Elementary Differential Equations and Boundary Value Problems

Eleventh Edition

WILLIAM E. BOYCE
Edward P. Hamilton Professor Emeritus
Department of Mathematical Sciences
Rensselaer Polytechnic Institute

RICHARD C. DIPRIMA
formerly Eliza Ricketts Foundation Professor
Department of Mathematical Sciences
Rensselaer Polytechnic Institute

DOUGLAS B. MEADE
Department of Mathematics
University of South Carolina - Columbia

WILEY

Contents

CHAPTER 2

First-Order Differential Equations

This chapter deals with differential equations of first order

$$\frac{dy}{dt} = f(t, y), \tag{1}$$

where f is a given function of two variables. Any differentiable function $y = \phi(t)$ that satisfies this equation for all t in some interval is called a solution, and our objective is to determine whether such functions exist and, if so, to develop methods for finding them. Unfortunately, for an arbitrary function f, there is no general method for solving the equation in terms of elementary functions. Instead, we will describe several methods, each of which is applicable to a certain subclass of first-order equations.

The most important of these are linear equations (Section 2.1), separable equations (Section 2.2), and exact equations (Section 2.6). Other sections of this chapter describe some of the important applications of first-order differential equations, introduce the idea of approximating a solution by numerical computation, and discuss some theoretical questions related to the existence and uniqueness of solutions. The final section includes an example of chaotic solutions in the context of first-order difference equations, which have some important points of similarity with differential equations and are simpler to investigate.

2.1 Linear Differential Equations; Method of Integrating Factors

If the function f in equation (1) depends linearly on the dependent variable y, then equation (1) is a first-order linear differential equation. In Sections 1.1 and 1.2 we discussed a restricted type of first-order linear differential equation in which the coefficients are constants. A typical example is

$$\frac{dy}{dt} = -ay + b, \tag{2}$$

where a and b are given constants. Recall that an equation of this form describes the motion of an object falling in the atmosphere.

Now we want to consider the most general first-order linear differential equation, which is obtained by replacing the coefficients a and b in equation (2) by arbitrary functions of t. We will usually write the general **first-order linear differential equation** in the standard form

$$\frac{dy}{dt} + p(t)y = g(t), \tag{3}$$

where p and g are given functions of the independent variable t. Sometimes it is more convenient to write the equation in the form

$$P(t)\frac{dy}{dt} + Q(t)y = G(t), \tag{4}$$

where P, Q, and G are given. Of course, as long as $P(t) \neq 0$, you can convert equation (4) to equation (3) by dividing both sides of equation (4) by $P(t)$.

In some cases it is possible to solve a first-order linear differential equation immediately by integrating the equation, as in the next example.

EXAMPLE 1

Solve the differential equation

$$(4+t^2)\frac{dy}{dt}+2ty=4t. \tag{5}$$

Solution:

The left-hand side of equation (5) is a linear combination of dy/dt and y, a combination that also appears in the rule from calculus for differentiating a product. In fact,

$$(4+t^2)\frac{dy}{dt}+2ty=\frac{d}{dt}\left((4+t^2)y\right);$$

it follows that equation (5) can be rewritten as

$$\frac{d}{dt}\left((4+t^2)y\right)=4t. \tag{6}$$

Thus, even though y is unknown, we can integrate both sides of equation (6) with respect to t, thereby obtaining

$$(4+t^2)y=2t^2+c, \tag{7}$$

where c is an arbitrary constant of integration. Solving for y, we find that

$$y=\frac{2t^2}{4+t^2}+\frac{c}{4+t^2}. \tag{8}$$

This is the general solution of equation (5).

Unfortunately, most first-order linear differential equations cannot be solved as illustrated in Example 1 because their left-hand sides are not the derivative of the product of y and some other function. However, Leibniz discovered that if the differential equation is multiplied by a certain function $\mu(t)$, then the equation is converted into one that is immediately integrable by using the product rule for derivatives, just as in Example 1. The function $\mu(t)$ is called an **integrating factor** and our main task in this section is to determine how to find it for a given equation. We will show how this method works first for an example and then for the general first-order linear differential equation in the standard form (3).

EXAMPLE 2

Find the general solution of the differential equation

$$\frac{dy}{dt}+\frac{1}{2}y=\frac{1}{2}e^{t/3}. \tag{9}$$

Draw some representative integral curves; that is, plot solutions corresponding to several values of the arbitrary constant c. Also find the particular solution whose graph contains the point $(0, 1)$.

Solution:

The first step is to multiply equation (9) by a function $\mu(t)$, as yet undetermined; thus

$$\mu(t)\frac{dy}{dt}+\frac{1}{2}\mu(t)y=\frac{1}{2}\mu(t)e^{t/3}. \tag{10}$$

The question now is whether we can choose $\mu(t)$ so that the left-hand side of equation (10) is the derivative of the product $\mu(t)y$. For any differentiable function $\mu(t)$ we have

$$\frac{d}{dt}(\mu(t)y)=\mu(t)\frac{dy}{dt}+\frac{d\mu(t)}{dt}y. \tag{11}$$

Thus the left-hand side of equation (10) and the right-hand side of equation (11) are identical, provided that we choose $\mu(t)$ to satisfy

$$\frac{d\mu(t)}{dt}=\frac{1}{2}\mu(t). \tag{12}$$

Our search for an integrating factor will be successful if we can find a solution of equation (12). Perhaps you can readily identify a function that satisfies equation (12): What well-known function from calculus has a derivative that is equal to one-half times the original function? More systematically, rewrite equation (12) as

$$\frac{1}{\mu(t)}\frac{d\mu(t)}{dt}=\frac{1}{2},$$

which is equivalent to

$$\frac{d}{dt}\ln|\mu(t)|=\frac{1}{2}. \tag{13}$$

Then it follows that

$$\ln|\mu(t)|=\frac{1}{2}t+C,$$

or

$$\mu(t)=ce^{t/2}. \tag{14}$$

The function $\mu(t)$ given by equation (14) is an integrating factor for equation (9). Since we do not need the most general integrating factor, we will choose c to be 1 in equation (14) and use $\mu(t)=e^{t/2}$.

Now we return to equation (9), multiply it by the integrating factor $e^{t/2}$, and obtain

$$e^{t/2}\frac{dy}{dt}+\frac{1}{2}e^{t/2}y=\frac{1}{2}e^{5t/6}. \tag{15}$$

By the choice we have made of the integrating factor, the left-hand side of equation (15) is the derivative of $e^{t/2}y$, so that equation (15) becomes

$$\frac{d}{dt}(e^{t/2}y)=\frac{1}{2}e^{5t/6}. \tag{16}$$

By integrating both sides of equation (16), we obtain

$$e^{t/2}y=\frac{3}{5}e^{5t/6}+c, \tag{17}$$

where c is an arbitrary constant. Finally, on solving equation (17) for y, we have the general solution of equation (9), namely,

$$y=\frac{3}{5}e^{t/3}+ce^{-t/2}. \tag{18}$$

To find the solution passing through the point $(0, 1)$, we set $t=0$ and $y=1$ in equation (18), obtaining $1=3/5+c$. Thus $c=2/5$, and the desired solution is

$$y=\frac{3}{5}e^{t/3}+\frac{2}{5}e^{-t/2}. \tag{19}$$

Figure 2.1.1 includes the graphs of equation (18) for several values of c with a direction field in the background. The solution satisfying $y(0)=1$ is shown by the green curve.

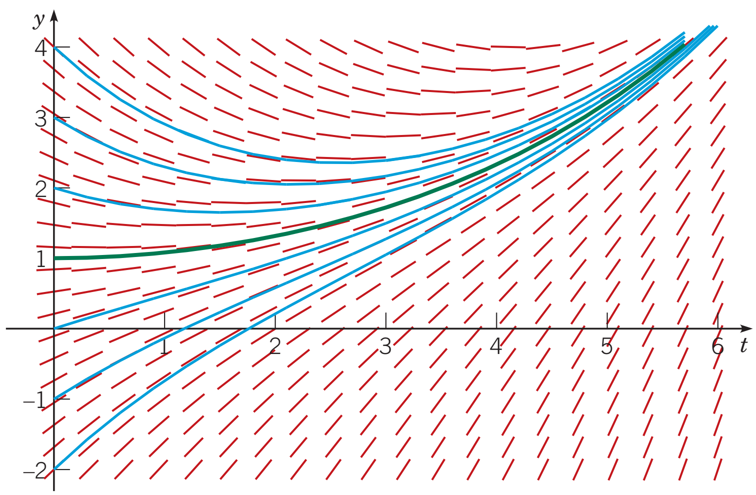


FIGURE 2.1.1 Direction field and integral curves of $y'+\frac{1}{2}y=\frac{1}{2}e^{t/3}$; the green curve passes through the point $(0, 1)$.

Let us now extend the method of integrating factors to equations of the form

$$\frac{dy}{dt} + ay = g(t), \tag{20}$$

where a is a given constant and $g(t)$ is a given function. Proceeding as in Example 2, we find that the integrating factor $\mu(t)$ must satisfy

$$\frac{d\mu}{dt} = a\mu, \tag{21}$$

rather than equation (12). Thus the integrating factor is $\mu(t) = e^{at}$. Multiplying equation (20) by $\mu(t)$, we obtain

$$e^{at}\frac{dy}{dt} + ae^{at}y = e^{at}g(t),$$

or

$$\frac{d}{dt}(e^{at}y) = e^{at}g(t). \tag{22}$$

By integrating both sides of equation (22), we find that

$$e^{at}y = \int e^{at}g(t)\,dt + c, \tag{23}$$

where c is an arbitrary constant. For many simple functions $g(t)$, we can evaluate the integral in equation (23) and express the solution y in terms of elementary functions, as in Example 2. However, for more complicated functions $g(t)$, it is necessary to leave the solution in integral form. In this case

$$y = e^{-at}\int_{t_0}^{t} e^{as}g(s)\,ds + ce^{-at}. \tag{24}$$

Note that in equation (24) we have used s to denote the integration variable to distinguish it from the independent variable t, and we have chosen some convenient value t_0 as the lower limit of integration. (See Theorem 2.4.1.) The choice of t_0 determines the specific value of the constant c but does not change the solution. For example, plugging $t = t_0$ into the solution formula (24) shows that $c = y(t_0)e^{at_0}$.

EXAMPLE 3

Find the general solution of the differential equation

$$\frac{dy}{dt} - 2y = 4 - t \tag{25}$$

and plot the graphs of several solutions. Discuss the behavior of solutions as $t \to \infty$.

Solution:

Equation (25) is of the form (20) with $a = -2$; therefore, the integrating factor is $\mu(t) = e^{-2t}$. Multiplying the differential equation (25) by $\mu(t)$, we obtain

$$e^{-2t}\frac{dy}{dt} - 2e^{-2t}y = 4e^{-2t} - te^{-2t},$$

or

$$\frac{d}{dt}\left(e^{-2t}y\right) = 4e^{-2t} - te^{-2t}. \tag{26}$$

Then, by integrating both sides of this equation, we have

$$e^{-2t}y = -2e^{-2t} + \frac{1}{2}te^{-2t} + \frac{1}{4}e^{-2t} + c,$$

where we have used integration by parts on the last term in equation (26). Thus the general solution of equation (25) is

$$y = -\frac{7}{4} + \frac{1}{2}t + ce^{2t}. \tag{27}$$

Figure 2.1.2 shows the direction field and graphs of the solution (27) for several values of c. The behavior of the solution for large values of t is determined by the term ce^{2t}. If $c \neq 0$, then the solution grows exponentially large in magnitude, with the same sign as c itself. Thus the solutions diverge as t becomes large. The boundary between solutions that ultimately grow positively and those that ultimately grow negatively occurs when $c = 0$. If we substitute $c = 0$ into equation (27) and then set $t = 0$, we find that $y = -7/4$ is the separation point on the y-axis. Note that for this initial value, the solution is $y = -\frac{7}{4} + \frac{1}{2}t$; it grows positively, but linearly rather than exponentially.

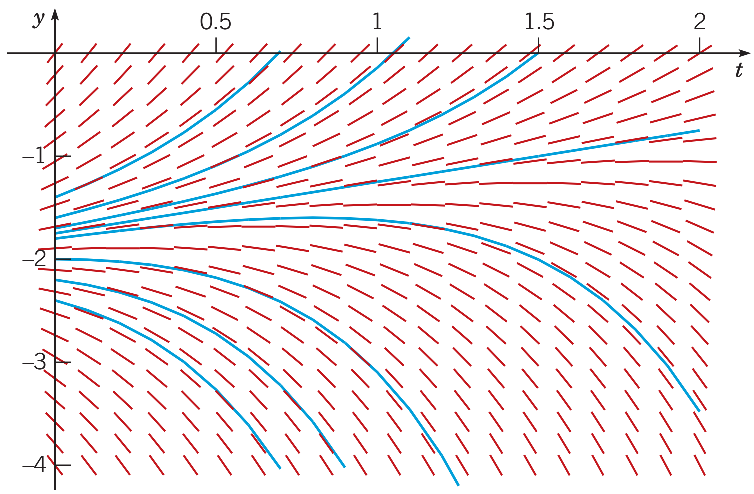


FIGURE 2.1.2 Direction field and integral curves of $y' - 2y = 4 - t$.

Now we return to the general first-order linear differential equation (3)

$$\frac{dy}{dt} + p(t)y = g(t),$$

where p and g are given functions. To determine an appropriate integrating factor, we multiply equation (3) by an as yet undetermined function $\mu(t)$, obtaining

$$\mu(t)\frac{dy}{dt} + p(t)\mu(t)y = \mu(t)g(t). \tag{28}$$

Following the same line of development as in Example 2, we see that the left-hand side of equation (28) is the derivative of the product $\mu(t)y$, provided that $\mu(t)$ satisfies the equation

$$\frac{d\mu(t)}{dt} = p(t)\mu(t). \tag{29}$$

If we assume temporarily that $\mu(t)$ is positive, then we have

$$\frac{1}{\mu(t)}\frac{d\mu(t)}{dt} = p(t),$$

and consequently

$$\ln|\mu(t)| = \int p(t)\,dt + k.$$

By choosing the arbitrary constant k to be zero, we obtain the simplest possible function for μ, namely,

$$\mu(t) = \exp\int p(t)\,dt. \tag{30}$$

Note that $\mu(t)$ is positive for all t, as we assumed. Returning to equation (28), we have

$$\frac{d}{dt}(\mu(t)y) = \mu(t)g(t). \tag{31}$$

Hence

$$\mu(t)y = \int \mu(t)g(t)\,dt + c, \tag{32}$$

where c is an arbitrary constant. Sometimes the integral in equation (32) can be evaluated in terms of elementary functions. However, in general this is not possible, so the general solution of equation (3) is

$$y = \frac{1}{\mu(t)}\left(\int_{t_0}^{t} \mu(s)g(s)\,ds + c\right), \tag{33}$$

where again t_0 is some convenient lower limit of integration. Observe that equation (33) involves two integrations, one to obtain $\mu(t)$ from equation (30) and the other to determine y from equation (33).

EXAMPLE 4

Solve the initial value problem

$$ty' + 2y = 4t^2, \tag{34}$$

$$y(1) = 2. \tag{35}$$

Solution:

In order to determine $p(t)$ and $g(t)$ correctly, we must first rewrite equation (34) in the standard form (3). Thus we have

$$y' + \frac{2}{t}y = 4t, \tag{36}$$

so $p(t) = 2/t$ and $g(t) = 4t$. To solve equation (36), we first compute the integrating factor $\mu(t)$:

$$\mu(t) = \exp\left(\int \frac{2}{t}\,dt\right) = e^{2\ln|t|} = t^2.$$

On multiplying equation (36) by $\mu(t) = t^2$, we obtain

$$t^2y' + 2ty = (t^2y)' = 4t^3,$$

and therefore

$$t^2y = \int 4t^3\,dt = t^4 + c,$$

where c is an arbitrary constant. It follows that, for $t > 0$,

$$y = t^2 + \frac{c}{t^2} \tag{37}$$

is the general solution of equation (34). Integral curves of equation (34) for several values of c are shown in Figure 2.1.3.

To satisfy initial condition (35), set $t = 1$ and $y = 2$ in equation (37): $2 = 1 + c$, so $c = 1$; thus

$$y = t^2 + \frac{1}{t^2}, \quad t > 0 \tag{38}$$

is the solution of the initial value problem (24), (25). This solution is shown by the green curve in Figure 2.1.3. Note that it becomes unbounded and is asymptotic to the positive y-axis as $t \to 0$ from the right. This is the effect of the infinite discontinuity in the coefficient $p(t)$ at the origin. It is important to note that while the function $y = t^2 + 1/t^2$ for $t < 0$ is part of the general solution of equation (34), it is not part of the solution of this initial value problem.

This is the first example in which the solution fails to exist for some values of t. Again, this is due to the infinite discontinuity in $p(t)$ at $t = 0$, which restricts the solution to the interval $0 < t < \infty$.

Looking again at Figure 2.1.3, we see that some solutions (those for which $c > 0$) are asymptotic to the positive y-axis as $t \to 0$ from the right, while other solutions (for which $c < 0$) are asymptotic to the negative y-axis. If we generalize the initial condition (35) to

$$y(1) = y_0, \tag{39}$$

then $c = y_0 - 1$ and the solution (38) becomes

$$y = t^2 + \frac{y_0 - 1}{t^2}, \quad t > 0 \tag{40}$$

Note that when $y_0 = 1$, so $c = 0$, the solution is $y = t^2$, which remains bounded and differentiable even at $t = 0$. (This is the red curve in Figure 2.1.3.)

As in Example 3, this is another instance where there is a critical initial value, namely, $y_0 = 1$, that separates solutions that behave in one way from others that behave quite differently.

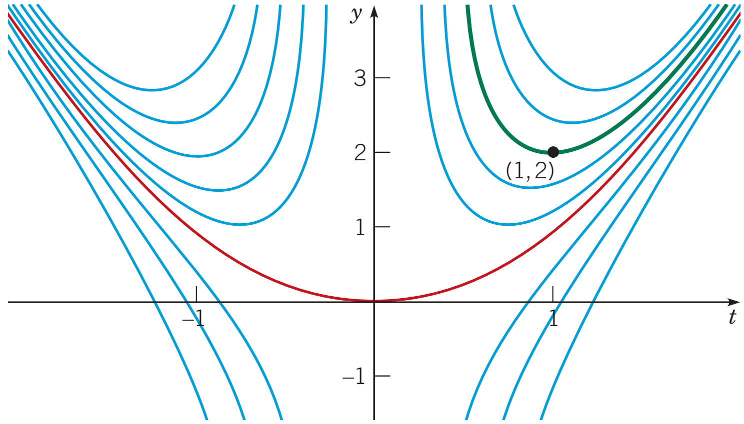


FIGURE 2.1.3 Integral curves of the differential equation $ty' + 2y = 4t^2$; the green curve is the particular solution with $y(1) = 2$. The red curve is the particular solution with $y(1) = 1$.

EXAMPLE 5

Solve the initial value problem

$$2y' + ty = 2, \tag{41}$$

$$y(0) = 1. \tag{42}$$

Solution:

To convert the differential equation (41) to the standard form (3), we must divide equation (41) by 2, obtaining

$$y' + \frac{t}{2}y = 1. \tag{43}$$

Thus $p(t) = t/2$, and the integrating factor is $\mu(t) = \exp(t^2/4)$. Then multiply equation (43) by $\mu(t)$, so that

$$e^{t^2/4}y' + \frac{t}{2}e^{t^2/4}y = e^{t^2/4}. \tag{44}$$

The left-hand side of equation (44) is the derivative of $e^{t^2/4}y$, so by integrating both sides of equation (44), we obtain

$$e^{t^2/4}y = \int e^{t^2/4}\,dt + c. \tag{45}$$

The integral on the right-hand side of equation (45) cannot be evaluated in terms of the usual elementary functions, so we leave the integral unevaluated. By choosing the lower limit of integration as the initial point $t = 0$, we can replace equation (45) by

$$e^{t^2/4}y = \int_0^t e^{s^2/4}\,ds + c, \tag{46}$$

where c is an arbitrary constant. It then follows that the general solution y of equation (41) is given by

$$y = e^{-t^2/4}\int_0^t e^{s^2/4}\,ds + ce^{-t^2/4}. \tag{47}$$

To determine the particular solution that satisfies the initial condition (42), set $t = 0$ and $y = 1$ in equation (47):

$$\begin{aligned} 1 &= e^0 \int_0^0 e^{-s^2/4}\,ds + ce^0 \\ &= 0 + c, \end{aligned}$$

so $c = 1$.

The main purpose of this example is to illustrate that sometimes the solution must be left in terms of an integral. This is usually at most a slight inconvenience, rather than a serious obstacle. For a given value of t, the integral in equation (47) is a definite integral and can be approximated to any desired degree of accuracy by using readily available numerical integrators. By repeating this process for many values of t and plotting the results, you can obtain a graph of a solution. Alternatively, you can use a numerical approximation method, such as those discussed in Chapter 8, that proceed directly from the differential equation and need no expression for the solution. Software packages such as Maple, Mathematica, MATLAB and Sage readily execute such procedures and produce graphs of solutions of differential equations.

Figure 2.1.4 displays graphs of the solution (47) for several values of c. The particular solution satisfying the initial condition $y(0) = 1$ is shown in black. From the figure it may be plausible to conjecture that all solutions approach a limit as $t \to \infty$. The limit can also be found analytically (see Problem 22).

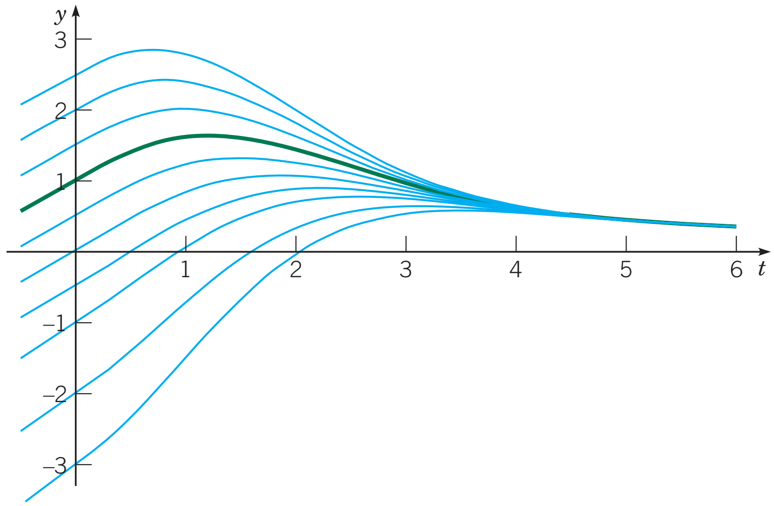


FIGURE 2.1.4 Integral curves of $2y' + ty = 2$; the green curve is the particular solution satisfying the initial condition $y(0) = 1$.

Problems

In each of Problems 1 through 8:

a. Draw a direction field for the given differential equation.
b. Based on an inspection of the direction field, describe how solutions behave for large t.
c. Find the general solution of the given differential equation, and use it to determine how solutions behave as $t \to \infty$.

1. $y' + 3y = t + e^{-2t}$

2. $y' - 2y = t^2 e^{2t}$

3. $y' + y = te^{-t} + 1$

4. $y' + \frac{1}{t}y = 3\cos(2t), \quad t > 0$

5. $y' - 2y = 3e^t$

6. $ty' - y = t^2 e^{-t}, \quad t > 0$

7. $y' + y = 5\sin(2t)$

8. $2y' + y = 3t^2$

In each of Problems 9 through 12, find the solution of the given initial value problem.

9. $y' - y = 2te^{2t}, \quad y(0) = 1$

10. $y' + 2y = te^{-2t}, \quad y(1) = 0$

11. $y' + \frac{2}{t}y = \frac{\cos t}{t^2}, \quad y(\pi) = 0, \quad t > 0$

12. $ty' + (t + 1)y = t, \quad y(\ln 2) = 1, \quad t > 0$

In each of Problems 13 and 14:

a. Draw a direction field for the given differential equation. How do solutions appear to behave as t becomes large? Does the behavior depend on the choice of the initial value a? Let a_0 be the value of a for which the transition from one type of behavior to another occurs. Estimate the value of a_0.
b. Solve the initial value problem and find the critical value a_0 exactly.
c. Describe the behavior of the solution corresponding to the initial value a_0.

13. $y' - \frac{1}{2}y = 2\cos t, \quad y(0) = a$

14. $3y' - 2y = e^{-\pi t/2}, \quad y(0) = a$

In each of Problems 15 and 16:

G **a.** Draw a direction field for the given differential equation. How do solutions appear to behave as $t \to 0$? Does the behavior depend on the choice of the initial value a? Let a_0 be the critical value of a, that is, the initial value such that the solutions for $a < a_0$ and the solutions for $a > a_0$ have different behaviors as $t \to \infty$. Estimate the value of a_0.

b. Solve the initial value problem and find the critical value a_0 exactly.

c. Describe the behavior of the solution corresponding to the initial value a_0.

15. $ty' + (t+1)y = 2te^{-t}, \quad y(1) = a, \quad t > 0$

16. $(\sin t)y' + (\cos t)y = e^t, \quad y(1) = a, \quad 0 < t < \pi$

G **17.** Consider the initial value problem

$$y' + \frac{1}{2}y = 2\cos t, \quad y(0) = -1.$$

Find the coordinates of the first local maximum point of the solution for $t > 0$.

N **18.** Consider the initial value problem

$$y' + \frac{2}{3}y = 1 - \frac{1}{2}t, \quad y(0) = y_0.$$

Find the value of y_0 for which the solution touches, but does not cross, the t-axis.

19. Consider the initial value problem

$$y' + \frac{1}{4}y = 3 + 2\cos(2t), \quad y(0) = 0.$$

a. Find the solution of this initial value problem and describe its behavior for large t.

N **b.** Determine the value of t for which the solution first intersects the line $y = 12$.

20. Find the value of y_0 for which the solution of the initial value problem

$$y' - y = 1 + 3\sin t, \quad y(0) = y_0$$

remains finite as $t \to \infty$.

21. Consider the initial value problem

$$y' - \frac{3}{2}y = 3t + 2e^t, \quad y(0) = y_0.$$

Find the value of y_0 that separates solutions that grow positively as $t \to \infty$ from those that grow negatively. How does the solution that corresponds to this critical value of y_0 behave as $t \to \infty$?

22. Show that all solutions of $2y' + ty = 2$ [equation (41) of the text] approach a limit as $t \to \infty$, and find the limiting value.
Hint: Consider the general solution, equation (47). Show that the first term in the solution (47) is indeterminate with form $0 \cdot \infty$. Then, use l'Hôpital's rule to compute the limit as $t \to \infty$.

23. Show that if a and λ are positive constants, and b is any real number, then every solution of the equation

$$y' + ay = be^{-\lambda t}$$

has the property that $y \to 0$ as $t \to \infty$.
Hint: Consider the cases $a = \lambda$ and $a \neq \lambda$ separately.

In each of Problems 24 through 27, construct a first-order linear differential equation whose solutions have the required behavior as $t \to \infty$. Then solve your equation and confirm that the solutions do indeed have the specified property.

24. All solutions have the limit 3 as $t \to \infty$.

25. All solutions are asymptotic to the line $y = 3 - t$ as $t \to \infty$.

26. All solutions are asymptotic to the line $y = 2t - 5$ as $t \to \infty$.

27. All solutions approach the curve $y = 4 - t^2$ as $t \to \infty$.

28. **Variation of Parameters.** Consider the following method of solving the general linear equation of first order:

$$y' + p(t)y = g(t). \tag{48}$$

a. If $g(t) = 0$ for all t, show that the solution is

$$y = A\exp\left(-\int p(t)\,dt\right), \tag{49}$$

where A is a constant.

b. If $g(t)$ is not everywhere zero, assume that the solution of equation (48) is of the form

$$y = A(t)\exp\left(-\int p(t)\,dt\right), \tag{50}$$

where A is now a function of t. By substituting for y in the given differential equation, show that $A(t)$ must satisfy the condition

$$A'(t) = g(t)\exp\left(\int p(t)\,dt\right). \tag{51}$$

c. Find $A(t)$ from equation (51). Then substitute for $A(t)$ in equation (50) and determine y. Verify that the solution obtained in this manner agrees with that of equation (33) in the text. This technique is known as the method of **variation of parameters**; it is discussed in detail in Section 3.6 in connection with second-order linear equations.

In each of Problems 29 and 30, use the method of Problem 28 to solve the given differential equation.

29. $y' - 2y = t^2e^{2t}$

30. $y' + \frac{1}{t}y = \cos(2t), \quad t > 0$

2.2 Separable Differential Equations

In Section 1.2 we used a process of direct integration to solve first-order linear differential equations of the form

$$\frac{dy}{dt} = ay + b, \tag{1}$$

where a and b are constants. We will now show that this process is actually applicable to a much larger class of nonlinear differential equations.

We will use x, rather than t, to denote the independent variable in this section for two reasons. In the first place, different letters are frequently used for the variables in a differential equation, and you should not become too accustomed to using a single pair. In particular, x often occurs as the independent variable. Further, we want to reserve t for another purpose later in the section.

The general first-order differential equation is

$$\frac{dy}{dx} = f(x, y). \tag{2}$$

Linear differential equations were considered in the preceding section, but if equation (2) is nonlinear, then there is no universally applicable method for solving the equation. Here, we consider a subclass of first-order equations that can be solved by direct integration.

To identify this class of equations, we first rewrite equation (2) in the form

$$M(x, y) + N(x, y)\frac{dy}{dx} = 0. \tag{3}$$

It is always possible to do this by setting $M(x, y) = -f(x, y)$ and $N(x, y) = 1$, but there may be other ways as well. When M is a function of x only and N is a function of y only, then equation (3) becomes

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

Such an equation is said to be **separable**, because if it is written in the **differential form**

$$M(x)\,dx + N(y)\,dy = 0, \tag{5}$$

then, if you wish, terms involving each variable may be placed on opposite sides of the equation. The differential form (5) is also more symmetric and tends to suppress the distinction between independent and dependent variables.

A separable equation can be solved by integrating the functions M and N. We illustrate the process by an example and then discuss it in general for equation (4).

EXAMPLE 1

Show that the equation

$$\frac{dy}{dx} = \frac{x^2}{1 - y^2} \tag{6}$$

is separable, and then find an equation for its integral curves.

Solution:

If we write equation (6) as

$$-x^2 + (1 - y^2)\,\frac{dy}{dx} = 0, \tag{7}$$

then it has the form (4) and is therefore separable. Recall from calculus that if y is a function of x, then by the chain rule,

$$\frac{d}{dx}f(y) = \frac{d}{dy}f(y)\,\frac{dy}{dx} = f'(y)\frac{dy}{dx}.$$

For example, if $f(y) = y - y^3/3$, then

$$\frac{d}{dx}\left(y - \frac{y^3}{3}\right) = (1 - y^2)\frac{dy}{dx}.$$

Thus the second term in equation (7) is the derivative with respect to x of $y - y^3/3$, and the first term is the derivative of $-x^3/3$. Thus equation (7) can be written as

$$\frac{d}{dx}\left(-\frac{x^3}{3}\right) + \frac{d}{dx}\left(y - \frac{y^3}{3}\right) = 0,$$

or

$$\frac{d}{dx}\left(-\frac{x^3}{3} + y - \frac{y^3}{3}\right) = 0.$$

Therefore, by integrating (and multiplying the result by 3), we obtain

$$-x^3 + 3y - y^3 = c, \tag{8}$$

where c is an arbitrary constant.

Equation (8) is an equation for the integral curves of equation (6). A direction field and several integral curves are shown in Figure 2.2.1. Any differentiable function $y = \phi(x)$ that satisfies equation (8) is a solution of equation (6). An equation of the integral curve passing through a particular point (x_0, y_0) can be found by substituting x_0 and y_0 for x and y, respectively, in equation (8) and determining the corresponding value of c.

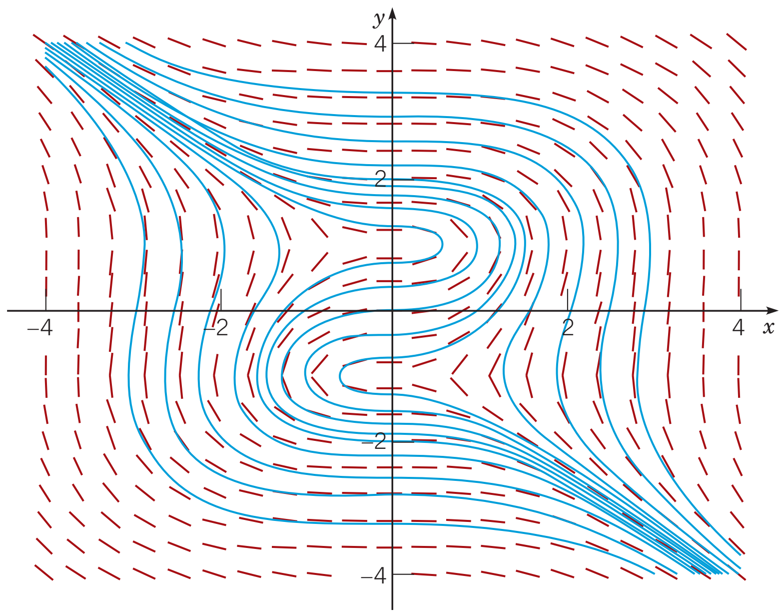


FIGURE 2.2.1 Direction field and integral curves of $y' = x^2/(1 - y^2)$.

Essentially the same procedure can be followed for any separable equation. Returning to equation (4), let H_1 and H_2 be any antiderivatives of M and N, respectively. Thus

$$H_1'(x) = M(x), \quad H_2'(y) = N(y), \tag{9}$$

and equation (4) becomes

$$H_1'(x) + H_2'(y)\frac{dy}{dx} = 0. \tag{10}$$

If y is regarded as a function of x, then according to the chain rule,

$$H_2'(y)\frac{dy}{dx} = \frac{d}{dy}H_2(y)\frac{dy}{dx} = \frac{d}{dx}H_2(y). \tag{11}$$

Consequently, we can write equation (10) as

$$\frac{d}{dx}\big(H_1(x) + H_2(y)\big) = 0. \tag{12}$$

By integrating equation (12) with respect to x, we obtain

$$H_1(x) + H_2(y) = c, \tag{13}$$

where c is an arbitrary constant. Any differentiable function $y = \phi(x)$ that satisfies equation (13) is a solution of equation (4); in other words, equation (13) defines the solution implicitly rather than explicitly. In practice, equation (13) is usually obtained from equation (5) by integrating the first term with respect to x and the second term with respect to y. The justification for this is the argument that we have just given.

The differential equation (4), together with an initial condition

$$y(x_0) = y_0, \tag{14}$$

forms an initial value problem. To solve this initial value problem, we must determine the appropriate value for the constant c in equation (13). We do this by setting $x = x_0$ and $y = y_0$ in equation (13) with the result that

$$c = H_1(x_0) + H_2(y_0). \tag{15}$$

Substituting this value of c in equation (13) and noting that

$$H_1(x) - H_1(x_0) = \int_{x_0}^{x} M(s)ds, \quad H_2(y) - H_2(y_0) = \int_{y_0}^{y} N(s)ds,$$

we obtain

$$\int_{x_0}^{x} M(s)ds + \int_{y_0}^{y} N(s)ds = 0. \tag{16}$$

Equation (16) is an implicit representation of the solution of the differential equation (4) that also satisfies the initial condition (14). Bear in mind that to determine an explicit formula for the solution, you need to solve equation (16) for y as a function of x. Unfortunately, it is often impossible to do this analytically; in such cases you can resort to numerical methods to find approximate values of y for given values of x.

EXAMPLE 2

Solve the initial value problem

$$\frac{dy}{dx} = \frac{3x^2 + 4x + 2}{2(y-1)}, \quad y(0) = -1, \tag{17}$$

and determine the interval in which the solution exists.

Solution:

The differential equation can be written as

$$2(y-1)dy = (3x^2 + 4x + 2)dx.$$

Integrating the left-hand side with respect to y and the right-hand side with respect to x gives

$$y^2 - 2y = x^3 + 2x^2 + 2x + c, \tag{18}$$

where c is an arbitrary constant. To determine the solution satisfying the prescribed initial condition, we substitute $x = 0$ and $y = -1$ in equation (18), obtaining $c = 3$. Hence the solution of the initial value problem is given implicitly by

$$y^2 - 2y = x^3 + 2x^2 + 2x + 3. \tag{19}$$

To obtain the solution explicitly, we must solve equation (19) for y in terms of x. That is a simple matter in this case, since equation (19) is quadratic in y, and we obtain

$$y = 1 \pm \sqrt{x^3 + 2x^2 + 2x + 4}. \tag{20}$$

Equation (20) gives two solutions of the differential equation, only one of which, however, satisfies the given initial condition. This is the solution corresponding to the minus sign in equation (20), so we finally obtain

$$y = \phi(x) = 1 - \sqrt{x^3 + 2x^2 + 2x + 4} \tag{21}$$

as the solution of the initial value problem (15). Note that if we choose the plus sign by mistake in equation (20), then we obtain the solution of the same differential equation that satisfies the initial condition $y(0) = 3$. Finally, to determine the interval in which the solution (21) is valid, we must find the interval in which the quantity under the radical is positive. The only real zero of this expression is $x = -2$, so the desired interval is $x > -2$. Some integral curves of the differential equation are shown in Figure 2.2.2. The green curve passes through the point $(0, -1)$ and thus is the solution of the initial value problem (15). Observe that the boundary of the interval of validity of the solution (21) is determined by the point $(-2, 1)$ at which the tangent line is vertical.

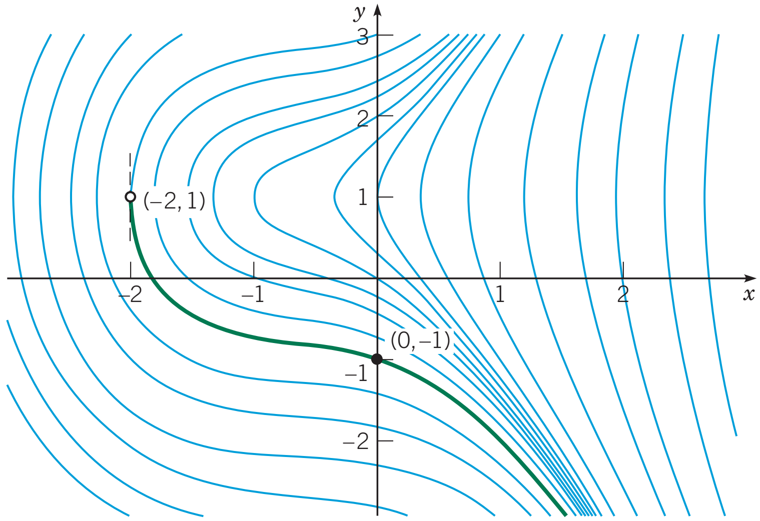


FIGURE 2.2.2 Integral curves of $y' = (3x^2 + 4x + 2)/2(y - 1)$; the solution satisfying $y(0) = -1$ is shown in green and is valid for $x > -2$.

EXAMPLE 3

Solve the separable differential equation

$$\frac{dy}{dx} = \frac{4x - x^3}{4 + y^3} \tag{22}$$

and draw graphs of several integral curves. Also find the solution passing through the point $(0, 1)$ and determine its interval of validity.

Solution:

Rewriting equation (22) as

$$(4 + y^3)dy = (4x - x^3)dx,$$

integrating each side, multiplying by 4, and rearranging the terms, we obtain

$$y^4 + 16y + x^4 - 8x^2 = c, \tag{23}$$

where c is an arbitrary constant. Any differentiable function $y = \phi(x)$ that satisfies equation (23) is a solution of the differential equation (22). Graphs of equation (23) for several values of c are shown in Figure 2.2.3.

To find the particular solution passing through $(0, 1)$, we set $x = 0$ and $y = 1$ in equation (23) with the result that $c = 17$. Thus the solution in question is given implicitly by

$$y^4 + 16y + x^4 - 8x^2 = 17. \tag{24}$$

It is shown by the green curve in Figure 2.2.3. The interval of validity of this solution extends on either side of the initial point as long as the function remains differentiable. From the figure we see that the interval ends when we reach points where the tangent line is vertical. It follows from the differential equation (22) that these are points where $4 + y^3 = 0$, or $y = (-4)^{1/3} \cong -1.5874$. From equation (24) the corresponding values of x are $x \cong \pm 3.3488$. These points are marked on the graph in Figure 2.2.3.

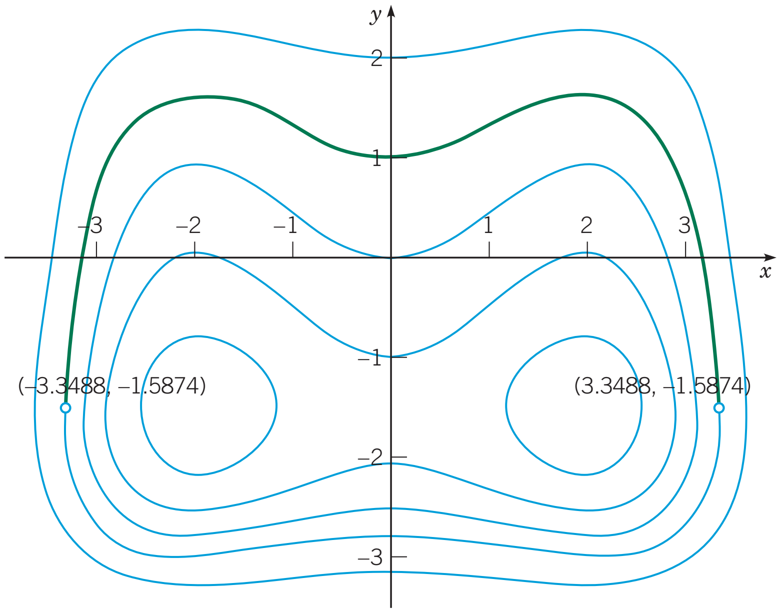


FIGURE 2.2.3 Integral curves of $y' = (4x - x^3)/(4 + y^3)$. The solution passing through $(0, 1)$ is shown by the green curve.

Note 1: Sometimes a differential equation of the form (2):

$$\frac{dy}{dx} = f(x, y)$$

has a constant solution $y = y_0$. Such a solution is usually easy to find because if $f(x, y_0) = 0$ for some value y_0 and for all x, then the constant function $y = y_0$ is a solution of the differential equation (2). For example, the equation

$$\frac{dy}{dx} = \frac{(y-3)\cos x}{1 + 2y^2} \tag{25}$$

has the constant solution $y = 3$. Other solutions of this equation can be found by separating the variables and integrating.

Note 2: The investigation of a first-order nonlinear differential equation can sometimes be facilitated by regarding both x and y as functions of a third variable t. Then

$$\frac{dy}{dx} = \frac{dy/dt}{dx/dt}. \tag{26}$$

If the differential equation is

$$\frac{dy}{dx} = \frac{F(x, y)}{G(x, y)}, \tag{27}$$

then, by comparing numerators and denominators in equations (26) and (27), we obtain the system

$$\frac{dx}{dt} = G(x, y), \quad \frac{dy}{dt} = F(x, y). \tag{28}$$

At first sight it may seem unlikely that a problem will be simplified by replacing a single equation by a pair of equations, but in fact, the system (28) may well be more amenable to investigation than the single equation (27). Chapter 9 is devoted to nonlinear systems of the form (28).

Note 3: In Example 2 it was not difficult to solve explicitly for y as a function of x. However, this situation is exceptional, and often it will be better to leave the solution in implicit form, as in Examples 1 and 3. Thus, in the problems below and in other sections where nonlinear equations appear, the words "solve the following differential equation" mean to find the solution explicitly if it is convenient to do so, but otherwise to find an equation defining the solution implicitly.

Problems

In each of Problems 1 through 8, solve the given differential equation.

1. $y' = \dfrac{x^2}{y}$
2. $y' + y^2 \sin x = 0$
3. $y' = \cos^2(x)\cos^2(2y)$
4. $xy' = (1 - y^2)^{1/2}$
5. $\dfrac{dy}{dx} = \dfrac{x - e^{-x}}{y + e^y}$
6. $\dfrac{dy}{dx} = \dfrac{x^2}{1 + y^2}$
7. $\dfrac{dy}{dx} = \dfrac{y}{x}$
8. $\dfrac{dy}{dx} = \dfrac{-x}{y}$

In each of Problems 9 through 16:

a. Find the solution of the given initial value problem in explicit form.
G **b.** Plot the graph of the solution.
c. Determine (at least approximately) the interval in which the solution is defined.

9. $y' = (1 - 2x)y^2, \quad y(0) = -1/6$
10. $y' = (1 - 2x)/y, \quad y(1) = -2$
11. $x\,dx + ye^{-x}dy = 0, \quad y(0) = 1$
12. $dr/d\theta = r^2/\theta, \quad r(1) = 2$
13. $y' = xy^3(1 + x^2)^{-1/2}, \quad y(0) = 1$
14. $y' = 2x/(1 + 2y), \quad y(2) = 0$
15. $y' = (3x^2 - e^x)/(2y - 5), \quad y(0) = 1$
16. $\sin(2x)\,dx + \cos(3y)\,dy = 0, \quad y(\pi/2) = \pi/3$

Some of the results requested in Problems 17 through 22 can be obtained either by solving the given equations analytically or by plotting numerically generated approximations to the solutions. Try to form an opinion about the advantages and disadvantages of each approach.

G **17.** Solve the initial value problem

$$y' = \frac{1 + 3x^2}{3y^2 - 6y}, \quad y(0) = 1$$

and determine the interval in which the solution is valid.
Hint: To find the interval of definition, look for points where the integral curve has a vertical tangent.

G **18.** Solve the initial value problem

$$y' = \frac{3x^2}{3y^2 - 4}, \quad y(1) = 0$$

and determine the interval in which the solution is valid.
Hint: To find the interval of definition, look for points where the integral curve has a vertical tangent.

G **19.** Solve the initial value problem

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

and determine where the solution attains its minimum value.

G **20.** Solve the initial value problem

$$y' = \frac{2 - e^x}{3 + 2y}, \quad y(0) = 0$$

and determine where the solution attains its maximum value.

G **21.** Consider the initial value problem

$$y' = \frac{ty(4 - y)}{3}, \quad y(0) = y_0.$$

a. Determine how the behavior of the solution as t increases depends on the initial value y_0.
b. Suppose that $y_0 = 0.5$. Find the time T at which the solution first reaches the value 3.98.

G **22.** Consider the initial value problem

$$y' = \frac{ty(4 - y)}{1 + t}, \quad y(0) = y_0 > 0.$$

a. Determine how the solution behaves as $t \to \infty$.
b. If $y_0 = 2$, find the time T at which the solution first reaches the value 3.99.
c. Find the range of initial values for which the solution lies in the interval $3.99 < y < 4.01$ by the time $t = 2$.

23. Solve the equation

$$\frac{dy}{dx} = \frac{ay + b}{cy + d},$$

where a, b, c, and d are constants.

24. Use separation of variables to solve the differential equation

$$\frac{dQ}{dt} = r(a + bQ), \quad Q(0) = Q_0,$$

where a, b, r, and Q_0 are constants. Determine how the solution behaves as $t \to \infty$

Homogeneous Equations. If the right-hand side of the equation $dy/dx = f(x, y)$ can be expressed as a function of the ratio y/x only, then the equation is said to be homogeneous.[1] Such equations can always be transformed into separable equations by a change of the dependent variable. Problem 25 illustrates how to solve first-order homogeneous equations.

[1]The word "homogeneous" has different meanings in different mathematical contexts. The homogeneous equations considered here have nothing to do with the homogeneous equations that will occur in Chapter 3 and elsewhere.

N **25.** Consider the equation

$$\frac{dy}{dx} = \frac{y - 4x}{x - y}. \tag{29}$$

a. Show that equation (29) can be rewritten as

$$\frac{dy}{dx} = \frac{(y/x) - 4}{1 - (y/x)}; \tag{30}$$

thus equation (29) is homogeneous.

b. Introduce a new dependent variable v so that $v = y/x$, or $y = xv(x)$. Express dy/dx in terms of x, v, and dv/dx.

c. Replace y and dy/dx in equation (30) by the expressions from part b that involve v and dv/dx. Show that the resulting differential equation is

$$v + x\frac{dv}{dx} = \frac{v - 4}{1 - v},$$

or

$$x\frac{dv}{dx} = \frac{v^2 - 4}{1 - v}. \tag{31}$$

Observe that equation (31) is separable.

d. Solve equation (31), obtaining v implicitly in terms of x.

e. Find the solution of equation (29) by replacing v by y/x in the solution in part d.

f. Draw a direction field and some integral curves for equation (29). Recall that the right-hand side of equation (29) actually depends only on the ratio y/x. This means that integral curves have the same slope at all points on any given straight line through the origin, although the slope changes from one line to another. Therefore, the direction field and the integral curves are symmetric with respect to the origin. Is this symmetry property evident from your plot?

The method outlined in Problem 25 can be used for any homogeneous equation. That is, the substitution $y = xv(x)$ transforms a homogeneous equation into a separable equation. The latter equation can be solved by direct integration, and then replacing v by y/x gives the solution to the original equation. In each of Problems 26 through 31:

a. Show that the given equation is homogeneous.

b. Solve the differential equation.

G **c.** Draw a direction field and some integral curves. Are they symmetric with respect to the origin?

26. $\dfrac{dy}{dx} = \dfrac{x^2 + xy + y^2}{x^2}$

27. $\dfrac{dy}{dx} = \dfrac{x^2 + 3y^2}{2xy}$

28. $\dfrac{dy}{dx} = \dfrac{4y - 3x}{2x - y}$

29. $\dfrac{dy}{dx} = -\dfrac{4x + 3y}{2x + y}$

30. $\dfrac{dy}{dx} = \dfrac{x^2 - 3y^2}{2xy}$

31. $\dfrac{dy}{dx} = \dfrac{3y^2 - x^2}{2xy}$

2.3 Modeling with First-Order Differential Equations

Differential equations are of interest to nonmathematicians primarily because of the possibility of using them to investigate a wide variety of problems in the physical, biological, and social sciences. One reason for this is that mathematical models and their solutions lead to equations relating the variables and parameters in the problem. These equations often enable you to make predictions about how the natural process will behave in various circumstances. It is often easy to vary parameters in the mathematical model over wide ranges, whereas this may be very time-consuming or expensive, if not impossible, in an experimental setting. Nevertheless, mathematical modeling and experiment or observation are both critically important and have somewhat complementary roles in scientific investigations. Mathematical models are validated by comparison of their predictions with experimental results. On the other hand, mathematical analyses may suggest the most promising directions to explore experimentally, and they may indicate fairly precisely what experimental data will be most helpful.

In Sections 1.1 and 1.2 we formulated and investigated a few simple mathematical models. We begin by recapitulating and expanding on some of the conclusions reached in those sections. Regardless of the specific field of application, there are three identifiable steps that are always present in the process of mathematical modeling.

Step 1: Construction of the Model. In this step the physical situation is translated into mathematical terms, often using the steps listed at the end of Section 1.1. Perhaps most critical at this stage is to state clearly the physical principle(s) that are believed to govern the process. For example, it has been observed that in some circumstances heat passes from a warmer to a cooler body at a rate proportional to the temperature difference, that objects move about in accordance with Newton's laws of motion, and that isolated insect populations grow at a rate proportional to the current population. Each of these statements involves a rate of change (derivative) and consequently, when expressed mathematically, leads to a differential equation. The differential equation is a mathematical model of the process.

It is important to realize that the mathematical equations are almost always only an approximate description of the actual process. For example, bodies moving at speeds comparable to the speed of light are not governed by Newton's laws, insect populations do not grow indefinitely as stated because of eventual lack of food or space, and heat transfer is affected by factors other than the temperature difference. Thus you should always be aware of the limitations of the model so that you will use it only when it is reasonable to believe that it is accurate. Alternatively, you can adopt the point of view that the mathematical equations exactly describe the operation of a simplified physical model, which has been constructed (or conceived of) so as to embody the most important features of the actual process. Sometimes, the process of mathematical modeling involves the conceptual replacement of a discrete process by a continuous one. For instance, the number of members in an insect population changes by discrete amounts; however, if the population is large, it seems reasonable to consider it as a continuous variable and even to speak of its derivative.

Step 2: Analysis of the Model. Once the problem has been formulated mathematically, you are often faced with the problem of solving one or more differential equations or, failing that, of finding out as much as possible about the properties of the solution. It may happen that this mathematical problem is quite difficult, and if so, further approximations may be indicated at this stage to make the problem mathematically tractable. For example, a nonlinear equation may be approximated by a linear one, or a slowly varying coefficient may be replaced by a constant. Naturally, any such approximations must also be examined from the physical point of view to make sure that the simplified mathematical problem still reflects the essential features of the physical process under investigation. At the same time, an intimate knowledge of the physics of the problem may suggest reasonable mathematical approximations that will make the mathematical problem more amenable to analysis. This interplay of understanding of physical phenomena and knowledge of mathematical techniques and their limitations is characteristic of applied mathematics at its best, and it is indispensable in successfully constructing useful mathematical models of intricate physical processes.

Step 3: Comparison with Experiment or Observation. Finally, having obtained the solution (or at least some information about it), you must interpret this information in the context in which the problem arose. In particular, you should always check that the mathematical solution appears physically reasonable. If possible, calculate the values of the solution at selected points and compare them with experimentally observed values. Or ask whether the behavior of the solution after a long time is consistent with observations. Or examine the solutions corresponding to certain special values of parameters in the problem. Of course, the fact that the mathematical solution appears to be reasonable does not guarantee that it is correct. However, if the predictions of the mathematical model are seriously inconsistent with observations of the physical system it purports to describe, this suggests that errors have been made in solving the mathematical problem, that the mathematical model itself needs refinement, or that observations must be made with greater care.

The examples in this section are typical of applications in which first-order differential equations arise.

EXAMPLE 1 | Mixing

At time $t = 0$ a tank contains Q_0 lb of salt dissolved in 100 gal of water; see Figure 2.3.1. Assume that water containing $\frac{1}{4}$ lb of salt per gallon is entering the tank at a rate of r gal/min and that the well-stirred mixture is draining from the tank at the same rate. Set up the initial value problem that describes this flow process. Find the amount of salt $Q(t)$ in the tank at any time, and also find the limiting amount Q_L that is present after a very long time. If $r = 3$ and $Q_0 = 2Q_L$, find the time T after which the salt level is within 2% of Q_L. Also find the flow rate that is required if the value of T is not to exceed 45 min.

Solution:

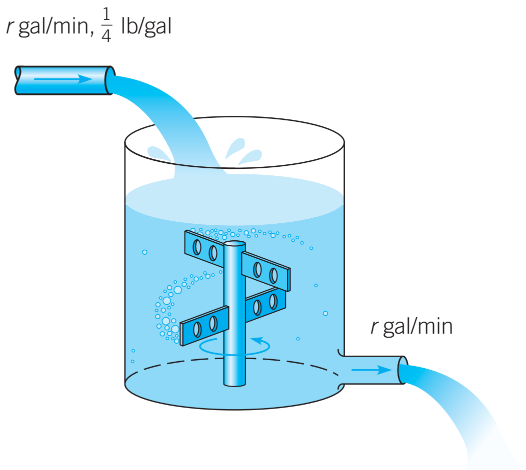


FIGURE 2.3.1 The water tank in Example 1.

We assume that salt is neither created nor destroyed in the tank. Therefore, variations in the amount of salt are due solely to the flows in and out of the tank. More precisely, the rate of change of salt in the tank, dQ/dt, is equal to the rate at which salt is flowing in minus the rate at which it is flowing out. In symbols,

$$\frac{dQ}{dt} = \text{rate in} - \text{rate out}. \tag{1}$$

The rate at which salt enters the tank is the concentration $\frac{1}{4}$ lb/gal times the flow rate r gal/min, or $r/4$ lb/min. To find the rate at which salt leaves the tank, we need to multiply the concentration of salt in the tank by the rate of outflow, r gal/min. Since the rates of flow in and out are equal, the volume of water in the tank remains constant at 100 gal, and since the mixture is "well-stirred," the concentration throughout the tank is the same, namely, $Q(t)/100$ lb/gal. Therefore, the rate at which salt leaves the tank is $rQ(t)/100$ lb/min. Thus the differential equation governing this process is

$$\frac{dQ}{dt} = \frac{r}{4} - \frac{rQ}{100}. \tag{2}$$

The initial condition is

$$Q(0) = Q_0. \tag{3}$$

Upon thinking about the problem physically, we might anticipate that eventually the mixture originally in the tank will be essentially replaced by the mixture flowing in, whose concentration is $\frac{1}{4}$ lb/gal. Consequently, we might expect that ultimately the amount of salt in the tank would be very close to 25 lb. We can also find the limiting amount $Q_L = 25$ by setting dQ/dt equal to zero in equation (2) and solving the resulting algebraic equation for Q.

To solve the initial value problem (2), (3) analytically, note that equation (2) is linear. (It is also separable, see Problem 24 in Section 2.2.) Rewriting the differential equation (2) in the standard form for a linear differential equation, we have

$$\frac{dQ}{dt} + \frac{rQ}{100} = \frac{r}{4}. \tag{4}$$

Thus the integrating factor is $e^{rt/100}$ and the general solution is

$$Q(t) = 25 + ce^{-rt/100}, \tag{5}$$

where c is an arbitrary constant. To satisfy the initial condition (3), we must choose $c = Q_0 - 25$. Therefore, the solution of the initial value problem (2), (3) is

$$Q(t) = 25 + (Q_0 - 25)e^{-rt/100}, \tag{6}$$

or

$$Q(t) = 25(1 - e^{-rt/100}) + Q_0 e^{-rt/100}. \tag{7}$$

From either form of the solution, (6) or (7), you can see that $Q(t) \to 25$ (lb) as $t \to \infty$, so the limiting value Q_L is 25, confirming our physical intuition.

Further, $Q(t)$ approaches the limit more rapidly as r increases. In interpreting the solution (7), note that the second term on the right-hand side is the portion of the original salt that remains at time t, while the first term gives the amount of salt in the tank as a consequence of the flow processes. Plots of the solution for $r = 3$ and for several values of Q_0 are shown in Figure 2.3.2.

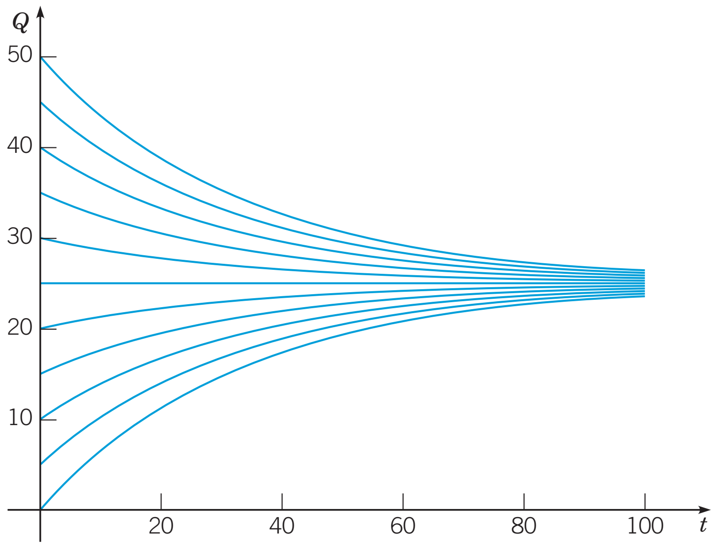


FIGURE 2.3.2 Solutions of the initial value problem (2): $dQ/dt = r/4 - rQ/100$, $Q(0) = Q_0$ for $r = 3$ and several values of Q_0.

Now suppose that $r = 3$ and $Q_0 = 2Q_L = 50$; then equation (6) becomes

$$Q(t) = 25 + 25e^{-0.03t}. \tag{8}$$

Since 2% of 25 is 0.5, we wish to find the time T at which $Q(t)$ has the value 25.5. Substituting $t = T$ and $Q = 25.5$ in equation (8) and solving for T, we obtain

$$T = \frac{\ln(50)}{0.03} \cong 130.4 \text{ (min)}. \tag{9}$$

To determine r so that $T = 45$, return to equation (6), set $t = 45$, $Q_0 = 50$, $Q(t) = 25.5$, and solve for r. The result is

$$r = \frac{100}{45} \ln 50 \cong 8.69 \text{ gal/min}. \tag{10}$$

Since this example is hypothetical, the validity of the model is not in question. If the flow rates are as stated, and if the concentration of salt in the tank is uniform, then the differential equation (1) is an accurate description of the flow process. Although this particular example has no special significance, models of this kind are often used in problems involving a pollutant in a lake, or a drug in an organ of the body, for example, rather than a tank of salt water. In such cases the flow rates may not be easy to determine or may vary with time. Similarly, the concentration may be far from uniform in some cases. Finally, the rates of inflow and outflow may be different, which means that the variation of the amount of liquid in the problem must also be taken into account.

EXAMPLE 2 | Compound Interest

Suppose that a sum of money, S_0, is deposited in a bank or money fund that pays interest at an annual rate r. The value $S(t)$ of the investment at any time t depends on the frequency with which interest is compounded as well as on the interest rate. Financial institutions have various policies concerning compounding: some compound monthly, some weekly, and some even daily. Assume that compounding takes place *continuously*. Set up an initial value problem that describes the growth of the investment.

Solution:

The rate of change of the value of the investment is dS/dt, and this quantity is equal to the rate at which interest accrues, which is the interest rate r times the current value of the investment $S(t)$. Thus

$$\frac{dS}{dt} = rS \tag{11}$$

is the differential equation that governs the process. If we let t denote the time, in years, since the original deposit, the corresponding initial condition is

$$S(0) = S_0. \tag{12}$$

Then the solution of the initial value problem (8) gives the balance $S(t)$ in the account at any time t. This initial value problem is readily solved, since the differential equation (11) is both linear and separable. Consequently, by solving equations (11) and (12), we find that

$$S(t) = S_0 e^{rt}. \tag{13}$$

Thus a bank account with continuously compounding interest grows exponentially.

The model in Example 2 is easily extended to situations involving deposits or withdrawals in addition to the accrual of interest, dividends, or annual capital gains. If we assume that the deposits or withdrawals take place at a constant rate k, then equation (11) is replaced by

$$\frac{dS}{dt} = rS + k,$$

or, in standard form,

$$\frac{dS}{dt} - rS = k, \tag{14}$$

where k is positive for deposits and negative for withdrawals.

Equation (14) is linear with the integrating factor e^{-rt}, so its general solution is

$$S(t) = ce^{rt} - \frac{k}{r},$$

where c is an arbitrary constant. To satisfy the initial condition (12), we must choose $c = S_0 + k/r$. Thus the solution of the initial value problem (10), (8) is

$$S(t) = S_0 e^{rt} + \frac{k}{r}(e^{rt} - 1). \tag{15}$$

The first term in expression (15) is the part of $S(t)$ that is due to the return accumulated on the initial amount S_0, and the second term is the part that is due to the deposit or withdrawal rate k.

The advantage of stating the problem in this general way without specific values for S_0, r, or k lies in the generality of the resulting formula (15) for $S(t)$. With this formula we can readily compare the results of different investment programs or different rates of return.

For instance, suppose that one opens an individual retirement account (IRA) at age 25 and makes annual investments of \$2000 thereafter in a continuous manner. Assuming a rate of return of 8%, what will be the balance in the IRA at age 65? We have $S_0 = 0$, $r = 0.08$, and $k = \$2000$, and we wish to determine $S(40)$. From equation (15) we have

$$S(40) = 25{,}000(e^{3.2} - 1) = \$588{,}313. \tag{16}$$

It is interesting to note that the total amount invested is \$80,000, so the remaining amount of \$508,313 results from the accumulated return on the investment. The balance after 40 years is also fairly sensitive to the assumed rate. For instance, $S(40) = \$508{,}948$ if $r = 0.075$ and $S(40) = \$681{,}508$ if $r = 0.085$.

Let us now examine the assumptions that have gone into the model. First, we have assumed that the return is compounded continuously and that additional capital is invested continuously. Neither of these is true in an actual financial situation. We have also assumed that the return rate r is constant for the entire period involved, whereas in fact it is likely to fluctuate considerably. Although we cannot reliably predict future rates, we can use solution (15) to determine the approximate effect of different rate projections. It is also possible to consider r and k in equation (14) to be functions of t rather than constants; in that case, of course, the solution may be much more complicated than equation (15).

The initial value problem (10), (8) and the solution (15) can also be used to analyze a number of other financial situations, including annuities, mortgages, and automobile loans.

Let us now compare the results from the model with continuously compounded interest (and no other deposits or withdrawals) with the corresponding situation in which compounding occurs at finite time intervals. If interest is compounded once a year, then after t years

$$S(t) = S_0(1+r)^t.$$

If interest is compounded twice a year, then at the end of 6 months the value of the investment is $S_0(1+(r/2))$, and at the end of 1 year it is $S_0(1+r/2)^2$. Thus, after t years, we have

$$S(t) = S_0\left(1+\frac{r}{2}\right)^{2t}.$$

In general, if interest is compounded m times per year, then

$$S(t) = S_0\left(1+\frac{r}{m}\right)^{mt}. \tag{17}$$

The relation between formulas (13) and (17) is clarified if we recall from calculus that

$$\lim_{m\to\infty} S_0\left(1+\frac{r}{m}\right)^{mt} = S_0 e^{rt}.$$

The same model applies equally well to more general investments in which dividends and perhaps capital gains can also accumulate, as well as interest. In recognition of this fact, we will from now on refer to r as the rate of return.

Table 2.3.1 shows the effect of changing the frequency of compounding for a return rate r of 8%. The second and third columns are calculated from equation (17) for quarterly and daily compounding, respectively, and the fourth column is calculated from equation (13) for continuous compounding. The results show that the frequency of compounding is not particularly important in most cases. For example, during a 10-year period the difference between quarterly and continuous compounding is \$17.50 per \$1000 invested, or less than \$2/year. The difference would be somewhat greater for higher rates of return and less for lower rates. From the first row in the table, we see that for the return rate $r = 8\%$, the annual yield for quarterly compounding is 8.24% and for daily or continuous compounding it is 8.33%.

TABLE 2.3.1 Growth of Capital at a Return Rate $r = 8\%$ for Several Modes of Compounding

	$S(t)/S(t_0)$ From Equation (17)		$S(t)/S(t_0)$
Years	$m = 4$	$m = 365$	From Equation (13)
1	1.0824	1.0833	1.0833
2	1.1717	1.1735	1.1735
5	1.4859	1.4918	1.4918
10	2.2080	2.2253	2.2255
20	4.8754	4.9522	4.9530
30	10.7652	11.0203	11.0232
40	23.7699	24.5239	24.5325

EXAMPLE 3 | Chemicals in a Pond

Consider a pond that initially contains 10 million gallons of fresh water. Water containing an undesirable chemical flows into the pond at the rate of 5 million gallons per year, and the mixture in the pond flows out at the same rate. The concentration $\gamma(t)$ of chemical in the incoming water varies periodically with time according to the expression $\gamma(t) = 2 + \sin(2t)$ g/gal. Construct a mathematical model of this flow process and determine the amount of chemical in the pond at any time. Plot the solution and describe in words the effect of the variation in the incoming concentration.

Solution:

Since the incoming and outgoing flows of water are the same, the amount of water in the pond remains constant at 10^7 gal. Let us denote time by t, measured in years, and the chemical by $Q(t)$, measured in grams. This example is similar to Example 1, and the same inflow/outflow principle applies. Thus

$$\frac{dQ}{dt} = \text{rate in} - \text{rate out},$$

where "rate in" and "rate out" refer to the rates at which the chemical flows into and out of the pond, respectively. The rate at which the chemical flows in is given by

$$\text{rate in} = (5 \times 10^6)\ \text{gal/yr}\ (2 + \sin(2t))\ \text{g/gal}. \tag{18}$$

The concentration of chemical in the pond is $Q(t)/10^7$ g/gal, so the rate of flow out is

$$\text{rate out} = (5 \times 10^6)\ \text{gal/year}\ (Q(t)/10^7)\ \text{g/gal} = Q(t)/2\ \text{g/yr}. \tag{19}$$

Thus we obtain the differential equation

$$\frac{dQ}{dt} = (5 \times 10^6)(2 + \sin(2t)) - \frac{Q(t)}{2}, \tag{20}$$

where each term has the units of g/yr.

To make the coefficients more manageable, it is convenient to introduce a new dependent variable defined by $q(t) = Q(t)/10^6$, or $Q(t) = 10^6 q(t)$. This means that $q(t)$ is measured in millions of grams, or megagrams (metric tons). If we make this substitution in equation (20), then each term contains the factor 10^6, which can be canceled. If we also transpose the term involving $q(t)$ to the left-hand side of the equation, we finally have

$$\frac{dq}{dt} + \frac{1}{2}q = 10 + 5\sin(2t). \tag{21}$$

Originally, there is no chemical in the pond, so the initial condition is

$$q(0) = 0. \tag{22}$$

Equation (21) is linear, and although the right-hand side is a function of time, the coefficient of q is a constant. Thus the integrating factor is $e^{t/2}$. Multiplying equation (21) by this factor and integrating the resulting equation, we obtain the general solution

$$q(t) = 20 - \frac{40}{17}\cos(2t) + \frac{10}{17}\sin(2t) + ce^{-t/2}. \tag{23}$$

The initial condition (22) requires that $c = -300/17$, so the solution of the initial value problem (17), (18) is

$$q(t) = 20 - \frac{40}{17}\cos(2t) + \frac{10}{17}\sin(2t) - \frac{300}{17}e^{-t/2}. \tag{24}$$

A plot of the solution (24) is shown in Figure 2.3.3, along with the line $q = 20$ (shown in black). The exponential term in the solution is important for small t, but it diminishes rapidly as t increases. Later, the solution consists of an oscillation, due to the $\sin(2t)$ and $\cos(2t)$ terms, about the constant level $q = 20$. Note that if the $\sin(2t)$ term were not present in equation (21), then $q = 20$ would be the equilibrium solution of that equation.

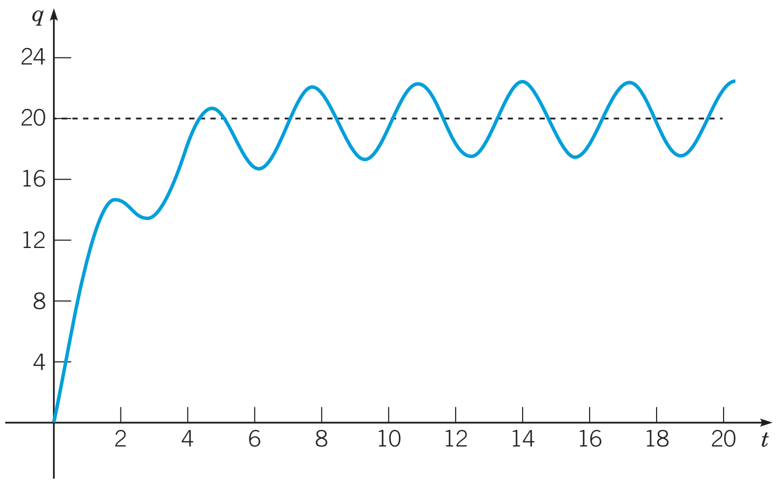


FIGURE 2.3.3 Solution of the initial value problem (17), (18): $dq/dt + q/2 = 10 + 5\sin(2t)$, $q(0) = 0$.

Let us now consider the adequacy of the mathematical model itself for this problem. The model rests on several assumptions that have not yet been stated explicitly. In the first place, the amount of water in the pond is controlled entirely by the rates of flow in and out—none is lost by evaporation or by seepage into the ground, and none is gained by rainfall. The same is true of the chemical; it flows into and out of the pond, but none is absorbed by fish or other organisms living in the pond. In addition, we assume that the concentration of chemical in the pond is uniform throughout the pond. Whether the results obtained from the model are accurate depends strongly on the validity of these simplifying assumptions.

EXAMPLE 4 | Escape Velocity

A body of constant mass m is projected away from the earth in a direction perpendicular to the earth's surface with an initial velocity v_0. Assuming that there is no air resistance, but taking into account the variation of the earth's gravitational field with distance, find an expression for the velocity during the ensuing motion. Also find the initial velocity that is required to lift the body to a given maximum altitude A_{max} above the surface of the earth, and find the least initial velocity for which the body will not return to the earth; the latter is the **escape velocity**.

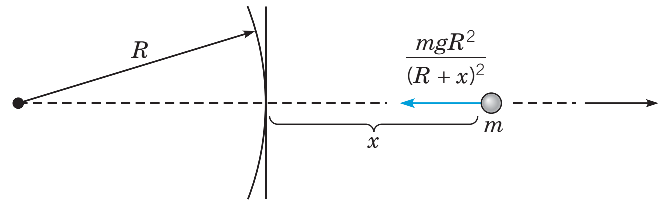


FIGURE 2.3.4 A body in the earth's gravitational field is pulled towards the center of the earth.

Solution:

Let the positive x-axis point away from the center of the earth along the line of motion with $x = 0$ lying on the earth's surface; see Figure 2.3.4. The figure is drawn horizontally to remind you that gravity is directed toward the center of the earth, which is not necessarily downward from a perspective away from the earth's surface. The gravitational force acting on the body (that is, its weight) is inversely proportional to the square of the distance from the center of the earth and is given by $w(x) = -k/(x+R)^2$, where k is a constant, R is the radius of the earth, and the minus sign signifies that $w(x)$ is directed in the negative x direction. We know that on the earth's surface $w(0)$ is given by $-mg$, where g is the acceleration due to gravity at sea level. Therefore, $k = mgR^2$ and

$$w(x) = -\frac{mgR^2}{(R+x)^2}. \tag{25}$$

Since there are no other forces acting on the body, the equation of motion is

$$m\frac{dv}{dt} = -\frac{mgR^2}{(R+x)^2}, \tag{26}$$

and the initial condition is

$$v(0) = v_0. \tag{27}$$

Unfortunately, equation (26) involves too many variables since it depends on t, x, and v. To remedy this situation, we can eliminate t from equation (26) by thinking of x, rather than t, as the independent variable. Then we can express dv/dt in terms of dv/dx by using the chain rule; hence

$$\frac{dv}{dt} = \frac{dv}{dx}\frac{dx}{dt} = v\frac{dv}{dx},$$

and equation (26) is replaced by

$$v\frac{dv}{dx} = -\frac{gR^2}{(R+x)^2}. \tag{28}$$

Equation (28) is separable but not linear, so by separating the variables and integrating, we obtain

$$\frac{v^2}{2} = \frac{gR^2}{R+x} + c. \tag{29}$$

Since $x = 0$ when $t = 0$, the initial condition (27) at $t = 0$ can be replaced by the condition that $v = v_0$ when $x = 0$. Hence $c = (v_0^2/2) - gR$ and

$$v = \pm\sqrt{v_0^2 - 2gR + \frac{2gR^2}{R+x}}. \tag{30}$$

Note that equation (30) gives the velocity as a function of altitude rather than as a function of time. The plus sign must be chosen if the body is rising, and the minus sign must be chosen if it is falling back to earth.

To determine the maximum altitude $A_{\max}$ that the body reaches, we set $v = 0$ and $x = A_{\max}$ in equation (30) and then solve for $A_{\max}$, obtaining

$$A_{\max} = \frac{v_0^2 R}{2gR - v_0^2}. \tag{31}$$

Solving equation (31) for v_0, we find the initial velocity required to lift the body to the altitude $A_{\max}$, namely,

$$v_0 = \sqrt{2gR\frac{A_{\max}}{R + A_{\max}}}. \tag{32}$$

The escape velocity v_e is then found by letting $A_{\max} \to \infty$. Consequently,

$$v_e = \sqrt{2gR}. \tag{33}$$

The numerical value of v_e is approximately 6.9 mi/s, or 11.1 km/s.

The preceding calculation of the escape velocity neglects the effect of air resistance, so the actual escape velocity (including the effect of air resistance) is somewhat higher. On the other hand, the effective escape velocity can be significantly reduced if the body is transported a considerable distance above sea level before being launched. Both gravitational and frictional forces are thereby reduced; air resistance, in particular, diminishes quite rapidly with increasing altitude. You should keep in mind also that it may well be impractical to impart too large an initial velocity instantaneously; space vehicles, for instance, receive their initial acceleration during a period of a few minutes.

Problems

1. Consider a tank used in certain hydrodynamic experiments. After one experiment the tank contains 200 L of a dye solution with a concentration of 1 g/L. To prepare for the next experiment, the tank is to be rinsed with fresh water flowing in at a rate of 2 L/min, the well-stirred solution flowing out at the same rate. Find the time that will elapse before the concentration of dye in the tank reaches 1% of its original value.

2. A tank initially contains 120 L of pure water. A mixture containing a concentration of γ g/L of salt enters the tank at a rate of 2 L/min, and the well-stirred mixture leaves the tank at the same rate. Find an expression in terms of γ for the amount of salt in the tank at any time t. Also find the limiting amount of salt in the tank as $t \to \infty$.

3. A tank contains 100 gal of water and 50 oz of salt. Water containing a salt concentration of $\frac{1}{4}\left(1 + \frac{1}{2}\sin t\right)$ oz/gal flows into the tank at a rate of 2 gal/min, and the mixture in the tank flows out at the same rate.

a. Find the amount of salt in the tank at any time.

G **b.** Plot the solution for a time period long enough so that you see the ultimate behavior of the graph.

c. The long-time behavior of the solution is an oscillation about a certain constant level. What is this level? What is the amplitude of the oscillation?

4. Suppose that a tank containing a certain liquid has an outlet near the bottom. Let $h(t)$ be the height of the liquid surface above the outlet at time t. Torricelli's[2] principle states that the outflow velocity v at the outlet is equal to the velocity of a particle falling freely (with no drag) from the height h.

a. Show that $v = \sqrt{2gh}$, where g is the acceleration due to gravity.

b. By equating the rate of outflow to the rate of change of liquid in the tank, show that $h(t)$ satisfies the equation

$$A(h)\frac{dh}{dt} = -\alpha a\sqrt{2gh}, \tag{34}$$

where $A(h)$ is the area of the cross section of the tank at height h and a is the area of the outlet. The constant α is a contraction coefficient that accounts for the observed fact that the cross section of the (smooth) outflow stream is smaller than a. The value of α for water is about 0.6.

c. Consider a water tank in the form of a right circular cylinder that is 3 m high above the outlet. The radius of the tank is 1 m, and the radius of the circular outlet is 0.1 m. If the tank is initially full of water, determine how long it takes to drain the tank down to the level of the outlet.

5. Suppose that a sum S_0 is invested at an annual rate of return r compounded continuously.

a. Find the time T required for the original sum to double in value as a function of r.

b. Determine T if $r = 7\%$.

c. Find the return rate that must be achieved if the initial investment is to double in 8 years.

6. A young person with no initial capital invests k dollars per year at an annual rate of return r. Assume that investments are made continuously and that the return is compounded continuously.

a. Determine the sum $S(t)$ accumulated at any time t.

b. If $r = 7.5\%$, determine k so that \$1 million will be available for retirement in 40 years.

c. If $k = \$2000$/year, determine the return rate r that must be obtained to have \$1 million available in 40 years.

[2] Evangelista Torricelli (1608–1647), successor to Galileo as court mathematician in Florence, published this result in 1644. In addition to this work in fluid dynamics, he is also known for constructing the first mercury barometer and for making important contributions to geometry.

7. A certain college graduate borrows \$8000 to buy a car. The lender charges interest at an annual rate of 10%. Assuming that interest is compounded continuously and that the borrower makes payments continuously at a constant annual rate k, determine the payment rate k that is required to pay off the loan in 3 years. Also determine how much interest is paid during the 3-year period.

N **8.** A recent college graduate borrows \$150,000 at an interest rate of 6% to purchase a condominium. Anticipating steady salary increases, the buyer expects to make payments at a monthly rate of $800 + 10t$, where t is the number of months since the loan was made.

a. Assuming that this payment schedule can be maintained, when will the loan be fully paid?
b. Assuming the same payment schedule, how large a loan could be paid off in exactly 20 years?

9. An important tool in archeological research is radiocarbon dating, developed by the American chemist Willard F. Libby.[3] This is a means of determining the age of certain wood and plant remains, and hence of animal or human bones or artifacts found buried at the same levels. Radiocarbon dating is based on the fact that some wood or plant remains contain residual amounts of carbon-14, a radioactive isotope of carbon. This isotope is accumulated during the lifetime of the plant and begins to decay at its death. Since the half-life of carbon-14 is long (approximately 5730 years),[4] measurable amounts of carbon-14 remain after many thousands of years. If even a tiny fraction of the original amount of carbon-14 is still present, then by appropriate laboratory measurements the *proportion* of the original amount of carbon-14 that remains can be accurately determined. In other words, if $Q(t)$ is the amount of carbon-14 at time t and Q_0 is the original amount, then the ratio $Q(t)/Q_0$ can be determined, as long as this quantity is not too small. Present measurement techniques permit the use of this method for time periods of 50,000 years or more.

a. Assuming that Q satisfies the differential equation $Q' = -rQ$, determine the decay constant r for carbon-14.
b. Find an expression for $Q(t)$ at any time t, if $Q(0) = Q_0$.
c. Suppose that certain remains are discovered in which the current residual amount of carbon-14 is 20% of the original amount. Determine the age of these remains.

N **10.** Suppose that a certain population has a growth rate that varies with time and that this population satisfies the differential equation

$$\frac{dy}{dt} = (0.5 + \sin t)\frac{y}{5}.$$

a. If $y(0) = 1$, find (or estimate) the time τ at which the population has doubled. Choose other initial conditions and determine whether the doubling time τ depends on the initial population.
b. Suppose that the growth rate is replaced by its average value $1/10$. Determine the doubling time τ in this case.
c. Suppose that the term $\sin t$ in the differential equation is replaced by $\sin 2\pi t$; that is, the variation in the growth rate has a substantially higher frequency. What effect does this have on the doubling time τ?
d. Plot the solutions obtained in parts a, b, and c on a single set of axes.

N **11.** Suppose that a certain population satisfies the initial value problem

$$dy/dt = r(t)y - k, \quad y(0) = y_0,$$

where the growth rate $r(t)$ is given by $r(t) = (1 + \sin t)/5$, and k represents the rate of predation.

G **a.** Suppose that $k = 1/5$. Plot y versus t for several values of y_0 between $1/2$ and 1.
b. Estimate the critical initial population y_c below which the population will become extinct.
c. Choose other values of k and find the corresponding y_c for each one.
G **d.** Use the data you have found in parts b and c to plot y_c versus k.

12. Newton's law of cooling states that the temperature of an object changes at a rate proportional to the difference between its temperature and that of its surroundings. Suppose that the temperature of a cup of coffee obeys Newton's law of cooling. If the coffee has a temperature of 200°F when freshly poured, and 1 min later has cooled to 190°F in a room at 70°F, determine when the coffee reaches a temperature of 150°F.

13. Heat transfer from a body to its surroundings by radiation, based on the Stefan–Boltzmann[5] law, is described by the differential equation

$$\frac{du}{dt} = -\alpha(u^4 - T^4), \tag{35}$$

where $u(t)$ is the absolute temperature of the body at time t, T is the absolute temperature of the surroundings, and α is a constant depending on the physical parameters of the body. However, if u is much larger than T, then solutions of equation (35) are well approximated by solutions of the simpler equation

$$\frac{du}{dt} = -\alpha u^4. \tag{36}$$

Suppose that a body with initial temperature 2000 K is surrounded by a medium with temperature 300 K and that $\alpha = 2.0 \times 10^{-12}\ \mathrm{K}^{-3}/\mathrm{s}$.

a. Determine the temperature of the body at any time by solving equation (36).
G **b.** Plot the graph of u versus t.
N **c.** Find the time τ at which $u(\tau) = 600$—that is, twice the ambient temperature. Up to this time the error in using equation (36) to approximate the solutions of equation (35) is no more than 1%.

N **14.** Consider an insulated box (a building, perhaps) with internal temperature $u(t)$. According to Newton's law of cooling, u satisfies the differential equation

$$\frac{du}{dt} = -k(u - T(t)), \tag{37}$$

where $T(t)$ is the ambient (external) temperature. Suppose that $T(t)$ varies sinusoidally; for example, assume that $T(t) = T_0 + T_1\cos(\omega t)$.

[3] Willard F. Libby (1908–1980) was born in rural Colorado and received his education at the University of California at Berkeley. He developed the method of radiocarbon dating beginning in 1947 while he was at the University of Chicago. For this work he was awarded the Nobel Prize in Chemistry in 1960.

[4] *McGraw-Hill Encyclopedia of Science and Technology* (8th ed.) (New York: McGraw-Hill, 1997), Vol. 5, p. 48.

[5] Jozef Stefan (1835–1893), professor of physics at Vienna, stated the radiation law on empirical grounds in 1879. His student Ludwig Boltzmann (1844–1906) derived it theoretically from the principles of thermodynamics in 1884. Boltzmann is best known for his pioneering work in statistical mechanics.

a. Solve equation (37) and express $u(t)$ in terms of t, k, T_0, T_1, and ω. Observe that part of your solution approaches zero as t becomes large; this is called the transient part. The remainder of the solution is called the steady state; denote it by $S(t)$.

G **b.** Suppose that t is measured in hours and that $\omega = \pi/12$, corresponding to a period of 24 h for $T(t)$. Further, let $T_0 = 60°\text{F}$, $T_1 = 15°\text{F}$, and $k = 0.2/\text{h}$. Draw graphs of $S(t)$ and $T(t)$ versus t on the same axes. From your graph estimate the amplitude R of the oscillatory part of $S(t)$. Also estimate the time lag τ between corresponding maxima of $T(t)$ and $S(t)$.

c. Let k, T_0, T_1, and ω now be unspecified. Write the oscillatory part of $S(t)$ in the form $R\cos(\omega(t-\tau))$. Use trigonometric identities to find expressions for R and τ. Let T_1 and ω have the values given in part b, and plot graphs of R and τ versus k.

15. Consider a lake of constant volume V containing at time t an amount $Q(t)$ of pollutant, evenly distributed throughout the lake with a concentration $c(t)$, where $c(t) = Q(t)/V$. Assume that water containing a concentration k of pollutant enters the lake at a rate r, and that water leaves the lake at the same rate. Suppose that pollutants are also added directly to the lake at a constant rate P. Note that the given assumptions neglect a number of factors that may, in some cases, be important—for example, the water added or lost by precipitation, absorption, and evaporation; the stratifying effect of temperature differences in a deep lake; the tendency of irregularities in the coastline to produce sheltered bays; and the fact that pollutants are deposited unevenly throughout the lake but (usually) at isolated points around its periphery. The results below must be interpreted in light of the neglect of such factors as these.

a. If at time $t = 0$ the concentration of pollutant is c_0, find an expression for the concentration $c(t)$ at any time. What is the limiting concentration as $t \to \infty$?

b. If the addition of pollutants to the lake is terminated ($k = 0$ and $P = 0$ for $t > 0$), determine the time interval T that must elapse before the concentration of pollutants is reduced to 50% of its original value; to 10% of its original value.

c. Table 2.3.2 contains data[6] for several of the Great Lakes. Using these data, determine from part b the time T that is needed to reduce the contamination of each of these lakes to 10% of the original value.

TABLE 2.3.2 Volume and Flow Data for the Great Lakes

Lake	$10^3 \times V$ (km^3)	r (km^3/year)
Superior	12.2	65.2
Michigan	4.9	158
Erie	0.46	175
Ontario	1.6	209

N **16.** A ball with mass 0.15 kg is thrown upward with initial velocity 20 m/s from the roof of a building 30 m high. Neglect air resistance.

a. Find the maximum height above the ground that the ball reaches.

b. Assuming that the ball misses the building on the way down, find the time that it hits the ground.

G **c.** Plot the graphs of velocity and position versus time.

[6] This problem is based on R. H. Rainey, "Natural Displacement of Pollution from the Great Lakes," *Science 155* (1967), pp. 1242–1243; the information in the table was taken from that source.

N **17.** Assume that the conditions are as in Problem 16 except that there is a force due to air resistance of magnitude $|v|/30$ directed opposite to the velocity, where the velocity v is measured in m/s.

a. Find the maximum height above the ground that the ball reaches.

b. Find the time that the ball hits the ground.

G **c.** Plot the graphs of velocity and position versus time. Compare these graphs with the corresponding ones in Problem 16.

N **18.** Assume that the conditions are as in Problem 16 except that there is a force due to air resistance of magnitude $v^2/1325$ directed opposite to the velocity, where the velocity v is measured in m/s.

a. Find the maximum height above the ground that the ball reaches.

b. Find the time that the ball hits the ground.

G **c.** Plot the graphs of velocity and position versus time. Compare these graphs with the corresponding ones in Problems 16 and 17.

19. A body of constant mass m is projected vertically upward with an initial velocity v_0 in a medium offering a resistance $k|v|$, where k is a constant. Neglect changes in the gravitational force.

a. Find the maximum height x_m attained by the body and the time t_m at which this maximum height is reached.

b. Show that if $kv_0/mg < 1$, then t_m and x_m can be expressed as

$$t_m = \frac{v_0}{g}\left(1 - \frac{1}{2}\frac{kv_0}{mg} + \frac{1}{3}\left(\frac{kv_0}{mg}\right)^2 - \cdots\right),$$

$$x_m = \frac{v_0^2}{2g}\left(1 - \frac{2}{3}\frac{kv_0}{mg} + \frac{1}{2}\left(\frac{kv_0}{mg}\right)^2 - \cdots\right).$$

c. Show that the quantity kv_0/mg is dimensionless.

20. A body of mass m is projected vertically upward with an initial velocity v_0 in a medium offering a resistance $k|v|$, where k is a constant. Assume that the gravitational attraction of the earth is constant.

a. Find the velocity $v(t)$ of the body at any time.

b. Use the result of part a to calculate the limit of $v(t)$ as $k \to 0$—that is, as the resistance approaches zero. Does this result agree with the velocity of a mass m projected upward with an initial velocity v_0 in a vacuum?

c. Use the result of part a to calculate the limit of $v(t)$ as $m \to 0$—that is, as the mass approaches zero.

21. A body falling in a relatively dense fluid, oil for example, is acted on by three forces (see Figure 2.3.5): a resistive force R, a buoyant force B, and its weight w due to gravity. The buoyant force is equal to the weight of the fluid displaced by the object. For a slowly moving spherical body of radius a, the resistive force is given by Stokes's law, $R = 6\pi\mu a|v|$, where v is the velocity of the body, and μ is the coefficient of viscosity of the surrounding fluid.[7]

[7] Sir George Gabriel Stokes (1819–1903) was born in Ireland but spent most of his life at Cambridge University, first as a student and later as a professor. Stokes was one of the foremost applied mathematicians of the nineteenth century, best known for his work in fluid dynamics and the wave theory of light. The basic equations of fluid mechanics (the Navier–Stokes equations) are named partly in his honor, and one of the fundamental theorems of vector calculus bears his name. He was also one of the pioneers in the use of divergent (asymptotic) series.

a. Find the limiting velocity of a solid sphere of radius a and density ρ falling freely in a medium of density ρ' and coefficient of viscosity μ.
b. In 1910 R. A. Millikan[8] studied the motion of tiny droplets of oil falling in an electric field. A field of strength E exerts a force Ee on a droplet with charge e. Assume that E has been adjusted so the droplet is held stationary ($v = 0$) and that w and B are as given above. Find an expression for e. Millikan repeated this experiment many times, and from the data that he gathered he was able to deduce the charge on an electron.

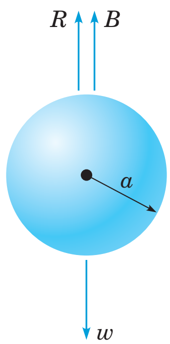


FIGURE 2.3.5 A body falling in a dense fluid (see Problem 21).

22. Let $v(t)$ and $w(t)$ be the horizontal and vertical components, respectively, of the velocity of a batted (or thrown) baseball. In the absence of air resistance, v and w satisfy the equations

$$\frac{dv}{dt} = 0, \qquad \frac{dw}{dt} = -g.$$

a. Show that

$$v = u\cos A, \qquad w = -gt + u\sin A,$$

where u is the initial speed of the ball and A is its initial angle of elevation.
b. Let $x(t)$ and $y(t)$ be the horizontal and vertical coordinates, respectively, of the ball at time t. If $x(0) = 0$ and $y(0) = h$, find $x(t)$ and $y(t)$ at any time t.
G **c.** Let $g = 32$ ft/s^2, $u = 125$ ft/s, and $h = 3$ ft. Plot the trajectory of the ball for several values of the angle A; that is, plot $x(t)$ and $y(t)$ parametrically.
d. Suppose the outfield wall is at a distance L and has height H. Find a relation between u and A that must be satisfied if the ball is to clear the wall.
e. Suppose that $L = 350$ ft and $H = 10$ ft. Using the relation in part (d), find (or estimate from a plot) the range of values of A that correspond to an initial velocity of $u = 110$ ft/s.
f. For $L = 350$ and $H = 10$, find the minimum initial velocity u and the corresponding optimal angle A for which the ball will clear the wall.

N **23.** A more realistic model (than that in Problem 22) of a baseball in flight includes the effect of air resistance. In this case the equations of motion are

$$\frac{dv}{dt} = -rv, \qquad \frac{dw}{dt} = -g - rw,$$

where r is the coefficient of resistance.

a. Determine $v(t)$ and $w(t)$ in terms of initial speed u and initial angle of elevation A.
b. Find $x(t)$ and $y(t)$ if $x(0) = 0$ and $y(0) = h$.
G **c.** Plot the trajectory of the ball for $r = 1/5$, $u = 125$, $h = 3$, and for several values of A. How do the trajectories differ from those in Problem 22 with $r = 0$?
d. Assuming that $r = 1/5$ and $h = 3$, find the minimum initial velocity u and the optimal angle A for which the ball will clear a wall that is 350 ft distant and 10 ft high. Compare this result with that in Problem 22f.

24. Brachistochrone Problem. One of the famous problems in the history of mathematics is the brachistochrone[9] problem: to find the curve along which a particle will slide without friction in the minimum time from one given point P to another Q, the second point being lower than the first but not directly beneath it (see Figure 2.3.6). This problem was posed by Johann Bernoulli in 1696 as a challenge problem to the mathematicians of his day. Correct solutions were found by Johann Bernoulli and his brother Jakob Bernoulli and by Isaac Newton, Gottfried Leibniz, and the Marquis de L'Hôpital. The brachistochrone problem is important in the development of mathematics as one of the forerunners of the calculus of variations.

In solving this problem, it is convenient to take the origin as the upper point P and to orient the axes as shown in Figure 2.3.6. The lower point Q has coordinates (x_0, y_0). It is then possible to show that the curve of minimum time is given by a function $y = \phi(x)$ that satisfies the differential equation

$$(1 + y'^2)y = k^2, \tag{38}$$

where k^2 is a certain positive constant to be determined later.
a. Solve equation (38) for y'. Why is it necessary to choose the positive square root?
b. Introduce the new variable t by the relation

$$y = k^2 \sin^2 t. \tag{39}$$

Show that the equation found in part a then takes the form

$$2k^2 \sin^2 t\, dt = dx. \tag{40}$$

c. Letting $\theta = 2t$, show that the solution of equation (40) for which $x = 0$ when $y = 0$ is given by

$$x = k^2(\theta - \sin\theta)/2, \qquad y = k^2(1 - \cos\theta)/2. \tag{41}$$

Equations (41) are parametric equations of the solution of equation (38) that passes through $(0, 0)$. The graph of equations (41) is called a **cycloid**.
d. If we make a proper choice of the constant k, then the cycloid also passes through the point (x_0, y_0) and is the solution of the brachistochrone problem. Find k if $x_0 = 1$ and $y_0 = 2$.

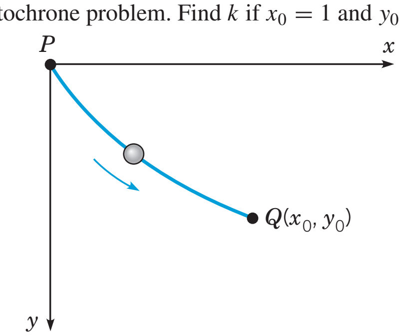


FIGURE 2.3.6 The brachistochrone (see Problem 24).

[8]Robert A. Millikan (1868–1953) was educated at Oberlin College and Columbia University. Later he was a professor at the University of Chicago and California Institute of Technology. His determination of the charge on an electron was published in 1910. For this work, and for other studies of the photoelectric effect, he was awarded the Nobel Prize for Physics in 1923.

[9]The word "brachistochrone" comes from the Greek words *brachistos*, meaning shortest, and *chronos*, meaning time.

2.4 Differences Between Linear and Nonlinear Differential Equations

Up to now, we have been primarily concerned with showing that first-order differential equations can be used to investigate many different kinds of problems in the natural sciences, and with presenting methods of solving such equations if they are either linear or separable. Now it is time to turn our attention to some more general questions about differential equations and to explore in more detail some important ways in which nonlinear equations differ from linear ones.

Existence and Uniqueness of Solutions. So far, we have discussed a number of initial value problems, each of which had a solution and apparently only one solution. That raises the question of whether this is true of all initial value problems for first-order equations. In other words, does every initial value problem have exactly one solution? This may be an important question even for nonmathematicians. If you encounter an initial value problem in the course of investigating some physical problem, you might want to know that it has a solution before spending very much time and effort in trying to find it. Further, if you are successful in finding one solution, you might be interested in knowing whether you should continue a search for other possible solutions or whether you can be sure that there are no other solutions. For linear equations, the answers to these questions are given by the following fundamental theorem.

Theorem 2.4.1 | Existence and Uniqueness Theorem for First-Order Linear Equations

If the functions p and g are continuous on an open interval $I: \alpha < t < \beta$ containing the point $t = t_0$, then there exists a unique function $y = \phi(t)$ that satisfies the differential equation

$$y' + p(t)y = g(t) \tag{1}$$

for each t in I, and that also satisfies the initial condition

$$y(t_0) = y_0, \tag{2}$$

where y_0 is an arbitrary prescribed initial value.

Observe that Theorem 2.4.1 states that the given initial value problem *has* a solution and also that the problem has *only one* solution. In other words, the theorem asserts both the *existence* and the *uniqueness* of the solution of the initial value problem (1). In addition, it states that the solution exists throughout any interval I containing the initial point t_0 in which the coefficients p and g are continuous. That is, the solution can be discontinuous or fail to exist only at points where at least one of p and g is discontinuous. Such points can often be identified at a glance.

The proof of this theorem is partly contained in the discussion in Section 2.1 leading to the formula (see equation (32) in Section 2.1)

$$\mu(t)y = \int \mu(t)g(t)\,dt + c, \tag{3}$$

where [equation (30) in Section 2.1]

$$\mu(t) = \exp \int p(t)\,dt. \tag{4}$$

The derivation in Section 2.1 shows that if equation (1) has a solution, then it must be given by equation (3). By looking slightly more closely at that derivation, we can also conclude that the differential equation (1) must indeed have a solution. Since p is continuous for $\alpha < t < \beta$, it follows that on the interval $\alpha < t < \beta$, μ is defined, is a differentiable function, and is nonzero. Upon multiplying equation (1) by $\mu(t)$, we obtain

$$(\mu(t)y)' = \mu(t)g(t). \tag{5}$$

Since both μ and g are continuous, the function μg is integrable, and equation (3) follows from equation (5). Further, the integral of μg is differentiable, so y as given by equation (3) exists and is differentiable throughout the interval $\alpha < t < \beta$. By substituting the expression for y from equation (3) into either equation (1) or equation (5), you can verify that this expression satisfies the differential equation throughout the interval $\alpha < t < \beta$. Finally, the initial condition (2) determines the constant c uniquely, so there is only one solution of the initial value problem; this completes the proof.

Equation (4) determines the integrating factor $\mu(t)$ only up to a multiplicative factor that depends on the lower limit of integration. If we choose this lower limit to be t_0, then

$$\mu(t) = \exp \int_{t_0}^{t} p(s)ds, \tag{6}$$

and it follows that $\mu(t_0) = 1$. Using the integrating factor given by equation (6), and choosing the lower limit of integration in equation (3) also to be t_0, we obtain the general solution of equation (1) in the form

$$y = \frac{1}{\mu(t)}\left(\int_{t_0}^{t} \mu(s)g(s)\,ds + c\right). \tag{7}$$

To satisfy the initial condition (2), we must choose $c = y_0$. Thus the solution of the initial value problem (1) is

$$y = \frac{1}{\mu(t)}\left(\int_{t_0}^{t} \mu(s)g(s)\,ds + y_0\right), \tag{8}$$

where $\mu(t)$ is given by equation (6).

Turning now to nonlinear differential equations, we must replace Theorem 2.4.1 by a more general theorem, such as the one that follows.

Theorem 2.4.2 | Existence and Uniqueness Theorem for First-Order Nonlinear Equations

Let the functions f and $\partial f/\partial y$ be continuous in some rectangle $\alpha < t < \beta, \gamma < y < \delta$ containing the point (t_0, y_0). Then, in some interval $t_0 - h < t < t_0 + h$ contained in $\alpha < t < \beta$, there is a unique solution $y = \phi(t)$ of the initial value problem

$$y' = f(t, y), \quad y(t_0) = y_0. \tag{9}$$

Observe that the hypotheses in Theorem 2.4.2 reduce to those in Theorem 2.4.1 if the differential equation is linear. In this case

$$f(t, y) = -p(t)y + g(t) \text{ and } \frac{\partial f(t, y)}{\partial y} = -p(t),$$

so the continuity of f and $\dfrac{\partial f}{\partial y}$ is equivalent to the continuity of p and g.

The proof of Theorem 2.4.1 was comparatively simple because it could be based on the expression (3) that gives the solution of an arbitrary linear equation. There is no corresponding expression for the solution of the differential equation (9), so the proof of Theorem 2.4.2 is much more difficult. It is discussed to some extent in Section 2.8 and in greater depth in more advanced books on differential equations.

We note that the conditions stated in Theorem 2.4.2 are sufficient to guarantee the existence of a unique solution of the initial value problem (6) in some interval $(t_0 - h, t_0 + h)$, but they are not necessary. That is, the conclusion remains true under slightly weaker hypotheses about the function f. In fact, the existence of a solution (but not its uniqueness) can be established on the basis of the continuity of f alone.

An important geometrical consequence of the uniqueness parts of Theorems 2.4.1 and 2.4.2 is that the graphs of two solutions cannot intersect each other. Otherwise, there would be two solutions that satisfy the initial condition corresponding to the point of intersection, in contradiction to Theorem 2.4.1 or 2.4.2.

We now consider some examples.

EXAMPLE 1

Use Theorem 2.4.1 to find an interval in which the initial value problem

$$ty' + 2y = 4t^2, \tag{10}$$

$$y(1) = 2 \tag{11}$$

has a unique solution. Then do the same when the initial condition is changed to $y(-1) = 2$.

Solution:

Rewriting equation (10) in the standard form (1), we have

$$y' + (2/t)y = 4t,$$

so $p(t) = 2/t$ and $g(t) = 4t$. Thus, for this equation, g is continuous for all t, while p is continuous only for $t < 0$ or for $t > 0$. The interval $t > 0$ contains the initial point; consequently, Theorem 2.4.1 guarantees that the problem (7), (8) has a unique solution on the interval $0 < t < \infty$. In Example 4 of Section 2.1 we found the solution of this initial value problem to be

$$y = t^2 + \frac{1}{t^2}, \quad t > 0. \tag{12}$$

Now suppose that the initial condition (11) is changed to $y(-1) = 2$. Then Theorem 2.4.1 asserts the existence of a unique solution for $t < 0$. As you can readily verify, the solution is again given by equation (12), but now on the interval $t < 0$.

EXAMPLE 2

Apply Theorem 2.4.2 to the initial value problem

$$\frac{dy}{dx} = \frac{3x^2 + 4x + 2}{2(y-1)}, \quad y(0) = -1. \tag{13}$$

Repeat this analysis when the initial condition is changed to $y(0) = 1$.

Solution:

Note that Theorem 2.4.1 is not applicable to this problem since the differential equation is nonlinear. To apply Theorem 2.4.2, observe that

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

Thus each of these functions is continuous everywhere except on the line $y = 1$. Consequently, a rectangle can be drawn about the initial point $(0, -1)$ in which both f and $\partial f/\partial y$ are continuous. Therefore, Theorem 2.4.2 guarantees that the initial value problem has a unique solution in some interval about $x = 0$. However, even though the rectangle can be stretched infinitely far in both the positive and the negative x directions, this does not necessarily mean that the solution exists for all x. Indeed, the initial value problem (9) was solved in Example 2 of Section 2.2, and the solution exists only for $x > -2$.

Now suppose we change the initial condition to $y(0) = 1$. The initial point now lies on the line $y = 1$, so no rectangle can be drawn about it within which f and $\partial f/\partial y$ are continuous. Consequently, Theorem 2.4.2 says nothing about possible solutions of this modified problem. However, if we separate the variables and integrate, as in Section 2.2, we find that

$$y^2 - 2y = x^3 + 2x^2 + 2x + c.$$

Further, if $x = 0$ and $y = 1$, then $c = -1$. Finally, by solving for y, we obtain

$$y = 1 \pm \sqrt{x^3 + 2x^2 + 2x}. \tag{14}$$

Equation (14) provides two functions that satisfy the given differential equation for $x > 0$ and also satisfy the initial condition $y(0) = 1$. The fact that there are two solutions to this initial value problem reinforces the conclusion that Theorem 2.4.2 does not apply to this initial value problem.

EXAMPLE 3

Consider the initial value problem

$$y' = y^{1/3}, \quad y(0) = 0 \tag{15}$$

for $t \geq 0$. Apply Theorem 2.4.2 to this initial value problem and then solve the problem.

Solution:

The function $f(t, y) = y^{1/3}$ is continuous everywhere, but $\dfrac{\partial f}{\partial y} = \dfrac{1}{3}y^{-2/3}$ does not exist when $y = 0$, and hence it is not continuous there. Thus Theorem 2.4.2 does not apply to this problem, and no conclusion can be drawn from it. However, by the remark following Theorem 2.4.2, the continuity of f does ensure the existence of solutions, though not their uniqueness.

To understand the situation more clearly, we must actually solve the problem, which is easy to do since the differential equation is separable. Thus we have

$$y^{-1/3}dy = dt,$$

so

$$\frac{3}{2}y^{2/3} = t + c$$

and

$$y = \left(\frac{2}{3}(t + c)\right)^{3/2}.$$

The initial condition is satisfied if $c = 0$, so

$$y = \phi_1(t) = \left(\frac{2}{3}t\right)^{3/2}, \quad t \geq 0 \tag{16}$$

satisfies both of equations (15). On the other hand, the function

$$y = \phi_2(t) = -\left(\frac{2}{3}t\right)^{3/2}, \quad t \geq 0 \tag{17}$$

is also a solution of the initial value problem. Moreover, the function

$$y = \psi(t) = 0, \quad t \geq 0 \tag{18}$$

is yet another solution. Indeed, for an arbitrary positive t_0, the functions

$$y = \chi(t) = \begin{cases} 0, & \text{if } 0 \leq t < t_0, \\ \pm\left(\frac{2}{3}(t - t_0)\right)^{3/2}, & \text{if } t \geq t_0 \end{cases} \tag{19}$$

are continuous, are differentiable (in particular at $t = t_0$), and are solutions of the initial value problem (11). Hence this problem has an infinite family of solutions; see Figure 2.4.1, where a few of these solutions are shown.

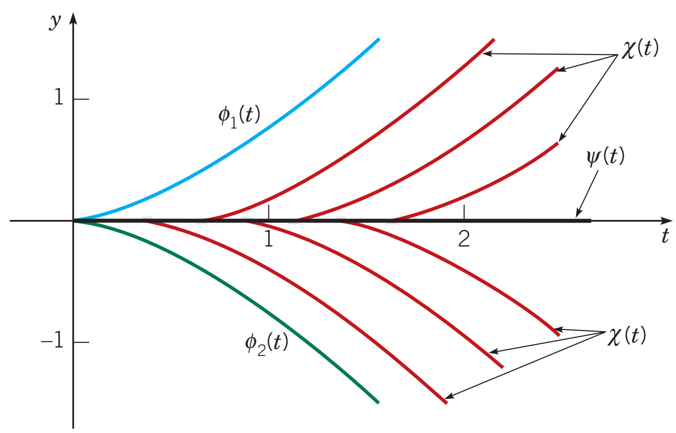


FIGURE 2.4.1 Several solutions of the initial value problem $y' = y^{1/3}$, $y(0) = 0$.

As already noted, the nonuniqueness of the solutions of the problem (11) does not contradict the existence and uniqueness theorem, since Theorem 2.4.2 is not applicable if the initial point lies on the t-axis. If (t_0, y_0) is any point not on the t-axis, however, then the theorem guarantees that there is a unique solution of the differential equation $y' = y^{1/3}$ passing through (t_0, y_0).

Interval of Existence. According to Theorem 2.4.1, the solution of a linear equation (1)

$$y' + p(t)y = g(t),$$

subject to the initial condition $y(t_0) = y_0$, exists throughout any interval about $t = t_0$ in which the functions p and g are continuous. Thus vertical asymptotes or other discontinuities in the solution can occur only at points of discontinuity of p or g. For instance, the solutions in Example 1 (with one exception) are asymptotic to the y-axis, corresponding to the discontinuity at $t = 0$ in the coefficient $p(t) = 2/t$, but none of the solutions has any other point where it fails to exist and to be differentiable. The one exceptional solution shows that solutions may sometimes remain continuous even at points of discontinuity of the coefficients.

On the other hand, for a nonlinear initial value problem satisfying the hypotheses of Theorem 2.4.2, the interval in which a solution exists may be difficult to determine. The solution $y = \phi(t)$ is certain to exist as long as the point $(t, \phi(t))$ remains within a region in which the hypotheses of Theorem 2.4.2 are satisfied. This is what determines the value of h in that theorem. However, since $\phi(t)$ is usually not known, it may be impossible to locate the point $(t, \phi(t))$ with respect to this region. In any case, the interval in which a solution exists may have no simple relationship to the function f in the differential equation $y' = f(t, y)$. This is illustrated by the following example.

EXAMPLE 4

Solve the initial value problem

$$y' = y^2, \quad y(0) = 1, \tag{20}$$

and determine the interval in which the solution exists.

Solution:

Theorem 2.4.2 guarantees that this problem has a unique solution since $f(t, y) = y^2$ and $\dfrac{\partial f}{\partial y} = 2y$ are continuous everywhere. To find the solution, we separate the variables and integrate with the result that

$$y^{-2}\,dy = dt \tag{21}$$

and

$$-y^{-1} = t + c.$$

Then, solving for y, we have

$$y = -\frac{1}{t + c}. \tag{22}$$

To satisfy the initial condition, we must choose $c = -1$, so

$$y = \frac{1}{1 - t} \tag{23}$$

is the solution of the given initial value problem. Clearly, the solution becomes unbounded as $t \to 1$; therefore, the solution exists only in the interval $-\infty < t < 1$. There is no indication from the differential equation itself, however, that the point $t = 1$ is in any way remarkable. Moreover, if the initial condition is replaced by

$$y(0) = y_0, \tag{24}$$

then the constant c in equation (22) must be chosen to be $c = -1/y_0$ $(y_0 \neq 0)$, and it follows that

$$y = \frac{y_0}{1 - y_0 t} \tag{25}$$

is the solution of the initial value problem with the initial condition (24). Observe that the solution (25) becomes unbounded as $t \to 1/y_0$, so the interval of existence of the solution is $-\infty < t < 1/y_0$ if $y_0 > 0$, and is $1/y_0 < t < \infty$ if $y_0 < 0$. This example illustrates another feature of initial value problems for nonlinear equations: the singularities of the solution may depend in an essential way on the initial conditions as well as on the differential equation.

General Solution. Another way in which linear and nonlinear equations differ concerns the concept of a general solution. For a first-order linear differential equation it is possible to obtain a solution containing one arbitrary constant, from which all possible solutions follow by specifying values for this constant. For nonlinear equations this may not be the case; even though a solution containing an arbitrary constant may be found, there may be other solutions that cannot be obtained by giving values to this constant. For instance, for the differential equation $y' = y^2$ in Example 4, the expression in equation (22) contains an arbitrary constant but does not include all solutions of the differential equation. To show this, observe that the function $y = 0$ for all t is certainly a solution of the differential equation, but it cannot be obtained from equation (22) by assigning a value to c. In this example we might anticipate that something of this sort might happen, because to rewrite the original differential equation in the form (21), we must require that y is not zero. However, the existence of "additional" solutions is not uncommon for nonlinear equations; a less obvious example is given in Problem 18. Thus we will use the term "general solution" only when discussing linear equations.

Implicit Solutions. Recall again that for an initial value problem for a first-order linear differential equation, equation (8) provides an explicit formula for the solution $y = \phi(t)$. As long as the necessary antiderivatives can be found, the value of the solution at any point can be determined merely by substituting the appropriate value of t into the equation. The situation for nonlinear equations is much less satisfactory. Usually, the best that we can hope for is to find an equation

$$F(t, y) = 0 \tag{26}$$

involving t and y that is satisfied by the solution $y = \phi(t)$. Even this can be done only for differential equations of certain particular types, of which separable equations are the most important. The equation (26) is called an integral, or first integral, of the differential equation, and (as we have already noted) its graph is an integral curve, or perhaps a family of integral curves. Equation (26), assuming it can be found, defines the solution implicitly; that is, for each value of t we must solve equation (26) to find the corresponding value of y. If equation (26) is simple enough, it may be possible to solve it for y by analytical means and thereby obtain an explicit formula for the solution. However, more frequently this will not be possible, and you will have to resort to a numerical calculation to determine (approximately) the value of y for a given value of t. Once several pairs of values of t and y have been calculated, it is often helpful to plot them and then to sketch the integral curve that passes through them. You should take advantage of the wide range of computational and graphical utilities available to carry out these calculations and to create the graph of one or more integral curves.

Examples 2, 3, and 4 involve nonlinear problems in which it is easy to solve for an explicit formula for the solution $y = \phi(t)$. On the other hand, Examples 1 and 3 in Section 2.2 are cases in which it is better to leave the solution in implicit form and to use numerical means to evaluate it for particular values of the independent variable. The latter situation is more typical; unless the implicit relation is quadratic in y or has some other particularly simple form, it is unlikely that it can be solved exactly by analytical methods. Indeed, more often than not, it is impossible even to find an implicit expression for the solution of a first-order nonlinear equation.

Graphical or Numerical Construction of Integral Curves. Because of the difficulty in obtaining exact analytical solutions of nonlinear differential equations, methods that yield approximate solutions or other qualitative information about solutions are of correspondingly greater importance. We have already described, in Section 1.1, how the direction field of a differential equation can be constructed. The direction field can often show the qualitative form of solutions and can also be helpful in identifying regions of the ty-plane where solutions exhibit interesting features that merit more detailed analytical or numerical investigation. Graphical methods for first-order differential equations are discussed further in Section 2.5.

An introduction to numerical methods for first-order equations is given in Section 2.7, and a systematic discussion of numerical methods appears in Chapter 8. However, it is not necessary to study the numerical algorithms themselves in order to use effectively one of the many software packages that generate and plot numerical approximations to solutions of initial value problems.

Summary. The linear equation $y' + p(t)y = g(t)$ has several nice properties that can be summarized in the following statements:

1. Assuming that the coefficients are continuous, there is a general solution, containing an arbitrary constant, that includes all solutions of the differential equation. A particular solution that satisfies a given initial condition can be picked out by choosing the proper value for the arbitrary constant.
2. There is an expression for the solution, namely, equation (7) or equation (8). Moreover, although it involves two integrations, the expression is an explicit one for the solution $y = \phi(t)$ rather than an equation that defines ϕ implicitly.
3. The possible points of discontinuity, or singularities, of the solution can be identified (without solving the problem) merely by finding the points of discontinuity of the coefficients. Thus, if the coefficients are continuous for all t, then the solution also exists and is differentiable for all t.

None of these statements are true, in general, of nonlinear equations. Although a nonlinear equation may well have a solution involving an arbitrary constant, there may also be other solutions. There is no general formula for solutions of nonlinear equations. If you are able to integrate a nonlinear equation, you are likely to obtain an equation defining solutions implicitly rather than explicitly. Finally, the singularities of solutions of nonlinear equations can usually be found only by solving the equation and examining the solution. It is likely that the singularities will depend on the initial condition as well as on the differential equation.

Problems

In each of Problems 1 through 4, determine (without solving the problem) an interval in which the solution of the given initial value problem is certain to exist.

1. $(t-3)y' + (\ln t)y = 2t, \quad y(1) = 2$
2. $y' + (\tan t)y = \sin t, \quad y(\pi) = 0$
3. $(4-t^2)y' + 2ty = 3t^2, \quad y(-3) = 1$
4. $(\ln t)y' + y = \cot t, \quad y(2) = 3$

In each of Problems 5 through 8, state where in the ty-plane the hypotheses of Theorem 2.4.2 are satisfied.

5. $y' = (1 - t^2 - y^2)^{1/2}$
6. $y' = \dfrac{\ln|ty|}{1 - t^2 + y^2}$
7. $y' = (t^2 + y^2)^{3/2}$
8. $y' = \dfrac{1+t^2}{3y - y^2}$

In each of Problems 9 through 12, solve the given initial value problem and determine how the interval in which the solution exists depends on the initial value y_0.

9. $y' = -4t/y, \quad y(0) = y_0$
10. $y' = 2ty^2, \quad y(0) = y_0$
11. $y' + y^3 = 0, \quad y(0) = y_0$
12. $y' = \dfrac{t^2}{y(1+t^3)}, \quad y(0) = y_0$

In each of Problems 13 through 16, draw a direction field and plot (or sketch) several solutions of the given differential equation. Describe how solutions appear to behave as t increases and how their behavior depends on the initial value y_0 when $t = 0$.

G 13. $y' = ty(3 - y)$

G 14. $y' = y(3 - ty)$

G 15. $y' = -y(3 - ty)$

G 16. $y' = t - 1 - y^2$

17. Consider the initial value problem $y' = y^{1/3}$, $y(0) = 0$ from Example 3 in the text.
    a. Is there a solution that passes through the point $(1, 1)$? If so, find it.
    b. Is there a solution that passes through the point $(2, 1)$? If so, find it.
    c. Consider all possible solutions of the given initial value problem. Determine the set of values that these solutions have at $t = 2$.
18. a. Verify that both $y_1(t) = 1 - t$ and $y_2(t) = -t^2/4$ are solutions of the initial value problem

$$y' = \frac{-t + \sqrt{t^2 + 4y}}{2}, \quad y(2) = -1.$$

Where are these solutions valid?

b. Explain why the existence of two solutions of the given problem does not contradict the uniqueness part of Theorem 2.4.2.

c. Show that $y = ct + c^2$, where c is an arbitrary constant, satisfies the differential equation in part a for $t \geq -2c$. If $c = -1$, the initial condition is also satisfied, and the solution $y = y_1(t)$ is obtained. Show that there is no choice of c that gives the second solution $y = y_2(t)$.

19. **a.** Show that $\phi(t) = e^{2t}$ is a solution of $y' - 2y = 0$ and that $y = c\phi(t)$ is also a solution of this equation for any value of the constant c.

b. Show that $\phi(t) = 1/t$ is a solution of $y' + y^2 = 0$ for $t > 0$, but that $y = c\phi(t)$ is not a solution of this equation unless $c = 0$ or $c = 1$. Note that the equation of part b is nonlinear, while that of part a is linear.

20. Show that if $y = \phi(t)$ is a solution of $y' + p(t)y = 0$, then $y = c\phi(t)$ is also a solution for any value of the constant c.

21. Let $y = y_1(t)$ be a solution of

$$y' + p(t)y = 0, \tag{27}$$

and let $y = y_2(t)$ be a solution of

$$y' + p(t)y = g(t). \tag{28}$$

Show that $y = y_1(t) + y_2(t)$ is also a solution of equation (28).

22. **a.** Show that the solution (7) of the general linear equation (1) can be written in the form

$$y = cy_1(t) + y_2(t), \tag{29}$$

where c is an arbitrary constant.

b. Show that y_1 is a solution of the differential equation

$$y' + p(t)y = 0, \tag{30}$$

corresponding to $g(t) = 0$.

c. Show that y_2 is a solution of the full linear equation (1). We see later (for example, in Section 3.5) that solutions of higher-order linear equations have a pattern similar to equation (29).

Bernoulli Equations. Sometimes it is possible to solve a nonlinear equation by making a change of the dependent variable that converts it into a linear equation. The most important such equation has the form

$$y' + p(t)y = q(t)y^n,$$

and is called a Bernoulli equation after Jakob Bernoulli. Problems 23 and 25 deal with equations of this type.

23. **a.** Solve Bernoulli's equation when $n = 0$; when $n = 1$.

b. Show that if $n \neq 0, 1$, then the substitution $v = y^{1-n}$ reduces Bernoulli's equation to a linear equation. This method of solution was formulated by Leibniz in 1696.

In each of Problems 24 through 25, the given equation is a Bernoulli equation. In each case solve it by using the substitution mentioned in Problem 23b.

24. $y' = ry - ky^2$, $r > 0$ and $k > 0$. This equation is important in population dynamics and is discussed in detail in Section 2.5.

25. $y' = \epsilon y - \sigma y^3$, $\epsilon > 0$ and $\sigma > 0$. This equation occurs in the study of the stability of fluid flow.

Discontinuous Coefficients. Linear differential equations sometimes occur in which one or both of the functions p and g have jump discontinuities. If t_0 is such a point of discontinuity, then it is necessary to solve the equation separately for $t < t_0$ and $t > t_0$. Afterward, the two solutions are matched so that y is continuous at t_0; this is accomplished by a proper choice of the arbitrary constants. The following two problems illustrate this situation. Note in each case that it is impossible also to make y' continuous at t_0.

26. Solve the initial value problem

$$y' + 2y = g(t), \quad y(0) = 0,$$

where

$$g(t) = \begin{cases} 1, & 0 \leq t \leq 1, \\ 0, & t > 1. \end{cases}$$

27. Solve the initial value problem

$$y' + p(t)y = 0, \quad y(0) = 1,$$

where

$$p(t) = \begin{cases} 2, & 0 \leq t \leq 1, \\ 1, & t > 1. \end{cases}$$

2.5 Autonomous Differential Equations and Population Dynamics

An important class of first-order equations consists of those in which the independent variable does not appear explicitly. Such equations are called **autonomous** and have the form

$$dy/dt = f(y). \tag{1}$$

We will discuss these equations in the context of the growth or decline of the population of a given species, an important issue in fields ranging from medicine to ecology to global economics. A number of other applications are mentioned in some of the problems. Recall that in Sections 1.1 and 1.2 we considered the special case of equation (1) in which $f(y) = ay + b$.

Equation (1) is separable, so the discussion in Section 2.2 is applicable to it, but the main purpose of this section is to show how geometric methods can be used to obtain important qualitative information directly from the differential equation without solving the equation. Of fundamental importance in this effort are the concepts of stability and instability of solutions of differential equations. These ideas were introduced informally in Chapter 1, but without using this terminology. They are discussed further here and will be examined in greater depth and in a more general setting in Chapter 9.

Exponential Growth. Let $y = \phi(t)$ be the population of the given species at time t. The simplest hypothesis concerning the variation of population is that the rate of change of y is proportional[10] to the current value of y; that is,

$$\frac{dy}{dt} = ry, \tag{2}$$

where the constant of proportionality r is called the **rate of growth** or **decline**, depending on whether r is positive or negative. Here, we assume that the population is growing, so $r > 0$.

Solving equation (2) subject to the initial condition[11]

$$y(0) = y_0, \tag{3}$$

we obtain

$$y = y_0 e^{rt}. \tag{4}$$

Thus the mathematical model consisting of the initial value problem (1), (2) with $r > 0$ predicts that the population will grow exponentially for all time, as shown in Figure 2.5.1 for several values of y_0. Under ideal conditions, equation (4) has been observed to be reasonably accurate for many populations, at least for limited periods of time. However, it is clear that such ideal conditions cannot continue indefinitely; eventually, limitations on space, food supply, or other resources will reduce the growth rate and bring an end to uninhibited exponential growth.

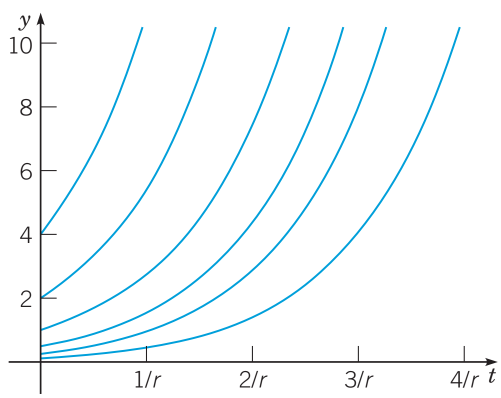


FIGURE 2.5.1 Exponential growth: y versus t for $dy/dt = ry$ $(r > 0)$.

Logistic Growth. To take account of the fact that the growth rate actually depends on the population, we replace the constant r in equation (2) by a function $h(y)$ and thereby obtain the modified equation

$$\frac{dy}{dt} = h(y)y. \tag{5}$$

We now want to choose $h(y)$ so that $h(y) \cong r > 0$ when y is small, $h(y)$ decreases as y grows larger, and $h(y) < 0$ when y is sufficiently large. The simplest function that has these properties is $h(y) = r - ay$, where a is also a positive constant. Using this function in equation (5), we obtain

$$\frac{dy}{dt} = (r - ay)y. \tag{6}$$

[10] It was apparently the British economist Thomas Malthus (1766–1834) who first observed that many biological populations increase at a rate proportional to the population. His first paper on populations appeared in 1798.

[11] In this section, because the unknown function is a population, we assume $y_0 > 0$.

Equation (6) is known as the Verhulst[12] equation or the **logistic equation**. It is often convenient to write the logistic equation in the equivalent form

$$\frac{dy}{dt} = r\left(1 - \frac{y}{K}\right)y, \tag{7}$$

where $K = r/a$. In this form, the constant r is called the **intrinsic growth rate**—that is, the growth rate in the absence of any limiting factors. The interpretation of K will become clear shortly.

We will investigate the solutions of equation (7) in some detail later in this section. Before doing that, however, we will show how you can easily draw a *qualitatively correct* sketch of the solutions. The same methods also apply to the more general equation (1).

We first seek solutions of equation (7) of the simplest possible type—that is, constant functions. For such a solution $dy/dt = 0$ for all t, so any constant solution of equation (7) must satisfy the algebraic equation

$$r\left(1 - \frac{y}{K}\right)y = 0.$$

Thus the constant solutions are $y = \phi_1(t) = 0$ and $y = \phi_2(t) = K$. These solutions are called **equilibrium solutions** of equation (7) because they correspond to no change or variation in the value of y as t increases. In the same way, any equilibrium solutions of the more general equation (1) can be found by locating the roots of $f(y) = 0$. The zeros of $f(y)$ are also called **critical points**.

To visualize other solutions of equation (7) and to sketch their graphs quickly, we start by drawing the graph of $f(y)$ versus y. In the case of equation (7), $f(y) = r(1 - y/K)y$, so the graph is the parabola shown in Figure 2.5.2. The intercepts are $(0, 0)$ and $(K, 0)$, corresponding to the critical points of equation (7), and the vertex of the parabola is $(K/2, rK/4)$. Observe that $dy/dt > 0$ for $0 < y < K$. Therefore, y is an increasing function of t when y is in this interval; this is indicated by the rightward-pointing arrows near the y-axis in Figure 2.5.2. Similarly, if $y > K$, then $dy/dt < 0$; hence y is decreasing, as indicated by the leftward-pointing arrow in Figure 2.5.2.

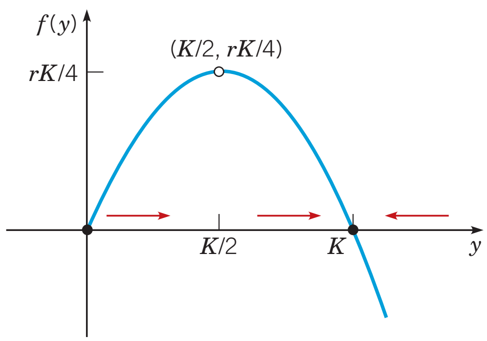


FIGURE 2.5.2 $f(y)$ versus y for $dy/dt = r(1-y/K)y$.

In this context the y-axis is often called the **phase line**, and it is reproduced in its more customary vertical orientation in Figure 2.5.3*a*. The dots at $y = 0$ and $y = K$ are the critical points, or equilibrium solutions. The arrows again indicate that y is increasing whenever $0 < y < K$ and that y is decreasing whenever $y > K$.

[12]Pierre F. Verhulst (1804–1849) was a Belgian mathematician who introduced equation (6) as a model for human population growth in 1838. He referred to it as logistic growth, so equation (6) is often called the logistic equation. He was unable to test the accuracy of his model because of inadequate census data, and it did not receive much attention until many years later. Reasonable agreement with experimental data was demonstrated by R. Pearl (1930) for *Drosophila melanogaster* (fruit fly) populations and by G. F. Gause (1935) for *Paramecium* and *Tribolium* (flour beetle) populations.

Further, from Figure 2.5.2, note that if y is near zero or K, then the slope $f(y)$ is near zero, so the solution curves are relatively flat. They become steeper as the value of y leaves the neighborhood of zero or K.

To sketch the graphs of solutions of equation (7) in the ty-plane, we start with the equilibrium solutions $y = \phi_1(t) = 0$ and $y = \phi_2(t) = K$; then we draw other curves that are increasing when $0 < y < K$, decreasing when $y > K$, and flatten out as y approaches either of the values 0 or K. Thus the graphs of solutions of equation (7) must have the general shape shown in Figure 2.5.3*b*, regardless of the values of r and K.

Figure 2.5.3*b* may seem to show that other solutions intersect the equilibrium solution $y = K$, but is this really possible? No, the uniqueness part of Theorem 2.4.2, the fundamental existence and uniqueness theorem, states that only one solution can pass through a given point in the ty-plane. Thus, although other solutions may be asymptotic to the equilibrium solution as $t \to \infty$, they cannot intersect it at any finite time. Consequently, a solution that starts in the interval $0 < y < K$ remains in this interval for all time, and similarly for a solution that starts in $K < y < \infty$.

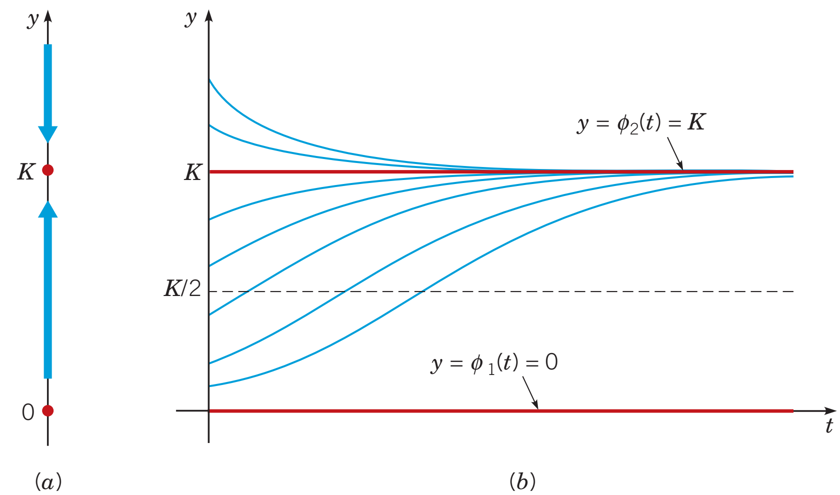


FIGURE 2.5.3 Logistic growth: $dy/dt = r(1 - y/K)y$. (*a*) The phase line. (*b*) Plots of y versus t.

To carry the investigation one step further, we can determine the concavity of the solution curves and the location of inflection points by finding d^2y/dt^2. From the differential equation (1), we obtain (using the chain rule)

$$\frac{d^2y}{dt^2} = \frac{d}{dt}\frac{dy}{dt} = \frac{d}{dt}f(y) = f'(y)\frac{dy}{dt} = f'(y)f(y). \tag{8}$$

The graph of y versus t is concave up when $y'' > 0$—that is, when f and f' have the same sign. Similarly, it is concave down when $y'' < 0$, which occurs when f and f' have opposite signs. The signs of f and f' can be easily identified from the graph of $f(y)$ versus y. Inflection points may occur when $f'(y) = 0$.

In the case of equation (7), solutions are concave up for $0 < y < K/2$ where f is positive and increasing (see Figure 2.5.2), so that both f and f' are positive. Solutions are also concave up for $y > K$ where f is negative and decreasing (both f and f' are negative). For $K/2 < y < K$, solutions are concave down since here f is positive and decreasing, so f is positive but f' is negative. There is an inflection point whenever the graph of y versus t crosses the line $y = K/2$. The graphs in Figure 2.5.3*b* exhibit these properties.

Finally, observe that K is the upper bound that is approached, but not exceeded, by growing populations starting below this value. Thus it is natural to refer to K as the **saturation level**, or the **environmental carrying capacity**, for the given species.

A comparison of Figures 2.5.1 and 2.5.3*b* reveals that solutions of the nonlinear equation (7) are strikingly different from those of the linear equation (1), at least for large values of t. Regardless of the value of K—that is, no matter how small the nonlinear term in equation (7)—solutions of that equation approach a finite value as $t \to \infty$, whereas solutions of equation (1) grow (exponentially) without bound as $t \to \infty$. Thus even a tiny nonlinear term in the differential equation (7) has a decisive effect on the solution for large t.

In many situations it is sufficient to have the qualitative information about a solution $y = \phi(t)$ of equation (7) that is shown in Figure 2.5.3*b*. This information was obtained entirely from the graph of $f(y)$ versus y and without solving the differential equation (7). However, if we wish to have a more detailed description of logistic growth—for example, if we wish to know the value of the population at some particular time—then we must solve equation (7) subject to the initial condition (3). Provided that $y \neq 0$ and $y \neq K$, we can write equation (7) in the form

$$\frac{dy}{(1 - y/K)y} = r\,dt.$$

Using a partial fraction expansion on the left-hand side, we have

$$\left(\frac{1}{y} + \frac{1/K}{1 - y/K}\right)dy = r\,dt.$$

Then, by integrating both sides, we obtain

$$\ln|y| - \ln\left|1 - \frac{y}{K}\right| = rt + c, \tag{9}$$

where c is an arbitrary constant of integration to be determined from the initial condition $y(0) = y_0$. We have already noted that if $0 < y_0 < K$, then y remains in this interval for all time. Thus in this case we can remove the absolute value bars in equation (9), and by taking the exponential of both sides, we find that

$$\frac{y}{1 - (y/K)} = Ce^{rt}, \tag{10}$$

where $C = e^c$. In order to satisfy the initial condition $y(0) = y_0$, we must choose $C = y_0/(1-(y_0/K))$. Using this value for C in equation (10) and solving for y (see Problem 10), we obtain

$$y = \frac{y_0 K}{y_0 + (K - y_0)e^{-rt}}. \tag{11}$$

We have derived the solution (11) under the assumption that $0 < y_0 < K$. If $y_0 > K$, then the details of dealing with equation (9) are only slightly different, and we leave it to you to show that equation (11) is also valid in this case. Finally, note that equation (11) also contains the equilibrium solutions $y = \phi_1(t) = 0$ and $y = \phi_2(t) = K$ corresponding to the initial conditions $y_0 = 0$ and $y_0 = K$, respectively.

All the qualitative conclusions that we reached earlier by geometrical reasoning can be confirmed by examining the solution (11). In particular, if $y_0 = 0$, then equation (11) requires that $y(t) = 0$ for all t. If $y_0 > 0$, and if we let $t \to \infty$ in equation (11), then we obtain

$$\lim_{t\to\infty} y(t) = \frac{y_0 K}{y_0} = K.$$

Thus, for each $y_0 > 0$, the solution approaches the equilibrium solution $y = \phi_2(t) = K$ asymptotically as $t \to \infty$. Therefore, we say that the constant solution $\phi_2(t) = K$ is an **asymptotically stable solution** of equation (7) or that the point $y = K$ is an asymptotically stable equilibrium or critical point. After a long time, the population is close to the saturation level K regardless of the initial population size, as long as it is positive. Other solutions approach the equilibrium solution more rapidly as r increases.

On the other hand, the situation for the equilibrium solution $y = \phi_1(t) = 0$ is quite different. Even solutions that start very near zero grow as t increases and, as we have seen, approach K as $t \to \infty$. We say that $\phi_1(t) = 0$ is an **unstable equilibrium solution** or that $y = 0$ is an unstable equilibrium or critical point. This means that the only way to guarantee that the solution remains near zero is to make sure its initial value is *exactly* equal to zero.

EXAMPLE 1

The logistic model has been applied to the natural growth of the halibut population in certain areas of the Pacific Ocean.[13] Let y, measured in kilograms, be the biomass, that is, the total mass, of the halibut population, at time t. The parameters in the logistic equation are estimated to have the values $r = 0.71/\text{year}$ and $K = 80.5 \times 10^6$ kg. If the initial biomass is $y_0 = 0.25K$, find the biomass 2 years later. Also find the time τ for which $y(\tau) = 0.75K$.

Solution:

It is convenient to scale the solution (11) to the carrying capacity K; thus we write equation (11) in the form

$$\frac{y}{K} = \frac{y_0/K}{(y_0/K) + (1 - y_0/K)e^{-rt}}. \tag{12}$$

Using the data given in the problem, we find that

$$\frac{y(2)}{K} = \frac{0.25}{0.25 + 0.75e^{-1.42}} \cong 0.5797.$$

Consequently, $y(2) \cong 46.7 \times 10^6$ kg.

To find τ, the time when $y_0/K = 0.75$ we first solve equation (12) for t, obtaining

$$e^{-rt} = \frac{(y_0/K)(1 - y/K)}{(y/K)\big(1 - y_0/K\big)};$$

hence

$$t = -\frac{1}{r}\ln\left(\frac{(y_0/K)(1 - y/K)}{(y/K)\big(1 - y_0/K\big)}\right). \tag{13}$$

Using the given values of r and y_0/K and setting $y/K = 0.75$, we find that

$$\tau = -\frac{1}{0.71}\ln\frac{(0.25)(0.25)}{(0.75)(0.75)} = \frac{1}{0.71}\ln 9 \cong 3.095 \text{ years}.$$

The graphs of y/K versus t for the given parameter values and for several initial conditions are shown in Figure 2.5.4. The green curve corresponds to the initial condition $y_0 = 0.25K$.

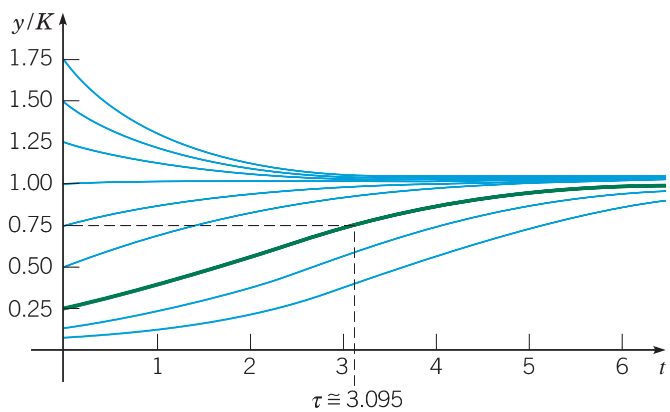


FIGURE 2.5.4 y/K versus t for population model of halibut in the Pacific Ocean. The green curve satisfies the initial condition $y(0)/K = 0.25$. The solution with $y(0) = 0.25$ reaches 75% of the carrying capacity at time $t = \tau 3.095$ years.

[13] A good source of information on the population dynamics and economics involved in making efficient use of a renewable resource, with particular emphasis on fisheries, is the book by Clark listed in the references at the end of this chapter. The parameter values used here are given on page 53 of this book and were obtained from a study by H. S. Mohring.

A Critical Threshold. We now turn to a consideration of the equation

$$\frac{dy}{dt} = -r\left(1 - \frac{y}{T}\right)y, \tag{14}$$

where r and T are given positive constants. Observe that (except for replacing the parameter K by T) this equation differs from the logistic equation (7) only in the presence of the minus sign on the right-hand side. However, as we will see, the solutions of equation (14) behave very differently from those of equation (7).

For equation (14) the graph of $f(y)$ versus y is the parabola shown in Figure 2.5.5. The intercepts on the y-axis are the critical points $y = 0$ and $y = T$, corresponding to the equilibrium solutions $y = \phi_1(t) = 0$ and $y = \phi_2(t) = T$. If $0 < y < T$, then $dy/dt < 0$, and y is positive and decreases as t increases. Thus $\phi_1(t) = 0$ is an asymptotically stable equilibrium solution. On the other hand, if $y > T$, then $dy/dt > 0$, so that y is positive and increasing as t increases; thus $\phi_2(t) = T$ is an unstable equilibrium solution.

Furthermore, the concavity of solutions can be determined by looking at the sign of $y'' = f'(y)f(y)$; see equation (8). Figure 2.5.5 clearly shows that $f'(y)$ is negative for $0 < y < T/2$ and positive for $T/2 < y < T$, so the graph of y versus t is concave up and concave down, respectively, in these intervals. Also, $f'(y)$ and $f(y)$ are both positive for $y > T$, so the graph of y versus t is also concave up there.

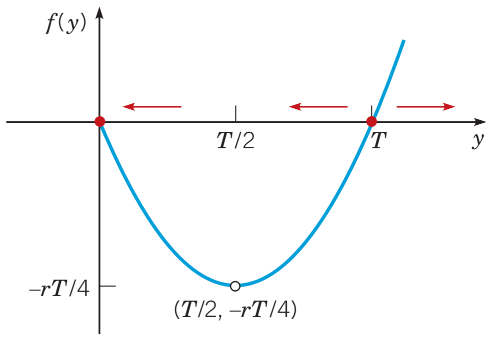


FIGURE 2.5.5 $f(y)$ versus y for $dy/dt = -r(1 - y/T)y$.

Figure 2.5.6(a) shows the phase line (the y-axis) for equation (14). The dots at $y = 0$ and $y = T$ are the critical points, or equilibrium solutions, and the arrows indicate where solutions are either increasing or decreasing.

Solution curves of equation (14) can now be sketched quickly, as follows. First draw the equilibrium solutions $y = \phi_1(t) = 0$ and $y = \phi_2(t) = T$. Then sketch curves in the strip $0 < y < T$ that are decreasing as t increases and change concavity as they cross the line $y = T/2$. Next draw some curves above $y = T$ that increase more and more steeply as t and y increase. Make sure that all curves become flatter as y approaches either zero or T. The result is Figure 2.5.6(b), which is a qualitatively accurate sketch of solutions of equation (14) for any r and T. From this figure it appears that as time increases, y either approaches zero or grows without bound, depending on whether the initial value y_0 is less than or greater than T. Thus T is a **threshold level**, below which growth does not occur.

We can confirm the conclusions that we have reached through geometrical reasoning by solving the differential equation (14). This can be done by separating the variables and integrating, just as we did for equation (7). However, if we note that equation (14) can be obtained from equation (7) by replacing K by T and r by $-r$, then we can make the same substitutions in the solution (11) and thereby obtain

$$y = \frac{y_0 T}{y_0 + (T - y_0)e^{rt}}, \tag{15}$$

which is the solution of equation (14) subject to the initial condition $y(0) = y_0$.

If $0 < y_0 < T$, then it follows from equation (15) that $y \to 0$ as $t \to \infty$. This agrees with our qualitative geometric analysis. If $y_0 > T$, then the denominator on the right-hand

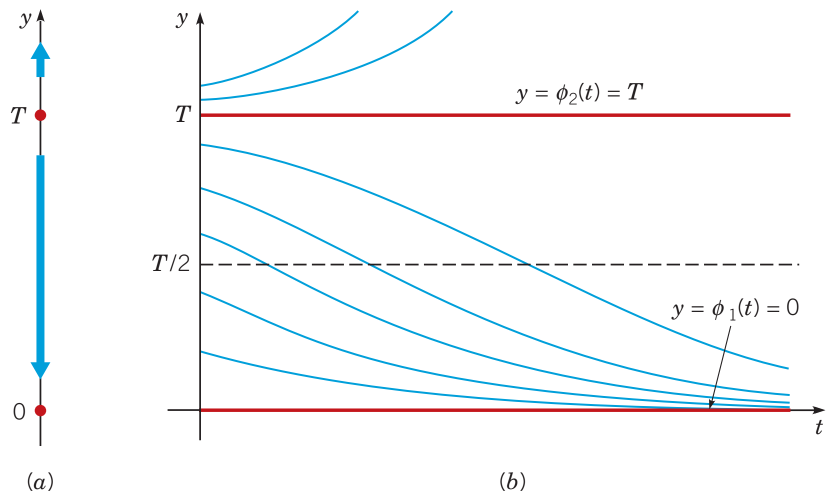


FIGURE 2.5.6 Growth with a threshold: $dy/dt = -r(1 - y/T)y$; $y = T$ is an asymptotically unstable equilibrium, while $y = 0$ is asymptotically stable. (a) The phase line. (b) Plots of y versus t.

side of equation (15) is zero for a certain finite value of t. We denote this value by t^* and calculate it from

$$y_0 - (y_0 - T)e^{rt^*} = 0,$$

which gives (see Problem 12)

$$t^* = \frac{1}{r} \ln \frac{y_0}{y_0 - T}. \tag{16}$$

Thus, if the initial population y_0 is above the threshold T, the threshold model predicts that the graph of y versus t has a vertical asymptote at $t = t^*$; in other words, the population becomes unbounded in a finite time, whose value depends on y_0, T, and r. The existence and location of this asymptote were not apparent from the geometric analysis, so in this case the explicit solution yields additional important qualitative, as well as quantitative, information.

The populations of some species exhibit the threshold phenomenon. If too few are present, then the species cannot propagate itself successfully and the population becomes extinct. However, if the population is larger than the threshold level, then further growth occurs. Of course, the population cannot become unbounded, so eventually equation (14) must be modified to take this into account.

Critical thresholds also occur in other circumstances. For example, in fluid mechanics, equations of the form (7) or (14) often govern the evolution of a small disturbance y in a *laminar* (or smooth) fluid flow. For instance, if equation (14) holds and $y < T$, then the disturbance is damped out and the laminar flow persists. However, if $y > T$, then the disturbance grows larger and the laminar flow breaks up into a turbulent one. In this case T is referred to as the *critical amplitude*. Experimenters speak of keeping the disturbance level in a wind tunnel low enough so that they can study laminar flow over an airfoil, for example.

Logistic Growth with a Threshold. As we mentioned in the last subsection, the threshold model (14) may need to be modified so that unbounded growth does not occur when y is above the threshold T. The simplest way to do this is to introduce another factor that will have the effect of making dy/dt negative when y is large. Thus we consider

$$\frac{dy}{dt} = -r\left(1 - \frac{y}{T}\right)\left(1 - \frac{y}{K}\right)y, \tag{17}$$

where $r > 0$ and $0 < T < K$.

The graph of $f(y)$ versus y is shown in Figure 2.5.7. In this problem there are three critical points, $y = 0$, $y = T$, and $y = K$, corresponding to the equilibrium solutions $y = \phi_1(t) = 0$, $y = \phi_2(t) = T$, and $y = \phi_3(t) = K$, respectively. From Figure 2.5.7 we observe that $dy/dt > 0$ for $T < y < K$, and consequently y is increasing there. Further, $dy/dt < 0$ for $y < T$ and for $y > K$, so y is decreasing in these intervals. Consequently, the equilibrium solutions $y = \phi_1(t) = 0$ and $y = \phi_3(t) = K$ are asymptotically stable, and the solution $y = \phi_2(t) = T$ is unstable.

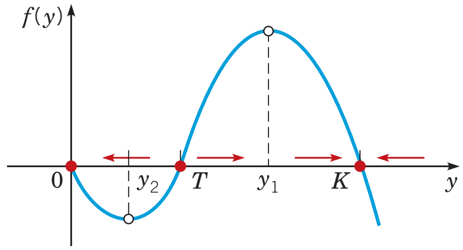


FIGURE 2.5.7 $f(y)$ versus y for $dy/dt = -r(1 - y/T)(1 - y/K)y$.

The phase line for equation (17) is shown in Figure 2.5.8*a*, and the graphs of some solutions are sketched in Figure 2.5.8*b*. You should make sure that you understand the relation between these two figures, as well as the relation between Figures 2.5.7 and 2.5.8*a*. From Figure 2.5.8*b* we see that if y starts below the threshold T, then y declines to ultimate extinction. On the other hand, if y starts above T, then y eventually approaches the carrying capacity K. The inflection points on the graphs of y versus t in Figure 2.5.8*b* correspond to the maximum and minimum points, y_1 and y_2, respectively, on the graph of $f(y)$ versus y in Figure 2.5.7. These values can be obtained by differentiating the right-hand side of equation (17) with respect to y, setting the result equal to zero, and solving for y. We obtain

$$y_{1,2} = (K + T \pm \sqrt{K^2 - KT + T^2})/3, \tag{18}$$

where the plus sign yields y_1 and the minus sign y_2.

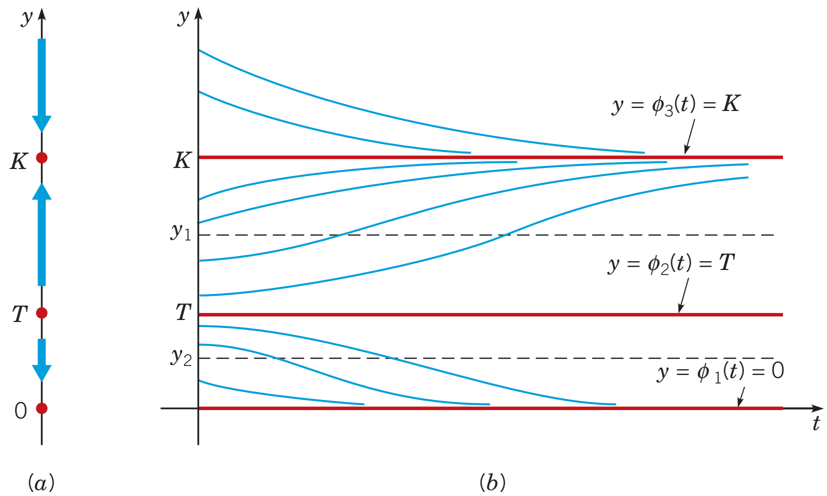


FIGURE 2.5.8 Logistic growth with a threshold: $dy/dt = -r(1 - y/T)(1 - y/K)y$; $y = \phi_1(t) = 0$ and $y = \phi_3(t) = K$ are asymptotically stable equilibria and $y = \phi_2(t) = T$ is an asymptotically unstable equilibrium. (*a*) The phase line. (*b*) Plots of y versus t.

A model of this general sort apparently describes the population of the passenger pigeon,[14] which was present in the United States in vast numbers until the late nineteenth century. It was heavily hunted for food and for sport, and consequently its numbers were drastically reduced by the 1880s. Unfortunately, the passenger pigeon could apparently breed successfully only when present in a large concentration, corresponding to a relatively high threshold T. Although a reasonably large number of individual birds remained alive in the late 1880s, there were not enough in any one place to permit successful breeding, and the population rapidly declined to extinction. The last passenger pigeon died in 1914. The precipitous decline in the passenger pigeon population from huge numbers to extinction in a few decades was one of the early factors contributing to a concern for conservation in this country.

[14] See, for example, Oliver L. Austin, Jr., *Birds of the World* (New York: Golden Press, 1983), pp. 143–145.

Problems

Problems 1 through 4 involve equations of the form $dy/dt = f(y)$. In each problem sketch the graph of $f(y)$ versus y, determine the critical (equilibrium) points, and classify each one as asymptotically stable or unstable. Draw the phase line, and sketch several graphs of solutions in the ty-plane.

1. $dy/dt = ay + by^2, \quad a > 0, \quad b > 0, \quad -\infty < y_0 < \infty$

2. $dy/dt = y(y-1)(y-2), \quad y_0 \geq 0$

3. $dy/dt = e^y - 1, \quad -\infty < y_0 < \infty$

4. $dy/dt = e^{-y} - 1, \quad -\infty < y_0 < \infty$

5. Semistable Equilibrium Solutions. Sometimes a constant equilibrium solution has the property that solutions lying on one side of the equilibrium solution tend to approach it, whereas solutions lying on the other side depart from it (see Figure 2.5.9). In this case the equilibrium solution is said to be **semistable**.

a. Consider the equation

$$dy/dt = k(1-y)^2, \tag{19}$$

where k is a positive constant. Show that $y = 1$ is the only critical point, with the corresponding equilibrium solution $\phi(t) = 1$.

b. Sketch $f(y)$ versus y. Show that y is increasing as a function of t for $y < 1$ and also for $y > 1$. The phase line has upward-pointing arrows both below and above $y = 1$. Thus solutions below the equilibrium solution approach it, and those above it grow farther away. Therefore, $\phi(t) = 1$ is semistable.

c. Solve equation (19) subject to the initial condition $y(0) = y_0$ and confirm the conclusions reached in part b.

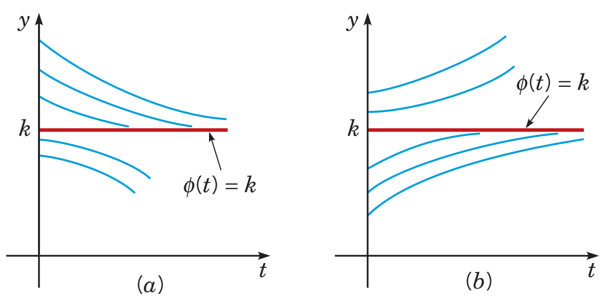


FIGURE 2.5.9 In both cases the equilibrium solution $\phi(t) = k$ is semistable. (*a*) $dy/dt \leq 0$; (*b*) $dy/dt \geq 0$.

Problems 6 through 9 involve equations of the form $dy/dt = f(y)$. In each problem sketch the graph of $f(y)$ versus y, determine the critical (equilibrium) points, and classify each one as asymptotically stable, unstable, or semistable (see Problem 5). Draw the phase line, and sketch several graphs of solutions in the ty-plane.

6. $dy/dt = y^2(y^2 - 1), \quad -\infty < y_0 < \infty$

7. $dy/dt = y(1 - y^2), \quad -\infty < y_0 < \infty$

8. $dy/dt = y^2(4 - y^2), \quad -\infty < y_0 < \infty$

9. $dy/dt = y^2(1 - y)^2, \quad -\infty < y_0 < \infty$

10. Complete the derivation of the explicit formula for the solution (11) of the logistic model by solving equation (10) for y.

11. In Example 1, complete the manipulations needed to arrive at equation (13). That is, solve the solution (11) for t.

12. Complete the derivation of the location of the vertical asymptote in the solution (15) when $y_0 > T$. That is, derive formula (16) by finding the value of t when the denominator of the right-hand side of equation (15) is zero.

13. Complete the derivation of formula (18) for the locations of the inflection points of the solution of the logistic growth model with a threshold (17). *Hint:* Follow the steps outlined on p. 66.

14. Consider the equation $dy/dt = f(y)$ and suppose that y_1 is a critical point—that is, $f(y_1) = 0$. Show that the constant equilibrium solution $\phi(t) = y_1$ is asymptotically stable if $f'(y_1) < 0$ and unstable if $f'(y_1) > 0$.

15. Suppose that a certain population obeys the logistic equation $dy/dt = ry(1 - (y/K))$.

a. If $y_0 = K/3$, find the time τ at which the initial population has doubled. Find the value of τ corresponding to $r = 0.025$ per year.

b. If $y_0/K = \alpha$, find the time T at which $y(T)/K = \beta$, where $0 < \alpha, \beta < 1$. Observe that $T \to \infty$ as $\alpha \to 0$ or as $\beta \to 1$. Find the value of T for $r = 0.025$ per year, $\alpha = 0.1$, and $\beta = 0.9$.

16. Another equation that has been used to model population growth is the Gompertz[15] equation

$$\frac{dy}{dt} = ry \ln\left(\frac{K}{y}\right),$$

where r and K are positive constants.

a. Sketch the graph of $f(y)$ versus y, find the critical points, and determine whether each is asymptotically stable or unstable.

b. For $0 \leq y \leq K$, determine where the graph of y versus t is concave up and where it is concave down.

c. For each y in $0 < y \leq K$, show that dy/dt as given by the Gompertz equation is never less than dy/dt as given by the logistic equation.

17. a. Solve the Gompertz equation

$$\frac{dy}{dt} = ry \ln\left(\frac{K}{y}\right),$$

subject to the initial condition $y(0) = y_0$.
Hint: You may wish to let $u = \ln(y/K)$.

b. For the data given in Example 1 in the text ($r = 0.71$ per year, $K = 80.5 \times 10^6$ kg, $y_0/K = 0.25$), use the Gompertz model to find the predicted value of $y(2)$.

c. For the same data as in part b, use the Gompertz model to find the time τ at which $y(\tau) = 0.75K$.

[15] Benjamin Gompertz (1779–1865) was an English actuary. He developed his model for population growth, published in 1825, in the course of constructing mortality tables for his insurance company.

18. A pond forms as water collects in a conical depression of radius a and depth h. Suppose that water flows in at a constant rate k and is lost through evaporation at a rate proportional to the surface area.

a. Show that the volume $V(t)$ of water in the pond at time t satisfies the differential equation

$$\frac{dV}{dt} = k - \alpha\pi(3a/\pi h)^{2/3}V^{2/3},$$

where α is the coefficient of evaporation.

b. Find the equilibrium depth of water in the pond. Is the equilibrium asymptotically stable?

c. Find a condition that must be satisfied if the pond is not to overflow.

Harvesting a Renewable Resource. Suppose that the population y of a certain species of fish (for example, tuna or halibut) in a given area of the ocean is described by the logistic equation

$$\frac{dy}{dt} = r\left(1 - \frac{y}{K}\right)y.$$

Although it is desirable to utilize this source of food, it is intuitively clear that if too many fish are caught, then the fish population may be reduced below a useful level and possibly even driven to extinction. Problems 19 and 20 explore some of the questions involved in formulating a rational strategy for managing the fishery.[16]

19. At a given level of effort, it is reasonable to assume that the rate at which fish are caught depends on the population y: the more fish there are, the easier it is to catch them. Thus we assume that the rate at which fish are caught is given by Ey, where E is a positive constant, with units of 1/time, that measures the total effort made to harvest the given species of fish. To include this effect, the logistic equation is replaced by

$$\frac{dy}{dt} = r\left(1 - \frac{y}{K}\right)y - Ey. \tag{20}$$

This equation is known as the **Schaefer model** after the biologist M. B. Schaefer, who applied it to fish populations.

a. Show that if $E < r$, then there are two equilibrium points, $y_1 = 0$ and $y_2 = K(1 - E/r) > 0$.

b. Show that $y = y_1$ is unstable and $y = y_2$ is asymptotically stable.

c. A sustainable yield Y of the fishery is a rate at which fish can be caught indefinitely. It is the product of the effort E and the asymptotically stable population y_2. Find Y as a function of the effort E; the graph of this function is known as the yield–effort curve.

d. Determine E so as to maximize Y and thereby find the **maximum sustainable yield** Y_m.

20. In this problem we assume that fish are caught at a constant rate h independent of the size of the fish population. Then y satisfies

$$\frac{dy}{dt} = r\left(1 - \frac{y}{K}\right)y - h. \tag{21}$$

The assumption of a constant catch rate h may be reasonable when y is large but becomes less so when y is small.

a. If $h < rK/4$, show that equation (21) has two equilibrium points y_1 and y_2 with $y_1 < y_2$; determine these points.

b. Show that y_1 is unstable and y_2 is asymptotically stable.

c. From a plot of $f(y)$ versus y, show that if the initial population $y_0 > y_1$, then $y \to y_2$ as $t \to \infty$, but that if $y_0 < y_1$, then y decreases as t increases. Note that $y = 0$ is not an equilibrium point, so if $y_0 < y_1$, then extinction will be reached in a finite time.

d. If $h > rK/4$, show that y decreases to zero as t increases, regardless of the value of y_0.

e. If $h = rK/4$, show that there is a single equilibrium point $y = K/2$ and that this point is semistable (see Problem 5). Thus the maximum sustainable yield is $h_m = rK/4$, corresponding to the equilibrium value $y = K/2$. Observe that h_m has the same value as Y_m in Problem 19d. The fishery is considered to be overexploited if y is reduced to a level below $K/2$.

Epidemics. The use of mathematical methods to study the spread of contagious diseases goes back at least to some work by Daniel Bernoulli in 1760 on smallpox. In more recent years many mathematical models have been proposed and studied for many different diseases.[17] Problems 21 through 23 deal with a few of the simpler models and the conclusions that can be drawn from them. Similar models have also been used to describe the spread of rumors and of consumer products.

21. Suppose that a given population can be divided into two parts: those who have a given disease and can infect others, and those who do not have it but are susceptible. Let x be the proportion of susceptible individuals and y the proportion of infectious individuals; then $x + y = 1$. Assume that the disease spreads by contact between sick and well members of the population and that the rate of spread dy/dt is proportional to the number of such contacts. Further, assume that members of both groups move about freely among each other, so the number of contacts is proportional to the product of x and y. Since $x = 1 - y$, we obtain the initial value problem

$$\frac{dy}{dt} = \alpha y(1 - y), \quad y(0) = y_0, \tag{22}$$

where α is a positive proportionality factor, and y_0 is the initial proportion of infectious individuals.

a. Find the equilibrium points for the differential equation (22) and determine whether each is asymptotically stable, semistable, or unstable.

b. Solve the initial value problem 22 and verify that the conclusions you reached in part a are correct. Show that $y(t) \to 1$ as $t \to \infty$, which means that ultimately the disease spreads through the entire population.

22. Some diseases (such as typhoid fever) are spread largely by *carriers*, individuals who can transmit the disease but who exhibit no overt symptoms. Let x and y denote the proportions of susceptibles and carriers, respectively, in the population. Suppose that carriers are identified and removed from the population at a rate β, so

$$\frac{dy}{dt} = -\beta y. \tag{23}$$

Suppose also that the disease spreads at a rate proportional to the product of x and y; thus

$$\frac{dx}{dt} = -\alpha xy. \tag{24}$$

a. Determine y at any time t by solving equation (23) subject to the initial condition $y(0) = y_0$.

b. Use the result of part a to find x at any time t by solving equation (24) subject to the initial condition $x(0) = x_0$.

c. Find the proportion of the population that escapes the epidemic by finding the limiting value of x as $t \to \infty$.

[16] An excellent treatment of this kind of problem, which goes far beyond what is outlined here, may be found in the book by Clark mentioned previously, especially in the first two chapters. Numerous additional references are given there.

[17] A standard source is the book by Bailey listed in the references. The models in Problems 21, 22, and 23 are discussed by Bailey in Chapters 5, 10, and 20, respectively.

23. Daniel Bernoulli's work in 1760 had the goal of appraising the effectiveness of a controversial inoculation program against smallpox, which at that time was a major threat to public health. His model applies equally well to any other disease that, once contracted and survived, confers a lifetime immunity.

Consider the cohort of individuals born in a given year ($t = 0$), and let $n(t)$ be the number of these individuals surviving t years later. Let $x(t)$ be the number of members of this cohort who have not had smallpox by year t and who are therefore still susceptible. Let β be the rate at which susceptibles contract smallpox, and let ν be the rate at which people who contract smallpox die from the disease. Finally, let $\mu(t)$ be the death rate from all causes other than smallpox. Then dx/dt, the rate at which the number of susceptibles declines, is given by

$$\frac{dx}{dt} = -(\beta + \mu(t))x. \tag{25}$$

The first term on the right-hand side of equation (25) is the rate at which susceptibles contract smallpox, and the second term is the rate at which they die from all other causes. Also

$$\frac{dn}{dt} = -\nu\beta x - \mu(t)n, \tag{26}$$

where dn/dt is the death rate of the entire cohort, and the two terms on the right-hand side are the death rates due to smallpox and to all other causes, respectively.

a. Let $z = x/n$, and show that z satisfies the initial value problem

$$\frac{dz}{dt} = -\beta z(1 - \nu z), \quad z(0) = 1. \tag{27}$$

Observe that the initial value problem (27) does not depend on $\mu(t)$.

b. Find $z(t)$ by solving equation (27).

c. Bernoulli estimated that $\nu = \beta = 1/8$. Using these values, determine the proportion of 20-year-olds who have not had smallpox.

Note: On the basis of the model just described and the best mortality data available at the time, Bernoulli calculated that if deaths due to smallpox could be eliminated ($\nu = 0$), then approximately 3 years could be added to the average life expectancy (in 1760) of 26 years, 7 months. He therefore supported the inoculation program.

Bifurcation Points. For an equation of the form

$$\frac{dy}{dt} = f(a, y), \tag{28}$$

where a is a real parameter, the critical points (equilibrium solutions) usually depend on the value of a. As a steadily increases or decreases, it often happens that at a certain value of a, called a **bifurcation point**, critical points come together, or separate, and equilibrium solutions may be either lost or gained. Bifurcation points are of great interest in many applications, because near them the nature of the solution of the underlying differential equation is undergoing an abrupt change. For example, in fluid mechanics a smooth (laminar) flow may break up and become turbulent. Or an axially loaded column may suddenly buckle and exhibit a large lateral displacement. Or, as the amount of one of the chemicals in a certain mixture is increased, spiral wave patterns of varying color may suddenly emerge in an originally quiescent fluid. Problems 24 through 26 describe three types of bifurcations that can occur in simple equations of the form (28).

24. Consider the equation

$$\frac{dy}{dt} = a - y^2. \tag{29}$$

a. Find all of the critical points for equation (29). Observe that there are no critical points if $a < 0$, one critical point if $a = 0$, and two critical points if $a > 0$.

G **b.** Draw the phase line in each case and determine whether each critical point is asymptotically stable, semistable, or unstable.

G **c.** In each case sketch several solutions of equation (29) in the ty-plane.

Note: If we plot the location of the critical points as a function of a in the ay-plane, we obtain Figure 2.5.10. This is called the **bifurcation diagram** for equation (29). The bifurcation at $a = 0$ is called a **saddle – node** bifurcation. This name is more natural in the context of second-order systems, which are discussed in Chapter 9.

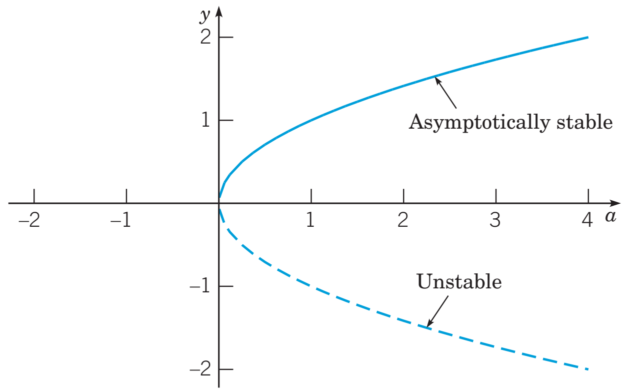


FIGURE 2.5.10 Bifurcation diagram for $y' = a - y^2$.

25. Consider the equation

$$\frac{dy}{dt} = ay - y^3 = y(a - y^2). \tag{30}$$

G **a.** Again consider the cases $a < 0$, $a = 0$, and $a > 0$. In each case find the critical points, draw the phase line, and determine whether each critical point is asymptotically stable, semistable, or unstable.

G **b.** In each case sketch several solutions of equation (30) in the ty-plane.

G **c.** Draw the bifurcation diagram for equation (30)—that is, plot the location of the critical points versus a.

Note: For equation (30) the bifurcation point at $a = 0$ is called a **pitchfork bifurcation**. Your diagram may suggest why this name is appropriate.

26. Consider the equation

$$\frac{dy}{dt} = ay - y^2 = y(a - y). \tag{31}$$

a. Again consider the cases $a < 0$, $a = 0$, and $a > 0$. In each case find the critical points, draw the phase line, and determine whether each critical point is asymptotically stable, semistable, or unstable.

b. In each case sketch several solutions of equation (31) in the ty-plane.

c. Draw the bifurcation diagram for equation (31).

Note: Observe that for equation (31) there are the same number of critical points for $a < 0$ and $a > 0$ but that their stability has changed. For $a < 0$ the equilibrium solution $y = 0$ is asymptotically stable and $y = a$ is unstable, while for $a > 0$ the situation is reversed. Thus there has been an **exchange of stability** as a passes through the bifurcation point $a = 0$. This type of bifurcation is called a **transcritical bifurcation**.

27. Chemical Reactions. A second-order chemical reaction involves the interaction (collision) of one molecule of a substance P with one molecule of a substance Q to produce one molecule of a new substance X; this is denoted by $P + Q \to X$. Suppose that p and q, where $p \neq q$, are the initial concentrations of P and Q, respectively, and let $x(t)$ be the concentration of X at time t. Then $p - x(t)$ and $q - x(t)$ are the concentrations of P and Q at time t, and the rate at which the reaction occurs is given by the equation

$$\frac{dx}{dt} = \alpha(p - x)(q - x), \tag{32}$$

where α is a positive constant.

a. If $x(0) = 0$, determine the limiting value of $x(t)$ as $t \to \infty$ without solving the differential equation. Then solve the initial value problem and find $x(t)$ for any t.

b. If the substances P and Q are the same, then $p = q$ and equation (32) is replaced by

$$\frac{dx}{dt} = \alpha(p - x)^2. \tag{33}$$

If $x(0) = 0$, determine the limiting value of $x(t)$ as $t \to \infty$ without solving the differential equation. Then solve the initial value problem and determine $x(t)$ for any t.

2.6 Exact Differential Equations and Integrating Factors

For first-order differential equations there are a number of integration methods that are applicable to various classes of problems. The most important of these are linear equations and separable equations, which we have discussed previously. Here, we consider a class of equations known as exact differential equations for which there is also a well-defined method of solution. Keep in mind, however, that the first-order differential equations that can be solved by elementary integration methods are rather special; most first-order equations cannot be solved in this way.

EXAMPLE 1

Solve the differential equation

$$2x + y^2 + 2xyy' = 0. \tag{1}$$

Solution:

The equation is neither linear nor separable, so the methods suitable for those types of equations are not applicable here. However, observe that the function $\psi(x, y) = x^2 + xy^2$ has the property that

$$2x + y^2 = \frac{\partial \psi}{\partial x}, \quad 2xy = \frac{\partial \psi}{\partial y}. \tag{2}$$

Therefore, the differential equation can be written as

$$\frac{\partial \psi}{\partial x} + \frac{\partial \psi}{\partial y}\frac{dy}{dx} = 0. \tag{3}$$

Assuming that y is a function of x, we can use the chain rule to write the left-hand side of equation (3) as $d\psi(x, y)/dx$. Then equation (3) has the form

$$\frac{d\psi}{dx}(x, y) = \frac{d}{dx}(x^2 + xy^2) = 0. \tag{4}$$

Integrating equation (4) we obtain

$$\psi(x, y) = x^2 + xy^2 = c, \tag{5}$$

where c is an arbitrary constant. The level curves of $\psi(x, y)$ are the integral curves of equation (1). Solutions of equation (1) are defined implicitly by equation (5).

In solving equation (1) the key step was the recognition that there is a function ψ that satisfies equations (2). More generally, let the differential equation

$$M(x, y) + N(x, y)y' = 0 \tag{6}$$

be given. Suppose that we can identify a function $\psi(x, y)$ such that

$$\frac{\partial \psi}{\partial x}(x, y) = M(x, y), \qquad \frac{\partial \psi}{\partial y}(x, y) = N(x, y), \tag{7}$$

and such that $\psi(x, y) = c$ defines $y = \phi(x)$ implicitly as a differentiable function of x.[18]

When there is a function $\psi(x, y)$ such that $\psi_x = M$ and $\psi_y = N$, we can write

$$M(x, y) + N(x, y)y' = \frac{\partial \psi}{\partial x} + \frac{\partial \psi}{\partial y}\frac{dy}{dx} = \frac{d}{dx}\psi(x, \phi(x))$$

and the differential equation (6) becomes

$$\frac{d}{dx}\psi(x, \phi(x)) = 0. \tag{8}$$

In this case equation (6) is said to be an **exact differential equation** because it can be expressed exactly as the derivative of a specific function. Solutions of equation (6), or the equivalent equation (8), are given implicitly by

$$\psi(x, y) = c, \tag{9}$$

where c is an arbitrary constant.

In Example 1 it was relatively easy to see that the differential equation was exact and, in fact, easy to find its solution, at least implicitly, by recognizing the required function ψ. For more complicated equations it may not be possible to do this so easily. How can we tell whether a given equation is exact, and if it is, how can we find the function $\psi(x, y)$? The following theorem answers the first question, and its proof provides a way of answering the second.

Theorem 2.6.1

Let the functions M, N, M_y, and N_x, where subscripts denote partial derivatives, be continuous in the rectangular[19] region $R: \alpha < x < \beta, \gamma < y < \delta$. Then equation (6)

$$M(x, y) + N(x, y)y' = 0$$

is an exact differential equation in R if and only if

$$M_y(x, y) = N_x(x, y) \tag{10}$$

at each point of R. That is, there exists a function ψ satisfying equations (7),

$$\psi_x(x, y) = M(x, y), \qquad \psi_y(x, y) = N(x, y),$$

if and only if M and N satisfy equation (10).

The proof of this theorem has two parts. First, we show that if there is a function ψ such that equations (7) are true, then it follows that equation (10) is satisfied. Computing M_y and N_x from equations (7), we obtain

$$M_y(x, y) = \psi_{xy}(x, y), \qquad N_x(x, y) = \psi_{yx}(x, y). \tag{11}$$

Since M_y and N_x are continuous, it follows that ψ_{xy} and ψ_{yx} are also continuous. This guarantees their equality, and equation (10) is valid.

We now show that if M and N satisfy equation (10), then equation (6) is exact. The proof involves the construction of a function ψ satisfying equations (7)

$$\psi_x(x, y) = M(x, y), \qquad \psi_y(x, y) = N(x, y).$$

[18] While a complete discussion of when $\psi(x, y) = c$ defines $y = \phi(x)$ implicitly as a differentiable function of x is beyond the scope and focus of this course, in general terms this condition is satisfied, locally, at points (x, y), where $\partial\psi/\partial y(x, y) \neq 0$. More details can be found in most books on advanced calculus.

[19] It is not essential that the region be rectangular, only that it be simply connected. In two dimensions this means that the region has no holes in its interior. Thus, for example, rectangular or circular regions are simply connected, but an annular region is not. More details can be found in most books on advanced calculus.

We begin by integrating the first of equations (7) with respect to x, holding y constant. We obtain

$$\psi(x, y) = Q(x, y) + h(y), \tag{12}$$

where $Q(x, y)$ is any differentiable function such that $Q_x = M$. For example, we might choose

$$Q(x, y) = \int_{x_0}^{x} M(s, y)\,ds, \tag{13}$$

where x_0 is some specified constant with $\alpha < x_0 < \beta$. The function h in equation (12) is an arbitrary differentiable function of y, playing the role of the arbitrary constant (with respect to x). Now we must show that it is always possible to choose $h(y)$ so that the second of equations (7) is satisfied—that is, $\psi_y = N$. By differentiating equation (12) with respect to y and setting the result equal to $N(x, y)$, we obtain

$$\psi_y(x, y) = \frac{\partial Q}{\partial y}(x, y) + h'(y) = N(x, y).$$

Then, solving for $h'(y)$, we have

$$h'(y) = N(x, y) - \frac{\partial Q}{\partial y}(x, y). \tag{14}$$

In order for us to determine $h(y)$ from equation (14), the right-hand side of equation (14), despite its appearance, must be a function of y only. One way to show that this is true is to show that its derivative with respect to x is zero. Thus we differentiate the right-hand side of equation (14) with respect to x, obtaining the expression

$$\frac{\partial N}{\partial x}(x, y) - \frac{\partial}{\partial x}\frac{\partial Q}{\partial y}(x, y). \tag{15}$$

By interchanging the order of differentiation in the second term of equation (15), we have

$$\frac{\partial N}{\partial x}(x, y) - \frac{\partial}{\partial y}\frac{\partial Q}{\partial x}(x, y),$$

or, since $Q_x = M$,

$$\frac{\partial N}{\partial x}(x, y) - \frac{\partial M}{\partial y}(x, y),$$

which is zero on account of equation (10). Hence, despite its apparent form, the right-hand side of equation (14) does not, in fact, depend on x. Then we find $h(y)$ by integrating equation (14) and, upon substituting this function in equation (12), we obtain the required function $\psi(x, y)$. This completes the proof of Theorem 2.6.1.

It is possible to obtain an explicit expression for $\psi(x, y)$ in terms of integrals (see Problem 13), but in solving specific exact equations, it is usually simpler and easier just to repeat the procedure used in the preceding proof. That is, after showing that $M_y = N_x$, integrate $\psi_x = M$ with respect to x, including an arbitrary function of $h(y)$ instead of an arbitrary constant, and then differentiate the result with respect to y and set it equal to N. Finally, use this last equation to solve for $h(y)$. The next example illustrates this procedure.

EXAMPLE 2

Solve the differential equation

$$(y\cos x + 2xe^y) + (\sin x + x^2e^y - 1)y' = 0. \tag{16}$$

Solution:

By calculating M_y and N_x, we find that

$$M_y(x, y) = \cos x + 2xe^y = N_x(x, y),$$

so the given equation is exact. Thus there is a $\psi(x, y)$ such that

$$\psi_x(x, y) = y \cos x + 2xe^y,$$
$$\psi_y(x, y) = \sin x + x^2 e^y - 1.$$

Integrating the first of these equations with respect to x, we obtain

$$\psi(x, y) = y \sin x + x^2 e^y + h(y). \tag{17}$$

Next, computing ψ_y from equation (17) and setting $\psi_y = N$ gives

$$\psi_y(x, y) = \sin x + x^2 e^y + h'(y) = \sin x + x^2 e^y - 1.$$

Thus $h'(y) = -1$ and $h(y) = -y$. The constant of integration can be omitted since any solution of the preceding differential equation is satisfactory; we do not require the most general one. Substituting for $h(y)$ in equation (17) gives

$$\psi(x, y) = y \sin x + x^2 e^y - y.$$

Hence solutions of equation (16) are given implicitly by

$$y \sin x + x^2 e^y - y = c. \tag{18}$$

EXAMPLE 3

Solve the differential equation

$$(3xy + y^2) + (x^2 + xy)y' = 0. \tag{19}$$

Solution:

We have

$$M_y(x, y) = 3x + 2y, \quad N_x(x, y) = 2x + y;$$

since $M_y \neq N_x$, the given equation is not exact. To see that it cannot be solved by the procedure described above, let us seek a function ψ such that

$$\psi_x(x, y) = 3xy + y^2, \quad \psi_y(x, y) = x^2 + xy. \tag{20}$$

Integrating the first of equations (20) with respect to x gives

$$\psi(x, y) = \frac{3}{2}x^2 y + xy^2 + h(y), \tag{21}$$

where h is an arbitrary function of y only. To try to satisfy the second of equations (20), we compute ψ_y from equation (21) and set it equal to N, obtaining

$$\frac{3}{2}x^2 + 2xy + h'(y) = x^2 + xy$$

or

$$h'(y) = -\frac{1}{2}x^2 - xy. \tag{22}$$

Since the right-hand side of equation (22) depends on x as well as y, it is impossible to solve equation (22) for $h(y)$. Thus there is no $\psi(x, y)$ satisfying both of equations (20).

Integrating Factors. It is sometimes possible to convert a differential equation that is not exact into an exact differential equation by multiplying the equation by a suitable integrating factor. Recall that this is the procedure that we used in solving linear differential equations in Section 2.1. To investigate the possibility of implementing this idea more generally, let us multiply the equation

$$M(x, y) + N(x, y)y' = 0 \tag{23}$$

by a function μ and then try to choose μ so that the resulting equation

$$\mu(x, y)M(x, y) + \mu(x, y)N(x, y)y' = 0 \tag{24}$$

is exact. By Theorem 2.6.1, equation (24) is exact if and only if

$$(\mu M)_y = (\mu N)_x. \tag{25}$$

Since M and N are given functions, equation (25) states that the integrating factor μ must satisfy the first-order partial differential equation

$$M\mu_y - N\mu_x + (M_y - N_x)\mu = 0. \tag{26}$$

If a function μ satisfying equation (26) can be found, then equation (24) will be exact. The solution of equation (24) can then be obtained by the method described in the first part of this section. The solution found in this way also satisfies equation (23), since the integrating factor μ can be canceled out of equation (24).

A partial differential equation of the form (26) may have more than one solution; if this is the case, any such solution may be used as an integrating factor of equation (23). This possible nonuniqueness of the integrating factor is illustrated in Example 4.

Unfortunately, equation (26), which determines the integrating factor μ, is ordinarily at least as hard to solve as the original equation (23). Therefore, although in principle integrating factors are powerful tools for solving differential equations, in practice they can be found only in special cases. The most important situations in which simple integrating factors can be found occur when μ is a function of only one of the variables x or y, instead of both.

Let us determine conditions on M and N so that equation (23) has an integrating factor μ that depends on x only. If we assume that μ is a function of x only, then the partial derivative μ_x reduces to the ordinary derivative $d\mu/dx$ and $\mu_y = 0$. Making these substitutions in equation (26), we find that

$$\frac{d\mu}{dx} = \frac{M_y - N_x}{N}\mu. \tag{27}$$

If $(M_y - N_x)/N$ is a function of x only, then there is an integrating factor μ that also depends only on x; further, $\mu(x)$ can be found by solving differential equation (27), which is both linear and separable.

A similar procedure can be used to determine a condition under which equation (23) has an integrating factor depending only on y; see Problem 17.

EXAMPLE 4

Find an integrating factor for the equation

$$(3xy + y^2) + (x^2 + xy)y' = 0 \tag{19}$$

and then solve the equation.

Solution:

In Example 3 we showed that this equation is not exact. Let us determine whether it has an integrating factor that depends on x only. On computing the quantity $(M_y - N_x)/N$, we find that

$$\frac{M_y(x, y) - N_x(x, y)}{N(x, y)} = \frac{3x + 2y - (2x + y)}{x^2 + xy} = \frac{1}{x}. \tag{28}$$

Thus there is an integrating factor μ that is a function of x only, and it satisfies the differential equation

$$\frac{d\mu}{dx} = \frac{\mu}{x}. \tag{29}$$

Hence (see Problem 7 in Section 2.2)

$$\mu(x) = x. \tag{30}$$

Multiplying equation (19) by this integrating factor, we obtain

$$(3x^2y + xy^2) + (x^3 + x^2y)y' = 0. \tag{31}$$

Equation (31) is exact, since

$$\frac{\partial}{\partial y}(3x^2y + xy^2) = 3x^2 + 2xy = \frac{\partial}{\partial x}(x^3 + x^2y).$$

Thus there is a function ψ such that

$$\psi_x(x, y) = 3x^2y + xy^2, \quad \psi_y(x, y) = x^3 + x^2y. \tag{32}$$

Integrating the first of equations (32) with respect to x, we obtain

$$\psi(x, y) = x^3 y + \frac{1}{2}x^2y^2 + h(y).$$

Substituting this expression for $\psi(x, y)$ in the second of equations (32), we find that

$$x^3 + x^2 y + h'(y) = x^3 + x^2 y,$$

so $h'(y) = 0$ and $h(y)$ is a constant. Thus the solutions of equation (31), and hence of equation (19), are given implicitly by

$$x^3 y + \frac{1}{2}x^2y^2 = c. \tag{33}$$

Solutions may also be found in explicit form since equation (33) is quadratic in y.

You may also verify that a second integrating factor for equation (19) is

$$\mu(x, y) = \frac{1}{xy(2x + y)}$$

and that the same solution is obtained, though with much greater difficulty, if this integrating factor is used (see Problem 22).

Problems

Determine whether each of the equations in Problems 1 through 8 is exact. If it is exact, find the solution.

1. $(2x + 3) + (2y - 2)y' = 0$

2. $(2x + 4y) + (2x - 2y)y' = 0$

3. $(3x^2 - 2xy + 2) + (6y^2 - x^2 + 3)y' = 0$

4. $\dfrac{dy}{dx} = -\dfrac{ax + by}{bx + cy}$

5. $\dfrac{dy}{dx} = -\dfrac{ax - by}{bx - cy}$

6. $(ye^{xy}\cos(2x) - 2e^{xy}\sin(2x) + 2x) + (xe^{xy}\cos(2x) - 3)y' = 0$

7. $(y/x + 6x) + (\ln x - 2)y' = 0, \quad x > 0$

8. $\dfrac{x}{(x^2 + y^2)^{3/2}} + \dfrac{y}{(x^2 + y^2)^{3/2}}\dfrac{dy}{dx} = 0$

In each of Problems 9 and 10, solve the given initial value problem and determine at least approximately where the solution is valid.

9. $(2x - y) + (2y - x)y' = 0, \quad y(1) = 3$

10. $(9x^2 + y - 1) - (4y - x)y' = 0, \quad y(1) = 0$

In each of Problems 11 and 12, find the value of b for which the given equation is exact, and then solve it using that value of b.

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

12. $(ye^{2xy} + x) + bxe^{2xy}y' = 0$

13. Assume that equation (6) meets the requirements of Theorem 2.6.1 in a rectangle R and is therefore exact. Show that a possible function $\psi(x, y)$ is

$$\psi(x, y) = \int_{x_0}^{x} M(s, y_0)\,ds + \int_{y_0}^{y} N(x, t)\,dt,$$

where (x_0, y_0) is a point in R.

14. Show that any separable equation

$$M(x) + N(y)y' = 0$$

is also exact.

In each of Problems 15 and 16, show that the given equation is not exact but becomes exact when multiplied by the given integrating factor. Then solve the equation.

15. $x^2y^3 + x(1 + y^2)y' = 0, \quad \mu(x, y) = 1/(xy^3)$

16. $(x + 2)\sin y + (x\cos y)y' = 0, \quad \mu(x, y) = xe^x$

17. Show that if $(N_x - M_y)/M = Q$, where Q is a function of y only, then the differential equation

$$M + Ny' = 0$$

has an integrating factor of the form

$$\mu(y) = \exp\int Q(y)\,dy.$$

In each of Problems 18 through 21, find an integrating factor and solve the given equation.

18. $(3x^2y + 2xy + y^3) + (x^2 + y^2)y' = 0$

19. $y' = e^{2x} + y - 1$

20. $1 + (x/y - \sin y)y' = 0$

21. $y + (2xy - e^{-2y})y' = 0$

22. Solve the differential equation

$$(3xy + y^2) + (x^2 + xy)y' = 0$$

using the integrating factor $\mu(x, y) = (xy(2x + y))^{-1}$. Verify that the solution is the same as that obtained in Example 4 with a different integrating factor.

2.7 Numerical Approximations: Euler's Method

Recall two important facts about the first-order initial value problem

$$\frac{dy}{dt} = f(t, y), \quad y(t_0) = y_0. \tag{1}$$

First, if f and $\partial f/\partial y$ are continuous, then the initial value problem (1) has a unique solution $y = \phi(t)$ in some interval surrounding the initial point $t = t_0$. Second, it is usually not possible to find the solution ϕ by symbolic manipulations of the differential equation. Up to now we have considered the main exceptions to the latter statement: differential equations that are linear, separable, or exact, or that can be transformed into one of these types. Nevertheless, it remains true that solutions of the vast majority of first-order initial value problems cannot be found by analytical means, such as those considered in the first part of this chapter.

Therefore, it is important to be able to approach the problem in other ways. As we have already seen, one of these ways is to draw a direction field for the differential equation (which does not involve solving the equation) and then to visualize the behavior of solutions from the direction field. This has the advantage of being a relatively simple process, even for complicated differential equations. However, it does not lend itself to quantitative computations or comparisons, and this is often a critical shortcoming.

For example, Figure 2.7.1 shows a direction field for the differential equation

$$\frac{dy}{dt} = 3 - 2t - 0.5y. \tag{2}$$

From the direction field you can visualize the behavior of solutions on the rectangle shown in the figure. On this rectangle a solution starting at a point on the y-axis initially increases with t, but it soon reaches a maximum value and then begins to decrease as t increases further.

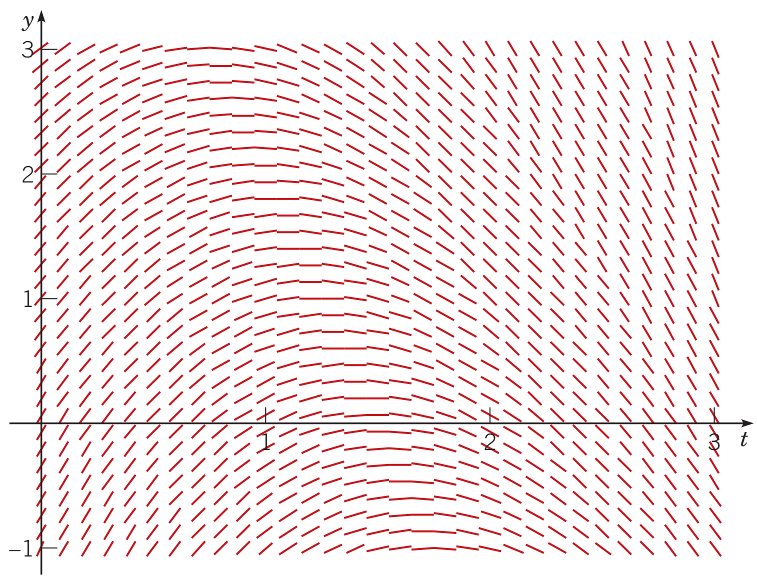


FIGURE 2.7.1 A direction field for equation (2): $dy/dt = 3 - 2t - 0.5y$.

You may also observe that in Figure 2.7.1 many tangent line segments at successive values of t almost touch each other. It takes only a bit of imagination to consider starting at a point on the y-axis and linking line segments for successive values of t in the grid, thereby producing a piecewise linear graph. Such a graph would apparently be an approximation to a solution of the differential equation. To convert this idea into a useful method for generating approximate solutions, we must answer several questions, including the following:

1. Can we carry out the linking of tangent lines in a systematic and straightforward manner?
2. If so, does the resulting piecewise linear function provide an approximation to an actual solution of the differential equation?
3. If so, can we assess the accuracy of the approximation? That is, can we estimate how far the approximation deviates from the solution itself?

It turns out that the answer to each of these questions is affirmative. The resulting method was originated by Euler about 1768 and is referred to as the **tangent line method** or the **Euler method**. We will deal with the first two questions in this section, but will defer a systematic discussion of the third question until Chapter 8.

To see how the Euler method works, let us consider how the tangent lines might be used to approximate the solution $y = \phi(t)$ of initial value problem (1) near $t = t_0$. We know that the solution passes through the initial point (t_0, y_0), and from the differential equation, we also know that its slope at this point is $f(t_0, y_0)$. Thus we can write down an equation for the line tangent to the solution curve at (t_0, y_0), namely,

$$y = y_0 + f(t_0, y_0)(t - t_0). \tag{3}$$

The tangent line is a good approximation to the actual solution curve on an interval short enough so that the slope of the solution does not change appreciably from its value at the initial point; see Figure 2.7.2. Thus, if t_1 is close enough to t_0, we can approximate $\phi(t_1)$ by the value y_1 determined by substituting $t = t_1$ into the tangent line approximation at $t = t_0$; thus

$$y_1 = y_0 + f(t_0, y_0)(t_1 - t_0). \tag{4}$$

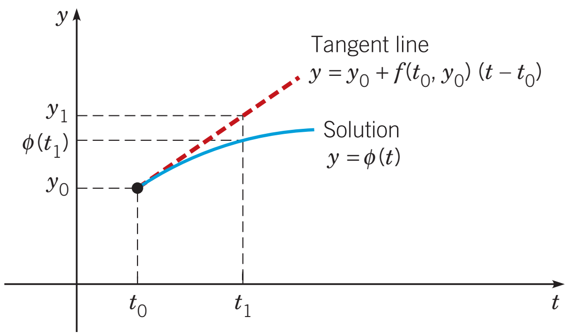


FIGURE 2.7.2 A tangent line approximation of $y' = f(t, y)$ at (t_0, y_0).

To proceed further, we can try to repeat the process. Unfortunately, we do not know the value $\phi(t_1)$ of the solution at t_1. The best we can do is to use the approximate value y_1 instead. Thus we construct the line through (t_1, y_1) with the slope $f(t_1, y_1)$,

$$y = y_1 + f(t_1, y_1)(t - t_1). \tag{5}$$

To approximate the value of $\phi(t)$ at a nearby point t_2, we use equation (5) instead of equation (3), obtaining

$$y_2 = y_1 + f(t_1, y_1)(t_2 - t_1). \tag{6}$$

Continuing in this manner, we use the value of y calculated at each step to determine the slope of the approximation for the next step. The general expression for the tangent line starting at (t_n, y_n) is

$$y = y_n + f(t_n, y_n)(t - t_n); \tag{7}$$

hence the approximate value y_{n+1} at t_{n+1} in terms of t_n, t_{n+1}, and y_n is

$$y_{n+1} = y_n + f(t_n, y_n)(t_{n+1} - t_n), \quad n = 0, 1, 2, \ldots . \tag{8}$$

If we introduce the notation $f_n = f(t_n, y_n)$, then we can rewrite equation (8) as

$$y_{n+1} = y_n + f_n \cdot (t_{n+1} - t_n), \quad n = 0, 1, 2, \ldots . \tag{9}$$

Finally, if we assume that there is a uniform step size h between the points $t_0, t_1, t_2, \ldots$, then $t_{n+1} = t_n + h$ for each n, and we obtain Euler's formula in the form

$$y_{n+1} = y_n + f_n h, \quad n = 0, 1, 2, \ldots . \tag{10}$$

To use Euler's method, you repeatedly evaluate equation (9) or equation (10), depending on whether or not the step size is constant, using the result of each step to execute the next step. In this way you generate a sequence of values $y_1, y_2, y_3, \ldots$ that approximate the values of the solution $\phi(t)$ at the points $t_1, t_2, t_3, \ldots$. If, instead of a sequence of points, you need a function to approximate the solution $\phi(t)$, then you can use the piecewise linear function constructed from the collection of tangent line segments. That is, let y be given in $[t_0, t_1]$ by equation (7) with $n = 0$, in $[t_1, t_2]$ by equation (7) with $n = 1$, and so on.

EXAMPLE 1

Consider the initial value problem

$$\frac{dy}{dt} = 3 - 2t - 0.5y, \quad y(0) = 1. \tag{11}$$

Use Euler's method with step size $h = 0.2$ to find approximate values of the solution of initial value problem (9) at $t = 0.2, 0.4, 0.6, 0.8$, and 1. Compare them with the corresponding values of the actual solution of the initial value problem.

Solution:

Note that the differential equation in the given initial value problem is the same as in equation (2); its direction field is shown in Figure 2.7.1. Before applying Euler's method, observe that this differential equation is linear, so it can be solved as in Section 2.1, using the integrating factor $e^{t/2}$. The resulting solution of the initial value problem (9) is

$$y = \phi(t) = 14 - 4t - 13e^{-t/2}. \tag{12}$$

We will use this information to assess how the approximate solution obtained by Euler's method compares with the exact solution.

To approximate this solution by Euler's method, note that $f(t, y) = 3 - 2t - 0.5y$. Using the initial values $t_0 = 0$ and $y_0 = 1$, we find that

$$f_0 = f(t_0, y_0) = f(0, 1) = 3 - 0 - 0.5 = 2.5$$

and then, from equation (3), the tangent line approximation near $t = 0$ is

$$y = 1 + 2.5(t - 0) = 1 + 2.5t. \tag{13}$$

Setting $t = 0.2$ in equation (13), we find the approximate value y_1 of the solution at $t = 0.2$, namely,

$$y_1 = 1 + (2.5)(0.2) = 1.5.$$

At the next step we have

$$f_1 = f(t_1, y_1) = f(0.2, 1.5) = 3 - 2(0.2) - (0.5)(1.5) = 3 - 0.4 - 0.75 = 1.85.$$

Then the tangent line approximation near $t = 0.2$ is

$$y = 1.5 + 1.85(t - 0.2) = 1.13 + 1.85t. \tag{14}$$

Evaluating the expression in equation (14) for $t = 0.4$, we obtain

$$y_2 = 1.13 + 1.85(0.4) = 1.87.$$

Repeating this computational procedure three more times, we obtain the results shown in Table 2.7.1.

TABLE 2.7.1 Results of Euler's Method with $h = 0.2$ for $y' = 3 - 2t - 0.5y, \quad y(0) = 1$

n	t_n	y_n	$f_n = f(t_n, y_n)$	Tangent Line	Exact $y(t_n)$
0	0.0	1.00000	2.5	$y = 1 + 2.5(t - 0)$	1.00000
1	0.2	1.50000	1.85	$y = 1.5 + 1.85(t - 0.2)$	1.43711
2	0.4	1.87000	1.265	$y = 1.87 + 1.265(t - 0.4)$	1.75650
3	0.6	2.12300	0.7385	$y = 2.123 + 0.7385(t - 0.6)$	1.96936
4	0.8	2.27070	0.26465	$y = 2.2707 + 0.26465(t - 0.8)$	2.08584
5	1.0	2.32363			2.11510

The second column contains the t-values separated by the step size $h = 0.2$. The third column shows the corresponding y-values computed from Euler's formula (10). Column four contains the slopes f_n of the tangent line at the current point, (t_n, y_n). In the fifth column are the tangent line approximations found from equation (7). The sixth column contains values of the solution (12) of the initial value problem (9), correct to five decimal places. The solution (12) and the tangent line approximation are also plotted in Figure 2.7.3.

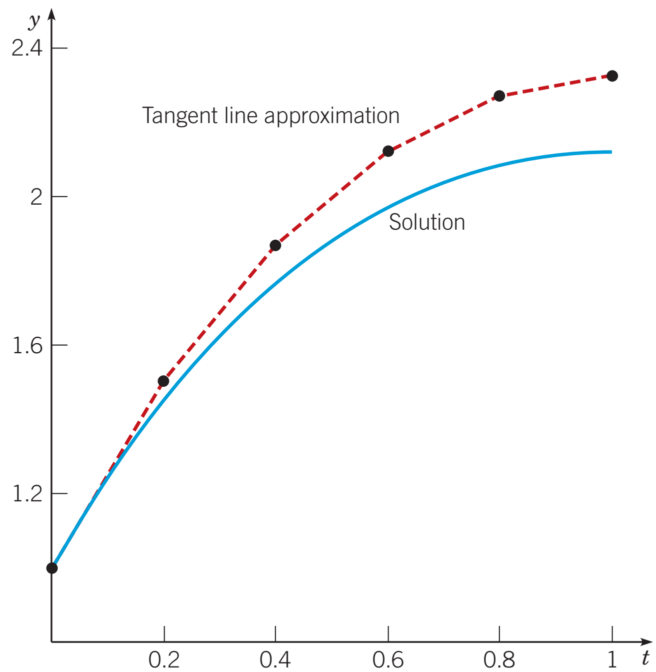


FIGURE 2.7.3 Plots of the solution and a tangent line approximation with $h = 0.2$ for the initial value problem (9): $dy/dt = 3 - 2t - 0.5y, \ y(0) = 1$.

From Table 2.7.1 and Figure 2.7.3 we see that the approximations given by Euler's method for this problem are greater than the corresponding values of the actual solution. This is because the graph of the solution is concave down and therefore the tangent line approximations lie above the graph.

The accuracy of the approximations in this example is not good enough to be satisfactory in a typical scientific or engineering application. For example, at $t = 1$ the error in the approximation is $2.32363 - 2.11510 = 0.20853$, which is a percentage error of about 9.86% relative to the exact solution. One way to achieve more accurate results is to use a smaller step size, with a corresponding increase in the number of computational steps. We explore this possibility in the next example.

Of course, computations such as those in Example 1 and in the other examples in this section are usually done on a computer. Some software packages include code for the Euler method, while others do not. In any case, it is straightforward to write a computer program that will carry out the calculations required to produce results such as those in Table 2.7.1.

Basically, what is required is a loop that will evaluate equation (10) repetitively, along with suitable instructions for input and output. The output can be a list of numbers, as in Table 2.7.1, or a plot, as in Figure 2.7.3. The specific instructions can be written in any high-level programming language with which you are familiar.

EXAMPLE 2

Consider again the initial value problem (9)

$$\frac{dy}{dt} = 3 - 2t - 0.5y, \quad y(0) = 1.$$

Use Euler's method with various step sizes to calculate approximate values of the solution for $0 \le t \le 5$. Compare the calculated results with the corresponding values of the exact solution (12)

$$y = 14 - 4t - 13e^{-t/2}.$$

Solution:

We used step sizes $h = 0.1$, 0.05, 0.025, and 0.01, corresponding to 50, 100, 200, and 500 steps, respectively, to go from $t = 0$ to $t = 5$. The results of these calculations, along with the values of the exact solution, are summarized in Table 2.7.2. All computed entries are rounded to four decimal places, although more digits were retained in the intermediate calculations.

TABLE 2.7.2 Comparison of the Exact Solution with Euler's Method for Several Step Sizes h for $y' = 3 - 2t - 0.5y, \quad y(0) = 1$

t	$h = 0.1$	$h = 0.05$	$h = 0.025$	$h = 0.01$	Exact
0.0	1.0000	1.0000	1.0000	1.0000	1.0000
1.0	2.2164	2.1651	2.1399	2.1250	2.1151
2.0	1.3397	1.2780	1.2476	1.2295	1.2176
3.0	−0.7903	−0.8459	−0.8734	−0.8898	−0.9007
4.0	−3.6707	−3.7152	−3.7373	−3.7506	−3.7594
5.0	−7.0003	−7.0337	−7.0504	−7.0604	−7.0671

What conclusions can we draw from the data in Table 2.7.2? The most important observation is that, for a fixed value of t, the computed approximate values become more accurate as the step size h decreases. You can see this by reading across a particular row in the table from left to right. This is what we would expect, of course, but it is encouraging that the data confirm our expectations. For example, for $t = 2$ the approximate value with $h = 0.1$ is too large by 0.1221 (about 10%), whereas the value with $h = 0.01$ is too large by only 0.0119 (about 1%). In this case, reducing the step size by a factor of 10 (and performing 10 times as many computations) also reduces the error by a factor of about 10. Comparing the errors for other pairs of values in the table confirms that this relation between step size and error holds for them also: reducing the step size by a given factor also reduces the error by approximately the same factor. Does this mean that for the Euler method the error is approximately proportional to the step size? Of course, one example does not establish such a general result, but it is at least an interesting conjecture.[20]

A second observation from Table 2.7.2 is that, for a fixed step size h, the approximations become more accurate as t increases, at least for $t > 2$. For instance, for $h = 0.1$ the error for $t = 5$ is only 0.0668, which is a little more than one-half of the error at $t = 2$. We will return to this matter later in this section.

All in all, Euler's method seems to work rather well for this problem. Reasonably good results are obtained even for a moderately large step size $h = 0.1$, and the approximation can be improved by decreasing h.

[20] A more detailed discussion of the errors in using the Euler method appears in Chapter 8.

Let us now look at another example.

EXAMPLE 3

Consider the initial value problem

$$\frac{dy}{dt} = 4 - t + 2y, \quad y(0) = 1. \tag{15}$$

The general solution of this differential equation was found in Example 2 of Section 2.1, and the solution of the initial value problem (11) is

$$y = -\frac{7}{4} + \frac{1}{2}t + \frac{11}{4}e^{2t}. \tag{16}$$

Use Euler's method with several step sizes to find approximate values of the solution on the interval $0 \le t \le 5$. Compare the results with the corresponding values of the solution (16).

Solution:

Using the same range of step sizes as in Example 2, we obtain the results presented in Table 2.7.3.

TABLE 2.7.3 Comparison of the Exact Solution with Euler's Method for Several Step Sizes h for $y' = 4 - t + 2y, \quad y(0) = 1$

t	$h = 0.1$	$h = 0.05$	$h = 0.025$	$h = 0.01$	Exact
0.0	1.000000	1.000000	1.000000	1.000000	1.000000
1.0	15.77728	17.25062	18.10997	18.67278	19.06990
2.0	104.6784	123.7130	135.5440	143.5835	149.3949
3.0	652.5349	837.0745	959.2580	1045.395	1109.179
4.0	4042.122	5633.351	6755.175	7575.577	8197.884
5.0	25026.95	37897.43	47555.35	54881.32	60573.53

The data in Table 2.7.3 again confirm our expectation that, for a given value of t, accuracy improves as the step size h is reduced. For example, for $t = 1$ the percentage error diminishes from 17.3% when $h = 0.1$ to 2.1% when $h = 0.01$. However, the error increases fairly rapidly as t increases for a fixed h. Even for $h = 0.01$, the error at $t = 5$ is 9.4%, and it is much greater for larger step sizes. Of course, the accuracy that is needed depends on the purpose for which the results are intended, but the errors in Table 2.7.3 are too large for most scientific or engineering applications. To improve the situation, we might either try even smaller step sizes or else restrict the computations to a rather short interval away from the initial point. Nevertheless, it is clear that Euler's method is much less effective in this example than in Example 2.

To understand better what is happening in these examples, let us look again at Euler's method for the general initial value problem (1)

$$\frac{dy}{dt} = f(t, y), \quad y(t_0) = y_0,$$

whose exact solution we denote by $\phi(t)$. Recall that a first-order differential equation has an infinite family of solutions, indexed by an arbitrary constant c, and that the initial condition picks out one member of this infinite family by determining the value of c. Thus in the infinite family of solutions, $\phi(t)$ is the one solution that satisfies the initial condition $\phi(t_0) = y_0$.

At the first step Euler's method uses the tangent line approximation to the graph of $y = \phi(t)$ passing through the initial point (t_0, y_0), and this produces the approximate value y_1 at t_1. Usually, $y_1 \neq \phi(t_1)$, so at the second step Euler's method uses the tangent line approximation not to $y = \phi(t)$, but to a nearby solution $y = \phi_1(t)$ that passes through the point (t_1, y_1). So it is at each subsequent step. Euler's method uses a succession of tangent line approximations to a sequence of different solutions $\phi(t), \phi_1(t), \phi_2(t), \ldots$ of the differential equation. At each step the tangent line is constructed to the solution passing through the point determined by the result of the preceding step, as shown in Figure 2.7.4. The quality of the approximation after many steps depends strongly on the behavior of the set of solutions that pass through the points (t_n, y_n) for $n = 1, 2, 3, \ldots$.

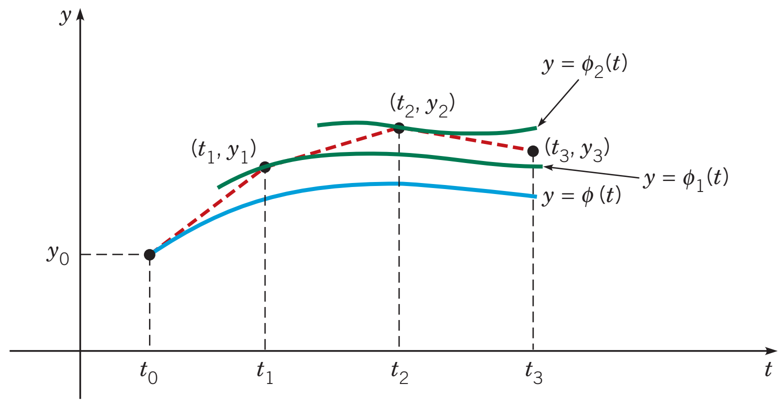


FIGURE 2.7.4 The Euler method.

In Example 2 the general solution of the differential equation is

$$y = 14 - 4t + ce^{-t/2} \tag{17}$$

and the solution of the initial value problem (9) corresponds to $c = -13$. The family of solutions (17) is a converging family since the term involving the arbitrary constant c approaches zero as $t \to \infty$. It does not matter very much which solutions we are approximating by tangent lines in the implementation of Euler's method, since all the solutions are getting closer and closer to each other as t increases.

On the other hand, in Example 3 the general solution of the differential equation is

$$y = -\frac{7}{4} + \frac{1}{2}t + ce^{2t}, \tag{18}$$

and, because the term involving the arbitrary constant c grows without bound as $t \to \infty$, this is a diverging family. Note that solutions corresponding to two nearby values of c become arbitrarily far apart as t increases. In Example 3 we are trying approximate the solution for $c = 11/4$, but in the use of Euler's method we are actually at each step following another solution that separates from the desired one faster and faster as t increases. This explains why the errors in Example 3 are so much larger than those in Example 2.

In using a numerical procedure such as the Euler method, you must always keep in mind the question of whether the results are accurate enough to be useful. In the preceding examples, the accuracy of the numerical results could be determined directly by a comparison with the solution obtained analytically. Of course, usually the analytical solution is not available if a numerical procedure is to be employed, so what we usually need are bounds for, or at least estimates of, the error that do not require a knowledge of the exact solution. You should also keep in mind that the best that we can expect, or hope for, from a numerical approximation is that it reflects the behavior of the actual solution. Thus a member of a diverging family of solutions will always be harder to approximate than a member of a converging family.

If you wish to read more about numerical approximations to solutions of initial value problems, you may go directly to Chapter 8 at this point. There, we present some information on the analysis of errors and also discuss several algorithms that are computationally much more efficient than the Euler method.

Problems

Note about Variations of Computed Results. Most of the problems in this section call for fairly extensive numerical computations. To handle these problems you need suitable computing hardware and software. Keep in mind that numerical results may vary somewhat, depending on how your program is constructed and on how your computer executes arithmetic steps, rounds off, and so forth. Minor variations in the last decimal place may be due to such causes and do not necessarily indicate that something is amiss. Answers in the back of the book are recorded to six digits in most cases, although more digits were retained in the intermediate calculations.

In each of Problems 1 through 4:

N **a.** Find approximate values of the solution of the given initial value problem at $t = 0.1, 0.2, 0.3$, and 0.4 using the Euler method with $h = 0.1$.

N **b.** Repeat part (a) with $h = 0.05$. Compare the results with those found in a.

N **c.** Repeat part a with $h = 0.025$. Compare the results with those found in a and b.

N **d.** Find the solution $y = \phi(t)$ of the given problem and evaluate $\phi(t)$ at $t = 0.1, 0.2, 0.3$, and 0.4. Compare these values with the results of a, b, and c.

1. $y' = 3 + t - y, \quad y(0) = 1$

2. $y' = 2y - 1, \quad y(0) = 1$

3. $y' = 0.5 - t + 2y, \quad y(0) = 1$

4. $y' = 3\cos t - 2y, \quad y(0) = 0$

In each of Problems 5 through 8, draw a direction field for the given differential equation and state whether you think that the solutions are converging or diverging.

G **5.** $y' = 5 - 3\sqrt{y}$

G **6.** $y' = y(3 - ty)$

G **7.** $y' = -ty + 0.1y^3$

G **8.** $y' = t^2 + y^2$

In each of Problems 9 and 10, use Euler's method to find approximate values of the solution of the given initial value problem at $t = 0.5$, 1, 1.5, 2, 2.5, and 3: (a) With $h = 0.1$, (b) With $h = 0.05$, (c) With $h = 0.025$, (d) With $h = 0.01$.

N **9.** $y' = 5 - 3\sqrt{y}, \quad y(0) = 2$

N **10.** $y' = y(3 - ty), \quad y(0) = 0.5$

11. Consider the initial value problem

$$y' = \frac{3t^2}{3y^2 - 4}, \quad y(1) = 0.$$

N **a.** Use Euler's method with $h = 0.1$ to obtain approximate values of the solution at $t = 1.2, 1.4, 1.6$, and 1.8.

N **b.** Repeat part a with $h = 0.05$.

c. Compare the results of parts a and b. Note that they are reasonably close for $t = 1.2, 1.4$, and 1.6 but are quite different for $t = 1.8$. Also note (from the differential equation) that the line tangent to the solution is parallel to the y-axis when $y = \pm 2/\sqrt{3} \cong \pm 1.155$. Explain how this might cause such a difference in the calculated values.

N **12.** Consider the initial value problem

$$y' = t^2 + y^2, \quad y(0) = 1.$$

Use Euler's method with $h = 0.1, 0.05, 0.025$, and 0.01 to explore the solution of this problem for $0 \le t \le 1$. What is your best estimate of the value of the solution at $t = 0.8$? At $t = 1$? Are your results consistent with the direction field in Problem 8?

13. Consider the initial value problem

$$y' = -ty + 0.1y^3, \quad y(0) = \alpha,$$

where α is a given number.

G **a.** Draw a direction field for the differential equation (or reexamine the one from Problem 7). Observe that there is a critical value of α in the interval $2 \le \alpha \le 3$ that separates converging solutions from diverging ones. Call this critical value α_0.

N **b.** Use Euler's method with $h = 0.01$ to estimate α_0. Do this by restricting α_0 to an interval $[a, b]$, where $b - a = 0.01$.

14. Consider the initial value problem

$$y' = y^2 - t^2, \quad y(0) = \alpha,$$

where α is a given number.

G **a.** Draw a direction field for the differential equation. Note that there is a critical value of α in the interval $0 \le \alpha \le 1$ that separates converging solutions from diverging ones. Call this critical value α_0.

N **b.** Use Euler's method with $h = 0.01$ to estimate α_0. Do this by restricting α_0 to an interval $[a, b]$, where $b - a = 0.01$.

15. Convergence of Euler's Method. It can be shown that under suitable conditions on f, the numerical approximation generated by the Euler method for the initial value problem $y' = f(t, y)$, $y(t_0) = y_0$ converges to the exact solution as the step size h decreases. This is illustrated by the following example. Consider the initial value problem

$$y' = 1 - t + y, \quad y(t_0) = y_0.$$

a. Show that the exact solution is $y = \phi(t) = (y_0 - t_0)e^{t - t_0} + t$.

N **b.** Using the Euler formula, show that

$$y_k = (1 + h)y_{k-1} + h - ht_{k-1}, \quad k = 1, 2, \ldots .$$

c. Noting that $y_1 = (1 + h)(y_0 - t_0) + t_1$, show by induction that

$$y_n = (1 + h)^n (y_0 - t_0) + t_n \tag{19}$$

for each positive integer n.

d. Consider a fixed point $t > t_0$ and for a given n choose $h = (t - t_0)/n$. Then $t_n = t$ for every n. Note also that $h \to 0$ as $n \to \infty$. By substituting for h in equation (19) and letting $n \to \infty$, show that $y_n \to \phi(t)$ as $n \to \infty$.

Hint: $\lim_{n \to \infty} (1 + a/n)^n = e^a$.

In each of Problems 16 and 17, use the technique discussed in Problem 15 to show that the approximation obtained by the Euler method converges to the exact solution at any fixed point as $h \to 0$.

16. $y' = y, \quad y(0) = 1$

17. $y' = 2y - 1, \quad y(0) = 1 \quad$ *Hint:* $y_1 = (1 + 2h)/2 + 1/2$

2.8 The Existence and Uniqueness Theorem

In this section we discuss the proof of Theorem 2.4.2, the fundamental existence and uniqueness theorem for first-order initial value problems. Recall that this theorem states that under certain conditions on $f(t, y)$, the initial value problem

$$y' = f(t, y), \quad y(t_0) = y_0 \tag{1}$$

has a unique solution in some interval containing the point t_0.

In some cases (for example, if the differential equation is linear), the existence of a solution of the initial value problem (1) can be established directly by actually solving the problem and exhibiting a formula for the solution. However, in general, this approach is not feasible because there is no method of solving the differential equation that applies in all cases. Therefore, for the general case, it is necessary to adopt an indirect approach that demonstrates the existence of a solution of initial value problem (1) but usually does not provide a practical means of finding it. The heart of this method is the construction of a sequence of functions that converges to a limit function satisfying the initial value problem, although the members of the sequence individually do not. As a rule, it is impossible to compute explicitly more than a few members of the sequence; therefore, the limit function can be determined only in rare cases. Nevertheless, under the restrictions on $f(t, y)$ stated in Theorem 2.4.2, it is possible to show that the sequence in question converges and that the limit function has the desired properties. The argument is fairly intricate and depends, in part, on techniques and results that are usually encountered for the first time in a course on advanced calculus. Consequently, we do not go into all the details of the proof here; we do, however, indicate its main features and point out some of the difficulties that must be overcome.

First of all, we note that it is sufficient to consider the problem in which the initial point (t_0, y_0) is the origin; that is, we consider the problem

$$y' = f(t, y), \quad y(0) = 0. \tag{2}$$

If some other initial point is given, then we can always make a preliminary change of variables, corresponding to a translation of the coordinate axes, that will take the given point (t_0, y_0) into the origin. The existence and uniqueness theorem can now be stated in the following way.

Theorem 2.8.1 | Existence and Uniqueness of Solutions of $y' = f(t, y), y(0) = 0$

If f and $\partial f/\partial y$ are continuous in a rectangle $R: |t| \leq a$, $|y| \leq b$, then there is some interval $|t| \leq h \leq a$ in which there exists a unique solution $y = \phi(t)$ of the initial value problem (2).

For the method of proof discussed here it is necessary to transform initial value problem (2) into a more convenient form. If we suppose temporarily that there is a differentiable function $y = \phi(t)$ that satisfies the initial value problem, then $f(t, \phi(t))$ is a continuous function of t only. Hence we can integrate $y' = f(t, y)$ from the initial point $t = 0$ to an arbitrary value of t, obtaining

$$\phi(t) = \int_0^t f(s, \phi(s))\, ds, \tag{3}$$

where we have made use of the initial condition $\phi(0) = 0$. We also denote the dummy variable of integration by s.

Since equation (3) contains an integral of the unknown function ϕ, it is called an **integral equation**. This integral equation is not a formula for the solution of the initial value problem, but it does provide another relation satisfied by any solution of equations (2). Conversely, suppose that there is a continuous function $y = \phi(t)$ that satisfies the integral equation (3); then this function also satisfies the initial value problem (2). To show this, we first substitute zero for t in equation (3), which shows that the initial condition is satisfied. Further, since the integrand in equation (3) is continuous, it follows from the fundamental theorem of calculus that ϕ is differentiable and that $\phi'(t) = f(t, \phi(t))$. Therefore, the initial value problem and the integral equation are equivalent in the sense that any solution of one is also a solution of the other. It is more convenient to show that there is a unique solution of the integral equation in a certain interval $|t| \leq h$. The same conclusion also holds for the initial value problem (2).

One method of showing that the integral equation (3) has a unique solution is known as the **method of successive approximations** or Picard's[21] **iteration method**. In using this method,

[21]Charles-Émile Picard (1856–1914) was appointed professor at the Sorbonne before the age of 30. Except for Henri Poincaré, he is perhaps the most distinguished French mathematician of his generation. He is known for important theorems in complex variables and algebraic geometry as well as differential equations. A special case of the method of successive approximations was first published by Liouville in 1838. However, the method is usually credited to Picard, who established it in a general and widely applicable form in a series of papers beginning in 1890.

we start by choosing an initial function ϕ_0, either arbitrarily or to approximate in some way the solution of the initial value problem. The simplest choice is

$$\phi_0(t) = 0; \tag{4}$$

then ϕ_0 at least satisfies the initial condition in equations (2), although presumably not the differential equation. The next approximation ϕ_1 is obtained by substituting $\phi_0(s)$ for $\phi(s)$ in the right-hand side of equation (3) and calling the result of this operation $\phi_1(t)$. Thus

$$\phi_1(t) = \int_0^t f\big(s, \phi_0(s)\big)\, ds. \tag{5}$$

Similarly, ϕ_2 is obtained from ϕ_1:

$$\phi_2(t) = \int_0^t f\big(s, \phi_1(s)\big)\, ds, \tag{6}$$

and, in general,

$$\phi_{n+1}(t) = \int_0^t f(s, \phi_n(s))\, ds. \tag{7}$$

In this manner we generate the sequence of functions $\{\phi_n\} = \{\phi_0, \phi_1, \phi_2, \ldots, \phi_n, \ldots\}$.

Each member of the sequence satisfies the initial condition, but in general none satisfies the differential equation. However, if at some stage, say, for $n = k$, we find that $\phi_{k+1}(t) = \phi_k(t)$, then it follows that ϕ_k is a solution of the integral equation (3). Hence ϕ_k is also a solution of the initial value problem (2), and the sequence is terminated at this point. In general, this does not occur, and it is necessary to consider the entire infinite sequence.

To establish Theorem 2.8.1, we must answer four principal questions:

1. Do all members of the sequence $\{\phi_n\}$ exist, or may the process break down at some stage?
2. Does the sequence converge?
3. What are the properties of the limit function? In particular, does it satisfy the integral equation (3) and hence the initial value problem (2)?
4. Is this the only solution, or may there be others?

We first show how these questions can be answered in a specific and relatively simple example and then comment on some of the difficulties that may be encountered in the general case.

EXAMPLE 1

Solve the initial value problem

$$y' = 2t(1 + y), \quad y(0) = 0 \tag{8}$$

by the method of successive approximations.

Solution:

Note first that if $y = \phi(t)$, then the corresponding integral equation is

$$\phi(t) = \int_0^t 2s(1 + \phi(s))\, ds. \tag{9}$$

If the initial approximation is $\phi_0(t) = 0$, it follows that

$$\phi_1(t) = \int_0^t 2s\big(1 + \phi_0(s)\big)\, ds = \int_0^t 2s ds = t^2. \tag{10}$$

Similarly,

$$\phi_2(t) = \int_0^t 2s\big(1 + \phi_1(s)\big)ds = \int_0^t 2s(1 + s^2)ds = t^2 + \frac{t^4}{2} \tag{11}$$

and

$$\phi_3(t) = \int_0^t 2s(1+\phi_2(s))\,ds = \int_0^t 2s\left(1+s^2+\frac{s^4}{2}\right)ds = t^2+\frac{t^4}{2}+\frac{t^6}{2\cdot 3}. \tag{12}$$

Equations (10), (11), and (12) suggest that

$$\phi_n(t) = t^2 + \frac{t^4}{2!} + \frac{t^6}{3!} + \cdots + \frac{t^{2n}}{n!} \tag{13}$$

for each $n \geq 1$, and this result can be established by mathematical induction, as follows. Equation (13) is certainly true for $n = 1$; see equation (10). We must show that if it is true for $n = k$, then it also holds for $n = k + 1$. We have

$$\begin{aligned} \phi_{k+1}(t) &= \int_0^t 2s(1+\phi_k(s))\,ds \\ &= \int_0^t 2s\left(1+s^2+\frac{s^4}{2!}+\cdots+\frac{s^{2k}}{k!}\right)ds \\ &= \int_0^t 2s+2s^3+\frac{2s^5}{2!}+\cdots+\frac{2s^{2k+1}}{k!}\,ds \\ &= t^2+\frac{t^4}{2!}+\frac{t^6}{3!}+\cdots+\frac{t^{2k+2}}{(k+1)!}, \end{aligned} \tag{14}$$

and the inductive proof is complete.

A plot of the first four iterates, $\phi_1(t)$, $\phi_2(t)$, $\phi_3(t)$, and $\phi_4(t)$, is shown in Figure 2.8.1. As k increases, the iterates seem to remain close over a gradually increasing interval, suggesting eventual convergence to a limit function.

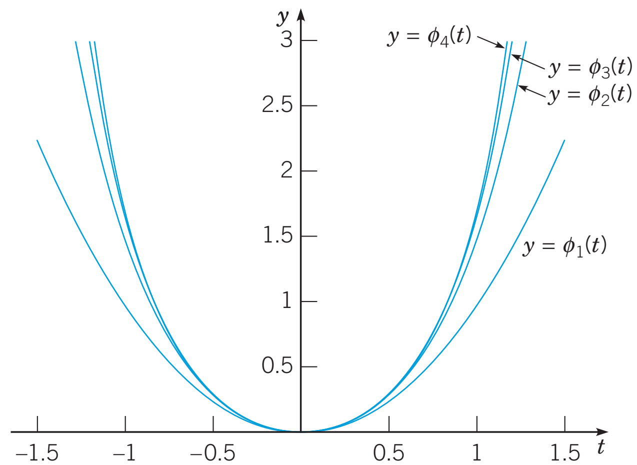


FIGURE 2.8.1 Plots of the first four Picard iterates $y = \phi_1(t), \ldots, y = \phi_4(t)$ for Example 1: $dy/dt = 2t(1+y),\ y(0) = 0$.

It follows from equation (13) that $\phi_n(t)$ is the n^{th} partial sum of the infinite series

$$\sum_{k=1}^{\infty} \frac{t^{2k}}{k!}; \tag{15}$$

hence $\lim_{n\to\infty} \phi_n(t)$ exists if and only if the series (15) converges. Applying the ratio test, we see that, for each t,

$$\left| \frac{t^{2k+2}}{(k+1)!}\,\frac{k!}{t^{2k}} \right| = \frac{t^2}{k+1} \to 0 \text{ as } k \to \infty. \tag{16}$$

Thus the interval of convergence for series (15) is the entire t-axis. This means its sum $\phi(t)$ is the limit of the sequence $\{\phi_n(t)\}$ for every value of t. Further, since the series (15) is a Taylor series, it can be differentiated or integrated term-by-term for all values of t. Therefore, we can verify by direct computation that $\phi(t) = \sum_{k=1}^{\infty} t^{2k}/k!$ is a solution of the integral equation (9). Alternatively, by substituting $\phi(t)$ for y in equations (8), we can verify that this function satisfies the initial value problem (6). In this example it is also possible, from the series (15), to identify the solution $\phi(t)$ in terms of elementary functions, namely, $\phi(t) = e^{t^2} - 1$. (See Problem 13.) However, this is not necessary for the discussion of existence and uniqueness.

Explicit knowledge of $\phi(t)$ does make it possible to visualize the convergence of the sequence of iterates more clearly by plotting the difference $e_k(t) = \phi(t) - \phi_k(t)$ for various values of k. Figure 2.8.2 shows this difference for $k = 1, 2, 3, 4$. This figure clearly illustrates the gradually increasing interval over which successive iterates provide a good approximation to the solution of the initial value problem.

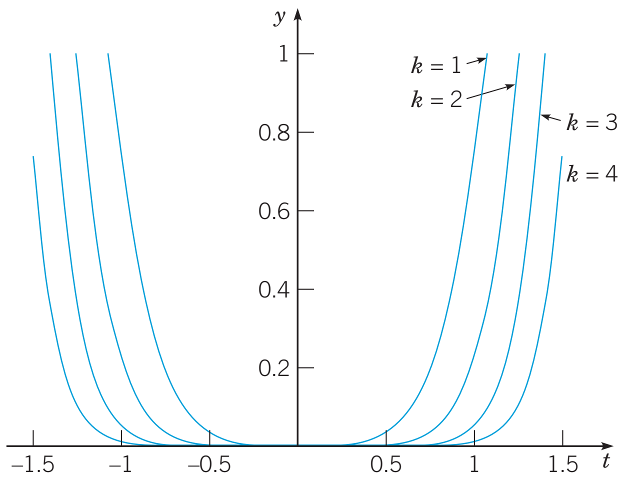


FIGURE 2.8.2 Plots of $y = e_k(t) = \phi(t) - \phi_k(t)$ for Example 1 for $k = 1, \ldots, 4$.

Finally, to deal with the question of uniqueness, let us suppose that the initial value problem has two different solutions ϕ and ψ. The assumption that ϕ and ψ are different means there is at least one value of t for which $\phi(t) - \psi(t) \neq 0$. Also, since ϕ and ψ both satisfy the integral equation (9), we have by subtraction (and the linearity of integration) that

$$\phi(t) - \psi(t) = \int_0^t 2s(\phi(s) - \psi(s))ds.$$

Taking absolute values of both sides, we have, if $t > 0$,

$$|\phi(t) - \psi(t)| = \left| \int_0^t 2s(\phi(s) - \psi(s))ds \right| \leq \int_0^t 2s|\phi(s) - \psi(s)|ds.$$

If we restrict t to lie in the interval $0 \leq t \leq A/2$, where A is arbitrary, then $2t \leq A$ and

$$|\phi(t) - \psi(t)| \leq A \int_0^t |\phi(s) - \psi(s)|ds \quad \text{for } 0 \leq t \leq A/2. \tag{17}$$

It is now convenient to introduce the function U defined by

$$U(t) = \int_0^t |\phi(s) - \psi(s)|ds. \tag{18}$$

Then it follows at once that

$$U(0) = 0, \tag{19}$$

$$U(t) \geq 0, \quad \text{for } t \geq 0. \tag{20}$$

Further, U is differentiable, and $U'(t) = |\phi(t) - \psi(t)|$. Hence, by equation (17),

$$U'(t) - AU(t) \leq 0 \quad \text{for } 0 \leq t \leq A/2. \tag{21}$$

Multiplying equation (21) by the positive quantity e^{-At} gives

$$\left(e^{-At}U(t)\right)' \leq 0 \quad \text{for } 0 \leq t \leq A/2. \tag{22}$$

Then, upon integrating equation (22) from zero to t and using equation (19), we obtain

$$e^{-At}U(t) \le 0 \quad \text{for} \quad 0 \le t \le A/2.$$

Hence $U(t) \le 0$ for $0 \le t \le A/2$. However, since A is arbitrary, we conclude that $U(t) \le 0$ for all nonnegative t. This result and equation (20) are compatible only if $U(t) = 0$ for each $t \ge 0$. Thus $U'(t) = 0$ and therefore $\psi(t) = \phi(t)$ for all $t \ge 0$. This contradicts the hypothesis that ϕ and ψ are two different solutions. Consequently, there cannot be two different solutions of the initial value problem for $t \ge 0$. A slight modification of this argument leads to the same conclusion for $t \le 0$.

Returning now to the general problem of solving the integral equation (3), let us consider briefly each of the questions raised earlier:

1. *Do all members of the sequence $\{\phi_n\}$ exist?*

In the example, f and $\partial f/\partial y$ were continuous in the whole ty-plane, and each member of the sequence could be explicitly calculated. In contrast, in the general case, f and $\partial f/\partial y$ are assumed to be continuous only in the rectangle R: $|t| \le a$, $|y| \le b$ (see Figure 2.8.3). Furthermore, the members of the sequence cannot as a rule be explicitly determined. The danger is that at some stage, say, for $n = k$, the graph of $y = \phi_k(t)$ may contain points that lie outside the rectangle R. More precisely, in the computation of $\phi_{k+1}(t)$ it would be necessary to evaluate $f(t, y)$ at points where it is not known to be continuous or even to exist. Thus the calculation of $\phi_{k+1}(t)$ might be impossible.

To avoid this danger, it may be necessary to restrict t to a smaller interval than $|t| \le a$. To find such an interval, we make use of the fact that a continuous function on a closed bounded region is bounded. Hence f is bounded on R; thus there exists a positive number M such that

$$|f(t, y)| \le M, \quad (t, y) \text{ in } R. \tag{23}$$

We have mentioned before that

$$\phi_n(0) = 0$$

for each n. Since $f(t, \phi_k(t))$ is equal to $\phi'_{k+1}(t)$, the maximum absolute slope of the graph of the equation $y = \phi_{k+1}(t)$ is M. Since this graph contains the point $(0, 0)$, it must lie in a bow tie-shaped shaded region as shown in Figure 2.8.4. Hence the point $(t, \phi_{k+1}(t))$ remains in R at least as long as R contains the bow tie-shaped region, which is for $|t| \le b/M$. We hereafter consider only the rectangle D: $|t| \le h$, $|y| \le b$, where h is equal either to a or to b/M, whichever is smaller. With this restriction, all members of the sequence $\{\phi_n(t)\}$ exist. Note that whenever $b/M < a$, you can try to obtain a larger value of h by finding a better (that is, smaller) bound M for $|f(t, y)|$, if this is possible.

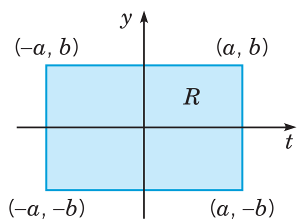


FIGURE 2.8.3 Region of definition for Theorem 2.8.1.

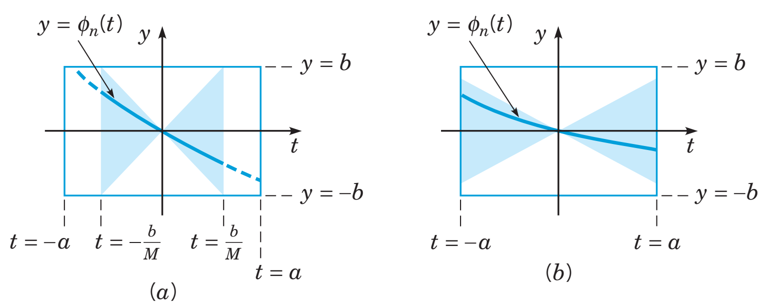


FIGURE 2.8.4 Bow-tie regions in which successive iterates lie. (*a*) if $b/M < a$ then $h = b/M$; (*b*) if $b/M > a$ then $h = a$.

2. *Does the sequence* $\{\phi_n(t)\}$ *converge?*

We can identify $\phi_n(t) = \phi_1(t) + (\phi_2(t) - \phi_1(t)) + \cdots + (\phi_n(t) - \phi_{n-1}(t))$ as the n^{th} partial sum of the series

$$\phi_1(t) + \sum_{k=1}^{\infty} (\phi_{k+1}(t) - \phi_k(t)). \tag{24}$$

The convergence of the sequence $\{\phi_n(t)\}$ is established by showing that the series (24) converges. To do this, it is necessary to estimate the magnitude $|\phi_{k+1}(t) - \phi_k(t)|$ of the general term. The argument by which this is done is indicated in Problems 14 through 17 and will be omitted here. Assuming that the sequence converges, denote the limit function by ϕ, and so

$$\phi(t) = \lim_{n\to\infty} \phi_n(t). \tag{25}$$

3. *What are the properties of the limit function* ϕ*?*

In the first place, we would like to know that ϕ is continuous. This is not, however, a necessary consequence of the convergence of the sequence $\{\phi_n(t)\}$, even though each member of the sequence is itself continuous. Sometimes a sequence of continuous functions converges to a limit function that is discontinuous. A simple example of this phenomenon is given in Problem 11. One way to show that ϕ is continuous is to show not only that the sequence $\{\phi_n\}$ converges, but also that it converges in a certain manner, known as **uniform convergence**. We do not take up this matter here, but note only that the argument referred to in the discussion of question 2 is sufficient to establish the uniform convergence of the sequence $\{\phi_n\}$ and, hence, the continuity of the limit function ϕ in the interval $|t| \le h$.

Now let us return to equation (7)

$$\phi_{n+1}(t) = \int_0^t f(s, \phi_n(s))\, ds.$$

Allowing n to approach ∞ on both sides, we obtain

$$\phi(t) = \lim_{n\to\infty} \int_0^t f(s, \phi_n(s))\, ds. \tag{26}$$

We would like to interchange the operations of integrating and taking the limit on the right-hand side of equation (26) so as to obtain

$$\phi(t) = \int_0^t \lim_{n\to\infty} f(s, \phi_n(s))\, ds. \tag{27}$$

In general, such an interchange is not permissible (see Problem 12, for example), but once again, the fact that the sequence $\{\phi_n(t)\}$ converges uniformly is sufficient to allow us to take the limiting operation inside the integral sign. Next, we wish to take the limit inside the function f, which would give

$$\phi(t) = \int_0^t f\left(s, \lim_{n\to\infty} \phi_n(s)\right) ds \tag{28}$$

and hence

$$\phi(t) = \int_0^t f(s, \phi(s))\, ds. \tag{29}$$

The statement that

$$\lim_{n\to\infty} f(s, \phi_n(s)) = f\left(s, \lim_{n\to\infty} \phi_n(s)\right)$$

is equivalent to the statement that f is continuous in its second variable, which is known by hypothesis. Hence equation (29) is valid, and the function ϕ satisfies the integral equation (3). Thus $y = \phi(t)$ is also a solution of the initial value problem (2).

4. *Are there other solutions of the integral equation (3) besides $y = \phi(t)$?*
To show the uniqueness of the solution $y = \phi(t)$, we can proceed much as in the example. First, assume the existence of another solution $y = \psi(t)$. It is then possible to show (see Problem 18) that the difference $\phi(t) - \psi(t)$ satisfies the inequality

$$|\phi(t) - \psi(t)| \leq A \int_0^t |\phi(s) - \psi(s)|\, ds \tag{30}$$

for $0 \leq t \leq h$ and a suitable positive number A. From this point the argument is identical to that given in the example, and we conclude that there is no solution of the initial value problem (2) other than the one generated by the method of successive approximations.

Problems

In each of Problems 1 and 2, transform the given initial value problem into an equivalent problem with the initial point at the origin.

1. $dy/dt = t^2 + y^2, \quad y(1) = 2$
2. $dy/dt = 1 - y^3, \quad y(-1) = 3$

In each of Problems 3 through 4, let $\phi_0(t) = 0$ and define $\{\phi_n(t)\}$ by the method of successive approximations.

a. Determine $\phi_n(t)$ for an arbitrary value of n.
G **b.** Plot $\phi_n(t)$ for $n = 1, \ldots, 4$. Observe whether the iterates appear to be converging.
c. Express $\lim_{n\to\infty} \phi_n(t) = \phi(t)$ in terms of elementary functions; that is, solve the given initial value problem.
G **d.** Plot $|\phi(t) - \phi_n(t)|$ for $n = 1, \ldots, 4$. For each of $\phi_1(t), \ldots, \phi_4(t)$, estimate the interval in which it is a reasonably good approximation to the actual solution.

N 3. $y' = 2(y + 1), \quad y(0) = 0$
N 4. $y' = -y/2 + t, \quad y(0) = 0$

In each of Problems 5 and 6, let $\phi_0(t) = 0$ and use the method of successive approximations to solve the given initial value problem.

a. Determine $\phi_n(t)$ for an arbitrary value of n.
G **b.** Plot $\phi_n(t)$ for $n = 1, \ldots, 4$. Observe whether the iterates appear to be converging.
c. Show that the sequence $\{\phi_n(t)\}$ converges.

5. $y' = ty + 1, \quad y(0) = 0$
6. $y' = t^2 y - t, \quad y(0) = 0$

In each of Problems 7 and 8, let $\phi_0(t) = 0$ and use the method of successive approximations to approximate the solution of the given initial value problem.

a. Calculate $\phi_1(t), \ldots, \phi_3(t)$.
G **b.** Plot $\phi_1(t), \ldots, \phi_3(t)$. Observe whether the iterates appear to be converging.

7. $y' = t^2 + y^2, \quad y(0) = 0$
8. $y' = 1 - y^3, \quad y(0) = 0$

In each of Problems 9 and 10, let $\phi_0(t) = 0$ and use the method of successive approximations to approximate the solution of the given initial value problem.

a. Calculate $\phi_1(t), \ldots, \phi_4(t)$, or (if necessary) Taylor approximations to these iterates. Keep terms up to order six.
G **b.** Plot the functions you found in part a and observe whether they appear to be converging.

9. $y' = -\sin y + 1, \quad y(0) = 0$
10. $y' = \dfrac{3t^2 + 4t + 2}{2(y - 1)}, \quad y(0) = 0$

11. Let $\phi_n(x) = x^n$ for $0 \leq x \leq 1$ and show that

$$\lim_{n\to\infty} \phi_n(x) = \begin{cases} 0, & 0 \leq x < 1, \\ 1, & x = 1. \end{cases}$$

This example shows that a sequence of continuous functions may converge to a limit function that is discontinuous.

12. Consider the sequence $\phi_n(x) = 2nxe^{-nx^2}, \quad 0 \leq x \leq 1$.

a. Show that $\lim_{n\to\infty} \phi_n(x) = 0$ for $0 \leq x \leq 1$; hence

$$\int_0^1 \lim_{n\to\infty} \phi_n(x)\,dx = 0.$$

b. Show that $\displaystyle\int_0^1 2nxe^{-nx^2}\,dx = 1 - e^{-n}$; hence

$$\lim_{n\to\infty} \int_0^1 \phi_n(x)\,dx = 1.$$

Thus, in this example,

$$\lim_{n\to\infty} \int_a^b \phi_n(x)\,dx \neq \int_a^b \lim_{n\to\infty} \phi_n(x)\,dx,$$

even though $\lim_{n\to\infty} \phi_n(x)$ exists and is continuous.

13. **a.** Verify that $\phi(t) = \displaystyle\sum_{k=1}^{\infty} \frac{t^{2k}}{k!}$ is a solution of the integral equation (9).
b. Verify that $\phi(t)$ is also a solution of the initial value problem (6).
c. Use the fact that $\displaystyle\sum_{k=0}^{\infty} \frac{t^k}{k!} = e^t$ to evaluate $\phi(t)$ in terms of elementary functions.
d. Solve initial value problem (6) as a separable equation.
e. Solve initial value problem (6) as a first order linear equation.

In Problems 14 through 17, we indicate how to prove that the sequence $\{\phi_n(t)\}$, defined by equations (4) through (7), converges.

a. Verify that $\phi(t) = \displaystyle\sum_{k=1}^{\infty} \frac{t^{2k}}{k!}$ is a solution of the integral equation (9).
b. Verify that $\phi(t)$ is also a solution of the initial value problem (6).
c. Use the fact that $\displaystyle\sum_{k=0}^{\infty} \frac{t^k}{k!} = e^t$ to evaluate $\phi(t)$ in terms of elementary functions.
d. Solve initial value problem (6) as a separable equation.
e. Solve initial value problem (6) as a first order linear equation.

14. If $\partial f/\partial y$ is continuous in the rectangle D, show that there is a positive constant K such that

$$|f(t, y_1) - f(t, y_2)| \le K|y_1 - y_2|, \tag{31}$$

where (t, y_1) and (t, y_2) are any two points in D having the same t coordinate. This inequality is known as a Lipschitz[22] condition.
Hint: Hold t fixed and use the mean value theorem on f as a function of y only. Choose K to be the maximum value of $|\partial f/\partial y|$ in D.

15. If $\phi_{n-1}(t)$ and $\phi_n(t)$ are members of the sequence $\{\phi_n(t)\}$, use the result of Problem 14 to show that

$$\left|f(t, \phi_n(t)) - f\big(t, \phi_{n-1}(t)\big)\right| \le K\left|\phi_n(t) - \phi_{n-1}(t)\right|.$$

16. **a.** Show that if $|t| \le h$, then

$$|\phi_1(t)| \le M|t|,$$

where M is chosen so that $|f(t, y)| \le M$ for (t, y) in D.

b. Use the results of Problem 15 and part a of Problem 16 to show that

$$|\phi_2(t) - \phi_1(t)| \le \frac{MK|t|^2}{2}.$$

c. Show, by mathematical induction, that

$$|\phi_n(t) - \phi_{n-1}(t)| \le \frac{MK^{n-1}|t|^n}{n!} \le \frac{MK^{n-1}h^n}{n!}.$$

17. Note that

$$\phi_n(t) = \phi_1(t) + \big(\phi_2(t) - \phi_1(t)\big) + \cdots + \big(\phi_n(t) - \phi_{n-1}(t)\big).$$

a. Show that

$$|\phi_n(t)| \le |\phi_1(t)| + |\phi_2(t) - \phi_1(t)| + \cdots + |\phi_n(t) - \phi_{n-1}(t)|.$$

b. Use the results of Problem 16 to show that

$$|\phi_n(t)| \le \frac{M}{K}\left(Kh + \frac{(Kh)^2}{2!} + \cdots + \frac{(Kh)^n}{n!}\right).$$

c. Show that the sum in part b converges as $n \to \infty$ and, hence, the sum in part a also converges as $n \to \infty$. Conclude therefore that the sequence $\{\phi_n(t)\}$ converges since it is the sequence of partial sums of a convergent infinite series.

18. In this problem we deal with the question of uniqueness of the solution of the integral equation (3)

$$\phi(t) = \int_0^t f(s, \phi(s))\, ds.$$

a. Suppose that ϕ and ψ are two solutions of equation (3). Show that, for $t \ge 0$,

$$\phi(t) - \psi(t) = \int_0^t \big(f(s, \phi(s)) - f(s, \psi(s))\big)\, ds.$$

b. Show that

$$|\phi(t) - \psi(t)| \le \int_0^t \big(f(s, \phi(s)) - f(s, \psi(s))\big)\, ds.$$

c. Use the result of Problem 14 to show that

$$|\phi(t) - \psi(t)| \le K \int_0^t |\phi(s) - \psi(s)|\, ds,$$

where K is an upper bound for $|\partial f/\partial y|$ in D. This is the same as equation (30), and the rest of the proof may be constructed as indicated in the text.

[22]The German mathematician Rudolf Lipschitz (1832–1903), professor at the University of Bonn for many years, worked in several areas of mathematics. The inequality (i) can replace the hypothesis that $\partial f/\partial y$ is continuous in Theorem 2.8.1; this results in a slightly stronger theorem.

2.9 First-Order Difference Equations

Although a continuous model leading to a differential equation is reasonable and attractive for many problems, there are some cases in which a discrete model may be more natural. For instance, the continuous model of compound interest used in Section 2.3 is only an approximation to the actual discrete process. Similarly, sometimes population growth may be described more accurately by a discrete model than by a continuous model. This is true, for example, of species whose generations do not overlap and that propagate at regular intervals, such as at particular times of the calendar year. Then the population y_{n+1} of the species in the year $n+1$ is some function of n and the population y_n in the preceding year; that is,

$$y_{n+1} = f(n, y_n), \quad n = 0, 1, 2, \ldots. \tag{1}$$

Equation (1) is called a **first-order difference equation**. It is first-order because the value of y_{n+1} depends on the value of y_n but not on earlier values y_{n-1}, y_{n-2}, and so forth. As for differential equations, the difference equation (1) is **linear** if f is a linear function of y_n; otherwise, it is **nonlinear**. A **solution** of the difference equation (1) is a sequence of numbers $y_0, y_1, y_2, \ldots$ that satisfy the equation for each n. In addition to the difference equation itself, there may also be an **initial condition**

$$y_0 = \alpha \tag{2}$$

that prescribes the value of the first term of the solution sequence.

We now assume temporarily that the function f in equation (1) depends only on y_n, but not on n. In this case

$$y_{n+1} = f(y_n), \quad n = 0, 1, 2, \ldots. \tag{3}$$

If y_0 is given, then successive terms of the solution can be found from equation (3). Thus

$$y_1 = f(y_0),$$

and

$$y_2 = f(y_1) = f(f(y_0)).$$

The quantity $f(f(y_0))$ is called the second iterate of the difference equation and is sometimes denoted by $f^2(y_0)$. Similarly, the third iterate y_3 is given by

$$y_3 = f(y_2) = f\big(f\big(f(y_0)\big)\big) = f^3(y_0),$$

and so on. In general, the n^{th} iterate y_n is

$$y_n = f(y_{n-1}) = f^n(y_0).$$

This procedure is referred to as iterating the difference equation. It is often of primary interest to determine the behavior of y_n as $n \to \infty$. In particular, does y_n approach a limit, and if so, what is it?

Solutions for which y_n has the same value for all n are called **equilibrium solutions**. They are frequently of special importance, just as in the study of differential equations. If equilibrium solutions exist, you can find them by setting y_{n+1} equal to y_n in equation (3) and solving the resulting equation

$$y_n = f(y_n) \tag{4}$$

for y_n.

Linear Equations. Suppose that the population of a certain species in a given region in year $n+1$, denoted by y_{n+1}, is a positive multiple ρ_n of the population y_n in year n; that is,

$$y_{n+1} = \rho_n y_n, \qquad n = 0, 1, 2, \ldots . \tag{5}$$

Note that the reproduction rate ρ_n may differ from year to year. The difference equation (5) is linear and can easily be solved by iteration. We obtain

$$y_1 = \rho_0 y_0,$$
$$y_2 = \rho_1 y_1 = \rho_1 \rho_0 y_0,$$

and, in general,

$$y_n = \rho_{n-1} \cdots \rho_0 y_0, \qquad n = 1, 2, \ldots . \tag{6}$$

Thus, if the initial population y_0 is given, then the population of each succeeding generation is determined by equation (6). Although for a population problem ρ_n is intrinsically positive, the solution (6) is also valid if ρ_n is negative for some or all values of n. Note, however, that if ρ_n is zero for some n, then y_{n+1} and all succeeding values of y are zero; in other words, the species has become extinct.

If the reproduction rate ρ_n has the same value ρ for each n, then the difference equation (5) becomes

$$y_{n+1} = \rho y_n \tag{7}$$

and its solution is

$$y_n = \rho^n y_0. \tag{8}$$

Equation (7) also has an equilibrium solution, namely, $y_n = 0$ for all n, corresponding to the initial value $y_0 = 0$. The limiting behavior of y_n is easy to determine from equation (8). In fact,

$$\lim_{n\to\infty} y_n = \begin{cases} 0, & \text{if } |\rho| < 1; \\ y_0, & \text{if } \rho = 1; \\ \text{does not exist}, & \text{otherwise.} \end{cases} \tag{9}$$

In other words, the equilibrium solution $y_n = 0$ is asymptotically stable for $|\rho| < 1$ and unstable for $|\rho| > 1$.

Now we will modify the population model represented by equation (5) to include the effect of immigration or emigration. If b_n is the net increase in population in year n due to immigration, then the population in year $n+1$ is the sum of the part of the population resulting from natural reproduction and the part due to immigration. Thus

$$y_{n+1} = \rho y_n + b_n, \quad n = 0, 1, 2, \dots, \tag{10}$$

where we are now assuming that the reproduction rate ρ is constant. We can solve equation (10) by iteration in the same manner as before. We have

$$\begin{aligned} y_1 &= \rho y_0 + b_0, \\ y_2 &= \rho(\rho y_0 + b_0) + b_1 = \rho^2 y_0 + \rho b_0 + b_1, \\ y_3 &= \rho(\rho^2 y_0 + \rho b_0 + b_1) + b_2 = \rho^3 y_0 + \rho^2 b_0 + \rho b_1 + b_2, \end{aligned}$$

and so forth. In general, we obtain

$$y_n = \rho^n y_0 + \rho^{n-1} b_0 + \cdots + \rho b_{n-2} + b_{n-1} = \rho^n y_0 + \sum_{j=0}^{n-1} \rho^{n-1-j} b_j. \tag{11}$$

Note that the first term on the right-hand side of equation (11) represents the descendants of the original population, while the other terms represent the population in year n resulting from immigration in all preceding years.

In the special case where $b_n = b \neq 0$ for all n, the difference equation is

$$y_{n+1} = \rho y_n + b, \tag{12}$$

and from equation (11) its solution is

$$y_n = \rho^n y_0 + (1 + \rho + \rho^2 + \cdots + \rho^{n-1})b. \tag{13}$$

If $\rho \neq 1$, we can write this solution in the more compact form

$$y_n = \rho^n y_0 + \frac{1-\rho^n}{1-\rho} b, \tag{14}$$

where again the two terms on the right-hand side are the effects of the original population and of immigration, respectively. Rewriting equation (14) as

$$y_n = \rho^n \left(y_0 - \frac{b}{1-\rho} \right) + \frac{b}{1-\rho} \tag{15}$$

makes the long-time behavior of y_n more evident. It follows from equation (15) that $y_n \to b/(1-\rho)$ if $|\rho| < 1$. If $|\rho| > 1$ or if $\rho = -1$ then y_n has no limit unless $y_0 = b/(1-\rho)$. The quantity $b/(1-\rho)$, for $\rho \neq 1$, is an equilibrium solution of equation (12), as can readily be seen directly from that equation. Of course, equation (14) is not valid for $\rho = 1$. To deal with that case, we must return to equation (13) and let $\rho = 1$ there. It follows that

$$y_n = y_0 + nb, \tag{16}$$

so in this case y_n becomes unbounded as $n \to \infty$.

The same model also provides a framework for solving many problems of a financial character. For such problems, y_n is the account balance in the nth time period, $\rho_n = 1 + r_n$, where r_n is the interest rate for that period, and b_n is the amount deposited or withdrawn. The following example is typical.

EXAMPLE 1

A recent college graduate takes out a \$10,000 loan to purchase a car. If the interest rate is 12%, what monthly payment is required to pay off the loan in 4 years?

Solution:

The relevant difference equation is equation (12), where y_n is the loan balance outstanding in the n^{th} month, $\rho = 1 + r$, where r is the interest rate per month, and b is the effect of the monthly payment.

Note that $\rho = 1.01$, corresponding to a monthly interest rate of 1%. Since payments reduce the loan balance, b must be negative; the actual payment is $|b|$.

The solution of the difference equation (12) with this value for ρ and the initial condition $y_0 = 10{,}000$ is given by equation (15); that is,

$$y_n = (1.01)^n(10{,}000 + 100b) - 100b. \tag{17}$$

The value of b needed to pay off the loan in 4 years is found by setting $y_{48} = 0$ and solving for b. This gives

$$b = -100\frac{(1.01)^{48}}{(1.01)^{48} - 1} = -263.34. \tag{18}$$

The total amount paid on the loan is 48 times $|b|$, or $12,640.32. Of this amount, $10,000 is repayment of the principal and the remaining $2640.32 is interest.

Nonlinear Equations. Nonlinear difference equations are much more complicated and have much more varied solutions than linear equations. We will restrict our attention to a single equation, the **logistic difference equation**

$$y_{n+1} = \rho y_n\left(1 - \frac{y_n}{k}\right), \tag{19}$$

which is analogous to the logistic differential equation

$$\frac{dy}{dt} = ry\left(1 - \frac{y}{K}\right) \tag{20}$$

that was discussed in Section 2.5. Note that if the derivative dy/dt in equation (20) is replaced by the difference quotient $(y_{n+1} - y_n)/h$, then equation (20) reduces to equation (19) with $\rho = 1 + hr$ and $k = (1 + hr)K/hr$. To simplify equation (19) a little more, we can scale the variable y_n by introducing the new variable $u_n = y_n/k$. Then equation (19) becomes

$$u_{n+1} = \rho u_n(1 - u_n), \tag{21}$$

where ρ is a positive parameter.

We begin our investigation of equation (21) by seeking the equilibrium, or constant, solutions. These can be found by setting u_{n+1} equal to u_n in equation (21), which corresponds to setting dy/dt equal to zero in equation (20). The resulting equation is

$$u_n = \rho u_n - \rho u_n^2, \tag{22}$$

so it follows that the equilibrium solutions of equation (21) are

$$u_n = 0, \quad u_n = \frac{\rho - 1}{\rho}. \tag{23}$$

The next question is whether the equilibrium solutions are asymptotically stable or unstable. That is, for an initial condition near one of the equilibrium solutions, does the resulting solution sequence approach or depart from the equilibrium solution?

One way to examine this question is by approximating equation (21) by a linear equation in the neighborhood of an equilibrium solution. For example, near the equilibrium solution $u_n = 0$, the quantity u_n^2 is small compared to u_n itself, so we assume that we can neglect the quadratic term in equation (21) in comparison with the linear terms. This leaves us with the linear difference equation

$$u_{n+1} = \rho u_n, \tag{24}$$

which is presumably a good approximation to equation (21) for u_n sufficiently near zero. However, equation (24) is the same as equation (7), and we have already concluded, in equation (9), that $u_n \to 0$ as $n \to \infty$ if and only if $|\rho| < 1$, or (since ρ must be positive) for $0 < \rho < 1$. Thus the equilibrium solution $u_n = 0$ is asymptotically stable for the linear approximation (24) for this set of ρ values, so we conclude that it is also asymptotically stable for the full nonlinear equation (21).

The previous conclusion is correct, although our argument is not complete. What is lacking is a theorem stating that the solutions of the nonlinear equation (21) resemble those of the linear equation (24) near the equilibrium solution $u_n = 0$. We will not take time to discuss this issue here; the same question is treated for differential equations in Section 9.3.

Now consider the other equilibrium solution $u_n = (\rho - 1)/\rho$. To study solutions in the neighborhood of this point, we write

$$u_n = \frac{\rho - 1}{\rho} + v_n, \tag{25}$$

where we assume that v_n is small. By substituting from equation (25) in equation (21) and simplifying the resulting equation, we eventually obtain

$$v_{n+1} = (2 - \rho)v_n - \rho v_n^2. \tag{26}$$

Since v_n is small, we again neglect the quadratic term in comparison with the linear terms and thereby obtain the linear equation

$$v_{n+1} = (2 - \rho)v_n. \tag{27}$$

Referring to equation (9) once more, we find that $v_n \to 0$ as $n \to \infty$ for $|2 - \rho| < 1$, or in other words for $1 < \rho < 3$. Therefore, we conclude that for this range of values of ρ, the equilibrium solution $u_n = (\rho - 1)/\rho$ is asymptotically stable.

Figure 2.9.1 contains the graphs of solutions of equation (21) for $\rho = 0.8$, $\rho = 1.5$, and $\rho = 2.8$, respectively. Observe that the solution converges to zero for $\rho = 0.8$ and to the nonzero equilibrium solution for $\rho = 1.5$ and $\rho = 2.8$. The convergence is (eventually) monotone for $\rho = 0.8$ and for $\rho = 1.5$ and is oscillatory for $\rho = 2.8$. The graphs shown are for particular initial conditions, but the graphs for other initial conditions are similar.

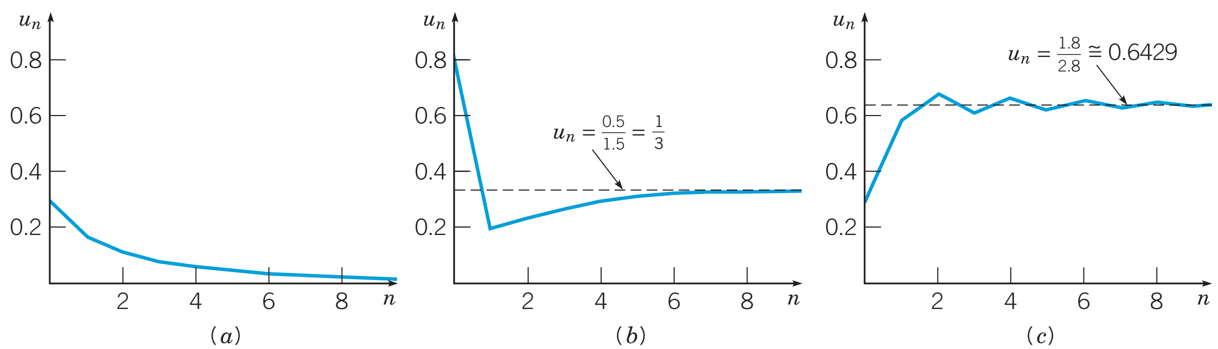


FIGURE 2.9.1 Solutions of $u_{n+1} = \rho u_n(1 - u_n)$: (*a*) $\rho = 0.8$; (*b*) $\rho = 1.5$; (*c*) $\rho = 2.8$.

Another way of displaying the solution of a difference equation is shown in Figure 2.9.2. In each part of this figure, the graphs of the parabola $y = \rho x(1 - x)$ and of the straight line $y = x$ are shown. The equilibrium solutions correspond to the points of intersection of these two curves. The piecewise linear graph consisting of successive vertical and horizontal line segments, sometimes called a stairstep or cobweb diagram, represents the solution sequence. The sequence starts at the point u_0 on the x-axis. The vertical line segment drawn upward to the parabola at u_0 corresponds to the calculation of $\rho u_0(1 - u_0) = u_1$. This value is then transferred from the y-axis to the x-axis; this step is represented by the horizontal line segment from the parabola to the line $y = x$. Then the process is repeated over and over again. Clearly, the sequence converges to the origin in Figure 2.9.2*a* and to the nonzero equilibrium solution in the other two cases.

To summarize our results so far: the difference equation (21) has two equilibrium solutions, $u_n = 0$ and $u_n = (\rho - 1)/\rho$; the former is asymptotically stable for $0 \le \rho < 1$, and the latter is asymptotically stable for $1 < \rho < 3$. When $\rho = 1$, the two equilibrium solutions coincide at $u = 0$; this solution can be shown to be asymptotically stable. In Figure 2.9.3 the parameter ρ is plotted on the horizontal axis and u on the vertical axis. The equilibrium solutions $u = 0$ and $u = (\rho - 1)/\rho$ are shown. The intervals in which each one is asymptotically stable are indicated by the solid portions of the curves. There is an **exchange of stability** from one equilibrium solution to the other at $\rho = 1$.

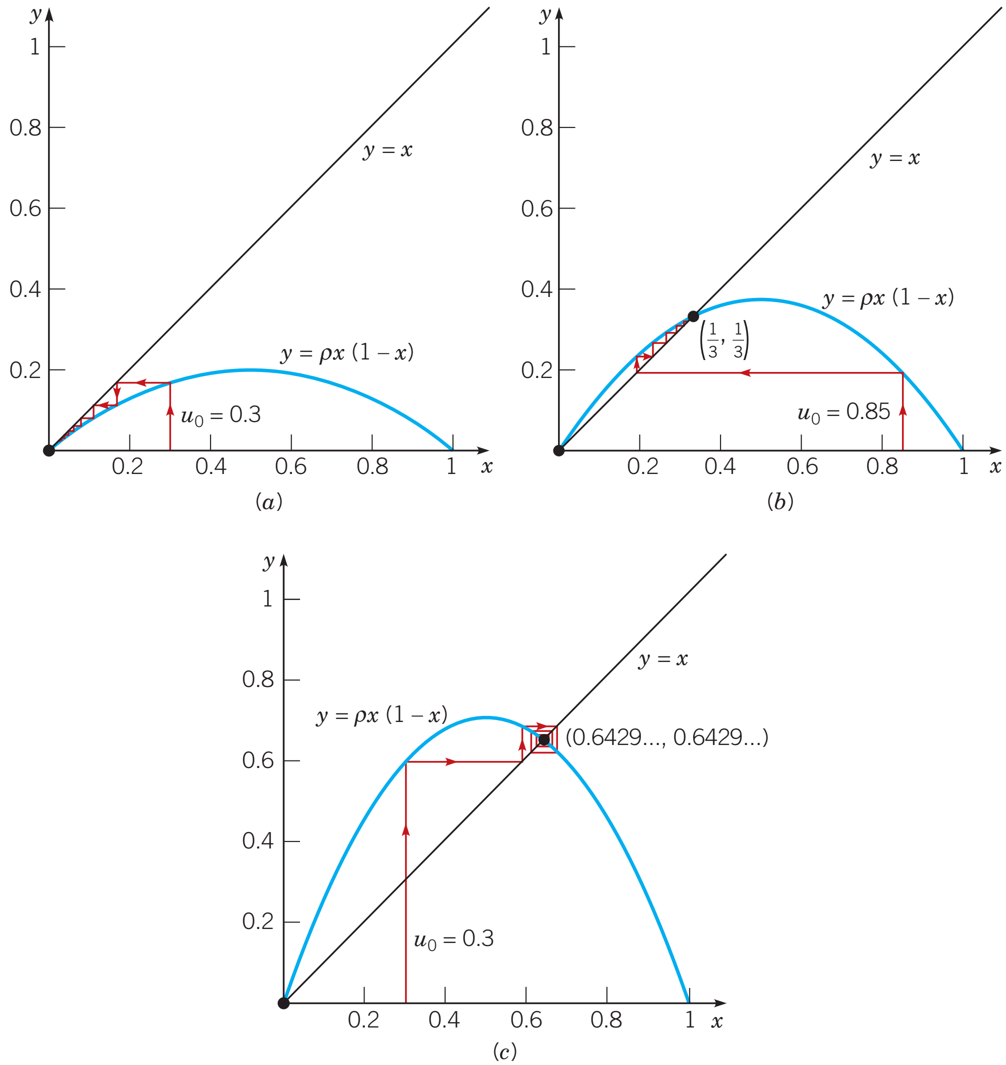


FIGURE 2.9.2 Iterates of $u_{n+1} = \rho u_n(1 - u_n)$: (*a*) $\rho = 0.8$; (*b*) $\rho = 1.5$; (*c*) $\rho = 2.8$.

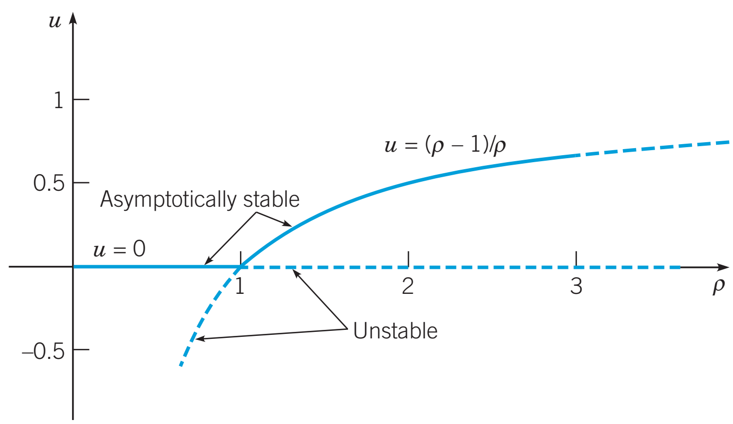


FIGURE 2.9.3 Exchange of stability for $u_{n+1} = \rho u_n(1 - u_n)$.

For $\rho > 3$, neither of the equilibrium solutions is stable, and the solutions of equation (21) exhibit increasing complexity as ρ increases. For ρ somewhat greater than 3, the sequence u_n rapidly approaches a steady oscillation of period 2; that is, u_n oscillates back and forth between two distinct values. For $\rho = 3.2$, a solution is shown in Figure 2.9.4. For n greater than about

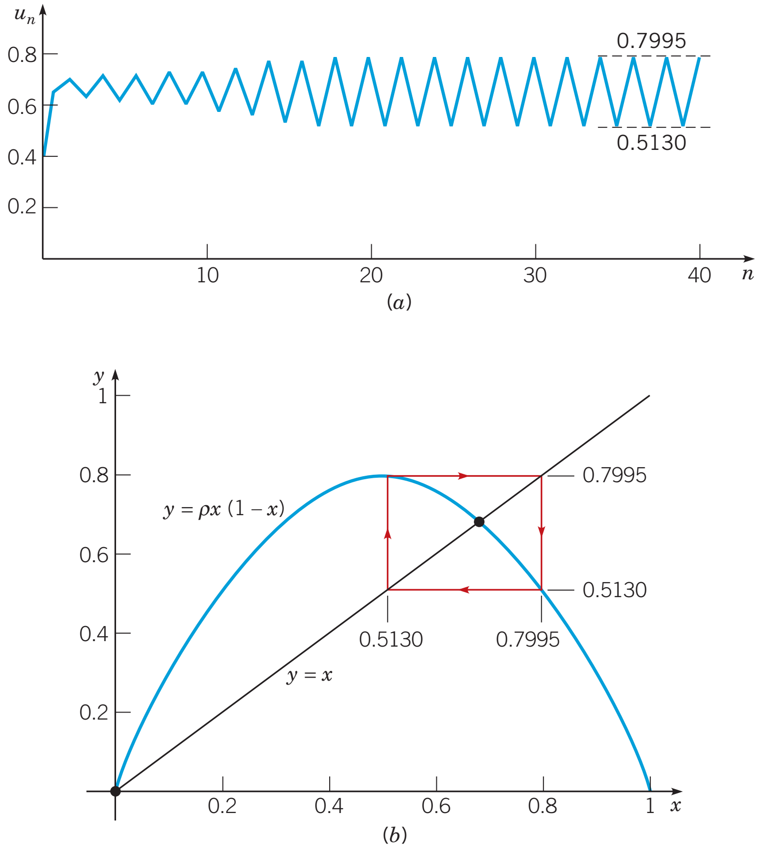
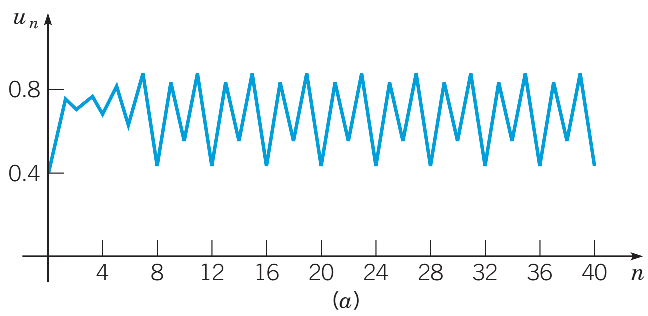


FIGURE 2.9.4 A solution of $u_{n+1} = \rho u_n(1 - u_n)$ for $\rho = 3.2$; period 2. (*a*) u_n versus n; (*b*) the cobweb diagram shows the iterates are in a two-cycle.

20, the solution alternates between the values 0.5130 and 0.7995. The graph is drawn for the particular initial condition $u_0 = 0.3$, but it is similar for all other initial values between 0 and 1. Figure 2.9.4*b* also shows the same steady oscillation as a rectangular path that is traversed repeatedly in the clockwise direction.

At about $\rho = 3.449$, each state in the oscillation of period 2 separates into two distinct states, and the solution becomes periodic with period 4; see Figure 2.9.5, which shows a solution of period 4 for $\rho = 3.5$. As ρ increases further, periodic solutions of period 8, 16, . . . appear. The transition from solutions with one period to solutions with a new period that occurs at a certain parameter value is called a **bifurcation**; the value of the parameter where the bifurcation occurs is called a **bifurcation value** of the parameter.

The values of ρ at which the successive period doublings occur approach a limit that is approximately 3.57. For $\rho > 3.57$, the solutions possess some regularity but no discernible detailed pattern for most values of ρ. For example, a solution for $\rho = 3.65$ is shown in Figure 2.9.6. It oscillates between approximately 0.3 and 0.9, but its fine structure is unpredictable. The term **chaotic** is used to describe this situation. One of the features of chaotic solutions is extreme sensitivity to the initial conditions. This is illustrated in Figure 2.9.7, where two solutions of equation (21) for $\rho = 3.65$ are shown. One solution is the same as that in Figure 2.9.6 and has the initial value $u_0 = 0.3$, while the other solution has the initial value $u_0 = 0.305$. For about 15 iterations the two solutions remain close and are hard to distinguish from each other in the figure. After that, although they continue to wander about in approximately the same set of values, their graphs are quite dissimilar. It would certainly not be possible to use one of these solutions to estimate the value of the other for values of n larger than about 15.



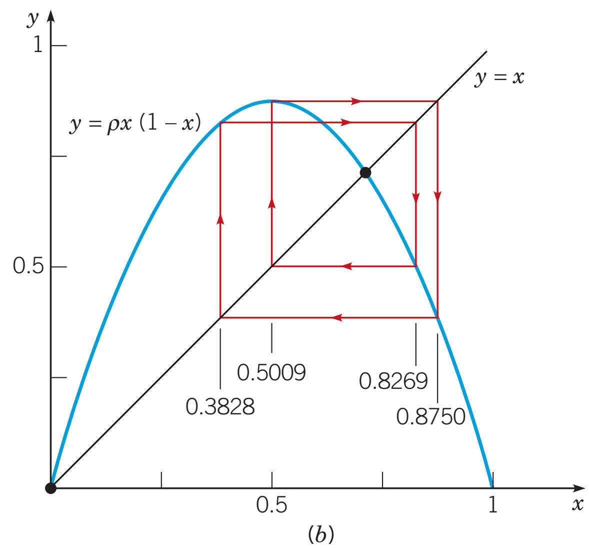


FIGURE 2.9.5 A solution of $u_{n+1} = \rho u_n(1 - u_n)$ for $\rho = 3.5$; period 4. (*a*) u_n versus n; (*b*) the cobweb diagram shows the iterates are in a four-cycle.

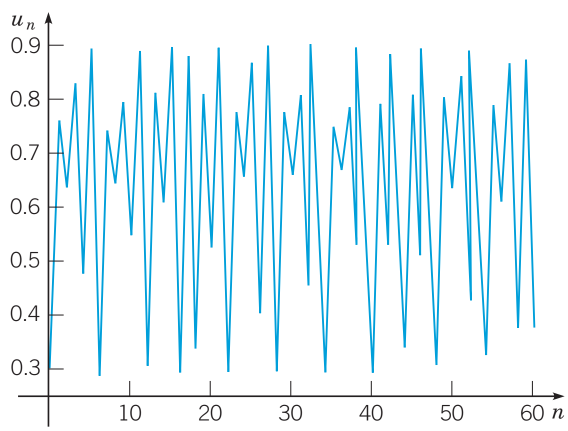


FIGURE 2.9.6 A solution of $u_{n+1} = \rho u_n(1 - u_n)$ for $\rho = 3.65$; a chaotic solution.

It is only comparatively recently that chaotic solutions of difference and differential equations have become widely known. Equation (20) was one of the first instances of mathematical chaos to be found and studied in detail, by Robert May[23] in 1974. On the basis

[23]Robert M. May (1936–) was born in Sydney, Australia, and received his education at the University of Sydney with a doctorate in theoretical physics in 1959. His interests soon turned to population dynamics and theoretical ecology; the work cited in the text is described in two papers listed in the References at the end of this chapter. He has held professorships at Sydney, at Princeton, at Imperial College (London), and (since 1988) at Oxford.

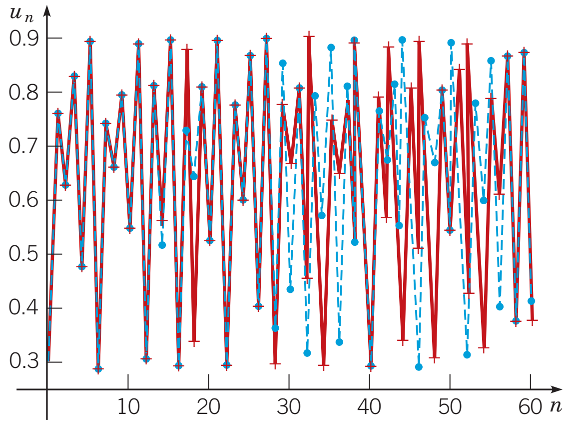


FIGURE 2.9.7 Two solutions of $u_{n+1} = \rho u_n(1 - u_n)$ for $\rho = 3.65$; $u_0 = 0.3$ and $u_0 = 0.305$.

of his analysis of this equation as a model of the population of certain insect species, May suggested that if the growth rate ρ is too large, then it will be impossible to make effective long-range predictions about these insect populations. The occurrence of chaotic solutions in seemingly simple problems has stimulated an enormous amount of research, but many questions remain unanswered. It is increasingly clear, however, that chaotic solutions are much more common than was suspected at first and that they may be a part of the investigation of a wide range of phenomena.

Problems

In each of Problems 1 through 4, solve the given difference equation in terms of the initial value y_0. Describe the behavior of the solution as $n \to \infty$.

1. $y_{n+1} = -0.9 y_n$

2. $y_{n+1} = \sqrt{\dfrac{n+3}{n+1}}\, y_n$

3. $y_{n+1} = (-1)^{n+1} y_n$

4. $y_{n+1} = 0.5 y_n + 6$

5. An investor deposits \$1000 in an account paying interest at a rate of 8%, compounded monthly, and also makes additional deposits of \$25 per month. Find the balance in the account after 3 years.

6. A certain college graduate borrows \$8000 to buy a car. The lender charges interest at an annual rate of 10%. What monthly payment rate is required to pay off the loan in 3 years? Compare your result with that of Problem 7 in Section 2.3.

7. A homebuyer takes out a mortgage of \$100,000 with an interest rate of 9%. What monthly payment is required to pay off the loan in 30 years? In 20 years? What is the total amount paid during the term of the loan in each of these cases?

8. If the interest rate on a 20-year mortgage is fixed at 10% and if a monthly payment of \$1000 is the maximum that the buyer can afford, what is the maximum mortgage loan that can be made under these conditions?

9. A homebuyer wishes to finance the purchase with a \$95,000 mortgage with a 20-year term. What is the maximum interest rate the buyer can afford if the monthly payment is not to exceed \$900?

The Logistic Difference Equation. Problems 10 through 15 deal with the difference equation (21), $u_{n+1} = \rho u_n(1 - u_n)$.

10. Carry out the details in the linear stability analysis of the equilibrium solution $u_n = (\rho - 1)/\rho$. That is, derive the difference equation (26) in the text for the perturbation v_n.

11. N **a.** For $\rho = 3.2$, plot or calculate the solution of the logistic equation (21) for several initial conditions, say, $u_0 = 0.2, 0.4, 0.6$, and 0.8. Observe that in each case the solution approaches a steady oscillation between the same two values. This illustrates that the long-term behavior of the solution is independent of the initial conditions.

N **b.** Make similar calculations and verify that the nature of the solution for large n is independent of the initial condition for other values of ρ, such as 2.6, 2.8, and 3.4.

12. Assume that $\rho > 1$ in equation (21).

G **a.** Draw a qualitatively correct stairstep diagram and thereby show that if $u_0 < 0$, then $u_n \to -\infty$ as $n \to \infty$.

G **b.** In a similar way, determine what happens as $n \to \infty$ if $u_0 > 1$.

13. The solutions of equation (21) change from convergent sequences to periodic oscillations of period 2 as the parameter ρ passes through the value 3. To see more clearly how this happens, carry out the following calculations.

N **a.** Plot or calculate the solution for $\rho = 2.9$, 2.95, and 2.99, respectively, using an initial value u_0 of your choice in the interval $(0, 1)$. In each case estimate how many iterations are required for the solution to get "very close" to the limiting value. Use any convenient interpretation of what "very close" means in the preceding sentence.

N **b.** Plot or calculate the solution for $\rho = 3.01$, 3.05, and 3.1, respectively, using the same initial condition as in part a. In each case estimate how many iterations are needed to reach a steady-state oscillation. Also find or estimate the two values in the steady-state oscillation.

N **14.** By calculating or plotting the solution of equation (21) for different values of ρ, estimate the value of ρ at which the solution changes from an oscillation of period 2 to one of period 4. In the same way, estimate the value of ρ at which the solution changes from period 4 to period 8.

N **15.** Let ρ_k be the value of ρ at which the solution of equation (21) changes from period 2^{k-1} to period 2^k. Thus, as noted in the text, $\rho_1 = 3$, $\rho_2 \cong 3.449$, and $\rho_3 \cong 3.544$.

a. Using these values of ρ_1, ρ_2, and ρ_3, or those you found in Problem 14, calculate $(\rho_2 - \rho_1)/(\rho_3 - \rho_2)$.

b. Let $\delta_n = (\rho_n - \rho_{n-1})/(\rho_{n+1} - \rho_n)$. It can be shown that δ_n approaches a limit δ as $n \to \infty$, where $\delta \cong 4.6692$ is known as the Feigenbaum[24] number. Determine the percentage difference between the limiting value δ and δ_2, as calculated in part a.

c. Assume that $\delta_3 = \delta$ and use this relation to estimate ρ_4, the value of ρ at which solutions of period 16 appear.

G **d.** By plotting or calculating solutions near the value of ρ_4 found in part c, try to detect the appearance of a period 16 solution.

e. Observe that

$$\rho_n = \rho_1 + (\rho_2 - \rho_1) + (\rho_3 - \rho_2) + \cdots + (\rho_n - \rho_{n-1}).$$

Assuming that

$$\rho_4 - \rho_3 = (\rho_3 - \rho_2)\delta^{-1}, \quad \rho_5 - \rho_4 = (\rho_3 - \rho_2)\delta^{-2},$$

and so forth, express ρ_n as a geometric sum. Then find the limit ρ_n as $n \to \infty$. This is an estimate of the value of ρ at which the onset of chaos occurs in the solution of the logistic equation (21).

[24]This result for the logistic difference equation was discovered in August 1975 by Mitchell Feigenbaum (1944–), while he was working at the Los Alamos National Laboratory. Within a few weeks he had established that the same limiting value also appears in a large class of period-doubling difference equations. Feigenbaum, who has a doctorate in physics from M.I.T., is now at Rockefeller University.

Chapter Review Problems

Miscellaneous Problems. One of the difficulties in solving first-order differential equations is that there are several methods of solution, each of which can be used on a certain type of equation. It may take some time to become proficient in matching solution methods with equations. The first 24 of the following problems are presented to give you some practice in identifying the method or methods applicable to a given equation. The remaining problems involve certain types of equations that can be solved by specialized methods.

In each of Problems 1 through 24, solve the given differential equation. If an initial condition is given, also find the solution that satisfies it.

1. $\dfrac{dy}{dx} = \dfrac{x^3 - 2y}{x}$

2. $\dfrac{dy}{dx} = \dfrac{1 + \cos x}{2 - \sin y}$

3. $\dfrac{dy}{dx} = \dfrac{2x + y}{3 + 3y^2 - x}, \quad y(0) = 0$

4. $\dfrac{dy}{dx} = 3 - 6x + y - 2xy$

5. $\dfrac{dy}{dx} = -\dfrac{2xy + y^2 + 1}{x^2 + 2xy}$

6. $x\dfrac{dy}{dx} + xy = 1 - y, \quad y(1) = 0$

7. $x\dfrac{dy}{dx} + 2y = \dfrac{\sin x}{x}, \quad y(2) = 1$

8. $\dfrac{dy}{dx} = -\dfrac{2xy + 1}{x^2 + 2y}$

9. $(x^2y + xy - y) + (x^2y - 2x^2)\dfrac{dy}{dx} = 0$

10. $(x^2 + y) + (x + e^y)\dfrac{dy}{dx} = 0$

11. $(x + y) + (x + 2y)\dfrac{dy}{dx} = 0, \quad y(2) = 3$

12. $(e^x + 1)\dfrac{dy}{dx} = y - ye^x$

13. $\dfrac{dy}{dx} = \dfrac{e^{-x}\cos y - e^{2y}\cos x}{-e^{-x}\sin y + 2e^{2y}\sin x}$

14. $\dfrac{dy}{dx} = e^{2x} + 3y$

15. $\dfrac{dy}{dx} + 2y = e^{-x^2-2x}, \quad y(0) = 3$

16. $\dfrac{dy}{dx} = \dfrac{3x^2 - 2y - y^3}{2x + 3xy^2}$

17. $y' = e^{x+y}$

18. $\dfrac{dy}{dx} + \dfrac{2y^2 + 6xy - 4}{3x^2 + 4xy + 3y^2} = 0$

19. $t\dfrac{dy}{dt} + (t + 1)y = e^{2t}$

20. $xy' = y + xe^{y/x}$

21. $\dfrac{dy}{dx} = \dfrac{x}{x^2y + y^3}$ *Hint:* Let $u = x^2$.

22. $\dfrac{dy}{dx} = \dfrac{x + y}{x - y}$

23. $(3y^2 + 2xy) - (2xy + x^2)\dfrac{dy}{dx} = 0$

24. $xy' + y - y^2e^{2x} = 0, \quad y(1) = 2$

25. Riccati Equations. The equation

$$\frac{dy}{dt} = q_1(t) + q_2(t)y + q_3(t)y^2$$

is known as a Riccati[25] equation. Suppose that some particular solution y_1 of this equation is known. A more general solution containing one arbitrary constant can be obtained through the substitution

$$y = y_1(t) + \frac{1}{v(t)}.$$

Show that $v(t)$ satisfies the first-order *linear* equation

$$\frac{dv}{dt} = -(q_2 + 2q_3 y_1)v - q_3.$$

Note that $v(t)$ will contain a single arbitrary constant.

26. Verify that the given function is a particular solution of the given Riccati equation. Then use the method of Problem 25 to solve the following Riccati equations:

a. $y' = 1 + t^2 - 2ty + y^2; \quad y_1(t) = t$

b. $y' = -\frac{1}{t^2} - \frac{y}{t} + y^2; \quad y_1(t) = \frac{1}{t}$

c. $\frac{dy}{dt} = \frac{2\cos^2 t - \sin^2 t + y^2}{2\cos t}; \quad y_1(t) = \sin t$

27. The propagation of a single action in a large population (for example, drivers turning on headlights at sunset) often depends partly on external circumstances (gathering darkness) and partly on a tendency to imitate others who have already performed the action in question. In this case the proportion $y(t)$ of people who have performed the action can be described[26] by the equation

$$dy/dt = (1 - y)(x(t) + by), \tag{28}$$

where $x(t)$ measures the external stimulus and b is the imitation coefficient.

a. Observe that equation (28) is a Riccati equation and that $y_1(t) = 1$ is one solution. Use the transformation suggested in Problem 25, and find the linear equation satisfied by $v(t)$.

b. Find $v(t)$ in the case that $x(t) = at$, where a is a constant. Leave your answer in the form of an integral.

[25]Riccati equations are named for Jacopo Francesco Riccati (1676–1754), a Venetian nobleman, who declined university appointments in Italy, Austria, and Russia to pursue his mathematical studies privately at home. Riccati studied these equations extensively; however, it was Euler (in 1760) who discovered the result stated in this problem.

[26]See Anatol Rapoport, "Contribution to the Mathematical Theory of Mass Behavior: I. The Propagation of Single Acts," *Bulletin of Mathematical Biophysics 14* (1952), pp. 159–169, and John Z. Hearon, "Note on the Theory of Mass Behavior," *Bulletin of Mathematical Biophysics 17* (1955), pp. 7–13.

Some Special Second-Order Differential Equations. Second-order differential equations involve the second derivative of the unknown function and have the general form $y'' = f(t, y, y')$. Usually, such equations cannot be solved by methods designed for first-order equations. However, there are two types of second-order equations that can be transformed into first-order equations by a suitable change of variable. The resulting equation can sometimes be solved by the methods presented in this chapter. Problems 28 through 37 deal with these types of equations.

Equations with the Dependent Variable Missing. For a second-order differential equation of the form $y'' = f(t, y')$, the substitution $v = y'$, $v' = y''$ leads to a first-order differential equation of the form $v' = f(t, v)$. If this equation can be solved for v, then y can be obtained by integrating $dy/dt = v$. Note that one arbitrary constant is obtained in solving the first-order equation for v, and a second is introduced in the integration for y. In each of Problems 28 through 31, use this substitution to solve the given equation.

28. $t^2 y'' + 2ty' - 1 = 0, \quad t > 0$

29. $ty'' + y' = 1, \quad t > 0$

30. $y'' + t(y')^2 = 0$

31. $2t^2 y'' + (y')^3 = 2ty', \quad t > 0$

Equations with the Independent Variable Missing. Consider second-order differential equations of the form $y'' = f(y, y')$, in which the independent variable t does not appear explicitly. If we let $v = y'$, then we obtain $dv/dt = f(y, v)$. Since the right-hand side of this equation depends on y and v, rather than on t and v, this equation contains too many variables. However, if we think of y as the independent variable, then by the chain rule, $dv/dt = (dv/dy)(dy/dt) = v(dv/dy)$. Hence the original differential equation can be written as $v(dv/dy) = f(y, v)$. Provided that this first-order equation can be solved, we obtain v as a function of y. A relation between y and t results from solving $dy/dt = v(y)$, which is a separable equation. Again, there are two arbitrary constants in the final result. In each of Problems 32 through 35, use this method to solve the given differential equation.

32. $yy'' + (y')^2 = 0$

33. $y'' + y = 0$

34. $yy'' - (y')^3 = 0$

35. $y'' + (y')^2 = 2e^{-y}$

Hint: In Problem 35 the transformed equation is a Bernoulli equation. See Problem 23 in Section 2.4.

In each of Problems 36 through 37, solve the given initial value problem using the methods of Problems 28 through 35.

36. $y'y'' = 2, \quad y(0) = 1, \quad y'(0) = 2$

37. $(1 + t^2)y'' + 2ty' + 3t^{-2} = 0, \quad y(1) = 2, \quad y'(1) = -1$

References

The two books mentioned in Section 2.5 are

Bailey, N. T. J., *The Mathematical Theory of Infectious Diseases and Its Applications* (2nd ed.) (New York: Hafner Press, 1975).

Clark, Colin W., *Mathematical Bioeconomics* (2nd ed.) (New York: Wiley-Interscience, 1990).

A good introduction to population dynamics, in general, is

Frauenthal, J. C., *Introduction to Population Modeling* (Boston: Birkhauser, 1980).

A fuller discussion of the proof of the fundamental existence and uniqueness theorem can be found in many more advanced books on differential equations. Two that are reasonably accessible to elementary readers are

Coddington, E. A., *An Introduction to Ordinary Differential Equations* (Englewood Cliffs, NJ: Prentice-Hall, 1961; New York: Dover, 1989).

Brauer, F., and Nohel, J., *The Qualitative Theory of Ordinary Differential Equations* (New York: Benjamin, 1969; New York: Dover, 1989).

A valuable compendium of methods for solving differential equations is

Zwillinger, D., *Handbook of Differential Equations* (3rd ed.) (San Diego: Academic Press, 1998).

For further discussion and examples of nonlinear phenomena, including bifurcations and chaos, see

Strogatz, Steven H., *Nonlinear Dynamics and Chaos* (Reading, MA: Addison-Wesley, 1994).

A general reference on difference equations is

Mickens, R. E., *Difference Equations, Theory and Applications* (2nd ed.) (New York: Van Nostrand Reinhold, 1990).

Two papers by Robert May cited in the text are

R. M. May, "Biological Populations with Nonoverlapping Generations: Stable Points, Stable Cycles, and Chaos," *Science 186* (1974), pp. 645–647; "Biological Populations Obeying Difference Equations: Stable Points, Stable Cycles, and Chaos," *Journal of Theoretical Biology 51* (1975), pp. 511–524.

An elementary treatment of chaotic solutions of difference equations is

Devaney, R. L., *Chaos, Fractals, and Dynamics* (Reading, MA: Addison-Wesley, 1990).