Experiments show that the time rate of change of the temperature T of a body B (which conducts heat well, for example, as a copper ball does) is proportional to the difference between T and the temperature of the surrounding medium **(Newton's law of cooling).**

Solution. ***Step 1. Setting up a model.*** Let $T(t)$ be the temperature inside the building and T_A the outside temperature (assumed to be constant in Newton's law). Then by Newton's law,

$$\frac{dT}{dt} = k(T - T_A). \tag{6}$$

Such experimental laws are derived under idealized assumptions that rarely hold exactly. However, even if a model seems to fit the reality only poorly (as in the present case), it may still give valuable qualitative information. To see how good a model is, the engineer will collect experimental data and compare them with calculations from the model.

Step 2. General solution. We cannot solve (6) because we do not know T_A, just that it varied between 50°F and 40°F, so we follow the ***Golden Rule:*** *If you cannot solve your problem, try to solve a simpler one.* We solve (6) with the unknown function T_A replaced with the average of the two known values, or 45°F. For physical reasons we may expect that this will give us a reasonable approximate value of T in the building at 6 A.M.

For constant $T_A = 45$ (or any other *constant* value) the ODE (6) is separable. Separation, integration, and taking exponents gives the general solution

$$\frac{dT}{T - 45} = k\,dt, \qquad \ln|T - 45| = kt + c^*, \qquad T(t) = 45 + ce^{kt} \qquad (c = e^{c^*}).$$

Step 3. Particular solution. We choose 10 P.M. to be $t = 0$. Then the given initial condition is $T(0) = 70$ and yields a particular solution, call it T_p. By substitution,

$$T(0) = 45 + ce^0 = 70, \qquad c = 70 - 45 = 25, \qquad T_p(t) = 45 + 25e^{kt}.$$

Step 4. Determination of k. We use $T(4) = 65$, where $t = 4$ is 2 A.M. Solving algebraically for k and inserting k into $T_p(t)$ gives (Fig. 12)

$$T_p(4) = 45 + 25e^{4k} = 65, \qquad e^{4k} = 0.8, \qquad k = \tfrac{1}{4}\ln 0.8 = -0.056, \qquad T_p(t) = 45 + 25e^{-0.056t}.$$

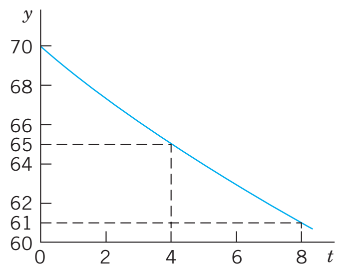


Fig. 12. Particular solution (temperature) in Example 6

---

[3]Sir ISAAC NEWTON (1642–1727), great English physicist and mathematician, became a professor at Cambridge in 1669 and Master of the Mint in 1699. He and the German mathematician and philosopher GOTTFRIED WILHELM LEIBNIZ (1646–1716) invented (independently) the differential and integral calculus. Newton discovered many basic physical laws and created the method of investigating physical problems by means of calculus. His *Philosophiae naturalis principia mathematica* (*Mathematical Principles of Natural Philosophy,* 1687) contains the development of classical mechanics. His work is of greatest importance to both mathematics and physics.

Step 5. Answer and interpretation. 6 A.M. is $t = 8$ (namely, 8 hours after 10 P.M.), and

$$T_p(8) = 45 + 25e^{-0.056 \cdot 8} = 61[°\text{F}].$$

Hence the temperature in the building dropped 9°F, a result that looks reasonable. ■

EXAMPLE 7 Leaking Tank. Outflow of Water Through a Hole (Torricelli's Law)

This is another prototype engineering problem that leads to an ODE. It concerns the outflow of water from a cylindrical tank with a hole at the bottom (Fig. 13). You are asked to find the height of the water in the tank at any time if the tank has diameter 2 m, the hole has diameter 1 cm, and the initial height of the water when the hole is opened is 2.25 m. When will the tank be empty?

Physical information. Under the influence of gravity the outflowing water has velocity

$$v(t) = 0.600\sqrt{2gh(t)} \qquad \textbf{(Torricelli's law[4])}, \tag{7}$$

where $h(t)$ is the height of the water above the hole at time t, and $g = 980 \text{ cm/sec}^2 = 32.17 \text{ ft/sec}^2$ is the acceleration of gravity at the surface of the earth.

Solution. ***Step 1. Setting up the model.*** To get an equation, we relate the decrease in water level $h(t)$ to the outflow. The volume ΔV of the outflow during a short time Δt is

$$\Delta V = Av\,\Delta t \qquad (A = \text{Area of hole}).$$

ΔV must equal the change ΔV^* of the volume of the water in the tank. Now

$$\Delta V^* = -B\,\Delta h \qquad (B = \text{Cross-sectional area of tank})$$

where $\Delta h\ (> 0)$ is the decrease of the height $h(t)$ of the water. The minus sign appears because the volume of the water in the tank decreases. Equating ΔV and ΔV^* gives

$$-B\,\Delta h = Av\,\Delta t.$$

We now express v according to Torricelli's law and then let Δt (the length of the time interval considered) approach 0—this is a ***standard way*** of obtaining an ODE as a model. That is, we have

$$\frac{\Delta h}{\Delta t} = -\frac{A}{B}v = -\frac{A}{B}0.600\sqrt{2gh(t)}$$

and by letting $\Delta t \to 0$ we obtain the ODE

$$\frac{dh}{dt} = -26.56\frac{A}{B}\sqrt{h},$$

where $26.56 = 0.600\sqrt{2 \cdot 980}$. This is our model, a first-order ODE.

Step 2. General solution. Our ODE is separable. A/B is constant. Separation and integration gives

$$\frac{dh}{\sqrt{h}} = -26.56\frac{A}{B}dt \qquad \text{and} \qquad 2\sqrt{h} = c^* - 26.56\frac{A}{B}t.$$

Dividing by 2 and squaring gives $h = (c - 13.28At/B)^2$. Inserting $13.28A/B = 13.28 \cdot 0.5^2\pi/100^2\pi = 0.000332$ yields the general solution

$$h(t) = (c - 0.000332t)^2.$$

[4] EVANGELISTA TORRICELLI (1608–1647), Italian physicist, pupil and successor of GALILEO GALILEI (1564–1642) at Florence. The "contraction factor" 0.600 was introduced by J. C. BORDA in 1766 because the stream has a smaller cross section than the area of the hole.

Step 3. Particular solution. The initial height (the initial condition) is $h(0) = 225$ cm. Substitution of $t = 0$ and $h = 225$ gives from the general solution $c^2 = 225$, $c = 15.00$ and thus the particular solution (Fig. 13)

$$h_p(t) = (15.00 - 0.000\,332t)^2.$$

Step 4. Tank empty. $h_p(t) = 0$ if $t = 15.00/0.000\,332 = 45{,}181$ [sec] $= 12.6$ [hours].

Here you see distinctly the ***importance of the choice of units***—we have been working with the cgs system, in which time is measured in seconds! We used $g = 980\ \text{cm/sec}^2$.

Step 5. Checking. Check the result. ■

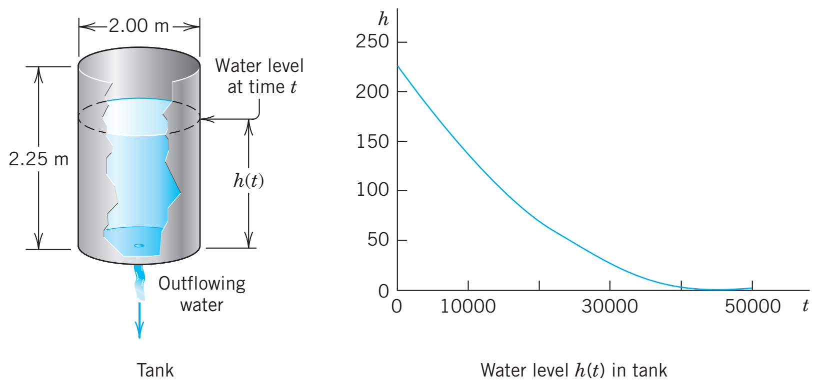


Fig. 13. Example 7. Outflow from a cylindrical tank ("leaking tank"). Torricelli's law

Extended Method: Reduction to Separable Form

Certain nonseparable ODEs can be made separable by transformations that introduce for y a new unknown function. We discuss this technique for a class of ODEs of practical importance, namely, for equations

$$y' = f\left(\frac{y}{x}\right). \tag{8}$$

Here, f is any (differentiable) function of y/x, such as $\sin(y/x)$, $(y/x)^4$, and so on. (Such an ODE is sometimes called a *homogeneous ODE,* a term we shall not use but reserve for a more important purpose in Sec. 1.5.)

The form of such an ODE suggests that we set $y/x = u$; thus,

$$y = ux \qquad \text{and by product differentiation} \qquad y' = u'x + u. \tag{9}$$

Substitution into $y' = f(y/x)$ then gives $u'x + u = f(u)$ or $u'x = f(u) - u$. We see that if $f(u) - u \neq 0$, this can be separated:

$$\frac{du}{f(u) - u} = \frac{dx}{x}. \tag{10}$$

EXAMPLE 8 Reduction to Separable Form

Solve

$$2xyy' = y^2 - x^2.$$

Solution. To get the usual explicit form, divide the given equation by $2xy$,

$$y' = \frac{y^2 - x^2}{2xy} = \frac{y}{2x} - \frac{x}{2y}.$$

Now substitute y and y' from (9) and then simplify by subtracting u on both sides,

$$u'x + u = \frac{u}{2} - \frac{1}{2u}, \qquad u'x = -\frac{u}{2} - \frac{1}{2u} = \frac{-u^2 - 1}{2u}.$$

You see that in the last equation you can now separate the variables,

$$\frac{2u\,du}{1 + u^2} = -\frac{dx}{x}. \qquad \text{By integration,} \qquad \ln(1 + u^2) = -\ln|x| + c^* = \ln\left|\frac{1}{x}\right| + c^*.$$

Take exponents on both sides to get $1 + u^2 = c/x$ or $1 + (y/x)^2 = c/x$. Multiply the last equation by x^2 to obtain (Fig. 14)

$$x^2 + y^2 = cx. \qquad \text{Thus} \qquad \left(x - \frac{c}{2}\right)^2 + y^2 = \frac{c^2}{4}.$$

This general solution represents a family of circles passing through the origin with centers on the x-axis. ■

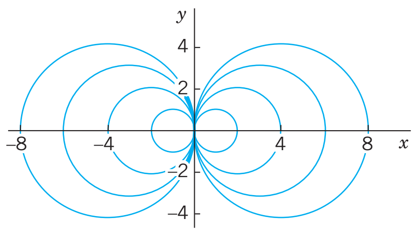


Fig. 14. General solution (family of circles) in Example 8

PROBLEM SET 1.3

1. **CAUTION! Constant of integration.** Why is it important to introduce the constant of integration immediately when you integrate?

2–10 GENERAL SOLUTION

Find a general solution. Show the steps of derivation. Check your answer by substitution.

2. $y^3 y' + x^3 = 0$
3. $y' = \sec^2 y$
4. $y' \sin 2\pi x = \pi y \cos 2\pi x$
5. $yy' + 36x = 0$
6. $y' = e^{2x-1} y^2$
7. $xy' = y + 2x^3 \sin^2 \dfrac{y}{x}$ (Set $y/x = u$)
8. $y' = (y + 4x)^2$ (Set $y + 4x = v$)
9. $xy' = y^2 + y$ (Set $y/x = u$)
10. $xy' = x + y$ (Set $y/x = u$)

11–17 INITIAL VALUE PROBLEMS (IVPs)

Solve the IVP. Show the steps of derivation, beginning with the general solution.

11. $xy' + y = 0, \quad y(4) = 6$
12. $y' = 1 + 4y^2, \quad y(1) = 0$
13. $y' \cosh^2 x = \sin^2 y, \quad y(0) = \frac{1}{2}\pi$
14. $dr/dt = -2tr, \quad r(0) = r_0$
15. $y' = -4x/y, \quad y(2) = 3$
16. $y' = (x + y - 2)^2, \quad y(0) = 2$ (Set $v = x + y - 2$)
17. $xy' = y + 3x^4 \cos^2(y/x), \quad y(1) = 0$ (Set $y/x = u$)
18. **Particular solution.** Introduce limits of integration in (3) such that y obtained from (3) satisfies the initial condition $y(x_0) = y_0$.

19–36 MODELING, APPLICATIONS

19. Exponential growth. If the growth rate of the number of bacteria at any time t is proportional to the number present at t and doubles in 1 week, how many bacteria can be expected after 2 weeks? After 4 weeks?

20. Another population model.

(a) If the birth rate and death rate of the number of bacteria are proportional to the number of bacteria present, what is the population as a function of time.

(b) What is the limiting situation for increasing time? Interpret it.

21. Radiocarbon dating. What should be the ${}^{14}_{6}C$ content (in percent of y_0) of a fossilized tree that is claimed to be 3000 years old? (See Example 4.)

22. Linear accelerators are used in physics for accelerating charged particles. Suppose that an alpha particle enters an accelerator and undergoes a constant acceleration that increases the speed of the particle from 10^3 m/sec to 10^4 m/sec in 10^{-3} sec. Find the acceleration a and the distance traveled during that period of 10^{-3} sec.

23. Boyle–Mariotte's law for ideal gases.[5] Experiments show for a gas at low pressure p (and constant temperature) the rate of change of the volume $V(p)$ equals $-V/p$. Solve the model.

24. Mixing problem. A tank contains 400 gal of brine in which 100 lb of salt are dissolved. Fresh water runs into the tank at a rate of 2 gal/min.The mixture, kept practically uniform by stirring, runs out at the same rate. How much salt will there be in the tank at the end of 1 hour?

25. Newton's law of cooling. A thermometer, reading 5°C, is brought into a room whose temperature is 22°C. One minute later the thermometer reading is 12°C. How long does it take until the reading is practically 22°C, say, 21.9°C?

26. Gompertz growth in tumors. The Gompertz model is $y' = -Ay \ln y$ $(A > 0)$, where $y(t)$ is the mass of tumor cells at time t. The model agrees well with clinical observations. The declining growth rate with increasing $y > 1$ corresponds to the fact that cells in the interior of a tumor may die because of insufficient oxygen and nutrients. Use the ODE to discuss the growth and decline of solutions (tumors) and to find constant solutions. Then solve the ODE.

27. Dryer. If a wet sheet in a dryer loses its moisture at a rate proportional to its moisture content, and if it loses half of its moisture during the first 10 min of drying, when will it be practically dry, say, when will it have lost 99% of its moisture? First guess, then calculate.

28. Estimation. Could you see, practically without calculation, that the answer in Prob. 27 must lie between 60 and 70 min? Explain.

29. Alibi? Jack, arrested when leaving a bar, claims that he has been inside for at least half an hour (which would provide him with an alibi). The police check the water temperature of his car (parked near the entrance of the bar) at the instant of arrest and again 30 min later, obtaining the values 190°F and 110°F, respectively. Do these results give Jack an alibi? (Solve by inspection.)

30. Rocket. A rocket is shot straight up from the earth, with a net acceleration (= acceleration by the rocket engine minus gravitational pullback) of $7t$ m/sec^2 during the initial stage of flight until the engine cut out at $t = 10$ sec. How high will it go, air resistance neglected?

31. Solution curves of $y' = g(y/x)$. Show that any (nonvertical) straight line through the origin of the xy-plane intersects all these curves of a given ODE at the same angle.

32. Friction. If a body slides on a surface, it experiences friction F (a force against the direction of motion). Experiments show that $|F| = \mu|N|$ (*Coulomb's*[6] *law of kinetic friction without lubrication*), where N is the normal force (force that holds the two surfaces together; see Fig. 15) and the constant of proportionality μ is called the *coefficient of kinetic friction*. In Fig. 15 assume that the body weighs 45 nt (about 10 lb; see front cover for conversion). $\mu = 0.20$ (corresponding to steel on steel), $a = 30°$, the slide is 10 m long, the initial velocity is zero, and air resistance is negligible. Find the velocity of the body at the end of the slide.

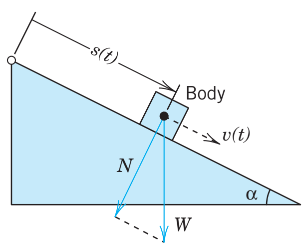


Fig. 15. Problem 32

---

[5] ROBERT BOYLE (1627–1691), English physicist and chemist, one of the founders of the Royal Society. EDME MARIOTTE (about 1620–1684), French physicist and prior of a monastry near Dijon. They found the law experimentally in 1662 and 1676, respectively.

[6] CHARLES AUGUSTIN DE COULOMB (1736–1806), French physicist and engineer.

33. Rope. To tie a boat in a harbor, how many times must a rope be wound around a bollard (a vertical rough cylindrical post fixed on the ground) so that a man holding one end of the rope can resist a force exerted by the boat 1000 times greater than the man can exert? First guess. Experiments show that the change ΔS of the force S in a small portion of the rope is proportional to S and to the small angle $\Delta\phi$ in Fig. 16. Take the proportionality constant 0.15. The result should surprise you!

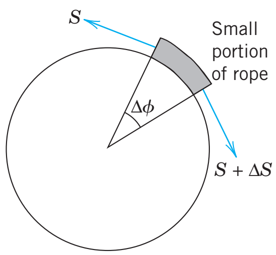


Fig. 16. Problem 33

34. TEAM PROJECT. Family of Curves. A family of curves can often be characterized as the general solution of $y' = f(x, y)$.

(a) Show that for the circles with center at the origin we get $y' = -x/y$.

(b) Graph some of the hyperbolas $xy = c$. Find an ODE for them.

(c) Find an ODE for the straight lines through the origin.

(d) You will see that the product of the right sides of the ODEs in (a) and (c) equals -1. Do you recognize this as the condition for the two families to be orthogonal (i.e., to intersect at right angles)? Do your graphs confirm this?

(e) Sketch families of curves of your own choice and find their ODEs. Can every family of curves be given by an ODE?

35. CAS PROJECT. Graphing Solutions. A CAS can usually graph solutions, even if they are integrals that cannot be evaluated by the usual analytical methods of calculus.

(a) Show this for the five initial value problems $y' = e^{-x^2}$, $y(0) = 0, \pm 1, \pm 2$ graphing all five curves on the same axes.

(b) Graph approximate solution curves, using the first few terms of the Maclaurin series (obtained by termwise integration of that of y') and compare with the exact curves.

(c) Repeat the work in (a) for another ODE and initial conditions of your own choice, leading to an integral that cannot be evaluated as indicated.

36. TEAM PROJECT. Torricelli's Law. Suppose that the tank in Example 7 is hemispherical, of radius R, initially full of water, and has an outlet of 5 cm^2 cross-sectional area at the bottom. (Make a sketch.) Set up the model for outflow. Indicate what portion of your work in Example 7 you can use (so that it can become part of the general method independent of the shape of the tank). Find the time t to empty the tank (a) for any R, (b) for $R = 1$ m. Plot t as function of R. Find the time when $h = R/2$ (a) for any R, (b) for $R = 1$ m.

1.4 Exact ODEs. Integrating Factors

We recall from calculus that if a function $u(x, y)$ has continuous partial derivatives, its **differential** (also called its *total differential*) is

$$du = \frac{\partial u}{\partial x}\,dx + \frac{\partial u}{\partial y}\,dy.$$

From this it follows that if $u(x, y) = c = \text{const}$, then $du = 0$.

For example, if $u = x + x^2y^3 = c$, then

$$du = (1 + 2xy^3)\,dx + 3x^2y^2\,dy = 0$$

or

$$y' = \frac{dy}{dx} = -\frac{1 + 2xy^3}{3x^2y^2},$$

an ODE that we can solve by going backward. This idea leads to a powerful solution method as follows.

A first-order ODE $M(x, y) + N(x, y)y' = 0$, written as (use $dy = y'dx$ as in Sec. 1.3)

$$\boxed{M(x, y)\,dx + N(x, y)\,dy = 0} \tag{1}$$

is called an **exact differential equation** if the **differential form** $M(x, y)\,dx + N(x, y)\,dy$ is **exact**, that is, this form is the differential

$$du = \frac{\partial u}{\partial x}dx + \frac{\partial u}{\partial y}dy \tag{2}$$

of some function $u(x, y)$. Then (1) can be written

$$du = 0.$$

By integration we immediately obtain the general solution of (1) in the form

$$\boxed{u(x, y) = c.} \tag{3}$$

This is called an **implicit solution**, in contrast to a solution $y = h(x)$ as defined in Sec. 1.1, which is also called an *explicit solution,* for distinction. Sometimes an implicit solution can be converted to explicit form. (Do this for $x^2 + y^2 = 1$.) If this is not possible, your CAS may graph a figure of the **contour lines** (3) of the function $u(x, y)$ and help you in understanding the solution.

Comparing (1) and (2), we see that (1) is an exact differential equation if there is some function $u(x, y)$ such that

$$\text{(a)}\quad \frac{\partial u}{\partial x} = M, \qquad \text{(b)}\quad \frac{\partial u}{\partial y} = N. \tag{4}$$

From this we can derive a formula for checking whether (1) is exact or not, as follows.

Let M and N be continuous and have continuous first partial derivatives in a region in the xy-plane whose boundary is a closed curve without self-intersections. Then by partial differentiation of (4) (see App. 3.2 for notation),

$$\frac{\partial M}{\partial y} = \frac{\partial^2 u}{\partial y\,\partial x},$$

$$\frac{\partial N}{\partial x} = \frac{\partial^2 u}{\partial x\,\partial y}.$$

By the assumption of continuity the two second partial derivaties are equal. Thus

$$\boxed{\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}.} \tag{5}$$

This condition is not only necessary but also sufficient for (1) to be an exact differential equation. (We shall prove this in Sec. 10.2 in another context. Some calculus books, for instance, [GenRef 12], also contain a proof.)

If (1) is exact, the function $u(x, y)$ can be found by inspection or in the following systematic way. From (4a) we have by integration with respect to x

$$u = \int M\,dx + k(y); \tag{6}$$

in this integration, y is to be regarded as a constant, and $k(y)$ plays the role of a "constant" of integration. To determine $k(y)$, we derive $\partial u/\partial y$ from (6), use (4b) to get dk/dy, and integrate dk/dy to get k. (See Example 1, below.)

Formula (6) was obtained from (4a). Instead of (4a) we may equally well use (4b). Then, instead of (6), we first have by integration with respect to y

$$u = \int N\,dy + l(x). \tag{6*}$$

To determine $l(x)$, we derive $\partial u/\partial x$ from (6*), use (4a) to get dl/dx, and integrate. We illustrate all this by the following typical examples.

EXAMPLE 1 **An Exact ODE**

Solve

$$\cos(x + y)\,dx + (3y^2 + 2y + \cos(x + y))\,dy = 0. \tag{7}$$

Solution. ***Step 1. Test for exactness.*** Our equation is of the form (1) with

$$M = \cos(x + y),$$
$$N = 3y^2 + 2y + \cos(x + y).$$

Thus

$$\frac{\partial M}{\partial y} = -\sin(x + y),$$
$$\frac{\partial N}{\partial x} = -\sin(x + y).$$

From this and (5) we see that (7) is exact.

Step 2. Implicit general solution. From (6) we obtain by integration

$$u = \int M\,dx + k(y) = \int \cos(x + y)\,dx + k(y) = \sin(x + y) + k(y). \tag{8}$$

To find $k(y)$, we differentiate this formula with respect to y and use formula (4b), obtaining

$$\frac{\partial u}{\partial y} = \cos(x + y) + \frac{dk}{dy} = N = 3y^2 + 2y + \cos(x + y).$$

Hence $dk/dy = 3y^2 + 2y$. By integration, $k = y^3 + y^2 + c^*$. Inserting this result into (8) and observing (3), we obtain the *answer*

$$u(x, y) = \sin(x + y) + y^3 + y^2 = c.$$

Step 3. Checking an implicit solution. We can check by differentiating the implicit solution $u(x, y) = c$ implicitly and see whether this leads to the given ODE (7):

$$du = \frac{\partial u}{\partial x}dx + \frac{\partial u}{\partial y}dy = \cos(x + y)\,dx + (\cos(x + y) + 3y^2 + 2y)\,dy = 0. \tag{9}$$

This completes the check. ■

EXAMPLE 2 **An Initial Value Problem**

Solve the initial value problem

$$(\cos y \sinh x + 1)\,dx - \sin y \cosh x\,dy = 0, \qquad y(1) = 2. \tag{10}$$

Solution. You may verify that the given ODE is exact. We find u. For a change, let us use (6*),

$$u = -\int \sin y \cosh x\,dy + l(x) = \cos y \cosh x + l(x).$$

From this, $\partial u/\partial x = \cos y \sinh x + dl/dx = M = \cos y \sinh x + 1$. Hence $dl/dx = 1$. By integration, $l(x) = x + c^*$. This gives the general solution $u(x, y) = \cos y \cosh x + x = c$. From the initial condition, $\cos 2 \cosh 1 + 1 = 0.358 = c$. Hence the answer is $\cos y \cosh x + x = 0.358$. Figure 17 shows the particular solutions for $c = 0, 0.358$ (thicker curve), 1, 2, 3. Check that the answer satisfies the ODE. (Proceed as in Example 1.) Also check that the initial condition is satisfied. ■

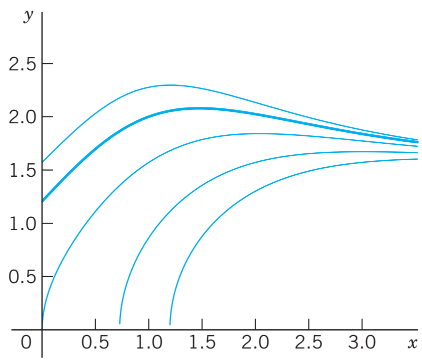


Fig. 17. Particular solutions in Example 2

EXAMPLE 3 **WARNING! Breakdown in the Case of Nonexactness**

The equation $-y\,dx + x\,dy = 0$ is not exact because $M = -y$ and $N = x$, so that in (5), $\partial M/\partial y = -1$ but $\partial N/\partial x = 1$. Let us show that in such a case the present method does not work. From (6),

$$u = \int M\,dx + k(y) = -xy + k(y), \qquad \text{hence} \qquad \frac{\partial u}{\partial y} = -x + \frac{dk}{dy}.$$

Now, $\partial u/\partial y$ should equal $N = x$, by (4b). However, this is impossible because $k(y)$ can depend only on y. Try (6*); it will also fail. Solve the equation by another method that we have discussed. ■

Reduction to Exact Form. Integrating Factors

The ODE in Example 3 is $-y\,dx + x\,dy = 0$. It is not exact. However, if we multiply it by $1/x^2$, we get an exact equation [check exactness by (5)!],

$$\frac{-y\,dx + x\,dy}{x^2} = -\frac{y}{x^2}dx + \frac{1}{x}dy = d\left(\frac{y}{x}\right) = 0. \tag{11}$$

Integration of (11) then gives the general solution $y/x = c = \text{const}$.

This example gives the idea. All we did was to multiply a given nonexact equation, say,

$$P(x, y)\,dx + Q(x, y)\,dy = 0, \tag{12}$$

by a function F that, in general, will be a function of both x and y. The result was an equation

$$FP\,dx + FQ\,dy = 0 \tag{13}$$

that is exact, so we can solve it as just discussed. Such a function $F(x, y)$ is then called an **integrating factor** of (12).

EXAMPLE 4 Integrating Factor

The integrating factor in (11) is $F = 1/x^2$. Hence in this case the exact equation (13) is

$$FP\,dx + FQ\,dy = \frac{-y\,dx + x\,dy}{x^2} = d\left(\frac{y}{x}\right) = 0. \qquad \text{Solution} \qquad \frac{y}{x} = c.$$

These are straight lines $y = cx$ through the origin. (Note that $x = 0$ is also a solution of $-y\,dx + x\,dy = 0$.)

It is remarkable that we can readily find other integrating factors for the equation $-y\,dx + x\,dy = 0$, namely, $1/y^2$, $1/(xy)$, and $1/(x^2 + y^2)$, because

$$\frac{-y\,dx + x\,dy}{y^2} = d\left(\frac{x}{y}\right), \quad \frac{-y\,dx + x\,dy}{xy} = -d\left(\ln\frac{x}{y}\right), \quad \frac{-y\,dx + x\,dy}{x^2 + y^2} = d\left(\arctan\frac{y}{x}\right). \tag{14}$$ ■

How to Find Integrating Factors

In simpler cases we may find integrating factors by inspection or perhaps after some trials, keeping (14) in mind. In the general case, the idea is the following.

For $M\,dx + N\,dy = 0$ the exactness condition (5) is $\partial M/\partial y = \partial N/\partial x$. Hence for (13), $FP\,dx + FQ\,dy = 0$, the exactness condition is

$$\frac{\partial}{\partial y}(FP) = \frac{\partial}{\partial x}(FQ). \tag{15}$$

By the product rule, with subscripts denoting partial derivatives, this gives

$$F_y P + F P_y = F_x Q + F Q_x.$$

In the general case, this would be complicated and useless. So we follow the ***Golden Rule:*** *If you cannot solve your problem, try to solve a simpler one*—the result may be useful (and may also help you later on). Hence we look for an integrating factor depending only on ***one*** variable: fortunately, in many practical cases, there are such factors, as we shall see. Thus, let $F = F(x)$. Then $F_y = 0$, and $F_x = F' = dF/dx$, so that (15) becomes

$$FP_y = F'Q + FQ_x.$$

Dividing by FQ and reshuffling terms, we have

$$\frac{1}{F}\frac{dF}{dx} = R, \qquad \text{where} \qquad R = \frac{1}{Q}\left(\frac{\partial P}{\partial y} - \frac{\partial Q}{\partial x}\right). \tag{16}$$

This proves the following theorem.

THEOREM 1

Integrating Factor $F(x)$

If (12) is such that the right side R of (16) depends only on x, then (12) has an integrating factor $F = F(x)$, which is obtained by integrating (16) and taking exponents on both sides.

$$F(x) = \exp \int R(x)\, dx. \tag{17}$$

Similarly, if $F^* = F^*(y)$, then instead of (16) we get

$$\frac{1}{F^*}\frac{dF^*}{dy} = R^*, \qquad \text{where} \qquad R^* = \frac{1}{P}\left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) \tag{18}$$

and we have the companion

THEOREM 2

Integrating Factor $F^*(y)$

If (12) is such that the right side R^* of (18) depends only on y, then (12) has an integrating factor $F^* = F^*(y)$, which is obtained from (18) in the form

$$F^*(y) = \exp \int R^*(y)\, dy. \tag{19}$$

EXAMPLE 5

Application of Theorems 1 and 2. Initial Value Problem

Using Theorem 1 or 2, find an integrating factor and solve the initial value problem

$$(e^{x+y} + ye^y)\, dx + (xe^y - 1)\, dy = 0, \quad y(0) = -1 \tag{20}$$

Solution. ***Step 1. Nonexactness.*** The exactness check fails:

$$\frac{\partial P}{\partial y} = \frac{\partial}{\partial y}(e^{x+y} + ye^y) = e^{x+y} + e^y + ye^y \qquad \text{but} \qquad \frac{\partial Q}{\partial x} = \frac{\partial}{\partial x}(xe^y - 1) = e^y.$$

Step 2. Integrating factor. General solution. Theorem 1 fails because R [the right side of (16)] depends on both x and y.

$$R = \frac{1}{Q}\left(\frac{\partial P}{\partial y} - \frac{\partial Q}{\partial x}\right) = \frac{1}{xe^y - 1}(e^{x+y} + e^y + ye^y - e^y).$$

Try Theorem 2. The right side of (18) is

$$R^* = \frac{1}{P}\left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) = \frac{1}{e^{x+y} + ye^y}(e^y - e^{x+y} - e^y - ye^y) = -1.$$

Hence (19) gives the integrating factor $F^*(y) = e^{-y}$. From this result and (20) you get the exact equation

$$(e^x + y)\, dx + (x - e^{-y})\, dy = 0.$$

Test for exactness; you will get 1 on both sides of the exactness condition. By integration, using (4a),

$$u = \int (e^x + y)\, dx = e^x + xy + k(y).$$

Differentiate this with respect to y and use (4b) to get

$$\frac{\partial u}{\partial y} = x + \frac{dk}{dy} = N = x - e^{-y}, \qquad \frac{dk}{dy} = -e^{-y}, \qquad k = e^{-y} + c^*.$$

Hence the general solution is

$$u(x, y) = e^x + xy + e^{-y} = c.$$

Setp 3. Particular solution. The initial condition $y(0) = -1$ gives $u(0, -1) = 1 + 0 + e = 3.72$. Hence the answer is $e^x + xy + e^{-y} = 1 + e = 3.72$. Figure 18 shows several particular solutions obtained as level curves of $u(x, y) = c$, obtained by a CAS, a convenient way in cases in which it is impossible or difficult to cast a solution into explicit form. Note the curve that (nearly) satisfies the initial condition.

Step 4. Checking. Check by substitution that the answer satisfies the given equation as well as the initial condition. ■

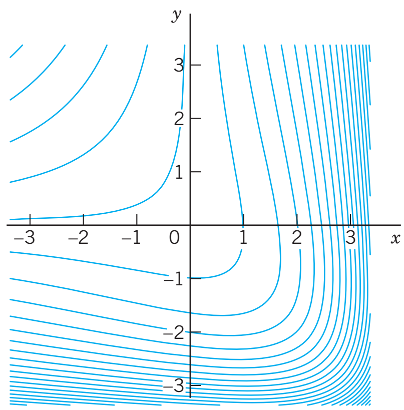


Fig. 18. Particular solutions in Example 5

PROBLEM SET 1.4

1–14 ODEs. INTEGRATING FACTORS

Test for exactness. If exact, solve. If not, use an integrating factor as given or obtained by inspection or by the theorems in the text. Also, if an initial condition is given, find the corresponding particular solution.

1. $2xy\, dx + x^2\, dy = 0$
2. $x^3 dx + y^3 dy = 0$
3. $\sin x \cos y\, dx + \cos x \sin y\, dy = 0$
4. $e^{3\theta}(dr + 3r\, d\theta) = 0$
5. $(x^2 + y^2)\, dx - 2xy\, dy = 0$
6. $3(y + 1)\, dx = 2x\, dy, \quad (y + 1)x^{-4}$
7. $2x \tan y\, dx + \sec^2 y\, dy = 0$
8. $e^x(\cos y\, dx - \sin y\, dy) = 0$
9. $e^{2x}(2 \cos y\, dx - \sin y\, dy) = 0, \quad y(0) = 0$
10. $y\, dx + [y + \tan(x + y)]\, dy = 0, \quad \cos(x + y)$
11. $2 \cosh x \cos y\, dx = \sinh x \sin y\, dy$
12. $(2xy\, dx + dy)e^{x^2} = 0, \quad y(0) = 2$
13. $e^{-y}\, dx + e^{-x}(-e^{-y} + 1)\, dy = 0, \quad F = e^{x+y}$
14. $(a + 1)y\, dx + (b + 1)x\, dy = 0, \quad y(1) = 1,$ $F = x^a y^b$
15. **Exactness.** Under what conditions for the constants a, b, k, l is $(ax + by)\, dx + (kx + ly)\, dy = 0$ exact? Solve the exact ODE.

16. TEAM PROJECT. Solution by Several Methods. Show this as indicated. Compare the amount of work.

(a) $e^y(\sinh x\,dx + \cosh x\,dy) = 0$ as an exact ODE and by separation.

(b) $(1 + 2x)\cos y\,dx + dy/\cos y = 0$ by Theorem 2 and by separation.

(c) $(x^2 + y^2)\,dx - 2xy\,dy = 0$ by Theorem 1 or 2 and by separation with $v = y/x$.

(d) $3x^2 y\,dx + 4x^3\,dy = 0$ by Theorems 1 and 2 and by separation.

(e) Search the text and the problems for further ODEs that can be solved by more than one of the methods discussed so far. Make a list of these ODEs. Find further cases of your own.

17. WRITING PROJECT. Working Backward. Working backward from the solution to the problem is useful in many areas. Euler, Lagrange, and other great masters did it. To get additional insight into the idea of integrating factors, start from a $u(x, y)$ of your choice, find $du = 0$, destroy exactness by division by some $F(x, y)$, and see what ODE's solvable by integrating factors you can get. Can you proceed systematically, beginning with the simplest $F(x, y)$?

18. CAS PROJECT. Graphing Particular Solutions. Graph particular solutions of the following ODE, proceeding as explained.

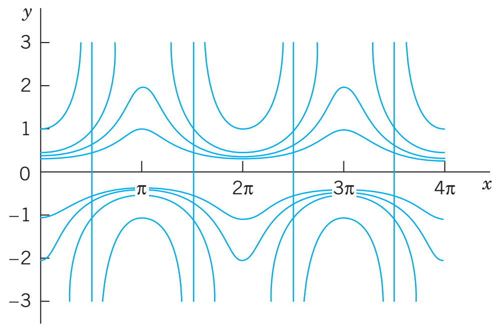
$$dy - y^2 \sin x\,dx = 0. \tag{21}$$

(a) Show that (21) is not exact. Find an integrating factor using either Theorem 1 or 2. Solve (21).

(b) Solve (21) by separating variables. Is this simpler than (a)?

(c) Graph the seven particular solutions satisfying the following initial conditions $y(0) = 1$, $y(\pi/2) = \pm\frac{1}{2}$, $\pm\frac{2}{3}$, ± 1 (see figure below).

(d) Which solution of (21) do we not get in (a) or (b)?



Particular solutions in CAS Project 18

1.5 Linear ODEs. Bernoulli Equation. Population Dynamics

Linear ODEs or ODEs that can be transformed to linear form are models of various phenomena, for instance, in physics, biology, population dynamics, and ecology, as we shall see. A first-order ODE is said to be **linear** if it can be brought into the form

$$y' + p(x)y = r(x), \tag{1}$$

by algebra, and **nonlinear** if it cannot be brought into this form.

The defining feature of the linear ODE (1) is that it is linear in both the unknown function y and its derivative $y' = dy/dx$, whereas p and r may be ***any*** given functions of x. If in an application the independent variable is time, we write t instead of x.

If the first term is $f(x)y'$ (instead of y'), divide the equation by $f(x)$ to get the **standard form** (1), with y' as the first term, which is practical.

For instance, $y' \cos x + y \sin x = x$ is a linear ODE, and its standard form is $y' + y \tan x = x \sec x$.

The function $r(x)$ on the right may be a force, and the solution $y(x)$ a displacement in a motion or an electrical current or some other physical quantity. In engineering, $r(x)$ is frequently called the **input**, and $y(x)$ is called the **output** or the *response* to the input (and, if given, to the initial condition).

Homogeneous Linear ODE. We want to solve (1) in some interval $a < x < b$, call it J, and we begin with the simpler special case that $r(x)$ is zero for all x in J. (This is sometimes written $r(x) \equiv 0$.) Then the ODE (1) becomes

$$y' + p(x)y = 0 \tag{2}$$

and is called **homogeneous**. By separating variables and integrating we then obtain

$$\frac{dy}{y} = -p(x)\,dx, \qquad \text{thus} \qquad \ln|y| = -\int p(x)\,dx + c^*.$$

Taking exponents on both sides, we obtain the general solution of the homogeneous ODE (2),

$$y(x) = ce^{-\int p(x)dx} \qquad (c = \pm e^{c^*} \quad \text{when} \quad y \gtrless 0); \tag{3}$$

here we may also choose $c = 0$ and obtain the **trivial solution** $y(x) = 0$ for all x in that interval.

Nonhomogeneous Linear ODE. We now solve (1) in the case that $r(x)$ in (1) is not everywhere zero in the interval J considered. Then the ODE (1) is called **nonhomogeneous**. It turns out that in this case, (1) has a pleasant property; namely, it has an integrating factor depending only on x. We can find this factor $F(x)$ by Theorem 1 in the previous section or we can proceed directly, as follows. We multiply (1) by $F(x)$, obtaining

$$Fy' + pFy = rF. \tag{1*}$$

The left side is the derivative $(Fy)' = F'y + Fy'$ of the product Fy if

$$pFy = F'y, \qquad \text{thus} \qquad pF = F'.$$

By separating variables, $dF/F = p\,dx$. By integration, writing $h = \int p\,dx$,

$$\ln|F| = h = \int p\,dx, \qquad \text{thus} \qquad F = e^h.$$

With this F and $h' = p$, Eq. (1*) becomes

$$e^h y' + h'e^h y = e^h y' + (e^h)'y = (e^h y)' = re^h.$$

By integration,

$$e^h y = \int e^h r\,dx + c.$$

Dividing by e^h, we obtain the desired solution formula

$$y(x) = e^{-h}\left(\int e^h r\,dx + c\right), \qquad h = \int p(x)\,dx. \tag{4}$$

This reduces solving (1) to the generally simpler task of evaluating integrals. For ODEs for which this is still difficult, you may have to use a numeric method for integrals from Sec. 19.5 or for the ODE itself from Sec. 21.1. We mention that h has nothing to do with $h(x)$ in Sec. 1.1 and that the constant of integration in h does not matter; see Prob. 2.

The structure of (4) is interesting. The only quantity depending on a given initial condition is c. Accordingly, writing (4) as a sum of two terms,

$$y(x) = e^{-h} \int e^{h} r \, dx + ce^{-h}, \tag{4*}$$

we see the following:

$$\text{Total Output} = \text{Response to the Input } r + \text{Response to the Initial Data.} \tag{5}$$

EXAMPLE 1 **First-Order ODE, General Solution, Initial Value Problem**

Solve the initial value problem

$$y' + y \tan x = \sin 2x, \qquad y(0) = 1.$$

Solution. Here $p = \tan x$, $r = \sin 2x = 2 \sin x \cos x$, and

$$h = \int p \, dx = \int \tan x \, dx = \ln |\sec x|.$$

From this we see that in (4),

$$e^{h} = \sec x, \qquad e^{-h} = \cos x, \qquad e^{h} r = (\sec x)(2 \sin x \cos x) = 2 \sin x,$$

and the general solution of our equation is

$$y(x) = \cos x \left(2 \int \sin x \, dx + c \right) = c \cos x - 2 \cos^2 x.$$

From this and the initial condition, $1 = c \cdot 1 - 2 \cdot 1^2$; thus $c = 3$ and the solution of our initial value problem is $y = 3 \cos x - 2 \cos^2 x$. Here $3 \cos x$ is the response to the initial data, and $-2 \cos^2 x$ is the response to the input $\sin 2x$. ■

EXAMPLE 2 **Electric Circuit**

Model the ***RL*-circuit** in Fig. 19 and solve the resulting ODE for the current $I(t)$ A (amperes), where t is time. Assume that the circuit contains as an **EMF** $E(t)$ (electromotive force) a battery of $E = 48$ V (volts), which is constant, a ***resistor*** of $R = 11\ \Omega$ (ohms), and an ***inductor*** of $L = 0.1$ H (henrys), and that the current is initially zero.

Physical Laws. A current I in the circuit causes a **voltage drop** RI across the resistor **(Ohm's law)** and a voltage drop $LI' = L\, dI/dt$ across the conductor, and the sum of these two voltage drops equals the EMF **(Kirchhoff's Voltage Law, KVL)**.

Remark. In general, KVL states that "The voltage (the electromotive force EMF) impressed on a closed loop is equal to the sum of the voltage drops across all the other elements of the loop." For Kirchoff's Current Law (KCL) and historical information, see footnote 7 in Sec. 2.9.

Solution. According to these laws the model of the RL-circuit is $LI' + RI = E(t)$, in standard form

$$I' + \frac{R}{L} I = \frac{E(t)}{L}. \tag{6}$$

We can solve this linear ODE by (4) with $x = t$, $y = I$, $p = R/L$, $h = (R/L)t$, obtaining the general solution

$$I = e^{-(R/L)t}\left(\int e^{(R/L)t}\frac{E(t)}{L}dt + c\right).$$

By integration,

$$I = e^{-(R/L)t}\left(\frac{E}{L}\frac{e^{(R/L)t}}{R/L} + c\right) = \frac{E}{R} + ce^{-(R/L)t}. \tag{7}$$

In our case, $R/L = 11/0.1 = 110$ and $E(t) = 48/0.1 = 480 = \text{const}$; thus,

$$I = \tfrac{48}{11} + ce^{-110t}.$$

In modeling, one often gets better insight into the nature of a solution (and smaller roundoff errors) by inserting given numeric data only near the end. Here, the general solution (7) shows that the current approaches the limit $E/R = 48/11$ faster the larger R/L is, in our case, $R/L = 11/0.1 = 110$, and the approach is very fast, from below if $I(0) < 48/11$ or from above if $I(0) > 48/11$. If $I(0) = 48/11$, the solution is constant (48/11 A). See Fig. 19.

The initial value $I(0) = 0$ gives $I(0) = E/R + c = 0$, $c = -E/R$ and the particular solution

$$I = \frac{E}{R}(1 - e^{-(R/L)t}), \qquad \text{thus} \qquad I = \frac{48}{11}(1 - e^{-110t}). \tag{8}$$

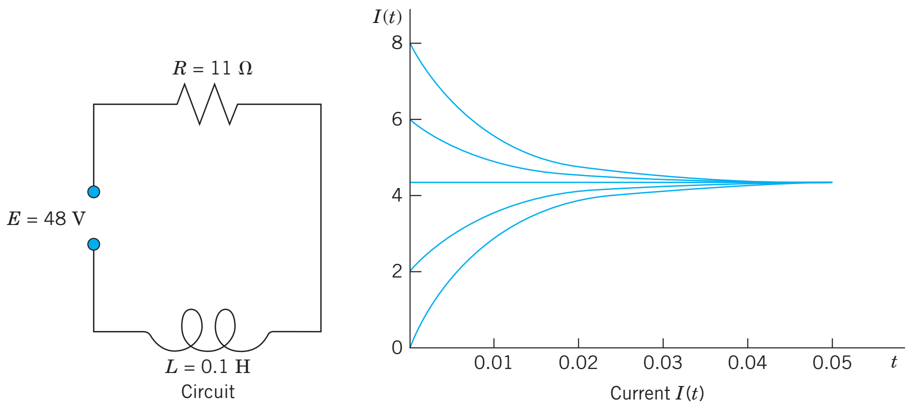


Fig. 19. *RL*-circuit

EXAMPLE 3 Hormone Level

Assume that the level of a certain hormone in the blood of a patient varies with time. Suppose that the time rate of change is the difference between a sinusoidal input of a 24-hour period from the thyroid gland and a continuous removal rate proportional to the level present. Set up a model for the hormone level in the blood and find its general solution. Find the particular solution satisfying a suitable initial condition.

Solution. ***Step 1. Setting up a model.*** Let $y(t)$ be the hormone level at time t. Then the removal rate is $Ky(t)$. The input rate is $A + B\cos\omega t$, where $\omega = 2\pi/24 = \pi/12$ and A is the average input rate; here $A \geqq B$ to make the input rate nonnegative. The constants A, B, K can be determined from measurements. Hence the model is the linear ODE

$$y'(t) = \text{In} - \text{Out} = A + B\cos\omega t - Ky(t), \qquad \text{thus} \qquad y' + Ky = A + B\cos\omega t.$$

The initial condition for a particular solution y_{part} is $y_{\text{part}}(0) = y_0$ with $t = 0$ suitably chosen, for example, 6:00 A.M.

Step 2. General solution. In (4) we have $p = K = \text{const}$, $h = Kt$, and $r = A + B\cos\omega t$. Hence (4) gives the general solution (evaluate $\int e^{Kt}\cos\omega t\, dt$ by integration by parts)

$$
\begin{aligned}
y(t) &= e^{-Kt}\int e^{Kt}\Big(A + B\cos\omega t\Big)\,dt + ce^{-Kt} \\
&= e^{-Kt}e^{Kt}\left[\frac{A}{K} + \frac{B}{K^2+\omega^2}\Big(K\cos\omega t + \omega\sin\omega t\Big)\right] + ce^{-Kt} \\
&= \frac{A}{K} + \frac{B}{K^2+(\pi/12)^2}\left(K\cos\frac{\pi t}{12} + \frac{\pi}{12}\sin\frac{\pi t}{12}\right) + ce^{-Kt}.
\end{aligned}
$$

The last term decreases to 0 as t increases, practically after a short time and regardless of c (that is, of the initial condition). The other part of $y(t)$ is called the **steady-state solution** because it consists of constant and periodic terms. The entire solution is called the **transient-state solution** because it models the transition from rest to the steady state. These terms are used quite generally for physical and other systems whose behavior depends on time.

Step 3. Particular solution. Setting $t = 0$ in $y(t)$ and choosing $y_0 = 0$, we have

$$
y(0) = \frac{A}{K} + \frac{B}{K^2+(\pi/12)^2}\,\frac{u}{\pi}\,K + c = 0, \qquad \text{thus} \qquad c = -\frac{A}{K} - \frac{KB}{K^2+(\pi/12)^2}.
$$

Inserting this result into $y(t)$, we obtain the particular solution

$$
y_{\text{part}}(t) = \frac{A}{K} + \frac{B}{K^2+(\pi/12)^2}\left(K\cos\frac{\pi t}{12} + \frac{\pi}{12}\sin\frac{\pi t}{12}\right) - \left(\frac{A}{K} + \frac{KB}{K^2+(\pi/12)^2}\right)e^{-K}
$$

with the steady-state part as before. To plot y_{part} we must specify values for the constants, say, $A = B = 1$ and $K = 0.05$. Figure 20 shows this solution. Notice that the transition period is relatively short (although K is small), and the curve soon looks sinusoidal; this is the response to the input $A + B\cos\left(\frac{1}{12}\pi t\right) = 1 + \cos\left(\frac{1}{12}\pi t\right)$. ■

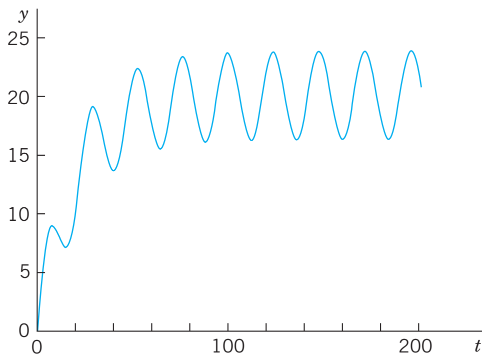


Fig. 20. Particular solution in Example 3

Reduction to Linear Form. Bernoulli Equation

Numerous applications can be modeled by ODEs that are nonlinear but can be transformed to linear ODEs. One of the most useful ones of these is the **Bernoulli equation**[7]

$$
y' + p(x)y = g(x)y^a \qquad (a \text{ any real number}). \tag{9}
$$

[7]JAKOB BERNOULLI (1654–1705), Swiss mathematician, professor at Basel, also known for his contribution to elasticity theory and mathematical probability. The method for solving Bernoulli's equation was discovered by Leibniz in 1696. Jakob Bernoulli's students included his nephew NIKLAUS BERNOULLI (1687–1759), who contributed to probability theory and infinite series, and his youngest brother JOHANN BERNOULLI (1667–1748), who had profound influence on the development of calculus, became Jakob's successor at Basel, and had among his students GABRIEL CRAMER (see Sec. 7.7) and LEONHARD EULER (see Sec. 2.5). His son DANIEL BERNOULLI (1700–1782) is known for his basic work in fluid flow and the kinetic theory of gases.

If $a = 0$ or $a = 1$, Equation (9) is linear. Otherwise it is nonlinear. Then we set

$$u(x) = [y(x)]^{1-a}.$$

We differentiate this and substitute y' from (9), obtaining

$$u' = (1 - a)y^{-a}y' = (1 - a)y^{-a}(gy^a - py).$$

Simplification gives

$$u' = (1 - a)(g - py^{1-a}),$$

where $y^{1-a} = u$ on the right, so that we get the linear ODE

$$u' + (1 - a)pu = (1 - a)g. \tag{10}$$

For further ODEs reducible to linear form, see lnce's classic [A11] listed in App. 1. See also Team Project 30 in Problem Set 1.5.

EXAMPLE 4 Logistic Equation

Solve the following Bernoulli equation, known as the **logistic equation** (or **Verhulst equation**[8]):

$$y' = Ay - By^2 \tag{11}$$

Solution. Write (11) in the form (9), that is,

$$y' - Ay = -By^2$$

to see that $a = 2$, so that $u = y^{1-a} = y^{-1}$. Differentiate this u and substitute y' from (11),

$$u' = -y^{-2}y' = -y^{-2}(Ay - By^2) = B - Ay^{-1}.$$

The last term is $-Ay^{-1} = -Au$. Hence we have obtained the linear ODE

$$u' + Au = B.$$

The general solution is [by (4)]

$$u = ce^{-At} + B/A.$$

Since $u = 1/y$, this gives the general solution of (11),

$$y = \frac{1}{u} = \frac{1}{ce^{-At} + B/A} \tag{12}$$ (Fig. 21)

Directly from (11) we see that $y \equiv 0$ ($y(t) = 0$ for all t) is also a solution. ■

[8]PIERRE-FRANÇOIS VERHULST, Belgian statistician, who introduced Eq. (8) as a model for human population growth in 1838.

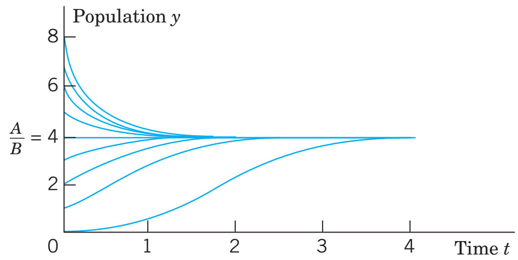


Fig. 21. Logistic population model. Curves (9) in Example 4 with $A/B = 4$

Population Dynamics

The logistic equation (11) plays an important role in **population dynamics**, a field that models the evolution of populations of plants, animals, or humans over time t. If $B = 0$, then (11) is $y' = dy/dt = Ay$. In this case its solution (12) is $y = (1/c)e^{At}$ and gives exponential growth, as for a small population in a large country (the United States in early times!). This is called ***Malthus's law***. (See also Example 3 in Sec. 1.1.)

The term $-By^2$ in (11) is a "braking term" that prevents the population from growing without bound. Indeed, if we write $y' = Ay[1 - (B/A)y]$, we see that if $y < A/B$, then $y' > 0$, so that an initially small population keeps growing as long as $y < A/B$. But if $y > A/B$, then $y' < 0$ and the population is decreasing as long as $y > A/B$. The limit is the same in both cases, namely, A/B. See Fig. 21.

We see that in the logistic equation (11) the independent variable t does not occur explicitly. An ODE $y' = f(t, y)$ in which t does not occur explicitly is of the form

$$y' = f(y) \tag{13}$$

and is called an **autonomous ODE**. Thus the logistic equation (11) is autonomous.

Equation (13) has constant solutions, called **equilibrium solutions** or **equilibrium points**. These are determined by the zeros of $f(y)$, because $f(y) = 0$ gives $y' = 0$ by (13); hence $y = \text{const}$. These zeros are known as **critical points** of (13). An equilibrium solution is called **stable** if solutions close to it for some t remain close to it for all further t. It is called **unstable** if solutions initially close to it do not remain close to it as t increases. For instance, $y = 0$ in Fig. 21 is an unstable equilibrium solution, and $y = 4$ is a stable one. Note that (11) has the critical points $y = 0$ and $y = A/B$.

EXAMPLE 5 Stable and Unstable Equilibrium Solutions. "Phase Line Plot"

The ODE $y' = (y - 1)(y - 2)$ has the stable equilibrium solution $y_1 = 1$ and the unstable $y_2 = 2$, as the direction field in Fig. 22 suggests. The values y_1 and y_2 are the zeros of the parabola $f(y) = (y - 1)(y - 2)$ in the figure. Now, since the ODE is autonomous, we can "condense" the direction field to a "phase line plot" giving y_1 and y_2, and the direction (upward or downward) of the arrows in the field, and thus giving information about the stability or instability of the equilibrium solutions. ■

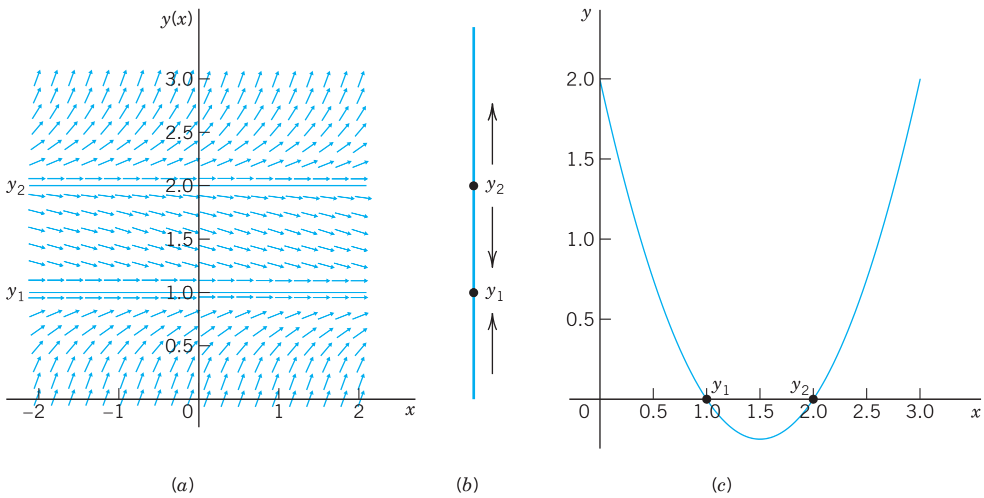


Fig. 22. Example 5. (A) Direction field. (B) "Phase line". (C) Parabola $f(y)$

A few further population models will be discussed in the problem set. For some more details of population dynamics, see C. W. Clark. *Mathematical Bioeconomics*: *The Mathematics of Conservation* 3rd ed. Hoboken, NJ, Wiley, 2010.

Further applications of linear ODEs follow in the next section.

PROBLEM SET 1.5

1. CAUTION! Show that $e^{-\ln x} = 1/x$ (not $-x$) and $e^{-\ln(\sec x)} = \cos x$.

2. Integration constant. Give a reason why in (4) you may choose the constant of integration in $\int p\,dx$ to be zero.

3–13 GENERAL SOLUTION. INITIAL VALUE PROBLEMS

Find the general solution. If an initial condition is given, find also the corresponding particular solution and graph or sketch it. (Show the details of your work.)

3. $y' - y = 5.2$

4. $y' = 2y - 4x$

5. $y' + ky = e^{-kx}$

6. $y' + 2y = 4\cos 2x, \quad y(\frac{1}{4}\pi) = 3$

7. $xy' = 2y + x^3e^x$

8. $y' + y\tan x = e^{-0.01x}\cos x, \quad y(0) = 0$

9. $y' + y\sin x = e^{\cos x}, \quad y(0) = -2.5$

10. $y'\cos x + (3y - 1)\sec x = 0, \quad y(\frac{1}{4}\pi) = 4/3$

11. $y' = (y - 2)\cot x$

12. $xy' + 4y = 8x^4, \quad y(1) = 2$

13. $y' = 6(y - 2.5)\tanh 1.5x$

14. CAS EXPERIMENT. (a) Solve the ODE $y' - y/x = -x^{-1}\cos(1/x)$. Find an initial condition for which the arbitrary constant becomes zero. Graph the resulting particular solution, experimenting to obtain a good figure near $x = 0$.

(b) Generalizing (a) from $n = 1$ to arbitrary n, solve the ODE $y' - ny/x = -x^{n-2}\cos(1/x)$. Find an initial condition as in (a) and experiment with the graph.

15–20 GENERAL PROPERTIES OF LINEAR ODEs

These properties are of practical and theoretical importance because they enable us to obtain new solutions from given ones. Thus in modeling, whenever possible, we prefer linear ODEs over nonlinear ones, which have no similar properties.

Show that nonhomogeneous linear ODEs (1) and homogeneous linear ODEs (2) have the following properties. Illustrate each property by a calculation for two or three equations of your choice. Give proofs.

15. The sum $y_1 + y_2$ of two solutions y_1 and y_2 of the homogeneous equation (2) is a solution of (2), and so is a scalar multiple ay_1 for any constant a. These properties are not true for (1)!

16. $y = 0$ (that is, $y(x) = 0$ for all x, also written $y(x) \equiv 0$) is a solution of (2) [not of (1) if $r(x) \neq 0$!], called the **trivial solution**.

17. The sum of a solution of (1) and a solution of (2) is a solution of (1).

18. The difference of two solutions of (1) is a solution of (2).

19. If y_1 is a solution of (1), what can you say about cy_1?

20. If y_1 and y_2 are solutions of $y_1' + py_1 = r_1$ and $y_2' + py_2 = r_2$, respectively (with the same p!), what can you say about the sum $y_1 + y_2$?

21. Variation of parameter. Another method of obtaining (4) results from the following idea. Write (3) as cy^*, where y^* is the exponential function, which is a solution of the homogeneous linear ODE $y^{*\prime} + py^* = 0$. Replace the arbitrary constant c in (3) with a function u to be determined so that the resulting function $y = uy^*$ is a solution of the nonhomogeneous linear ODE $y' + py = r$.

22–28 NONLINEAR ODEs

Using a method of this section or separating variables, find the general solution. If an initial condition is given, find also the particular solution and sketch or graph it.

22. $y' + y = y^2, \quad y(0) = -\frac{1}{3}$

23. $y' + xy = xy^{-1}, \quad y(0) = 3$

24. $y' + y = -x/y$

25. $y' = 3.2y - 10y^2$

26. $y' = (\tan y)/(x - 1), \quad y(0) = \frac{1}{2}\pi$

27. $y' = 1/(6e^y - 2x)$

28. $2xyy' + (x - 1)y^2 = x^2e^x \quad (\text{Set } y^2 = z)$

29. REPORT PROJECT. Transformation of ODEs. We have transformed ODEs to separable form, to exact form, and to linear form. The purpose of such transformations is an extension of solution methods to larger classes of ODEs. Describe the key idea of each of these transformations and give three typical examples of your choice for each transformation. Show each step (not just the transformed ODE).

30. TEAM PROJECT. Riccati Equation. Clairaut Equation. Singular Solution.
A **Riccati equation** is of the form

$$y' + p(x)y = g(x)y^2 + h(x). \tag{14}$$

A **Clairaut equation** is of the form

$$y = xy' + g(y'). \tag{15}$$

(a) Apply the transformation $y = Y + 1/u$ to the Riccati equation (14), where Y is a solution of (14), and obtain for u the linear ODE $u' + (2Yg - p)u = -g$. Explain the effect of the transformation by writing it as $y = Y + v$, $v = 1/u$.

(b) Show that $y = Y = x$ is a solution of the ODE $y' - (2x^3 + 1)\, y = -x^2y^2 - x^4 - x + 1$ and solve this Riccati equation, showing the details.

(c) Solve the Clairaut equation $y'^2 - xy' + y = 0$ as follows. Differentiate it with respect to x, obtaining $y''(2y' - x) = 0$. Then solve (A) $y'' = 0$ and (B) $2y' - x = 0$ separately and substitute the two solutions (a) and (b) of (A) and (B) into the given ODE. Thus obtain (a) a general solution (straight lines) and (b) a parabola for which those lines (a) are tangents (Fig. 6 in Prob. Set 1.1); so (b) is the envelope of (a). Such a solution (b) that cannot be obtained from a general solution is called a **singular solution**.

(d) Show that the Clairaut equation (15) has as solutions a family of straight lines $y = cx + g(c)$ and a singular solution determined by $g'(s) = -x$, where $s = y'$, that forms the envelope of that family.

31–40 MODELING. FURTHER APPLICATIONS

31. Newton's law of cooling. If the temperature of a cake is 300°F when it leaves the oven and is 200°F ten minutes later, when will it be practically equal to the room temperature of 60°F, say, when will it be 61°F?

32. Heating and cooling of a building. Heating and cooling of a building can be modeled by the ODE

$$T' = k_1(T - T_a) + k_2(T - T_\omega) + P,$$

where $T = T(t)$ is the temperature in the building at time t, T_a the outside temperature, T_w the temperature wanted in the building, and P the rate of increase of T due to machines and people in the building, and k_1 and k_2 are (negative) constants. Solve this ODE, assuming $P = \text{const}$, $T_w = \text{const}$, and T_a varying sinusoidally over 24 hours, say, $T_a = A - C\cos(2\pi/24)t$. Discuss the effect of each term of the equation on the solution.

33. Drug injection. Find and solve the model for drug injection into the bloodstream if, beginning at $t = 0$, a constant amount A g/min is injected and the drug is simultaneously removed at a rate proportional to the amount of the drug present at time t.

34. Epidemics. A model for the spread of contagious diseases is obtained by assuming that the rate of spread is proportional to the number of contacts between infected and noninfected persons, who are assumed to move freely among each other. Set up the model. Find the equilibrium solutions and indicate their stability or instability. Solve the ODE. Find the limit of the proportion of infected persons as $t \to \infty$ and explain what it means.

35. Lake Erie. Lake Erie has a water volume of about 450 km^3 and a flow rate (in and out) of about 175 km^2

per year. If at some instant the lake has pollution concentration $p = 0.04\%$, how long, approximately, will it take to decrease it to $p/2$, assuming that the inflow is much cleaner, say, it has pollution concentration $p/4$, and the mixture is uniform (an assumption that is only imperfectly true)? First guess.

36. Harvesting renewable resources. Fishing. Suppose that the population $y(t)$ of a certain kind of fish is given by the logistic equation (11), and fish are caught at a rate Hy proportional to y. Solve this so-called *Schaefer model.* Find the equilibrium solutions y_1 and y_2 (> 0) when $H < A$. The expression $Y = Hy_2$ is called the **equilibrium harvest** or **sustainable yield** *corresponding to H.* Why?

37. Harvesting. In Prob. 36 find and graph the solution satisfying $y(0) = 2$ when (for simplicity) $A = B = 1$ and $H = 0.2$. What is the limit? What does it mean? What if there were no fishing?

38. Intermittent harvesting. In Prob. 36 assume that you fish for 3 years, then fishing is banned for the next 3 years. Thereafter you start again. And so on. This is called *intermittent harvesting.* Describe qualitatively how the population will develop if intermitting is continued periodically. Find and graph the solution for the first 9 years, assuming that $A = B = 1$, $H = 0.2$, and $y(0) = 2$.

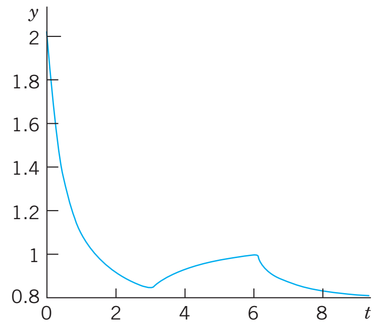


Fig. 23. Fish population in Problem 38

39. Extinction vs. unlimited growth. If in a population $y(t)$ the death rate is proportional to the population, and the birth rate is proportional to the chance encounters of meeting mates for reproduction, what will the model be? Without solving, find out what will eventually happen to a small initial population. To a large one. Then solve the model.

40. Air circulation. In a room containing 20,000 ft^3 of air, 600 ft^3 of fresh air flows in per minute, and the mixture (made practically uniform by circulating fans) is exhausted at a rate of 600 cubic feet per minute (cfm). What is the amount of fresh air $y(t)$ at any time if $y(0) = 0$? After what time will 90% of the air be fresh?

1.6 Orthogonal Trajectories. *Optional*

An important type of problem in physics or geometry is to find a family of curves that intersects a given family of curves at right angles. The new curves are called **orthogonal trajectories** of the given curves (and conversely). Examples are curves of equal temperature **(isotherms)** and curves of heat flow, curves of equal altitude **(contour lines)** on a map and curves of steepest descent on that map, curves of equal potential (**equipotential curves**, curves of equal voltage—the ellipses in Fig. 24) and curves of electric force (the parabolas in Fig. 24).

Here the **angle of intersection** between two curves is defined to be the angle between the tangents of the curves at the intersection point. *Orthogonal* is another word for *perpendicular.*

In many cases orthogonal trajectories can be found using ODEs. In general, if we consider $G(x, y, c) = 0$ to be a given family of curves in the xy-plane, then each value of c gives a particular curve. Since c is one parameter, such a family is called a **one-parameter family of curves**.

In detail, let us explain this method by a family of ellipses

$$\tfrac{1}{2}x^2 + y^2 = c \qquad (c > 0) \tag{1}$$

and illustrated in Fig. 24. We assume that this family of ellipses represents electric equipotential curves between the two black ellipses (equipotential surfaces between two elliptic cylinders in space, of which Fig. 24 shows a cross-section). We seek the orthogonal trajectories, the curves of electric force. Equation (1) is a *one-parameter family* with *parameter* c. Each value of c (> 0) corresponds to one of these ellipses.

Step 1. Find an ODE for which the given family is a general solution. Of course, this ODE must no longer contain the parameter c. Differentiating (1), we have $x + 2yy' = 0$. Hence the ODE of the given curves is

$$y' = f(x, y) = -\frac{x}{2y}. \tag{2}$$

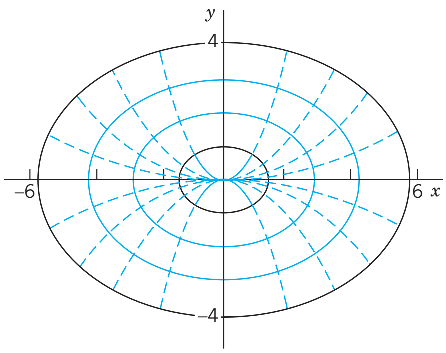


Fig. 24. Electrostatic field between two ellipses (elliptic cylinders in space): Elliptic equipotential curves (equipotential surfaces) and orthogonal trajectories (parabolas)

Step 2. Find an ODE for the orthogonal trajectories $\tilde{y} = \tilde{y}(x)$. This ODE is

$$\tilde{y}' = -\frac{1}{f(x, \tilde{y})} = +\frac{2\tilde{y}}{x} \tag{3}$$

with the same f as in (2). Why? Well, a given curve passing through a point (x_0, y_0) has slope $f(x_0, y_0)$ at that point, by (2). The trajectory through (x_0, y_0) has slope $-1/f(x_0, y_0)$ by (3). The product of these slopes is -1, as we see. From calculus it is known that this is the condition for orthogonality (perpendicularity) of two straight lines (the tangents at (x_0, y_0)), hence of the curve and its orthogonal trajectory at (x_0, y_0).

Step 3. Solve (3) by separating variables, integrating, and taking exponents:

$$\frac{d\tilde{y}}{\tilde{y}} = 2\frac{dx}{x}, \qquad \ln|\tilde{y}| = 2\ln x + c, \qquad \tilde{y} = c^* x^2.$$

This is the family of orthogonal trajectories, the quadratic parabolas along which electrons or other charged particles (of very small mass) would move in the electric field between the black ellipses (elliptic cylinders).

PROBLEM SET 1.6

1–3 FAMILIES OF CURVES

Represent the given family of curves in the form $G(x, y; c) = 0$ and sketch some of the curves.

1. All ellipses with foci -3 and 3 on the x-axis.

2. All circles with centers on the cubic parabola $y = x^3$ and passing through the origin $(0, 0)$.

3. The catenaries obtained by translating the catenary $y = \cosh x$ in the direction of the straight line $y = x$.

4–10 ORTHOGONAL TRAJECTORIES (OTs)

Sketch or graph some of the given curves. Guess what their OTs may look like. Find these OTs.

4. $y = x^2 + c$

5. $y = cx$

6. $xy = c$

7. $y = c/x^2$

8. $y = \sqrt{x + c}$

9. $y = ce^{-x^2}$

10. $x^2 + (y - c)^2 = c^2$

11–16 APPLICATIONS, EXTENSIONS

11. Electric field. Let the electric **equipotential lines** (curves of constant potential) between two concentric cylinders with the z-axis in space be given by $u(x, y) = x^2 + y^2 = c$ (these are circular cylinders in the xyz-space). Using the method in the text, find their orthogonal trajectories (the curves of electric force).

12. Electric field. The lines of electric force of two opposite charges of the same strength at $(-1, 0)$ and $(1, 0)$ are the circles through $(-1, 0)$ and $(1, 0)$. Show that these circles are given by $x^2 + (y - c)^2 = 1 + c^2$. Show that the **equipotential lines** (which are orthogonal trajectories of those circles) are the circles given by $(x + c^*)^2 + \tilde{y}^2 = c^{*2} - 1$ (dashed in Fig. 25).

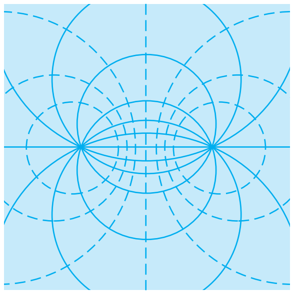


Fig. 25. Electric field in Problem 12

13. Temperature field. Let the **isotherms** (curves of constant temperature) in a body in the upper half-plane $y > 0$ be given by $4x^2 + 9y^2 = c$. Find the orthogonal trajectories (the curves along which heat will flow in regions filled with heat-conducting material and free of heat sources or heat sinks).

14. Conic sections. Find the conditions under which the orthogonal trajectories of families of ellipses $x^2/a^2 + y^2/b^2 = c$ are again conic sections. Illustrate your result graphically by sketches or by using your CAS. What happens if $a \to 0$? If $b \to 0$?

15. Cauchy–Riemann equations. Show that for a family $u(x, y) = c =$ const the orthogonal trajectories $v(x, y) = c^* =$ const can be obtained from the following *Cauchy–Riemann equations* (which are basic in complex analysis in Chap. 13) and use them to find the orthogonal trajectories of $e^x \sin y =$ const. (Here, subscripts denote partial derivatives.)

$$u_x = v_y, \qquad u_y = -v_x$$

16. Congruent OTs. If $y' = f(x)$ with f independent of y, show that the curves of the corresponding family are congruent, and so are their OTs.

1.7 Existence and Uniqueness of Solutions for Initial Value Problems

The initial value problem

$$|y'| + |y| = 0, \qquad y(0) = 1$$

has no solution because $y = 0$ (that is, $y(x) = 0$ for all x) is the only solution of the ODE. The initial value problem

$$y' = 2x, \qquad y(0) = 1$$

has precisely one solution, namely, $y = x^2 + 1$. The initial value problem

$$xy' = y - 1, \qquad y(0) = 1$$

has infinitely many solutions, namely, $y = 1 + cx$, where c is an arbitrary constant because $y(0) = 1$ for all c.

From these examples we see that an **initial value problem**

$$y' = f(x, y), \qquad y(x_0) = y_0 \tag{1}$$

may have no solution, precisely one solution, or more than one solution. This fact leads to the following two fundamental questions.

Problem of Existence

Under what conditions does an initial value problem of the form (1) *have at least one solution (hence one or several solutions)?*

Problem of Uniqueness

Under what conditions does that problem have at most one solution (hence excluding the case that is has more than one solution)?

Theorems that state such conditions are called **existence theorems** and **uniqueness theorems**, respectively.

Of course, for our simple examples, we need no theorems because we can solve these examples by inspection; however, for complicated ODEs such theorems may be of considerable practical importance. Even when you are sure that your physical or other system behaves uniquely, occasionally your model may be oversimplified and may not give a faithful picture of reality.

THEOREM 1

Existence Theorem

Let the right side $f(x, y)$ of the ODE in the initial value problem

$$y' = f(x, y), \qquad y(x_0) = y_0 \tag{1}$$

be continuous at all points (x, y) in some rectangle

$$R\colon |x - x_0| < a, \qquad |y - y_0| < b \qquad \text{(Fig. 26)}$$

and **bounded** *in R; that is, there is a number K such that*

$$|f(x, y)| \leqq K \qquad \textit{for all } (x, y) \textit{ in } R. \tag{2}$$

Then the initial value problem (1) *has at least one solution $y(x)$. This solution exists at least for all x in the subinterval $|x - x_0| < \alpha$ of the interval $|x - x_0| < a$; here, α is the smaller of the two numbers a and b/K.*

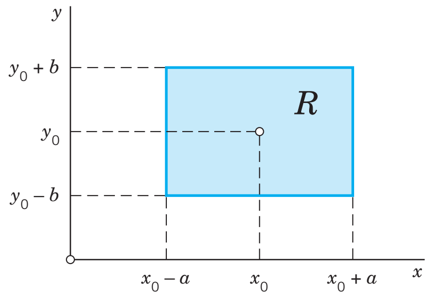


Fig. 26. Rectangle R in the existence and uniqueness theorems

(*Example of Boundedness.* The function $f(x, y) = x^2 + y^2$ is bounded (with $K = 2$) in the square $|x| < 1, |y| < 1$. The function $f(x, y) = \tan(x + y)$ is not bounded for $|x + y| < \pi/2$. Explain!)

THEOREM 2

Uniqueness Theorem

Let f and its partial derivative $f_y = \partial f/\partial y$ be continuous for all (x, y) in the rectangle R (Fig. 26) and bounded, say,

$$(3) \qquad \text{(a)} \quad |f(x, y)| \leqq K, \qquad \text{(b)} \quad |f_y(x, y)| \leqq M \qquad \textit{for all } (x, y) \textit{ in } R.$$

Then the initial value problem (1) has at most one solution $y(x)$. Thus, by Theorem 1, the problem has precisely one solution. This solution exists at least for all x in that subinterval $|x - x_0| < \alpha$.

Understanding These Theorems

These two theorems take care of almost all practical cases. Theorem 1 says that if $f(x, y)$ is continuous in some region in the xy-plane containing the point (x_0, y_0), then the initial value problem (1) has at least one solution.

Theorem 2 says that if, moreover, the partial derivative $\partial f/\partial y$ of f with respect to y exists and is continuous in that region, then (1) can have at most one solution; hence, by Theorem 1, it has precisely one solution.

Read again what you have just read—these are entirely new ideas in our discussion.

Proofs of these theorems are beyond the level of this book (see Ref. [A11] in App. 1); however, the following remarks and examples may help you to a good understanding of the theorems.

Since $y' = f(x, y)$, the condition (2) implies that $|y'| \leqq K$; that is, the slope of any solution curve $y(x)$ in R is at least $-K$ and at most K. Hence a solution curve that passes through the point (x_0, y_0) must lie in the colored region in Fig. 27 bounded by the lines l_1 and l_2 whose slopes are $-K$ and K, respectively. Depending on the form of R, two different cases may arise. In the first case, shown in Fig. 27a, we have $b/K \geqq a$ and therefore $\alpha = a$ in the existence theorem, which then asserts that the solution exists for all x between $x_0 - a$ and $x_0 + a$. In the second case, shown in Fig. 27b, we have $b/K < a$. Therefore, $\alpha = b/K < a$, and all we can conclude from the theorems is that the solution exists for all x between $x_0 - b/K$ and $x_0 + b/K$. For larger or smaller x's the solution curve may leave the rectangle R, and since nothing is assumed about f outside R, nothing can be concluded about the solution for those larger or amaller x's; that is, for such x's the solution may or may not exist—we don't know.

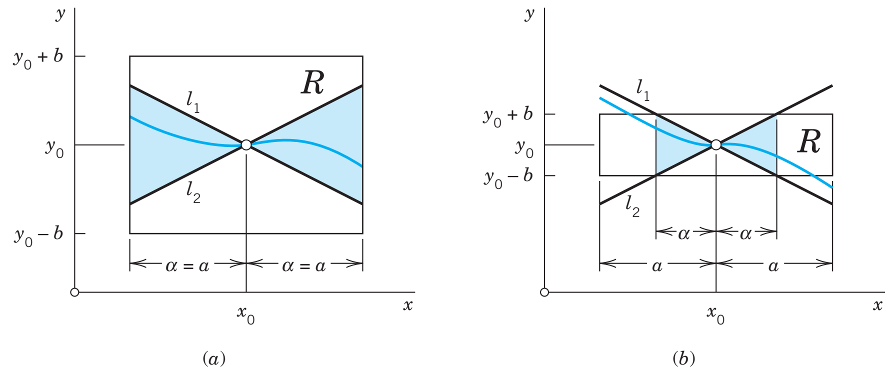


Fig. 27. The condition (2) of the existence theorem. (a) First case. (b) Second case

Let us illustrate our discussion with a simple example. We shall see that our choice of a rectangle R with a large base (a long x-interval) will lead to the case in Fig. 27b.

EXAMPLE 1 **Choice of a Rectangle**

Consider the initial value problem

$$y' = 1 + y^2, \qquad y(0) = 0$$

and take the rectangle R; $|x| < 5$, $|y| < 3$. Then $a = 5$, $b = 3$, and

$$|f(x, y)| = |1 + y^2| \leqq K = 10,$$

$$\left|\frac{\partial f}{\partial y}\right| = 2|y| \leqq M = 6,$$

$$\alpha = \frac{b}{K} = 0.3 < a.$$

Indeed, the solution of the problem is $y = \tan x$ (see Sec. 1.3, Example 1). This solution is discontinuous at $\pm\pi/2$, and there is no *continuous* solution valid in the entire interval $|x| < 5$ from which we started. ■

The conditions in the two theorems are sufficient conditions rather than necessary ones, and can be lessened. In particular, by the mean value theorem of differential calculus we have

$$f(x, y_2) - f(x, y_1) = (y_2 - y_1) \left.\frac{\partial f}{\partial y}\right|_{y=\tilde{y}}$$

where (x, y_1) and (x, y_2) are assumed to be in R, and $\tilde{y}$ is a suitable value between y_1 and y_2. From this and (3b) it follows that

$$|f(x, y_2) - f(x, y_1)| \leqq M|y_2 - y_1|. \tag{4}$$

It can be shown that (3b) may be replaced by the weaker condition (4), which is known as a **Lipschitz condition**.[9] However, continuity of $f(x, y)$ is not enough to guarantee the *uniqueness* of the solution. This may be illustrated by the following example.

EXAMPLE 2 Nonuniqueness

The initial value problem

$$y' = \sqrt{|y|}. \qquad y(0) = 0$$

has the two solutions

$$y = 0 \qquad \text{and} \qquad y^* = \begin{cases} x^2/4 & \text{if } x \geqq 0 \\ -x^2/4 & \text{if } x < 0 \end{cases}$$

although $f(x, y) = \sqrt{|y|}$ is continuous for all y. The Lipschitz condition (4) is violated in any region that includes the line $y = 0$, because for $y_1 = 0$ and positive y_2 we have

$$\frac{|f(x, y_2) - f(x, y_1)|}{|y_2 - y_1|} = \frac{\sqrt{y_2}}{y_2} = \frac{1}{\sqrt{y_2}}, \qquad (\sqrt{y_2} > 0) \tag{5}$$

and this can be made as large as we please by choosing y_2 sufficiently small, whereas (4) requires that the quotient on the left side of (5) should not exceed a fixed constant M. ■

PROBLEM SET 1.7

1. **Linear ODE.** If p and r in $y' + p(x)y = r(x)$ are continuous for all x in an interval $|x - x_0| \leq a$, show that $f(x, y)$ in this ODE satisfies the conditions of our present theorems, so that a corresponding initial value problem has a unique solution. Do you actually need these theorems for this ODE?
2. **Existence?** Does the initial value problem $(x - 2)y' = y$, $y(2) = 1$ have a solution? Does your result contradict our present theorems?
3. **Vertical strip.** If the assumptions of Theorems 1 and 2 are satisfied not merely in a rectangle but in a vertical infinite strip $|x - x_0| < a$, in what interval will the solution of (1) exist?
4. **Change of initial condition.** What happens in Prob. 2 if you replace $y(2) = 1$ with $y(2) = k$?
5. **Length of x-interval.** In most cases the solution of an initial value problem (1) exists in an x-interval larger than that guaranteed by the present theorems. Show this fact for $y' = 2y^2$, $y(1) = 1$ by finding the best possible a (choosing b optimally) and comparing the result with the actual solution.
6. **CAS PROJECT. Picard Iteration. (a)** Show that by integrating the ODE in (1) and observing the initial condition you obtain

$$y(x) = y_0 + \int_{x_0}^{x} f(t, y(t))\, dt. \tag{6}$$

This form (6) of (1) suggests **Picard's Iteration Method**[10] which is defined by

$$y_n(x) = y_0 + \int_{x_0}^{x} f(t, y_{n-1}(t)\, dt, \quad n = 1, 2, \cdots. \tag{7}$$

It gives approximations $y_1, y_2, y_3, \ldots$of the unknown solution y of (1). Indeed, you obtain y_1 by substituting $y = y_0$ on the right and integrating—this is the first step—then y_2 by substituting $y = y_1$ on the right and integrating—this is the second step—and so on. Write

[9]RUDOLF LIPSCHITZ (1832–1903), German mathematician. Lipschitz and similar conditions are important in modern theories, for instance, in partial differential equations.

[10]EMILE PICARD (1856–1941). French mathematician, also known for his important contributions to complex analysis (see Sec. 16.2 for his famous theorem). Picard used his method to prove Theorems 1 and 2 as well as the convergence of the sequence (7) to the solution of (1). In precomputer times, the iteration was of little *practical* value because of the integrations.

a program of the iteration that gives a printout of the first approximations $y_0, y_1, \ldots, y_N$ as well as their graphs on common axes. Try your program on two initial value problems of your own choice.

(b) Apply the iteration to $y' = x + y$, $y(0) = 0$. Also solve the problem exactly.

(c) Apply the iteration to $y' = 2y^2$, $y(0) = 1$. Also solve the problem exactly.

(d) Find all solutions of $y' = 2\sqrt{y}$, $y(1) = 0$. Which of them does Picard's iteration approximate?

(e) Experiment with the conjecture that Picard's iteration converges to the solution of the problem for any initial choice of y in the integrand in (7) (leaving y_0 outside the integral as it is). Begin with a simple ODE and see what happens. When you are reasonably sure, take a slightly more complicated ODE and give it a try.

7. Maximum α. What is the largest possible α in Example 1 in the text?

8. Lipschitz condition. Show that for a linear ODE $y' + p(x)y = r(x)$ with continuous p and r in $|x - x_0| \leqq a$ a Lipschitz condition holds. This is remarkable because it means that for a *linear* ODE the continuity of $f(x, y)$ guarantees not only the existence but also the uniqueness of the solution of an initial value problem. (Of course, this also follows directly from (4) in Sec. 1.5.)

9. Common points. Can two solution curves of the same ODE have a common point in a rectangle in which the assumptions of the present theorems are satisfied?

10. Three possible cases. Find all initial conditions such that $(x^2 - x)y' = (2x - 1)y$ has no solution, precisely one solution, and more than one solution.

CHAPTER 1 REVIEW QUESTIONS AND PROBLEMS

1. Explain the basic concepts ordinary and partial differential equations (ODEs, PDEs), order, general and particular solutions, initial value problems (IVPs). Give examples.
2. What is a linear ODE? Why is it easier to solve than a nonlinear ODE?
3. Does every first-order ODE have a solution? A solution formula? Give examples.
4. What is a direction field? A numeric method for first-order ODEs?
5. What is an exact ODE? Is $f(x)\,dx + g(y)\,dy = 0$ always exact?
6. Explain the idea of an integrating factor. Give two examples.
7. What other solution methods did we consider in this chapter?
8. Can an ODE sometimes be solved by several methods? Give three examples.
9. What does modeling mean? Can a CAS solve a model given by a first-order ODE? Can a CAS set up a model?
10. Give problems from mechanics, heat conduction, and population dynamics that can be modeled by first-order ODEs.

11–16 DIRECTION FIELD: NUMERIC SOLUTION

Graph a direction field (by a CAS or by hand) and sketch some solution curves. Solve the ODE exactly and compare. In Prob. 16 use Euler's method.

11. $y' + 2y = 0$

12. $y' = 1 - y^2$

13. $y' = y - 4y^2$

14. $xy' = y + x^2$

15. $y' + y = 1.01 \cos 10x$

16. Solve $y' = y - y^2$, $y(0) = 0.2$ by Euler's method (10 steps, $h = 0.1$). Solve exactly and compute the error.

17–21 GENERAL SOLUTION

Find the general solution. Indicate which method in this chapter you are using. Show the details of your work.

17. $y' + 2.5y = 1.6x$

18. $y' - 0.4y = 29 \sin x$

19. $25yy' - 4x = 0$

20. $y' = ay + by^2 \quad (a \neq 0)$

21. $(3xe^y + 2y)\,dx + (x^2e^y + x)\,dy = 0$

22–26 INITIAL VALUE PROBLEM (IVP)

Solve the IVP. Indicate the method used. Show the details of your work.

22. $y' + 4xy = e^{-2x^2}, \quad y(0) = -4.3$

23. $y' = \sqrt{1 - y^2}, \quad y(0) = 1/\sqrt{2}$

24. $y' + \frac{1}{2}y = y^3, \quad y(0) = \frac{1}{3}$

25. $3 \sec y\,dx + \frac{1}{3} \sec x\,dy = 0, \quad y(0) = 0$

26. $x \sinh y\,dy = \cosh y\,dx, \quad y(3) = 0$

27–30 MODELING, APPLICATIONS

27. Exponential growth. If the growth rate of a culture of bacteria is proportional to the number of bacteria present and after 1 day is 1.25 times the original number, within what interval of time will the number of bacteria (a) double, (b) triple?

28. Mixing problem. The tank in Fig. 28 contains 80 lb of salt dissolved in 500 gal of water. The inflow per minute is 20 lb of salt dissolved in 20 gal of water. The outflow is 20 gal/min of the uniform mixture. Find the time when the salt content $y(t)$ in the tank reaches 95% of its limiting value (as $t \to \infty$).

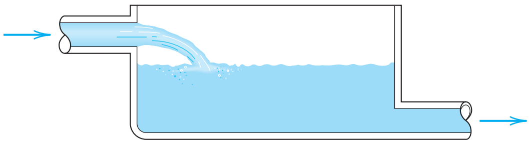


Fig. 28. Tank in Problem 28

29. Half-life. If in a reactor, uranium ${}^{237}_{97}\mathrm{U}$ loses 10% of its weight within one day, what is its half-life? How long would it take for 99% of the original amount to disappear?

30. Newton's law of cooling. A metal bar whose temperature is 20°C is placed in boiling water. How long does it take to heat the bar to practically 100°C, say, to 99.9°C, if the temperature of the bar after 1 min of heating is 51.5°C? First guess, then calculate.

SUMMARY OF CHAPTER 1
First-Order ODEs

This chapter concerns **ordinary differential equations (ODEs) of first order** and their applications. These are equations of the form

$$F(x, y, y') = 0 \qquad \text{or in explicit form} \qquad y' = f(x, y) \tag{1}$$

involving the derivative $y' = dy/dx$ of an unknown function y, given functions of x, and, perhaps, y itself. If the independent variable x is time, we denote it by t.

In Sec. 1.1 we explained the basic concepts and the process of **modeling**, that is, of expressing a physical or other problem in some mathematical form and solving it. Then we discussed the method of direction fields (Sec. 1.2), solution methods and models (Secs. 1.3–1.6), and, finally, ideas on existence and uniqueness of solutions (Sec. 1.7).

A first-order ODE usually has a **general solution**, that is, a solution involving an arbitrary constant, which we denote by c. In applications we usually have to find a unique solution by determining a value of c from an **initial condition** $y(x_0) = y_0$. Together with the ODE this is called an **initial value problem**

$$y' = f(x, y), \qquad y(x_0) = y_0 \qquad (x_0, y_0 \text{ given numbers}) \tag{2}$$

and its solution is a **particular solution** of the ODE. Geometrically, a general solution represents a family of curves, which can be graphed by using **direction fields** (Sec. 1.2). And each particular solution corresponds to one of these curves.

A **separable ODE** is one that we can put into the form

$$g(y)\, dy = f(x)\, dx \qquad \text{(Sec. 1.3)} \tag{3}$$

by algebraic manipulations (possibly combined with transformations, such as $y/x = u$) and solve by integrating on both sides.

An **exact ODE** is of the form

$$M(x, y)\,dx + N(x, y)\,dy = 0 \tag{4}$$ (Sec. 1.4)

where $M\,dx + N\,dy$ is the **differential**

$$du = u_x\,dx + u_y\,dy$$

of a function $u(x, y)$, so that from $du = 0$ we immediately get the implicit general solution $u(x, y) = c$. This method extends to nonexact ODEs that can be made exact by multiplying them by some function $F(x, y,)$, called an **integrating factor** (Sec. 1.4).

Linear ODEs

$$y' + p(x)y = r(x) \tag{5}$$

are very important. Their solutions are given by the integral formula (4), Sec. 1.5. Certain nonlinear ODEs can be transformed to linear form in terms of new variables. This holds for the **Bernoulli equation**

$$y' + p(x)y = g(x)y^a$$ (Sec. 1.5).

Applications and ***modeling*** are discussed throughout the chapter, in particular in Secs. 1.1, 1.3, 1.5 (***population dynamics***, etc.), and 1.6 (***trajectories***).

Picard's ***existence*** and ***uniqueness theorems*** are explained in Sec. 1.7 (and *Picard's iteration* in Problem Set 1.7).

Numeric methods for first-order ODEs can be studied in Secs. 21.1 and 21.2 immediately after this chapter, as indicated in the chapter opening.

CHAPTER 2

Second-Order Linear ODEs

Many important applications in mechanical and electrical engineering, as shown in Secs. 2.4, 2.8, and 2.9, are modeled by linear ordinary differential equations (linear ODEs) of the second order. Their theory is representative of all linear ODEs as is seen when compared to linear ODEs of third and higher order, respectively. However, the solution formulas for second-order linear ODEs are simpler than those of higher order, so it is a natural progression to study ODEs of second order first in this chapter and then of higher order in Chap. 3.

Although ordinary differential equations (ODEs) can be grouped into linear and nonlinear ODEs, nonlinear ODEs are difficult to solve in contrast to linear ODEs for which many beautiful standard methods exist.

Chapter 2 includes the derivation of general and particular solutions, the latter in connection with initial value problems.

For those interested in solution methods for Legendre's, Bessel's, and the hypergeometric equations consult Chap. 5 and for Sturm–Liouville problems Chap. 11.

COMMENT. ***Numerics for second-order ODEs can be studied immediately after this chapter.*** See Sec. 21.3, which is independent of other sections in Chaps. 19–21.

Prerequisite: Chap. 1, in particular, Sec. 1.5.
Sections that may be omitted in a shorter course: 2.3, 2.9, 2.10.
References and Answers to Problems: App. 1 Part A, and App. 2.

2.1 Homogeneous Linear ODEs of Second Order

We have already considered first-order linear ODEs (Sec. 1.5) and shall now define and discuss linear ODEs of second order. These equations have important engineering applications, especially in connection with mechanical and electrical vibrations (Secs. 2.4, 2.8, 2.9) as well as in wave motion, heat conduction, and other parts of physics, as we shall see in Chap. 12.

A second-order ODE is called **linear** if it can be written

$$y'' + p(x)y' + q(x)y = r(x) \tag{1}$$

and **nonlinear** if it cannot be written in this form.

The distinctive feature of this equation is that it is *linear in y and its derivatives,* whereas the functions p, q, and r on the right may be any given functions of x. If the equation begins with, say, $f(x)y''$, then divide by $f(x)$ to have the **standard form** (1) with y'' as the first term.

The definitions of homogeneous and nonhomogenous second-order linear ODEs are very similar to those of first-order ODEs discussed in Sec. 1.5. Indeed, if $r(x) \equiv 0$ (that is, $r(x) = 0$ for all x considered; read "$r(x)$ is identically zero"), then (1) reduces to

$$y'' + p(x)y' + q(x)y = 0 \tag{2}$$

and is called **homogeneous**. If $r(x) \not\equiv 0$, then (1) is called **nonhomogeneous**. This is similar to Sec. 1.5.

An example of a nonhomogeneous linear ODE is

$$y'' + 25y = e^{-x} \cos x,$$

and a homogeneous linear ODE is

$$xy'' + y' + xy = 0, \qquad \text{written in standard form} \qquad y'' + \frac{1}{x}y' + y = 0.$$

Finally, an example of a nonlinear ODE is

$$y''y + y'^2 = 0.$$

The functions p and q in (1) and (2) are called the **coefficients** of the ODEs.

Solutions are defined similarly as for first-order ODEs in Chap. 1. A function

$$y = h(x)$$

is called a *solution* of a (linear or nonlinear) second-order ODE on some open interval I if h is defined and twice differentiable throughout that interval and is such that the ODE becomes an identity if we replace the unknown y by h, the derivative y' by h', and the second derivative y'' by h''. Examples are given below.

Homogeneous Linear ODEs: Superposition Principle

Sections 2.1–2.6 will be devoted to **homogeneous** linear ODEs (2) and the remaining sections of the chapter to nonhomogeneous linear ODEs.

Linear ODEs have a rich solution structure. For the homogeneous equation the backbone of this structure is the *superposition principle* or *linearity principle,* which says that we can obtain further solutions from given ones by adding them or by multiplying them with any constants. Of course, this is a great advantage of homogeneous linear ODEs. Let us first discuss an example.

EXAMPLE 1 **Homogeneous Linear ODEs: Superposition of Solutions**

The functions $y = \cos x$ and $y = \sin x$ are solutions of the homogeneous linear ODE

$$y'' + y = 0$$

for all x. We verify this by differentiation and substitution. We obtain $(\cos x)'' = -\cos x$; hence

$$y'' + y = (\cos x)'' + \cos x = -\cos x + \cos x = 0.$$

Similarly for $y = \sin x$ (verify!). We can go an important step further. We multiply $\cos x$ by any constant, for instance, 4.7, and $\sin x$ by, say, -2, and take the sum of the results, claiming that it is a solution. Indeed, differentiation and substitution gives

$$(4.7 \cos x - 2 \sin x)'' + (4.7 \cos x - 2 \sin x) = -4.7 \cos x + 2 \sin x + 4.7 \cos x - 2 \sin x = 0.$$

In this example we have obtained from y_1 ($= \cos x$) and y_2 ($= \sin x$) a function of the form

$$y = c_1 y_1 + c_2 y_2 \qquad (c_1, c_2 \text{ arbitrary constants}). \tag{3}$$

This is called a **linear combination** of y_1 and y_2. In terms of this concept we can now formulate the result suggested by our example, often called the **superposition principle** or **linearity principle**.

THEOREM 1

Fundamental Theorem for the Homogeneous Linear ODE (2)

For a homogeneous linear ODE (2), *any linear combination of two solutions on an open interval I is again a solution of* (2) *on I. In particular, for such an equation, sums and constant multiples of solutions are again solutions.*

PROOF Let y_1 and y_2 be solutions of (2) on I. Then by substituting $y = c_1 y_1 + c_2 y_2$ and its derivatives into (2), and using the familiar rule $(c_1 y_1 + c_2 y_2)' = c_1 y_1' + c_2 y_2'$, etc., we get

$$\begin{aligned} y'' + py' + qy &= (c_1 y_1 + c_2 y_2)'' + p(c_1 y_1 + c_2 y_2)' + q(c_1 y_1 + c_2 y_2) \\ &= c_1 y_1'' + c_2 y_2'' + p(c_1 y_1' + c_2 y_2') + q(c_1 y_1 + c_2 y_2) \\ &= c_1(y_1'' + py_1' + qy_1) + c_2(y_2'' + py_2' + qy_2) = 0, \end{aligned}$$

since in the last line, $(\cdots) = 0$ because y_1 and y_2 are solutions, by assumption. This shows that y is a solution of (2) on I.

CAUTION! Don't forget that this highly important theorem holds for *homogeneous linear* ODEs only but ***does not hold*** for nonhomogeneous linear or nonlinear ODEs, as the following two examples illustrate.

EXAMPLE 2 **A Nonhomogeneous Linear ODE**

Verify by substitution that the functions $y = 1 + \cos x$ and $y = 1 + \sin x$ are solutions of the nonhomogeneous linear ODE

$$y'' + y = 1,$$

but their sum is not a solution. Neither is, for instance, $2(1 + \cos x)$ or $5(1 + \sin x)$.

EXAMPLE 3 **A Nonlinear ODE**

Verify by substitution that the functions $y = x^2$ and $y = 1$ are solutions of the nonlinear ODE

$$y''y - xy' = 0,$$

but their sum is not a solution. Neither is $-x^2$, so you cannot even multiply by -1!

Initial Value Problem. Basis. General Solution

Recall from Chap. 1 that for a first-order ODE, an *initial value problem* consists of the ODE and one *initial condition* $y(x_0) = y_0$. The initial condition is used to determine the *arbitrary constant c* in the *general solution* of the ODE. This results in a unique solution, as we need it in most applications. That solution is called a *particular solution* of the ODE. These ideas extend to second-order ODEs as follows.

For a second-order homogeneous linear ODE (2) an **initial value problem** consists of (2) and two **initial conditions**

$$y(x_0) = K_0, \qquad y'(x_0) = K_1. \tag{4}$$

These conditions prescribe given values K_0 and K_1 of the solution and its first derivative (the slope of its curve) at the same given $x = x_0$ in the open interval considered.

The conditions (4) are used to determine the two arbitrary constants c_1 and c_2 in a **general solution**

$$y = c_1 y_1 + c_2 y_2 \tag{5}$$

of the ODE; here, y_1 and y_2 are suitable solutions of the ODE, with "suitable" to be explained after the next example. This results in a unique solution, passing through the point (x_0, K_0) with K_1 as the tangent direction (the slope) at that point. That solution is called a **particular solution** of the ODE (2).

EXAMPLE 4 Initial Value Problem

Solve the initial value problem

$$y'' + y = 0, \qquad y(0) = 3.0, \qquad y'(0) = -0.5.$$

Solution. ***Step 1. General solution.*** The functions $\cos x$ and $\sin x$ are solutions of the ODE (by Example 1), and we take

$$y = c_1 \cos x + c_2 \sin x.$$

This will turn out to be a general solution as defined below.

Step 2. Particular solution. We need the derivative $y' = -c_1 \sin x + c_2 \cos x$. From this and the initial values we obtain, since $\cos 0 = 1$ and $\sin 0 = 0$,

$$y(0) = c_1 = 3.0 \qquad \text{and} \qquad y'(0) = c_2 = -0.5.$$

This gives as the solution of our initial value problem the particular solution

$$y = 3.0 \cos x - 0.5 \sin x.$$

Figure 29 shows that at $x = 0$ it has the value 3.0 and the slope -0.5, so that its tangent intersects the x-axis at $x = 3.0/0.5 = 6.0$. (The scales on the axes differ!) ■

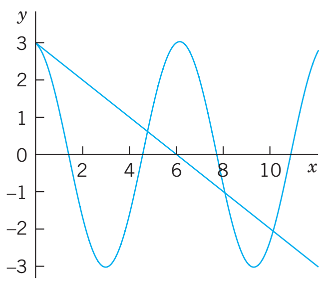


Fig. 29. Particular solution and initial tangent in Example 4

Observation. Our choice of y_1 and y_2 was general enough to satisfy both initial conditions. Now let us take instead two proportional solutions $y_1 = \cos x$ and $y_2 = k \cos x$, so that $y_1/y_2 = 1/k = \text{const}$. Then we can write $y = c_1 y_1 + c_2 y_2$ in the form

$$y = c_1 \cos x + c_2(k \cos x) = C \cos x \qquad \text{where} \qquad C = c_1 + c_2 k.$$

Hence we are no longer able to satisfy two initial conditions with only one arbitrary constant C. Consequently, in defining the concept of a general solution, we must exclude proportionality. And we see at the same time why the concept of a general solution is of importance in connection with initial value problems.

DEFINITION

General Solution, Basis, Particular Solution

A **general solution** of an ODE (2) on an open interval I is a solution (5) in which y_1 and y_2 are solutions of (2) on I that are not proportional, and c_1 and c_2 are arbitrary constants. These y_1, y_2 are called a **basis** (or a **fundamental system**) of solutions of (2) on I.

A **particular solution** of (2) on I is obtained if we assign specific values to c_1 and c_2 in (5).

For the definition of an *interval* see Sec. 1.1. Furthermore, as usual, y_1 and y_2 are called *proportional* on I if for all x on I,

$$(6) \qquad \text{(a)} \quad y_1 = ky_2 \qquad \text{or} \qquad \text{(b)} \quad y_2 = ly_1$$

where k and l are numbers, zero or not. (Note that (a) implies (b) if and only if $k \neq 0$).

Actually, we can reformulate our definition of a basis by using a concept of general importance. Namely, two functions y_1 and y_2 are called **linearly independent** on an interval I where they are defined if

$$(7) \qquad k_1y_1(x) + k_2y_2(x) = 0 \qquad \text{everywhere on } I \text{ implies} \qquad k_1 = 0 \text{ and } k_2 = 0.$$

And y_1 and y_2 are called **linearly dependent** on I if (7) also holds for some constants k_1, k_2 not both zero. Then, if $k_1 \neq 0$ or $k_2 \neq 0$, we can divide and see that y_1 and y_2 are proportional,

$$y_1 = -\frac{k_2}{k_1}y_2 \qquad \text{or} \qquad y_2 = -\frac{k_1}{k_2}y_1.$$

In contrast, in the case of linear *independence* these functions are not proportional because then we cannot divide in (7). This gives the following

DEFINITION

Basis (Reformulated)

A **basis** of solutions of (2) on an open interval I is a pair of linearly independent solutions of (2) on I.

If the coefficients p and q of (2) are continuous on some open interval I, then (2) has a general solution. It yields the unique solution of any initial value problem (2), (4). It includes all solutions of (2) on I; hence (2) has no *singular solutions* (solutions not obtainable from of a general solution; see also Problem Set 1.1). All this will be shown in Sec. 2.6.

EXAMPLE 5 **Basis, General Solution, Particular Solution**

$\cos x$ and $\sin x$ in Example 4 form a basis of solutions of the ODE $y'' + y = 0$ for all x because their quotient is $\cot x \neq \text{const}$ (or $\tan x \neq \text{const}$). Hence $y = c_1 \cos x + c_2 \sin x$ is a general solution. The solution $y = 3.0 \cos x - 0.5 \sin x$ of the initial value problem is a particular solution. ■

EXAMPLE 6 **Basis, General Solution, Particular Solution**

Verify by substitution that $y_1 = e^x$ and $y_2 = e^{-x}$ are solutions of the ODE $y'' - y = 0$. Then solve the initial value problem

$$y'' - y = 0, \qquad y(0) = 6, \qquad y'(0) = -2.$$

Solution. $(e^x)'' - e^x = 0$ and $(e^{-x})'' - e^{-x} = 0$ show that e^x and e^{-x} are solutions. They are not proportional, $e^x/e^{-x} = e^{2x} \neq \text{const}$. Hence e^x, e^{-x} form a basis for all x. We now write down the corresponding general solution and its derivative and equate their values at 0 to the given initial conditions,

$$y = c_1 e^x + c_2 e^{-x}, \qquad y' = c_1 e^x - c_2 e^{-x}, \qquad y(0) = c_1 + c_2 = 6, \qquad y'(0) = c_1 - c_2 = -2.$$

By addition and subtraction, $c_1 = 2$, $c_2 = 4$, so that the *answer* is $y = 2e^x + 4e^{-x}$. This is the particular solution satisfying the two initial conditions. ■

Find a Basis if One Solution Is Known. Reduction of Order

It happens quite often that one solution can be found by inspection or in some other way. Then a second linearly independent solution can be obtained by solving a first-order ODE. This is called the method of **reduction of order**.[1] We first show how this method works in an example and then in general.

EXAMPLE 7 **Reduction of Order if a Solution Is Known. Basis**

Find a basis of solutions of the ODE

$$(x^2 - x)y'' - xy' + y = 0.$$

Solution. Inspection shows that $y_1 = x$ is a solution because $y_1' = 1$ and $y_1'' = 0$, so that the first term vanishes identically and the second and third terms cancel. The idea of the method is to substitute

$$y = uy_1 = ux, \qquad y' = u'x + u, \qquad y'' = u''x + 2u'$$

into the ODE. This gives

$$(x^2 - x)(u''x + 2u') - x(u'x + u) + ux = 0.$$

ux and $-xu$ cancel and we are left with the following ODE, which we divide by x, order, and simplify,

$$(x^2 - x)(u''x + 2u') - x^2u' = 0, \qquad (x^2 - x)u'' + (x - 2)u' = 0.$$

This ODE is of first order in $v = u'$, namely, $(x^2 - x)v' + (x - 2)v = 0$. Separation of variables and integration gives

$$\frac{dv}{v} = -\frac{x-2}{x^2 - x}\,dx = \left(\frac{1}{x-1} - \frac{2}{x}\right)dx, \qquad \ln|v| = \ln|x - 1| - 2\ln|x| = \ln\frac{|x-1|}{x^2}.$$

---

[1]Credited to the great mathematician JOSEPH LOUIS LAGRANGE (1736–1813), who was born in Turin, of French extraction, got his first professorship when he was 19 (at the Military Academy of Turin), became director of the mathematical section of the Berlin Academy in 1766, and moved to Paris in 1787. His important major work was in the calculus of variations, celestial mechanics, general mechanics (*Mécanique analytique,* Paris, 1788), differential equations, approximation theory, algebra, and number theory.

We need no constant of integration because we want to obtain a particular solution; similarly in the next integration. Taking exponents and integrating again, we obtain

$$v = \frac{x-1}{x^2} = \frac{1}{x} - \frac{1}{x^2}, \qquad u = \int v\, dx = \ln|x| + \frac{1}{x}, \qquad \text{hence} \qquad y_2 = ux = x\ln|x| + 1.$$

Since $y_1 = x$ and $y_2 = x\ln|x| + 1$ are linearly independent (their quotient is not constant), we have obtained a basis of solutions, valid for all positive x. ■

In this example we applied **reduction of order** to a homogeneous linear ODE [see (2)]

$$y'' + p(x)y' + q(x)y = 0.$$

Note that we now take the ODE in standard form, with y'', not $f(x)y''$—this is essential in applying our subsequent formulas. We assume a solution y_1 of (2), on an open interval I, to be known and want to find a basis. For this we need a second linearly independent solution y_2 of (2) on I. To get y_2, we substitute

$$y = y_2 = uy_1, \qquad y' = y_2' = u'y_1 + uy_1', \qquad y'' = y_2'' = u''y_1 + 2u'y_1' + uy_1''$$

into (2). This gives

$$u''y_1 + 2u'y_1' + uy_1'' + p(u'y_1 + uy_1') + quy_1 = 0. \tag{8}$$

Collecting terms in u'', u', and u, we have

$$u''y_1 + u'(2y_1' + py_1) + u(y_1'' + py_1' + qy_1) = 0.$$

Now comes the main point. Since y_1 is a solution of (2), the expression in the last parentheses is zero. Hence u is gone, and we are left with an ODE in u' and u''. We divide this remaining ODE by y_1 and set $u' = U$, $u'' = U'$,

$$u'' + u'\frac{2y_1' + py_1}{y_1} = 0, \qquad \text{thus} \qquad U' + \left(\frac{2y_1'}{y_1} + p\right)U = 0.$$

This is the desired first-order ODE, the reduced ODE. Separation of variables and integration gives

$$\frac{dU}{U} = -\left(\frac{2y_1'}{y_1} + p\right)dx \qquad \text{and} \qquad \ln|U| = -2\ln|y_1| - \int p\, dx.$$

By taking exponents we finally obtain

$$U = \frac{1}{y_1^2}e^{-\int p\, dx}. \tag{9}$$

Here $U = u'$, so that $u = \int U\, dx$. Hence the desired second solution is

$$y_2 = y_1 u = y_1 \int U\, dx.$$

The quotient $y_2/y_1 = u = \int U\, dx$ cannot be constant (since $U > 0$), so that y_1 and y_2 form a basis of solutions.

PROBLEM SET 2.1

REDUCTION OF ORDER is important because it gives a simpler ODE. A general second-order ODE $F(x, y, y', y'') = 0$, linear or not, can be reduced to first order if y does not occur explicitly (Prob. 1) or if x does not occur explicitly (Prob. 2) or if the ODE is homogeneous linear and we know a solution (see the text).

1. **Reduction.** Show that $F(x, y', y'') = 0$ can be reduced to first order in $z = y'$ (from which y follows by integration). Give two examples of your own.
2. **Reduction.** Show that $F(y, y', y'') = 0$ can be reduced to a first-order ODE with y as the independent variable and $y'' = (dz/dy)z$, where $z = y'$; derive this by the chain rule. Give two examples.

3–10 REDUCTION OF ORDER

Reduce to first order and solve, showing each step in detail.

3. $y'' + y' = 0$
4. $2xy'' = 3y'$
5. $yy'' = 3y'^2$
6. $xy'' + 2y' + xy = 0, \quad y_1 = (\cos x)/x$
7. $y'' + y'^3 \sin y = 0$
8. $y'' = 1 + y'^2$
9. $x^2y'' - 5xy' + 9y = 0, \quad y_1 = x^3$
10. $y'' + (1 + 1/y)y'^2 = 0$

11–14 APPLICATIONS OF REDUCIBLE ODEs

11. **Curve.** Find the curve through the origin in the xy-plane which satisfies $y'' = 2y'$ and whose tangent at the origin has slope 1.
12. **Hanging cable.** It can be shown that the curve $y(x)$ of an inextensible flexible homogeneous cable hanging between two fixed points is obtained by solving $y'' = k\sqrt{1 + y'^2}$, where the constant k depends on the weight. This curve is called *catenary* (from Latin *catena* = the chain). Find and graph $y(x)$, assuming that $k = 1$ and those fixed points are $(-1, 0)$ and $(1, 0)$ in a vertical xy-plane.
13. **Motion.** If, in the motion of a small body on a straight line, the sum of velocity and acceleration equals a positive constant, how will the distance $y(t)$ depend on the initial velocity and position?
14. **Motion.** In a straight-line motion, let the velocity be the reciprocal of the acceleration. Find the distance $y(t)$ for arbitrary initial position and velocity.

15–19 GENERAL SOLUTION. INITIAL VALUE PROBLEM (IVP)

(More in the next set.) **(a)** Verify that the given functions are linearly independent and form a basis of solutions of the given ODE. **(b)** Solve the IVP. Graph or sketch the solution.

15. $4y'' + 25y = 0, \quad y(0) = 3.0, \quad y'(0) = -2.5,$
 $\cos 2.5x, \sin 2.5x$
16. $y'' + 0.6y' + 0.09y = 0, \quad y(0) = 2.2, \quad y'(0) = 0.14,$
 $e^{-0.3x}, xe^{-0.3x}$
17. $4x^2y'' - 3y = 0, \quad y(1) = -3, \quad y'(1) = 0,$
 $x^{3/2}, x^{-1/2}$
18. $x^2y'' - xy' + y = 0, \quad y(1) = 4.3, \quad y'(1) = 0.5,$
 $x, x \ln x$
19. $y'' + 2y' + 2y = 0, \quad y(0) = 0, \quad y'(0) = 15,$
 $e^{-x} \cos x, e^{-x} \sin x$
20. **CAS PROJECT. Linear Independence.** Write a program for testing linear independence and dependence. Try it out on some of the problems in this and the next problem set and on examples of your own.

2.2 Homogeneous Linear ODEs with Constant Coefficients

We shall now consider second-order homogeneous linear ODEs whose coefficients a and b are constant,

$$y'' + ay' + by = 0. \tag{1}$$

These equations have important applications in mechanical and electrical vibrations, as we shall see in Secs. 2.4, 2.8, and 2.9.

To solve (1), we recall from Sec. 1.5 that the solution of the first-order linear ODE with a constant coefficient k

$$y' + ky = 0$$

is an exponential function $y = ce^{-kx}$. This gives us the idea to try as a solution of (1) the function

$$y = e^{\lambda x}. \tag{2}$$

Substituting (2) and its derivatives

$$y' = \lambda e^{\lambda x} \qquad \text{and} \qquad y'' = \lambda^2 e^{\lambda x}$$

into our equation (1), we obtain

$$(\lambda^2 + a\lambda + b)e^{\lambda x} = 0.$$

Hence if λ is a solution of the important **characteristic equation** (or *auxiliary equation*)

$$\boxed{\lambda^2 + a\lambda + b = 0} \tag{3}$$

then the exponential function (2) is a solution of the ODE (1). Now from algebra we recall that the roots of this quadratic equation (3) are

$$\lambda_1 = \tfrac{1}{2}\left(-a + \sqrt{a^2 - 4b}\right), \qquad \lambda_2 = \tfrac{1}{2}\left(-a - \sqrt{a^2 - 4b}\right). \tag{4}$$

(3) and (4) will be basic because our derivation shows that the functions

$$y_1 = e^{\lambda_1 x} \qquad \text{and} \qquad y_2 = e^{\lambda_2 x} \tag{5}$$

are solutions of (1). Verify this by substituting (5) into (1).

From algebra we further know that the quadratic equation (3) may have three kinds of roots, depending on the sign of the discriminant $a^2 - 4b$, namely,

(Case I)	*Two real roots if* $a^2 - 4b > 0$,
(Case II)	*A real double root if* $a^2 - 4b = 0$,
(Case III)	*Complex conjugate roots if* $a^2 - 4b < 0$.

Case I. Two Distinct Real-Roots λ_1 and λ_2

In this case, a basis of solutions of (1) on any interval is

$$y_1 = e^{\lambda_1 x} \qquad \text{and} \qquad y_2 = e^{\lambda_2 x}$$

because y_1 and y_2 are defined (and real) for all x and their quotient is not constant. The corresponding general solution is

$$\boxed{y = c_1 e^{\lambda_1 x} + c_2 e^{\lambda_2 x}.} \tag{6}$$

EXAMPLE 1 General Solution in the Case of Distinct Real Roots

We can now solve $y'' - y = 0$ in Example 6 of Sec. 2.1 systematically. The characteristic equation is $\lambda^2 - 1 = 0$. Its roots are $\lambda_1 = 1$ and $\lambda_2 = -1$. Hence a basis of solutions is e^x and e^{-x} and gives the same general solution as before,

$$y = c_1 e^x + c_2 e^{-x}.$$

EXAMPLE 2 Initial Value Problem in the Case of Distinct Real Roots

Solve the initial value problem

$$y'' + y' - 2y = 0, \qquad y(0) = 4, \qquad y'(0) = -5.$$

Solution. ***Step 1. General solution.*** The characteristic equation is

$$\lambda^2 + \lambda - 2 = 0.$$

Its roots are

$$\lambda_1 = \tfrac{1}{2}(-1 + \sqrt{9}) = 1 \qquad \text{and} \qquad \lambda_2 = \tfrac{1}{2}(-1 - \sqrt{9}) = -2$$

so that we obtain the general solution

$$y = c_1 e^x + c_2 e^{-2x}.$$

Step 2. Particular solution. Since $y'(x) = c_1 e^x - 2c_2 e^{-2x}$, we obtain from the general solution and the initial conditions

$$y(0) = c_1 + c_2 = 4,$$
$$y'(0) = c_1 - 2c_2 = -5.$$

Hence $c_1 = 1$ and $c_2 = 3$. This gives the *answer* $y = e^x + 3e^{-2x}$. Figure 30 shows that the curve begins at $y = 4$ with a negative slope (-5, but note that the axes have different scales!), in agreement with the initial conditions.

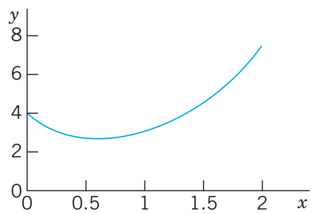


Fig. 30. Solution in Example 2

Case II. Real Double Root $\lambda = -a/2$

If the discriminant $a^2 - 4b$ is zero, we see directly from (4) that we get only one root, $\lambda = \lambda_1 = \lambda_2 = -a/2$, hence only one solution,

$$y_1 = e^{-(a/2)x}.$$

To obtain a second independent solution y_2 (needed for a basis), we use the method of reduction of order discussed in the last section, setting $y_2 = uy_1$. Substituting this and its derivatives $y_2' = u'y_1 + uy_1'$ and y_2'' into (1), we first have

$$(u''y_1 + 2u'y_1' + uy_1'') + a(u'y_1 + uy_1') + buy_1 = 0.$$

Collecting terms in u'', u', and u, as in the last section, we obtain

$$u''y_1 + u'(2y_1' + ay_1) + u(y_1'' + ay_1' + by_1) = 0.$$

The expression in the last parentheses is zero, since y_1 is a solution of (1). The expression in the first parentheses is zero, too, since

$$2y_1' = -ae^{-ax/2} = -ay_1.$$

We are thus left with $u''y_1 = 0$. Hence $u'' = 0$. By two integrations, $u = c_1 x + c_2$. To get a second independent solution $y_2 = uy_1$, we can simply choose $c_1 = 1$, $c_2 = 0$ and take $u = x$. Then $y_2 = xy_1$. Since these solutions are not proportional, they form a basis. Hence in the case of a double root of (3) a basis of solutions of (1) on any interval is

$$e^{-ax/2}, \qquad xe^{-ax/2}.$$

The corresponding general solution is

$$y = (c_1 + c_2 x)e^{-ax/2}. \tag{7}$$

WARNING! If λ is a *simple* root of (4), then $(c_1 + c_2 x)e^{\lambda x}$ with $c_2 \neq 0$ is ***not*** a solution of (1).

EXAMPLE 3 General Solution in the Case of a Double Root

The characteristic equation of the ODE $y'' + 6y' + 9y = 0$ is $\lambda^2 + 6\lambda + 9 = (\lambda + 3)^2 = 0$. It has the double root $\lambda = -3$. Hence a basis is e^{-3x} and xe^{-3x}. The corresponding general solution is $y = (c_1 + c_2 x)e^{-3x}$. ■

EXAMPLE 4 Initial Value Problem in the Case of a Double Root

Solve the initial value problem

$$y'' + y' + 0.25y = 0, \qquad y(0) = 3.0, \quad y'(0) = -3.5.$$

Solution. The characteristic equation is $\lambda^2 + \lambda + 0.25 = (\lambda + 0.5)^2 = 0$. It has the double root $\lambda = -0.5$. This gives the general solution

$$y = (c_1 + c_2 x)e^{-0.5x}.$$

We need its derivative

$$y' = c_2 e^{-0.5x} - 0.5(c_1 + c_2 x)e^{-0.5x}.$$

From this and the initial conditions we obtain

$$y(0) = c_1 = 3.0, \qquad y'(0) = c_2 - 0.5c_1 = 3.5; \qquad \text{hence} \qquad c_2 = -2.$$

The particular solution of the initial value problem is $y = (3 - 2x)e^{-0.5x}$. See Fig. 31. ■

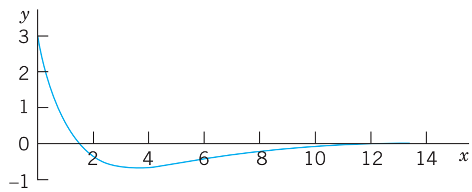


Fig. 31. Solution in Example 4

Case III. Complex Roots $-\frac{1}{2}a + i\omega$ and $-\frac{1}{2}a - i\omega$

This case occurs if the discriminant $a^2 - 4b$ of the characteristic equation (3) is negative. In this case, the roots of (3) are the complex $\lambda = -\frac{1}{2}a \pm i\omega$ that give the complex solutions of the ODE (1). However, we will show that we can obtain a basis of *real* solutions

$$y_1 = e^{-ax/2} \cos \omega x, \qquad y_2 = e^{-ax/2} \sin \omega x \qquad (\omega > 0) \tag{8}$$

where $\omega^2 = b - \frac{1}{4}a^2$. It can be verified by substitution that these are solutions in the present case. We shall derive them systematically after the two examples by using the complex exponential function. They form a basis on any interval since their quotient $\cot \omega x$ is not constant. Hence a real general solution in Case III is

$$\boxed{y = e^{-ax/2}(A \cos \omega x + B \sin \omega x)} \qquad (A, B \text{ arbitrary}). \tag{9}$$

EXAMPLE 5 **Complex Roots. Initial Value Problem**

Solve the initial value problem

$$y'' + 0.4y' + 9.04y = 0, \qquad y(0) = 0, \qquad y'(0) = 3.$$

Solution. ***Step 1. General solution.*** The characteristic equation is $\lambda^2 + 0.4\lambda + 9.04 = 0$. It has the roots $-0.2 \pm 3i$. Hence $\omega = 3$, and a general solution (9) is

$$y = e^{-0.2x}(A \cos 3x + B \sin 3x).$$

Step 2. Particular solution. The first initial condition gives $y(0) = A = 0$. The remaining expression is $y = Be^{-0.2x} \sin 3x$. We need the derivative (chain rule!)

$$y' = B(-0.2e^{-0.2x} \sin 3x + 3e^{-0.2x} \cos 3x).$$

From this and the second initial condition we obtain $y'(0) = 3B = 3$. Hence $B = 1$. Our solution is

$$y = e^{-0.2x} \sin 3x.$$

Figure 32 shows y and the curves of $e^{-0.2x}$ and $-e^{-0.2x}$ (dashed), between which the curve of y oscillates. Such "damped vibrations" (with $x = t$ being time) have important mechanical and electrical applications, as we shall soon see (in Sec. 2.4). ■

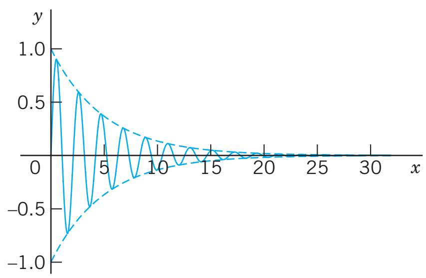


Fig. 32. Solution in Example 5

EXAMPLE 6 **Complex Roots**

A general solution of the ODE

$$y'' + \omega^2 y = 0 \qquad (\omega \text{ constant, not zero})$$

is

$$y = A \cos \omega x + B \sin \omega x.$$

With $\omega = 1$ this confirms Example 4 in Sec. 2.1. ■

Summary of Cases I–III

Case	Roots of (2)	Basis of (1)	General Solution of (1)
I	Distinct real λ_1, λ_2	$e^{\lambda_1 x}, e^{\lambda_2 x}$	$y = c_1 e^{\lambda_1 x} + c_2 e^{\lambda_2 x}$
II	Real double root $\lambda = -\frac{1}{2}a$	$e^{-ax/2}, xe^{-ax/2}$	$y = (c_1 + c_2 x)e^{-ax/2}$
III	Complex conjugate $\lambda_1 = -\frac{1}{2}a + i\omega$, $\lambda_2 = -\frac{1}{2}a - i\omega$	$e^{-ax/2} \cos \omega x$ $e^{-ax/2} \sin \omega x$	$y = e^{-ax/2}(A \cos \omega x + B \sin \omega x)$

It is very interesting that in applications to mechanical systems or electrical circuits, these three cases correspond to three different forms of motion or flows of current, respectively. We shall discuss this basic relation between theory and practice in detail in Sec. 2.4 (and again in Sec. 2.8).

Derivation in Case III. Complex Exponential Function

If verification of the solutions in (8) satisfies you, skip the systematic derivation of these real solutions from the complex solutions by means of the complex exponential function e^z of a complex variable $z = r + it$. We write $r + it$, not $x + iy$ because x and y occur in the ODE. The definition of e^z in terms of the real functions e^r, $\cos t$, and $\sin t$ is

$$e^z = e^{r+it} = e^r e^{it} = e^r(\cos t + i \sin t). \tag{10}$$

This is motivated as follows. For real $z = r$, hence $t = 0$, $\cos 0 = 1$, $\sin 0 = 0$, we get the real exponential function e^r. It can be shown that $e^{z_1+z_2} = e^{z_1}e^{z_2}$, just as in real. (Proof in Sec. 13.5.) Finally, if we use the Maclaurin series of e^z with $z = it$ as well as $i^2 = -1$, $i^3 = -i$, $i^4 = 1$, etc., and reorder the terms as shown (this is permissible, as can be proved), we obtain the series

$$\begin{aligned} e^{it} &= 1 + it + \frac{(it)^2}{2!} + \frac{(it)^3}{3!} + \frac{(it)^4}{4!} + \frac{(it)^5}{5!} + \cdots \\ &= 1 - \frac{t^2}{2!} + \frac{t^4}{4!} - + \cdots + i\left(t - \frac{t^3}{3!} + \frac{t^5}{5!} - + \cdots\right) \\ &= \cos t + i \sin t. \end{aligned}$$

(Look up these real series in your calculus book if necessary.) We see that we have obtained the formula

$$\boxed{e^{it} = \cos t + i \sin t,} \tag{11}$$

called the **Euler formula**. Multiplication by e^r gives (10).

For later use we note that $e^{-it} = \cos(-t) + i \sin(-t) = \cos t - i \sin t$, so that by addition and subtraction of this and (11),

$$\cos t = \tfrac{1}{2}(e^{it} + e^{-it}), \qquad \sin t = \frac{1}{2i}(e^{it} - e^{-it}). \tag{12}$$

After these comments on the definition (10), let us now turn to Case III.

In Case III the radicand $a^2 - 4b$ in (4) is negative. Hence $4b - a^2$ is positive and, using $\sqrt{-1} = i$, we obtain in (4)

$$\tfrac{1}{2}\sqrt{a^2 - 4b} = \tfrac{1}{2}\sqrt{-(4b - a^2)} = \sqrt{-(b - \tfrac{1}{4}a^2)} = i\sqrt{b - \tfrac{1}{4}a^2} = i\omega$$

with ω defined as in (8). Hence in (4),

$$\lambda_1 = \tfrac{1}{2}a + i\omega \qquad \text{and, similarly,} \qquad \lambda_2 = \tfrac{1}{2}a - i\omega.$$

Using (10) with $r = -\tfrac{1}{2}ax$ and $t = \omega x$, we thus obtain

$$e^{\lambda_1 x} = e^{-(a/2)x + i\omega x} = e^{-(a/2)x}(\cos \omega x + i \sin \omega x)$$

$$e^{\lambda_2 x} = e^{-(a/2)x - i\omega x} = e^{-(a/2)x}(\cos \omega x - i \sin \omega x).$$

We now add these two lines and multiply the result by $\tfrac{1}{2}$. This gives y_1 as in (8). Then we subtract the second line from the first and multiply the result by $1/(2i)$. This gives y_2 as in (8). These results obtained by addition and multiplication by constants are again solutions, as follows from the superposition principle in Sec. 2.1. This concludes the derivation of these real solutions in Case III.

PROBLEM SET 2.2

1–15 GENERAL SOLUTION

Find a general solution. Check your answer by substitution. ODEs of this kind have important applications to be discussed in Secs. 2.4, 2.7, and 2.9.

1. $4y'' - 25y = 0$
2. $y'' + 36y = 0$
3. $y'' + 6y' + 8.96y = 0$
4. $y'' + 4y' + (\pi^2 + 4)y = 0$
5. $y'' + 2\pi y' + \pi^2 y = 0$
6. $10y'' - 32y' + 25.6y = 0$
7. $y'' + 4.5y' = 0$
8. $y'' + y' + 3.25y = 0$
9. $y'' + 1.8y' - 2.08y = 0$
10. $100y'' + 240y' + (196\pi^2 + 144)y = 0$
11. $4y'' - 4y' - 3y = 0$
12. $y'' + 9y' + 20y = 0$
13. $9y'' - 30y' + 25y = 0$
14. $y'' + 2k^2y' + k^4y = 0$
15. $y'' + 0.54y' + (0.0729 + \pi)y = 0$

16–20 FIND AN ODE

$y'' + ay' + by = 0$ for the given basis.

16. $e^{2.6x}$, $e^{-4.3x}$
17. $e^{-\sqrt{5}x}$, $xe^{-\sqrt{5}x}$
18. $\cos 2\pi x$, $\sin 2\pi x$
19. $e^{(-2+i)x}$, $e^{(-2-i)x}$
20. $e^{-3.1x} \cos 2.1x$, $e^{-3.1x} \sin 2.1x$

21–30 INITIAL VALUES PROBLEMS

Solve the IVP. Check that your answer satisfies the ODE as well as the initial conditions. Show the details of your work.

21. $y'' + 25y = 0, \quad y(0) = 4.6, \quad y'(0) = -1.2$
22. The ODE in Prob. 4, $y(\tfrac{1}{2}) = 1, \quad y'(\tfrac{1}{2}) = -2$
23. $y'' + y' - 6y = 0, \quad y(0) = 10, \quad y'(0) = 0$
24. $4y'' - 4y' - 3y = 0, \quad y(-2) = e, \quad y'(-2) = -e/2$
25. $y'' - y = 0, \quad y(0) = 2, \quad y'(0) = -2$
26. $y'' - k^2y = 0\ (k \neq 0), \quad y(0) = 1, \quad y'(0) = 1$

27. The ODE in Prob. 5, $y(0) = 4.5, \quad y'(0) = -4.5\pi - 1 = 13.137$

28. $8y'' - 2y' - y = 0, \quad y(0) = -0.2, \quad y'(0) = -0.325$

29. The ODE in Prob. 15, $y(0) = 0, \quad y'(0) = 1$

30. $9y'' - 30y' + 25y = 0, \quad y(0) = 3.3, \quad y'(0) = 10.0$

31–36 LINEAR INDEPENDENCE is of basic importance, in this chapter, in connection with general solutions, as explained in the text. Are the following functions linearly independent on the given interval? Show the details of your work.

31. $e^{kx}, xe^{kx}, \quad$ any interval

32. $e^{ax}, e^{-ax}, \quad x > 0$

33. $x^2, x^2 \ln x, \quad x > 1$

34. $\ln x, \ln (x^3), \quad x > 1$

35. $\sin 2x, \cos x \sin x, \quad x < 0$

36. $e^{-x} \cos \frac{1}{2}x, 0, \quad -1 \leqq x \leqq 1$

37. Instability. Solve $y'' - y = 0$ for the initial conditions $y(0) = 1, y'(0) = -1$. Then change the initial conditions to $y(0) = 1.001, y'(0) = -0.999$ and explain why this small change of 0.001 at $t = 0$ causes a large change later, e.g., 22 at $t = 10$. This is instability: a small initial difference in setting a quantity (a current, for instance) becomes larger and larger with time t. This is undesirable.

38. TEAM PROJECT. General Properties of Solutions

(a) Coefficient formulas. Show how a and b in (1) can be expressed in terms of λ_1 and λ_2. Explain how these formulas can be used in constructing equations for given bases.

(b) Root zero. Solve $y'' + 4y' = 0$ (i) by the present method, and (ii) by reduction to first order. Can you explain why the result must be the same in both cases? Can you do the same for a general ODE $y'' + ay' = 0$?

(c) Double root. Verify directly that $xe^{\lambda x}$ with $\lambda = -a/2$ is a solution of (1) in the case of a double root. Verify and explain why $y = e^{-2x}$ is a solution of $y'' - y' - 6y = 0$ but xe^{-2x} is not.

(d) Limits. Double roots should be limiting cases of distinct roots λ_1, λ_2 as, say, $\lambda_2 \to \lambda_1$. Experiment with this idea. (Remember l'Hôpital's rule from calculus.) Can you arrive at $xe^{\lambda_1 x}$? Give it a try.

2.3 Differential Operators. *Optional*

This short section can be omitted without interrupting the flow of ideas. It will not be used subsequently, except for the notations Dy, D^2y, etc. to stand for y', y'', etc.

Operational calculus means the technique and application of operators. Here, an **operator** is a transformation that transforms a function into another function. Hence differential calculus involves an operator, the **differential operator** D, which transforms a (differentiable) function into its derivative. In operator notation we write $D = \frac{d}{dx}$ and

$$Dy = y' = \frac{dy}{dx}. \tag{1}$$

Similarly, for the higher derivatives we write $D^2y = D(Dy) = y''$, and so on. For example, $D \sin = \cos$, $D^2 \sin = -\sin$, etc.

For a homogeneous linear ODE $y'' + ay' + by = 0$ with constant coefficients we can now introduce the **second-order differential operator**

$$L = P(D) = D^2 + aD + bI,$$

where I is the **identity operator** defined by $Iy = y$. Then we can write that ODE as

$$Ly = P(D)y = (D^2 + aD + bI)y = 0. \tag{2}$$

P suggests "polynomial." L is a **linear operator**. By definition this means that if Ly and Lw exist (this is the case if y and w are twice differentiable), then $L(cy + kw)$ exists for any constants c and k, and

$$L(cy + kw) = cLy + kLw.$$

Let us show that from (2) we reach agreement with the results in Sec. 2.2. Since $(De^{\lambda})(x) = \lambda e^{\lambda x}$ and $(D^2e^{\lambda})(x) = \lambda^2 e^{\lambda x}$, we obtain

$$\begin{aligned} Le^{\lambda}(x) &= P(D)e^{\lambda}(x) = (D^2 + aD + bI)e^{\lambda}(x) \\ &= (\lambda^2 + a\lambda + b)e^{\lambda x} = P(\lambda)e^{\lambda x} = 0. \end{aligned} \tag{3}$$

This confirms our result of Sec. 2.2 that $e^{\lambda x}$ *is a solution of the ODE* (2) *if and only if* λ *is a solution of the characteristic equation* $P(\lambda) = 0$.

$P(\lambda)$ is a polynomial in the usual sense of algebra. If we replace λ by the operator D, we obtain the "operator polynomial" $P(D)$. *The point of this operational calculus is that* $P(D)$ *can be treated just like an algebraic quantity.* In particular, we can factor it.

EXAMPLE 1 Factorization, Solution of an ODE

Factor $P(D) = D^2 - 3D - 40I$ and solve $P(D)y = 0$.

Solution. $D^2 - 3D - 40I = (D - 8I)(D + 5I)$ because $I^2 = I$. Now $(D - 8I)y = y' - 8y = 0$ has the solution $y_1 = e^{8x}$. Similarly, the solution of $(D + 5I)y = 0$ is $y_2 = e^{-5x}$. This is a basis of $P(D)y = 0$ on any interval. From the factorization we obtain the ODE, as expected,

$$\begin{aligned} (D - 8I)(D + 5I)y &= (D - 8I)(y' + 5y) = D(y' + 5y) - 8(y' + 5y) \\ &= y'' + 5y' - 8y' - 40y = y'' - 3' - 40y = 0. \end{aligned}$$

Verify that this agrees with the result of our method in Sec. 2.2. This is not unexpected because we factored $P(D)$ in the same way as the characteristic polynomial $P(\lambda) = \lambda^2 - 3\lambda - 40$. ■

It was essential that L in (2) had *constant* coefficients. Extension of operator methods to variable-coefficient ODEs is more difficult and will not be considered here.

If operational methods were limited to the simple situations illustrated in this section, it would perhaps not be worth mentioning. Actually, the power of the operator approach appears in more complicated engineering problems, as we shall see in Chap. 6.

PROBLEM SET 2.3

1–5 APPLICATION OF DIFFERENTIAL OPERATORS

Apply the given operator to the given functions. Show all steps in detail.

1. $D^2 + 2D$; $\cosh 2x$, $e^{-x} + e^{2x}$, $\cos x$
2. $D - 3I$; $3x^2 + 3x$, $3e^{3x}$, $\cos 4x - \sin 4x$
3. $(D - 2I)^2$; e^{2x}, xe^{2x}, e^{-2x}
4. $(D + 6I)^2$; $6x + \sin 6x$, xe^{-6x}
5. $(D - 2I)(D + 3I)$; e^{2x}, xe^{2x}, e^{-3x}

6–12 GENERAL SOLUTION

Factor as in the text and solve.

6. $(D^2 + 4.00D + 3.36I)y = 0$
7. $(4D^2 - I)y = 0$
8. $(D^2 + 3I)y = 0$
9. $(D^2 - 4.20D + 4.41I)y = 0$
10. $(D^2 + 4.80D + 5.76I)y = 0$
11. $(D^2 - 4.00D + 3.84I)y = 0$
12. $(D^2 + 3.0D + 2.5I)y = 0$

13. Linear operator. Illustrate the linearity of L in (2) by taking $c = 4, k = -6, y = e^{2x}$, and $w = \cos 2x$. Prove that L is linear.

14. Double root. If $D^2 + aD + bI$ has distinct roots μ and λ, show that a particular solution is $y = (e^{\mu x} - e^{\lambda x})/(\mu - \lambda)$. Obtain from this a solution $xe^{\lambda x}$ by letting $\mu \to \lambda$ and applying l'Hôpital's rule.

15. Definition of linearity. Show that the definition of linearity in the text is equivalent to the following. If $L[y]$ and $L[w]$ exist, then $L[y + w]$ exists and $L[cy]$ and $L[kw]$ exist for all constants c and k, and $L[y + w] = L[y] + L[w]$ as well as $L[cy] = cL[y]$ and $L[kw] = kL[w]$.

2.4 Modeling of Free Oscillations of a Mass–Spring System

Linear ODEs with constant coefficients have important applications in mechanics, as we show in this section as well as in Sec. 2.8, and in electrical circuits as we show in Sec. 2.9. In this section we model and solve a basic mechanical system consisting of a mass on an elastic spring (a so-called "mass–spring system," Fig. 33), which moves up and down.

Setting Up the Model

We take an ordinary coil spring that resists extension as well as compression. We suspend it vertically from a fixed support and attach a body at its lower end, for instance, an iron ball, as shown in Fig. 33. We let $y = 0$ denote the position of the ball when the system is at rest (Fig. 33b). Furthermore, we choose ***the downward direction as positive***, thus regarding downward forces as *positive* and upward forces as *negative*.

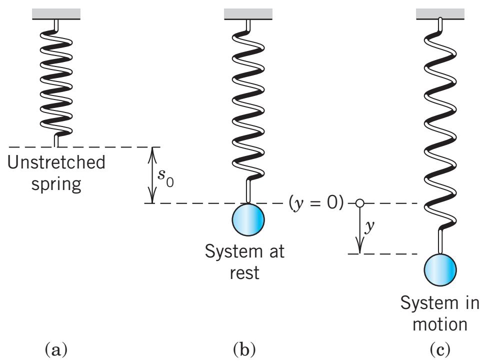


Fig. 33. Mechanical mass–spring system

We now let the ball move, as follows. We pull it down by an amount $y > 0$ (Fig. 33c). This causes a spring force

$$F_1 = -ky \qquad \textbf{(1)} \qquad \textbf{(Hooke's law}^2\textbf{)}$$

proportional to the stretch y, with $k\,(>0)$ called the **spring constant**. The minus sign indicates that F_1 points upward, against the displacement. It is a *restoring force*: It wants to restore the system, that is, to pull it back to $y = 0$. Stiff springs have large k.

[2]ROBERT HOOKE (1635–1703), English physicist, a forerunner of Newton with respect to the law of gravitation.

Note that an additional force $-F_0$ is present in the spring, caused by stretching it in fastening the ball, but F_0 has no effect on the motion because it is in equilibrium with the weight W of the ball, $-F_0 = W = mg$, where $g = 980 \text{ cm/sec}^2 = 9.8 \text{ m/sec}^2 = 32.17 \text{ ft/sec}^2$ is the **constant of gravity at the Earth's surface** (not to be confused with the *universal gravitational constant* $G = gR^2/M = 6.67 \cdot 10^{-11} \text{ nt m}^2/\text{kg}^2$, which we shall not need; here $R = 6.37 \cdot 10^6$ m and $M = 5.98 \cdot 10^{24}$ kg are the Earth's radius and mass, respectively).

The motion of our mass–spring system is determined by **Newton's second law**

$$\text{Mass} \times \text{Acceleration} = my'' = \text{Force} \tag{2}$$

where $y'' = d^2y/dt^2$ and "Force" is the resultant of all the forces acting on the ball. (For systems of units, see the inside of the front cover.)

ODE of the Undamped System

Every system has damping. Otherwise it would keep moving forever. But if the damping is small and the motion of the system is considered over a relatively short time, we may disregard damping. Then Newton's law with $F = -F_1$ gives the model $my'' = -F_1 = -ky$; thus

$$my'' + ky = 0. \tag{3}$$

This is a homogeneous linear ODE with constant coefficients. A general solution is obtained as in Sec. 2.2, namely (see Example 6 in Sec. 2.2)

$$y(t) = A \cos \omega_0 t + B \sin \omega_0 t \qquad \omega_0 = \sqrt{\frac{k}{m}}. \tag{4}$$

This motion is called a **harmonic oscillation** (Fig. 34). Its *frequency* is $f = \omega_0/2\pi$ Hertz[3] (= cycles/sec) because cos and sin in (4) have the period $2\pi/\omega_0$. The frequency f is called the **natural frequency** of the system. (We write ω_0 to reserve ω for Sec. 2.8.)

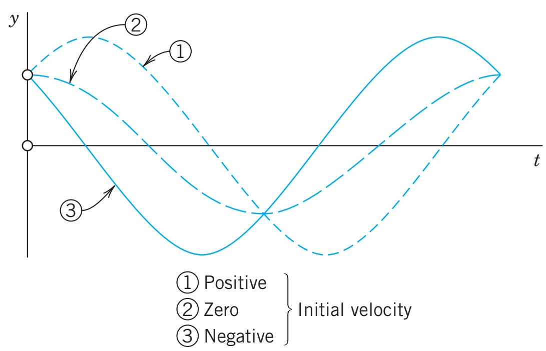


Fig. 34. Typical harmonic oscillations (4) and (4*) with the same $y(0) = A$ and different initial velocities $y'(0) = \omega_0 B$, positive 1, zero 2, negative 3

[3]**HEINRICH HERTZ (1857–1894)**, German physicist, who discovered electromagnetic waves, as the basis of wireless communication developed by GUGLIELMO MARCONI (1874–1937), Italian physicist (Nobel prize in 1909).

An alternative representation of (4), which shows the physical characteristics of amplitude and phase shift of (4), is

(4*) $$y(t) = C\cos(\omega_0 t - \delta)$$

with $C = \sqrt{A^2 + B^2}$ and phase angle δ, where $\tan\delta = B/A$. This follows from the addition formula (6) in App. 3.1.

EXAMPLE 1 Harmonic Oscillation of an Undamped Mass–Spring System

If a mass–spring system with an iron ball of weight $W = 98$ nt (about 22 lb) can be regarded as undamped, and the spring is such that the ball stretches it 1.09 m (about 43 in.), how many cycles per minute will the system execute? What will its motion be if we pull the ball down from rest by 16 cm (about 6 in.) and let it start with zero initial velocity?

Solution. Hooke's law (1) with W as the force and 1.09 meter as the stretch gives $W = 1.09k$; thus $k = W/1.09 = 98/1.09 = 90$ [kg/sec^2] $= 90$ [nt/meter]. The mass is $m = W/g = 98/9.8 = 10$ [kg]. This gives the frequency $\omega_0/(2\pi) = \sqrt{k/m}/(2\pi) = 3/(2\pi) = 0.48$ [Hz] $= 29$ [cycles/min].

From (4) and the initial conditions, $y(0) = A = 0.16$ [meter] and $y'(0) = \omega_0 B = 0$. Hence the motion is

$$y(t) = 0.16\cos 3t \text{ [meter]} \qquad \text{or} \qquad 0.52\cos 3t \text{ [ft]} \qquad \text{(Fig. 35).}$$

If you have a chance of experimenting with a mass–spring system, don't miss it. You will be surprised about the good agreement between theory and experiment, usually within a fraction of one percent if you measure carefully. ■

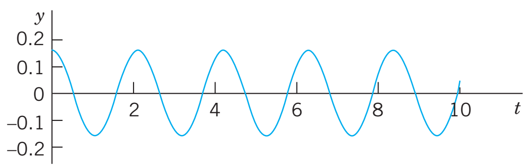


Fig. 35. Harmonic oscillation in Example 1

ODE of the Damped System

To our model $my'' = -ky$ we now add a damping force

$$F_2 = -cy',$$

obtaining $my'' = -ky - cy'$; thus the ODE of the damped mass–spring system is

(5) $$my'' + cy' + ky = 0. \qquad \text{(Fig. 36)}$$

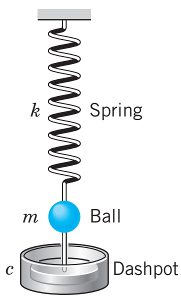


Fig. 36. Damped system

Physically this can be done by connecting the ball to a dashpot; see Fig. 36. We assume this damping force to be proportional to the velocity $y' = dy/dt$. This is generally a good approximation for small velocities.

The constant c is called the *damping constant*. Let us show that c is positive. Indeed, the damping force $F_2 = -cy'$ acts *against* the motion; hence for a downward motion we have $y' > 0$ which for positive c makes F negative (an upward force), as it should be. Similarly, for an upward motion we have $y' < 0$ which, for $c > 0$ makes F_2 positive (a downward force).

The ODE (5) is homogeneous linear and has constant coefficients. Hence we can solve it by the method in Sec. 2.2. The characteristic equation is (divide (5) by m)

$$\lambda^2 + \frac{c}{m}\lambda + \frac{k}{m} = 0.$$

By the usual formula for the roots of a quadratic equation we obtain, as in Sec. 2.2,

$$\text{(6)} \quad \lambda_1 = -\alpha + \beta, \quad \lambda_2 = -\alpha - \beta, \quad \text{where} \quad \alpha = \frac{c}{2m} \quad \text{and} \quad \beta = \frac{1}{2m}\sqrt{c^2 - 4mk}.$$

It is now interesting that depending on the amount of damping present—whether a lot of damping, a medium amount of damping or little damping—three types of motions occur, respectively:

Case I.	$c^2 > 4mk$.	*Distinct real roots* λ_1, λ_2.	**(Overdamping)**
Case II.	$c^2 = 4mk$.	*A real double root.*	**(Critical damping)**
Case III.	$c^2 < 4mk$.	*Complex conjugate roots.*	**(Underdamping)**

They correspond to the three Cases I, II, III in Sec. 2.2.

Discussion of the Three Cases

Case I. Overdamping

If the damping constant c is so large that $c^2 > 4mk$, then λ_1 and λ_2 are distinct real roots. In this case the corresponding general solution of (5) is

$$\textbf{(7)} \qquad \boxed{y(t) = c_1 e^{-(\alpha-\beta)t} + c_2 e^{-(\alpha+\beta)t}.}$$

We see that in this case, damping takes out energy so quickly that the body does not oscillate. For $t > 0$ both exponents in (7) are negative because $\alpha > 0, \beta > 0$, and $\beta^2 = \alpha^2 - k/m < \alpha^2$. Hence both terms in (7) approach zero as $t \to \infty$. Practically speaking, after a sufficiently long time the mass will be at rest at the *static equilibrium position* ($y = 0$). Figure 37 shows (7) for some typical initial conditions.

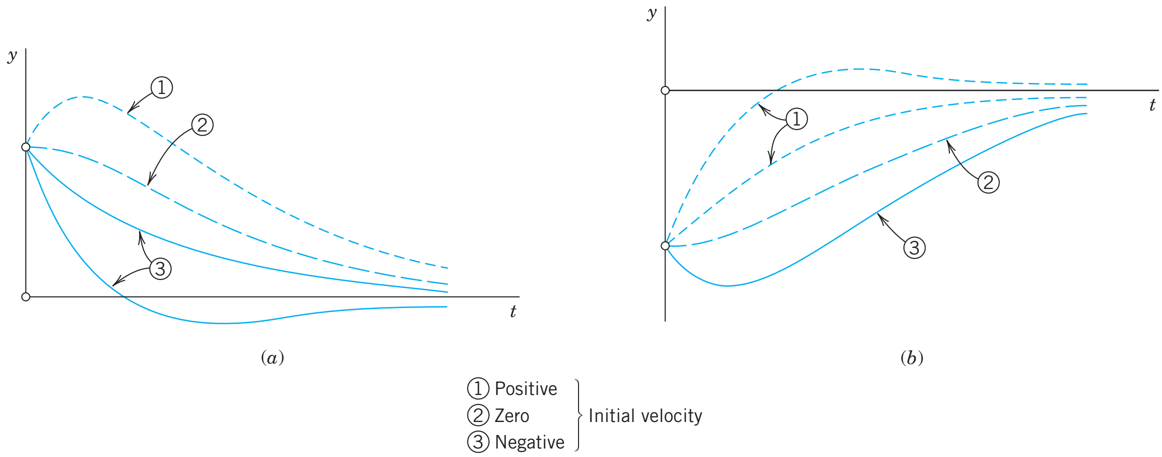


Fig. 37. Typical motions (7) in the overdamped case
(*a*) Positive initial displacement
(*b*) Negative initial displacement

Case II. Critical Damping

Critical damping is the border case between nonoscillatory motions (Case I) and oscillations (Case III). It occurs if the characteristic equation has a double root, that is, if $c^2 = 4mk$, so that $\beta = 0$, $\lambda_1 = \lambda_2 = -\alpha$. Then the corresponding general solution of (5) is

$$y(t) = (c_1 + c_2 t)e^{-\alpha t}. \tag{8}$$

This solution can pass through the equilibrium position $y = 0$ at most once because $e^{-\alpha t}$ is never zero and $c_1 + c_2 t$ can have at most one positive zero. If both c_1 and c_2 are positive (or both negative), it has no positive zero, so that y does not pass through 0 at all. Figure 38 shows typical forms of (8). Note that they look almost like those in the previous figure.

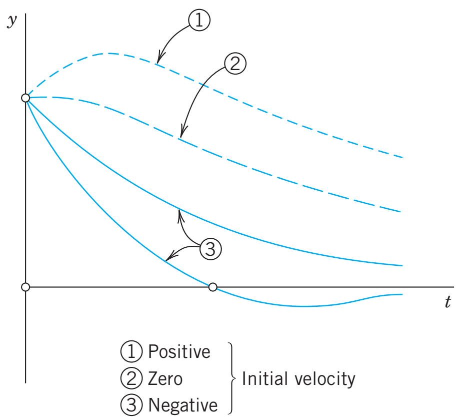


Fig. 38. Critical damping [see (8)]

Case III. Underdamping

This is the most interesting case. It occurs if the damping constant c is so small that $c^2 < 4mk$. Then β in (6) is no longer real but pure imaginary, say,

$$\beta = i\omega^* \qquad \text{where} \qquad \omega^* = \frac{1}{2m}\sqrt{4mk - c^2} = \sqrt{\frac{k}{m} - \frac{c^2}{4m^2}} \qquad (>0). \tag{9}$$

(We now write ω^* to reserve ω for driving and electromotive forces in Secs. 2.8 and 2.9.) The roots of the characteristic equation are now complex conjugates,

$$\lambda_1 = -\alpha + i\omega^*, \qquad \lambda_2 = -\alpha - i\omega^*$$

with $\alpha = c/(2m)$, as given in (6). Hence the corresponding general solution is

$$y(t) = e^{-\alpha t}(A \cos \omega^* t + B \sin \omega^* t) = Ce^{-\alpha t} \cos (\omega^* t - \delta) \tag{10}$$

where $C^2 = A^2 + B^2$ and $\tan \delta = B/A$, as in (4*).

This represents **damped oscillations**. Their curve lies between the dashed curves $y = Ce^{-\alpha t}$ and $y = -Ce^{-\alpha t}$ in Fig. 39, touching them when $\omega^* t - \delta$ is an integer multiple of π because these are the points at which $\cos (\omega^* t - \delta)$ equals 1 or -1.

The frequency is $\omega^*/(2\pi)$ Hz (hertz, cycles/sec). From (9) we see that the smaller c (>0) is, the larger is ω^* and the more rapid the oscillations become. If c approaches 0, then ω^* approaches $\omega_0 = \sqrt{k/m}$, giving the harmonic oscillation (4), whose frequency $\omega_0/(2\pi)$ is the natural frequency of the system.

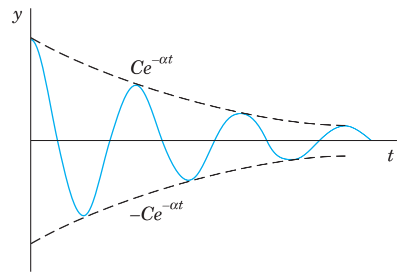


Fig. 39. Damped oscillation in Case III [see (10)]

EXAMPLE 2 The Three Cases of Damped Motion

How does the motion in Example 1 change if we change the damping constant c from one to another of the following three values, with $y(0) = 0.16$ and $y'(0) = 0$ as before?

$$\text{(I) } c = 100 \text{ kg/sec}, \qquad \text{(II) } c = 60 \text{ kg/sec}, \qquad \text{(III) } c = 10 \text{ kg/sec}.$$

Solution. It is interesting to see how the behavior of the system changes due to the effect of the damping, which takes energy from the system, so that the oscillations decrease in amplitude (Case III) or even disappear (Cases II and I).

(I) With $m = 10$ and $k = 90$, as in Example 1, the model is the initial value problem

$$10y'' + 100y' + 90y = 0, \qquad y(0) = 0.16 \text{ [meter]}, \qquad y'(0) = 0.$$

The characteristic equation is $10\lambda^2 + 100\lambda + 90 = 10(\lambda + 9)(\lambda + 1) = 0$. It has the roots -9 and -1. This gives the general solution

$$y = c_1 e^{-9t} + c_2 e^{-t}. \qquad \text{We also need} \qquad y' = -9c_1 e^{-9t} - c_2 e^{-t}.$$

The initial conditions give $c_1 + c_2 = 0.16$, $-9c_1 - c_2 = 0$. The solution is $c_1 = -0.02$, $c_2 = 0.18$. Hence in the overdamped case the solution is

$$y = -0.02e^{-9t} + 0.18e^{-t}.$$

It approaches 0 as $t \to \infty$. The approach is rapid; after a few seconds the solution is practically 0, that is, the iron ball is at rest.

(II) The model is as before, with $c = 60$ instead of 100. The characteristic equation now has the form $10\lambda^2 + 60\lambda + 90 = 10(\lambda + 3)^2 = 0$. It has the double root -3. Hence the corresponding general solution is

$$y = (c_1 + c_2 t)e^{-3t}. \qquad \text{We also need} \qquad y' = (c_2 - 3c_1 - 3c_2 t)e^{-3t}.$$

The initial conditions give $y(0) = c_1 = 0.16$, $y'(0) = c_2 - 3c_1 = 0$, $c_2 = 0.48$. Hence in the critical case the solution is

$$y = (0.16 + 0.48t)e^{-3t}.$$

It is always positive and decreases to 0 in a monotone fashion.

(III) The model now is $10y'' + 10y' + 90y = 0$. Since $c = 10$ is smaller than the critical c, we shall get oscillations. The characteristic equation is $10\lambda^2 + 10\lambda + 90 = 10[(\lambda + \frac{1}{2})^2 + 9 - \frac{1}{4}] = 0$. It has the complex roots [see (4) in Sec. 2.2 with $a = 1$ and $b = 9$]

$$\lambda = -0.5 \pm \sqrt{0.5^2 - 9} = -0.5 \pm 2.96i.$$

This gives the general solution

$$y = e^{-0.5t}(A \cos 2.96t + B \sin 2.96t).$$

Thus $y(0) = A = 0.16$. We also need the derivative

$$y' = e^{-0.5t}(-0.5A \cos 2.96t - 0.5B \sin 2.96t - 2.96A \sin 2.96t + 2.96B \cos 2.96t).$$

Hence $y'(0) = -0.5A + 2.96B = 0$, $\quad B = 0.5A/2.96 = 0.027$. This gives the solution

$$y = e^{-0.5t}(0.16 \cos 2.96t + 0.027 \sin 2.96t) = 0.162e^{-0.5t} \cos (2.96t - 0.17).$$

We see that these damped oscillations have a smaller frequency than the harmonic oscillations in Example 1 by about 1% (since 2.96 is smaller than 3.00 by about 1%). Their amplitude goes to zero. See Fig. 40. ■

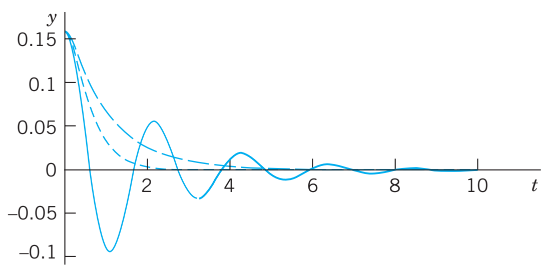


Fig. 40. The three solutions in Example 2

This section concerned ***free motions*** of mass–spring systems. Their models are *homogeneous* linear ODEs. *Nonhomogeneous* linear ODEs will arise as models of **forced motions**, that is, motions under the influence of a "driving force." We shall study them in Sec. 2.8, after we have learned how to solve those ODEs.

PROBLEM SET 2.4

1–10 HARMONIC OSCILLATIONS (UNDAMPED MOTION)

1. **Initial value problem.** Find the harmonic motion (4) that starts from y_0 with initial velocity v_0. Graph or sketch the solutions for $\omega_0 = \pi$, $y_0 = 1$, and various v_0 of your choice on common axes. At what t-values do all these curves intersect? Why?
2. **Frequency.** If a weight of 20 nt (about 4.5 lb) stretches a certain spring by 2 cm, what will the frequency of the corresponding harmonic oscillation be? The period?
3. **Frequency.** How does the frequency of the harmonic oscillation change if we (i) double the mass, (ii) take a spring of twice the modulus? First find qualitative answers by physics, then look at formulas.
4. **Initial velocity.** Could you make a harmonic oscillation move faster by giving the body a greater initial push?
5. **Springs in parallel.** What are the frequencies of vibration of a body of mass $m = 5$ kg (i) on a spring of modulus $k_1 = 20$ nt/m, (ii) on a spring of modulus $k_2 = 45$ nt/m, (iii) on the two springs in parallel? See Fig. 41.

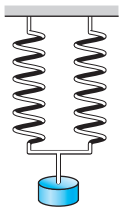


Fig. 41. Parallel springs (Problem 5)

6. **Spring in series.** If a body hangs on a spring s_1 of modulus $k_1 = 8$, which in turn hangs on a spring s_2 of modulus $k_2 = 12$, what is the modulus k of this combination of springs?
7. **Pendulum.** Find the frequency of oscillation of a pendulum of length L (Fig. 42), neglecting air resistance and the weight of the rod, and assuming θ to be so small that $\sin\theta$ practically equals θ.

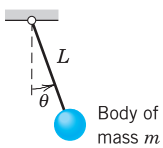


Fig. 42. Pendulum (Problem 7)

8. **Archimedian principle.** This principle states that the buoyancy force equals the weight of the water displaced by the body (partly or totally submerged). The cylindrical buoy of diameter 60 cm in Fig. 43 is floating in water with its axis vertical. When depressed downward in the water and released, it vibrates with period 2 sec. What is its weight?

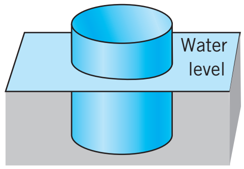


Fig. 43. Buoy (Problem 8)

9. **Vibration of water in a tube.** If 1 liter of water (about 1.06 US quart) is vibrating up and down under the influence of gravitation in a U-shaped tube of diameter 2 cm (Fig. 44), what is the frequency? Neglect friction. First guess.

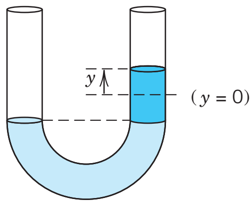


Fig. 44. Tube (Problem 9)

10. **TEAM PROJECT. Harmonic Motions of Similar Models.** The ***unifying power of mathematical methods*** results to a large extent from the fact that different physical (or other) systems may have the same or very similar models. Illustrate this for the following three systems

    (a) Pendulum clock. A clock has a 1-meter pendulum. The clock ticks once for each time the pendulum completes a full swing, returning to its original position. How many times a minute does the clock tick?

    (b) Flat spring (Fig. 45). The harmonic oscillations of a flat spring with a body attached at one end and horizontally clamped at the other are also governed by (3). Find its motions, assuming that the body weighs 8 nt (about 1.8 lb), the system has its static equilibrium 1 cm below the horizontal line, and we let it start from this position with initial velocity 10 cm/sec.

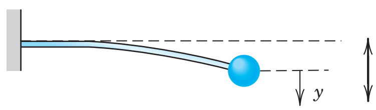


Fig. 45. Flat spring

(c) Torsional vibrations (Fig. 46). Undamped torsional vibrations (rotations back and forth) of a wheel attached to an elastic thin rod or wire are governed by the equation $I_0\theta'' + K\theta = 0$, where θ is the angle measured from the state of equilibrium. Solve this equation for $K/I_0 = 13.69 \text{ sec}^{-2}$, initial angle $30°(= 0.5235 \text{ rad})$ and initial angular velocity $20° \text{ sec}^{-1} (= 0.349 \text{ rad} \cdot \text{sec}^{-1})$.

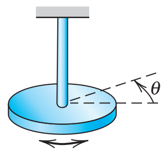


Fig. 46. Torsional vibrations

11–20 DAMPED MOTION

11. Overdamping. Show that for (7) to satisfy initial conditions $y(0) = y_0$ and $v(0) = v_0$ we must have $c_1 = [(1 + \alpha/\beta)y_0 + v_0/\beta]/2$ and $c_2 = [(1 - \alpha/\beta)y_0 - v_0/\beta]/2$.

12. Overdamping. Show that in the overdamped case, the body can pass through $y = 0$ at most once (Fig. 37).

13. Initial value problem. Find the critical motion (8) that starts from y_0 with initial velocity v_0. Graph solution curves for $\alpha = 1$, $y_0 = 1$ and several v_0 such that (i) the curve does not intersect the t-axis, (ii) it intersects it at $t = 1, 2, \dots, 5$, respectively.

14. Shock absorber. What is the smallest value of the damping constant of a shock absorber in the suspension of a wheel of a car (consisting of a spring and an absorber) that will provide (theoretically) an oscillation-free ride if the mass of the car is 2000 kg and the spring constant equals 4500 kg/sec^2?

15. Frequency. Find an approximation formula for ω^* in terms of ω_0 by applying the binomial theorem in (9) and retaining only the first two terms. How good is the approximation in Example 2, III?

16. Maxima. Show that the maxima of an underdamped motion occur at equidistant t-values and find the distance.

17. Underdamping. Determine the values of t corresponding to the maxima and minima of the oscillation $y(t) = e^{-t} \sin t$. Check your result by graphing $y(t)$.

18. Logarithmic decrement. Show that the ratio of two consecutive maximum amplitudes of a damped oscillation (10) is constant, and the natural logarithm of this ratio called the *logarithmic decrement*, equals $\Delta = 2\pi\alpha/\omega^*$. Find Δ for the solutions of $y'' + 2y' + 5y = 0$.

19. Damping constant. Consider an underdamped motion of a body of mass $m = 0.5$ kg. If the time between two consecutive maxima is 3 sec and the maximum amplitude decreases to $\frac{1}{2}$ its initial value after 10 cycles, what is the damping constant of the system?

20. CAS PROJECT. Transition Between Cases I, II, III. Study this transition in terms of graphs of typical solutions. (Cf. Fig. 47.)

(a) ***Avoiding unnecessary generality is part of good modeling.*** Show that the initial value problems (A) and (B),

$$\text{(A)} \quad y'' + cy' + y = 0, \quad y(0) = 1, \quad y'(0) = 0$$

(B) the same with different c and $y'(0) = -2$ (instead of 0), will give practically as much information as a problem with other m, k, $y(0)$, $y'(0)$.

(b) ***Consider*** **(A).** Choose suitable values of c, perhaps better ones than in Fig. 47, for the transition from Case III to II and I. Guess c for the curves in the figure.

(c) ***Time to go to rest.*** Theoretically, this time is infinite (why?). Practically, the system is at rest when its motion has become very small, say, less than 0.1% of the initial displacement (this choice being up to us), that is in our case,

$$(11) \quad |y(t)| < 0.001 \quad \text{for all } t \text{ greater than some } t_1.$$

In engineering constructions, damping can often be varied without too much trouble. Experimenting with your graphs, find empirically a relation between t_1 and c.

(d) ***Solve*** **(A)** ***analytically.*** Give a reason why the solution c of $y(t_2) = -0.001$, with t_2 the solution of $y'(t) = 0$, will give you the best possible c satisfying (11).

(e) Consider (B) empirically as in (a) and (b). What is the main difference between (B) and (A)?

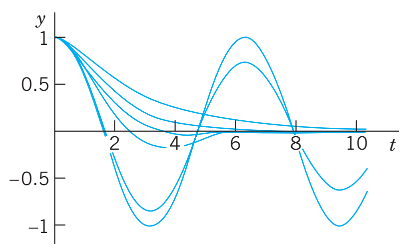


Fig. 47. CAS Project 20

2.5 Euler–Cauchy Equations

Euler–Cauchy equations[4] are ODEs of the form

$$x^2y'' + axy' + by = 0 \tag{1}$$

with given constants a and b and unknown function $y(x)$. We substitute

$$y = x^m, \qquad y' = mx^{m-1}, \qquad y'' = m(m-1)x^{m-2}$$

into (1). This gives

$$x^2m(m-1)x^{m-2} + axmx^{m-1} + bx^m = 0$$

and we now see that $y = x^m$ was a rather natural choice because we have obtained a common factor x^m. Dropping it, we have the auxiliary equation $m(m-1) + am + b = 0$ or

$$m^2 + (a-1)m + b = 0. \qquad \text{(Note: } a-1, \text{ not } a.) \tag{2}$$

Hence $y = x^m$ is a solution of (1) if and only if m is a root of (2). The roots of (2) are

$$m_1 = \tfrac{1}{2}(1-a) + \sqrt{\tfrac{1}{4}(1-a)^2 - b}, \qquad m_2 = \tfrac{1}{2}(1-a) - \sqrt{\tfrac{1}{4}(1-a)^2 - b}. \tag{3}$$

Case I. Real different roots m_1 and m_2 give two real solutions

$$y_1(x) = x^{m_1} \qquad \text{and} \qquad y_2(x) = x^{m_2}.$$

These are linearly independent since their quotient is not constant. Hence they constitute a basis of solutions of (1) for all x for which they are real. The corresponding general solution for all these x is

$$y = c_1x^{m_1} + c_2x^{m_2} \qquad (c_1, c_2 \text{ arbitrary}). \tag{4}$$

EXAMPLE 1 General Solution in the Case of Different Real Roots

The Euler–Cauchy equation $x^2y'' + 1.5xy' - 0.5y = 0$ has the auxiliary equation $m^2 + 0.5m - 0.5 = 0$. The roots are 0.5 and -1. Hence a basis of solutions for all positive x is $y_1 = x^{0.5}$ and $y_2 = 1/x$ and gives the general solution

$$y = c_1\sqrt{x} + \frac{c_2}{x} \qquad (x > 0). \qquad \blacksquare$$

[4]**LEONHARD EULER (1707–1783)** was an enormously creative Swiss mathematician. He made fundamental contributions to almost all branches of mathematics and its application to physics. His important books on algebra and calculus contain numerous basic results of his own research. The great French mathematician AUGUSTIN LOUIS CAUCHY (1789–1857) is the father of modern analysis. He is the creator of complex analysis and had great influence on ODEs, PDEs, infinite series, elasticity theory, and optics.

Case II. A real double root $m_1 = \frac{1}{2}(1 - a)$ occurs if and only if $b = \frac{1}{4}(a - 1)^2$ because then (2) becomes $[m + \frac{1}{2}(a - 1)]^2$, as can be readily verified. Then a solution is $y_1 = x^{(1-a)/2}$, and (1) is of the form

$$(5) \qquad x^2y'' + axy' + \tfrac{1}{4}(1 - a)^2 y = 0 \qquad \text{or} \qquad y'' + \frac{a}{x}y' + \frac{(1-a)^2}{4x^2}y = 0.$$

A second linearly independent solution can be obtained by the method of reduction of order from Sec. 2.1, as follows. Starting from $y_2 = uy_1$, we obtain for u the expression (9) Sec. 2.1, namely,

$$u = \int U\,dx \qquad \text{where} \qquad U = \frac{1}{y_1^2}\exp\left(-\int p\,dx\right).$$

From (5) in standard form (second ODE) we see that $p = a/x$ (not ax; this is essential!). Hence $\exp\int(-p\,dx) = \exp(-a \ln x) = \exp(\ln x^{-a}) = 1/x^a$. Division by $y_1^2 = x^{1-a}$ gives $U = 1/x$, so that $u = \ln x$ by integration. Thus, $y_2 = uy_1 = y_1 \ln x$, and y_1 and y_2 are linearly independent since their quotient is not constant. The general solution corresponding to this basis is

$$\textbf{(6)} \qquad \boxed{y = (c_1 + c_2 \ln x)\,x^m,} \qquad m = \tfrac{1}{2}(1 - a).$$

EXAMPLE 2 General Solution in the Case of a Double Root

The Euler–Cauchy equation $x^2y'' - 5xy' + 9y = 0$ has the auxiliary equation $m^2 - 6m + 9 = 0$. It has the double root $m = 3$, so that a general solution for all positive x is

$$y = (c_1 + c_2 \ln x)\,x^3.$$

■

Case III. Complex conjugate roots are of minor practical importance, and we discuss the derivation of real solutions from complex ones just in terms of a typical example.

EXAMPLE 3 Real General Solution in the Case of Complex Roots

The Euler–Cauchy equation $x^2y'' + 0.6xy' + 16.04y = 0$ has the auxiliary equation $m^2 - 0.4m + 16.04 = 0$. The roots are complex conjugate, $m_1 = 0.2 + 4i$ and $m_2 = 0.2 - 4i$, where $i = \sqrt{-1}$. We now use the trick of writing $x = e^{\ln x}$ and obtain

$$x^{m_1} = x^{0.2+4i} = x^{0.2}(e^{\ln x})^{4i} = x^{0.2}e^{(4 \ln x)i},$$
$$x^{m_2} = x^{0.2-4i} = x^{0.2}(e^{\ln x})^{-4i} = x^{0.2}e^{-(4 \ln x)i}.$$

Next we apply Euler's formula (11) in Sec. 2.2 with $t = 4 \ln x$ to these two formulas. This gives

$$x^{m_1} = x^{0.2}[\cos(4 \ln x) + i \sin(4 \ln x)],$$
$$x^{m_2} = x^{0.2}[\cos(4 \ln x) - i \sin(4 \ln x)].$$

We now add these two formulas, so that the sine drops out, and divide the result by 2. Then we subtract the second formula from the first, so that the cosine drops out, and divide the result by $2i$. This yields

$$x^{0.2}\cos(4 \ln x) \qquad \text{and} \qquad x^{0.2}\sin(4 \ln x)$$

respectively. By the superposition principle in Sec. 2.2 these are solutions of the Euler–Cauchy equation (1). Since their quotient $\cot(4 \ln x)$ is not constant, they are linearly independent. Hence they form a basis of solutions, and the corresponding real general solution for all positive x is

$$(8) \qquad y = x^{0.2}[A \cos(4 \ln x) + B \sin(4 \ln x)].$$

Figure 48 shows typical solution curves in the three cases discussed, in particular the real basis functions in Examples 1 and 3. ■

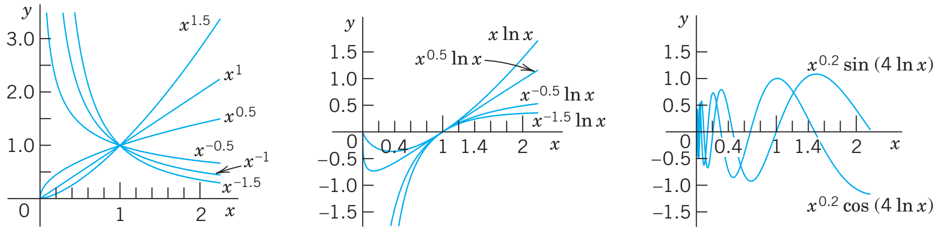


Fig. 48. Euler–Cauchy equations

EXAMPLE 4 **Boundary Value Problem. Electric Potential Field Between Two Concentric Spheres**

Find the electrostatic potential $v = v(r)$ between two concentric spheres of radii $r_1 = 5$ cm and $r_2 = 10$ cm kept at potentials $v_1 = 110$ V and $v_2 = 0$, respectively.

Physical Information. $v(r)$ is a solution of the Euler–Cauchy equation $rv'' + 2v' = 0$, where $v' = dv/dr$.

Solution. The auxiliary equation is $m^2 + m = 0$. It has the roots 0 and -1. This gives the general solution $v(r) = c_1 + c_2/r$. From the "boundary conditions" (the potentials on the spheres) we obtain

$$v(5) = c_1 + \frac{c_2}{5} = 110. \qquad v(10) = c_1 + \frac{c_2}{10} = 0.$$

By subtraction, $c_2/10 = 110$, $c_2 = 1100$. From the second equation, $c_1 = -c_2/10 = -110$. *Answer:* $v(r) = -110 + 1100/r$ V. Figure 49 shows that the potential is not a straight line, as it would be for a potential between two parallel plates. For example, on the sphere of radius 7.5 cm it is not $110/2 = 55$ V, but considerably less. (What is it?) ■

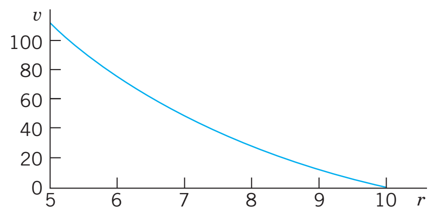


Fig. 49. Potential $v(r)$ in Example 4

PROBLEM SET 2.5

1. **Double root.** Verify directly by substitution that $x^{(1-a)/2} \ln x$ is a solution of (1) if (2) has a double root, but $x^{m_1} \ln x$ and $x^{m_2} \ln x$ are not solutions of (1) if the roots m_1 and m_2 of (2) are different.

2–11 GENERAL SOLUTION

Find a real general solution. Show the details of your work.

2. $x^2y'' - 20y = 0$
3. $5x^2y'' + 23xy' + 16.2y = 0$
4. $xy'' + 2y' = 0$
5. $4x^2y'' + 5y = 0$
6. $x^2y'' + 0.7xy' - 0.1y = 0$
7. $(x^2D^2 - 4xD + 6I)y = C$
8. $(x^2D^2 - 3xD + 4I)y = 0$
9. $(x^2D^2 - 0.2xD + 0.36I)y = 0$
10. $(x^2D^2 - xD + 5I)y = 0$
11. $(x^2D^2 - 3xD + 10I)y = 0$

12–19 **INITIAL VALUE PROBLEM**

Solve and graph the solution. Show the details of your work.

12. $x^2y'' - 4xy' + 6y = 0, \quad y(1) = 0.4, \quad y'(1) = 0$

13. $x^2y'' + 3xy' + 0.75y = 0, \quad y(1) = 1,$ $y'(1) = -1.5$

14. $x^2y'' + xy' + 9y = 0, \quad y(1) = 0, \quad y'(1) = 2.5$

15. $x^2y'' + 3xy' + y = 0, \quad y(1) = 3.6, \quad y'(1) = 0.4$

16. $(x^2D^2 - 3xD + 4I)y = 0, \quad y(1) = -\pi, \quad y'(1) = 2\pi$

17. $(x^2D^2 + xD + I)y = 0, \quad y(1) = 1, \quad y'(1) = 1$

18. $(9x^2D^2 + 3xD + I)y = 0, \quad y(1) = 1, \quad y'(1) = 0$

19. $(x^2D^2 - xD - 15I)y = 0, \quad y(1) = 0.1,$ $y'(1) = -4.5$

20. TEAM PROJECT. Double Root

(a) Derive a second linearly independent solution of (1) by reduction of order; but instead of using (9), Sec. 2.1, perform all steps directly for the present ODE (1).

(b) Obtain $x^m \ln x$ by considering the solutions x^m and x^{m+s} of a suitable Euler–Cauchy equation and letting $s \to 0$.

(c) Verify by substitution that $x^m \ln x$, $m = (1 - a)/2$, is a solution in the critical case.

(d) Transform the Euler–Cauchy equation (1) into an ODE with constant coefficients by setting $x = e^t$ $(x > 0)$.

(e) Obtain a second linearly independent solution of the Euler–Cauchy equation in the "critical case" from that of a constant-coefficient ODE.

2.6 Existence and Uniqueness of Solutions. Wronskian

In this section we shall discuss the general theory of homogeneous linear ODEs

$$y'' + p(x)y' + q(x)y = 0 \tag{1}$$

with continuous, but otherwise arbitrary, ***variable coefficients*** p and q. This will concern the existence and form of a general solution of (1) as well as the uniqueness of the solution of initial value problems consisting of such an ODE and two initial conditions

$$y(x_0) = K_0, \qquad y'(x_0) = K_1 \tag{2}$$

with given x_0, K_0, and K_1.

The two main results will be Theorem 1, stating that such an initial value problem always has a solution which is unique, and Theorem 4, stating that a general solution

$$y = c_1y_1 + c_2y_2 \qquad (c_1, c_2 \text{ arbitrary}) \tag{3}$$

includes all solutions. Hence *linear* ODEs with continuous coefficients have no *"singular solutions"* (solutions not obtainable from a general solution).

Clearly, no such theory was needed for constant-coefficient or Euler–Cauchy equations because everything resulted explicitly from our calculations.

Central to our present discussion is the following theorem.

THEOREM 1

Existence and Uniqueness Theorem for Initial Value Problems

If $p(x)$ and $q(x)$ are continuous functions on some open interval I (see Sec. 1.1) *and x_0 is in I, then the initial value problem consisting of* (1) *and* (2) *has a unique solution $y(x)$ on the interval I.*

The proof of existence uses the same prerequisites as the existence proof in Sec. 1.7 and will not be presented here; it can be found in Ref. [A11] listed in App. 1. Uniqueness proofs are usually simpler than existence proofs. But for Theorem 1, even the uniqueness proof is long, and we give it as an additional proof in App. 4.

Linear Independence of Solutions

Remember from Sec. 2.1 that a general solution on an open interval I is made up from a **basis** y_1, y_2 on I, that is, from a pair of linearly independent solutions on I. Here we call y_1, y_2 **linearly independent** on I if the equation

$$k_1 y_1(x) + k_2 y_2(x) = 0 \quad \text{on } I \qquad \text{implies} \qquad k_1 = 0, \quad k_2 = 0. \tag{4}$$

We call y_1, y_2 **linearly dependent** on I if this equation also holds for constants k_1, k_2 not both 0. In this case, and only in this case, y_1 and y_2 are proportional on I, that is (see Sec. 2.1),

$$\text{(a)} \quad y_1 = k y_2 \qquad \text{or} \qquad \text{(b)} \quad y_2 = l y_1 \qquad \text{for all on } I. \tag{5}$$

For our discussion the following criterion of linear independence and dependence of solutions will be helpful.

THEOREM 2

Linear Dependence and Independence of Solutions

Let the ODE (1) *have continuous coefficients $p(x)$ and $q(x)$ on an open interval I. Then two solutions y_1 and y_2 of* (1) *on I are linearly dependent on I if and only if their* "**Wronskian**"

$$W(y_1, y_2) = y_1 y_2' - y_2 y_1' \tag{6}$$

is 0 *at some x_0 in I. Furthermore, if $W = 0$ at an $x = x_0$ in I, then $W = 0$ on I; hence, if there is an x_1 in I at which W is not* 0*, then y_1, y_2 are linearly independent on I.*

PROOF **(a)** Let y_1 and y_2 be linearly dependent on I. Then (5a) or (5b) holds on I. If (5a) holds, then

$$W(y_1, y_2) = y_1 y_2' - y_2 y_1' = k y_2 y_2' - y_2 k y_2' = 0.$$

Similarly if (5b) holds.

(b) Conversely, we let $W(y_1, y_2) = 0$ for some $x = x_0$ and show that this implies linear dependence of y_1 and y_2 on I. We consider the linear system of equations in the unknowns k_1, k_2

$$\begin{aligned} k_1 y_1(x_0) + k_2 y_2(x_0) &= 0 \\ k_1 y_1'(x_0) + k_2 y_2'(x_0) &= 0. \end{aligned} \tag{7}$$

To eliminate k_2, multiply the first equation by y_2' and the second by $-y_2$ and add the resulting equations. This gives

$$k_1 y_1(x_0) y_2'(x_0) - k_1 y_1'(x_0) y_2(x_0) = k_1 W(y_1(x_0), y_2(x_0)) = 0.$$

Similarly, to eliminate k_1, multiply the first equation by $-y_1'$ and the second by y_1 and add the resulting equations. This gives

$$k_2 W(y_1(x_0), y_2(x_0)) = 0.$$

If W were not 0 at x_0, we could divide by W and conclude that $k_1 = k_2 = 0$. Since W is 0, division is not possible, and the system has a solution for which k_1 and k_2 are not both 0. Using *these numbers* k_1, k_2, we introduce the function

$$y(x) = k_1 y_1(x) + k_2 y_2(x).$$

Since (1) is homogeneous linear, Fundamental Theorem 1 in Sec. 2.1 (the superposition principle) implies that this function is a solution of (1) on I. From (7) we see that it satisfies the initial conditions $y(x_0) = 0$, $y'(x_0) = 0$. Now another solution of (1) satisfying the same initial conditions is $y^* \equiv 0$. Since the coefficients p and q of (1) are continuous, Theorem 1 applies and gives uniqueness, that is, $y \equiv y^*$, written out

$$k_1 y_1 + k_2 y_2 \equiv 0 \qquad \text{on } I.$$

Now since k_1 and k_2 are not both zero, this means linear dependence of y_1, y_2 on I.

(c) We prove the last statement of the theorem. If $W(x_0) = 0$ at an x_0 in I, we have linear dependence of y_1, y_2 on I by part (b), hence $W \equiv 0$ by part (a) of this proof. Hence in the case of linear dependence it cannot happen that $W(x_1) \neq 0$ at an x_1 in I. If it does happen, it thus implies linear independence as claimed. ■

For calculations, the following formulas are often simpler than (6).

$$\textbf{(6*)} \quad W(y_1, y_2) = (a) \quad \left(\frac{y_2}{y_1}\right)' y_1^2 \quad (y_1 \neq 0) \qquad \text{or} \qquad (b) \quad -\left(\frac{y_1}{y_2}\right)' y_2^2 \quad (y_2 \neq 0).$$

These formulas follow from the quotient rule of differentiation.

Remark. Determinants. Students familiar with second-order determinants may have noticed that

$$W(y_1, y_2) = \begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix} = y_1 y_2' - y_2 y_1'.$$

This determinant is called the *Wronski determinant*[5] or, briefly, the **Wronskian**, of two solutions y_1 and y_2 of (1), as has already been mentioned in (6). Note that its four entries occupy the same positions as in the linear system (7).

[5]Introduced by WRONSKI (JOSEF MARIA HÖNE, 1776–1853), Polish mathematician.

EXAMPLE 1 **Illustration of Theorem 2**

The functions $y_1 = \cos \omega x$ and $y_2 = \sin \omega x$ are solutions of $y'' + \omega^2 y = 0$. Their Wronskian is

$$W(\cos \omega x, \sin \omega x) = \begin{vmatrix} \cos \omega x & \sin \omega x \\ -\omega \sin \omega x & \omega \cos \omega x \end{vmatrix} = y_1 y_2' - y_2 y_1' = \omega \cos^2 \omega x + \omega \sin^2 \omega x = \omega.$$

Theorem 2 shows that these solutions are linearly independent if and only if $\omega \neq 0$. Of course, we can see this directly from the quotient $y_2/y_1 = \tan \omega x$. For $\omega = 0$ we have $y_2 = 0$, which implies linear dependence (why?). ■

EXAMPLE 2 **Illustration of Theorem 2 for a Double Root**

A general solution of $y'' - 2y' + y = 0$ on any interval is $y = (c_1 + c_2 x)e^x$. (Verify!). The corresponding Wronskian is not 0, which shows linear independence of e^x and xe^x on any interval. Namely,

$$W(x, xe^x) = \begin{vmatrix} e^x & xe^x \\ e^x & (x+1)e^x \end{vmatrix} = (x+1)e^{2x} - xe^{2x} = e^{2x} \neq 0. \quad ■$$

A General Solution of (1) Includes All Solutions

This will be our second main result, as announced at the beginning. Let us start with existence.

THEOREM 3 **Existence of a General Solution**

If $p(x)$ and $q(x)$ are continuous on an open interval I, then (1) *has a general solution on I.*

PROOF By Theorem 1, the ODE (1) has a solution $y_1(x)$ on I satisfying the initial conditions

$$y_1(x_0) = 1, \qquad y_1'(x_0) = 0$$

and a solution $y_2(x)$ on I satisfying the initial conditions

$$y_2(x_0) = 0, \qquad y_2'(x_0) = 1.$$

The Wronskian of these two solutions has at $x = x_0$ the value

$$W(y_1(0), y_2(0)) = y_1(x_0)y_2'(x_0) - y_2(x_0)y_1'(x_0) = 1.$$

Hence, by Theorem 2, these solutions are linearly independent on I. They form a basis of solutions of (1) on I, and $y = c_1 y_1 + c_2 y_2$ with arbitrary c_1, c_2 is a general solution of (1) on I, whose existence we wanted to prove. ■

We finally show that a general solution is as general as it can possibly be.

THEOREM 4

A General Solution Includes All Solutions

If the ODE (1) *has continuous coefficients* $p(x)$ *and* $q(x)$ *on some open interval* I, *then every solution* $y = Y(x)$ *of* (1) *on* I *is of the form*

$$Y(x) = C_1 y_1(x) + C_2 y_2(x) \tag{8}$$

where y_1, y_2 *is any basis of solutions of* (1) *on* I *and* C_1, C_2 *are suitable constants.*

Hence (1) *does not have* **singular solutions** *(that is, solutions not obtainable from a general solution).*

PROOF Let $y = Y(x)$ be any solution of (1) on I. Now, by Theorem 3 the ODE (1) has a general solution

$$y(x) = c_1 y_1(x) + c_2 y_2(x) \tag{9}$$

on I. We have to find suitable values of c_1, c_2 such that $y(x) = Y(x)$ on I. We choose any x_0 in I and show first that we can find values of c_1, c_2 such that we reach agreement at x_0, that is, $y(x_0) = Y(x_0)$ and $y'(x_0) = Y'(x_0)$. Written out in terms of (9), this becomes

$$\begin{aligned} &\text{(a)} \quad c_1 y_1(x_0) + c_2 y_2(x_0) = Y(x_0) \\ &\text{(b)} \quad c_1 y_1'(x_0) + c_2 y_2'(x_0) = Y'(x_0). \end{aligned} \tag{10}$$

We determine the unknowns c_1 and c_2. To eliminate c_2, we multiply (10a) by $y_2'(x_0)$ and (10b) by $-y_2(x_0)$ and add the resulting equations. This gives an equation for c_1. Then we multiply (10a) by $-y_1'(x_0)$ and (10b) by $y_1(x_0)$ and add the resulting equations. This gives an equation for c_2. These new equations are as follows, where we take the values of $y_1, y_1', y_2, y_2', Y, Y'$ at x_0.

$$c_1(y_1 y_2' - y_2 y_1') = c_1 W(y_1, y_2) = Y y_2' - y_2 Y'$$

$$c_2(y_1 y_2' - y_2 y_1') = c_2 W(y_1, y_2) = y_1 Y' - Y y_1'.$$

Since y_1, y_2 is a basis, the Wronskian W in these equations is not 0, and we can solve for c_1 and c_2. We call the (unique) solution $c_1 = C_1, c_2 = C_2$. By substituting it into (9) we obtain from (9) the particular solution

$$y^*(x) = C_1 y_1(x) + C_2 y_2(x).$$

Now since C_1, C_2 is a solution of (10), we see from (10) that

$$y^*(x_0) = Y(x_0), \qquad y^{*\prime}(x_0) = Y'(x_0).$$

From the uniqueness stated in Theorem 1 this implies that y^* and Y must be equal everywhere on I, and the proof is complete. ■

Reflecting on this section, we note that homogeneous linear ODEs with continuous variable coefficients have a conceptually and structurally rather transparent existence and uniqueness theory of solutions. Important in itself, this theory will also provide the foundation for our study of nonhomogeneous linear ODEs, whose theory and engineering applications form the content of the remaining four sections of this chapter.

PROBLEM SET 2.6

1. Derive (6*) from (6).

2–8 BASIS OF SOLUTIONS. WRONSKIAN

Find the Wronskian. Show linear independence by using quotients and confirm it by Theorem 2.

2. $e^{4.0x}, e^{-1.5x}$

3. $e^{-0.4x}, e^{-2.6x}$

4. $x, 1/x$

5. x^3, x^2

6. $e^{-x}\cos \omega x, e^{-x}\sin \omega x$

7. $\cosh ax, \sinh ax$

8. $x^k \cos(\ln x), x^k \sin(\ln x)$

9–15 ODE FOR GIVEN BASIS. WRONSKIAN. IVP

(a) Find a second-order homogeneous linear ODE for which the given functions are solutions. **(b)** Show linear independence by the Wronskian. **(c)** Solve the initial value problem.

9. $\cos 5x, \sin 5x, \quad y(0) = 3, \quad y'(0) = -5$

10. $x^{m_1}, x^{m_2}, \quad y(1) = -2, \quad y'(1) = 2m_1 - 4m_2$

11. $e^{-2.5x}\cos 0.3x, e^{-2.5x}\sin 0.3x, \quad y(0) = 3,$
$y'(0) = -7.5$

12. $x^2, x^2 \ln x, \quad y(1) = 4, \quad y'(1) = 6$

13. $1, e^{-2x}, \quad y(0) = 1, \quad y'(0) = -1$

14. $e^{-kx}\cos \pi x, e^{-kx}\sin \pi x, \quad y(0) = 1,$
$y'(0) = -k - \pi$

15. $\cosh 1.8x, \sinh 1.8x, \quad y(0) = 14.20, \quad y'(0) = 16.38$

16. TEAM PROJECT. Consequences of the Present Theory. This concerns some noteworthy general properties of solutions. Assume that the coefficients p and q of the ODE (1) are continuous on some open interval I, to which the subsequent statements refer.

(a) Solve $y'' - y = 0$ (a) by exponential functions, (b) by hyperbolic functions. How are the constants in the corresponding general solutions related?

(b) Prove that the solutions of a basis cannot be 0 at the same point.

(c) Prove that the solutions of a basis cannot have a maximum or minimum at the same point.

(d) Why is it likely that formulas of the form (6*) should exist?

(e) Sketch $y_1(x) = x^3$ if $x \geqq 0$ and 0 if $x < 0$, $y_2(x) = 0$ if $x \geqq 0$ and x^3 if $x < 0$. Show linear independence on $-1 < x < 1$. What is their Wronskian? What Euler–Cauchy equation do y_1, y_2 satisfy? Is there a contradiction to Theorem 2?

(f) Prove **Abel's formula**[6]

$$W(y_1(x), y_2(x)) = c \exp\left[-\int_{x_0}^{x} p(t)\,dt\right]$$

where $c = W(y_1(x_0), y_2(x_0))$. Apply it to Prob. 6. *Hint:* Write (1) for y_1 and for y_2. Eliminate q algebraically from these two ODEs, obtaining a first-order linear ODE. Solve it.

2.7 Nonhomogeneous ODEs

We now advance from homogeneous to nonhomogeneous linear ODEs.

Consider the second-order nonhomogeneous linear ODE

$$y'' + p(x)y' + q(x)y = r(x) \tag{1}$$

where $r(x) \not\equiv 0$. We shall see that a "general solution" of (1) is the sum of a general solution of the corresponding homogeneous ODE

[6]**NIELS HENRIK ABEL** (1802–1829), Norwegian mathematician.

(2) $$y'' + p(x)y' + q(x)y = 0$$

and a "particular solution" of (1). These two new terms "general solution of (1)" and "particular solution of (1)" are defined as follows.

DEFINITION

General Solution, Particular Solution

A **general solution** of the nonhomogeneous ODE (1) on an open interval I is a solution of the form

(3) $$y(x) = y_h(x) + y_p(x);$$

here, $y_h = c_1y_1 + c_2y_2$ is a general solution of the homogeneous ODE (2) on I and y_p is any solution of (1) on I containing no arbitrary constants.

A **particular solution** of (1) on I is a solution obtained from (3) by assigning specific values to the arbitrary constants c_1 and c_2 in y_h.

Our task is now twofold, first to justify these definitions and then to develop a method for finding a solution y_p of (1).

Accordingly, we first show that a general solution as just defined satisfies (1) and that the solutions of (1) and (2) are related in a very simple way.

THEOREM 1

Relations of Solutions of (1) to Those of (2)

(a) *The sum of a solution y of* (1) *on some open interval I and a solution $\tilde{y}$ of* (2) *on I is a solution of* (1) *on I. In particular,* (3) *is a solution of* (1) *on I.*

(b) *The difference of two solutions of* (1) *on I is a solution of* (2) *on I.*

PROOF **(a)** Let $L[y]$ denote the left side of (1). Then for any solutions y of (1) and $\tilde{y}$ of (2) on I,

$$L[y + \tilde{y}] = L[y] + L[\tilde{y}] = r + 0 = r.$$

(b) For any solutions y and y^* of (1) on I we have $L[y - y^*] = L[y] - L[y^*] = r - r = 0.$ ■

Now for *homogeneous ODEs* (2) we know that general solutions include all solutions. We show that the same is true for nonhomogeneous ODEs (1).

THEOREM 2

A General Solution of a Nonhomogeneous ODE Includes All Solutions

If the coefficients $p(x)$, $q(x)$, and the function $r(x)$ in (1) *are continuous on some open interval I, then every solution of* (1) *on I is obtained by assigning suitable values to the arbitrary constants c_1 and c_2 in a general solution* (3) *of* (1) *on I.*

PROOF Let y^* be any solution of (1) on I and x_0 any x in I. Let (3) be any general solution of (1) on I. This solution exists. Indeed, $y_h = c_1y_1 + c_2y_2$ exists by Theorem 3 in Sec. 2.6

because of the continuity assumption, and y_p exists according to a construction to be shown in Sec. 2.10. Now, by Theorem 1(b) just proved, the difference $Y = y^* - y_p$ is a solution of (2) on I. At x_0 we have

$$Y(x_0) = y^*(x_0) - y_p(x_0). \qquad Y'(x_0) = y^{*\prime}(x_0) - y_p'(x_0).$$

Theorem 1 in Sec. 2.6 implies that for these conditions, as for any other initial conditions in I, there exists a unique particular solution of (2) obtained by assigning suitable values to c_1, c_2 in y_h. From this and $y^* = Y + y_p$ the statement follows. ■

Method of Undetermined Coefficients

Our discussion suggests the following. *To solve the nonhomogeneous ODE* (1) *or an initial value problem for* (1), *we have to solve the homogeneous ODE* (2) *and find any solution* y_p *of* (1), so that we obtain a general solution (3) of (1).

How can we find a solution y_p of (1)? One method is the so-called **method of undetermined coefficients**. It is much simpler than another, more general, method (given in Sec. 2.10). Since it applies to models of vibrational systems and electric circuits to be shown in the next two sections, it is frequently used in engineering.

More precisely, the method of undetermined coefficients is suitable for linear ODEs with ***constant coefficients a and b***

$$y'' + ay' + by = r(x) \tag{4}$$

when $r(x)$ is an exponential function, a power of x, a cosine or sine, or sums or products of such functions. These functions have derivatives similar to $r(x)$ itself. This gives the idea. We choose a form for y_p similar to $r(x)$, but with unknown coefficients to be determined by substituting that y_p and its derivatives into the ODE. Table 2.1 on p. 82 shows the choice of y_p for practically important forms of $r(x)$. Corresponding rules are as follows.

Choice Rules for the Method of Undetermined Coefficients

(a) Basic Rule. *If* $r(x)$ *in* (4) *is one of the functions in the first column in Table* 2.1, *choose* y_p *in the same line and determine its undetermined coefficients by substituting* y_p *and its derivatives into* (4).

(b) Modification Rule. *If a term in your choice for* y_p *happens to be a solution of the homogeneous ODE corresponding to* (4), *multiply this term by* x (*or by* x^2 *if this solution corresponds to a double root of the characteristic equation of the homogeneous ODE*).

(c) Sum Rule. *If* $r(x)$ *is a sum of functions in the first column of Table* 2.1, *choose for* y_p *the sum of the functions in the corresponding lines of the second column.*

The Basic Rule applies when $r(x)$ is a single term. The Modification Rule helps in the indicated case, and to recognize such a case, we have to solve the homogeneous ODE first. The Sum Rule follows by noting that the sum of two solutions of (1) with $r = r_1$ and $r = r_2$ (and the same left side!) is a solution of (1) with $r = r_1 + r_2$. (Verify!)

The method is self-correcting. A false choice for y_p or one with too few terms will lead to a contradiction. A choice with too many terms will give a correct result, with superfluous coefficients coming out zero.

Let us illustrate Rules (a)–(c) by the typical Examples 1–3.

Table 2.1 Method of Undetermined Coefficients

Term in $r(x)$	Choice for $y_p(x)$
$ke^{\gamma x}$	$Ce^{\gamma x}$
$kx^n \ (n = 0, 1, \cdots)$	$K_n x^n + K_{n-1}x^{n-1} + \cdots + K_1 x + K_0$
$k \cos \omega x$ $k \sin \omega x$	$K \cos \omega x + M \sin \omega x$
$ke^{\alpha x} \cos \omega x$ $ke^{\alpha x} \sin \omega x$	$e^{\alpha x}(K \cos \omega x + M \sin \omega x)$

EXAMPLE 1 Application of the Basic Rule (a)

Solve the initial value problem

$$y'' + y = 0.001x^2, \qquad y(0) = 0, \qquad y'(0) = 1.5. \tag{5}$$

Solution. ***Step 1. General solution of the homogeneous ODE.*** The ODE $y'' + y = 0$ has the general solution

$$y_h = A \cos x + B \sin x.$$

Step 2. Solution y_p of the nonhomogeneous ODE. We first try $y_p = Kx^2$. Then $y_p'' = 2K$. By substitution, $2K + Kx^2 = 0.001x^2$. For this to hold for all x, the coefficient of each power of x (x^2 and x^0) must be the same on both sides; thus $K = 0.001$ and $2K = 0$, a contradiction.

The second line in Table 2.1 suggests the choice

$$y_p = K_2x^2 + K_1x + K_0. \qquad \text{Then} \qquad y_p'' + y_p = 2K_2 + K_2x^2 + K_1x + K_0 = 0.001x^2.$$

Equating the coefficients of x^2, x, x^0 on both sides, we have $K_2 = 0.001, K_1 = 0, 2K_2 + K_0 = 0$. Hence $K_0 = -2K_2 = -0.002$. This gives $y_p = 0.001x^2 - 0.002$, and

$$y = y_h + y_p = A \cos x + B \sin x + 0.001x^2 - 0.002.$$

Step 3. Solution of the initial value problem. Setting $x = 0$ and using the first initial condition gives $y(0) = A - 0.002 = 0$, hence $A = 0.002$. By differentiation and from the second initial condition,

$$y' = y_h' + y_p' = -A \sin x + B \cos x + 0.002x \qquad \text{and} \qquad y'(0) = B = 1.5.$$

This gives the answer (Fig. 50)

$$y = 0.002 \cos x + 1.5 \sin x + 0.001x^2 - 0.002.$$

Figure 50 shows y as well as the quadratic parabola y_p about which y is oscillating, practically like a sine curve since the cosine term is smaller by a factor of about $1/1000$. ■

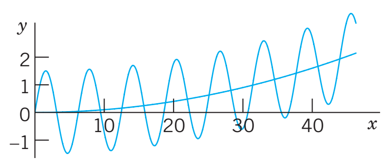


Fig. 50. Solution in Example 1

EXAMPLE 2 Application of the Modification Rule (b)

Solve the initial value problem

$$y'' + 3y' + 2.25y = -10e^{-1.5x}, \qquad y(0) = 1, \qquad y'(0) = 0. \tag{6}$$

Solution. ***Step 1. General solution of the homogeneous ODE.*** The characteristic equation of the homogeneous ODE is $\lambda^2 + 3\lambda + 2.25 = (\lambda + 1.5)^2 = 0$. Hence the homogeneous ODE has the general solution

$$y_h = (c_1 + c_2 x)e^{-1.5x}.$$

Step 2. Solution y_p of the nonhomogeneous ODE. The function $e^{-1.5x}$ on the right would normally require the choice $Ce^{-1.5x}$. But we see from y_h that this function is a solution of the homogeneous ODE, which corresponds to a *double root* of the characteristic equation. Hence, according to the Modification Rule we have to multiply our choice function by x^2. That is, we choose

$$y_p = Cx^2e^{-1.5x}. \qquad \text{Then} \qquad y_p' = C(2x - 1.5x^2)e^{-1.5x}, \qquad y_p'' = C(2 - 3x - 3x + 2.25x^2)e^{-1.5x}.$$

We substitute these expressions into the given ODE and omit the factor $e^{-1.5x}$. This yields

$$C(2 - 6x + 2.25x^2) + 3C(2x - 1.5x^2) + 2.25Cx^2 = -10.$$

Comparing the coefficients of x^2, x, x^0 gives $0 = 0, 0 = 0, 2C = -10$, hence $C = -5$. This gives the solution $y_p = -5x^2e^{-1.5x}$. Hence the given ODE has the general solution

$$y = y_h + y_p = (c_1 + c_2 x)e^{-1.5x} - 5x^2e^{-1.5x}.$$

Step 3. Solution of the initial value problem. Setting $x = 0$ in y and using the first initial condition, we obtain $y(0) = c_1 = 1$. Differentiation of y gives

$$y' = (c_2 - 1.5c_1 - 1.5c_2 x)e^{-1.5x} - 10xe^{-1.5x} + 7.5x^2e^{-1.5x}.$$

From this and the second initial condition we have $y'(0) = c_2 - 1.5c_1 = 0$. Hence $c_2 = 1.5c_1 = 1.5$. This gives the answer (Fig. 51)

$$y = (1 + 1.5x)e^{-1.5x} - 5x^2e^{-1.5x} = (1 + 1.5x - 5x^2)e^{-1.5x}.$$

The curve begins with a horizontal tangent, crosses the x-axis at $x = 0.6217$ (where $1 + 1.5x - 5x^2 = 0$) and approaches the axis from below as x increases. ■

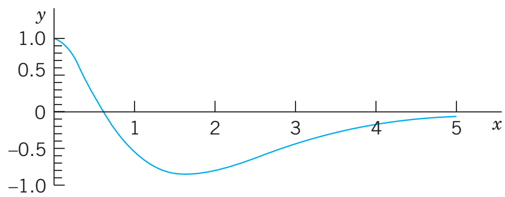


Fig. 51. Solution in Example 2

EXAMPLE 3 Application of the Sum Rule (c)

Solve the initial value problem

$$y'' + 2y' + 0.75y = 2\cos x - 0.25\sin x + 0.09x, \qquad y(0) = 2.78, \qquad y'(0) = -0.43. \tag{7}$$

Solution. ***Step 1. General solution of the homogeneous ODE.*** The characteristic equation of the homogeneous ODE is

$$\lambda^2 + 2\lambda + 0.75 = (\lambda + \tfrac{1}{2})(\lambda + \tfrac{3}{2}) = 0$$

which gives the general solution $y_h = c_1 e^{-x/2} + c_2 e^{-3x/2}$.

Step 2. Particular solution of the nonhomogeneous ODE. We write $y_p = y_{p1} + y_{p2}$ and, following Table 2.1, (C) and (B),

$$y_{p1} = K\cos x + M\sin x \qquad \text{and} \qquad y_{p2} = K_1 x + K_0.$$

Differentiation gives $y'_{p1} = -K\sin x + M\cos x$, $y''_{p1} = -K\cos x - M\sin x$ and $y'_{p2} = 1$, $y''_{p2} = 0$. Substitution of y_{p1} into the ODE in (7) gives, by comparing the cosine and sine terms,

$$-K + 2M + 0.75K = 2, \qquad -M - 2K + 0.75M = -0.25,$$

hence $K = 0$ and $M = 1$. Substituting y_{p2} into the ODE in (7) and comparing the x- and x^0-terms gives

$$0.75K_1 = 0.09, \quad 2K_1 + 0.75K_0 = 0, \qquad \text{thus} \qquad K_1 = 0.12, \quad K_0 = -0.32.$$

Hence a general solution of the ODE in (7) is

$$y = c_1 e^{-x/2} + c_2 e^{-3x/2} + \sin x + 0.12x - 0.32.$$

Step 3. Solution of the initial value problem. From y, y' and the initial conditions we obtain

$$y(0) = c_1 + c_2 - 0.32 = 2.78, \qquad y'(0) = -\tfrac{1}{2}c_1 - \tfrac{3}{2}c_2 + 1 + 0.12 = -0.4.$$

Hence $c_1 = 3.1$, $c_2 = 0$. This gives the solution of the IVP (Fig. 52)

$$y = 3.1e^{-x/2} + \sin x + 0.12x - 0.32.$$

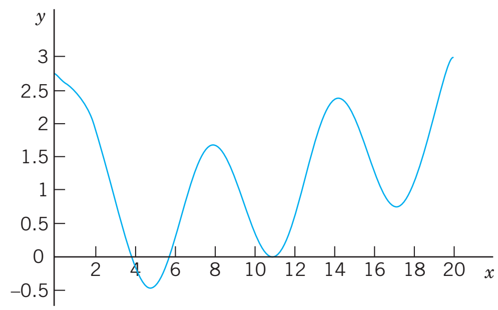


Fig. 52. Solution in Example 3

Stability. The following is important. If (and only if) all the roots of the characteristic equation of the homogeneous ODE $y'' + ay' + by = 0$ in (4) are negative, or have a negative real part, then a general solution y_h of this ODE goes to 0 as $x \to \infty$, so that the "**transient solution**" $y = y_h + y_p$ of (4) approaches the "**steady-state solution**" y_p. In this case the nonhomogeneous ODE and the physical or other system modeled by the ODE are called **stable**; otherwise they are called **unstable**. For instance, the ODE in Example 1 is unstable.

Applications follow in the next two sections.

PROBLEM SET 2.7

1–10 NONHOMOGENEOUS LINEAR ODEs: GENERAL SOLUTION

Find a (real) general solution. State which rule you are using. Show each step of your work.

1. $y'' + 5y' + 4y = 10e^{-3x}$

2. $10y'' + 50y' + 57.6y = \cos x$

3. $y'' + 3y' + 2y = 12x^2$

4. $y'' - 9y = 18\cos \pi x$

5. $y'' + 4y' + 4y = e^{-x}\cos x$

6. $y'' + y' + (\pi^2 + \tfrac{1}{4})y = e^{-x/2}\sin \pi x$

7. $(D^2 + 2D + \frac{3}{4}I)y = 3e^x + \frac{9}{2}x$

8. $(3D^2 + 27I)y = 3\cos x + \cos 3x$

9. $(D^2 - 16I)y = 9.6e^{4x} + 30e^x$

10. $(D^2 + 2D + I)y = 2x \sin x$

11–18 NONHOMOGENEOUS LINEAR ODEs: IVPs

Solve the initial value problem. State which rule you are using. Show each step of your calculation in detail.

11. $y'' + 3y = 18x^2, \quad y(0) = -3, \quad y'(0) = 0$

12. $y'' + 4y = -12\sin 2x, \quad y(0) = 1.8, \quad y'(0) = 5.0$

13. $8y'' - 6y' + y = 6\cosh x, \quad y(0) = 0.2, \quad y'(0) = 0.05$

14. $y'' + 4y' + 4y = e^{-2x}\sin 2x, \quad y(0) = 1, \quad y'(0) = -1.5$

15. $(x^2D^2 - 3xD + 3I)y = 3\ln x - 4, \quad y(1) = 0, \quad y'(1) = 1; \quad y_p = \ln x$

16. $(D^2 - 2D)y = 6e^{2x} - 4e^{-2x}, \quad y(0) = -1, \quad y'(0) = 6$

17. $(D^2 + 0.2D + 0.26I)y = 1.22e^{0.5x}, \quad y(0) = 3.5, \quad y'(0) = 0.35$

18. $(D^2 + 2D + 10I)y = 17\sin x - 37\sin 3x, \quad y(0) = 6.6, \quad y'(0) = -2.2$

19. CAS PROJECT. Structure of Solutions of Initial Value Problems. Using the present method, find, graph, and discuss the solutions y of initial value problems of your own choice. Explore effects on solutions caused by changes of initial conditions. Graph $y_p, y, y - y_p$ separately, to see the separate effects. Find a problem in which **(a)** the part of y resulting from y_h decreases to zero, **(b)** increases, **(c)** is not present in the answer y. Study a problem with $y(0) = 0, y'(0) = 0$. Consider a problem in which you need the Modification Rule (a) for a simple root, (b) for a double root. Make sure that your problems cover all three Cases I, II, III (see Sec. 2.2).

20. TEAM PROJECT. Extensions of the Method of Undetermined Coefficients. (a) Extend the method to products of the function in Table 2.1, **(b)** Extend the method to Euler–Cauchy equations. Comment on the practical significance of such extensions.

2.8 Modeling: Forced Oscillations. Resonance

In Sec. 2.4 we considered vertical motions of a mass–spring system (vibration of a mass m on an elastic spring, as in Figs. 33 and 53) and modeled it by the ***homogeneous*** linear ODE

$$my'' + cy' + ky = 0. \tag{1}$$

Here $y(t)$ as a function of time t is the displacement of the body of mass m from rest.

The mass–spring system of Sec. 2.4 exhibited only free motion. This means no external forces (outside forces) but only internal forces controlled the motion. The internal forces are forces within the system. They are the force of inertia my'', the damping force cy' (if $c > 0$), and the spring force ky, a restoring force.

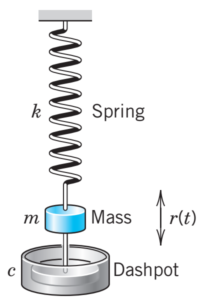


Fig. 53. Mass on a spring

We now extend our model by including an additional force, that is, the external force $r(t)$, on the right. Then we have

$$my'' + cy' + ky = r(t). \tag{2*}$$

Mechanically this means that at each instant t the resultant of the internal forces is in equilibrium with $r(t)$. The resulting motion is called a **forced motion** with **forcing function** $r(t)$, which is also known as **input** or **driving force**, and the solution $y(t)$ to be obtained is called the **output** or the **response** *of the system to the driving force.*

Of special interest are periodic external forces, and we shall consider a driving force of the form

$$r(t) = F_0 \cos \omega t \qquad (F_0 > 0, \omega > 0).$$

Then we have the nonhomogeneous ODE

$$\boxed{my'' + cy' + ky = F_0 \cos \omega t.} \tag{2}$$

Its solution will reveal facts that are fundamental in engineering mathematics and allow us to model resonance.

Solving the Nonhomogeneous ODE (2)

From Sec. 2.7 we know that a general solution of (2) is the sum of a general solution y_h of the homogeneous ODE (1) plus any solution y_p of (2). To find y_p, we use the method of undetermined coefficients (Sec. 2.7), starting from

$$\boxed{y_p(t) = a \cos \omega t + b \sin \omega t.} \tag{3}$$

By differentiating this function (chain rule!) we obtain

$$\begin{aligned} y_p' &= -\omega a \sin \omega t + \omega b \cos \omega t, \\ y_p'' &= -\omega^2 a \cos \omega t - \omega^2 b \sin \omega t. \end{aligned}$$

Substituting y_p, y_p', and y_p'' into (2) and collecting the cosine and the sine terms, we get

$$[(k - m\omega^2)a + \omega cb] \cos \omega t + [-\omega ca + (k - m\omega^2)b] \sin \omega t = F_0 \cos \omega t.$$

The cosine terms on both sides must be equal, and the coefficient of the sine term on the left must be zero since there is no sine term on the right. This gives the two equations

$$\begin{aligned} (k - m\omega^2)a + \quad \omega cb \quad &= F_0 \\ -\omega ca \quad + (k - m\omega^2)b &= 0 \end{aligned} \tag{4}$$

for determining the unknown coefficients a and b. This is a linear system. We can solve it by elimination. To eliminate b, multiply the first equation by $k - m\omega^2$ and the second by $-\omega c$ and add the results, obtaining

$$(k - m\omega^2)^2 a + \omega^2 c^2 a = F_0(k - m\omega^2).$$

Similarly, to eliminate a, multiply (the first equation by ωc and the second by $k - m\omega^2$ and add to get

$$\omega^2 c^2 b + (k - m\omega^2)^2 b = F_0 \omega c.$$

If the factor $(k - m\omega^2)^2 + \omega^2 c^2$ is not zero, we can divide by this factor and solve for a and b,

$$a = F_0 \frac{k - m\omega^2}{(k - m\omega^2)^2 + \omega^2 c^2}, \qquad b = F_0 \frac{\omega c}{(k - m\omega^2)^2 + \omega^2 c^2}.$$

If we set $\sqrt{k/m} = \omega_0 \ (> 0)$ as in Sec. 2.4, then $k = m\omega_0^2$ and we obtain

(5) $$a = F_0 \frac{m(\omega_0^2 - \omega^2)}{m^2(\omega_0^2 - \omega^2)^2 + \omega^2 c^2}, \qquad b = F_0 \frac{\omega c}{m^2(\omega_0^2 - \omega^2)^2 + \omega^2 c^2}.$$

We thus obtain the general solution of the nonhomogeneous ODE (2) in the form

(6) $$y(t) = y_h(t) + y_p(t).$$

Here y_h is a general solution of the homogeneous ODE (1) and y_p is given by (3) with coefficients (5).

We shall now discuss the behavior of the mechanical system, distinguishing between the two cases $c = 0$ (no damping) and $c > 0$ (damping). These cases will correspond to two basically different types of output.

Case 1. Undamped Forced Oscillations. Resonance

If the damping of the physical system is so small that its effect can be neglected over the time interval considered, we can set $c = 0$. Then (5) reduces to $a = F_0/[m(\omega_0^2 - \omega^2)]$ and $b = 0$. Hence (3) becomes (use $\omega_0{}^2 = k/m$)

(7) $$y_p(t) = \frac{F_0}{m(\omega_0^2 - \omega^2)} \cos \omega t = \frac{F_0}{k[1 - (\omega/\omega_0)^2]} \cos \omega t.$$

Here we must assume that $\omega^2 \neq \omega_0{}^2$; physically, the frequency $\omega/(2\pi)$ [cycles/sec] of the driving force is different from the *natural frequency* $\omega_0/(2\pi)$ of the system, which is the frequency of the free undamped motion [see (4) in Sec. 2.4]. From (7) and from (4*) in Sec. 2.4 we have the general solution of the "undamped system"

(8) $$y(t) = C \cos(\omega_0 t - \delta) + \frac{F_0}{m(\omega_0^2 - \omega^2)} \cos \omega t.$$

We see that this output is a ***superposition of two harmonic oscillations*** *of the frequencies just mentioned.*

Resonance. We discuss (7). We see that the maximum amplitude of y_p is (put $\cos \omega t = 1$)

$$a_0 = \frac{F_0}{k}\rho \qquad \text{where} \qquad \rho = \frac{1}{1 - (\omega/\omega_0)^2}. \tag{9}$$

a_0 depends on ω and ω_0. If $\omega \to \omega_0$, then ρ and a_0 tend to infinity. This excitation of large oscillations by matching input and natural frequencies ($\omega = \omega_0$) is called **resonance**. ρ is called the **resonance factor** (Fig. 54), and from (9) we see that $\rho/k = a_0/F_0$ is the ratio of the amplitudes of the particular solution y_p and of the input $F_0 \cos \omega t$. We shall see later in this section that resonance is of basic importance in the study of vibrating systems.

In the case of resonance the nonhomogeneous ODE (2) becomes

$$y'' + \omega_0^2 y = \frac{F_0}{m} \cos \omega_0 t. \tag{10}$$

Then (7) is no longer valid, and, from the Modification Rule in Sec. 2.7, we conclude that a particular solution of (10) is of the form

$$y_p(t) = t(a \cos \omega_0 t + b \sin \omega_0 t).$$

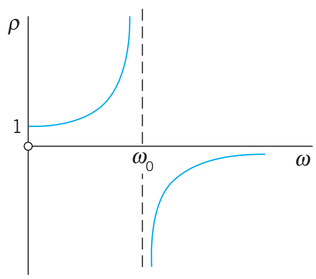


Fig. 54. Resonance factor $\rho(\omega)$

By substituting this into (10) we find $a = 0$ and $b = F_0/(2m\omega_0)$. Hence (Fig. 55)

$$\boxed{y_p(t) = \frac{F_0}{2m\omega_0} t \sin \omega_0 t.} \tag{11}$$

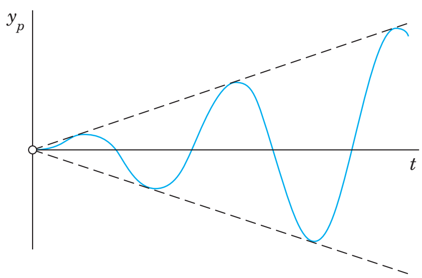


Fig. 55. Particular solution in the case of resonance

We see that, because of the factor t, the amplitude of the vibration becomes larger and larger. Practically speaking, systems with very little damping may undergo large vibrations that can destroy the system. We shall return to this practical aspect of resonance later in this section.

Beats. Another interesting and highly important type of oscillation is obtained if ω is close to ω_0. Take, for example, the particular solution [see (8)]

$$y(t) = \frac{F_0}{m(\omega_0^2 - \omega^2)}(\cos \omega t - \cos \omega_0 t) \qquad (\omega \neq \omega_0). \tag{12}$$

Using (12) in App. 3.1, we may write this as

$$y(t) = \frac{2F_0}{m(\omega_0^2 - \omega^2)} \sin\left(\frac{\omega_0 + \omega}{2}t\right) \sin\left(\frac{\omega_0 - \omega}{2}t\right).$$

Since ω is close to ω_0, the difference $\omega_0 - \omega$ is small. Hence the period of the last sine function is large, and we obtain an oscillation of the type shown in Fig. 56, the dashed curve resulting from the first sine factor. This is what musicians are listening to when they ***tune*** their instruments.

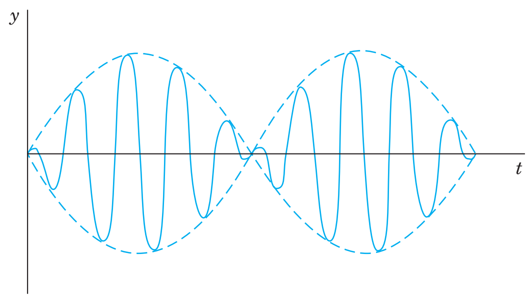


Fig. 56. Forced undamped oscillation when the difference of the input and natural frequencies is small (**"beats"**)

Case 2. Damped Forced Oscillations

If the damping of the mass–spring system is not negligibly small, we have $c > 0$ and a damping term cy' in (1) and (2). Then the general solution y_h of the homogeneous ODE (1) approaches zero as t goes to infinity, as we know from Sec. 2.4. Practically, it is zero after a sufficiently long time. Hence the "**transient solution**" (6) of (2), given by $y = y_h + y_p$, approaches the "**steady-state solution**" y_p. This proves the following.

THEOREM 1

Steady-State Solution

After a sufficiently long time the output of a damped vibrating system under a purely sinusoidal driving force [*see* (2)] *will practically be a harmonic oscillation whose frequency is that of the input.*

Amplitude of the Steady-State Solution. Practical Resonance

Whereas in the undamped case the amplitude of y_p approaches infinity as ω approaches ω_0, this will not happen in the damped case. In this case the amplitude will always be finite. But it may have a maximum for some ω depending on the damping constant c. This may be called **practical resonance**. It is of great importance because if c is not too large, then some input may excite oscillations large enough to damage or even destroy the system. Such cases happened, in particular in earlier times when less was known about resonance. Machines, cars, ships, airplanes, bridges, and high-rising buildings are vibrating mechanical systems, and it is sometimes rather difficult to find constructions that are completely free of undesired resonance effects, caused, for instance, by an engine or by strong winds.

To study the amplitude of y_p as a function of ω, we write (3) in the form

$$y_p(t) = C^* \cos(\omega t - \eta). \tag{13}$$

C^* is called the **amplitude** of y_p and η the **phase angle** or **phase lag** because it measures the lag of the output behind the input. According to (5), these quantities are

$$\begin{aligned} C^*(\omega) &= \sqrt{a^2 + b^2} = \frac{F_0}{\sqrt{m^2(\omega_0^2 - \omega^2)^2 + \omega^2 c^2}}, \\ \tan \eta(\omega) &= \frac{b}{a} = \frac{\omega c}{m(\omega_0^2 - \omega^2)}. \end{aligned} \tag{14}$$

Let us see whether $C^*(\omega)$ has a maximum and, if so, find its location and then its size. We denote the radicand in the second root in C^* by R. Equating the derivative of C^* to zero, we obtain

$$\frac{dC^*}{d\omega} = F_0\left(-\frac{1}{2}R^{-3/2}\right)[2m^2(\omega_0^2 - \omega^2)(-2\omega) + 2\omega c^2].$$

The expression in the brackets [. . .] is zero if

$$c^2 = 2m^2(\omega_0^2 - \omega^2) \qquad (\omega_0^2 = k/m). \tag{15}$$

By reshuffling terms we have

$$2m^2\omega^2 = 2m^2\omega_0{}^2 - c^2 = 2mk - c^2.$$

The right side of this equation becomes negative if $c^2 > 2mk$, so that then (15) has no real solution and C^* decreases monotone as ω increases, as the lowest curve in Fig. 57 shows. If c is smaller, $c^2 < 2mk$, then (15) has a real solution $\omega = \omega_{\max}$, where

$$\omega_{\max}^2 = \omega_0^2 - \frac{c^2}{2m^2}. \tag{15*}$$

From (15*) we see that this solution increases as c decreases and approaches ω_0 as c approaches zero. See also Fig. 57.

The size of $C^*(\omega_{\max})$ is obtained from (14), with $\omega^2 = \omega_{\max}^2$ given by (15*). For this ω^2 we obtain in the second radicand in (14) from (15*)

$$m^2(\omega_0^2 - \omega_{\max}^2)^2 = \frac{c^4}{4m^2} \quad \text{and} \quad \omega_{\max}^2 c^2 = \left(\omega_0^2 - \frac{c^2}{2m^2}\right)c^2.$$

The sum of the right sides of these two formulas is

$$(c^4 + 4m^2\omega_0^2 c^2 - 2c^4)/(4m^2) = c^2(4m^2\omega_0^2 - c^2)/(4m^2).$$

Substitution into (14) gives

$$C^*(\omega_{\max}) = \frac{2mF_0}{c\sqrt{4m^2\omega_0^2 - c^2}}. \tag{16}$$

We see that $C^*(\omega_{\max})$ is always finite when $c > 0$. Furthermore, since the expression

$$c^2 4m^2\omega_0^2 - c^4 = c^2(4mk - c^2)$$

in the denominator of (16) decreases monotone to zero as c^2 ($<2mk$) goes to zero, the maximum amplitude (16) increases monotone to infinity, in agreement with our result in Case 1. Figure 57 shows the **amplification** C^*/F_0 (ratio of the amplitudes of output and input) as a function of ω for $m = 1$, $k = 1$, hence $\omega_0 = 1$, and various values of the damping constant c.

Figure 58 shows the phase angle (the lag of the output behind the input), which is less than $\pi/2$ when $\omega < \omega_0$, and greater than $\pi/2$ for $\omega > \omega_0$.

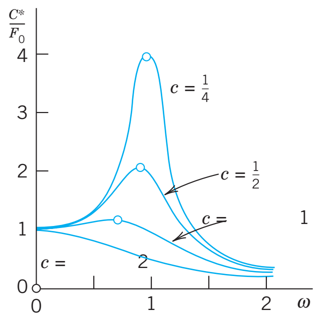


Fig. 57. Amplification C^*/F_0 as a function of ω for $m = 1$, $k = 1$, and various values of the damping constant c

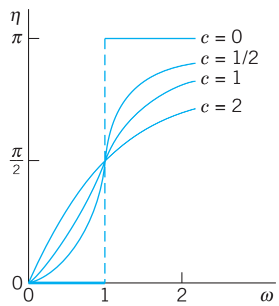


Fig. 58. Phase lag η as a function of ω for $m = 1$, $k = 1$, thus $\omega_0 = 1$, and various values of the damping constant c

PROBLEM SET 2.8

1. **WRITING REPORT. Free and Forced Vibrations.** Write a condensed report of 2–3 pages on the most important similarities and differences of free and forced vibrations, with examples of your own. No proofs.

2. **Which of Probs.** 1–18 in Sec. 2.7 (with x = time t) can be models of mass–spring systems with a harmonic oscillation as steady-state solution?

3–7 STEADY-STATE SOLUTIONS

Find the steady-state motion of the mass–spring system modeled by the ODE. Show the details of your work.

3. $y'' + 6y' + 8y = 42.5 \cos 2t$

4. $y'' + 2.5y' + 10y = -13.6 \sin 4t$

5. $(D^2 + D + 4.25I)y = 22.1 \cos 4.5t$

6. $(D^2 + 4D + 3I)y = \cos t + \frac{1}{3}\cos 3t$

7. $(4D^2 + 12D + 9I)y = 225 - 75 \sin 3t$

8–15 TRANSIENT SOLUTIONS

Find the transient motion of the mass–spring system modeled by the ODE. Show the details of your work.

8. $2y'' + 4y' + 6.5y = 4 \sin 1.5t$

9. $y'' + 3y' + 3.25y = 3 \cos t - 1.5 \sin t$

10. $y'' + 16y = 56 \cos 4t$

11. $(D^2 + 2I)y = \cos \sqrt{2}t + \sin \sqrt{2}t$

12. $(D^2 + 2D + 5I)y = 4 \cos t + 8 \sin t$

13. $(D^2 + I)y = \cos \omega t, \omega^2 \neq 1$

14. $(D^2 + I)y = 5e^{-t} \cos t$

15. $(D^2 + 4D + 8I)y = 2 \cos 2t + \sin 2t$

16–20 INITIAL VALUE PROBLEMS

Find the motion of the mass–spring system modeled by the ODE and the initial conditions. Sketch or graph the solution curve. In addition, sketch or graph the curve of $y - y_p$ to see when the system practically reaches the steady state.

16. $y'' + 25y = 24 \sin t$, $y(0) = 1$, $y'(0) = 1$

17. $(D^2 + 4I)y = \sin t + \frac{1}{3}\sin 3t + \frac{1}{5}\sin 5t$, $y(0) = 0$, $y'(0) = \frac{3}{35}$

18. $(D^2 + 8D + 17I)y = 474.5 \sin 0.5t$, $y(0) = -5.4$, $y'(0) = 9.4$

19. $(D^2 + 2D + 2I)y = e^{-t/2} \sin \frac{1}{2}t$, $y(0) = 0$, $y'(0) = 1$

20. $(D^2 + 5I)y = \cos \pi t - \sin \pi t$, $y(0) = 0$, $y'(0) = 0$

21. Beats. Derive the formula after (12) from (12). Can we have beats in a damped system?

22. Beats. Solve $y'' + 25y = 99 \cos 4.9t$, $y(0) = 2$, $y'(0) = 0$. How does the graph of the solution change if you change **(a)** $y(0)$, **(b)** the frequency of the driving force?

23. TEAM EXPERIMENT. Practical Resonance.
(a) Derive, in detail, the crucial formula (16).

(b) By considering dC^*/dc show that $C^*(\omega_{max})$ increases as c $(\leqq \sqrt{2mk})$ decreases.

(c) Illustrate practical resonance with an ODE of your own in which you vary c, and sketch or graph corresponding curves as in Fig. 57.

(d) Take your ODE with c fixed and an input of two terms, one with frequency close to the practical resonance frequency and the other not. Discuss and sketch or graph the output.

(e) Give other applications (not in the book) in which resonance is important.

24. Gun barrel. Solve $y'' + y = 1 - t^2/\pi^2$ if $0 \leqq t \leqq \pi$ and 0 if $t \to \infty$; here, $y(0) = 0$, $y'(0) = 0$. This models an undamped system on which a force F acts during some interval of time (see Fig. 59), for instance, the force on a gun barrel when a shell is fired, the barrel being braked by heavy springs (and then damped by a dashpot, which we disregard for simplicity). *Hint:* At π both y and y' must be continuous.

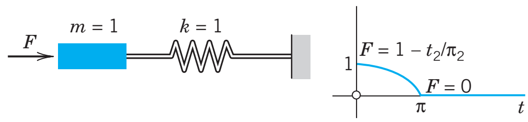


Fig. 59. Problem 24

25. CAS EXPERIMENT. Undamped Vibrations.
(a) Solve the initial value problem $y'' + y = \cos \omega t$, $\omega^2 \neq 1$, $y(0) = 0$, $y'(0) = 0$. Show that the solution can be written

$$y(t) = \frac{2}{1 - \omega^2} \sin\left[\tfrac{1}{2}(1 + \omega)t\right] \sin\left[\tfrac{1}{2}(1 - \omega)t\right].$$

(b) Experiment with the solution by changing ω to see the change of the curves from those for small ω (>0) to beats, to resonance, and to large values of ω (see Fig. 60).

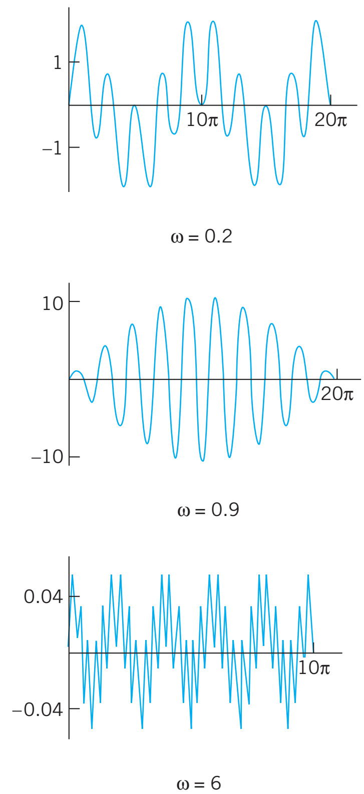


Fig. 60. Typical solution curves in CAS Experiment 25

2.9 Modeling: Electric Circuits

Designing good models is a task the computer cannot do. Hence setting up models has become an important task in modern applied mathematics. The best way to gain experience in successful modeling is to carefully examine the modeling process in various fields and applications. Accordingly, modeling electric circuits will be *profitable for all students,* not just for electrical engineers and computer scientists.

Figure 61 shows an ***RLC*-circuit**, as it occurs as a basic building block of large electric networks in computers and elsewhere. An *RLC*-circuit is obtained from an *RL*-circuit by adding a capacitor. Recall Example 2 on the *RL*-circuit in Sec. 1.5: The model of the *RL*-circuit is $LI' + RI = E(t)$. It was obtained by **KVL** (Kirchhoff's Voltage Law)[7] by equating the voltage drops across the resistor and the inductor to the EMF (electromotive force). Hence we obtain the model of the *RLC*-circuit simply by adding the voltage drop Q/C across the capacitor. Here, C F (farads) is the capacitance of the capacitor. Q coulombs is the charge on the capacitor, related to the current by

$$I(t) = \frac{dQ}{dt}, \qquad \text{equivalently} \qquad Q(t) = \int I(t)\,dt.$$

See also Fig. 62. Assuming a sinusoidal EMF as in Fig. 61, we thus have the model of the *RLC*-circuit

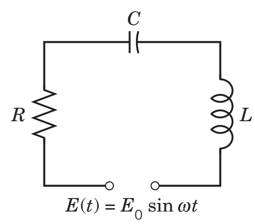


Fig. 61. *RLC*-circuit

Name	Symbol	Notation		Unit	Voltage Drop
Ohm's Resistor		R	Ohm's Resistance	ohms (Ω)	RI
Inductor		L	Inductance	henrys (H)	$L\dfrac{dI}{dt}$
Capacitor		C	Capacitance	farads (F)	Q/C

Fig. 62. Elements in an *RLC*-circuit

[7]GUSTAV ROBERT KIRCHHOFF (1824–1887), German physicist. Later we shall also need **Kirchhoff's Current Law (KCL)**:

At any point of a circuit, the sum of the inflowing currents is equal to the sum of the outflowing currents.

The units of measurement of electrical quantities are named after ANDRÉ MARIE AMPÈRE (1775–1836), French physicist, CHARLES AUGUSTIN DE COULOMB (1736–1806), French physicist and engineer, MICHAEL FARADAY (1791–1867), English physicist, JOSEPH HENRY (1797–1878), American physicist, GEORG SIMON OHM (1789–1854), German physicist, and ALESSANDRO VOLTA (1745–1827), Italian physicist.

$$(1') \qquad LI' + RI + \frac{1}{C}\int I\,dt = E(t) = E_0 \sin \omega t.$$

This is an "integro-differential equation." To get rid of the integral, we differentiate $(1')$ with respect to t, obtaining

$$\textbf{(1)} \qquad \boxed{LI'' + RI' + \frac{1}{C} I = E'(t) = E_0 \omega \cos \omega t.}$$

This shows that the current in an *RLC*-circuit is obtained as the solution of this nonhomogeneous second-order ODE (1) with constant coefficients.

In connection with initial value problems, we shall occasionally use

$$(1'') \qquad LQ'' + RQ'' + \frac{1}{C} Q = E(t),$$

obtained from $(1')$ and $I = Q'$.

Solving the ODE (1) for the Current in an *RLC*-Circuit

A general solution of (1) is the sum $I = I_h + I_p$, where I_h is a general solution of the homogeneous ODE corresponding to (1) and I_p is a particular solution of (1). We first determine I_p by the method of undetermined coefficients, proceeding as in the previous section. We substitute

$$(2) \qquad \begin{aligned} I_p &= a \cos \omega t + b \sin \omega t \\ I_p' &= \omega(-a \sin \omega t + b \cos \omega t) \\ I_p'' &= \omega^2(-a \cos \omega t - b \sin \omega t) \end{aligned}$$

into (1). Then we collect the cosine terms and equate them to $E_0 \omega \cos \omega t$ on the right, and we equate the sine terms to zero because there is no sine term on the right,

$$\begin{aligned} L\omega^2(-a) + R\omega b + a/C &= E_0 \omega && \text{(Cosine terms)} \\ L\omega^2(-b) + R\omega(-a) + b/C &= 0 && \text{(Sine terms).} \end{aligned}$$

Before solving this system for a and b, we first introduce a combination of L and C, called the **reactance**

$$\textbf{(3)} \qquad \boxed{S = \omega L - \frac{1}{\omega C}.}$$

Dividing the previous two equations by ω, ordering them, and substituting S gives

$$\begin{aligned} -Sa + Rb &= E_0 \\ -Ra - Sb &= 0. \end{aligned}$$

We now eliminate b by multiplying the first equation by S and the second by R, and adding. Then we eliminate a by multiplying the first equation by R and the second by $-S$, and adding. This gives

$$-(S^2 + R^2)a = E_0 S, \qquad (R^2 + S^2)b = E_0 R.$$

We can solve for a and b,

$$a = \frac{-E_0 S}{R^2 + S^2}, \qquad b = \frac{E_0 R}{R^2 + S^2}. \tag{4}$$

Equation (2) with coefficients a and b given by (4) is the desired particular solution I_p of the nonhomogeneous ODE (1) governing the current I in an RLC-circuit with sinusoidal electromotive force.

Using (4), we can write I_p in terms of "physically visible" quantities, namely, amplitude I_0 and phase lag θ of the current behind the EMF, that is,

$$I_p(t) = I_0 \sin(\omega t - \theta) \tag{5}$$

where [see (14) in App. A3.1]

$$I_0 = \sqrt{a^2 + b^2} = \frac{E_0}{\sqrt{R^2 + S^2}}, \qquad \tan\theta = -\frac{a}{b} = \frac{S}{R}.$$

The quantity $\sqrt{R^2 + S^2}$ is called the **impedance**. Our formula shows that the impedance equals the ratio E_0/I_0. This is somewhat analogous to $E/I = R$ (Ohm's law) and, because of this analogy, the impedance is also known as the **apparent resistance**.

A general solution of the homogeneous equation corresponding to (1) is

$$I_h = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t}$$

where λ_1 and λ_2 are the roots of the characteristic equation

$$\lambda^2 + \frac{R}{L}\lambda + \frac{1}{LC} = 0.$$

We can write these roots in the form $\lambda_1 = -\alpha + \beta$ and $\lambda_2 = -\alpha - \beta$, where

$$\alpha = \frac{R}{2L}, \qquad \beta = \sqrt{\frac{R^2}{4L^2} - \frac{1}{LC}} = \frac{1}{2L}\sqrt{R^2 - \frac{4L}{C}}.$$

Now in an actual circuit, R is never zero (hence $R > 0$). From this it follows that I_h approaches zero, theoretically as $t \to \infty$, but practically after a relatively short time. Hence the transient current $I = I_h + I_p$ tends to the steady-state current I_p, and after some time the output will practically be a harmonic oscillation, which is given by (5) and whose frequency is that of the input (of the electromotive force).

EXAMPLE 1 ***RLC*-Circuit**

Find the current $I(t)$ in an RLC-circuit with $R = 11\ \Omega$ (ohms), $L = 0.1$ H (henry), $C = 10^{-2}$ F (farad), which is connected to a source of EMF $E(t) = 110 \sin (60 \cdot 2\pi t) = 110 \sin 377\, t$ (hence 60 Hz = 60 cycles/sec, the usual in the U.S. and Canada; in Europe it would be 220 V and 50 Hz). Assume that current and capacitor charge are 0 when $t = 0$.

Solution. ***Step 1. General solution of the homogeneous ODE.*** Substituting R, L, C and the derivative $E'(t)$ into (1), we obtain

$$0.1I'' + 11I' + 100I = 110 \cdot 377 \cos 377t.$$

Hence the homogeneous ODE is $0.1I'' + 11I' + 100I = 0$. Its characteristic equation is

$$0.1\lambda^2 + 11\lambda + 100 = 0.$$

The roots are $\lambda_1 = -10$ and $\lambda_2 = -100$. The corresponding general solution of the homogeneous ODE is

$$I_h(t) = c_1e^{-10t} + c_2e^{-100t}.$$

Step 2. Particular solution I_p of (1). We calculate the reactance $S = 37.7 - 0.3 = 37.4$ and the steady-state current

$$I_p(t) = a \cos 377t + b \sin 377t$$

with coefficients obtained from (4) (and rounded)

$$a = \frac{-110 \cdot 37.4}{11^2 + 37.4^2} = -2.71, \qquad b = \frac{110 \cdot 11}{11^2 + 37.4^2} = 0.796.$$

Hence in our present case, a general solution of the nonhomogeneous ODE (1) is

$$I(t) = c_1e^{-10t} + c_2e^{-100t} - 2.71 \cos 377t + 0.796 \sin 377t. \tag{6}$$

Step 3. Particular solution satisfying the initial conditions. **How to use $Q(0) = 0$?** We finally determine c_1 and c_2 from the in initial conditions $I(0) = 0$ and $Q(0) = 0$. From the first condition and (6) we have

$$I(0) = c_1 + c_2 - 2.71 = 0, \qquad \text{hence} \qquad c_2 = 2.71 - c_1. \tag{7}$$

We turn to $Q(0) = 0$. The integral in (1′) equals $\int I\, dt = Q(t)$; see near the beginning of this section. Hence for $t = 0$, Eq. (1′) becomes

$$LI'(0) + R \cdot 0 = 0, \qquad \text{so that} \qquad I'(0) = 0.$$

Differentiating (6) and setting $t = 0$, we thus obtain

$$I'(0) = -10c_1 - 100c_2 + 0 + 0.796 \cdot 377 = 0, \qquad \text{hence by (7),} \qquad -10c_1 = 100(2.71 - c_1) - 300.1.$$

The solution of this and (7) is $c_1 = -0.323$, $c_2 = 3.033$. Hence the answer is

$$I(t) = -0.323e^{-10t} + 3.033e^{-100t} - 2.71 \cos 377t + 0.796 \sin 377t.$$

You may get slightly different values depending on the rounding. Figure 63 shows $I(t)$ as well as $I_p(t)$, which practically coincide, except for a very short time near $t = 0$ because the exponential terms go to zero very rapidly. Thus after a very short time the current will practically execute harmonic oscillations of the input frequency 60 Hz = 60 cycles/sec. Its maximum amplitude and phase lag can be seen from (5), which here takes the form

$$I_p(t) = 2.824 \sin (377t - 1.29).$$

■

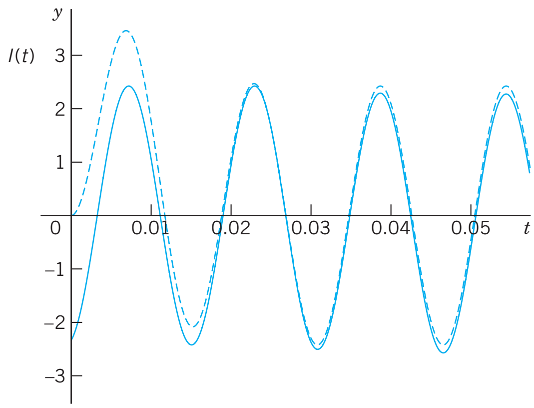


Fig. 63. Transient (upper curve) and steady-state currents in Example 1

Analogy of Electrical and Mechanical Quantities

Entirely different physical or other systems may have the same mathematical model. For instance, we have seen this from the various applications of the ODE $y' = ky$ in Chap. 1. Another impressive demonstration of this ***unifying power of mathematics*** is given by the ODE (1) for an electric *RLC*-circuit and the ODE (2) in the last section for a mass–spring system. Both equations

$$LI'' + RI' + \frac{1}{C}I = E_0\omega \cos \omega t \qquad \text{and} \qquad my'' + cy' + ky = F_0 \cos \omega t$$

are of the same form. Table 2.2 shows the analogy between the various quantities involved. The inductance L corresponds to the mass m and, indeed, an inductor opposes a change in current, having an "inertia effect" similar to that of a mass. The resistance R corresponds to the damping constant c, and a resistor causes loss of energy, just as a damping dashpot does. And so on.

This analogy is ***strictly quantitative*** in the sense that to a given mechanical system we can construct an electric circuit whose current will give the exact values of the displacement in the mechanical system when suitable scale factors are introduced.

The ***practical importance*** of this analogy is almost obvious. The analogy may be used for constructing an "electrical model" of a given mechanical model, resulting in substantial savings of time and money because electric circuits are easy to assemble, and electric quantities can be measured much more quickly and accurately than mechanical ones.

Table 2.2 Analogy of Electrical and Mechanical Quantities

Electrical System	Mechanical System
Inductance L	Mass m
Resistance R	Damping constant c
Reciprocal $1/C$ of capacitance	Spring modulus k
Derivative $E_0\omega \cos \omega t$ of electromotive force	Driving force $F_0 \cos \omega t$
Current $I(t)$	Displacement $y(t)$

Related to this analogy are **transducers**, devices that convert changes in a mechanical quantity (for instance, in a displacement) into changes in an electrical quantity that can be monitored; see Ref. [GenRef11] in App. 1.

PROBLEM SET 2.9

1–6 *RLC*-CIRCUITS: SPECIAL CASES

1. ***RC*-Circuit.** Model the *RC*-circuit in Fig. 64. Find the current due to a constant E.

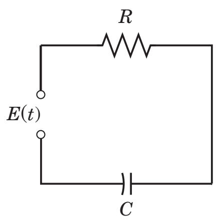


Fig. 64. *RC*-circuit

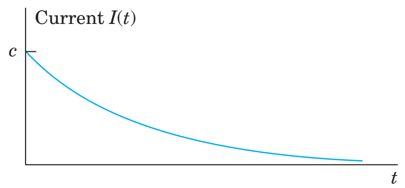


Fig. 65. Current 1 in Problem 1

2. ***RC*-Circuit.** Solve Prob. 1 when $E = E_0 \sin \omega t$ and R, C, E_0, and ω are arbitrary.
3. ***RL*-Circuit.** Model the *RL*-circuit in Fig. 66. Find a general solution when R, L, E are any constants. Graph or sketch solutions when $L = 0.25$ H, $R = 10\ \Omega$, and $E = 48$ V.

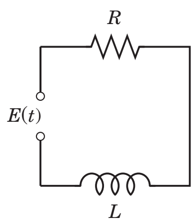


Fig. 66. *RL*-circuit

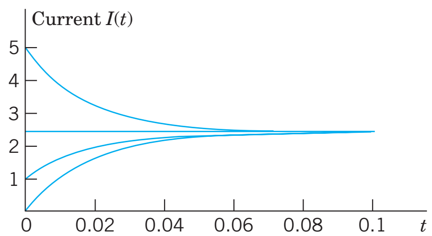


Fig. 67. Currents in Problem 3

4. ***RL*-Circuit.** Solve Prob. 3 when $E = E_0 \sin \omega t$ and R, L, E_0, and are arbitrary. Sketch a typical solution.

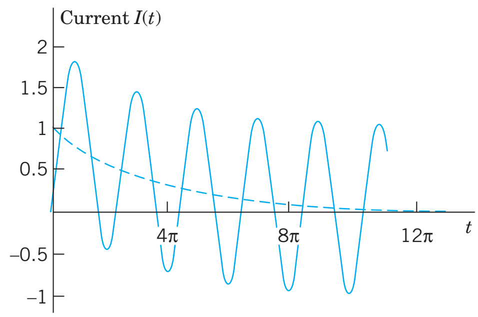


Fig. 68. Typical current $I = e^{-0.1t} + \sin(t - \frac{1}{4}\pi)$ in Problem 4

5. ***LC*-Circuit.** This is an *RLC*-circuit with negligibly small R (analog of an undamped mass–spring system). Find the current when $L = 0.5$ H, $C = 0.005$ F, and $E = \sin t$ V, assuming zero initial current and charge.

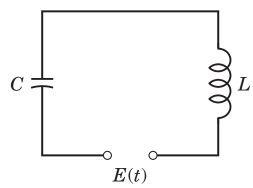


Fig. 69. *LC*-circuit

6. ***LC*-Circuit.** Find the current when $L = 0.5$ H, $C = 0.005$ F, $E = 2t^2$ V, and initial current and charge zero.

7–18 GENERAL *RLC*-CIRCUITS

7. **Tuning.** In tuning a stereo system to a radio station, we adjust the tuning control (turn a knob) that changes C (or perhaps L) in an *RLC*-circuit so that the amplitude of the steady-state current (5) becomes maximum. For what C will this happen?

8–14 Find the **steady-state current** in the *RLC*-circuit in Fig. 61 for the given data. Show the details of your work.

8. $R = 4\ \Omega$, $L = 0.5$ H, $C = 0.1$ F, $E = 500 \sin 2t$ V
9. $R = 4\ \Omega$, $L = 0.1$ H, $C = 0.05$ F, $E = 110$ V
10. $R = 2\ \Omega$, $L = 1$ H, $C = \frac{1}{20}$ F, $E = 157 \sin 3t$ V

11. $R = 12\ \Omega$, $L = 0.4$ H, $C = \frac{1}{80}$ F, $E = 220 \sin 10t$ V

12. $R = 0.2\ \Omega$, $L = 0.1$ H, $C = 2$ F, $E = 220 \sin 314t$ V

13. $R = 12$, $L = 1.2$ H, $C = \frac{20}{3} \cdot 10^{-3}$ F, $E = 12{,}000 \sin 25t$ V

14. Prove the claim in the text that if $R \neq 0$ (hence $R > 0$), then the transient current approaches I_p as $t \to \infty$.

15. Cases of damping. What are the conditions for an *RLC*-circuit to be (I) overdamped, (II) critically damped, (III) underdamped? What is the critical resistance R_{crit} (the analog of the critical damping constant $2\sqrt{mk}$)?

16–18 Solve the **initial value problem** for the *RLC*-circuit in Fig. 61 with the given data, assuming zero initial current and charge. Graph or sketch the solution. Show the details of your work.

16. $R = 8\ \Omega$, $L = 0.2$ H, $C = 12.5 \cdot 10^{-3}$ F, $E = 100 \sin 10t$ V

17. $R = 6\ \Omega$, $L = 1$ H, $C = 0.04$ F, $E = 600 (\cos t + 4 \sin t)$ V

18. $R = 18\ \Omega$, $L = 1$ H, $C = 12.5 \cdot 10^{-3}$ F, $E = 820 \cos 10t$ V

19. WRITING REPORT. Mechanic-Electric Analogy. Explain Table 2.2 in a 1–2 page report with examples, e.g., the analog (with $L = 1$ H) of a mass–spring system of mass 5 kg, damping constant 10 kg/sec, spring constant 60 kg/sec^2, and driving force $220 \cos 10t$ kg/sec.

20. Complex Solution Method. Solve $L\tilde{I}'' + R\tilde{I}' + \tilde{I}/C = E_0 e^{i\omega t}$, $i = \sqrt{-1}$, by substituting $I_p = Ke^{i\omega t}$ (K unknown) and its derivatives and taking the real part I_p of the solution $\tilde{I}_p$. Show agreement with (2), (4). *Hint*: Use (11) $e^{i\omega t} = \cos \omega t + i \sin \omega t$; cf. Sec. 2.2, and $i^2 = -1$.

2.10 Solution by Variation of Parameters

We continue our discussion of nonhomogeneous linear ODEs, that is

$$y'' + p(x)y' + q(x)y = r(x). \tag{1}$$

In Sec. 2.6 we have seen that a general solution of (1) is the sum of a general solution y_h of the corresponding homogeneous ODE and any particular solution y_p of (1). To obtain y_p when $r(x)$ is not too complicated, we can often use the *method of undetermined coefficients,* as we have shown in Sec. 2.7 and applied to basic engineering models in Secs. 2.8 and 2.9.

However, since this method is restricted to functions $r(x)$ whose derivatives are of a form similar to $r(x)$ itself (powers, exponential functions, etc.), it is desirable to have a method valid for more general ODEs (1), which we shall now develop. It is called the **method of variation of parameters** and is credited to Lagrange (Sec. 2.1). Here p, q, r in (1) may be variable (given functions of x), but we assume that they are continuous on some open interval I.

Lagrange's method gives a particular solution y_p of (1) on I in the form

$$y_p(x) = -y_1 \int \frac{y_2 r}{W}\, dx + y_2 \int \frac{y_1 r}{W}\, dx \tag{2}$$

where y_1, y_2 form a basis of solutions of the corresponding homogeneous ODE

$$y'' + p(x)y' + q(x)y = 0 \tag{3}$$

on I, and W is the Wronskian of y_1, y_2,

$$W = y_1 y_2' - y_2 y_1' \qquad \text{(see Sec. 2.6).} \tag{4}$$

CAUTION! The solution formula (2) is obtained under the assumption that the ODE is written in standard form, with y'' as the first term as shown in (1). If it starts with $f(x)y''$, divide first by $f(x)$.

The integration in (2) may often cause difficulties, and so may the determination of y_1, y_2 if (1) has variable coefficients. If you have a choice, use the previous method. It is simpler. Before deriving (2) let us work an example for which you *do need* the new method. (Try otherwise.)

EXAMPLE 1 Method of Variation of Parameters

Solve the nonhomogeneous ODE

$$y'' + y = \sec x = \frac{1}{\cos x}.$$

Solution. A basis of solutions of the homogeneous ODE on any interval is $y_1 = \cos x$, $y_2 = \sin x$. This gives the Wronskian

$$W(y_1, y_2) = \cos x \cos x - \sin x\,(-\sin x) = 1.$$

From (2), choosing zero constants of integration, we get the particular solution of the given ODE

$$\begin{aligned} y_p &= -\cos x \int \sin x \sec x \, dx + \sin x \int \cos x \sec x \, dx \\ &= \cos x \ln |\cos x| + x \sin x \end{aligned} \qquad \text{(Fig. 70)}$$

Figure 70 shows y_p and its first term, which is small, so that $x \sin x$ essentially determines the shape of the curve of y_p. (Recall from Sec. 2.8 that we have seen $x \sin x$ in connection with resonance, except for notation.) From y_p and the general solution $y_h = c_1 y_1 + c_2 y_2$ of the homogeneous ODE we obtain the *answer*

$$y = y_h + y_p = (c_1 + \ln |\cos x|) \cos x + (c_2 + x) \sin x.$$

Had we included integration constants $-c_1, c_2$ in (2), then (2) would have given the additional $c_1 \cos x + c_2 \sin x = c_1 y_1 + c_2 y_2$, that is, a general solution of the given ODE directly from (2). This will always be the case. ■

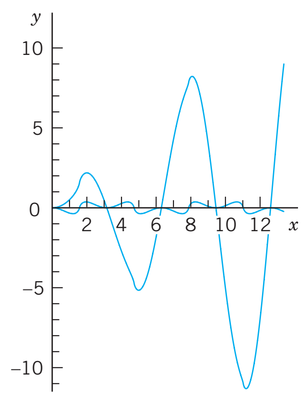


Fig. 70. Particular solution y_p and its first term in Example 1

Idea of the Method. Derivation of (2)

What idea did Lagrange have? What gave the method the name? Where do we use the continuity assumptions?

The idea is to start from a general solution

$$y_h(x) = c_1 y_1(x) + c_2 y_2(x)$$

of the homogeneous ODE (3) on an open interval I and to replace the constants ("the parameters") c_1 and c_2 by functions $u(x)$ and $v(x)$; this suggests the name of the method. We shall determine u and v so that the resulting function

$$y_p(x) = u(x)y_1(x) + v(x)y_2(x) \tag{5}$$

is a particular solution of the nonhomogeneous ODE (1). Note that y_h exists by Theorem 3 in Sec. 2.6 because of the continuity of p and q on I. (The continuity of r will be used later.)

We determine u and v by substituting (5) and its derivatives into (1). Differentiating (5), we obtain

$$y_p' = u'y_1 + uy_1' + v'y_2 + vy_2'.$$

Now y_p must satisfy (1). This is *one* condition for *two* functions u and v. It seems plausible that we may impose a *second* condition. Indeed, our calculation will show that we can determine u and v such that y_p satisfies (1) and u and v satisfy as a second condition the equation

$$u'y_1 + v'y_2 = 0. \tag{6}$$

This reduces the first derivative y_p' to the simpler form

$$y_p' = uy_1' + vy_2'. \tag{7}$$

Differentiating (7), we obtain

$$y_p'' = u'y_1' + uy_1'' + v'y_2' + vy_2''. \tag{8}$$

We now substitute y_p and its derivatives according to (5), (7), (8) into (1). Collecting terms in u and terms in v, we obtain

$$u(y_1'' + py_1' + qy_1) + v(y_2'' + py_2' + qy_2) + u'y_1' + v'y_2' = r.$$

Since y_1 and y_2 are solutions of the homogeneous ODE (3), this reduces to

$$u'y_1' + v'y_2' = r. \tag{9a}$$

Equation (6) is

$$u'y_1 + v'y_2 = 0. \tag{9b}$$

This is a linear system of two algebraic equations for the unknown functions u' and v'. We can solve it by elimination as follows (or by Cramer's rule in Sec. 7.6). To eliminate v', we multiply (9a) by $-y_2$ and (9b) by y_2' and add, obtaining

$$u'(y_1y_2' - y_2y_1') = -y_2r, \qquad \text{thus} \qquad u'W = -y_2r.$$

Here, W is the Wronskian (4) of y_1, y_2. To eliminate u' we multiply (9a) by y_1, and (9b) by $-y_1'$ and add, obtaining

$$v'(y_1y_2' - y_2y_1') = -y_1r, \qquad \text{thus} \qquad v'W = y_1r.$$

Since y_1, y_2 form a basis, we have $W \neq 0$ (by Theorem 2 in Sec. 2.6) and can divide by W,

$$u' = -\frac{y_2r}{W}, \qquad v' = \frac{y_1r}{W}. \tag{10}$$

By integration,

$$u = -\int \frac{y_2r}{W}\,dx, \qquad v = \int \frac{y_1r}{W}\,dx.$$

These integrals exist because $r(x)$ is continuous. Inserting them into (5) gives (2) and completes the derivation. ■

PROBLEM SET 2.10

1–13 GENERAL SOLUTION

Solve the given nonhomogeneous linear ODE by variation of parameters or undetermined coefficients. Show the details of your work.

1. $y'' + 9y = \sec 3x$
2. $y'' + 9y = \csc 3x$
3. $x^2y'' - 2xy' + 2y = x^3 \sin x$
4. $y'' - 4y' + 5y = e^{2x} \csc x$
5. $y'' + y = \cos x - \sin x$
6. $(D^2 + 6D + 9I)y = 16e^{-3x}/(x^2 + 1)$
7. $(D^2 - 4D + 4I)y = 6e^{2x}/x^4$
8. $(D^2 + 4I)y = \cosh 2x$
9. $(D^2 - 2D + I)y = 35x^{3/2}e^x$
10. $(D^2 + 2D + 2I)y = 4e^{-x} \sec^3 x$
11. $(x^2D^2 - 4xD + 6I)y = 21x^{-4}$
12. $(D^2 - I)y = 1/\cosh x$
13. $(x^2D^2 + xD - 9I)y = 48x^5$
14. **TEAM PROJECT. Comparison of Methods. Invention.** The undetermined-coefficient method should be used whenever possible because it is simpler. Compare it with the present method as follows.

    (a) Solve $y'' + 4y' + 3y = 65 \cos 2x$ by both methods, showing all details, and compare.

    (b) Solve $y'' - 2y' + y = r_1 + r_2$, $r_1 = 35x^{3/2}e^x$ $r_2 = x^2$ by applying each method to a suitable function on the right.

    (c) Experiment to invent an undetermined-coefficient method for nonhomogeneous Euler–Cauchy equations.

CHAPTER 2 REVIEW QUESTIONS AND PROBLEMS

1. Why are linear ODEs preferable to nonlinear ones in modeling?
2. What does an initial value problem of a second-order ODE look like? Why must you have a general solution to solve it?
3. By what methods can you get a general solution of a nonhomogeneous ODE from a general solution of a homogeneous one?
4. Describe applications of ODEs in mechanical systems. What are the electrical analogs of the latter?
5. What is resonance? How can you remove undesirable resonance of a construction, such as a bridge, a ship, or a machine?
6. What do you know about existence and uniqueness of solutions of linear second-order ODEs?

7–18 GENERAL SOLUTION

Find a general solution. Show the details of your calculation.

7. $4y'' + 32y' + 63y = 0$
8. $y'' + y' - 12y = 0$
9. $y'' + 6y' + 34y = 0$
10. $y'' + 0.20y' + 0.17y = 0$
11. $(100D^2 - 160D + 64I)y = 0$
12. $(D^2 + 4\pi D + 4\pi^2 I)y = 0$
13. $(x^2D^2 + 2xD - 12I)y = 0$
14. $(x^2D^2 + xD - 9I)y = 0$
15. $(2D^2 - 3D - 2I)y = 13 - 2x^2$
16. $(D^2 + 2D + 2I)y = 3e^{-x} \cos 2x$
17. $(4D^2 - 12D + 9I)y = 2e^{1.5x}$
18. $yy'' = 2y'^2$

19–22 INITIAL VALUE PROBLEMS

Solve the problem, showing the details of your work. Sketch or graph the solution.

19. $y'' + 16y = 17e^x, \quad y(0) = 6, \quad y'(0) = -2$

20. $y'' - 3y' + 2y = 10 \sin x, \quad y(0) = 1, \quad y'(0) = -6$

21. $(x^2D^2 + xD - I)y = 16x^3, \quad y(1) = -1, \quad y'(1) = 1$

22. $(x^2D^2 + 15xD + 49I)y = 0, \quad y(1) = 2, \quad y'(1) = -11$

23–30 APPLICATIONS

23. Find the steady-state current in the *RLC*-circuit in Fig. 71 when $R = 2\ \text{k}\Omega\ (2000\ \Omega)$, $L = 1$ H, $C = 4 \cdot 10^{-3}$ F, and $E = 110 \sin 415t$ V (66 cycles/sec).

24. Find a general solution of the homogeneous linear ODE corresponding to the ODE in Prob. 23.

25. Find the steady-state current in the *RLC*-circuit in Fig. 71 when $R = 50\ \Omega$, $L = 30$ H, $C = 0.025$ F, $E = 200 \sin 4t$ V.

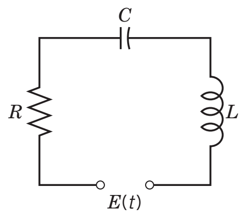


Fig. 71. *RLC*-circuit

26. Find the current in the *RLC*-circuit in Fig. 71 when $R = 40\ \Omega$, $L = 0.4$ H, $C = 10^{-4}$ F, $E = 220 \sin 314t$ V (50 cycles/sec).

27. Find an electrical analog of the mass–spring system with mass 4 kg, spring constant 10 kg/sec^2, damping constant 20 kg/sec, and driving force $100 \sin 4t$ nt.

28. Find the motion of the mass–spring system in Fig. 72 with mass 0.125 kg, damping 0, spring constant 1.125 kg/sec^2, and driving force $\cos t - 4 \sin t$ nt, assuming zero initial displacement and velocity. For what frequency of the driving force would you get resonance?

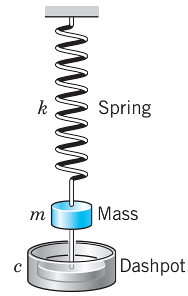


Fig. 72. Mass–spring system

29. Show that the system in Fig. 72 with $m = 4$, $c = 0$, $k = 36$, and driving force $61 \cos 3.1t$ exhibits beats. *Hint:* Choose zero initial conditions.

30. In Fig. 72, let $m = 1$ kg, $c = 4$ kg/sec, $k = 24$ kg/sec^2, and $r(t) = 10 \cos \omega t$ nt. Determine w such that you get the steady-state vibration of maximum possible amplitude. Determine this amplitude. Then find the general solution with this ω and check whether the results are in agreement.

SUMMARY OF CHAPTER 2
Second-Order Linear ODEs

Second-order linear ODEs are particularly important in applications, for instance, in mechanics (Secs. 2.4, 2.8) and electrical engineering (Sec. 2.9). A second-order ODE is called **linear** if it can be written

$$y'' + p(x)y' + q(x)y = r(x) \qquad \text{(Sec. 2.1).} \tag{1}$$

(If the first term is, say, $f(x)y''$, divide by $f(x)$ to get the "**standard form**" (1) with y'' as the first term.) Equation (1) is called **homogeneous** if $r(x)$ is zero for all x considered, usually in some open interval; this is written $r(x) \equiv 0$. Then

$$y'' + p(x)y' + q(x)y = 0. \tag{2}$$

Equation (1) is called **nonhomogeneous** if $r(x) \not\equiv 0$ (meaning $r(x)$ is not zero for some x considered).

For the homogeneous ODE (2) we have the important **superposition principle** (Sec. 2.1) that a linear combination $y = ky_1 + ly_2$ of two solutions y_1, y_2 is again a solution.

Two *linearly independent* solutions y_1, y_2 of (2) on an open interval I form a **basis** (or **fundamental system**) of solutions on I. and $y = c_1y_1 + c_2y_2$ with arbitrary constants c_1, c_2 a **general solution** of (2) on I. From it we obtain a **particular solution** if we specify numeric values (numbers) for c_1 and c_2, usually by prescribing two **initial conditions**

$$(3) \qquad y(x_0) = K_0, \qquad y'(x_0) = K_1 \qquad (x_0, K_0, K_1 \text{ given numbers; Sec. 2.1}).$$

(2) and (3) together form an **initial value problem**. Similarly for (1) and (3).

For a nonhomogeneous ODE (1) a **general solution** is of the form

$$(4) \qquad y = y_h + y_p \qquad \text{(Sec. 2.7)}.$$

Here y_h is a general solution of (2) and y_p is a particular solution of (1). Such a y_p can be determined by a general method (***variation of parameters***, Sec. 2.10) or in many practical cases by the ***method of undetermined coefficients***. The latter applies when (1) has constant coefficients p and q, and $r(x)$ is a power of x, sine, cosine, etc. (Sec. 2.7). Then we write (1) as

$$(5) \qquad y'' + ay' + by = r(x) \qquad \text{(Sec. 2.7)}.$$

The corresponding homogeneous ODE $y' + ay' + by = 0$ has solutions $y = e^{\lambda x}$, where λ is a root of

$$(6) \qquad \lambda^2 + a\lambda + b = 0.$$

Hence there are three cases (Sec. 2.2):

Case	Type of Roots	General Solution
I	Distinct real λ_1, λ_2	$y = c_1e^{\lambda_1 x} + c_2e^{\lambda_2 x}$
II	Double $-\frac{1}{2}a$	$y = (c_1 + c_2x)e^{-ax/2}$
III	Complex $-\frac{1}{2}a \pm i\omega^*$	$y = e^{-ax/2}(A \cos \omega^* x + B \sin \omega^* x)$

Here ω^* is used since ω is needed in driving forces.

Important applications of (5) in mechanical and electrical engineering in connection with ***vibrations*** and ***resonance*** are discussed in Secs. 2.4, 2.7, and 2.8.

Another large class of ODEs solvable "algebraically" consists of the **Euler–Cauchy equations**

$$(7) \qquad x^2y'' + axy' + by = 0 \qquad \text{(Sec. 2.5)}.$$

These have solutions of the form $y = x^m$, where m is a solution of the auxiliary equation

$$(8) \qquad m^2 + (a - 1)m + b = 0.$$

Existence and uniqueness of solutions of (1) and (2) is discussed in Secs. 2.6 and 2.7, and ***reduction of order*** in Sec. 2.1.

CHAPTER 3

Higher Order Linear ODEs

The concepts and methods of solving linear ODEs of order $n = 2$ extend nicely to linear ODEs of higher order n, that is, $n = 3, 4$, etc. This shows that the theory explained in Chap. 2 for second-order linear ODEs is attractive, since it can be extended in a straightforward way to arbitrary n. We do so in this chapter and notice that the formulas become more involved, the variety of roots of the characteristic equation (in Sec. 3.2) becomes much larger with increasing n, and the Wronskian plays a more prominent role.

The concepts and methods of solving second-order linear ODEs extend readily to linear ODEs of higher order.

This chapter follows Chap. 2 naturally, since the results of Chap. 2 can be readily extended to that of Chap. 3.

Prerequisite: Secs. 2.1, 2.2, 2.6, 2.7, 2.10.
References and Answers to Problems: App. 1 Part A, and App. 2.

3.1 Homogeneous Linear ODEs

Recall from Sec. 1.1 that an ODE is of ***n*th order** if the nth derivative $y^{(n)} = d^n y/dx^n$ of the unknown function $y(x)$ is the highest occurring derivative. Thus the ODE is of the form

$$F(x, y, y', \cdots, y^{(n)}) = 0$$

where lower order derivatives and y itself may or may not occur. Such an ODE is called **linear** if it can be written

$$y^{(n)} + p_{n-1}(x)y^{(n-1)} + \cdots + p_1(x)y' + p_0(x)y = r(x). \tag{1}$$

(For $n = 2$ this is (1) in Sec. 2.1 with $p_1 = p$ and $p_0 = q$.) The **coefficients** $p_0, \cdots, p_{n-1}$ and the function r on the right are any given functions of x, and y is unknown. $y^{(n)}$ has coefficient 1. We call this the **standard form**. (If you have $p_n(x)y^{(n)}$, divide by $p_n(x)$ to get this form.) An nth-order ODE that cannot be written in the form (1) is called **nonlinear**.

If $r(x)$ is identically zero, $r(x) \equiv 0$ (zero for all x considered, usually in some open interval I), then (1) becomes

$$y^{(n)} + p_{n-1}(x)y^{(n-1)} + \cdots + p_1(x)y' + p_0(x)y = 0 \tag{2}$$

and is called **homogeneous**. If $r(x)$ is not identically zero, then the ODE is called **nonhomogeneous**. This is as in Sec. 2.1.

A **solution** of an nth-order (linear or nonlinear) ODE on some open interval I is a function $y = h(x)$ that is defined and n times differentiable on I and is such that the ODE becomes an identity if we replace the unknown function y and its derivatives by h and its corresponding derivatives.

Sections 3.1–3.2 will be devoted to homogeneous linear ODEs and Section 3.3 to nonhomogeneous linear ODEs.

Homogeneous Linear ODE: Superposition Principle, General Solution

The basic **superposition or linearity principle** of Sec. 2.1 extends to nth order homogeneous linear ODEs as follows.

THEOREM 1

Fundamental Theorem for the Homogeneous Linear ODE (2)

For a homogeneous linear ODE (2), *sums and constant multiples of solutions on some open interval I are again solutions on I.* (This does ***not*** hold for a nonhomogeneous or nonlinear ODE!)

The proof is a simple generalization of that in Sec. 2.1 and we leave it to the student.

Our further discussion parallels and extends that for second-order ODEs in Sec. 2.1. So we next define a general solution of (2), which will require an extension of linear independence from 2 to n functions.

DEFINITION

General Solution, Basis, Particular Solution

A **general solution** of (2) on an open interval I is a solution of (2) on I of the form

$$y(x) = c_1 y_1(x) + \cdots + c_n y_n(x) \qquad (c_1, \cdots, c_n \text{ arbitrary}) \tag{3}$$

where $y_1, \cdots, y_n$ is a **basis** (or **fundamental system**) of solutions of (2) on I; that is, these solutions are linearly independent on I, as defined below.

A **particular solution** of (2) on I is obtained if we assign specific values to the n constants $c_1, \cdots, c_n$ in (3).

DEFINITION

Linear Independence and Dependence

Consider n functions $y_1(x), \cdots, y_n(x)$ defined *on some interval I.*
These functions are called **linearly independent** on I if the equation

$$k_1 y_1(x) + \cdots + k_n y_n(x) = 0 \qquad \text{on } I \tag{4}$$

implies that all $k_1, \cdots, k_n$ are zero. These functions are called **linearly dependent** on I if this equation also holds on I for some $k_1, \cdots, k_n$ not all zero.

If and only if $y_1, \cdots, y_n$ are linearly dependent on I, we can express (at least) one of these functions on I as a "**linear combination**" of the other $n - 1$ functions, that is, as a sum of those functions, each multiplied by a constant (zero or not). This motivates the term "linearly dependent." For instance, if (4) holds with $k_1 \neq 0$, we can divide by k_1 and express y_1 as the linear combination

$$y_1 = -\frac{1}{k_1}(k_2 y_2 + \cdots + k_n y_n).$$

Note that when $n = 2$, these concepts reduce to those defined in Sec. 2.1.

EXAMPLE 1 **Linear Dependence**

Show that the functions $y_1 = x^2$, $y_2 = 5x$, $y_3 = 2x$ are linearly dependent on any interval.

Solution. $y_2 = 0y_1 + 2.5y_3$. This proves linear dependence on any interval. ■

EXAMPLE 2 **Linear Independence**

Show that $y_1 = x$, $y_2 = x^2$, $y_3 = x^3$ are linearly independent on any interval, for instance, on $-1 \leqq x \leqq 2$.

Solution. Equation (4) is $k_1 x + k_2 x^2 + k_3 x^3 = 0$. Taking (a) $x = -1$, (b) $x = 1$, (c) $x = 2$, we get

$$\text{(a) } -k_1 + k_2 - k_3 = 0, \qquad \text{(b) } k_1 + k_2 + k_3 = 0, \qquad \text{(c) } 2k_1 + 4k_2 + 8k_3 = 0.$$

$k_2 = 0$ from (a) + (b). Then $k_3 = 0$ from (c) -2(b). Then $k_1 = 0$ from (b). This proves linear independence.

A better method for testing linear independence of solutions of ODEs will soon be explained. ■

EXAMPLE 3 **General Solution. Basis**

Solve the fourth-order ODE

$$y^{\text{iv}} - 5y'' + 4y = 0 \qquad \text{(where } y^{\text{iv}} = d^4y/dx^4\text{)}.$$

Solution. As in Sec. 2.2 we substitute $y = e^{\lambda x}$. Omitting the common factor $e^{\lambda x}$, we obtain the characteristic equation

$$\lambda^4 - 5\lambda^2 + 4 = 0.$$

This is a quadratic equation in $\mu = \lambda^2$, namely,

$$\mu^2 - 5\mu + 4 = (\mu - 1)(\mu - 4) = 0.$$

The roots are $\mu = 1$ and 4. Hence $\lambda = -2, -1, 1, 2$. This gives four solutions. A general solution on any interval is

$$y = c_1 e^{-2x} + c_2 e^{-x} + c_3 e^{x} + c_4 e^{2x}$$

provided those four solutions are linearly independent. This is true but will be shown later. ■

Initial Value Problem. Existence and Uniqueness

An **initial value problem** for the ODE (2) consists of (2) and n **initial conditions**

$$\text{(5)} \qquad y(x_0) = K_0, \qquad y'(x_0) = K_1, \qquad \cdots, \qquad y^{(n-1)}(x_0) = K_{n-1}$$

with given x_0 in the open interval I considered, and given $K_0, \cdots, K_{n-1}$.

In extension of the existence and uniqueness theorem in Sec. 2.6 we now have the following.

THEOREM 2

Existence and Uniqueness Theorem for Initial Value Problems

If the coefficients $p_0(x), \cdots, p_{n-1}(x)$ of (2) *are continuous on some open interval I and x_0 is in I, then the initial value problem* (2), (5) *has a unique solution $y(x)$ on I.*

Existence is proved in Ref. [A11] in App. 1. Uniqueness can be proved by a slight generalization of the uniqueness proof at the beginning of App. 4.

EXAMPLE 4 **Initial Value Problem for a Third-Order Euler–Cauchy Equation**

Solve the following initial value problem on any open interval I on the positive x-axis containing $x = 1$.

$$x^3y''' - 3x^2y'' + 6xy' - 6y = 0, \qquad y(1) = 2, \qquad y'(1) = 1, \qquad y''(1) = -4.$$

Solution. ***Step 1. General solution.*** As in Sec. 2.5 we try $y = x^m$. By differentiation and substitution,

$$m(m-1)(m-2)x^m - 3m(m-1)x^m + 6mx^m - 6x^m = 0.$$

Dropping x^m and ordering gives $m^3 - 6m^2 + 11m - 6 = 0$. If we can guess the root $m = 1$. We can divide by $m - 1$ and find the other roots 2 and 3, thus obtaining the solutions x, x^2, x^3, which are linearly independent on I (see Example 2). [In general one shall need a root-finding method, such as Newton's (Sec. 19.2), also available in a CAS (Computer Algebra System).] Hence a general solution is

$$y = c_1x + c_2x^2 + c_3x^3$$

valid on any interval I, even when it includes $x = 0$ where the coefficients of the ODE divided by x^3 (to have the standard form) are not continuous.

Step 2. Particular solution. The derivatives are $y' = c_1 + 2c_2x + 3c_3x^2$ and $y'' = 2c_2 + 6c_3x$. From this, and y and the initial conditions, we get by setting $x = 1$

$$\begin{aligned}
&\text{(a)}\ \ y(1) &&= c_1 + c_2 + c_3 = 2\\
&\text{(b)}\ \ y'(1) &&= c_1 + 2c_2 + 3c_3 = 1\\
&\text{(c)}\ \ y''(1) &&= 2c_2 + 6c_3 = -4.
\end{aligned}$$

This is solved by Cramer's rule (Sec. 7.6), or by elimination, which is simple, as follows. (b) − (a) gives (d) $c_2 + 2c_3 = -1$. Then (c) − 2(d) gives $c_3 = -1$. Then (c) gives $c_2 = 1$. Finally $c_1 = 2$ from (a). *Answer:* $y = 2x + x^2 - x^3$. ■

Linear Independence of Solutions. Wronskian

Linear independence of solutions is crucial for obtaining general solutions. Although it can often be seen by inspection, it would be good to have a criterion for it. Now Theorem 2 in Sec. 2.6 extends from order $n = 2$ to any n. This extended criterion uses the **Wronskian** W of n solutions $y_1, \cdots, y_n$ defined as the nth-order determinant

$$W(y_1, \cdots, y_n) = \begin{vmatrix} y_1 & y_2 & \cdots & y_n \\ y_1' & y_2' & \cdots & y_n' \\ \cdot & \cdot & \cdots & \cdot \\ y_1^{(n-1)} & y_2^{(n-1)} & \cdots & y_n^{(n-1)} \end{vmatrix}. \tag{6}$$

Note that W depends on x since $y_1, \cdots, y_n$ do. The criterion states that these solutions form a basis if and only if W is not zero; more precisely:

THEOREM 3

Linear Dependence and Independence of Solutions

Let the ODE (2) have continuous coefficients $p_0(x), \cdots, p_{n-1}(x)$ on an open interval I. Then n solutions $y_1, \cdots, y_n$ of (2) on I are linearly dependent on I if and only if their Wronskian is zero for some $x = x_0$ in I. Furthermore, if W is zero for $x = x_0$, then W is identically zero on I. Hence if there is an x_1 in I at which W is not zero, then $y_1, \cdots, y_n$ are linearly independent on I, so that they form a basis of solutions of (2) on I.

PROOF **(a)** Let $y_1, \cdots, y_n$ be linearly dependent solutions of (2) on I. Then, by definition, there are constants $k_1, \cdots, k_n$ *not all zero*, such that for all x in I,

$$k_1 y_1 + \cdots + k_n y_n = 0. \tag{7}$$

By $n - 1$ differentiations of (7) we obtain for all x in I

$$\begin{aligned} k_1 y_1' + \cdots + k_n y_n' &= 0 \\ &\vdots \\ k_1 y_1^{(n-1)} + \cdots + k_n y_n^{(n-1)} &= 0. \end{aligned} \tag{8}$$

(7), (8) is a homogeneous linear system of algebraic equations with a nontrivial solution $k_1, \cdots, k_n$. Hence its coefficient determinant must be zero for every x on I, by Cramer's theorem (Sec. 7.7). But that determinant is the Wronskian W, as we see from (6). Hence W is zero for every x on I.

(b) Conversely, if W is zero at an x_0 in I, then the system (7), (8) with $x = x_0$ has a solution $k_1^*, \cdots, k_n^*$, not all zero, by the same theorem. With these constants we define the solution $y^* = k_1^* y_1 + \cdots + k_n^* y_n$ of (2) on I. By (7), (8) this solution satisfies the initial conditions $y^*(x_0) = 0, \cdots, y^{*(n-1)}(x_0) = 0$. But another solution satisfying the same conditions is $y \equiv 0$. Hence $y^* \equiv y$ by Theorem 2, which applies since the coefficients of (2) are continuous. Together, $y^* = k_1^* y_1 + \cdots + k_n^* y_n \equiv 0$ on I. This means linear dependence of $y_1, \cdots, y_n$ on I.

(c) If W is zero at an x_0 in I, we have linear dependence by (b) and then $W \equiv 0$ by (a). Hence if W is not zero at an x_1 in I, the solutions $y_1, \cdots, y_n$ must be linearly independent on I. ∎

EXAMPLE 5 **Basis, Wronskian**

We can now prove that in Example 3 we do have a basis. In evaluating W, pull out the exponential functions columnwise. In the result, subtract Column 1 from Columns 2, 3, 4 (without changing Column 1). Then expand by Row 1. In the resulting third-order determinant, subtract Column 1 from Column 2 and expand the result by Row 2:

$$W = \begin{vmatrix} e^{-2x} & e^{-x} & e^{x} & e^{2x} \\ -2e^{-2x} & -e^{-x} & e^{x} & 2e^{2x} \\ 4e^{-2x} & e^{-x} & e^{x} & 4e^{2x} \\ -8e^{-2x} & -e^{-x} & e^{x} & 8e^{2x} \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 & 1 \\ -2 & -1 & 1 & 2 \\ 4 & 1 & 1 & 4 \\ -8 & -1 & 1 & 8 \end{vmatrix} = \begin{vmatrix} 1 & 3 & 4 \\ -3 & -3 & 0 \\ 7 & 9 & 16 \end{vmatrix} = 72.$$ ∎

A General Solution of (2) Includes All Solutions

Let us first show that general solutions always exist. Indeed, Theorem 3 in Sec. 2.6 extends as follows.

THEOREM 4

Existence of a General Solution

If the coefficients $p_0(x), \cdots, p_{n-1}(x)$ of (2) *are continuous on some open interval I, then* (2) *has a general solution on I.*

PROOF We choose any fixed x_0 in I. By Theorem 2 the ODE (2) has n solutions $y_1, \cdots, y_n$, where y_j satisfies initial conditions (5) with $K_{j-1} = 1$ and all other K's equal to zero. Their Wronskian at x_0 equals 1. For instance, when $n = 3$, then $y_1(x_0) = 1, y_2'(x_0) = 1$, $y_3''(x_0) = 1$, and the other initial values are zero. Thus, as claimed,

$$W(y_1(x_0), y_2(x_0), y_3(x_0)) = \begin{vmatrix} y_1(x_0) & y_2(x_0) & y_3(x_0) \\ y_1'(x_0) & y_2'(x_0) & y_3'(x_0) \\ y_1''(x_0) & y_2''(x_0) & y_3''(x_0) \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1.$$

Hence for any n those solutions $y_1, \cdots, y_n$ are linearly independent on I, by Theorem 3. They form a basis on I, and $y = c_1 y_1 + \cdots + c_n y_n$ is a general solution of (2) on I. ■

We can now prove the basic property that, from a general solution of (2), every solution of (2) can be obtained by choosing suitable values of the arbitrary constants. Hence an nth-order ***linear*** ODE has no **singular solutions**, that is, solutions that cannot be obtained from a general solution.

THEOREM 5

General Solution Includes All Solutions

If the ODE (2) *has continuous coefficients $p_0(x), \cdots, p_{n-1}(x)$ on some open interval I, then every solution $y = Y(x)$ of* (2) *on I is of the form*

$$Y(x) = C_1 y_1(x) + \cdots + C_n y_n(x) \tag{9}$$

where $y_1, \cdots, y_n$ is a basis of solutions of (2) *on I and $C_1, \cdots, C_n$ are suitable constants.*

PROOF Let Y be a given solution and $y = c_1 y_1 + \cdots + c_n y_n$ a general solution of (2) on I. We choose any fixed x_0 in I and show that we can find constants $c_1, \cdots, c_n$ for which y and its first $n - 1$ derivatives agree with Y and its corresponding derivatives at x_0. That is, we should have at $x = x_0$

$$\begin{aligned} c_1 y_1 + \cdots + c_n y_n &= Y \\ c_1 y_1' + \cdots + c_n y_n' &= Y' \\ &\vdots \\ c_1 y_1^{(n-1)} + \cdots + c_n y_n^{(n-1)} &= Y^{(n-1)}. \end{aligned} \tag{10}$$

But this is a linear system of equations in the unknowns $c_1, \cdots, c_n$. Its coefficient determinant is the Wronskian W of $y_1, \cdots, y_n$ at x_0. Since $y_1, \cdots, y_n$ form a basis, they are linearly independent, so that W is not zero by Theorem 3. Hence (10) has a unique solution $c_1 = C_1, \cdots, c_n = C_n$ (by Cramer's theorem in Sec. 7.7). With these values we obtain the particular solution

$$y^*(x) = C_1 y_1(x) + \cdots + C_n y_n(x)$$

on I. Equation (10) shows that y^* and its first $n - 1$ derivatives agree at x_0 with Y and its corresponding derivatives. That is, y^* and Y satisfy, at x_0, the same initial conditions. The uniqueness theorem (Theorem 2) now implies that $y^* \equiv Y$ on I. This proves the theorem. ■

This completes our theory of the homogeneous linear ODE (2). Note that for $n = 2$ it is identical with that in Sec. 2.6. This had to be expected.

PROBLEM SET 3.1

1–6 BASES: TYPICAL EXAMPLES

To get a feel for higher order ODEs, show that the given functions are solutions and form a basis on any interval. Use Wronskians. In Prob. 6, $x > 0$,

1. $1, x, x^2, x^3, \quad y^{\text{iv}} = 0$

2. $e^x, e^{-x}, e^{2x}, \quad y''' - 2y'' - y' + 2y = 0$

3. $\cos x, \sin x, x \cos x, x \sin x, \quad y^{\text{iv}} + 2y'' + y = 0$

4. $e^{-4x}, xe^{-4x}, x^2e^{-4x}, \quad y''' + 12y'' + 48y' + 64y = 0$

5. $1, e^{-x} \cos 2x, e^{-x} \sin 2x, \quad y''' + 2y'' + 5y' = 0$

6. $1, x^2, x^4, \quad x^2y''' - 3xy'' + 3y' = 0$

7. TEAM PROJECT. General Properties of Solutions of Linear ODEs. These properties are important in obtaining new solutions from given ones. Therefore extend Team Project 38 in Sec. 2.2 to nth-order ODEs. Explore statements on sums and multiples of solutions of (1) and (2) systematically and with proofs. Recognize clearly that no new ideas are needed in this extension from $n = 2$ to general n.

8–15 LINEAR INDEPENDENCE

Are the given functions linearly independent or dependent on the half-axis $x \geq 0$? Give reason.

8. $x^2, 1/x^2, 0$

9. $\tan x, \cot x, 1$

10. $e^{2x}, xe^{2x}, x^2e^{2x}$

11. $e^x \cos x, e^x \sin x, e^x$

12. $\sin^2 x, \cos^2 x, \cos 2x$

13. $\sin x, \cos x, \sin 2x$

14. $\cos^2 x, \sin^2 x, 2\pi$

15. $\cosh 2x, \sinh 2x, e^{2x}$

16. TEAM PROJECT. Linear Independence and Dependence. (a) Investigate the given question about a set S of functions on an interval I. Give an example. Prove your answer.

(1) If S contains the zero function, can S be linearly independent?

(2) If S is linearly independent on a subinterval J of I, is it linearly independent on I?

(3) If S is linearly dependent on a subinterval J of I, is it linearly dependent on I?

(4) If S is linearly independent on I, is it linearly independent on a subinterval J?

(5) If S is linearly dependent on I, is it linearly independent on a subinterval J?

(6) If S is linearly dependent on I, and if T contains S, is T linearly dependent on I?

(b) In what cases can you use the Wronskian for testing linear independence? By what other means can you perform such a test?

3.2 Homogeneous Linear ODEs with Constant Coefficients

We proceed along the lines of Sec. 2.2, and generalize the results from $n = 2$ to arbitrary n. We want to solve an nth-order homogeneous linear ODE with constant coefficients, written as

$$y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1y' + a_0y = 0 \tag{1}$$

where $y^{(n)} = d^n y/dx^n$, etc. As in Sec. 2.2, we substitute $y = e^{\lambda x}$ to obtain the characteristic equation

$$\lambda^{(n)} + a_{n-1}\lambda^{(n-1)} + \cdots + a_1\lambda + a_0 y = 0 \tag{2}$$

of (1). If λ is a root of (2), then $y = e^{\lambda x}$ is a solution of (1). To find these roots, you may need a numeric method, such as Newton's in Sec. 19.2, also available on the usual CASs. For general n there are more cases than for $n = 2$. We can have distinct real roots, simple complex roots, multiple roots, and multiple complex roots, respectively. This will be shown next and illustrated by examples.

Distinct Real Roots

If all the n roots $\lambda_1, \cdots, \lambda_n$ of (2) are real and different, then the n solutions

$$y_1 = e^{\lambda_1 x}, \quad \cdots, \quad y_n = e^{\lambda_n x}. \tag{3}$$

constitute a basis for all x. The corresponding general solution of (1) is

$$y = c_1 e^{\lambda_1 x} + \cdots + c_n e^{\lambda_n x}. \tag{4}$$

Indeed, the solutions in (3) are linearly independent, as we shall see after the example.

EXAMPLE 1 Distinct Real Roots

Solve the ODE $y''' - 2y'' - y' + 2y = 0$.

Solution. The characteristic equation is $\lambda^3 - 2\lambda^2 - \lambda + 2 = 0$. It has the roots $-1, 1, 2$; if you find one of them by inspection, you can obtain the other two roots by solving a quadratic equation (explain!). The corresponding general solution (4) is $y = c_1 e^{-x} + c_2 e^x + c_3 e^{2x}$. ■

Linear Independence of (3). Students familiar with nth-order determinants may verify that, by pulling out all exponential functions from the columns and denoting their product by $E = \exp[\lambda_1 + \cdots + \lambda_n)x]$, the Wronskian of the solutions in (3) becomes

$$\begin{aligned} W &= \begin{vmatrix} e^{\lambda_1 x} & e^{\lambda_2 x} & \cdots & e^{\lambda_n x} \\ \lambda_1 e^{\lambda_1 x} & \lambda_2 e^{\lambda_2 x} & \cdots & \lambda_n e^{\lambda_n x} \\ \lambda_1^2 e^{\lambda_1 x} & \lambda_2^2 e^{\lambda_2 x} & \cdots & \lambda_n^2 e^{\lambda_n x} \\ \cdot & \cdot & \cdots & \cdot \\ \lambda_1^{n-1} e^{\lambda_1 x} & \lambda_2^{n-1} e^{\lambda_2 x} & \cdots & \lambda_n^{n-1} e^{\lambda_n x} \end{vmatrix} \\ &= E \begin{vmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_n \\ \lambda_1^2 & \lambda_2^2 & \cdots & \lambda_n^2 \\ \cdot & \cdot & \cdots & \cdot \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \cdots & \lambda_n^{n-1} \end{vmatrix}. \end{aligned} \tag{5}$$

The exponential function E is never zero. Hence $W = 0$ if and only if the determinant on the right is zero. This is a so-called **Vandermonde** or **Cauchy determinant**.[1] It can be shown that it equals

$$(-1)^{n(n-1)/2}V \tag{6}$$

where V is the product of all factors $\lambda_j - \lambda_k$ with $j < k\ (\leqq n)$; for instance, when $n = 3$ we get $-V = -(\lambda_1 - \lambda_2)(\lambda_1 - \lambda_3)(\lambda_2 - \lambda_3)$. This shows that the Wronskian is not zero if and only if all the n roots of (2) are different and thus gives the following.

THEOREM 1

Basis

Solutions $y_1 = e^{\lambda_1 x}, \cdots, y_n = e^{\lambda_n x}$ of (1) *(with any real or complex λ_j's) form a basis of solutions of* (1) *on any open interval if and only if all n roots of* (2) *are different.*

Actually, Theorem 1 is an important special case of our more general result obtained from (5) and (6):

THEOREM 2

Linear Independence

Any number of solutions of (1) *of the form $e^{\lambda x}$ are linearly independent on an open interval I if and only if the corresponding λ are all different.*

Simple Complex Roots

If complex roots occur, they must occur in conjugate pairs since the coefficients of (1) are real. Thus, if $\lambda = \gamma + i\omega$ is a simple root of (2), so is the conjugate $\bar{\lambda} = \gamma - i\omega$, and two corresponding linearly independent solutions are (as in Sec. 2.2, except for notation)

$$y_1 = e^{\gamma x} \cos \omega x, \qquad y_2 = e^{\gamma x} \sin \omega x.$$

EXAMPLE 2 **Simple Complex Roots. Initial Value Problem**

Solve the initial value problem

$$y''' - y'' + 100y' - 100y = 0, \qquad y(0) = 4, \qquad y'(0) = 11, \qquad y''(0) = -299.$$

Solution. The characteristic equation is $\lambda^3 - \lambda^2 + 100\lambda - 100 = 0$. It has the root 1, as can perhaps be seen by inspection. Then division by $\lambda - 1$ shows that the other roots are $\pm 10i$. Hence a general solution and its derivatives (obtained by differentiation) are

$$\begin{aligned} y &= c_1 e^x + A \cos 10x + B \sin 10x, \\ y' &= c_1 e^x - 10A \sin 10x + 10B \cos 10x, \\ y'' &= c_1 e^x - 100A \cos 10x - 100B \sin 10x. \end{aligned}$$

[1]ALEXANDRE THÉOPHILE VANDERMONDE (1735–1796), French mathematician, who worked on solution of equations by determinants. For CAUCHY see footnote 4, in Sec. 2.5.

From this and the initial conditions we obtain, by setting $x = 0$,

$$\text{(a)}\ c_1 + A = 4, \qquad \text{(b)}\ c_1 + 10B = 11, \qquad \text{(c)}\ c_1 - 100A = -299.$$

We solve this system for the unknowns A, B, c_1. Equation (a) minus Equation (c) gives $101A = 303$, $A = 3$. Then $c_1 = 1$ from (a) and $B = 1$ from (b). The solution is (Fig. 73)

$$y = e^x + 3 \cos 10x + \sin 10x.$$

This gives the solution curve, which oscillates about e^x (dashed in Fig. 73). ■

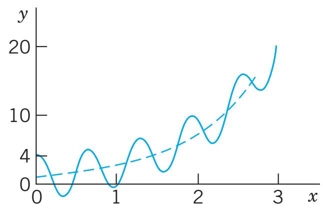


Fig. 73. Solution in Example 2

Multiple Real Roots

If a real double root occurs, say, $\lambda_1 = \lambda_2$, then $y_1 = y_2$ in (3), and we take y_1 and xy_1 as corresponding linearly independent solutions. This is as in Sec. 2.2.

More generally, if λ is a real root of order m, then m corresponding linearly independent solutions are

$$e^{\lambda x}, \quad xe^{\lambda x}, \quad x^2e^{\lambda x}, \quad \cdots, \quad x^{m-1}e^{\lambda x}. \tag{7}$$

We derive these solutions after the next example and indicate how to prove their linear independence.

EXAMPLE 3 Real Double and Triple Roots

Solve the ODE $y^v - 3y^{iv} + 3y''' - y'' = 0$.

Solution. The characteristic equation $\lambda^5 - 3\lambda^4 + 3\lambda^3 - \lambda^2 = 0$ has the roots $\lambda_1 = \lambda_2 = 0$, and $\lambda_3 = \lambda_4 = \lambda_5 = 1$, and the answer is

$$y = c_1 + c_2x + (c_3 + c_4x + c_5x^2)e^x. \tag{8}$$ ■

Derivation of **(7).** We write the left side of (1) as

$$L[y] = y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_0y.$$

Let $y = e^{\lambda x}$. Then by performing the differentiations we have

$$L[e^{\lambda x}] = (\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_0)e^{\lambda x}.$$

Now let λ_1 be a root of mth order of the polynomial on the right, where $m \leqq n$. For $m < n$ let $\lambda_{m+1}, \cdots, \lambda_n$ be the other roots, all different from λ_1. Writing the polynomial in product form, we then have

$$L[e^{\lambda x}] = (\lambda - \lambda_1)^m h(\lambda) e^{\lambda x}$$

with $h(\lambda) = 1$ if $m = n$, and $h(\lambda) = (\lambda - \lambda_{m+1}) \cdots (\lambda - \lambda_n)$ if $m < n$. Now comes the key idea: We differentiate on both sides with respect to λ,

$$\frac{\partial}{\partial \lambda} L[e^{\lambda x}] = m(\lambda - \lambda_1)^{m-1} h(\lambda) e^{\lambda x} + (\lambda - \lambda_1)^m \frac{\partial}{\partial \lambda} [h(\lambda) e^{\lambda x}]. \tag{9}$$

The differentiations with respect to x and λ are independent and the resulting derivatives are continuous, so that we can interchange their order on the left:

$$\frac{\partial}{\partial \lambda} L[e^{\lambda x}] = L\left[\frac{\partial}{\partial \lambda} e^{\lambda x}\right] = L[xe^{\lambda x}]. \tag{10}$$

The right side of (9) is zero for $\lambda = \lambda_1$ because of the factors $\lambda - \lambda_1$ (and $m \geqq 2$ since we have a multiple root!). Hence $L[xe^{\lambda_1 x}] = 0$ by (9) and (10). This proves that $xe^{\lambda_1 x}$ is a solution of (1).

We can repeat this step and produce $x^2 e^{\lambda_1 x}, \cdots, x^{m-1} e^{\lambda_1 x}$ by another $m - 2$ such differentiations with respect to λ. Going one step further would no longer give zero on the right because the lowest power of $\lambda - \lambda_1$ would then be $(\lambda - \lambda_1)^0$, multiplied by $m!h(\lambda)$ and $h(\lambda_1) \neq 0$ because $h(\lambda)$ has no factors $\lambda - \lambda_1$; so we get *precisely* the solutions in (7).

We finally show that the solutions (7) are linearly independent. For a specific n this can be seen by calculating their Wronskian, which turns out to be nonzero. For arbitrary m we can pull out the exponential functions from the Wronskian. This gives $(e^{\lambda x})^m = e^{\lambda m x}$ times a determinant which by "row operations" can be reduced to the Wronskian of 1, $x, \cdots, x^{m-1}$. The latter is constant and different from zero (equal to $1!2! \cdots (m-1)!$). These functions are solutions of the ODE $y^{(m)} = 0$, so that linear independence follows from Theroem 3 in Sec. 3.1.

Multiple Complex Roots

In this case, real solutions are obtained as for complex simple roots above. Consequently, if $\lambda = \gamma + i\omega$ is a **complex double root**, so is the conjugate $\bar{\lambda} = \gamma - i\omega$. Corresponding linearly independent solutions are

$$e^{\gamma x} \cos \omega x, \quad e^{\gamma x} \sin \omega x, \quad xe^{\gamma x} \cos \omega x, \quad xe^{\gamma x} \sin \omega x. \tag{11}$$

The first two of these result from $e^{\lambda x}$ and $e^{\bar{\lambda} x}$ as before, and the second two from $xe^{\lambda x}$ and $xe^{\bar{\lambda} x}$ in the same fashion. Obviously, the corresponding general solution is

$$y = e^{\gamma x}[(A_1 + A_2 x) \cos \omega x + (B_1 + B_2 x) \sin \omega x]. \tag{12}$$

For *complex triple roots* (which hardly ever occur in applications), one would obtain two more solutions $x^2 e^{\gamma x} \cos \omega x$, $x^2 e^{\gamma x} \sin \omega x$, and so on.

PROBLEM SET 3.2

1–6 GENERAL SOLUTION

Solve the given ODE. Show the details of your work.

1. $y''' + 25y' = 0$
2. $y^{\text{iv}} + 2y'' + y = 0$
3. $y^{\text{iv}} + 4y'' = 0$
4. $(D^3 - D^2 - D + I)y = 0$
5. $(D^4 + 10D^2 + 9I)y = 0$
6. $(D^5 + 8D^3 + 16D)y = 0$

7–13 INITIAL VALUE PROBLEM

Solve the IVP by a CAS, giving a general solution and the particular solution and its graph.

7. $y''' + 3.2y'' + 4.81y' = 0$, $y(0) = 3.4$, $y'(0) = -4.6$, $y''(0) = 9.91$
8. $y''' + 7.5y'' + 14.25y' - 9.125y = 0$, $y(0) = 10.05$, $y'(0) = -54.975$, $y''(0) = 257.5125$
9. $4y''' + 8y'' + 41y' + 37y = 0$, $y(0) = 9$, $y'(0) = -6.5$, $y''(0) = -39.75$
10. $y^{\text{iv}} + 4y = 0$, $y(0) = \frac{1}{2}$, $y'(0) = -\frac{3}{2}$, $y''(0) = \frac{5}{2}$, $y'''(0) = -\frac{7}{2}$
11. $y^{\text{iv}} - 9y'' - 400y = 0$, $y(0) = 0$, $y'(0) = 0$, $y''(0) = 41$, $y'''(0) = 0$
12. $y^{\text{v}} - 5y''' + 4y' = 0$, $y(0) = 3$, $y'(0) = -5$, $y''(0) = 11$, $y'''(0) = -23$, $y^{\text{iv}}(0) = 47$
13. $y^{\text{iv}} + 0.45y''' - 0.165y'' + 0.0045y' - 0.00175y = 0$, $y(0) = 17.4$, $y'(0) = -2.82$, $y''(0) = 2.0485$, $y'''(0) = -1.458675$
14. **PROJECT. Reduction of Order.** This is of practical interest since a single solution of an ODE can often be guessed. For second order, see Example 7 in Sec. 2.1.

    (a) How could you reduce the order of a linear constant-coefficient ODE if a solution is known?

    (b) Extend the method to a variable-coefficient ODE

    $$y''' + p_2(x)y'' + p_1(x)y' + p_0(x)y = 0.$$

    Assuming a solution y_1 to be known, show that another solution is $y_2(x) = u(x)y_1(x)$ with $u(x) = \int z(x)\,dx$ and z obtained by solving

    $$y_1 z'' + (3y_1' + p_2 y_1)z' + (3y_1'' + 2p_2 y_1' + p_1 y_1)z = 0.$$

    (c) Reduce

    $$x^3 y''' - 3x^2 y'' + (6 - x^2)xy' - (6 - x^2)y = 0,$$

    using $y_1 = x$ (perhaps obtainable by inspection).
15. **CAS EXPERIMENT. Reduction of Order.** Starting with a basis, find third-order linear ODEs with variable coefficients for which the reduction to second order turns out to be relatively simple.

3.3 Nonhomogeneous Linear ODEs

We now turn from homogeneous to nonhomogeneous linear ODEs of nth order. We write them in standard form

$$y^{(n)} + p_{n-1}(x)y^{(n-1)} + \cdots + p_1(x)y' + p_0(x)y = r(x) \tag{1}$$

with $y^{(n)} = d^n y/dx^n$ as the first term, and $r(x) \not\equiv 0$. As for second-order ODEs, a general solution of (1) on an open interval I of the x-axis is of the form

$$y(x) = y_h(x) + y_p(x). \tag{2}$$

Here $y_h(x) = c_1 y_1(x) + \cdots + c_n y_n(x)$ is a general solution of the corresponding homogeneous ODE

$$y^{(n)} + p_{n-1}(x)y^{(n-1)} + \cdots + p_1(x)y' + p_0(x)y = 0 \tag{3}$$

on I. Also, y_p is any solution of (1) on I containing no arbitrary constants. If (1) has continuous coefficients and a continuous $r(x)$ on I, then a general solution of (1) exists and includes all solutions. Thus (1) has no singular solutions.

An **initial value problem** for (1) consists of (1) and n **initial conditions**

$$y(x_0) = K_0, \qquad y'(x_0) = K_1, \qquad \cdots, \qquad y^{(n-1)}(x_0) = K_{n-1} \tag{4}$$

with x_0 in I. Under those continuity assumptions it has a unique solution. The ideas of proof are the same as those for $n = 2$ in Sec. 2.7.

Method of Undetermined Coefficients

Equation (2) shows that for solving (1) we have to determine a particular solution of (1). For a constant-coefficient equation

$$y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1y' + a_0y = r(x) \tag{5}$$

($a_0, \cdots, a_{n-1}$ constant) and special $r(x)$ as in Sec. 2.7, such a $y_p(x)$ can be determined by the **method of undetermined coefficients,** as in Sec. 2.7, using the following rules.

(A) Basic Rule as in Sec. 2.7.

(B) Modification Rule. *If a term in your choice for* $y_p(x)$ *is a solution of the homogeneous equation* (3), *then multiply this term by* x^k, *where* k *is the smallest positive integer such that this term times* x^k *is not a solution of* (3).

(C) Sum Rule as in Sec. 2.7.

The practical application of the method is the same as that in Sec. 2.7. It suffices to illustrate the typical steps of solving an initial value problem and, in particular, the new Modification Rule, which includes the old Modification Rule as a particular case (with $k = 1$ or 2). We shall see that the technicalities are the same as for $n = 2$, except perhaps for the more involved determination of the constants.

EXAMPLE 1 Initial Value Problem. Modification Rule

Solve the initial value problem

$$y''' + 3y'' + 3y' + y = 30e^{-x}, \qquad y(0) = 3, \qquad y'(0) = -3, \qquad y''(0) = -47. \tag{6}$$

Solution. ***Step 1.*** The characteristic equation is $\lambda^3 + 3\lambda^2 + 3\lambda + 1 = (\lambda + 1)^3 = 0$. It has the triple root $\lambda = -1$. Hence a general solution of the homogeneous ODE is

$$\begin{aligned} y_h &= c_1e^{-x} + c_2xe^{-x} + c_3x^2e^{-x} \\ &= (c_1 + c_2x + c_3x^2)e^{-x}. \end{aligned}$$

Step 2. If we try $y_p = Ce^{-x}$, we get $-C + 3C - 3C + C = 30$, which has no solution. Try Cxe^{-x} and Cx^2e^{-x}. The Modification Rule calls for

$$y_p = Cx^3e^{-x}.$$

Then

$$\begin{aligned} y_p' &= C(3x^2 - x^3)e^{-x}, \\ y_p'' &= C(6x - 6x^2 + x^3)e^{-x}, \\ y_p''' &= C(6 - 18x + 9x^2 - x^3)e^{-x}. \end{aligned}$$

Substitution of these expressions into (6) and omission of the common factor e^{-x} gives

$$C(6 - 18x + 9x^2 - x^3) + 3C(6x - 6x^2 + x^3) + 3C(3x^2 - x^3) + Cx^3 = 30.$$

The linear, quadratic, and cubic terms drop out, and $6C = 30$. Hence $C = 5$. This gives $y_p = 5x^3e^{-x}$.

Step 3. We now write down $y = y_h + y_p$, the general solution of the given ODE. From it we find c_1 by the first initial condition. We insert the value, differentiate, and determine c_2 from the second initial condition, insert the value, and finally determine c_3 from $y''(0)$ and the third initial condition:

$$\begin{aligned}
y &= y_h + y_p = (c_1 + c_2x + c_3x^2)e^{-x} + 5x^3e^{-x}, & y(0) &= c_1 = 3 & &\\
y' &= [-3 + c_2 + (-c_2 + 2c_3)x + (15 - c_3)x^2 - 5x^3]e^{-x}, & y'(0) &= -3 + c_2 = -3, & c_2 &= 0\\
y'' &= [3 + 2c_3 + (30 - 4c_3)x + (-30 + c_3)x^2 + 5x^3]e^{-x}, & y''(0) &= 3 + 2c_3 = -47, & c_3 &= -25.
\end{aligned}$$

Hence the *answer* to our problem is (Fig. 73)

$$y = (3 - 25x^2)e^{-x} + 5x^3e^{-x}.$$

The curve of y begins at (0, 3) with a negative slope, as expected from the initial values, and approaches zero as $x \to \infty$. The dashed curve in Fig. 74 is y_p. ■

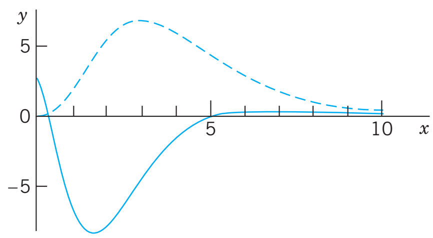


Fig. 74. y and y_p (dashed) in Example 1

Method of Variation of Parameters

The method of variation of parameters (see Sec. 2.10) also extends to arbitrary order n. It gives a particular solution y_p for the nonhomogeneous equation (1) (in standard form with $y^{(n)}$ as the first term!) by the formula

$$\begin{aligned}
y_p(x) &= \sum_{k=1}^{n} y_k(x) \int \frac{W_k(x)}{W(x)}\, r(x)\, dx \\
&= y_1(x) \int \frac{W_1(x)}{W(x)}\, r(x)\, dx + \cdots + y_n(x) \int \frac{W_n(x)}{W(x)}\, r(x)\, dx
\end{aligned} \tag{7}$$

on an open interval I on which the coefficients of (1) and $r(x)$ are continuous. In (7) the functions $y_1, \cdots, y_n$ form a basis of the homogeneous ODE (3), with Wronskian W, and W_j ($j = 1, \cdots, n$) is obtained from W by replacing the jth column of W by the column $[0 \quad 0 \quad \cdots \quad 0 \quad 1]^\mathsf{T}$. Thus, when $n = 2$, this becomes identical with (2) in Sec. 2.10,

$$W = \begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix}, \qquad W_1 = \begin{vmatrix} 0 & y_2 \\ 1 & y_2' \end{vmatrix} = -y_2, \qquad W_2 = \begin{vmatrix} y_1 & 0 \\ y_1' & 1 \end{vmatrix} = y_1.$$

The proof of (7) uses an extension of the idea of the proof of (2) in Sec. 2.10 and can be found in Ref [A11] listed in App. 1.

EXAMPLE 2 **Variation of Parameters. Nonhomogeneous Euler–Cauchy Equation**

Solve the nonhomogeneous Euler–Cauchy equation

$$x^3y''' - 3x^2y'' + 6xy' - 6y = x^4 \ln x \qquad (x > 0).$$

Solution. ***Step 1. General solution of the homogeneous ODE.*** Substitution of $y = x^m$ and the derivatives into the homogeneous ODE and deletion of the factor x^m gives

$$m(m-1)(m-2) - 3m(m-1) + 6m - 6 = 0.$$

The roots are 1, 2, 3 and give as a basis

$$y_1 = x, \qquad y_2 = x^2, \qquad y_3 = x^3.$$

Hence the corresponding general solution of the homogeneous ODE is

$$y_h = c_1x + c_2x^2 + c_3x^3.$$

***Step 2. Determinants needed in* (7).** These are

$$W = \begin{vmatrix} x & x^2 & x^3 \\ 1 & 2x & 3x^2 \\ 0 & 2 & 6x \end{vmatrix} = 2x^3$$

$$W_1 = \begin{vmatrix} 0 & x^2 & x^3 \\ 0 & 2x & 3x^2 \\ 1 & 2 & 6x \end{vmatrix} = x^4$$

$$W_2 = \begin{vmatrix} x & 0 & x^3 \\ 1 & 0 & 3x^2 \\ 0 & 1 & 6x \end{vmatrix} = -2x^3$$

$$W_3 = \begin{vmatrix} x & x^2 & 0 \\ 1 & 2x & 0 \\ 0 & 2 & 1 \end{vmatrix} = x^2.$$

Step 3. Integration. In (7) we also need the right side $r(x)$ of our ODE in standard form, obtained by division of the given equation by the coefficient x^3 of y'''; thus, $r(x) = (x^4 \ln x)/x^3 = x \ln x$. In (7) we have the simple quotients $W_1/W = x/2$, $W_2/W = -1$, $W_3/W = 1/(2x)$. Hence (7) becomes

$$\begin{aligned} y_p &= x \int \frac{x}{2} x \ln x \, dx - x^2 \int x \ln x \, dx + x^3 \int \frac{1}{2x} x \ln x \, dx \\ &= \frac{x}{2}\left(\frac{x^3}{3} \ln x - \frac{x^3}{9}\right) - x^2\left(\frac{x^2}{2} \ln x - \frac{x^2}{4}\right) + \frac{x^3}{2}(x \ln x - x). \end{aligned}$$

Simplification gives $y_p = \frac{1}{6}x^4 (\ln x - \frac{11}{6})$. Hence the answer is

$$y = y_h + y_p = c_1x + c_2x^2 + c_3x^3 + \tfrac{1}{6}x^4 (\ln x - \tfrac{11}{6}).$$

Figure 75 shows y_p. Can you explain the shape of this curve? Its behavior near $x = 0$? The occurrence of a minimum? Its rapid increase? Why would the method of undetermined coefficients not have given the solution? ■

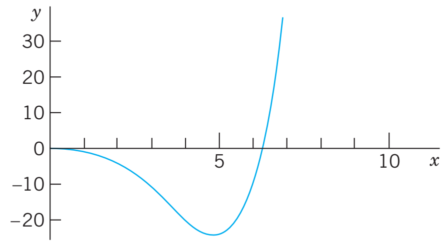


Fig. 75. Particular solution y_p of the nonhomogeneous Euler–Cauchy equation in Example 2

Application: Elastic Beams

Whereas second-order ODEs have various applications, of which we have discussed some of the more important ones, higher order ODEs have much fewer engineering applications. An important fourth-order ODE governs the bending of elastic beams, such as wooden or iron girders in a building or a bridge.

A related application of vibration of beams does not fit in here since it leads to PDEs and will therefore be discussed in Sec. 12.3.

EXAMPLE 3 Bending of an Elastic Beam under a Load

We consider a beam B of length L and constant (e.g., **rectangular**) cross section and homogeneous elastic material (e.g., steel); see Fig. 76. We assume that under its own weight the beam is bent so little that it is practically straight. If we apply a load to B in a vertical plane through the axis of symmetry (the x-axis in Fig. 76), B is bent. Its axis is curved into the so-called **elastic curve** C (or **deflection curve**). It is shown in elasticity theory that the bending moment $M(x)$ is proportional to the curvature $k(x)$ of C. We assume the bending to be small, so that the deflection $y(x)$ and its derivative $y'(x)$ (determining the tangent direction of C) are small. Then, by calculus, $k = y''/(1 + y'^2)^{3/2} \approx y''$. Hence

$$M(x) = EIy''(x).$$

EI is the constant of proportionality. E is *Young's modulus of elasticity* of the material of the beam. I is the moment of inertia of the cross section about the (horizontal) z-axis in Fig. 76.

Elasticity theory shows further that $M''(x) = f(x)$, where $f(x)$ is the load per unit length. Together,

$$EIy^{\text{iv}} = f(x). \tag{8}$$

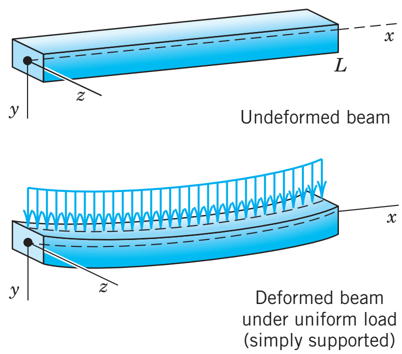


Fig. 76. Elastic beam

In applications the most important supports and corresponding boundary conditions are as follows and shown in Fig. 77.

(A) Simply supported $y = y'' = 0$ at $x = 0$ and L

(B) Clamped at both ends $y = y' = 0$ at $x = 0$ and L

(C) Clamped at $x = 0$, free at $x = L$ $y(0) = y'(0) = 0,\ y''(L) = y'''(L) = 0.$

The boundary condition $y = 0$ means no displacement at that point, $y' = 0$ means a horizontal tangent, $y'' = 0$ means no bending moment, and $y''' = 0$ means no shear force.

Let us apply this to the uniformly loaded simply supported beam in Fig. 76. The load is $f(x) \equiv f_0 = \text{const}$. Then (8) is

$$y^{\text{iv}} = k, \qquad k = \frac{f_0}{EI}. \tag{9}$$

This can be solved simply by calculus. Two integrations give

$$y'' = \frac{k}{2}x^2 + c_1x + c_2.$$

$y''(0) = 0$ gives $c_2 = 0$. Then $y''(L) = L(\frac{1}{2}kL + c_1) = 0$, $c_1 = -kL/2$ (since $L \neq 0$). Hence

$$y'' = \frac{k}{2}(x^2 - Lx).$$

Integrating this twice, we obtain

$$y = \frac{k}{2}\left(\frac{1}{12}x^4 - \frac{L}{6}x^3 + c_3x + c_4\right)$$

with $c_4 = 0$ from $y(0) = 0$. Then

$$y(L) = \frac{kL}{2}\left(\frac{L^3}{12} - \frac{L^3}{6} + c_3\right) = 0, \qquad c_3 = \frac{L^3}{12}.$$

Inserting the expression for k, we obtain as our solution

$$y = \frac{f_0}{24EI}(x^4 - 2Lx^3 + L^3x).$$

Since the boundary conditions at both ends are the same, we expect the deflection $y(x)$ to be "symmetric" with respect to $L/2$, that is, $y(x) = y(L - x)$. Verify this directly or set $x = u + L/2$ and show that y becomes an even function of u,

$$y = \frac{f_0}{24EI}\left(u^2 - \frac{1}{4}L^2\right)\left(u^2 - \frac{5}{4}L^2\right).$$

From this we can see that the maximum deflection in the middle at $u = 0\,(x = L/2)$ is $5f_0L^4/(16 \cdot 24EI)$. Recall that the positive direction points downward. ■

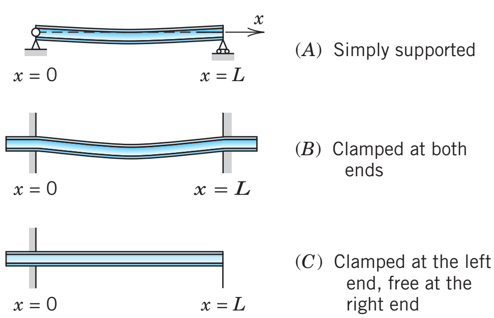


Fig. 77. Supports of a beam

PROBLEM SET 3.3

1–7 GENERAL SOLUTION

Solve the following ODEs, showing the details of your work.

1. $y''' + 3y'' + 3y' + y = e^x - x - 1$
2. $y''' + 2y'' - y' - 2y = 1 - 4x^3$
3. $(D^4 + 10D^2 + 9I)y = 6.5 \sinh 2x$
4. $(D^3 + 3D^2 - 5D - 39I)y = -300 \cos x$
5. $(x^3D^3 + x^2D^2 - 2xD + 2I)y = x^{-2}$
6. $(D^3 + 4D)y = \sin x$
7. $(D^3 - 9D^2 + 27D - 27I)y = 27 \sin 3x$

8–13 INITIAL VALUE PROBLEM

Solve the given IVP, showing the details of your work.

8. $y^{\text{iv}} - 5y'' + 4y = 10e^{-3x}$, $y(0) = 1$, $y'(0) = 0$, $y''(0) = 0$, $y'''(0) = 0$
9. $y^{\text{iv}} + 5y'' + 4y = 90 \sin 4x$, $y(0) = 1$, $y'(0) = 2$, $y''(0) = -1$, $y'''(0) = -32$
10. $x^3y''' + xy' - y = x^2$, $y(1) = 1$, $y'(1) = 3$, $y''(1) = 14$
11. $(D^3 - 2D^2 - 3D)y = 74e^{-3x} \sin x$, $y(0) = -1.4$, $y'(0) = 3.2$, $y''(0) = -5.2$
12. $(D^3 - 2D^2 - 9D + 18I)y = e^{2x}$, $y(0) = 4.5$, $y'(0) = 8.8$, $y''(0) = 17.2$
13. $(D^3 - 4D)y = 10 \cos x + 5 \sin x$, $y(0) = 3$, $y'(0) = -2$, $y''(0) = -1$

14. **CAS EXPERIMENT. Undetermined Coefficients.** Since variation of parameters is generally complicated, it seems worthwhile to try to extend the other method. Find out experimentally for what ODEs this is possible and for what not. *Hint:* Work backward, solving ODEs with a CAS and then looking whether the solution could be obtained by undetermined coefficients. For example, consider

$$y''' - 3y'' + 3y' - y = x^{1/2}e^x$$

and

$$x^3y''' + x^2y'' - 2xy' + 2y = x^3 \ln x.$$

15. **WRITING REPORT. Comparison of Methods.** Write a report on the method of undetermined coefficients and the method of variation of parameters, discussing and comparing the advantages and disadvantages of each method. Illustrate your findings with typical examples. Try to show that the method of undetermined coefficients, say, for a third-order ODE with constant coefficients and an exponential function on the right, can be derived from the method of variation of parameters.

CHAPTER 3 REVIEW QUESTIONS AND PROBLEMS

1. What is the superposition or linearity principle? For what nth-order ODEs does it hold?
2. List some other basic theorems that extend from second-order to nth-order ODEs.
3. If you know a general solution of a homogeneous linear ODE, what do you need to obtain from it a general solution of a corresponding nonhomogeneous linear ODE?
4. What form does an initial value problem for an nth-order linear ODE have?
5. What is the Wronskian? What is it used for?

6–15 GENERAL SOLUTION

Solve the given ODE. Show the details of your work.

6. $y^{\text{iv}} - 3y'' - 4y = 0$
7. $y''' + 4y'' + 13y' = 0$
8. $y''' - 4y'' - y' + 4y = 30e^{2x}$
9. $(D^4 - 16I)y = -15 \cosh x$
10. $x^2y''' + 3xy'' - 2y' = 0$
11. $y''' + 4.5y'' + 6.75y' + 3.375y = 0$
12. $(D^3 - D)y = \sinh 0.8x$
13. $(D^3 + 6D^2 + 12D + 8I)y = 8x^2$
14. $(D^4 - 13D^2 + 36I)y = 12e^x$
15. $4x^3y''' + 3xy' - 3y = 10$

16–20 INITIAL VALUE PROBLEM

Solve the IVP. Show the details of your work.

16. $(D^3 - D^2 - D + I)y = 0$, $y(0) = 0$, $Dy(0) = 1$, $D^2y(0) = 0$
17. $y''' + 5y'' + 24y' + 20y = x$, $y(0) = 1.94$, $y'(0) = -3.95$, $y'' = -24$
18. $(D^4 - 26D^2 + 25I)y = 50(x + 1)^2$, $y(0) = 12.16$, $Dy(0) = -6$, $D^2y(0) = 34$, $D^3y(0) = -130$
19. $(D^3 + 9D^2 + 23D + 15I)y = 12\exp(-4x)$, $y(0) = 9$, $Dy(0) = -41$, $D^2y(0) = 189$
20. $(D^3 + 3D^2 + 3D + I)y = 8 \sin x$, $y(0) = -1$, $y'(0) = -3$, $y''(0) = 5$

SUMMARY OF CHAPTER 3
Higher Order Linear ODEs

Compare with the similar Summary of Chap. 2 (the case n = 2).
Chapter 3 extends Chap. 2 from order $n = 2$ to arbitrary order n. An ***n*th-order linear ODE** is an ODE that can be written

$$(1) \qquad y^{(n)} + p_{n-1}(x)y^{(n-1)} + \cdots + p_1(x)y' + p_0(x)y = r(x)$$

with $y^{(n)} = d^n y/dx^n$ as the first term; we again call this the **standard form**. Equation (1) is called **homogeneous** if $r(x) \equiv 0$ on a given open interval I considered, **nonhomogeneous** if $r(x) \not\equiv 0$ on I. For the homogeneous ODE

$$(2) \qquad y^{(n)} + p_{n-1}(x)y^{(n-1)} + \cdots + p_1(x)y' + p_0(x)y = 0$$

the ***superposition principle*** (Sec. 3.1) holds, just as in the case $n = 2$. A **basis** or **fundamental system** of solutions of (2) on I consists of n linearly independent solutions $y_1, \cdots, y_n$ of (2) on I. A **general solution** of (2) on I is a linear combination of these,

$$(3) \qquad y = c_1 y_1 + \cdots + c_n y_n \qquad (c_1, \cdots, c_n \text{ arbitrary constants}).$$

A **general solution** of the nonhomogeneous ODE (1) on I is of the form

$$(4) \qquad y = y_h + y_p \qquad \text{(Sec. 3.3)}.$$

Here, y_p is a particular solution of (1) and is obtained by two methods (***undetermined coefficients*** or ***variation of parameters***) explained in Sec. 3.3.

An **initial value problem** for (1) or (2) consists of one of these ODEs and n initial conditions (Secs. 3.1, 3.3)

$$(5) \qquad y(x_0) = K_0, \qquad y'(x_0) = K_1, \qquad \cdots, \qquad y^{(n-1)}(x_0) = K_{n-1}$$

with given x_0 in I and given $K_0, \cdots, K_{n-1}$. If $p_0, \cdots, p_{n-1}, r$ are continuous on I, then general solutions of (1) and (2) on I exist, and initial value problems (1), (5) or (2), (5) have a unique solution.

CHAPTER 4

Systems of ODEs. Phase Plane. Qualitative Methods

Tying in with Chap. 3, we present another method of solving higher order ODEs in Sec. 4.1. This converts any nth-order ODE into a system of n first-order ODEs. We also show some applications. Moreover, in the same section we solve systems of first-order ODEs that occur directly in applications, that is, not derived from an nth-order ODE but dictated by the application such as two tanks in mixing problems and two circuits in electrical networks. (The elementary aspects of vectors and matrices needed in this chapter are reviewed in Sec. 4.0 and are probably familiar to most students.)

In Sec. 4.3 we introduce a totally different way of looking at systems of ODEs. The method consists of examining the general behavior of whole families of solutions of ODEs in the ***phase plane***, and aptly is called the phase plane method. It gives information on the **stability** of solutions. (*Stability of a physical system* is desirable and means roughly that a small change at some instant causes only a small change in the behavior of the system at later times.) This approach to systems of ODEs is a **qualitative method** because it depends only on the nature of the ODEs and does not require the actual solutions. This can be very useful because it is often difficult or even impossible to solve systems of ODEs. In contrast, the approach of actually solving a system is known as a *quantitative method*.

The phase plane method has many applications in control theory, circuit theory, population dynamics and so on. Its use in linear systems is discussed in Secs. 4.3, 4.4, and 4.6 and its even more important use in nonlinear systems is discussed in Sec. 4.5 with applications to the pendulum equation and the Lokta–Volterra population model. The chapter closes with a discussion of nonhomogeneous linear systems of ODEs.

NOTATION. We continue to denote unknown functions by y; thus, $y_1(t), y_2(t)$—analogous to Chaps. 1–3. (Note that some authors use x for functions, $x_1(t), x_2(t)$ when dealing with systems of ODEs.)

Prerequisite: Chap. 2.
References and Answers to Problems: App. 1 Part A, and App. 2.

4.0 For Reference: Basics of Matrices and Vectors

For clarity and simplicity of notation, we use matrices and vectors in our discussion of linear systems of ODEs. We need only a few elementary facts (and not the bulk of the material of Chaps. 7 and 8). Most students will very likely be already familiar with these facts. Thus ***this section is for reference only***. *Begin with Sec. 4.1 and consult 4.0 as needed.*

Most of our linear systems will consist of two linear ODEs in two unknown functions $y_1(t)$, $y_2(t)$,

$$\begin{aligned} y_1' &= a_{11}y_1 + a_{12}y_2, \\ y_2' &= a_{21}y_1 + a_{22}y_2, \end{aligned} \qquad \text{for example,} \qquad \begin{aligned} y_1' &= -5y_1 + 2y_2 \\ y_2' &= \;\; 13y_1 + \tfrac{1}{2}y_2 \end{aligned} \tag{1}$$

(perhaps with additional *given* functions $g_1(t)$, $g_2(t)$ on the right in the two ODEs).

Similarly, a linear system of n first-order ODEs in n unknown functions $y_1(t), \cdots, y_n(t)$ is of the form

$$\begin{aligned} y_1' &= a_{11}y_1 + a_{12}y_2 + \cdots + a_{1n}y_n \\ y_2' &= a_{21}y_1 + a_{22}y_2 + \cdots + a_{2n}y_n \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ y_n' &= a_{n1}y_1 + a_{n2}y_2 + \cdots + a_{nn}y_n \end{aligned} \tag{2}$$

(perhaps with an additional given function on the right in each ODE).

Some Definitions and Terms

Matrices. In (1) the (constant or variable) coefficients form a **2 × 2 matrix A**, that is, an array

$$\mathbf{A} = [a_{jk}] = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \qquad \text{for example,} \qquad \mathbf{A} = \begin{bmatrix} -5 & 2 \\ 13 & \frac{1}{2} \end{bmatrix}. \tag{3}$$

Similarly, the coefficients in (2) form an ***n* × *n* matrix**

$$\mathbf{A} = [a_{jk}] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}. \tag{4}$$

The $a_{11}, a_{12}, \cdots$ are called **entries**, the horizontal lines **rows**, and the vertical lines **columns**. Thus, in (3) the first row is $[a_{11} \quad a_{12}]$, the second row is $[a_{21} \quad a_{22}]$, and the first and second columns are

$$\begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix}.$$

In the "***double subscript notation***" for entries, the first subscript denotes the *row* and the second the *column* in which the entry stands. Similarly in (4). The **main diagonal** is the diagonal $a_{11} \quad a_{22} \quad \cdots \quad a_{nn}$ in (4), hence $a_{11} \quad a_{22}$ in (3).

We shall need only **square matrices**, that is, matrices with the same number of rows and columns, as in (3) and (4).

Vectors. A **column vector x** with n **components** $x_1, \cdots, x_n$ is of the form

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \qquad \text{thus if } n = 2, \qquad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Similarly, a **row vector v** is of the form

$$\mathbf{v} = [v_1 \quad \cdots \quad v_n], \qquad \text{thus if } n = 2, \text{ then} \qquad \mathbf{v} = [v_1 \quad v_2].$$

Calculations with Matrices and Vectors

Equality. Two $n \times n$ matrices are *equal* if and only if corresponding entries are equal. Thus for $n = 2$, let

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \qquad \text{and} \qquad \mathbf{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}.$$

Then $\mathbf{A} = \mathbf{B}$ if and only if

$$\begin{aligned} a_{11} &= b_{11}, & a_{12} &= b_{12} \\ a_{21} &= b_{21}, & a_{22} &= b_{22}. \end{aligned}$$

Two column vectors (or two row vectors) are *equal* if and only if they both have n components and corresponding components are equal. Thus, let

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \quad \text{and} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \qquad \text{Then} \qquad \mathbf{v} = \mathbf{x} \quad \text{if and only if} \quad \begin{aligned} v_1 &= x_1 \\ v_2 &= x_2. \end{aligned}$$

Addition is performed by adding corresponding entries (or components); here, matrices must both be $n \times n$, and vectors must both have the same number of components. Thus for $n = 2$,

$$\text{(5)} \qquad \mathbf{A} + \mathbf{B} = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \end{bmatrix}, \qquad \mathbf{v} + \mathbf{x} = \begin{bmatrix} v_1 + x_1 \\ v_2 + x_2 \end{bmatrix}.$$

Scalar multiplication (multiplication by a number c) is performed by multiplying each entry (or component) by c. For example, if

$$\mathbf{A} = \begin{bmatrix} 9 & 3 \\ -2 & 0 \end{bmatrix}, \qquad \text{then} \qquad -7\mathbf{A} = \begin{bmatrix} -63 & -21 \\ 14 & 0 \end{bmatrix}.$$

If

$$\mathbf{v} = \begin{bmatrix} 0.4 \\ -13 \end{bmatrix}, \qquad \text{then} \qquad 10\mathbf{v} = \begin{bmatrix} 4 \\ -130 \end{bmatrix}.$$

Matrix Multiplication. The product $\mathbf{C} = \mathbf{AB}$ (in this order) of two $n \times n$ matrices $\mathbf{A} = [a_{jk}]$ and $\mathbf{B} = [b_{jk}]$ is the $n \times n$ matrix $\mathbf{C} = [c_{jk}]$ with entries

$$\textbf{(6)} \qquad c_{jk} = \sum_{m=1}^{n} a_{jm} b_{mk} \qquad \begin{array}{l} j = 1, \cdots, n \\ k = 1, \cdots, n, \end{array}$$

that is, multiply each entry in the jth *row* of $\mathbf{A}$ by the corresponding entry in the kth *column* of $\mathbf{B}$ and then add these n products. One says briefly that this is a "multiplication of rows into columns." For example,

$$\begin{bmatrix} 9 & 3 \\ -2 & 0 \end{bmatrix} \begin{bmatrix} 1 & -4 \\ 2 & 5 \end{bmatrix} = \begin{bmatrix} 9 \cdot 1 + 3 \cdot 2 & 9 \cdot (-4) + 3 \cdot 5 \\ -2 \cdot 1 + 0 \cdot 2 & (-2) \cdot (-4) + 0 \cdot 5 \end{bmatrix},$$

$$= \begin{bmatrix} 15 & -21 \\ -2 & 8 \end{bmatrix}.$$

CAUTION! Matrix multiplication is ***not commutative,*** $\mathbf{AB} \neq \mathbf{BA}$ in general. In our example,

$$\begin{bmatrix} 1 & -4 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} 9 & 3 \\ -2 & 0 \end{bmatrix} = \begin{bmatrix} 1 \cdot 9 + (-4) \cdot (-2) & 1 \cdot 3 + (-4) \cdot 0 \\ 2 \cdot 9 + 5 \cdot (-2) & 2 \cdot 3 + 5 \cdot 0 \end{bmatrix}$$

$$= \begin{bmatrix} 17 & 3 \\ 8 & 6 \end{bmatrix}.$$

Multiplication of an $n \times n$ matrix $\mathbf{A}$ by a vector $\mathbf{x}$ with n components is defined by the same rule: $\mathbf{v} = \mathbf{Ax}$ is the vector with the n components

$$v_j = \sum_{m=1}^{n} a_{jm} x_m \qquad j = 1, \cdots, n.$$

For example,

$$\begin{bmatrix} 12 & 7 \\ -8 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 12x_1 + 7x_2 \\ -8x_1 + 3x_2 \end{bmatrix}.$$

Systems of ODEs as Vector Equations

Differentiation. The *derivative* of a matrix (or vector) with variable entries (or components) is obtained by differentiating each entry (or component). Thus, if

$$\mathbf{y}(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = \begin{bmatrix} e^{-2t} \\ \sin t \end{bmatrix}, \qquad \text{then} \qquad \mathbf{y}'(t) = \begin{bmatrix} y_1'(t) \\ y_2'(t) \end{bmatrix} = \begin{bmatrix} -2e^{-2t} \\ \cos t \end{bmatrix}.$$

Using matrix multiplication and differentiation, we can now write (1) as

$$\text{(7)} \qquad \mathbf{y}' = \begin{bmatrix} y_1' \\ y_2' \end{bmatrix} = \mathbf{Ay} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \quad \text{e.g.,} \quad \mathbf{y}' = \begin{bmatrix} -5 & 2 \\ 13 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

Similarly for (2) by means of an $n \times n$ matrix $\mathbf{A}$ and a column vector $\mathbf{y}$ with n components, namely, $\mathbf{y}' = \mathbf{Ay}$. The vector equation (7) is equivalent to two equations for the components, and these are precisely the two ODEs in (1).

Some Further Operations and Terms

Transposition is the operation of writing columns as rows and conversely and is indicated by T. Thus the transpose $\mathbf{A}^{\mathsf{T}}$ of the 2×2 matrix

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} -5 & 2 \\ 13 & \frac{1}{2} \end{bmatrix} \qquad \text{is} \qquad \mathbf{A}^{\mathsf{T}} = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix} = \begin{bmatrix} -5 & 13 \\ 2 & \frac{1}{2} \end{bmatrix}.$$

The transpose of a column vector, say,

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}, \qquad \text{is a row vector,} \qquad \mathbf{v}^{\mathsf{T}} = [v_1 \quad v_2],$$

and conversely.

Inverse of a Matrix. The $n \times n$ **unit matrix I** is the $n \times n$ matrix with main diagonal $1, 1, \cdots, 1$ and all other entries zero. If, for a given $n \times n$ matrix $\mathbf{A}$, there is an $n \times n$ matrix $\mathbf{B}$ such that $\mathbf{AB} = \mathbf{BA} = \mathbf{I}$, then $\mathbf{A}$ is called **nonsingular** and $\mathbf{B}$ is called the **inverse** of $\mathbf{A}$ and is denoted by $\mathbf{A}^{-1}$; thus

$$\text{(8)} \qquad \mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}.$$

The inverse exists if the determinant $\det \mathbf{A}$ of $\mathbf{A}$ is not zero.

If $\mathbf{A}$ has no inverse, it is called **singular**. For $n = 2$,

$$\text{(9)} \qquad \mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix},$$

where the **determinant** of $\mathbf{A}$ is

$$\text{(10)} \qquad \det \mathbf{A} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

(For general n, see Sec. 7.7, but this will not be needed in this chapter.)

Linear Independence. r given vectors $\mathbf{v}^{(1)}, \cdots, \mathbf{v}^{(r)}$ with n components are called a *linearly independent set* or, more briefly, **linearly independent**, if

$$\text{(11)} \qquad c_1\mathbf{v}^{(1)} + \cdots + c_r\mathbf{v}^{(r)} = \mathbf{0}$$

implies that all scalars $c_1, \cdots, c_r$ must be zero; here, $\mathbf{0}$ denotes the **zero vector**, whose n components are all zero. If (11) also holds for scalars not all zero (so that at least one of these scalars is not zero), then these vectors are called a *linearly dependent set* or, briefly, **linearly dependent**, because then at least one of them can be expressed as a **linear combination** of the others; that is, if, for instance, $c_1 \neq 0$ in (11), then we can obtain

$$\mathbf{v}^{(1)} = -\frac{1}{c_1}(c_2\mathbf{v}^{(2)} + \cdots + c_r\mathbf{v}^{(r)}).$$

Eigenvalues, Eigenvectors

Eigenvalues and eigenvectors will be very important in this chapter (and, as a matter of fact, throughout mathematics).

Let $\mathbf{A} = [a_{jk}]$ be an $n \times n$ matrix. Consider the equation

$$\mathbf{Ax} = \lambda\mathbf{x} \tag{12}$$

where λ is a scalar (a real or complex number) to be determined and $\mathbf{x}$ is a vector to be determined. Now, for every λ, a solution is $\mathbf{x} = \mathbf{0}$. A scalar λ such that (12) holds for some vector $\mathbf{x} \neq \mathbf{0}$ is called an **eigenvalue** of $\mathbf{A}$, and this vector is called an **eigenvector** of $\mathbf{A}$ corresponding to this eigenvalue λ.

We can write (12) as $\mathbf{Ax} - \lambda\mathbf{x} = \mathbf{0}$ or

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}. \tag{13}$$

These are n linear algebraic equations in the n unknowns $x_1, \cdots, x_n$ (the components of $\mathbf{x}$). For these equations to have a solution $\mathbf{x} \neq \mathbf{0}$, the determinant of the coefficient matrix $\mathbf{A} - \lambda\mathbf{I}$ must be zero. This is proved as a basic fact in linear algebra (Theorem 4 in Sec. 7.7). In this chapter we need this only for $n = 2$. Then (13) is

$$\begin{bmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}; \tag{14}$$

in components,

$$\begin{aligned} (a_{11} - \lambda)x_1 + \quad a_{12}x_2 \quad &= 0 \\ a_{21}x_1 \quad + (a_{22} - \lambda)x_2 &= 0. \end{aligned} \tag{14*}$$

Now $\mathbf{A} - \lambda\mathbf{I}$ is singular if and only if its determinant $\det(\mathbf{A} - \lambda\mathbf{I})$, called the **characteristic determinant** of $\mathbf{A}$ (also for general n), is zero. This gives

$$\begin{aligned} \det(\mathbf{A} - \lambda\mathbf{I}) &= \begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{vmatrix} \\ &= (a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21} \\ &= \lambda^2 - (a_{11} + a_{22})\lambda + a_{11}a_{22} - a_{12}a_{21} = 0. \end{aligned} \tag{15}$$

This quadratic equation in λ is called the **characteristic equation** of $\mathbf{A}$. Its solutions are the eigenvalues λ_1 and λ_2 of $\mathbf{A}$. First determine these. Then use (14*) with $\lambda = \lambda_1$ to determine an eigenvector $\mathbf{x}^{(1)}$ of $\mathbf{A}$ corresponding to λ_1. Finally use (14*) with $\lambda = \lambda_2$ to find an eigenvector $\mathbf{x}^{(2)}$ of $\mathbf{A}$ corresponding to λ_2. Note that if $\mathbf{x}$ is an eigenvector of $\mathbf{A}$, so is $k\mathbf{x}$ with any $k \neq 0$.

EXAMPLE 1 Eigenvalue Problem

Find the eigenvalues and eigenvectors of the matrix

$$\mathbf{A} = \begin{bmatrix} -4.0 & 4.0 \\ -1.6 & 1.2 \end{bmatrix} \tag{16}$$

Solution. The characteristic equation is the quadratic equation

$$\det|\mathbf{A} - \lambda\mathbf{I}| = \begin{vmatrix} -4 - \lambda & 4 \\ -1.6 & 1.2 - \lambda \end{vmatrix} = \lambda^2 + 2.8\lambda + 1.6 = 0.$$

It has the solutions $\lambda_1 = -2$ and $\lambda_2 = -0.8$. These are the eigenvalues of $\mathbf{A}$.

Eigenvectors are obtained from (14*). For $\lambda = \lambda_1 = -2$ we have from (14*)

$$\begin{aligned} (-4.0 + 2.0)x_1 + \quad 4.0x_2 \quad &= 0 \\ -1.6x_1 \quad + (1.2 + 2.0)x_2 &= 0. \end{aligned}$$

A solution of the first equation is $x_1 = 2, x_2 = 1$. This also satisfies the second equation. (Why?) Hence an eigenvector of $\mathbf{A}$ corresponding to $\lambda_1 = -2.0$ is

$$\mathbf{x}^{(1)} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}. \qquad \text{Similarly,} \qquad \mathbf{x}^{(2)} = \begin{bmatrix} 1 \\ 0.8 \end{bmatrix} \tag{17}$$

is an eigenvector of $\mathbf{A}$ corresponding to $\lambda_2 = -0.8$, as obtained from (14*) with $\lambda = \lambda_2$. Verify this. ■

4.1 Systems of ODEs as Models in Engineering Applications

We show how systems of ODEs are of practical importance as follows. We first illustrate how systems of ODEs can serve as models in various applications. Then we show how a higher order ODE (with the highest derivative standing alone on one side) can be reduced to a first-order system.

EXAMPLE 1 Mixing Problem Involving Two Tanks

A mixing problem involving a single tank is modeled by a single ODE, and you may first review the corresponding Example 3 in Sec. 1.3 because the principle of modeling will be the same for two tanks. The model will be a system of two first-order ODEs.

Tank T_1 and T_2 in Fig. 78 contain initially 100 gal of water each. In T_1 the water is pure, whereas 150 lb of fertilizer are dissolved in T_2. By circulating liquid at a rate of 2 gal/min and stirring (to keep the mixture uniform) the amounts of fertilizer $y_1(t)$ in T_1 and $y_2(t)$ in T_2 change with time t. How long should we let the liquid circulate so that T_1 will contain at least half as much fertilizer as there will be left in T_2?

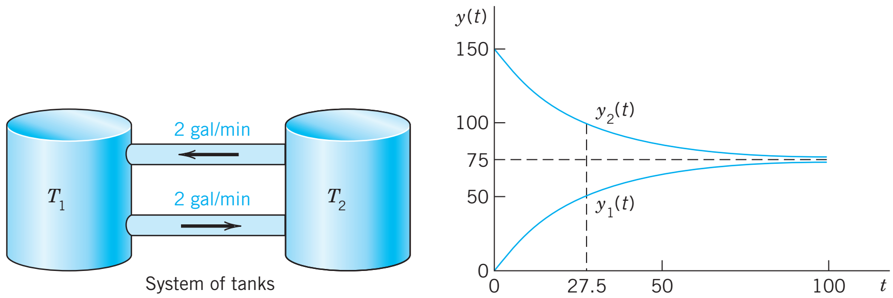


Fig. 78. Fertilizer content in Tanks T_1 (lower curve) and T_2

Solution. ***Step 1.*** **Setting up the model.** As for a single tank, the time rate of change $y_1'(t)$ of $y_1(t)$ equals inflow minus outflow. Similarly for tank T_2. From Fig. 78 we see that

$$y_1' = \text{Inflow/min} - \text{Outflow/min} = \frac{2}{100}y_2 - \frac{2}{100}y_1 \qquad \text{(Tank } T_1\text{)}$$

$$y_2' = \text{Inflow/min} - \text{Outflow/min} = \frac{2}{100}y_1 - \frac{2}{100}y_2 \qquad \text{(Tank } T_2\text{).}$$

Hence the mathematical model of our mixture problem is the system of first-order ODEs

$$\begin{aligned} y_1' &= -0.02y_1 + 0.02y_2 && \text{(Tank } T_1\text{)} \\ y_2' &= 0.02y_1 - 0.02y_2 && \text{(Tank } T_2\text{).} \end{aligned}$$

As a vector equation with column vector $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ and matrix $\mathbf{A}$ this becomes

$$\mathbf{y}' = \mathbf{A}\mathbf{y}, \qquad \text{where} \qquad \mathbf{A} = \begin{bmatrix} -0.02 & 0.02 \\ 0.02 & -0.02 \end{bmatrix}.$$

Step 2. **General solution.** As for a single equation, we try an exponential function of t,

$$\text{(1)} \qquad \mathbf{y} = \mathbf{x}e^{\lambda t}. \qquad \text{Then} \qquad \mathbf{y}' = \lambda\mathbf{x}e^{\lambda t} = \mathbf{A}\mathbf{x}e^{\lambda t}.$$

Dividing the last equation $\lambda\mathbf{x}e^{\lambda t} = \mathbf{A}\mathbf{x}e^{\lambda t}$ by $e^{\lambda t}$ and interchanging the left and right sides, we obtain

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}.$$

We need nontrivial solutions (solutions that are not identically zero). Hence we have to look for eigenvalues and eigenvectors of $\mathbf{A}$. The eigenvalues are the solutions of the characteristic equation

$$\text{(2)} \qquad \det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} -0.02 - \lambda & 0.02 \\ 0.02 & -0.02 - \lambda \end{vmatrix} = (-0.02 - \lambda)^2 - 0.02^2 = \lambda(\lambda + 0.04) = 0.$$

We see that $\lambda_1 = 0$ (which can very well happen—don't get mixed up—it is eigen*vectors* that must not be zero) and $\lambda_2 = -0.04$. Eigenvectors are obtained from (14*) in Sec. 4.0 with $\lambda = 0$ and $\lambda = -0.04$. For our present $\mathbf{A}$ this gives [we need only the first equation in (14*)]

$$-0.02x_1 + 0.02x_2 = 0 \qquad \text{and} \qquad (-0.02 + 0.04)x_1 + 0.02x_2 = 0,$$

respectively. Hence $x_1 = x_2$ and $x_1 = -x_2$, respectively, and we can take $x_1 = x_2 = 1$ and $x_1 = -x_2 = 1$. This gives two eigenvectors corresponding to $\lambda_1 = 0$ and $\lambda_2 = -0.04$, respectively, namely,

$$\mathbf{x}^{(1)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{and} \quad \mathbf{x}^{(2)} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

From (1) and the superposition principle (which continues to hold for systems of homogeneous linear ODEs) we thus obtain a solution

$$\mathbf{y} = c_1\mathbf{x}^{(1)}e^{\lambda_1 t} + c_2\mathbf{x}^{(2)}e^{\lambda_2 t} = c_1\begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2\begin{bmatrix} 1 \\ -1 \end{bmatrix}e^{-0.04t} \tag{3}$$

where c_1 and c_2 are arbitrary constants. Later we shall call this a **general solution**.

Step 3. **Use of initial conditions.** The initial conditions are $y_1(0) = 0$ (no fertilizer in tank T_1) and $y_2(0) = 150$. From this and (3) with $t = 0$ we obtain

$$\mathbf{y}(0) = c_1\begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} c_1 + c_2 \\ c_1 - c_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 150 \end{bmatrix}.$$

In components this is $c_1 + c_2 = 0$, $c_1 - c_2 = 150$. The solution is $c_1 = 75$, $c_2 = -75$. This gives the answer

$$\mathbf{y} = 75\mathbf{x}^{(1)} - 75\mathbf{x}^{(2)}e^{-0.04t} = 75\begin{bmatrix} 1 \\ 1 \end{bmatrix} - 75\begin{bmatrix} 1 \\ -1 \end{bmatrix}e^{-0.04t}.$$

In components,

$$y_1 = 75 - 75e^{-0.04t} \qquad \text{(Tank } T_1\text{, lower curve)}$$
$$y_2 = 75 + 75e^{-0.04t} \qquad \text{(Tank } T_2\text{, upper curve).}$$

Figure 78 shows the exponential increase of y_1 and the exponential decrease of y_2 to the common limit 75 lb. Did you expect this for physical reasons? Can you physically explain why the curves look "symmetric"? Would the limit change if T_1 initially contained 100 lb of fertilizer and T_2 contained 50 lb?

Step 4. **Answer.** T_1 contains half the fertilizer amount of T_2 if it contains 1/3 of the total amount, that is, 50 lb. Thus

$$y_1 = 75 - 75e^{-0.04t} = 50, \qquad e^{-0.04t} = \tfrac{1}{3}, \qquad t = (\ln 3)/0.04 = 27.5.$$

Hence the fluid should circulate for at least about half an hour. ■

EXAMPLE 2 Electrical Network

Find the currents $I_1(t)$ and $I_2(t)$ in the network in Fig. 79. Assume all currents and charges to be zero at $t = 0$, the instant when the switch is closed.

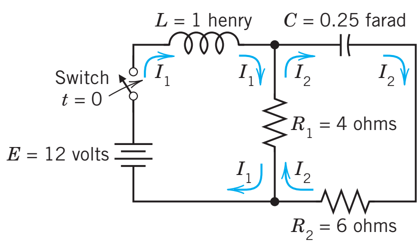


Fig. 79. Electrical network in Example 2

Solution. ***Step 1.*** **Setting up the mathematical model.** The model of this network is obtained from Kirchhoff's Voltage Law, as in Sec. 2.9 (where we considered single circuits). Let $I_1(t)$ and $I_2(t)$ be the currents in the left and right loops, respectively. In the left loop, the voltage drops are $LI_1' = I_1'$ [V] over the inductor and $R_1(I_1 - I_2) = 4(I_1 - I_2)$ [V] over the resistor, the difference because I_1 and I_2 flow through the resistor in opposite directions. By Kirchhoff's Voltage Law the sum of these drops equals the voltage of the battery; that is, $I_1' + 4(I_1 - I_2) = 12$, hence

$$I_1' = -4I_1 + 4I_2 + 12. \tag{4a}$$

In the right loop, the voltage drops are $R_2I_2 = 6I_2$ [V] and $R_1(I_2 - I_1) = 4(I_2 - I_1)$ [V] over the resistors and $(I/C)\int I_2\,dt = 4\int I_2\,dt$ [V] over the capacitor, and their sum is zero,

$$6I_2 + 4(I_2 - I_1) + 4\int I_2\,dt = 0 \qquad \text{or} \qquad 10I_2 - 4I_1 + 4\int I_2\,dt = 0.$$

Division by 10 and differentiation gives $I_2' - 0.4I_1' + 0.4I_2 = 0$.

To simplify the solution process, we first get rid of $0.4I_1'$, which by (4a) equals $0.4(-4I_1 + 4I_2 + 12)$. Substitution into the present ODE gives

$$I_2' = 0.4I_1' - 0.4I_2 = 0.4(-4I_1 + 4I_2 + 12) - 0.4I_2$$

and by simplification

$$I_2' = -1.6I_1 + 1.2I_2 + 4.8. \tag{4b}$$

In matrix form, (4) is (we write $\mathbf{J}$ since $\mathbf{I}$ is the unit matrix)

$$\mathbf{J}' = \mathbf{AJ} + \mathbf{g}, \qquad \text{where} \qquad \mathbf{J} = \begin{bmatrix} I_1 \\ I_2 \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} -4.0 & 4.0 \\ -1.6 & 1.2 \end{bmatrix}, \quad \mathbf{g} = \begin{bmatrix} 12.0 \\ 4.8 \end{bmatrix}. \tag{5}$$

Step 2. **Solving (5).** Because of the vector $\mathbf{g}$ this is a *nonhomogeneous* system, and we try to proceed as for a single ODE, solving first the *homogeneous* system $\mathbf{J}' = \mathbf{AJ}$ (thus $\mathbf{J}' - \mathbf{AJ} = \mathbf{0}$) by substituting $\mathbf{J} = \mathbf{x}e^{\lambda t}$. This gives

$$\mathbf{J}' = \lambda\mathbf{x}e^{\lambda t} = \mathbf{Ax}e^{\lambda t}, \qquad \text{hence} \qquad \mathbf{Ax} = \lambda\mathbf{x}.$$

Hence, to obtain a nontrivial solution, we again need the eigenvalues and eigenvectors. For the present matrix $\mathbf{A}$ they are derived in Example 1 in Sec. 4.0:

$$\lambda_1 = -2, \qquad \mathbf{x}^{(1)} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}; \qquad \lambda_2 = -0.8, \qquad \mathbf{x}^{(2)} = \begin{bmatrix} 1 \\ 0.8 \end{bmatrix}.$$

Hence a "general solution" of the homogeneous system is

$$\mathbf{J}_h = c_1\mathbf{x}^{(1)}e^{-2t} + c_2\mathbf{x}^{(2)}e^{-0.8t}.$$

For a particular solution of the nonhomogeneous system (5), since $\mathbf{g}$ is constant, we try a constant column vector $\mathbf{J}_p = \mathbf{a}$ with components a_1, a_2. Then $\mathbf{J}_p' = \mathbf{0}$, and substitution into (5) gives $\mathbf{Aa} + \mathbf{g} = \mathbf{0}$; in components,

$$\begin{aligned} -4.0a_1 + 4.0a_2 + 12.0 &= 0 \\ -1.6a_1 + 1.2a_2 + 4.8 &= 0. \end{aligned}$$

The solution is $a_1 = 3$, $a_2 = 0$; thus $\mathbf{a} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$. Hence

$$\mathbf{J} = \mathbf{J}_h + \mathbf{J}_p = c_1\mathbf{x}^{(1)}e^{-2t} + c_2\mathbf{x}^{(2)}e^{-0.8t} + \mathbf{a}; \tag{6}$$

in components,

$$\begin{aligned} I_1 &= 2c_1e^{-2t} + c_2e^{-0.8t} + 3 \\ I_2 &= c_1e^{-2t} + 0.8c_2e^{-0.8t}. \end{aligned}$$

The initial conditions give

$$I_1(0) = 2c_1 + \quad c_2 + 3 = 0$$
$$I_2(0) = \quad c_1 + 0.8c_2 \quad = 0.$$

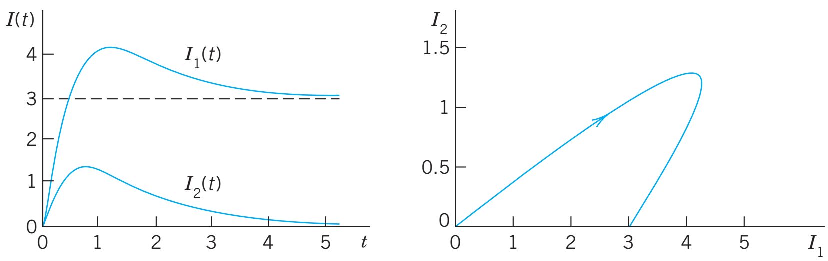
Hence $c_1 = -4$ and $c_2 = 5$. As the solution of our problem we thus obtain

$$\mathbf{J} = -4\mathbf{x}^{(1)}e^{-2t} + 5\mathbf{x}^{(2)}e^{-0.8t} + \mathbf{a}. \tag{7}$$

In components (Fig. 80b),

$$I_1 = -8e^{-2t} + 5e^{-0.8t} + 3$$
$$I_2 = -4e^{-2t} + 4e^{-0.8t}.$$

Now comes an important idea, on which we shall elaborate further, beginning in Sec. 4.3. Figure 80a shows $I_1(t)$ and $I_2(t)$ as two separate curves. Figure 80b shows these two currents as a single curve $[I_1(t), I_2(t)]$ in the I_1I_2-plane. This is a parametric representation with time t as the parameter. It is often important to know in which sense such a curve is traced. This can be indicated by an arrow in the sense of increasing t, as is shown. The I_1I_2-plane is called the **phase plane** of our system (5), and the curve in Fig. 80b is called a **trajectory**. We shall see that such "**phase plane representations**" are far more important than graphs as in Fig. 80a because they will give a much better qualitative overall impression of the general behavior of whole families of solutions, not merely of one solution as in the present case. ■



(a) Currents I_1 (upper curve) and I_2

(b) Trajectory $[I_1(t), I_2(t)]^T$ in the I_1I_2-plane (the "phase plane")

Fig. 80. Currents in Example 2

Remark. In both examples, by growing the dimension of the problem (from one tank to two tanks or one circuit to two circuits) we also increased the number of ODEs (from one ODE to two ODEs). This "growth" in the problem being reflected by an "increase" in the mathematical model is attractive and affirms the quality of our mathematical modeling and theory.

Conversion of an *n*th-Order ODE to a System

We show that an nth-order ODE of the general form (8) (see Theorem 1) can be converted to a system of n first-order ODEs. This is practically and theoretically important—practically because it permits the study and solution of single ODEs by methods for systems, and theoretically because it opens a way of including the theory of higher order ODEs into that of first-order systems. This conversion is another reason for the importance of systems, in addition to their use as models in various basic applications. The idea of the conversion is simple and straightforward, as follows.

THEOREM 1

Conversion of an ODE

An nth-order ODE

$$y^{(n)} = F(t, y, y', \cdots, y^{(n-1)}) \tag{8}$$

can be converted to a system of n first-order ODEs by setting

$$y_1 = y, \quad y_2 = y', \quad y_3 = y'', \cdots, y_n = y^{(n-1)}. \tag{9}$$

This system is of the form

$$\begin{aligned} y_1' &= y_2 \\ y_2' &= y_3 \\ &\vdots \\ y_{n-1}' &= y_n \\ y_n' &= F(t, y_1, y_2, \cdots, y_n). \end{aligned} \tag{10}$$

PROOF The first $n - 1$ of these n ODEs follows immediately from (9) by differentiation. Also, $y_n' = y^{(n)}$ by (9), so that the last equation in (10) results from the given ODE (8). ■

EXAMPLE 3 Mass on a Spring

To gain confidence in the conversion method, let us apply it to an old friend of ours, modeling the free motions of a mass on a spring (see Sec. 2.4)

$$my'' + cy' + ky = 0 \qquad \text{or} \qquad y'' = -\frac{c}{m}y' - \frac{k}{m}y.$$

For this ODE (8) the system (10) is linear and homogeneous,

$$\begin{aligned} y_1' &= y_2 \\ y_2' &= -\frac{k}{m}y_1 - \frac{c}{m}y_2. \end{aligned}$$

Setting $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$, we get in matrix form

$$\mathbf{y}' = \mathbf{A}\mathbf{y} = \begin{bmatrix} 0 & 1 \\ -\dfrac{k}{m} & -\dfrac{c}{m} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

The characteristic equation is

$$\det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} -\lambda & 1 \\ -\dfrac{k}{m} & -\dfrac{c}{m} - \lambda \end{vmatrix} = \lambda^2 + \frac{c}{m}\lambda + \frac{k}{m} = 0.$$

It agrees with that in Sec. 2.4. For an illustrative computation, let $m = 1, c = 2$, and $k = 0.75$. Then

$$\lambda^2 + 2\lambda + 0.75 = (\lambda + 0.5)(\lambda + 1.5) = 0.$$

This gives the eigenvalues $\lambda_1 = -0.5$ and $\lambda_2 = -1.5$. Eigenvectors follow from the first equation in $\mathbf{A} - \lambda\mathbf{I} = \mathbf{0}$, which is $-\lambda x_1 + x_2 = 0$. For λ_1 this gives $0.5x_1 + x_2 = 0$, say, $x_1 = 2$, $x_2 = -1$. For $\lambda_2 = -1.5$ it gives $1.5x_1 + x_2 = 0$, say, $x_1 = 1$, $x_2 = -1.5$. These eigenvectors

$$\mathbf{x}^{(1)} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \quad \mathbf{x}^{(2)} = \begin{bmatrix} 1 \\ -1.5 \end{bmatrix} \quad \text{give} \quad \mathbf{y} = c_1 \begin{bmatrix} 2 \\ -1 \end{bmatrix} e^{-0.5t} + c_2 \begin{bmatrix} 1 \\ -1.5 \end{bmatrix} e^{-1.5t}.$$

This vector solution has the first component

$$y = y_1 = 2c_1 e^{-0.5t} + c_2 e^{-1.5t}$$

which is the expected solution. The second component is its derivative

$$y_2 = y_1' = y' = -c_1 e^{-0.5t} - 1.5c_2 e^{-1.5t}.$$ ■

PROBLEM SET 4.1

1–6 MIXING PROBLEMS

1. Find out, without calculation, whether doubling the flow rate in Example 1 has the same effect as halfing the tank sizes. (Give a reason.)

2. What happens in Example 1 if we replace T_1 by a tank containing 200 gal of water and 150 lb of fertilizer dissolved in it?

3. Derive the eigenvectors in Example 1 without consulting this book.

4. In Example 1 find a "general solution" for any ratio $a = (\text{flow rate})/(\text{tank size})$, tank sizes being equal. Comment on the result.

5. If you extend Example 1 by a tank T_3 of the same size as the others and connected to T_2 by two tubes with flow rates as between T_1 and T_2, what system of ODEs will you get?

6. Find a "general solution" of the system in Prob. 5.

7–9 ELECTRICAL NETWORK

In Example 2 find the currents:

7. If the initial currents are 0 A and -3 A (minus meaning that $I_2(0)$ flows against the direction of the arrow).

8. If the capacitance is changed to $C = 5/27$ F. (General solution only.)

9. If the initial currents in Example 2 are 28 A and 14 A.

10–13 CONVERSION TO SYSTEMS

Find a general solution of the given ODE **(a)** by first converting it to a system, **(b)**, as given. Show the details of your work.

10. $y'' + 3y' + 2y = 0$

11. $4y'' - 15y' - 4y = 0$

12. $y''' + 2y'' - y' - 2y = 0$

13. $y'' + 2y' - 24y = 0$

14. TEAM PROJECT. Two Masses on Springs. **(a)** Set up the model for the (undamped) system in Fig. 81. **(b)** Solve the system of ODEs obtained. *Hint.* Try $\mathbf{y} = \mathbf{x}e^{\omega t}$ and set $\omega^2 = \lambda$. Proceed as in Example 1 or 2. **(c)** Describe the influence of initial conditions on the possible kind of motions.

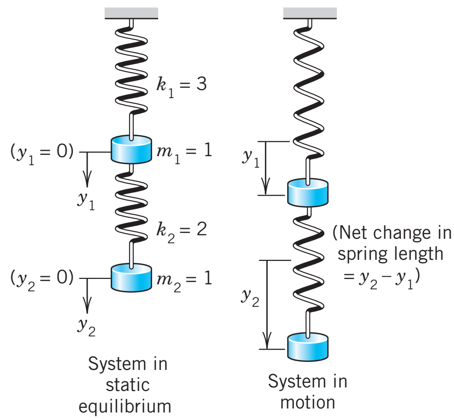


Fig. 81. Mechanical system in Team Project

15. CAS EXPERIMENT. Electrical Network. **(a)** In Example 2 choose a sequence of values of C that increases beyond bound, and compare the corresponding sequences of eigenvalues of $\mathbf{A}$. What limits of these sequences do your numeric values (approximately) suggest?

(b) Find these limits analytically.

(c) Explain your result physically.

(d) Below what value (approximately) must you decrease C to get vibrations?

4.2 Basic Theory of Systems of ODEs. Wronskian

In this section we discuss some basic concepts and facts about system of ODEs that are quite similar to those for single ODEs.

The first-order systems in the last section were special cases of the more general system

$$\begin{aligned} y_1' &= f_1(t, y_1, \cdots, y_n) \\ y_2' &= f_2(t, y_1, \cdots, y_n) \\ &\cdots \\ y_n' &= f_n(t, y_1, \cdots, y_n). \end{aligned} \tag{1}$$

We can write the system (1) as a vector equation by introducing the column vectors $\mathbf{y} = [y_1 \quad \cdots \quad y_n]^\mathsf{T}$ and $\mathbf{f} = [f_1 \quad \cdots \quad f_n]^\mathsf{T}$ (where T means *transposition* and saves us the space that would be needed for writing $\mathbf{y}$ and $\mathbf{f}$ as columns). This gives

$$\mathbf{y}' = \mathbf{f}(t, \mathbf{y}). \tag{1}$$

This system (1) includes almost all cases of practical interest. For $n = 1$ it becomes $y_1' = f_1(t, y_1)$ or, simply, $y' = f(t, y)$, well known to us from Chap. 1.

A **solution** of (1) on some interval $a < t < b$ is a set of n differentiable functions

$$y_1 = h_1(t), \quad \cdots, \quad y_n = h_n(t)$$

on $a < t < b$ that satisfy (1) throughout this interval. In vector from, introducing the "*solution vector*" $\mathbf{h} = [h_1 \quad \cdots \quad h_n]^\mathsf{T}$ (a column vector!) we can write

$$\mathbf{y} = \mathbf{h}(t).$$

An **initial value problem** for (1) consists of (1) and n given **initial conditions**

$$y_1(t_0) = K_1, \qquad y_2(t_0) = K_2, \qquad \cdots, \qquad y_n(t_0) = K_n, \tag{2}$$

in vector form, $\mathbf{y}(t_0) = \mathbf{K}$, where t_0 is a specified value of t in the interval considered and the components of $\mathbf{K} = [K_1 \quad \cdots \quad K_n]^\mathsf{T}$ are given numbers. Sufficient conditions for the existence and uniqueness of a solution of an initial value problem (1), (2) are stated in the following theorem, which extends the theorems in Sec. 1.7 for a single equation. (For a proof, see Ref. [A7].)

THEOREM 1

Existence and Uniqueness Theorem

Let $f_1, \cdots, f_n$ in (1) *be continuous functions having continuous partial derivatives $\partial f_1/\partial y_1, \cdots, \partial f_1/\partial y_n, \cdots, \partial f_n/\partial y_n$ in some domain R of $ty_1y_2\cdots y_n$-space containing the point $(t_0, K_1, \cdots, K_n)$. Then* (1) *has a solution on some interval $t_0 - \alpha < t < t_0 + \alpha$ satisfying* (2), *and this solution is unique.*

Linear Systems

Extending the notion of a *linear* ODE, we call (1) a **linear system** if it is linear in $y_1, \cdots, y_n$; that is, if it can be written

$$\begin{aligned} y_1' &= a_{11}(t)y_1 + \cdots + a_{1n}(t)y_n + g_1(t) \\ &\vdots \\ y_n' &= a_{n1}(t)y_1 + \cdots + a_{nn}(t)y_n + g_n(t). \end{aligned} \tag{3}$$

As a vector equation this becomes

$$\mathbf{y}' = \mathbf{A}\mathbf{y} + \mathbf{g} \tag{3}$$

where $\mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ . & \cdots & . \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}, \quad \mathbf{g} = \begin{bmatrix} g_1 \\ \vdots \\ g_n \end{bmatrix}.$

This system is called **homogeneous** if $\mathbf{g} = \mathbf{0}$, so that it is

$$\mathbf{y}' = \mathbf{A}\mathbf{y}. \tag{4}$$

If $\mathbf{g} \neq \mathbf{0}$, then (3) is called **nonhomogeneous**. For example, the systems in Examples 1 and 3 of Sec. 4.1 are homogeneous. The system in Example 2 of that section is nonhomogeneous.

For a linear system (3) we have $\partial f_1/\partial y_1 = a_{11}(t), \cdots, \partial f_n/\partial y_n = a_{nn}(t)$ in Theorem 1. Hence for a linear system we simply obtain the following.

THEOREM 2

Existence and Uniqueness in the Linear Case

Let the a_{jk}'s and g_j's in (3) *be continuous functions of t on an open interval $\alpha < t < \beta$ containing the point $t = t_0$. Then* (3) *has a solution $\mathbf{y}(t)$ on this interval satisfying* (2), *and this solution is unique.*

As for a single homogeneous linear ODE we have

THEOREM 3

Superposition Principle or Linearity Principle

If $\mathbf{y}^{(1)}$ and $\mathbf{y}^{(2)}$ are solutions of the ***homogeneous linear*** *system* (4) *on some interval, so is any linear combination $\mathbf{y} = c_1\mathbf{y}^{(1)} + c_1\mathbf{y}^{(2)}$.*

PROOF Differentiating and using (4), we obtain

$$\begin{aligned} \mathbf{y}' &= [c_1\mathbf{y}^{(1)} + c_1\mathbf{y}^{(2)}]' \\ &= c_1\mathbf{y}^{(1)\prime} + c_2\mathbf{y}^{(2)\prime} \\ &= c_1\mathbf{A}\mathbf{y}^{(1)} + c_2\mathbf{A}\mathbf{y}^{(2)} \\ &= \mathbf{A}(c_1\mathbf{y}^{(1)} + c_2\mathbf{y}^{(2)}) = \mathbf{A}\mathbf{y}. \end{aligned}$$ ∎

The general theory of linear systems of ODEs is quite similar to that of a single linear ODE in Secs. 2.6 and 2.7. To see this, we explain the most basic concepts and facts. For proofs we refer to more advanced texts, such as [A7].

Basis. General Solution. Wronskian

By a **basis** or a **fundamental system** of solutions of the homogeneous system (4) on some interval J we mean a linearly independent set of n solutions $\mathbf{y}^{(1)}, \cdots, \mathbf{y}^{(n)}$ of (4) on that interval. (We write J because we need $\mathbf{I}$ to denote the unit matrix.) We call a corresponding linear combination

$$\mathbf{y} = c_1\mathbf{y}^{(1)} \cdots + c_n\mathbf{y}^{(n)} \qquad (c_1, \cdots, c_n \text{ arbitrary}) \tag{5}$$

a **general solution** of (4) on J. It can be shown that if the $a_{jk}(t)$ in (4) are continuous on J, then (4) has a basis of solutions on J, hence a general solution, which includes every solution of (4) on J.

We can write n solutions $\mathbf{y}^{(1)}, \cdots, \mathbf{y}^{(n)}$ of (4) on some interval J as columns of an $n \times n$ matrix

$$\mathbf{Y} = [\mathbf{y}^{(1)} \quad \cdots \quad \mathbf{y}^{(n)}]. \tag{6}$$

The determinant of $\mathbf{Y}$ is called the **Wronskian** of $\mathbf{y}^{(1)}, \cdots, \mathbf{y}^{(n)}$, written

$$W(\mathbf{y}^{(1)}, \cdots, \mathbf{y}^{(n)}) = \begin{vmatrix} y_1^{(1)} & y_1^{(2)} & \cdots & y_1^{(n)} \\ y_2^{(1)} & y_2^{(2)} & \cdots & y_2^{(n)} \\ \cdot & \cdot & \cdots & \cdot \\ y_n^{(1)} & y_n^{(2)} & \cdots & y_n^{(n)} \end{vmatrix}. \tag{7}$$

The columns are these solutions, each in terms of components. These solutions form a basis on J if and only if W is not zero at any t_1 in this interval. W is either identically zero or nowhere zero in J. (This is similar to Secs. 2.6 and 3.1.)

If the solutions $\mathbf{y}^{(1)}, \cdots, \mathbf{y}^{(n)}$ in (5) form a basis (a fundamental system), then (6) is often called a **fundamental matrix**. Introducing a column vector $\mathbf{c} = [c_1 \quad c_2 \quad \cdots \quad c_n]^\mathsf{T}$, we can now write (5) simply as

$$\mathbf{y} = \mathbf{Y}\mathbf{c}. \tag{8}$$

Furthermore, we can relate (7) to Sec. 2.6, as follows. If y and z are solutions of a second-order homogeneous linear ODE, their Wronskian is

$$W(y, z) = \begin{vmatrix} y & z \\ y' & z' \end{vmatrix}.$$

To write this ODE as a system, we have to set $y = y_1, y' = y_1' = y_2$ and similarly for z (see Sec. 4.1). But then $W(y, z)$ becomes (7), except for notation.

4.3 Constant-Coefficient Systems. Phase Plane Method

Continuing, we now assume that our homogeneous linear system

$$\mathbf{y}' = \mathbf{A}\mathbf{y} \tag{1}$$

under discussion has ***constant coefficients***, so that the $n \times n$ matrix $\mathbf{A} = [a_{jk}]$ has entries not depending on t. We want to solve (1). Now a single ODE $y' = ky$ has the solution $y = Ce^{kt}$. So let us try

$$\mathbf{y} = \mathbf{x}e^{\lambda t}. \tag{2}$$

Substitution into (1) gives $\mathbf{y}' = \lambda\mathbf{x}e^{\lambda t} = \mathbf{A}\mathbf{y} = \mathbf{A}\mathbf{x}e^{\lambda t}$. Dividing by $e^{\lambda t}$, we obtain the **eigenvalue problem**

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}. \tag{3}$$

Thus the nontrivial solutions of (1) (solutions that are not zero vectors) are of the form (2), where λ is an eigenvalue of $\mathbf{A}$ and $\mathbf{x}$ is a corresponding eigenvector.

We assume that $\mathbf{A}$ has a linearly independent set of n eigenvectors. This holds in most applications, in particular if $\mathbf{A}$ is symmetric ($a_{kj} = a_{jk}$) or skew-symmetric ($a_{kj} = -a_{jk}$) or has n *different* eigenvalues.

Let those eigenvectors be $\mathbf{x}^{(1)}, \cdots, \mathbf{x}^{(n)}$ and let them correspond to eigenvalues $\lambda_1, \cdots, \lambda_n$ (which may be all different, or some—or even all—may be equal). Then the corresponding solutions (2) are

$$\mathbf{y}^{(4)} = \mathbf{x}^{(1)}e^{\lambda_1 t}, \quad \cdots, \quad \mathbf{y}^{(n)} = \mathbf{x}^{(n)}e^{\lambda_n t}. \tag{4}$$

Their Wronskian $W = W(\mathbf{y}^{(1)}, \cdots, \mathbf{y}^{(n)})$ [(7) in Sec. 4.2] is given by

$$W = (\mathbf{y}^{(1)}, \cdots, \mathbf{y}^{(n)}) = \begin{vmatrix} x_1^{(1)}e^{\lambda_1 t} & \cdots & x_1^{(n)}e^{\lambda_n t} \\ x_2^{(1)}e^{\lambda_1 t} & \cdots & x_2^{(n)}e^{\lambda_n t} \\ \cdot & \cdots & \cdot \\ x_n^{(1)}e^{\lambda_1 t} & \cdots & x_n^{(n)}e^{\lambda_n t} \end{vmatrix} = e^{\lambda_1 t + \cdots + \lambda_n t} \begin{vmatrix} x_1^{(1)} & \cdots & x_1^{(n)} \\ x_2^{(1)} & \cdots & x_2^{(n)} \\ \cdot & \cdots & \cdot \\ x_n^{(1)} & \cdots & x_n^{(n)} \end{vmatrix}.$$

On the right, the exponential function is never zero, and the determinant is not zero either because its columns are the n linearly independent eigenvectors. This proves the following theorem, whose assumption is true if the matrix $\mathbf{A}$ is symmetric or skew-symmetric, or if the n eigenvalues of $\mathbf{A}$ are all different.

THEOREM 1

General Solution

If the constant matrix $\mathbf{A}$ *in the system* (1) *has a linearly independent set of* n *eigenvectors, then the corresponding solutions* $\mathbf{y}^{(1)}, \cdots, \mathbf{y}^{(n)}$ *in* (4) *form a basis of solutions of* (1), *and the corresponding general solution is*

$$\mathbf{y} = c_1\mathbf{x}^{(1)}e^{\lambda_1 t} + \cdots + c_n\mathbf{x}^{(n)}e^{\lambda_n t}. \tag{5}$$

How to Graph Solutions in the Phase Plane

We shall now concentrate on systems (1) with constant coefficients consisting of two ODEs

$$\mathbf{y}' = \mathbf{A}\mathbf{y}; \qquad \text{in components,} \qquad \begin{aligned} y_1' &= a_{11}y_1 + a_{12}y_2 \\ y_2' &= a_{21}y_1 + a_{22}y_2. \end{aligned} \tag{6}$$

Of course, we can graph solutions of (6),

$$\mathbf{y}(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix}, \tag{7}$$

as two curves over the t-axis, one for each component of $\mathbf{y}(t)$. (Figure 80a in Sec. 4.1 shows an example.) But we can also graph (7) as a single curve in the y_1y_2-plane. This is a *parametric representation* (*parametric equation*) with parameter t. (See Fig. 80b for an example. Many more follow. Parametric equations also occur in calculus.) Such a curve is called a **trajectory** (or sometimes an *orbit* or *path*) of (6). The y_1y_2-plane is called the **phase plane**.[1] If we fill the phase plane with trajectories of (6), we obtain the so-called **phase portrait** of (6).

Studies of solutions in the phase plane have become quite important, along with advances in computer graphics, because a phase portrait gives a good general qualitative impression of the entire family of solutions. Consider the following example, in which we develop such a phase portrait.

EXAMPLE 1 **Trajectories in the Phase Plane (Phase Portrait)**

Find and graph solutions of the system.

In order to see what is going on, let us find and graph solutions of the system

$$\mathbf{y}' = \mathbf{A}\mathbf{y} = \begin{bmatrix} -3 & 1 \\ 1 & -3 \end{bmatrix}\mathbf{y}, \qquad \text{thus} \qquad \begin{aligned} y_1' &= -3y_1 + y_2 \\ y_2' &= y_1 - 3y_2. \end{aligned} \tag{8}$$

---

[1]A name that comes from physics, where it is the y-(mv)-plane, used to plot a motion in terms of position y and velocity $y' = v$ (m = mass); but the name is now used quite generally for the y_1y_2-plane.

The use of the phase plane is a **qualitative method**, a method of obtaining general qualitative information on solutions without actually solving an ODE or a system. This method was created by HENRI POINCARÉ (1854–1912), a great French mathematician, whose work was also fundamental in complex analysis, divergent series, topology, and astronomy.

Solution. By substituting $\mathbf{y} = \mathbf{x}e^{\lambda t}$ and $\mathbf{y}' = \lambda\mathbf{x}e^{\lambda t}$ and dropping the exponential function we get $\mathbf{Ax} = \lambda\mathbf{x}$. The characteristic equation is

$$\det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} -3-\lambda & 1 \\ 1 & -3-\lambda \end{vmatrix} = \lambda^2 + 6\lambda + 8 = 0.$$

This gives the eigenvalues $\lambda_1 = -2$ and $\lambda_2 = -4$. Eigenvectors are then obtained from

$$(-3-\lambda)x_1 + x_2 = 0.$$

For $\lambda_1 = -2$ this is $-x_1 + x_2 = 0$. Hence we can take $\mathbf{x}^{(1)} = [1 \quad 1]^{\mathsf{T}}$. For $\lambda_2 = -4$ this becomes $x_1 + x_2 = 0$, and an eigenvector is $\mathbf{x}^{(2)} = [1 \quad -1]^{\mathsf{T}}$. This gives the general solution

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = c_1\mathbf{y}^{(1)} + c_2\mathbf{y}^{(2)} = c_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{-2t} + c_2 \begin{bmatrix} 1 \\ -1 \end{bmatrix} e^{-4t}.$$

Figure 82 shows a phase portrait of some of the trajectories (to which more trajectories could be added if so desired). The two straight trajectories correspond to $c_1 = 0$ and $c_2 = 0$ and the others to other choices of c_1, c_2. ■

The method of the phase plane is particularly valuable in the frequent cases when solving an ODE or a system is inconvenient of impossible.

Critical Points of the System (6)

The point $\mathbf{y} = \mathbf{0}$ in Fig. 82 seems to be a common point of all trajectories, and we want to explore the reason for this remarkable observation. The answer will follow by calculus. Indeed, from (6) we obtain

$$\frac{dy_2}{dy_1} = \frac{y_2'\,dt}{y_1'\,dt} = \frac{y_2'}{y_1'} = \frac{a_{21}y_1 + a_{22}y_2}{a_{11}y_1 + a_{12}y_2}. \tag{9}$$

This associates with every point $P\colon (y_1, y_2)$ a unique tangent direction dy_2/dy_1 of the trajectory passing through P, except for the point $P = P_0\colon (0, 0)$, where the right side of (9) becomes $0/0$. This point P_0, at which dy_2/dy_1 becomes undetermined, is called a **critical point** of (6).

Five Types of Critical Points

There are five types of critical points depending on the geometric shape of the trajectories near them. They are called **improper nodes, proper nodes, saddle points, centers**, and **spiral points**. We define and illustrate them in Examples 1–5.

EXAMPLE 1 **(*Continued*) Improper Node (Fig. 82)**

An **improper node** is a critical point P_0 at which all the trajectories, except for two of them, have the same limiting direction of the tangent. The two exceptional trajectories also have a limiting direction of the tangent at P_0 which, however, is different.

The system (8) has an improper node at $\mathbf{0}$, as its phase portrait Fig. 82 shows. The common limiting direction at $\mathbf{0}$ is that of the eigenvector $\mathbf{x}^{(1)} = [1 \quad 1]^{\mathsf{T}}$ because e^{-4t} goes to zero faster than e^{-2t} as t increases. The two exceptional limiting tangent directions are those of $\mathbf{x}^{(2)} = [1 \quad -1]^{\mathsf{T}}$ and $-\mathbf{x}^{(2)} = [-1 \quad 1]^{\mathsf{T}}$. ■

EXAMPLE 2 Proper Node (Fig. 83)

A **proper node** is a critical point P_0 at which every trajectory has a definite limiting direction and for any given direction $\mathbf{d}$ at P_0 there is a trajectory having $\mathbf{d}$ as its limiting direction.

The system

$$\mathbf{y}' = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \mathbf{y}, \qquad \text{thus} \qquad \begin{aligned} y_1' &= y_1 \\ y_2' &= y_2 \end{aligned} \tag{10}$$

has a proper node at the origin (see Fig. 83). Indeed, the matrix is the unit matrix. Its characteristic equation $(1-\lambda)^2 = 0$ has the root $\lambda = 1$. Any $\mathbf{x} \neq \mathbf{0}$ is an eigenvector, and we can take $[1 \quad 0]^\mathsf{T}$ and $[0 \quad 1]^\mathsf{T}$. Hence a general solution is

$$\mathbf{y} = c_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} e^t + c_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} e^t \qquad \text{or} \qquad \begin{aligned} y_1 &= c_1 e^t \\ y_2 &= c_2 e^t \end{aligned} \qquad \text{or} \qquad c_1 y_2 = c_2 y_1. \qquad \blacksquare$$

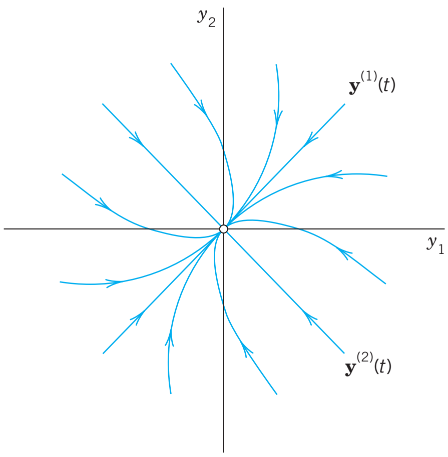


Fig. 82. Trajectories of the system (8) (Improper node)

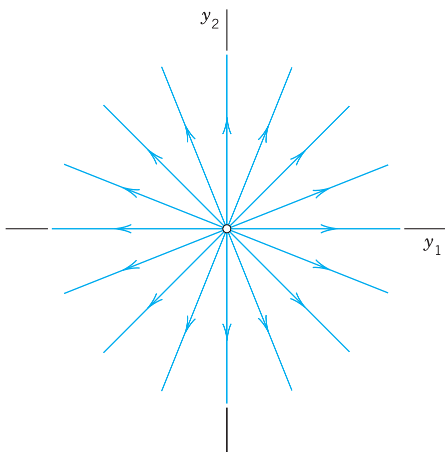


Fig. 83. Trajectories of the system (10) (Proper node)

EXAMPLE 3 Saddle Point (Fig. 84)

A **saddle point** is a critical point P_0 at which there are two incoming trajectories, two outgoing trajectories, and all the other trajectories in a neighborhood of P_0 bypass P_0.

The system

$$\mathbf{y}' = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \mathbf{y}, \qquad \text{thus} \qquad \begin{aligned} y_1' &= y_1 \\ y_1' &= -y_2 \end{aligned} \tag{11}$$

has a saddle point at the origin. Its characteristic equation $(1-\lambda)(-1-\lambda) = 0$ has the roots $\lambda_1 = 1$ and $\lambda_2 = -1$. For $\lambda = 1$ an eigenvector $[1 \quad 0]^\mathsf{T}$ is obtained from the second row of $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$, that is, $0x_1 + (-1-1)x_2 = 0$. For $\lambda_2 = -1$ the first row gives $[0 \quad 1]^\mathsf{T}$. Hence a general solution is

$$\mathbf{y} = c_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} e^t + c_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} e^{-t} \qquad \text{or} \qquad \begin{aligned} y_1 &= c_1 e^t \\ y_2 &= c_2 e^{-t} \end{aligned} \qquad \text{or} \qquad y_1 y_2 = \text{const.}$$

This is a family of hyperbolas (and the coordinate axes); see Fig. 84. $\blacksquare$

EXAMPLE 4 Center (Fig. 85)

A **center** is a critical point that is enclosed by infinitely many closed trajectories.

The system

$$\mathbf{y}' = \begin{bmatrix} 0 & 1 \\ -4 & 0 \end{bmatrix}\mathbf{y}, \qquad \text{thus} \qquad \begin{aligned} &\text{(a)} \quad y_1' = y_2 \\ &\text{(b)} \quad y_2' = -4y_1 \end{aligned} \tag{12}$$

has a center at the origin. The characteristic equation $\lambda^2 + 4 = 0$ gives the eigenvalues $2i$ and $-2i$. For $2i$ an eigenvector follows from the first equation $-2ix_1 + x_2 = 0$ of $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$, say, $[1 \quad 2i]^\mathsf{T}$. For $\lambda = -2i$ that equation is $-(-2i)x_1 + x_2 = 0$ and gives, say, $[1 \quad -2i]^\mathsf{T}$. Hence a complex general solution is

$$\mathbf{y} = c_1 \begin{bmatrix} 1 \\ 2i \end{bmatrix} e^{2it} + c_2 \begin{bmatrix} 1 \\ -2i \end{bmatrix} e^{-2it}, \qquad \text{thus} \qquad \begin{aligned} y_1 &= c_1 e^{2it} + c_2 e^{-2it} \\ y_2 &= 2ic_1 e^{2it} - 2ic_2 e^{-2it}. \end{aligned} \tag{12*}$$

A real solution is obtained from (12*) by the Euler formula or directly from (12) by a trick. (Remember the trick and call it a method when you apply it again.) Namely, the left side of (a) times the right side of (b) is $-4y_1y_1'$. This must equal the left side of (b) times the right side of (a). Thus,

$$-4y_1y_1' = y_2y_2'. \qquad \text{By integration,} \qquad 2y_1^2 + \tfrac{1}{2}y_2^2 = \text{const.}$$

This is a family of ellipses (see Fig. 85) enclosing the center at the origin. ■

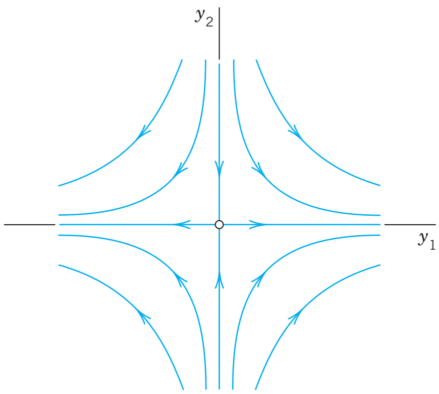


Fig. 84. Trajectories of the system (11) (Saddle point)

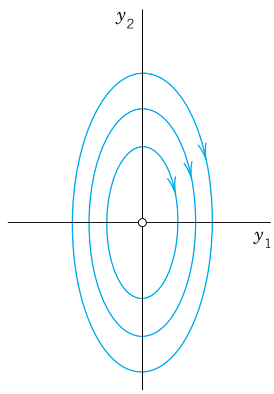


Fig. 85. Trajectories of the system (12) (Center)

EXAMPLE 5 Spiral Point (Fig. 86)

A **spiral point** is a critical point P_0 about which the trajectories spiral, approaching P_0 as $t \to \infty$ (or tracing these spirals in the opposite sense, away from P_0).

The system

$$\mathbf{y}' = \begin{bmatrix} -1 & 1 \\ -1 & -1 \end{bmatrix}\mathbf{y}, \qquad \text{thus} \qquad \begin{aligned} y_1' &= -y_1 + y_2 \\ y_2' &= -y_1 - y_2 \end{aligned} \tag{13}$$

has a spiral point at the origin, as we shall see. The characteristic equation is $\lambda^2 + 2\lambda + 2 = 0$. It gives the eigenvalues $-1 + i$ and $-1 - i$. Corresponding eigenvectors are obtained from $(-1 - \lambda)x_1 + x_2 = 0$. For

$\lambda = -1 + i$ this becomes $-ix_1 + x_2 = 0$ and we can take $[1 \quad i]^\mathsf{T}$ as an eigenvector. Similarly, an eigenvector corresponding to $-1 - i$ is $[1 \quad -i]^\mathsf{T}$. This gives the complex general solution

$$\mathbf{y} = c_1 \begin{bmatrix} 1 \\ i \end{bmatrix} e^{(-1+i)t} + c_2 \begin{bmatrix} 1 \\ -i \end{bmatrix} e^{(-1-i)t}.$$

The next step would be the transformation of this complex solution to a real general solution by the Euler formula. But, as in the last example, we just wanted to see what eigenvalues to expect in the case of a spiral point. Accordingly, we start again from the beginning and instead of that rather lengthy systematic calculation we use a shortcut. We multiply the first equation in (13) by y_1, the second by y_2, and add, obtaining

$$y_1 y_1' + y_2 y_2' = -(y_1^2 + y_2^2).$$

We now introduce polar coordinates r, t, where $r^2 = y_1^2 + y_2^2$. Differentiating this with respect to t gives $2rr' = 2y_1 y_1' + 2y_2 y_2'$. Hence the previous equation can be written

$$rr' = -r^2, \qquad \text{Thus,} \qquad r' = -r, \qquad dr/r = -dt, \qquad \ln|r| = -t + c^*, \qquad r = ce^{-t}.$$

For each real c this is a spiral, as claimed (see Fig. 86). ■

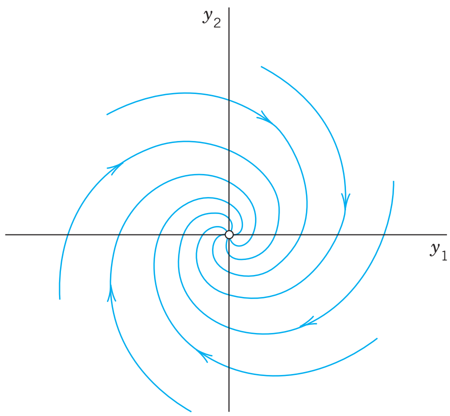


Fig. 86. Trajectories of the system (13) (Spiral point)

EXAMPLE 6 No Basis of Eigenvectors Available. Degenerate Node (Fig. 87)

This cannot happen if $\mathbf{A}$ in (1) is symmetric ($a_{kj} = a_{jk}$, as in Examples 1–3) or skew-symmetric ($a_{kj} = -a_{jk}$, thus $a_{jj} = 0$). And it does not happen in many other cases (see Examples 4 and 5). Hence it suffices to explain the method to be used by an example.

Find and graph a general solution of

$$\mathbf{y}' = \mathbf{A}\mathbf{y} = \begin{bmatrix} 4 & 1 \\ -1 & 2 \end{bmatrix} \mathbf{y}. \tag{14}$$

Solution. $\mathbf{A}$ is not skew-symmetric! Its characteristic equation is

$$\det(\mathbf{A} - \lambda \mathbf{I}) = \begin{vmatrix} 4 - \lambda & 1 \\ -1 & 2 - \lambda \end{vmatrix} = \lambda^2 - 6\lambda + 9 = (\lambda - 3)^2 = 0.$$

It has a double root $\lambda = 3$. Hence eigenvectors are obtained from $(4 - \lambda)x_1 + x_2 = 0$, thus from $x_1 + x_2 = 0$, say, $\mathbf{x}^{(1)} = [1 \quad -1]^\mathsf{T}$ and nonzero multiples of it (which do not help). The method now is to substitute

$$\mathbf{y}^{(2)} = \mathbf{x}te^{\lambda t} + \mathbf{u}e^{\lambda t}$$

with constant $\mathbf{u} = [u_1 \quad u_2]^\mathsf{T}$ into (14). (The $\mathbf{x}t$-term alone, the analog of what we did in Sec. 2.2 in the case of a double root, would not be enough. Try it.) This gives

$$\mathbf{y}^{(2)\prime} = \mathbf{x}e^{\lambda t} + \lambda\mathbf{x}te^{\lambda t} + \lambda\mathbf{u}e^{\lambda t} = \mathbf{A}\mathbf{y}^{(2)} = \mathbf{A}\mathbf{x}te^{\lambda t} + \mathbf{A}\mathbf{u}e^{\lambda t}.$$

On the right, $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$. Hence the terms $\lambda\mathbf{x}te^{\lambda t}$ cancel, and then division by $e^{\lambda t}$ gives

$$\mathbf{x} + \lambda\mathbf{u} = \mathbf{A}\mathbf{u}, \qquad \text{thus} \qquad (\mathbf{A} - \lambda\mathbf{I})\mathbf{u} = \mathbf{x}.$$

Here $\lambda = 3$ and $\mathbf{x} = [1 \quad -1]^\mathsf{T}$, so that

$$(\mathbf{A} - 3\mathbf{I})\mathbf{u} = \begin{bmatrix} 4-3 & 1 \\ -1 & 2-3 \end{bmatrix}\mathbf{u} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \qquad \text{thus} \qquad \begin{aligned} u_1 + u_2 &= 1 \\ -u_1 - u_2 &= -1. \end{aligned}$$

A solution, linearly independent of $\mathbf{x} = [1 \quad -1]^\mathsf{T}$, is $\mathbf{u} = [0 \quad 1]^\mathsf{T}$. This yields the answer (Fig. 87)

$$\mathbf{y} = c_1\mathbf{y}^{(1)} + c_2\mathbf{y}^{(2)} = c_1\begin{bmatrix} 1 \\ -1 \end{bmatrix}e^{3t} + c_2\left(\begin{bmatrix} 1 \\ -1 \end{bmatrix}t + \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right)e^{3t}.$$

The critical point at the origin is often called a **degenerate node**. $c_1\mathbf{y}^{(1)}$ gives the heavy straight line, with $c_1 > 0$ the lower part and $c_1 < 0$ the upper part of it. $\mathbf{y}^{(2)}$ gives the right part of the heavy curve from 0 through the second, first, and—finally—fourth quadrants. $-\mathbf{y}^{(2)}$ gives the other part of that curve. ■

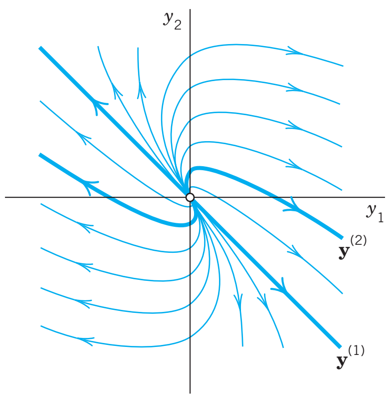


Fig. 87. Degenerate node in Example 6

We mention that for a system (1) with three or more equations and a triple eigenvalue with only one linearly independent eigenvector, one will get two solutions, as just discussed, and a third linearly independent one from

$$\mathbf{y}^{(3)} = \tfrac{1}{2}\mathbf{x}t^2e^{\lambda t} + \mathbf{u}te^{\lambda t} + \mathbf{v}e^{\lambda t} \qquad \text{with } \mathbf{v} \text{ from} \qquad \mathbf{u} + \lambda\mathbf{v} = \mathbf{A}\mathbf{v}.$$

PROBLEM SET 4.3

1–9 GENERAL SOLUTION

Find a real general solution of the following systems. Show the details.

1. $y_1' = y_1 + y_2$
$y_2' = 3y_1 - y_2$

2. $y_1' = 6y_1 + 9y_2$
$y_2' = y_1 + 6y_2$

3. $y_1' = y_1 + 2y_2$
$y_2' = \frac{1}{2}y_1 + y_2$

4. $y_1' = -8y_1 - 2y_2$
$y_2' = 2y_1 - 4y_2$

5. $y_1' = 2y_1 + 5y_2$
$y_2' = 5y_1 + 12.5y_2$

6. $y_1' = 2y_1 - 2y_2$
$y_2' = 2y_1 + 2y_2$

7. $y_1' = y_2$
$y_2' = -y_1 + y_3$
$y_3' = -y_2$

8. $y_1' = 8y_1 - y_2$
$y_2' = y_1 + 10y_2$

9. $y_1' = 10y_1 - 10y_2 - 4y_3$
$y_2' = -10y_1 + y_2 - 14y_3$
$y_3' = -4y_1 - 14y_2 - 2y_3$

10–15 IVPs

Solve the following initial value problems.

10. $y_1' = 2y_1 + 2y_2$
$y_2' = 5y_1 - y_2$
$y_1(0) = 0, \quad y_2(0) = 7$

11. $y_1' = 2y_1 + 5y_2$
$y_2' = -\frac{1}{2}y_1 - \frac{3}{2}y_2$
$y_1(0) = -12, \quad y_2(0) = 0$

12. $y_1' = y_1 + 3y_2$
$y_2' = \frac{1}{3}y_1 + y_2$
$y_1(0) = 12, \quad y_2(0) = 2$

13. $y_1' = y_2$
$y_2' = y_1$
$y_1(0) = 0, \quad y_2(0) = 2$

14. $y_1' = -y_1 - y_2$
$y_2' = y_1 - y_2$
$y_1(0) = 1, \quad y_2(0) = 0$

15. $y_1' = 3y_1 + 2y_2$
$y_2' = 2y_1 + 3y_2$
$y_1(0) = 0.5, \quad y_2(0) = -0.5$

16–17 CONVERSION

Find a general solution by conversion to a single ODE.

16. The system in Prob. 8.

17. The system in Example 5 of the text.

18. **Mixing problem, Fig. 88.** Each of the two tanks contains 200 gal of water, in which initially 100 lb (Tank T_1) and 200 lb (Tank T_2) of fertilizer are dissolved. The inflow, circulation, and outflow are shown in Fig. 88. The mixture is kept uniform by stirring. Find the fertilizer contents $y_1(t)$ in T_1 and $y_2(t)$ in T_2.

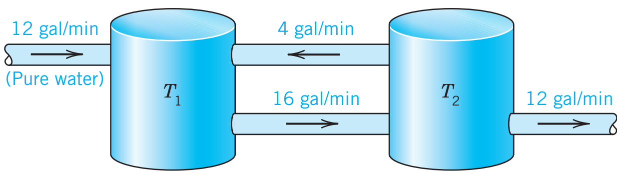


Fig. 88. Tanks in Problem 18

19. **Network.** Show that a model for the currents $I_1(t)$ and $I_2(t)$ in Fig. 89 is

$$\frac{1}{C}\int I_1\,dt + R(I_1 - I_2) = 0, \quad LI_2' + R(I_2 - I_1) = 0.$$

Find a general solution, assuming that $R = 3\ \Omega$, $L = 4$ H, $C = 1/12$ F.

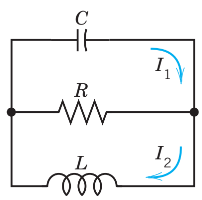


Fig. 89. Network in Problem 19

20. **CAS PROJECT. Phase Portraits.** Graph some of the figures in this section, in particular Fig. 87 on the degenerate node, in which the vector $\mathbf{y}^{(2)}$ depends on t. In each figure highlight a trajectory that satisfies an initial condition of your choice.

4.4 Criteria for Critical Points. Stability

We continue our discussion of homogeneous linear systems with constant coefficients (1). Let us review where we are. From Sec. 4.3 we have

$$\text{(1)}\qquad \mathbf{y}' = \mathbf{A}\mathbf{y} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\mathbf{y}, \qquad \text{in components,} \qquad \begin{aligned} y_1' &= a_{11}y_1 + a_{12}y_2 \\ y_2' &= a_{21}y_1 + a_{22}y_2. \end{aligned}$$

From the examples in the last section, we have seen that we can obtain an overview of families of solution curves if we represent them parametrically as $\mathbf{y}(t) = [y_1(t) \quad y_2(t)]^{\mathsf{T}}$ and graph them as curves in the y_1y_2-plane, called the **phase plane**. Such a curve is called a **trajectory** of (1), and their totality is known as the **phase portrait** of (1).

Now we have seen that solutions are of the form

$$\mathbf{y}(t) = \mathbf{x}e^{\lambda t}. \qquad \text{Substitution into (1) gives} \qquad \mathbf{y}'(t) = \lambda\mathbf{x}e^{\lambda t} = \mathbf{A}\mathbf{y} = \mathbf{A}\mathbf{x}e^{\lambda t}.$$

Dropping the common factor $e^{\lambda t}$, we have

$$\textbf{(2)}\qquad \mathbf{A}\mathbf{x} = \lambda\mathbf{x}.$$

Hence $\mathbf{y}(t)$ is a (nonzero) solution of (1) if λ is an eigenvalue of $\mathbf{A}$ and $\mathbf{x}$ a corresponding eigenvector.

Our examples in the last section show that the general form of the phase portrait is determined to a large extent by the type of **critical point** of the system (1) defined as a point at which dy_2/dy_1 becomes undetermined, $0/0$; here [see (9) in Sec. 4.3]

$$\textbf{(3)}\qquad \frac{dy_2}{dy_1} = \frac{y_2'\,dt}{y_1'\,dt} = \frac{a_{21}y_1 + a_{22}y_2}{a_{11}y_1 + a_{12}y_2}.$$

We also recall from Sec. 4.3 that there are various types of critical points.

What is now new, is that we shall see how these types of critical points are related to the eigenvalues. The latter are solutions $\lambda = \lambda_1$ and λ_2 of the characteristic equation

$$\text{(4)}\qquad \det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{vmatrix} = \lambda^2 - (a_{11} + a_{22})\lambda + \det\mathbf{A} = 0.$$

This is a quadratic equation $\lambda^2 - p\lambda + q = 0$ with coefficients p, q and discriminant Δ given by

$$\text{(5)}\qquad p = a_{11} + a_{22}, \qquad q = \det\mathbf{A} = a_{11}a_{22} - a_{12}a_{21}, \qquad \Delta = p^2 - 4q.$$

From algebra we know that the solutions of this equation are

$$\text{(6)}\qquad \lambda_1 = \tfrac{1}{2}(p + \sqrt{\Delta}), \qquad \lambda_2 = \tfrac{1}{2}(p - \sqrt{\Delta}).$$

Furthermore, the product representation of the equation gives

$$\lambda^2 - p\lambda + q = (\lambda - \lambda_1)(\lambda - \lambda_2) = \lambda^2 - (\lambda_1 + \lambda_2)\lambda + \lambda_1\lambda_2.$$

Hence p is the sum and q the product of the eigenvalues. Also $\lambda_1 - \lambda_2 = \sqrt{\Delta}$ from (6). Together,

$$p = \lambda_1 + \lambda_2, \qquad q = \lambda_1\lambda_2, \qquad \Delta = (\lambda_1 - \lambda_2)^2. \tag{7}$$

This gives the criteria in Table 4.1 for classifying critical points. A derivation will be indicated later in this section.

Table 4.1 Eigenvalue Criteria for Critical Points (Derivation after Table 4.2)

Name	$p = \lambda_1 + \lambda_2$	$q = \lambda_1\lambda_2$	$\Delta = (\lambda_1 - \lambda_2)^2$	Comments on λ_1, λ_2
(a) Node		$q > 0$	$\Delta \geqq 0$	Real, same sign
(b) Saddle point		$q < 0$		Real, opposite signs
(c) Center	$p = 0$	$q > 0$		Pure imaginary
(d) Spiral point	$p \neq 0$		$\Delta < 0$	Complex, not pure imaginary

Stability

Critical points may also be classified in terms of their stability. Stability concepts are basic in engineering and other applications. They are suggested by physics, where **stability** means, roughly speaking, that a small change (a small disturbance) of a physical system at some instant changes the behavior of the system only slightly at all future times t. For critical points, the following concepts are appropriate.

DEFINITIONS

Stable, Unstable, Stable and Attractive

A critical point P_0 of (1) is called **stable**[2] if, roughly, all trajectories of (1) that at some instant are close to P_0 remain close to P_0 at all future times; precisely: if for every disk D_ϵ of radius $\epsilon > 0$ with center P_0 there is a disk D_δ of radius $\delta > 0$ with center P_0 such that every trajectory of (1) that has a point P_1 (corresponding to $t = t_1$, say) in D_δ has all its points corresponding to $t \geqq t_1$ in D_ϵ. See Fig. 90.

P_0 is called **unstable** if P_0 is not stable.

P_0 is called **stable and attractive** (or *asymptotically stable*) if P_0 is stable and every trajectory that has a point in D_δ approaches P_0 as $t \to \infty$. See Fig. 91.

Classification criteria for critical points in terms of stability are given in Table 4.2. Both tables are summarized in the **stability chart** in Fig. 92. In this chart region of instability is dark blue.

[2]In the sense of the Russian mathematician ALEXANDER MICHAILOVICH LJAPUNOV (1857–1918), whose work was fundamental in stability theory for ODEs. This is perhaps the most appropriate definition of stability (and the only we shall use), but there are others, too.

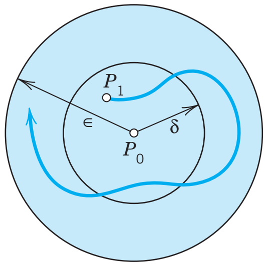


Fig. 90. Stable critical point P_0 of (1) (The trajectory initiating at P_1 stays in the disk of radius ϵ.)

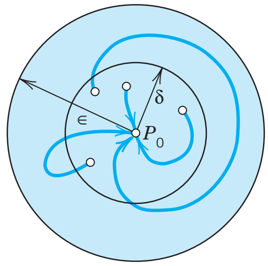


Fig. 91. Stable and attractive critical point P_0 of (1)

Table 4.2 Stability Criteria for Critical Points

Type of Stability	$p = \lambda_1 + \lambda_2$		$q = \lambda_1 \lambda_2$
(a) Stable and attractive	$p < 0$		$q > 0$
(b) Stable	$p \leqq 0$		$q > 0$
(c) Unstable	$p > 0$	OR	$q < 0$

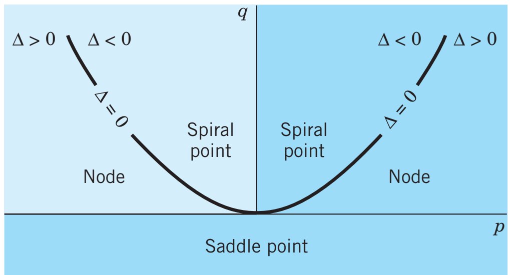


Fig. 92. Stability chart of the system (1) with p, q, Δ defined in (5). Stable and attractive: The second quadrant without the q-axis. Stability also on the positive q-axis (which corresponds to centers). Unstable: Dark blue region

We indicate how the criteria in Tables 4.1 and 4.2 are obtained. If $q = \lambda_1 \lambda_2 > 0$, both of the eigenvalues are positive or both are negative or complex conjugates. If also $p = \lambda_1 + \lambda_2 < 0$, both are negative or have a negative real part. Hence P_0 is stable and attractive. The reasoning for the other two lines in Table 4.2 is similar.

If $\Delta < 0$, the eigenvalues are complex conjugates, say, $\lambda_1 = \alpha + i\beta$ and $\lambda_2 = \alpha - i\beta$. If also $p = \lambda_1 + \lambda_2 = 2\alpha < 0$, this gives a spiral point that is stable and attractive. If $p = 2\alpha > 0$, this gives an unstable spiral point.

If $p = 0$, then $\lambda_2 = -\lambda_1$ and $q = \lambda_1 \lambda_2 = -\lambda_1^2$. If also $q > 0$, then $\lambda_1^2 = -q < 0$, so that λ_1, and thus λ_2, must be pure imaginary. This gives periodic solutions, their trajectories being closed curves around P_0, which is a center.

EXAMPLE 1 Application of the Criteria in Tables 4.1 and 4.2

In Example 1, Sec 4.3, we have $\mathbf{y}' = \begin{bmatrix} -3 & 1 \\ 1 & -3 \end{bmatrix} \mathbf{y}$, $p = -6$, $q = 8$, $\Delta = 4$, a node by Table 4.1(a), which is stable and attractive by Table 4.2(a). ■

EXAMPLE 2 Free Motions of a Mass on a Spring

What kind of critical point does $my'' + cy' + ky = 0$ in Sec. 2.4 have?

Solution. Division by m gives $y'' = -(k/m)y - (c/m)y'$. To get a system, set $y_1 = y$, $y_2 = y'$ (see Sec. 4.1). Then $y_2' = y'' = -(k/m)y_1 - (c/m)y_2$. Hence

$$\mathbf{y}' = \begin{bmatrix} 0 & 1 \\ -k/m & -c/m \end{bmatrix} \mathbf{y}, \qquad \det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} -\lambda & 1 \\ -k/m & -c/m - \lambda \end{vmatrix} = \lambda^2 + \frac{c}{m}\lambda + \frac{k}{m} = 0.$$

We see that $p = -c/m$, $q = k/m$, $\Delta = (c/m)^2 - 4k/m$. From this and Tables 4.1 and 4.2 we obtain the following results. Note that in the last three cases the discriminant Δ plays an essential role.

No damping. $c = 0$, $p = 0$, $q > 0$, a center.
Underdamping. $c^2 < 4mk$, $p < 0$, $q > 0$, $\Delta < 0$, a stable and attractive spiral point.
Critical damping. $c^2 = 4mk$, $p < 0$, $q > 0$, $\Delta = 0$, a stable and attractive node.
Overdamping. $c^2 > 4mk$, $p < 0$, $q > 0$, $\Delta > 0$, a stable and attractive node. ■

PROBLEM SET 4.4

1–10 TYPE AND STABILITY OF CRITICAL POINT

Determine the type and stability of the critical point. Then find a real general solution and sketch or graph some of the trajectories in the phase plane. Show the details of your work.

1. $y_1' = y_1$, $y_2' = 2y_2$

2. $y_1' = -4y_1$, $y_2' = -3y_2$

3. $y_1' = y_2$, $y_2' = -9y_1$

4. $y_1' = 2y_1 + y_2$, $y_2' = 5y_1 - 2y_2$

5. $y_1' = -2y_1 + 2y_2$, $y_2' = -2y_1 - 2y_2$

6. $y_1' = -6y_1 - y_2$, $y_2' = -9y_1 - 6y_2$

7. $y_1' = y_1 + 2y_2$, $y_2' = 2y_1 + y_2$

8. $y_1' = -y_1 + 4y_2$, $y_2' = 3y_1 - 2y_2$

9. $y_1' = 4y_1 + y_2$, $y_2' = 4y_1 + 4y_2$

10. $y_1' = y_2$, $y_2' = -5y_1 - 2y_2$

11–18 TRAJECTORIES OF SYSTEMS AND SECOND-ORDER ODEs. CRITICAL POINTS

11. Damped oscillations. Solve $y'' + 2y' + 2y = 0$. What kind of curves are the trajectories?

12. Harmonic oscillations. Solve $y'' + \frac{1}{9}y = 0$. Find the trajectories. Sketch or graph some of them.

13. Types of critical points. Discuss the critical points in (10)–(13) of Sec. 4.3 by using Tables 4.1 and 4.2.

14. Transformation of parameter. What happens to the critical point in Example 1 if you introduce $\tau = -t$ as a new independent variable?

15. Perturbation of center. What happens in Example 4 of Sec. 4.3 if you change $\mathbf{A}$ to $\mathbf{A} + 0.1\mathbf{I}$, where $\mathbf{I}$ is the unit matrix?

16. Perturbation of center. If a system has a center as its critical point, what happens if you replace the matrix $\mathbf{A}$ by $\tilde{\mathbf{A}} = \mathbf{A} + k\mathbf{I}$ with any real number $k \neq 0$ (representing measurement errors in the diagonal entries)?

17. Perturbation. The system in Example 4 in Sec. 4.3 has a center as its critical point. Replace each a_{jk} in Example 4, Sec. 4.3, by $a_{jk} + b$. Find values of b such that you get **(a)** a saddle point, **(b)** a stable and attractive node, **(c)** a stable and attractive spiral, **(d)** an unstable spiral, **(e)** an unstable node.

18. CAS EXPERIMENT. Phase Portraits. Graph phase portraits for the systems in Prob. 17 with the values of b suggested in the answer. Try to illustrate how the phase portrait changes "continuously" under a continuous change of b.

19. WRITING PROBLEM. Stability. Stability concepts are basic in physics and engineering. Write a two-part report of 3 pages each (A) on general applications in which stability plays a role (be as precise as you can), and (B) on material related to stability in this section. Use your own formulations and examples; do not copy.

20. Stability chart. Locate the critical points of the systems (10)–(14) in Sec. 4.3 and of Probs. 1, 3, 5 in this problem set on the stability chart.

4.5 Qualitative Methods for Nonlinear Systems

Qualitative methods are methods of obtaining qualitative information on solutions without actually solving a system. These methods are particularly valuable for systems whose solution by analytic methods is difficult or impossible. This is the case for many practically important **nonlinear systems**

$$\mathbf{y}' = \mathbf{f}(\mathbf{y}), \qquad \text{thus} \qquad \begin{aligned} y_1' &= f_1(y_1, y_2) \\ y_2' &= f_2(y_1, y_2). \end{aligned} \tag{1}$$

In this section we extend **phase plane methods**, as just discussed, from linear systems to nonlinear systems (1). We assume that (1) is **autonomous**, that is, the independent variable t does not occur explicitly. (All examples in the last section are autonomous.) We shall again exhibit entire families of solutions. This is an advantage over numeric methods, which give only one (approximate) solution at a time.

Concepts needed from the last section are the **phase plane** (the $y_1 y_2$-plane), **trajectories** (solution curves of (1) in the phase plane), the **phase portrait** of (1) (the totality of these trajectories), and **critical points** of (1) (points (y_1, y_2) at which both $f_1(y_1, y_2)$ and $f_2(y_1, y_2)$ are zero).

Now (1) may have several critical points. Our approach shall be to discuss one critical point after another. If a critical point P_0 is not at the origin, then, for technical convenience, we shall move this point to the origin before analyzing the point. More formally, if P_0: (a, b) is a critical point with (a, b) not at the origin $(0, 0)$, then we apply the translation

$$\tilde{y}_1 = y_1 - a, \qquad \tilde{y}_2 = y_2 - b$$

which moves P_0 to $(0, 0)$ as desired. Thus we can assume P_0 to be the origin $(0, 0)$, and for simplicity we continue to write y_1, y_2 (instead of $\tilde{y}_1, \tilde{y}_2$). We also assume that P_0 is **isolated**, that is, it is the only critical point of (1) within a (sufficiently small) disk with center at the origin. If (1) has only finitely many critical points, that is automatically true. (Explain!)

Linearization of Nonlinear Systems

How can we determine the kind and stability property of a critical point P_0: $(0, 0)$ of (1)? In most cases this can be done by **linearization** of (1) near P_0, writing (1) as $\mathbf{y}' = \mathbf{f}(\mathbf{y}) = \mathbf{A}\mathbf{y} + \mathbf{h}(\mathbf{y})$ and dropping $\mathbf{h}(\mathbf{y})$, as follows.

Since P_0 is critical, $f_1(0, 0) = 0$, $f_2(0, 0) = 0$, so that f_1 and f_2 have no constant terms and we can write

$$\mathbf{y}' = \mathbf{A}\mathbf{y} + \mathbf{h}(\mathbf{y}), \qquad \text{thus} \qquad \begin{aligned} y_1' &= a_{11} y_1 + a_{12} y_2 + h_1(y_1, y_2) \\ y_2' &= a_{21} y_1 + a_{22} y_2 + h_2(y_1, y_2). \end{aligned} \tag{2}$$

$\mathbf{A}$ is constant (independent of t) since (1) is autonomous. One can prove the following (proof in Ref. [A7], pp. 375–388, listed in App. 1).

THEOREM 1

Linearization

If f_1 and f_2 in (1) *are continuous and have continuous partial derivatives in a neighborhood of the critical point P_0:* (0, 0), *and if* $\det \mathbf{A} \neq 0$ *in* (2), *then the kind and stability of the critical point of* (1) *are the same as those of the* **linearized system**

$$\mathbf{y}' = \mathbf{A}\mathbf{y}, \qquad \text{thus} \qquad \begin{aligned} y_1' &= a_{11}y_1 + a_{12}y_2 \\ y_2' &= a_{21}y_1 + a_{22}y_2. \end{aligned} \tag{3}$$

Exceptions occur if $\mathbf{A}$ *has equal or pure imaginary eigenvalues; then* (1) *may have the same kind of critical point as* (3) *or a spiral point.*

EXAMPLE 1

Free Undamped Pendulum. Linearization

Figure 93a shows a pendulum consisting of a body of mass m (the bob) and a rod of length L. Determine the locations and types of the critical points. Assume that the mass of the rod and air resistance are negligible.

Solution. ***Step 1. Setting up the mathematical model.*** Let θ denote the angular displacement, measured counterclockwise from the equilibrium position. The weight of the bob is mg (g the acceleration of gravity). It causes a restoring force $mg \sin\theta$ tangent to the curve of motion (circular arc) of the bob. By Newton's second law, at each instant this force is balanced by the force of acceleration $mL\theta''$, where $L\theta''$ is the acceleration; hence the resultant of these two forces is zero, and we obtain as the mathematical model

$$mL\theta'' + mg \sin\theta = 0.$$

Dividing this by mL, we have

$$\theta'' + k \sin\theta = 0 \qquad \left(k = \frac{g}{L}\right). \tag{4}$$

When θ is very small, we can approximate $\sin\theta$ rather accurately by θ and obtain as an *approximate* solution $A \cos\sqrt{k}t + B \sin\sqrt{k}t$, but the *exact* solution for any θ is not an elementary function.

Step 2. Critical points $(\mathbf{0}, \mathbf{0}), (\pm 2\pi, \mathbf{0}), (\pm 4\pi, \mathbf{0}), \cdots$, ***Linearization.*** To obtain a system of ODEs, we set $\theta = y_1$, $\theta' = y_2$. Then from (4) we obtain a nonlinear system (1) of the form

$$\begin{aligned} y_1' &= f_1(y_1, y_2) = y_2 \\ y_2' &= f_2(y_1, y_2) = -k \sin y_1. \end{aligned} \tag{4*}$$

The right sides are both zero when $y_2 = 0$ and $\sin y_1 = 0$. This gives infinitely many critical points $(n\pi, 0)$, where $n = 0, \pm 1, \pm 2, \cdots$. We consider (0, 0). Since the Maclaurin series is

$$\sin y_1 = y_1 - \tfrac{1}{6}y_1^3 + - \cdots \approx y_1,$$

the linearized system at (0, 0) is

$$\mathbf{y}' = \mathbf{A}\mathbf{y} = \begin{bmatrix} 0 & 1 \\ -k & 0 \end{bmatrix} \mathbf{y}, \qquad \text{thus} \qquad \begin{aligned} y_1' &= y_2 \\ y_2' &= -ky_1. \end{aligned}$$

To apply our criteria in Sec. 4.4 we calculate $p = a_{11} + a_{22} = 0$, $q = \det \mathbf{A} = k = g/L\,(>0)$, and $\Delta = p^2 - 4q = -4k$. From this and Table 4.1(c) in Sec. 4.4 we conclude that (0, 0) is a center, which is always stable. Since $\sin\theta = \sin y_1$ is periodic with period 2π, the critical points $(n\pi, 0)$, $n = \pm 2, \pm 4, \cdots$, are all centers.

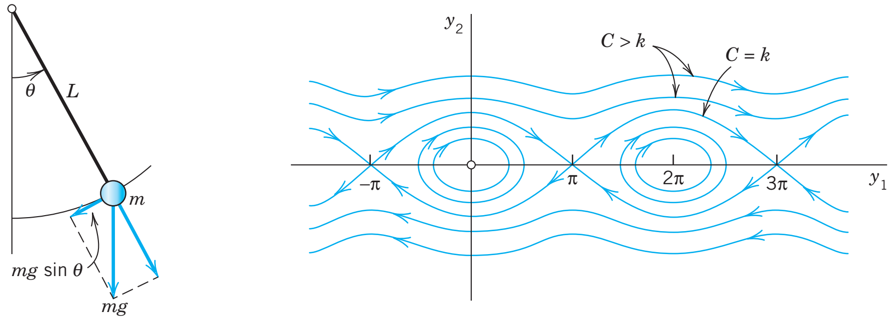
Step 3. Critical points $(\pm\pi, \mathbf{0}), (\pm 3\pi, \mathbf{0}), (\pm 5\pi, \mathbf{0}), \cdots$, ***Linearization.*** We now consider the critical point $(\pi, 0)$, setting $\theta - \pi = y_1$ and $(\theta - \pi)' = \theta' = y_2$. Then in (4),

$$\sin\theta = \sin(y_1 + \pi) = -\sin y_1 = -y_1 + \tfrac{1}{6}y_1^3 - + \cdots \approx -y_1$$

and the linearized system at $(\pi, 0)$ is now

$$\mathbf{y}' = \mathbf{A}\mathbf{y} = \begin{bmatrix} 0 & 1 \\ k & 0 \end{bmatrix} \mathbf{y}, \qquad \text{thus} \qquad \begin{aligned} y_1' &= y_2 \\ y_2' &= k y_1. \end{aligned}$$

We see that $p = 0$, $q = -k\ (<0)$, and $\Delta = -4q = 4k$. Hence, by Table 4.1(b), this gives a saddle point, which is always unstable. Because of periodicity, the critical points $(n\pi, 0)$, $n = \pm 1, \pm 3, \cdots$, are all saddle points. These results agree with the impression we get from Fig. 93b. ■



(a) Pendulum

(b) Solution curves $y_2(y_1)$ of (4) in the phase plane

Fig. 93. Example 1 (C will be explained in Example 4.)

EXAMPLE 2 Linearization of the Damped Pendulum Equation

To gain further experience in investigating critical points, as another practically important case, let us see how Example 1 changes when we add a damping term $c\theta'$ (damping proportional to the angular velocity) to equation (4), so that it becomes

$$\boxed{\theta'' + c\theta' + k \sin\theta = 0} \tag{5}$$

where $k > 0$ and $c \geqq 0$ (which includes our previous case of no damping, $c = 0$). Setting $\theta = y_1$, $\theta' = y_2$, as before, we obtain the nonlinear system (use $\theta'' = y_2'$)

$$\begin{aligned} y_1' &= y_2 \\ y_2' &= -k \sin y_1 - c y_2. \end{aligned}$$

We see that the critical points have the same locations as before, namely, $(0, 0)$, $(\pm\pi, 0)$, $(\pm 2\pi, 0), \cdots$. We consider $(0, 0)$. Linearizing $\sin y_1 \approx y_1$ as in Example 1, we get the linearized system at $(0, 0)$

$$\mathbf{y}' = \mathbf{A}\mathbf{y} = \begin{bmatrix} 0 & 1 \\ -k & -c \end{bmatrix} \mathbf{y}, \qquad \text{thus} \qquad \begin{aligned} y_1' &= y_2 \\ y_2' &= -k y_1 - c y_2. \end{aligned} \tag{6}$$

This is identical with the system in Example 2 of Sec. 4.4, except for the (positive!) factor m (and except for the physical meaning of y_1). Hence for $c = 0$ (no damping) we have a center (see Fig. 93b), for small damping we have a spiral point (see Fig. 94), and so on.

We now consider the critical point $(\pi, 0)$. We set $\theta - \pi = y_1$, $(\theta - \pi)' = \theta' = y_2$ and linearize

$$\sin\theta = \sin(y_1 + \pi) = -\sin y_1 \approx -y_1.$$

This gives the new linearized system at $(\pi, 0)$

$$\mathbf{y}' = \mathbf{A}\mathbf{y} = \begin{bmatrix} 0 & 1 \\ k & -c \end{bmatrix} \mathbf{y}, \qquad \text{thus} \qquad \begin{aligned} y_1' &= y_2 \\ y_2' &= k y_1 - c y_2. \end{aligned} \tag{6*}$$

For our criteria in Sec. 4.4 we calculate $p = a_{11} + a_{22} = -c$, $q = \det \mathbf{A} = -k$, and $\Delta = p^2 - 4q = c^2 + 4k$. This gives the following results for the critical point at $(\pi, 0)$.

No damping. $c = 0, p = 0, q < 0, \Delta > 0$, a saddle point. See Fig. 93b.
Damping. $c > 0, p < 0, q < 0, \Delta > 0$, a saddle point. See Fig. 94.

Since $\sin y_1$ is periodic with period 2π, the critical points $(\pm 2\pi, 0), (\pm 4\pi, 0), \cdots$ are of the same type as $(0, 0)$, and the critical points $(-\pi, 0), (\pm 3\pi, 0), \cdots$ are of the same type as $(\pi, 0)$, so that our task is finished.

Figure 94 shows the trajectories in the case of damping. What we see agrees with our physical intuition. Indeed, damping means loss of energy. Hence instead of the closed trajectories of periodic solutions in Fig. 93b we now have trajectories spiraling around one of the critical points $(0, 0), (\pm 2\pi, 0), \cdots$. Even the wavy trajectories corresponding to whirly motions eventually spiral around one of these points. Furthermore, there are no more trajectories that connect critical points (as there were in the undamped case for the saddle points). ■

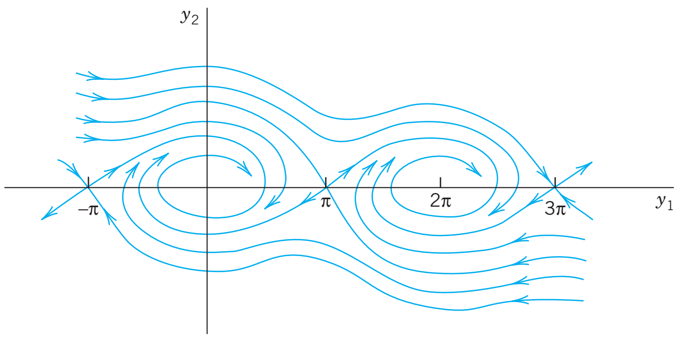


Fig. 94. Trajectories in the phase plane for the damped pendulum in Example 2

Lotka–Volterra Population Model

EXAMPLE 3 **Predator–Prey Population Model**[3]

This model concerns two species, say, rabbits and foxes, and the foxes prey on the rabbits.

Step 1. Setting up the model. We assume the following.

1. Rabbits have unlimited food supply. Hence, if there were no foxes, their number $y_1(t)$ would grow exponentially, $y_1' = ay_1$.
2. Actually, y_1 is decreased because of the kill by foxes, say, at a rate proportional to $y_1 y_2$, where $y_2(t)$ is the number of foxes. Hence $y_1' = ay_1 - by_1 y_2$, where $a > 0$ and $b > 0$.
3. If there were no rabbits, then $y_2(t)$ would exponentially decrease to zero, $y_2' = -ly_2$. However, y_2 is increased by a rate proportional to the number of encounters between predator and prey; together we have $y_2' = -ly_2 + ky_1 y_2$, where $k > 0$ and $l > 0$.

This gives the (nonlinear!) Lotka–Volterra system

$$\begin{aligned} y_1' &= f_1(y_1, y_2) = ay_1 - by_1 y_2 \\ y_2' &= f_2(y_1, y_2) = ky_1 y_2 - ly_2. \end{aligned} \tag{7}$$

[3]Introduced by ALFRED J. LOTKA (1880–1949), American biophysicist, and VITO VOLTERRA (1860–1940), Italian mathematician, the initiator of functional analysis (see [GR7] in App. 1).

Step 2. Critical point* (0, 0)*, Linearization. We see from (7) that the critical points are the solutions of

$$f_1(y_1, y_2) = y_1(a - by_2) = 0, \qquad f_2(y_1, y_2) = y_2(ky_1 - l) = 0. \tag{7*}$$

The solutions are $(y_1, y_2) = (0, 0)$ and $\left(\frac{l}{k}, \frac{a}{b}\right)$. We consider (0, 0). Dropping $-by_1y_2$ and ky_1y_2 from (7) gives the linearized system

$$\mathbf{y}' = \begin{bmatrix} a & 0 \\ 0 & -l \end{bmatrix} \mathbf{y}.$$

Its eigenvalues are $\lambda_1 = a > 0$ and $\lambda_2 = -l < 0$. They have opposite signs, so that we get a saddle point.

Step. 3. Critical point (l/k, a/b), Linearization. We set $y_1 = \tilde{y}_1 + l/k$, $y_2 = \tilde{y}_2 + a/b$. Then the critical point $(l/k, a/b)$ corresponds to $(\tilde{y}_1, \tilde{y}_2) = (0, 0)$. Since $\tilde{y}_1' = y_1'$, $\tilde{y}_2' = y_2'$, we obtain from (7) [factorized as in (7*)]

$$\tilde{y}_1' = \left(\tilde{y}_1 + \frac{l}{k}\right)\left[a - b\left(\tilde{y}_2 + \frac{a}{b}\right)\right] = \left(\tilde{y}_1 + \frac{l}{k}\right)(-b\tilde{y}_2)$$

$$\tilde{y}_2' = \left(\tilde{y}_2 + \frac{a}{b}\right)\left[k\left(\tilde{y}_1 + \frac{l}{k}\right) - l\right] = \left(\tilde{y}_2 + \frac{a}{b}\right)k\tilde{y}_1.$$

Dropping the two nonlinear terms $-b\tilde{y}_1\tilde{y}_2$ and $k\tilde{y}_1\tilde{y}_2$, we have the linearized system

$$\text{(7**)} \qquad \begin{aligned} &\text{(a)} \quad \tilde{y}_1' = -\frac{lb}{k}\tilde{y}_2 \\ &\text{(b)} \quad \tilde{y}_2' = \frac{ak}{b}\tilde{y}_1. \end{aligned}$$

The left side of (a) times the right side of (b) must equal the right side of (a) times the left side of (b),

$$\frac{ak}{b}\tilde{y}_1\tilde{y}_1' = -\frac{lb}{k}\tilde{y}_2\tilde{y}_2'. \qquad \text{By integration,} \qquad \frac{ak}{b}\tilde{y}_1^2 + \frac{lb}{k}\tilde{y}_2^2 = \text{const.}$$

This is a family of ellipses, so that the critical point $(l/k, a/b)$ of the linearized system (7**) is a center (Fig. 95). It can be shown, by a complicated analysis, that the nonlinear system (7) also has a center (rather than a spiral point) at $(l/k, a/b)$ surrounded by closed trajectories (not ellipses).

We see that the predators and prey have a cyclic variation about the critical point. Let us move counterclockwise around the ellipse, beginning at the right vertex, where the rabbits have a maximum number. Foxes are sharply increasing in number until they reach a maximum at the upper vertex, and the number of rabbits is then sharply decreasing until it reaches a minimum at the left vertex, and so on. Cyclic variations of this kind have been observed in nature, for example, for lynx and snowshoe hare near the Hudson Bay, with a cycle of about 10 years.

For models of more complicated situations and a systematic discussion, see C. W. Clark, *Mathematical Bioeconomics: The Mathematics of Conservation,* 3rd ed. Hoboken, NJ, Wiley, 2010. ■

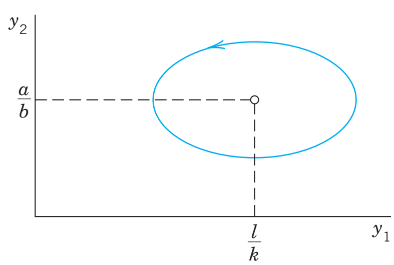


Fig. 95. Ecological equilibrium point and trajectory of the linearized Lotka–Volterra system (7**)

Transformation to a First-Order Equation in the Phase Plane

Another phase plane method is based on the idea of transforming a second-order **autonomous ODE** (an ODE in which t does not occur explicitly)

$$F(y, y', y'') = 0$$

to first order by taking $y = y_1$ as the independent variable, setting $y' = y_2$ and transforming y'' by the chain rule,

$$y'' = y_2' = \frac{dy_2}{dt} = \frac{dy_2}{dy_1}\frac{dy_1}{dt} = \frac{dy_2}{dy_1} y_2.$$

Then the ODE becomes of first order,

$$\boxed{F\left(y_1, y_2, \frac{dy_2}{dy_1} y_2\right) = 0} \tag{8}$$

and can sometimes be solved or treated by direction fields. We illustrate this for the equation in Example 1 and shall gain much more insight into the behavior of solutions.

EXAMPLE 4 An ODE (8) for the Free Undamped Pendulum

If in (4) $\theta'' + k \sin\theta = 0$ we set $\theta = y_1, \theta' = y_2$ (the angular velocity) and use

$$\theta'' = \frac{dy_2}{dt} = \frac{dy_2}{dy_1}\frac{dy_1}{dt} = \frac{dy_2}{dy_1} y_2, \qquad \text{we get} \qquad \frac{dy_2}{dy_1} y_2 = -k \sin y_1.$$

Separation of variables gives $y_2\, dy_2 = -k \sin y_1\, dy_1$. By integration,

$$\tfrac{1}{2} y_2^2 = k \cos y_1 + C \qquad (C \text{ constant}). \tag{9}$$

Multiplying this by mL^2, we get

$$\tfrac{1}{2} m(Ly_2)^2 - mL^2 k \cos y_1 = mL^2 C.$$

We see that these three terms are **energies**. Indeed, y_2 is the angular velocity, so that Ly_2 is the velocity and the first term is the kinetic energy. The second term (including the minus sign) is the potential energy of the pendulum, and mL^2C is its total energy, which is constant, as expected from the law of conservation of energy, because there is no damping (no loss of energy). The type of motion depends on the total energy, hence on C, as follows.

Figure 93b shows trajectories for various values of C. These graphs continue periodically with period 2π to the left and to the right. We see that some of them are ellipse-like and closed, others are wavy, and there are two trajectories (passing through the saddle points $(n\pi, 0), n = \pm 1, \pm 3, \cdots$) that separate those two types of trajectories. From (9) we see that the smallest possible C is $C = -k$; then $y_2 = 0$, and $\cos y_1 = 1$, so that the pendulum is at rest. The pendulum will change its direction of motion if there are points at which $y_2 = \theta' = 0$. Then $k \cos y_1 + C = 0$ by (9). If $y_1 = \pi$, then $\cos y_1 = -1$ and $C = k$. Hence if $-k < C < k$, then the pendulum reverses its direction for a $|y_1| = |\theta| < \pi$, and for these values of C with $|C| < k$ the pendulum oscillates. This corresponds to the closed trajectories in the figure. However, if $C > k$, then $y_2 = 0$ is impossible and the pendulum makes a whirly motion that appears as a wavy trajectory in the $y_1 y_2$-plane. Finally, the value $C = k$ corresponds to the two "separating trajectories" in Fig. 93b connecting the saddle points. ■

The phase plane method of deriving a single first-order equation (8) may be of practical interest not only when (8) can be solved (as in Example 4) but also when a solution is not possible and we have to utilize fields (Sec. 1.2). We illustrate this with a very famous example:

EXAMPLE 5 Self-Sustained Oscillations. Van der Pol Equation

There are physical systems such that for small oscillations, energy is fed into the system, whereas for large oscillations, energy is taken from the system. In other words, large oscillations will be damped, whereas for small oscillations there is "negative damping" (feeding of energy into the system). For physical reasons we expect such a system to approach a periodic behavior, which will thus appear as a closed trajectory in the phase plane, called a **limit cycle**. A differential equation describing such vibrations is the famous **van der Pol equation**[4]

$$y'' - \mu(1 - y^2)y' + y = 0 \qquad (\mu > 0, \text{ constant}). \tag{10}$$

It first occurred in the study of electrical circuits containing vacuum tubes. For $\mu = 0$ this equation becomes $y'' + y = 0$ and we obtain harmonic oscillations. Let $\mu > 0$. The damping term has the factor $-\mu(1 - y^2)$. This is negative for small oscillations, when $y^2 < 1$, so that we have "negative damping," is zero for $y^2 = 1$ (no damping), and is positive if $y^2 > 1$ (positive damping, loss of energy). If μ is small, we expect a limit cycle that is almost a circle because then our equation differs but little from $y'' + y = 0$. If μ is large, the limit cycle will probably look different.

Setting $y = y_1$, $y' = y_2$ and using $y'' = (dy_2/dy_1)y_2$ as in (8), we have from (10)

$$\frac{dy_2}{dy_1} y_2 - \mu(1 - y_1^2)y_2 + y_1 = 0. \tag{11}$$

The isoclines in the y_1y_2-plane (the phase plane) are the curves $dy_2/dy_1 = K = \text{const}$, that is,

$$\frac{dy_2}{dy_1} = \mu(1 - y_1^2) - \frac{y_1}{y_2} = K.$$

Solving algebraically for y_2, we see that the isoclines are given by

$$y_2 = \frac{y_1}{\mu(1 - y_1^2) - K} \qquad \text{(Figs. 96, 97).}$$

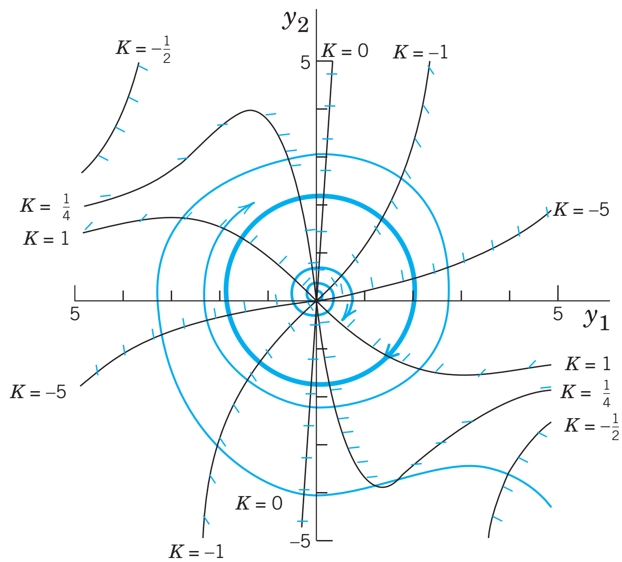


Fig. 96. Direction field for the van der Pol equation with $\mu = 0.1$ in the phase plane, showing also the limit cycle and two trajectories. See also Fig. 8 in Sec. 1.2

[4]BALTHASAR VAN DER POL (1889–1959), Dutch physicist and engineer.

Figure 96 shows some isoclines when μ is small, $\mu = 0.1$, the limit cycle (almost a circle), and two (blue) trajectories approaching it, one from the outside and the other from the inside, of which only the initial portion, a small spiral, is shown. Due to this approach by trajectories, a limit cycle differs conceptually from a closed curve (a trajectory) surrounding a center, which is not approached by trajectories. For larger μ the limit cycle no longer resembles a circle, and the trajectories approach it more rapidly than for smaller μ. Figure 97 illustrates this for $\mu = 1$. ■

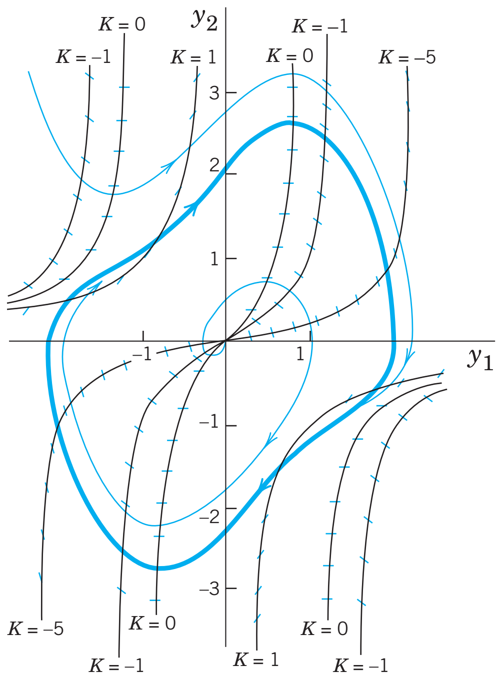


Fig. 97. Direction field for the van der Pol equation with $\mu = 1$ in the phase plane, showing also the limit cycle and two trajectories approaching it

PROBLEM SET 4.5

1. **Pendulum.** To what state (position, speed, direction of motion) do the four points of intersection of a closed trajectory with the axes in Fig. 93b correspond? The point of intersection of a wavy curve with the y_2-axis?

2. **Limit cycle.** What is the essential difference between a limit cycle and a closed trajectory surrounding a center?

3. **CAS EXPERIMENT. Deformation of Limit Cycle.** Convert the van der Pol equation to a system. Graph the limit cycle and some approaching trajectories for $\mu = 0.2, 0.4, 0.6, 0.8, 1.0, 1.5, 2.0$. Try to observe how the limit cycle changes its form continuously if you vary μ continuously. Describe in words how the limit cycle is deformed with growing μ.

4–8 CRITICAL POINTS. LINEARIZATION

Find the location and type of all critical points by linearization. Show the details of your work.

4. $y_1' = 4y_1 - y_1^2$
   $y_2' = y_2$

5. $y_1' = y_2$
   $y_2' = -y_1 + \frac{1}{2}y_1^2$

6. $y_1' = y_2$
   $y_2' = -y_1 - y_1^2$

7. $y_1' = -y_1 + y_2 - y_2^2$
   $y_2' = -y_1 - y_2$

8. $y_1' = y_2 - y_2^2$
   $y_2' = y_1 - y_1^2$

9–13 CRITICAL POINTS OF ODEs

Find the location and type of all critical points by first converting the ODE to a system and then linearizing it.

9. $y'' - 9y + y^3 = 0$

10. $y'' + y - y^3 = 0$

11. $y'' + \cos y = 0$

12. $y'' + 9y + y^2 = 0$

13. $y'' + \sin y = 0$

14. TEAM PROJECT. Self-sustained oscillations.
(a) Van der Pol equation. Determine the type of the critical point at $(0, 0)$ when $\mu > 0, \mu = 0, \mu < 0$.
(b) Rayleigh equation. Show that the Rayleigh equation[5]

$$Y'' - \mu(1 - \tfrac{1}{3}Y'^2)Y' + Y = 0 \quad (\mu > 0)$$

also describes self-sustained oscillations and that by differentiating it and setting $y = Y'$ one obtains the van der Pol equation.
(c) Duffing equation. The Duffing equation is

$$y'' + \omega_0^2 y + \beta y^3 = 0$$

where usually $|\beta|$ is small, thus characterizing a small deviation of the restoring force from linearity. $\beta > 0$ and $\beta < 0$ are called the cases of a *hard spring* and a *soft spring*, respectively. Find the equation of the trajectories in the phase plane. (Note that for $\beta > 0$ all these curves are closed.)

15. Trajectories. Write the ODE $y'' - 4y + y^3 = 0$ as a system, solve it for y_2 as a function of y_1, and sketch or graph some of the trajectories in the phase plane.

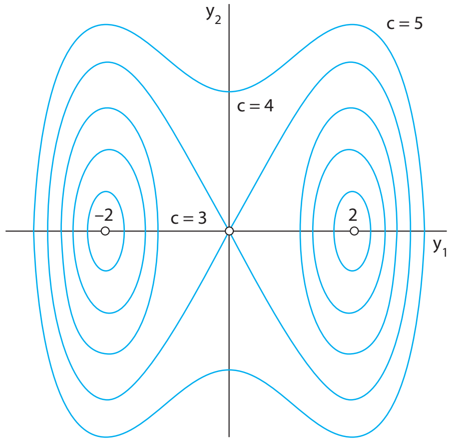


Fig. 98. Trajectories in Problem 15

4.6 Nonhomogeneous Linear Systems of ODEs

In this section, the last one of Chap. 4, we discuss methods for solving nonhomogeneous linear systems of ODEs

$$\mathbf{y}' = \mathbf{A}\mathbf{y} + \mathbf{g} \tag{1}$$ (see Sec. 4.2)

where the vector $\mathbf{g}(t)$ is not identically zero. We assume $\mathbf{g}(t)$ and the entries of the $n \times n$ matrix $\mathbf{A}(t)$ to be continuous on some interval J of the t-axis. From a general solution $\mathbf{y}^{(h)}(t)$ of the homogeneous system $\mathbf{y}' = \mathbf{A}\mathbf{y}$ on J and a **particular solution** $\mathbf{y}^{(p)}(t)$ of (1) on J [i.e., a solution of (1) containing no arbitrary constants], we get a solution of (1),

$$\mathbf{y} = \mathbf{y}^{(h)} + \mathbf{y}^{(p)}. \tag{2}$$

$\mathbf{y}$ is called a **general solution** of (1) on J because it includes every solution of (1) on J. This follows from Theorem 2 in Sec. 4.2 (see Prob. 1 of this section).

Having studied homogeneous linear systems in Secs. 4.1–4.4, our present task will be to explain methods for obtaining particular solutions of (1). We discuss the method of

[5]LORD RAYLEIGH (JOHN WILLIAM STRUTT) (1842–1919), English physicist and mathematician, professor at Cambridge and London, known by his important contributions to the theory of waves, elasticity theory, hydrodynamics, and various other branches of applied mathematics and theoretical physics. In 1904 he was awarded the Nobel Prize in physics.

undetermined coefficients and the method of the variation of parameters; these have counterparts for a single ODE, as we know from Secs. 2.7 and 2.10.

Method of Undetermined Coefficients

Just as for a single ODE, this method is suitable if the entries of $\mathbf{A}$ are constants and the components of $\mathbf{g}$ are constants, positive integer powers of t, exponential functions, or cosines and sines. In such a case a particular solution $\mathbf{y}^{(p)}$ is assumed in a form similar to $\mathbf{g}$; for instance, $\mathbf{y}^{(p)} = \mathbf{u} + \mathbf{v}t + \mathbf{w}t^2$ if $\mathbf{g}$ has components quadratic in t, with $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ to be determined by substitution into (1). This is similar to Sec. 2.7, except for the Modification Rule. It suffices to show this by an example.

EXAMPLE 1 Method of Undetermined Coefficients. Modification Rule

Find a general solution of

$$\mathbf{y}' = \mathbf{A}\mathbf{y} + \mathbf{g} = \begin{bmatrix} -3 & 1 \\ 1 & -3 \end{bmatrix} \mathbf{y} + \begin{bmatrix} -6 \\ 2 \end{bmatrix} e^{-2t}. \tag{3}$$

Solution. A general equation of the homogeneous system is (see Example 1 in Sec. 4.3)

$$\mathbf{y}^{(h)} = c_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{-2t} + c_2 \begin{bmatrix} 1 \\ -1 \end{bmatrix} e^{-4t}. \tag{4}$$

Since $\lambda = -2$ is an eigenvalue of $\mathbf{A}$, the function e^{-2t} on the right side also appears in $\mathbf{y}^{(h)}$, and we must apply the Modification Rule by setting

$$\mathbf{y}^{(p)} = \mathbf{u}te^{-2t} + \mathbf{v}e^{-2t} \qquad \text{(rather than } \mathbf{u}e^{-2t}\text{)}.$$

Note that the first of these two terms is the analog of the modification in Sec. 2.7, but it would not be sufficient here. (Try it.) By substitution,

$$\mathbf{y}^{(p)\prime} = \mathbf{u}e^{-2t} - 2\mathbf{u}te^{-2t} - 2\mathbf{v}e^{-2t} = \mathbf{A}\mathbf{u}te^{-2t} + \mathbf{A}\mathbf{v}e^{-2t} + \mathbf{g}.$$

Equating the te^{-2t}-terms on both sides, we have $-2\mathbf{u} = \mathbf{A}\mathbf{u}$. Hence $\mathbf{u}$ is an eigenvector of $\mathbf{A}$ corresponding to $\lambda = -2$; thus [see (5)] $\mathbf{u} = a[1 \quad 1]^{\mathsf{T}}$ with any $a \neq 0$. Equating the other terms gives

$$\mathbf{u} - 2\mathbf{v} = \mathbf{A}\mathbf{v} + \begin{bmatrix} -6 \\ 2 \end{bmatrix} \qquad \text{thus} \qquad \begin{bmatrix} a \\ a \end{bmatrix} - \begin{bmatrix} 2v_1 \\ 2v_2 \end{bmatrix} = \begin{bmatrix} -3v_1 + v_2 \\ v_1 - 3v_2 \end{bmatrix} + \begin{bmatrix} -6 \\ 2 \end{bmatrix}.$$

Collecting terms and reshuffling gives

$$\begin{aligned} v_1 - v_2 &= -a - 6 \\ -v_1 + v_2 &= -a + 2. \end{aligned}$$

By addition, $0 = -2a - 4$, $a = -2$, and then $v_2 = v_1 + 4$, say, $v_1 = k$, $v_2 = k + 4$, thus, $\mathbf{v} = [k \quad k+4]^{\mathsf{T}}$. We can simply choose $k = 0$. This gives the *answer*

$$\mathbf{y} = \mathbf{y}^{(h)} + \mathbf{y}^{(p)} = c_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{-2t} + c_2 \begin{bmatrix} 1 \\ -1 \end{bmatrix} e^{-4t} - 2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} te^{-2t} + \begin{bmatrix} 0 \\ 4 \end{bmatrix} e^{-2t}. \tag{5}$$

For other k we get other $\mathbf{v}$; for instance, $k = -2$ gives $\mathbf{v} = [-2 \quad 2]^{\mathsf{T}}$, so that the *answer* becomes

$$\mathbf{y} = c_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{-2t} + c_2 \begin{bmatrix} 1 \\ -1 \end{bmatrix} e^{-4t} - 2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} te^{-2t} + \begin{bmatrix} -2 \\ 2 \end{bmatrix} e^{-2t}, \qquad \text{etc.} \tag{5*}$$

■

Method of Variation of Parameters

This method can be applied to nonhomogeneous linear systems

$$\mathbf{y}' = \mathbf{A}(t)\mathbf{y} + \mathbf{g}(t) \tag{6}$$

with variable $\mathbf{A} = \mathbf{A}(t)$ and general $\mathbf{g}(t)$. It yields a particular solution $\mathbf{y}^{(p)}$ of (6) on some open interval J on the t-axis if a general solution of the homogeneous system $\mathbf{y}' = \mathbf{A}(t)\mathbf{y}$ on J is known. We explain the method in terms of the previous example.

EXAMPLE 2 Solution by the Method of Variation of Parameters

Solve (3) in Example 1.

Solution. A basis of solutions of the homogeneous system is $[e^{-2t} \quad e^{-2t}]^\mathsf{T}$ and $[e^{-4t} \quad -e^{-4t}]^\mathsf{T}$. Hence the general solution (4) of the homogeneous system may be written

$$\mathbf{y}^{(h)} = \begin{bmatrix} e^{-2t} & e^{-4t} \\ e^{-2t} & -e^{-4t} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \mathbf{Y}(t)\,\mathbf{c}. \tag{7}$$

Here, $\mathbf{Y}(t) = [\mathbf{y}^{(1)} \quad \mathbf{y}^{(2)}]^\mathsf{T}$ is the fundamental matrix (see Sec. 4.2). As in Sec. 2.10 we replace the constant vector $\mathbf{c}$ by a variable vector $\mathbf{u}(t)$ to obtain a particular solution

$$\mathbf{y}^{(p)} = \mathbf{Y}(t)\mathbf{u}(t).$$

Substitution into (3) $\mathbf{y}' = \mathbf{A}\mathbf{y} + \mathbf{g}$ gives

$$\mathbf{Y}'\mathbf{u} + \mathbf{Y}\mathbf{u}' = \mathbf{A}\mathbf{Y}\mathbf{u} + \mathbf{g}. \tag{8}$$

Now since $\mathbf{y}^{(1)}$ and $\mathbf{y}^{(2)}$ are solutions of the homogeneous system, we have

$$\mathbf{y}^{(1)\prime} = \mathbf{A}\mathbf{y}^{(1)}, \qquad \mathbf{y}^{(2)\prime} = \mathbf{A}\mathbf{y}^{(2)}, \qquad \text{thus} \qquad \mathbf{Y}' = \mathbf{A}\mathbf{Y}.$$

Hence $\mathbf{Y}'\mathbf{u} = \mathbf{A}\mathbf{Y}\mathbf{u}$, so that (8) reduces to

$$\mathbf{Y}\mathbf{u}' = \mathbf{g}. \qquad \text{The solution is} \qquad \mathbf{u}' = \mathbf{Y}^{-1}\mathbf{g};$$

here we use that the inverse $\mathbf{Y}^{-1}$ of $\mathbf{Y}$ (Sec. 4.0) exists because the determinant of $\mathbf{Y}$ is the Wronskian W, which is not zero for a basis. Equation (9) in Sec. 4.0 gives the form of $\mathbf{Y}^{-1}$,

$$\mathbf{Y}^{-1} = \frac{1}{-2e^{-6t}} \begin{bmatrix} -e^{-4t} & -e^{-4t} \\ -e^{-2t} & e^{-2t} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} e^{2t} & e^{2t} \\ e^{4t} & -e^{4t} \end{bmatrix}.$$

We multiply this by $\mathbf{g}$, obtaining

$$\mathbf{u}' = \mathbf{Y}^{-1}\mathbf{g} = \frac{1}{2} \begin{bmatrix} e^{2t} & e^{2t} \\ e^{4t} & -e^{4t} \end{bmatrix} \begin{bmatrix} -6e^{-2t} \\ 2e^{-2t} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} -4 \\ -8e^{2t} \end{bmatrix} = \begin{bmatrix} -2 \\ -4e^{2t} \end{bmatrix}.$$

Integration is done componentwise (just as differentiation) and gives

$$\mathbf{u}(t) = \int_0^t \begin{bmatrix} -2 \\ -4e^{2\tilde{t}} \end{bmatrix} d\tilde{t} = \begin{bmatrix} -2t \\ -2e^{2t} + 2 \end{bmatrix}$$

(where $+\,2$ comes from the lower limit of integration). From this and $\mathbf{Y}$ in (7) we obtain

$$\mathbf{Y}\mathbf{u} = \begin{bmatrix} e^{-2t} & e^{-4t} \\ e^{-2t} & -e^{-4t} \end{bmatrix} \begin{bmatrix} -2t \\ -2e^{2t} + 2 \end{bmatrix} = \begin{bmatrix} -2te^{-2t} - 2e^{-2t} + 2e^{-4t} \\ -2te^{-2t} + 2e^{-2t} - 2e^{-4t} \end{bmatrix} = \begin{bmatrix} -2t - 2 \\ -2t + 2 \end{bmatrix} e^{-2t} + \begin{bmatrix} 2 \\ -2 \end{bmatrix} e^{-4t}.$$

The last term on the right is a solution of the homogeneous system. Hence we can absorb it into $\mathbf{y}^{(h)}$. We thus obtain as a general solution of the system (3), in agreement with (5*).

$$\mathbf{y} = c_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{-2t} + c_2 \begin{bmatrix} 1 \\ -1 \end{bmatrix} e^{-4t} - 2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} te^{-2t} + \begin{bmatrix} -2 \\ 2 \end{bmatrix} e^{-2t}. \tag{9}$$

PROBLEM SET 4.6

1. Prove that (2) includes every solution of (1).

2–7 GENERAL SOLUTION

Find a general solution. Show the details of your work.

2. $y_1' = y_1 + y_2 + 10 \cos t$
$y_2' = 3y_1 - y_2 - 10 \sin t$

3. $y_1' = y_2 + e^{3t}$
$y_2' = y_1 - 3e^{3t}$

4. $y_1' = 4y_1 - 8y_2 + 2 \cosh t$
$y_2' = 2y_1 - 6y_2 + \cosh t + 2 \sinh t$

5. $y_1' = 4y_1 + y_2 + 0.6t$
$y_2' = 2y_1 + 3y_2 - 2.5t$

6. $y_1' = 4y_2$
$y_2' = 4y_1 - 16t^2 + 2$

7. $y_1' = -3y_1 - 4y_2 + 11t + 15$
$y_2' = 5y_1 + 6y_2 + 3e^{-t} - 15t - 20$

8. CAS EXPERIMENT. Undetermined Coefficients. Find out experimentally how general you must choose $\mathbf{y}^{(p)}$, in particular when the components of $\mathbf{g}$ have a different form (e.g., as in Prob. 7). Write a short report, covering also the situation in the case of the modification rule.

9. Undetermined Coefficients. Explain why, in Example 1 of the text, we have some freedom in choosing the vector $\mathbf{v}$.

10–15 INITIAL VALUE PROBLEM

Solve, showing details:

10. $y_1' = -3y_1 - 4y_2 + 5e^t$
$y_2' = 5y_1 + 6y_2 - 6e^t$
$y_1(0) = 19, \quad y_2(0) = -23$

11. $y_1' = y_2 + 6e^{2t}$
$y_2' = y_1 - e^{2t}$
$y_1(0) = 1, \quad y_2(0) = 0$

12. $y_1' = y_1 + 4y_2 - t^2 + 6t$
$y_2' = y_1 + y_2 - t^2 + t - 1$
$y_1(0) = 2, \quad y_2(0) = -1$

13. $y_1' = y_2 - 5 \sin t$
$y_2' = -4y_1 + 17 \cos t$
$y_1(0) = 5, \quad y_2(0) = 2$

14. $y_1' = 4y_2 + 5e^t$
$y_2' = -y_1 - 20e^{-t}$
$y_1(0) = 1, \quad y_2(0) = 0$

15. $y_1' = y_1 + 2y_2 + e^{2t} - 2t$
$y_2' = -y_2 + 1 + t$
$y_1(0) = 1, \quad y_2(0) = -4$

16. WRITING PROJECT. Undetermined Coefficients. Write a short report in which you compare the application of the method of undetermined coefficients to a single ODE and to a system of ODEs, using ODEs and systems of your choice.

17–20 NETWORK

Find the currents in Fig. 99 (Probs. 17–19) and Fig. 100 (Prob. 20) for the following data, showing the details of your work.

17. $R_1 = 2\ \Omega, R_2 = 8\ \Omega, L = 1$ H, $C = 0.5$ F, $E = 200$ V

18. Solve Prob. 17 with $E = 440 \sin t$ V and the other data as before.

19. In Prob. 17 find the particular solution when currents and charge at $t = 0$ are zero.

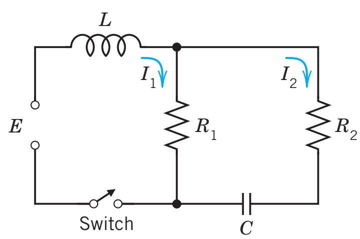


Fig. 99. Problems 17–19

20. $R_1 = 1\ \Omega$, $R_2 = 1.4\ \Omega$, $L_1 = 0.8$ H, $L_2 = 1$ H, $E = 100$ V, $I_1(0) = I_2(0) = 0$

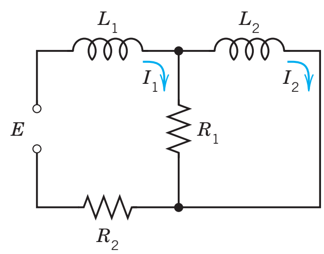


Fig. 100. Problem 20

CHAPTER 4 REVIEW QUESTIONS AND PROBLEMS

1. State some applications that can be modeled by systems of ODEs.

2. What is population dynamics? Give examples.

3. How can you transform an ODE into a system of ODEs?

4. What are qualitative methods for systems? Why are they important?

5. What is the phase plane? The phase plane method? A trajectory? The phase portrait of a system of ODEs?

6. What are critical points of a system of ODEs? How did we classify them? Why are they important?

7. What are eigenvalues? What role did they play in this chapter?

8. What does stability mean in general? In connection with critical points? Why is stability important in engineering?

9. What does linearization of a system mean?

10. Review the pendulum equations and their linearizations.

11–17 GENERAL SOLUTION. CRITICAL POINTS

Find a general solution. Determine the kind and stability of the critical point.

11. $y_1' = 2y_2$, $y_2' = 8y_1$

12. $y_1' = 5y_1$, $y_2' = y_2$

13. $y_1' = -2y_1 + 5y_2$, $y_2' = -y_1 - 6y_2$

14. $y_1' = 3y_1 + 4y_2$, $y_2' = 3y_1 + 2y_2$

15. $y_1' = -3y_1 - 2y_2$, $y_2' = -2y_1 - 3y_2$

16. $y_1' = 4y_2$, $y_2' = -4y_1$

17. $y_1' = -y_1 + 2y_2$, $y_2' = -2y_1 - y_2$

18–19 CRITICAL POINT

What kind of critical point does $\mathbf{y}' = \mathbf{A}\mathbf{y}$ have if $\mathbf{A}$ has the eigenvalues

18. -4 and 2

19. $2 + 3i,\ 2 - 3i$

20–23 NONHOMOGENEOUS SYSTEMS

Find a general solution. Show the details of your work.

20. $y_1' = 2y_1 + 2y_2 + e^t$, $y_2' = -2y_1 - 3y_2 + e^t$

21. $y_1' = 4y_2$, $y_2' = 4y_1 + 32t^2$

22. $y_1' = y_1 + y_2 + \sin t$, $y_2' = 4y_1 + y_2$

23. $y_1' = y_1 + 4y_2 - 2\cos t$, $y_2' = y_1 + y_2 - \cos t + \sin t$

24. Mixing problem. Tank T_1 in Fig. 101 initially contains 200 gal of water in which 160 lb of salt are dissolved. Tank T_2 initially contains 100 gal of pure water. Liquid is pumped through the system as indicated, and the mixtures are kept uniform by stirring. Find the amounts of salt $y_1(t)$ and $y_2(t)$ in T_1 and T_2, respectively.

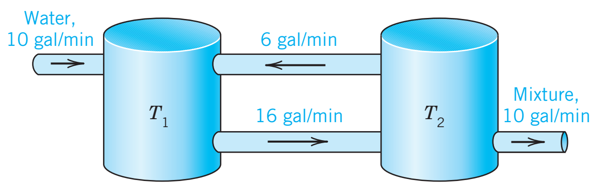


Fig. 101. Tanks in Problem 24

25. Network. Find the currents in Fig. 102 when $R = 2.5\ \Omega$, $L = 1$ H, $C = 0.04$ F, $E(t) = 169 \sin t$ V, $I_1(0) = 0$, $I_2(0) = 0$.

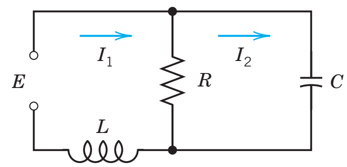


Fig. 102. Network in Problem 25

26. Network. Find the currents in Fig. 103 when $R = 1\ \Omega$, $L = 1.25$ H, $C = 0.2$ F, $I_1(0) = 1$ A, $I_2(0) = 1$ A.

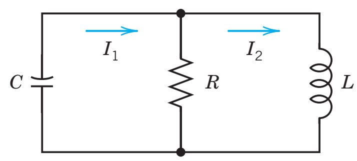


Fig. 103. Network in Problem 26

27–30 LINEARIZATION

Find the location and kind of all critical points of the given nonlinear system by linearization.

27. $y_1' = y_2$, $y_2' = y_1 - y_1^3$

28. $y_1' = \cos y_2$, $y_2' = 3y_1$

29. $y_1' = -4y_2$, $y_2' = \sin y_1$

30. $y_1' = 2y_2 + 2y_2^2$, $y_2' = -8y_1$

SUMMARY OF CHAPTER 4

Systems of ODEs. Phase Plane. Qualitative Methods

Whereas single electric circuits or single mass–spring systems are modeled by single ODEs (Chap. 2), networks of several circuits, systems of several masses and springs, and other engineering problems lead to **systems of ODEs**, involving several unknown functions $y_1(t), \cdots, y_n(t)$. Of central interest are **first-order systems** (Sec. 4.2):

$$\mathbf{y}' = \mathbf{f}(t, \mathbf{y}), \qquad \text{in components,} \qquad \begin{aligned} y_1' &= f_1(t, y_1, \cdots, y_n) \\ &\vdots \\ y_n' &= f_n(t, y_1, \cdots, y_n), \end{aligned}$$

to which higher order ODEs and systems of ODEs can be reduced (Sec. 4.1). In this summary we let $n = 2$, so that

$$(1) \qquad \mathbf{y}' = \mathbf{f}(t, \mathbf{y}), \qquad \text{in components,} \qquad \begin{aligned} y_1' &= f_1(t, y_1, y_2) \\ y_2' &= f_2(t, y_1, y_2). \end{aligned}$$

Then we can represent solution curves as ***trajectories*** in ***the phase plane*** (the y_1y_2-plane), investigate their totality [the "***phase portrait***" of (1)], and study the kind and ***stability*** of the ***critical points*** (points at which both f_1 and f_2 are zero), and classify them as ***nodes, saddle points, centers,*** or ***spiral points*** (Secs. 4.3, 4.4). These phase plane methods are **qualitative**; with their use we can discover various general properties of solutions without actually solving the system. They are primarily used for **autonomous systems**, that is, systems in which t does not occur explicitly.

A **linear system** is of the form

$$(2) \qquad \mathbf{y}' = \mathbf{A}\mathbf{y} + \mathbf{g}, \quad \text{where} \quad \mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \quad \mathbf{g} = \begin{bmatrix} g_1 \\ g_2 \end{bmatrix}.$$

If $\mathbf{g} = \mathbf{0}$, the system is called **homogeneous** and is of the form

$$(3) \qquad \mathbf{y}' = \mathbf{A}\mathbf{y}.$$

If $a_{11}, \cdots, a_{22}$ are constants, it has solutions $\mathbf{y} = \mathbf{x}e^{\lambda t}$, where λ is a solution of the quadratic equation

$$\begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{vmatrix} = (a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21} = 0$$

and $\mathbf{x} \neq \mathbf{0}$ has components x_1, x_2 determined up to a multiplicative constant by

$$(a_{11} - \lambda)x_1 + a_{12}x_2 = 0.$$

(These λ's are called the **eigenvalues** and these vectors $\mathbf{x}$ **eigenvectors** of the matrix $\mathbf{A}$. Further explanation is given in Sec. 4.0.)

A system (2) with $\mathbf{g} \neq \mathbf{0}$ is called **nonhomogeneous**. Its general solution is of the form $\mathbf{y} = \mathbf{y}_h + \mathbf{y}_p$, where $\mathbf{y}_h$ is a general solution of (3) and $\mathbf{y}_p$ a particular solution of (2). Methods of determining the latter are discussed in Sec. 4.6.

The discussion of critical points of linear systems based on eigenvalues is summarized in Tables 4.1 and 4.2 in Sec. 4.4. It also applies to nonlinear systems if the latter are first linearized. The key theorem for this is Theorem 1 in Sec. 4.5, which also includes three famous applications, namely the pendulum and van der Pol equations and the Lotka–Volterra predator–prey population model.

CHAPTER 5

Series Solutions of ODEs. Special Functions

In the previous chapters, we have seen that linear ODEs with *constant coefficients* can be solved by algebraic methods, and that their solutions are elementary functions known from calculus. For ODEs with *variable coefficients* the situation is more complicated, and their solutions may be nonelementary functions. *Legendre's, Bessel's*, and the *hypergeometric equations* are important ODEs of this kind. Since these ODEs and their solutions, the *Legendre polynomials, Bessel functions*, and *hypergeometric functions*, play an important role in engineering modeling, we shall consider the two standard methods for solving such ODEs.

The first method is called the **power series method** because it gives solutions in the form of a power series $a_0 + a_1x + a_2x^2 + a_3x^3 + \cdots$.

The second method is called the **Frobenius method** and generalizes the first; it gives solutions in power series, multiplied by a logarithmic term $\ln x$ or a fractional power x^r, in cases such as Bessel's equation, in which the first method is not general enough.

All those more advanced solutions and various other functions not appearing in calculus are known as *higher functions* or **special functions**, which has become a technical term. Each of these functions is important enough to give it a name and investigate its properties and relations to other functions in great detail (take a look into Refs. [GenRef1], [GenRef10], or [All] in App. 1). Your CAS knows practically all functions you will ever need in industry or research labs, but it is up to you to find your way through this vast terrain of formulas. The present chapter may give you some help in this task.

COMMENT. You can study this chapter directly after Chap. 2 because it needs no material from Chaps. 3 or 4.

Prerequisite: Chap. 2.
Section that may be omitted in a shorter course: 5.5.
References and Answers to Problems: App. 1 Part A, and App. 2.

5.1 Power Series Method

The **power series method** is the standard method for solving linear ODEs with ***variable*** coefficients. It gives solutions in the form of power series. These series can be used for computing values, graphing curves, proving formulas, and exploring properties of solutions, as we shall see. In this section we begin by explaining the idea of the power series method.

From calculus we remember that a **power series** (in powers of $x - x_0$) is an infinite series of the form

$$\sum_{m=0}^{\infty} a_m(x - x_0)^m = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \cdots. \tag{1}$$

Here, x is a variable. $a_0, a_1, a_2, \cdots$ are constants, called the **coefficients** of the series. x_0 is a constant, called the **center** of the series. In particular, if $x_0 = 0$, we obtain a **power series in powers of *x***

$$\sum_{m=0}^{\infty} a_m x^m = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots. \tag{2}$$

We shall assume that all variables and constants are real.

We note that the term "power series" usually refers to a series of the form (1) [or (2)] but ***does not include*** series of negative or fractional powers of x. We use m as the summation letter, reserving n as a standard notation in the Legendre and Bessel equations for integer values of the parameter.

EXAMPLE 1 **Familiar Power Series** are the Maclaurin series

$$\frac{1}{1-x} = \sum_{m=0}^{\infty} x^m = 1 + x + x^2 + \cdots \qquad (|x| < 1, \textbf{ geometric series})$$

$$e^x = \sum_{m=0}^{\infty} \frac{x^m}{m!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

$$\cos x = \sum_{m=0}^{\infty} \frac{(-1)^m x^{2m}}{(2m)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - + \cdots$$

$$\sin x = \sum_{m=0}^{\infty} \frac{(-1)^m x^{2m+1}}{(2m+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - + \cdots.$$

Idea and Technique of the Power Series Method

The idea of the power series method for solving linear ODEs seems natural, once we know that the most important ODEs in applied mathematics have solutions of this form. We explain the idea by an ODE that can readily be solved otherwise.

EXAMPLE 2 **Power Series Solution.** Solve $y' - y = 0$.

Solution. In the first step we insert

$$y = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots = \sum_{m=0}^{\infty} a_m x^m \tag{2}$$

and the series obtained by termwise differentiation

$$(3) \qquad y' = a_1 + 2a_2 x + 3a_3 x^2 + \cdots = \sum_{m=1}^{\infty} m a_m x^{m-1}$$

into the ODE:

$$(a_1 + 2a_2 x + 3a_3 x^2 + \cdots) - (a_0 + a_1 x + a_2 x^2 + \cdots) = 0.$$

Then we collect like powers of x, finding

$$(a_1 - a_0) + (2a_2 - a_1)x + (3a_3 - a_2)x^2 + \cdots = 0.$$

Equating the coefficient of each power of x to zero, we have

$$a_1 - a_0 = 0, \qquad 2a_2 - a_1 = 0, \qquad 3a_3 - a_2 = 0, \cdots.$$

Solving these equations, we may express $a_1, a_2, \cdots$ in terms of a_0, which remains arbitrary:

$$a_1 = a_0, \qquad a_2 = \frac{a_1}{2} = \frac{a_0}{2!}, \qquad a_3 = \frac{a_2}{3} = \frac{a_0}{3!}, \cdots.$$

With these values of the coefficients, the series solution becomes the familiar general solution

$$y = a_0 + a_0 x + \frac{a_0}{2!}x^2 + \frac{a_0}{3!}x^3 + \cdots = a_0\left(1 + x + \frac{x^2}{2!} + \frac{x^3}{3!}\right) = a_0 e^x.$$

Test your comprehension by solving $y'' + y = 0$ by power series. You should get the result $y = a_0 \cos x + a_1 \sin x$. ■

We now describe the method in general and justify it after the next example. For a given ODE

$$(4) \qquad y'' + p(x)y' + q(x)y = 0$$

we first represent $p(x)$ and $q(x)$ by power series in powers of x (or of $x - x_0$ if solutions in powers of $x - x_0$ are wanted). Often $p(x)$ and $q(x)$ are polynomials, and then nothing needs to be done in this first step. Next we assume a solution in the form of a power series (2) with unknown coefficients and insert it as well as (3) and

$$(5) \qquad y'' = 2a_2 + 3 \cdot 2a_3 x + 4 \cdot 3a_4 x^2 + \cdots = \sum_{m=2}^{\infty} m(m-1)a_m x^{m-2}$$

into the ODE. Then we collect like powers of x and equate the sum of the coefficients of each occurring power of x to zero, starting with the constant terms, then taking the terms containing x, then the terms in x^2, and so on. This gives equations from which we can determine the unknown coefficients of (3) successively.

EXAMPLE 3 **A Special Legendre Equation.** The ODE

$$(1 - x^2)y'' - 2xy' + 2y = 0$$

occurs in models exhibiting spherical symmetry. Solve it.

Solution. Substitute (2), (3), and (5) into the ODE. $(1-x^2)y''$ gives two series, one for y'' and one for $-x^2y''$. In the term $-2xy'$ use (3) and in $2y$ use (2). Write like powers of x vertically aligned. This gives

$$\begin{aligned}
y'' &= 2a_2 + 6a_3x + 12a_4x^2 + 20a_5x^3 + 30a_6x^4 + \cdots \\
-x^2y'' &= \qquad\qquad\qquad - 2a_2x^2 - 6a_3x^3 - 12a_4x^4 - \cdots \\
-2xy' &= \qquad - 2a_1x - 4a_2x^2 - 6a_3x^3 - 8a_4x^4 - \cdots \\
2y &= 2a_0 + 2a_1x + 2a_2x^2 + 2a_3x^3 + 2a_4x^4 + \cdots .
\end{aligned}$$

Add terms of like powers of x. For each power $x^0, x, x^2, \cdots$ equate the sum obtained to zero. Denote these sums by [0] (constant terms), [1] (first power of x), and so on:

Sum	Power	Equations	
[0]	$[x^0]$	$a_2 = -a_0$	
[1]	$[x]$	$a_3 = 0$	
[2]	$[x^2]$	$12a_4 = 4a_2,$	$a_4 = \frac{4}{12}a_2 = -\frac{1}{3}a_0$
[3]	$[x^3]$	$a_5 = 0$	since $a_3 = 0$
[4]	$[x^4]$	$30a_6 = 18a_4,$	$a_6 = \frac{18}{30}a_4 = \frac{18}{30}(-\frac{1}{3})a_0 = -\frac{1}{5}a_0.$

This gives the solution

$$y = a_1x + a_0(1 - x^2 - \tfrac{1}{3}x^4 - \tfrac{1}{5}x^6 - \cdots).$$

a_0 and a_1 remain arbitrary. Hence, this is a general solution that consists of two solutions: x and $1 - x^2 - \frac{1}{3}x^4 - \frac{1}{5}x^6 - \cdots$. These two solutions are members of families of functions called *Legendre polynomials* $P_n(x)$ and *Legendre functions* $Q_n(x)$; here we have $x = P_1(x)$ and $1 - x^2 - \frac{1}{3}x^4 - \frac{1}{5}x^6 - \cdots = -Q_1(x)$. The minus is by convention. The index 1 is called the *order* of these two functions and here the order is 1. More on Legendre polynomials in the next section. ■

Theory of the Power Series Method

The ***n*th partial sum** of (1) is

$$s_n(x) = a_0 + a_1(x-x_0) + a_2(x-x_0)^2 + \cdots + a_n(x-x_0)^n \tag{6}$$

where $n = 0, 1, \cdots$. If we omit the terms of s_n from (1), the remaining expression is

$$R_n(x) = a_{n+1}(x-x_0)^{n+1} + a_{n+2}(x-x_0)^{n+2} + \cdots . \tag{7}$$

This expression is called the **remainder** *of* (1) *after the term* $a_n(x-x_0)^n$.

For example, in the case of the geometric series

$$1 + x + x^2 + \cdots + x^n + \cdots$$

we have

$$\begin{aligned}
s_0 &= 1, & R_0 &= x + x^2 + x^3 + \cdots, \\
s_1 &= 1 + x, & R_1 &= x^2 + x^3 + x^4 + \cdots, \\
s_2 &= 1 + x + x^2, & R_2 &= x^3 + x^4 + x^5 + \cdots, \qquad \text{etc.}
\end{aligned}$$

In this way we have now associated with (1) the sequence of the partial sums $s_0(x), s_1(x), s_2(x), \cdots$. If for some $x = x_1$ this sequence converges, say,

$$\lim_{n\to\infty} s_n(x_1) = s(x_1),$$

then the series (1) is called **convergent** *at* $x = x_1$, the number $s(x_1)$ is called the **value** or *sum* of (1) at x_1, and we write

$$s(x_1) = \sum_{m=0}^{\infty} a_m(x_1 - x_0)^m.$$

Then we have for every n,

(8) $$s(x_1) = s_n(x_1) + R_n(x_1).$$

If that sequence diverges at $x = x_1$, the series (1) is called **divergent** at $x = x_1$.

In the case of convergence, for any positive ϵ there is an N (depending on ϵ) such that, by (8)

(9) $$|R_n(x_1)| = |s(x_1) - s_n(x_1)| < \epsilon \qquad \text{for all } n > N.$$

Geometrically, this means that all $s_n(x_1)$ with $n > N$ lie between $s(x_1) - \epsilon$ and $s(x_1) + \epsilon$ (Fig. 104). Practically, this means that in the case of convergence we can approximate the sum $s(x_1)$ of (1) at x_1 by $s_n(x_1)$ as accurately as we please, by taking n large enough.

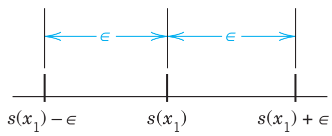


Fig. 104. Inequality (9)

Where does a power series converge? Now if we choose $x = x_0$ in (1), the series reduces to the single term a_0 because the other terms are zero. Hence the series converges at x_0. In some cases this may be the only value of x for which (1) converges. If there are other values of x for which the series converges, these values form an interval, the **convergence interval**. This interval may be finite, as in Fig. 105, with midpoint x_0. Then the series (1) converges for all x in the interior of the interval, that is, for all x for which

(10) $$|x - x_0| < R$$

and diverges for $|x - x_0| > R$. The interval may also be infinite, that is, the series may converge for all x.

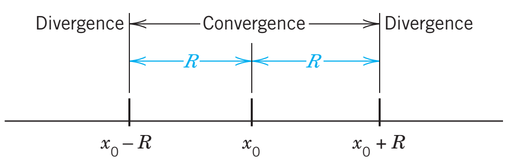


Fig. 105. Convergence interval (10) of a power series with center x_0

The quantity R in Fig. 105 is called the **radius of convergence** (because for a *complex* power series it is the radius of *disk* of convergence). If the series converges for all x, we set $R = \infty$ (and $1/R = 0$).

The radius of convergence can be determined from the coefficients of the series by means of each of the formulas

$$\text{(11)} \qquad \text{(a)} \quad R = 1\Big/\lim_{m\to\infty} \sqrt[m]{|a_m|} \qquad \text{(b)} \quad R = 1\Big/\lim_{m\to\infty} \left|\frac{a_{m+1}}{a_m}\right|$$

provided these limits exist and are not zero. [If these limits are infinite, then (1) converges only at the center x_0.]

EXAMPLE 4 Convergence Radius $R = \infty, 1, 0$

For all three series let $m \to \infty$

$$e^x = \sum_{m=0}^{\infty} \frac{x^m}{m!} = 1 + x + \frac{x^2}{2!} + \cdots, \qquad \left|\frac{a_{m+1}}{a_m}\right| = \frac{1/(m+1)!}{1/m!} = \frac{1}{m+1} \to 0, \qquad R = \infty$$

$$\frac{1}{1-x} = \sum_{m=0}^{\infty} x^m = 1 + x + x^2 + \cdots, \qquad \left|\frac{a_{m+1}}{a_m}\right| = \frac{1}{1} = 1, \qquad R = 1$$

$$\sum_{m=0}^{\infty} m!x^m = 1 + x + 2x^2 + \cdots, \qquad \left|\frac{a_{m+1}}{a_m}\right| = \frac{(m+1)!}{m!} = m + 1 \to \infty, \qquad R = 0.$$

Convergence for all x ($R = \infty$) is the best possible case, convergence in some finite interval the usual, and convergence only at the center ($R = 0$) is useless. ■

When do power series solutions exist? *Answer:* if p, q, r in the ODEs

$$\text{(12)} \qquad y'' + p(x)y' + q(x)y = r(x)$$

have power series representations (Taylor series). More precisely, a function $f(x)$ is called **analytic** at a point $x = x_0$ if it can be represented by a power series in powers of $x - x_0$ with positive radius of convergence. Using this concept, we can state the following basic theorem, in which the ODE (12) is **in standard form**, that is, it begins with the y''. If your ODE begins with, say, $h(x)y''$, divide it first by $h(x)$ and then apply the theorem to the resulting new ODE.

THEOREM 1 Existence of Power Series Solutions

If p, q, and r in (12) are analytic at $x = x_0$, then every solution of (12) is analytic at $x = x_0$ and can thus be represented by a power series in powers of $x - x_0$ with radius of convergence $R > 0$.

The proof of this theorem requires advanced complex analysis and can be found in Ref. [A11] listed in App. 1.

We mention that the radius of convergence R in Theorem 1 is at least equal to the distance from the point $x = x_0$ to the point (or points) closest to x_0 at which one of the functions p, q, r, as functions of a *complex variable*, is not analytic. (Note that that point may not lie on the x-axis but somewhere in the complex plane.)

Further Theory: Operations on Power Series

In the power series method we differentiate, add, and multiply power series, and we obtain coefficient recursions (as, for instance, in Example 3) by equating the sum of the coefficients of each occurring power of x to zero. These four operations are permissible in the sense explained in what follows. Proofs can be found in Sec. 15.3.

1. Termwise Differentiation. *A power series may be differentiated term by term.* More precisely: if

$$y(x) = \sum_{m=0}^{\infty} a_m(x - x_0)^m$$

converges for $|x - x_0| < R$, where $R > 0$, then the series obtained by differentiating term by term also converges for those x and represents the derivative y' of y for those x:

$$y'(x) = \sum_{m=1}^{\infty} m a_m(x - x_0)^{m-1} \qquad (|x - x_0| < R).$$

Similarly for the second and further derivatives.

2. Termwise Addition. *Two power series may be added term by term.* More precisely: if the series

$$\sum_{m=0}^{\infty} a_m(x - x_0)^m \quad \text{and} \quad \sum_{m=0}^{\infty} b_m(x - x_0)^m \tag{13}$$

have positive radii of convergence and their sums are $f(x)$ and $g(x)$, then the series

$$\sum_{m=0}^{\infty} (a_m + b_m)(x - x_0)^m$$

converges and represents $f(x) + g(x)$ for each x that lies in the interior of the convergence interval common to each of the two given series.

3. Termwise Multiplication. *Two power series may be multiplied term by term.* More precisely: Suppose that the series (13) have positive radii of convergence and let $f(x)$ and $g(x)$ be their sums. Then the series obtained by multiplying each term of the first series by each term of the second series and collecting like powers of $x - x_0$, that is,

$$a_0b_0 + (a_0b_1 + a_1b_0)(x - x_0) + (a_0b_2 + a_1b_1 + a_2b_0)(x - x_0)^2 + \cdots$$

$$= \sum_{m=0}^{\infty} (a_0b_m + a_1b_{m-1} + \cdots + a_mb_0)(x - x_0)^m$$

converges and represents $f(x)g(x)$ for each x in the interior of the convergence interval of each of the two given series.

4. Vanishing of All Coefficients (*"Identity Theorem for Power Series."*) *If a power series has a positive radius of convergent convergence and a sum that is identically zero throughout its interval of convergence, then each coefficient of the series must be zero.*

PROBLEM SET 5.1

1. WRITING AND LITERATURE PROJECT. Power Series in Calculus. (a) Write a review (2–3 pages) on power series in calculus. Use your own formulations and examples—do not just copy from textbooks. No proofs. **(b)** Collect and arrange Maclaurin series in a systematic list that you can use for your work.

2–5 REVIEW: RADIUS OF CONVERGENCE

Determine the radius of convergence. Show the details of your work.

2. $\sum_{m=0}^{\infty} (m+1)mx^m$

3. $\sum_{m=0}^{\infty} \frac{(-1)^m}{k^m} x^{2m}$

4. $\sum_{m=0}^{\infty} \frac{x^{2m+1}}{(2m+1)!}$

5. $\sum_{m=0}^{\infty} \left(\frac{2}{3}\right)^m x^{2m}$

6–9 SERIES SOLUTIONS BY HAND

Apply the power series method. Do this by hand, not by a CAS, to get a feel for the method, e.g., why a series may terminate, or has even powers only, etc. Show the details.

6. $(1+x)y' = y$

7. $y' = -2xy$

8. $xy' - 3y = k\ (= \text{const})$

9. $y'' + y = 0$

10–14 SERIES SOLUTIONS

Find a power series solution in powers of x. Show the details.

10. $y'' - y' + xy = 0$

11. $y'' - y' + x^2y = 0$

12. $(1 - x^2)y'' - 2xy' + 2y = 0$

13. $y'' + (1 + x^2)y = 0$

14. $y'' - 4xy' + (4x^2 - 2)y = 0$

15. Shifting summation indices is often convenient or necessary in the power series method. Shift the index so that the power under the summation sign is x^m. Check by writing the first few terms explicity.

$$\sum_{s=2}^{\infty} \frac{s(s+1)}{s^2+1} x^{s-1}, \qquad \sum_{p=1}^{\infty} \frac{p^2}{(p+1)!} x^{p+4}$$

16–19 CAS PROBLEMS. IVPs

Solve the initial value problem by a power series. Graph the partial sums of the powers up to and including x^5. Find the value of the sum s (5 digits) at x_1.

16. $y' + 4y = 1, \quad y(0) = 1.25, \quad x_1 = 0.2$

17. $y'' + 3xy' + 2y = 0, \quad y(0) = 1, \quad y'(0) = 1,$ $x = 0.5$

18. $(1 - x^2)y'' - 2xy' + 30y = 0, \quad y(0) = 0,$ $y'(0) = 1.875, \quad x_1 = 0.5$

19. $(x - 2)y' = xy, \quad y(0) = 4, \quad x_1 = 2$

20. CAS Experiment. Information from Graphs of Partial Sums. In numerics we use partial sums of power series. To get a feel for the accuracy for various x, experiment with $\sin x$. Graph partial sums of the Maclaurin series of an increasing number of terms, describing qualitatively the "breakaway points" of these graphs from the graph of $\sin x$. Consider other Maclaurin series of your choice.

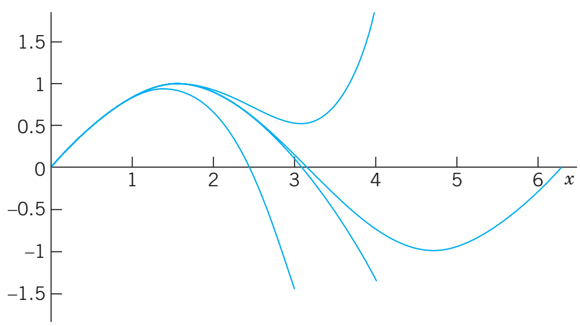


Fig. 106. CAS Experiment 20. $\sin x$ and partial sums s_3, s_5, s_7

5.2 Legendre's Equation. Legendre Polynomials $P_n(x)$

Legendre's differential equation[1]

$$(1 - x^2)y'' - 2xy' + n(n + 1)y = 0 \qquad (n \text{ constant}) \tag{1}$$

is one of the most important ODEs in physics. It arises in numerous problems, particularly in boundary value problems for spheres (take a quick look at Example 1 in Sec. 12.10).

The equation involves a **parameter** n, whose value depends on the physical or engineering problem. So (1) is actually a whole family of ODEs. For $n = 1$ we solved it in Example 3 of Sec. 5.1 (look back at it). Any solution of (1) is called a **Legendre function**. The study of these and other "higher" functions not occurring in calculus is called the **theory of special functions**. Further special functions will occur in the next sections.

Dividing (1) by $1 - x^2$, we obtain the standard form needed in Theorem 1 of Sec. 5.1 and we see that the coefficients $-2x/(1 - x^2)$ and $n(n + 1)/(1 - x^2)$ of the new equation are analytic at $x = 0$, so that we may apply the power series method. Substituting

$$y = \sum_{m=0}^{\infty} a_m x^m \tag{2}$$

and its derivatives into (1), and denoting the constant $n(n + 1)$ simply by k, we obtain

$$(1 - x^2) \sum_{m=2}^{\infty} m(m - 1)a_m x^{m-2} - 2x \sum_{m=1}^{\infty} m a_m x^{m-1} + k \sum_{m=0}^{\infty} a_m x^m = 0.$$

By writing the first expression as two separate series we have the equation

$$\sum_{m=2}^{\infty} m(m - 1)a_m x^{m-2} - \sum_{m=2}^{\infty} m(m - 1)a_m x^m - \sum_{m=1}^{\infty} 2m a_m x^m + \sum_{m=0}^{\infty} k a_m x^m = 0.$$

It may help you to write out the first few terms of each series explicitly, as in Example 3 of Sec. 5.1; or you may continue as follows. To obtain the same general power x^s in all four series, set $m - 2 = s$ (thus $m = s + 2$) in the first series and simply write s instead of m in the other three series. This gives

$$\sum_{s=0}^{\infty} (s + 2)(s + 1)a_{s+2} x^s - \sum_{s=2}^{\infty} s(s - 1)a_s x^s - \sum_{s=1}^{\infty} 2s a_s x^s + \sum_{s=0}^{\infty} k a_s x^s = 0.$$

---

[1]ADRIEN-MARIE LEGENDRE (1752–1833), French mathematician, who became a professor in Paris in 1775 and made important contributions to special functions, elliptic integrals, number theory, and the calculus of variations. His book *Éléments de géométrie* (1794) became very famous and had 12 editions in less than 30 years.

Formulas on Legendre functions may be found in Refs. [GenRef1] and [GenRef10].

(Note that in the first series the summation begins with $s = 0$.) Since this equation with the right side 0 must be an identity in x if (2) is to be a solution of (1), the sum of the coefficients of each power of x on the left must be zero. Now x^0 occurs in the first and fourth series only, and gives [remember that $k = n(n + 1)$]

$$2 \cdot 1 a_2 + n(n + 1)a_0 = 0. \tag{3a}$$

x^1 occurs in the first, third, and fourth series and gives

$$3 \cdot 2 a_3 + [-2 + n(n + 1)]a_1 = 0. \tag{3b}$$

The higher powers $x^2, x^3, \cdots$ occur in all four series and give

$$(s + 2)(s + 1)a_{s+2} + [-s(s - 1) - 2s + n(n + 1)]a_s = 0. \tag{3c}$$

The expression in the brackets $[\cdots]$ can be written $(n - s)(n + s + 1)$, as you may readily verify. Solving (3a) for a_2 and (3b) for a_3 as well as (3c) for a_{s+2}, we obtain the general formula

$$\boxed{a_{s+2} = -\frac{(n - s)(n + s + 1)}{(s + 2)(s + 1)} a_s} \qquad (s = 0, 1, \cdots). \tag{4}$$

This is called a **recurrence relation** or **recursion formula**. (Its derivation you may verify with your CAS.) It gives each coefficient in terms of the second one preceding it, except for a_0 and a_1, which are left as arbitrary constants. We find successively

$$a_2 = -\frac{n(n + 1)}{2!} a_0 \qquad\qquad a_3 = -\frac{(n - 1)(n + 2)}{3!} a_1$$

$$a_4 = -\frac{(n - 2)(n + 3)}{4 \cdot 3} a_2 \qquad\qquad a_5 = -\frac{(n - 3)(n + 4)}{5 \cdot 4} a_3$$

$$= \frac{(n - 2)n(n + 1)(n + 3)}{4!} a_0 \qquad\qquad = \frac{(n - 3)(n - 1)(n + 2)(n + 4)}{5!} a_1$$

and so on. By inserting these expressions for the coefficients into (2) we obtain

$$y(x) = a_0 y_1(x) + a_1 y_2(x) \tag{5}$$

where

$$\boxed{y_1(x) = 1 - \frac{n(n + 1)}{2!} x^2 + \frac{(n - 2)n(n + 1)(n + 3)}{4!} x^4 - + \cdots} \tag{6}$$

$$\boxed{y_2(x) = x - \frac{(n - 1)(n + 2)}{3!} x^3 + \frac{(n - 3)(n - 1)(n + 2)(n + 4)}{5!} x^5 - + \cdots.} \tag{7}$$

These series converge for $|x| < 1$ (see Prob. 4; or they may terminate, see below). Since (6) contains even powers of x only, while (7) contains odd powers of x only, the ratio y_1/y_2 is not a constant, so that y_1 and y_2 are not proportional and are thus linearly independent solutions. Hence (5) is a general solution of (1) on the interval $-1 < x < 1$.

Note that $x = \pm 1$ are the points at which $1 - x^2 = 0$, so that the coefficients of the standardized ODE are no longer analytic. So it should not surprise you that we do not get a longer convergence interval of (6) and (7), unless these series terminate after finitely many powers. In that case, the series become polynomials.

Polynomial Solutions. Legendre Polynomials $P_n(x)$

The reduction of power series to polynomials is a great advantage because then we have solutions for all x, without convergence restrictions. For special functions arising as solutions of ODEs this happens quite frequently, leading to various important families of polynomials; see Refs. [GenRef1], [GenRef10] in App. 1. For Legendre's equation this happens when the parameter n is a nonnegative integer because then the right side of (4) is zero for $s = n$, so that $a_{n+2} = 0$, $a_{n+4} = 0$, $a_{n+6} = 0, \cdots$. Hence if n is even, $y_1(x)$ reduces to a polynomial of degree n. If n is odd, the same is true for $y_2(x)$. These polynomials, multiplied by some constants, are called **Legendre polynomials** and are denoted by $P_n(x)$. The standard choice of such constants is done as follows. We choose the coefficient a_n of the highest power x^n as

$$a_n = \frac{(2n)!}{2^n (n!)^2} = \frac{1 \cdot 3 \cdot 5 \cdots (2n-1)}{n!} \qquad (n \text{ a positive integer}) \tag{8}$$

(and $a_n = 1$ if $n = 0$). Then we calculate the other coefficients from (4), solved for a_s in terms of a_{s+2}, that is,

$$a_s = -\frac{(s+2)(s+1)}{(n-s)(n+s+1)} a_{s+2} \qquad (s \leqq n-2). \tag{9}$$

The choice (8) makes $p_n(1) = 1$ for every n (see Fig. 107); this motivates (8). From (9) with $s = n - 2$ and (8) we obtain

$$a_{n-2} = -\frac{n(n-1)}{2(2n-1)} a_n = -\frac{n(n-1)}{2(2n-1)} \cdot \frac{(2n)!}{2^n (n!)^2}$$

Using $(2n)! = 2n(2n-1)(2n-2)!$ in the numerator and $n! = n(n-1)!$ and $n! = n(n-1)(n-2)!$ in the denominator, we obtain

$$a_{n-2} = -\frac{n(n-1)2n(2n-1)(2n-2)!}{2(2n-1)2^n n(n-1)!\, n(n-1)(n-2)!}.$$

$n(n-1)2n(2n-1)$ cancels, so that we get

$$a_{n-2} = -\frac{(2n-2)!}{2^n (n-1)!\,(n-2)!}.$$

Similarly,

$$a_{n-4} = -\frac{(n-2)(n-3)}{4(2n-3)} a_{n-2}$$
$$= \frac{(2n-4)!}{2^n 2!\,(n-2)!\,(n-4)!}$$

and so on, and in general, when $n - 2m \geqq 0$,

$$a_{n-2m} = (-1)^m \frac{(2n-2m)!}{2^n m!\,(n-m)!\,(n-2m)!}. \tag{10}$$

The resulting solution of Legendre's differential equation (1) is called the **Legendre polynomial** *of degree n* and is denoted by $P_n(x)$.

From (10) we obtain

$$\begin{aligned} P_n(x) &= \sum_{m=0}^{M} (-1)^m \frac{(2n-2m)!}{2^n m!\,(n-m)!\,(n-2m)!} x^{n-2m} \\ &= \frac{(2n)!}{2^n (n!)^2} x^n - \frac{(2n-2)!}{2^n 1!\,(n-1)!\,(n-2)!} x^{n-2} + - \cdots \end{aligned} \tag{11}$$

where $M = n/2$ or $(n-1)/2$, whichever is an integer. The first few of these functions are (Fig. 107)

$$\begin{aligned} &P_0(x) = 1, && P_1(x) = x \\ &P_2(x) = \tfrac{1}{2}(3x^2 - 1), && P_3(x) = \tfrac{1}{2}(5x^3 - 3x) \\ &P_4(x) = \tfrac{1}{8}(35x^4 - 30x^2 + 3), && P_5(x) = \tfrac{1}{8}(63x^5 - 70x^3 + 15x) \end{aligned} \tag{11$'$}$$

and so on. You may now program (11) on your CAS and calculate $P_n(x)$ as needed.

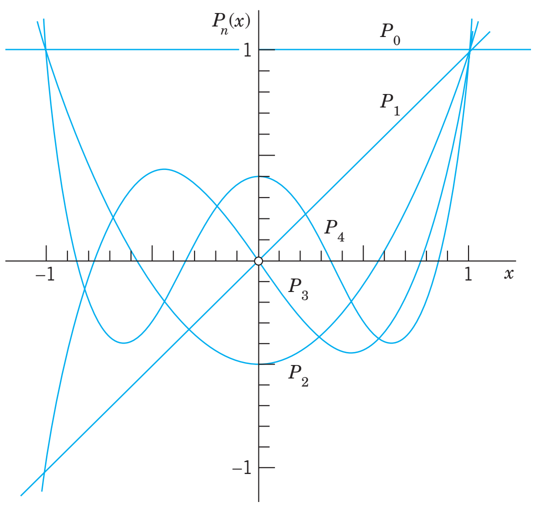


Fig. 107. Legendre polynomials

The Legendre polynomials $P_n(x)$ are **orthogonal** on the interval $-1 \leqq x \leqq 1$, a basic property to be defined and used in making up "Fourier–Legendre series" in the chapter on Fourier series (see Secs. 11.5–11.6).

PROBLEM SET 5.2

1–5 LEGENDRE POLYNOMIALS AND FUNCTIONS

1. **Legendre functions for $n = 0$.** Show that (6) with $n = 0$ gives $P_0(x) = 1$ and (7) gives (use $\ln(1 + x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 + \cdots$)

$$y_2(x) = x + \frac{1}{3}x^3 + \frac{1}{5}x^5 + \cdots = \frac{1}{2}\ln\frac{1+x}{1-x}.$$

Verify this by solving (1) with $n = 0$, setting $z = y'$ and separating variables.

2. **Legendre functions for $n = 1$.** Show that (7) with $n = 1$ gives $y_2(x) = P_1(x) = x$ and (6) gives

$$y_1 = 1 - x^2 - \frac{1}{3}x^4 - \frac{1}{5}x^6 - \cdots$$
$$= 1 - \frac{1}{2}x\ln\frac{1+x}{1-x}.$$

3. **Special n.** Derive (11′) from (11).

4. **Legendre's ODE.** Verify that the polynomials in (11′) satisfy (1).

5. Obtain P_6 and P_7.

6–9 CAS PROBLEMS

6. Graph $P_2(x), \cdots, P_{10}(x)$ on common axes. For what x (approximately) and $n = 2, \cdots, 10$ is $|P_n(x)| < \frac{1}{2}$?

7. From what n on will your CAS no longer produce faithful graphs of $P_n(x)$? Why?

8. Graph $Q_0(x), Q_1(x)$, and some further Legendre functions.

9. Substitute $a_s x^s + a_{s+1}x^{s+1} + a_{s+2}x^{s+2}$ into Legendre's equation and obtain the coefficient recursion (4).

10. **TEAM PROJECT. Generating Functions.** Generating functions play a significant role in modern applied mathematics (see [GenRef5]). The idea is simple. If we want to study a certain sequence $(f_n(x))$ and can find a function

$$G(u, x) = \sum_{n=0}^{\infty} f_n(x)u^n,$$

we may obtain properties of $(f_n(x))$ from those of G, which "generates" this sequence and is called a **generating function** of the sequence.

(a) Legendre polynomials. Show that

$$G(u, x) = \frac{1}{\sqrt{1 - 2xu + u^2}} = \sum_{n=0}^{\infty} P_n(x)u^n \tag{12}$$

is a generating function of the Legendre polynomials. *Hint:* Start from the binomial expansion of $1/\sqrt{1 - v}$, then set $v = 2xu - u^2$, multiply the powers of $2xu - u^2$ out, collect all the terms involving u^n, and verify that the sum of these terms is $P_n(x)u^n$.

(b) Potential theory. Let A_1 and A_2 be two points in space (Fig. 108, $r_2 > 0$). Using (12), show that

$$\frac{1}{r} = \frac{1}{\sqrt{r_1^2 + r_2^2 - 2r_1r_2\cos\theta}}$$
$$= \frac{1}{r_2}\sum_{m=0}^{\infty} P_m(\cos\theta)\left(\frac{r_1}{r_2}\right)^m.$$

This formula has applications in potential theory. (Q/r is the electrostatic potential at A_2 due to a charge Q located at A_1. And the series expresses $1/r$ in terms of the distances of A_1 and A_2 from any origin O and the angle θ between the segments OA_1 and OA_2.)

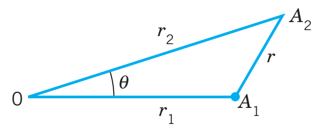


Fig. 108. Team Project 10

(c) Further applications of (12). Show that $P_n(1) = 1$, $P_n(-1) = (-1)^n$, $P_{2n+1}(0) = 0$, and $P_{2n}(0) = (-1)^n \cdot 1 \cdot 3 \cdots (2n - 1)/[2 \cdot 4 \cdots (2n)]$.

11–15 FURTHER FORMULAS

11. **ODE.** Find a solution of $(a^2 - x^2)y'' - 2xy' + n(n + 1)y = 0$, $a \neq 0$, by reduction to the Legendre equation.

12. **Rodrigues's formula (13)**[2] Applying the binomial theorem to $(x^2 - 1)^n$, differentiating it n times term by term, and comparing the result with (11), show that

$$P_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}[(x^2 - 1)^n]. \tag{13}$$

[2]OLINDE RODRIGUES (1794–1851), French mathematician and economist.

13. Rodrigues's formula. Obtain (11′) from (13).

14. Bonnet's recursion.[3] Differentiating (13) with respect to u, using (13) in the resulting formula, and comparing coefficients of u^n, obtain the *Bonnet recursion.*

(14) $(n+1)P_{n+1}(x) = (2n+1)xP_n(x) - np_{n-1}(x),$

where $n = 1, 2, \cdots$. This formula is useful for computations, the loss of significant digits being small (except near zeros). Try (14) out for a few computations of your own choice.

15. Associated Legendre functions $P_n^k(x)$ are needed, e.g., in quantum physics. They are defined by

$$P_n^k(x) = (1 - x^2)^{k/2} \frac{d^k p_n(x)}{dx^k} \qquad \textbf{(15)}$$

and are solutions of the ODE

$$(1 - x^2)y'' - 2xy' + q(x)y = 0 \qquad \textbf{(16)}$$

where $q(x) = n(n+1) - k^2/(1 - x^2)$. Find $P_1^1(x)$, $P_2^1(x)$, $P_2^2(x)$, and $P_4^2(x)$ and verify that they satisfy (16).

5.3 Extended Power Series Method: Frobenius Method

Several second-order ODEs of considerable practical importance—the famous Bessel equation among them—have coefficients that are not analytic (definition in Sec. 5.1), but are "not too bad," so that these ODEs can still be solved by series (power series times a logarithm or times a fractional power of x, etc.). Indeed, the following theorem permits an extension of the power series method. The new method is called the **Frobenius method.**[4] Both methods, that is, the power series method and the Frobenius method, have gained in significance due to the use of software in actual calculations.

THEOREM 1

Frobenius Method

Let $b(x)$ and $c(x)$ be any functions that are analytic at $x = 0$. Then the ODE

$$y'' + \frac{b(x)}{x} y' + \frac{c(x)}{x^2} y = 0 \qquad \textbf{(1)}$$

has at least one solution that can be represented in the form

$$y(x) = x^r \sum_{m=0}^{\infty} a_m x^m = x^r(a_0 + a_1 x + a_2 x^2 + \cdots) \qquad (a_0 \neq 0) \qquad (2)$$

where the exponent r may be any (real or complex) number (and r is chosen so that $a_0 \neq 0$).

The ODE (1) *also has a second solution (such that these two solutions are linearly independent) that may be similar to* (2) *(with a different r and different coefficients) or may contain a logarithmic term.* (Details in Theorem 2 below.)

[3]OSSIAN BONNET (1819–1892), French mathematician, whose main work was in differential geometry.

[4]GEORG FROBENIUS (1849–1917), German mathematician, professor at ETH Zurich and University of Berlin, student of Karl Weierstrass (see footnote, Sect. 15.5). He is also known for his work on matrices and in group theory.

In this theorem we may replace x by $x - x_0$ with any number x_0. The condition $a_0 \neq 0$ is no restriction; it simply means that we factor out the highest possible power of x.

The singular point of (1) at $x = 0$ is often called a **regular singular point**, a term confusing to the student, which we shall not use.

For example, Bessel's equation (to be discussed in the next section)

$$y'' + \frac{1}{x}y' + \left(\frac{x^2 - \nu^2}{x^2}\right)y = 0 \qquad (\nu \text{ a parameter})$$

is of the form (1) with $b(x) = 1$ and $c(x) = x^2 - \nu^2$ analytic at $x = 0$, so that the theorem applies. This ODE could not be handled in full generality by the power series method.

Similarly, the so-called hypergeometric differential equation (see Problem Set 5.3) also requires the Frobenius method.

The point is that in (2) we have a power series times a single power of x whose exponent r is not restricted to be a nonnegative integer. (The latter restriction would make the whole expression a power series, by definition; see Sec. 5.1.)

The proof of the theorem requires advanced methods of complex analysis and can be found in Ref. [A11] listed in App. 1.

Regular and Singular Points. The following terms are practical and commonly used. A **regular point** of the ODE

$$y'' + p(x)y' + q(x)y = 0$$

is a point x_0 at which the coefficients p and q are analytic. Similarly, a **regular point** of the ODE

$$\tilde{h}(x)y'' + \tilde{p}(x)y'(x) + \tilde{q}(x)y = 0$$

is an x_0 at which $\tilde{h}, \tilde{p}, \tilde{q}$ are analytic and $\tilde{h}(x_0) \neq 0$ (so what we can divide by $\tilde{h}$ and get the previous standard form). Then the power series method can be applied. If x_0 is not a regular point, it is called a **singular point**.

Indicial Equation, Indicating the Form of Solutions

We shall now explain the Frobenius method for solving (1). Multiplication of (1) by x^2 gives the more convenient form

$$x^2y'' + xb(x)y' + c(x)y = 0. \tag{1'}$$

We first expand $b(x)$ and $c(x)$ in power series,

$$b(x) = b_0 + b_1x + b_2x^2 + \cdots, \qquad c(x) = c_0 + c_1x + c_2x^2 + \cdots$$

or we do nothing if $b(x)$ and $c(x)$ are polynomials. Then we differentiate (2) term by term, finding

$$\begin{aligned} y'(x) &= \sum_{m=0}^{\infty} (m + r)a_m x^{m+r-1} = x^{r-1}[ra_0 + (r + 1)a_1x + \cdots] \\ y''(x) &= \sum_{m=0}^{\infty} (m + r)(m + r - 1)a_m x^{m+r-2} \\ &= x^{r-2}[r(r - 1)a_0 + (r + 1)ra_1x + \cdots]. \end{aligned} \tag{2*}$$

By inserting all these series into (1′) we obtain

$$x^r[r(r-1)a_0 + \cdots] + (b_0 + b_1x + \cdots)x^r(ra_0 + \cdots) + (c_0 + c_1x + \cdots)x^r(a_0 + a_1x + \cdots) = 0. \tag{3}$$

We now equate the sum of the coefficients of each power $x^r, x^{r+1}, x^{r+2}, \cdots$ to zero. This yields a system of equations involving the unknown coefficients a_m. The smallest power is x^r and the corresponding equation is

$$[r(r-1) + b_0r + c_0]a_0 = 0.$$

Since by assumption $a_0 \neq 0$, the expression in the brackets $[\cdots]$ must be zero. This gives

$$r(r-1) + b_0r + c_0 = 0. \tag{4}$$

This important quadratic equation is called the **indicial equation** of the ODE (1). Its role is as follows.

The Frobenius method yields a basis of solutions. One of the two solutions will always be of the form (2), where r is a root of (4). The other solution will be of a form indicated by the indicial equation. There are three cases:

Case 1. Distinct roots not differing by an integer $1, 2, 3, \cdots$.

Case 2. A double root.

Case 3. Roots differing by an integer $1, 2, 3, \cdots$.

Cases 1 and 2 are not unexpected because of the Euler–Cauchy equation (Sec. 2.5), the simplest ODE of the form (1). Case 1 includes complex conjugate roots r_1 and $r_2 = \bar{r}_1$ because $r_1 - r_2 = r_1 - \bar{r}_1 = 2i \operatorname{Im} r_1$ is imaginary, so it cannot be a *real* integer. The form of a basis will be given in Theorem 2 (which is proved in App. 4), without a general theory of convergence, but convergence of the occurring series can be tested in each individual case as usual. Note that in Case 2 *we **must*** have a logarithm, whereas in Case 3 we *may* or *may not*.

THEOREM 2

Frobenius Method. Basis of Solutions. Three Cases

Suppose that the ODE (1) *satisfies the assumptions in Theorem* 1. *Let r_1 and r_2 be the roots of the indicial equation* (4). *Then we have the following three cases.*

***Case* 1. *Distinct Roots Not Differing by an Integer*.** *A basis is*

$$y_1(x) = x^{r_1}(a_0 + a_1x + a_2x^2 + \cdots) \tag{5}$$

and

$$y_2(x) = x^{r_2}(A_0 + A_1x + A_2x^2 + \cdots) \tag{6}$$

with coefficients obtained successively from (3) *with $r = r_1$ and $r = r_2$, respectively.*

Case 2. Double Root $r_1 = r_2 = r$. *A basis is*

$$y_1(x) = x^r(a_0 + a_1x + a_2x^2 + \cdots) \qquad [r = \tfrac{1}{2}(1 - b_0)] \tag{7}$$

(*of the same general form as before*) *and*

$$y_2(x) = y_1(x)\ln x + x^r(A_1x + A_2x^2 + \cdots) \qquad (x > 0). \tag{8}$$

Case 3. Roots Differing by an Integer. *A basis is*

$$y_1(x) = x^{r_1}(a_0 + a_1x + a_2x^2 + \cdots) \tag{9}$$

(*of the same general form as before*) *and*

$$y_2(x) = ky_1(x)\ln x + x^{r_2}(A_0 + A_1x + A_2x^2 + \cdots), \tag{10}$$

where the roots are so denoted that $r_1 - r_2 > 0$ *and* k *may turn out to be zero.*

Typical Applications

Technically, the Frobenius method is similar to the power series method, once the roots of the indicial equation have been determined. However, (5)–(10) merely indicate the general form of a basis, and a second solution can often be obtained more rapidly by reduction of order (Sec. 2.1).

EXAMPLE 1 Euler–Cauchy Equation, Illustrating Cases 1 and 2 and Case 3 without a Logarithm

For the Euler–Cauchy equation (Sec. 2.5)

$$x^2y'' + b_0xy' + c_0y = 0 \qquad (b_0, c_0 \text{ constant})$$

substitution of $y = x^r$ gives the auxiliary equation

$$r(r-1) + b_0r + c_0 = 0,$$

which is the indicial equation [and $y = x^r$ is a very special form of (2)!]. For different roots r_1, r_2 we get a basis $y_1 = x^{r_1}$, $y_2 = x^{r_2}$, and for a double root r we get a basis $x^r, x^r \ln x$. Accordingly, for this simple ODE, Case 3 plays no extra role. ■

EXAMPLE 2 Illustration of Case 2 (Double Root)

Solve the ODE

$$x(x-1)y'' + (3x-1)y' + y = 0. \tag{11}$$

(This is a special hypergeometric equation, as we shall see in the problem set.)

Solution. Writing (11) in the standard form (1), we see that it satisfies the assumptions in Theorem 1. [What are $b(x)$ and $c(x)$ in (11)?] By inserting (2) and its derivatives (2*) into (11) we obtain

$$\begin{aligned}\sum_{m=0}^{\infty}(m+r)(m+r-1)a_mx^{m+r} &- \sum_{m=0}^{\infty}(m+r)(m+r-1)a_mx^{m+r-1}\\ + 3\sum_{m=0}^{\infty}(m+r)a_mx^{m+r} &- \sum_{m=0}^{\infty}(m+r)a_mx^{m+r-1} + \sum_{m=0}^{\infty}a_mx^{m+r} = 0.\end{aligned} \tag{12}$$

The smallest power is x^{r-1}, occurring in the second and the fourth series; by equating the sum of its coefficients to zero we have

$$[-r(r-1)-r]a_0 = 0, \qquad \text{thus} \qquad r^2 = 0.$$

Hence this indicial equation has the double root $r = 0$.

First Solution. We insert this value $r = 0$ into (12) and equate the sum of the coefficients of the power x^s to zero, obtaining

$$s(s-1)a_s - (s+1)sa_{s+1} + 3sa_s - (s+1)a_{s+1} + a_s = 0$$

thus $a_{s+1} = a_s$. Hence $a_0 = a_1 = a_2 = \cdots$, and by choosing $a_0 = 1$ we obtain the solution

$$y_1(x) = \sum_{m=0}^{\infty} x^m = \frac{1}{1-x} \qquad (|x| < 1).$$

Second Solution. We get a second independent solution y_2 by the method of reduction of order (Sec. 2.1), substituting $y_2 = uy_1$ and its derivatives into the equation. This leads to (9), Sec. 2.1, which we shall use in this example, instead of starting reduction of order from scratch (as we shall do in the next example). In (9) of Sec. 2.1 we have $p = (3x-1)/(x^2-x)$, the coefficient of y' in (11) ***in standard form***. By partial fractions,

$$-\int p\,dx = -\int \frac{3x-1}{x(x-1)}\,dx = -\int\left(\frac{2}{x-1} + \frac{1}{x}\right)dx = -2\ln(x-1) - \ln x.$$

Hence (9), Sec. 2.1, becomes

$$u' = U = y_1^{-2}e^{-\int p\,dx} = \frac{(x-1)^2}{(x-1)^2 x} = \frac{1}{x}, \qquad u = \ln x, \qquad y_2 = uy_1 = \frac{\ln x}{1-x}.$$

y_1 and y_2 are shown in Fig. 109. These functions are linearly independent and thus form a basis on the interval $0 < x < 1$ (as well as on $1 < x < \infty$). ■

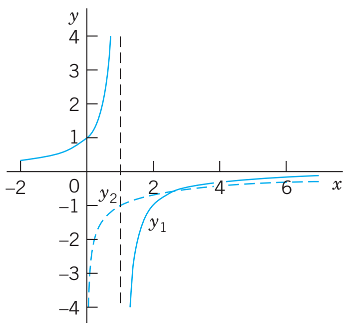


Fig. 109. Solutions in Example 2

EXAMPLE 3 Case 3, Second Solution with Logarithmic Term

Solve the ODE

$$(x^2-x)y'' - xy' + y = 0. \tag{13}$$

Solution. Substituting (2) and (2*) into (13), we have

$$(x^2-x)\sum_{m=0}^{\infty}(m+r)(m+r-1)a_m x^{m+r-2} - x\sum_{m=0}^{\infty}(m+r)a_m x^{m+r-1} + \sum_{m=0}^{\infty}a_m x^{m+r} = 0.$$

We now take x^2, x, and x inside the summations and collect all terms with power x^{m+r} and simplify algebraically,

$$\sum_{m=0}^{\infty} (m + r - 1)^2 a_m x^{m+r} - \sum_{m=0}^{\infty} (m + r)(m + r - 1) a_m x^{m+r-1} = 0.$$

In the first series we set $m = s$ and in the second $m = s + 1$, thus $s = m - 1$. Then

$$\sum_{s=0}^{\infty} (s + r - 1)^2 a_s x^{s+r} - \sum_{s=-1}^{\infty} (s + r + 1)(s + r) a_{s+1} x^{s+r} = 0. \tag{14}$$

The lowest power is x^{r-1} (take $s = -1$ in the second series) and gives the indicial equation

$$r(r - 1) = 0.$$

The roots are $r_1 = 1$ and $r_2 = 0$. They differ by an integer. This is Case 3.

First Solution. From (14) with $r = r_1 = 1$ we have

$$\sum_{s=0}^{\infty} [s^2 a_s - (s + 2)(s + 1) a_{s+1}] x^{s+1} = 0.$$

This gives the recurrence relation

$$a_{s+1} = \frac{s^2}{(s + 2)(s + 1)} a_s \qquad (s = 0, 1, \cdots).$$

Hence $a_1 = 0, a_2 = 0, \cdots$ successively. Taking $a_0 = 1$, we get as a first solution $y_1 = x^{r_1} a_0 = x$.

Second Solution. Applying reduction of order (Sec. 2.1), we substitute $y_2 = y_1 u = xu$, $y_2' = xu' + u$ and $y_2'' = xu'' + 2u'$ into the ODE, obtaining

$$(x^2 - x)(xu'' + 2u') - x(xu' + u) + xu = 0.$$

xu drops out. Division by x and simplification give

$$(x^2 - x)u'' + (x - 2)u' = 0.$$

From this, using partial fractions and integrating (taking the integration constant zero), we get

$$\frac{u''}{u'} = -\frac{x - 2}{x^2 - x} = -\frac{2}{x} + \frac{1}{1 - x}, \qquad \ln u' = \ln \left| \frac{x - 1}{x^2} \right|.$$

Taking exponents and integrating (again taking the integration constant zero), we obtain

$$u' = \frac{x - 1}{x^2} = \frac{1}{x} - \frac{1}{x^2}, \qquad u = \ln x + \frac{1}{x}, \qquad y_2 = xu = x \ln x + 1.$$

y_1 and y_2 are linearly independent, and y_2 has a logarithmic term. Hence y_1 and y_2 constitute a basis of solutions for positive x. ■

The Frobenius method solves the **hypergeometric equation**, whose solutions include many known functions as special cases (see the problem set). In the next section we use the method for solving Bessel's equation.

PROBLEM SET 5.3

1. **WRITING PROJECT. Power Series Method and Frobenius Method.** Write a report of 2–3 pages explaining the difference between the two methods. No proofs. Give simple examples of your own.

2–13 FROBENIUS METHOD

Find a basis of solutions by the Frobenius method. Try to identify the series as expansions of known functions. Show the details of your work.

2. $(x + 2)^2y'' + (x + 2)y' - y = 0$
3. $xy'' + 2y' + xy = 0$
4. $xy'' + y = 0$
5. $xy'' + (2x + 1)y' + (x + 1)y = 0$
6. $xy'' + 2x^3y' + (x^2 - 2)y = 0$
7. $y'' + (x - 1)y = 0$
8. $xy'' + y' - xy = 0$
9. $2x(x - 1)y'' - (x + 1)y' + y = 0$
10. $xy'' + 2y' + 4xy = 0$
11. $xy'' + (2 - 2x)y' + (x - 2)y = 0$
12. $x^2y'' + 6xy' + (4x^2 + 6)y = 0$
13. $xy'' + (1 - 2x)y' + (x - 1)y = 0$

14. **TEAM PROJECT. Hypergeometric Equation, Series, and Function. Gauss's hypergeometric ODE**[5] is

$$x(1 - x)y'' + [c - (a + b + 1)x]y' - aby = 0. \tag{15}$$

Here, a, b, c are constants. This ODE is of the form $p_2y'' + p_1y' + p_0y = 0$, where p_2, p_1, p_0 are polynomials of degree 2, 1, 0, respectively. These polynomials are written so that the series solution takes a most practical form, namely,

$$\begin{aligned} y_1(x) = 1 &+ \frac{ab}{1!\,c}x + \frac{a(a+1)b(b+1)}{2!\,c(c+1)}x^2 \\ &+ \frac{a(a+1)(a+2)b(b+1)(b+2)}{3!\,c(c+1)(c+2)}x^3 + \cdots. \end{aligned} \tag{16}$$

This series is called the **hypergeometric series**. Its sum $y_1(x)$ is called the **hypergeometric function** and is denoted by $F(a, b, c; x)$. Here, $c \neq 0, -1, -2, \cdots$. By choosing specific values of a, b, c we can obtain an incredibly large number of special functions as solutions of (15) [see the small sample of elementary functions in part (c)]. This accounts for the importance of (15).

(a) Hypergeometric series and function. Show that the indicial equation of (15) has the roots $r_1 = 0$ and $r_2 = 1 - c$. Show that for $r_1 = 0$ the Frobenius method gives (16). Motivate the name for (16) by showing that

$$F(1, 1, 1; x) = F(1, b, b; x) = F(a, 1, a; x) = \frac{1}{1 - x}.$$

(b) Convergence. For what a or b will (16) reduce to a polynomial? Show that for any other a, b, c ($c \neq 0, -1, -2, \cdots$) the series (16) converges when $|x| < 1$.

(c) Special cases. Show that

$$\begin{aligned} (1 + x)^n &= F(-n, b, b; -x), \\ (1 - x)^n &= 1 - nxF(1 - n, 1, 2; x), \\ \arctan x &= xF(\tfrac{1}{2}, 1, \tfrac{3}{2}; -x^2) \\ \arcsin x &= xF(\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{3}{2}; x^2), \\ \ln(1 + x) &= xF(1, 1, 2; -x), \\ \ln\frac{1 + x}{1 - x} &= 2xF(\tfrac{1}{2}, 1, \tfrac{3}{2}; x^2). \end{aligned}$$

Find more such relations from the literature on special functions, for instance, from [GenRef1] in App. 1.

(d) Second solution. Show that for $r_2 = 1 - c$ the Frobenius method yields the following solution (where $c \neq 2, 3, 4, \cdots$):

$$\begin{aligned} y_2(x) = x^{1-c}\biggl(1 &+ \frac{(a-c+1)(b-c+1)}{1!\,(-c+2)}x \\ &+ \frac{(a-c+1)(a-c+2)(b-c+1)(b-c+2)}{2!\,(-c+2)(-c+3)}x^2 \\ &+ \cdots\biggr). \end{aligned} \tag{17}$$

Show that

$$y_2(x) = x^{1-c}F(a - c + 1, b - c + 1, 2 - c; x).$$

(e) On the generality of the hypergeometric equation. Show that

$$(t^2 + At + B)\ddot{y} + (Ct + D)\dot{y} + Ky = 0 \tag{18}$$

[5] CARL FRIEDRICH GAUSS (1777–1855), great German mathematician. He already made the first of his great discoveries as a student at Helmstedt and Göttingen. In 1807 he became a professor and director of the Observatory at Göttingen. His work was of basic importance in algebra, number theory, differential equations, differential geometry, non-Euclidean geometry, complex analysis, numeric analysis, astronomy, geodesy, electromagnetism, and theoretical mechanics. He also paved the way for a general and systematic use of complex numbers.

with $\dot{y} = dy/dt$, etc., constant A, B, C, D, K, and $t^2 + At + B = (t - t_1)(t - t_2)$, $t_1 \neq t_2$, can be reduced to the hypergeometric equation with independent variable

$$x = \frac{t - t_1}{t_2 - t_1}$$

and parameters related by $Ct_1 + D = -c(t_2 - t_1)$, $C = a + b + 1$, $K = ab$. From this you see that (15) is a "normalized form" of the more general (18) and that various cases of (18) can thus be solved in terms of hypergeometric functions.

15–20 HYPERGEOMETRIC ODE

Find a general solution in terms of hypergeometric functions.

15. $2x(1 - x)y'' - (1 + 6x)y' - 2y = 0$

16. $x(1 - x)y'' + (\frac{1}{2} + 2x)y' - 2y = 0$

17. $4x(1 - x)y'' + y' + 8y = 0$

18. $4(t^2 - 3t + 2)\ddot{y} - 2\dot{y} + y = 0$

19. $2(t^2 - 5t + 6)\ddot{y} + (2t - 3)\dot{y} - 8y = 0$

20. $3t(1 + t)\ddot{y} + t\dot{y} - y = 0$

5.4 Bessel's Equation. Bessel Functions $J_\nu(x)$

One of the most important ODEs in applied mathematics in **Bessel's equation**,[6]

$$x^2y'' + xy' + (x^2 - \nu^2)y = 0 \tag{1}$$

where the parameter ν (nu) is a given real number which is positive or zero. Bessel's equation often appears if a problem shows cylindrical symmetry, for example, as the membranes in Sec.12.9. The equation satisfies the assumptions of Theorem 1. To see this, divide (1) by x^2 to get the standard form $y'' + y'/x + (1 - \nu^2/x^2)y = 0$. Hence, according to the Frobenius theory, it has a solution of the form

$$y(x) = \sum_{m=0}^{\infty} a_m x^{m+r} \qquad (a_0 \neq 0). \tag{2}$$

Substituting (2) and its first and second derivatives into Bessel's equation, we obtain

$$\sum_{m=0}^{\infty} (m + r)(m + r - 1)a_m x^{m+r} + \sum_{m=0}^{\infty} (m + r)a_m x^{m+r}$$
$$+ \sum_{m=0}^{\infty} a_m x^{m+r+2} - \nu^2 \sum_{m=0}^{\infty} a_m x^{m+r} = 0.$$

We equate the sum of the coefficients of x^{s+r} to zero. Note that this power x^{s+r} corresponds to $m = s$ in the first, second, and fourth series, and to $m = s - 2$ in the third series. Hence for $s = 0$ and $s = 1$, the third series does not contribute since $m \geqq 0$.

[6]FRIEDRICH WILHELM BESSEL (1784–1846), German astronomer and mathematician, studied astronomy on his own in his spare time as an apprentice of a trade company and finally became director of the new Königsberg Observatory.

Formulas on Bessel functions are contained in Ref. [GenRef10] and the standard treatise [A13].

For $s = 2, 3, \cdots$ all four series contribute, so that we get a general formula for all these s. We find

$$\text{(3)}\quad \begin{aligned} &\text{(a)} && r(r-1)a_0 + ra_0 - \nu^2 a_0 = 0 && (s = 0)\\ &\text{(b)} && (r+1)ra_1 + (r+1)a_1 - \nu^2 a_1 = 0 && (s = 1)\\ &\text{(c)} && (s+r)(s+r-1)a_s + (s+r)a_s + a_{s-2} - \nu^2 a_s = 0 && (s = 2, 3, \cdots). \end{aligned}$$

From (3a) we obtain the **indicial equation** by dropping a_0,

$$\textbf{(4)}\qquad (r+\nu)(r-\nu) = 0.$$

The roots are $r_1 = \nu\ (\geqq 0)$ and $r_2 = -\nu$.

Coefficient Recursion for $r = r_1 = \nu$. For $r = \nu$, Eq. (3b) reduces to $(2\nu + 1)a_1 = 0$. Hence $a_1 = 0$ since $\nu \geqq 0$. Substituting $r = \nu$ in (3c) and combining the three terms containing a_s gives simply

$$\text{(5)}\qquad (s + 2\nu)sa_s + a_{s-2} = 0.$$

Since $a_1 = 0$ and $\nu \geqq 0$, it follows from (5) that $a_3 = 0, a_5 = 0, \cdots$. Hence we have to deal only with *even-numbered* coefficients a_s with $s = 2m$. For $s = 2m$, Eq. (5) becomes

$$(2m + 2\nu)2ma_{2m} + a_{2m-2} = 0.$$

Solving for a_{2m} gives the recursion formula

$$\textbf{(6)}\qquad a_{2m} = -\frac{1}{2^2 m(\nu + m)}\, a_{2m-2}, \qquad m = 1, 2, \cdots.$$

From (6) we can now determine $a_2, a_4, \cdots$ successively. This gives

$$a_2 = -\frac{a_0}{2^2(\nu + 1)}$$

$$a_4 = -\frac{a_2}{2^2 2(\nu + 2)} = \frac{a_0}{2^4 2!\,(\nu + 1)(\nu + 2)}$$

and so on, and in general

$$\textbf{(7)}\qquad a_{2m} = \frac{(-1)^m a_0}{2^{2m} m!\,(\nu + 1)(\nu + 2)\cdots(\nu + m)}, \qquad m = 1, 2, \cdots.$$

Bessel Functions $J_n(x)$ for Integer $\nu = n$

Integer values of ν are denoted by n. This is standard. For $\nu = n$ the relation (7) becomes

$$\text{(8)}\qquad a_{2m} = \frac{(-1)^m a_0}{2^{2m} m!\,(n + 1)(n + 2)\cdots(n + m)}, \qquad m = 1, 2, \cdots.$$

a_0 is still arbitrary, so that the series (2) with these coefficients would contain this arbitrary factor a_0. This would be a highly impractical situation for developing formulas or computing values of this new function. Accordingly, we have to make a choice. The choice $a_0 = 1$ would be possible. A simpler series (2) could be obtained if we could absorb the growing product $(n + 1)(n + 2) \cdots (n + m)$ into a factorial function $(n + m)!$ What should be our choice? Our choice should be

$$a_0 = \frac{1}{2^n n!} \tag{9}$$

because then $n!\,(n + 1) \cdots (n + m) = (n + m)!$ in (8), so that (8) simply becomes

$$a_{2m} = \frac{(-1)^m}{2^{2m+n} m!\,(n + m)!}, \qquad m = 1, 2, \cdots. \tag{10}$$

By inserting these coefficients into (2) and remembering that $c_1 = 0, c_3 = 0, \cdots$ we obtain a particular solution of Bessel's equation that is denoted by $J_n(x)$:

$$\boxed{J_n(x) = x^n \sum_{m=0}^{\infty} \frac{(-1)^m x^{2m}}{2^{2m+n} m!\,(n + m)!}} \qquad (n \geqq 0). \tag{11}$$

$J_n(x)$ is called the **Bessel function of the first kind** *of order n.* The series (11) converges for all x, as the ratio test shows. Hence $J_n(x)$ is defined for all x. The series converges very rapidly because of the factorials in the denominator.

EXAMPLE 1 Bessel Functions $J_0(x)$ and $J_1(x)$

For $n = 0$ we obtain from (11) the **Bessel function of order 0**

$$\boxed{J_0(x) = \sum_{m=0}^{\infty} \frac{(-1)^m x^{2m}}{2^{2m} (m!)^2} = 1 - \frac{x^2}{2^2(1!)^2} + \frac{x^4}{2^4(2!)^2} - \frac{x^6}{2^6(3!)^2} + - \cdots} \tag{12}$$

which looks similar to a cosine (Fig. 110). For $n = 1$ we obtain the **Bessel function of order 1**

$$\boxed{J_1(x) = \sum_{m=0}^{\infty} \frac{(-1)^m x^{2m+1}}{2^{2m+1} m!\,(m + 1)!} = \frac{x}{2} - \frac{x^3}{2^3 1!\,2!} + \frac{x^5}{2^5 2!\,3!} - \frac{x^7}{2^7 3!\,4!} + - \cdots,} \tag{13}$$

which looks similar to a sine (Fig. 110). But the zeros of these functions are not completely regularly spaced (see also Table A1 in App. 5) and the height of the "waves" decreases with increasing x. Heuristically, n^2/x^2 in (1) in standard form [(1) divided by x^2] is zero (if $n = 0$) or small in absolute value for large x, and so is y'/x, so that then Bessel's equation comes close to $y'' + y = 0$, the equation of $\cos x$ and $\sin x$; also y'/x acts as a "damping term," in part responsible for the decrease in height. One can show that for large x,

$$\boxed{J_n(x) \sim \sqrt{\frac{2}{\pi x}} \cos\left(x - \frac{n\pi}{2} - \frac{\pi}{4}\right)} \tag{14}$$

where $\sim$ is read "**asymptotically equal**" and means that *for fixed n* the quotient of the two sides approaches 1 as $x \to \infty$.

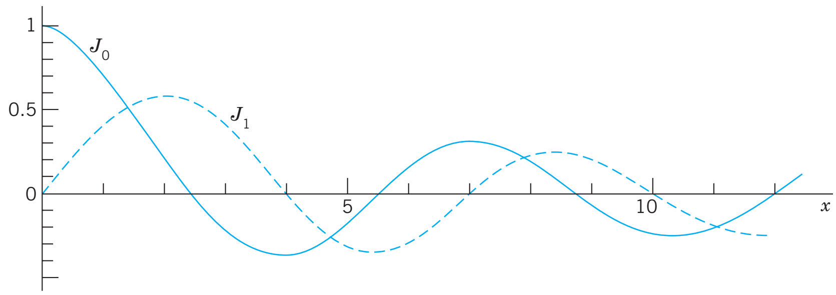


Fig. 110. Bessel functions of the first kind J_0 and J_1

Formula (14) is surprisingly accurate even for smaller $x\,(>0)$. For instance, it will give you good starting values in a computer program for the basic task of computing zeros. For example, for the first three zeros of J_0 you obtain the values 2.356 (2.405 exact to 3 decimals, error 0.049), 5.498 (5.520, error 0.022), 8.639 (8.654, error 0.015), etc. ■

Bessel Functions $J_\nu(x)$ for any $\nu \geqq 0$. Gamma Function

We now proceed from integer $\nu = n$ to any $\nu \geqq 0$. We had $a_0 = 1/(2^n n!)$ in (9). So we have to extend the factorial function $n!$ to any $\nu \geqq 0$. For this we choose

$$a_0 = \frac{1}{2^\nu \Gamma(\nu + 1)} \tag{15}$$

with the **gamma function** $\Gamma(\nu + 1)$ defined by

$$\boxed{\Gamma(\nu + 1) = \int_0^\infty e^{-t} t^\nu \, dt} \qquad (\nu > -1). \tag{16}$$

(**CAUTION!** Note the convention $\nu + 1$ on the left but ν in the integral.) Integration by parts gives

$$\Gamma(\nu + 1) = -e^{-t} t^\nu \Big|_0^\infty + \nu \int_0^\infty e^{-t} t^{\nu-1} \, dt = 0 + \nu\Gamma(\nu).$$

This is the basic functional relation of the gamma function

$$\boxed{\Gamma(\nu + 1) = \nu\Gamma(\nu).} \tag{17}$$

Now from (16) with $\nu = 0$ and then by (17) we obtain

$$\Gamma(1) = \int_0^\infty e^{-t} \, dt = -e^{-t} \Big|_0^\infty = 0 - (-1) = 1$$

and then $\Gamma(2) = 1 \cdot \Gamma(1) = 1!$, $\Gamma(3) = 2\Gamma(1) = 2!$ and in general

$$\boxed{\Gamma(n + 1) = n!} \qquad (n = 0, 1, \cdots). \tag{18}$$

Hence ***the gamma function generalizes the factorial function*** **to arbitrary positive** ν. Thus (15) with $\nu = n$ agrees with (9).

Furthermore, from (7) with a_0 given by (15) we first have

$$a_{2m} = \frac{(-1)^m}{2^{2m} m!\,(\nu+1)(\nu+2)\cdots(\nu+m)2^\nu\Gamma(\nu+1)}.$$

Now (17) gives $(\nu+1)\Gamma(\nu+1) = \Gamma(\nu+2)$, $(\nu+2)\Gamma(\nu+2) = \Gamma(\nu+3)$ and so on, so that

$$(\nu+1)(\nu+2)\cdots(\nu+m)\Gamma(\nu+1) = \Gamma(\nu+m+1).$$

Hence because of our (standard!) choice (15) of a_0 the coefficients (7) are simply

$$a_{2m} = \frac{(-1)^m}{2^{2m+\nu} m!\,\Gamma(\nu+m+1)}. \tag{19}$$

With these coefficients and $r = r_1 = \nu$ we get from (2) a particular solution of (1), denoted by $J_\nu(x)$ and given by

$$J_\nu(x) = x^\nu \sum_{m=0}^{\infty} \frac{(-1)^m x^{2m}}{2^{2m+\nu} m!\,\Gamma(\nu+m+1)}. \tag{20}$$

$J_\nu(x)$ is called the **Bessel function of the first kind of order ν**. The series (20) converges for all x, as one can verify by the ratio test.

Discovery of Properties from Series

Bessel functions are a model case for showing how to discover properties and relations of functions from series by which they are *defined*. Bessel functions satisfy an incredibly large number of relationships—look at Ref. [A13] in App. 1; also, find out what your CAS knows. In Theorem 3 we shall discuss four formulas that are backbones in applications and theory.

THEOREM 1

Derivatives, Recursions

The derivative of $J_\nu(x)$ with respect to x can be expressed by $J_{\nu-1}(x)$ or $J_{\nu+1}(x)$ by the formulas

$$\text{(21)}\qquad \begin{aligned} &\textbf{(a)}\quad [x^\nu J_\nu(x)]' = x^\nu J_{\nu-1}(x)\\ &\textbf{(b)}\quad [x^{-\nu} J_\nu(x)]' = -x^{-\nu} J_{\nu+1}(x). \end{aligned}$$

Furthermore, $J_\nu(x)$ and its derivative satisfy the recurrence relations

$$\text{(21)}\qquad \begin{aligned} &\textbf{(c)}\quad J_{\nu-1}(x) + J_{\nu+1}(x) = \frac{2\nu}{x} J_\nu(x)\\ &\textbf{(d)}\quad J_{\nu-1}(x) - J_{\nu+1}(x) = 2J_\nu'(x). \end{aligned}$$

PROOF **(a)** We multiply (20) by x^{ν} and take $x^{2\nu}$ under the summation sign. Then we have

$$x^{\nu}J_{\nu}(x) = \sum_{m=0}^{\infty} \frac{(-1)^m x^{2m+2\nu}}{2^{2m+\nu} m!\, \Gamma(\nu + m + 1)}.$$

We now differentiate this, cancel a factor 2, pull $x^{2\nu-1}$ out, and use the functional relationship $\Gamma(\nu + m + 1) = (\nu + m)\Gamma(\nu + m)$ [see (17)]. Then (20) with $\nu - 1$ instead of ν shows that we obtain the right side of (21a). Indeed,

$$(x^{\nu}J_{\nu})' = \sum_{m=0}^{\infty} \frac{(-1)^m 2(m + \nu)x^{2m+2\nu-1}}{2^{2m+\nu} m!\, \Gamma(\nu + m + 1)} = x^{\nu}x^{\nu-1} \sum_{m=0}^{\infty} \frac{(-1)^m x^{2m}}{2^{2m+\nu-1} m!\, \Gamma(\nu + m)}.$$

(b) Similarly, we multiply (20) by $x^{-\nu}$, so that x^{ν} in (20) cancels. Then we differentiate, cancel $2m$, and use $m! = m(m - 1)!$. This gives, with $m = s + 1$,

$$(x^{-\nu}J_{\nu})' = \sum_{m=1}^{\infty} \frac{(-1)^m x^{2m-1}}{2^{2m+\nu-1}(m - 1)!\, \Gamma(\nu + m + 1)} = \sum_{s=0}^{\infty} \frac{(-1)^{s+1} x^{2s+1}}{2^{2s+\nu+1} s!\, \Gamma(\nu + s + 2)}.$$

Equation (20) with $\nu + 1$ instead of ν and s instead of m shows that the expression on the right is $-x^{-\nu}J_{\nu+1}(x)$. This proves (21b).

(c), **(d)** We perform the differentiation in (21a). Then we do the same in (21b) and multiply the result on both sides by $x^{2\nu}$. This gives

$$\begin{aligned} &\text{(a*)} \qquad \nu x^{\nu-1}J_{\nu} + x^{\nu}J_{\nu}' = x^{\nu}J_{\nu-1} \\ &\text{(b*)} \qquad -\nu x^{\nu-1}J_{\nu} + x^{\nu}J_{\nu}' = -x^{\nu}J_{\nu+1}. \end{aligned}$$

Substracting (b*) from (a*) and dividing the result by x^{ν} gives (21c). Adding (a*) and (b*) and dividing the result by x^{ν} gives (21d). ■

EXAMPLE 2 **Application of Theorem 1 in Evaluation and Integration**

Formula (21c) can be used recursively in the form

$$J_{\nu+1}(x) = \frac{2\nu}{x} J_{\nu}(x) - J_{\nu-1}(x)$$

for calculating Bessel functions of higher order from those of lower order. For instance, $J_2(x) = 2J_1(x)/x - J_0(x)$, so that J_2 can be obtained from tables of J_0 and J_1 (in App. 5 or, more accurately, in Ref. [GenRef1] in App. 1).

To illustrate how Theorem 1 helps in integration, we use (21b) with $\nu = 3$ integrated on both sides. This evaluates, for instance, the integral

$$I = \int_1^2 x^{-3}J_4(x)\, dx = -x^{-3}J_3(x)\Big|_1^2 = -\frac{1}{8} J_3(2) + J_3(1).$$

A table of J_3 (on p. 398 of Ref. [GenRef1]) or your CAS will give you

$$-\tfrac{1}{8} \cdot 0.128943 + 0.019563 = 0.003445.$$

Your CAS (or a human computer in precomputer times) obtains J_3 from (21), first using (21c) with $\nu = 2$, that is, $J_3 = 4x^{-1}J_2 - J_1$, then (21c) with $\nu = 1$, that is, $J_2 = 2x^{-1}J_1 - J_0$. Together,

$$I = x^{-3}(4x^{-1}(2x^{-1}J_1 - J_0) - J_1)\Big|_1^2$$
$$= -\tfrac{1}{8}[2J_1(2) - 2J_0(2) - J_1(2)] + [8J_1(1) - 4J_0(1) - J_1(1)]$$
$$= -\tfrac{1}{8}J_1(2) + \tfrac{1}{4}J_0(2) + 7J_1(1) - 4J_0(1).$$

This is what you get, for instance, with Maple if you type int$(\cdots)$. And if you type evalf(int$(\cdots)$), you obtain 0.003445448, in agreement with the result near the beginning of the example. ■

Bessel Functions J_ν with Half-Integer ν Are Elementary

We discover this remarkable fact as another property obtained from the series (20) and confirm it in the problem set by using Bessel's ODE.

EXAMPLE 3 **Elementary Bessel Functions J_ν with $\nu = \pm\frac{1}{2}, \pm\frac{3}{2}, \pm\frac{5}{2}, \cdots$. The Value $\Gamma(\frac{1}{2})$**

We first prove (Fig. 111)

$$\textbf{(22)} \qquad \text{(a)} \quad J_{1/2}(x) = \sqrt{\frac{2}{\pi x}}\sin x, \qquad \text{(b)} \quad J_{-1/2}(x) = \sqrt{\frac{2}{\pi x}}\cos x.$$

The series (20) with $\nu = \frac{1}{2}$ is

$$J_{1/2}(x) = \sqrt{x}\sum_{m=0}^{\infty}\frac{(-1)^m x^{2m}}{2^{2m+1/2}m!\,\Gamma(m+\frac{3}{2})} = \sqrt{\frac{2}{x}}\sum_{m=0}^{\infty}\frac{(-1)^m x^{2m+1}}{2^{2m+1}m!\,\Gamma(m+\frac{3}{2})}.$$

The denominator can be written as a product AB, where (use (16) in B)

$$A = 2^m m! = 2m(2m-2)(2m-4)\cdots 4\cdot 2,$$
$$B = 2^{m+1}\Gamma(m+\tfrac{3}{2}) = 2^{m+1}(m+\tfrac{1}{2})(m-\tfrac{1}{2})\cdots\tfrac{3}{2}\cdot\tfrac{1}{2}\Gamma(\tfrac{1}{2})$$
$$= (2m+1)(2m-1)\cdots 3\cdot 1\cdot\sqrt{\pi};$$

here we used (proof below)

$$\textbf{(23)} \qquad \Gamma(\tfrac{1}{2}) = \sqrt{\pi}.$$

The product of the right sides of A and B can be written

$$AB = (2m+1)2m(2m-1)\cdots 3\cdot 2\cdot 1\sqrt{\pi} = (2m+1)!\sqrt{\pi}.$$

Hence

$$J_{1/2}(x) - \sqrt{\frac{2}{\pi x}}\sum_{m=0}^{\infty}\frac{(-1)^m x^{2m+1}}{(2m+1)!} = \sqrt{\frac{2}{\pi x}}\sin x.$$

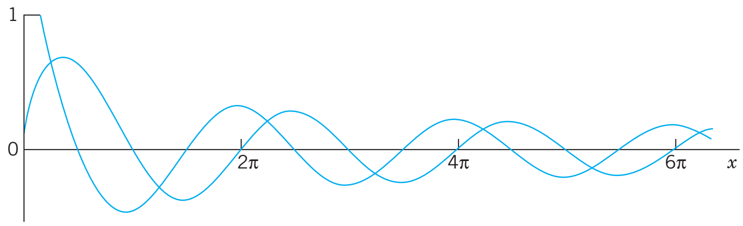


Fig. 111. Bessel functions $J_{1/2}$ and $J_{-1/2}$

This proves (22a). Differentiation and the use of (21a) with $\nu = \frac{1}{2}$ now gives

$$[\sqrt{x} J_{1/2}(x)]' = \sqrt{\frac{2}{\pi}} \cos x = x^{1/2} J_{-1/2}(x).$$

This proves (22b). From (22) follow further formulas successively by (21c), used as in Example 2.

We finally prove $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ by a standard trick worth remembering. In (15) we set $t = u^2$. Then $dt = 2u\, du$ and

$$\Gamma\left(\frac{1}{2}\right) = \int_0^\infty e^{-t} t^{-1/2}\, dt = 2\int_0^\infty e^{-u^2}\, du.$$

We square on both sides, write v instead of u in the second integral, and then write the product of the integrals as a double integral:

$$\Gamma\left(\frac{1}{2}\right)^2 = 4\int_0^\infty e^{-u^2}\, du \int_0^\infty e^{-v^2}\, dv = 4\int_0^\infty \int_0^\infty e^{-(u^2+v^2)}\, du\, dv.$$

We now use polar coordinates r, θ by setting $u = r\cos\theta$, $v = r\sin\theta$. Then the element of area is $du\, dv = r\, dr\, d\theta$ and we have to integrate over r from 0 to ∞ and over θ from 0 to $\pi/2$ (that is, over the first quadrant of the uv-plane):

$$\Gamma\left(\frac{1}{2}\right)^2 = 4\int_0^{\pi/2} \int_0^\infty e^{-r^2} r\, dr\, d\theta = 4 \cdot \frac{\pi}{2} \int_0^\infty e^{-r^2} r\, dr = 2\left(-\frac{1}{2}\right) e^{-r^2} \Big|_0^\infty = \pi.$$

By taking the square root on both sides we obtain (23). ■

General Solution. Linear Dependence

For a general solution of Bessel's equation (1) in addition to J_ν we need a second linearly independent solution. For ν not an integer this is easy. Replacing ν by $-\nu$ in (20), we have

$$J_{-\nu}(x) = x^{-\nu} \sum_{m=0}^{\infty} \frac{(-1)^m x^{2m}}{2^{2m-\nu} m!\, \Gamma(m - \nu + 1)}. \tag{24}$$

Since Bessel's equation involves ν^2, the functions J_ν and $J_{-\nu}$ are solutions of the equation for the same ν. If ν is not an integer, they are linearly independent, because the first terms in (20) and in (24) are finite nonzero multiples of x^ν and $x^{-\nu}$. Thus, if ν is not an integer, a general solution of Bessel's equation for all $x \neq 0$ is

$$y(x) = c_1 J_\nu(x) + c_2 J_{-\nu}(x)$$

This cannot be the general solution for an integer $\nu = n$ because, in that case, we have linear dependence. It can be seen that the first terms in (20) and (24) are finite nonzero multiples of x^ν and $x^{-\nu}$, respectively. This means that, for any integer $\nu = n$, we have linear dependence because

$$\boxed{J_{-n}(x) = (-1)^n J_n(x)} \qquad (n = 1, 2, \cdots). \tag{25}$$

PROOF To prove (25), we use (24) and let ν approach a positive integer n. Then the gamma function in the coefficients of the first n terms becomes infinite (see Fig. 553 in App. A3.1), the coefficients become zero, and the summation starts with $m = n$. Since in this case $\Gamma(m - n + 1) = (m - n)!$ by (18), we obtain

$$J_{-n}(x) = \sum_{m=n}^{\infty} \frac{(-1)^m x^{2m-n}}{2^{2m-n} m!\,(m-n)!} = \sum_{s=0}^{\infty} \frac{(-1)^{n+s} x^{2s+n}}{2^{2s+n}(n+s)!\,s!} \qquad (m = n + s). \tag{26}$$

The last series represents $(-1)^n J_n(x)$, as you can see from (11) with m replaced by s. This completes the proof. ∎

The difficulty caused by (25) will be overcome in the next section by introducing further Bessel functions, called *of the second kind* and denoted by Y_ν.

PROBLEM SET 5.4

1. Convergence. Show that the series (11) converges for all x. Why is the convergence very rapid?

2–10 ODEs REDUCIBLE TO BESSEL'S ODE

This is just a sample of such ODEs; some more follow in the next problem set. Find a general solution in terms of J_ν and $J_{-\nu}$ or indicate when this is not possible. Use the indicated substitutions. Show the details of your work.

2. $x^2y'' + xy' + (x^2 - \frac{4}{49})y = 0$

3. $xy'' + y' + \frac{1}{4}y = 0 \quad (\sqrt{x} = z)$

4. $y'' + (e^{-2x} - \frac{1}{9})y = 0 \quad (e^{-x} = z)$

5. Two-parameter ODE
$x^2y'' + xy' + (\lambda^2x^2 - \nu^2)y = 0 \quad (\lambda x = z)$

6. $x^2y'' + \frac{1}{4}(x + \frac{3}{4})\,y = 0 \quad (y = u\sqrt{x},\ \sqrt{x} = z)$

7. $x^2y'' + xy' + \frac{1}{4}(x^2 - 1)y = 0 \quad (x = 2z)$

8. $(2x + 1)^2y'' + 2(2x + 1)y' + 16x(x + 1)y = 0$
$(2x + 1 = z)$

9. $xy'' + (2\nu + 1)y' + xy = 0 \quad (y = x^{-\nu}u)$

10. $x^2y'' + (1 - 2\nu)xy' + \nu^2(x^{2\nu} + 1 - \nu^2)y = 0$
$(y = x^\nu u,\ x^\nu = z)$

11. CAS EXPERIMENT. Change of Coefficient. Find and graph (on common axes) the solutions of

$$y'' + kx^{-1}y' + y = 0,\ y(0) = 1,\ y'(0) = 0,$$

for $k = 0, 1, 2, \cdots, 10$ (or as far as you get useful graphs). For what k do you get elementary functions? Why? Try for noninteger k, particularly between 0 and 2, to see the continuous change of the curve. Describe the change of the location of the zeros and of the extrema as k increases from 0. Can you interpret the ODE as a model in mechanics, thereby explaining your observations?

12. CAS EXPERIMENT. Bessel Functions for Large x.

(a) Graph $J_n(x)$ for $n = 0, \cdots, 5$ on common axes.

(b) Experiment with (14) for integer n. Using graphs, find out from which $x = x_n$ on the curves of (11) and (14) practically coincide. How does x_n change with n?

(c) What happens in (b) if $n = \pm\frac{1}{2}$? (Our usual notation in this case would be ν.)

(d) How does the error of (14) behave as a function of x for fixed n? [Error = exact value minus approximation (14).]

(e) Show from the graphs that $J_0(x)$ has extrema where $J_1(x) = 0$. Which formula proves this? Find further relations between zeros and extrema.

13–15 ZEROS of Bessel functions play a key role in modeling (e.g. of vibrations; see Sec. 12.9).

13. Interlacing of zeros. Using (21) and Rolle's theorem, show that between any two consecutive positive zeros of $J_n(x)$ there is precisely one zero of $J_{n+1}(x)$.

14. Zeros. Compute the first four positive zeros of $J_0(x)$ and $J_1(x)$ from (14). Determine the error and comment.

15. Interlacing of zeros. Using (21) and Rolle's theorem, show that between any two consecutive zeros of $J_0(x)$ there is precisely one zero of $J_1(x)$.

16–18 HALF-INTEGER PARAMETER: APPROACH BY THE ODE

16. Elimination of first derivative. Show that $y = uv$ with $v(x) = \exp\left(-\frac{1}{2}\int p(x)\,dx\right)$ gives from the ODE $y'' + p(x)y' + q(x)y = 0$ the ODE

$$u'' + \left[q(x) - \tfrac{1}{4}p(x)^2 - \tfrac{1}{2}p'(x)\right]u = 0,$$

not containing the first derivative of u.

17. Bessel's equation. Show that for (1) the substitution in Prob. 16 is $y = ux^{-1/2}$ and gives

$$x^2u'' + (x^2 + \tfrac{1}{4} - \nu^2)u = 0. \tag{27}$$

18. Elementary Bessel functions. Derive (22) in Example 3 from (27).

19–25 APPLICATION OF (21): DERIVATIVES, INTEGRALS

Use the powerful formulas (21) to do Probs. 19–25. Show the details of your work.

19. Derivatives. Show that $J_0'(x) = -J_1(x)$, $J_1'(x) = J_0(x) - J_1(x)/x$, $J_2'(x) = \frac{1}{2}[J_1(x) - J_3(x)]$.

20. Bessel's equation. Derive (1) from (21).

21. Basic integral formula. Show that

$$\int x^{\nu}J_{\nu-1}(x)\,dx = x^{\nu}J_{\nu}(x) + c.$$

22. Basic integral formulas. Show that

$$\int x^{-\nu}J_{\nu+1}(x)\,dx = -x^{-\nu}J_{\nu}(x) + c,$$

$$\int J_{\nu+1}(x)\,dx = \int J_{\nu-1}(x)\,dx - 2J_{\nu}(x).$$

23. Integration. Show that $\int x^2J_0(x)\,dx = x^2J_1(x) + xJ_0(x) - \int J_0(x)\,dx$. (The last integral is nonelementary; tables exist, e.g., in Ref. [A13] in App. 1.)

24. Integration. Evaluate $\int x^{-1}J_4(x)\,dx$.

25. Integration. Evaluate $\int J_5(x)\,dx$.

5.5 Bessel Functions $Y_\nu(x)$. General Solution

To obtain a general solution of Bessel's equation (1), Sec. 5.4, for any ν, we now introduce **Bessel functions of the second kind** $Y_\nu(x)$, beginning with the case $\nu = n = 0$.

When $n = 0$, Bessel's equation can be written (divide by x)

$$xy'' + y' + xy = 0. \tag{1}$$

Then the indicial equation (4) in Sec. 5.4 has a double root $r = 0$. This is Case 2 in Sec. 5.3. In this case we first have only one solution, $J_0(x)$. From (8) in Sec. 5.3 we see that the desired second solution must be of the form

$$y_2(x) = J_0(x)\ln x + \sum_{m=1}^{\infty} A_m x^m. \tag{2}$$

We substitute y_2 and its derivatives

$$y_2' = J_0'\ln x + \frac{J_0}{x} + \sum_{m=1}^{\infty} mA_m x^{m-1}$$

$$y_2'' = J_0''\ln x + \frac{2J_0'}{x} - \frac{J_0}{x^2} + \sum_{m=1}^{\infty} m(m-1)A_m x^{m-2}$$

into (1). Then the sum of the three logarithmic terms $xJ_0''\ln x$, $J_0'\ln x$, and $xJ_0\ln x$ is zero because J_0 is a solution of (1). The terms $-J_0/x$ and J_0/x (from xy'' and y') cancel. Hence we are left with

$$2J_0' + \sum_{m=1}^{\infty} m(m-1)A_m x^{m-1} + \sum_{m=1}^{\infty} mA_m x^{m-1} + \sum_{m=1}^{\infty} A_m x^{m+1} = 0.$$

Addition of the first and second series gives $\Sigma m^2 A_m x^{m-1}$. The power series of $J_0'(x)$ is obtained from (12) in Sec. 5.4 and the use of $m!/m = (m-1)!$ in the form

$$J_0'(x) = \sum_{m=1}^{\infty} \frac{(-1)^m 2m x^{2m-1}}{2^{2m}(m!)^2} = \sum_{m=1}^{\infty} \frac{(-1)^m x^{2m-1}}{2^{2m-1} m!\,(m-1)!}.$$

Together with $\Sigma m^2 A_m x^{m-1}$ and $\Sigma A_m x^{m+1}$ this gives

$$\sum_{m=1}^{\infty} \frac{(-1)^m x^{2m-1}}{2^{2m-2} m!\,(m-1)!} + \sum_{m=1}^{\infty} m^2 A_m x^{m-1} + \sum_{m=1}^{\infty} A_m x^{m+1} = 0. \tag{3*}$$

First, we show that the A_m with odd subscripts are all zero. The power x^0 occurs only in the second series, with coefficient A_1. Hence $A_1 = 0$. Next, we consider the even powers x^{2s}. The first series contains none. In the second series, $m - 1 = 2s$ gives the term $(2s+1)^2 A_{2s+1} x^{2s}$. In the third series, $m + 1 = 2s$. Hence by equating the sum of the coefficients of x^{2s} to zero we have

$$(2s+1)^2 A_{2s+1} + A_{2s-1} = 0, \qquad s = 1, 2, \cdots.$$

Since $A_1 = 0$, we thus obtain $A_3 = 0, A_5 = 0, \cdots$, successively.

We now equate the sum of the coefficients of x^{2s+1} to zero. For $s = 0$ this gives

$$-1 + 4A_2 = 0, \qquad \text{thus} \qquad A_2 = \tfrac{1}{4}.$$

For the other values of s we have in the first series in (3*) $2m - 1 = 2s + 1$, hence $m = s + 1$, in the second $m - 1 = 2s + 1$, and in the third $m + 1 = 2s + 1$. We thus obtain

$$\frac{(-1)^{s+1}}{2^{2s}(s+1)!\,s!} + (2s+2)^2 A_{2s+2} + A_{2s} = 0.$$

For $s = 1$ this yields

$$\tfrac{1}{8} + 16A_4 + A_2 = 0, \qquad \text{thus} \qquad A_4 = -\tfrac{3}{128}$$

and in general

$$A_{2m} = \frac{(-1)^{m-1}}{2^{2m}(m!)^2}\left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{m}\right), \qquad m = 1, 2, \cdots. \tag{3}$$

Using the short notations

$$h_1 = 1 \qquad h_m = 1 + \frac{1}{2} + \cdots + \frac{1}{m} \qquad m = 2, 3, \cdots \tag{4}$$

and inserting (4) and $A_1 = A_3 = \cdots = 0$ into (2), we obtain the result

$$y_2(x) = J_0(x) \ln x + \sum_{m=1}^{\infty} \frac{(-1)^{m-1} h_m}{2^{2m}(m!)^2} x^{2m}$$

$$= J_0(x) \ln x + \frac{1}{4}x^2 - \frac{3}{128}x^4 + \frac{11}{13{,}824}x^6 - + \cdots. \tag{5}$$

Since J_0 and y_2 are linearly independent functions, they form a basis of (1) for $x > 0$. Of course, another basis is obtained if we replace y_2 by an independent particular solution of the form $a(y_2 + bJ_0)$, where $a\ (\neq 0)$ and b are constants. It is customary to choose $a = 2/\pi$ and $b = \gamma - \ln 2$, where the number $\gamma = 0.57721566490 \cdots$ is the so-called **Euler constant**, which is defined as the limit of

$$1 + \frac{1}{2} + \cdots + \frac{1}{s} - \ln s$$

as s approaches infinity. The standard particular solution thus obtained is called the **Bessel function of the second kind** *of order zero* (Fig. 112) or **Neumann's function** *of order zero* and is denoted by $Y_0(x)$. Thus [see (4)]

$$\textbf{(6)} \qquad Y_0(x) = \frac{2}{\pi}\left[J_0(x)\left(\ln\frac{x}{2} + \gamma\right) + \sum_{m=1}^{\infty} \frac{(-1)^{m-1}h_m}{2^{2m}(m!)^2}\, x^{2m}\right].$$

For small $x > 0$ the function $Y_0(x)$ behaves about like $\ln x$ (see Fig. 112, why?), and $Y_0(x) \to -\infty$ as $x \to 0$.

Bessel Functions of the Second Kind $Y_n(x)$

For $\nu = n = 1, 2, \cdots$ a second solution can be obtained by manipulations similar to those for $n = 0$, starting from (10), Sec. 5.4. It turns out that in these cases the solution also contains a logarithmic term.

The situation is not yet completely satisfactory, because the second solution is defined differently, depending on whether the order ν is an integer or not. To provide uniformity of formalism, it is desirable to adopt a form of the second solution that is valid for all values of the order. For this reason we introduce a standard second solution $Y_\nu(x)$ defined for all ν by the formula

$$\textbf{(7)} \qquad \begin{aligned} &\textbf{(a)} \qquad Y_\nu(x) = \frac{1}{\sin \nu\pi}\,[J_\nu(x)\cos \nu\pi - J_{-\nu}(x)] \\ &\textbf{(b)} \qquad Y_n(x) = \lim_{\nu \to n} Y_\nu(x). \end{aligned}$$

This function is called the **Bessel function of the second kind** *of order* ν or **Neumann's function**[7] *of order* ν. Figure 112 shows $Y_0(x)$ and $Y_1(x)$.

Let us show that J_ν and Y_ν are indeed linearly independent for all ν (and $x > 0$).

For noninteger order ν, the function $Y_\nu(x)$ is evidently a solution of Bessel's equation because $J_\nu(x)$ and $J_{-\nu}(x)$ are solutions of that equation. Since for those ν the solutions J_ν and $J_{-\nu}$ are linearly independent and Y_ν involves $J_{-\nu}$, the functions J_ν and Y_ν are

---

[7] CARL NEUMANN (1832–1925), German mathematician and physicist. His work on potential theory using integer equation methods inspired VITO VOLTERRA (1800–1940) of Rome, ERIK IVAR FREDHOLM (1866–1927) of Stockholm, and DAVID HILBERT (1962–1943) of Göttingen (see the footnote in Sec. 7.9) to develop the field of integral equations. For details see Birkhoff, G. and E. Kreyszig, The Establishment of Functional Analysis, *Historia Mathematica* 11 (1984), pp. 258–321.

The solutions $Y_\nu(x)$ are sometimes denoted by $N_\nu(x)$; in Ref. [A13] they are called **Weber's functions**; Euler's constant in (6) is often denoted by C or $\ln \gamma$.

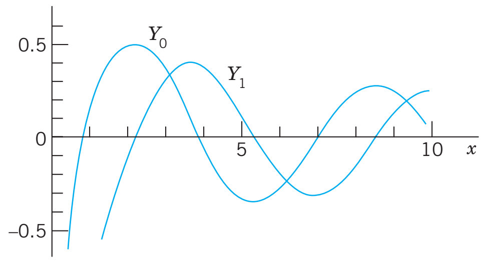


Fig. 112. Bessel functions of the second kind Y_0 and Y_1.
(For a small table, see App. 5.)

linearly independent. Furthermore, it can be shown that the limit in (7b) exists and Y_n is a solution of Bessel's equation for integer order; see Ref. [A13] in App. 1. We shall see that the series development of $Y_n(x)$ contains a logarithmic term. Hence $J_n(x)$ and $Y_n(x)$ are linearly independent solutions of Bessel's equation. The series development of $Y_n(x)$ can be obtained if we insert the series (20) in Sec. 5.4 and (2) in this section for $J_\nu(x)$ and $J_{-\nu}(x)$ into (7a) and then let ν approach n; for details see Ref. [A13]. The result is

$$
\textbf{(8)} \quad \begin{aligned} Y_n(x) = \frac{2}{\pi} J_n(x)\left(\ln\frac{x}{2} + \gamma\right) + \frac{x^n}{\pi}\sum_{m=0}^{\infty} \frac{(-1)^{m-1}(h_m + h_{m+n})}{2^{2m+n}m!\,(m+n)!} x^{2m} \\ - \frac{x^{-n}}{\pi}\sum_{m=0}^{n-1} \frac{(n-m-1)!}{2^{2m-n}m!} x^{2m} \end{aligned}
$$

where $x > 0$, $n = 0, 1, \cdots$, and [as in (4)] $h_0 = 0$, $h_1 = 1$,

$$
h_m = 1 + \frac{1}{2} + \cdots + \frac{1}{m}, \qquad h_{m+n} = 1 + \frac{1}{2} + \cdots + \frac{1}{m+n}.
$$

For $n = 0$ the last sum in (8) is to be replaced by 0 [giving agreement with (6)].

Furthermore, it can be shown that

$$
Y_{-n}(x) = (-1)^n Y_n(x).
$$

Our main result may now be formulated as follows.

THEOREM 1

General Solution of Bessel's Equation

A general solution of Bessel's equation for all values of ν *(and* $x > 0$*) is*

$$
\textbf{(9)} \qquad y(x) = C_1 J_\nu(x) + C_2 Y_\nu(x).
$$

We finally mention that there is a practical need for solutions of Bessel's equation that are complex for real values of x. For this purpose the solutions

$$
(10) \qquad \begin{aligned} H_\nu^{(1)}(x) &= J_\nu(x) + iY_\nu(x) \\ H_\nu^{(2)}(x) &= J_\nu(x) - iY_\nu(x) \end{aligned}
$$

are frequently used. These linearly independent functions are called **Bessel functions of the third kind** *of order* ν or *first and second* **Hankel functions**[8] *of order* ν.

This finishes our discussion on Bessel functions, except for their "orthogonality," which we explain in Sec. 11.6. Applications to vibrations follow in Sec. 12.10.

PROBLEM SET 5.5

1–9 FURTHER ODE's REDUCIBLE TO BESSEL'S ODE

Find a general solution in terms of J_ν and Y_ν. Indicate whether you could also use $J_{-\nu}$ instead of Y_ν. Use the indicated substitution. Show the details of your work.

1. $x^2y'' + xy' + (x^2 - 16)\,y = 0$

2. $xy'' + 5y' + xy = 0 \quad (y = u/x^2)$

3. $9x^2y'' + 9xy' + (36x^4 - 16)y = 0 \quad (x^2 = z)$

4. $y'' + xy = 0 \quad (y = u\sqrt{x}, \frac{2}{3}x^{3/2} = z)$

5. $4xy'' + 4y' + y = 0 \quad (\sqrt{x} = z)$

6. $xy'' + y' + 36y = 0 \quad (12\sqrt{x} = z)$

7. $y'' + k^2x^2y = 0 \quad (y = u\sqrt{x}, \frac{1}{2}kx^2 = z)$

8. $y'' + k^2x^4y = 0 \quad (y = u\sqrt{x}, \frac{1}{3}kx^3 = z)$

9. $xy'' - 5y' + xy = 0 \quad (y = x^3u)$

10. CAS EXPERIMENT. Bessel Functions for Large x. It can be shown that for large x,

$$Y_n(x) \sim \sqrt{2/(\pi x)}\,\sin\left(x - \tfrac{1}{2}n\pi - \tfrac{1}{4}\pi\right) \tag{11}$$

with $\sim$ defined as in (14) of Sec. 5.4.

(a) Graph $Y_n(x)$ for $n = 0, \cdots, 5$ on common axes. Are there relations between zeros of one function and extrema of another? For what functions?

(b) Find out from graphs from which $x = x_n$ on the curves of (8) and (11) (both obtained from your CAS) practically coincide. How does x_n change with n?

(c) Calculate the first ten zeros x_m, $m = 1, \cdots, 10$, of $Y_0(x)$ from your CAS and from (11). How does the error behave as m increases?

(d) Do (c) for $Y_1(x)$ and $Y_2(x)$. How do the errors compare to those in (c)?

11–15 HANKEL AND MODIFIED BESSEL FUNCTIONS

11. Hankel functions. Show that the Hankel functions (10) form a basis of solutions of Bessel's equation for any ν.

12. Modified Bessel functions of the first kind of order ν are defined by $I_\nu(x) = i^{-\nu}J_\nu(ix)$, $i = \sqrt{-1}$. Show that I_ν satisfies the ODE

$$x^2y'' + xy' - (x^2 + \nu^2)\,y = 0. \tag{12}$$

13. Modified Bessel functions. Show that $I_\nu(x)$ has the representation

$$I_\nu(x) = \sum_{m=0}^{\infty} \frac{x^{2m+\nu}}{2^{2m+\nu}m!\,\Gamma(m + \nu + 1)}. \tag{13}$$

14. Reality of I_ν. Show that $I_\nu(x)$ is real for all real x (and real ν), $I_\nu(x) \neq 0$ for all real $x \neq 0$, and $I_{-n}(x) = I_n(x)$, where n is any integer.

15. Modified Bessel functions of the third kind (sometimes called *of the second kind*) are defined by the formula (14) below. Show that they satisfy the ODE (12).

$$K_\nu(x) = \frac{\pi}{2\sin\nu\pi}\left[I_{-\nu}(x) - I_\nu(x)\right]. \tag{14}$$

CHAPTER 5 REVIEW QUESTIONS AND PROBLEMS

1. Why are we looking for power series solutions of ODEs?

2. What is the difference between the two methods in this chapter? Why do we need two methods?

3. What is the indicial equation? Why is it needed?

4. List the three cases of the Frobenius method, and give examples of your own.

5. Write down the most important ODEs in this chapter from memory.

6. Can a power series solution reduce to a polynomial? When? Why is this important?

7. What is the hypergeometric equation? Where does the name come from?

8. List some properties of the Legendre polynomials.

9. Why did we introduce two kinds of Bessel functions?

10. Can a Bessel function reduce to an elementary function? When?

[8]HERMANN HANKEL (1839–1873), German mathematician.

11–20 **POWER SERIES METHOD OR FROBENIUS METHOD**

Find a basis of solutions. Try to identify the series as expansions of known functions. Show the details of your work.

11. $y'' + 4y = 0$

12. $xy'' + (1 - 2x)y' + (x - 1)y = 0$

13. $(x - 1)^2 y'' - (x - 1)y' - 35y = 0$

14. $16(x + 1)^2 y'' + 3y = 0$

15. $x^2 y'' + xy' + (x^2 - 5)y = 0$

16. $x^2 y'' + 2x^3 y' + (x^2 - 2)y = 0$

17. $xy'' - (x + 1)y' + y = 0$

18. $xy'' + 3y' + 4x^3 y = 0$

19. $y'' + \dfrac{1}{4x} y = 0$

20. $xy'' + y' - xy = 0$

SUMMARY OF CHAPTER 5

Series Solution of ODEs. Special Functions

The **power series method** gives solutions of linear ODEs

$$y'' + p(x)y' + q(x)y = 0 \tag{1}$$

with ***variable coefficients*** p and q in the form of a power series (with any center x_0, e.g., $x_0 = 0$)

$$y(x) = \sum_{m=0}^{\infty} a_m (x - x_0)^m = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \cdots. \tag{2}$$

Such a solution is obtained by substituting (2) and its derivatives into (1). This gives a ***recurrence formula*** for the coefficients. You may program this formula (or even obtain and graph the whole solution) on your CAS.

If p and q are **analytic** at x_0 (that is, representable by a power series in powers of $x - x_0$ with positive radius of convergence; Sec. 5.1), then (1) has solutions of this form (2). The same holds if $\tilde{h}$, $\tilde{p}$, $\tilde{q}$ in

$$\tilde{h}(x)y'' + \tilde{p}(x)y' + \tilde{q}(x)y = 0$$

are analytic at x_0 and $\tilde{h}(x_0) \neq 0$, so that we can divide by $\tilde{h}$ and obtain the standard form (1). ***Legendre's equation*** is solved by the power series method in Sec. 5.2.

The ***Frobenius method*** (Sec. 5.3) extends the power series method to ODEs

$$y'' + \frac{a(x)}{x - x_0} y' + \frac{b(x)}{(x - x_0)^2} y = 0 \tag{3}$$

whose coefficients are **singular** (i.e., not analytic) at x_0, but are "not too bad," namely, such that a and b are analytic at x_0. Then (3) has at least one solution of the form

$$y(x) = (x - x_0)^r \sum_{m=0}^{\infty} a_m (x - x_0)^m = a_0(x - x_0)^r + a_1(x - x_0)^{r+1} + \cdots \tag{4}$$

where r can be any real (or even complex) number and is determined by substituting (4) into (3) from the ***indicial equation*** (Sec. 5.3), along with the coefficients of (4). A second linearly independent solution of (3) may be of a similar form (with different r and a_m's) or may involve a logarithmic term. ***Bessel's equation*** is solved by the Frobenius method in Secs. 5.4 and 5.5.

"**Special functions**" is a common name for higher functions, as opposed to the usual functions of calculus. Most of them arise either as nonelementary integrals [see (24)–(44) in App. 3.1] or as solutions of (1) or (3). They get a name and notation and are included in the usual CASs if they are important in application or in theory. Of this kind, and particularly useful to the engineer and physicist, are ***Legendre's equation and polynomials*** $P_0, P_1, \cdots$ (Sec. 5.2), ***Gauss's hypergeometric equation and functions*** $F(a, b, c; x)$ (Sec. 5.3), and ***Bessel's equation and functions*** J_ν ***and*** Y_ν (Secs. 5.4, 5.5).

CHAPTER 6

Laplace Transforms

Laplace transforms are invaluable for any engineer's mathematical toolbox as they make solving linear ODEs and related initial value problems, as well as systems of linear ODEs, much easier. Applications abound: electrical networks, springs, mixing problems, signal processing, and other areas of engineering and physics.

The process of solving an ODE using the Laplace transform method consists of three steps, shown schematically in Fig. 113:

Step 1. The given ODE is transformed into an algebraic equation, called the **subsidiary equation**.

Step 2. The subsidiary equation is solved by purely algebraic manipulations.

Step 3. The solution in Step 2 is transformed back, resulting in the solution of the given problem.

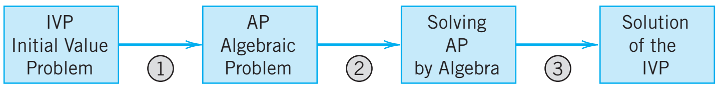


Fig. 113. Solving an IVP by Laplace transforms

The key motivation for learning about Laplace transforms is that the process of solving an ODE is simplified to an algebraic problem (and transformations). This type of mathematics that converts problems of calculus to algebraic problems is known as *operational calculus*. The Laplace transform method has two main advantages over the methods discussed in Chaps. 1–4:

I. Problems are solved more directly: Initial value problems are solved without first determining a general solution. Nonhomogenous ODEs are solved without first solving the corresponding homogeneous ODE.

II. More importantly, the use of the ***unit step function*** (***Heaviside function*** in Sec. 6.3) and ***Dirac's delta*** (in Sec. 6.4) make the method particularly powerful for problems with inputs (driving forces) that have discontinuities or represent short impulses or complicated periodic functions.

The following chart shows where to find information on the Laplace transform in this book.

Topic	Where to find it
ODEs, engineering applications and Laplace transforms	Chapter 6
PDEs, engineering applications and Laplace transforms	Section 12.11
List of general formulas of Laplace transforms	Section 6.8
List of Laplace transforms and inverses	Section 6.9
Note: Your CAS can handle most Laplace transforms.	

Prerequisite: Chap. 2
Sections that may be omitted in a shorter course: 6.5, 6.7
References and Answers to Problems: App. 1 Part A, App. 2.

6.1 Laplace Transform. Linearity. First Shifting Theorem (s-Shifting)

In this section, we learn about Laplace transforms and some of their properties. Because Laplace transforms are of basic importance to the engineer, the student should pay close attention to the material. Applications to ODEs follow in the next section.

Roughly speaking, the Laplace transform, when applied to a function, changes that function into a new function by using a process that involves integration. Details are as follows.

If $f(t)$ is a function defined for all $t \geqq 0$, its **Laplace transform**[1] is the integral of $f(t)$ times e^{-st} from $t = 0$ to ∞. It is a function of s, say, $F(s)$, and is denoted by $\mathcal{L}(f)$; thus

(1)
$$F(s) = \mathcal{L}(f) = \int_0^\infty e^{-st} f(t)\, dt.$$

Here we must assume that $f(t)$ is such that the integral exists (that is, has some finite value). This assumption is usually satisfied in applications—we shall discuss this near the end of the section.

[1] PIERRE SIMON MARQUIS DE LAPLACE (1749–1827), great French mathematician, was a professor in Paris. He developed the foundation of potential theory and made important contributions to celestial mechanics, astronomy in general, special functions, and probability theory. Napoléon Bonaparte was his student for a year. For Laplace's interesting political involvements, see Ref. [GenRef2], listed in App. 1.

The powerful practical Laplace transform techniques were developed over a century later by the English electrical engineer OLIVER HEAVISIDE (1850–1925) and were often called "Heaviside calculus."

We shall drop variables when this simplifies formulas without causing confusion. For instance, in (1) we wrote $\mathcal{L}(f)$ instead of $\mathcal{L}(f)(s)$ and in (1*) $\mathcal{L}^{-1}(F)$ instead of $\mathcal{L}^{-1}(F)(t)$.

Not only is the result $F(s)$ called the Laplace transform, but the operation just described, which yields $F(s)$ from a given $f(t)$, is also called the **Laplace transform**. It is an "**integral transform**"

$$F(s) = \int_0^\infty k(s, t) f(t)\, dt$$

with "**kernel**" $k(s, t) = e^{-st}$.

Note that the Laplace transform is called an integral transform because it transforms (changes) a function in one space to a function in another space by a *process of integration* that involves a kernel. The kernel or kernel function is a function of the variables in the two spaces and defines the integral transform.

Furthermore, the given function $f(t)$ in (1) is called the **inverse transform** of $F(s)$ and is denoted by $\mathcal{L}^{-1}(F)$; that is, we shall write

(1*) $$f(t) = \mathcal{L}^{-1}(F).$$

Note that (1) and (1*) together imply $\mathcal{L}^{-1}(\mathcal{L}(f)) = f$ and $\mathcal{L}(\mathcal{L}^{-1}(F)) = F$.

Notation

Original functions depend on t and their transforms on s—keep this in mind! Original functions are denoted by ***lowercase letters*** and their transforms by the same ***letters in capital***, so that $F(s)$ denotes the transform of $f(t)$, and $Y(s)$ denotes the transform of $y(t)$, and so on.

EXAMPLE 1 Laplace Transform

Let $f(t) = 1$ when $t \geqq 0$. Find $F(s)$.

Solution. From (1) we obtain by integration

$$\mathcal{L}(f) = \mathcal{L}(1) = \int_0^\infty e^{-st}\, dt = -\frac{1}{s} e^{-st} \Big|_0^\infty = \frac{1}{s} \qquad (s > 0).$$

Such an integral is called an **improper integral** and, by definition, is evaluated according to the rule

$$\int_0^\infty e^{-st} f(t)\, dt = \lim_{T\to\infty} \int_0^T e^{-st} f(t)\, dt.$$

Hence our convenient notation means

$$\int_0^\infty e^{-st}\, dt = \lim_{T\to\infty} \left[-\frac{1}{s} e^{-st} \right]_0^T = \lim_{T\to\infty} \left[-\frac{1}{s} e^{-sT} + \frac{1}{s} e^0 \right] = \frac{1}{s} \qquad (s > 0).$$

We shall use this notation throughout this chapter. ■

EXAMPLE 2 Laplace Transform $\mathcal{L}(e^{at})$ of the Exponential Function e^{at}

Let $f(t) = e^{at}$ when $t \geqq 0$, where a is a constant. Find $\mathcal{L}(f)$.

Solution. Again by (1),

$$\mathcal{L}(e^{at}) = \int_0^\infty e^{-st} e^{at}\, dt = \frac{1}{a - s} e^{-(s-a)t} \Big|_0^\infty;$$

hence, when $s - a > 0$,

$$\mathcal{L}(e^{at}) = \frac{1}{s - a}.$$ ■

Must we go on in this fashion and obtain the transform of one function after another directly from the definition? No! We can obtain new transforms from known ones by the use of the many general properties of the Laplace transform. Above all, the Laplace transform is a "linear operation," just as are differentiation and integration. By this we mean the following.

THEOREM 1

Linearity of the Laplace Transform

The Laplace transform is a linear operation; that is, for any functions $f(t)$ and $g(t)$ whose transforms exist and any constants a and b the transform of $af(t) + bg(t)$ exists, and

$$\mathcal{L}\{af(t) + bg(t)\} = a\mathcal{L}\{f(t)\} + b\mathcal{L}\{g(t)\}.$$

PROOF This is true because integration is a linear operation so that (1) gives

$$\begin{aligned}\mathcal{L}\{af(t) + bg(t)\} &= \int_0^\infty e^{-st}[af(t) + bg(t)]\,dt \\ &= a\int_0^\infty e^{-st} f(t)\,dt + b\int_0^\infty e^{-st} g(t)\,dt = a\mathcal{L}\{f(t)\} + b\mathcal{L}\{g(t)\}. \qquad \blacksquare\end{aligned}$$

EXAMPLE 3 **Application of Theorem 1: Hyperbolic Functions**

Find the transforms of $\cosh at$ and $\sinh at$.

Solution. Since $\cosh at = \frac{1}{2}(e^{at} + e^{-at})$ and $\sinh at = \frac{1}{2}(e^{at} - e^{-at})$, we obtain from Example 2 and Theorem 1

$$\mathcal{L}(\cosh at) = \frac{1}{2}(\mathcal{L}(e^{at}) + \mathcal{L}(e^{-at})) = \frac{1}{2}\left(\frac{1}{s-a} + \frac{1}{s+a}\right) = \frac{s}{s^2 - a^2}$$

$$\mathcal{L}(\sinh at) = \frac{1}{2}(\mathcal{L}(e^{at}) - \mathcal{L}(e^{-at})) = \frac{1}{2}\left(\frac{1}{s-a} - \frac{1}{s+a}\right) = \frac{a}{s^2 - a^2}. \qquad \blacksquare$$

EXAMPLE 4 **Cosine and Sine**

Derive the formulas

$$\mathcal{L}(\cos \omega t) = \frac{s}{s^2 + \omega^2}, \qquad \mathcal{L}(\sin \omega t) = \frac{\omega}{s^2 + \omega^2}.$$

Solution. We write $L_c = \mathcal{L}(\cos \omega t)$ and $L_s = \mathcal{L}(\sin \omega t)$. Integrating by parts and noting that the integral-free parts give no contribution from the upper limit ∞, we obtain

$$L_c = \int_0^\infty e^{-st}\cos \omega t\,dt = \frac{e^{-st}}{-s}\cos \omega t\,\Big|_0^\infty - \frac{\omega}{s}\int_0^\infty e^{-st}\sin \omega t\,dt = \frac{1}{s} - \frac{\omega}{s}L_s,$$

$$L_s = \int_0^\infty e^{-st}\sin \omega t\,dt = \frac{e^{-st}}{-s}\sin \omega t\,\Big|_0^\infty + \frac{\omega}{s}\int_0^\infty e^{-st}\cos \omega t\,dt = \frac{\omega}{s}L_c.$$

By substituting L_s into the formula for L_c on the right and then by substituting L_c into the formula for L_s on the right, we obtain

$$L_c = \frac{1}{s} - \frac{\omega}{s}\left(\frac{\omega}{s}L_c\right), \qquad L_c\left(1 + \frac{\omega^2}{s^2}\right) = \frac{1}{s}, \qquad L_c = \frac{s}{s^2 + \omega^2},$$

$$L_s = \frac{\omega}{s}\left(\frac{1}{s} - \frac{\omega}{s}L_s\right), \qquad L_s\left(1 + \frac{\omega^2}{s^2}\right) = \frac{\omega}{s^2}, \qquad L_s = \frac{\omega}{s^2 + \omega^2}.$$

Basic transforms are listed in Table 6.1. We shall see that from these almost all the others can be obtained by the use of the general properties of the Laplace transform. Formulas 1–3 are special cases of formula 4, which is proved by induction. Indeed, it is true for $n = 0$ because of Example 1 and $0! = 1$. We make the induction hypothesis that it holds for any integer $n \geqq 0$ and then get it for $n + 1$ directly from (1). Indeed, integration by parts first gives

$$\mathcal{L}(t^{n+1}) = \int_0^\infty e^{-st}t^{n+1}\,dt = -\frac{1}{s}e^{-st}t^{n+1}\Big|_0^\infty + \frac{n+1}{s}\int_0^\infty e^{-st}t^n\,dt.$$

Now the integral-free part is zero and the last part is $(n + 1)/s$ times $\mathcal{L}(t^n)$. From this and the induction hypothesis,

$$\mathcal{L}(t^{n+1}) = \frac{n+1}{s}\mathcal{L}(t^n) = \frac{n+1}{s}\cdot\frac{n!}{s^{n+1}} = \frac{(n+1)!}{s^{n+2}}.$$

This proves formula 4.

Table 6.1 Some Functions $f(t)$ and Their Laplace Transforms $\mathcal{L}(f)$

	$f(t)$	$\mathcal{L}(f)$		$f(t)$	$\mathcal{L}(f)$
1	1	$1/s$	7	$\cos \omega t$	$\dfrac{s}{s^2 + \omega^2}$
2	t	$1/s^2$	8	$\sin \omega t$	$\dfrac{\omega}{s^2 + \omega^2}$
3	t^2	$2!/s^3$	9	$\cosh at$	$\dfrac{s}{s^2 - a^2}$
4	t^n $(n = 0, 1, \cdots)$	$\dfrac{n!}{s^{n+1}}$	10	$\sinh at$	$\dfrac{a}{s^2 - a^2}$
5	t^a (a positive)	$\dfrac{\Gamma(a+1)}{s^{a+1}}$	11	$e^{at}\cos \omega t$	$\dfrac{s - a}{(s - a)^2 + \omega^2}$
6	e^{at}	$\dfrac{1}{s - a}$	12	$e^{at}\sin \omega t$	$\dfrac{\omega}{(s - a)^2 + \omega^2}$

$\Gamma(a + 1)$ in formula 5 is the so-called *gamma function* [(15) in Sec. 5.5 or (24) in App. A3.1]. We get formula 5 from (1), setting $st = x$:

$$\mathcal{L}(t^a) = \int_0^\infty e^{-st} t^a \, dt = \int_0^\infty e^{-x} \left(\frac{x}{s}\right)^a \frac{dx}{s} = \frac{1}{s^{a+1}} \int_0^\infty e^{-x} x^a \, dx$$

where $s > 0$. The last integral is precisely that defining $\Gamma(a + 1)$, so we have $\Gamma(a + 1)/s^{a+1}$, as claimed. (**CAUTION!** $\Gamma(a + 1)$ has x^a in the integral, not x^{a+1}.)

Note the formula 4 also follows from 5 because $\Gamma(n + 1) = n!$ for integer $n \geqq 0$.

Formulas 6–10 were proved in Examples 2–4. Formulas 11 and 12 will follow from 7 and 8 by "shifting," to which we turn next.

s-Shifting: Replacing s by $s - a$ in the Transform

The Laplace transform has the very useful property that, if we know the transform of $f(t)$, we can immediately get that of $e^{at} f(t)$, as follows.

THEOREM 2

First Shifting Theorem, s-Shifting

If $f(t)$ has the transform $F(s)$ (where $s > k$ for some k), then $e^{at} f(t)$ has the transform $F(s - a)$ (where $s - a > k$). In formulas,

$$\mathcal{L}\{e^{at} f(t)\} = F(s - a)$$

or, if we take the inverse on both sides,

$$e^{at} f(t) = \mathcal{L}^{-1}\{F(s - a)\}.$$

PROOF We obtain $F(s - a)$ by replacing s with $s - a$ in the integral in (1), so that

$$F(s - a) = \int_0^\infty e^{-(s-a)t} f(t) \, dt = \int_0^\infty e^{-st} [e^{at} f(t)] \, dt = \mathcal{L}\{e^{at} f(t)\}.$$

If $F(s)$ exists (i.e., is finite) for s greater than some k, then our first integral exists for $s - a > k$. Now take the inverse on both sides of this formula to obtain the second formula in the theorem. (**CAUTION!** $-a$ in $F(s - a)$ but $+a$ in $e^{at} f(t)$.) ■

EXAMPLE 5 **s-Shifting: Damped Vibrations. Completing the Square**

From Example 4 and the first shifting theorem we immediately obtain formulas 11 and 12 in Table 6.1,

$$\mathcal{L}\{e^{at} \cos \omega t\} = \frac{s - a}{(s - a)^2 + \omega^2}, \qquad \mathcal{L}\{e^{at} \sin \omega t\} = \frac{\omega}{(s - a)^2 + \omega^2}.$$

For instance, use these formulas to find the inverse of the transform

$$\mathcal{L}(f) = \frac{3s - 137}{s^2 + 2s + 401}.$$

Solution. Applying the inverse transform, using its linearity (Prob. 24), and completing the square, we obtain

$$f = \mathcal{L}^{-1}\left\{\frac{3(s+1) - 140}{(s+1)^2 + 400}\right\} = 3\mathcal{L}^{-1}\left\{\frac{s+1}{(s+1)^2 + 20^2}\right\} - 7\mathcal{L}^{-1}\left\{\frac{20}{(s+1)^2 + 20^2}\right\}.$$

We now see that the inverse of the right side is the damped vibration (Fig. 114)

$$f(t) = e^{-t}(3\cos 20t - 7\sin 20t).$$

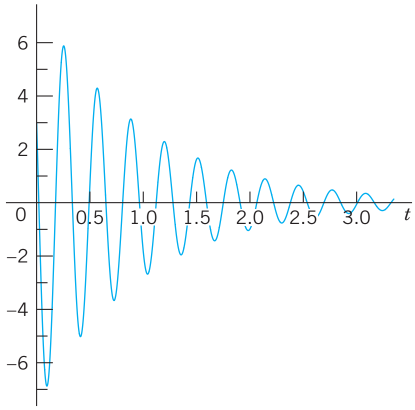


Fig. 114. Vibrations in Example 5

Existence and Uniqueness of Laplace Transforms

This is not a big *practical* problem because in most cases we can check the solution of an ODE without too much trouble. Nevertheless we should be aware of some basic facts.

A function $f(t)$ has a Laplace transform if it does not grow too fast, say, if for all $t \geqq 0$ and some constants M and k it satisfies the "**growth restriction**"

$$|f(t)| \leqq Me^{kt}. \tag{2}$$

(The growth restriction (2) is sometimes called "growth of exponential order," which may be misleading since it hides that the exponent must be kt, not kt^2 or similar.)

$f(t)$ need not be continuous, but it should not be too bad. The technical term (generally used in mathematics) is *piecewise continuity*. $f(t)$ is **piecewise continuous** on a finite interval $a \leqq t \leqq b$ where f is defined, if this interval can be divided into *finitely many* subintervals in each of which f is continuous and has finite limits as t approaches either endpoint of such a subinterval from the interior. This then gives **finite jumps** as in Fig. 115 as the only possible discontinuities, but this suffices in most applications, and so does the following theorem.

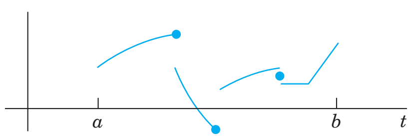


Fig. 115. Example of a piecewise continuous function $f(t)$. (The dots mark the function values at the jumps.)

THEOREM 3

Existence Theorem for Laplace Transforms

If $f(t)$ is defined and piecewise continuous on every finite interval on the semi-axis $t \geqq 0$ and satisfies (2) for all $t \geqq 0$ and some constants M and k, then the Laplace transform $\mathscr{L}(f)$ exists for all $s > k$.

PROOF Since $f(t)$ is piecewise continuous, $e^{-st}f(t)$ is integrable over any finite interval on the t-axis. From (2), assuming that $s > k$ (to be needed for the existence of the last of the following integrals), we obtain the proof of the existence of $\mathscr{L}(f)$ from

$$|\mathscr{L}(f)| = \left|\int_0^\infty e^{-st}f(t)\,dt\right| \leqq \int_0^\infty |f(t)|e^{-st}\,dt \leqq \int_0^\infty Me^{kt}e^{-st}\,dt = \frac{M}{s-k}. \quad \blacksquare$$

Note that (2) can be readily checked. For instance, $\cosh t < e^t$, $t^n < n!e^t$ (because $t^n/n!$ is a single term of the Maclaurin series), and so on. A function that does not satisfy (2) for any M and k is e^{t^2} (take logarithms to see it). We mention that the conditions in Theorem 3 are sufficient rather than necessary (see Prob. 22).

Uniqueness. If the Laplace transform of a given function exists, it is uniquely determined. Conversely, it can be shown that if two functions (both defined on the positive real axis) have the same transform, these functions cannot differ over an interval of positive length, although they may differ at isolated points (see Ref. [A14] in App. 1). Hence we may say that the inverse of a given transform is essentially unique. In particular, if two *continuous* functions have the same transform, they are completely identical.

PROBLEM SET 6.1

1–16 LAPLACE TRANSFORMS

Find the transform. Show the details of your work. Assume that a, b, ω, θ are constants.

1. $3t + 12$

2. $(a - bt)^2$

3. $\cos \pi t$

4. $\cos^2 \omega t$

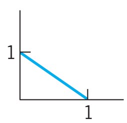
5. $e^{2t} \sinh t$

6. $e^{-t} \sinh 4t$

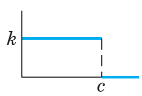
7. $\sin(\omega t + \theta)$

8. $1.5 \sin(3t - \pi/2)$

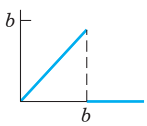
9.



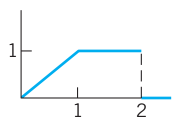
10.



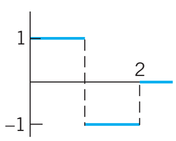
11.



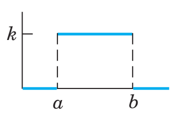
12.



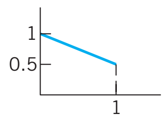
13.



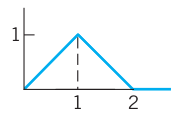
14.



15.



16.



17–24 SOME THEORY

17. Table 6.1. Convert this table to a table for finding inverse transforms (with obvious changes, e.g., $\mathscr{L}^{-1}(1/s^n) = t^{n-1}/(n-1)$, etc).

18. Using $\mathscr{L}(f)$ in Prob. 10, find $\mathscr{L}(f_1)$, where $f_1(t) = 0$ if $t \leqq 2$ and $f_1(t) = 1$ if $t > 2$.

19. Table 6.1. Derive formula 6 from formulas 9 and 10.

20. Nonexistence. Show that e^{t^2} does not satisfy a condition of the form (2).

21. Nonexistence. Give simple examples of functions (defined for all $t \geqq 0$) that have no Laplace transform.

22. Existence. Show that $\mathscr{L}(1/\sqrt{t}) = \sqrt{\pi/s}$. [Use (30) $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ in App. 3.1.] Conclude from this that the conditions in Theorem 3 are sufficient but not necessary for the existence of a Laplace transform.

23. Change of scale. If $\mathcal{L}(f(t)) = F(s)$ and c is any positive constant, show that $\mathcal{L}(f(ct)) = F(s/c)/c$ (***Hint:*** Use (1).) Use this to obtain $\mathcal{L}(\cos \omega t)$ from $\mathcal{L}(\cos t)$.

24. Inverse transform. Prove that $\mathcal{L}^{-1}$ is linear. *Hint:* Use the fact that $\mathcal{L}$ is linear.

25–32 INVERSE LAPLACE TRANSFORMS

Given $F(s) = \mathcal{L}(f)$, find $f(t)$. a, b, L, n are constants. Show the details of your work.

25. $\dfrac{0.2s + 1.8}{s^2 + 3.24}$ **26.** $\dfrac{5s + 1}{s^2 - 25}$

27. $\dfrac{s}{L^2s^2 + n^2\pi^2}$ **28.** $\dfrac{1}{(s + \sqrt{2})(s - \sqrt{3})}$

29. $\dfrac{12}{s^4} - \dfrac{228}{s^6}$ **30.** $\dfrac{4s + 32}{s^2 - 16}$

31. $\dfrac{s + 10}{s^2 - s - 2}$ **32.** $\dfrac{1}{(s + a)(s + b)}$

33–45 APPLICATION OF *s*-SHIFTING

In Probs. 33–36 find the transform. In Probs. 37–45 find the inverse transform. Show the details of your work.

33. t^2e^{-3t} **34.** $ke^{-at}\cos \omega t$

35. $0.5e^{-4.5t}\sin 2\pi t$ **36.** $\sinh t \cos t$

37. $\dfrac{\pi}{(s + \pi)^2}$ **38.** $\dfrac{6}{(s + 1)^3}$

39. $\dfrac{21}{(s + \sqrt{2})^4}$ **40.** $\dfrac{4}{s^2 - 2s - 3}$

41. $\dfrac{\pi}{s^2 + 10\pi s + 24\pi^2}$

42. $\dfrac{a_0}{s + 1} + \dfrac{a_1}{(s + 1)^2} + \dfrac{a_2}{(s + 1)^3}$

43. $\dfrac{2s - 1}{s^2 - 6s + 18}$ **44.** $\dfrac{a(s + k) + b\pi}{(s + k)^2 + \pi^2}$

45. $\dfrac{k_0(s + a) + k_1}{(s + a)^2}$

6.2 Transforms of Derivatives and Integrals. ODEs

The Laplace transform is a method of solving ODEs and initial value problems. The crucial idea is that ***operations of calculus on functions are replaced by operations of algebra on transforms***. Roughly, *differentiation* of $f(t)$ will correspond to *multiplication* of $\mathcal{L}(f)$ by s (see Theorems 1 and 2) and *integration* of $f(t)$ to *division* of $\mathcal{L}(f)$ by s. To solve ODEs, we must first consider the Laplace transform of derivatives. You have encountered such an idea in your study of logarithms. Under the application of the natural logarithm, a product of numbers becomes a sum of their logarithms, a division of numbers becomes their difference of logarithms (see Appendix 3, formulas (2), (3)). To simplify calculations was one of the main reasons that logarithms were invented in pre-computer times.

THEOREM 1

Laplace Transform of Derivatives

The transforms of the first and second derivatives of $f(t)$ satisfy

(1) $$\mathcal{L}(f') = s\mathcal{L}(f) - f(0)$$

(2) $$\mathcal{L}(f'') = s^2\mathcal{L}(f) - sf(0) - f'(0).$$

Formula (1) *holds if $f(t)$ is continuous for all $t \geqq 0$ and satisfies the growth restriction* (2) *in Sec.* 6.1 *and $f'(t)$ is piecewise continuous on every finite interval on the semi-axis $t \geqq 0$. Similarly,* (2) *holds if f and f' are continuous for all $t \geqq 0$ and satisfy the growth restriction and f'' is piecewise continuous on every finite interval on the semi-axis $t \geqq 0$.*

PROOF We prove (1) first under the *additional assumption* that f' is continuous. Then, by the definition and integration by parts,

$$\mathscr{L}(f') = \int_0^\infty e^{-st} f'(t)\,dt = \left[e^{-st}f(t)\right]\Big|_0^\infty + s\int_0^\infty e^{-st}f(t)\,dt.$$

Since f satisfies (2) in Sec. 6.1, the integrated part on the right is zero at the upper limit when $s > k$, and at the lower limit it contributes $-f(0)$. The last integral is $\mathscr{L}(f)$. It exists for $s > k$ because of Theorem 3 in Sec. 6.1. Hence $\mathscr{L}(f')$ exists when $s > k$ and (1) holds.

If f' is merely piecewise continuous, the proof is similar. In this case the interval of integration of f' must be broken up into parts such that f' is continuous in each such part.

The proof of (2) now follows by applying (1) to f'' and then substituting (1), that is

$$\mathscr{L}(f'') = s\mathscr{L}(f') - f'(0) = s[s\mathscr{L}(f) - f(0)] = s^2\mathscr{L}(f) - sf(0) - f'(0).$$ ■

Continuing by substitution as in the proof of (2) and using induction, we obtain the following extension of Theorem 1.

THEOREM 2

Laplace Transform of the Derivative $f^{(n)}$ of Any Order

Let $f, f', \cdots, f^{(n-1)}$ be continuous for all $t \geqq 0$ and satisfy the growth restriction (2) in Sec. 6.1. Furthermore, let $f^{(n)}$ be piecewise continuous on every finite interval on the semi-axis $t \geqq 0$. Then the transform of $f^{(n)}$ satisfies

$$\mathscr{L}(f^{(n)}) = s^n\mathscr{L}(f) - s^{n-1}f(0) - s^{n-2}f'(0) - \cdots - f^{(n-1)}(0). \tag{3}$$

EXAMPLE 1 Transform of a Resonance Term (Sec. 2.8)

Let $f(t) = t \sin \omega t$. Then $f(0) = 0$, $f'(t) = \sin \omega t + \omega t \cos \omega t$, $f'(0) = 0$, $f'' = 2\omega \cos \omega t - \omega^2 t \sin \omega t$. Hence by (2),

$$\mathscr{L}(f'') = 2\omega\frac{s}{s^2+\omega^2} - \omega^2\mathscr{L}(f) = s^2\mathscr{L}(f), \qquad \text{thus} \qquad \mathscr{L}(f) = \mathscr{L}(t\sin\omega t) = \frac{2\omega s}{(s^2+\omega^2)^2}.$$ ■

EXAMPLE 2 Formulas 7 and 8 in Table 6.1, Sec. 6.1

This is a third derivation of $\mathscr{L}(\cos \omega t)$ and $\mathscr{L}(\sin \omega t)$; cf. Example 4 in Sec. 6.1. Let $f(t) = \cos \omega t$. Then $f(0) = 1$, $f'(0) = 0$, $f''(t) = -\omega^2 \cos \omega t$. From this and (2) we obtain

$$\mathscr{L}(f'') = s^2\mathscr{L}(f) - s = -\omega^2\mathscr{L}(f). \qquad \text{By algebra,} \qquad \mathscr{L}(\cos\omega t) = \frac{s}{s^2+\omega^2}.$$

Similarly, let $g = \sin \omega t$. Then $g(0) = 0$, $g' = \omega \cos \omega t$. From this and (1) we obtain

$$\mathscr{L}(g') = s\mathscr{L}(g) = \omega\mathscr{L}(\cos\omega t). \qquad \text{Hence,} \qquad \mathscr{L}(\sin\omega t) = \frac{\omega}{s}\mathscr{L}(\cos\omega t) = \frac{\omega}{s^2+\omega^2}.$$ ■

Laplace Transform of the Integral of a Function

Differentiation and integration are inverse operations, and so are multiplication and division. Since differentiation of a function $f(t)$ (roughly) corresponds to multiplication of its transform $\mathscr{L}(f)$ by s, we expect integration of $f(t)$ to correspond to division of $\mathscr{L}(f)$ by s:

THEOREM 3

Laplace Transform of Integral

Let $F(s)$ denote the transform of a function $f(t)$ which is piecewise continuous for $t \geqq 0$ and satisfies a growth restriction (2), Sec. 6.1. Then, for $s > 0$, $s > k$, and $t > 0$,

$$\mathcal{L}\left\{\int_0^t f(\tau)\, d\tau\right\} = \frac{1}{s}F(s), \qquad \text{thus} \qquad \int_0^t f(\tau)\, d\tau = \mathcal{L}^{-1}\left\{\frac{1}{s}F(s)\right\}. \tag{4}$$

PROOF Denote the integral in (4) by $g(t)$. Since $f(t)$ is piecewise continuous, $g(t)$ is continuous, and (2), Sec. 6.1, gives

$$|g(t)| = \left|\int_0^t f(\tau)\, d\tau\right| \leqq \int_0^t |f(\tau)|\, d\tau \leqq M\int_0^t e^{k\tau}\, d\tau = \frac{M}{k}(e^{kt} - 1) \leqq \frac{M}{k}e^{kt} \qquad (k > 0).$$

This shows that $g(t)$ also satisfies a growth restriction. Also, $g'(t) = f(t)$, except at points at which $f(t)$ is discontinuous. Hence $g'(t)$ is piecewise continuous on each finite interval and, by Theorem 1, since $g(0) = 0$ (the integral from 0 to 0 is zero)

$$\mathcal{L}\{f(t)\} = \mathcal{L}\{g'(t)\} = s\mathcal{L}\{g(t)\} - g(0) = s\mathcal{L}\{g(t)\}.$$

Division by s and interchange of the left and right sides gives the first formula in (4), from which the second follows by taking the inverse transform on both sides. ■

EXAMPLE 3 **Application of Theorem 3: Formulas 19 and 20 in the Table of Sec. 6.9**

Using Theorem 3, find the inverse of $\dfrac{1}{s(s^2 + \omega^2)}$ and $\dfrac{1}{s^2(s^2 + \omega^2)}$.

Solution. From Table 6.1 in Sec. 6.1 and the integration in (4) (second formula with the sides interchanged) we obtain

$$\mathcal{L}^{-1}\left\{\frac{1}{s^2 + \omega^2}\right\} = \frac{\sin \omega t}{\omega}, \qquad \mathcal{L}^{-1}\left\{\frac{1}{s(s^2 + \omega^2)}\right\} = \int_0^t \frac{\sin \omega \tau}{\omega}\, d\tau = \frac{1}{\omega^2}(1 - \cos \omega t).$$

This is formula 19 in Sec. 6.9. Integrating this result again and using (4) as before, we obtain formula 20 in Sec. 6.9:

$$\mathcal{L}^{-1}\left\{\frac{1}{s^2(s^2 + \omega^2)}\right\} = \frac{1}{\omega^2}\int_0^t (1 - \cos \omega\tau)\, d\tau = \left[\frac{\tau}{\omega^2} - \frac{\sin \omega\tau}{\omega^3}\right]_0^t = \frac{t}{\omega^2} - \frac{\sin \omega t}{\omega^3}.$$

It is typical that results such as these can be found in several ways. In this example, try partial fraction reduction. ■

Differential Equations, Initial Value Problems

Let us now discuss how the Laplace transform method solves ODEs and initial value problems. We consider an initial value problem

$$y'' + ay' + by = r(t), \qquad y(0) = K_0, \qquad y'(0) = K_1 \tag{5}$$

where a and b are constant. Here $r(t)$ is the given **input** (*driving force*) applied to the mechanical or electrical system and $y(t)$ is the **output** (*response to the input*) to be obtained. In Laplace's method we do three steps:

Step 1. Setting up the subsidiary equation. This is an algebraic equation for the transform $Y = \mathcal{L}(y)$ obtained by transforming (5) by means of (1) and (2), namely,

$$[s^2Y - sy(0) - y'(0)] + a[sY - y(0)] + bY = R(s)$$

where $R(s) = \mathcal{L}(r)$. Collecting the Y-terms, we have the subsidiary equation

$$(s^2 + as + b)Y = (s + a)y(0) + y'(0) + R(s).$$

Step 2. Solution of the subsidiary equation by algebra. We divide by $s^2 + as + b$ and use the so-called **transfer function**

$$Q(s) = \frac{1}{s^2 + as + b} = \frac{1}{(s + \frac{1}{2}a)^2 + b - \frac{1}{4}a^2}. \tag{6}$$

(Q is often denoted by H, but we need H much more frequently for other purposes.) This gives the solution

$$Y(s) = [(s + a)y(0) + y'(0)]Q(s) + R(s)Q(s). \tag{7}$$

If $y(0) = y'(0) = 0$, this is simply $Y = RQ$; hence

$$Q = \frac{Y}{R} = \frac{\mathcal{L}(\text{output})}{\mathcal{L}(\text{input})}$$

and this explains the name of Q. Note that ***Q depends neither on $r(t)$ nor on the initial conditions*** (but only on a and b).

Step 3. Inversion of Y to obtain $y = \mathcal{L}^{-1}(Y)$. We reduce (7) (usually by ***partial fractions*** as in calculus) to a sum of terms whose inverses can be found from the tables (e.g., in Sec. 6.1 or Sec. 6.9) or by a CAS, so that we obtain the solution $y(t) = \mathcal{L}^{-1}(Y)$ *of* (5).

EXAMPLE 4 Initial Value Problem: The Basic Laplace Steps

Solve

$$y'' - y = t, \qquad y(0) = 1, \qquad y'(0) = 1.$$

Solution. ***Step 1.*** From (2) and Table 6.1 we get the subsidiary equation [with $Y = \mathcal{L}(y)$]

$$s^2Y - sy(0) - y'(0) - Y = 1/s^2, \qquad \text{thus} \qquad (s^2 - 1)Y = s + 1 + 1/s^2.$$

Step 2. The transfer function is $Q = 1/(s^2 - 1)$, and (7) becomes

$$Y = (s + 1)Q + \frac{1}{s^2}Q = \frac{s + 1}{s^2 - 1} + \frac{1}{s^2(s^2 - 1)}.$$

Simplification of the first fraction and an expansion of the last fraction gives

$$Y = \frac{1}{s - 1} + \left(\frac{1}{s^2 - 1} - \frac{1}{s^2}\right).$$

Step 3. From this expression for Y and Table 6.1 we obtain the solution

$$y(t) = \mathcal{L}^{-1}(Y) = \mathcal{L}^{-1}\left\{\frac{1}{s-1}\right\} + \mathcal{L}^{-1}\left\{\frac{1}{s^2-1}\right\} - \mathcal{L}^{-1}\left\{\frac{1}{s^2}\right\} = e^t + \sinh t - t.$$

The diagram in Fig. 116 summarizes our approach. ■

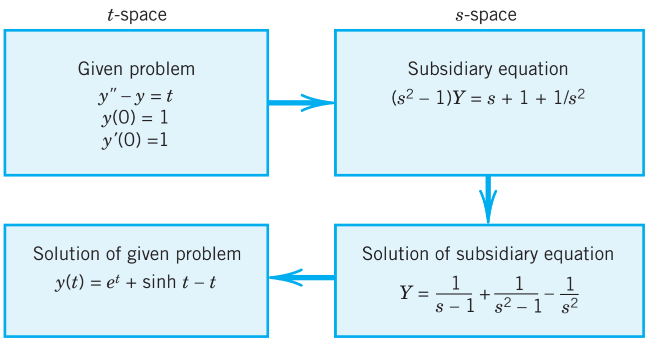


Fig. 116. Steps of the Laplace transform method

EXAMPLE 5 Comparison with the Usual Method

Solve the initial value problem

$$y'' + y' + 9y = 0. \qquad y(0) = 0.16, \qquad y'(0) = 0.$$

Solution. From (1) and (2) we see that the subsidiary equation is

$$s^2Y - 0.16s + sY - 0.16 + 9Y = 0, \qquad \text{thus} \qquad (s^2 + s + 9)Y = 0.16(s + 1).$$

The solution is

$$Y = \frac{0.16(s+1)}{s^2 + s + 9} = \frac{0.16(s + \frac{1}{2}) + 0.08}{(s + \frac{1}{2})^2 + \frac{35}{4}}.$$

Hence by the first shifting theorem and the formulas for cos and sin in Table 6.1 we obtain

$$\begin{aligned} y(t) = \mathcal{L}^{-1}(Y) &= e^{-t/2}\left(0.16 \cos \sqrt{\frac{35}{4}}\, t + \frac{0.08}{\frac{1}{2}\sqrt{35}} \sin \sqrt{\frac{35}{4}}\, t\right) \\ &= e^{-0.5t}(0.16 \cos 2.96t + 0.027 \sin 2.96t). \end{aligned}$$

This agrees with Example 2, Case (III) in Sec. 2.4. The work was less. ■

Advantages of the Laplace Method

1. *Solving a nonhomogeneous ODE does not require first solving the homogeneous ODE.* See Example 4.
2. *Initial values are automatically taken care of.* See Examples 4 and 5.
3. *Complicated inputs* $r(t)$ (right sides of linear ODEs) *can be handled very efficiently,* as we show in the next sections.

EXAMPLE 6 Shifted Data Problems

This means initial value problems with initial conditions given at some $t = t_0 > 0$ instead of $t = 0$. For such a problem set $t = \tilde{t} + t_0$, so that $t = t_0$ gives $\tilde{t} = 0$ and the Laplace transform can be applied. For instance, solve

$$y'' + y = 2t, \quad y(\tfrac{1}{4}\pi) = \tfrac{1}{2}\pi, \quad y'(\tfrac{1}{4}\pi) = 2 - \sqrt{2}.$$

Solution. We have $t_0 = \frac{1}{4}\pi$ and we set $t = \tilde{t} + \frac{1}{4}\pi$. Then the problem is

$$\tilde{y}'' + \tilde{y} = 2(\tilde{t} + \tfrac{1}{4}\pi), \quad \tilde{y}(0) = \tfrac{1}{2}\pi, \quad \tilde{y}'(0) = 2 - \sqrt{2}$$

where $\tilde{y}(\tilde{t}) = y(t)$. Using (2) and Table 6.1 and denoting the transform of $\tilde{y}$ by $\tilde{Y}$, we see that the subsidiary equation of the "shifted" initial value problem is

$$s^2\tilde{Y} - s \cdot \tfrac{1}{2}\pi - (2 - \sqrt{2}) + \tilde{Y} = \frac{2}{s^2} + \frac{\frac{1}{2}\pi}{s}, \qquad \text{thus} \qquad (s^2 + 1)\tilde{Y} = \frac{2}{s^2} + \frac{\frac{1}{2}\pi}{s} + \frac{1}{2}\pi s + 2 - \sqrt{2}.$$

Solving this algebraically for $\tilde{Y}$, we obtain

$$\tilde{Y} = \frac{2}{(s^2 + 1)s^2} + \frac{\frac{1}{2}\pi}{(s^2 + 1)s} + \frac{\frac{1}{2}\pi s}{s^2 + 1} + \frac{2 - \sqrt{2}}{s^2 + 1}.$$

The inverse of the first two terms can be seen from Example 3 (with $\omega = 1$), and the last two terms give cos and sin,

$$\begin{aligned} \tilde{y} = \mathcal{L}^{-1}(\tilde{Y}) &= 2(\tilde{t} - \sin \tilde{t}) + \tfrac{1}{2}\pi(1 - \cos \tilde{t}) + \tfrac{1}{2}\pi \cos \tilde{t} + (2 - \sqrt{2}) \sin \tilde{t} \\ &= 2\tilde{t} + \tfrac{1}{2}\pi - \sqrt{2} \sin \tilde{t}. \end{aligned}$$

Now $\tilde{t} = t - \frac{1}{4}\pi$, $\sin \tilde{t} = \dfrac{1}{\sqrt{2}}(\sin t - \cos t)$, so that the answer (the solution) is

$$y = 2t - \sin t + \cos t.$$

PROBLEM SET 6.2

1–11 INITIAL VALUE PROBLEMS (IVPS)

Solve the IVPs by the Laplace transform. If necessary, use partial fraction expansion as in Example 4 of the text. Show all details.

1. $y' + 5.2y = 19.4 \sin 2t, \quad y(0) = 0$
2. $y' + 2y = 0, \quad y(0) = 1.5$
3. $y'' - y' - 6y = 0, \quad y(0) = 11, \quad y'(0) = 28$
4. $y'' + 9y = 10e^{-t}, \quad y(0) = 0, \quad y'(0) = 0$
5. $y'' - \frac{1}{4}y = 0, \quad y(0) = 12, \quad y'(0) = 0$
6. $y'' - 6y' + 5y = 29 \cos 2t, \quad y(0) = 3.2,$ $y'(0) = 6.2$
7. $y'' + 7y' + 12y = 21e^{3t}, \quad y(0) = 3.5,$ $y'(0) = -10$
8. $y'' - 4y' + 4y = 0, \quad y(0) = 8.1, \quad y'(0) = 3.9$
9. $y'' - 4y' + 3y = 6t - 8, \quad y(0) = 0, \quad y'(0) = 0$
10. $y'' + 0.04y = 0.02t^2, \quad y(0) = -25, \quad y'(0) = 0$
11. $y'' + 3y' + 2.25y = 9t^3 + 64, \quad y(0) = 1,$ $y'(0) = 31.5$

12–15 SHIFTED DATA PROBLEMS

Solve the shifted data IVPs by the Laplace transform. Show the details.

12. $y'' - 2y' - 3y = 0, \quad y(4) = -3,$ $y'(4) = -17$
13. $y' - 6y = 0, \quad y(-1) = 4$
14. $y'' + 2y' + 5y = 50t - 100, \quad y(2) = -4,$ $y'(2) = 14$
15. $y'' + 3y' - 4y = 6e^{2t-3}, \quad y(1.5) = 4,$ $y'(1.5) = 5$

16–21 OBTAINING TRANSFORMS BY DIFFERENTIATION

Using (1) or (2), find $\mathcal{L}(f)$ if $f(t)$ equals:

16. $t \cos 4t$
17. te^{-at}
18. $\cos^2 2t$
19. $\sin^2 \omega t$
20. $\sin^4 t$. Use Prob. 19.
21. $\cosh^2 t$

22. PROJECT. Further Results by Differentiation. Proceeding as in Example 1, obtain

(a) $\mathcal{L}(t \cos \omega t) = \dfrac{s^2 - \omega^2}{(s^2 + \omega^2)^2}$

and from this and Example 1: **(b)** formula 21, **(c)** 22, **(d)** 23 in Sec. 6.9,

(e) $\mathcal{L}(t \cosh at) = \dfrac{s^2 + a^2}{(s^2 - a^2)^2}$,

(f) $\mathcal{L}(t \sinh at) = \dfrac{2as}{(s^2 - a^2)^2}$.

23–29 INVERSE TRANSFORMS BY INTEGRATION

Using Theorem 3, find $f(t)$ if $\mathcal{L}(F)$ equals:

23. $\dfrac{3}{s^2 + s/4}$

24. $\dfrac{20}{s^3 - 2\pi s^2}$

25. $\dfrac{1}{s(s^2 + \omega^2)}$

26. $\dfrac{1}{s^4 - s^2}$

27. $\dfrac{s + 1}{s^4 + 9s^2}$

28. $\dfrac{3s + 4}{s^4 + k^2 s^2}$

29. $\dfrac{1}{s^3 + as^2}$

30. PROJECT. Comments on Sec. 6.2. (a) Give reasons why Theorems 1 and 2 are more important than Theorem 3.

(b) Extend Theorem 1 by showing that if $f(t)$ is continuous, except for an ordinary discontinuity (finite jump) at some $t = a$ (>0), the other conditions remaining as in Theorem 1, then (see Fig. 117)

(1*) $\mathcal{L}(f') = s\mathcal{L}(f) - f(0) - [f(a + 0) - f(a - 0)]e^{-as}$.

(c) Verify (1*) for $f(t) = e^{-t}$ if $0 < t < 1$ and 0 if $t > 1$.

(d) Compare the Laplace transform of solving ODEs with the method in Chap. 2. Give examples of your own to illustrate the advantages of the present method (to the extent we have seen them so far).

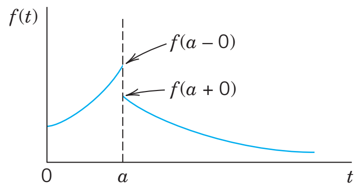


Fig. 117. Formula (1*)

6.3 Unit Step Function (Heaviside Function). Second Shifting Theorem (t-Shifting)

This section and the next one are extremely important because we shall now reach the point where the Laplace transform method shows its real power in applications and its superiority over the classical approach of Chap. 2. The reason is that we shall introduce two auxiliary functions, the *unit step function* or *Heaviside function* $u(t - a)$ (below) and *Dirac's delta* $\delta(t - a)$ (in Sec. 6.4). These functions are suitable for solving ODEs with complicated right sides of considerable engineering interest, such as single waves, inputs (driving forces) that are discontinuous or act for some time only, periodic inputs more general than just cosine and sine, or impulsive forces acting for an instant (hammerblows, for example).

Unit Step Function (Heaviside Function) $u(t - a)$

The **unit step function** or **Heaviside function** $u(t - a)$ is 0 for $t < a$, has a jump of size 1 at $t = a$ (where we can leave it undefined), and is 1 for $t > a$, in a formula:

$$u(t - a) = \begin{cases} 0 & \text{if } t < a \\ 1 & \text{if } t > a \end{cases} \qquad (a \geqq 0). \tag{1}$$

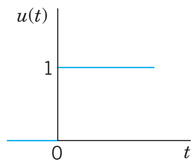


Fig. 118. Unit step function $u(t)$

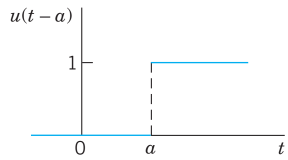


Fig. 119. Unit step function $u(t - a)$

Figure 118 shows the special case $u(t)$, which has its jump at zero, and Fig. 119 the general case $u(t - a)$ for an arbitrary positive a. (For Heaviside, see Sec. 6.1.)

The transform of $u(t - a)$ follows directly from the defining integral in Sec. 6.1,

$$\mathcal{L}\{u(t-a)\} = \int_0^\infty e^{-st} u(t-a)\, dt = \int_0^\infty e^{-st} \cdot 1\, dt = -\frac{e^{-st}}{s}\bigg|_{t=a}^{\infty};$$

here the integration begins at $t = a\,(\geqq 0)$ because $u(t - a)$ is 0 for $t < a$. Hence

$$\mathcal{L}\{u(t-a)\} = \frac{e^{-as}}{s} \qquad (s > 0). \tag{2}$$

The unit step function is a typical "engineering function" made to measure for engineering applications, which often involve functions (mechanical or electrical driving forces) that are either "off" or "on." Multiplying functions $f(t)$ with $u(t - a)$, we can produce all sorts of effects. The simple basic idea is illustrated in Figs. 120 and 121. In Fig. 120 the given function is shown in (A). In (B) it is switched off between $t = 0$ and $t = 2$ (because $u(t - 2) = 0$ when $t < 2$) and is switched on beginning at $t = 2$. In (C) it is shifted to the right by 2 units, say, for instance, by 2 sec, so that it begins 2 sec later in the same fashion as before. More generally we have the following.

Let $f(t) = 0$ for all negative t. Then $f(t - a)u(t - a)$ with $a > 0$ is $f(t)$ ***shifted*** *(translated) to the right by the amount a.*

Figure 121 shows the effect of many unit step functions, three of them in (A) and infinitely many in (B) when continued periodically to the right; this is the effect of a rectifier that clips off the negative half-waves of a sinusoidal voltage. **CAUTION!** Make sure that you fully understand these figures, in particular the difference between parts (B) and (C) of Fig. 120. Figure 120(C) will be applied next.

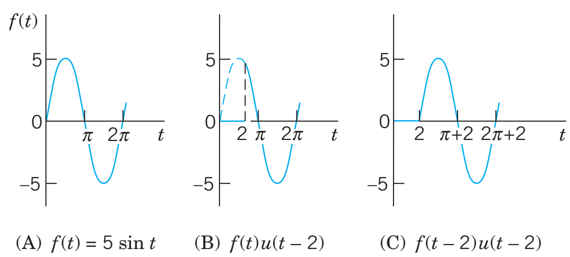


Fig. 120. Effects of the unit step function: (A) Given function. (B) Switching off and on. (C) Shift.

(A) $k[u(t-1) - 2u(t-4) + u(t-6)]$ (B) $4 \sin(\frac{1}{2}\pi t)[u(t) - u(t-2) + u(t-4) - + \cdots]$

Fig. 121. Use of many unit step functions.

Time Shifting (t-Shifting): Replacing t by $t - a$ in $f(t)$

The first shifting theorem ("s-shifting") in Sec. 6.1 concerned transforms $F(s) = \mathcal{L}\{f(t)\}$ and $F(s - a) = \mathcal{L}\{e^{at}f(t)\}$. The second shifting theorem will concern functions $f(t)$ and $f(t - a)$. Unit step functions are just tools, and the theorem will be needed to apply them in connection with any other functions.

THEOREM 1

Second Shifting Theorem; Time Shifting

If $f(t)$ has the transform $F(s)$, then the "**shifted function**"

$$\tilde{f}(t) = f(t-a)u(t-a) = \begin{cases} 0 & \text{if } t < a \\ f(t-a) & \text{if } t > a \end{cases} \tag{3}$$

has the transform $e^{-as}F(s)$. *That is, if* $\mathcal{L}\{f(t)\} = F(s)$, *then*

$$\mathcal{L}\{f(t-a)u(t-a)\} = e^{-as}F(s). \tag{4}$$

Or, if we take the inverse on both sides, we can write

$$f(t-a)u(t-a) = \mathcal{L}^{-1}\{e^{-as}F(s)\}. \tag{4*}$$

Practically speaking, if we know $F(s)$, we can obtain the transform of (3) by multiplying $F(s)$ by e^{-as}. In Fig. 120, the transform of $5 \sin t$ is $F(s) = 5/(s^2 + 1)$, hence the shifted function $5 \sin(t - 2)u(t - 2)$ shown in Fig. 120(C) has the transform

$$e^{-2s}F(s) = 5e^{-2s}/(s^2 + 1).$$

PROOF We prove Theorem 1. In (4), on the right, we use the definition of the Laplace transform, writing τ for t (to have t available later). Then, taking e^{-as} inside the integral, we have

$$e^{-as}F(s) = e^{-as}\int_0^\infty e^{-s\tau}f(\tau)\,d\tau = \int_0^\infty e^{-s(\tau+a)}f(\tau)\,d\tau.$$

Substituting $\tau + a = t$, thus $\tau = t - a$, $d\tau = dt$ in the integral (**CAUTION**, the lower limit changes!), we obtain

$$e^{-as}F(s) = \int_a^\infty e^{-st}f(t-a)\,dt.$$

To make the right side into a Laplace transform, we must have an integral from 0 to ∞, not from a to ∞. But this is easy. We multiply the integrand by $u(t-a)$. Then for t from 0 to a the integrand is 0, and we can write, with $\tilde{f}$ as in (3),

$$e^{-as}F(s) = \int_0^\infty e^{-st} f(t-a)u(t-a)\,dt = \int_0^\infty e^{-st}\tilde{f}(t)\,dt.$$

(Do you now see why $u(t-a)$ appears?) This integral is the left side of (4), the Laplace transform of $\tilde{f}(t)$ in (3). This completes the proof. ■

EXAMPLE 1 **Application of Theorem 1. Use of Unit Step Functions**

Write the following function using unit step functions and find its transform.

$$f(t) = \begin{cases} 2 & \text{if } 0 < t < 1 \\ \tfrac{1}{2}t^2 & \text{if } 1 < t < \tfrac{1}{2}\pi \\ \cos t & \text{if } \quad t > \tfrac{1}{2}\pi. \end{cases} \qquad \text{(Fig. 122)}$$

Solution. ***Step 1.*** In terms of unit step functions,

$$f(t) = 2(1 - u(t-1)) + \tfrac{1}{2}t^2(u(t-1) - u(t - \tfrac{1}{2}\pi)) + (\cos t)u(t - \tfrac{1}{2}\pi).$$

Indeed, $2(1 - u(t-1))$ gives $f(t)$ for $0 < t < 1$, and so on.

Step 2. To apply Theorem 1, we must write each term in $f(t)$ in the form $f(t-a)u(t-a)$. Thus, $2(1 - u(t-1))$ remains as it is and gives the transform $2(1 - e^{-s})/s$. Then

$$\mathcal{L}\left\{\frac{1}{2}t^2 u(t-1)\right\} = \mathcal{L}\left(\frac{1}{2}(t-1)^2 + (t-1) + \frac{1}{2}\right)u(t-1)\right\} = \left(\frac{1}{s^3} + \frac{1}{s^2} + \frac{1}{2s}\right)e^{-s}$$

$$\mathcal{L}\left\{\frac{1}{2}t^2 u\left(t - \frac{1}{2}\pi\right)\right\} = \mathcal{L}\left\{\frac{1}{2}\left(t - \frac{1}{2}\pi\right)^2 + \frac{\pi}{2}\left(t - \frac{1}{2}\pi\right) + \frac{\pi^2}{8}\right)u\left(t - \frac{1}{2}\pi\right)\right\}$$

$$= \left(\frac{1}{s^3} + \frac{\pi}{2s^2} + \frac{\pi^2}{8s}\right)e^{-\pi s/2}$$

$$\mathcal{L}\left\{(\cos t)\,u\left(t - \frac{1}{2}\pi\right)\right\} = \mathcal{L}\left\{-\left(\sin\left(t - \frac{1}{2}\pi\right)\right)u\left(t - \frac{1}{2}\pi\right)\right\} = -\frac{1}{s^2+1}e^{-\pi s/2}.$$

Together,

$$\mathcal{L}(f) = \frac{2}{s} - \frac{2}{s}e^{-s} + \left(\frac{1}{s^3} + \frac{1}{s^2} + \frac{1}{2s}\right)e^{-s} - \left(\frac{1}{s^3} + \frac{\pi}{2s^2} + \frac{\pi^2}{8s}\right)e^{-\pi s/2} - \frac{1}{s^2+1}e^{-\pi s/2}.$$

If the conversion of $f(t)$ to $f(t-a)$ is inconvenient, replace it by

(4)** $$\mathcal{L}\{f(t)u(t-a)\} = e^{-as}\mathcal{L}\{f(t+a)\}.$$

(4**) follows from (4) by writing $f(t-a) = g(t)$, hence $f(t) = g(t+a)$ and then again writing f for g. Thus,

$$\mathcal{L}\left\{\frac{1}{2}t^2 u(t-1)\right\} = e^{-s}\mathcal{L}\left\{\frac{1}{2}(t+1)^2\right\} = e^{-s}\mathcal{L}\left\{\frac{1}{2}t^2 + t + \frac{1}{2}\right\} = e^{-s}\left(\frac{1}{s^3} + \frac{1}{s^2} + \frac{1}{2s}\right)$$

as before. Similarly for $\mathcal{L}\{\frac{1}{2}t^2 u(t - \frac{1}{2}\pi)\}$. Finally, by (4**),

$$\mathcal{L}\left\{\cos t\, u\left(t - \frac{1}{2}\pi\right)\right\} = e^{-\pi s/2}\mathcal{L}\left\{\cos\left(t + \frac{1}{2}\pi\right)\right\} = e^{-\pi s/2}\mathcal{L}\{-\sin t\} = -e^{-\pi s/2}\frac{1}{s^2+1}.$$ ■

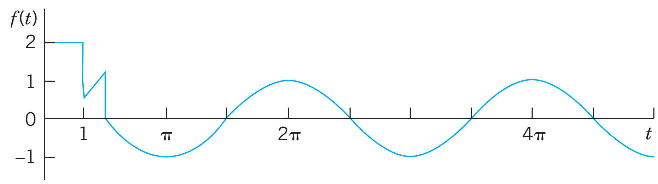


Fig. 122. $f(t)$ in Example 1

EXAMPLE 2 Application of Both Shifting Theorems. Inverse Transform

Find the inverse transform $f(t)$ of

$$F(s) = \frac{e^{-s}}{s^2 + \pi^2} + \frac{e^{-2s}}{s^2 + \pi^2} + \frac{e^{-3s}}{(s + 2)^2}.$$

Solution. Without the exponential functions in the numerator the three terms of $F(s)$ would have the inverses $(\sin \pi t)/\pi$, $(\sin \pi t)/\pi$, and te^{-2t} because $1/s^2$ has the inverse t, so that $1/(s+2)^2$ has the inverse te^{-2t} by the first shifting theorem in Sec. 6.1. Hence by the second shifting theorem (t-shifting),

$$f(t) = \frac{1}{\pi} \sin(\pi(t-1))\, u(t-1) + \frac{1}{\pi} \sin(\pi(t-2))\, u(t-2) + (t-3)e^{-2(t-3)}\, u(t-3).$$

Now $\sin(\pi t - \pi) = -\sin \pi t$ and $\sin(\pi t - 2\pi) = \sin \pi t$, so that the first and second terms cancel each other when $t > 2$. Hence we obtain $f(t) = 0$ if $0 < t < 1$, $-(\sin \pi t)/\pi$ if $1 < t < 2$, 0 if $2 < t < 3$, and $(t-3)e^{-2(t-3)}$ if $t > 3$. See Fig. 123. ■

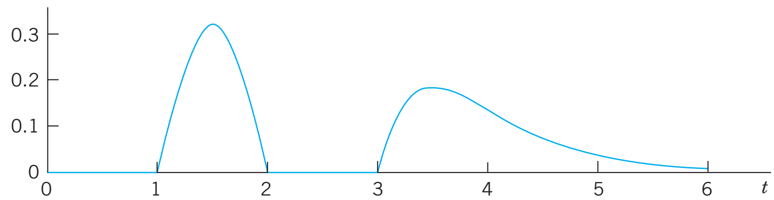


Fig. 123. $f(t)$ in Example 2

EXAMPLE 3 Response of an *RC*-Circuit to a Single Rectangular Wave

Find the current $i(t)$ in the RC-circuit in Fig. 124 if a single rectangular wave with voltage V_0 is applied. The circuit is assumed to be quiescent before the wave is applied.

Solution. The input is $V_0[u(t-a) - u(t-b)]$. Hence the circuit is modeled by the integro-differential equation (see Sec. 2.9 and Fig. 124)

$$Ri(t) + \frac{q(t)}{C} = Ri(t) + \frac{1}{C}\int_0^t i(\tau)\, d\tau = v(t) = V_0[u(t-a) - u(t-b)].$$

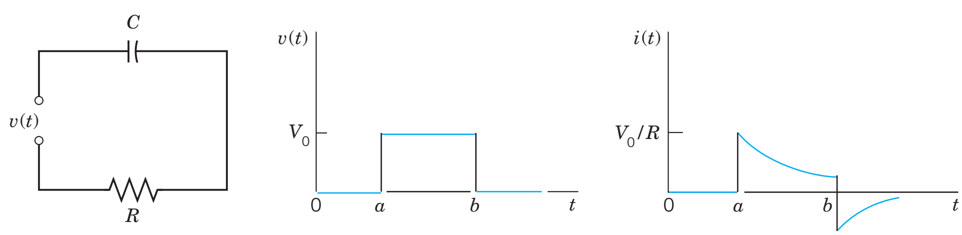


Fig. 124. RC-circuit, electromotive force $v(t)$, and current in Example 3

Using Theorem 3 in Sec. 6.2 and formula (1) in this section, we obtain the subsidiary equation

$$RI(s) + \frac{I(s)}{sC} = \frac{V_0}{s}[e^{-as} - e^{-bs}].$$

Solving this equation algebraically for $I(s)$, we get

$$I(s) = F(s)(e^{-as} - e^{-bs}) \qquad \text{where} \qquad F(s) = \frac{V_0 IR}{s + 1/(RC)} \qquad \text{and} \qquad \mathcal{L}^{-1}(F) = \frac{V_0}{R} e^{-t/(RC)},$$

the last expression being obtained from Table 6.1 in Sec. 6.1. Hence Theorem 1 yields the solution (Fig. 124)

$$i(t) = \mathcal{L}^{-1}(I) = \mathcal{L}^{-1}\{e^{-as}F(s)\} - \mathcal{L}^{-1}\{e^{-bs}F(s)\} = \frac{V_0}{R}[e^{-(t-a)/(RC)}u(t-a) - e^{-(t-b)/(RC)}u(t-b)];$$

that is, $i(t) = 0$ if $t < a$, and

$$i(t) = \begin{cases} K_1 e^{-t/(RC)} & \text{if } a < t < b \\ (K_1 - K_2)e^{-t/(RC)} & \text{if } a > b \end{cases}$$

where $K_1 = V_0 e^{a/(RC)}/R$ and $K_2 = V_0 e^{b/(RC)}/R$. ■

EXAMPLE 4 Response of an *RLC*-Circuit to a Sinusoidal Input Acting Over a Time Interval

Find the response (the current) of the *RLC*-circuit in Fig. 125, where $E(t)$ is sinusoidal, acting for a short time interval only, say,

$$E(t) = 100 \sin 400t \quad \text{if } 0 < t < 2\pi \qquad \text{and} \qquad E(t) = 0 \text{ if } t > 2\pi$$

and current and charge are initially zero.

Solution. The electromotive force $E(t)$ can be represented by $(100 \sin 400t)(1 - u(t - 2\pi))$. Hence the model for the current $i(t)$ in the circuit is the integro-differential equation (see Sec. 2.9)

$$0.1i' + 11i + 100\int_0^t i(\tau)\,d\tau = (100 \sin 400t)(1 - u(t - 2\pi)). \qquad i(0) = 0, \quad i'(0) = 0.$$

From Theorems 2 and 3 in Sec. 6.2 we obtain the subsidiary equation for $I(s) = \mathcal{L}(i)$

$$0.1sI + 11I + 100\frac{I}{s} = \frac{100 \cdot 400s}{s^2 + 400^2}\left(\frac{1}{s} - \frac{e^{-2\pi s}}{s}\right).$$

Solving it algebraically and noting that $s^2 + 110s + 1000 = (s + 10)(s + 100)$, we obtain

$$I(s) = \frac{1000 \cdot 400}{(s+10)(s+100)}\left(\frac{s}{s^2 + 400^2} - \frac{se^{-2\pi s}}{s^2 + 400^2}\right).$$

For the first term in the parentheses $(\cdots)$ times the factor in front of them we use the partial fraction expansion

$$\frac{400{,}000s}{(s+10)(s+100)(s^2 + 400^2)} = \frac{A}{s+10} + \frac{B}{s+100} + \frac{Ds + K}{s^2 + 400^2}.$$

Now determine A, B, D, K by your favorite method or by a CAS or as follows. Multiplication by the common denominator gives

$$400{,}000s = A(s+100)(s^2 + 400^2) + B(s+10)(s^2 + 400^2) + (Ds + K)(s+10)(s+100).$$

We set $s = -10$ and -100 and then equate the sums of the s^3 and s^2 terms to zero, obtaining (all values rounded)

$$
\begin{aligned}
&(s=-10) && -4{,}000{,}000 = 90(10^2 + 400^2)A, && A = -0.27760\\
&(s=-100) && -40{,}000{,}000 = -90(100^2 + 400^2)B, && B = 2.6144\\
&(s^3\text{-terms}) && 0 = A + B + D, && D = -2.3368\\
&(s^2\text{-terms}) && 0 = 100A + 10B + 110D + K, && K = 258.66.
\end{aligned}
$$

Since $K = 258.66 = 0.6467 \cdot 400$, we thus obtain for the first term I_1 in $I = I_1 - I_2$

$$I_1 = -\frac{0.2776}{s+10} + \frac{2.6144}{s+100} - \frac{2.3368s}{s^2+400^2} + \frac{0.6467 \cdot 400}{s^2+400^2}.$$

From Table 6.1 in Sec. 6.1 we see that its inverse is

$$i_1(t) = -0.2776e^{-10t} + 2.6144e^{-100t} - 2.3368 \cos 400t + 0.6467 \sin 400t.$$

This is the current $i(t)$ when $0 < t < 2\pi$. It agrees for $0 < t < 2\pi$ with that in Example 1 of Sec. 2.9 (except for notation), which concerned the same *RLC*-circuit. Its graph in Fig. 63 in Sec. 2.9 shows that the exponential terms decrease very rapidly. Note that the present amount of work was substantially less.

The second term I_1 of I differs from the first term by the factor $e^{-2\pi s}$. Since $\cos 400(t - 2\pi) = \cos 400t$ and $\sin 400(t - 2\pi) = \sin 400t$, the second shifting theorem (Theorem 1) gives the inverse $i_2(t) = 0$ if $0 < t < 2\pi$, and for $> 2\pi$ it gives

$$i_2(t) = -0.2776e^{-10(t-2\pi)} + 2.6144e^{-100(t-2\pi)} - 2.3368 \cos 400t + 0.6467 \sin 400t.$$

Hence in $i(t)$ the cosine and sine terms cancel, and the current for $t > 2\pi$ is

$$i(t) = -0.2776(e^{-10t} - e^{-10(t-2\pi)}) + 2.6144(e^{-100t} - e^{-100(t-2\pi)}).$$

It goes to zero very rapidly, practically within 0.5 sec. ■

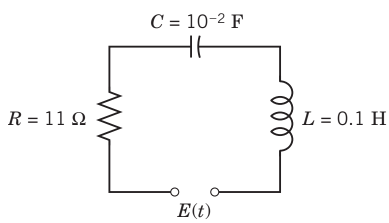


Fig. 125. *RLC*-circuit in Example 4

PROBLEM SET 6.3

1. **Report on Shifting Theorems.** Explain and compare the different roles of the two shifting theorems, using your own formulations and simple examples. Give no proofs.

2–11 SECOND SHIFTING THEOREM, UNIT STEP FUNCTION

Sketch or graph the given function, which is assumed to be zero outside the given interval. Represent it, using unit step functions. Find its transform. Show the details of your work.

2. $t\ (0 < t < 2)$ **3.** $t - 2\ (t > 2)$

4. $\cos 4t\ (0 < t < \pi)$ **5.** $e^t\ (0 < t < \pi/2)$

6. $\sin \pi t\ (2 < t < 4)$ **7.** $e^{-\pi t}\ (2 < t < 4)$

8. $t^2\ (1 < t < 2)$ **9.** $t^2\ (t > \frac{3}{2})$

10. $\sinh t\ (0 < t < 2)$ **11.** $\sin t\ (\pi/2 < t < \pi)$

12–17 INVERSE TRANSFORMS BY THE 2ND SHIFTING THEOREM

Find and sketch or graph $f(t)$ if $\mathcal{L}(f)$ equals

12. $e^{-3s}/(s-1)^3$ **13.** $6(1 - e^{-\pi s})/(s^2 + 9)$

14. $4(e^{-2s} - 2e^{-5s})/s$ **15.** e^{-3s}/s^4

16. $2(e^{-s} - e^{-3s})/(s^2 - 4)$

17. $(1 + e^{-2\pi(s+1)})(s+1)/((s+1)^2 + 1)$

18–27 IVPs, SOME WITH DISCONTINUOUS INPUT

Using the Laplace transform and showing the details, solve

18. $9y'' - 6y' + y = 0, \quad y(0) = 3, y'(0) = 1$

19. $y'' + 6y' + 8y = e^{-3t} - e^{-5t}, \quad y(0) = 0, y'(0) = 0$

20. $y'' + 10y' + 24y = 144t^2, \quad y(0) = 19/12, y'(0) = -5$

21. $y'' + 9y = 8 \sin t$ if $0 < t < \pi$ and 0 if $t > \pi$; $y(0) = 0, y'(0) = 4$

22. $y'' + 3y' + 2y = 4t$ if $0 < t < 1$ and 8 if $t > 1$; $y(0) = 0, y'(0) = 0$

23. $y'' + y' - 2y = 3 \sin t - \cos t$ if $0 < t < 2\pi$ and $3 \sin 2t - \cos 2t$ if $t > 2\pi$; $\quad y(0) = 1, y'(0) = 0$

24. $y'' + 3y' + 2y = 1$ if $0 < t < 1$ and 0 if $t > 1$; $y(0) = 0, y'(0) = 0$

25. $y'' + y = t$ if $0 < t < 1$ and 0 if $t > 1$; $\quad y(0) = 0, y'(0) = 0$

26. Shifted data. $y'' + 2y' + 5y = 10 \sin t$ if $0 < t < 2\pi$ and 0 if $t > 2\pi$; $\quad y(\pi) = 1, y'(\pi) = 2e^{-\pi} - 2$

27. Shifted data. $y'' + 4y = 8t^2$ if $0 < t < 5$ and 0 if $t > 5$; $\quad y(1) = 1 + \cos 2, y'(1) = 4 - 2 \sin 2$

28–40 MODELS OF ELECTRIC CIRCUITS

28–30 *RL*-CIRCUIT

Using the Laplace transform and showing the details, find the current $i(t)$ in the circuit in Fig. 126, assuming $i(0) = 0$ and:

28. $R = 1$ kΩ ($=1000\ \Omega$), $L = 1$ H, $v = 0$ if $0 < t < \pi$, and $40 \sin t$ V if $t > \pi$

29. $R = 25\ \Omega$, $L = 0.1$ H, $v = 490\, e^{-5t}$ V if $0 < t < 1$ and 0 if $t > 1$

30. $R = 10\ \Omega$, $L = 0.5$ H, $v = 200t$ V if $0 < t < 2$ and 0 if $t > 2$

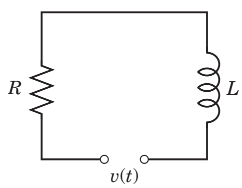


Fig. 126. Problems 28–30

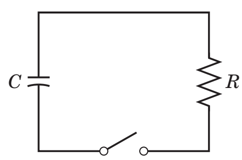


Fig. 127. Problem 31

31. Discharge in *RC*-circuit. Using the Laplace transform, find the charge $q(t)$ on the capacitor of capacitance C in Fig. 127 if the capacitor is charged so that its potential is V_0 and the switch is closed at $t = 0$.

32–34 *RC*-CIRCUIT

Using the Laplace transform and showing the details, find the current $i(t)$ in the circuit in Fig. 128 with $R = 10\ \Omega$ and $C = 10^{-2}$ F, where the current at $t = 0$ is assumed to be zero, and:

32. $v = 0$ if $t < 4$ and $14 \cdot 10^6 e^{-3t}$ V if $t > 4$

33. $v = 0$ if $t < 2$ and $100(t - 2)$ V if $t > 2$

34. $v(t) = 100$ V if $0.5 < t < 0.6$ and 0 otherwise. Why does $i(t)$ have jumps?

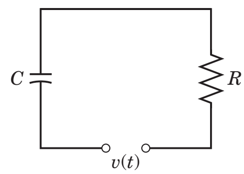


Fig. 128. Problems 32–34

35–37 *LC*-CIRCUIT

Using the Laplace transform and showing the details, find the current $i(t)$ in the circuit in Fig. 129, assuming zero initial current and charge on the capacitor and:

35. $L = 1$ H, $C = 10^{-2}$ F, $v = -9900 \cos t$ V if $\pi < t < 3\pi$ and 0 otherwise

36. $L = 1$ H, $C = 0.25$ F, $v = 200\,(t - \frac{1}{3}t^3)$ V if $0 < t < 1$ and 0 if $t > 1$

37. $L = 0.5$ H, $C = 0.05$ F, $v = 78 \sin t$ V if $0 < t < \pi$ and 0 if $t > \pi$

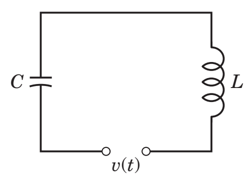


Fig. 129. Problems 35–37

38–40 *RLC*-CIRCUIT

Using the Laplace transform and showing the details, find the current $i(t)$ in the circuit in Fig. 130, assuming zero initial current and charge and:

38. $R = 4\ \Omega$, $L = 1$ H, $C = 0.05$ F, $v = 34e^{-t}$ V if $0 < t < 4$ and 0 if $t > 4$

39. $R = 2\ \Omega$, $L = 1$ H, $C = 0.5$ F, $v(t) = 1$ kV if $0 < t < 2$ and 0 if $t > 2$

40. $R = 2\ \Omega$, $L = 1$ H, $C = 0.1$ F, $v = 255 \sin t$ V if $0 < t < 2\pi$ and 0 if $t > 2\pi$

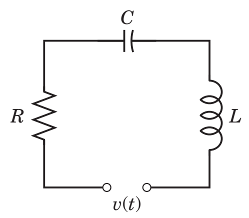


Fig. 130. Problems 38–40

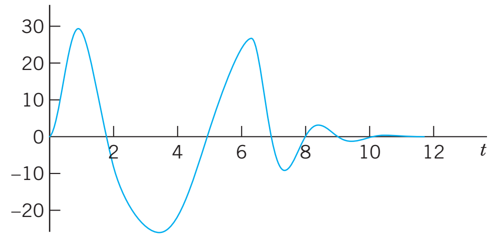


Fig. 131. Current in Problem 40

6.4 Short Impulses. Dirac's Delta Function. Partial Fractions

An airplane making a "hard" landing, a mechanical system being hit by a hammerblow, a ship being hit by a single high wave, a tennis ball being hit by a racket, and many other similar examples appear in everyday life. They are phenomena of an impulsive nature where actions of forces—mechanical, electrical, etc.—are applied over short intervals of time.

We can model such phenomena and problems by "Dirac's delta function," and solve them very effecively by the Laplace transform.

To model situations of that type, we consider the function

$$f_k(t-a) = \begin{cases} 1/k & \text{if } a \leqq t \leqq a+k \\ 0 & \text{otherwise} \end{cases} \tag{1}$$ (Fig. 132)

(and later its limit as $k \to 0$). This function represents, for instance, a force of magnitude $1/k$ acting from $t = a$ to $t = a + k$, where k is positive and small. In mechanics, the integral of a force acting over a time interval $a \leqq t \leqq a + k$ is called the **impulse** of the force; similarly for electromotive forces $E(t)$ acting on circuits. Since the blue rectangle in Fig. 132 has area 1, the impulse of f_k in (1) is

$$I_k = \int_0^\infty f_k(t-a)\,dt = \int_a^{a+k} \frac{1}{k}\,dt = 1. \tag{2}$$

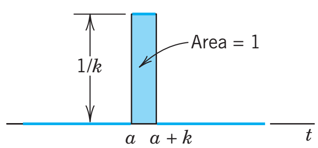


Fig. 132. The function $f_k(t-a)$ in (1)

To find out what will happen if k becomes smaller and smaller, we take the limit of f_k as $k \to 0$ $(k > 0)$. This limit is denoted by $\delta(t - a)$, that is,

$$\delta(t - a) = \lim_{k \to 0} f_k(t - a).$$

$\delta(t - a)$ is called the **Dirac delta function**[2] or the **unit impulse function**.

$\delta(t - a)$ is not a function in the ordinary sense as used in calculus, but a so-called *generalized function.*[2] To see this, we note that the impulse I_k of f_k is 1, so that from (1) and (2) by taking the limit as $k \to 0$ we obtain

$$\text{(3)} \qquad \delta(t - a) = \begin{cases} \infty & \text{if } t = a \\ 0 & \text{otherwise} \end{cases} \qquad \text{and} \qquad \int_0^\infty \delta(t - a)\, dt = 1,$$

but from calculus we know that a function which is everywhere 0 except at a single point must have the integral equal to 0. Nevertheless, in impulse problems, it is convenient to operate on $\delta(t - a)$ as though it were an ordinary function. In particular, for a *continuous* function $g(t)$ one uses the property [often called the **sifting property** of $\delta(t - a)$, not to be confused with *shifting*]

$$\textbf{(4)} \qquad \boxed{\int_0^\infty g(t)\delta(t - a)\, dt = g(a)}$$

which is plausible by (2).

To obtain the Laplace transform of $\delta(t - a)$, we write

$$f_k(t - a) = \frac{1}{k}\left[u(t - a) - u(t - (a + k))\right]$$

and take the transform [see (2)]

$$\mathcal{L}\{f_k(t - a)\} = \frac{1}{ks}\left[e^{-as} - e^{-(a+k)s}\right] = e^{-as}\,\frac{1 - e^{-ks}}{ks}.$$

We now take the limit as $k \to 0$. By l'Hôpital's rule the quotient on the right has the limit 1 (differentiate the numerator and the denominator separately with respect to k, obtaining se^{-ks} and s, respectively, and use $se^{-ks}/s \to 1$ as $k \to 0$). Hence the right side has the limit e^{-as}. This suggests defining the transform of $\delta(t - a)$ by this limit, that is,

$$\textbf{(5)} \qquad \boxed{\mathcal{L}\{\delta(t - a)\} = e^{-as}.}$$

The unit step and unit impulse functions can now be used on the right side of ODEs modeling mechanical or electrical systems, as we illustrate next.

---

[2] PAUL DIRAC (1902–1984), English physicist, was awarded the Nobel Prize [jointly with the Austrian ERWIN SCHRÖDINGER (1887–1961)] in 1933 for his work in quantum mechanics.

Generalized functions are also called **distributions**. Their theory was created in 1936 by the Russian mathematician SERGEI L'VOVICH SOBOLEV (1908–1989), and in 1945, under wider aspects, by the French mathematician LAURENT SCHWARTZ (1915–2002).

EXAMPLE 1 Mass–Spring System Under a Square Wave

Determine the response of the damped mass–spring system (see Sec. 2.8) under a square wave, modeled by (see Fig. 133)

$$y'' + 3y' + 2y = r(t) = u(t-1) - u(t-2), \qquad y(0) = 0, \qquad y'(0) = 0.$$

Solution. From (1) and (2) in Sec. 6.2 and (2) and (4) in this section we obtain the subsidiary equation

$$s^2Y + 3sY + 2Y = \frac{1}{s}(e^{-s} - e^{-2s}). \qquad \text{Solution} \qquad Y(s) = \frac{1}{s(s^2+3s+2)}(e^{-s} - e^{-2s}).$$

Using the notation $F(s)$ and partial fractions, we obtain

$$F(s) = \frac{1}{s(s^2+3s+2)} = \frac{1}{s(s+1)(s+2)} = \frac{\frac{1}{2}}{s} - \frac{1}{s+1} + \frac{\frac{1}{2}}{s+2}.$$

From Table 6.1 in Sec. 6.1, we see that the inverse is

$$f(t) = \mathcal{L}^{-1}(F) = \tfrac{1}{2} - e^{-t} + \tfrac{1}{2}e^{-2t}.$$

Therefore, by Theorem 1 in Sec. 6.3 (t-shifting) we obtain the square-wave response shown in Fig. 133,

$$\begin{aligned} y &= \mathcal{L}^{-1}(F(s)e^{-s} - F(s)e^{-2s}) \\ &= f(t-1)u(t-1) - f(t-2)u(t-2) \\ &= \begin{cases} 0 & (0 < t < 1) \\ \frac{1}{2} - e^{-(t-1)} + \frac{1}{2}e^{-2(t-1)} & (1 < t < 2) \\ -e^{-(t-1)} + e^{-(t-2)} + \frac{1}{2}e^{-2(t-1)} - \frac{1}{2}e^{-2(t-2)} & (t > 2). \end{cases} \end{aligned}$$

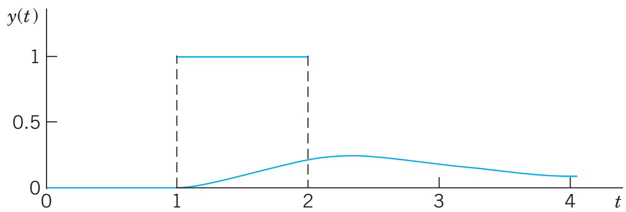


Fig. 133. Square wave and response in Example 1

EXAMPLE 2 Hammerblow Response of a Mass–Spring System

Find the response of the system in Example 1 with the square wave replaced by a unit impulse at time $t = 1$.

Solution. We now have the ODE and the subsidiary equation

$$y'' + 3y' + 2y = \delta(t-1), \qquad \text{and} \qquad (s^2 + 3s + 2)Y = e^{-s}.$$

Solving algebraically gives

$$Y(s) = \frac{e^{-s}}{(s+1)(s+2)} = \left(\frac{1}{s+1} - \frac{1}{s+2}\right)e^{-s}.$$

By Theorem 1 the inverse is

$$y(t) = \mathcal{L}^{-1}(Y) = \begin{cases} 0 & \text{if } 0 < t < 1 \\ e^{-(t-1)} - e^{-2(t-1)} & \text{if } \quad t > 1. \end{cases}$$

$y(t)$ is shown in Fig. 134. Can you imagine how Fig. 133 approaches Fig. 134 as the wave becomes shorter and shorter, the area of the rectangle remaining 1? ■

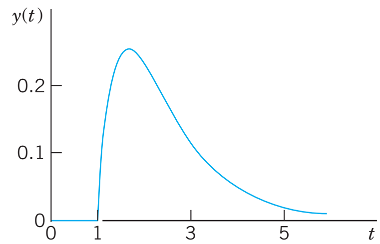


Fig. 134. Response to a hammerblow in Example 2

EXAMPLE 3 Four-Terminal *RLC*-Network

Find the output voltage response in Fig. 135 if $R = 20\ \Omega$, $L = 1$ H, $C = 10^{-4}$ F, the input is $\delta(t)$ (a unit impulse at time $t = 0$), and current and charge are zero at time $t = 0$.

Solution. To understand what is going on, note that the network is an *RLC*-circuit to which two wires at A and B are attached for recording the voltage $v(t)$ on the capacitor. Recalling from Sec. 2.9 that current $i(t)$ and charge $q(t)$ are related by $i = q' = dq/dt$, we obtain the model

$$Li' + Ri + \frac{q}{C} = Lq'' + Rq' + \frac{q}{C} = q'' + 20q' + 10{,}000q = \delta(t).$$

From (1) and (2) in Sec. 6.2 and (5) in this section we obtain the subsidiary equation for $Q(s) = \mathcal{L}(q)$

$$(s^2 + 20s + 10{,}000)Q = 1. \qquad \text{Solution} \qquad Q = \frac{1}{(s + 10)^2 + 9900}.$$

By the first shifting theorem in Sec. 6.1 we obtain from Q damped oscillations for q and v; rounding $9900 \approx 99.50^2$, we get (Fig. 135)

$$q = \mathcal{L}^{-1}(Q) = \frac{1}{99.50} e^{-10t} \sin 99.50t \qquad \text{and} \qquad v = \frac{q}{C} = 100.5 e^{-10t} \sin 99.50t. \quad ■$$

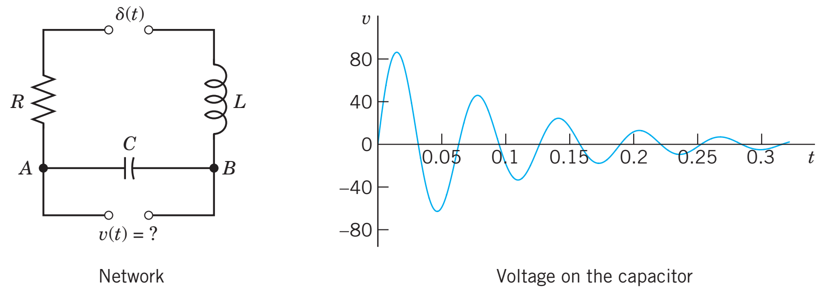


Fig. 135. Network and output voltage in Example 3

More on Partial Fractions

We have seen that the solution Y of a subsidiary equation usually appears as a quotient of polynomials $Y(s) = F(s)/G(s)$, so that a partial fraction representation leads to a sum of expressions whose inverses we can obtain from a table, aided by the first shifting theorem (Sec. 6.1). These representations are sometimes called **Heaviside expansions**.

An *unrepeated factor* $s - a$ in $G(s)$ requires a single partial fraction $A/(s - a)$. See Examples 1 and 2. *Repeated real factors* $(s - a)^2$, $(s - a)^3$, etc., require partial fractions

$$\frac{A_2}{(s-a)^2} + \frac{A_1}{s-a}, \qquad \frac{A_3}{(s-a)^3} + \frac{A_2}{(s-a)^2} + \frac{A_1}{s-a}, \quad \text{etc.,}$$

The inverses are $(A_2 t + A_1)e^{at}$, $(\frac{1}{2}A_3 t^2 + A_2 t + A_1)e^{at}$, etc.

Unrepeated complex factors $(s - a)(s - \bar{a})$, $a = \alpha + i\beta$, $\bar{a} = \alpha - i\beta$, require a partial fraction $(As + B)/[(s - \alpha)^2 + \beta^2]$. For an application, see Example 4 in Sec. 6.3. A further one is the following.

EXAMPLE 4 Unrepeated Complex Factors. Damped Forced Vibrations

Solve the initial value problem for a damped mass–spring system acted upon by a sinusoidal force for some time interval (Fig. 136),

$$y'' + 2y' + 2y = r(t), \quad r(t) = 10 \sin 2t \text{ if } 0 < t < \pi \text{ and } 0 \text{ if } t > \pi; \qquad y(0) = 1, \quad y'(0) = -5.$$

Solution. From Table 6.1, (1), (2) in Sec. 6.2, and the second shifting theorem in Sec. 6.3, we obtain the subsidiary equation

$$(s^2 Y - s + 5) + 2(sY - 1) + 2Y = 10 \frac{2}{s^2 + 4}(1 - e^{-\pi s}).$$

We collect the Y-terms, $(s^2 + 2s + 2)Y$, take $-s + 5 - 2 = -s + 3$ to the right, and solve,

$$Y = \frac{20}{(s^2 + 4)(s^2 + 2s + 2)} - \frac{20e^{-\pi s}}{(s^2 + 4)(s^2 + 2s + 2)} + \frac{s - 3}{s^2 + 2s + 2}. \tag{6}$$

For the last fraction we get from Table 6.1 and the first shifting theorem

$$\mathcal{L}^{-1}\left\{\frac{s + 1 - 4}{(s + 1)^2 + 1}\right\} = e^{-t}(\cos t - 4 \sin t). \tag{7}$$

In the first fraction in (6) we have unrepeated complex roots, hence a partial fraction representation

$$\frac{20}{(s^2 + 4)(s^2 + 2s + 2)} = \frac{As + B}{s^2 + 4} + \frac{Ms + N}{s^2 + 2s + 2}.$$

Multiplication by the common denominator gives

$$20 = (As + B)(s^2 + 2s + 2) + (Ms + N)(s^2 + 4).$$

We determine A, B, M, N. Equating the coefficients of each power of s on both sides gives the four equations

$$\begin{array}{llll} \text{(a)} \ [s^3]: & 0 = A + M & \text{(b)} \ [s^2]: & 0 = 2A + B + N \\ \text{(c)} \ [s]: & 0 = 2A + 2B + 4M & \text{(d)} \ [s^0]: & 20 = 2B + 4N. \end{array}$$

We can solve this, for instance, obtaining $M = -A$ from (a), then $A = B$ from (c), then $N = -3A$ from (b), and finally $A = -2$ from (d). Hence $A = -2$, $B = -2$, $M = 2$, $N = 6$, and the first fraction in (6) has the representation

$$\frac{-2s - 2}{s^2 + 4} + \frac{2(s + 1) + 6 - 2}{(s + 1)^2 + 1}. \quad \text{Inverse transform:} \quad -2\cos 2t - \sin 2t + e^{-t}(2\cos t + 4 \sin t). \tag{8}$$

The sum of this inverse and (7) is the solution of the problem for $0 < t < \pi$, namely (the sines cancel),

$$y(t) = 3e^{-t}\cos t - 2\cos 2t - \sin 2t \qquad \text{if } 0 < t < \pi. \tag{9}$$

In the second fraction in (6), taken with the minus sign, we have the factor $e^{-\pi s}$, so that from (8) and the second shifting theorem (Sec. 6.3) we get the inverse transform of this fraction for $t > 0$ in the form

$$\begin{aligned} &+2\cos(2t - 2\pi) + \sin(2t - 2\pi) - e^{-(t-\pi)}[2\cos(t-\pi) + 4\sin(t-\pi)] \\ &= 2\cos 2t + \sin 2t + e^{-(t-\pi)}(2\cos t + 4\sin t). \end{aligned}$$

The sum of this and (9) is the solution for $t > \pi$,

$$y(t) = e^{-t}[(3 + 2e^{\pi})\cos t + 4e^{\pi}\sin t] \qquad \text{if } t > \pi. \tag{10}$$

Figure 136 shows (9) (for $0 < t < \pi$) and (10) (for $t > \pi$), a beginning vibration, which goes to zero rapidly because of the damping and the absence of a driving force after $t = \pi$. ■

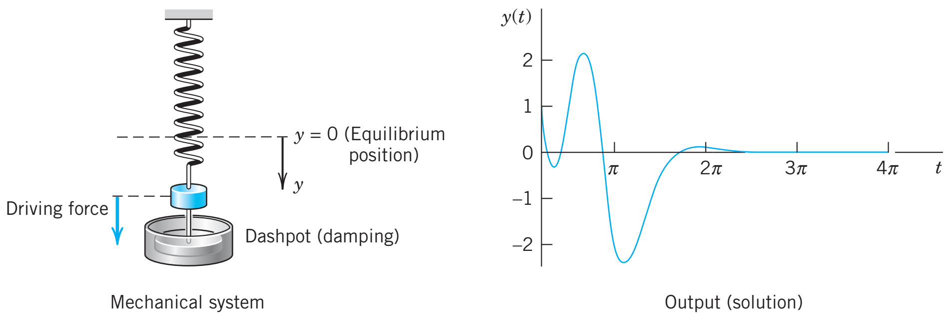


Fig. 136. Example 4

The case of repeated complex factors $[(s - a)(s - \bar{a})]^2$, which is important in connection with resonance, will be handled by "convolution" in the next section.

PROBLEM SET 6.4

1. **CAS PROJECT. Effect of Damping.** Consider a vibrating system of your choice modeled by

$$y'' + cy' + ky = \delta(t).$$

(a) Using graphs of the solution, describe the effect of continuously decreasing the damping to 0, keeping k constant.

(b) What happens if c is kept constant and k is continuously increased, starting from 0?

(c) Extend your results to a system with two δ-functions on the right, acting at different times.

2. **CAS EXPERIMENT. Limit of a Rectangular Wave. Effects of Impulse.**

(a) In Example 1 in the text, take a rectangular wave of area 1 from 1 to $1 + k$. Graph the responses for a sequence of values of k approaching zero, illustrating that for smaller and smaller k those curves approach the curve shown in Fig. 134. *Hint:* If your CAS gives no solution for the differential equation, involving k, take specific k's from the beginning.

(b) Experiment on the response of the ODE in Example 1 (or of another ODE of your choice) to an impulse $\delta(t - a)$ for various systematically chosen a (> 0); choose initial conditions $y(0) \neq 0$, $y'(0) = 0$. Also consider the solution if no impulse is applied. Is there a dependence of the response on a? On b if you choose $b\delta(t - a)$? Would $-\delta(t - \tilde{a})$ with $\tilde{a} > a$ annihilate the effect of $\delta(t - a)$? Can you think of other questions that one could consider experimentally by inspecting graphs?

3–12 EFFECT OF DELTA (IMPULSE) ON VIBRATING SYSTEMS

Find and graph or sketch the solution of the IVP. Show the details.

3. $y'' + 4y = \delta(t - \pi), \quad y(0) = 8, y'(0) = 0$

4. $y'' + 16y = 4\delta(t - 3\pi), \quad y(0) = 2, y'(0) = 0$

5. $y'' + y = \delta(t - \pi) - \delta(t - 2\pi),$
$y(0) = 0, y'(0) = 1$

6. $y'' + 4y' + 5y = \delta(t - 1), \quad y(0) = 0, y'(0) = 3$

7. $4y'' + 24y' + 37y = 17e^{-t} + \delta(t - \frac{1}{2}),$
$y(0) = 1, y'(0) = 1$

8. $y'' + 3y' + 2y = 10(\sin t + \delta(t - 1)), \quad y(0) = 1,$
$y'(0) = -1$

9. $y'' + 4y' + 5y = [1 - u(t - 10)]e^t - e^{10}\delta(t - 10),$
$y(0) = 0, y'(0) = 1$

10. $y'' + 5y' + 6y = \delta(t - \frac{1}{2}\pi) + u(t - \pi)\cos t,$
$y(0) = 0, y'(0) = 0$

11. $y'' + 5y' + 6y = u(t - 1) + \delta(t - 2),$
$y(0) = 0, y'(0) = 1$

12. $y'' + 2y' + 5y = 25t - 100\delta(t - \pi), \quad y(0) = -2,$
$y'(0) = 5$

13. PROJECT. Heaviside Formulas. (a) Show that for a simple root a and fraction $A/(s - a)$ in $F(s)/G(s)$ we have the *Heaviside formula*

$$A = \lim_{s \to a} \frac{(s - a)F(s)}{G(s)}.$$

(b) Similarly, show that for a root a of order m and fractions in

$$\frac{F(s)}{G(s)} = \frac{A_m}{(s - a)^m} + \frac{A_{m-1}}{(s - a)^{m-1}} + \cdots$$
$$+ \frac{A_1}{s - a} + \text{further fractions}$$

we have the *Heaviside formulas* for the first coefficient

$$A_m = \lim_{s \to a} \frac{(s - a)^m F(s)}{G(s)}$$

and for the other coefficients

$$A_k = \frac{1}{(m - k)!} \lim_{s \to a} \frac{d^{m-k}}{ds^{m-k}} \left[\frac{(s - a)^m F(s)}{G(s)}\right],$$
$$k = 1, \cdots, m - 1.$$

14. TEAM PROJECT. Laplace Transform of Periodic Functions

(a) Theorem. *The Laplace transform of a piecewise continuous function $f(t)$ with period p is*

$$\mathscr{L}(f) = \frac{1}{1 - e^{-ps}} \int_0^p e^{-st} f(t)\, dt \qquad (s > 0). \tag{11}$$

Prove this theorem. *Hint:* Write $\int_0^\infty = \int_0^p + \int_p^{2p} + \cdots$. Set $t = (n - 1)p$ in the nth integral. Take out $e^{-(n-1)p}$ from under the integral sign. Use the sum formula for the geometric series.

(b) Half-wave rectifier. Using (11), show that the half-wave rectification of $\sin \omega t$ in Fig. 137 has the Laplace transform

$$\mathscr{L}(f) = \frac{\omega(1 + e^{-\pi s/\omega})}{(s^2 + \omega^2)(1 - e^{-2\pi s/\omega})}$$
$$= \frac{\omega}{(s^2 + \omega^2)(1 - e^{-\pi s/\omega})}.$$

(A *half-wave rectifier* clips the negative portions of the curve. A *full-wave rectifier* converts them to positive; see Fig. 138.)

(c) Full-wave rectifier. Show that the Laplace transform of the full-wave rectification of $\sin \omega t$ is

$$\frac{\omega}{s^2 + \omega^2} \coth \frac{\pi s}{2\omega}.$$

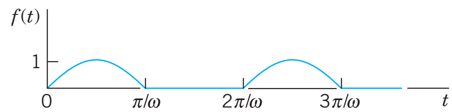


Fig. 137. Half-wave rectification

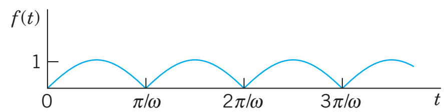


Fig. 138. Full-wave rectification

(d) Saw-tooth wave. Find the Laplace transform of the saw-tooth wave in Fig. 139.

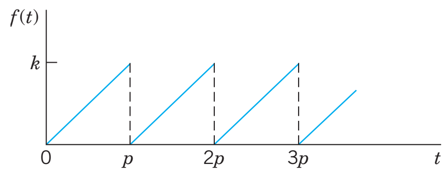


Fig. 139. Saw-tooth wave

15. Staircase function. Find the Laplace transform of the staircase function in Fig. 140 by noting that it is the difference of kt/p and the function in 14(d).

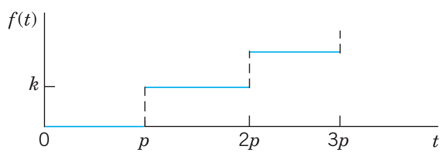


Fig. 140. Staircase function

6.5 Convolution. Integral Equations

Convolution has to do with the multiplication of transforms. The situation is as follows. *Addition* of transforms provides no problem; we know that $\mathcal{L}(f+g) = \mathcal{L}(f) + \mathcal{L}(g)$. Now **multiplication of transforms** occurs frequently in connection with ODEs, integral equations, and elsewhere. Then we usually know $\mathcal{L}(f)$ and $\mathcal{L}(g)$ and would like to know the function whose transform is the product $\mathcal{L}(f)\mathcal{L}(g)$. We might perhaps guess that it is fg, but this is false. *The transform of a product is generally different from the product of the transforms of the factors,*

$$\mathcal{L}(fg) \neq \mathcal{L}(f)\mathcal{L}(g) \qquad \text{in general.}$$

To see this take $f = e^t$ and $g = 1$. Then $fg = e^t$, $\mathcal{L}(fg) = 1/(s-1)$, but $\mathcal{L}(f) = 1/(s-1)$ and $\mathcal{L}(1) = 1/s$ give $\mathcal{L}(f)\mathcal{L}(g) = 1/(s^2 - s)$.

According to the next theorem, the correct answer is that $\mathcal{L}(f)\mathcal{L}(g)$ is the transform of the **convolution** of f and g, denoted by the standard notation $f * g$ and defined by the integral

$$h(t) = (f * g)(t) = \int_0^t f(\tau)g(t-\tau)\,d\tau. \tag{1}$$

THEOREM 1

Convolution Theorem

If two functions f and g satisfy the assumption in the existence theorem in Sec. 6.1, *so that their transforms F and G exist, the product $H = FG$ is the transform of h given by* (1). (Proof after Example 2.)

EXAMPLE 1 **Convolution**

Let $H(s) = 1/[(s-a)s]$. Find $h(t)$.

Solution. $1/(s-a)$ has the inverse $f(t) = e^{at}$, and $1/s$ has the inverse $g(t) = 1$. With $f(\tau) = e^{a\tau}$ and $g(t-\tau) \equiv 1$ we thus obtain from (1) the answer

$$h(t) = e^{at} * 1 = \int_0^t e^{a\tau} \cdot 1\,d\tau = \frac{1}{a}(e^{at} - 1).$$

To check, calculate

$$H(s) = \mathcal{L}(h)(s) = \frac{1}{a}\left(\frac{1}{s-a} - \frac{1}{s}\right) = \frac{1}{a} \cdot \frac{a}{s^2 - as} = \frac{1}{s-a} \cdot \frac{1}{s} = \mathcal{L}(e^{at})\mathcal{L}(1). \qquad \blacksquare$$

EXAMPLE 2 **Convolution**

Let $H(s) = 1/(s^2 + \omega^2)^2$. Find $h(t)$.

Solution. The inverse of $1/(s^2 + \omega^2)$ is $(\sin \omega t)/\omega$. Hence from (1) and the first formula in (11) in App. 3.1 we obtain

$$\begin{aligned} h(t) = \frac{\sin \omega t}{\omega} * \frac{\sin \omega t}{\omega} &= \frac{1}{\omega^2}\int_0^t \sin \omega\tau \sin \omega(t-\tau)\,d\tau \\ &= \frac{1}{2\omega^2}\int_0^t [-\cos \omega t + \cos(2\omega\tau - \omega t)]\,d\tau \end{aligned}$$

$$= \frac{1}{2\omega^2}\left[-\tau\cos\omega t + \frac{\sin\omega\tau}{\omega}\right]_{\tau=0}^{t}$$

$$= \frac{1}{2\omega^2}\left[-t\cos\omega t + \frac{\sin\omega t}{\omega}\right]$$

in agreement with formula 21 in the table in Sec. 6.9. ■

PROOF We prove the Convolution Theorem 1. CAUTION! Note which ones are the variables of integration! We can denote them as we want, for instance, by τ and p, and write

$$F(s) = \int_0^\infty e^{-s\tau} f(\tau)\, d\tau \qquad \text{and} \qquad G(s) = \int_0^\infty e^{-sp} g(p)\, dp.$$

We now set $t = p + \tau$, where τ is at first constant. Then $p = t - \tau$, and t varies from τ to ∞. Thus

$$G(s) = \int_\tau^\infty e^{-s(t-\tau)} g(t-\tau)\, dt = e^{s\tau}\int_\tau^\infty e^{-st} g(t-\tau)\, dt.$$

τ in F and t in G vary independently. Hence we can insert the G-integral into the F-integral. Cancellation of $e^{-s\tau}$ and $e^{s\tau}$ then gives

$$F(s)G(s) = \int_0^\infty e^{-s\tau} f(\tau) e^{s\tau} \int_\tau^\infty e^{-st} g(t-\tau)\, dt\, d\tau = \int_0^\infty f(\tau)\int_\tau^\infty e^{-st} g(t-\tau)\, dt\, d\tau.$$

Here we integrate for fixed τ over t from τ to ∞ and then over τ from 0 to ∞. This is the blue region in Fig. 141. Under the assumption on f and g the order of integration can be reversed (see Ref. [A5] for a proof using uniform convergence). We then integrate first over τ from 0 to t and then over t from 0 to ∞, that is,

$$F(s)G(s) = \int_0^\infty e^{-st}\int_0^t f(\tau)g(t-\tau)\, d\tau\, dt = \int_0^\infty e^{-st} h(t)\, dt = \mathcal{L}(h) = H(s).$$

This completes the proof. ■

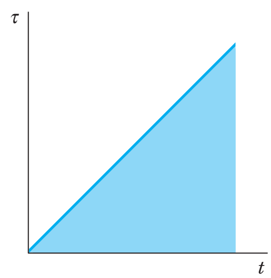


Fig. 141. Region of integration in the $t\tau$-plane in the proof of Theorem 1

From the definition it follows almost immediately that convolution has the properties

$$f * g = g * f \qquad \text{(commutative law)}$$

$$f * (g_1 + g_2) = f * g_1 + f * g_2 \qquad \text{(distributive law)}$$

$$(f * g) * v = f * (g * v) \qquad \text{(associative law)}$$

$$f * 0 = 0 * f = 0$$

similar to those of the multiplication of numbers. However, there are differences of which you should be aware.

EXAMPLE 3 Unusual Properties of Convolution

$f * 1 \neq f$ in general. For instance,

$$t * 1 = \int_0^t \tau \cdot 1 \, d\tau = \frac{1}{2} t^2 \neq t.$$

$(f * f)(t) \geqq 0$ may not hold. For instance, Example 2 with $\omega = 1$ gives

$$\sin t * \sin t = -\tfrac{1}{2} t \cos t + \tfrac{1}{2} \sin t \qquad \text{(Fig. 142).}$$

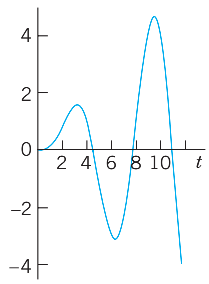


Fig. 142. Example 3

We shall now take up the case of a complex double root (left aside in the last section in connection with partial fractions) and find the solution (the inverse transform) directly by convolution.

EXAMPLE 4 Repeated Complex Factors. Resonance

In an undamped mass–spring system, resonance occurs if the frequency of the driving force equals the natural frequency of the system. Then the model is (see Sec. 2.8)

$$y'' + \omega_0^2 y = K \sin \omega_0 t$$

where $\omega_0^2 = k/m$, k is the spring constant, and m is the mass of the body attached to the spring. We assume $y(0) = 0$ and $y'(0) = 0$, for simplicity. Then the subsidiary equation is

$$s^2 Y + \omega_0^2 Y = \frac{K\omega_0}{s^2 + \omega_0^2}. \qquad \text{Its solution is} \qquad Y = \frac{K\omega_0}{(s^2 + \omega_0^2)^2}.$$

This is a transform as in Example 2 with $\omega = \omega_0$ and multiplied by $K\omega_0$. Hence from Example 2 we can see directly that the solution of our problem is

$$y(t) = \frac{K\omega_0}{2\omega_0^2}\left(-t\cos\omega_0 t + \frac{\sin\omega_0 t}{\omega_0}\right) = \frac{K}{2\omega_0^2}(-\omega_0 t\cos\omega_0 t + \sin\omega_0 t).$$

We see that the first term grows without bound. Clearly, in the case of resonance such a term must occur. (See also a similar kind of solution in Fig. 55 in Sec. 2.8.) ■

Application to Nonhomogeneous Linear ODEs

Nonhomogeneous linear ODEs can now be solved by a general method based on convolution by which the solution is obtained in the form of an integral. To see this, recall from Sec. 6.2 that the subsidiary equation of the ODE

$$y'' + ay' + by = r(t) \qquad (a, b \text{ constant}) \tag{2}$$

has the solution [(7) in Sec. 6.2]

$$Y(s) = [(s + a)y(0) + y'(0)]Q(s) + R(s)Q(s)$$

with $R(s) = \mathcal{L}(r)$ and $Q(s) = 1/(s^2 + as + b)$ the transfer function. Inversion of the first term $[\cdots]$ provides no difficulty; depending on whether $\frac{1}{4}a^2 - b$ is positive, zero, or negative, its inverse will be a linear combination of two exponential functions, or of the form $(c_1 + c_2 t)e^{-at/2}$, or a damped oscillation, respectively. The interesting term is $R(s)Q(s)$ because $r(t)$ can have various forms of practical importance, as we shall see. If $y(0) = 0$ and $y'(0) = 0$, then $Y = RQ$, and the convolution theorem gives the solution

$$\boxed{y(t) = \int_0^t q(t - \tau)r(\tau)\,d\tau.} \tag{3}$$

EXAMPLE 5 Response of a Damped Vibrating System to a Single Square Wave

Using convolution, determine the response of the damped mass–spring system modeled by

$$y'' + 3y' + 2y = r(t), \qquad r(t) = 1 \text{ if } 1 < t < 2 \text{ and } 0 \text{ otherwise}, \qquad y(0) = y'(0) = 0.$$

This system with an **input** (a driving force) ***that acts for some time only*** (Fig. 143) has been solved by partial fraction reduction in Sec. 6.4 (Example 1).

Solution by Convolution. The transfer function and its inverse are

$$Q(s) = \frac{1}{s^2 + 3s + 2} = \frac{1}{(s + 1)(s + 2)} = \frac{1}{s + 1} - \frac{1}{s + 2}, \qquad \text{hence} \qquad q(t) = e^{-t} - e^{-2t}.$$

Hence the convolution integral (3) is (except for the limits of integration)

$$y(t) = \int q(t - \tau)\cdot 1\,d\tau = \int \left[e^{-(t-\tau)} - e^{-2(t-\tau)}\right] d\tau = e^{-(t-\tau)} - \frac{1}{2}e^{-2(t-\tau)}.$$

Now comes an important point in handling convolution. $r(\tau) = 1$ if $1 < \tau < 2$ only. Hence if $t < 1$, the integral is zero. If $1 < t < 2$, we have to integrate from $\tau = 1$ (not 0) to t. This gives (with the first two terms from the upper limit)

$$y(t) = e^{-0} - \tfrac{1}{2}e^{-0} - (e^{-(t-1)} - \tfrac{1}{2}e^{-2(t-1)}) = \tfrac{1}{2} - e^{-(t-1)} + \tfrac{1}{2}e^{-2(t-1)}.$$

If $t > 2$, we have to integrate from $\tau = 1$ to 2 (not to t). This gives

$$y(t) = e^{-(t-2)} - \tfrac{1}{2}e^{-2(t-2)} - (e^{-(t-1)} - \tfrac{1}{2}e^{-2(t-1)}).$$

Figure 143 shows the input (the square wave) and the interesting output, which is zero from 0 to 1, then increases, reaches a maximum (near 2.6) after the input has become zero (why?), and finally decreases to zero in a monotone fashion. ■

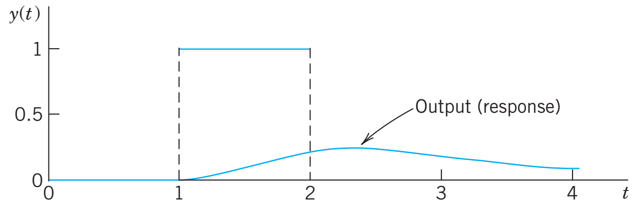


Fig. 143. Square wave and response in Example 5

Integral Equations

Convolution also helps in solving certain **integral equations**, that is, equations in which the unknown function $y(t)$ appears in an integral (and perhaps also outside of it). This concerns equations with an integral of the form of a convolution. Hence these are special and it suffices to explain the idea in terms of two examples and add a few problems in the problem set.

EXAMPLE 6 A Volterra Integral Equation of the Second Kind

Solve the Volterra integral equation of the second kind[3]

$$y(t) - \int_0^t y(\tau)\sin(t - \tau)\, d\tau = t.$$

Solution. From (1) we see that the given equation can be written as a convolution, $y - y * \sin t = t$. Writing $Y = \mathcal{L}(y)$ and applying the convolution theorem, we obtain

$$Y(s) - Y(s)\frac{1}{s^2 + 1} = Y(s)\frac{s^2}{s^2 + 1} = \frac{1}{s^2}.$$

The solution is

$$Y(s) = \frac{s^2 + 1}{s^4} = \frac{1}{s^2} + \frac{1}{s^4} \qquad \text{and gives the answer} \qquad y(t) = t + \frac{t^3}{6}.$$

Check the result by a CAS or by substitution and repeated integration by parts (which will need patience). ■

EXAMPLE 7 Another Volterra Integral Equation of the Second Kind

Solve the Volterra integral equation

$$y(t) - \int_0^t (1 + \tau)\, y(t - \tau)\, d\tau = 1 - \sinh t.$$

[3]If the upper limit of integration is *variable*, the equation is named after the Italian mathematician VITO VOLTERRA (1860–1940), and if that limit is *constant*, the equation is named after the Swedish mathematician ERIK IVAR FREDHOLM (1866–1927). "Of the second kind (first kind)" indicates that y occurs (does not occur) outside of the integral.

Solution. By (1) we can write $y - (1 + t) * y = 1 - \sinh t$. Writing $Y = \mathcal{L}(y)$, we obtain by using the convolution theorem and then taking common denominators

$$Y(s)\left[1 - \left(\frac{1}{s} + \frac{1}{s^2}\right)\right] = \frac{1}{s} - \frac{1}{s^2 - 1}, \qquad \text{hence} \qquad Y(s) \cdot \frac{s^2 - s - 1}{s^2} = \frac{s^2 - 1 - s}{s(s^2 - 1)}.$$

$(s^2 - s - 1)/s$ cancels on both sides, so that solving for Y simply gives

$$Y(s) = \frac{s}{s^2 - 1} \qquad \text{and the solution is} \qquad y(t) = \cosh t.$$

PROBLEM SET 6.5

1–7 CONVOLUTIONS BY INTEGRATION

Find:

1. $1 * 1$

2. $1 * \sin \omega t$

3. $e^t * e^{-t}$

4. $(\cos \omega t) * (\cos \omega t)$

5. $(\sin \omega t) * (\cos \omega t)$

6. $e^{at} * e^{bt} \; (a \neq b)$

7. $t * e^t$

8–14 INTEGRAL EQUATIONS

Solve by the Laplace transform, showing the details:

8. $y(t) + 4\int_0^t y(\tau)(t - \tau)\, d\tau = 2t$

9. $y(t) - \int_0^t y(\tau)\, d\tau = 1$

10. $y(t) - \int_0^t y(\tau) \sin 2(t - \tau)\, d\tau = \sin 2t$

11. $y(t) + \int_0^t (t - \tau) y(\tau)\, d\tau = 1$

12. $y(t) + \int_0^t y(\tau) \cosh (t - \tau)\, d\tau = t + e^t$

13. $y(t) + 2e^t \int_0^t y(\tau) e^{-\tau}\, d\tau = te^t$

14. $y(t) - \int_0^t y(\tau)(t - \tau)\, d\tau = 2 - \frac{1}{2}t^2$

15. CAS EXPERIMENT. Variation of a Parameter. **(a)** Replace 2 in Prob. 13 by a parameter k and investigate graphically how the solution curve changes if you vary k, in particular near $k = -2$.

(b) Make similar experiments with an integral equation of your choice whose solution is oscillating.

16. TEAM PROJECT. Properties of Convolution. Prove:

(a) Commutativity, $f * g = g * f$

(b) Associativity, $(f * g) * v = f * (g * v)$

(c) Distributivity, $f * (g_1 + g_2) = f * g_1 + f * g_2$

(d) Dirac's delta. Derive the sifting formula (4) in Sec. 6.4 by using f_k with $a = 0$ [(1), Sec. 6.4] and applying the mean value theorem for integrals.

(e) Unspecified driving force. Show that forced vibrations governed by

$$y'' + \omega^2 y = r(t), \quad y(0) = K_1, \quad y'(0) = K_2$$

with $\omega \neq 0$ and an unspecified driving force $r(t)$ can be written in convolution form,

$$y = \frac{1}{\omega} \sin \omega t * r(t) + K_1 \cos \omega t + \frac{K_2}{\omega} \sin \omega t.$$

17–26 INVERSE TRANSFORMS BY CONVOLUTION

Showing details, find $f(t)$ if $\mathcal{L}(f)$ equals:

17. $\dfrac{5.5}{(s + 1.5)(s - 4)}$

18. $\dfrac{1}{(s - a)^2}$

19. $\dfrac{2\pi s}{(s^2 + \pi^2)^2}$

20. $\dfrac{9}{s(s + 3)}$

21. $\dfrac{\omega}{s^2(s^2 + \omega^2)}$

22. $\dfrac{e^{-as}}{s(s - 2)}$

23. $\dfrac{40.5}{s(s^2 - 9)}$

24. $\dfrac{240}{(s^2 + 1)(s^2 + 25)}$

25. $\dfrac{18s}{(s^2 + 36)^2}$

26. Partial Fractions. Solve Probs. 17, 21, and 23 by partial fraction reduction.

6.6 Differentiation and Integration of Transforms. ODEs with Variable Coefficients

The variety of methods for obtaining transforms and inverse transforms and their application in solving ODEs is surprisingly large. We have seen that they include direct integration, the use of linearity (Sec. 6.1), shifting (Secs. 6.1, 6.3), convolution (Sec. 6.5), and differentiation and integration of functions $f(t)$ (Sec. 6.2). In this section, we shall consider operations of somewhat lesser importance. They are the differentiation and integration of transforms $F(s)$ and corresponding operations for functions $f(t)$. We show how they are applied to ODEs with variable coefficients.

Differentiation of Transforms

It can be shown that, if a function $f(t)$ satisfies the conditions of the existence theorem in Sec. 6.1, then the derivative $F'(s) = dF/ds$ of the transform $F(s) = \mathcal{L}(f)$ can be obtained by differentiating $F(s)$ under the integral sign with respect to s (proof in Ref. [GenRef4] listed in App. 1). Thus, if

$$F(s) = \int_0^\infty e^{-st} f(t)\, dt, \qquad \text{then} \qquad F'(s) = -\int_0^\infty e^{-st} t f(t)\, dt.$$

Consequently, if $\mathcal{L}(f) = F(s)$, then

(1) $$\mathcal{L}\{tf(t)\} = -F'(s), \qquad \text{hence} \qquad \mathcal{L}^{-1}\{F'(s)\} = -tf(t)$$

where the second formula is obtained by applying $\mathcal{L}^{-1}$ on both sides of the first formula. In this way, *differentiation of the transform of a function corresponds to the multiplication of the function by* $-t$.

EXAMPLE 1 Differentiation of Transforms. Formulas 21–23 in Sec. 6.9

We shall derive the following three formulas.

	$\mathcal{L}(f)$	$f(t)$
(2)	$\dfrac{1}{(s^2+\beta^2)^2}$	$\dfrac{1}{2\beta^3}(\sin\beta t - \beta t\cos\beta t)$
(3)	$\dfrac{s}{(s^2+\beta^2)^2}$	$\dfrac{1}{2\beta}\sin\beta t$
(4)	$\dfrac{s^2}{(s^2+\beta^2)^2}$	$\dfrac{1}{2\beta}(\sin\beta t + \beta t\cos\beta t)$

Solution. From (1) and formula 8 (with $\omega = \beta$) in Table 6.1 of Sec. 6.1 we obtain by differentiation (CAUTION! Chain rule!)

$$\mathcal{L}(t\sin\beta t) = \frac{2\beta s}{(s^2+\beta^2)^2}.$$

Dividing by 2β and using the linearity of $\mathcal{L}$, we obtain (3).

Formulas (2) and (4) are obtained as follows. From (1) and formula 7 (with $\omega = \beta$) in Table 6.1 we find

$$\mathcal{L}(t \cos \beta t) = -\frac{(s^2+\beta^2) - 2s^2}{(s^2+\beta^2)^2} = \frac{s^2-\beta^2}{(s^2+\beta^2)^2}. \tag{5}$$

From this and formula 8 (with $\omega = \beta$) in Table 6.1 we have

$$\mathcal{L}\left(t \cos \beta t \pm \frac{1}{\beta} \sin \beta t\right) = \frac{s^2-\beta^2}{(s^2+\beta^2)^2} \pm \frac{1}{s^2+\beta^2}.$$

On the right we now take the common denominator. Then we see that for the plus sign the numerator becomes $s^2 - \beta^2 + s^2 + \beta^2 = 2s^2$, so that (4) follows by division by 2. Similarly, for the minus sign the numerator takes the form $s^2 - \beta^2 - s^2 - \beta^2 = -2\beta^2$, and we obtain (2). This agrees with Example 2 in Sec. 6.5. ■

Integration of Transforms

Similarly, if $f(t)$ satisfies the conditions of the existence theorem in Sec. 6.1 and the limit of $f(t)/t$, as t approaches 0 from the right, exists, then for $s > k$,

(6) $$\boxed{\mathcal{L}\left\{\frac{f(t)}{t}\right\} = \int_s^\infty F(\tilde{s})\, d\tilde{s}} \qquad \text{hence} \qquad \boxed{\mathcal{L}^{-1}\left\{\int_s^\infty F(\tilde{s})\, d\tilde{s}\right\} = \frac{f(t)}{t}}.$$

In this way, *integration of the transform of a function $f(t)$ corresponds to the division of $f(t)$ by t.*

We indicate how (6) is obtained. From the definition it follows that

$$\int_s^\infty F(\tilde{s})\, d\tilde{s} = \int_s^\infty \left[\int_0^\infty e^{-\tilde{s}t} f(t)\, dt\right] d\tilde{s},$$

and it can be shown (see Ref. [GenRef4] in App. 1) that under the above assumptions we may reverse the order of integration, that is,

$$\int_s^\infty F(\tilde{s})\, d\tilde{s} = \int_0^\infty \left[\int_s^\infty e^{-\tilde{s}t} f(t)\, d\tilde{s}\right] dt = \int_0^\infty f(t) \left[\int_s^\infty e^{-\tilde{s}t}\, d\tilde{s}\right] dt.$$

Integration of $e^{-\tilde{s}t}$ with respect to $\tilde{s}$ gives $e^{-\tilde{s}t}/(-t)$. Here the integral over $\tilde{s}$ on the right equals e^{-st}/t. Therefore,

$$\int_s^\infty F(\tilde{s})\, d\tilde{s} = \int_0^\infty e^{-st} \frac{f(t)}{t}\, dt = \mathcal{L}\left\{\frac{f(t)}{t}\right\} \qquad (s > k). \qquad ■$$

EXAMPLE 2 Differentiation and Integration of Transforms

Find the inverse transform of $\ln\left(1 + \frac{\omega^2}{s^2}\right) = \ln \frac{s^2+\omega^2}{s^2}$.

Solution. Denote the given transform by $F(s)$. Its derivative is

$$F'(s) = \frac{d}{ds}(\ln(s^2+\omega^2) - \ln s^2) = \frac{2s}{s^2+\omega^2} - \frac{2s}{s^2}.$$

Taking the inverse transform and using (1), we obtain

$$\mathcal{L}^{-}\{F'(s)\} = \mathcal{L}^{-1}\left\{\frac{2s}{s^2+\omega^2} - \frac{2}{s}\right\} = 2\cos\omega t - 2 = -tf(t).$$

Hence the inverse $f(t)$ of $F(s)$ is $f(t) = 2(1 - \cos\omega t)/t$. This agrees with formula 42 in Sec. 6.9.

Alternatively, if we let

$$G(s) = \frac{2s}{s^2+\omega^2} - \frac{2}{s}, \qquad \text{then} \qquad g(t) = \mathcal{L}^{-1}(G) - 2(\cos\omega t - 1).$$

From this and (6) we get, in agreement with the answer just obtained,

$$\mathcal{L}^{-1}\left\{\ln\frac{s^2+\omega^2}{s^2}\right\} = \mathcal{L}^{-1}\left\{\int_s^\infty G(s)\,ds\right\} = -\frac{g(t)}{t} = \frac{2}{t}(1 - \cos\omega t),$$

the minus occurring since s is the lower limit of integration.

In a similar way we obtain formula 43 in Sec. 6.9,

$$\mathcal{L}^{-1}\left\{\ln\left(1 - \frac{a^2}{s^2}\right)\right\} = \frac{2}{t}(1 - \cosh at).$$

Special Linear ODEs with Variable Coefficients

Formula (1) can be used to solve certain ODEs with variable coefficients. The idea is this. Let $\mathcal{L}(y) = Y$. Then $\mathcal{L}(y') = sY - y(0)$ (see Sec. 6.2). Hence by (1),

$$\mathcal{L}(ty') = -\frac{d}{ds}[sY - y(0)] = -Y - s\frac{dY}{ds}. \tag{7}$$

Similarly, $\mathcal{L}(y'') = s^2Y - sy(0) - y'(0)$ and by (1)

$$\mathcal{L}(ty'') = -\frac{d}{ds}[s^2Y - sy(0) - y'(0)] = -2sY - s^2\frac{dY}{ds} + y(0). \tag{8}$$

Hence if an ODE has coefficients such as $at + b$, the subsidiary equation is a first-order ODE for Y, which is sometimes simpler than the given second-order ODE. But if the latter has coefficients $at^2 + bt + c$, then two applications of (1) would give a second-order ODE for Y, and this shows that the present method works well only for rather special ODEs with variable coefficients. An important ODE for which the method is advantageous is the following.

EXAMPLE 3 **Laguerre's Equation. Laguerre Polynomials**

Laguerre's ODE is

$$ty'' + (1 - t)y' + ny = 0. \tag{9}$$

We determine a solution of (9) with $n = 0, 1, 2, \cdots$. From (7)–(9) we get the subsidiary equation

$$\left[-2sY - s^2\frac{dY}{ds} + y(0)\right] + sY - y(0) - \left(-Y - s\frac{dY}{ds}\right) + nY = 0.$$

Simplification gives

$$(s - s^2)\frac{dY}{ds} + (n + 1 - s)Y = 0.$$

Separating variables, using partial fractions, integrating (with the constant of integration taken to be zero), and taking exponentials, we get

$$\text{(10*)} \qquad \frac{dY}{Y} = -\frac{n+1-s}{s-s^2}\,ds = \left(\frac{n}{s-1} - \frac{n+1}{s}\right)ds \qquad \text{and} \qquad Y = \frac{(s-1)^n}{s^{n+1}}.$$

We write $l_n = \mathcal{L}^{-1}(Y)$ and prove **Rodrigues's formula**

$$\text{(10)} \qquad l_0 = 1, \qquad l_n(t) = \frac{e^t}{n!}\frac{d^n}{dt^n}(t^n e^{-t}), \qquad n = 1, 2, \cdots.$$

These are polynomials because the exponential terms cancel if we perform the indicated differentiations. They are called **Laguerre polynomials** and are usually denoted by L_n (see Problem Set 5.7, but we continue to reserve capital letters for transforms). We prove (10). By Table 6.1 and the first shifting theorem (s-shifting),

$$\mathcal{L}(t^n e^{-t}) = \frac{n!}{(s+1)^{n+1}}, \qquad \text{hence by (3) in Sec. 6.2} \qquad \mathcal{L}\left\{\frac{d^n}{dt^n}(t^n e^{-t})\right\} = \frac{n!s^n}{(s+1)^{n+1}}$$

because the derivatives up to the order $n - 1$ are zero at 0. Now make another shift and divide by $n!$ to get [see (10) and then (10*)]

$$\mathcal{L}(l_n) = \frac{(s-1)^n}{s^{n+1}} = Y. \qquad \blacksquare$$

PROBLEM SET 6.6

1. **REVIEW REPORT. Differentiation and Integration of Functions and Transforms.** Make a draft of these four operations from memory. Then compare your draft with the text and write a 2- to 3-page report on these operations and their significance in applications.

2–11 TRANSFORMS BY DIFFERENTIATION

Showing the details of your work, find $\mathcal{L}(f)$ if $f(t)$ equals:

2. $3t \sinh 4t$
3. $\frac{1}{2}te^{-3t}$
4. $te^{-t}\cos t$
5. $t \cos \omega t$
6. $t^2 \sin 3t$
7. $t^2 \cosh 2t$
8. $te^{-kt}\sin t$
9. $\frac{1}{2}t^2 \sin \pi t$
10. $t^n e^{kt}$
11. $4t \cos \frac{1}{2}\pi t$
12. **CAS PROJECT. Laguerre Polynomials. (a)** Write a CAS program for finding $l_n(t)$ in explicit form from (10). Apply it to calculate $l_0, \cdots, l_{10}$. Verify that $l_0, \cdots, l_{10}$ satisfy Laguerre's differential equation (9).

 (b) Show that

$$l_n(t) = \sum_{m=0}^{n} \frac{(-1)^m}{m!}\binom{n}{m}t^m$$

 and calculate $l_0, \cdots, l_{10}$ from this formula.

 (c) Calculate $l_0, \cdots, l_{10}$ recursively from $l_0 = 1$, $l_1 = 1 - t$ by

$$(n+1)l_{n+1} = (2n + 1 - t)l_n - nl_{n-1}.$$

 (d) A **generating function** (definition in Problem Set 5.2) for the Laguerre polynomials is

$$\sum_{n=0}^{\infty} l_n(t)x^n = (1-x)^{-1}e^{tx/(x-1)}.$$

 Obtain $l_0, \cdots, l_{10}$ from the corresponding partial sum of this power series in x and compare the l_n with those in (a), (b), or (c).

13. **CAS EXPERIMENT. Laguerre Polynomials.** Experiment with the graphs of $l_0, \cdots, l_{10}$, finding out empirically how the first maximum, first minimum, $\cdots$ is moving with respect to its location as a function of n. Write a short report on this.

14–20 INVERSE TRANSFORMS

Using differentiation, integration, s-shifting, or convolution, and showing the details, find $f(t)$ if $\mathcal{L}(f)$ equals:

14. $\dfrac{s}{(s^2+16)^2}$

15. $\dfrac{s}{(s^2-9)^2}$

16. $\dfrac{2s+6}{(s^2+6s+10)^2}$

17. $\ln \dfrac{s}{s-1}$

18. $\operatorname{arccot} \dfrac{s}{\pi}$

19. $\ln \dfrac{s^2+1}{(s-1)^2}$

20. $\ln \dfrac{s+a}{s+b}$

6.7 Systems of ODEs

The Laplace transform method may also be used for solving systems of ODEs, as we shall explain in terms of typical applications. We consider a first-order linear system with constant coefficients (as discussed in Sec. 4.1)

$$
\begin{aligned}
y_1' &= a_{11}y_1 + a_{12}y_2 + g_1(t) \\
y_2' &= a_{21}y_1 + a_{22}y_2 + g_2(t).
\end{aligned}
\tag{1}
$$

Writing $Y_1 = \mathcal{L}(y_1)$, $Y_2 = \mathcal{L}(y_2)$, $G_1 = \mathcal{L}(g_1)$, $G_2 = \mathcal{L}(g_2)$, we obtain from (1) in Sec. 6.2 the subsidiary system

$$
\begin{aligned}
sY_1 - y_1(0) &= a_{11}Y_1 + a_{12}Y_2 + G_1(s) \\
sY_2 - y_2(0) &= a_{21}Y_1 + a_{22}Y_2 + G_2(s).
\end{aligned}
$$

By collecting the Y_1- and Y_2-terms we have

$$
\begin{aligned}
(a_{11}-s)Y_1 + \quad a_{12}Y_2 \qquad &= -y_1(0) - G_1(s) \\
a_{21}Y_1 \qquad + (a_{22}-s)Y_2 &= -y_2(0) - G_2(s).
\end{aligned}
\tag{2}
$$

By solving this system algebraically for $Y_1(s)$, $Y_2(s)$ and taking the inverse transform we obtain the solution $y_1 = \mathcal{L}^{-1}(Y_1)$, $y_2 = \mathcal{L}^{-1}(Y_2)$ of the given system (1).

Note that (1) and (2) may be written in vector form (and similarly for the systems in the examples); thus, setting $\mathbf{y} = [y_1 \quad y_2]^{\mathsf{T}}$, $\mathbf{A} = [a_{jk}]$, $\mathbf{g} = [g_1 \quad g_2]^{\mathsf{T}}$, $\mathbf{Y} = [Y_1 \quad Y_2]^{\mathsf{T}}$, $\mathbf{G} = [G_1 \quad G_2]^{\mathsf{T}}$ we have

$$\mathbf{y}' = \mathbf{A}\mathbf{y} + \mathbf{g} \qquad \text{and} \qquad (\mathbf{A} - s\mathbf{I})\mathbf{Y} = -\mathbf{y}(0) - \mathbf{G}.$$

EXAMPLE 1 Mixing Problem Involving Two Tanks

Tank T_1 in Fig. 144 initially contains 100 gal of pure water. Tank T_2 initially contains 100 gal of water in which 150 lb of salt are dissolved. The inflow into T_1 is 2 gal/min from T_2 and 6 gal/min containing 6 lb of salt from the outside. The inflow into T_2 is 8 gal/min from T_1. The outflow from T_2 is $2 + 6 = 8$ gal/min, as shown in the figure. The mixtures are kept uniform by stirring. Find and plot the salt contents $y_1(t)$ and $y_2(t)$ in T_1 and T_2, respectively.

Solution. The model is obtained in the form of two equations

$$\text{Time rate of change} = \text{Inflow/min} - \text{Outflow/min}$$

for the two tanks (see Sec. 4.1). Thus,

$$y_1' = -\tfrac{8}{100}y_1 + \tfrac{2}{100}y_2 + 6. \qquad y_2' = \tfrac{8}{100}y_1 - \tfrac{8}{100}y_2.$$

The initial conditions are $y_1(0) = 0$, $y_2(0) = 150$. From this we see that the subsidiary system (2) is

$$\begin{aligned} (-0.08 - s)Y_1 + \quad 0.02Y_2 \quad &= -\frac{6}{s} \\ 0.08Y_1 \quad + (-0.08 - s)Y_2 &= -150. \end{aligned}$$

We solve this algebraically for Y_1 and Y_2 by elimination (or by Cramer's rule in Sec. 7.7), and we write the solutions in terms of partial fractions,

$$\begin{aligned} Y_1 &= \frac{9s + 0.48}{s(s + 0.12)(s + 0.04)} = \frac{100}{s} - \frac{62.5}{s + 0.12} - \frac{37.5}{s + 0.04} \\ Y_2 &= \frac{150s^2 + 12s + 0.48}{s(s + 0.12)(s + 0.04)} = \frac{100}{s} + \frac{125}{s + 0.12} - \frac{75}{s + 0.04}. \end{aligned}$$

By taking the inverse transform we arrive at the solution

$$\begin{aligned} y_1 &= 100 - 62.5e^{-0.12t} - 37.5e^{-0.04t} \\ y_2 &= 100 + 125e^{-0.12t} - 75e^{-0.04t}. \end{aligned}$$

Figure 144 shows the interesting plot of these functions. Can you give physical explanations for their main features? Why do they have the limit 100? Why is y_2 not monotone, whereas y_1 is? Why is y_1 from some time on suddenly larger than y_2? Etc. ■

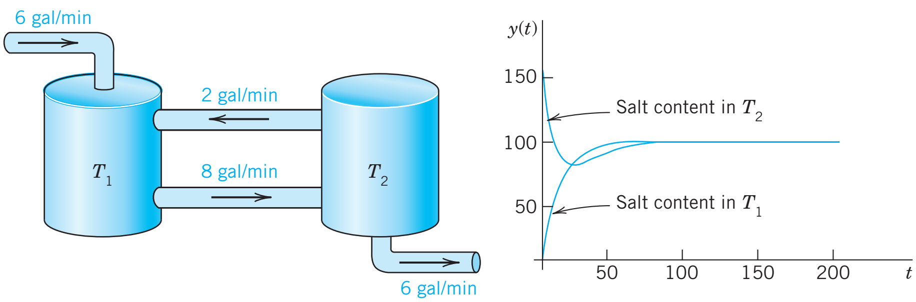


Fig. 144. Mixing problem in Example 1

Other systems of ODEs of practical importance can be solved by the Laplace transform method in a similar way, and eigenvalues and eigenvectors, as we had to determine them in Chap. 4, will come out automatically, as we have seen in Example 1.

EXAMPLE 2 Electrical Network

Find the currents $i_1(t)$ and $i_2(t)$ in the network in Fig. 145 with L and R measured in terms of the usual units (see Sec. 2.9), $v(t) = 100$ volts if $0 \leqq t \leqq 0.5$ sec and 0 thereafter, and $i(0) = 0$, $i'(0) = 0$.

Solution. The model of the network is obtained from Kirchhoff's Voltage Law as in Sec. 2.9. For the lower circuit we obtain

$$0.8i_1' + 1(i_1 - i_2) + 1.4i_1 = 100[1 - u(t - \tfrac{1}{2})]$$

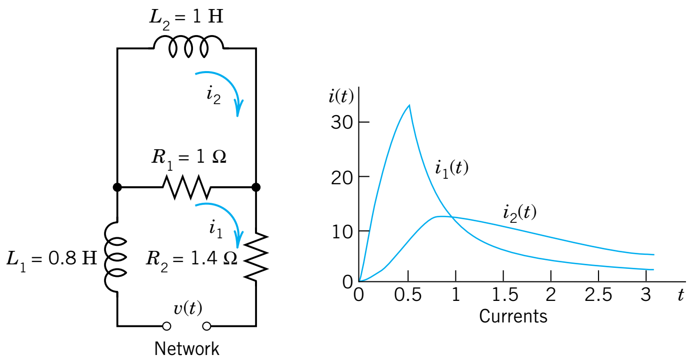


Fig. 145. Electrical network in Example 2

and for the upper

$$1 \cdot i_2' + 1(i_2 - i_1) \qquad\qquad = 0.$$

Division by 0.8 and ordering gives for the lower circuit

$$i_1' + 3i_1 - 1.25i_2 = 125[1 - u(t - \tfrac{1}{2})]$$

and for the upper

$$i_2' - i_1 + \qquad i_2 = 0.$$

With $i_1(0) = 0, i_2(0) = 0$ we obtain from (1) in Sec. 6.2 and the second shifting theorem the subsidiary system

$$\begin{aligned}(s + 3)I_1 - 1.25I_2 &= 125\left(\frac{1}{s} - \frac{e^{-s/2}}{s}\right)\\ -I_1 + (s + 1)I_2 &= 0.\end{aligned}$$

Solving algebraically for I_1 and I_2 gives

$$\begin{aligned}I_1 &= \frac{125(s + 1)}{s(s + \frac{1}{2})(s + \frac{7}{2})}(1 - e^{-s/2}),\\ I_2 &= \frac{125}{s(s + \frac{1}{2})(s + \frac{7}{2})}(1 - e^{-s/2}).\end{aligned}$$

The right sides, without the factor $1 - e^{-s/2}$, have the partial fraction expansions

$$\frac{500}{7s} - \frac{125}{3(s + \frac{1}{2})} - \frac{625}{21(s + \frac{7}{2})}$$

and

$$\frac{500}{7s} - \frac{250}{3(s + \frac{1}{2})} + \frac{250}{21(s + \frac{7}{2})},$$

respectively. The inverse transform of this gives the solution for $0 \leqq t \leqq \frac{1}{2}$,

$$\left.\begin{aligned}i_1(t) &= -\tfrac{125}{3}e^{-t/2} - \tfrac{625}{21}e^{-7t/2} + \tfrac{500}{7}\\ i_2(t) &= -\tfrac{250}{3}e^{-t/2} + \tfrac{250}{21}e^{-7t/2} + \tfrac{500}{7}\end{aligned}\right\} \qquad (0 \leqq t \leqq \tfrac{1}{2}).$$

According to the second shifting theorem the solution for $t > \frac{1}{2}$ is $i_1(t) - i_1(t - \frac{1}{2})$ and $i_2(t) - i_2(t - \frac{1}{2})$, that is,

$$\begin{aligned} i_1(t) &= -\tfrac{125}{3}(1 - e^{1/4})e^{-t/2} - \tfrac{625}{21}(1 - e^{7/4})e^{-7t/2} \\ i_2(t) &= -\tfrac{250}{3}(1 - e^{1/4})e^{-t/2} + \tfrac{250}{21}(1 - e^{7/4})e^{-7t/2} \end{aligned} \qquad (t > \tfrac{1}{2}).$$

Can you explain physically why both currents eventually go to zero, and why $i_1(t)$ has a sharp cusp whereas $i_2(t)$ has a continuous tangent direction at $t = \frac{1}{2}$? ■

Systems of ODEs of higher order can be solved by the Laplace transform method in a similar fashion. As an important application, typical of many similar mechanical systems, we consider coupled vibrating masses on springs.

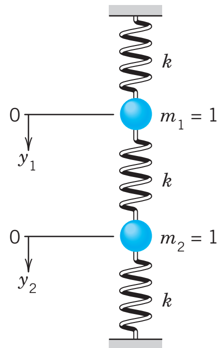


Fig. 146. Example 3

EXAMPLE 3 Model of Two Masses on Springs (Fig. 146)

The mechanical system in Fig. 146 consists of two bodies of mass 1 on three springs of the same spring constant k and of negligibly small masses of the springs. Also damping is assumed to be practically zero. Then the model of the physical system is the system of ODEs

$$\begin{aligned} y_1'' &= -ky_1 + k(y_2 - y_1) \\ y_2'' &= -k(y_2 - y_1) - ky_2. \end{aligned} \tag{3}$$

Here y_1 and y_2 are the displacements of the bodies from their positions of static equilibrium. These ODEs follow from **Newton's second law**, *Mass* $\times$ *Acceleration* $=$ *Force,* as in Sec. 2.4 for a single body. We again regard downward forces as positive and upward as negative. On the upper body, $-ky_1$ is the force of the upper spring and $k(y_2 - y_1)$ that of the middle spring, $y_2 - y_1$ being the net change in spring length—think this over before going on. On the lower body, $-k(y_2 - y_1)$ is the force of the middle spring and $-ky_2$ that of the lower spring.

We shall determine the solution corresponding to the initial conditions $y_1(0) = 1$, $y_2(0) = 1$, $y_1'(0) = \sqrt{3k}$, $y_2'(0) = -\sqrt{3k}$. Let $Y_1 = \mathscr{L}(y_1)$ and $Y_2 = \mathscr{L}(y_2)$. Then from (2) in Sec. 6.2 and the initial conditions we obtain the subsidiary system

$$\begin{aligned} s^2Y_1 - s - \sqrt{3k} &= -kY_1 + k(Y_2 - Y_1) \\ s^2Y_2 - s + \sqrt{3k} &= -k(Y_2 - Y_1) - kY_2. \end{aligned}$$

This system of linear algebraic equations in the unknowns Y_1 and Y_2 may be written

$$\begin{aligned} (s^2 + 2k)Y_1 - \qquad kY_2 \qquad &= s + \sqrt{3k} \\ -ky_1 \qquad + (s^2 + 2k)Y_2 &= s - \sqrt{3k}. \end{aligned}$$

Elimination (or Cramer's rule in Sec. 7.7) yields the solution, which we can expand in terms of partial fractions,

$$Y_1 = \frac{(s+\sqrt{3k})(s^2+2k)+k(s-\sqrt{3k})}{(s^2+2k)^2-k^2} = \frac{s}{s^2+k} + \frac{\sqrt{3k}}{s^2+3k}$$

$$Y_2 = \frac{(s^2+2k)(s-\sqrt{3k})+k(s+\sqrt{3k})}{(s^2+2k)^2-k^2} = \frac{s}{s^2+k} - \frac{\sqrt{3k}}{s^2+3k}.$$

Hence the solution of our initial value problem is (Fig. 147)

$$y_1(t) = \mathcal{L}^{-1}(Y_1) = \cos\sqrt{k}t + \sin\sqrt{3k}t$$

$$y_2(t) = \mathcal{L}^{-1}(Y_2) = \cos\sqrt{k}t - \sin\sqrt{3k}t.$$

We see that the motion of each mass is harmonic (the system is undamped!), being the superposition of a "slow" oscillation and a "rapid" oscillation. ■

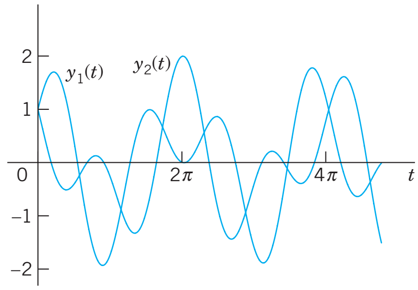


Fig. 147. Solutions in Example 3

PROBLEM SET 6.7

1. **TEAM PROJECT. Comparison of Methods for Linear Systems of ODEs**

   (a) Models. Solve the models in Examples 1 and 2 of Sec. 4.1 by Laplace transforms and compare the amount of work with that in Sec. 4.1. Show the details of your work.

   (b) Homogeneous Systems. Solve the systems (8), (11)–(13) in Sec. 4.3 by Laplace transforms. Show the details.

   (c) Nonhomogeneous System. Solve the system (3) in Sec. 4.6 by Laplace transforms. Show the details.

2–15 SYSTEMS OF ODES

Using the Laplace transform and showing the details of your work, solve the IVP:

2. $y_1' + y_2 = 0,\quad y_1 + y_2' = 2\cos t,$ $y_1(0) = 1,\quad y_2(0) = 0$
3. $y_1' = -y_1 + 4y_2,\quad y_2' = 3y_1 - 2y_2,$ $y_1(0) = 3,\quad y_2(0) = 4$
4. $y_1' = 4y_2 - 8\cos 4t,\quad y_2' = -3y_1 - 9\sin 4t,$ $y_1(0) = 0,\quad y_2(0) = 3$
5. $y_1' = y_2 + 1 - u(t-1),\quad y_2' = -y_1 + 1 - u(t-1),$ $y_1(0) = 0,\quad y_2(0) = 0$
6. $y_1' = 5y_1 + y_2,\quad y_2' = y_1 + 5y_2,$ $y_1(0) = 1,\quad y_2(0) = -3$
7. $y_1' = 2y_1 - 4y_2 + u(t-1)e^t,$ $y_2' = y_1 - 3y_2 + u(t-1)e^t,\quad y_1(0) = 3,\quad y_2(0) = 0$
8. $y_1' = -2y_1 + 3y_2,\quad y_2' = 4y_1 - y_2,$ $y_1(0) = 4,\quad y_2(0) = 3$
9. $y_1' = 4y_1 + y_2,\quad y_2' = -y_1 + 2y_2,\quad y_1(0) = 3,$ $y_2(0) = 1$
10. $y_1' = -y_2,\quad y_2' = -y_1 + 2[1 - u(t-2\pi)]\cos t,$ $y_1(0) = 1,\quad y_2(0) = 0$
11. $y_1'' = y_1 + 3y_2,\quad y_2'' = 4y_1 - 4e^t,$ $y_1(0) = 2,\quad y_1'(0) = 3,\quad y_2(0) = 1,\quad y_2'(0) = 2$
12. $y_1'' = -2y_1 + 2y_2,\quad y_2'' = 2y_1 - 5y_2,$ $y_1(0) = 1,\quad y_1'(0) = 0,\quad y_2(0) = 3,\quad y_2'(0) = 0$
13. $y_1'' + y_2 = -101\sin 10t,\quad y_2'' + y_1 = 101\sin 10t,$ $y_1(0) = 0,\quad y_1'(0) = 6,\quad y_2(0) = 8,\quad y_2'(0) = -6$

14. $4y_1' + y_2' - 2y_3' = 0, \quad -2y_1' + y_3' = 1,$
$2y_2' - 4y_3' = -16t$
$y_1(0) = 2, \quad y_2(0) = 0, \quad y_3(0) = 0$

15. $y_1' + y_2' = 2 \sinh t, \quad y_2' + y_3' = e^t,$
$y_3' + y_1' = 2e^t + e^{-t}, \quad y_1(0) = 1, \quad y_2(0) = 1,$
$y_3(0) = 0$

FURTHER APPLICATIONS

16. Forced vibrations of two masses. Solve the model in Example 3 with $k = 4$ and initial conditions $y_1(0) = 1$, $y_1'(0) = 1$, $y_2(0) = 1$, $y_2' = -1$ under the assumption that the force $11 \sin t$ is acting on the first body and the force $-11 \sin t$ on the second. Graph the two curves on common axes and explain the motion physically.

17. CAS Experiment. Effect of Initial Conditions. In Prob. 16, vary the initial conditions systematically, describe and explain the graphs physically. The great variety of curves will surprise you. Are they always periodic? Can you find empirical laws for the changes in terms of continuous changes of those conditions?

18. Mixing problem. What will happen in Example 1 if you double all flows (in particular, an increase to 12 gal/min containing 12 lb of salt from the outside), leaving the size of the tanks and the initial conditions as before? First guess, then calculate. Can you relate the new solution to the old one?

19. Electrical network. Using Laplace transforms, find the currents $i_1(t)$ and $i_2(t)$ in Fig. 148, where $v(t) = 390 \cos t$ and $i_1(0) = 0$, $i_2(0) = 0$. How soon will the currents practically reach their steady state?

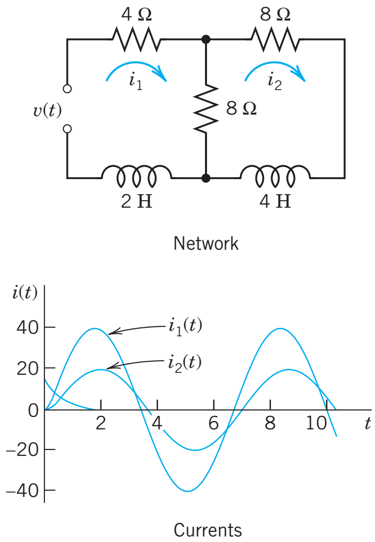


Fig. 148. Electrical network and currents in Problem 19

20. Single cosine wave. Solve Prob. 19 when the EMF (electromotive force) is acting from 0 to 2π only. Can you do this just by looking at Prob. 19, practically without calculation?

6.8 Laplace Transform: General Formulas

Formula	Name, Comments	Sec.
$F(s) = \mathcal{L}\{f(t)\} = \int_0^{\infty} e^{-st} f(t)\, dt$ $f(t) = \mathcal{L}^{-1}\{F(s)\}$	Definition of Transform Inverse Transform	6.1
$\mathcal{L}\{af(t) + bg(t)\} = a\mathcal{L}\{f(t)\} + b\mathcal{L}\{g(t)\}$	Linearity	6.1
$\mathcal{L}\{e^{at} f(t)\} = F(s - a)$ $\mathcal{L}^{-1}\{F(s - a)\} = e^{at} f(t)$	s-Shifting (First Shifting Theorem)	6.1
$\mathcal{L}(f') = s\mathcal{L}(f) - f(0)$ $\mathcal{L}(f'') = s^2\mathcal{L}(f) - sf(0) - f'(0)$ $\mathcal{L}(f^{(n)}) = s^n\mathcal{L}(f) - s^{(n-1)}f(0) - \cdots$ $\cdots - f^{(n-1)}(0)$ $\mathcal{L}\left\{\int_0^t f(\tau)\, d\tau\right\} = \frac{1}{s}\mathcal{L}(f)$	Differentiation of Function Integration of Function	6.2
$(f * g)(t) = \int_0^t f(\tau) g(t - \tau)\, d\tau$ $= \int_0^t f(t - \tau) g(\tau)\, d\tau$ $\mathcal{L}(f * g) = \mathcal{L}(f)\mathcal{L}(g)$	Convolution	6.5
$\mathcal{L}\{f(t - a)\, u(t - a)\} = e^{-as}F(s)$ $\mathcal{L}^{-1}\{e^{-as}F(s)\} = f(t - a)\, u(t - a)$	t-Shifting (Second Shifting Theorem)	6.3
$\mathcal{L}\{tf(t)\} = -F'(s)$ $\mathcal{L}\left\{\frac{f(t)}{t}\right\} = \int_s^{\infty} F(\tilde{s})\, d\tilde{s}$	Differentiation of Transform Integration of Transform	6.6
$\mathcal{L}(f) = \frac{1}{1 - e^{-ps}} \int_0^p e^{-st} f(t)\, dt$	f Periodic with Period p	6.4 Project 16

6.9 Table of Laplace Transforms

For more extensive tables, see Ref.* [A9] *in Appendix 1.

	$F(s) = \mathcal{L}\{f(t)\}$	$f(t)$	Sec.
1	$1/s$	1	6.1
2	$1/s^2$	t	
3	$1/s^n \qquad (n = 1, 2, \cdots)$	$t^{n-1}/(n-1)!$	
4	$1/\sqrt{s}$	$1/\sqrt{\pi t}$	
5	$1/s^{3/2}$	$2\sqrt{t/\pi}$	
6	$1/s^a \qquad (a > 0)$	$t^{a-1}/\Gamma(a)$	
7	$\dfrac{1}{s-a}$	e^{at}	6.1
8	$\dfrac{1}{(s-a)^2}$	te^{at}	
9	$\dfrac{1}{(s-a)^n} \qquad (n = 1, 2, \cdots)$	$\dfrac{1}{(n-1)!} t^{n-1}e^{at}$	
10	$\dfrac{1}{(s-a)^k} \qquad (k > 0)$	$\dfrac{1}{\Gamma(k)} t^{k-1}e^{at}$	
11	$\dfrac{1}{(s-a)(s-b)} \qquad (a \neq b)$	$\dfrac{1}{a-b}(e^{at} - e^{bt})$	
12	$\dfrac{s}{(s-a)(s-b)} \qquad (a \neq b)$	$\dfrac{1}{a-b}(ae^{at} - be^{bt})$	
13	$\dfrac{1}{s^2 + \omega^2}$	$\dfrac{1}{\omega} \sin \omega t$	6.1
14	$\dfrac{s}{s^2 + \omega^2}$	$\cos \omega t$	
15	$\dfrac{1}{s^2 - a^2}$	$\dfrac{1}{a} \sinh at$	
16	$\dfrac{s}{s^2 - a^2}$	$\cosh at$	
17	$\dfrac{1}{(s-a)^2 + \omega^2}$	$\dfrac{1}{\omega} e^{at} \sinh \omega t$	
18	$\dfrac{s-a}{(s-a)^2 + \omega^2}$	$e^{at} \cos \omega t$	
19	$\dfrac{1}{s(s^2 + \omega^2)}$	$\dfrac{1}{\omega^2}(1 - \cos \omega t)$	6.2
20	$\dfrac{1}{s^2(s^2 + \omega^2)}$	$\dfrac{1}{\omega^3}(\omega t - \sin \omega t)$	

(continued)

Table of Laplace Transforms (*continued*)

	$F(s) = \mathcal{L}\{f(t)\}$	$f(t)$	Sec.
21	$\dfrac{1}{(s^2+\omega^2)^2}$	$\dfrac{1}{2\omega^3}(\sin\omega t - \omega t\cos\omega t)$	6.6
22	$\dfrac{s}{(s^2+\omega^2)^2}$	$\dfrac{t}{2\omega}\sin\omega t$	6.6
23	$\dfrac{s^2}{(s^2+\omega^2)^2}$	$\dfrac{1}{2\omega}(\sin\omega t + \omega t\cos\omega t)$	6.6
24	$\dfrac{s}{(s^2+a^2)(s^2+b^2)} \quad (a^2 \neq b^2)$	$\dfrac{1}{b^2-a^2}(\cos at - \cos bt)$	
25	$\dfrac{1}{s^4+4k^4}$	$\dfrac{1}{4k^3}(\sin kt\cos kt - \cos kt\sinh kt)$	
26	$\dfrac{s}{s^4+4k^4}$	$\dfrac{1}{2k^2}\sin kt\sinh kt$	
27	$\dfrac{1}{s^4-k^4}$	$\dfrac{1}{2k^3}(\sinh kt - \sin kt)$	
28	$\dfrac{s}{s^4-k^4}$	$\dfrac{1}{2k^2}(\cosh kt - \cos kt)$	
29	$\sqrt{s-a} - \sqrt{s-b}$	$\dfrac{1}{2\sqrt{\pi t^3}}(e^{bt} - e^{at})$	
30	$\dfrac{1}{\sqrt{s+a}\,\sqrt{s+b}}$	$e^{-(a+b)t/2}I_0\left(\dfrac{a-b}{2}t\right)$	I 5.5
31	$\dfrac{1}{\sqrt{s^2+a^2}}$	$J_0(at)$	J 5.4
32	$\dfrac{s}{(s-a)^{3/2}}$	$\dfrac{1}{\sqrt{\pi t}}e^{at}(1+2at)$	
33	$\dfrac{1}{(s^2-a^2)^k} \quad (k>0)$	$\dfrac{\sqrt{\pi}}{\Gamma(k)}\left(\dfrac{t}{2a}\right)^{k-1/2} I_{k-1/2}(at)$	I 5.5
34	e^{-as}/s	$u(t-a)$	6.3
35	e^{-as}	$\delta(t-a)$	6.4
36	$\dfrac{1}{s}e^{-k/s}$	$J_0(2\sqrt{kt})$	J 5.4
37	$\dfrac{1}{\sqrt{s}}e^{-k/s}$	$\dfrac{1}{\sqrt{\pi t}}\cos 2\sqrt{kt}$	
38	$\dfrac{1}{s^{3/2}}e^{k/s}$	$\dfrac{1}{\sqrt{\pi k}}\sinh 2\sqrt{kt}$	
39	$e^{-k\sqrt{s}} \quad (k>0)$	$\dfrac{k}{2\sqrt{\pi t^3}}e^{-k^2/4t}$	

(continued)

Table of Laplace Transforms (*continued*)

	$F(s) = \mathcal{L}\{f(t)\}$	$f(t)$	Sec.
40	$\dfrac{1}{s}\ln s$	$-\ln t - \gamma \quad (\gamma \approx 0.5772)$	γ 5.5
41	$\ln \dfrac{s-a}{s-b}$	$\dfrac{1}{t}(e^{bt} - e^{at})$	
42	$\ln \dfrac{s^2+\omega^2}{s^2}$	$\dfrac{2}{t}(1 - \cos \omega t)$	6.6
43	$\ln \dfrac{s^2-a^2}{s^2}$	$\dfrac{2}{t}(1 - \cosh at)$	
44	$\arctan \dfrac{\omega}{s}$	$\dfrac{1}{t}\sin \omega t$	
45	$\dfrac{1}{s}\operatorname{arccot} s$	$\mathrm{Si}(t)$	App. A3.1

CHAPTER 6 REVIEW QUESTIONS AND PROBLEMS

1. State the Laplace transforms of a few simple functions from memory.

2. What are the steps of solving an ODE by the Laplace transform?

3. In what cases of solving ODEs is the present method preferable to that in Chap. 2?

4. What property of the Laplace transform is crucial in solving ODEs?

5. Is $\mathcal{L}\{f(t) + g(t)\} = \mathcal{L}\{f(t)\} + \mathcal{L}\{g(t)\}$? $\mathcal{L}\{f(t)g(t)\} = \mathcal{L}\{f(t)\}\mathcal{L}\{g(t)\}$? Explain.

6. When and how do you use the unit step function and Dirac's delta?

7. If you know $f(t) = \mathcal{L}^{-1}\{F(s)\}$, how would you find $\mathcal{L}^{-1}\{F(s)/s^2\}$?

8. Explain the use of the two shifting theorems from memory.

9. Can a discontinuous function have a Laplace transform? Give reason.

10. If two different continuous functions have transforms, the latter are different. Why is this practically important?

11–19 LAPLACE TRANSFORMS

Find the transform, indicating the method used and showing the details.

11. $5 \cosh 2t - 3 \sinh t$

12. $e^{-t}(\cos 4t - 2 \sin 4t)$

13. $\sin^2(\frac{1}{2}\pi t)$

14. $16t^2 u(t - \frac{1}{4})$

15. $e^{t/2}u(t - 3)$

16. $u(t - 2\pi) \sin t$

17. $t \cos t + \sin t$

18. $(\sin \omega t) * (\cos \omega t)$

19. $12t * e^{-3t}$

20–28 INVERSE LAPLACE TRANSFORM

Find the inverse transform, indicating the method used and showing the details:

20. $\dfrac{7.5}{s^2 - 2s - 8}$

21. $\dfrac{s+1}{s^2}e^{-s}$

22. $\dfrac{\frac{1}{16}}{s^2 + s + \frac{1}{2}}$

23. $\dfrac{\omega \cos \theta + s \sin \theta}{s^2 + \omega^2}$

24. $\dfrac{s^2 - 6.25}{(s^2 + 6.25)^2}$

25. $\dfrac{6(s+1)}{s^4}$

26. $\dfrac{2s - 10}{s^3}e^{-5s}$

27. $\dfrac{3s + 4}{s^2 + 4s + 5}$

28. $\dfrac{3s}{s^2 - 2s + 2}$

29–37 ODEs AND SYSTEMS

Solve by the Laplace transform, showing the details and graphing the solution:

29. $y'' + 4y' + 5y = 50t, \quad y(0) = 5, \quad y'(0) = -5$

30. $y'' + 16y = 4\delta(t - \pi), \quad y(0) = -1, \quad y'(0) = 0$

31. $y'' - y' - 2y = 12u(t - \pi) \sin t, \quad y(0) = 1, \quad y'(0) = -1$

32. $y'' + 4y = \delta(t - \pi) - \delta(t - 2\pi), \quad y(0) = 1, \quad y'(0) = 0$

33. $y'' + 3y' + 2y = 2u(t - 2), \quad y(0) = 0, \quad y'(0) = 0$

34. $y_1' = y_2, \quad y_2' = -4y_1 + \delta(t - \pi), \quad y_1(0) = 0, \quad y_2(0) = 0$

35. $y_1' = 2y_1 - 4y_2, \quad y_2' = y_1 - 3y_2, \quad y_1(0) = 3, \quad y_2(0) = 0$

36. $y_1' = 2y_1 + 4y_2, \quad y_2' = y_1 + 2y_2, \quad y_1(0) = -4, \quad y_2(0) = -4$

37. $y_1' = y_2 + u(t - \pi), \quad y_2' = -y_1 + u(t - 2\pi), \quad y_1(0) = 1, \quad y_2(0) = 0$

38–45 MASS–SPRING SYSTEMS, CIRCUITS, NETWORKS

Model and solve by the Laplace transform:

38. Show that the model of the mechanical system in Fig. 149 (no friction, no damping) is

$$m_1 y_1'' = -k_1 y_1 + k_2(y_2 - y_1)$$
$$m_2 y_2'' = -k_2(y_2 - y_1) - k_3 y_2.$$

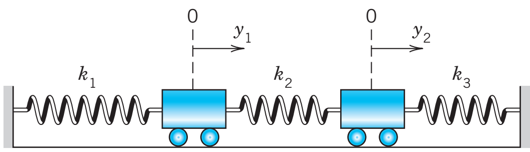


Fig. 149. System in Problems 38 and 39

39. In Prob. 38, let $m_1 = m_2 = 10$ kg, $k_1 = k_3 = 20$ kg/sec^2, $k_2 = 40$ kg/sec^2. Find the solution satisfying the initial conditions $y_1(0) = y_2(0) = 0$, $y_1'(0) = 1$ meter/sec, $y_2'(0) = -1$ meter/sec.

40. Find the model (the system of ODEs) in Prob. 38 extended by adding another mass m_3 and another spring of modulus k_4 in series.

41. Find the current $i(t)$ in the *RC*-circuit in Fig. 150, where $R = 10\ \Omega$, $C = 0.1$ F, $v(t) = 10t$ V if $0 < t < 4$, $v(t) = 40$ V if $t > 4$, and the initial charge on the capacitor is 0.

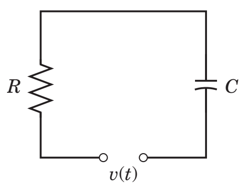


Fig. 150. *RC*-circuit

42. Find and graph the charge $q(t)$ and the current $i(t)$ in the *LC*-circuit in Fig. 151, assuming $L = 1$ H, $C = 1$ F, $v(t) = 1 - e^{-t}$ if $0 < t < \pi$, $v(t) = 0$ if $t > \pi$, and zero initial current and charge.

43. Find the current $i(t)$ in the *RLC*-circuit in Fig. 152, where $R = 160\ \Omega$, $L = 20$ H, $C = 0.002$ F, $v(t) = 37 \sin 10t$ V, and current and charge at $t = 0$ are zero.

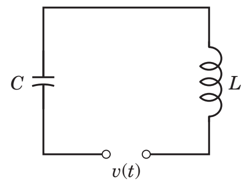


Fig. 151. *LC*-circuit

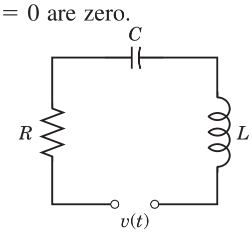


Fig. 152. *RLC*-circuit

44. Show that, by Kirchhoff's Voltage Law (Sec. 2.9), the currents in the network in Fig. 153 are obtained from the system

$$Li_1' + R(i_1 - i_2) = v(t)$$
$$R(i_2' - i_1') + \frac{1}{C} i_2 = 0.$$

Solve this system, assuming that $R = 10\ \Omega$, $L = 20$ H, $C = 0.05$ F, $v = 20$ V, $i_1(0) = 0$, $i_2(0) = 2$ A.

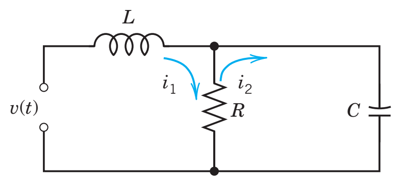


Fig. 153. Network in Problem 44

45. Set up the model of the network in Fig. 154 and find the solution, assuming that all charges and currents are 0 when the switch is closed at $t = 0$. Find the limits of $i_1(t)$ and $i_2(t)$ as $t \to \infty$, (i) from the solution, (ii) directly from the given network.

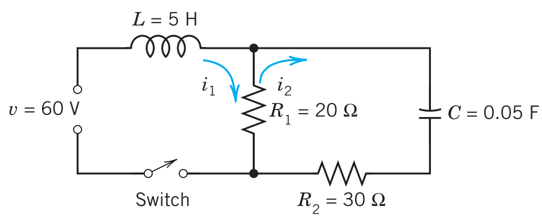


Fig. 154. Network in Problem 45

SUMMARY OF CHAPTER 6
Laplace Transforms

The main purpose of Laplace transforms is the solution of differential equations and systems of such equations, as well as corresponding initial value problems. The **Laplace transform** $F(s) = \mathcal{L}(f)$ of a function $f(t)$ is defined by

$$F(s) = \mathcal{L}(f) = \int_0^\infty e^{-st} f(t)\, dt \qquad \text{(Sec. 6.1).} \tag{1}$$

This definition is motivated by the property that the differentiation of f with respect to t corresponds to the multiplication of the transform F by s; more precisely,

$$\begin{aligned} \mathcal{L}(f') &= s\mathcal{L}(f) - f(0) \\ \mathcal{L}(f'') &= s^2\mathcal{L}(f) - sf(0) - f'(0) \end{aligned} \qquad \text{(Sec. 6.2)} \tag{2}$$

etc. Hence by taking the transform of a given differential equation

$$y'' + ay' + by = r(t) \qquad (a, b \text{ constant}) \tag{3}$$

and writing $\mathcal{L}(y) = Y(s)$, we obtain the **subsidiary equation**

$$(s^2 + as + b)Y = \mathcal{L}(r) + sf(0) + f'(0) + af(0). \tag{4}$$

Here, in obtaining the transform $\mathcal{L}(r)$ we can get help from the small table in Sec. 6.1 or the larger table in Sec. 6.9. This is the first step. In the second step we solve the subsidiary equation *algebraically* for $Y(s)$. In the third step we determine the **inverse transform** $y(t) = \mathcal{L}^{-1}(Y)$, that is, the solution of the problem. This is generally the hardest step, and in it we may again use one of those two tables. $Y(s)$ will often be a rational function, so that we can obtain the inverse $\mathcal{L}^{-1}(Y)$ by partial fraction reduction (Sec. 6.4) if we see no simpler way.

The Laplace method avoids the determination of a general solution of the homogeneous ODE, and we also need not determine values of arbitrary constants in a general solution from initial conditions; instead, we can insert the latter directly into (4). Two further facts account for the practical importance of the Laplace transform. First, it has some basic properties and resulting techniques that simplify the determination of transforms and inverses. The most important of these properties are listed in Sec. 6.8, together with references to the corresponding sections. More on the use of unit step functions and Dirac's delta can be found in Secs. 6.3 and 6.4, and more on convolution in Sec. 6.5. Second, due to these properties, the present method is particularly suitable for handling right sides $r(t)$ given by different expressions over different intervals of time, for instance, when $r(t)$ is a square wave or an impulse or of a form such as $r(t) = \cos t$ if $0 \leqq t \leqq 4\pi$ and 0 elsewhere.

The application of the Laplace transform to systems of ODEs is shown in Sec. 6.7. (The application to PDEs follows in Sec. 12.12.)

PART B

Linear Algebra. Vector Calculus

Matrices and *vectors*, which underlie **linear algebra** (Chaps. 7 and 8), allow us to represent numbers or functions in an ordered and compact form. Matrices can hold enormous amounts of data—think of a network of millions of computer connections or cell phone connections—in a form that can be rapidly processed by computers. The main topic of Chap. 7 is *how to solve systems of linear equations using matrices*. Concepts of rank, basis, linear transformations, and vector spaces are closely related. Chapter 8 deals with eigenvalue problems. Linear algebra is an active field that has many applications in engineering physics, numerics (see Chaps. 20–22), economics, and others.

Chapters 9 and 10 extend calculus to ***vector calculus***. We start with vectors from linear algebra and develop vector differential calculus. We differentiate functions of several variables and discuss vector differential operations such as grad, div, and curl. Chapter 10 extends regular integration to integration over curves, surfaces, and solids, thereby obtaining new types of integrals. Ingenious theorems by Gauss, Green, and Stokes allow us to transform these integrals into one another.

Software suitable for linear algebra (Lapack, Maple, Mathematica, Matlab) can be found in the list at the opening of Part E of the book if needed.

Numeric linear algebra (Chap. 20) ***can be studied directly after Chap. 7 or 8*** because Chap. 20 is independent of the other chapters in Part E on numerics.

CHAPTER 7

Linear Algebra: Matrices, Vectors, Determinants. Linear Systems

Linear algebra is a fairly extensive subject that covers vectors and matrices, determinants, systems of linear equations, vector spaces and linear transformations, eigenvalue problems, and other topics. As an area of study it has a broad appeal in that it has many applications in engineering, physics, geometry, computer science, economics, and other areas. It also contributes to a deeper understanding of mathematics itself.

Matrices, which are rectangular arrays of numbers or functions, and **vectors** are the main tools of linear algebra. Matrices are important because they let us express large amounts of data and functions in an organized and concise form. Furthermore, since matrices are single objects, we denote them by single letters and calculate with them directly. All these features have made matrices and vectors very popular for expressing scientific and mathematical ideas.

The chapter keeps a good mix between applications (electric networks, Markov processes, traffic flow, etc.) and theory. Chapter 7 is structured as follows: Sections 7.1 and 7.2 provide an intuitive introduction to matrices and vectors and their operations, including matrix multiplication. The next block of sections, that is, Secs. 7.3–7.5 provide the most important method for ***solving systems of linear equations*** by the Gauss elimination method. This method is a cornerstone of linear algebra, and the method itself and variants of it appear in different areas of mathematics and in many applications. It leads to a consideration of the behavior of solutions and concepts such as rank of a matrix, linear independence, and bases. We shift to determinants, a topic that has declined in importance, in Secs. 7.6 and 7.7. Section 7.8 covers inverses of matrices. The chapter ends with vector spaces, inner product spaces, linear transformations, and composition of linear transformations. Eigenvalue problems follow in Chap. 8.

COMMENT. ***Numeric linear algebra (Secs. 20.1–20.5) can be studied immediately after this chapter.***

Prerequisite: None.
Sections that may be omitted in a short course: 7.5, 7.9.
References and Answers to Problems: App. 1 Part B, and App. 2.

7.1 Matrices, Vectors: Addition and Scalar Multiplication

The basic concepts and rules of matrix and vector algebra are introduced in Secs. 7.1 and 7.2 and are followed by **linear systems** (systems of linear equations), a main application, in Sec. 7.3.

Let us first take a leisurely look at matrices before we formalize our discussion. A **matrix** is a rectangular array of numbers or functions which we will enclose in brackets. For example,

$$
(1) \qquad \begin{bmatrix} 0.3 & 1 & -5 \\ 0 & -0.2 & 16 \end{bmatrix}, \qquad \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix},
$$
$$
\begin{bmatrix} e^{-x} & 2x^2 \\ e^{6x} & 4x \end{bmatrix}, \qquad [a_1 \quad a_2 \quad a_3], \qquad \begin{bmatrix} 4 \\ \frac{1}{2} \end{bmatrix}
$$

are matrices. The numbers (or functions) are called **entries** or, less commonly, *elements* of the matrix. The first matrix in (1) has two **rows**, which are the horizontal lines of entries. Furthermore, it has three **columns**, which are the vertical lines of entries. The second and third matrices are **square matrices**, which means that each has as many rows as columns—3 and 2, respectively. The entries of the second matrix have two indices, signifying their location within the matrix. The first index is the number of the row and the second is the number of the column, so that together the entry's position is uniquely identified. For example, a_{23} (read *a two three*) is in Row 2 and Column 3, etc. The notation is standard and applies to all matrices, including those that are not square.

Matrices having just a single row or column are called **vectors**. Thus, the fourth matrix in (1) has just one row and is called a **row vector**. The last matrix in (1) has just one column and is called a **column vector**. Because the goal of the indexing of entries was to uniquely identify the position of an element within a matrix, one index suffices for vectors, whether they are row or column vectors. Thus, the third entry of the row vector in (1) is denoted by a_3.

Matrices are handy for storing and processing data in applications. Consider the following two common examples.

EXAMPLE 1 Linear Systems, a Major Application of Matrices

We are given a system of linear equations, briefly a **linear system**, such as

$$
\begin{aligned}
4x_1 + 6x_2 + 9x_3 &= 6 \\
6x_1 \qquad\quad - 2x_3 &= 20 \\
5x_1 - 8x_2 + x_3 &= 10
\end{aligned}
$$

where x_1, x_2, x_3 are the **unknowns**. We form the **coefficient matrix**, call it **A**, by listing the coefficients of the unknowns in the position in which they appear in the linear equations. In the second equation, there is no unknown x_2, which means that the coefficient of x_2 is 0 and hence in matrix **A**, $a_{22} = 0$, Thus,

$$\mathbf{A} = \begin{bmatrix} 4 & 6 & 9 \\ 6 & 0 & -2 \\ 5 & -8 & 1 \end{bmatrix}. \qquad \text{We form another matrix} \qquad \tilde{\mathbf{A}} = \begin{bmatrix} 4 & 6 & 9 & 6 \\ 6 & 0 & -2 & 20 \\ 5 & -8 & 1 & 10 \end{bmatrix}$$

by augmenting **A** with the right sides of the linear system and call it the augmented matrix of the system.

Since we can go back and recapture the system of linear equations directly from the augmented matrix $\tilde{\mathbf{A}}$, $\tilde{\mathbf{A}}$ contains all the information of the system and can thus be used to solve the linear system. This means that we can just use the augmented matrix to do the calculations needed to solve the system. We shall explain this in detail in Sec. 7.3. Meanwhile you may verify by substitution that the solution is $x_1 = 3, x_2 = \frac{1}{2}, x_3 = -1$.

The notation x_1, x_2, x_3 for the unknowns is practical but not essential; we could choose x, y, z or some other letters. ■

EXAMPLE 2 Sales Figures in Matrix Form

Sales figures for three products I, II, III in a store on Monday (Mon), Tuesday (Tues), $\cdots$ may for each week be arranged in a matrix

$$\begin{array}{cc} & \begin{array}{ccccccc} \text{Mon} & \text{Tues} & \text{Wed} & \text{Thur} & \text{Fri} & \text{Sat} & \text{Sun} \end{array} \\ \mathbf{A} = & \begin{bmatrix} 40 & 33 & 81 & 0 & 21 & 47 & 33 \\ 0 & 12 & 78 & 50 & 50 & 96 & 90 \\ 10 & 0 & 0 & 27 & 43 & 78 & 56 \end{bmatrix} \cdot \begin{array}{c} \text{I} \\ \text{II} \\ \text{III} \end{array} \end{array}$$

If the company has 10 stores, we can set up 10 such matrices, one for each store. Then, by adding corresponding entries of these matrices, we can get a matrix showing the total sales of each product on each day. Can you think of other data which can be stored in matrix form? For instance, in transportation or storage problems? Or in listing distances in a network of roads? ■

General Concepts and Notations

Let us formalize what we just have discussed. We shall denote matrices by capital boldface letters **A**, **B**, **C**, $\cdots$, or by writing the general entry in brackets; thus $\mathbf{A} = [a_{jk}]$, and so on. By an ***m* × *n* matrix** (read *m by n matrix*) we mean a matrix with m rows and n columns—rows always come first! $m \times n$ is called the **size** of the matrix. Thus an $m \times n$ matrix is of the form

$$\textbf{(2)} \qquad \mathbf{A} = [a_{jk}] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

The matrices in (1) are of sizes 2×3, 3×3, 2×2, 1×3, and 2×1, respectively.

Each entry in (2) has two subscripts. The first is the *row number* and the second is the *column number*. Thus a_{21} is the entry in Row 2 and Column 1.

If $m = n$, we call **A** an $n \times n$ **square matrix**. Then its diagonal containing the entries $a_{11}, a_{22}, \cdots, a_{nn}$ is called the **main diagonal** of **A**. Thus the main diagonals of the two square matrices in (1) are a_{11}, a_{22}, a_{33} and e^{-x}, $4x$, respectively.

Square matrices are particularly important, as we shall see. A matrix of any size $m \times n$ is called a **rectangular matrix**; this includes square matrices as a special case.

Vectors

A **vector** is a matrix with only one row or column. Its entries are called the **components** of the vector. We shall denote vectors by ***lowercase*** boldface letters $\mathbf{a}, \mathbf{b}, \cdots$ or by its general component in brackets, $\mathbf{a} = [a_j]$, and so on. Our special vectors in (1) suggest that a (general) **row vector** is of the form

$$\mathbf{a} = [a_1 \quad a_2 \quad \cdots \quad a_n]. \qquad \text{For instance,} \qquad \mathbf{a} = [-2 \quad 5 \quad 0.8 \quad 0 \quad 1].$$

A **column vector** is of the form

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}. \qquad \text{For instance,} \qquad \mathbf{b} = \begin{bmatrix} 4 \\ 0 \\ -7 \end{bmatrix}.$$

Addition and Scalar Multiplication of Matrices and Vectors

What makes matrices and vectors really useful and particularly suitable for computers is the fact that we can calculate with them almost as easily as with numbers. Indeed, we now introduce rules for addition and for scalar multiplication (multiplication by numbers) that were suggested by practical applications. (Multiplication of matrices by matrices follows in the next section.) We first need the concept of equality.

DEFINITION

Equality of Matrices

Two matrices $\mathbf{A} = [a_{jk}]$ and $\mathbf{B} = [b_{jk}]$ are **equal**, written $\mathbf{A} = \mathbf{B}$, if and only if they have the same size and the corresponding entries are equal, that is, $a_{11} = b_{11}$, $a_{12} = b_{12}$, and so on. Matrices that are not equal are called **different**. Thus, matrices of different sizes are always different.

EXAMPLE 3 **Equality of Matrices**

Let

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \qquad \text{and} \qquad \mathbf{B} = \begin{bmatrix} 4 & 0 \\ 3 & -1 \end{bmatrix}.$$

Then

$$\mathbf{A} = \mathbf{B} \qquad \text{if and only if} \qquad \begin{matrix} a_{11} = 4, & a_{12} = 0, \\ a_{21} = 3, & a_{22} = -1. \end{matrix}$$

The following matrices are all different. Explain!

$$\begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix} \qquad \begin{bmatrix} 4 & 2 \\ 1 & 3 \end{bmatrix} \qquad \begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix} \qquad \begin{bmatrix} 1 & 3 & 0 \\ 4 & 2 & 0 \end{bmatrix} \qquad \begin{bmatrix} 0 & 1 & 3 \\ 0 & 4 & 2 \end{bmatrix}$$ ■

DEFINITION

Addition of Matrices

The **sum** of two matrices $\mathbf{A} = [a_{jk}]$ and $\mathbf{B} = [b_{jk}]$ ***of the same size*** is written $\mathbf{A} + \mathbf{B}$ and has the entries $a_{jk} + b_{jk}$ obtained by adding the corresponding entries of $\mathbf{A}$ and $\mathbf{B}$. Matrices of different sizes cannot be added.

As a special case, the **sum** $\mathbf{a} + \mathbf{b}$ of two row vectors or two column vectors, which must have the same number of components, is obtained by adding the corresponding components.

EXAMPLE 4 **Addition of Matrices and Vectors**

$$\text{If} \quad \mathbf{A} = \begin{bmatrix} -4 & 6 & 3 \\ 0 & 1 & 2 \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} 5 & -1 & 0 \\ 3 & 1 & 0 \end{bmatrix}, \quad \text{then} \quad \mathbf{A} + \mathbf{B} = \begin{bmatrix} 1 & 5 & 3 \\ 3 & 2 & 2 \end{bmatrix}.$$

$\mathbf{A}$ in Example 3 and our present $\mathbf{A}$ cannot be added. If $\mathbf{a} = [5 \quad 7 \quad 2]$ and $\mathbf{b} = [-6 \quad 2 \quad 0]$, then $\mathbf{a} + \mathbf{b} = [-1 \quad 9 \quad 2]$.

An application of matrix addition was suggested in Example 2. Many others will follow. ∎

DEFINITION

Scalar Multiplication (Multiplication by a Number)

The **product** of any $m \times n$ matrix $\mathbf{A} = [a_{jk}]$ and any **scalar** c (number c) is written $c\mathbf{A}$ and is the $m \times n$ matrix $c\mathbf{A} = [ca_{jk}]$ obtained by multiplying each entry of $\mathbf{A}$ by c.

Here $(-1)\mathbf{A}$ is simply written $-\mathbf{A}$ and is called the **negative** of $\mathbf{A}$. Similarly, $(-k)\mathbf{A}$ is written $-k\mathbf{A}$. Also, $\mathbf{A} + (-\mathbf{B})$ is written $\mathbf{A} - \mathbf{B}$ and is called the **difference** of $\mathbf{A}$ and $\mathbf{B}$ (which must have the same size!).

EXAMPLE 5 **Scalar Multiplication**

$$\text{If} \quad \mathbf{A} = \begin{bmatrix} 2.7 & -1.8 \\ 0 & 0.9 \\ 9.0 & -4.5 \end{bmatrix}, \quad \text{then} \quad -\mathbf{A} = \begin{bmatrix} -2.7 & 1.8 \\ 0 & -0.9 \\ -9.0 & 4.5 \end{bmatrix}, \quad \frac{10}{9}\mathbf{A} = \begin{bmatrix} 3 & -2 \\ 0 & 1 \\ 10 & -5 \end{bmatrix}, \quad 0\mathbf{A} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

If a matrix $\mathbf{B}$ shows the distances between some cities in miles, $1.609\mathbf{B}$ gives these distances in kilometers. ∎

Rules for Matrix Addition and Scalar Multiplication. From the familiar laws for the addition of numbers we obtain similar laws for the addition of matrices of the same size $m \times n$, namely,

$$\textbf{(3)} \quad \begin{aligned} &\text{(a)} \quad && \mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A} \\ &\text{(b)} \quad && (\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C}) \qquad \text{(written } \mathbf{A} + \mathbf{B} + \mathbf{C}\text{)} \\ &\text{(c)} \quad && \mathbf{A} + \mathbf{0} = \mathbf{A} \\ &\text{(d)} \quad && \mathbf{A} + (-\mathbf{A}) = \mathbf{0}. \end{aligned}$$

Here $\mathbf{0}$ denotes the **zero matrix** (of size $m \times n$), that is, the $m \times n$ matrix with all entries zero. If $m = 1$ or $n = 1$, this is a vector, called a **zero vector**.

Hence matrix addition is *commutative* and *associative* [by (3a) and (3b)].
Similarly, for scalar multiplication we obtain the rules

$$\textbf{(4)}\qquad \begin{aligned} &\text{(a)} \quad c(\mathbf{A}+\mathbf{B}) = c\mathbf{A} + c\mathbf{B} \\ &\text{(b)} \quad (c+k)\mathbf{A} = c\mathbf{A} + k\mathbf{A} \\ &\text{(c)} \quad c(k\mathbf{A}) = (ck)\mathbf{A} \qquad \text{(written } ck\mathbf{A}\text{)} \\ &\text{(d)} \quad 1\mathbf{A} = \mathbf{A}. \end{aligned}$$

PROBLEM SET 7.1

1–7 GENERAL QUESTIONS

1. **Equality.** Give reasons why the five matrices in Example 3 are all different.
2. **Double subscript notation.** If you write the matrix in Example 2 in the form $\mathbf{A} = [a_{jk}]$, what is a_{31}? a_{13}? a_{26}? a_{33}?
3. **Sizes.** What sizes do the matrices in Examples 1, 2, 3, and 5 have?
4. **Main diagonal.** What is the main diagonal of **A** in Example 1? Of **A** and **B** in Example 3?
5. **Scalar multiplication.** If **A** in Example 2 shows the number of items sold, what is the matrix **B** of units sold if a unit consists of **(a)** 5 items and **(b)** 10 items?
6. If a 12×12 matrix **A** shows the distances between 12 cities in kilometers, how can you obtain from **A** the matrix **B** showing these distances in miles?
7. **Addition of vectors.** Can you add: A row and a column vector with different numbers of components? With the same number of components? Two row vectors with the same number of components but different numbers of zeros? A vector and a scalar? A vector with four components and a 2×2 matrix?

8–16 ADDITION AND SCALAR MULTIPLICATION OF MATRICES AND VECTORS

Let

$$\mathbf{A} = \begin{bmatrix} 0 & 2 & 4 \\ 6 & 5 & 5 \\ 1 & 0 & -3 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 0 & 5 & 2 \\ 5 & 3 & 4 \\ -2 & 4 & -2 \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} 5 & 2 \\ -2 & 4 \\ 1 & 0 \end{bmatrix}, \qquad \mathbf{D} = \begin{bmatrix} -4 & 1 \\ 5 & 0 \\ 2 & -1 \end{bmatrix},$$

$$\mathbf{E} = \begin{bmatrix} 0 & 2 \\ 3 & 4 \\ 3 & -1 \end{bmatrix}$$

$$\mathbf{u} = \begin{bmatrix} 1.5 \\ 0 \\ -3.0 \end{bmatrix}, \qquad \mathbf{v} = \begin{bmatrix} -1 \\ 3 \\ 2 \end{bmatrix}, \qquad \mathbf{w} = \begin{bmatrix} -5 \\ -30 \\ 10 \end{bmatrix}.$$

Find the following expressions, indicating which of the rules in (3) or (4) they illustrate, or give reasons why they are not defined.

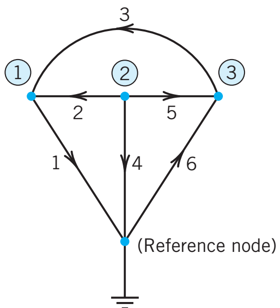
8. $2\mathbf{A} + 4\mathbf{B}$, $4\mathbf{B} + 2\mathbf{A}$, $0\mathbf{A} + \mathbf{B}$, $0.4\mathbf{B} - 4.2\mathbf{A}$
9. $3\mathbf{A}$, $0.5\mathbf{B}$, $3\mathbf{A} + 0.5\mathbf{B}$, $3\mathbf{A} + 0.5\mathbf{B} + \mathbf{C}$
10. $(4 \cdot 3)\mathbf{A}$, $4(3\mathbf{A})$, $14\mathbf{B} - 3\mathbf{B}$, $11\mathbf{B}$
11. $8\mathbf{C} + 10\mathbf{D}$, $2(5\mathbf{D} + 4\mathbf{C})$, $0.6\mathbf{C} - 0.6\mathbf{D}$, $0.6(\mathbf{C} - \mathbf{D})$
12. $(\mathbf{C} + \mathbf{D}) + \mathbf{E}$, $(\mathbf{D} + \mathbf{E}) + \mathbf{C}$, $0(\mathbf{C} - \mathbf{E}) + 4\mathbf{D}$, $\mathbf{A} - 0\mathbf{C}$
13. $(2 \cdot 7)\mathbf{C}$, $2(7\mathbf{C})$, $-\mathbf{D} + 0\mathbf{E}$, $\mathbf{E} - \mathbf{D} + \mathbf{C} + \mathbf{u}$
14. $(5\mathbf{u} + 5\mathbf{v}) - \frac{1}{2}\mathbf{w}$, $-20(\mathbf{u} + \mathbf{v}) + 2\mathbf{w}$, $\mathbf{E} - (\mathbf{u} + \mathbf{v})$, $10(\mathbf{u} + \mathbf{v}) + \mathbf{w}$
15. $(\mathbf{u} + \mathbf{v}) - \mathbf{w}$, $\mathbf{u} + (\mathbf{v} - \mathbf{w})$, $\mathbf{C} + 0\mathbf{w}$, $0\mathbf{E} + \mathbf{u} - \mathbf{v}$
16. $15\mathbf{v} - 3\mathbf{w} - 0\mathbf{u}$, $-3\mathbf{w} + 15\mathbf{v}$, $\mathbf{D} - \mathbf{u} + 3\mathbf{C}$, $8.5\mathbf{w} - 11.1\mathbf{u} + 0.4\mathbf{v}$
17. **Resultant of forces.** If the above vectors **u**, **v**, **w** represent forces in space, their sum is called their *resultant*. Calculate it.
18. **Equilibrium.** By definition, forces are *in equilibrium* if their resultant is the zero vector. Find a force **p** such that the above **u**, **v**, **w**, and **p** are in equilibrium.
19. **General rules.** Prove (3) and (4) for general 2×3 matrices and scalars c and k.

20. TEAM PROJECT. Matrices for Networks. Matrices have various engineering applications, as we shall see. For instance, they can be used to characterize connections in electrical networks, in nets of roads, in production processes, etc., as follows.

(a) Nodal Incidence Matrix. The network in Fig. 155 consists of six *branches* (connections) and four *nodes* (points where two or more branches come together). One node is the *reference node* (grounded node, whose voltage is zero). We number the other nodes and number and direct the branches. This we do arbitrarily. The network can now be described by a matrix $\mathbf{A} = [a_{jk}]$, where

$$a_{jk} = \begin{cases} +1 \text{ if branch } k \text{ leaves node } (j) \\ -1 \text{ if branch } k \text{ enters node } (j) \\ \quad 0 \text{ if branch } k \text{ does not touch node } (j). \end{cases}$$

$\mathbf{A}$ is called the *nodal incidence matrix* of the network. Show that for the network in Fig. 155 the matrix $\mathbf{A}$ has the given form.



$$\begin{array}{c} \text{Branch} \\ \text{Node } (1) \\ \text{Node } (2) \\ \text{Node } (3) \end{array} \begin{array}{c} \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 \end{matrix} \\ \begin{bmatrix} 1 & -1 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & -1 & -1 \end{bmatrix} \end{array}$$

Fig. 155. Network and nodal incidence matrix in Team Project 20(a)

(b) Find the nodal incidence matrices of the networks in Fig. 156.

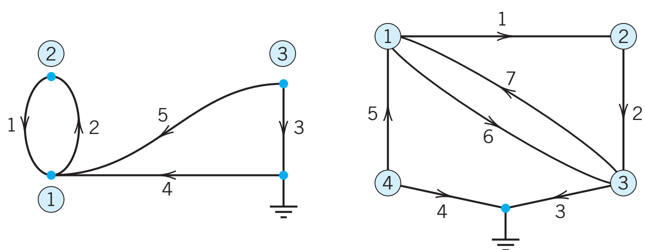


Fig. 156. Electrical networks in Team Project 20(b)

(c) Sketch the three networks corresponding to the nodal incidence matrices

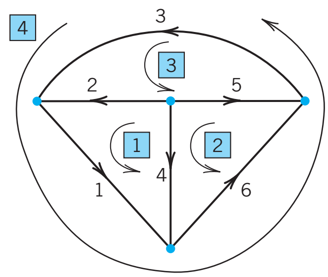
$$\begin{bmatrix} 1 & 0 & 0 & 1 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & -1 & 0 & 0 & 1 \\ -1 & 1 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 \end{bmatrix},$$

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ -1 & 1 & 0 & 1 & 0 \\ 0 & -1 & -1 & 0 & 1 \end{bmatrix}.$$

(d) Mesh Incidence Matrix. A network can also be characterized by the *mesh incidence matrix* $\mathbf{M} = [m_{jk}]$, where

$$m_{jk} = \begin{cases} +1 \text{ if branch } k \text{ is in mesh } \boxed{j} \\ \qquad \text{and has the same orientation} \\ -1 \text{ if branch } k \text{ is in mesh } \boxed{j} \\ \qquad \text{and has the opposite orientation} \\ \quad 0 \text{ if branch } k \text{ is not in mesh } \boxed{j} \end{cases}$$

and a mesh is a loop with no branch in its interior (or in its exterior). Here, the meshes are numbered and directed (oriented) in an arbitrary fashion. Show that for the network in Fig. 157, the matrix $\mathbf{M}$ has the given form, where Row 1 corresponds to mesh 1, etc.



$$\mathbf{M} = \begin{bmatrix} 1 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 1 \\ 0 & -1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$

Fig. 157. Network and matrix $\mathbf{M}$ in Team Project 20(d)

7.2 Matrix Multiplication

Matrix multiplication means that one multiplies matrices by matrices. Its definition is standard but it looks artificial. *Thus you have to study matrix multiplication carefully*, multiply a few matrices together for practice until you can understand how to do it. Here then is the definition. (Motivation follows later.)

DEFINITION

Multiplication of a Matrix by a Matrix

The **product** $\mathbf{C} = \mathbf{AB}$ (in this order) of an $m \times n$ matrix $\mathbf{A} = [a_{jk}]$ times an $r \times p$ matrix $\mathbf{B} = [b_{jk}]$ is defined if and only if $r = n$ and is then the $m \times p$ matrix $\mathbf{C} = [c_{jk}]$ with entries

$$c_{jk} = \sum_{l=1}^{n} a_{jl} b_{lk} = a_{j1}b_{1k} + a_{j2}b_{2k} + \cdots + a_{jn}b_{nk} \qquad \begin{aligned} j &= 1, \cdots, m \\ k &= 1, \cdots, p. \end{aligned} \tag{1}$$

The condition $r = n$ means that the second factor, $\mathbf{B}$, must have as many rows as the first factor has columns, namely n. A diagram of sizes that shows when matrix multiplication is possible is as follows:

$$\begin{array}{ccccc} \mathbf{A} & \mathbf{B} & = & \mathbf{C} \\ [m \times n] & [n \times p] & = & [m \times p]. \end{array}$$

The entry c_{jk} in (1) is obtained by multiplying each entry in the jth row of $\mathbf{A}$ by the corresponding entry in the kth column of $\mathbf{B}$ and then adding these n products. For instance, $c_{21} = a_{21}b_{11} + a_{22}b_{21} + \cdots + a_{2n}b_{n1}$, and so on. One calls this briefly a ***multiplication of rows into columns***. For $n = 3$, this is illustrated by

$$m = 4 \left\{ \overbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix}}^{n=3} \right. \overbrace{\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix}}^{p=2} = \left. \overbrace{\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \\ c_{31} & c_{32} \\ c_{41} & c_{42} \end{bmatrix}}^{p=2} \right\} m = 4$$

Notations in a product $\mathbf{AB} = \mathbf{C}$

where we shaded the entries that contribute to the calculation of entry c_{21} just discussed.

Matrix multiplication will be motivated by its use in *linear transformations* in this section and more fully in Sec. 7.9.

Let us illustrate the main points of matrix multiplication by some examples. Note that matrix multiplication also includes multiplying a matrix by a vector, since, after all, a vector is a special matrix.

EXAMPLE 1 **Matrix Multiplication**

$$\mathbf{AB} = \begin{bmatrix} 3 & 5 & -1 \\ 4 & 0 & 2 \\ -6 & -3 & 2 \end{bmatrix} \begin{bmatrix} 2 & -2 & 3 & 1 \\ 5 & 0 & 7 & 8 \\ 9 & -4 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 22 & -2 & 43 & 42 \\ 26 & -16 & 14 & 6 \\ -9 & 4 & -37 & -28 \end{bmatrix}$$

Here $c_{11} = 3 \cdot 2 + 5 \cdot 5 + (-1) \cdot 9 = 22$, and so on. The entry in the box is $c_{23} = 4 \cdot 3 + 0 \cdot 7 + 2 \cdot 1 = 14$. The product $\mathbf{BA}$ is not defined.

EXAMPLE 2 Multiplication of a Matrix and a Vector

$$\begin{bmatrix} 4 & 2 \\ 1 & 8 \end{bmatrix}\begin{bmatrix} 3 \\ 5 \end{bmatrix} = \begin{bmatrix} 4 \cdot 3 + 2 \cdot 5 \\ 1 \cdot 3 + 8 \cdot 5 \end{bmatrix} = \begin{bmatrix} 22 \\ 43 \end{bmatrix} \quad \text{whereas} \quad \begin{bmatrix} 3 \\ 5 \end{bmatrix}\begin{bmatrix} 4 & 2 \\ 1 & 8 \end{bmatrix} \quad \text{is undefined.}$$

EXAMPLE 3 Products of Row and Column Vectors

$$[3 \quad 6 \quad 1]\begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix} = [19], \qquad \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}[3 \quad 6 \quad 1] = \begin{bmatrix} 3 & 6 & 1 \\ 6 & 12 & 2 \\ 12 & 24 & 4 \end{bmatrix}.$$

EXAMPLE 4 CAUTION! Matrix Multiplication Is Not Commutative, AB ≠ BA in General

This is illustrated by Examples 1 and 2, where one of the two products is not even defined, and by Example 3, where the two products have different sizes. But it also holds for square matrices. For instance,

$$\begin{bmatrix} 1 & 1 \\ 100 & 100 \end{bmatrix}\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{but} \quad \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 100 & 100 \end{bmatrix} = \begin{bmatrix} 99 & 99 \\ -99 & -99 \end{bmatrix}.$$

It is interesting that this also shows that $\mathbf{AB} = \mathbf{0}$ does ***not*** necessarily imply $\mathbf{BA} = \mathbf{0}$ or $\mathbf{A} = \mathbf{0}$ or $\mathbf{B} = \mathbf{0}$. We shall discuss this further in Sec. 7.8, along with reasons when this happens.

Our examples show that in matrix products ***the order of factors must always be observed very carefully***. Otherwise matrix multiplication satisfies rules similar to those for numbers, namely.

$$\text{(2)} \qquad \begin{aligned} &\text{(a)} \quad (k\mathbf{A})\mathbf{B} = k(\mathbf{AB}) = \mathbf{A}(k\mathbf{B}) \quad \textit{written } k\mathbf{AB} \textit{ or } \mathbf{A}k\mathbf{B} \\ &\text{(b)} \quad \mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C} \qquad\qquad \textit{written } \mathbf{ABC} \\ &\text{(c)} \quad (\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC} \\ &\text{(d)} \quad \mathbf{C}(\mathbf{A} + \mathbf{B}) = \mathbf{CA} + \mathbf{CB} \end{aligned}$$

provided **A**, **B**, and **C** are such that the expressions on the left are defined; here, k is any scalar. (2b) is called the **associative law**. (2c) and (2d) are called the **distributive laws**.

Since matrix multiplication is a multiplication of rows into columns, we can write the defining formula (1) more compactly as

$$\textbf{(3)} \qquad \boxed{c_{jk} = \mathbf{a}_j\mathbf{b}_k,} \qquad j = 1, \cdots, m; \quad k = 1, \cdots, p,$$

where $\mathbf{a}_j$ is the jth row vector of **A** and $\mathbf{b}_k$ is the kth column vector of **B**, so that in agreement with (1),

$$\mathbf{a}_j\mathbf{b}_k = [a_{j1} \quad a_{j2} \quad \cdots \quad a_{jn}]\begin{bmatrix} b_{1k} \\ \vdots \\ b_{nk} \end{bmatrix} = a_{j1}b_{1k} + a_{j2}b_{2k} + \cdots + a_{jn}b_{nk}.$$

EXAMPLE 5 Product in Terms of Row and Column Vectors

If $\mathbf{A} = [a_{jk}]$ is of size 3×3 and $\mathbf{B} = [b_{jk}]$ is of size 3×4, then

$$\mathbf{AB} = \begin{bmatrix} \mathbf{a}_1\mathbf{b}_1 & \mathbf{a}_1\mathbf{b}_2 & \mathbf{a}_1\mathbf{b}_3 & \mathbf{a}_1\mathbf{b}_4 \\ \mathbf{a}_2\mathbf{b}_1 & \mathbf{a}_2\mathbf{b}_2 & \mathbf{a}_2\mathbf{b}_3 & \mathbf{a}_2\mathbf{b}_4 \\ \mathbf{a}_3\mathbf{b}_1 & \mathbf{a}_3\mathbf{b}_2 & \mathbf{a}_3\mathbf{b}_3 & \mathbf{a}_3\mathbf{b}_4 \end{bmatrix}. \tag{4}$$

Taking $\mathbf{a}_1 = [3 \quad 5 \quad -1]$, $\mathbf{a}_2 = [4 \quad 0 \quad 2]$, etc., verify (4) for the product in Example 1. ■

Parallel processing of products on the computer is facilitated by a variant of (3) for computing $\mathbf{C} = \mathbf{AB}$, which is used by standard algorithms (such as in Lapack). In this method, $\mathbf{A}$ is used as given, $\mathbf{B}$ is taken in terms of its column vectors, and the product is computed columnwise; thus,

$$\mathbf{AB} = \mathbf{A}[\mathbf{b}_1 \quad \mathbf{b}_2 \quad \cdots \quad \mathbf{b}_p] = [\mathbf{Ab}_1 \quad \mathbf{Ab}_2 \quad \cdots \quad \mathbf{Ab}_p]. \tag{5}$$

Columns of $\mathbf{B}$ are then assigned to different processors (individually or several to each processor), which simultaneously compute the columns of the product matrix $\mathbf{Ab}_1$, $\mathbf{Ab}_2$, etc.

EXAMPLE 6 Computing Products Columnwise by (5)

To obtain

$$\mathbf{AB} = \begin{bmatrix} 4 & 1 \\ -5 & 2 \end{bmatrix}\begin{bmatrix} 3 & 0 & 7 \\ -1 & 4 & 6 \end{bmatrix} = \begin{bmatrix} 11 & 4 & 34 \\ -17 & 8 & -23 \end{bmatrix}$$

from (5), calculate the columns

$$\begin{bmatrix} 4 & 1 \\ -5 & 2 \end{bmatrix}\begin{bmatrix} 3 \\ -1 \end{bmatrix} = \begin{bmatrix} 11 \\ -17 \end{bmatrix}, \quad \begin{bmatrix} 4 & 1 \\ -5 & 2 \end{bmatrix}\begin{bmatrix} 0 \\ 4 \end{bmatrix} = \begin{bmatrix} 4 \\ 8 \end{bmatrix}, \quad \begin{bmatrix} 4 & 1 \\ -5 & 2 \end{bmatrix}\begin{bmatrix} 7 \\ 6 \end{bmatrix} = \begin{bmatrix} 34 \\ -23 \end{bmatrix}$$

of $\mathbf{AB}$ and then write them as a single matrix, as shown in the first formula on the right. ■

Motivation of Multiplication by Linear Transformations

Let us now motivate the "unnatural" matrix multiplication by its use in **linear transformations**. For $n = 2$ variables these transformations are of the form

$$\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 \\ y_2 &= a_{21}x_1 + a_{22}x_2 \end{aligned} \tag{6*}$$

and suffice to explain the idea. (For general n they will be discussed in Sec. 7.9.) For instance, (6*) may relate an x_1x_2-coordinate system to a y_1y_2-coordinate system in the plane. In vectorial form we can write (6*) as

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \mathbf{Ax} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{bmatrix}. \tag{6}$$

Now suppose further that the x_1x_2-system is related to a w_1w_2-system by another linear transformation, say,

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{Bw} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} b_{11}w_1 + b_{12}w_2 \\ b_{21}w_1 + b_{22}w_2 \end{bmatrix}. \tag{7}$$

Then the y_1y_2-system is related to the w_1w_2-system indirectly via the x_1x_2-system, and we wish to express this relation directly. Substitution will show that this direct relation is a linear transformation, too, say,

$$\mathbf{y} = \mathbf{Cw} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} c_{11}w_1 + c_{12}w_2 \\ c_{21}w_1 + c_{22}w_2 \end{bmatrix}. \tag{8}$$

Indeed, substituting (7) into (6), we obtain

$$\begin{aligned} y_1 &= a_{11}(b_{11}w_1 + b_{12}w_2) + a_{12}(b_{21}w_1 + b_{22}w_2) \\ &= (a_{11}b_{11} + a_{12}b_{21})w_1 + (a_{11}b_{12} + a_{12}b_{22})w_2 \\ y_2 &= a_{21}(b_{11}w_1 + b_{12}w_2) + a_{22}(b_{21}w_1 + b_{22}w_2) \\ &= (a_{21}b_{11} + a_{22}b_{21})w_1 + (a_{21}b_{12} + a_{22}b_{22})w_2. \end{aligned}$$

Comparing this with (8), we see that

$$\begin{aligned} c_{11} &= a_{11}b_{11} + a_{12}b_{21} \qquad & c_{12} &= a_{11}b_{12} + a_{12}b_{22} \\ c_{21} &= a_{21}b_{11} + a_{22}b_{21} \qquad & c_{22} &= a_{21}b_{12} + a_{22}b_{22}. \end{aligned}$$

This proves that $\mathbf{C} = \mathbf{AB}$ with the product defined as in (1). For larger matrix sizes the idea and result are exactly the same. Only the number of variables changes. We then have m variables y and n variables x and p variables w. The matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C} = \mathbf{AB}$ then have sizes $m \times n$, $n \times p$, and $m \times p$, respectively. And the requirement that $\mathbf{C}$ be the product $\mathbf{AB}$ leads to formula (1) in its general form. *This motivates matrix multiplication.*

Transposition

We obtain the transpose of a matrix by writing its rows as columns (or equivalently its columns as rows). This also applies to the transpose of vectors. Thus, a row vector becomes a column vector and vice versa. In addition, for square matrices, we can also "reflect" the elements along the main diagonal, that is, interchange entries that are symmetrically positioned with respect to the main diagonal to obtain the transpose. Hence a_{12} becomes a_{21}, a_{31} becomes a_{13}, and so forth. Example 7 illustrates these ideas. Also note that, if $\mathbf{A}$ is the given matrix, then we denote its transpose by $\mathbf{A}^{\mathsf{T}}$.

EXAMPLE 7 Transposition of Matrices and Vectors

If $$\mathbf{A} = \begin{bmatrix} 5 & -8 & 1 \\ 4 & 0 & 0 \end{bmatrix}, \quad \text{then} \quad \mathbf{A}^{\mathsf{T}} = \begin{bmatrix} 5 & 4 \\ -8 & 0 \\ 1 & 0 \end{bmatrix}.$$

A little more compactly, we can write

$$\begin{bmatrix} 5 & -8 & 1 \\ 4 & 0 & 0 \end{bmatrix}^{\mathsf{T}} = \begin{bmatrix} 5 & 4 \\ -8 & 0 \\ 1 & 0 \end{bmatrix}, \qquad \begin{bmatrix} 3 & 0 & 7 \\ 8 & -1 & 5 \\ 1 & -9 & 4 \end{bmatrix}^{\mathsf{T}} = \begin{bmatrix} 3 & 8 & 1 \\ 0 & -1 & -9 \\ 7 & 5 & 4 \end{bmatrix},$$

Furthermore, the transpose $[6 \quad 2 \quad 3]^{\mathsf{T}}$ of the row vector $[6 \quad 2 \quad 3]$ is the column vector

$$[6 \quad 2 \quad 3]^{\mathsf{T}} = \begin{bmatrix} 6 \\ 2 \\ 3 \end{bmatrix} \cdot \qquad \text{Conversely,} \qquad \begin{bmatrix} 6 \\ 2 \\ 3 \end{bmatrix}^{\mathsf{T}} = [6 \quad 2 \quad 3].$$

■

DEFINITION

Transposition of Matrices and Vectors

The transpose of an $m \times n$ matrix $\mathbf{A} = [a_{jk}]$ is the $n \times m$ matrix $\mathbf{A}^{\mathsf{T}}$ (read *A transpose*) that has the first *row* of $\mathbf{A}$ as its first *column*, the second *row* of $\mathbf{A}$ as its second *column*, and so on. Thus the transpose of $\mathbf{A}$ in (2) is $\mathbf{A}^{\mathsf{T}} = [a_{kj}]$, written out

$$\mathbf{A}^{\mathsf{T}} = [a_{kj}] = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \cdot & \cdot & \cdots & \cdot \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix}. \tag{9}$$

As a special case, transposition converts row vectors to column vectors and conversely.

Transposition gives us a choice in that we can work either with the matrix or its transpose, whichever is more convenient.

Rules for transposition are

(10)

$$\begin{aligned} &\text{(a)} & (\mathbf{A}^{\mathsf{T}})^{\mathsf{T}} &= \mathbf{A} \\ &\text{(b)} & (\mathbf{A} + \mathbf{B})^{\mathsf{T}} &= \mathbf{A}^{\mathsf{T}} + \mathbf{B}^{\mathsf{T}} \\ &\text{(c)} & (c\mathbf{A})^{\mathsf{T}} &= c\mathbf{A}^{\mathsf{T}} \\ &\text{(d)} & (\mathbf{A}\mathbf{B})^{\mathsf{T}} &= \mathbf{B}^{\mathsf{T}}\mathbf{A}^{\mathsf{T}}. \end{aligned}$$

CAUTION! Note that in (10d) the transposed matrices are ***in reversed order***. We leave the proofs as an exercise in Probs. 9 and 10.

Special Matrices

Certain kinds of matrices will occur quite frequently in our work, and we now list the most important ones of them.

Symmetric and Skew-Symmetric Matrices. Transposition gives rise to two useful classes of matrices. **Symmetric** matrices are square matrices whose transpose equals the matrix itself. **Skew-symmetric** matrices are square matrices whose transpose equals ***minus*** the matrix. Both cases are defined in (11) and illustrated by Example 8.

$$\mathbf{A}^\mathsf{T} = \mathbf{A} \quad (\text{thus } a_{kj} = a_{jk}), \qquad \mathbf{A}^\mathsf{T} = -\mathbf{A} \quad (\text{thus } a_{kj} = -a_{jk}, \text{ hence } a_{jj} = 0). \tag{11}$$

Symmetric Matrix | Skew-Symmetric Matrix

EXAMPLE 8 Symmetric and Skew-Symmetric Matrices

$$\mathbf{A} = \begin{bmatrix} 20 & 120 & 200 \\ 120 & 10 & 150 \\ 200 & 150 & 30 \end{bmatrix} \quad \text{is symmetric, and} \quad \mathbf{B} = \begin{bmatrix} 0 & 1 & -3 \\ -1 & 0 & -2 \\ 3 & 2 & 0 \end{bmatrix} \quad \text{is skew-symmetric.}$$

For instance, if a company has three building supply centers C_1, C_2, C_3, then $\mathbf{A}$ could show costs, say, a_{jj} for handling 1000 bags of cement at center C_j, and a_{jk} $(j \neq k)$ the cost of shipping 1000 bags from C_j to C_k. Clearly, $a_{jk} = a_{kj}$ if we assume shipping in the opposite direction will cost the same.

Symmetric matrices have several general properties which make them important. This will be seen as we proceed. ■

Triangular Matrices. **Upper triangular matrices** are square matrices that can have nonzero entries only on and ***above*** the main diagonal, whereas any entry below the diagonal must be zero. Similarly, **lower triangular matrices** can have nonzero entries only on and ***below*** the main diagonal. Any entry on the main diagonal of a triangular matrix may be zero or not.

EXAMPLE 9 Upper and Lower Triangular Matrices

$$\underbrace{\begin{bmatrix} 1 & 3 \\ 0 & 2 \end{bmatrix}, \quad \begin{bmatrix} 1 & 4 & 2 \\ 0 & 3 & 2 \\ 0 & 0 & 6 \end{bmatrix}}_{\text{Upper triangular}}, \quad \underbrace{\begin{bmatrix} 2 & 0 & 0 \\ 8 & -1 & 0 \\ 7 & 6 & 8 \end{bmatrix}, \quad \begin{bmatrix} 3 & 0 & 0 & 0 \\ 9 & -3 & 0 & 0 \\ 1 & 0 & 2 & 0 \\ 1 & 9 & 3 & 6 \end{bmatrix}}_{\text{Lower triangular}}.$$ ■

Diagonal Matrices. These are square matrices that can have nonzero entries only on the main diagonal. Any entry above or below the main diagonal must be zero.

If all the diagonal entries of a diagonal matrix $\mathbf{S}$ are equal, say, c, we call $\mathbf{S}$ a **scalar matrix** because multiplication of any square matrix $\mathbf{A}$ of the same size by $\mathbf{S}$ has the same effect as the multiplication by a scalar, that is,

$$\mathbf{AS} = \mathbf{SA} = c\mathbf{A}. \tag{12}$$

In particular, a scalar matrix, whose entries on the main diagonal are all 1, is called a **unit matrix** (or **identity matrix**) and is denoted by $\mathbf{I}_n$ or simply by $\mathbf{I}$. For $\mathbf{I}$, formula (12) becomes

$$\mathbf{AI} = \mathbf{IA} = \mathbf{A}. \tag{13}$$

EXAMPLE 10 Diagonal Matrix D. Scalar Matrix S. Unit Matrix I

$$\mathbf{D} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{S} = \begin{bmatrix} c & 0 & 0 \\ 0 & c & 0 \\ 0 & 0 & c \end{bmatrix}, \quad \mathbf{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$ ■

Some Applications of Matrix Multiplication

EXAMPLE 11 Computer Production. Matrix Times Matrix

Supercomp Ltd produces two computer models PC1086 and PC1186. The matrix **A** shows the cost per computer (in thousands of dollars) and **B** the production figures for the year 2010 (in multiples of 10,000 units.) Find a matrix **C** that shows the shareholders the cost per quarter (in millions of dollars) for raw material, labor, and miscellaneous.

$$\mathbf{A} = \begin{array}{c} \begin{array}{cc} \text{PC1086} & \text{PC1186} \end{array} \\ \begin{bmatrix} 1.2 & 1.6 \\ 0.3 & 0.4 \\ 0.5 & 0.6 \end{bmatrix} \end{array} \begin{array}{l} \\ \text{Raw Components} \\ \text{Labor} \\ \text{Miscellaneous} \end{array} \qquad \mathbf{B} = \begin{array}{c} \text{Quarter} \\ \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} 3 & 8 & 6 & 9 \\ 6 & 2 & 4 & 3 \end{bmatrix} \end{array} \begin{array}{l} \\ \\ \text{PC1086} \\ \text{PC1186} \end{array}$$

Solution.

$$\mathbf{C} = \mathbf{AB} = \begin{array}{c} \text{Quarter} \\ \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} 13.2 & 12.8 & 13.6 & 15.6 \\ 3.3 & 3.2 & 3.4 & 3.9 \\ 5.1 & 5.2 & 5.4 & 6.3 \end{bmatrix} \end{array} \begin{array}{l} \\ \\ \text{Raw Components} \\ \text{Labor} \\ \text{Miscellaneous} \end{array}$$

Since cost is given in multiples of \$1000 and production in multiples of 10,000 units, the entries of **C** are multiples of \$10 millions; thus $c_{11} = 13.2$ means \$132 million, etc. ■

EXAMPLE 12 Weight Watching. Matrix Times Vector

Suppose that in a weight-watching program, a person of 185 lb burns 350 cal/hr in walking (3 mph), 500 in bicycling (13 mph), and 950 in jogging (5.5 mph). Bill, weighing 185 lb, plans to exercise according to the matrix shown. Verify the calculations (W = Walking, B = Bicycling, J = Jogging).

$$\begin{array}{l} \\ \text{MON} \\ \text{WED} \\ \text{FRI} \\ \text{SAT} \end{array} \begin{array}{c} \begin{array}{ccc} \text{W} & \text{B} & \text{J} \end{array} \\ \begin{bmatrix} 1.0 & 0 & 0.5 \\ 1.0 & 1.0 & 0.5 \\ 1.5 & 0 & 0.5 \\ 2.0 & 1.5 & 1.0 \end{bmatrix} \end{array} \begin{bmatrix} 350 \\ 500 \\ 950 \end{bmatrix} = \begin{bmatrix} 825 \\ 1325 \\ 1000 \\ 2400 \end{bmatrix} \begin{array}{l} \text{MON} \\ \text{WED} \\ \text{FRI} \\ \text{SAT} \end{array}$$

■

EXAMPLE 13 Markov Process. Powers of a Matrix. Stochastic Matrix

Suppose that the 2004 state of land use in a city of 60 mi^2 of built-up area is

C: Commercially Used 25% I: Industrially Used 20% R: Residentially Used 55%.

Find the states in 2009, 2014, and 2019, assuming that the transition probabilities for 5-year intervals are given by the matrix **A** and remain practically the same over the time considered.

$$\mathbf{A} = \begin{array}{c} \begin{array}{ccc} \text{From C} & \text{From I} & \text{From R} \end{array} \\ \begin{bmatrix} 0.7 & 0.1 & 0 \\ 0.2 & 0.9 & 0.2 \\ 0.1 & 0 & 0.8 \end{bmatrix} \end{array} \begin{array}{l} \\ \text{To C} \\ \text{To I} \\ \text{To R} \end{array}$$

$\mathbf{A}$ is a **stochastic matrix**, that is, a square matrix with all entries nonnegative and all column sums equal to 1. Our example concerns a **Markov process**,[1] that is, a process for which the probability of entering a certain state depends only on the last state occupied (and the matrix $\mathbf{A}$), not on any earlier state.

Solution. From the matrix $\mathbf{A}$ and the 2004 state we can compute the 2009 state,

$$\begin{array}{c} \text{C} \\ \text{I} \\ \text{R} \end{array} \begin{bmatrix} 0.7 \cdot 25 + 0.1 \cdot 20 + 0 \cdot 55 \\ 0.2 \cdot 25 + 0.9 \cdot 20 + 0.2 \cdot 55 \\ 0.1 \cdot 25 + 0 \cdot 20 + 0.8 \cdot 55 \end{bmatrix} = \begin{bmatrix} 0.7 & 0.1 & 0 \\ 0.2 & 0.9 & 0.2 \\ 0.1 & 0 & 0.8 \end{bmatrix} \begin{bmatrix} 25 \\ 20 \\ 55 \end{bmatrix} = \begin{bmatrix} 19.5 \\ 34.0 \\ 46.5 \end{bmatrix}.$$

To explain: The 2009 figure for C equals 25% times the probability 0.7 that C goes into C, plus 20% times the probability 0.1 that I goes into C, plus 55% times the probability 0 that R goes into C. Together,

$$25 \cdot 0.7 + 20 \cdot 0.1 + 55 \cdot 0 = 19.5\,[\%]. \qquad \text{Also} \qquad 25 \cdot 0.2 + 20 \cdot 0.9 + 55 \cdot 0.2 = 34\,[\%].$$

Similarly, the new R is 46.5%. We see that the 2009 state vector is the column vector

$$\mathbf{y} = [19.5 \quad 34.0 \quad 46.5]^\mathsf{T} = \mathbf{Ax} = \mathbf{A}\,[25 \quad 20 \quad 55]^\mathsf{T}$$

where the column vector $\mathbf{x} = [25 \quad 20 \quad 55]^\mathsf{T}$ is the given 2004 state vector. Note that the sum of the entries of $\mathbf{y}$ is 100 [%]. Similarly, you may verify that for 2014 and 2019 we get the state vectors

$$\mathbf{z} = \mathbf{Ay} = \mathbf{A}(\mathbf{Ax}) = \mathbf{A}^2\mathbf{x} = [17.05 \quad 43.80 \quad 39.15]^\mathsf{T}$$
$$\mathbf{u} = \mathbf{Az} = \mathbf{A}^2\mathbf{y} = \mathbf{A}^3\mathbf{x} = [16.315 \quad 50.660 \quad 33.025]^\mathsf{T}.$$

Answer. In 2009 the commercial area will be 19.5% (11.7 mi^2), the industrial 34% (20.4 mi^2), and the residential 46.5% (27.9 mi^2). For 2014 the corresponding figures are 17.05%, 43.80%, and 39.15%. For 2019 they are 16.315%, 50.660%, and 33.025%. (In Sec. 8.2 we shall see what happens in the limit, assuming that those probabilities remain the same. In the meantime, can you experiment or guess?) ■

PROBLEM SET 7.2

1–10 GENERAL QUESTIONS

1. **Multiplication.** Why is multiplication of matrices restricted by conditions on the factors?
2. **Square matrix.** What form does a 3×3 matrix have if it is symmetric as well as skew-symmetric?
3. **Product of vectors.** Can every 3×3 matrix be represented by two vectors as in Example 3?
4. **Skew-symmetric matrix.** How many different entries can a 4×4 skew-symmetric matrix have? An $n \times n$ skew-symmetric matrix?
5. Same questions as in Prob. 4 for symmetric matrices.
6. **Triangular matrix.** If $\mathbf{U}_1, \mathbf{U}_2$ are upper triangular and $\mathbf{L}_1, \mathbf{L}_2$ are lower triangular, which of the following are triangular?

$$\mathbf{U}_1 + \mathbf{U}_2, \quad \mathbf{U}_1\mathbf{U}_2, \quad \mathbf{U}_1^2, \quad \mathbf{U}_1 + \mathbf{L}_1, \quad \mathbf{U}_1\mathbf{L}_1, \quad \mathbf{L}_1 + \mathbf{L}_2$$

7. **Idempotent matrix,** defined by $\mathbf{A}^2 = \mathbf{A}$. Can you find four 2×2 idempotent matrices?
8. **Nilpotent matrix,** defined by $\mathbf{B}^m = \mathbf{0}$ for some m. Can you find three 2×2 nilpotent matrices?
9. **Transposition.** Can you prove (10a)–(10c) for 3×3 matrices? For $m \times n$ matrices?
10. **Transposition.** **(a)** Illustrate (10d) by simple examples. **(b)** Prove (10d).

11–20 MULTIPLICATION, ADDITION, AND TRANSPOSITION OF MATRICES AND VECTORS

Let

$$\mathbf{A} = \begin{bmatrix} 4 & -2 & 3 \\ -2 & 1 & 6 \\ 1 & 2 & 2 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 1 & -3 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} 0 & 1 \\ 3 & 2 \\ -2 & 0 \end{bmatrix}, \qquad \mathbf{a} = [1 \quad -2 \quad 0], \quad \mathbf{b} = \begin{bmatrix} 3 \\ 1 \\ -1 \end{bmatrix}.$$

[1]ANDREI ANDREJEVITCH MARKOV (1856–1922), Russian mathematician, known for his work in probability theory.

Showing all intermediate results, calculate the following expressions or give reasons why they are undefined:

11. $\mathbf{AB}$, $\mathbf{AB}^\mathsf{T}$, $\mathbf{BA}$, $\mathbf{B}^\mathsf{T}\mathbf{A}$

12. $\mathbf{AA}^\mathsf{T}$, $\mathbf{A}^2$, $\mathbf{BB}^\mathsf{T}$, $\mathbf{B}^2$

13. $\mathbf{CC}^\mathsf{T}$, $\mathbf{BC}$, $\mathbf{CB}$, $\mathbf{C}^\mathsf{T}\mathbf{B}$

14. $3\mathbf{A} - 2\mathbf{B}$, $(3\mathbf{A} - 2\mathbf{B})^\mathsf{T}$, $3\mathbf{A}^\mathsf{T} - 2\mathbf{B}^\mathsf{T}$, $(3\mathbf{A} - 2\mathbf{B})^\mathsf{T}\mathbf{a}^\mathsf{T}$

15. $\mathbf{Aa}$, $\mathbf{Aa}^\mathsf{T}$, $(\mathbf{Ab})^\mathsf{T}$, $\mathbf{b}^\mathsf{T}\mathbf{A}^\mathsf{T}$

16. $\mathbf{BC}$, $\mathbf{BC}^\mathsf{T}$, $\mathbf{Bb}$, $\mathbf{b}^\mathsf{T}\mathbf{B}$

17. $\mathbf{ABC}$, $\mathbf{ABa}$, $\mathbf{ABb}$, $\mathbf{Ca}^\mathsf{T}$

18. $\mathbf{ab}$, $\mathbf{ba}$, $\mathbf{aA}$, $\mathbf{Bb}$

19. $1.5\mathbf{a} + 3.0\mathbf{b}$, $1.5\mathbf{a}^\mathsf{T} + 3.0\mathbf{b}$, $(\mathbf{A} - \mathbf{B})\mathbf{b}$, $\mathbf{Ab} - \mathbf{Bb}$

20. $\mathbf{b}^\mathsf{T}\mathbf{Ab}$, $\mathbf{aBa}^\mathsf{T}$, $\mathbf{aCC}^\mathsf{T}$, $\mathbf{C}^\mathsf{T}\mathbf{ba}$

21. General rules. Prove (2) for 2×2 matrices $\mathbf{A} = [a_{jk}]$, $\mathbf{B} = [b_{jk}]$, $\mathbf{C} = [c_{jk}]$, and a general scalar.

22. Product. Write $\mathbf{AB}$ in Prob. 11 in terms of row and column vectors.

23. Product. Calculate $\mathbf{AB}$ in Prob. 11 columnwise. See Example 1.

24. Commutativity. Find all 2×2 matrices $\mathbf{A} = [a_{jk}]$ that commute with $\mathbf{B} = [b_{jk}]$, where $b_{jk} = j + k$.

25. TEAM PROJECT. Symmetric and Skew-Symmetric Matrices. These matrices occur quite frequently in applications, so it is worthwhile to study some of their most important properties.

(a) Verify the claims in (11) that $a_{kj} = a_{jk}$ for a symmetric matrix, and $a_{kj} = -a_{jk}$ for a skew-symmetric matrix. Give examples.

(b) Show that for every square matrix $\mathbf{C}$ the matrix $\mathbf{C} + \mathbf{C}^\mathsf{T}$ is symmetric and $\mathbf{C} - \mathbf{C}^\mathsf{T}$ is skew-symmetric. Write $\mathbf{C}$ in the form $\mathbf{C} = \mathbf{S} + \mathbf{T}$, where $\mathbf{S}$ is symmetric and $\mathbf{T}$ is skew-symmetric and find $\mathbf{S}$ and $\mathbf{T}$ in terms of $\mathbf{C}$. Represent $\mathbf{A}$ and $\mathbf{B}$ in Probs. 11–20 in this form.

(c) A **linear combination** of matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}, \cdots, \mathbf{M}$ of the same size is an expression of the form

$$a\mathbf{A} + b\mathbf{B} + c\mathbf{C} + \cdots + m\mathbf{M}, \tag{14}$$

where $a, \cdots, m$ are any scalars. Show that if these matrices are square and symmetric, so is (14); similarly, if they are skew-symmetric, so is (14).

(d) Show that $\mathbf{AB}$ with symmetric $\mathbf{A}$ and $\mathbf{B}$ is symmetric if and only if $\mathbf{A}$ and $\mathbf{B}$ **commute**, that is, $\mathbf{AB} = \mathbf{BA}$.

(e) Under what condition is the product of skew-symmetric matrices skew-symmetric?

26–30 FURTHER APPLICATIONS

26. Production. In a production process, let N mean "no trouble" and T "trouble." Let the transition probabilities from one day to the next be 0.8 for $N \to N$, hence 0.2 for $N \to T$, and 0.5 for $T \to N$, hence 0.5 for $T \to T$.

If today there is no trouble, what is the probability of N two days after today? Three days after today?

27. CAS Experiment. Markov Process. Write a program for a Markov process. Use it to calculate further steps in Example 13 of the text. Experiment with other stochastic 3×3 matrices, also using different starting values.

28. Concert subscription. In a community of 100,000 adults, subscribers to a concert series tend to renew their subscription with probability 90% and persons presently not subscribing will subscribe for the next season with probability 0.2%. If the present number of subscribers is 1200, can one predict an increase, decrease, or no change over each of the next three seasons?

29. Profit vector. Two factory outlets F_1 and F_2 in New York and Los Angeles sell sofas (S), chairs (C), and tables (T) with a profit of \$35, \$62, and \$30, respectively. Let the sales in a certain week be given by the matrix

$$\mathbf{A} = \begin{array}{c} \begin{matrix} \text{S} & \text{C} & \text{T} \end{matrix} \\ \begin{bmatrix} 400 & 60 & 240 \\ 100 & 120 & 500 \end{bmatrix} \end{array} \begin{matrix} F_1 \\ F_2 \end{matrix}$$

Introduce a "profit vector" $\mathbf{p}$ such that the components of $\mathbf{v} = \mathbf{Ap}$ give the total profits of F_1 and F_2.

30. TEAM PROJECT. Special Linear Transformations. Rotations have various applications. We show in this project how they can be handled by matrices.

(a) Rotation in the plane. Show that the linear transformation $\mathbf{y} = \mathbf{Ax}$ with

$$\mathbf{A} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

is a counterclockwise rotation of the Cartesian x_1x_2-coordinate system in the plane about the origin, where θ is the angle of rotation.

(b) Rotation through $n\theta$. Show that in (a)

$$\mathbf{A}^n = \begin{bmatrix} \cos n\theta & -\sin n\theta \\ \sin n\theta & \cos n\theta \end{bmatrix}.$$

Is this plausible? Explain this in words.

(c) Addition formulas for cosine and sine. By geometry we should have

$$\begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}\begin{bmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{bmatrix} = \begin{bmatrix} \cos(\alpha+\beta) & -\sin(\alpha+\beta) \\ \sin(\alpha+\beta) & \cos(\alpha+\beta) \end{bmatrix}.$$

Derive from this the addition formulas (6) in App. A3.1.

(d) Computer graphics. To visualize a three-dimensional object with plane faces (e.g., a cube), we may store the position vectors of the vertices with respect to a suitable $x_1x_2x_3$-coordinate system (and a list of the connecting edges) and then obtain a two-dimensional image on a video screen by projecting the object onto a coordinate plane, for instance, onto the x_1x_2-plane by setting $x_3 = 0$. To change the appearance of the image, we can impose a linear transformation on the position vectors stored. Show that a diagonal matrix $\mathbf{D}$ with main diagonal entries 3, 1, $\frac{1}{2}$ gives from an $\mathbf{x} = [x_j]$ the new position vector $\mathbf{y} = \mathbf{D}\mathbf{x}$, where $y_1 = 3x_1$ (stretch in the x_1-direction by a factor 3), $y_2 = x_2$ (unchanged), $y_3 = \frac{1}{2}x_3$ (contraction in the x_3-direction). What effect would a scalar matrix have?

(e) Rotations in space. Explain $\mathbf{y} = \mathbf{A}\mathbf{x}$ geometrically when $\mathbf{A}$ is one of the three matrices

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix},$$

$$\begin{bmatrix} \cos\varphi & 0 & -\sin\varphi \\ 0 & 1 & 0 \\ \sin\varphi & 0 & \cos\varphi \end{bmatrix}, \quad \begin{bmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

What effect would these transformations have in situations such as that described in (d)?

7.3 Linear Systems of Equations. Gauss Elimination

We now come to one of the most important use of matrices, that is, using matrices to solve systems of linear equations. We showed informally, in Example 1 of Sec. 7.1, how to represent the information contained in a system of linear equations by a matrix, called the augmented matrix. This matrix will then be used in solving the linear system of equations. Our approach to solving linear systems is called the Gauss elimination method. Since this method is so fundamental to linear algebra, the student should be alert.

A shorter term for systems of linear equations is just **linear systems**. Linear systems model many applications in engineering, economics, statistics, and many other areas. Electrical networks, traffic flow, and commodity markets may serve as specific examples of applications.

Linear System, Coefficient Matrix, Augmented Matrix

A **linear system of m equations in n unknowns** $x_1, \cdots, x_n$ is a set of equations of the form

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n &= b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots&\cdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m. \end{aligned} \tag{1}$$

The system is called *linear* because each variable x_j appears in the first power only, just as in the equation of a straight line. $a_{11}, \cdots, a_{mn}$ are given numbers, called the **coefficients** of the system. $b_1, \cdots, b_m$ on the right are also given numbers. If all the b_j are zero, then (1) is called a **homogeneous system**. If at least one b_j is not zero, then (1) is called a **nonhomogeneous system**.

A **solution** of (1) is a set of numbers $x_1, \cdots, x_n$ that satisfies all the m equations. A **solution vector** of (1) is a vector $\mathbf{x}$ whose components form a solution of (1). If the system (1) is homogeneous, it always has at least the **trivial solution** $x_1 = 0, \cdots, x_n = 0$.

Matrix Form of the Linear System (1). From the definition of matrix multiplication we see that the m equations of (1) may be written as a single vector equation

$$\mathbf{Ax} = \mathbf{b} \tag{2}$$

where the **coefficient matrix** $\mathbf{A} = [a_{jk}]$ is the $m \times n$ matrix

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad \text{and} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

are column vectors. We assume that the coefficients a_{jk} are not all zero, so that $\mathbf{A}$ is not a zero matrix. Note that $\mathbf{x}$ has n components, whereas $\mathbf{b}$ has m components. The matrix

$$\tilde{\mathbf{A}} = \left[\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdots & \cdot & \cdot \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right]$$

is called the **augmented matrix** of the system (1). The dashed vertical line could be omitted, as we shall do later. It is merely a reminder that the last column of $\tilde{\mathbf{A}}$ did not come from matrix $\mathbf{A}$ but came from vector $\mathbf{b}$. Thus, we *augmented* the matrix $\mathbf{A}$.

*Note that **the augmented matrix** $\tilde{\mathbf{A}}$ **determines the system (1) completely*** because it contains all the given numbers appearing in (1).

EXAMPLE 1 **Geometric Interpretation. Existence and Uniqueness of Solutions**

If $m = n = 2$, we have two equations in two unknowns x_1, x_2

$$a_{11}x_1 + a_{12}x_2 = b_1$$
$$a_{21}x_1 + a_{22}x_2 = b_2.$$

If we interpret x_1, x_2 as coordinates in the x_1x_2-plane, then each of the two equations represents a straight line, and (x_1, x_2) is a solution if and only if the point P with coordinates x_1, x_2 lies on both lines. Hence there are three possible cases (see Fig. 158 on next page):

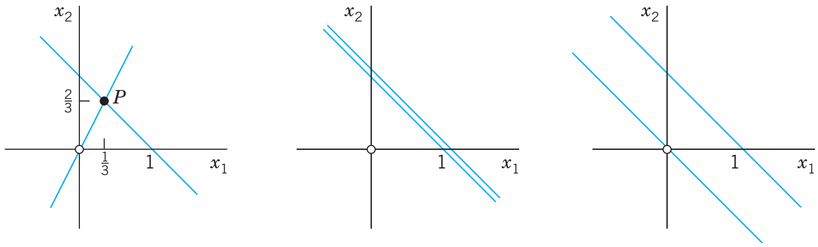
(*a*) Precisely one solution if the lines intersect

(*b*) Infinitely many solutions if the lines coincide

(*c*) No solution if the lines are parallel

For instance,

$x_1 + x_2 = 1$	$x_1 + x_2 = 1$	$x_1 + x_2 = 1$
$2x_1 - x_2 = 0$	$2x_1 + 2x_2 = 2$	$x_1 + x_2 = 0$
Case (a)	Case (b)	Case (c)



If the system is homogenous, Case (c) cannot happen, because then those two straight lines pass through the origin, whose coordinates (0, 0) constitute the trivial solution. Similarly, our present discussion can be extended from two equations in two unknowns to three equations in three unknowns. We give the geometric interpretation of three possible cases concerning solutions in Fig. 158. Instead of straight lines we have planes and the solution depends on the positioning of these planes in space relative to each other. The student may wish to come up with some specific examples. ■

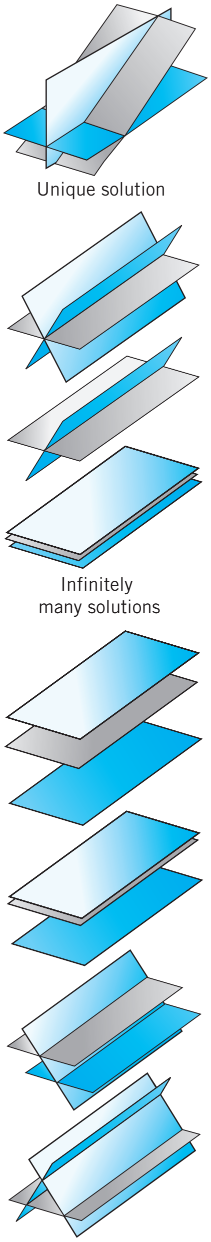


Fig. 158. Three equations in three unknowns interpreted as planes in space

Our simple example illustrated that a system (1) may have no solution. This leads to such questions as: Does a given system (1) have a solution? Under what conditions does it have precisely one solution? If it has more than one solution, how can we characterize the set of all solutions? We shall consider such questions in Sec. 7.5.

First, however, let us discuss an important systematic method for solving linear systems.

Gauss Elimination and Back Substitution

The Gauss elimination method can be motivated as follows. Consider a linear system that is in ***triangular form*** (in full, *upper* triangular form) such as

$$\begin{aligned} 2x_1 + 5x_2 &= 2 \\ 13x_2 &= -26 \end{aligned}$$

(*Triangular* means that all the nonzero entries of the corresponding coefficient matrix lie above the diagonal and form an upside-down 90° triangle.) Then we can solve the system by **back substitution**, that is, we solve the last equation for the variable, $x_2 = -26/13 = -2$, and then work backward, substituting $x_2 = -2$ into the first equation and solving it for x_1, obtaining $x_1 = \frac{1}{2}(2 - 5x_2) = \frac{1}{2}(2 - 5 \cdot (-2)) = 6$. This gives us the idea of first reducing a general system to triangular form. For instance, let the given system be

$$\begin{aligned} 2x_1 + 5x_2 &= 2 \\ -4x_1 + 3x_2 &= -30. \end{aligned} \qquad \text{Its augmented matrix is} \qquad \begin{bmatrix} 2 & 5 & 2 \\ -4 & 3 & -30 \end{bmatrix}.$$

We leave the first equation as it is. We eliminate x_1 from the second equation, to get a triangular system. For this we add twice the first equation to the second, and we do the same operation on the ***rows*** of the augmented matrix. This gives $-4x_1 + 4x_1 + 3x_2 + 10x_2 = -30 + 2 \cdot 2$, that is,

$$\begin{aligned} 2x_1 + 5x_2 &= \quad 2 \\ 13x_2 &= -26 \end{aligned} \qquad \begin{matrix} \\ \text{Row 2 + 2 Row 1} \end{matrix} \qquad \begin{bmatrix} 2 & 5 & 2 \\ 0 & 13 & -26 \end{bmatrix}$$

where Row 2 + 2 Row 1 means "Add twice Row 1 to Row 2" in the original matrix. This is the **Gauss elimination** (for 2 equations in 2 unknowns) giving the triangular form, from which back substitution now yields $x_2 = -2$ and $x_1 = 6$, as before.

Since a linear system is completely determined by its augmented matrix, ***Gauss elimination can be done by merely considering the matrices***, as we have just indicated. We do this again in the next example, emphasizing the matrices by writing them first and the equations behind them, just as a help in order not to lose track.

EXAMPLE 2 Gauss Elimination. Electrical Network

Solve the linear system

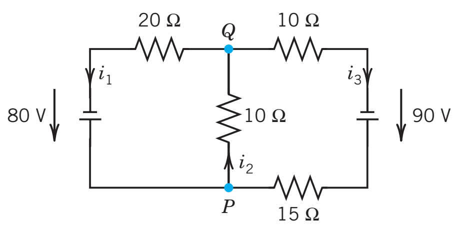
$$\begin{aligned} x_1 - \quad x_2 + \quad x_3 &= 0 \\ -x_1 + \quad x_2 - \quad x_3 &= 0 \\ 10x_2 + 25x_3 &= 90 \\ 20x_1 + 10x_2 \qquad &= 80. \end{aligned}$$

***Derivation from the circuit in Fig.* 159 (*Optional*).** This is the system for the unknown currents $x_1 = i_1$, $x_2 = i_2$, $x_3 = i_3$ in the electrical network in Fig. 159. To obtain it, we label the currents as shown, choosing directions arbitrarily; if a current will come out negative, this will simply mean that the current flows against the direction of our arrow. The current entering each battery will be the same as the current leaving it. The equations for the currents result from Kirchhoff's laws:

Kirchhoff's Current Law (KCL). *At any point of a circuit, the sum of the inflowing currents equals the sum of the outflowing currents.*

Kirchhoff's Voltage Law (KVL). *In any closed loop, the sum of all voltage drops equals the impressed electromotive force.*

Node P gives the first equation, node Q the second, the right loop the third, and the left loop the fourth, as indicated in the figure.



Node P:	$i_1 - i_2 + i_3 = 0$
Node Q:	$-i_1 + i_2 - i_3 = 0$
Right loop:	$10i_2 + 25i_3 = 90$
Left loop:	$20i_1 + 10i_2 = 80$

Fig. 159. Network in Example 2 and equations relating the currents

Solution by Gauss Elimination. This system could be solved rather quickly by noticing its particular form. But this is not the point. The point is that the Gauss elimination is systematic and will work in general, also for large systems. We apply it to our system and then do back substitution. As indicated, let us write the augmented matrix of the system first and then the system itself:

$$\begin{array}{r}
\text{Pivot 1} \longrightarrow \\
\\
\text{Eliminate} \longrightarrow \\
\\
\end{array}
\overset{\text{Augmented Matrix } \tilde{\mathbf{A}}}{\left[\begin{array}{rrr|r}
1 & -1 & 1 & 0 \\
-1 & 1 & -1 & 0 \\
0 & 10 & 25 & 90 \\
20 & 10 & 0 & 80
\end{array}\right]}
\qquad
\begin{array}{r}
\text{Pivot 1} \longrightarrow \\
\\
\text{Eliminate} \longrightarrow \\
\\
\end{array}
\overset{\text{Equations}}{\begin{array}{rrrrrcl}
x_1 & - & x_2 & + & x_3 & = & 0 \\
-x_1 & + & x_2 & - & x_3 & = & 0 \\
 & & 10x_2 & + & 25x_3 & = & 90 \\
20x_1 & + & 10x_2 & & & = & 80.
\end{array}}$$

Step 1. Elimination of x_1

Call the first row of **A** the **pivot row** and the first equation the **pivot equation**. Call the coefficient 1 of its x_1-term the **pivot** in this step. Use this equation to eliminate x_1 (get rid of x_1) in the other equations. For this, do:

Add 1 times the pivot equation to the second equation.

Add -20 times the pivot equation to the fourth equation.

This corresponds to **row operations** on the augmented matrix as indicated in BLUE behind the ***new matrix*** in (3). So the operations are performed on the ***preceding matrix***. The result is

$$\left[\begin{array}{rrr|r}
1 & -1 & 1 & 0 \\
0 & 0 & 0 & 0 \\
0 & 10 & 25 & 90 \\
0 & 30 & -20 & 80
\end{array}\right]
\begin{array}{l}
\\
\text{Row 2 + Row 1} \\
\\
\text{Row 4} - \text{20 Row 1}
\end{array}
\qquad
\begin{array}{rrrrrcl}
x_1 & - & x_2 & + & x_3 & = & 0 \\
 & & & & 0 & = & 0 \\
 & & 10x_2 & + & 25x_3 & = & 90 \\
 & & 30x_2 & - & 20x_3 & = & 80.
\end{array}
\tag{3}$$

Step 2. Elimination of x_2

The first equation remains as it is. We want the new second equation to serve as the next pivot equation. But since it has no x_2-term (in fact, it is $0 = 0$), we must first change the order of the equations and the corresponding rows of the new matrix. We put $0 = 0$ at the end and move the third equation and the fourth equation one place up. This is called **partial pivoting** (as opposed to the rarely used *total pivoting*, in which the order of the unknowns is also changed). It gives

$$\begin{array}{r}
\\
\text{Pivot 10} \longrightarrow \\
\text{Eliminate 30} \longrightarrow \\
\\
\end{array}
\left[\begin{array}{rrr|r}
1 & -1 & 1 & 0 \\
0 & 10 & 25 & 90 \\
0 & 30 & -20 & 80 \\
0 & 0 & 0 & 0
\end{array}\right]
\qquad
\begin{array}{r}
\\
\text{Pivot 10} \longrightarrow \\
\text{Eliminate } 30x_2 \longrightarrow \\
\\
\end{array}
\begin{array}{rrrrrcl}
x_1 & - & x_2 & + & x_3 & = & 0 \\
 & & 10x_2 & + & 25x_3 & = & 90 \\
 & & 30x_2 & - & 20x_3 & = & 80 \\
 & & & & 0 & = & 0.
\end{array}$$

To eliminate x_2, do:

Add -3 times the pivot equation to the third equation.

The result is

$$\left[\begin{array}{rrr|r}
1 & -1 & 1 & 0 \\
0 & 10 & 25 & 90 \\
0 & 0 & -95 & -190 \\
0 & 0 & 0 & 0
\end{array}\right]
\begin{array}{l}
\\
\\
\text{Row 3} - \text{3 Row 2} \\
\\
\end{array}
\qquad
\begin{array}{rrrrrcr}
x_1 & - & x_2 & + & x_3 & = & 0 \\
 & & 10x_2 & + & 25x_3 & = & 90 \\
 & & & - & 95x_3 & = & -190 \\
 & & & & 0 & = & 0.
\end{array}
\tag{4}$$

Back Substitution. ***Determination of*** x_3, x_2, x_1 ***(in this order)***

Working backward from the last to the first equation of this "triangular" system (4), we can now readily find x_3, then x_2, and then x_1:

$$\begin{array}{rrrrrcr}
 & & & - & 95x_3 & = & -190 \\
 & & 10x_2 & + & 25x_3 & = & 90 \\
x_1 & - & x_2 & + & x_3 & = & 0
\end{array}
\qquad
\begin{array}{r}
x_3 = i_3 = 2\ [\text{A}] \\
x_2 = \tfrac{1}{10}(90 - 25x_3) = i_2 = 4\ [\text{A}] \\
x_1 = x_2 - x_3 = i_1 = 2\ [\text{A}]
\end{array}$$

where A stands for "amperes." This is the answer to our problem. The solution is unique. ■

Elementary Row Operations. Row-Equivalent Systems

Example 2 illustrates the operations of the Gauss elimination. These are the first two of three operations, which are called

Elementary Row Operations for Matrices:

Interchange of two rows

Addition of a constant multiple of one row to another row

*Multiplication of a row by a **nonzero** constant c*

CAUTION! These operations are for rows, ***not for columns***! They correspond to the following

Elementary Operations for Equations:

Interchange of two equations

Addition of a constant multiple of one equation to another equation

*Multiplication of an equation by a **nonzero** constant c*

Clearly, the interchange of two equations does not alter the solution set. Neither does their addition because we can undo it by a corresponding subtraction. Similarly for their multiplication, which we can undo by multiplying the new equation by $1/c$ (since $c \neq 0$), producing the original equation.

We now call a linear system S_1 **row-equivalent** to a linear system S_2 if S_1 can be obtained from S_2 by (finitely many!) row operations. This justifies Gauss elimination and establishes the following result.

THEOREM 1

Row-Equivalent Systems

Row-equivalent linear systems have the same set of solutions.

Because of this theorem, systems having the same solution sets are often called *equivalent systems*. But note well that we are dealing with ***row operations***. No column operations on the augmented matrix are permitted in this context because they would generally alter the solution set.

A linear system (1) is called **overdetermined** if it has more equations than unknowns, as in Example 2, **determined** if $m = n$, as in Example 1, and **underdetermined** if it has fewer equations than unknowns.

Furthermore, a system (1) is called **consistent** if it has at least one solution (thus, one solution or infinitely many solutions), but **inconsistent** if it has no solutions at all, as $x_1 + x_2 = 1, x_1 + x_2 = 0$ in Example 1, Case (c).

Gauss Elimination: The Three Possible Cases of Systems

We have seen, in Example 2, that Gauss elimination can solve linear systems that have a unique solution. This leaves us to apply Gauss elimination to a system with infinitely many solutions (in Example 3) and one with no solution (in Example 4).

EXAMPLE 3 Gauss Elimination if Infinitely Many Solutions Exist

Solve the following linear system of three equations in four unknowns whose augmented matrix is

$$
(5) \quad \left[\begin{array}{cccc|c} 3.0 & 2.0 & 2.0 & -5.0 & 8.0 \\ 0.6 & 1.5 & 1.5 & -5.4 & 2.7 \\ 1.2 & -0.3 & -0.3 & 2.4 & 2.1 \end{array}\right]. \qquad \text{Thus,} \qquad \begin{aligned} 3.0x_1 + 2.0x_2 + 2.0x_3 - 5.0x_4 &= 8.0 \\ \boxed{0.6x_1} + 1.5x_2 + 1.5x_3 - 5.4x_4 &= 2.7 \\ \boxed{1.2x_1} - 0.3x_2 - 0.3x_3 + 2.4x_4 &= 2.1. \end{aligned}
$$

Solution. As in the previous example, we circle pivots and box terms of equations and corresponding entries to be eliminated. We indicate the operations in terms of equations and operate on both equations and matrices.

Step 1. Elimination of x_1 from the second and third equations by adding

$$
-0.6/3.0 = -0.2 \text{ times the first equation to the second equation,}
$$

$$
-1.2/3.0 = -0.4 \text{ times the first equation to the third equation.}
$$

This gives the following, in which the pivot of the next step is circled.

$$
(6) \quad \left[\begin{array}{cccc|c} 3.0 & 2.0 & 2.0 & -5.0 & 8.0 \\ 0 & 1.1 & 1.1 & -4.4 & 1.1 \\ 0 & -1.1 & -1.1 & 4.4 & -1.1 \end{array}\right] \begin{array}{l} \\ \text{Row 2} - 0.2 \text{ Row 1} \\ \text{Row 3} - 0.4 \text{ Row 1} \end{array} \qquad \begin{aligned} 3.0x_1 + 2.0x_2 + 2.0x_3 - 5.0x_4 &= 8.0 \\ 1.1x_2 + 1.1x_3 - 4.4x_4 &= 1.1 \\ \boxed{-1.1x_2} - 1.1x_3 + 4.4x_4 &= -1.1. \end{aligned}
$$

Step 2. Elimination of x_2 from the third equation of (6) by adding

$$
1.1/1.1 = 1 \text{ times the second equation to the third equation.}
$$

This gives

$$
(7) \quad \left[\begin{array}{cccc|c} 3.0 & 2.0 & 2.0 & -5.0 & 8.0 \\ 0 & 1.1 & 1.1 & -4.4 & 1.1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \begin{array}{l} \\ \\ \text{Row 3} + \text{Row 2} \end{array} \qquad \begin{aligned} 3.0x_1 + 2.0x_2 + 2.0x_3 - 5.0x_4 &= 8.0 \\ 1.1x_2 + 1.1x_3 - 4.4x_4 &= 1.1 \\ 0 &= 0. \end{aligned}
$$

Back Substitution. From the second equation, $x_2 = 1 - x_3 + 4x_4$. From this and the first equation, $x_1 = 2 - x_4$. Since x_3 and x_4 remain arbitrary, we have infinitely many solutions. If we choose a value of x_3 and a value of x_4, then the corresponding values of x_1 and x_2 are uniquely determined.

On Notation. If unknowns remain arbitrary, it is also customary to denote them by other letters $t_1, t_2, \cdots$. In this example we may thus write $x_1 = 2 - x_4 = 2 - t_2$, $x_2 = 1 - x_3 + 4x_4 = 1 - t_1 + 4t_2$, $x_3 = t_1$ (first arbitrary unknown), $x_4 = t_2$ (second arbitrary unknown). ■

EXAMPLE 4 Gauss Elimination if no Solution Exists

What will happen if we apply the Gauss elimination to a linear system that has no solution? The answer is that in this case the method will show this fact by producing a contradiction. For instance, consider

$$
\left[\begin{array}{ccc|c} 3 & 2 & 1 & 3 \\ 2 & 1 & 1 & 0 \\ 6 & 2 & 4 & 6 \end{array}\right] \qquad \begin{aligned} 3x_1 + 2x_2 + x_3 &= 3 \\ \boxed{2x_1} + x_2 + x_3 &= 0 \\ \boxed{6x_1} + 2x_2 + 4x_3 &= 6. \end{aligned}
$$

Step 1. Elimination of x_1 from the second and third equations by adding

$$
-\tfrac{2}{3} \text{ times the first equation to the second equation,}
$$

$$
-\tfrac{6}{3} = -2 \text{ times the first equation to the third equation.}
$$

This gives

$$\left[\begin{array}{ccc|c} 3 & 2 & 1 & 3 \\ 0 & -\frac{1}{3} & \frac{1}{3} & -2 \\ 0 & -2 & 2 & 0 \end{array}\right] \quad \begin{array}{l} \\ \text{Row 2} - \frac{2}{3}\,\text{Row 1} \\ \text{Row 3} - 2\,\text{Row 1} \end{array} \qquad \begin{aligned} 3x_1 + 2x_2 + x_3 &= 3 \\ -\tfrac{1}{3}x_2 + \tfrac{1}{3}x_3 &= -2 \\ \boxed{-\,2x_2} + 2x_3 &= 0. \end{aligned}$$

Step 2. Elimination of x_2 from the third equation gives

$$\left[\begin{array}{ccc|c} 3 & 2 & 1 & 3 \\ 0 & -\frac{1}{3} & \frac{1}{3} & -2 \\ 0 & 0 & 0 & 12 \end{array}\right] \quad \begin{array}{l} \\ \\ \text{Row 3} - 6\,\text{Row 2} \end{array} \qquad \begin{aligned} 3x_1 + 2x_2 + x_3 &= 3 \\ -\tfrac{1}{3}x_2 + \tfrac{1}{3}x_3 &= -\,2 \\ 0 &= 12. \end{aligned}$$

The false statement $0 = 12$ shows that the system has no solution. ■

Row Echelon Form and Information From It

At the end of the Gauss elimination the form of the coefficient matrix, the augmented matrix, and the system itself are called the **row echelon form**. In it, rows of zeros, if present, are the last rows, and, in each nonzero row, the leftmost nonzero entry is farther to the right than in the previous row. For instance, in Example 4 the coefficient matrix and its augmented in row echelon form are

$$\begin{bmatrix} 3 & 2 & 1 \\ 0 & -\frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \left[\begin{array}{ccc|c} 3 & 2 & 1 & 3 \\ 0 & -\frac{1}{3} & \frac{1}{3} & -2 \\ 0 & 0 & 0 & 12 \end{array}\right]. \tag{8}$$

Note that we do not require that the leftmost nonzero entries be 1 since this would have no theoretic or numeric advantage. (The so-called *reduced echelon form*, in which those entries *are* 1, will be discussed in Sec. 7.8.)

The original system of m equations in n unknowns has augmented matrix $[\mathbf{A} \,|\, \mathbf{b}]$. This is to be row reduced to matrix $[\mathbf{R} \,|\, \mathbf{f}]$. The two systems $\mathbf{Ax} = \mathbf{b}$ and $\mathbf{Rx} = \mathbf{f}$ are equivalent: if either one has a solution, so does the other, and the solutions are identical.

At the end of the Gauss elimination (before the back substitution), the row echelon form of the augmented matrix will be

$$\left[\begin{array}{ccccc|c} r_{11} & r_{12} & \cdots\cdots & & r_{1n} & f_1 \\ & r_{22} & \cdots\cdots & & r_{2n} & f_2 \\ & & \ddots & & \vdots & \vdots \\ & & & r_{rr} \;\cdots & r_{rn} & f_r \\ & & & & & f_{r+1} \\ & & & & & \vdots \\ & & & & & f_m \end{array}\right]. \tag{9}$$

Here, $r \leqq m$, $r_{11} \neq 0$, and all entries in the blue triangle and blue rectangle are zero.

The number of nonzero rows, r, in the row-reduced coefficient matrix $\mathbf{R}$ is called the **rank of R** and also the **rank of A**. Here is the method for determining whether $\mathbf{Ax} = \mathbf{b}$ has solutions and what they are:

(a) No solution. If r is less than m (meaning that $\mathbf{R}$ actually has at least one row of all 0s) *and* at least one of the numbers $f_{r+1}, f_{r+2}, \cdots, f_m$ is not zero, then the system

$\mathbf{Rx} = \mathbf{f}$ is inconsistent: No solution is possible. Therefore the system $\mathbf{Ax} = \mathbf{b}$ is inconsistent as well. See Example 4, where $r = 2 < m = 3$ and $f_{r+1} = f_3 = 12$.

If the system is consistent (either $r = m$, or $r < m$ *and* all the numbers $f_{r+1}, f_{r+2}, \cdots, f_m$ are zero), then there are solutions.

(b) Unique solution. If the system is consistent and $r = n$, there is exactly one solution, which can be found by back substitution. See Example 2, where $r = n = 3$ and $m = 4$.

(c) Infinitely many solutions. To obtain any of these solutions, choose values of $x_{r+1}, \cdots, x_n$ arbitrarily. Then solve the rth equation for x_r (in terms of those arbitrary values), then the $(r - 1)$st equation for x_{r-1}, and so on up the line. See Example 3.

Orientation. Gauss elimination is reasonable in computing time and storage demand. We shall consider those aspects in Sec. 20.1 in the chapter on numeric linear algebra. Section 7.4 develops fundamental concepts of linear algebra such as linear independence and rank of a matrix. These in turn will be used in Sec. 7.5 to fully characterize the behavior of linear systems in terms of existence and uniqueness of solutions.

PROBLEM SET 7.3

1–14 GAUSS ELIMINATION

Solve the linear system given explicitly or by its augmented matrix. Show details.

1. $$\begin{aligned} 4x - 6y &= -11 \\ -3x + 8y &= 10 \end{aligned}$$

2. $$\begin{bmatrix} 3.0 & -0.5 & 0.6 \\ 1.5 & 4.5 & 6.0 \end{bmatrix}$$

3. $$\begin{aligned} x + y - z &= 9 \\ 8y + 6z &= -6 \\ -2x + 4y - 6z &= 40 \end{aligned}$$

4. $$\begin{bmatrix} 4 & 1 & 0 & 4 \\ 5 & -3 & 1 & 2 \\ -9 & 2 & -1 & 5 \end{bmatrix}$$

5. $$\begin{bmatrix} 13 & 12 & -6 \\ -4 & 7 & -73 \\ 11 & -13 & 157 \end{bmatrix}$$

6. $$\begin{bmatrix} 4 & -8 & 3 & 16 \\ -1 & 2 & -5 & -21 \\ 3 & -6 & 1 & 7 \end{bmatrix}$$

7. $$\begin{bmatrix} 2 & 4 & 1 & 0 \\ -1 & 1 & -2 & 0 \\ 4 & 0 & 6 & 0 \end{bmatrix}$$

8. $$\begin{aligned} 4y + 3z &= 8 \\ 2x - z &= 2 \\ 3x + 2y &= 5 \end{aligned}$$

9. $$\begin{aligned} -2y - 2z &= -8 \\ 3x + 4y - 5z &= 13 \end{aligned}$$

10. $$\begin{bmatrix} 5 & -7 & 3 & 17 \\ -15 & 21 & -9 & 50 \end{bmatrix}$$

11. $$\begin{bmatrix} 0 & 5 & 5 & -10 & 0 \\ 2 & -3 & -3 & 6 & 2 \\ 4 & 1 & 1 & -2 & 4 \end{bmatrix}$$

12. $$\begin{bmatrix} 2 & -2 & 4 & 0 & 0 \\ -3 & 3 & -6 & 5 & 15 \\ 1 & -1 & 2 & 0 & 0 \end{bmatrix}$$

13. $$\begin{aligned} 10x + 4y - 2z &= -4 \\ -3w - 17x + y + 2z &= 2 \\ w + x + y &= 6 \\ 8w - 34x + 16y - 10z &= 4 \end{aligned}$$

14. $$\begin{bmatrix} 2 & 3 & 1 & -11 & 1 \\ 5 & -2 & 5 & -4 & 5 \\ 1 & -1 & 3 & -3 & 3 \\ 3 & 4 & -7 & 2 & -7 \end{bmatrix}$$

15. Equivalence relation. By definition, an *equivalence relation* on a set is a relation satisfying three conditions: (named as indicated)

(i) Each element A of the set is equivalent to itself (*Reflexivity*).

(ii) If A is equivalent to B, then B is equivalent to A (*Symmetry*).

(iii) If A is equivalent to B and B is equivalent to C, then A is equivalent to C (*Transitivity*).

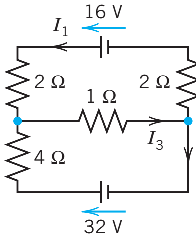
Show that row equivalence of matrices satisfies these three conditions. *Hint.* Show that for each of the three elementary row operations these conditions hold.

16. CAS PROJECT. Gauss Elimination and Back Substitution. Write a program for Gauss elimination and back substitution **(a)** that does not include pivoting and **(b)** that does include pivoting. Apply the programs to Probs. 11–14 and to some larger systems of your choice.

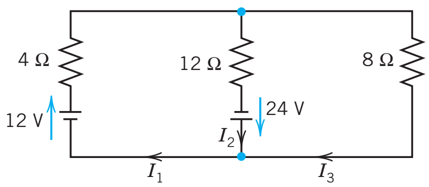
17–21 MODELS OF NETWORKS

In Probs. 17–19, using Kirchhoff's laws (see Example 2) and showing the details, find the currents:

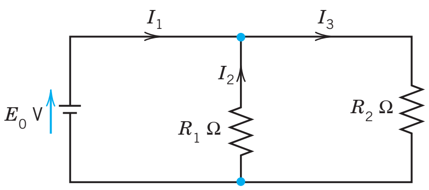
17.



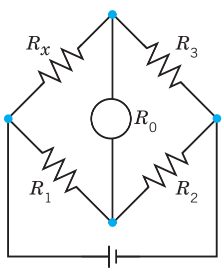
18.



19.

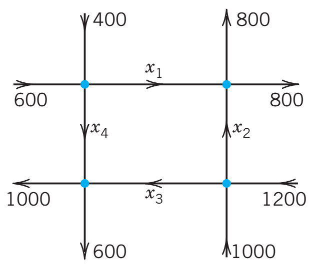


20. Wheatstone bridge. Show that if $R_x/R_3 = R_1/R_2$ in the figure, then $I = 0$. (R_0 is the resistance of the instrument by which I is measured.) This bridge is a method for determining R_x. R_1, R_2, R_3 are known. R_3 is variable. To get R_x, make $I = 0$ by varying R_3. Then calculate $R_x = R_3R_1/R_2$.



Wheatstone bridge

Problem 20



Net of one-way streets

Problem 21

21. Traffic flow. Methods of electrical circuit analysis have applications to other fields. For instance, applying the analog of Kirchhoff's Current Law, find the traffic flow (cars per hour) in the net of one-way streets (in the directions indicated by the arrows) shown in the figure. Is the solution unique?

22. Models of markets. Determine the equilibrium solution ($D_1 = S_1, D_2 = S_2$) of the two-commodity market with linear model (D, S, P = demand, supply, price; index 1 = first commodity, index 2 = second commodity)

$$D_1 = 40 - 2P_1 - P_2, \qquad S_1 = 4P_1 - P_2 + 4,$$

$$D_2 = 5P_1 - 2P_2 + 16, \qquad S_2 = 3P_2 - 4.$$

23. Balancing a chemical equation $x_1C_3H_8 + x_2O_2 \rightarrow x_3CO_2 + x_4H_2O$ means finding integer x_1, x_2, x_3, x_4 such that the numbers of atoms of carbon (C), hydrogen (H), and oxygen (O) are the same on both sides of this reaction, in which propane C_3H_8 and O_2 give carbon dioxide and water. Find the smallest positive integers $x_1, \cdots, x_4$.

24. PROJECT. Elementary Matrices. The idea is that elementary operations can be accomplished by matrix multiplication. If $\mathbf{A}$ is an $m \times n$ matrix on which we want to do an elementary operation, then there is a matrix $\mathbf{E}$ such that $\mathbf{EA}$ is the new matrix after the operation. Such an $\mathbf{E}$ is called an **elementary matrix**. This idea can be helpful, for instance, in the design of algorithms. (*Computationally*, it is generally preferable to do row operations *directly*, rather than by multiplication by $\mathbf{E}$.)

(a) Show that the following are elementary matrices, for interchanging Rows 2 and 3, for adding -5 times the first row to the third, and for multiplying the fourth row by 8.

$$\mathbf{E}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$\mathbf{E}_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -5 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$\mathbf{E}_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 8 \end{bmatrix}.$$

Apply $\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3$ to a vector and to a 4×3 matrix of your choice. Find $\mathbf{B} = \mathbf{E}_3\mathbf{E}_2\mathbf{E}_1\mathbf{A}$, where $\mathbf{A} = [a_{jk}]$ is the general 4×2 matrix. Is $\mathbf{B}$ equal to $\mathbf{C} = \mathbf{E}_1\mathbf{E}_2\mathbf{E}_3\mathbf{A}$?

(b) Conclude that $\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3$ are obtained by doing the corresponding elementary operations on the 4×4 unit matrix. Prove that *if* $\mathbf{M}$ *is obtained from* $\mathbf{A}$ *by an elementary row operation, then*

$$\mathbf{M} = \mathbf{EA},$$

where $\mathbf{E}$ *is obtained from the* $n \times n$ *unit matrix* $\mathbf{I}_n$ *by the same row operation.*

7.4 Linear Independence. Rank of a Matrix. Vector Space

Since our next goal is to fully characterize the behavior of linear systems in terms of existence and uniqueness of solutions (Sec. 7.5), we have to introduce new fundamental linear algebraic concepts that will aid us in doing so. Foremost among these are ***linear independence*** and the ***rank of a matrix***. Keep in mind that these concepts are intimately linked with the important Gauss elimination method and how it works.

Linear Independence and Dependence of Vectors

Given any set of m vectors $\mathbf{a}_{(1)}, \cdots, \mathbf{a}_{(m)}$ (with the same number of components), a **linear combination** of these vectors is an expression of the form

$$c_1\mathbf{a}_{(1)} + c_2\mathbf{a}_{(2)} + \cdots + c_m\mathbf{a}_{(m)}$$

where $c_1, c_2, \cdots, c_m$ are any scalars. Now consider the equation

$$c_1\mathbf{a}_{(1)} + c_2\mathbf{a}_{(2)} + \cdots + c_m\mathbf{a}_{(m)} = \mathbf{0}. \tag{1}$$

Clearly, this vector equation (1) holds if we choose all c_j's zero, because then it becomes $\mathbf{0} = \mathbf{0}$. If this is the only m-tuple of scalars for which (1) holds, then our vectors $\mathbf{a}_{(1)}, \cdots, \mathbf{a}_{(m)}$ are said to form a *linearly independent set* or, more briefly, we call them **linearly independent**. Otherwise, if (1) also holds with scalars not all zero, we call these vectors **linearly dependent**. This means that we can express at least one of the vectors as a linear combination of the other vectors. For instance, if (1) holds with, say, $c_1 \neq 0$, we can solve (1) for $\mathbf{a}_{(1)}$:

$$\mathbf{a}_{(1)} = k_2\mathbf{a}_{(2)} + \cdots + k_m\mathbf{a}_{(m)} \qquad \text{where } k_j = -c_j/c_1.$$

(Some k_j's may be zero. Or even all of them, namely, if $\mathbf{a}_{(1)} = \mathbf{0}$.)

Why is linear independence important? Well, if a set of vectors is linearly dependent, then we can get rid of at least one or perhaps more of the vectors until we get a linearly independent set. This set is then the smallest "truly essential" set with which we can work. Thus, we cannot express any of the vectors, of this set, linearly in terms of the others.

EXAMPLE 1 **Linear Independence and Dependence**

The three vectors

$$\begin{aligned} \mathbf{a}_{(1)} &= [\ 3 \quad 0 \quad 2 \quad 2] \\ \mathbf{a}_{(2)} &= [-6 \quad 42 \quad 24 \quad 54] \\ \mathbf{a}_{(3)} &= [\ 21 \quad -21 \quad 0 \quad -15] \end{aligned}$$

are linearly dependent because

$$6\mathbf{a}_{(1)} - \tfrac{1}{2}\mathbf{a}_{(2)} - \mathbf{a}_{(3)} = \mathbf{0}.$$

Although this is easily checked by vector arithmetic (do it!), it is not so easy to discover. However, a systematic method for finding out about linear independence and dependence follows below.

The first two of the three vectors are linearly independent because $c_1\mathbf{a}_{(1)} + c_2\mathbf{a}_{(2)} = \mathbf{0}$ implies $c_2 = 0$ (from the second components) and then $c_1 = 0$ (from any other component of $\mathbf{a}_{(1)}$. ■

Rank of a Matrix

DEFINITION

The **rank** of a matrix **A** is the maximum number of linearly independent row vectors of **A**. It is denoted by rank **A**.

Our further discussion will show that the rank of a matrix is an important key concept for understanding general properties of matrices and linear systems of equations.

EXAMPLE 2 **Rank**

The matrix

$$\mathbf{A} = \begin{bmatrix} 3 & 0 & 2 & 2 \\ -6 & 42 & 24 & 54 \\ 21 & -21 & 0 & -15 \end{bmatrix} \tag{2}$$

has rank 2, because Example 1 shows that the first two row vectors are linearly independent, whereas all three row vectors are linearly dependent.

Note further that rank $\mathbf{A} = 0$ if and only if $\mathbf{A} = \mathbf{0}$. This follows directly from the definition. ■

We call a matrix $\mathbf{A}_1$ **row-equivalent** to a matrix $\mathbf{A}_2$ if $\mathbf{A}_1$ can be obtained from $\mathbf{A}_2$ by (finitely many!) elementary row operations.

Now the maximum number of linearly independent row vectors of a matrix does not change if we change the order of rows or multiply a row by a nonzero c or take a linear combination by adding a multiple of a row to another row. This shows that rank is **invariant** under elementary row operations:

THEOREM 1

Row-Equivalent Matrices

Row-equivalent matrices have the same rank.

Hence we can determine the rank of a matrix by reducing the matrix to row-echelon form, as was done in Sec. 7.3. Once the matrix is in row-echelon form, we count the number of nonzero rows, which is precisely the rank of the matrix.

EXAMPLE 3 **Determination of Rank**

For the matrix in Example 2 we obtain successively

$$\mathbf{A} = \begin{bmatrix} 3 & 0 & 2 & 2 \\ -6 & 42 & 24 & 54 \\ 21 & -21 & 0 & -15 \end{bmatrix} \quad \text{(given)}$$

$$\begin{bmatrix} 3 & 0 & 2 & 2 \\ 0 & 42 & 28 & 58 \\ 0 & -21 & -14 & -29 \end{bmatrix} \quad \begin{matrix} \\ \text{Row 2} + 2 \text{ Row 1} \\ \text{Row 3} - 7 \text{ Row 1} \end{matrix}$$

$$\begin{bmatrix} 3 & 0 & 2 & 2 \\ 0 & 42 & 28 & 58 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \begin{matrix} \\ \\ \text{Row 3} + \frac{1}{2} \text{ Row 2.} \end{matrix}$$

The last matrix is in row-echelon form and has two nonzero rows. Hence rank $\mathbf{A} = 2$, as before. ■

Examples 1–3 illustrate the following useful theorem (with $p = 3$, $n = 3$, and the rank of the matrix $= 2$).

THEOREM 2 **Linear Independence and Dependence of Vectors**

Consider p vectors that each have n components. Then these vectors are linearly independent if the matrix formed, with these vectors as row vectors, has rank p. However, these vectors are linearly dependent if that matrix has rank less than p.

Further important properties will result from the basic

THEOREM 3 **Rank in Terms of Column Vectors**

The rank r of a matrix $\mathbf{A}$ equals the maximum number of linearly independent ***column*** *vectors of $\mathbf{A}$.*

Hence $\mathbf{A}$ and its transpose $\mathbf{A}^{\mathsf{T}}$ have the same rank.

PROOF In this proof we write simply "rows" and "columns" for row and column vectors. Let $\mathbf{A}$ be an $m \times n$ matrix of rank $\mathbf{A} = r$. Then by definition of rank, $\mathbf{A}$ has r linearly independent rows which we denote by $\mathbf{v}_{(1)}, \cdots, \mathbf{v}_{(r)}$ (regardless of their position in $\mathbf{A}$), and all the rows $\mathbf{a}_{(1)}, \cdots, \mathbf{a}_{(m)}$ of $\mathbf{A}$ are linear combinations of those, say,

$$\begin{aligned} \mathbf{a}_{(1)} &= c_{11}\mathbf{v}_{(1)} + c_{12}\mathbf{v}_{(2)} + \cdots + c_{1r}\mathbf{v}_{(r)} \\ \mathbf{a}_{(2)} &= c_{21}\mathbf{v}_{(1)} + c_{22}\mathbf{v}_{(2)} + \cdots + c_{2r}\mathbf{v}_{(r)} \\ &\vdots \\ \mathbf{a}_{(m)} &= c_{m1}\mathbf{v}_{(1)} + c_{m2}\mathbf{v}_{(2)} + \cdots + c_{mr}\mathbf{v}_{(r)}. \end{aligned} \tag{3}$$

These are vector equations for rows. To switch to columns, we write (3) in terms of components as n such systems, with $k = 1, \cdots, n$,

$$
\begin{aligned}
a_{1k} &= c_{11}v_{1k} + c_{12}v_{2k} + \cdots + c_{1r}v_{rk} \\
a_{2k} &= c_{21}v_{1k} + c_{22}v_{2k} + \cdots + c_{2r}v_{rk} \\
&\;\;\vdots \\
a_{mk} &= c_{m1}v_{1k} + c_{m2}v_{2k} + \cdots + c_{mr}v_{rk}
\end{aligned}
\tag{4}
$$

and collect components in columns. Indeed, we can write (4) as

$$
\begin{bmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{mk} \end{bmatrix} = v_{1k}\begin{bmatrix} c_{11} \\ c_{21} \\ \vdots \\ c_{m1} \end{bmatrix} + v_{2k}\begin{bmatrix} c_{12} \\ c_{22} \\ \vdots \\ c_{m2} \end{bmatrix} + \cdots + v_{rk}\begin{bmatrix} c_{1r} \\ c_{2r} \\ \vdots \\ c_{mr} \end{bmatrix}
\tag{5}
$$

where $k = 1, \cdots, n$. Now the vector on the left is the kth column vector of $\mathbf{A}$. We see that each of these n columns is a linear combination of the same r columns on the right. Hence $\mathbf{A}$ cannot have more linearly independent columns than rows, whose number is rank $\mathbf{A} = r$. Now rows of $\mathbf{A}$ are columns of the transpose $\mathbf{A}^\mathsf{T}$. For $\mathbf{A}^\mathsf{T}$ our conclusion is that $\mathbf{A}^\mathsf{T}$ cannot have more linearly independent columns than rows, so that $\mathbf{A}$ cannot have more linearly independent rows than columns. Together, the number of linearly independent columns of $\mathbf{A}$ must be r, the rank of $\mathbf{A}$. This completes the proof. ■

EXAMPLE 4 Illustration of Theorem 3

The matrix in (2) has rank 2. From Example 3 we see that the first two row vectors are linearly independent and by "working backward" we can verify that Row 3 $= 6$ Row 1 $- \frac{1}{2}$ Row 2. Similarly, the first two columns are linearly independent, and by reducing the last matrix in Example 3 by columns we find that

$$\text{Column 3} = \tfrac{2}{3}\,\text{Column 1} + \tfrac{2}{3}\,\text{Column 2} \qquad \text{and} \qquad \text{Column 4} = \tfrac{2}{3}\,\text{Column 1} + \tfrac{29}{21}\,\text{Column 2}.$$ ■

Combining Theorems 2 and 3 we obtain

THEOREM 4

Linear Dependence of Vectors

Consider p vectors each having n components. If $n < p$, then *these vectors are linearly dependent.*

PROOF The matrix $\mathbf{A}$ with those p vectors as row vectors has p rows and $n < p$ columns; hence by Theorem 3 it has rank $\mathbf{A} \leqq n < p$, which implies linear dependence by Theorem 2. ■

Vector Space

The following related concepts are of general interest in linear algebra. In the present context they provide a clarification of essential properties of matrices and their role in connection with linear systems.

Consider a nonempty set V of vectors where each vector has the same number of components. If, for any two vectors $\mathbf{a}$ and $\mathbf{b}$ in V, we have that all their linear combinations $\alpha\mathbf{a} + \beta\mathbf{b}$ (α, β any real numbers) are also elements of V, and if, furthermore, $\mathbf{a}$ and $\mathbf{b}$ satisfy the laws (3a), (3c), (3d), and (4) in Sec. 7.1, as well as any vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ in V satisfy (3b) then V is a vector space. Note that here we wrote laws (3) and (4) of Sec. 7.1 in lowercase letters $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, which is our notation for vectors. More on vector spaces in Sec. 7.9.

The maximum number of linearly independent vectors in V is called the **dimension** of V and is denoted by dim V. Here we assume the dimension to be finite; infinite dimension will be defined in Sec. 7.9.

A linearly independent set in V consisting of a maximum possible number of vectors in V is called a **basis** for V. In other words, any largest possible set of independent vectors in V forms basis for V. That means, if we add one or more vector to that set, the set will be linearly dependent. (See also the beginning of Sec. 7.4 on linear independence and dependence of vectors.) Thus, the number of vectors of a basis for V equals dim V.

The set of all linear combinations of given vectors $\mathbf{a}_{(1)}, \cdots, \mathbf{a}_{(p)}$ with the same number of components is called the **span** of these vectors. Obviously, a span is a vector space. If in addition, the given vectors $\mathbf{a}_{(1)}, \cdots, \mathbf{a}_{(p)}$ are linearly independent, then they form a basis for that vector space.

This then leads to another equivalent definition of basis. A set of vectors is a **basis** for a vector space V if (1) the vectors in the set are linearly independent, and if (2) any vector in V can be expressed as a linear combination of the vectors in the set. If (2) holds, we also say that the set of vectors **spans** the vector space V.

By a **subspace** of a vector space V we mean a nonempty subset of V (including V itself) that forms a vector space with respect to the two algebraic operations (addition and scalar multiplication) defined for the vectors of V.

EXAMPLE 5 **Vector Space, Dimension, Basis**

The span of the three vectors in Example 1 is a vector space of dimension 2. A basis of this vector space consists of any two of those three vectors, for instance, $\mathbf{a}_{(1)}$, $\mathbf{a}_{(2)}$, or $\mathbf{a}_{(1)}$, $\mathbf{a}_{(3)}$, etc. ■

We further note the simple

THEOREM 5 **Vector Space R^n**

The vector space R^n consisting of all vectors with n components (n real numbers) has dimension n.

PROOF A basis of n vectors is $\mathbf{a}_{(1)} = [1 \quad 0 \quad \cdots \quad 0]$, $\mathbf{a}_{(2)} = [0 \quad 1 \quad 0 \quad \cdots \quad 0]$, $\cdots$, $\mathbf{a}_{(n)} = [0 \quad \cdots \quad 0 \quad 1]$. ■

For a matrix $\mathbf{A}$, we call the span of the row vectors the **row space** of $\mathbf{A}$. Similarly, the span of the column vectors of $\mathbf{A}$ is called the **column space** of $\mathbf{A}$.

Now, Theorem 3 shows that a matrix $\mathbf{A}$ has as many linearly independent rows as columns. By the definition of dimension, their number is the dimension of the row space or the column space of $\mathbf{A}$. This proves

THEOREM 6 **Row Space and Column Space**

The row space and the column space of a matrix $\mathbf{A}$ have the same dimension, equal to rank $\mathbf{A}$.

Finally, for a given matrix $\mathbf{A}$ the solution set of the homogeneous system $\mathbf{Ax} = 0$ is a vector space, called the **null space** of $\mathbf{A}$, and its dimension is called the **nullity** of $\mathbf{A}$. In the next section we motivate and prove the basic relation

$$\text{rank } \mathbf{A} + \text{nullity } \mathbf{A} = \text{Number of columns of } \mathbf{A}. \tag{6}$$

PROBLEM SET 7.4

1–10 RANK, ROW SPACE, COLUMN SPACE

Find the rank. Find a basis for the row space. Find a basis for the column space. *Hint.* Row-reduce the matrix and its transpose. (You may omit obvious factors from the vectors of these bases.)

1. $\begin{bmatrix} 4 & -2 & 6 \\ -2 & 1 & -3 \end{bmatrix}$

2. $\begin{bmatrix} a & b \\ b & a \end{bmatrix}$

3. $\begin{bmatrix} 0 & 3 & 5 \\ 3 & 5 & 0 \\ 5 & 0 & 10 \end{bmatrix}$

4. $\begin{bmatrix} 6 & -4 & 0 \\ -4 & 0 & 2 \\ 0 & 2 & 6 \end{bmatrix}$

5. $\begin{bmatrix} 0.2 & -0.1 & 0.4 \\ 0 & 1.1 & -0.3 \\ 0.1 & 0 & -2.1 \end{bmatrix}$

6. $\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & -4 \\ 0 & 4 & 0 \end{bmatrix}$

7. $\begin{bmatrix} 8 & 0 & 4 & 0 \\ 0 & 2 & 0 & 4 \\ 4 & 0 & 2 & 0 \end{bmatrix}$

8. $\begin{bmatrix} 2 & 4 & 8 & 16 \\ 16 & 8 & 4 & 2 \\ 4 & 8 & 16 & 2 \\ 2 & 16 & 8 & 4 \end{bmatrix}$

9. $\begin{bmatrix} 9 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$

10. $\begin{bmatrix} 5 & -2 & 1 & 0 \\ -2 & 0 & -4 & 1 \\ 1 & -4 & -11 & 2 \\ 0 & 1 & 2 & 0 \end{bmatrix}$

11. CAS Experiment. Rank. (a) Show experimentally that the $n \times n$ matrix $\mathbf{A} = [a_{jk}]$ with $a_{jk} = j + k - 1$ has rank 2 for any n. (Problem 20 shows $n = 4$.) Try to prove it.

(b) Do the same when $a_{jk} = j + k + c$, where c is any positive integer.

(c) What is rank $\mathbf{A}$ if $a_{jk} = 2^{j+k-2}$? Try to find other large matrices of low rank independent of n.

12–16 GENERAL PROPERTIES OF RANK

Show the following:

12. rank $\mathbf{B}^\mathsf{T}\mathbf{A}^\mathsf{T}$ = rank $\mathbf{AB}$. (Note the order!)

13. rank $\mathbf{A}$ = rank $\mathbf{B}$ does *not* imply rank $\mathbf{A}^2$ = rank $\mathbf{B}^2$. (Give a counterexample.)

14. If $\mathbf{A}$ is not square, either the row vectors or the column vectors of $\mathbf{A}$ are linearly dependent.

15. If the row vectors of a square matrix are linearly independent, so are the column vectors, and conversely.

16. Give examples showing that the rank of a product of matrices cannot exceed the rank of either factor.

17–25 LINEAR INDEPENDENCE

Are the following sets of vectors linearly independent? Show the details of your work.

17. $[3 \;\; 4 \;\; 0 \;\; 2]$, $[2 \;\; -1 \;\; 3 \;\; 7]$, $[1 \;\; 16 \;\; -12 \;\; -22]$

18. $[1 \;\; \tfrac{1}{2} \;\; \tfrac{1}{3} \;\; \tfrac{1}{4}]$, $[\tfrac{1}{2} \;\; \tfrac{1}{3} \;\; \tfrac{1}{4} \;\; \tfrac{1}{5}]$, $[\tfrac{1}{3} \;\; \tfrac{1}{4} \;\; \tfrac{1}{5} \;\; \tfrac{1}{6}]$, $[\tfrac{1}{4} \;\; \tfrac{1}{5} \;\; \tfrac{1}{6} \;\; \tfrac{1}{7}]$

19. $[0 \;\; 1 \;\; 1]$, $[1 \;\; 1 \;\; 1]$, $[0 \;\; 0 \;\; 1]$

20. $[1 \;\; 2 \;\; 3 \;\; 4]$, $[2 \;\; 3 \;\; 4 \;\; 5]$, $[3 \;\; 4 \;\; 5 \;\; 6]$, $[4 \;\; 5 \;\; 6 \;\; 7]$

21. $[2 \;\; 0 \;\; 0 \;\; 7]$, $[2 \;\; 0 \;\; 0 \;\; 8]$, $[2 \;\; 0 \;\; 0 \;\; 9]$, $[2 \;\; 0 \;\; 1 \;\; 0]$

22. $[0.4 \;\; -0.2 \;\; 0.2]$, $[0 \;\; 0 \;\; 0]$, $[3.0 \;\; -0.6 \;\; 1.5]$

23. $[9 \;\; 8 \;\; 7 \;\; 6 \;\; 5]$, $[9 \;\; 7 \;\; 5 \;\; 3 \;\; 1]$

24. $[4 \;\; -1 \;\; 3]$, $[0 \;\; 8 \;\; 1]$, $[1 \;\; 3 \;\; -5]$, $[2 \;\; 6 \;\; 1]$

25. $[6 \;\; 0 \;\; -1 \;\; 3]$, $[2 \;\; 2 \;\; 5 \;\; 0]$, $[-4 \;\; -4 \;\; -4 \;\; -4]$

26. Linearly independent subset. Beginning with the last of the vectors $[3 \;\; 0 \;\; 1 \;\; 2]$, $[6 \;\; 1 \;\; 0 \;\; 0]$, $[12 \;\; 1 \;\; 2 \;\; 4]$, $[6 \;\; 0 \;\; 2 \;\; 4]$, and $[9 \;\; 0 \;\; 1 \;\; 2]$, omit one after another until you get a linearly independent set.

27–35 VECTOR SPACE

Is the given set of vectors a vector space? Give reasons. If your answer is yes, determine the dimension and find a basis. ($v_1, v_2, \cdots$ denote components.)

27. All vectors in R^3 with $v_1 - v_2 + 2v_3 = 0$

28. All vectors in R^3 with $3v_2 + v_3 = k$

29. All vectors in R^2 with $v_1 \geqq v_2$

30. All vectors in R^n with the first $n - 2$ components zero

31. All vectors in R^5 with positive components

32. All vectors in R^3 with $3v_1 - 2v_2 + v_3 = 0$, $4v_1 + 5v_2 = 0$

33. All vectors in R^3 with $3v_1 - v_3 = 0$, $2v_1 + 3v_2 - 4v_3 = 0$

34. All vectors in R^n with $|v_j| = 1$ for $j = 1, \cdots, n$

35. All vectors in R^4 with $v_1 = 2v_2 = 3v_3 = 4v_4$

7.5 Solutions of Linear Systems: Existence, Uniqueness

Rank, as just defined, gives complete information about existence, uniqueness, and general structure of the solution set of linear systems as follows.

A linear system of equations in n unknowns has a unique solution if the coefficient matrix and the augmented matrix have the same rank n, and infinitely many solutions if that common rank is less than n. The system has no solution if those two matrices have different rank.

To state this precisely and prove it, we shall use the generally important concept of a **submatrix** of $\mathbf{A}$. By this we mean any matrix obtained from $\mathbf{A}$ by omitting some rows or columns (or both). By definition this includes $\mathbf{A}$ itself (as the matrix obtained by omitting no rows or columns); this is practical.

THEOREM 1

Fundamental Theorem for Linear Systems

(a) Existence. *A linear system of m equations in n unknowns* $x_1, \cdots, x_n$

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m
\end{aligned}
\tag{1}
$$

is **consistent**, *that is, has solutions, if and only if the coefficient matrix* $\mathbf{A}$ *and the augmented matrix* $\tilde{\mathbf{A}}$ *have the same rank. Here,*

$$
\mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \quad \textit{and} \quad \tilde{\mathbf{A}} = \left[\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdots & \cdot & \cdot \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right]
$$

(b) Uniqueness. *The system* (1) *has precisely one solution if and only if this common rank r of* $\mathbf{A}$ *and* $\tilde{\mathbf{A}}$ *equals n.*

(c) Infinitely many solutions. *If this common rank r is less than n, the system* (1) *has infinitely many solutions. All of these solutions are obtained by determining r suitable unknowns* (*whose submatrix of coefficients must have rank r*) *in terms of the remaining $n - r$ unknowns, to which arbitrary values can be assigned.* (See Example 3 in Sec. 7.3.)

(d) Gauss elimination (Sec. 7.3). *If solutions exist, they can all be obtained by the Gauss elimination.* (This method will automatically reveal whether or not solutions exist; see Sec. 7.3.)

PROOF **(a)** We can write the system (1) in vector form $\mathbf{Ax} = \mathbf{b}$ or in terms of column vectors $\mathbf{c}_{(1)}, \cdots, \mathbf{c}_{(n)}$ of $\mathbf{A}$:

$$\mathbf{c}_{(1)}x_1 + \mathbf{c}_{(2)}x_2 + \cdots + \mathbf{c}_{(n)}x_n = \mathbf{b}. \tag{2}$$

$\widetilde{\mathbf{A}}$ is obtained by augmenting $\mathbf{A}$ by a single column $\mathbf{b}$. Hence, by Theorem 3 in Sec. 7.4, rank $\widetilde{\mathbf{A}}$ equals rank $\mathbf{A}$ or rank $\mathbf{A} + 1$. Now if (1) has a solution $\mathbf{x}$, then (2) shows that $\mathbf{b}$ must be a linear combination of those column vectors, so that $\widetilde{\mathbf{A}}$ and $\mathbf{A}$ have the same maximum number of linearly independent column vectors and thus the same rank.

Conversely, if rank $\widetilde{\mathbf{A}} =$ rank $\mathbf{A}$, then $\mathbf{b}$ must be a linear combination of the column vectors of $\mathbf{A}$, say,

$$\mathbf{b} = \alpha_1\mathbf{c}_{(1)} + \cdots + \alpha_n\mathbf{c}_{(n)} \tag{2*}$$

since otherwise rank $\widetilde{\mathbf{A}} =$ rank $\mathbf{A} + 1$. But (2*) means that (1) has a solution, namely, $x_1 = \alpha_1, \cdots, x_n = \alpha_n$, as can be seen by comparing (2*) and (2).

(b) If rank $\mathbf{A} = n$, the n column vectors in (2) are linearly independent by Theorem 3 in Sec. 7.4. We claim that then the representation (2) of $\mathbf{b}$ is unique because otherwise

$$\mathbf{c}_{(1)}x_1 + \cdots + \mathbf{c}_{(n)}x_n = \mathbf{c}_{(1)}\tilde{x}_1 + \cdots + \mathbf{c}_{(n)}\tilde{x}_n.$$

This would imply (take all terms to the left, with a minus sign)

$$(x_1 - \tilde{x}_1)\mathbf{c}_{(1)} + \cdots + (x_n - \tilde{x}_n)\mathbf{c}_{(n)} = \mathbf{0}$$

and $x_1 - \tilde{x}_1 = 0, \cdots, x_n - \tilde{x}_n = 0$ by linear independence. But this means that the scalars $x_1, \cdots, x_n$ in (2) are uniquely determined, that is, the solution of (1) is unique.

(c) If rank $\mathbf{A} =$ rank $\widetilde{\mathbf{A}} = r < n$, then by Theorem 3 in Sec. 7.4 there is a linearly independent set K of r column vectors of $\mathbf{A}$ such that the other $n - r$ column vectors of $\mathbf{A}$ are linear combinations of those vectors. We renumber the columns and unknowns, denoting the renumbered quantities by $\hat{\ }$, so that $\{\hat{\mathbf{c}}_{(1)}, \cdots, \hat{\mathbf{c}}_{(r)}\}$ is that linearly independent set K. Then (2) becomes

$$\hat{\mathbf{c}}_{(1)}\hat{x}_1 + \cdots + \hat{\mathbf{c}}_{(r)}\hat{x}_r + \hat{\mathbf{c}}_{(r+1)}\hat{x}_{r+1} + \cdots + \hat{\mathbf{c}}_{(n)}\hat{x}_n = \mathbf{b},$$

$\hat{\mathbf{c}}_{(r+1)}, \cdots, \hat{\mathbf{c}}_{(n)}$ are linear combinations of the vectors of K, and so are the vectors $\hat{x}_{r+1}\hat{\mathbf{c}}_{(r+1)}, \cdots, \hat{x}_n\hat{\mathbf{c}}_{(n)}$. Expressing these vectors in terms of the vectors of K and collecting terms, we can thus write the system in the form

$$\hat{\mathbf{c}}_{(1)}y_1 + \cdots + \hat{\mathbf{c}}_{(r)}y_r = \mathbf{b} \tag{3}$$

with $y_j = \hat{x}_j + \beta_j$, where β_j results from the $n - r$ terms $\hat{\mathbf{c}}_{(r+1)}\hat{x}_{r+1}, \cdots, \hat{\mathbf{c}}_{(n)}\hat{x}_n$; here, $j = 1, \cdots, r$. Since the system has a solution, there are $y_1, \cdots, y_r$ satisfying (3). These scalars are unique since K is linearly independent. Choosing $\hat{x}_{r+1}, \cdots, \hat{x}_n$ fixes the β_j and corresponding $\hat{x}_j = y_j - \beta_j$, where $j = 1, \cdots, r$.

(d) This was discussed in Sec. 7.3 and is restated here as a reminder. ■

The theorem is illustrated in Sec. 7.3. In Example 2 there is a unique solution since rank $\tilde{\mathbf{A}}$ = rank $\mathbf{A} = n = 3$ (as can be seen from the last matrix in the example). In Example 3 we have rank $\tilde{\mathbf{A}}$ = rank $\mathbf{A} = 2 < n = 4$ and can choose x_3 and x_4 arbitrarily. In Example 4 there is no solution because rank $\mathbf{A} = 2 <$ rank $\tilde{\mathbf{A}} = 3$.

Homogeneous Linear System

Recall from Sec. 7.3 that a linear system (1) is called **homogeneous** if all the b_j's are zero, and **nonhomogeneous** if one or several b_j's are not zero. For the homogeneous system we obtain from the Fundamental Theorem the following results.

THEOREM 2

Homogeneous Linear System

A homogeneous linear system

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\
\cdots\cdots\cdots\cdots\cdots\cdots & \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0
\end{aligned}
\tag{4}
$$

always has the **trivial solution** $x_1 = 0, \cdots, x_n = 0$. *Nontrivial solutions exist if and only if rank* $\mathbf{A} < n$. *If* rank $\mathbf{A} = r < n$, *these solutions, together with* $\mathbf{x} = \mathbf{0}$, *form a vector space* (see Sec. 7.4) *of dimension* $n - r$ called the **solution space** of (4).

In particular, if $\mathbf{x}_{(1)}$ *and* $\mathbf{x}_{(2)}$ *are solution vectors of* (4), *then* $\mathbf{x} = c_1\mathbf{x}_{(1)} + c_2\mathbf{x}_{(2)}$ *with any scalars* c_1 *and* c_2 *is a solution vector of* (4). (This ***does not hold*** for nonhomogeneous systems. Also, the term *solution space is used for homogeneous* systems only.)

PROOF The first proposition can be seen directly from the system. It agrees with the fact that $\mathbf{b} = \mathbf{0}$ implies that rank $\tilde{\mathbf{A}}$ = rank $\mathbf{A}$, so that a homogeneous system is always ***consistent***. If rank $\mathbf{A} = n$, the trivial solution is the unique solution according to (b) in Theorem 1. If rank $\mathbf{A} < n$, there are nontrivial solutions according to (c) in Theorem 1. The solutions form a vector space because if $\mathbf{x}_{(1)}$ and $\mathbf{x}_{(2)}$ are any of them, then $\mathbf{A}\mathbf{x}_{(1)} = \mathbf{0}$, $\mathbf{A}\mathbf{x}_{(2)} = \mathbf{0}$, and this implies $\mathbf{A}(\mathbf{x}_{(1)} + \mathbf{x}_{(2)}) = \mathbf{A}\mathbf{x}_{(1)} + \mathbf{A}\mathbf{x}_{(2)} = \mathbf{0}$ as well as $\mathbf{A}(c\mathbf{x}_{(1)}) = c\mathbf{A}\mathbf{x}_{(1)} = \mathbf{0}$, where c is arbitrary. If rank $\mathbf{A} = r < n$, Theorem 1 (c) implies that we can choose $n - r$ suitable unknowns, call them $x_{r+1}, \cdots, x_n$, in an arbitrary fashion, and every solution is obtained in this way. Hence a basis for the solution space, briefly called a **basis of solutions** of (4), is $\mathbf{y}_{(1)}, \cdots, \mathbf{y}_{(n-r)}$, where the basis vector $\mathbf{y}_{(j)}$ is obtained by choosing $x_{r+j} = 1$ and the other $x_{r+1}, \cdots, x_n$ zero; the corresponding first r components of this solution vector are then determined. Thus the solution space of (4) has dimension $n - r$. This proves Theorem 2. ■

The solution space of (4) is also called the **null space** of $\mathbf{A}$ because $\mathbf{Ax} = \mathbf{0}$ for every $\mathbf{x}$ in the solution space of (4). Its dimension is called the **nullity** of $\mathbf{A}$. Hence Theorem 2 states that

$$\text{rank}\ \mathbf{A} + \text{nullity}\ \mathbf{A} = n \tag{5}$$

where n is the number of unknowns (number of columns of $\mathbf{A}$).

Furthermore, by the definition of rank we have rank $\mathbf{A} \leqq m$ in (4). Hence if $m < n$, then rank $\mathbf{A} < n$. By Theorem 2 this gives the practically important

THEOREM 3

Homogeneous Linear System with Fewer Equations Than Unknowns

A homogeneous linear system with fewer equations than unknowns always has nontrivial solutions.

Nonhomogeneous Linear Systems

The characterization of all solutions of the linear system (1) is now quite simple, as follows.

THEOREM 4

Nonhomogeneous Linear System

If a nonhomogeneous linear system (1) *is consistent, then all of its solutions are obtained as*

$$\mathbf{x} = \mathbf{x}_0 + \mathbf{x}_h \tag{6}$$

where $\mathbf{x}_0$ *is any* (fixed) *solution of* (1) *and* $\mathbf{x}_h$ *runs through all the solutions of the corresponding homogeneous system* (4).

PROOF The difference $\mathbf{x}_h = \mathbf{x} - \mathbf{x}_0$ of any two solutions of (1) is a solution of (4) because $\mathbf{Ax}_h = \mathbf{A}(\mathbf{x} - \mathbf{x}_0) = \mathbf{Ax} - \mathbf{Ax}_0 = \mathbf{b} - \mathbf{b} = \mathbf{0}$. Since $\mathbf{x}$ is any solution of (1), we get all the solutions of (1) if in (6) we take any solution $\mathbf{x}_0$ of (1) and let $\mathbf{x}_h$ vary throughout the solution space of (4). ■

This covers a main part of our discussion of characterizing the solutions of systems of linear equations. Our next main topic is determinants and their role in linear equations.

7.6 For Reference: Second- and Third-Order Determinants

We created this section as a quick general reference section on second- and third-order determinants. It is completely independent of the theory in Sec. 7.7 and suffices as a reference for many of our examples and problems. Since this section is for reference, ***go on to the next section, consulting this material only when needed***.

A **determinant of second order** is denoted and defined by

$$D = \det \mathbf{A} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}. \tag{1}$$

So here we have ***bars*** (whereas a matrix has ***brackets***).

Cramer's rule for solving linear systems of two equations in two unknowns

$$\text{(2)}\qquad \begin{aligned} &\text{(a)}\quad a_{11}x_1 + a_{12}x_2 = b_1 \\ &\text{(b)}\quad a_{21}x_1 + a_{22}x_2 = b_2 \end{aligned}$$

is

$$\text{(3)}\qquad \begin{aligned} x_1 &= \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{D} = \frac{b_1a_{22} - a_{12}b_2}{D}, \\ x_2 &= \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{D} = \frac{a_{11}b_2 - b_1a_{21}}{D} \end{aligned}$$

with D as in (1), provided

$$D \neq 0.$$

The value $D = 0$ appears for homogeneous systems with nontrivial solutions.

PROOF We prove (3). To eliminate x_2 multiply (2a) by a_{22} and (2b) by $-a_{12}$ and add,

$$(a_{11}a_{22} - a_{12}a_{21})x_1 = b_1a_{22} - a_{12}b_2.$$

Similarly, to eliminate x_1 multiply (2a) by $-a_{21}$ and (2b) by a_{11} and add,

$$(a_{11}a_{22} - a_{12}a_{21})x_2 = a_{11}b_2 - b_1a_{21}.$$

Assuming that $D = a_{11}a_{22} - a_{12}a_{21} \neq 0$, dividing, and writing the right sides of these two equations as determinants, we obtain (3). ■

EXAMPLE 1 Cramer's Rule for Two Equations

If $\begin{aligned} 4x_1 + 3x_2 &= 12 \\ 2x_1 + 5x_2 &= -8 \end{aligned}$ then $x_1 = \dfrac{\begin{vmatrix} 12 & 3 \\ -8 & 5 \end{vmatrix}}{\begin{vmatrix} 4 & 3 \\ 2 & 5 \end{vmatrix}} = \dfrac{84}{14} = 6,\quad x_2 = \dfrac{\begin{vmatrix} 4 & 12 \\ 2 & -8 \end{vmatrix}}{\begin{vmatrix} 4 & 3 \\ 2 & 5 \end{vmatrix}} = \dfrac{-56}{14} = -4.$ ■

Third-Order Determinants

A **determinant of third order** can be defined by

$$\text{(4)}\qquad D = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31}\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}.$$

Note the following. The signs on the right are $+ - +$. Each of the three terms on the right is an entry in the first column of D times its **minor**, that is, the second-order determinant obtained from D by deleting the row and column of that entry; thus, for a_{11} delete the first row and first column, and so on.

If we write out the minors in (4), we obtain

$$D = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{21}a_{13}a_{32} - a_{21}a_{12}a_{33} + a_{31}a_{12}a_{23} - a_{31}a_{13}a_{22}. \tag{4*}$$

Cramer's Rule for Linear Systems of Three Equations

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3 \end{aligned} \tag{5}$$

is

$$x_1 = \frac{D_1}{D}, \quad x_2 = \frac{D_2}{D}, \quad x_3 = \frac{D_3}{D} \qquad (D \neq 0) \tag{6}$$

with the *determinant D of the system* given by (4) and

$$D_1 = \begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}, \quad D_2 = \begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix}, \quad D_3 = \begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix}.$$

Note that D_1, D_2, D_3 are obtained by replacing Columns 1, 2, 3, respectively, by the column of the right sides of (5).

Cramer's rule (6) can be derived by eliminations similar to those for (3), but it also follows from the general case (Theorem 4) in the next section.

7.7 Determinants. Cramer's Rule

Determinants were originally introduced for solving linear systems. Although ***impractical in computations***, they have important engineering applications in eigenvalue problems (Sec. 8.1), differential equations, vector algebra (Sec. 9.3), and in other areas. They can be introduced in several equivalent ways. Our definition is particularly for dealing with linear systems.

A **determinant of order** n is a scalar associated with an $n \times n$ (hence ***square!***) matrix $\mathbf{A} = [a_{jk}]$, and is denoted by

$$D = \det \mathbf{A} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}. \tag{1}$$

For $n = 1$, this determinant is defined by

$$D = a_{11}. \tag{2}$$

For $n \geqq 2$ by

$$D = a_{j1}C_{j1} + a_{j2}C_{j2} + \cdots + a_{jn}C_{jn} \qquad (j = 1, 2, \cdots, \text{ or } n) \tag{3a}$$

or

$$D = a_{1k}C_{1k} + a_{2k}C_{2k} + \cdots + a_{nk}C_{nk} \quad (k = 1, 2, \cdots, \text{ or } n). \tag{3b}$$

Here,

$$C_{jk} = (-1)^{j+k}M_{jk}$$

and M_{jk} is a determinant of order $n - 1$, namely, the determinant of the submatrix of **A** obtained from **A** by omitting the row and column of the entry a_{jk}, that is, the jth row and the kth column.

In this way, D is defined in terms of n determinants of order $n - 1$, each of which is, in turn, defined in terms of $n - 1$ determinants of order $n - 2$, and so on—until we finally arrive at second-order determinants, in which those submatrices consist of single entries whose determinant is defined to be the entry itself.

From the definition it follows that *we may **expand** D by any row or column,* that is, choose in (3) the entries in any row or column, similarly when expanding the C_{jk}'s in (3), and so on.

This definition is unambiguous, that is, it yields the same value for D no matter which columns or rows we choose in expanding. A proof is given in App. 4.

Terms used in connection with determinants are taken from matrices. In D we have n^2 **entries** a_{jk}, also n **rows** and n **columns**, and a **main diagonal** on which $a_{11}, a_{22}, \cdots, a_{nn}$ stand. Two terms are new:

M_{jk} is called the **minor** *of* a_{jk} *in* D, and C_{jk} the **cofactor** *of* a_{jk} *in* D.

For later use we note that (3) may also be written in terms of minors

$$D = \sum_{k=1}^{n} (-1)^{j+k} a_{jk} M_{jk} \qquad (j = 1, 2, \cdots, \text{ or } n) \tag{4a}$$

$$D = \sum_{j=1}^{n} (-1)^{j+k} a_{jk} M_{jk} \qquad (k = 1, 2, \cdots, \text{ or } n). \tag{4b}$$

EXAMPLE 1 **Minors and Cofactors of a Third-Order Determinant**

In (4) of the previous section the minors and cofactors of the entries in the first column can be seen directly. For the entries in the second row the minors are

$$M_{21} = \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix}, \qquad M_{22} = \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix}, \qquad M_{23} = \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix}$$

and the cofactors are $C_{21} = -M_{21}$, $C_{22} = +M_{22}$, and $C_{23} = -M_{23}$. Similarly for the third row—write these down yourself. And verify that the signs in C_{jk} form a **checkerboard pattern**

$$\begin{matrix} + & - & + \\ - & + & - \\ + & - & + \end{matrix}$$ ■

EXAMPLE 2 Expansions of a Third-Order Determinant

$$D = \begin{vmatrix} 1 & 3 & 0 \\ 2 & 6 & 4 \\ -1 & 0 & 2 \end{vmatrix} = 1\begin{vmatrix} 6 & 4 \\ 0 & 2 \end{vmatrix} - 3\begin{vmatrix} 2 & 4 \\ -1 & 2 \end{vmatrix} + 0\begin{vmatrix} 2 & 6 \\ -1 & 0 \end{vmatrix}$$

$$= 1(12 - 0) - 3(4 + 4) + 0(0 + 6) = -12.$$

This is the expansion by the first row. The expansion by the third column is

$$D = 0\begin{vmatrix} 2 & 6 \\ -1 & 0 \end{vmatrix} - 4\begin{vmatrix} 1 & 3 \\ -1 & 0 \end{vmatrix} + 2\begin{vmatrix} 1 & 3 \\ 2 & 6 \end{vmatrix} = 0 - 12 + 0 = -12.$$

Verify that the other four expansions also give the value -12. ■

EXAMPLE 3 Determinant of a Triangular Matrix

$$\begin{vmatrix} -3 & 0 & 0 \\ 6 & 4 & 0 \\ -1 & 2 & 5 \end{vmatrix} = -3\begin{vmatrix} 4 & 0 \\ 2 & 5 \end{vmatrix} = -3 \cdot 4 \cdot 5 = -60.$$

Inspired by this, can you formulate a little theorem on determinants of triangular matrices? Of diagonal matrices? ■

General Properties of Determinants

There is an attractive way of finding determinants (1) that consists of applying elementary row operations to (1). By doing so we obtain an "upper triangular" determinant (see Sec. 7.1, for definition with "matrix" replaced by "determinant") whose value is then very easy to compute, being just the product of its diagonal entries. This approach is ***similar*** (***but not the same!***) to what we did to matrices in Sec. 7.3. *In particular, be aware that interchanging two rows in a determinant introduces a multiplicative factor of* -1 *to the value of the determinant*! Details are as follows.

THEOREM 1

Behavior of an *n*th-Order Determinant under Elementary Row Operations

(a) *Interchange of two rows multiplies the value of the determinant by* -1.

(b) *Addition of a multiple of a row to another row does not alter the value of the determinant.*

(c) *Multiplication of a row by a nonzero constant c multiplies the value of the determinant by c.* (This holds also when $c = 0$, but no longer gives an elementary row operation.)

PROOF **(a)** By induction. The statement holds for $n = 2$ because

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc, \qquad \text{but} \qquad \begin{vmatrix} c & d \\ a & b \end{vmatrix} = bc - ad.$$

We now make the induction hypothesis that (a) holds for determinants of order $n - 1 \geqq 2$ and show that it then holds for determinants of order n. Let D be of order n. Let E be obtained from D by the interchange of two rows. Expand D and E by a row that is ***not*** one of those interchanged, call it the jth row. Then by (4a),

$$D = \sum_{k=1}^{n} (-1)^{j+k} a_{jk} M_{jk}, \qquad E = \sum_{k=1}^{n} (-1)^{j+k} a_{jk} N_{jk} \tag{5}$$

where N_{jk} is obtained from the minor M_{jk} of a_{jk} in D by the interchange of those two rows which have been interchanged in D (and which N_{jk} must both contain because we expand by another row!). Now these minors are of order $n - 1$. Hence the induction hypothesis applies and gives $N_{jk} = -M_{jk}$. Thus $E = -D$ by (5).

(b) Add c times Row i to Row j. Let $\tilde{D}$ be the new determinant. Its entries in Row j are $a_{jk} + ca_{ik}$. If we expand $\tilde{D}$ by this Row j, we see that we can write it as $\tilde{D} = D_1 + cD_2$, where $D_1 = D$ has in Row j the a_{jk}, whereas D_2 has in that Row j the a_{jk} from the addition. Hence D_2 has a_{jk} in both Row i and Row j. Interchanging these two rows gives D_2 back, but on the other hand it gives $-D_2$ by (a). Together $D_2 = -D_2 = 0$, so that $\tilde{D} = D_1 = D$.

(c) Expand the determinant by the row that has been multiplied.

CAUTION! $\det(c\mathbf{A}) = c^n \det \mathbf{A}$ (not $c \det \mathbf{A}$). Explain why. ■

EXAMPLE 4 Evaluation of Determinants by Reduction to Triangular Form

Because of Theorem 1 we may evaluate determinants by reduction to triangular form, as in the Gauss elimination for a matrix. For instance (with the blue explanations always referring to the ***preceding determinant***)

$$D = \begin{vmatrix} 2 & 0 & -4 & 6 \\ 4 & 5 & 1 & 0 \\ 0 & 2 & 6 & -1 \\ -3 & 8 & 9 & 1 \end{vmatrix}$$

$$= \begin{vmatrix} 2 & 0 & -4 & 6 \\ 0 & 5 & 9 & -12 \\ 0 & 2 & 6 & -1 \\ 0 & 8 & 3 & 10 \end{vmatrix} \begin{array}{l} \\ \text{Row 2} - 2 \text{ Row 1} \\ \\ \text{Row 4} + 1.5 \text{ Row 1} \end{array}$$

$$= \begin{vmatrix} 2 & 0 & -4 & 6 \\ 0 & 5 & 9 & -12 \\ 0 & 0 & 2.4 & 3.8 \\ 0 & 0 & -11.4 & 29.2 \end{vmatrix} \begin{array}{l} \\ \\ \text{Row 3} - 0.4 \text{ Row 2} \\ \text{Row 4} - 1.6 \text{ Row 2} \end{array}$$

$$= \begin{vmatrix} 2 & 0 & -4 & 6 \\ 0 & 5 & 9 & -12 \\ 0 & 0 & 2.4 & 3.8 \\ 0 & 0 & -0 & 47.25 \end{vmatrix} \begin{array}{l} \\ \\ \\ \text{Row 4} + 4.75 \text{ Row 3} \end{array}$$

$$= 2 \cdot 5 \cdot 2.4 \cdot 47.25 = 1134.$$ ■

THEOREM 2

Further Properties of *n*th-Order Determinants

(a)–(c) *in Theorem* 1 *hold also for columns.*

(d) ***Transposition*** *leaves the value of a determinant unaltered.*

(e) ***A zero row or column*** *renders the value of a determinant zero.*

(f) ***Proportional rows or columns*** *render the value of a determinant zero. In particular, a determinant with two identical rows or columns has the value zero.*

PROOF **(a)–(e)** follow directly from the fact that a determinant can be expanded by any row column. In **(d)**, transposition is defined as for matrices, that is, the jth row becomes the jth column of the transpose.

(f) If Row $j = c$ times Row i, then $D = cD_1$, where D_1 has Row $j =$ Row i. Hence an interchange of these rows reproduces D_1, but it also gives $-D_1$ by Theorem 1(a). Hence $D_1 = 0$ and $D = cD_1 = 0$. Similarly for columns. ■

It is quite remarkable that the important concept of the rank of a matrix $\mathbf{A}$, which is the maximum number of linearly independent row or column vectors of $\mathbf{A}$ (see Sec. 7.4), can be related to determinants. Here we may assume that rank $\mathbf{A} > 0$ because the only matrices with rank 0 are the zero matrices (see Sec. 7.4).

THEOREM 3

Rank in Terms of Determinants

Consider an $m \times n$ *matrix* $\mathbf{A} = [a_{jk}]$:

(1) $\mathbf{A}$ *has rank* $r \geqq 1$ *if and only if* $\mathbf{A}$ *has an* $r \times r$ *submatrix with a nonzero determinant.*

(2) *The determinant of any square submatrix with more than r rows, contained in* $\mathbf{A}$ *(if such a matrix exists!) has a value equal to zero.*

Furthermore, if $m = n$, *we have:*

(3) *An* $n \times n$ *square matrix* $\mathbf{A}$ *has rank n if and only if*

$$\det \mathbf{A} \neq 0.$$

PROOF The key idea is that elementary row operations (Sec. 7.3) alter neither rank (by Theorem 1 in Sec. 7.4) nor the property of a determinant being nonzero (by Theorem 1 in this section). The echelon form $\hat{\mathbf{A}}$ of $\mathbf{A}$ (see Sec. 7.3) has r nonzero row vectors (which are the first r row vectors) if and only if rank $\mathbf{A} = r$. Without loss of generality, we can assume that $r \geqq 1$. Let $\hat{\mathbf{R}}$ be the $r \times r$ submatrix in the left upper corner of $\hat{\mathbf{A}}$ (so that the entries of $\hat{\mathbf{R}}$ are in both the first r rows and r columns of $\hat{\mathbf{A}}$). Now $\hat{\mathbf{R}}$ is triangular, with all diagonal entries r_{jj} nonzero. Thus, $\det \hat{\mathbf{R}} = r_{11} \cdots r_{rr} \neq 0$. Also $\det \mathbf{R} \neq 0$ for the corresponding $r \times r$ submatrix $\mathbf{R}$ of $\mathbf{A}$ because $\hat{\mathbf{R}}$ results from $\mathbf{R}$ by elementary row operations. This proves part (1).

Similarly, $\det \mathbf{S} = 0$ for any square submatrix $\mathbf{S}$ of $r + 1$ or more rows perhaps contained in $\mathbf{A}$ because the corresponding submatrix $\hat{\mathbf{S}}$ of $\hat{\mathbf{A}}$ must contain a row of zeros (otherwise we would have rank $\mathbf{A} \geqq r + 1$), so that $\det \hat{\mathbf{S}} = 0$ by Theorem 2. This proves part (2). Furthermore, we have proven the theorem for an $m \times n$ matrix.

For an $n \times n$ square matrix $\mathbf{A}$ we proceed as follows. To prove (3), we apply part (1) (already proven!). This gives us that rank $\mathbf{A} = n \geqq 1$ if and only if $\mathbf{A}$ contains an $n \times n$ submatrix with nonzero determinant. But the only such submatrix contained in our square matrix $\mathbf{A}$, is $\mathbf{A}$ itself, hence det $\mathbf{A} \neq 0$. This proves part (3). ■

Cramer's Rule

Theorem 3 opens the way to the classical solution formula for linear systems known as **Cramer's rule**,[2] which gives solutions as quotients of determinants. ***Cramer's rule is not practical in computations*** for which the methods in Secs. 7.3 and 20.1–20.3 are suitable. However, Cramer's rule is of ***theoretical interest*** in differential equations (Secs. 2.10 and 3.3) and in other theoretical work that has engineering applications.

THEOREM 4

Cramer's Theorem (Solution of Linear Systems by Determinants)

(a) *If a linear system of n equations in the same number of unknowns* $x_1, \cdots, x_n$

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ &\cdot\ \cdot\ \cdot\ \cdot \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n \end{aligned} \tag{6}$$

has a nonzero coefficient determinant $D = \det \mathbf{A}$, *the system has precisely one solution. This solution is given by the formulas*

$$x_1 = \frac{D_1}{D}, \quad x_2 = \frac{D_2}{D}, \cdots, \quad x_n = \frac{D_n}{D} \qquad \textbf{(Cramer's rule)} \tag{7}$$

where D_k *is the determinant obtained from D by replacing in D the kth column by the column with the entries* $b_1, \cdots, b_n$.

(b) *Hence if the system* (6) *is* ***homogeneous*** *and* $D \neq 0$, *it has only the trivial solution* $x_1 = 0, x_2 = 0, \cdots, x_n = 0$. *If* $D = 0$, *the homogeneous system also has nontrivial solutions.*

PROOF The augmented matrix $\tilde{\mathbf{A}}$ of the system (6) is of size $n \times (n+1)$. Hence its rank can be at most n. Now if

$$D = \det \mathbf{A} = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} \neq 0, \tag{8}$$

[2] GABRIEL CRAMER (1704–1752), Swiss mathematician.

then rank $\mathbf{A} = n$ by Theorem 3. Thus rank $\tilde{\mathbf{A}} = \text{rank}\,\mathbf{A}$. Hence, by the Fundamental Theorem in Sec. 7.5, the system (6) has a unique solution.

Let us now prove (7). Expanding D by its kth column, we obtain

$$D = a_{1k}C_{1k} + a_{2k}C_{2k} + \cdots + a_{nk}C_{nk}, \tag{9}$$

where C_{ik} is the cofactor of entry a_{ik} in D. If we replace the entries in the kth column of D by any other numbers, we obtain a new determinant, say, $\hat{D}$. Clearly, its expansion by the kth column will be of the form (9), with $a_{1k}, \cdots, a_{nk}$ replaced by those new numbers and the cofactors C_{ik} as before. In particular, if we choose as new numbers the entries $a_{1l}, \cdots, a_{nl}$ of the lth column of D (where $l \neq k$), we have a new determinant $\hat{D}$ which has the column $[a_{1l} \quad \cdots \quad a_{nl}]^{\mathsf{T}}$ twice, once as its lth column, and once as its kth because of the replacement. Hence $\hat{D} = 0$ by Theorem 2(f). If we now expand $\hat{D}$ by the column that has been replaced (the kth column), we thus obtain

$$a_{1l}C_{1k} + a_{2l}C_{2k} + \cdots + a_{nl}C_{nk} = 0 \qquad (l \neq k). \tag{10}$$

We now multiply the first equation in (6) by C_{1k} on both sides, the second by $C_{2k}, \cdots$, the last by C_{nk}, and add the resulting equations. This gives

$$\begin{aligned} C_{1k}(a_{11}x_1 + \cdots + a_{1n}x_n) + \cdots + C_{nk}(a_{n1}x_1 + \cdots + a_{nn}x_n) \\ = b_1C_{1k} + \cdots + b_nC_{nk}. \end{aligned} \tag{11}$$

Collecting terms with the same x_j, we can write the left side as

$$x_1(a_{11}C_{1k} + a_{21}C_{2k} + \cdots + a_{n1}C_{nk}) + \cdots + x_n(a_{1n}C_{1k} + a_{2n}C_{2k} + \cdots + a_{nn}C_{nk}).$$

From this we see that x_k is multiplied by

$$a_{1k}C_{1k} + a_{2k}C_{2k} + \cdots + a_{nk}C_{nk}.$$

Equation (9) shows that this equals D. Similarly, x_1 is multiplied by

$$a_{1l}C_{1k} + a_{2l}C_{2k} + \cdots + a_{nl}C_{nk}.$$

Equation (10) shows that this is zero when $l \neq k$. Accordingly, the left side of (11) equals simply x_kD, so that (11) becomes

$$x_kD = b_1C_{1k} + b_2C_{2k} + \cdots + b_nC_{nk}.$$

Now the right side of this is D_k as defined in the theorem, expanded by its kth column, so that division by D gives (7). This proves Cramer's rule.

If (6) is homogeneous and $D \neq 0$, then each D_k has a column of zeros, so that $D_k = 0$ by Theorem 2(e), and (7) gives the trivial solution.

Finally, if (6) is homogeneous and $D = 0$, then rank $\mathbf{A} < n$ by Theorem 3, so that nontrivial solutions exist by Theorem 2 in Sec. 7.5. ■

EXAMPLE 5 Illustration of Cramer's Rule (Theorem 4)

For $n = 2$, see Example 1 of Sec. 7.6. Also, at the end of that section, we give Cramer's rule for a general linear system of three equations. ■

Finally, an important application for Cramer's rule dealing with inverse matrices will be given in the next section.

PROBLEM SET 7.7

1–6 GENERAL PROBLEMS

1. **General Properties of Determinants.** Illustrate each statement in Theorems 1 and 2 with an example of your choice.

2. **Second-Order Determinant.** Expand a general second-order determinant in four possible ways and show that the results agree.

3. **Third-Order Determinant.** Do the task indicated in Theorem 2. Also evaluate D by reduction to triangular form.

4. **Expansion Numerically Impractical.** Show that the computation of an nth-order determinant by expansion involves $n!$ multiplications, which if a multiplication takes 10^{-9} sec would take these times:

n	10	15	20	25
Time	0.004 sec	22 min	77 years	$0.5 \cdot 10^9$ years

5. **Multiplication by Scalar.** Show that $\det(k\mathbf{A}) = k^n \det \mathbf{A}$ (not $k \det \mathbf{A}$). Give an example.

6. **Minors, cofactors.** Complete the list in Example 1.

7–15 EVALUATION OF DETERMINANTS

Showing the details, evaluate:

7. $\begin{vmatrix} \cos\alpha & \sin\alpha \\ \sin\beta & \cos\beta \end{vmatrix}$

8. $\begin{vmatrix} 0.4 & 4.9 \\ 1.5 & -1.3 \end{vmatrix}$

9. $\begin{vmatrix} \cos n\theta & \sin n\theta \\ -\sin n\theta & \cos n\theta \end{vmatrix}$

10. $\begin{vmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{vmatrix}$

11. $\begin{vmatrix} 4 & -1 & 8 \\ 0 & 2 & 3 \\ 0 & 0 & 5 \end{vmatrix}$

12. $\begin{vmatrix} a & b & c \\ c & a & b \\ b & c & a \end{vmatrix}$

13. $\begin{vmatrix} 0 & 4 & -1 & 5 \\ -4 & 0 & 3 & -2 \\ 1 & -3 & 0 & 1 \\ -5 & 2 & -1 & 0 \end{vmatrix}$

14. $\begin{vmatrix} 4 & 7 & 0 & 0 \\ 2 & 8 & 0 & 0 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & -2 & 2 \end{vmatrix}$

15. $\begin{vmatrix} 1 & 2 & 0 & 0 \\ 2 & 4 & 2 & 0 \\ 0 & 2 & 9 & 2 \\ 0 & 0 & 2 & 16 \end{vmatrix}$

16. **CAS EXPERIMENT. Determinant of Zeros and Ones.** Find the value of the determinant of the $n \times n$ matrix A_n with main diagonal entries all 0 and all others 1. Try to find a formula for this. Try to prove it by induction. Interpret A_3 and A_4 as *incidence matrices* (as in Problem Set 7.1 but without the minuses) of a triangle and a tetrahedron, respectively; similarly for an *n-simplex*, having n vertices and $n(n-1)/2$ edges (and spanning R^{n-1}, $n = 5, 6, \cdots$).

17–19 RANK BY DETERMINANTS

Find the rank by Theorem 3 (which is not very practical) and check by row reduction. Show details.

17. $\begin{bmatrix} 4 & 9 \\ -8 & -6 \\ 16 & 12 \end{bmatrix}$

18. $\begin{bmatrix} 0 & 4 & -6 \\ 4 & 0 & 10 \\ -6 & 10 & 0 \end{bmatrix}$

19. $\begin{bmatrix} 1 & 5 & 2 & 2 \\ 1 & 3 & 2 & 6 \\ 4 & 0 & 8 & 48 \end{bmatrix}$

20. **TEAM PROJECT. Geometric Applications: Curves and Surfaces Through Given Points.** The idea is to get an equation from the vanishing of the determinant of a homogeneous linear system as the condition for a nontrivial solution in Cramer's theorem. We explain the trick for obtaining such a system for the case of a line L through two given points P_1: (x_1, y_1) and P_2: (x_2, y_2). The unknown line is $ax + by = -c$, say. We write it as $ax + by + c \cdot 1 = 0$. To get a nontrivial solution a, b, c, the determinant of the "coefficients" $x, y, 1$ must be zero. The system is

$$\begin{aligned} ax + by + c \cdot 1 &= 0 \quad &&(\text{Line } L) \\ ax_1 + by_1 + c \cdot 1 &= 0 \quad &&(P_1 \text{ on } L) \\ ax_2 + by_2 + c \cdot 1 &= 0 \quad &&(P_2 \text{ on } L). \end{aligned} \tag{12}$$

(a) Line through two points. Derive from $D = 0$ in (12) the familiar formula

$$\frac{x - x_1}{x_1 - x_2} = \frac{y - y_1}{y_1 - y_2}.$$

(b) Plane. Find the analog of (12) for a plane through three given points. Apply it when the points are (1, 1, 1), (3, 2, 6), (5, 0, 5).

(c) Circle. Find a similar formula for a circle in the plane through three given points. Find and sketch the circle through (2, 6), (6, 4), (7, 1).

(d) Sphere. Find the analog of the formula in (c) for a sphere through four given points. Find the sphere through (0, 0, 5), (4, 0, 1), (0, 4, 1), (0, 0, −3) by this formula or by inspection.

(e) General conic section. Find a formula for a general conic section (the vanishing of a determinant of 6th order). Try it out for a quadratic parabola and for a more general conic section of your own choice.

21–25 CRAMER'S RULE

Solve by Cramer's rule. Check by Gauss elimination and back substitution. Show details.

21. $\begin{aligned} 3x - 5y &= 15.5 \\ 6x + 16y &= 5.0 \end{aligned}$

22. $\begin{aligned} 2x - 4y &= -24 \\ 5x + 2y &= 0 \end{aligned}$

23. $\begin{aligned} 3y - 4z &= 16 \\ 2x - 5y + 7z &= -27 \\ -x \qquad - 9z &= 9 \end{aligned}$

24. $\begin{aligned} 3x - 2y + z &= 13 \\ -2x + y + 4z &= 11 \\ x + 4y - 5z &= -31 \end{aligned}$

25. $\begin{aligned} -4w + x + y \qquad &= -10 \\ w - 4x \qquad + z &= 1 \\ w \qquad - 4y + z &= -7 \\ x + y - 4z &= 10 \end{aligned}$

7.8 Inverse of a Matrix. Gauss–Jordan Elimination

***In this section we consider* square *matrices exclusively*.**

The **inverse** of an $n \times n$ matrix $\mathbf{A} = [a_{jk}]$ is denoted by $\mathbf{A}^{-1}$ and is an $n \times n$ matrix such that

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I} \tag{1}$$

where $\mathbf{I}$ is the $n \times n$ unit matrix (see Sec. 7.2).

If $\mathbf{A}$ has an inverse, then $\mathbf{A}$ is called a **nonsingular matrix**. If $\mathbf{A}$ has no inverse, then $\mathbf{A}$ is called a **singular matrix**.

If $\mathbf{A}$ *has an inverse, the inverse is unique.*

Indeed, if both $\mathbf{B}$ and $\mathbf{C}$ are inverses of $\mathbf{A}$, then $\mathbf{AB} = \mathbf{I}$ and $\mathbf{CA} = \mathbf{I}$, so that we obtain the uniqueness from

$$\mathbf{B} = \mathbf{IB} = (\mathbf{CA})\mathbf{B} = \mathbf{C}(\mathbf{AB}) = \mathbf{CI} = \mathbf{C}.$$

We prove next that $\mathbf{A}$ has an inverse (is nonsingular) if and only if it has maximum possible rank n. The proof will also show that $\mathbf{Ax} = \mathbf{b}$ implies $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ provided $\mathbf{A}^{-1}$ exists, and will thus give a motivation for the inverse as well as a relation to linear systems. (But this will ***not*** give a good method of solving $\mathbf{Ax} = \mathbf{b}$ ***numerically*** because the Gauss elimination in Sec. 7.3 requires fewer computations.)

THEOREM 1

Existence of the Inverse

The inverse $\mathbf{A}^{-1}$ *of an* $n \times n$ *matrix* $\mathbf{A}$ *exists if and only if* rank $\mathbf{A} = n$, *thus* (by Theorem 3, Sec. 7.7) *if and only if* $\det \mathbf{A} \neq 0$. *Hence* $\mathbf{A}$ *is nonsingular if* rank $\mathbf{A} = n$, *and is singular if* rank $\mathbf{A} < n$.

PROOF Let $\mathbf{A}$ be a given $n \times n$ matrix and consider the linear system

$$\mathbf{Ax} = \mathbf{b}. \tag{2}$$

If the inverse $\mathbf{A}^{-1}$ exists, then multiplication from the left on both sides and use of (1) gives

$$\mathbf{A}^{-1}\mathbf{Ax} = \mathbf{x} = \mathbf{A}^{-1}\mathbf{b}.$$

This shows that (2) has a solution $\mathbf{x}$, which is unique because, for another solution $\mathbf{u}$, we have $\mathbf{Au} = \mathbf{b}$, so that $\mathbf{u} = \mathbf{A}^{-1}\mathbf{b} = \mathbf{x}$. Hence $\mathbf{A}$ must have rank n by the Fundamental Theorem in Sec. 7.5.

Conversely, let rank $\mathbf{A} = n$. Then by the same theorem, the system (2) has a unique solution $\mathbf{x}$ for any $\mathbf{b}$. Now the back substitution following the Gauss elimination (Sec. 7.3) shows that the components x_j of $\mathbf{x}$ are linear combinations of those of $\mathbf{b}$. Hence we can write

$$\mathbf{x} = \mathbf{Bb} \tag{3}$$

with $\mathbf{B}$ to be determined. Substitution into (2) gives

$$\mathbf{Ax} = \mathbf{A}(\mathbf{Bb}) = (\mathbf{AB})\mathbf{b} = \mathbf{Cb} = \mathbf{b} \qquad (\mathbf{C} = \mathbf{AB})$$

for any $\mathbf{b}$. Hence $\mathbf{C} = \mathbf{AB} = \mathbf{I}$, the unit matrix. Similarly, if we substitute (2) into (3) we get

$$\mathbf{x} = \mathbf{Bb} = \mathbf{B}(\mathbf{Ax}) = (\mathbf{BA})\mathbf{x}$$

for any $\mathbf{x}$ (and $\mathbf{b} = \mathbf{Ax}$). Hence $\mathbf{BA} = \mathbf{I}$. Together, $\mathbf{B} = \mathbf{A}^{-1}$ exists. ∎

Determination of the Inverse by the Gauss–Jordan Method

To actually determine the inverse $\mathbf{A}^{-1}$ of a nonsingular $n \times n$ matrix $\mathbf{A}$, we can use a variant of the Gauss elimination (Sec. 7.3), called the **Gauss–Jordan elimination**.[3] The idea of the method is as follows.

Using $\mathbf{A}$, we form n linear systems

$$\mathbf{Ax}_{(1)} = \mathbf{e}_{(1)}, \quad \cdots, \quad \mathbf{Ax}_{(n)} = \mathbf{e}_{(n)}$$

where the vectors $\mathbf{e}_{(1)}, \cdots, \mathbf{e}_{(n)}$ are the columns of the $n \times n$ unit matrix $\mathbf{I}$; thus, $\mathbf{e}_{(1)} = [1 \quad 0 \quad \cdots \quad 0]^{\mathsf{T}}$, $\mathbf{e}_{(2)} = [0 \quad 1 \quad 0 \quad \cdots \quad 0]^{\mathsf{T}}$, etc. These are n vector equations in the unknown vectors $\mathbf{x}_{(1)}, \cdots, \mathbf{x}_{(n)}$. We combine them into a single matrix equation

[3] WILHELM JORDAN (1842–1899), German geodesist and mathematician. He did important geodesic work in Africa, where he surveyed oases. [See Althoen, S.C. and R. McLaughlin, Gauss–Jordan reduction: A brief history. *American Mathematical Monthly*, Vol. **94**, No. 2 (1987), pp. 130–142.]

We do ***not recommend*** it as a method for solving systems of linear equations, since the number of operations in addition to those of the Gauss elimination is larger than that for back substitution, which the Gauss–Jordan elimination avoids. See also Sec. 20.1.

$\mathbf{AX} = \mathbf{I}$, with the unknown matrix $\mathbf{X}$ having the columns $\mathbf{x}_{(1)}, \cdots, \mathbf{x}_{(n)}$. Correspondingly, we combine the n augmented matrices $[\mathbf{A} \quad \mathbf{e}_{(1)}], \cdots, [\mathbf{A} \quad \mathbf{e}_{(n)}]$ into one wide $n \times 2n$ "augmented matrix" $\widetilde{\mathbf{A}} = [\mathbf{A} \quad \mathbf{I}]$. Now multiplication of $\mathbf{AX} = \mathbf{I}$ by $\mathbf{A}^{-1}$ from the left gives $\mathbf{X} = \mathbf{A}^{-1}\mathbf{I} = \mathbf{A}^{-1}$. Hence, to solve $\mathbf{AX} = \mathbf{I}$ for $\mathbf{X}$, we can apply the Gauss elimination to $\widetilde{\mathbf{A}} = [\mathbf{A} \quad \mathbf{I}]$. This gives a matrix of the form $[\mathbf{U} \quad \mathbf{H}]$ with upper triangular $\mathbf{U}$ because the Gauss elimination triangularizes systems. The Gauss–Jordan method reduces $\mathbf{U}$ by further elementary row operations to diagonal form, in fact to the unit matrix $\mathbf{I}$. This is done by eliminating the entries of $\mathbf{U}$ above the main diagonal and making the diagonal entries all 1 by multiplication (see Example 1). Of course, the method operates on the entire matrix $[\mathbf{U} \quad \mathbf{H}]$, transforming $\mathbf{H}$ into some matrix $\mathbf{K}$, hence the entire $[\mathbf{U} \quad \mathbf{H}]$ to $[\mathbf{I} \quad \mathbf{K}]$. This is the "augmented matrix" of $\mathbf{IX} = \mathbf{K}$. Now $\mathbf{IX} = \mathbf{X} = \mathbf{A}^{-1}$, as shown before. By comparison, $\mathbf{K} = \mathbf{A}^{-1}$, so that we can read $\mathbf{A}^{-1}$ directly from $[\mathbf{I} \quad \mathbf{K}]$.

The following example illustrates the practical details of the method.

EXAMPLE 1 Finding the Inverse of a Matrix by Gauss–Jordan Elimination

Determine the inverse $\mathbf{A}^{-1}$ of

$$\mathbf{A} = \begin{bmatrix} -1 & 1 & 2 \\ 3 & -1 & 1 \\ -1 & 3 & 4 \end{bmatrix}.$$

Solution. We apply the Gauss elimination (Sec. 7.3) to the following $n \times 2n = 3 \times 6$ matrix, where BLUE always refers to the previous matrix.

$$[\mathbf{A} \quad \mathbf{I}] = \left[\begin{array}{rrr|rrr} -1 & 1 & 2 & 1 & 0 & 0 \\ 3 & -1 & 1 & 0 & 1 & 0 \\ -1 & 3 & 4 & 0 & 0 & 1 \end{array}\right]$$

$$\left[\begin{array}{rrr|rrr} -1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 2 & 7 & 3 & 1 & 0 \\ 0 & 2 & 2 & -1 & 0 & 1 \end{array}\right] \begin{array}{l} \\ \text{Row 2} + 3 \text{ Row 1} \\ \text{Row 3} - \text{Row 1} \end{array}$$

$$\left[\begin{array}{rrr|rrr} -1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 2 & 7 & 3 & 1 & 0 \\ 0 & 0 & -5 & -4 & -1 & 1 \end{array}\right] \begin{array}{l} \\ \\ \text{Row 3} - \text{Row 2} \end{array}$$

This is $[\mathbf{U} \quad \mathbf{H}]$ as produced by the Gauss elimination. Now follow the additional Gauss–Jordan steps, reducing $\mathbf{U}$ to $\mathbf{I}$, that is, to diagonal form with entries 1 on the main diagonal.

$$\left[\begin{array}{rrr|rrr} 1 & -1 & -2 & -1 & 0 & 0 \\ 0 & 1 & 3.5 & 1.5 & 0.5 & 0 \\ 0 & 0 & 1 & 0.8 & 0.2 & -0.2 \end{array}\right] \begin{array}{l} -\text{Row 1} \\ 0.5 \text{ Row 2} \\ -0.2 \text{ Row 3} \end{array}$$

$$\left[\begin{array}{rrr|rrr} 1 & -1 & 0 & 0.6 & 0.4 & -0.4 \\ 0 & 1 & 0 & -1.3 & -0.2 & 0.7 \\ 0 & 0 & 1 & 0.8 & 0.2 & -0.2 \end{array}\right] \begin{array}{l} \text{Row 1} + 2 \text{ Row 3} \\ \text{Row 2} - 3.5 \text{ Row 3} \\ \\ \end{array}$$

$$\left[\begin{array}{rrr|rrr} 1 & 0 & 0 & -0.7 & 0.2 & 0.3 \\ 0 & 1 & 0 & -1.3 & -0.2 & 0.7 \\ 0 & 0 & 1 & 0.8 & 0.2 & -0.2 \end{array}\right] \begin{array}{l} \text{Row 1} + \text{Row 2} \\ \\ \\ \end{array}$$

The last three columns constitute $\mathbf{A}^{-1}$. Check:

$$\begin{bmatrix} -1 & 1 & 2 \\ 3 & -1 & 1 \\ -1 & 3 & 4 \end{bmatrix} \begin{bmatrix} -0.7 & 0.2 & 0.3 \\ -1.3 & -0.2 & 0.7 \\ 0.8 & 0.2 & -0.2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Hence $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$. Similarly, $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$. ■

Formulas for Inverses

Since finding the inverse of a matrix is really a problem of solving a system of linear equations, it is not surprising that Cramer's rule (Theorem 4, Sec. 7.7) might come into play. And similarly, as Cramer's rule was useful for theoretical study but not for computation, so too is the explicit formula (4) in the following theorem useful for theoretical considerations but not recommended for actually determining inverse matrices, except for the frequently occurring 2×2 case as given in (4*).

THEOREM 2

Inverse of a Matrix by Determinants

The inverse of a nonsingular $n \times n$ matrix $\mathbf{A} = [a_{jk}]$ is given by

$$\mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} [C_{jk}]^{\mathsf{T}} = \frac{1}{\det \mathbf{A}} \begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \cdot & \cdot & \cdots & \cdot \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix}, \tag{4}$$

where C_{jk} is the cofactor of a_{jk} in $\det \mathbf{A}$ (see Sec. 7.7). (CAUTION! Note well that in $\mathbf{A}^{-1}$, the cofactor C_{jk} occupies the same place as a_{kj} (not a_{jk}) does in $\mathbf{A}$.)

In particular, the inverse of

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad \textit{is} \quad \mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}. \tag{4*}$$

PROOF We denote the right side of (4) by $\mathbf{B}$ and show that $\mathbf{BA} = \mathbf{I}$. We first write

$$\mathbf{BA} = \mathbf{G} = [g_{kl}] \tag{5}$$

and then show that $\mathbf{G} = \mathbf{I}$. Now by the definition of matrix multiplication and because of the form of $\mathbf{B}$ in (4), we obtain (CAUTION! C_{sk}, not C_{ks})

$$g_{kl} = \sum_{s=1}^{n} \frac{C_{sk}}{\det \mathbf{A}} a_{sl} = \frac{1}{\det \mathbf{A}} (a_{1l}C_{1k} + \cdots + a_{nl}C_{nk}). \tag{6}$$

Now (9) and (10) in Sec. 7.7 show that the sum $(\cdots)$ on the right is $D = \det \mathbf{A}$ when $l = k$, and is zero when $l \neq k$. Hence

$$g_{kk} = \frac{1}{\det \mathbf{A}} \det \mathbf{A} = 1,$$

$$g_{kl} = 0 \quad (l \neq k).$$

In particular, for $n = 2$ we have in (4), in the first row, $C_{11} = a_{22}$, $C_{21} = -a_{12}$ and, in the second row, $C_{12} = -a_{21}$, $C_{22} = a_{11}$. This gives (4*). ■

The special case $n = 2$ occurs quite frequently in geometric and other applications. You may perhaps want to memorize formula (4*). Example 2 gives an illustration of (4*).

EXAMPLE 2 Inverse of a 2 × 2 Matrix by Determinants

$$\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 2 & 4 \end{bmatrix}, \quad \mathbf{A}^{-1} = \frac{1}{10}\begin{bmatrix} 4 & -1 \\ -2 & 3 \end{bmatrix} = \begin{bmatrix} 0.4 & -0.1 \\ -0.2 & 0.3 \end{bmatrix}$$ ■

EXAMPLE 3 Further Illustration of Theorem 2

Using (4), find the inverse of

$$\mathbf{A} = \begin{bmatrix} -1 & 1 & 2 \\ 3 & -1 & 1 \\ -1 & 3 & 4 \end{bmatrix}.$$

Solution. We obtain $\det \mathbf{A} = -1(-7) - 1 \cdot 13 + 2 \cdot 8 = 10$, and in (4),

$$C_{11} = \begin{vmatrix} -1 & 1 \\ 3 & 4 \end{vmatrix} = -7, \qquad C_{21} = -\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = 2, \qquad C_{31} = \begin{vmatrix} 1 & 2 \\ -1 & 1 \end{vmatrix} = 3,$$

$$C_{12} = -\begin{vmatrix} 3 & 1 \\ -1 & 4 \end{vmatrix} = -13, \qquad C_{22} = \begin{vmatrix} -1 & 2 \\ -1 & 4 \end{vmatrix} = -2, \qquad C_{32} = -\begin{vmatrix} -1 & 2 \\ 3 & 1 \end{vmatrix} = 7,$$

$$C_{13} = \begin{vmatrix} 3 & -1 \\ -1 & 3 \end{vmatrix} = 8, \qquad C_{23} = -\begin{vmatrix} -1 & 1 \\ -1 & 3 \end{vmatrix} = 2, \qquad C_{33} = \begin{vmatrix} -1 & 1 \\ 3 & -1 \end{vmatrix} = -2,$$

so that by (4), in agreement with Example 1,

$$\mathbf{A}^{-1} = \begin{bmatrix} -0.7 & 0.2 & 0.3 \\ -1.3 & -0.2 & 0.7 \\ 0.8 & 0.2 & -0.2 \end{bmatrix}.$$ ■

Diagonal matrices $\mathbf{A} = [a_{jk}]$, $a_{jk} = 0$ when $j \neq k$, have an inverse if and only if all $a_{jj} \neq 0$. Then $\mathbf{A}^{-1}$ is diagonal, too, with entries $1/a_{11}, \cdots, 1/a_{nn}$.

PROOF For a diagonal matrix we have in (4)

$$\frac{C_{11}}{D} = \frac{a_{22} \cdots a_{nn}}{a_{11}a_{22} \cdots a_{nn}} = \frac{1}{a_{11}}, \qquad \text{etc.}$$ ■

EXAMPLE 4 Inverse of a Diagonal Matrix

Let

$$\mathbf{A} = \begin{bmatrix} -0.5 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Then we obtain the inverse $\mathbf{A}^{-1}$ by inverting each individual diagonal element of $\mathbf{A}$, that is, by taking $1/(-0.5)$, $\frac{1}{4}$, and $\frac{1}{1}$ as the diagonal entries of $\mathbf{A}^{-1}$, that is,

$$\mathbf{A}^{-1} = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 0.25 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$ ■

Products can be inverted by taking the inverse of each factor and multiplying these inverses ***in reverse order***,

$$\textbf{(7)} \qquad \boxed{(\mathbf{AC})^{-1} = \mathbf{C}^{-1}\mathbf{A}^{-1}.}$$

Hence for more than two factors,

$$(8) \qquad (\mathbf{AC} \cdots \mathbf{PQ})^{-1} = \mathbf{Q}^{-1}\mathbf{P}^{-1} \cdots \mathbf{C}^{-1}\mathbf{A}^{-1}.$$

PROOF The idea is to start from (1) for $\mathbf{AC}$ instead of $\mathbf{A}$, that is, $\mathbf{AC}(\mathbf{AC})^{-1} = \mathbf{I}$, and multiply it on both sides from the left, first by $\mathbf{A}^{-1}$, which because of $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$ gives

$$\begin{aligned} \mathbf{A}^{-1}\mathbf{AC}(\mathbf{AC})^{-1} &= \mathbf{C}(\mathbf{AC})^{-1} \\ &= \mathbf{A}^{-1}\mathbf{I} = \mathbf{A}^{-1}, \end{aligned}$$

and then multiplying this on both sides from the left, this time by $\mathbf{C}^{-1}$ and by using $\mathbf{C}^{-1}\mathbf{C} = \mathbf{I}$,

$$\mathbf{C}^{-1}\mathbf{C}(\mathbf{AC})^{-1} = (\mathbf{AC})^{-1} = \mathbf{C}^{-1}\mathbf{A}^{-1}.$$

This proves (7), and from it, (8) follows by induction. ■

We also note that *the inverse of the inverse is the given matrix*, as you may prove,

$$(9) \qquad (\mathbf{A}^{-1})^{-1} = \mathbf{A}.$$

Unusual Properties of Matrix Multiplication. Cancellation Laws

Section 7.2 contains warnings that some properties of matrix multiplication deviate from those for numbers, and we are now able to explain the restricted validity of the so-called **cancellation laws** [2] and [3] below, using rank and inverse, concepts that were not yet available in Sec. 7.2. The deviations from the usual are of great practical importance and must be carefully observed. They are as follows.

[1] Matrix multiplication is not commutative, that is, in general we have

$$\mathbf{AB} \neq \mathbf{BA}.$$

[2] $\mathbf{AB} = \mathbf{0}$ does not generally imply $\mathbf{A} = \mathbf{0}$ or $\mathbf{B} = \mathbf{0}$ (or $\mathbf{BA} = \mathbf{0}$); for example,

$$\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

[3] $\mathbf{AC} = \mathbf{AD}$ does not generally imply $\mathbf{C} = \mathbf{D}$ (even when $\mathbf{A} \neq \mathbf{0}$).

Complete answers to [2] and [3] are contained in the following theorem.

THEOREM 3

Cancellation Laws

Let $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ *be* $n \times n$ *matrices. Then:*

(a) *If* rank $\mathbf{A} = n$ *and* $\mathbf{AB} = \mathbf{AC}$, *then* $\mathbf{B} = \mathbf{C}$.

(b) *If* rank $\mathbf{A} = n$, *then* $\mathbf{AB} = \mathbf{0}$ *implies* $\mathbf{B} = \mathbf{0}$. *Hence if* $\mathbf{AB} = \mathbf{0}$, *but* $\mathbf{A} \neq \mathbf{0}$ *as well as* $\mathbf{B} \neq \mathbf{0}$, *then* rank $\mathbf{A} < n$ *and* rank $\mathbf{B} < n$.

(c) *If* $\mathbf{A}$ *is singular, so are* $\mathbf{BA}$ *and* $\mathbf{AB}$.

PROOF **(a)** The inverse of $\mathbf{A}$ exists by Theorem 1. Multiplication by $\mathbf{A}^{-1}$ from the left gives $\mathbf{A}^{-1}\mathbf{AB} = \mathbf{A}^{-1}\mathbf{AC}$, hence $\mathbf{B} = \mathbf{C}$.

(b) Let rank $\mathbf{A} = n$. Then $\mathbf{A}^{-1}$ exists, and $\mathbf{AB} = \mathbf{0}$ implies $\mathbf{A}^{-1}\mathbf{AB} = \mathbf{B} = \mathbf{0}$. Similarly when rank $\mathbf{B} = n$. This implies the second statement in (b).

($\mathbf{c}_1$) Rank $\mathbf{A} < n$ by Theorem 1. Hence $\mathbf{Ax} = \mathbf{0}$ has nontrivial solutions by Theorem 2 in Sec. 7.5. Multiplication by $\mathbf{B}$ shows that these solutions are also solutions of $\mathbf{BAx} = \mathbf{0}$, so that rank $(\mathbf{BA}) < n$ by Theorem 2 in Sec. 7.5 and $\mathbf{BA}$ is singular by Theorem 1.

($\mathbf{c}_2$) $\mathbf{A}^{\mathsf{T}}$ is singular by Theorem 2(d) in Sec. 7.7. Hence $\mathbf{B}^{\mathsf{T}}\mathbf{A}^{\mathsf{T}}$ is singular by part (c_1), and is equal to $(\mathbf{AB})^{\mathsf{T}}$ by (10d) in Sec. 7.2. Hence $\mathbf{AB}$ is singular by Theorem 2(d) in Sec. 7.7. ■

Determinants of Matrix Products

The determinant of a matrix product $\mathbf{AB}$ or $\mathbf{BA}$ can be written as the product of the determinants of the factors, and it is interesting that det $\mathbf{AB}$ = det $\mathbf{BA}$, although $\mathbf{AB} \neq \mathbf{BA}$ in general. The corresponding formula (10) is needed occasionally and can be obtained by Gauss–Jordan elimination (see Example 1) and from the theorem just proved.

THEOREM 4

Determinant of a Product of Matrices

For any $n \times n$ *matrices* $\mathbf{A}$ *and* $\mathbf{B}$,

$$\det(\mathbf{AB}) = \det(\mathbf{BA}) = \det \mathbf{A} \det \mathbf{B}. \tag{10}$$

PROOF If **A** or **B** is singular, so are **AB** and **BA** by Theorem 3(c), and (10) reduces to $0 = 0$ by Theorem 3 in Sec. 7.7.

Now let **A** and **B** be nonsingular. Then we can reduce **A** to a diagonal matrix $\hat{\mathbf{A}} = [a_{jk}]$ by Gauss–Jordan steps. Under these operations, det **A** retains its value, by Theorem 1 in Sec. 7.7, (a) and (b) [not (c)] except perhaps for a sign reversal in row interchanging when pivoting. But the same operations reduce **AB** to $\hat{\mathbf{A}}\mathbf{B}$ with the same effect on det (**AB**). Hence it remains to prove (10) for $\hat{\mathbf{A}}\mathbf{B}$; written out,

$$\hat{\mathbf{A}}\mathbf{B} = \begin{bmatrix} \hat{a}_{11} & 0 & \cdots & 0 \\ 0 & \hat{a}_{22} & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \cdots & \hat{a}_{nn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ & & \vdots & \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{bmatrix}$$

$$= \begin{bmatrix} \hat{a}_{11}b_{11} & \hat{a}_{11}b_{12} & \cdots & \hat{a}_{11}b_{1n} \\ \hat{a}_{22}b_{21} & \hat{a}_{22}b_{22} & \cdots & \hat{a}_{22}b_{2n} \\ & & \vdots & \\ \hat{a}_{nn}b_{n1} & \hat{a}_{nn}b_{n2} & \cdots & \hat{a}_{nn}b_{nn} \end{bmatrix}.$$

We now take the determinant det ($\hat{\mathbf{A}}\mathbf{B}$). On the right we can take out a factor $\hat{a}_{11}$ from the first row, $\hat{a}_{22}$ from the second, $\cdots$, $\hat{a}_{nn}$ from the nth. But this product $\hat{a}_{11}\hat{a}_{22}\cdots\hat{a}_{nn}$ equals det $\hat{\mathbf{A}}$ because $\hat{\mathbf{A}}$ is diagonal. The remaining determinant is det **B**. This proves (10) for det (**AB**), and the proof for det (**BA**) follows by the same idea. ■

This completes our discussion of linear systems (Secs. 7.3–7.8). Section 7.9 on vector spaces and linear transformations is optional. ***Numeric methods*** are discussed in Secs. 20.1–20.4, which are independent of other sections on numerics.

PROBLEM SET 7.8

1–10 INVERSE

Find the inverse by Gauss–Jordan (or by (4*) if $n = 2$). Check by using (1).

1. $\begin{bmatrix} 1.80 & -2.32 \\ -0.25 & 0.60 \end{bmatrix}$

2. $\begin{bmatrix} \cos 2\theta & \sin 2\theta \\ -\sin 2\theta & \cos 2\theta \end{bmatrix}$

3. $\begin{bmatrix} 0.3 & -0.1 & 0.5 \\ 2 & 6 & 4 \\ 5 & 0 & 9 \end{bmatrix}$

4. $\begin{bmatrix} 0 & 0 & 0.1 \\ 0 & -0.4 & 0 \\ 2.5 & 0 & 0 \end{bmatrix}$

5. $\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 5 & 4 & 1 \end{bmatrix}$

6. $\begin{bmatrix} -4 & 0 & 0 \\ 0 & 8 & 13 \\ 0 & 3 & 5 \end{bmatrix}$

7. $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

8. $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$

9. $\begin{bmatrix} 0 & 8 & 0 \\ 0 & 0 & 4 \\ 2 & 0 & 0 \end{bmatrix}$

10. $\begin{bmatrix} \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \end{bmatrix}$

11–18 SOME GENERAL FORMULAS

11. Inverse of the square. Verify $(\mathbf{A}^2)^{-1} = (\mathbf{A}^{-1})^2$ for **A** in Prob. 1.

12. Prove the formula in Prob. 11.

13. Inverse of the transpose. Verify $(\mathbf{A}^\mathsf{T})^{-1} = (\mathbf{A}^{-1})^\mathsf{T}$ for $\mathbf{A}$ in Prob. 1.

14. Prove the formula in Prob. 13.

15. Inverse of the inverse. Prove that $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$.

16. Rotation. Give an application of the matrix in Prob. 2 that makes the form of the inverse obvious.

17. Triangular matrix. Is the inverse of a triangular matrix always triangular (as in Prob. 5)? Give reason.

18. Row interchange. Same task as in Prob. 16 for the matrix in Prob. 7.

19–20 FORMULA (4)

Formula (4) is occasionally needed in theory. To understand it, apply it and check the result by Gauss–Jordan:

19. In Prob. 3

20. In Prob. 6

7.9 Vector Spaces, Inner Product Spaces, Linear Transformations *Optional*

We have captured the essence of vector spaces in Sec. 7.4. There we dealt with ***special vector spaces*** that arose quite naturally in the context of matrices and linear systems. The elements of these vector spaces, called *vectors*, satisfied rules (3) and (4) of Sec. 7.1 (which were similar to those for numbers). These special vector spaces were generated by *spans,* that is, linear combination of finitely many vectors. Furthermore, each such vector had n real numbers as *components*. Review this material before going on.

We can generalize this idea by taking *all* vectors with n real numbers as components and obtain the very important ***real n-dimensional vector space*** R^n. The vectors are known as "real vectors." Thus, each vector in R^n is an ordered n-tuple of real numbers.

Now we can consider special values for n. For $n = 2$, we obtain R^2, the vector space of all ordered pairs, which correspond to the **vectors in the plane**. For $n = 3$, we obtain R^3, the vector space of all ordered triples, which are the **vectors in 3-space**. These vectors have wide applications in mechanics, geometry, and calculus and are basic to the engineer and physicist.

Similarly, if we take all ordered n-tuples of *complex numbers* as vectors and complex numbers as scalars, we obtain the **complex vector space** C^n, which we shall consider in Sec. 8.5.

Furthermore, there are other sets of practical interest consisting of matrices, functions, transformations, or others for which addition and scalar multiplication can be defined in an almost natural way so that they too form vector spaces.

It is perhaps not too great an intellectual jump to create, from the ***concrete model*** R^n, the ***abstract concept*** of a ***real vector space*** V by taking the basic properties (3) and (4) in Sec. 7.1 as axioms. In this way, the definition of a real vector space arises.

DEFINITION

Real Vector Space

A nonempty set V of elements $\mathbf{a}, \mathbf{b}, \cdots$ is called a **real vector space** (or *real linear space*), and these elements are called **vectors** (regardless of their nature, which will come out from the context or will be left arbitrary) if, in V, there are defined two algebraic operations (called *vector addition* and *scalar multiplication*) as follows.

I. Vector addition associates with every pair of vectors $\mathbf{a}$ and $\mathbf{b}$ of V a unique vector of V, called the *sum* of $\mathbf{a}$ and $\mathbf{b}$ and denoted by $\mathbf{a} + \mathbf{b}$, such that the following axioms are satisfied.

I.1 *Commutativity.* For any two vectors $\mathbf{a}$ and $\mathbf{b}$ of V,

$$\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}.$$

I.2 *Associativity.* For any three vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ of V,

$$(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c}) \qquad (\text{written } \mathbf{a} + \mathbf{b} + \mathbf{c}).$$

I.3 There is a unique vector in V, called the *zero vector* and denoted by $\mathbf{0}$, such that for every $\mathbf{a}$ in V,

$$\mathbf{a} + \mathbf{0} = \mathbf{a}.$$

I.4 For every $\mathbf{a}$ in V there is a unique vector in V that is denoted by $-\mathbf{a}$ and is such that

$$\mathbf{a} + (-\mathbf{a}) = \mathbf{0}.$$

II. Scalar multiplication. The real numbers are called **scalars**. Scalar multiplication associates with every $\mathbf{a}$ in V and every scalar c a unique vector of V, called the *product* of c and $\mathbf{a}$ and denoted by $c\mathbf{a}$ (or $\mathbf{a}c$) such that the following axioms are satisfied.

II.1 *Distributivity.* For every scalar c and vectors $\mathbf{a}$ and $\mathbf{b}$ in V,

$$c(\mathbf{a} + \mathbf{b}) = c\mathbf{a} + c\mathbf{b}.$$

II.2 *Distributivity.* For all scalars c and k and every $\mathbf{a}$ in V,

$$(c + k)\mathbf{a} = c\mathbf{a} + k\mathbf{a}.$$

II.3 *Associativity.* For all scalars c and k and every $\mathbf{a}$ in V,

$$c(k\mathbf{a}) = (ck)\mathbf{a} \qquad (\text{written } ck\mathbf{a}).$$

II.4 For every $\mathbf{a}$ in V,

$$1\mathbf{a} = \mathbf{a}.$$

If, in the above definition, we take complex numbers as scalars instead of real numbers, we obtain the axiomatic definition of a **complex vector space**.

Take a look at the axioms in the above definition. Each axiom stands on its own: It is concise, useful, and it expresses a simple property of V. There are as few axioms as possible and together they express ***all*** the desired properties of V. Selecting good axioms is a process of trial and error that often extends over a long period of time. But once agreed upon, axioms become *standard* such as the ones in the definition of a real vector space.

The following concepts related to a vector space are exactly defined as those given in Sec. 7.4. Indeed, a **linear combination** of vectors $a_{(1)}, \cdots, a_{(m)}$ in a vector space V is an expression

$$c_1\mathbf{a}_{(1)} + \cdots + c_m\mathbf{a}_m \qquad (c_1, \cdots, c_m \text{ any scalars}).$$

These vectors form a **linearly independent set** (briefly, they are called **linearly independent**) if

$$c_1\mathbf{a}_{(1)} + \cdots + c_m\mathbf{a}_{(m)} = \mathbf{0} \tag{1}$$

implies that $c_1 = 0, \cdots, c_m = 0$. Otherwise, if (1) also holds with scalars not all zero, the vectors are called **linearly dependent**.

Note that (1) with $m = 1$ is $c\mathbf{a} = \mathbf{0}$ and shows that a single vector $\mathbf{a}$ is linearly independent if and only if $\mathbf{a} \neq \mathbf{0}$.

V has **dimension *n***, or is ***n*-dimensional**, if it contains a linearly independent set of n vectors, whereas any set of more than n vectors in V is linearly dependent. That set of n linearly independent vectors is called a **basis** for V. Then every vector in V can be written as a linear combination of the basis vectors. Furthermore, for a given basis, this representation is unique (see Prob. 2).

EXAMPLE 1 Vector Space of Matrices

The real 2×2 matrices form a four-dimensional real vector space. A basis is

$$\mathbf{B}_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{B}_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{B}_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad \mathbf{B}_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

because any 2×2 matrix $\mathbf{A} = [a_{jk}]$ has a unique representation $\mathbf{A} = a_{11}\mathbf{B}_{11} + a_{12}\mathbf{B}_{12} + a_{21}\mathbf{B}_{21} + a_{22}\mathbf{B}_{22}$. Similarly, the real $m \times n$ matrices with fixed m and n form an mn-dimensional vector space. What is the dimension of the vector space of all 3×3 skew-symmetric matrices? Can you find a basis? ■

EXAMPLE 2 Vector Space of Polynomials

The set of all constant, linear, and quadratic polynomials in x together is a vector space of dimension 3 with basis $\{1, x, x^2\}$ under the usual addition and multiplication by real numbers because these two operations give polynomials not exceeding degree 2. What is the dimension of the vector space of all polynomials of degree not exceeding a given fixed n? Can you find a basis? ■

If a vector space V contains a linearly independent set of n vectors for every n, no matter how large, then V is called **infinite dimensional**, as opposed to a *finite dimensional* (n-dimensional) vector space just defined. An example of an infinite dimensional vector space is the space of all continuous functions on some interval $[a, b]$ of the x-axis, as we mention without proof.

Inner Product Spaces

If $\mathbf{a}$ and $\mathbf{b}$ are vectors in R^n, regarded as column vectors, we can form the product $\mathbf{a}^\mathsf{T}\mathbf{b}$. This is a 1×1 matrix, which we can identify with its single entry, that is, with a number.

This product is called the **inner product** or **dot product** of $\mathbf{a}$ and $\mathbf{b}$. Other notations for it are $(\mathbf{a}, \mathbf{b})$ and $\mathbf{a} \cdot \mathbf{b}$. Thus

$$\mathbf{a}^\mathsf{T}\mathbf{b} = (\mathbf{a}, \mathbf{b}) = \mathbf{a} \cdot \mathbf{b} = [a_1 \cdots a_n] \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = \sum_{i=1}^{n} a_l b_l = a_1 b_1 + \cdots + a_n b_n.$$

We now extend this concept to general real vector spaces by taking basic properties of $(\mathbf{a}, \mathbf{b})$ as axioms for an "abstract inner product" $(\mathbf{a}, \mathbf{b})$ as follows.

DEFINITION

Real Inner Product Space

A real vector space V is called a **real inner product space** (or *real pre-Hilbert*[4] *space*) if it has the following property. With every pair of vectors $\mathbf{a}$ and $\mathbf{b}$ in V there is associated a real number, which is denoted by $(\mathbf{a}, \mathbf{b})$ and is called the **inner product** of $\mathbf{a}$ and $\mathbf{b}$, such that the following axioms are satisfied.

I. For all scalars q_1 and q_2 and all vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ in V,

$$(q_1\mathbf{a} + q_2\mathbf{b}, \mathbf{c}) = q_1(\mathbf{a}, \mathbf{c}) + q_2(\mathbf{b}, \mathbf{c}) \qquad (\textit{Linearity}).$$

II. For all vectors $\mathbf{a}$ and $\mathbf{b}$ in V,

$$(\mathbf{a}, \mathbf{b}) = (\mathbf{b}, \mathbf{a}) \qquad (\textit{Symmetry}).$$

III. For every $\mathbf{a}$ in V,

$$\left.\begin{aligned} (\mathbf{a}, \mathbf{a}) &\geqq 0, \\ (\mathbf{a}, \mathbf{a}) = 0 \quad &\text{if and only if} \quad \mathbf{a} = \mathbf{0} \end{aligned}\right\} \qquad (\textit{Positive-definiteness}).$$

Vectors whose inner product is zero are called **orthogonal**.
The *length* or **norm** of a vector in V is defined by

$$\|\mathbf{a}\| = \sqrt{(\mathbf{a}, \mathbf{a})} \quad (\geqq 0). \tag{2}$$

A vector of norm 1 is called a **unit vector**.

[4]DAVID HILBERT (1862–1943), great German mathematician, taught at Königsberg and Göttingen and was the creator of the famous Göttingen mathematical school. He is known for his basic work in algebra, the calculus of variations, integral equations, functional analysis, and mathematical logic. His "Foundations of Geometry" helped the axiomatic method to gain general recognition. His famous 23 problems (presented in 1900 at the International Congress of Mathematicians in Paris) considerably influenced the development of modern mathematics.

If V is finite dimensional, it is actually a so-called *Hilbert space*; see [GenRef7], p. 128, listed in App. 1.

From these axioms and from (2) one can derive the basic inequality

$$|(\mathbf{a}, \mathbf{b})| \leqq \|\mathbf{a}\| \, \|\mathbf{b}\| \qquad (3)$$

(*Cauchy–Schwarz*[5] *inequality*).

From this follows

$$\|\mathbf{a} + \mathbf{b}\| \leqq \|\mathbf{a}\| + \|\mathbf{b}\| \qquad (4)$$

(*Triangle inequality*).

A simple direct calculation gives

$$\|\mathbf{a} + \mathbf{b}\|^2 + \|\mathbf{a} - \mathbf{b}\|^2 = 2(\|\mathbf{a}\|^2 + \|\mathbf{b}\|^2) \qquad (5)$$

(*Parallelogram equality*).

EXAMPLE 3 ***n*-Dimensional Euclidean Space**

R^n with the inner product

$$(\mathbf{a}, \mathbf{b}) = \mathbf{a}^\mathsf{T}\mathbf{b} = a_1 b_1 + \cdots + a_n b_n \qquad (6)$$

(where both **a** and **b** are *column* vectors) is called the ***n*-dimensional Euclidean space** and is denoted by E^n or again simply by R^n. Axioms I–III hold, as direct calculation shows. Equation (2) gives the "**Euclidean norm**"

$$\|\mathbf{a}\| = \sqrt{(\mathbf{a}, \mathbf{a})} = \sqrt{\mathbf{a}^\mathsf{T}\mathbf{a}} = \sqrt{a_1^2 + \cdots + a_n^2}. \qquad (7)$$

■

EXAMPLE 4 **An Inner Product for Functions. Function Space**

The set of all real-valued continuous functions $f(x), g(x), \cdots$ on a given interval $\alpha \leqq x \leqq \beta$ is a real vector space under the usual addition of functions and multiplication by scalars (real numbers). On this "**function space**" we can define an inner product by the integral

$$(f, g) = \int_\alpha^\beta f(x)\, g(x)\, dx. \qquad (8)$$

Axioms I–III can be verified by direct calculation. Equation (2) gives the norm

$$\|f\| = \sqrt{(f, f)} = \sqrt{\int_\alpha^\beta f(x)^2\, dx}. \qquad (9)$$

■

Our examples give a first impression of the great generality of the abstract concepts of vector spaces and inner product spaces. Further details belong to more advanced courses (on functional analysis, meaning abstract modern analysis; see [GenRef7] listed in App. 1) and cannot be discussed here. Instead we now take up a related topic where matrices play a central role.

Linear Transformations

Let X and Y be any vector spaces. To each vector **x** in X we assign a unique vector **y** in Y. Then we say that a **mapping** (or **transformation** or **operator**) of X into Y is given. Such a mapping is denoted by a capital letter, say F. The vector **y** in Y assigned to a vector **x** in X is called the **image** of **x** under F and is denoted by $F(\mathbf{x})$ [or $F\mathbf{x}$, without parentheses].

[5]HERMANN AMANDUS SCHWARZ (1843–1921). German mathematician, known by his work in complex analysis (conformal mapping) and differential geometry. For Cauchy see Sec. 2.5.

F is called a **linear mapping** or **linear transformation** if, for all vectors $\mathbf{v}$ and $\mathbf{x}$ in X and scalars c,

$$\begin{aligned} F(\mathbf{v}+\mathbf{x}) &= F(\mathbf{v})+F(\mathbf{x}) \\ F(c\mathbf{x}) &= cF(\mathbf{x}). \end{aligned} \tag{10}$$

Linear Transformation of Space R^n into Space R^m

From now on we let $X = R^n$ and $Y = R^m$. Then any real $m \times n$ matrix $\mathbf{A} = [a_{jk}]$ gives a transformation of R^n into R^m,

$$\mathbf{y} = \mathbf{A}\mathbf{x}. \tag{11}$$

Since $\mathbf{A}(\mathbf{u}+\mathbf{x}) = \mathbf{A}\mathbf{u} + \mathbf{A}\mathbf{x}$ and $\mathbf{A}(c\mathbf{x}) = c\mathbf{A}\mathbf{x}$, this transformation is linear.

We show that, conversely, every linear transformation F of R^n into R^m can be given in terms of an $m \times n$ matrix $\mathbf{A}$, after a basis for R^n and a basis for R^m have been chosen. This can be proved as follows.

Let $\mathbf{e}_{(1)}, \cdots, \mathbf{e}_{(n)}$ be any basis for R^n. Then every $\mathbf{x}$ in R^n has a unique representation

$$\mathbf{x} = x_1\mathbf{e}_{(1)} + \cdots + x_n\mathbf{e}_{(n)}.$$

Since F is linear, this representation implies for the image $F(\mathbf{x})$:

$$F(\mathbf{x}) = F(x_1\mathbf{e}_{(1)} + \cdots + x_n\mathbf{e}_{(n)}) = x_1F(\mathbf{e}_{(1)}) + \cdots + x_nF(\mathbf{e}_{(n)}).$$

Hence F is uniquely determined by the images of the vectors of a basis for R^n. We now choose for R^n the "**standard basis**"

$$\mathbf{e}_{(1)} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathbf{e}_{(2)} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \cdots, \quad \mathbf{e}_{(n)} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \tag{12}$$

where $\mathbf{e}_{(j)}$ has its jth component equal to 1 and all others 0. We show that we can now determine an $m \times n$ matrix $\mathbf{A} = [a_{jk}]$ such that for every $\mathbf{x}$ in R^n and image $\mathbf{y} = F(\mathbf{x})$ in R^m,

$$\mathbf{y} = F(\mathbf{x}) = \mathbf{A}\mathbf{x}.$$

Indeed, from the image $\mathbf{y}^{(1)} = F(\mathbf{e}_{(1)})$ of $\mathbf{e}_{(1)}$ we get the condition

$$\mathbf{y}^{(1)} = \begin{bmatrix} y_1^{(1)} \\ y_2^{(1)} \\ \vdots \\ y_m^{(1)} \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mm} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

from which we can determine the first column of $\mathbf{A}$, namely $a_{11} = y_1^{(1)}$, $a_{21} = y_2^{(1)}$, $\cdots$, $a_{m1} = y_m^{(1)}$. Similarly, from the image of $\mathbf{e}_{(2)}$ we get the second column of $\mathbf{A}$, and so on. This completes the proof. ■

We say that $\mathbf{A}$ **represents** F, or *is a representation of* F, with respect to the bases for R^n and R^m. Quite generally, the purpose of a "**representation**" is the replacement of one object of study by another object whose properties are more readily apparent.

In three-dimensional Euclidean space E^3 the standard basis is usually written $\mathbf{e}_{(1)} = \mathbf{i}$, $\mathbf{e}_{(2)} = \mathbf{j}$, $\mathbf{e}_{(3)} = \mathbf{k}$. Thus,

$$\mathbf{i} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{j} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{k} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}. \tag{13}$$

These are the three unit vectors in the positive directions of the axes of the **Cartesian coordinate system in space**, that is, the usual coordinate system with the same scale of measurement on the three mutually perpendicular coordinate axes.

EXAMPLE 5 **Linear Transformations**

Interpreted as transformations of Cartesian coordinates in the plane, the matrices

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \begin{bmatrix} a & 0 \\ 0 & 1 \end{bmatrix}$$

represent a reflection in the line $x_2 = x_1$, a reflection in the x_1-axis, a reflection in the origin, and a stretch (when $a > 1$, or a contraction when $0 < a < 1$) in the x_1-direction, respectively. ■

EXAMPLE 6 **Linear Transformations**

Our discussion preceding Example 5 is simpler than it may look at first sight. To see this, find $\mathbf{A}$ representing the linear transformation that maps (x_1, x_2) onto $(2x_1 - 5x_2, 3x_1 + 4x_2)$.

Solution. Obviously, the transformation is

$$\begin{aligned} y_1 &= 2x_1 - 5x_2 \\ y_2 &= 3x_1 + 4x_2. \end{aligned}$$

From this we can directly see that the matrix is

$$\mathbf{A} = \begin{bmatrix} 2 & -5 \\ 3 & 4 \end{bmatrix}. \qquad \text{Check:} \quad \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 2 & -5 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2x_1 - 5x_2 \\ 3x_1 + 4x_2 \end{bmatrix}.$$ ■

If $\mathbf{A}$ in (11) is square, $n \times n$, then (11) maps R^n into R^n. If this $\mathbf{A}$ is nonsingular, so that $\mathbf{A}^{-1}$ exists (see Sec. 7.8), then multiplication of (11) by $\mathbf{A}^{-1}$ from the left and use of $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$ gives the **inverse transformation**

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{y}. \tag{14}$$

It maps every $\mathbf{y} = \mathbf{y}_0$ onto that $\mathbf{x}$, which by (11) is mapped onto $\mathbf{y}_0$. *The inverse of a linear transformation is itself linear,* because it is given by a matrix, as (14) shows.

Composition of Linear Transformations

We want to give you a flavor of how linear transformations in general vector spaces work. You will notice, if you read carefully, that definitions and verifications (Example 7) strictly follow the given rules and you can think your way through the material by going in a slow systematic fashion.

The last operation we want to discuss is composition of linear transformations. Let X, Y, W be general vector spaces. As before, let F be a linear transformation from X to Y. Let G be a linear transformation from W to X. Then we denote, by H, the **composition** of F and G, that is,

$$H = F \circ G = FG = F(G),$$

which means *we take transformation G and then apply transformation F to it* (***in that order!***, *i.e. you go from left to right*).

Now, to give this a more concrete meaning, if we let $\mathbf{w}$ be a vector in W, then $G(\mathbf{w})$ is a vector in X and $F(G(\mathbf{w}))$ is a vector in Y. Thus, H maps W to Y, and we can write

$$H(\mathbf{w}) = (F \circ G)(\mathbf{w}) = (FG)(\mathbf{w}) = F(G(\mathbf{w})), \tag{15}$$

which completes the definition of composition in a general vector space setting. But is composition really linear? To check this we have to verify that $H,$ as defined in (15), obeys the two equations of (10).

EXAMPLE 7 The Composition of Linear Transformations Is Linear

To show that H is indeed linear we must show that (10) holds. We have, for two vectors $\mathbf{w}_1$, $\mathbf{w}_2$ in W,

$$\begin{aligned} H(\mathbf{w}_1 + \mathbf{w}_2) &= (F \circ G)(\mathbf{w}_1 + \mathbf{w}_2) \\ &= F(G(\mathbf{w}_1 + \mathbf{w}_2)) \\ &= F(G(\mathbf{w}_1) + G(\mathbf{w}_2)) && \text{(by linearity of } G) \\ &= F(G(\mathbf{w}_1)) + F(G(\mathbf{w}_2)) && \text{(by linearity of } F) \\ &= (F \circ G)(\mathbf{w}_1) + (F \circ G)(\mathbf{w}_2) && \text{(by (15))} \\ &= H(\mathbf{w}_1) + H(\mathbf{w}_2) && \text{(by definition of } H). \end{aligned}$$

Similarly, $H(c\mathbf{w}_2) = (F \circ G)(c\mathbf{w}_2) = F(G(c\mathbf{w}_2)) = F(c(G(\mathbf{w}_2))$

$$= cF(G(\mathbf{w}_2)) = c(F \circ G)(\mathbf{w}_2) = cH(\mathbf{w}_2).$$ ■

We defined composition as a linear transformation in a general vector space setting and showed that the composition of linear transformations is indeed linear.

Next we want to relate composition of linear transformations to matrix multiplication.

To do so we let $X = R^n$, $Y = R^m$, and $W = R^p$. This choice of particular vector spaces allows us to represent the linear transformations as matrices and form matrix equations, as was done in (11). Thus F can be represented by a general real $m \times n$ matrix $\mathbf{A} = [a_{jk}]$ and G by an $n \times p$ matrix $\mathbf{B} = [b_{jk}]$. Then we can write for $F,$ with column vectors $\mathbf{x}$ with n entries, and resulting vector $\mathbf{y}$, with m entries

$$\mathbf{y} = \mathbf{A}\mathbf{x} \tag{16}$$

and similarly for G, with column vector $\mathbf{w}$ with p entries,

$$\mathbf{x} = \mathbf{B}\mathbf{w}. \tag{17}$$

Substituting (17) into (16) gives

$$\mathbf{y} = \mathbf{A}\mathbf{x} = \mathbf{A}(\mathbf{B}\mathbf{w}) = (\mathbf{A}\mathbf{B})\mathbf{w} = \mathbf{A}\mathbf{B}\mathbf{w} = \mathbf{C}\mathbf{w} \qquad \text{where } \mathbf{C} = \mathbf{A}\mathbf{B}. \tag{18}$$

This is (15) in a matrix setting, this is, ***we can define the composition of linear transformations in the Euclidean spaces as multiplication by matrices***. Hence, the real $m \times p$ matrix $\mathbf{C}$ represents a linear transformation H which maps R^p to R^n with vector $\mathbf{w}$, a column vector with p entries.

Remarks. Our discussion is similar to the one in Sec. 7.2, where we motivated the "unnatural" matrix multiplication of matrices. Look back and see that our current, more general, discussion is written out there for the case of dimension $m = 2$, $n = 2$, and $p = 2$. (You may want to write out our development by picking small *distinct* dimensions, such as $m = 2$, $n = 3$, and $p = 4$, and writing down the matrices and vectors. This is a trick of the trade of mathematicians in that we like to develop and test theories on smaller examples to see that they work.)

EXAMPLE 8 Linear Transformations. Composition

In Example 5 of Sec. 7.9, let $\mathbf{A}$ be the first matrix and $\mathbf{B}$ be the fourth matrix with $a > 1$. Then, applying $\mathbf{B}$ to a vector $\mathbf{w} = [w_1 \quad w_2]^\mathsf{T}$, stretches the element w_1 by a in the x_1 direction. Next, when we apply $\mathbf{A}$ to the "stretched" vector, we reflect the vector along the line $x_1 = x_2$, resulting in a vector $\mathbf{y} = [w_2 \quad aw_1]^\mathsf{T}$. But this represents, precisely, a geometric description for the composition H of two linear transformations F and G represented by matrices $\mathbf{A}$ and $\mathbf{B}$. We now show that, for this example, our result can be obtained by straightforward matrix multiplication, that is,

$$\mathbf{A}\mathbf{B} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ a & 0 \end{bmatrix}$$

and as in (18) calculate

$$\mathbf{A}\mathbf{B}\mathbf{w} = \begin{bmatrix} 0 & 1 \\ a & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} w_2 \\ aw_1 \end{bmatrix},$$

which is the same as before. This shows that indeed $\mathbf{A}\mathbf{B} = \mathbf{C}$, and we see the composition of linear transformations can be represented by a linear transformation. It also shows that the order of matrix multiplication is important (!). You may want to try applying $\mathbf{A}$ first and then $\mathbf{B}$, resulting in $\mathbf{B}\mathbf{A}$. What do you see? Does it make geometric sense? Is it the same result as $\mathbf{A}\mathbf{B}$? ■

We have learned several abstract concepts such as vector space, inner product space, and linear transformation. *The introduction of such concepts allows engineers and scientists to communicate in a concise and common language.* For example, the concept of a vector space encapsulated a lot of ideas in a very concise manner. For the student, learning such concepts provides a foundation for more advanced studies in engineering.

This concludes Chapter 7. The central theme was *the Gaussian elimination of* Sec. 7.3 from which most of the other concepts and theory flowed. The next chapter again has a central theme, that is, *eigenvalue problems,* an area very rich in applications such as in engineering, modern physics, and other areas.

PROBLEM SET 7.9

1. **Basis.** Find three bases of R^2.
2. **Uniqueness.** Show that the representation $\mathbf{v} = c_1\mathbf{a}_{(1)} + \cdots + c_n\mathbf{a}_{(n)}$ of any given vector in an n-dimensional vector space V in terms of a given basis $\mathbf{a}_{(1)}, \cdots, \mathbf{a}_{(n)}$ for V is unique. ***Hint***. Take two representations and consider the difference.

3–10 VECTOR SPACE

(More problems in Problem Set 9.4.) Is the given set, taken with the usual addition and scalar multiplication, a vector space? Give reason. If your answer is yes, find the dimension and a basis.

3. All vectors in R^3 satisfying $-v_1 + 2v_2 + 3v_3 = 0$, $-4v_1 + v_2 + v_3 = 0$.
4. All skew-symmetric 3×3 matrices.
5. All polynomials in x of degree 4 or less with nonnegative coefficients.
6. All functions $y(x) = a\cos 2x + b\sin 2x$ with arbitrary constants a and b.
7. All functions $y(x) = (ax + b)e^{-x}$ with any constant a and b.
8. All $n \times n$ matrices $\mathbf{A}$ with fixed n and $\det \mathbf{A} = 0$.
9. All 2×2 matrices $[a_{jk}]$ with $a_{11} + a_{22} = 0$.
10. All 3×2 matrices $[a_{jk}]$ with first column any multiple of $[3 \quad 0 \quad -5]^\mathsf{T}$.

11–14 LINEAR TRANSFORMATIONS

Find the inverse transformation. Show the details.

11. $y_1 = 0.5x_1 - 0.5x_2$
    $y_2 = 1.5x_1 - 2.5x_2$
12. $y_1 = 3x_1 + 2x_2$
    $y_2 = 4x_1 + x_2$
13. $y_1 = 5x_1 + 3x_2 - 3x_3$
    $y_2 = 3x_1 + 2x_2 - 2x_3$
    $y_3 = 2x_1 - x_2 + 2x_3$
14. $y_1 = 0.2x_1 - 0.1x_2$
    $y_2 = -0.2x_2 + 0.1x_3$
    $y_3 = 0.1x_1 + 0.1x_3$

15–20 EUCLIDEAN NORM

Find the Euclidean norm of the vectors:

15. $[3 \quad 1 \quad -4]^\mathsf{T}$
16. $[\frac{1}{2} \quad \frac{1}{3} \quad -\frac{1}{2} \quad -\frac{1}{3}]^\mathsf{T}$
17. $[1 \quad 0 \quad 0 \quad 1 \quad -1 \quad 0 \quad -1 \quad 1]^\mathsf{T}$
18. $[-4 \quad 8 \quad -1]^\mathsf{T}$
19. $[\frac{2}{3} \quad \frac{2}{3} \quad \frac{1}{3} \quad 0]^\mathsf{T}$
20. $[\frac{1}{2} \quad -\frac{1}{2} \quad -\frac{1}{2} \quad \frac{1}{2}]^\mathsf{T}$

21–25 INNER PRODUCT. ORTHOGONALITY

21. **Orthogonality.** For what value(s) of k are the vectors $[2 \quad \frac{1}{2} \quad -4 \quad 0]^\mathsf{T}$ and $[5 \quad k \quad 0 \quad \frac{1}{4}]^\mathsf{T}$ orthogonal?
22. **Orthogonality.** Find all vectors in R^3 orthogonal to $[2 \quad 0 \quad 1]$. Do they form a vector space?
23. **Triangle inequality.** Verify (4) for the vectors in Probs. 15 and 18.
24. **Cauchy–Schwarz inequality.** Verify (3) for the vectors in Probs. 16 and 19.
25. **Parallelogram equality.** Verify (5) for the first two column vectors of the coefficient matrix in Prob. 13.

CHAPTER 7 REVIEW QUESTIONS AND PROBLEMS

1. What properties of matrix multiplication differ from those of the multiplication of numbers?
2. Let $\mathbf{A}$ be a 100×100 matrix and $\mathbf{B}$ a 100×50 matrix. Are the following expressions defined or not? $\mathbf{A} + \mathbf{B}$, $\mathbf{A}^2$, $\mathbf{B}^2$, $\mathbf{AB}$, $\mathbf{BA}$, $\mathbf{AA}^\mathsf{T}$, $\mathbf{B}^\mathsf{T}\mathbf{A}$, $\mathbf{B}^\mathsf{T}\mathbf{B}$, $\mathbf{BB}^\mathsf{T}$, $\mathbf{B}^\mathsf{T}\mathbf{AB}$. Give reasons.
3. Are there any linear systems without solutions? With one solution? With more than one solution? Give simple examples.
4. Let $\mathbf{C}$ be 10×10 matrix and $\mathbf{a}$ a column vector with 10 components. Are the following expressions defined or not? $\mathbf{Ca}$, $\mathbf{C}^\mathsf{T}\mathbf{a}$, $\mathbf{Ca}^\mathsf{T}$, $\mathbf{aC}$, $\mathbf{a}^\mathsf{T}\mathbf{C}$, $(\mathbf{Ca}^\mathsf{T})^\mathsf{T}$.
5. Motivate the definition of matrix multiplication.
6. Explain the use of matrices in linear transformations.
7. How can you give the rank of a matrix in terms of row vectors? Of column vectors? Of determinants?
8. What is the role of rank in connection with solving linear systems?
9. What is the idea of Gauss elimination and back substitution?
10. What is the inverse of a matrix? When does it exist? How would you determine it?

11–20 MATRIX AND VECTOR CALCULATIONS

Showing the details, calculate the following expressions or give reason why they are not defined, when

$$\mathbf{A} = \begin{bmatrix} 3 & 1 & -3 \\ 1 & 4 & 2 \\ -3 & 2 & 5 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 & 4 & 1 \\ -4 & 0 & -2 \\ -1 & 2 & 0 \end{bmatrix},$$

$$\mathbf{u} = \begin{bmatrix} 2 \\ 0 \\ -5 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} 7 \\ -3 \\ 3 \end{bmatrix}$$

11. $\mathbf{AB}$, $\mathbf{BA}$

12. $\mathbf{A}^\mathsf{T}$, $\mathbf{B}^\mathsf{T}$

13. $\mathbf{Au}$, $\mathbf{u}^\mathsf{T}\mathbf{A}$

14. $\mathbf{u}^\mathsf{T}\mathbf{v}$, $\mathbf{uv}^\mathsf{T}$

15. $\mathbf{u}^\mathsf{T}\mathbf{Au}$, $\mathbf{v}^\mathsf{T}\mathbf{Bv}$

16. $\mathbf{A}^{-1}$, $\mathbf{B}^{-1}$

17. $\det \mathbf{A}$, $\det \mathbf{A}^2$, $(\det \mathbf{A})^2$, $\det \mathbf{B}$

18. $(\mathbf{A}^2)^{-1}$, $(\mathbf{A}^{-1})^2$

19. $\mathbf{AB} - \mathbf{BA}$

20. $(\mathbf{A} + \mathbf{A}^\mathsf{T})(\mathbf{B} - \mathbf{B}^\mathsf{T})$

21–28 LINEAR SYSTEMS

Showing the details, find all solutions or indicate that no solution exists.

21.
$$\begin{aligned} 4y + z &= 0 \\ 12x - 5y - 3z &= 34 \\ -6x + 4z &= 8 \end{aligned}$$

22.
$$\begin{aligned} 5x - 3y + z &= 7 \\ 2x + 3y - z &= 0 \\ 8x + 9y - 3z &= 2 \end{aligned}$$

23.
$$\begin{aligned} 9x + 3y - 6z &= 60 \\ 2x - 4y + 8z &= 4 \end{aligned}$$

24.
$$\begin{aligned} -6x + 39y - 9z &= -12 \\ 2x - 13y + 3z &= 4 \end{aligned}$$

25.
$$\begin{aligned} 0.3x - 0.7y + 1.3z &= 3.24 \\ 0.9y - 0.8z &= -2.53 \\ 0.7z &= 1.19 \end{aligned}$$

26.
$$\begin{aligned} 2x + 3y - 7z &= 3 \\ -4x - 6y + 14z &= 7 \end{aligned}$$

27.
$$\begin{aligned} x + 2y &= 6 \\ 3x + 5y &= 20 \\ -4x + y &= -42 \end{aligned}$$

28.
$$\begin{aligned} -8x + 2z &= 1 \\ 6y + 4z &= 3 \\ 12x + 2y &= 2 \end{aligned}$$

29–32 RANK

Determine the ranks of the coefficient matrix and the augmented matrix and state how many solutions the linear system will have.

29. In Prob. 23

30. In Prob. 24

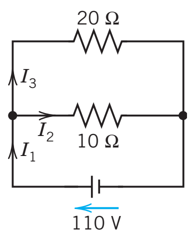
31. In Prob. 27

32. In Prob. 26

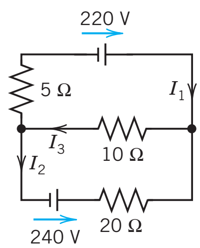
33–35 NETWORKS

Find the currents.

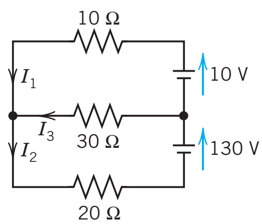
33.



34.



35.



SUMMARY OF CHAPTER 7

Linear Algebra: Matrices, Vectors, Determinants. Linear Systems

An $m \times n$ **matrix** $\mathbf{A} = [a_{jk}]$ is a rectangular array of numbers or functions ("entries," "elements") arranged in m horizontal **rows** and n vertical **columns**. If $m = n$, the matrix is called **square**. A $1 \times n$ matrix is called a **row vector** and an $m \times 1$ matrix a **column vector** (Sec. 7.1).

The **sum** $\mathbf{A} + \mathbf{B}$ of matrices of the same **size** (i.e., both $m \times n$) is obtained by adding corresponding entries. The **product** of $\mathbf{A}$ by a scalar c is obtained by multiplying each a_{jk} by c (Sec. 7.1).

The **product** $\mathbf{C} = \mathbf{AB}$ of an $m \times n$ matrix $\mathbf{A}$ by an $r \times p$ matrix $\mathbf{B} = [b_{jk}]$ is defined only when $r = n$, and is the $m \times p$ matrix $\mathbf{C} = [c_{jk}]$ with entries

$$c_{jk} = a_{j1}b_{1k} + a_{j2}b_{2k} + \cdots + a_{jn}b_{nk} \qquad \text{(row } j \text{ of } \mathbf{A} \text{ times column } k \text{ of } \mathbf{B}\text{).} \tag{1}$$

This multiplication is motivated by the composition of **linear transformations** (Secs. 7.2, 7.9). It is associative, but is ***not commutative:*** if $\mathbf{AB}$ is defined, $\mathbf{BA}$ may not be defined, but even if $\mathbf{BA}$ is defined, $\mathbf{AB} \neq \mathbf{BA}$ in general. Also $\mathbf{AB} = \mathbf{0}$ may not imply $\mathbf{A} = \mathbf{0}$ or $\mathbf{B} = \mathbf{0}$ or $\mathbf{BA} = \mathbf{0}$ (Secs. 7.2, 7.8). Illustrations:

$$\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}$$

$$[1 \quad 2]\begin{bmatrix} 3 \\ 4 \end{bmatrix} = [11], \qquad \begin{bmatrix} 3 \\ 4 \end{bmatrix}[1 \quad 2] = \begin{bmatrix} 3 & 6 \\ 4 & 8 \end{bmatrix}.$$

The **transpose** $\mathbf{A}^\mathsf{T}$ of a matrix $\mathbf{A} = [a_{jk}]$ is $\mathbf{A}^\mathsf{T} = [a_{kj}]$; rows become columns and conversely (Sec. 7.2). Here, $\mathbf{A}$ need not be square. If it is and $\mathbf{A} = \mathbf{A}^\mathsf{T}$, then $\mathbf{A}$ is called **symmetric;** if $\mathbf{A} = -\mathbf{A}^\mathsf{T}$, it is called **skew-symmetric**. For a product, $(\mathbf{AB})^\mathsf{T} = \mathbf{B}^\mathsf{T}\mathbf{A}^\mathsf{T}$ (Sec. 7.2).

A main application of matrices concerns **linear systems of equations**

$$\mathbf{Ax} = \mathbf{b} \qquad \text{(Sec. 7.3)} \tag{2}$$

(m equations in n unknowns $x_1, \cdots, x_n$; $\mathbf{A}$ and $\mathbf{b}$ given). The most important method of solution is the **Gauss elimination** (Sec. 7.3), which reduces the system to "triangular" form by *elementary row operations,* which leave the set of solutions unchanged. (Numeric aspects and variants, such as *Doolittle's* and *Cholesky's methods,* are discussed in Secs. 20.1 and 20.2.)

Cramer's rule (Secs. 7.6, 7.7) represents the unknowns in a system (2) of n equations in n unknowns as quotients of determinants; for numeric work it is impractical. **Determinants** (Sec. 7.7) have decreased in importance, but will retain their place in eigenvalue problems, elementary geometry, etc.

The **inverse** $\mathbf{A}^{-1}$ of a square matrix satisfies $\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$. It exists if and only if $\det \mathbf{A} \neq 0$. It can be computed by the *Gauss–Jordan elimination* (Sec. 7.8).

The **rank** r of a matrix $\mathbf{A}$ is the maximum number of linearly independent rows or columns of $\mathbf{A}$ or, equivalently, the number of rows of the largest square submatrix of $\mathbf{A}$ with nonzero determinant (Secs. 7.4, 7.7).

The system (2) has solutions if and only if $\operatorname{rank} \mathbf{A} = \operatorname{rank} [\mathbf{A} \quad \mathbf{b}]$, where $[\mathbf{A} \quad \mathbf{b}]$ is the **augmented matrix** (Fundamental Theorem, Sec. 7.5).

The **homogeneous system**

$$\mathbf{A}\mathbf{x} = \mathbf{0} \tag{3}$$

has solutions $\mathbf{x} \neq \mathbf{0}$ ("nontrivial solutions") if and only if $\operatorname{rank} \mathbf{A} < n$, in the case $m = n$ equivalently if and only if $\det \mathbf{A} = \mathbf{0}$ (Secs. 7.6, 7.7).

Vector spaces, inner product spaces, and linear transformations are discussed in Sec. 7.9. See also Sec. 7.4.

CHAPTER 8

Linear Algebra: Matrix Eigenvalue Problems

A matrix eigenvalue problem considers the vector equation

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}. \tag{1}$$

Here $\mathbf{A}$ is a given square matrix, λ an unknown scalar, and $\mathbf{x}$ an unknown vector. In a matrix eigenvalue problem, the task is to determine λ's and $\mathbf{x}$'s that satisfy (1). Since $\mathbf{x} = \mathbf{0}$ is always a solution for any λ and thus not interesting, we only admit solutions with $\mathbf{x} \neq \mathbf{0}$.

The solutions to (1) are given the following names: The λ's that satisfy (1) are called **eigenvalues of A** and the corresponding nonzero $\mathbf{x}$'s that also satisfy (1) are called **eigenvectors of A**.

From this rather innocent looking vector equation flows an amazing amount of relevant theory and an incredible richness of applications. Indeed, eigenvalue problems come up all the time in engineering, physics, geometry, numerics, theoretical mathematics, biology, environmental science, urban planning, economics, psychology, and other areas. Thus, in your career you are likely to encounter eigenvalue problems.

We start with a basic and thorough introduction to eigenvalue problems in Sec. 8.1 and explain (1) with several simple matrices. This is followed by a section devoted entirely to applications ranging from mass–spring systems of physics to population control models of environmental science. We show you these diverse examples to train your skills in modeling and solving eigenvalue problems. Eigenvalue problems for real symmetric, skew-symmetric, and orthogonal matrices are discussed in Sec. 8.3 and their complex counterparts (which are important in modern physics) in Sec. 8.5. In Sec. 8.4 we show how by diagonalizing a matrix, we obtain its eigenvalues.

COMMENT. ***Numerics for eigenvalues* (Secs. 20.6–20.9) *can be studied immediately after this chapter*.**

Prerequisite: Chap. 7.
Sections that may be omitted in a shorter course: 8.4, 8.5.
References and Answers to Problems: App. 1 Part B, App. 2.

The following chart identifies where different types of eigenvalue problems appear in the book.

Topic	Where to find it
Matrix Eigenvalue Problem (algebraic eigenvalue problem)	Chap. 8
Eigenvalue Problems in Numerics	Secs. 20.6–20.9
Eigenvalue Problem for ODEs (Sturm–Liouville problems)	Secs. 11.5, 11.6
Eigenvalue Problems for Systems of ODEs	Chap. 4
Eigenvalue Problems for PDEs	Secs. 12.3–12.11

8.1 The Matrix Eigenvalue Problem. Determining Eigenvalues and Eigenvectors

Consider multiplying nonzero vectors by a given square matrix, such as

$$\begin{bmatrix} 6 & 3 \\ 4 & 7 \end{bmatrix}\begin{bmatrix} 5 \\ 1 \end{bmatrix} = \begin{bmatrix} 33 \\ 27 \end{bmatrix}, \qquad \begin{bmatrix} 6 & 3 \\ 4 & 7 \end{bmatrix}\begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 30 \\ 40 \end{bmatrix}.$$

We want to see what influence the multiplication of the given matrix has on the vectors. In the first case, we get a totally new vector with a different direction and different length when compared to the original vector. This is what usually happens and is of no interest here. In the second case something interesting happens. The multiplication produces a vector $[30 \quad 40]^{\mathsf{T}} = 10\,[3 \quad 4]^{\mathsf{T}}$, which means the new vector has the same direction as the original vector. The scale constant, which we denote by λ is 10. *The problem of systematically finding such* λ's *and nonzero vectors for a given square matrix will be the theme of this chapter*. It is called the *matrix eigenvalue* problem or, more commonly, the *eigenvalue* problem.

We formalize our observation. Let $\mathbf{A} = [a_{jk}]$ be a given nonzero square matrix of dimension $n \times n$. Consider the following vector equation:

$$\mathbf{Ax} = \lambda\mathbf{x}. \tag{1}$$

The problem of finding nonzero $\mathbf{x}$'s and λ's that satisfy equation (1) is called an eigenvalue problem.

Remark. So $\mathbf{A}$ is a given square (!) matrix, $\mathbf{x}$ is an unknown vector, and λ is an unknown scalar. Our task is to find λ's and nonzero $\mathbf{x}$'s that satisfy (1). Geometrically, we are looking for vectors, $\mathbf{x}$, for which the multiplication by $\mathbf{A}$ has the same effect as the multiplication by a scalar λ; in other words, $\mathbf{Ax}$ should be proportional to $\mathbf{x}$. Thus, the multiplication has the effect of producing, from the original vector $\mathbf{x}$, a new vector $\lambda\mathbf{x}$ that has the same or opposite (minus sign) direction as the original vector. (This was all demonstrated in our intuitive opening example. Can you see that the second equation in that example satisfies (1) with $\lambda = 10$ and $\mathbf{x} = [3 \quad 4]^{\mathsf{T}}$, and $\mathbf{A}$ the given 2×2 matrix? Write it out.) Now why do we require $\mathbf{x}$ to be nonzero? The reason is that $\mathbf{x} = \mathbf{0}$ is always a solution of (1) for any value of λ, because $\mathbf{A0} = \mathbf{0}$. This is of no interest.

We introduce more terminology. A value of λ, for which (1) has a solution $\mathbf{x} \neq \mathbf{0}$, is called an **eigenvalue** or *characteristic value* of the matrix $\mathbf{A}$. Another term for λ is a *latent* root. ("Eigen" is German and means "proper" or "characteristic."). The corresponding solutions $\mathbf{x} \neq \mathbf{0}$ of (1) are called the **eigenvectors** or *characteristic vectors* of $\mathbf{A}$ corresponding to that eigenvalue λ. The set of all the eigenvalues of $\mathbf{A}$ is called the **spectrum** of $\mathbf{A}$. We shall see that the spectrum consists of at least one eigenvalue and at most of n numerically different eigenvalues. The largest of the absolute values of the eigenvalues of $\mathbf{A}$ is called the *spectral radius* of $\mathbf{A}$, a name to be motivated later.

How to Find Eigenvalues and Eigenvectors

Now, with the new terminology for (1), we can just say that the problem of determining the eigenvalues and eigenvectors of a matrix is called an eigenvalue problem. (However, more precisely, we are considering an algebraic eigenvalue problem, as opposed to an eigenvalue problem involving an ODE or PDE, as considered in Secs. 11.5 and 12.3, or an integral equation.)

Eigenvalues have a very large number of applications in diverse fields such as in engineering, geometry, physics, mathematics, biology, environmental science, economics, psychology, and other areas. You will encounter applications for elastic membranes, Markov processes, population models, and others in this chapter.

Since, from the viewpoint of engineering applications, eigenvalue problems are the most important problems in connection with matrices, the student should carefully follow our discussion.

Example 1 demonstrates how to systematically solve a simple eigenvalue problem.

EXAMPLE 1 Determination of Eigenvalues and Eigenvectors

We illustrate all the steps in terms of the matrix

$$\mathbf{A} = \begin{bmatrix} -5 & 2 \\ 2 & -2 \end{bmatrix}.$$

Solution. **(a)** ***Eigenvalues.*** These must be determined ***first***. Equation (1) is

$$\mathbf{Ax} = \begin{bmatrix} -5 & 2 \\ 2 & -2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}; \qquad \text{in components,} \qquad \begin{aligned} -5x_1 + 2x_2 &= \lambda x_1 \\ 2x_1 - 2x_2 &= \lambda x_2. \end{aligned}$$

Transferring the terms on the right to the left, we get

$$(2^*) \qquad \begin{aligned} (-5-\lambda)x_1 + \quad 2x_2 \quad &= 0 \\ 2x_1 \quad + (-2-\lambda)x_2 &= 0. \end{aligned}$$

This can be written in matrix notation

$$(3^*) \qquad (\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$$

because (1) is $\mathbf{Ax} - \lambda\mathbf{x} = \mathbf{Ax} - \lambda\mathbf{Ix} = (\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$, which gives (3*). We see that this is a ***homogeneous*** linear system. By Cramer's theorem in Sec. 7.7 it has a nontrivial solution $\mathbf{x} \neq \mathbf{0}$ (an eigenvector of $\mathbf{A}$ we are looking for) if and only if its coefficient determinant is zero, that is,

$$(4^*) \qquad D(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} -5-\lambda & 2 \\ 2 & -2-\lambda \end{vmatrix} = (-5-\lambda)(-2-\lambda) - 4 = \lambda^2 + 7\lambda + 6 = 0.$$

We call $D(\lambda)$ the **characteristic determinant** or, if expanded, the **characteristic polynomial**, and $D(\lambda) = 0$ the **characteristic equation** of $\mathbf{A}$. The solutions of this quadratic equation are $\lambda_1 = -1$ and $\lambda_2 = -6$. These are the eigenvalues of $\mathbf{A}$.

(b_1) ***Eigenvector of*** $\mathbf{A}$ ***corresponding to*** λ_1. This vector is obtained from (2*) with $\lambda = \lambda_1 = -1$, that is,

$$\begin{aligned} -4x_1 + 2x_2 &= 0 \\ 2x_1 - x_2 &= 0. \end{aligned}$$

A solution is $x_2 = 2x_1$, as we see from either of the two equations, so that we need only one of them. This determines an eigenvector corresponding to $\lambda_1 = -1$ up to a scalar multiple. If we choose $x_1 = 1$, we obtain the eigenvector

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \qquad \text{Check:} \qquad \mathbf{A}\mathbf{x}_1 = \begin{bmatrix} -5 & 2 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 \\ -2 \end{bmatrix} = (-1)\mathbf{x}_1 = \lambda_1 \mathbf{x}_1.$$

(b_2) ***Eigenvector of*** $\mathbf{A}$ ***corresponding to*** λ_2. For $\lambda = \lambda_2 = -6$, equation (2*) becomes

$$\begin{aligned} x_1 + 2x_2 &= 0 \\ 2x_1 + 4x_2 &= 0. \end{aligned}$$

A solution is $x_2 = -x_1/2$ with arbitrary x_1. If we choose $x_1 = 2$, we get $x_2 = -1$. Thus an eigenvector of $\mathbf{A}$ corresponding to $\lambda_2 = -6$ is

$$\mathbf{x}_2 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \qquad \text{Check:} \qquad \mathbf{A}\mathbf{x}_2 = \begin{bmatrix} -5 & 2 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} -12 \\ 6 \end{bmatrix} = (-6)\mathbf{x}_2 = \lambda_2 \mathbf{x}_2.$$

For the matrix in the intuitive opening example at the start of Sec. 8.1, the characteristic equation is $\lambda^2 - 13\lambda + 30 = (\lambda - 10)(\lambda - 3) = 0$. The eigenvalues are $\{10, \ 3\}$. Corresponding eigenvectors are $[3 \quad 4]^T$ and $[-1 \quad 1]^T$, respectively. The reader may want to verify this. ■

This example illustrates the general case as follows. Equation (1) written in components is

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= \lambda x_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n &= \lambda x_2 \\ \cdots\cdots\cdots\cdots&\cdots\cdots\cdots \\ a_{n1}x_1 + \cdots + a_{nn}x_n &= \lambda x_n. \end{aligned}$$

Transferring the terms on the right side to the left side, we have

$$\begin{aligned} (a_{11} - \lambda)x_1 + \quad a_{12}x_2 \quad + \cdots + \quad a_{1n}x_n \quad &= 0 \\ a_{21}x_1 \quad + (a_{22} - \lambda)x_2 + \cdots + \quad a_{2n}x_n \quad &= 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots&\cdots\cdots \\ a_{n1}x_1 \quad + \quad a_{n2}x_2 \quad + \cdots + (a_{nn} - \lambda)x_n &= 0. \end{aligned} \tag{2}$$

In matrix notation,

$$\boxed{(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}.} \tag{3}$$

By Cramer's theorem in Sec. 7.7, this homogeneous linear system of equations has a nontrivial solution if and only if the corresponding determinant of the coefficients is zero:

$$D(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix} = 0. \tag{4}$$

$\mathbf{A} - \lambda\mathbf{I}$ is called the **characteristic matrix** and $D(\lambda)$ the **characteristic determinant** of $\mathbf{A}$. Equation (4) is called the **characteristic equation** of $\mathbf{A}$. By developing $D(\lambda)$ we obtain a polynomial of nth degree in λ. This is called the **characteristic polynomial** of $\mathbf{A}$.

This proves the following important theorem.

THEOREM 1

Eigenvalues

The eigenvalues of a square matrix $\mathbf{A}$ *are the roots of the characteristic equation* (4) *of* $\mathbf{A}$.

Hence an $n \times n$ *matrix has at least one eigenvalue and at most* n *numerically different eigenvalues.*

For larger n, the actual computation of eigenvalues will, in general, require the use of Newton's method (Sec. 19.2) or another numeric approximation method in Secs. 20.7–20.9.

The eigenvalues must be determined first. Once these are known, corresponding eigen*vectors* are obtained from the system (2), for instance, by the Gauss elimination, where λ is the eigenvalue for which an eigenvector is wanted. This is what we did in Example 1 and shall do again in the examples below. (To prevent misunderstandings: *numeric approximation methods*, such as in Sec. 20.8, may determine eigen*vectors first*.)

Eigenvectors have the following properties.

THEOREM 2

Eigenvectors, Eigenspace

If $\mathbf{w}$ *and* $\mathbf{x}$ *are eigenvectors of a matrix* $\mathbf{A}$ *corresponding to* ***the same*** *eigenvalue* λ, *so are* $\mathbf{w} + \mathbf{x}$ (provided $\mathbf{x} \neq -\mathbf{w}$) *and* $k\mathbf{x}$ *for any* $k \neq 0$.

Hence the eigenvectors corresponding to one and the same eigenvalue λ *of* $\mathbf{A}$, *together with* $\mathbf{0}$, *form a vector space* (cf. Sec. 7.4), *called the* **eigenspace** *of* $\mathbf{A}$ *corresponding to that* λ.

PROOF $\mathbf{A}\mathbf{w} = \lambda\mathbf{w}$ and $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ *imply* $\mathbf{A}(\mathbf{w} + \mathbf{x}) = \mathbf{A}\mathbf{w} + \mathbf{A}\mathbf{x} = \lambda\mathbf{w} + \lambda\mathbf{x} = \lambda(\mathbf{w} + \mathbf{x})$ and $\mathbf{A}(k\mathbf{w}) = k(\mathbf{A}\mathbf{w}) = k(\lambda\mathbf{w}) = \lambda(k\mathbf{w})$; hence $\mathbf{A}(k\mathbf{w} + \ell\mathbf{x}) = \lambda(k\mathbf{w} + \ell\mathbf{x})$. ■

In particular, *an eigenvector* $\mathbf{x}$ *is determined only up to a constant factor.* Hence we can **normalize** $\mathbf{x}$, that is, multiply it by a scalar to get a unit vector (see Sec. 7.9). For instance, $\mathbf{x}_1 = [1 \quad 2]^\mathsf{T}$ in Example 1 has the length $\|\mathbf{x}_1\| = \sqrt{1^2 + 2^2} = \sqrt{5}$; hence $[1/\sqrt{5} \quad 2/\sqrt{5}]^\mathsf{T}$ is a normalized eigenvector (a unit eigenvector).

Examples 2 and 3 will illustrate that an $n \times n$ matrix may have n linearly independent eigenvectors, or it may have fewer than n. In Example 4 we shall see that a *real* matrix may have *complex* eigenvalues and eigenvectors.

EXAMPLE 2 Multiple Eigenvalues

Find the eigenvalues and eigenvectors of

$$\mathbf{A} = \begin{bmatrix} -2 & 2 & -3 \\ 2 & 1 & -6 \\ -1 & -2 & 0 \end{bmatrix}.$$

Solution. For our matrix, the characteristic determinant gives the characteristic equation

$$-\lambda^3 - \lambda^2 + 21\lambda + 45 = 0.$$

The roots (eigenvalues of $\mathbf{A}$) are $\lambda_1 = 5, \lambda_2 = \lambda_3 = -3$. (If you have trouble finding roots, you may want to use a root finding algorithm such as Newton's method (Sec. 19.2). Your CAS or scientific calculator can find roots. However, to really learn and remember this material, you have to do some exercises with paper and pencil.) To find eigenvectors, we apply the Gauss elimination (Sec. 7.3) to the system $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$, first with $\lambda = 5$ and then with $\lambda = -3$. For $\lambda = 5$ the characteristic matrix is

$$\mathbf{A} - \lambda\mathbf{I} = \mathbf{A} - 5\mathbf{I} = \begin{bmatrix} -7 & 2 & -3 \\ 2 & -4 & -6 \\ -1 & -2 & -5 \end{bmatrix}. \quad \text{It row-reduces to} \quad \begin{bmatrix} -7 & 2 & -3 \\ 0 & -\frac{24}{7} & -\frac{48}{7} \\ 0 & 0 & 0 \end{bmatrix}.$$

Hence it has rank 2. Choosing $x_3 = -1$ we have $x_2 = 2$ from $-\frac{24}{7}x_2 - \frac{48}{7}x_3 = 0$ and then $x_1 = 1$ from $-7x_1 + 2x_2 - 3x_3 = 0$. Hence an eigenvector of $\mathbf{A}$ corresponding to $\lambda = 5$ is $\mathbf{x}_1 = [1 \quad 2 \quad -1]^\mathsf{T}$.

For $\lambda = -3$ the characteristic matrix

$$\mathbf{A} - \lambda\mathbf{I} = \mathbf{A} + 3\mathbf{I} = \begin{bmatrix} 1 & 2 & -3 \\ 2 & 4 & -6 \\ -1 & -2 & 3 \end{bmatrix} \quad \text{row-reduces to} \quad \begin{bmatrix} 1 & 2 & -3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Hence it has rank 1. From $x_1 + 2x_2 - 3x_3 = 0$ we have $x_1 = -2x_2 + 3x_3$. Choosing $x_2 = 1, x_3 = 0$ and $x_2 = 0, x_3 = 1$, we obtain two linearly independent eigenvectors of $\mathbf{A}$ corresponding to $\lambda = -3$ [as they must exist by (5), Sec. 7.5, with rank = 1 and $n = 3$],

$$\mathbf{x}_2 = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$$

and

$$\mathbf{x}_3 = \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}.$$

■

The order M_λ of an eigenvalue λ as a root of the characteristic polynomial is called the **algebraic multiplicity** of λ. The number m_λ of linearly independent eigenvectors corresponding to λ is called the **geometric multiplicity** of λ. Thus m_λ is the dimension of the eigenspace corresponding to this λ.

Since the characteristic polynomial has degree n, the sum of all the algebraic multiplicities must equal n. In Example 2 for $\lambda = -3$ we have $m_\lambda = M_\lambda = 2$. In general, $m_\lambda \leqq M_\lambda$, as can be shown. The difference $\Delta_\lambda = M_\lambda - m_\lambda$ is called the **defect** of λ. Thus $\Delta_{-3} = 0$ in Example 2, but positive defects Δ_λ can easily occur:

EXAMPLE 3 Algebraic Multiplicity, Geometric Multiplicity. Positive Defect

The characteristic equation of the matrix

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad \text{is} \quad \det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} -\lambda & 1 \\ 0 & -\lambda \end{vmatrix} = \lambda^2 = 0.$$

Hence $\lambda = 0$ is an eigenvalue of algebraic multiplicity $M_0 = 2$. But its geometric multiplicity is only $m_0 = 1$, since eigenvectors result from $-0x_1 + x_2 = 0$, hence $x_2 = 0$, in the form $[x_1 \quad 0]^\mathsf{T}$. Hence for $\lambda = 0$ the defect is $\Delta_0 = 1$.

Similarly, the characteristic equation of the matrix

$$\mathbf{A} = \begin{bmatrix} 3 & 2 \\ 0 & 3 \end{bmatrix} \quad \text{is} \quad \det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} 3-\lambda & 2 \\ 0 & 3-\lambda \end{vmatrix} = (3-\lambda)^2 = 0.$$

Hence $\lambda = 3$ is an eigenvalue of algebraic multiplicity $M_3 = 2$, but its geometric multiplicity is only $m_3 = 1$, since eigenvectors result from $0x_1 + 2x_2 = 0$ in the form $[x_1 \quad 0]^\mathsf{T}$. ■

EXAMPLE 4 Real Matrices with Complex Eigenvalues and Eigenvectors

Since real polynomials may have complex roots (which then occur in conjugate pairs), a real matrix may have complex eigenvalues and eigenvectors. For instance, the characteristic equation of the skew-symmetric matrix

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad \text{is} \quad \det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} -\lambda & 1 \\ -1 & -\lambda \end{vmatrix} = \lambda^2 + 1 = 0.$$

It gives the eigenvalues $\lambda_1 = i\,(= \sqrt{-1})$, $\lambda_2 = -i$. Eigenvectors are obtained from $-ix_1 + x_2 = 0$ and $ix_1 + x_2 = 0$, respectively, and we can choose $x_1 = 1$ to get

$$\begin{bmatrix} 1 \\ i \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 \\ -i \end{bmatrix}.$$ ■

In the next section we shall need the following simple theorem.

THEOREM 3 Eigenvalues of the Transpose

The transpose $\mathbf{A}^\mathsf{T}$ of a square matrix $\mathbf{A}$ has the same eigenvalues as $\mathbf{A}$.

PROOF Transposition does not change the value of the characteristic determinant, as follows from Theorem 2d in Sec. 7.7. ■

Having gained a first impression of matrix eigenvalue problems, we shall illustrate their importance with some typical applications in Sec. 8.2.

PROBLEM SET 8.1

1–16 EIGENVALUES, EIGENVECTORS

Find the eigenvalues. Find the corresponding eigenvectors. Use the given λ or factor in Probs. 11 and 15.

1. $\begin{bmatrix} 3.0 & 0 \\ 0 & -0.6 \end{bmatrix}$

2. $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

3. $\begin{bmatrix} 5 & -2 \\ 9 & -6 \end{bmatrix}$

4. $\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$

5. $\begin{bmatrix} 0 & 3 \\ -3 & 0 \end{bmatrix}$

6. $\begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$

7. $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$

8. $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$

9. $\begin{bmatrix} 0.8 & -0.6 \\ 0.6 & 0.8 \end{bmatrix}$

10. $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$

11. $\begin{bmatrix} 6 & 2 & -2 \\ 2 & 5 & 0 \\ -2 & 0 & 7 \end{bmatrix}, \quad \lambda = 3$

12. $\begin{bmatrix} 3 & 5 & 3 \\ 0 & 4 & 6 \\ 0 & 0 & 1 \end{bmatrix}$

13. $\begin{bmatrix} 13 & 5 & 2 \\ 2 & 7 & -8 \\ 5 & 4 & 7 \end{bmatrix}$

14. $\begin{bmatrix} 2 & 0 & -1 \\ 0 & \frac{1}{2} & 0 \\ 1 & 0 & 4 \end{bmatrix}$

15. $\begin{bmatrix} -1 & 0 & 12 & 0 \\ 0 & -1 & 0 & 12 \\ 0 & 0 & -1 & -4 \\ 0 & 0 & -4 & -1 \end{bmatrix}, \quad (\lambda + 1)^2$

16. $\begin{bmatrix} -3 & 0 & 4 & 2 \\ 0 & 1 & -2 & 4 \\ 2 & 4 & -1 & -2 \\ 0 & 2 & -2 & 3 \end{bmatrix}$

17–20 LINEAR TRANSFORMATIONS AND EIGENVALUES

Find the matrix $\mathbf{A}$ in the linear transformation $\mathbf{y} = \mathbf{A}\mathbf{x}$, where $\mathbf{x} = [x_1 \;\; x_2]^\mathsf{T}$ ($\mathbf{x} = [x_1 \;\; x_2 \;\; x_3]^\mathsf{T}$) are Cartesian coordinates. Find the eigenvalues and eigenvectors and explain their geometric meaning.

17. Counterclockwise rotation through the angle $\pi/2$ about the origin in R^2.

18. Reflection about the x_1-axis in R^2.

19. Orthogonal projection (perpendicular projection) of R^2 onto the x_2-axis.

20. Orthogonal projection of R^3 onto the plane $x_2 = x_1$.

21–25 GENERAL PROBLEMS

21. Nonzero defect. Find further 2×2 and 3×3 matrices with positive defect. See Example 3.

22. Multiple eigenvalues. Find further 2×2 and 3×3 matrices with multiple eigenvalues. See Example 2.

23. Complex eigenvalues. Show that the eigenvalues of a real matrix are real or complex conjugate in pairs.

24. Inverse matrix. Show that $\mathbf{A}^{-1}$ exists if and only if the eigenvalues $\lambda_1, \cdots, \lambda_n$ are all nonzero, and then $\mathbf{A}^{-1}$ has the eigenvalues $1/\lambda_1, \cdots, 1/\lambda_n$.

25. Transpose. Illustrate Theorem 3 with examples of your own.

8.2 Some Applications of Eigenvalue Problems

We have selected some typical examples from the wide range of applications of matrix eigenvalue problems. The last example, that is, Example 4, shows an application involving vibrating springs and ODEs. It falls into the domain of Chapter 4, which covers matrix eigenvalue problems related to ODE's modeling mechanical systems and electrical networks. Example 4 is included to keep our discussion independent of Chapter 4. (However, the reader not interested in ODEs may want to skip Example 4 without loss of continuity.)

EXAMPLE 1 Stretching of an Elastic Membrane

An elastic membrane in the x_1x_2-plane with boundary circle $x_1^2 + x_2^2 = 1$ (Fig. 160) is stretched so that a point P: (x_1, x_2) goes over into the point Q: (y_1, y_2) given by

$$(1) \qquad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \mathbf{Ax} = \begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}; \qquad \text{in components,} \qquad \begin{aligned} y_1 &= 5x_1 + 3x_2 \\ y_2 &= 3x_1 + 5x_2. \end{aligned}$$

Find the **principal directions**, that is, the directions of the position vector $\mathbf{x}$ of P for which the direction of the position vector $\mathbf{y}$ of Q is the same or exactly opposite. What shape does the boundary circle take under this deformation?

Solution. We are looking for vectors $\mathbf{x}$ such that $\mathbf{y} = \lambda\mathbf{x}$. Since $\mathbf{y} = \mathbf{Ax}$, this gives $\mathbf{Ax} = \lambda\mathbf{x}$, the equation of an eigenvalue problem. In components, $\mathbf{Ax} = \lambda\mathbf{x}$ is

$$(2) \qquad \begin{aligned} 5x_1 + 3x_2 &= \lambda x_1 \\ 3x_1 + 5x_2 &= \lambda x_2 \end{aligned} \qquad \text{or} \qquad \begin{aligned} (5 - \lambda)x_1 + \quad 3x_2 \quad &= 0 \\ 3x_1 \quad + (5 - \lambda)x_2 &= 0. \end{aligned}$$

The characteristic equation is

$$(3) \qquad \begin{vmatrix} 5 - \lambda & 3 \\ 3 & 5 - \lambda \end{vmatrix} = (5 - \lambda)^2 - 9 = 0.$$

Its solutions are $\lambda_1 = 8$ and $\lambda_2 = 2$. These are the eigenvalues of our problem. For $\lambda = \lambda_1 = 8$, our system (2) becomes

$$\begin{aligned} -3x_1 + 3x_2 &= 0, \\ 3x_1 - 3x_2 &= 0. \end{aligned} \qquad \left| \qquad \begin{aligned} &\text{Solution } x_2 = x_1,\ x_1 \text{ arbitrary,} \\ &\text{for instance, } x_1 = x_2 = 1. \end{aligned} \right.$$

For $\lambda_2 = 2$, our system (2) becomes

$$\begin{aligned} 3x_1 + 3x_2 &= 0, \\ 3x_1 + 3x_2 &= 0. \end{aligned} \qquad \left| \qquad \begin{aligned} &\text{Solution } x_2 = -x_1,\ x_1 \text{ arbitrary,} \\ &\text{for instance, } x_1 = 1, x_2 = -1. \end{aligned} \right.$$

We thus obtain as eigenvectors of $\mathbf{A}$, for instance, $[1 \quad 1]^\mathsf{T}$ corresponding to λ_1 and $[1 \quad -1]^\mathsf{T}$ corresponding to λ_2 (or a nonzero scalar multiple of these). These vectors make 45° and 135° angles with the positive x_1-direction. They give the principal directions, the answer to our problem. The eigenvalues show that in the principal directions the membrane is stretched by factors 8 and 2, respectively; see Fig. 160.

Accordingly, if we choose the principal directions as directions of a new Cartesian u_1u_2-coordinate system, say, with the positive u_1-semi-axis in the first quadrant and the positive u_2-semi-axis in the second quadrant of the x_1x_2-system, and if we set $u_1 = r\cos\phi$, $u_2 = r\sin\phi$, then a boundary point of the unstretched circular membrane has coordinates $\cos\phi$, $\sin\phi$. Hence, after the stretch we have

$$z_1 = 8\cos\phi, \qquad z_2 = 2\sin\phi.$$

Since $\cos^2\phi + \sin^2\phi = 1$, this shows that the deformed boundary is an ellipse (Fig. 160)

$$(4) \qquad \frac{z_1^2}{8^2} + \frac{z_2^2}{2^2} = 1.$$

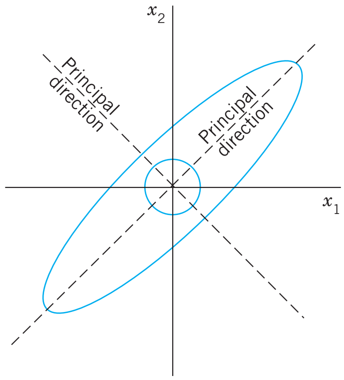


Fig. 160. Undeformed and deformed membrane in Example 1

EXAMPLE 2 Eigenvalue Problems Arising from Markov Processes

Markov processes as considered in Example 13 of Sec. 7.2 lead to eigenvalue problems if we ask for the limit state of the process in which the state vector $\mathbf{x}$ is reproduced under the multiplication by the stochastic matrix $\mathbf{A}$ governing the process, that is, $\mathbf{Ax} = \mathbf{x}$. Hence $\mathbf{A}$ should have the eigenvalue 1, and $\mathbf{x}$ should be a corresponding eigenvector. This is of practical interest because it shows the long-term tendency of the development modeled by the process.

In that example,

$$\mathbf{A} = \begin{bmatrix} 0.7 & 0.1 & 0 \\ 0.2 & 0.9 & 0.2 \\ 0.1 & 0 & 0.8 \end{bmatrix}. \qquad \text{For the transpose,} \qquad \begin{bmatrix} 0.7 & 0.2 & 0.1 \\ 0.1 & 0.9 & 0 \\ 0 & 0.2 & 0.8 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Hence $\mathbf{A}^{\mathsf{T}}$ has the eigenvalue 1, and the same is true for $\mathbf{A}$ by Theorem 3 in Sec. 8.1. An eigenvector $\mathbf{x}$ of $\mathbf{A}$ for $\lambda = 1$ is obtained from

$$\mathbf{A} - \mathbf{I} = \begin{bmatrix} -0.3 & 0.1 & 0 \\ 0.2 & -0.1 & 0.2 \\ 0.1 & 0 & -0.2 \end{bmatrix}, \qquad \text{row-reduced to} \qquad \begin{bmatrix} -\frac{3}{10} & \frac{1}{10} & 0 \\ 0 & -\frac{1}{30} & \frac{1}{5} \\ 0 & 0 & 0 \end{bmatrix}.$$

Taking $x_3 = 1$, we get $x_2 = 6$ from $-x_2/30 + x_3/5 = 0$ and then $x_1 = 2$ from $-3x_1/10 + x_2/10 = 0$. This gives $\mathbf{x} = [2 \quad 6 \quad 1]^{\mathsf{T}}$. It means that in the long run, the ratio Commercial:Industrial:Residential will approach 2:6:1, provided that the probabilities given by $\mathbf{A}$ remain (about) the same. (We switched to ordinary fractions to avoid rounding errors.) ■

EXAMPLE 3 Eigenvalue Problems Arising from Population Models. Leslie Model

The Leslie model describes age-specified population growth, as follows. Let the oldest age attained by the females in some animal population be 9 years. Divide the population into three age classes of 3 years each. Let the "*Leslie matrix*" be

$$\mathbf{L} = [l_{jk}] = \begin{bmatrix} 0 & 2.3 & 0.4 \\ 0.6 & 0 & 0 \\ 0 & 0.3 & 0 \end{bmatrix} \tag{5}$$

where l_{1k} is the average number of daughters born to a single female during the time she is in age class k, and $l_{j,j-1}(j = 2, 3)$ is the fraction of females in age class $j - 1$ that will survive and pass into class j. (a) What is the number of females in each class after 3, 6, 9 years if each class initially consists of 400 females? (b) For what initial distribution will the number of females in each class change by the same proportion? What is this rate of change?

Solution. **(a)** Initially, $\mathbf{x}_{(0)}^{\mathsf{T}} = [400 \quad 400 \quad 400]$. After 3 years,

$$\mathbf{x}_{(3)} = \mathbf{L}\mathbf{x}_{(0)} = \begin{bmatrix} 0 & 2.3 & 0.4 \\ 0.6 & 0 & 0 \\ 0 & 0.3 & 0 \end{bmatrix} \begin{bmatrix} 400 \\ 400 \\ 400 \end{bmatrix} = \begin{bmatrix} 1080 \\ 240 \\ 120 \end{bmatrix}.$$

Similarly, after 6 years the number of females in each class is given by $\mathbf{x}_{(6)}^{\mathsf{T}} = (\mathbf{L}\mathbf{x}_{(3)})^{\mathsf{T}} = [600 \quad 648 \quad 72]$, and after 9 years we have $\mathbf{x}_{(9)}^{\mathsf{T}} = (\mathbf{L}\mathbf{x}_{(6)})^{\mathsf{T}} = [1519.2 \quad 360 \quad 194.4]$.

(b) Proportional change means that we are looking for a distribution vector $\mathbf{x}$ such that $\mathbf{L}\mathbf{x} = \lambda\mathbf{x}$, where λ is the rate of change (growth if $\lambda > 1$, decrease if $\lambda < 1$). The characteristic equation is (develop the characteristic determinant by the first column)

$$\det(\mathbf{L} - \lambda\mathbf{I}) = -\lambda^3 - 0.6(-2.3\lambda - 0.3 \cdot 0.4) = -\lambda^3 + 1.38\lambda + 0.072 = 0.$$

A positive root is found to be (for instance, by Newton's method, Sec. 19.2) $\lambda = 1.2$. A corresponding eigenvector $\mathbf{x}$ can be determined from the characteristic matrix

$$\mathbf{A} - 1.2\mathbf{I} = \begin{bmatrix} -1.2 & 2.3 & 0.4 \\ 0.6 & -1.2 & 0 \\ 0 & 0.3 & -1.2 \end{bmatrix}, \qquad \text{say,} \qquad \mathbf{x} = \begin{bmatrix} 1 \\ 0.5 \\ 0.125 \end{bmatrix}$$

where $x_3 = 0.125$ is chosen, $x_2 = 0.5$ then follows from $0.3x_2 - 1.2x_3 = 0$, and $x_1 = 1$ from $-1.2x_1 + 2.3x_2 + 0.4x_3 = 0$. To get an initial population of 1200 as before, we multiply $\mathbf{x}$ by $1200/(1 + 0.5 + 0.125) = 738$. *Answer:* Proportional growth of the numbers of females in the three classes will occur if the initial values are 738, 369, 92 in classes 1, 2, 3, respectively. The growth rate will be 1.2 per 3 years. ■

EXAMPLE 4 Vibrating System of Two Masses on Two Springs (Fig. 161)

Mass–spring systems involving several masses and springs can be treated as eigenvalue problems. For instance, the mechanical system in Fig. 161 is governed by the system of ODEs

$$\begin{aligned} y_1'' &= -3y_1 - 2(y_1 - y_2) = -5y_1 + 2y_2 \\ y_2'' &= \qquad\quad -2(y_2 - y_1) = 2y_1 - 2y_2 \end{aligned} \tag{6}$$

where y_1 and y_2 are the displacements of the masses from rest, as shown in the figure, and primes denote derivatives with respect to time t. In vector form, this becomes

$$\mathbf{y}'' = \begin{bmatrix} y_1'' \\ y_2'' \end{bmatrix} = \mathbf{A}\mathbf{y} = \begin{bmatrix} -5 & 2 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}. \tag{7}$$

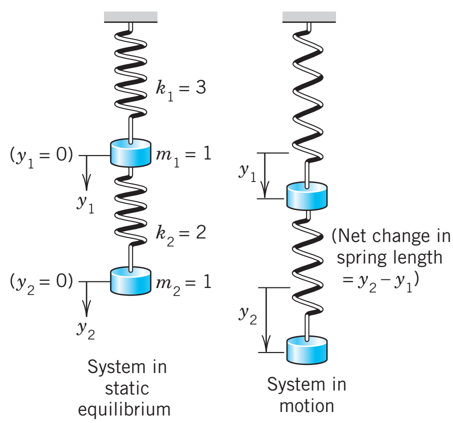


Fig. 161. Masses on springs in Example 4

We try a vector solution of the form

$$\mathbf{y} = \mathbf{x}e^{\omega t}. \tag{8}$$

This is suggested by a mechanical system of a single mass on a spring (Sec. 2.4), whose motion is given by exponential functions (and sines and cosines). Substitution into (7) gives

$$\omega^2\mathbf{x}e^{\omega t} = A\mathbf{x}e^{\omega t}.$$

Dividing by $e^{\omega t}$ and writing $\omega^2 = \lambda$, we see that our mechanical system leads to the eigenvalue problem

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x} \qquad \text{where } \lambda = \omega^2. \tag{9}$$

From Example 1 in Sec. 8.1 we see that $\mathbf{A}$ has the eigenvalues $\lambda_1 = -1$ and $\lambda_2 = -6$. Consequently, $\omega = \pm\sqrt{-1} = \pm i$ and $\sqrt{-6} = \pm i\sqrt{6}$, respectively. Corresponding eigenvectors are

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \qquad \text{and} \qquad \mathbf{x}_2 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}. \tag{10}$$

From (8) we thus obtain the four complex solutions [see (10), Sec. 2.2]

$$\mathbf{x}_1 e^{\pm it} = \mathbf{x}_1(\cos t \pm i \sin t),$$
$$\mathbf{x}_2 e^{\pm i\sqrt{6}t} = \mathbf{x}_2(\cos \sqrt{6}\, t \pm i \sin \sqrt{6}\, t).$$

By addition and subtraction (see Sec. 2.2) we get the four real solutions

$$\mathbf{x}_1 \cos t, \qquad \mathbf{x}_1 \sin t, \qquad \mathbf{x}_2 \cos \sqrt{6}\, t, \qquad \mathbf{x}_2 \sin \sqrt{6}\, t.$$

A general solution is obtained by taking a linear combination of these,

$$\mathbf{y} = \mathbf{x}_1(a_1 \cos t + b_1 \sin t) + \mathbf{x}_2(a_2 \cos \sqrt{6}\, t + b_2 \sin \sqrt{6}\, t)$$

with arbitrary constants a_1, b_1, a_2, b_2 (to which values can be assigned by prescribing initial displacement and initial velocity of each of the two masses). By (10), the components of $\mathbf{y}$ are

$$y_1 = a_1 \cos t + b_1 \sin t + 2a_2 \cos \sqrt{6}\, t + 2b_2 \sin \sqrt{6}\, t$$
$$y_2 = 2a_1 \cos t + 2b_1 \sin t - a_2 \cos \sqrt{6}\, t - b_2 \sin \sqrt{6}\, t.$$

These functions describe harmonic oscillations of the two masses. Physically, this had to be expected because we have neglected damping. ■

PROBLEM SET 8.2

1–6 ELASTIC DEFORMATIONS

Given $\mathbf{A}$ in a deformation $\mathbf{y} = \mathbf{A}\mathbf{x}$, find the principal directions and corresponding factors of extension or contraction. Show the details.

1. $\begin{bmatrix} 3.0 & 1.5 \\ 1.5 & 3.0 \end{bmatrix}$

2. $\begin{bmatrix} 2.0 & 0.4 \\ 0.4 & 2.0 \end{bmatrix}$

3. $\begin{bmatrix} 7 & \sqrt{6} \\ \sqrt{6} & 2 \end{bmatrix}$

4. $\begin{bmatrix} 5 & 2 \\ 2 & 13 \end{bmatrix}$

5. $\begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix}$

6. $\begin{bmatrix} 1.25 & 0.75 \\ 0.75 & 1.25 \end{bmatrix}$

7–9 MARKOV PROCESSES

Find the limit state of the Markov process modeled by the given matrix. Show the details.

7. $\begin{bmatrix} 0.2 & 0.5 \\ 0.8 & 0.5 \end{bmatrix}$

8. $\begin{bmatrix} 0.4 & 0.3 & 0.3 \\ 0.3 & 0.6 & 0.1 \\ 0.3 & 0.1 & 0.6 \end{bmatrix}$

9. $\begin{bmatrix} 0.6 & 0.1 & 0.2 \\ 0.4 & 0.1 & 0.4 \\ 0 & 0.8 & 0.4 \end{bmatrix}$

10–12 AGE-SPECIFIC POPULATION

Find the growth rate in the Leslie model (see Example 3) with the matrix as given. Show the details.

10. $\begin{bmatrix} 0 & 9.0 & 5.0 \\ 0.4 & 0 & 0 \\ 0 & 0.4 & 0 \end{bmatrix}$ **11.** $\begin{bmatrix} 0 & 3.45 & 0.60 \\ 0.90 & 0 & 0 \\ 0 & 0.45 & 0 \end{bmatrix}$

12. $\begin{bmatrix} 0 & 3.0 & 2.0 & 2.0 \\ 0.5 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0.1 & 0 \end{bmatrix}$

13–15 LEONTIEF MODELS[1]

13. Leontief input–output model. Suppose that three industries are interrelated so that their outputs are used as inputs by themselves, according to the 3×3 **consumption matrix**

$$\mathbf{A} = [a_{jk}] = \begin{bmatrix} 0.1 & 0.5 & 0 \\ 0.8 & 0 & 0.4 \\ 0.1 & 0.5 & 0.6 \end{bmatrix}$$

where a_{jk} is the fraction of the output of industry k consumed (purchased) by industry j. Let p_j be the price charged by industry j for its total output. A problem is to find prices so that for each industry, total expenditures equal total income. Show that this leads to $\mathbf{Ap} = \mathbf{p}$, where $\mathbf{p} = [p_1 \quad p_2 \quad p_3]^\mathsf{T}$, and find a solution $\mathbf{p}$ with nonnegative p_1, p_2, p_3.

14. Show that a consumption matrix as considered in Prob. 13 must have column sums 1 and always has the eigenvalue 1.

15. Open Leontief input–output model. If not the whole output but only a portion of it is consumed by the industries themselves, then instead of $\mathbf{Ax} = \mathbf{x}$ (as in Prob. 13), we have $\mathbf{x} - \mathbf{Ax} = \mathbf{y}$, where $\mathbf{x} = [x_1 \quad x_2 \quad x_3]^\mathsf{T}$ is produced, $\mathbf{Ax}$ is consumed by the industries, and, thus, $\mathbf{y}$ is the net production available for other consumers. Find for what production $\mathbf{x}$ a given **demand vector** $\mathbf{y} = [0.1 \quad 0.3 \quad 0.1]^\mathsf{T}$ can be achieved if the consumption matrix is

$$\mathbf{A} = \begin{bmatrix} 0.1 & 0.4 & 0.2 \\ 0.5 & 0 & 0.1 \\ 0.1 & 0.4 & 0.4 \end{bmatrix}.$$

16–20 GENERAL PROPERTIES OF EIGENVALUE PROBLEMS

Let $\mathbf{A} = [a_{jk}]$ be an $n \times n$ matrix with (not necessarily distinct) eigenvalues $\lambda_1, \cdots, \lambda_n$. Show.

16. Trace. The sum of the main diagonal entries, called the *trace* of $\mathbf{A}$, equals the sum of the eigenvalues of $\mathbf{A}$.

17. "Spectral shift." $\mathbf{A} - k\mathbf{I}$ has the eigenvalues $\lambda_1 - k, \cdots, \lambda_n - k$ and the same eigenvectors as $\mathbf{A}$.

18. Scalar multiples, powers. $k\mathbf{A}$ has the eigenvalues $k\lambda_1, \cdots, k\lambda_n$. $\mathbf{A}^m (m = 1, 2, \cdots)$ has the eigenvalues $\lambda_1^m, \cdots, \lambda_n^m$. The eigenvectors are those of $\mathbf{A}$.

19. Spectral mapping theorem. The "**polynomial matrix**"

$$p(\mathbf{A}) = k_m\mathbf{A}^m + k_{m-1}\mathbf{A}^{m-1} + \cdots + k_1\mathbf{A} + k_0\mathbf{I}$$

has the eigenvalues

$$p(\lambda_j) = k_m\lambda_j^m + k_{m-1}\lambda_j^{m-1} + \cdots + k_1\lambda_j + k_0$$

where $j = 1, \cdots, n$, and the same eigenvectors as $\mathbf{A}$.

20. Perron's theorem. A Leslie matrix $\mathbf{L}$ with positive $l_{12}, l_{13}, l_{21}, l_{32}$ has a positive eigenvalue. (This is a special case of the Perron–Frobenius theorem in Sec. 20.7, which is difficult to prove in its general form.)

8.3 Symmetric, Skew-Symmetric, and Orthogonal Matrices

We consider three classes of real square matrices that, because of their remarkable properties, occur quite frequently in applications. The first two matrices have already been mentioned in Sec. 7.2. The goal of Sec. 8.3 is to show their remarkable properties.

[1] WASSILY LEONTIEF (1906–1999). American economist at New York University. For his input–output analysis he was awarded the Nobel Prize in 1973.

DEFINITIONS

Symmetric, Skew-Symmetric, and Orthogonal Matrices

A ***real*** square matrix $\mathbf{A} = [a_{jk}]$ is called

symmetric if transposition leaves it unchanged,

(1) $$\mathbf{A}^\mathsf{T} = \mathbf{A}, \qquad \text{thus} \qquad a_{kj} = a_{jk},$$

skew-symmetric if transposition gives the negative of **A**,

(2) $$\mathbf{A}^\mathsf{T} = -\mathbf{A}, \qquad \text{thus} \qquad a_{kj} = -a_{jk},$$

orthogonal if transposition gives the inverse of **A**,

(3) $$\mathbf{A}^\mathsf{T} = \mathbf{A}^{-1}.$$

EXAMPLE 1 **Symmetric, Skew-Symmetric, and Orthogonal Matrices**

The matrices

$$\begin{bmatrix} -3 & 1 & 5 \\ 1 & 0 & -2 \\ 5 & -2 & 4 \end{bmatrix}, \quad \begin{bmatrix} 0 & 9 & -12 \\ -9 & 0 & 20 \\ 12 & -20 & 0 \end{bmatrix}, \quad \begin{bmatrix} \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \end{bmatrix}$$

are symmetric, skew-symmetric, and orthogonal, respectively, as you should verify. Every skew-symmetric matrix has all main diagonal entries zero. (Can you prove this?) ∎

Any real square matrix **A** may be written as the sum of a symmetric matrix **R** and a skew-symmetric matrix **S**, where

(4) $$\mathbf{R} = \tfrac{1}{2}(\mathbf{A} + \mathbf{A}^\mathsf{T}) \qquad \text{and} \qquad \mathbf{S} = \tfrac{1}{2}(\mathbf{A} - \mathbf{A}^\mathsf{T}).$$

EXAMPLE 2 **Illustration of Formula (4)**

$$\mathbf{A} = \begin{bmatrix} 9 & 5 & 2 \\ 2 & 3 & -8 \\ 5 & 4 & 3 \end{bmatrix} = \mathbf{R} + \mathbf{S} = \begin{bmatrix} 9.0 & 3.5 & 3.5 \\ 3.5 & 3.0 & -2.0 \\ 3.5 & -2.0 & 3.0 \end{bmatrix} + \begin{bmatrix} 0 & 1.5 & -1.5 \\ -1.5 & 0 & -6.0 \\ 1.5 & 6.0 & 0 \end{bmatrix}$$ ∎

THEOREM 1

Eigenvalues of Symmetric and Skew-Symmetric Matrices

(a) *The eigenvalues of a symmetric matrix are real.*

(b) *The eigenvalues of a skew-symmetric matrix are pure imaginary or zero.*

This basic theorem (and an extension of it) will be proved in Sec. 8.5.

EXAMPLE 3 Eigenvalues of Symmetric and Skew-Symmetric Matrices

The matrices in (1) and (7) of Sec. 8.2 are symmetric and have real eigenvalues. The skew-symmetric matrix in Example 1 has the eigenvalues 0, $-25i$, and $25i$. (Verify this.) The following matrix has the real eigenvalues 1 and 5 but is not symmetric. Does this contradict Theorem 1?

$$\begin{bmatrix} 3 & 4 \\ 1 & 3 \end{bmatrix}$$ ■

Orthogonal Transformations and Orthogonal Matrices

Orthogonal transformations are transformations

(5) $$\mathbf{y} = \mathbf{A}\mathbf{x} \qquad \text{where } \mathbf{A} \text{ is an orthogonal matrix.}$$

With each vector $\mathbf{x}$ in R^n such a transformation assigns a vector $\mathbf{y}$ in R^n. For instance, the plane rotation through an angle θ

(6) $$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

is an orthogonal transformation. It can be shown that any orthogonal transformation in the plane or in three-dimensional space is a **rotation** (possibly combined with a reflection in a straight line or a plane, respectively).

The main reason for the importance of orthogonal matrices is as follows.

THEOREM 2 Invariance of Inner Product

An orthogonal transformation preserves the value of the **inner product** *of vectors* $\mathbf{a}$ *and* $\mathbf{b}$ *in* R^n*, defined by*

(7) $$\mathbf{a} \bullet \mathbf{b} = \mathbf{a}^{\mathsf{T}}\mathbf{b} = [a_1 \quad \cdots \quad a_n] \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}.$$

That is, for any $\mathbf{a}$ *and* $\mathbf{b}$ *in* R^n*, orthogonal* $n \times n$ *matrix* $\mathbf{A}$*, and* $\mathbf{u} = \mathbf{A}\mathbf{a}, \mathbf{v} = \mathbf{A}\mathbf{b}$ *we have* $\mathbf{u} \bullet \mathbf{v} = \mathbf{a} \bullet \mathbf{b}$.

Hence the transformation also preserves the **length** *or* **norm** *of any vector* $\mathbf{a}$ *in* R^n *given by*

(8) $$\|\mathbf{a}\| = \sqrt{\mathbf{a} \bullet \mathbf{a}} = \sqrt{\mathbf{a}^{\mathsf{T}}\mathbf{a}}.$$

PROOF Let $\mathbf{A}$ be orthogonal. Let $\mathbf{u} = \mathbf{A}\mathbf{a}$ and $\mathbf{v} = \mathbf{A}\mathbf{b}$. We must show that $\mathbf{u} \bullet \mathbf{v} = \mathbf{a} \bullet \mathbf{b}$. Now $(\mathbf{A}\mathbf{a})^{\mathsf{T}} = \mathbf{a}^{\mathsf{T}}\mathbf{A}^{\mathsf{T}}$ by (10d) in Sec. 7.2 and $\mathbf{A}^{\mathsf{T}}\mathbf{A} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$ by (3). Hence

(9) $$\mathbf{u} \bullet \mathbf{v} = \mathbf{u}^{\mathsf{T}}\mathbf{v} = (\mathbf{A}\mathbf{a})^{\mathsf{T}}\mathbf{A}\mathbf{b} = \mathbf{a}^{\mathsf{T}}\mathbf{A}^{\mathsf{T}}\mathbf{A}\mathbf{b} = \mathbf{a}^{\mathsf{T}}\mathbf{I}\mathbf{b} = \mathbf{a}^{\mathsf{T}}\mathbf{b} = \mathbf{a} \bullet \mathbf{b}.$$

From this the invariance of $\|\mathbf{a}\|$ follows if we set $\mathbf{b} = \mathbf{a}$. ■

Orthogonal matrices have further interesting properties as follows.

THEOREM 3

Orthonormality of Column and Row Vectors

A real square matrix is orthogonal if and only if its column vectors $\mathbf{a}_1, \cdots, \mathbf{a}_n$ *(and also its row vectors) form an* **orthonormal system**, *that is,*

$$\mathbf{a}_j \bullet \mathbf{a}_k = \mathbf{a}_j^\mathsf{T}\mathbf{a}_k = \begin{cases} 0 & \text{if } \ j \neq k \\ 1 & \text{if } \ j = k. \end{cases} \tag{10}$$

PROOF **(a)** Let $\mathbf{A}$ be orthogonal. Then $\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}^\mathsf{T}\mathbf{A} = \mathbf{I}$. In terms of column vectors $\mathbf{a}_1, \cdots, \mathbf{a}_n$,

$$\mathbf{I} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{A}^\mathsf{T}\mathbf{A} = \begin{bmatrix} \mathbf{a}_1^\mathsf{T} \\ \vdots \\ \mathbf{a}_n^\mathsf{T} \end{bmatrix} [\mathbf{a}_1 \cdots \mathbf{a}_n] = \begin{bmatrix} \mathbf{a}_1^\mathsf{T}\mathbf{a}_1 & \mathbf{a}_1^\mathsf{T}\mathbf{a}_2 & \cdots & \mathbf{a}_1^\mathsf{T}\mathbf{a}_n \\ \cdot & \cdot & \cdots & \cdot \\ \mathbf{a}_n^\mathsf{T}\mathbf{a}_1 & \mathbf{a}_n^\mathsf{T}\mathbf{a}_2 & \cdots & \mathbf{a}_n^\mathsf{T}\mathbf{a}_n \end{bmatrix}. \tag{11}$$

The last equality implies (10), by the definition of the $n \times n$ unit matrix $\mathbf{I}$. From (3) it follows that the inverse of an orthogonal matrix is orthogonal (see CAS Experiment 12). Now the column vectors of $\mathbf{A}^{-1}(=\mathbf{A}^\mathsf{T})$ are the row vectors of $\mathbf{A}$. Hence the row vectors of $\mathbf{A}$ also form an orthonormal system.

(b) Conversely, if the column vectors of $\mathbf{A}$ satisfy (10), the off-diagonal entries in (11) must be 0 and the diagonal entries 1. Hence $\mathbf{A}^\mathsf{T}\mathbf{A} = \mathbf{I}$, as (11) shows. Similarly, $\mathbf{A}\mathbf{A}^\mathsf{T} = \mathbf{I}$. This implies $\mathbf{A}^\mathsf{T} = \mathbf{A}^{-1}$ because also $\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$ and the inverse is unique. Hence $\mathbf{A}$ is orthogonal. Similarly when the row vectors of $\mathbf{A}$ form an orthonormal system, by what has been said at the end of part (**a**). ■

THEOREM 4

Determinant of an Orthogonal Matrix

The determinant of an orthogonal matrix has the value $+1$ *or* -1.

PROOF From $\det \mathbf{AB} = \det \mathbf{A} \det \mathbf{B}$ (Sec. 7.8, Theorem 4) and $\det \mathbf{A}^\mathsf{T} = \det \mathbf{A}$ (Sec. 7.7, Theorem 2d), we get for an orthogonal matrix

$$1 = \det \mathbf{I} = \det(\mathbf{A}\mathbf{A}^{-1}) = \det(\mathbf{A}\mathbf{A}^\mathsf{T}) = \det \mathbf{A} \det \mathbf{A}^\mathsf{T} = (\det \mathbf{A})^2. \quad ■$$

EXAMPLE 4 **Illustration of Theorems 3 and 4**

The last matrix in Example 1 and the matrix in (6) illustrate Theorems 3 and 4 because their determinants are -1 and $+1$, as you should verify. ■

THEOREM 5

Eigenvalues of an Orthogonal Matrix

The eigenvalues of an orthogonal matrix $\mathbf{A}$ *are real or complex conjugates in pairs and have absolute value* 1.

PROOF The first part of the statement holds for any real matrix **A** because its characteristic polynomial has real coefficients, so that its zeros (the eigenvalues of **A**) must be as indicated. The claim that $|\lambda| = 1$ will be proved in Sec. 8.5. ■

EXAMPLE 5 Eigenvalues of an Orthogonal Matrix

The orthogonal matrix in Example 1 has the characteristic equation

$$-\lambda^3 + \tfrac{2}{3}\lambda^2 + \tfrac{2}{3}\lambda - 1 = 0.$$

Now one of the eigenvalues must be real (why?), hence $+1$ or -1. Trying, we find -1. Division by $\lambda + 1$ gives $-(\lambda^2 - 5\lambda/3 + 1) = 0$ and the two eigenvalues $(5 + i\sqrt{11})/6$ and $(5 - i\sqrt{11})/6$, which have absolute value 1. Verify all of this. ■

Looking back at this section, you will find that the numerous basic results it contains have relatively short, straightforward proofs. This is typical of large portions of matrix eigenvalue theory.

PROBLEM SET 8.3

1–10 SPECTRUM

Are the following matrices symmetric, skew-symmetric, or orthogonal? Find the spectrum of each, thereby illustrating Theorems 1 and 5. Show your work in detail.

1. $\begin{bmatrix} 0.8 & 0.6 \\ -0.6 & 0.8 \end{bmatrix}$ **2.** $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$

3. $\begin{bmatrix} 2 & 8 \\ -8 & 2 \end{bmatrix}$ **4.** $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$

5. $\begin{bmatrix} 6 & 0 & 0 \\ 0 & 2 & -2 \\ 0 & -2 & 5 \end{bmatrix}$ **6.** $\begin{bmatrix} a & k & k \\ k & a & k \\ k & k & a \end{bmatrix}$

7. $\begin{bmatrix} 0 & 9 & -12 \\ -9 & 0 & 20 \\ 12 & -20 & 0 \end{bmatrix}$ **8.** $\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}$

9. $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix}$ **10.** $\begin{bmatrix} \frac{4}{9} & \frac{8}{9} & \frac{1}{9} \\ -\frac{7}{9} & \frac{4}{9} & -\frac{4}{9} \\ -\frac{4}{9} & \frac{1}{9} & \frac{8}{9} \end{bmatrix}$

11. WRITING PROJECT. Section Summary. Summarize the main concepts and facts in this section, giving illustrative examples of your own.

12. CAS EXPERIMENT. Orthogonal Matrices.

(a) Products. Inverse. Prove that the product of two orthogonal matrices is orthogonal, and so is the inverse of an orthogonal matrix. What does this mean in terms of rotations?

(b) Rotation. Show that (6) is an orthogonal transformation. Verify that it satisfies Theorem 3. Find the inverse transformation.

(c) Powers. Write a program for computing powers $\mathbf{A}^m$ $(m = 1, 2, \cdots)$ of a 2×2 matrix **A** and their spectra. Apply it to the matrix in Prob. 1 (call it **A**). To what rotation does **A** correspond? Do the eigenvalues of $\mathbf{A}^m$ have a limit as $m \to \infty$?

(d) Compute the eigenvalues of $(0.9\mathbf{A})^m$, where **A** is the matrix in Prob. 1. Plot them as points. What is their limit? Along what kind of curve do these points approach the limit?

(e) Find **A** such that $\mathbf{y} = \mathbf{Ax}$ is a counterclockwise rotation through 30° in the plane.

13–20 GENERAL PROPERTIES

13. Verification. Verify the statements in Example 1.

14. Verify the statements in Examples 3 and 4.

15. Sum. Are the eigenvalues of $\mathbf{A} + \mathbf{B}$ sums of the eigenvalues of **A** and of **B**?

16. Orthogonality. Prove that eigenvectors of a symmetric matrix corresponding to different eigenvalues are orthogonal. Give examples.

17. Skew-symmetric matrix. Show that the inverse of a skew-symmetric matrix is skew-symmetric.

18. Do there exist nonsingular skew-symmetric $n \times n$ matrices with odd n?

19. Orthogonal matrix. Do there exist skew-symmetric orthogonal 3×3 matrices?

20. Symmetric matrix. Do there exist nondiagonal symmetric 3×3 matrices that are orthogonal?

8.4 Eigenbases. Diagonalization. Quadratic Forms

So far we have emphasized properties of eigen*values*. We now turn to general properties of eigen*vectors*. Eigenvectors of an $n \times n$ matrix $\mathbf{A}$ may (or may not!) form a basis for R^n. If we are interested in a transformation $\mathbf{y} = \mathbf{A}\mathbf{x}$, such an "**eigenbasis**" (basis of eigenvectors)—if it exists—is of great advantage because then we can represent any $\mathbf{x}$ in R^n uniquely as a linear combination of the eigenvectors $\mathbf{x}_1, \cdots, \mathbf{x}_n$, say,

$$\mathbf{x} = c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_n\mathbf{x}_n.$$

And, denoting the corresponding (not necessarily distinct) eigenvalues of the matrix $\mathbf{A}$ by $\lambda_1, \cdots, \lambda_n$, we have $\mathbf{A}\mathbf{x}_j = \lambda_j\mathbf{x}_j$, so that we simply obtain

$$\begin{aligned} \mathbf{y} = \mathbf{A}\mathbf{x} &= \mathbf{A}(c_1\mathbf{x}_1 + \cdots + c_n\mathbf{x}_n) \\ &= c_1\mathbf{A}\mathbf{x}_1 + \cdots + c_n\mathbf{A}\mathbf{x}_n \\ &= c_1\lambda_1\mathbf{x}_1 + \cdots + c_n\lambda_n\mathbf{x}_n. \end{aligned} \tag{1}$$

This shows that we have decomposed the complicated action of $\mathbf{A}$ on an arbitrary vector $\mathbf{x}$ into a sum of simple actions (multiplication by scalars) on the eigenvectors of $\mathbf{A}$. This is the point of an eigenbasis.

Now if the n eigenvalues are all different, we do obtain a basis:

THEOREM 1

Basis of Eigenvectors

If an $n \times n$ matrix $\mathbf{A}$ has n ***distinct*** *eigenvalues, then $\mathbf{A}$ has a basis of eigenvectors $\mathbf{x}_1, \cdots, \mathbf{x}_n$ for R^n.*

PROOF All we have to show is that $\mathbf{x}_1, \cdots, \mathbf{x}_n$ are linearly independent. Suppose they are not. Let r be the largest integer such that $\{\mathbf{x}_1, \cdots, \mathbf{x}_r\}$ is a linearly independent set. Then $r < n$ and the set $\{\mathbf{x}_1, \cdots, \mathbf{x}_r, \mathbf{x}_{r+1}\}$ is linearly dependent. Thus there are scalars $c_1, \cdots, c_{r+1}$, not all zero, such that

$$c_1\mathbf{x}_1 + \cdots + c_{r+1}\mathbf{x}_{r+1} = \mathbf{0} \tag{2}$$

(see Sec. 7.4). Multiplying both sides by $\mathbf{A}$ and using $\mathbf{A}\mathbf{x}_j = \lambda_j\mathbf{x}_j$, we obtain

$$\mathbf{A}(c_1\mathbf{x}_1 + \cdots + c_{r+1}\mathbf{x}_{r+1}) = c_1\lambda_1\mathbf{x}_1 + \cdots + c_{r+1}\lambda_{r+1}\mathbf{x}_{r+1} = \mathbf{A}\mathbf{0} = \mathbf{0}. \tag{3}$$

To get rid of the last term, we subtract λ_{r+1} times (2) from this, obtaining

$$c_1(\lambda_1 - \lambda_{r+1})\mathbf{x}_1 + \cdots + c_r(\lambda_r - \lambda_{r+1})\mathbf{x}_r = \mathbf{0}.$$

Here $c_1(\lambda_1 - \lambda_{r+1}) = 0, \cdots, c_r(\lambda_r - \lambda_{r+1}) = 0$ since $\{x_1, \cdots, x_r\}$ is linearly independent. Hence $c_1 = \cdots = c_r = 0$, since all the eigenvalues are distinct. But with this, (2) reduces to $c_{r+1}\mathbf{x}_{r+1} = \mathbf{0}$, hence $c_{r+1} = 0$, since $\mathbf{x}_{r+1} \neq \mathbf{0}$ (an eigenvector!). This contradicts the fact that not all scalars in (2) are zero. Hence the conclusion of the theorem must hold. ■

EXAMPLE 1 **Eigenbasis. Nondistinct Eigenvalues. Nonexistence**

The matrix $\mathbf{A} = \begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix}$ has a basis of eigenvectors $\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ corresponding to the eigenvalues $\lambda_1 = 8$, $\lambda_2 = 2$. (See Example 1 in Sec. 8.2.)

Even if not all n eigenvalues are different, a matrix $\mathbf{A}$ may still provide an eigenbasis for R^n. See Example 2 in Sec. 8.1, where $n = 3$.

On the other hand, $\mathbf{A}$ ***may not have enough linearly independent eigenvectors to make up a basis***. For instance, $\mathbf{A}$ in Example 3 of Sec. 8.1 is

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad \text{and has only one eigenvector} \qquad \begin{bmatrix} k \\ 0 \end{bmatrix} \qquad (k \neq 0, \text{ arbitrary}).$$

Actually, eigenbases exist under much more general conditions than those in Theorem 1. An important case is the following.

THEOREM 2 **Symmetric Matrices**

A symmetric matrix has an orthonormal basis of eigenvectors for R^n.

For a proof (which is involved) see Ref. [B3], vol. 1, pp. 270–272.

EXAMPLE 2 **Orthonormal Basis of Eigenvectors**

The first matrix in Example 1 is symmetric, and an orthonormal basis of eigenvectors is $[1/\sqrt{2} \quad 1/\sqrt{2}]^{\mathsf{T}}$, $[1/\sqrt{2} \quad -1/\sqrt{2}]^{\mathsf{T}}$.

Similarity of Matrices. Diagonalization

Eigenbases also play a role in reducing a matrix $\mathbf{A}$ to a diagonal matrix whose entries are the eigenvalues of $\mathbf{A}$. This is done by a "similarity transformation," which is defined as follows (and will have various applications in numerics in Chap. 20).

DEFINITION **Similar Matrices. Similarity Transformation**

An $n \times n$ matrix $\hat{\mathbf{A}}$ is called **similar** to an $n \times n$ matrix $\mathbf{A}$ if

$$\hat{\mathbf{A}} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P} \tag{4}$$

for some (nonsingular!) $n \times n$ matrix $\mathbf{P}$. This transformation, which gives $\hat{\mathbf{A}}$ from $\mathbf{A}$, is called a **similarity transformation**.

The key property of this transformation is that it preserves the eigenvalues of $\mathbf{A}$:

THEOREM 3 **Eigenvalues and Eigenvectors of Similar Matrices**

If $\hat{\mathbf{A}}$ *is similar to* $\mathbf{A}$, *then* $\hat{\mathbf{A}}$ *has the same eigenvalues as* $\mathbf{A}$.

Furthermore, if $\mathbf{x}$ *is an eigenvector of* $\mathbf{A}$, *then* $\mathbf{y} = \mathbf{P}^{-1}\mathbf{x}$ *is an eigenvector of* $\hat{\mathbf{A}}$ *corresponding to the same eigenvalue.*

PROOF From $\mathbf{Ax} = \lambda\mathbf{x}$ (λ an eigenvalue, $\mathbf{x} \neq \mathbf{0}$) we get $\mathbf{P}^{-1}\mathbf{Ax} = \lambda\mathbf{P}^{-1}\mathbf{x}$. Now $\mathbf{I} = \mathbf{PP}^{-1}$. By this ***identity trick*** the equation $\mathbf{P}^{-1}\mathbf{Ax} = \lambda\mathbf{P}^{-1}\mathbf{x}$ gives

$$\mathbf{P}^{-1}\mathbf{Ax} = \mathbf{P}^{-1}\mathbf{AIx} = \mathbf{P}^{-1}\mathbf{APP}^{-1}\mathbf{x} = (\mathbf{P}^{-1}\mathbf{AP})\mathbf{P}^{-1}\mathbf{x} = \hat{\mathbf{A}}(\mathbf{P}^{-1}\mathbf{x}) = \lambda\mathbf{P}^{-1}\mathbf{x}.$$

Hence λ is an eigenvalue of $\hat{\mathbf{A}}$ and $\mathbf{P}^{-1}\mathbf{x}$ a corresponding eigenvector. Indeed, $\mathbf{P}^{-1}\mathbf{x} \neq \mathbf{0}$ because $\mathbf{P}^{-1}\mathbf{x} = \mathbf{0}$ would give $\mathbf{x} = \mathbf{Ix} = \mathbf{PP}^{-1}\mathbf{x} = \mathbf{P0} = \mathbf{0}$, contradicting $\mathbf{x} \neq \mathbf{0}$. ■

EXAMPLE 3 Eigenvalues and Vectors of Similar Matrices

Let,
$$\mathbf{A} = \begin{bmatrix} 6 & -3 \\ 4 & -1 \end{bmatrix} \quad \text{and} \quad \mathbf{P} = \begin{bmatrix} 1 & 3 \\ 1 & 4 \end{bmatrix}.$$

Then
$$\hat{\mathbf{A}} = \begin{bmatrix} 4 & -3 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 6 & -3 \\ 4 & -1 \end{bmatrix}\begin{bmatrix} 1 & 3 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}.$$

Here $\mathbf{P}^{-1}$ was obtained from (4*) in Sec. 7.8 with det $\mathbf{P} = 1$. We see that $\hat{\mathbf{A}}$ has the eigenvalues $\lambda_1 = 3$, $\lambda_2 = 2$. The characteristic equation of $\mathbf{A}$ is $(6 - \lambda)(-1 - \lambda) + 12 = \lambda^2 - 5\lambda + 6 = 0$. It has the roots (the eigenvalues of $\mathbf{A}$) $\lambda_1 = 3$, $\lambda_2 = 2$, confirming the first part of Theorem 3.

We confirm the second part. From the first component of $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$ we have $(6 - \lambda)x_1 - 3x_2 = 0$. For $\lambda = 3$ this gives $3x_1 - 3x_2 = 0$, say, $\mathbf{x}_1 = [1 \quad 1]^\mathsf{T}$. For $\lambda = 2$ it gives $4x_1 - 3x_2 = 0$, say, $\mathbf{x}_2 = [3 \quad 4]^\mathsf{T}$. In Theorem 3 we thus have

$$\mathbf{y}_1 = \mathbf{P}^{-1}\mathbf{x}_1 = \begin{bmatrix} 4 & -3 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad \mathbf{y}_2 = \mathbf{P}^{-1}\mathbf{x}_2 = \begin{bmatrix} 4 & -3 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Indeed, these are eigenvectors of the diagonal matrix $\hat{\mathbf{A}}$.

Perhaps we see that $\mathbf{x}_1$ and $\mathbf{x}_2$ are the columns of $\mathbf{P}$. This suggests the general method of transforming a matrix $\mathbf{A}$ to diagonal form $\mathbf{D}$ by using $\mathbf{P} = \mathbf{X}$, the matrix with eigenvectors as columns. ■

By a suitable similarity transformation we can now transform a matrix $\mathbf{A}$ to a diagonal matrix $\mathbf{D}$ whose diagonal entries are the eigenvalues of $\mathbf{A}$:

THEOREM 4 Diagonalization of a Matrix

If an $n \times n$ matrix $\mathbf{A}$ has a basis of eigenvectors, then

$$\mathbf{D} = \mathbf{X}^{-1}\mathbf{AX} \tag{5}$$

is diagonal, with the eigenvalues of $\mathbf{A}$ as the entries on the main diagonal. Here $\mathbf{X}$ is the matrix with these eigenvectors as column vectors. Also,

$$\mathbf{D}^m = \mathbf{X}^{-1}\mathbf{A}^m\mathbf{X} \qquad (m = 2, 3, \cdots). \tag{5*}$$

PROOF Let $\mathbf{x}_1, \cdots, \mathbf{x}_n$ be a basis of eigenvectors of $\mathbf{A}$ for R^n. Let the corresponding eigenvalues of $\mathbf{A}$ be $\lambda_1, \cdots, \lambda_n$, respectively, so that $\mathbf{A}\mathbf{x}_1 = \lambda_1\mathbf{x}_1, \cdots, \mathbf{A}\mathbf{x}_n = \lambda_n\mathbf{x}_n$. Then $\mathbf{X} = [\mathbf{x}_1 \cdots \mathbf{x}_n]$ has rank n, by Theorem 3 in Sec. 7.4. Hence $\mathbf{X}^{-1}$ exists by Theorem 1 in Sec. 7.8. We claim that

$$\mathbf{A}\mathbf{X} = \mathbf{A}[\mathbf{x}_1 \quad \cdots \quad \mathbf{x}_n] = [\mathbf{A}\mathbf{x}_1 \quad \cdots \quad \mathbf{A}\mathbf{x}_n] = [\lambda_1\mathbf{x}_1 \quad \cdots \quad \lambda_n\mathbf{x}_n] = \mathbf{X}\mathbf{D} \tag{6}$$

where $\mathbf{D}$ is the diagonal matrix as in (5). The fourth equality in (6) follows by direct calculation. (Try it for $n = 2$ and then for general n.) The third equality uses $\mathbf{A}\mathbf{x}_k = \lambda_k\mathbf{x}_k$. The second equality results if we note that the first column of $\mathbf{AX}$ is $\mathbf{A}$ times the first column of $\mathbf{X}$, which is $\mathbf{x}_1$, and so on. For instance, when $n = 2$ and we write $\mathbf{x}_1 = [x_{11} \quad x_{21}]$, $\mathbf{x}_2 = [x_{12} \quad x_{22}]$, we have

$$\begin{aligned}\mathbf{AX} = \mathbf{A}[\mathbf{x}_1 \quad \mathbf{x}_2] &= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} \\ &= \underset{\text{Column 1} \qquad\qquad\qquad \text{Column 2}}{\begin{bmatrix} a_{11}x_{11} + a_{12}x_{21} & a_{11}x_{12} + a_{12}x_{22} \\ a_{21}x_{11} + a_{22}x_{21} & a_{21}x_{12} + a_{22}x_{22} \end{bmatrix}} = [\mathbf{A}\mathbf{x}_1 \quad \mathbf{A}\mathbf{x}_2].\end{aligned}$$

If we multiply (6) by $\mathbf{X}^{-1}$ from the left, we obtain (5). Since (5) is a similarity transformation, Theorem 3 implies that $\mathbf{D}$ has the same eigenvalues as $\mathbf{A}$. Equation (5*) follows if we note that

$$\mathbf{D}^2 = \mathbf{DD} = (\mathbf{X}^{-1}\mathbf{AX})(\mathbf{X}^{-1}\mathbf{AX}) = \mathbf{X}^{-1}\mathbf{A}(\mathbf{XX}^{-1})\mathbf{AX} = \mathbf{X}^{-1}\mathbf{AAX} = \mathbf{X}^{-1}\mathbf{A}^2\mathbf{X}, \quad \text{etc.} \quad ■$$

EXAMPLE 4 **Diagonalization**

Diagonalize

$$\mathbf{A} = \begin{bmatrix} 7.3 & 0.2 & -3.7 \\ -11.5 & 1.0 & 5.5 \\ 17.7 & 1.8 & -9.3 \end{bmatrix}.$$

Solution. The characteristic determinant gives the characteristic equation $-\lambda^3 - \lambda^2 + 12\lambda = 0$. The roots (eigenvalues of $\mathbf{A}$) are $\lambda_1 = 3$, $\lambda_2 = -4$, $\lambda_3 = 0$. By the Gauss elimination applied to $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$ with $\lambda = \lambda_1, \lambda_2, \lambda_3$ we find eigenvectors and then $\mathbf{X}^{-1}$ by the Gauss–Jordan elimination (Sec. 7.8, Example 1). The results are

$$\begin{bmatrix} -1 \\ 3 \\ -1 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}, \quad \begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix}, \qquad \mathbf{X} = \begin{bmatrix} -1 & 1 & 2 \\ 3 & -1 & 1 \\ -1 & 3 & 4 \end{bmatrix}, \qquad \mathbf{X}^{-1} = \begin{bmatrix} -0.7 & 0.2 & 0.3 \\ -1.3 & -0.2 & 0.7 \\ 0.8 & 0.2 & -0.2 \end{bmatrix}.$$

Calculating $\mathbf{AX}$ and multiplying by $\mathbf{X}^{-1}$ from the left, we thus obtain

$$\mathbf{D} = \mathbf{X}^{-1}\mathbf{AX} = \begin{bmatrix} -0.7 & 0.2 & 0.3 \\ -1.3 & -0.2 & 0.7 \\ 0.8 & 0.2 & -0.2 \end{bmatrix}\begin{bmatrix} -3 & -4 & 0 \\ 9 & 4 & 0 \\ -3 & -12 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \quad ■$$

Quadratic Forms. Transformation to Principal Axes

By definition, a **quadratic form** Q in the components $x_1, \cdots, x_n$ of a vector $\mathbf{x}$ is a sum of n^2 terms, namely,

$$
\textbf{(7)} \qquad
\begin{aligned}
Q = \mathbf{x}^{\mathsf{T}}\mathbf{A}\mathbf{x} &= \sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk}x_j x_k \\
&= \quad a_{11}x_1^2 \quad + a_{12}x_1x_2 + \cdots + a_{1n}x_1x_n \\
&\quad + a_{21}x_2x_1 + a_{22}x_2^2 \quad + \cdots + a_{2n}x_2x_n \\
&\quad + \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
&\quad + a_{n1}x_nx_1 + a_{n2}x_nx_2 + \cdots + a_{nn}x_n^2.
\end{aligned}
$$

$\mathbf{A} = [a_{jk}]$ is called the **coefficient matrix** of the form. We may assume that $\mathbf{A}$ is ***symmetric***, because we can take off-diagonal terms together in pairs and write the result as a sum of two equal terms; see the following example.

EXAMPLE 5 Quadratic Form. Symmetric Coefficient Matrix

Let

$$\mathbf{x}^{\mathsf{T}}\mathbf{A}\mathbf{x} = [x_1 \quad x_2]\begin{bmatrix} 3 & 4 \\ 6 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 3x_1^2 + 4x_1x_2 + 6x_2x_1 + 2x_2^2 = 3x_1^2 + 10x_1x_2 + 2x_2^2.$$

Here $4 + 6 = 10 = 5 + 5$. From the corresponding ***symmetric*** matrix $\mathbf{C} = [c_{jk}]$, where $c_{jk} = \frac{1}{2}(a_{jk} + a_{kj})$, thus $c_{11} = 3$, $c_{12} = c_{21} = 5$, $c_{22} = 2$, we get the same result; indeed,

$$\mathbf{x}^{\mathsf{T}}\mathbf{C}\mathbf{x} = [x_1 \quad x_2]\begin{bmatrix} 3 & 5 \\ 5 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 3x_1^2 + 5x_1x_2 + 5x_2x_1 + 2x_2^2 = 3x_1^2 + 10x_1x_2 + 2x_2^2. \qquad \blacksquare$$

Quadratic forms occur in physics and geometry, for instance, in connection with conic sections (ellipses $x_1^2/a^2 + x_2^2/b^2 = 1$, etc.) and quadratic surfaces (cones, etc.). Their transformation to principal axes is an important practical task related to the diagonalization of matrices, as follows.

By Theorem 2, the *symmetric* coefficient matrix $\mathbf{A}$ of (7) has an orthonormal basis of eigenvectors. Hence if we take these as column vectors, we obtain a matrix $\mathbf{X}$ that is orthogonal, so that $\mathbf{X}^{-1} = \mathbf{X}^{\mathsf{T}}$. From (5) we thus have $\mathbf{A} = \mathbf{X}\mathbf{D}\mathbf{X}^{-1} = \mathbf{X}\mathbf{D}\mathbf{X}^{\mathsf{T}}$. Substitution into (7) gives

$$(8) \qquad Q = \mathbf{x}^{\mathsf{T}}\mathbf{X}\mathbf{D}\mathbf{X}^{\mathsf{T}}\mathbf{x}.$$

If we set $\mathbf{X}^{\mathsf{T}}\mathbf{x} = \mathbf{y}$, then, since $\mathbf{X}^{\mathsf{T}} = \mathbf{X}^{-1}$, we have $\mathbf{X}^{-1}\mathbf{x} = \mathbf{y}$ and thus obtain

$$(9) \qquad \mathbf{x} = \mathbf{X}\mathbf{y}.$$

Furthermore, in (8) we have $\mathbf{x}^{\mathsf{T}}\mathbf{X} = (\mathbf{X}^{\mathsf{T}}\mathbf{x})^{\mathsf{T}} = \mathbf{y}^{\mathsf{T}}$ and $\mathbf{X}^{\mathsf{T}}\mathbf{x} = \mathbf{y}$, so that Q becomes simply

$$\textbf{(10)} \qquad Q = \mathbf{y}^{\mathsf{T}}\mathbf{D}\mathbf{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2.$$

This proves the following basic theorem.

THEOREM 5

Principal Axes Theorem

The substitution (9) *transforms a quadratic form*

$$Q = \mathbf{x}^{\mathsf{T}}\mathbf{A}\mathbf{x} = \sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk}x_jx_k \qquad (a_{kj} = a_{jk})$$

to the principal axes form or **canonical form** (10), *where* $\lambda_1, \cdots, \lambda_n$ *are the* (*not necessarily distinct*) *eigenvalues of the* (*symmetric!*) *matrix* $\mathbf{A}$, *and* $\mathbf{X}$ *is an orthogonal matrix with corresponding eigenvectors* $\mathbf{x}_1, \cdots, \mathbf{x}_n$, *respectively, as column vectors.*

EXAMPLE 6

Transformation to Principal Axes. Conic Sections

Find out what type of conic section the following quadratic form represents and transform it to principal axes:

$$Q = 17x_1^2 - 30x_1x_2 + 17x_2^2 = 128.$$

Solution. We have $Q = \mathbf{x}^{\mathsf{T}}\mathbf{A}\mathbf{x}$, where

$$\mathbf{A} = \begin{bmatrix} 17 & -15 \\ -15 & 17 \end{bmatrix}, \qquad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

This gives the characteristic equation $(17 - \lambda)^2 - 15^2 = 0$. It has the roots $\lambda_1 = 2$, $\lambda_2 = 32$. Hence (10) becomes

$$Q = 2y_1^2 + 32y_2^2.$$

We see that $Q = 128$ represents the ellipse $2y_1^2 + 32y_2^2 = 128$, that is,

$$\frac{y_1^2}{8^2} + \frac{y_2^2}{2^2} = 1.$$

If we want to know the direction of the principal axes in the x_1x_2-coordinates, we have to determine normalized eigenvectors from $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$ with $\lambda = \lambda_1 = 2$ and $\lambda = \lambda_2 = 32$ and then use (9). We get

$$\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix},$$

hence

$$\mathbf{x} = \mathbf{X}\mathbf{y} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \qquad \begin{aligned} x_1 &= y_1/\sqrt{2} - y_2/\sqrt{2} \\ x_2 &= y_1/\sqrt{2} + y_2/\sqrt{2}. \end{aligned}$$

This is a 45° rotation. Our results agree with those in Sec. 8.2, Example 1, except for the notations. See also Fig. 160 in that example. ■

PROBLEM SET 8.4

1–5 SIMILAR MATRICES HAVE EQUAL EIGENVALUES

Verify this for $\mathbf{A}$ and $\hat{\mathbf{A}} = \mathbf{P}^{-1}\mathbf{AP}$. If $\mathbf{y}$ is an eigenvector of $\mathbf{P}$, show that $\mathbf{x} = \mathbf{Py}$ are eigenvectors of $\mathbf{A}$. Show the details of your work.

1. $\mathbf{A} = \begin{bmatrix} 3 & 4 \\ 4 & -3 \end{bmatrix}, \quad \mathbf{P} = \begin{bmatrix} -4 & 2 \\ 3 & -1 \end{bmatrix}$

2. $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix}, \quad \mathbf{P} = \begin{bmatrix} 7 & -5 \\ 10 & -7 \end{bmatrix}$

3. $\mathbf{A} = \begin{bmatrix} 8 & -4 \\ 2 & 2 \end{bmatrix}, \quad \mathbf{P} = \begin{bmatrix} 0.28 & 0.96 \\ -0.96 & 0.28 \end{bmatrix}$

4. $\mathbf{A} = \begin{bmatrix} 0 & 0 & 2 \\ 0 & 3 & 2 \\ 1 & 0 & 1 \end{bmatrix}, \quad \mathbf{P} = \begin{bmatrix} 2 & 0 & 3 \\ 0 & 1 & 0 \\ 3 & 0 & 5 \end{bmatrix},$
$\lambda_1 = 3$

5. $\mathbf{A} = \begin{bmatrix} -5 & 0 & 15 \\ 3 & 4 & -9 \\ -5 & 0 & 15 \end{bmatrix}, \quad \mathbf{P} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

6. PROJECT. Similarity of Matrices. Similarity is basic, for instance, in designing numeric methods.

(a) Trace. By definition, the **trace** of an $n \times n$ matrix $\mathbf{A} = [a_{jk}]$ is the sum of the diagonal entries,

$$\text{trace } \mathbf{A} = a_{11} + a_{22} + \cdots + a_{nn}.$$

Show that the trace equals the sum of the eigenvalues, each counted as often as its algebraic multiplicity indicates. Illustrate this with the matrices $\mathbf{A}$ in Probs. 1, 3, and 5.

(b) Trace of product. Let $\mathbf{B} = [b_{jk}]$ be $n \times n$. Show that similar matrices have equal traces, by first proving

$$\text{trace } \mathbf{AB} = \sum_{i=1}^{n} \sum_{l=1}^{n} a_{il} b_{li} = \text{trace } \mathbf{BA}.$$

(c) Find a relationship between $\hat{\mathbf{A}}$ in (4) and $\hat{\hat{\mathbf{A}}} = \mathbf{PAP}^{-1}$.

(d) Diagonalization. What can you do in (5) if you want to change the order of the eigenvalues in $\mathbf{D}$, for instance, interchange $d_{11} = \lambda_1$ and $d_{22} = \lambda_2$?

7. No basis. Find further 2×2 and 3×3 matrices without eigenbasis.

8. Orthonormal basis. Illustrate Theorem 2 with further examples.

9–16 DIAGONALIZATION OF MATRICES

Find an eigenbasis (a basis of eigenvectors) and diagonalize. Show the details.

9. $\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$

10. $\begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix}$

11. $\begin{bmatrix} -19 & 7 \\ -42 & 16 \end{bmatrix}$

12. $\begin{bmatrix} -4.3 & 7.7 \\ 1.3 & 9.3 \end{bmatrix}$

13. $\begin{bmatrix} 4 & 0 & 0 \\ 12 & -2 & 0 \\ 21 & -6 & 1 \end{bmatrix}$

14. $\begin{bmatrix} -5 & -6 & 6 \\ -9 & -8 & 12 \\ -12 & -12 & 16 \end{bmatrix}, \quad \lambda_1 = -2$

15. $\begin{bmatrix} 4 & 3 & 3 \\ 3 & 6 & 1 \\ 3 & 1 & 6 \end{bmatrix}, \quad \lambda_1 = 10$

16. $\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & -4 \end{bmatrix}$

17–23 PRINCIPAL AXES. CONIC SECTIONS

What kind of conic section (or pair of straight lines) is given by the quadratic form? Transform it to principal axes. Express $\mathbf{x}^\mathsf{T} = [x_1 \quad x_2]$ in terms of the new coordinate vector $\mathbf{y}^\mathsf{T} = [y_1 \quad y_2]$, as in Example 6.

17. $7x_1^2 + 6x_1x_2 + 7x_2^2 = 200$

18. $3x_1^2 + 8x_1x_2 - 3x_2^2 = 10$

19. $3x_1^2 + 22x_1x_2 + 3x_2^2 = 0$

20. $9x_1^2 + 6x_1x_2 + x_2^2 = 10$

21. $x_1^2 - 12x_1x_2 + x_2^2 = 70$

22. $4x_1^2 + 12x_1x_2 + 13x_2^2 = 16$

23. $-11x_1^2 + 84x_1x_2 + 24x_2^2 = 156$

24. Definiteness. A quadratic form $Q(\mathbf{x}) = \mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x}$ and its (symmetric!) matrix $\mathbf{A}$ are called **(a) positive definite** if $Q(\mathbf{x}) > 0$ for all $\mathbf{x} \neq \mathbf{0}$, **(b) negative definite** if $Q(\mathbf{x}) < 0$ for all $\mathbf{x} \neq \mathbf{0}$, **(c) indefinite** if $Q(\mathbf{x})$ takes both positive and negative values. (See Fig. 162.) [$Q(\mathbf{x})$ and $\mathbf{A}$ are called *positive semidefinite* (*negative semidefinite*) if $Q(\mathbf{x}) \geqq 0$ ($Q(\mathbf{x}) \leqq 0$) for all $\mathbf{x}$.] Show that a necessary and sufficient condition for (a), (b), and (c) is that the eigenvalues of $\mathbf{A}$ are (a) all positive, (b) all negative, and (c) both positive and negative.

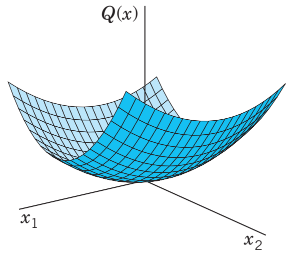
Hint. Use Theorem 5.

25. Definiteness. A necessary and sufficient condition for positive definiteness of a quadratic form $Q(\mathbf{x}) = \mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x}$ with *symmetric* matrix $\mathbf{A}$ is that all the **principal minors** are positive (see Ref. [B3], vol. 1, p. 306), that is,

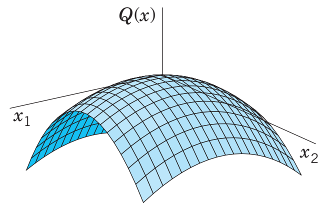
$$a_{11} > 0, \qquad \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix} > 0,$$

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{vmatrix} > 0, \qquad \cdots, \qquad \det \mathbf{A} > 0.$$

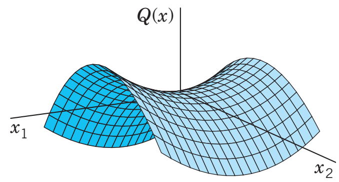
Show that the form in Prob. 22 is positive definite, whereas that in Prob. 23 is indefinite.



(*a*) Positive definite form



(*b*) Negative definite form



(*c*) Indefinite form

Fig. 162. Quadratic forms in two variables (Problem 24)

8.5 Complex Matrices and Forms. *Optional*

The three classes of matrices in Sec. 8.3 have complex counterparts which are of practical interest in certain applications, for instance, in quantum mechanics. This is mainly because of their spectra as shown in Theorem 1 in this section. The second topic is about extending quadratic forms of Sec. 8.4 to complex numbers. (The reader who wants to brush up on complex numbers may want to consult Sec. 13.1.)

Notations

$\overline{\mathbf{A}} = [\overline{a}_{jk}]$ is obtained from $\mathbf{A} = [a_{jk}]$ by replacing each entry $a_{jk} = \alpha + i\beta$ (α, β real) with its complex conjugate $\overline{a}_{jk} = \alpha - i\beta$. Also, $\overline{\mathbf{A}}^\mathsf{T} = [\overline{a}_{kj}]$ is the transpose of $\overline{\mathbf{A}}$, hence the conjugate transpose of $\mathbf{A}$.

EXAMPLE 1 Notations

$$\text{If } \mathbf{A} = \begin{bmatrix} 3+4i & 1-i \\ 6 & 2-5i \end{bmatrix}, \quad \text{then} \quad \overline{\mathbf{A}} = \begin{bmatrix} 3-4i & 1+i \\ 6 & 2+5i \end{bmatrix} \quad \text{and} \quad \overline{\mathbf{A}}^\mathsf{T} = \begin{bmatrix} 3-4i & 6 \\ 1+i & 2+5i \end{bmatrix}.$$

DEFINITION

Hermitian, Skew-Hermitian, and Unitary Matrices

A square matrix $\mathbf{A} = [a_{kj}]$ is called

Hermitian	if $\overline{\mathbf{A}}^{\mathsf{T}} = \mathbf{A}$,	that is,	$\overline{a}_{kj} = a_{jk}$
skew-Hermitian	if $\overline{\mathbf{A}}^{\mathsf{T}} = -\mathbf{A}$,	that is,	$\overline{a}_{kj} = -a_{jk}$
unitary	if $\overline{\mathbf{A}}^{\mathsf{T}} = \mathbf{A}^{-1}$.		

The first two classes are named after Hermite (see footnote 13 in Problem Set 5.8).

From the definitions we see the following. If $\mathbf{A}$ is Hermitian, the entries on the main diagonal must satisfy $\overline{a}_{jj} = a_{jj}$; that is, they are real. Similarly, if $\mathbf{A}$ is skew-Hermitian, then $\overline{a}_{jj} = -a_{jj}$. If we set $a_{jj} = \alpha + i\beta$, this becomes $\alpha - i\beta = -(\alpha + i\beta)$. Hence $\alpha = 0$, so that a_{jj} must be pure imaginary or 0.

EXAMPLE 2 Hermitian, Skew-Hermitian, and Unitary Matrices

$$\mathbf{A} = \begin{bmatrix} 4 & 1 - 3i \\ 1 + 3i & 7 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} 3i & 2 + i \\ -2 + i & -i \end{bmatrix} \qquad \mathbf{C} = \begin{bmatrix} \frac{1}{2}i & \frac{1}{2}\sqrt{3} \\ \frac{1}{2}\sqrt{3} & \frac{1}{2}i \end{bmatrix}$$

are Hermitian, skew-Hermitian, and unitary matrices, respectively, as you may verify by using the definitions. ■

If a Hermitian matrix is real, then $\overline{\mathbf{A}}^{\mathsf{T}} = \mathbf{A}^{\mathsf{T}} = \mathbf{A}$. Hence a real Hermitian matrix is a symmetric matrix (Sec. 8.3).

Similarly, if a skew-Hermitian matrix is real, then $\overline{\mathbf{A}}^{\mathsf{T}} = \mathbf{A}^{\mathsf{T}} = -\mathbf{A}$. Hence a real skew-Hermitian matrix is a skew-symmetric matrix.

Finally, if a unitary matrix is real, then $\overline{\mathbf{A}}^{\mathsf{T}} = \mathbf{A}^{\mathsf{T}} = \mathbf{A}^{-1}$. Hence a real unitary matrix is an orthogonal matrix.

This shows that *Hermitian, skew-Hermitian, and unitary matrices generalize symmetric, skew-symmetric, and orthogonal matrices, respectively.*

Eigenvalues

It is quite remarkable that the matrices under consideration have spectra (sets of eigenvalues; see Sec. 8.1) that can be characterized in a general way as follows (see Fig. 163).

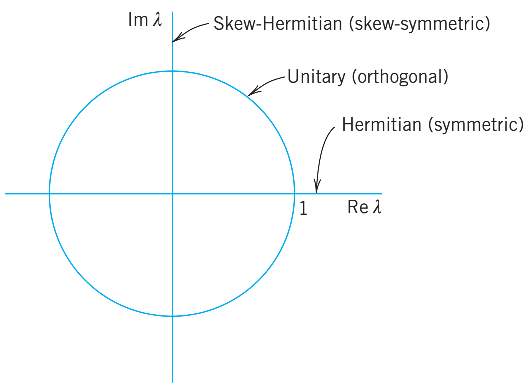


Fig. 163. Location of the eigenvalues of Hermitian, skew-Hermitian, and unitary matrices in the complex λ-plane

THEOREM 1

Eigenvalues

(a) *The eigenvalues of a Hermitian matrix (and thus of a symmetric matrix) are real.*

(b) *The eigenvalues of a skew-Hermitian matrix (and thus of a skew-symmetric matrix) are pure imaginary or zero.*

(c) *The eigenvalues of a unitary matrix (and thus of an orthogonal matrix) have absolute value* 1.

EXAMPLE 3

Illustration of Theorem 1

For the matrices in Example 2 we find by direct calculation

Matrix		Characteristic Equation	Eigenvalues
$\mathbf{A}$	Hermitian	$\lambda^2 - 11\lambda + 18 = 0$	$9, \quad 2$
$\mathbf{B}$	Skew-Hermitian	$\lambda^2 - 2i\lambda + 8 = 0$	$4i, \quad -2i$
$\mathbf{C}$	Unitary	$\lambda^2 - i\lambda - 1 = 0$	$\frac{1}{2}\sqrt{3} + \frac{1}{2}i, \quad -\frac{1}{2}\sqrt{3} + \frac{1}{2}i$

and $|\pm\frac{1}{2}\sqrt{3} + \frac{1}{2}i|^2 = \frac{3}{4} + \frac{1}{4} = 1$. ■

PROOF We prove Theorem 1. Let λ be an eigenvalue and $\mathbf{x}$ an eigenvector of $\mathbf{A}$. Multiply $\mathbf{Ax} = \lambda\mathbf{x}$ from the left by $\bar{\mathbf{x}}^\mathsf{T}$, thus $\bar{\mathbf{x}}^\mathsf{T}\mathbf{Ax} = \lambda\bar{\mathbf{x}}^\mathsf{T}\mathbf{x}$, and divide by $\bar{\mathbf{x}}^\mathsf{T}\mathbf{x} = \bar{x}_1 x_1 + \cdots + \bar{x}_n x_n = |x_1|^2 + \cdots + |x_n|^2$, which is real and not 0 because $\mathbf{x} \neq \mathbf{0}$. This gives

$$\lambda = \frac{\bar{\mathbf{x}}^\mathsf{T}\mathbf{Ax}}{\bar{\mathbf{x}}^\mathsf{T}\mathbf{x}}. \tag{1}$$

(a) If $\mathbf{A}$ is Hermitian, $\bar{\mathbf{A}}^\mathsf{T} = \mathbf{A}$ or $\mathbf{A}^\mathsf{T} = \bar{\mathbf{A}}$ and we show that then the numerator in (1) is real, which makes λ real. $\bar{\mathbf{x}}^\mathsf{T}\mathbf{Ax}$ is a scalar; hence taking the transpose has no effect. Thus

$$\bar{\mathbf{x}}^\mathsf{T}\mathbf{Ax} = (\bar{\mathbf{x}}^\mathsf{T}\mathbf{Ax})^\mathsf{T} = \mathbf{x}^\mathsf{T}\mathbf{A}^\mathsf{T}\bar{\mathbf{x}} = \mathbf{x}^\mathsf{T}\bar{\mathbf{A}}\bar{\mathbf{x}} = (\overline{\bar{\mathbf{x}}^\mathsf{T}\mathbf{Ax}}). \tag{2}$$

Hence, $\bar{\mathbf{x}}^\mathsf{T}\mathbf{Ax}$ equals its complex conjugate, so that it must be real. ($a + ib = a - ib$ implies $b = 0$.)

(b) If $\mathbf{A}$ is skew-Hermitian, $\mathbf{A}^\mathsf{T} = -\bar{\mathbf{A}}$ and instead of (2) we obtain

$$\bar{\mathbf{x}}^\mathsf{T}\mathbf{Ax} = -(\overline{\bar{\mathbf{x}}^\mathsf{T}\mathbf{Ax}}) \tag{3}$$

so that $\bar{\mathbf{x}}^\mathsf{T}\mathbf{Ax}$ equals minus its complex conjugate and is pure imaginary or 0. ($a + ib = -(a - ib)$ implies $a = 0$.)

(c) Let $\mathbf{A}$ be unitary. We take $\mathbf{Ax} = \lambda\mathbf{x}$ and its conjugate transpose

$$(\overline{\mathbf{Ax}})^\mathsf{T} = (\overline{\lambda\mathbf{x}})^\mathsf{T} = \bar{\lambda}\bar{\mathbf{x}}^\mathsf{T}$$

and multiply the two left sides and the two right sides,

$$(\overline{\mathbf{Ax}})^\mathsf{T}\mathbf{Ax} = \bar{\lambda}\lambda\bar{\mathbf{x}}^\mathsf{T}\mathbf{x} = |\lambda|^2\bar{\mathbf{x}}^\mathsf{T}\mathbf{x}.$$

But $\mathbf{A}$ is unitary, $\overline{\mathbf{A}}^{\mathsf{T}} = \mathbf{A}^{-1}$, so that on the left we obtain

$$(\overline{\mathbf{A}}\overline{\mathbf{x}})^{\mathsf{T}}\mathbf{A}\mathbf{x} = \overline{\mathbf{x}}^{\mathsf{T}}\overline{\mathbf{A}}^{\mathsf{T}}\mathbf{A}\mathbf{x} = \overline{\mathbf{x}}^{\mathsf{T}}\mathbf{A}^{-1}\mathbf{A}\mathbf{x} = \overline{\mathbf{x}}^{\mathsf{T}}\mathbf{I}\mathbf{x} = \overline{\mathbf{x}}^{\mathsf{T}}\mathbf{x}.$$

Together, $\overline{\mathbf{x}}^{\mathsf{T}}\mathbf{x} = |\lambda|^2\overline{\mathbf{x}}^{\mathsf{T}}\mathbf{x}$. We now divide by $\overline{\mathbf{x}}^{\mathsf{T}}\mathbf{x}\ (\neq 0)$ to get $|\lambda|^2 = 1$. Hence $|\lambda| = 1$.
This proves Theorem 1 as well as Theorems 1 and 5 in Sec. 8.3. ■

Key properties of orthogonal matrices (invariance of the inner product, orthonormality of rows and columns; see Sec. 8.3) generalize to unitary matrices in a remarkable way.

To see this, instead of R^n we now use the **complex vector space** C^n of all complex vectors with n complex numbers as components, and complex numbers as scalars. For such complex vectors the **inner product** is defined by (note the overbar for the complex conjugate)

$$\mathbf{a} \bullet \mathbf{b} = \overline{\mathbf{a}}^{\mathsf{T}}\mathbf{b}. \tag{4}$$

The **length** or **norm** of such a complex vector is a *real* number defined by

$$\|\mathbf{a}\| = \sqrt{\mathbf{a} \bullet \mathbf{a}} = \sqrt{\overline{\mathbf{a}}_j^{\mathsf{T}}\mathbf{a}} = \sqrt{\overline{a}_1 a_1 + \cdots + \overline{a}_n a_n} = \sqrt{|a_1|^2 + \cdots + |a_n|^2}. \tag{5}$$

THEOREM 2

Invariance of Inner Product

A **unitary transformation**, *that is,* $\mathbf{y} = \mathbf{A}\mathbf{x}$ *with a unitary matrix* $\mathbf{A}$, *preserves the value of the inner product* (4), *hence also the norm* (5).

PROOF The proof is the same as that of Theorem 2 in Sec. 8.3, which the theorem generalizes. In the analog of (9), Sec. 8.3, we now have bars,

$$\mathbf{u} \bullet \mathbf{v} = \overline{\mathbf{u}}^{\mathsf{T}}\mathbf{v} = (\overline{\mathbf{A}}\overline{\mathbf{a}})^{\mathsf{T}}\mathbf{A}\mathbf{b} = \overline{\mathbf{a}}^{\mathsf{T}}\overline{\mathbf{A}}^{\mathsf{T}}\mathbf{A}\mathbf{b} = \overline{\mathbf{a}}^{\mathsf{T}}\mathbf{I}\mathbf{b} = \overline{\mathbf{a}}^{\mathsf{T}}\mathbf{b} = \mathbf{a} \bullet \mathbf{b}.$$

The complex analog of an orthonormal system of real vectors (see Sec. 8.3) is defined as follows.

DEFINITION

Unitary System

A *unitary system* is a set of complex vectors satisfying the relationships

$$\mathbf{a}_j \bullet \mathbf{a}_k = \overline{\mathbf{a}}_j^{\mathsf{T}}\mathbf{a}_k = \begin{cases} 0 & \text{if} \quad j \neq k \\ 1 & \text{if} \quad j = k. \end{cases} \tag{6}$$

Theorem 3 in Sec. 8.3 extends to complex as follows.

THEOREM 3

Unitary Systems of Column and Row Vectors

A complex square matrix is unitary if and only if its column vectors (and also its row vectors) form a unitary system.

PROOF The proof is the same as that of Theorem 3 in Sec. 8.3, except for the bars required in $\bar{\mathbf{A}}^{\mathsf{T}} = \mathbf{A}^{-1}$ and in (4) and (6) of the present section.

THEOREM 4

Determinant of a Unitary Matrix

Let $\mathbf{A}$ *be a unitary matrix. Then its determinant has absolute value one, that is,* $|\det \mathbf{A}| = 1$.

PROOF Similarly, as in Sec. 8.3, we obtain

$$\begin{aligned} 1 &= \det(\mathbf{A}\mathbf{A}^{-1}) = \det(\mathbf{A}\bar{\mathbf{A}}^{\mathsf{T}}) = \det \mathbf{A} \det \bar{\mathbf{A}}^{\mathsf{T}} = \det \mathbf{A} \det \bar{\mathbf{A}} \\ &= \det \mathbf{A}\, \overline{\det \mathbf{A}} = |\det \mathbf{A}|^2. \end{aligned}$$

Hence $|\det \mathbf{A}| = 1$ (where $\det \mathbf{A}$ may now be complex). ■

EXAMPLE 4 **Unitary Matrix Illustrating Theorems 1c and 2–4**

For the vectors $\mathbf{a}^{\mathsf{T}} = [2 \quad -i]$ and $\mathbf{b}^{\mathsf{T}} = [1 + i \quad 4i]$ we get $\bar{\mathbf{a}}^{\mathsf{T}} = [2 \quad i]^{\mathsf{T}}$ and $\bar{\mathbf{a}}^{\mathsf{T}}\mathbf{b} = 2(1 + i) - 4 = -2 + 2i$ and with

$$\mathbf{A} = \begin{bmatrix} 0.8i & 0.6 \\ 0.6 & 0.8i \end{bmatrix} \quad \text{also} \quad \mathbf{A}\mathbf{a} = \begin{bmatrix} i \\ 2 \end{bmatrix} \quad \text{and} \quad \mathbf{A}\mathbf{b} = \begin{bmatrix} -0.8 + 3.2i \\ -2.6 + 0.6i \end{bmatrix},$$

as one can readily verify. This gives $(\overline{\mathbf{A}\mathbf{a}})^{\mathsf{T}}\mathbf{A}\mathbf{b} = -2 + 2i$, illustrating Theorem 2. The matrix is unitary. Its columns form a unitary system,

$$\bar{\mathbf{a}}_1^{\mathsf{T}}\mathbf{a}_1 = -0.8i \cdot 0.8i + 0.6^2 = 1, \qquad \bar{\mathbf{a}}_1^{\mathsf{T}}\mathbf{a}_2 = -0.8i \cdot 0.6 + 0.6 \cdot 0.8i = 0,$$
$$\bar{\mathbf{a}}_2^{\mathsf{T}}\mathbf{a}_2 = 0.6^2 + (-0.8i)0.8i = 1$$

and so do its rows. Also, $\det \mathbf{A} = -1$. The eigenvalues are $0.6 + 0.8i$ and $-0.6 + 0.8i$, with eigenvectors $[1 \quad 1]^{\mathsf{T}}$ and $[1 \quad -1]^{\mathsf{T}}$, respectively. ■

Theorem 2 in Sec. 8.4 on the existence of an eigenbasis extends to complex matrices as follows.

THEOREM 5

Basis of Eigenvectors

A Hermitian, skew-Hermitian, or unitary matrix has a basis of eigenvectors for C^n *that is a unitary system.*

For a proof see Ref. [B3], vol. 1, pp. 270–272 and p. 244 (Definition 2).

EXAMPLE 5 **Unitary Eigenbases**

The matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ in Example 2 have the following unitary systems of eigenvectors, as you should verify.

$$\mathbf{A}: \quad \frac{1}{\sqrt{35}}[1 - 3i \quad 5]^{\mathsf{T}} \quad (\lambda = 9), \qquad \frac{1}{\sqrt{14}}[1 - 3i \quad -2]^{\mathsf{T}} \quad (\lambda = 2)$$

$$\mathbf{B}: \quad \frac{1}{\sqrt{30}}[1 - 2i \quad -5]^{\mathsf{T}} \quad (\lambda = -2i), \qquad \frac{1}{\sqrt{30}}[5 \quad 1 + 2i]^{\mathsf{T}} \quad (\lambda = 4i)$$

$$\mathbf{C}: \quad \frac{1}{\sqrt{2}}[1 \quad 1]^{\mathsf{T}} \quad (\lambda = \tfrac{1}{2}(i + \sqrt{3})), \qquad \frac{1}{\sqrt{2}}[1 \quad -1]^{\mathsf{T}} \quad (\lambda = \tfrac{1}{2}(i - \sqrt{3})).$$

■

Hermitian and Skew-Hermitian Forms

The concept of a quadratic form (Sec. 8.4) can be extended to complex. We call the numerator $\bar{\mathbf{x}}^{\mathsf{T}}\mathbf{A}\mathbf{x}$ in (1) a **form** in the components $x_1, \cdots, x_n$ of $\mathbf{x}$, which may now be complex. This form is again a sum of n^2 terms

$$
\begin{aligned}
\bar{\mathbf{x}}^{\mathsf{T}}\mathbf{A}\mathbf{x} &= \sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk}\bar{x}_j x_k \\
&= \quad a_{11}\bar{x}_1x_1 + \cdots + a_{1n}\bar{x}_1x_n \\
&\quad + a_{21}\bar{x}_2x_1 + \cdots + a_{2n}\bar{x}_2x_n \\
&\quad + \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
&\quad + a_{n1}\bar{x}_nx_1 + \cdots + a_{nn}\bar{x}_nx_n.
\end{aligned} \tag{7}
$$

$\mathbf{A}$ is called its **coefficient matrix**. The form is called a **Hermitian** or **skew-Hermitian form** if $\mathbf{A}$ is Hermitian or skew-Hermitian, respectively. *The value of a Hermitian form is real, and that of a skew-Hermitian form is pure imaginary or zero.* This can be seen directly from (2) and (3) and accounts for the importance of these forms in physics. Note that (2) and (3) are valid for any vectors because, in the proof of (2) and (3), we did not use that $\mathbf{x}$ is an eigenvector but only that $\bar{\mathbf{x}}^{\mathsf{T}}\mathbf{x}$ is real and not 0.

EXAMPLE 6 **Hermitian Form**

For $\mathbf{A}$ in Example 2 and, say, $\mathbf{x} = [1+i \quad 5i]^{\mathsf{T}}$ we get

$$\bar{\mathbf{x}}^{\mathsf{T}}\mathbf{A}\mathbf{x} = [1-i \quad -5i]\begin{bmatrix} 4 & 1-3i \\ 1+3i & 7 \end{bmatrix}\begin{bmatrix} 1+i \\ 5i \end{bmatrix} = [1-i \quad -5i]\begin{bmatrix} 4(1+i)+(1-3i)\cdot 5i \\ (1+3i)(1+i)+7\cdot 5i \end{bmatrix} = 223. \quad \blacksquare$$

Clearly, if $\mathbf{A}$ and $\mathbf{x}$ in (4) are real, then (7) reduces to a quadratic form, as discussed in the last section.

PROBLEM SET 8.5

1–6 EIGENVALUES AND VECTORS

Is the given matrix Hermitian? Skew-Hermitian? Unitary? Find its eigenvalues and eigenvectors.

1. $\begin{bmatrix} 6 & i \\ -i & 6 \end{bmatrix}$

2. $\begin{bmatrix} i & 1+i \\ -1+i & 0 \end{bmatrix}$

3. $\begin{bmatrix} \frac{1}{2} & i\sqrt{\frac{3}{4}} \\ i\sqrt{\frac{3}{4}} & \frac{1}{2} \end{bmatrix}$

4. $\begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}$

5. $\begin{bmatrix} i & 0 & 0 \\ 0 & 0 & i \\ 0 & i & 0 \end{bmatrix}$

6. $\begin{bmatrix} 0 & 2+2i & 0 \\ 2-2i & 0 & 2+2i \\ 0 & 2-2i & 0 \end{bmatrix}$

7. Pauli spin matrices. Find the eigenvalues and eigenvectors of the so-called *Pauli spin matrices* and show that $\mathbf{S}_x\mathbf{S}_y = i\mathbf{S}_z$, $\mathbf{S}_y\mathbf{S}_x = -i\mathbf{S}_z$, $\mathbf{S}_x^2 = \mathbf{S}_y^2 = \mathbf{S}_z^2 = \mathbf{I}$, where

$$\mathbf{S}_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad \mathbf{S}_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix},$$

$$\mathbf{S}_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

8. Eigenvectors. Find eigenvectors of $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ in Examples 2 and 3.

9–12 COMPLEX FORMS

Is the matrix $\mathbf{A}$ Hermitian or skew-Hermitian? Find $\overline{\mathbf{x}}^{\mathsf{T}}\mathbf{A}\mathbf{x}$. Show the details.

9. $\mathbf{A} = \begin{bmatrix} 4 & 3-2i \\ 3+2i & -4 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} -4i \\ 2+2i \end{bmatrix}$

10. $\mathbf{A} = \begin{bmatrix} i & -2+3i \\ 2+3i & 0 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} 2i \\ 8 \end{bmatrix}$

11. $\mathbf{A} = \begin{bmatrix} i & 1 & 2+i \\ -1 & 0 & 3i \\ -2+i & 3i & i \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} 1 \\ i \\ -i \end{bmatrix}$

12. $\mathbf{A} = \begin{bmatrix} 1 & i & 4 \\ -i & 3 & 0 \\ 4 & 0 & 2 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} 1 \\ i \\ -i \end{bmatrix}$

13–20 GENERAL PROBLEMS

13. Product. Show that $\overline{(\mathbf{ABC})}^{\mathsf{T}} = -\mathbf{C}^{-1}\mathbf{BA}$ for any $n \times n$ Hermitian $\mathbf{A}$, skew-Hermitian $\mathbf{B}$, and unitary $\mathbf{C}$.

14. Product. Show $\overline{(\mathbf{BA})}^{\mathsf{T}} = -\mathbf{AB}$ for $\mathbf{A}$ and $\mathbf{B}$ in Example 2. For any $n \times n$ Hermitian $\mathbf{A}$ and skew-Hermitian $\mathbf{B}$.

15. Decomposition. Show that any square matrix may be written as the sum of a Hermitian and a skew-Hermitian matrix. Give examples.

16. Unitary matrices. Prove that the product of two unitary $n \times n$ matrices and the inverse of a unitary matrix are unitary. Give examples.

17. Powers of unitary matrices in applications may sometimes be very simple. Show that $\mathbf{C}^{12} = \mathbf{I}$ in Example 2. Find further examples.

18. Normal matrix. This important concept denotes a matrix that commutes with its conjugate transpose, $\mathbf{A}\overline{\mathbf{A}}^{\mathsf{T}} = \overline{\mathbf{A}}^{\mathsf{T}}\mathbf{A}$. Prove that Hermitian, skew-Hermitian, and unitary matrices are normal. Give corresponding examples of your own.

19. Normality criterion. Prove that $\mathbf{A}$ is normal if and only if the Hermitian and skew-Hermitian matrices in Prob. 18 commute.

20. Find a simple matrix that is not normal. Find a normal matrix that is not Hermitian, skew-Hermitian, or unitary.

CHAPTER 8 REVIEW QUESTIONS AND PROBLEMS

1. In solving an eigenvalue problem, what is given and what is sought?
2. Give a few typical applications of eigenvalue problems.
3. Do there exist square matrices without eigenvalues?
4. Can a real matrix have complex eigenvalues? Can a complex matrix have real eigenvalues?
5. Does a 5×5 matrix always have a real eigenvalue?
6. What is algebraic multiplicity of an eigenvalue? Defect?
7. What is an eigenbasis? When does it exist? Why is it important?
8. When can we expect orthogonal eigenvectors?
9. State the definitions and main properties of the three classes of real matrices and of complex matrices that we have discussed.
10. What is diagonalization? Transformation to principal axes?

11–15 SPECTRUM

Find the eigenvalues. Find the eigenvectors.

11. $\begin{bmatrix} 2.5 & 0.5 \\ 0.5 & 2.5 \end{bmatrix}$ **12.** $\begin{bmatrix} -7 & 4 \\ -12 & 7 \end{bmatrix}$

13. $\begin{bmatrix} 8 & -1 \\ 5 & 2 \end{bmatrix}$

14. $\begin{bmatrix} 7 & 2 & -1 \\ 2 & 7 & 1 \\ -1 & 1 & 8.5 \end{bmatrix}$

15. $\begin{bmatrix} 0 & -3 & -6 \\ 3 & 0 & -6 \\ 6 & 6 & 0 \end{bmatrix}$

16–17 SIMILARITY

Verify that $\mathbf{A}$ and $\hat{\mathbf{A}} = \mathbf{p}^{-1}\mathbf{AP}$ have the same spectrum.

16. $\mathbf{A} = \begin{bmatrix} 19 & 12 \\ 12 & 1 \end{bmatrix}, \quad \mathbf{P} = \begin{bmatrix} 2 & 4 \\ 4 & 2 \end{bmatrix}$

17. $\mathbf{A} = \begin{bmatrix} 7 & -4 \\ 12 & -7 \end{bmatrix}, \quad \mathbf{P} = \begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix}$

18. $\mathbf{A} = \begin{bmatrix} -4 & 6 & 6 \\ 0 & 2 & 0 \\ -1 & 1 & 1 \end{bmatrix}, \quad \mathbf{P} = \begin{bmatrix} 1 & 8 & -7 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix}$

19–21 DIAGONALIZATION

Find an eigenbasis and diagonalize.

9. $\begin{bmatrix} -1.4 & 1.0 \\ -1.0 & 1.1 \end{bmatrix}$

20. $\begin{bmatrix} 72 & -56 \\ -56 & 513 \end{bmatrix}$

21. $\begin{bmatrix} -12 & 22 & 6 \\ 8 & 2 & 6 \\ -8 & 20 & 16 \end{bmatrix}$

22–25 CONIC SECTIONS. PRINCIPAL AXES

Transform to canonical form (to principal axes). Express $[x_1 \quad x_2]^\mathsf{T}$ in terms of the new variables $[y_1 \quad y_2]^\mathsf{T}$.

22. $9x_1^2 - 6x_1x_2 + 17x_2^2 = 36$

23. $4x_1^2 + 24x_1x_2 - 14x_2^2 = 20$

24. $5x_1^2 + 24x_1x_2 - 5x_2^2 = 0$

25. $3.7x_1^2 + 3.2x_1x_2 + 1.3x_2^2 = 4.5$

SUMMARY OF CHAPTER 8

Linear Algebra: Matrix Eigenvalue Problems

The practical importance of matrix eigenvalue problems can hardly be overrated. The problems are defined by the vector equation

$$\mathbf{Ax} = \lambda\mathbf{x}. \tag{1}$$

A is a given square matrix. All matrices in this chapter are ***square***. λ is a scalar. To *solve* the problem (1) means to determine values of λ, called **eigenvalues** (or **characteristic values**) of **A**, such that (1) has a nontrivial solution **x** (that is, $\mathbf{x} \neq \mathbf{0}$), called an **eigenvector** of **A** corresponding to that λ. An $n \times n$ matrix has at least one and at most n numerically different eigenvalues. These are the solutions of the **characteristic equation** (Sec. 8.1)

$$D(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix} = 0. \tag{2}$$

$D(\lambda)$ is called the **characteristic determinant** of **A**. By expanding it we get the **characteristic polynomial** of **A**, which is of degree n in λ. Some typical applications are shown in Sec. 8.2.

Section 8.3 is devoted to eigenvalue problems for **symmetric** ($\mathbf{A}^\mathsf{T} = \mathbf{A}$), **skew-symmetric** ($\mathbf{A}^\mathsf{T} = -\mathbf{A}$), and **orthogonal matrices** ($\mathbf{A}^\mathsf{T} = \mathbf{A}^{-1}$). Section 8.4 concerns the diagonalization of matrices and the transformation of quadratic forms to principal axes and its relation to eigenvalues.

Section 8.5 extends Sec. 8.3 to the complex analogs of those real matrices, called **Hermitian** ($\bar{\mathbf{A}}^\mathsf{T} = \mathbf{A}$), **skew-Hermitian** ($\bar{\mathbf{A}}^\mathsf{T} = -\mathbf{A}$), and **unitary matrices** ($\bar{\mathbf{A}}^\mathsf{T} = \mathbf{A}^{-1}$). All the eigenvalues of a Hermitian matrix (and a symmetric one) are real. For a skew-Hermitian (and a skew-symmetric) matrix they are pure imaginary or zero. For a unitary (and an orthogonal) matrix they have absolute value 1.

CHAPTER 9

Vector Differential Calculus. Grad, Div, Curl

Engineering, physics, and computer sciences, in general, but particularly solid mechanics, aerodynamics, aeronautics, fluid flow, heat flow, electrostatics, quantum physics, laser technology, robotics as well as other areas have applications that require an understanding of **vector calculus**. This field encompasses vector differential calculus and vector integral calculus. Indeed, the engineer, physicist, and mathematician need a good grounding in these areas as provided by the carefully chosen material of Chaps. 9 and 10.

Forces, velocities, and various other quantities may be thought of as vectors. Vectors appear frequently in the applications above and also in the biological and social sciences, so it is natural that problems are modeled in **3-space**. This is the space of three dimensions with the usual measurement of distance, as given by the Pythagorean theorem. Within that realm, **2-space** (the plane) is a special case. Working in 3-space requires that we extend the common differential calculus to vector differential calculus, that is, the calculus that deals with vector functions and vector fields and is explained in this chapter.

Chapter 9 is arranged in three groups of sections. Sections 9.1–9.3 extend the basic algebraic operations of vectors into 3-space. These operations include the inner product and the cross product. Sections 9.4 and 9.5 form the heart of vector differential calculus. Finally, Secs. 9.7–9.9 discuss three physically important concepts related to scalar and vector fields: gradient (Sec. 9.7), divergence (Sec. 9.8), and curl (Sec. 9.9). They are expressed in Cartesian coordinates in this chapter and, if desired, expressed in ***curvilinear coordinates*** in a short section in App. A3.4.

We shall keep this chapter ***independent of Chaps. 7 and 8***. Our present approach is in harmony with Chap. 7, with the restriction to two and three dimensions providing for a richer theory with basic physical, engineering, and geometric applications.

Prerequisite: Elementary use of second- and third-order determinants in Sec. 9.3.
Sections that may be omitted in a shorter course: 9.5, 9.6.
References and Answers to Problems: App. 1 Part B, App. 2.

9.1 Vectors in 2-Space and 3-Space

In engineering, physics, mathematics, and other areas we encounter two kinds of quantities. They are scalars and vectors.

A **scalar** is a quantity that is determined by its magnitude. It takes on a numerical value, i.e., a number. Examples of scalars are time, temperature, length, distance, speed, density, energy, and voltage.

In contrast, a **vector** is a quantity that has both magnitude and direction. We can say that a vector is an ***arrow*** or a ***directed line segment***. For example, a velocity vector has length or magnitude, which is speed, and direction, which indicates the direction of motion. Typical examples of vectors are displacement, velocity, and force, see Fig. 164 as an illustration.

More formally, we have the following. We denote vectors by lowercase boldface letters **a**, **b**, **v**, etc. In handwriting you may use arrows, for instance, $\vec{a}$ (in place of **a**), $\vec{b}$, etc.

A vector (arrow) has a tail, called its **initial point**, and a tip, called its **terminal point**. This is motivated in the **translation** (displacement without rotation) of the triangle in Fig. 165, where the initial point P of the vector **a** is the original position of a point, and the terminal point Q is the terminal position of that point, its position *after* the translation. The length of the arrow equals the distance between P and Q. This is called the **length** (or *magnitude*) of the vector **a** and is denoted by $|\mathbf{a}|$. Another name for *length* is **norm** (or *Euclidean norm*).

A vector of length 1 is called a **unit vector**.

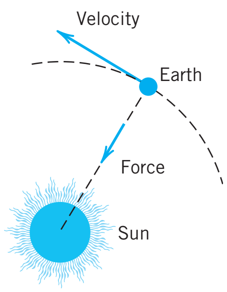


Fig. 164. Force and velocity

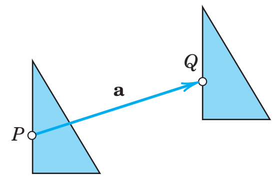


Fig. 165. Translation

Of course, we would like to calculate with vectors. For instance, we want to find the resultant of forces or compare parallel forces of different magnitude. This motivates our next ideas: to define *components* of a vector, and then the two basic algebraic operations of *vector addition* and *scalar multiplication.*

For this we must first define *equality of vectors* in a way that is practical in connection with forces and other applications.

DEFINITION

Equality of Vectors

Two vectors **a** and **b** are equal, written **a** = **b**, if they have the same length and the same direction [as explained in Fig. 166; in particular, note (B)]. Hence a vector can be arbitrarily translated; that is, its initial point can be chosen arbitrarily.

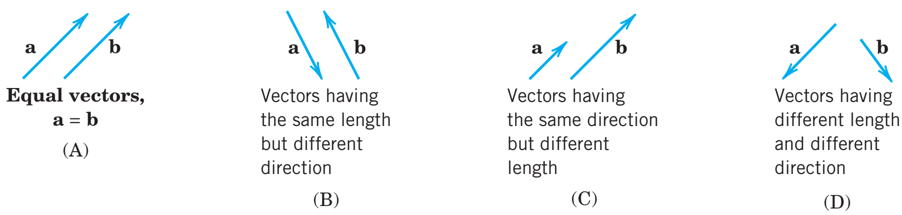


Fig. 166. (A) Equal vectors. (B)–(D) Different vectors

Components of a Vector

We choose an xyz **Cartesian coordinate system**[1] in space (Fig. 167), that is, a usual rectangular coordinate system with the same scale of measurement on the three mutually perpendicular coordinate axes. Let **a** be a given vector with initial point P: (x_1, y_1, z_1) and terminal point Q: (x_2, y_2, z_2). Then the three coordinate differences

$$a_1 = x_2 - x_1, \qquad a_2 = y_2 - y_1, \qquad a_3 = z_2 - z_1 \tag{1}$$

are called the **components** of the vector **a** with respect to that coordinate system, and we write simply $\mathbf{a} = [a_1, a_2, a_3]$. See Fig. 168.

The **length** $|\mathbf{a}|$ of **a** can now readily be expressed in terms of components because from (1) and the Pythagorean theorem we have

$$|\mathbf{a}| = \sqrt{a_1^2 + a_2^2 + a_3^2}. \tag{2}$$

EXAMPLE 1 Components and Length of a Vector

The vector **a** with initial point P: (4, 0, 2) and terminal point Q: (6, −1, 2) has the components

$$a_1 = 6 - 4 = 2, \qquad a_2 = -1 - 0 = -1, \qquad a_3 = 2 - 2 = 0.$$

Hence $\mathbf{a} = [2, -1, 0]$. (Can you sketch **a**, as in Fig. 168?) Equation (2) gives the length

$$|\mathbf{a}| = \sqrt{2^2 + (-1)^2 + 0^2} = \sqrt{5}.$$

If we choose (−1, 5, 8) as the initial point of **a**, the corresponding terminal point is (1, 4, 8).

If we choose the origin (0, 0, 0) as the initial point of **a**, the corresponding terminal point is (2, −1, 0); its coordinates equal the components of **a**. This suggests that we can determine each point in space by a vector, called the *position vector* of the point, as follows. ■

A Cartesian coordinate system being given, the **position vector r** of a point A: (x, y, z) is the vector with the origin (0, 0, 0) as the initial point and A as the terminal point (see Fig. 169). Thus in components, $\mathbf{r} = [x, y, z]$. This can be seen directly from (1) with $x_1 = y_1 = z_1 = 0$.

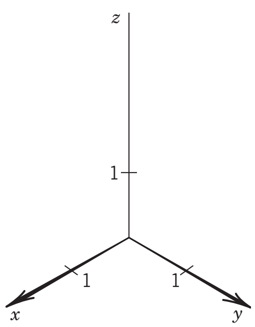


Fig. 167. Cartesian coordinate system

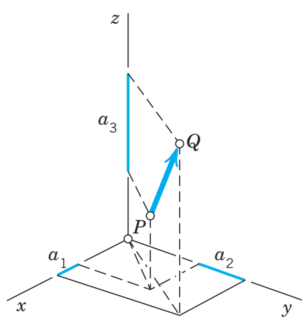


Fig. 168. Components of a vector

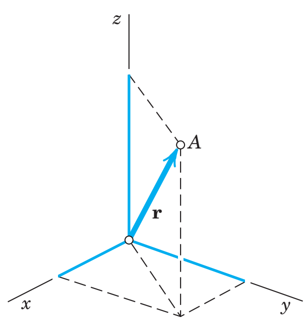


Fig. 169. Position vector **r** of a point A: (x, y, z)

[1]Named after the French philosopher and mathematician RENATUS CARTESIUS, latinized for RENÉ DESCARTES (1596–1650), who invented analytic geometry. His basic work *Géométrie* appeared in 1637, as an appendix to his *Discours de la méthode.*

Furthermore, if we translate a vector **a**, with initial point P and terminal point Q, then corresponding coordinates of P and Q change by the same amount, so that the differences in (1) remain unchanged. This proves

THEOREM 1

Vectors as Ordered Triples of Real Numbers

A fixed Cartesian coordinate system being given, each vector is uniquely determined by its ordered triple of corresponding components. Conversely, to each ordered triple of real numbers (a_1, a_2, a_3) *there corresponds precisely one vector* $\mathbf{a} = [a_1, a_2, a_3]$, *with* $(0, 0, 0)$ *corresponding to the* **zero vector 0**, *which has length* 0 *and no direction.*

Hence a vector equation $\mathbf{a} = \mathbf{b}$ *is equivalent to the three equations* $a_1 = b_1$, $a_2 = b_2$, $a_3 = b_3$ *for the components.*

We now see that from our "geometric" definition of a vector as an arrow we have arrived at an "algebraic" characterization of a vector by Theorem 1. We could have started from the latter and reversed our process. This shows that the two approaches are equivalent.

Vector Addition, Scalar Multiplication

Calculations with vectors are very useful and are almost as simple as the arithmetic for real numbers. Vector arithmetic follows almost naturally from applications. We first define how to add vectors and later on how to multiply a vector by a number.

DEFINITION

Addition of Vectors

The **sum** $\mathbf{a} + \mathbf{b}$ of two vectors $\mathbf{a} = [a_1, a_2, a_3]$ and $\mathbf{b} = [b_1, b_2, b_3]$ is obtained by adding the corresponding components,

$$\mathbf{a} + \mathbf{b} = [a_1 + b_1, \quad a_2 + b_2, \quad a_3 + b_3]. \tag{3}$$

Geometrically, place the vectors as in Fig. 170 (the initial point of **b** at the terminal point of **a**); then $\mathbf{a} + \mathbf{b}$ is the vector drawn from the initial point of **a** to the terminal point of **b**.

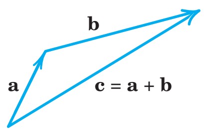


Fig. 170. Vector addition

For forces, this addition is the parallelogram law by which we obtain the **resultant** of two forces in mechanics. See Fig. 171.

Figure 172 shows (for the plane) that the "algebraic" way and the "geometric way" of vector addition give the same vector.

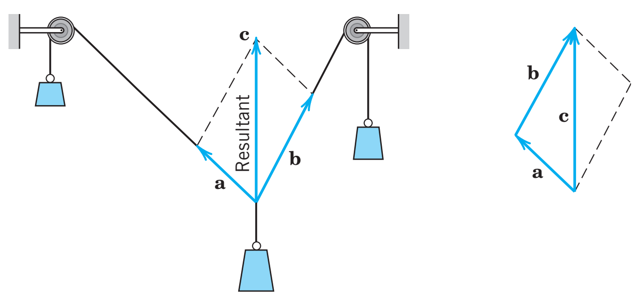


Fig. 171. Resultant of two forces (**parallelogram law**)

Basic Properties of Vector Addition. Familiar laws for real numbers give immediately

$$
\begin{aligned}
&\text{(a)} \quad && \mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a} && \textit{(Commutativity)} \\
&\text{(b)} \quad && (\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w}) && \textit{(Associativity)} \\
&\text{(c)} \quad && \mathbf{a} + \mathbf{0} = \mathbf{0} + \mathbf{a} = \mathbf{a} && \\
&\text{(d)} \quad && \mathbf{a} + (-\mathbf{a}) = \mathbf{0}. &&
\end{aligned} \tag{4}
$$

Properties (a) and (b) are verified geometrically in Figs. 173 and 174. Furthermore, $-\mathbf{a}$ denotes the vector having the length $|\mathbf{a}|$ and the direction opposite to that of $\mathbf{a}$.

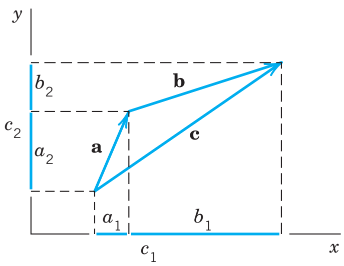


Fig. 172. Vector addition

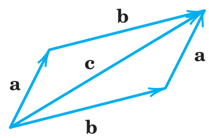


Fig. 173. Cummutativity of vector addition

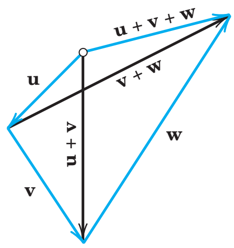


Fig. 174. Associativity of vector addition

In (4b) we may simply write $\mathbf{u} + \mathbf{v} + \mathbf{w}$, and similarly for sums of more than three vectors. Instead of $\mathbf{a} + \mathbf{a}$ we also write $2\mathbf{a}$, and so on. This (and the notation $-\mathbf{a}$ used just before) motivates defining the second algebraic operation for vectors as follows.

DEFINITION

Scalar Multiplication (Multiplication by a Number)

The product $c\mathbf{a}$ of any vector $\mathbf{a} = [a_1, a_2, a_3]$ and any scalar c (real number c) is the vector obtained by multiplying each component of $\mathbf{a}$ by c,

$$c\mathbf{a} = [ca_1, ca_2, ca_3]. \tag{5}$$

Geometrically, if $\mathbf{a} \neq \mathbf{0}$, then $c\mathbf{a}$ with $c > 0$ has the direction of $\mathbf{a}$ and with $c < 0$ the direction opposite to $\mathbf{a}$. In any case, the length of $c\mathbf{a}$ is $|c\mathbf{a}| = |c||\mathbf{a}|$, and $c\mathbf{a} = \mathbf{0}$ if $\mathbf{a} = \mathbf{0}$ or $c = 0$ (or both). (See Fig. 175.)

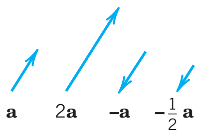


Fig. 175. Scalar multiplication [multiplication of vectors by scalars (numbers)]

Basic Properties of Scalar Multiplication. From the definitions we obtain directly

$$
\begin{aligned}
&\text{(a)} \quad && c(\mathbf{a} + \mathbf{b}) = c\mathbf{a} + c\mathbf{b} && \\
&\text{(b)} \quad && (c + k)\mathbf{a} = c\mathbf{a} + k\mathbf{a} && \\
&\text{(c)} \quad && c(k\mathbf{a}) = (ck)\mathbf{a} && \text{(written } ck\mathbf{a}\text{)} \\
&\text{(d)} \quad && 1\mathbf{a} = \mathbf{a}. &&
\end{aligned} \tag{6}
$$

You may prove that (4) and (6) imply for any vector **a**

$$\text{(7)}\qquad \begin{aligned} &\text{(a)} \quad 0\mathbf{a} = \mathbf{0} \\ &\text{(b)} \quad (-1)\mathbf{a} = -\mathbf{a}. \end{aligned}$$

Instead of $\mathbf{b} + (-\mathbf{a})$ we simply write $\mathbf{b} - \mathbf{a}$ (Fig. 176).

EXAMPLE 2 Vector Addition. Multiplication by Scalars

With respect to a given coordinate system, let

$$\mathbf{a} = [4, 0, 1] \qquad \text{and} \qquad \mathbf{b} = [2, -5, \tfrac{1}{3}].$$

Then $-\mathbf{a} = [-4, 0, -1]$, $7\mathbf{a} = [28, 0, 7]$, $\mathbf{a} + \mathbf{b} = [6, -5, \tfrac{4}{3}]$, and

$$2(\mathbf{a} - \mathbf{b}) = 2[2, 5, \tfrac{2}{3}] = [4, 10, \tfrac{4}{3}] = 2\mathbf{a} - 2\mathbf{b}.$$ ■

Unit Vectors i, j, k. Besides $\mathbf{a} = [a_1, a_2, a_3]$ another popular way of writing vectors is

$$\text{(8)}\qquad \mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}.$$

In this representation, **i**, **j**, **k** are the unit vectors in the positive directions of the axes of a Cartesian coordinate system (Fig. 177). Hence, in components,

$$\text{(9)}\qquad \mathbf{i} = [1, 0, 0], \qquad \mathbf{j} = [0, 1, 0], \qquad \mathbf{k} = [0, 0, 1]$$

and the right side of (8) is a sum of three vectors parallel to the three axes.

EXAMPLE 3 ijk Notation for Vectors

In Example 2 we have $\mathbf{a} = 4\mathbf{i} + \mathbf{k}$, $\mathbf{b} = 2\mathbf{i} - 5\mathbf{j} + \tfrac{1}{3}\mathbf{k}$, and so on. ■

All the vectors $\mathbf{a} = [a_1, a_2, a_3] = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$ (with real numbers as components) form the **real vector space** R^3 with the two *algebraic operations* of vector addition and scalar multiplication as just defined. R^3 has **dimension** 3. The triple of vectors **i**, **j**, **k** is called a **standard basis** of R^3. Given a Cartesian coordinate system, the representation (8) of a given vector is unique.

Vector space R^3 is a model of a general vector space, as discussed in Sec. 7.9, but is not needed in this chapter.

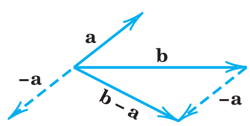


Fig. 176. Difference of vectors

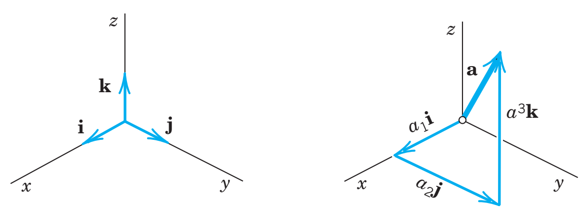


Fig. 177. The unit vectors **i**, **j**, **k** and the representation (8)

PROBLEM SET 9.1

1–5 COMPONENTS AND LENGTH

Find the components of the vector $\mathbf{v}$ with initial point P and terminal point Q. Find $|\mathbf{v}|$. Sketch $|\mathbf{v}|$. Find the unit vector $\mathbf{u}$ in the direction of $\mathbf{v}$.

1. P: (1, 1, 0), Q: (6, 2, 0)
2. P: (1, 1, 1), Q: (2, 2, 0)
3. P: (−3.0, 4,0, −0.5), Q: (5.5, 0, 1.2)
4. P: (1, 4, 2), Q: (−1, −4, −2)
5. P: (0, 0, 0), Q: (2, 1, −2)

6–10

Find the terminal point Q of the vector $\mathbf{v}$ with components as given and initial point P. Find $|\mathbf{v}|$.

6. 4, 0, 0; P: (0, 2, 13)
7. $\frac{1}{2}, 3, -\frac{1}{4}$; P: $(\frac{7}{2}, -3, \frac{3}{4})$
8. 13.1, 0.8, −2.0; P: (0, 0, 0)
9. 6, 1, −4; P: (−6, −1, −4)
10. 0, −3, 3; P: (0, 3, −3)

11–18 ADDITION, SCALAR MULTIPLICATION

Let $\mathbf{a} = [3, 2, 0] = 3\mathbf{i} + 2\mathbf{j}$; $\mathbf{b} = [-4, 6, 0] = 4\mathbf{i} + 6\mathbf{j}$, $\mathbf{c} = [5, -1, 8] = 5\mathbf{i} - \mathbf{j} + 8\mathbf{k}$, $\mathbf{d} = [0, 0, 4] = 4\mathbf{k}$.
Find:

11. $2\mathbf{a}$, $\frac{1}{2}\mathbf{a}$, $-\mathbf{a}$
12. $(\mathbf{a} + \mathbf{b}) + \mathbf{c}$, $\mathbf{a} + (\mathbf{b} + \mathbf{c})$
13. $\mathbf{b} + \mathbf{c}$, $\mathbf{c} + \mathbf{b}$
14. $3\mathbf{c} - 6\mathbf{d}$, $3(\mathbf{c} - 2\mathbf{d})$
15. $7(\mathbf{c} - \mathbf{b})$, $7\mathbf{c} - 7\mathbf{b}$
16. $\frac{9}{2}\mathbf{a} - 3\mathbf{c}$, $9(\frac{1}{2}\mathbf{a} - \frac{1}{3}\mathbf{c})$
17. $(7 - 3)\mathbf{a}$, $7\mathbf{a} - 3\mathbf{a}$
18. $4\mathbf{a} + 3\mathbf{b}$, $-4\mathbf{a} - 3\mathbf{b}$
19. What laws do Probs. 12–16 illustrate?
20. Prove Eqs. (4) and (6).

21–25 FORCES, RESULTANT

Find the resultant in terms of components and its magnitude.

21. $\mathbf{p} = [2, 3, 0]$, $\mathbf{q} = [0, 6, 1]$, $\mathbf{u} = [2, 0, -4]$
22. $\mathbf{p} = [1, -2, 3]$, $\mathbf{q} = [3, 21, -16]$, $\mathbf{u} = [-4, -19, 13]$
23. $\mathbf{u} = [8, -1, 0]$, $\mathbf{v} = [\frac{1}{2}, 0, \frac{4}{3}]$, $\mathbf{w} = [-\frac{17}{2}, 1, \frac{11}{3}]$
24. $\mathbf{p} = [-1, 2, -3]$, $\mathbf{q} = [1, 1, 1]$, $\mathbf{u} = [1, -2, 2]$
25. $\mathbf{u} = [3, 1, -6]$, $\mathbf{v} = [0, 2, 5]$, $\mathbf{w} = [3, -1, -13]$

26–37 FORCES, VELOCITIES

26. Equilibrium. Find $\mathbf{v}$ such that $\mathbf{p}$, $\mathbf{q}$, $\mathbf{u}$ in Prob. 21 and $\mathbf{v}$ are in equilibrium.

27. Find $\mathbf{p}$ such that $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ in Prob. 23 and $\mathbf{p}$ are in equilibrium.

28. Unit vector. Find the unit vector in the direction of the resultant in Prob. 24.

29. Restricted resultant. Find all $\mathbf{v}$ such that the resultant of $\mathbf{v}$, $\mathbf{p}$, $\mathbf{q}$, $\mathbf{u}$ with $\mathbf{p}$, $\mathbf{q}$, $\mathbf{u}$ as in Prob. 21 is parallel to the xy-plane.

30. Find $\mathbf{v}$ such that the resultant of $\mathbf{p}$, $\mathbf{q}$, $\mathbf{u}$, $\mathbf{v}$ with $\mathbf{p}$, $\mathbf{q}$, $\mathbf{u}$ as in Prob. 24 has no components in x- and y-directions.

31. For what k is the resultant of [2, 0, −7], [1, 2, −3], and [0, 3, k] parallel to the xy-plane?

32. If $|\mathbf{p}| = 6$ and $|\mathbf{q}| = 4$, what can you say about the magnitude and direction of the resultant? Can you think of an application to robotics?

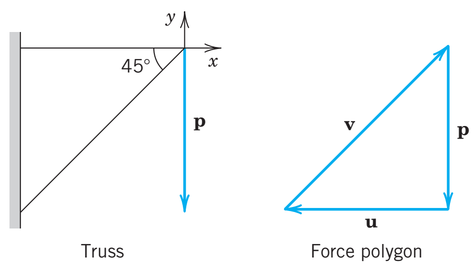
33. Same question as in Prob. 32 if $|\mathbf{p}| = 9$, $|\mathbf{q}| = 6$, $|\mathbf{u}| = 3$.

34. Relative velocity. If airplanes A and B are moving southwest with speed $|\mathbf{v}_A| = 550$ mph, and northwest with speed $|\mathbf{v}_B| = 450$ mph, respectively, what is the relative velocity $\mathbf{v} = \mathbf{v}_B - \mathbf{v}_A$ of B with respect to A?

35. Same question as in Prob. 34 for two ships moving northeast with speed $|\mathbf{v}_A| = 22$ knots and west with speed $|\mathbf{v}_B| = 19$ knots.

36. Reflection. If a ray of light is reflected once in each of two mutually perpendicular mirrors, what can you say about the reflected ray?

37. Force polygon. Truss. Find the forces in the system of two rods (*truss*) in the figure, where $|\mathbf{p}| = 1000$ nt. *Hint.* Forces in equilibrium form a polygon, the *force polygon*.



Problem 37

38. TEAM PROJECT. Geometric Applications. To increase your skill in dealing with vectors, use vectors to prove the following (see the figures).

(a) The diagonals of a parallelogram bisect each other.

(b) The line through the midpoints of adjacent sides of a parallelogram bisects one of the diagonals in the ratio 1 : 3.

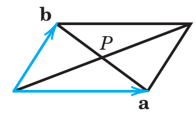
(c) Obtain (b) from (a).

(d) The three medians of a triangle (the segments from a vertex to the midpoint of the opposite side) meet at a single point, which divides the medians in the ratio 2 : 1.

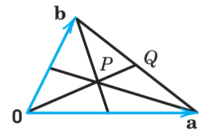
(e) The quadrilateral whose vertices are the mid-points of the sides of an arbitrary quadrilateral is a parallelogram.

(f) The four space diagonals of a parallelepiped meet and bisect each other.

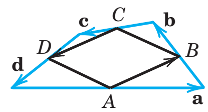
(g) The sum of the vectors drawn from the center of a regular polygon to its vertices is the zero vector.



Team Project 38(a)



Team Project 38(d)



Team Project 38(e)

9.2 Inner Product (Dot Product)

Orthogonality

The inner product or dot product can be motivated by calculating work done by a constant force, determining components of forces, or other applications. It involves the length of vectors and the angle between them. The inner product is a kind of multiplication of two vectors, defined in such a way that the outcome is a scalar. Indeed, another term for inner product is scalar product, a term we shall not use here. The definition of the inner product is as follows.

DEFINITION

Inner Product (Dot Product) of Vectors

The **inner product** or **dot product** $\mathbf{a} \cdot \mathbf{b}$ (read "**a** dot **b**") of two vectors **a** and **b** is the product of their lengths times the cosine of their angle (see Fig. 178),

$$(1) \qquad \begin{aligned} \mathbf{a} \cdot \mathbf{b} &= |\mathbf{a}||\mathbf{b}| \cos \gamma \qquad && \text{if} \quad \mathbf{a} \neq \mathbf{0}, \mathbf{b} \neq \mathbf{0} \\ \mathbf{a} \cdot \mathbf{b} &= 0 && \text{if} \quad \mathbf{a} = \mathbf{0} \text{ or } \mathbf{b} = \mathbf{0}. \end{aligned}$$

The angle γ, $0 \leqq \gamma \leqq \pi$, between **a** and **b** is measured when the initial points of the vectors coincide, as in Fig. 178. In components, $\mathbf{a} = [a_1, a_2, a_3]$, $\mathbf{b} = [b_1, b_2, b_3]$, and

$$(2) \qquad \mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2 + a_3 b_3.$$

The second line in (1) is needed because γ is undefined when $\mathbf{a} = \mathbf{0}$ or $\mathbf{b} = \mathbf{0}$. The derivation of (2) from (1) is shown below.

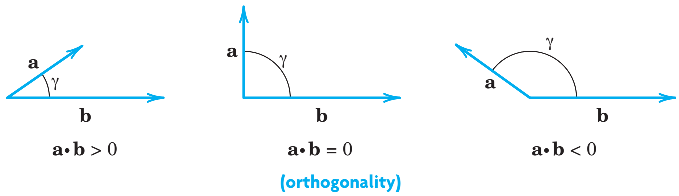


Fig. 178. Angle between vectors and value of inner product

Orthogonality. Since the cosine in (1) may be positive, 0, or negative, so may be the inner product (Fig. 178). The case that the inner product is zero is of particular practical interest and suggests the following concept.

A vector $\mathbf{a}$ is called **orthogonal** to a vector $\mathbf{b}$ if $\mathbf{a} \cdot \mathbf{b} = 0$. Then $\mathbf{b}$ is also orthogonal to $\mathbf{a}$, and we call $\mathbf{a}$ and $\mathbf{b}$ **orthogonal vectors**. Clearly, this happens for nonzero vectors if and only if $\cos \gamma = 0$; thus $\gamma = \pi/2$ (90°). This proves the important

THEOREM 1

Orthogonality Criterion

The inner product of two nonzero vectors is 0 *if and only if these vectors are perpendicular.*

Length and Angle. Equation (1) with $\mathbf{b} = \mathbf{a}$ gives $\mathbf{a} \cdot \mathbf{a} = |\mathbf{a}|^2$. Hence

$$|\mathbf{a}| = \sqrt{\mathbf{a} \cdot \mathbf{a}}. \tag{3}$$

From (3) and (1) we obtain for the angle γ between two nonzero vectors

$$\cos \gamma = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}||\mathbf{b}|} = \frac{\mathbf{a} \cdot \mathbf{b}}{\sqrt{\mathbf{a} \cdot \mathbf{a}}\sqrt{\mathbf{b} \cdot \mathbf{b}}}. \tag{4}$$

EXAMPLE 1

Inner Product. Angle Between Vectors

Find the inner product and the lengths of $\mathbf{a} = [1, 2, 0]$ and $\mathbf{b} = [3, -2, 1]$ as well as the angle between these vectors.

Solution. $\mathbf{a} \cdot \mathbf{b} = 1 \cdot 3 + 2 \cdot (-2) + 0 \cdot 1 = -1$, $|\mathbf{a}| = \sqrt{\mathbf{a} \cdot \mathbf{a}} = \sqrt{5}$, $|\mathbf{b}| = \sqrt{\mathbf{b} \cdot \mathbf{b}} = \sqrt{14}$, and (4) gives the angle

$$\gamma = \arccos \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}||\mathbf{b}|} = \arccos(-0.11952) = 1.69061 = 96.865°.$$ ■

From the definition we see that the inner product has the following properties. For any vectors **a**, **b**, **c** and scalars q_1, q_2,

$$\textbf{(5)}\quad \begin{array}{lll} \text{(a)} & (q_1\mathbf{a} + q_2\mathbf{b}) \cdot \mathbf{c} = q_1\mathbf{a} \cdot \mathbf{c} + q_1\mathbf{b} \cdot \mathbf{c} & (\textit{Linearity}) \\ \text{(b)} & \mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a} & (\textit{Symmetry}) \\ \text{(c)} & \left.\begin{array}{c} \mathbf{a} \cdot \mathbf{a} \geqq 0 \\ \mathbf{a} \cdot \mathbf{a} = 0 \quad \text{if and only if} \quad \mathbf{a} = \mathbf{0} \end{array}\right\} & (\textit{Positive-definiteness}). \end{array}$$

Hence *dot multiplication is commutative* as shown by (5b). Furthermore, it is *distributive with respect to vector addition.* This follows from (5a) with $q_1 = 1$ and $q_2 = 1$:

$$\textbf{(5a*)}\qquad (\mathbf{a} + \mathbf{b}) \cdot \mathbf{c} = \mathbf{a} \cdot \mathbf{c} + \mathbf{b} \cdot \mathbf{c} \qquad (\textit{Distributivity}).$$

Furthermore, from (1) and $|\cos \gamma| \leqq 1$ we see that

$$\textbf{(6)}\qquad |\mathbf{a} \cdot \mathbf{b}| \leqq |\mathbf{a}||\mathbf{b}| \qquad (\textit{Cauchy–Schwarz inequality}).$$

Using this and (3), you may prove (see Prob. 16)

$$\textbf{(7)}\qquad |\mathbf{a} + \mathbf{b}| \leqq |\mathbf{a}| + |\mathbf{b}| \qquad (\textbf{Triangle inequality}).$$

Geometrically, (7) with $<$ says that one side of a triangle must be shorter than the other two sides together; this motivates the name of (7).

A simple direct calculation with inner products shows that

$$(8)\qquad |\mathbf{a} + \mathbf{b}|^2 + |\mathbf{a} - \mathbf{b}|^2 = 2(|\mathbf{a}|^2 + |\mathbf{b}|^2) \quad (\textit{Parallelogram equality}).$$

Equations (6)–(8) play a basic role in so-called *Hilbert spaces*, which are abstract inner product spaces. Hilbert spaces form the basis of quantum mechanics, for details see [GenRef7] listed in App. 1.

Derivation of (2) from (1). We write $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$ and $\mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$, as in (8) of Sec. 9.1. If we substitute this into $\mathbf{a} \cdot \mathbf{b}$ and use (5a*), we first have a sum of $3 \times 3 = 9$ products

$$\mathbf{a} \cdot \mathbf{b} = a_1b_1\mathbf{i} \cdot \mathbf{i} + a_1b_2\mathbf{i} \cdot \mathbf{j} + \cdots + a_3b_3\mathbf{k} \cdot \mathbf{k}.$$

Now **i**, **j**, **k** are unit vectors, so that $\mathbf{i} \cdot \mathbf{i} = \mathbf{j} \cdot \mathbf{j} = \mathbf{k} \cdot \mathbf{k} = 1$ by (3). Since the coordinate axes are perpendicular, so are **i**, **j**, **k**, and Theorem 1 implies that the other six of those nine products are 0, namely, $\mathbf{i} \cdot \mathbf{j} = \mathbf{j} \cdot \mathbf{i} = \mathbf{j} \cdot \mathbf{k} = \mathbf{k} \cdot \mathbf{j} = \mathbf{k} \cdot \mathbf{i} = \mathbf{i} \cdot \mathbf{k} = 0$. But this reduces our sum for $\mathbf{a} \cdot \mathbf{b}$ to (2). ■

Applications of Inner Products

Typical applications of inner products are shown in the following examples and in Problem Set 9.2.

EXAMPLE 2 Work Done by a Force Expressed as an Inner Product

This is a major application. It concerns a body on which a *constant* force $\mathbf{p}$ acts. (For a *variable* force, see Sec. 10.1.) Let the body be given a displacement $\mathbf{d}$. Then the work done by $\mathbf{p}$ in the displacement is defined as

$$W = |\mathbf{p}||\mathbf{d}| \cos \alpha = \mathbf{p} \cdot \mathbf{d}, \tag{9}$$

that is, magnitude $|\mathbf{p}|$ of the force times length $|\mathbf{d}|$ of the displacement times the cosine of the angle α between $\mathbf{p}$ and $\mathbf{d}$ (Fig. 179). If $\alpha < 90°$, as in Fig. 179, then $W > 0$. If $\mathbf{p}$ and $\mathbf{d}$ are orthogonal, then the work is zero (why?). If $\alpha > 90°$, then $W < 0$, which means that in the displacement one has to do work against the force. For example, think of swimming across a river at some angle α against the current.

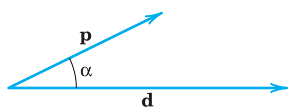


Fig. 179. Work done by a force

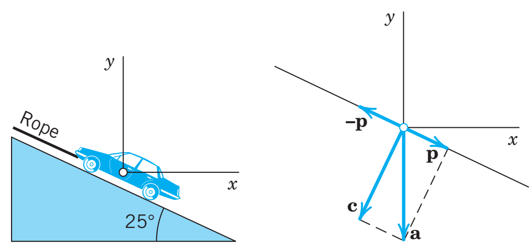


Fig. 180. Example 3

EXAMPLE 3 Component of a Force in a Given Direction

What force in the rope in Fig. 180 will hold a car of 5000 lb in equilibrium if the ramp makes an angle of 25° with the horizontal?

Solution. Introducing coordinates as shown, the weight is $\mathbf{a} = [0, -5000]$ because this force points downward, in the negative y-direction. We have to represent $\mathbf{a}$ as a sum (resultant) of two forces, $\mathbf{a} = \mathbf{c} + \mathbf{p}$, where $\mathbf{c}$ is the force the car exerts on the ramp, which is of no interest to us, and $\mathbf{p}$ is parallel to the rope. A vector in the direction of the rope is (see Fig. 180)

$$\mathbf{b} = [-1, \tan 25°] = [-1, 0.46631], \qquad \text{thus} \qquad |\mathbf{b}| = 1.10338,$$

The direction of the unit vector $\mathbf{u}$ is opposite to the direction of the rope so that

$$\mathbf{u} = -\frac{1}{|\mathbf{b}|}\mathbf{b} = [0.90631, -0.42262].$$

Since $|\mathbf{u}| = 1$ and $\cos \gamma > 0$, we see that we can write our result as

$$|\mathbf{p}| = (|\mathbf{a}| \cos \gamma)|\mathbf{u}| = \mathbf{a} \cdot \mathbf{u} = -\frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{b}|} = \frac{5000 \cdot 0.46631}{1.10338} = 2113 \text{ [lb]}.$$

We can also note that $\gamma = 90° - 25° = 65°$ is the angle between $\mathbf{a}$ and $\mathbf{p}$ so that

$$|\mathbf{p}| = |\mathbf{a}| \cos \gamma = 5000 \cos 65° = 2113 \text{ [lb]}.$$

Answer: About 2100 lb. ■

Example 3 is typical of applications that deal with the **component** or **projection** *of a vector* $\mathbf{a}$ *in the direction of a vector* $\mathbf{b}\,(\neq \mathbf{0})$. If we denote by p the length of the orthogonal projection of $\mathbf{a}$ on a straight line l parallel to $\mathbf{b}$ as shown in Fig. 181, then

$$p = |\mathbf{a}| \cos \gamma. \tag{10}$$

Here p is taken with the plus sign if $p\mathbf{b}$ has the direction of $\mathbf{b}$ and with the minus sign if $p\mathbf{b}$ has the direction opposite to $\mathbf{b}$.

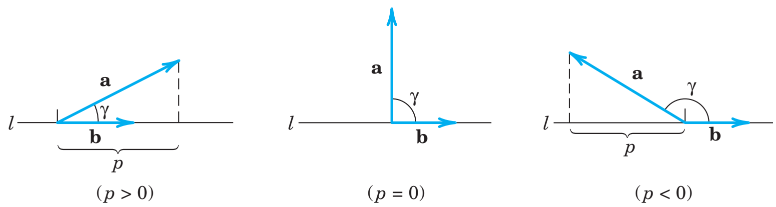


Fig. 181. Component of a vector **a** in the direction of a vector **b**

Multiplying (10) by $|\mathbf{b}|/|\mathbf{b}| = 1$, we have $\mathbf{a} \cdot \mathbf{b}$ in the numerator and thus

$$p = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{b}|} \qquad (\mathbf{b} \neq \mathbf{0}). \tag{11}$$

If $\mathbf{b}$ is a unit vector, as it is often used for fixing a direction, then (11) simply gives

$$p = \mathbf{a} \cdot \mathbf{b} \qquad (|\mathbf{b}| = 1). \tag{12}$$

Figure 182 shows the projection p of $\mathbf{a}$ in the direction of $\mathbf{b}$ (as in Fig. 181) and the projection $q = |\mathbf{b}| \cos \gamma$ of $\mathbf{b}$ in the direction of $\mathbf{a}$.

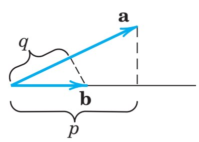


Fig. 182. Projections p of **a** on **b** and q of **b** on **a**

EXAMPLE 4 Orthonormal Basis

By definition, an *orthonormal basis* for 3-space is a basis $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ consisting of orthogonal unit vectors. It has the great advantage that the determination of the coefficients in representations $\mathbf{v} = l_1\mathbf{a} + l_2\mathbf{b} + l_3\mathbf{c}$ of a given vector $\mathbf{v}$ is very simple. We claim that $l_1 = \mathbf{a} \cdot \mathbf{v}$, $l_2 = \mathbf{b} \cdot \mathbf{v}$, $l_3 = \mathbf{c} \cdot \mathbf{v}$. Indeed, this follows simply by taking the inner products of the representation with $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, respectively, and using the orthonormality of the basis, $\mathbf{a} \cdot \mathbf{v} = l_1\mathbf{a} \cdot \mathbf{a} + l_2\mathbf{a} \cdot \mathbf{b} + l_3\mathbf{a} \cdot \mathbf{c} = l_1$, etc.

For example, the unit vectors $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ in (8), Sec. 9.1, associated with a Cartesian coordinate system form an orthonormal basis, called the **standard basis** with respect to the given coordinate system. ■

EXAMPLE 5 Orthogonal Straight Lines in the Plane

Find the straight line L_1 through the point P: (1, 3) in the xy-plane and perpendicular to the straight line $L_2 : x - 2y + 2 = 0$; see Fig. 183.

Solution. The idea is to write a general straight line $L_1 : a_1x + a_2y = c$ as $\mathbf{a} \bullet \mathbf{r} = c$ with $\mathbf{a} = [a_1, a_2] \neq \mathbf{0}$ and $\mathbf{r} = [x, y]$, according to (2). Now the line L_1^* through the origin and parallel to L_1 is $\mathbf{a} \bullet \mathbf{r} = 0$. Hence, by Theorem 1, the vector $\mathbf{a}$ is perpendicular to $\mathbf{r}$. Hence it is perpendicular to L_1^* and also to L_1 because L_1 and L_1^* are parallel. $\mathbf{a}$ is called a **normal vector** of L_1 (and of L_1^*).

Now a normal vector of the given line $x - 2y + 2 = 0$ is $\mathbf{b} = [1, -2]$. Thus L_1 is perpendicular to L_2 if $\mathbf{b} \bullet \mathbf{a} = a_1 - 2a_2 = 0$, for instance, if $\mathbf{a} = [2, 1]$. Hence L_1 is given by $2x + y = c$. It passes through P: (1, 3) when $2 \cdot 1 + 3 = c = 5$. *Answer:* $y = -2x + 5$. Show that the point of intersection is $(x, y) = (1.6, 1.8)$. ■

EXAMPLE 6 Normal Vector to a Plane

Find a unit vector perpendicular to the plane $4x + 2y + 4z = -7$.

Solution. Using (2), we may write any plane in space as

$$\mathbf{a} \bullet \mathbf{r} = a_1x + a_2y + a_3z = c \tag{13}$$

where $\mathbf{a} = [a_1, a_2, a_3] \neq \mathbf{0}$ and $\mathbf{r} = [x, y, z]$. The unit vector in the direction of $\mathbf{a}$ is (Fig. 184)

$$\mathbf{n} = \frac{1}{|\mathbf{a}|}\mathbf{a}.$$

Dividing by $|\mathbf{a}|$, we obtain from (13)

$$\mathbf{n} \bullet \mathbf{r} = p \qquad \text{where} \qquad p = \frac{c}{|\mathbf{a}|}. \tag{14}$$

From (12) we see that p is the projection of $\mathbf{r}$ in the direction of $\mathbf{n}$. This projection has the same constant value $c/|\mathbf{a}|$ for the position vector $\mathbf{r}$ of any point in the plane. Clearly this holds if and only if $\mathbf{n}$ is perpendicular to the plane. $\mathbf{n}$ is called a **unit normal vector** of the plane (the other being $-\mathbf{n}$).

Furthermore, from this and the definition of projection, it follows that $|p|$ is the distance of the plane from the origin. Representation (14) is called **Hesse's[2] normal form** of a plane. In our case, $\mathbf{a} = [4, 2, 4]$, $c = -7$, $|\mathbf{a}| = 6$, $\mathbf{n} = \frac{1}{6}\mathbf{a} = [\frac{2}{3}, \frac{1}{3}, \frac{2}{3}]$, and the plane has the distance $\frac{7}{6}$ from the origin. ■

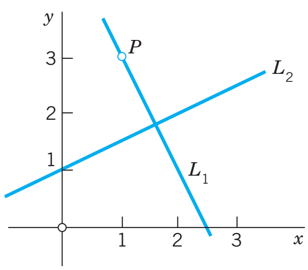


Fig. 183. Example 5

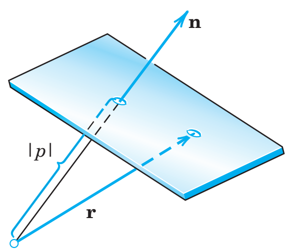


Fig. 184. Normal vector to a plane

[2]LUDWIG OTTO HESSE (1811–1874), German mathematician who contributed to the theory of curves and surfaces.

PROBLEM SET 9.2

1–10 INNER PRODUCT

Let $\mathbf{a} = [1, -3, 5]$, $\mathbf{b} = [4, 0, 8]$, $\mathbf{c} = [-2, 9, 1]$. Find:

1. $\mathbf{a} \cdot \mathbf{b}$, $\mathbf{b} \cdot \mathbf{a}$, $\mathbf{b} \cdot \mathbf{c}$

2. $(-3\mathbf{a} + 5\mathbf{c}) \cdot \mathbf{b}$, $15(\mathbf{a} - \mathbf{c}) \cdot \mathbf{b}$

3. $|\mathbf{a}|$, $|2\mathbf{b}|$, $|-\mathbf{c}|$

4. $|\mathbf{a} + \mathbf{b}|$, $|\mathbf{a}| + |\mathbf{b}|$

5. $|\mathbf{b} + \mathbf{c}|$, $|\mathbf{b}| + |\mathbf{c}|$

6. $|\mathbf{a} + \mathbf{c}|^2 + |\mathbf{a} - \mathbf{c}|^2 - 2(|\mathbf{a}|^2 + |\mathbf{c}|^2)$

7. $|\mathbf{a} \cdot \mathbf{c}|$, $|\mathbf{a}||\mathbf{c}|$

8. $5\mathbf{a} \cdot 13\mathbf{b}$, $65\mathbf{a} \cdot \mathbf{b}$

9. $15\mathbf{a} \cdot \mathbf{b} + 15\mathbf{a} \cdot \mathbf{c}$, $15\mathbf{a} \cdot (\mathbf{b} + \mathbf{c})$

10. $\mathbf{a} \cdot (\mathbf{b} - \mathbf{c})$, $(\mathbf{a} - \mathbf{b}) \cdot \mathbf{c}$

11–16 GENERAL PROBLEMS

11. What laws do Probs. 1 and 4–7 illustrate?

12. What does $\mathbf{u} \cdot \mathbf{v} = \mathbf{u} \cdot \mathbf{w}$ imply if $\mathbf{u} = \mathbf{0}$? If $\mathbf{u} \neq \mathbf{0}$?

13. Prove the Cauchy–Schwarz inequality.

14. Verify the Cauchy–Schwarz and triangle inequalities for the above **a** and **b**.

15. Prove the parallelogram equality. Explain its name.

16. Triangle inequality. Prove Eq. (7). *Hint.* Use Eq. (3) for $|\mathbf{a} + \mathbf{b}|$ and Eq. (6) to prove the square of Eq. (7), then take roots.

17–20 WORK

Find the work done by a force **p** acting on a body if the body is displaced along the straight segment $\overline{AB}$ from A to B. Sketch $\overline{AB}$ and **p**. Show the details.

17. $\mathbf{p} = [2, 5, 0]$, A: (1, 3, 3), B: (3, 5, 5)

18. $\mathbf{p} = [-1, -2, 4]$, A: (0, 0, 0), B: (6, 7, 5)

19. $\mathbf{p} = [0, 4, 3]$, A: (4, 5, −1), B: (1, 3, 0)

20. $\mathbf{p} = [6, -3, -3]$, A: (1, 5, 2), B: (3, 4, 1)

21. Resultant. Is the work done by the resultant of two forces in a displacement the sum of the work done by each of the forces separately? Give proof or counterexample.

22–30 ANGLE BETWEEN VECTORS

Let $\mathbf{a} = [1, 1, 0]$, $\mathbf{b} = [3, 2, 1]$, and $\mathbf{c} = [1, 0, 2]$. Find the angle between:

22. **a**, **b**

23. **b**, **c**

24. $\mathbf{a} + \mathbf{c}$, $\mathbf{b} + \mathbf{c}$

25. What will happen to the angle in Prob. 24 if we replace **c** by $n\mathbf{c}$ with larger and larger n?

26. Cosine law. Deduce the law of cosines by using vectors **a**, **b**, and $\mathbf{a} - \mathbf{b}$.

27. Addition law. $\cos(\alpha - \beta) = \cos\alpha \cos\beta + \sin\alpha \sin\beta$. Obtain this by using $\mathbf{a} = [\cos\alpha, \sin\alpha]$, $\mathbf{b} = [\cos\beta, \sin\beta]$ where $0 \leqq \alpha \leqq \beta \leqq 2\pi$.

28. Triangle. Find the angles of the triangle with vertices A: (0, 0, 2), B: (3, 0, 2), and C: (1, 1, 1). Sketch the triangle.

29. Parallelogram. Find the angles if the vertices are (0, 0), (6, 0), (8, 3), and (2, 3).

30. Distance. Find the distance of the point A: (1, 0, 2) from the plane P: $3x + y + z = 9$. Make a sketch.

31–35 ORTHOGONALITY

ORTHOGONALITY is particularly important, mainly because of orthogonal coordinates, such as ***Cartesian coordinates***, whose ***natural basis*** [Eq. (9), Sec. 9.1], consists of three orthogonal unit vectors.

31. For what values of a_1 are $[a_1, 4, 3]$ and $[3, -2, 12]$ orthogonal?

32. Planes. For what c are $3x + z = 5$ and $8x - y + cz = 9$ orthogonal?

33. Unit vectors. Find all unit vectors $\mathbf{a} = [a_1, a_2]$ in the plane orthogonal to [4, 3].

34. Corner reflector. Find the angle between a light ray and its reflection in three orthogonal plane mirrors, known as *corner reflector*.

35. Parallelogram. When will the diagonals be orthogonal? Give a proof.

36–40 COMPONENT IN THE DIRECTION OF A VECTOR

Find the component of **a** in the direction of **b**. Make a sketch.

36. $\mathbf{a} = [1, 1, 1]$, $\mathbf{b} = [2, 1, 3]$

37. $\mathbf{a} = [3, 4, 0]$, $\mathbf{b} = [4, -3, 2]$

38. $\mathbf{a} = [8, 2, 0]$, $\mathbf{b} = [-4, -1, 0]$

39. When will the component (the projection) of **a** in the direction of **b** be equal to the component (the projection) of **b** in the direction of **a**? First guess.

40. What happens to the component of **a** in the direction of **b** if you change the length of **b**?

9.3 Vector Product (Cross Product)

We shall define another form of multiplication of vectors, inspired by applications, whose result will be a *vector*. This is in contrast to the dot product of Sec. 9.2 where multiplication resulted in a *scalar*. We can construct a vector $\mathbf{v}$ that is perpendicular to two vectors $\mathbf{a}$ and $\mathbf{b}$, which are two sides of a parallelogram on a plane in space as indicated in Fig. 185, such that the length $|\mathbf{v}|$ is numerically equal to the area of that parallelogram. Here then is the new concept.

DEFINITION

Vector Product (Cross Product, Outer Product) of Vectors

The **vector product** or **cross product** $\mathbf{a} \times \mathbf{b}$ (read "$\mathbf{a}$ cross $\mathbf{b}$") of two vectors $\mathbf{a}$ and $\mathbf{b}$ is the vector $\mathbf{v}$ denoted by

$$\mathbf{v} = \mathbf{a} \times \mathbf{b}$$

I. If $\mathbf{a} = \mathbf{0}$ or $\mathbf{b} = \mathbf{0}$, then we define $\mathbf{v} = \mathbf{a} \times \mathbf{b} = \mathbf{0}$.

II. If both vectors are nonzero vectors, then vector $\mathbf{v}$ has the length

$$|\mathbf{v}| = |\mathbf{a} \times \mathbf{b}| = |\mathbf{a}||\mathbf{b}| \sin \gamma, \tag{1}$$

where γ is the angle between $\mathbf{a}$ and $\mathbf{b}$ as in Sec. 9.2.

Furthermore, by design, $\mathbf{a}$ and $\mathbf{b}$ form the sides of a parallelogram on a plane in space. The parallelogram is shaded in blue in Fig. 185. The area of this blue parallelogram is precisely given by Eq. (1), so that the length $|\mathbf{v}|$ of the vector $\mathbf{v}$ is equal to the area of that parallelogram.

III. If $\mathbf{a}$ and $\mathbf{b}$ lie in the same straight line, i.e., $\mathbf{a}$ and $\mathbf{b}$ have the same or opposite directions, then γ is $0°$ or $180°$ so that $\sin \gamma = 0$. In that case $|\mathbf{v}| = 0$ so that $\mathbf{v} = \mathbf{a} \times \mathbf{b} = \mathbf{0}$.

IV. If cases I and III do not occur, then $\mathbf{v}$ is a nonzero vector. The direction of $\mathbf{v} = \mathbf{a} \times \mathbf{b}$ is perpendicular to both $\mathbf{a}$ and $\mathbf{b}$ such that $\mathbf{a}$, $\mathbf{b}$, $\mathbf{v}$—precisely in this order (!)—form a right-handed triple as shown in Figs. 185–187 and explained below.

Another term for vector product is outer product.

Remark. Note that I and III completely characterize the exceptional case when the cross product is equal to the zero vector, and II and IV the regular case where the cross product is perpendicular to two vectors.

Just as we did with the dot product, we would also like to express the cross product in components. Let $\mathbf{a} = [a_1, a_2, a_3]$ and $\mathbf{b} = [b_1, b_2, b_3]$. Then $\mathbf{v} = [v_1, v_2, v_3] = \mathbf{a} \times \mathbf{b}$ has the components

$$v_1 = a_2b_3 - a_3b_2, \qquad v_2 = a_3b_1 - a_1b_3, \qquad v_3 = a_1b_2 - a_2b_1. \tag{2}$$

Here the Cartesian coordinate system is *right-handed*, as explained below (see also Fig. 188). (For a left-handed system, each component of $\mathbf{v}$ must be multiplied by -1. Derivation of (2) in App. 4.)

Right-Handed Triple. A triple of vectors **a**, **b**, **v** is *right-handed* if the vectors in the given order assume the same sort of orientation as the thumb, index finger, and middle finger of the right hand when these are held as in Fig. 186. We may also say that if **a** is rotated into the direction of **b** through the angle $\gamma\,(<\pi)$, then **v** advances in the same direction as a right-handed screw would if turned in the same way (Fig. 187).

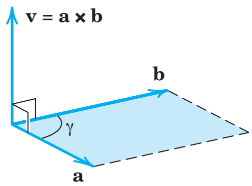


Fig. 185. Vector product

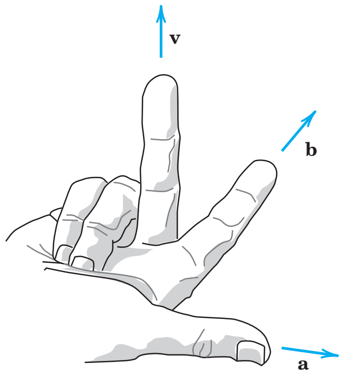


Fig. 186. Right-handed triple of vectors **a, b, v**

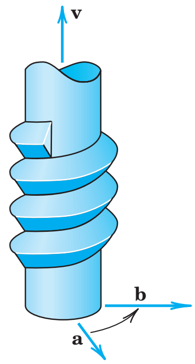
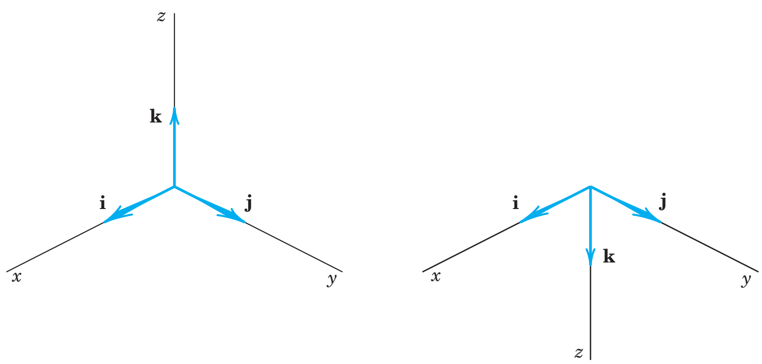


Fig. 187. Right-handed screw

Right-Handed Cartesian Coordinate System. The system is called **right-handed** if the corresponding unit vectors **i**, **j**, **k** in the positive directions of the axes (see Sec. 9.1) form a right-handed triple as in Fig. 188a. The system is called **left-handed** if the sense of **k** is reversed, as in Fig. 188b. In applications, we prefer right-handed systems.



(a) Right-handed (b) Left-handed

Fig. 188. The two types of Cartesian coordinate systems

How to Memorize (2). If you know second- and third-order determinants, you see that (2) can be written

$$(2^*) \qquad v_1 = \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}, \qquad v_2 = -\begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} = +\begin{vmatrix} a_3 & a_1 \\ b_3 & b_1 \end{vmatrix}, \qquad v_3 = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}$$

and $\mathbf{v} = [v_1, v_2, v_3] = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}$ is the expansion of the following symbolic determinant by its first row. (We call the determinant "symbolic" because the first row consists of vectors rather than of numbers.)

$$(2^{**}) \qquad \mathbf{v} = \mathbf{a} \times \mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}\mathbf{i} - \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix}\mathbf{j} + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}\mathbf{k}.$$

For a left-handed system the determinant has a minus sign in front.

EXAMPLE 1 Vector Product

For the vector product $\mathbf{v} = \mathbf{a} \times \mathbf{b}$ of $\mathbf{a} = [1, 1, 0]$ and $\mathbf{b} = [3, 0, 0]$ in right-handed coordinates we obtain from (2)

$$v_1 = 0, \qquad v_2 = 0, \qquad v_3 = 1 \cdot 0 - 1 \cdot 3 = -3.$$

We confirm this by (2**):

$$\mathbf{v} = \mathbf{a} \times \mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 1 & 0 \\ 3 & 0 & 0 \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ 0 & 0 \end{vmatrix}\mathbf{i} - \begin{vmatrix} 1 & 0 \\ 3 & 0 \end{vmatrix}\mathbf{j} + \begin{vmatrix} 1 & 1 \\ 3 & 0 \end{vmatrix}\mathbf{k} = -3\mathbf{k} = [0, 0, -3].$$

To check the result in this simple case, sketch **a**, **b**, and **v**. Can you see that two vectors in the xy-plane must always have their vector product parallel to the z-axis (or equal to the zero vector)? ■

EXAMPLE 2 Vector Products of the Standard Basis Vectors

$$(3) \qquad \begin{aligned} \mathbf{i} \times \mathbf{j} &= \mathbf{k}, & \mathbf{j} \times \mathbf{k} &= \mathbf{i}, & \mathbf{k} \times \mathbf{i} &= \mathbf{j} \\ \mathbf{j} \times \mathbf{i} &= -\mathbf{k}, & \mathbf{k} \times \mathbf{j} &= -\mathbf{i}, & \mathbf{i} \times \mathbf{k} &= -\mathbf{j}. \end{aligned}$$

We shall use this in the next proof. ■

THEOREM 1 General Properties of Vector Products

(a) *For every scalar l,*

$$(4) \qquad (l\mathbf{a}) \times \mathbf{b} = l(\mathbf{a} \times \mathbf{b}) = \mathbf{a} \times (l\mathbf{b}).$$

(b) *Cross multiplication is distributive with respect to vector addition; that is,*

$$(5) \qquad \begin{aligned} &(\alpha) \quad \mathbf{a} \times (\mathbf{b} + \mathbf{c}) = (\mathbf{a} \times \mathbf{b}) + (\mathbf{a} \times \mathbf{c}), \\ &(\beta) \quad (\mathbf{a} + \mathbf{b}) \times \mathbf{c} = (\mathbf{a} \times \mathbf{c}) + (\mathbf{b} \times \mathbf{c}). \end{aligned}$$

(c) *Cross multiplication is **not commutative** but **anticommutative;** that is,*

$$(6) \qquad \mathbf{b} \times \mathbf{a} = -(\mathbf{a} \times \mathbf{b}) \qquad \text{(Fig. 189).}$$

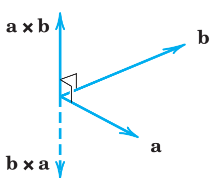


Fig. 189. Anticommutativity of cross multiplication

(d) *Cross multiplication is* ***not associative***; *that is, in general,*

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) \neq (\mathbf{a} \times \mathbf{b}) \times \mathbf{c} \tag{7}$$

so that the parentheses cannot be omitted.

PROOF Equation (4) follows directly from the definition. In (5α), formula (2*) gives for the first component on the left

$$\begin{vmatrix} a_2 & a_3 \\ b_2 + c_2 & b_3 + c_3 \end{vmatrix} = a_2(b_3 + c_3) - a_3(b_2 + c_2)$$
$$= (a_2b_3 - a_3b_2) + (a_2c_3 - a_3c_2)$$
$$= \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} + \begin{vmatrix} a_2 & a_3 \\ c_2 & c_3 \end{vmatrix}.$$

By (2*) the sum of the two determinants is the first component of $(\mathbf{a} \times \mathbf{b}) + (\mathbf{a} \times \mathbf{c})$, the right side of (5$\alpha$). For the other components in (5α) and in 5(β), equality follows by the same idea.

Anticommutativity (6) follows from (2**) by noting that the interchange of Rows 2 and 3 multiplies the determinant by -1. We can confirm this geometrically if we set $\mathbf{a} \times \mathbf{b} = \mathbf{v}$ and $\mathbf{b} \times \mathbf{a} = \mathbf{w}$; then $|\mathbf{v}| = |\mathbf{w}|$ by (1), and for $\mathbf{b}$, $\mathbf{a}$, $\mathbf{w}$ to form a *right-handed* triple, we must have $\mathbf{w} = -\mathbf{v}$.

Finally, $\mathbf{i} \times (\mathbf{i} \times \mathbf{j}) = \mathbf{i} \times \mathbf{k} = -\mathbf{j}$, whereas $(\mathbf{i} \times \mathbf{i}) \times \mathbf{j} = \mathbf{0} \times \mathbf{j} = \mathbf{0}$ (see Example 2). This proves (7). ∎

Typical Applications of Vector Products

EXAMPLE 3 **Moment of a Force**

In mechanics the moment m of a force $\mathbf{p}$ about a point Q is defined as the product $m = |\mathbf{p}|d$, where d is the (perpendicular) distance between Q and the line of action L of $\mathbf{p}$ (Fig. 190). If $\mathbf{r}$ is the vector from Q to any point A on L, then $d = |\mathbf{r}| \sin \gamma$, as shown in Fig. 190, and

$$m = |\mathbf{r}||\mathbf{p}| \sin \gamma.$$

Since γ is the angle between $\mathbf{r}$ and $\mathbf{p}$, we see from (1) that $m = |\mathbf{r} \times \mathbf{p}|$. The vector

$$\mathbf{m} = \mathbf{r} \times \mathbf{p} \tag{8}$$

is called the **moment vector** or **vector moment** of $\mathbf{p}$ about Q. Its magnitude is m. If $\mathbf{m} \neq \mathbf{0}$, its direction is that of the axis of the rotation about Q that $\mathbf{p}$ has the tendency to produce. This axis is perpendicular to both $\mathbf{r}$ and $\mathbf{p}$. ∎

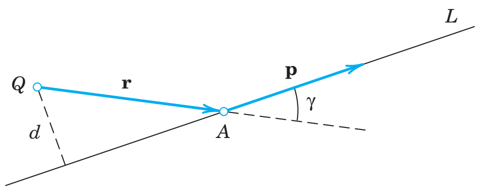


Fig. 190. Moment of a force $\mathbf{p}$