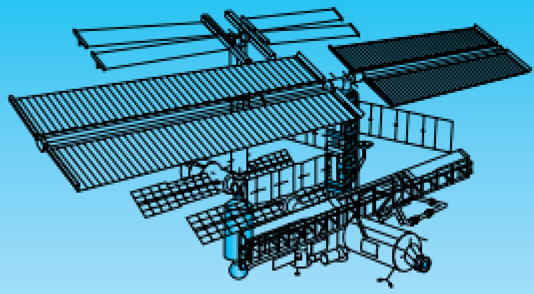
CHAPTER 5



First Order Differential Equations

Differential equations are fundamental to the study of engineering and physics, and this chapter marks the start of our discussion of this important topic. Typically, in an electrical problem, the dependent variable $i(t)$ in an ordinary differential equation might be the current flowing in a circuit at time t, in which case the independent variable would be the time. In all such examples, the nature of $i(t)$ depends on the current flow at the start, and the specification of information of this type is called an **initial condition** for the differential equation. Similarly, in chemical engineering, a dependent variable $m(t)$ might be the amount of a chemical produced by a reaction at time t. Here also the independent variable would be the time t, and to determine $m(t)$ in any particular case it would be necessary to specify the amount of $m(t)$ present at the start, that for convenience is usually taken to be when $t = 0$.

Many physical problems are capable of description in terms of a single first order ordinary differential equation, while other more complicated problems involve coupled first order differential equations, that after the elimination of all but one of the independent variables, can be replaced by a single higher order equation for the remaining dependent variable. This happens, for example, when determining the current in an R-L-C electrical circuit.

Thus first order ordinary differential equations can be considered as the building blocks in the study of higher order equations, and their properties are particularly important and easy to obtain when the equations are linear. The study and properties of the specially simple class of equations called **constant coefficient equations** is very important, as it forms the foundation of the study of higher order constant coefficient equations that will be developed later and have many and varied applications.

Motivation for the study of ordinary differential equations in general is provided by considering a number of typical problems that give rise to different types of differential equation. The first application involves the determination of orthogonal trajectories. A typical example of orthogonal trajectories arises in steady state two-dimensional temperature distributions, where one family of trajectories corresponds to the lines along which the temperature is constant, while the other family corresponds to lines along which heat flows. Other examples considered are the radioactive decay of a substance, the logistic equation and its connection with population growth, damped oscillations, the shape of a suspended power line, and the bending of beams.

The chapter starts by defining an ***m*th order** ordinary differential equation, of which a first order equation is a special case. Various important terms are defined, and the physical significance of **initial** and **boundary conditions** for differential equations are introduced and explained.

The geometrical interpretation of the derivative dy/dx as the slope of a curve is used in Section 5.3 to develop the concept of the **direction field** associated with the first order equation $dy/dx = f(x, y)$. This concept is particularly useful as it leads to a geometrical picture showing the qualitative behavior of all solutions of the differential equation. It will be seen later that the idea underlying a direction field forms the basis of the simple Euler method for the numerical solution of an initial value problem.

First order equations are considered, **separable equations** are defined and solved, and some other special types of equation are introduced that arise in applications, of which the most important is the general **linear first order differential equation**. Its solution is found by using what is called an **integrating factor**. The first order linear differential equation is important, because the structure of its solution is typical of linear differential equations of all orders.

Another special first order equation that is considered is the Bernoulli equation. The Bernoulli equation is an important type of nonlinear equation with many applications, and in a sense it stands on the border between linear and nonlinear first order differential equations. An application of the Bernoulli equation is outlined in the text, and another more detailed one is to be found in the Exercise set at the end of Section 5.8.

The chapter ends by considering the important and practical questions concerning the existence and uniqueness of solutions of $dy/dx = f(x, y)$.

5.1 Background to Ordinary Differential Equations

An **ordinary differential equation (ODE)** is an equation that relates a function $y(x)$ to some of its derivatives $y^{(r)}(x) = d^r y/dx^r$. It is usual to call x the **independent** variable and y the **dependent variable**, and to write the most general ordinary differential equation as

$$F\left(x, y, y^{(1)}, y^{(2)}, \ldots, y^{(n)}\right) = 0. \tag{1}$$

The number n in (1) is called the **order** of the ordinary differential equation, and it is the order of the highest derivative of y that occurs in the equation. A class of ODEs of particular importance in engineering and science, because of their frequency of occurrence and the extensive analytical methods that are available for their solution, are the linear ordinary differential equations.

***n*th order linear variable coefficient equation**

The most general ***n*th order linear differential equation** can be written

$$a_0(x)\frac{d^n y}{dx^n} + a_1(x)\frac{d^{n-1} y}{dx^{n-1}} + \cdots + a_{n-1}(x)\frac{dy}{dx} + a_n(x)y = f(x), \tag{2}$$

with $a_0(x) \neq 0$ and we will consider it to be defined over some interval $a \leq x \leq b$. The functions $a_0(x), a_1(x), \ldots, a_n(x)$, called the **coefficients** of the equation, are known functions, and the known function $f(x)$ is called the **nonhomogeneous term**. The name **forcing function** is also sometimes given to $f(x)$, because in applications it represents the influence of an external input that drives a physical system represented by the differential equation. Equation (2) is called **homogeneous** if $f(x) \equiv 0$.

It will be seen later that the solution of the nonhomogeneous equation (2) is related in a fundamental manner to the solution of its associated homogeneous equation.

When one or more of the coefficients of (2) depend on x, it is called a **variable coefficient** equation. Simpler than **variable coefficient** linear equations, but still of considerable importance, are the linear equations in which the coefficients are the constants $a_0, a_1, \ldots, a_n$, so that (2) becomes

***n*th order linear constant coefficient equation**

$$a_0 \frac{d^n y}{dx^n} + a_1 \frac{d^{n-1} y}{dx^{n-1}} + \cdots + a_{n-1} \frac{dy}{dx} + a_n y = f(x) \quad \text{for} \quad a \leq x \leq b. \tag{3}$$

Equations of this type are called **constant coefficient** linear equations.

If the interval $a \leq x \leq b$ on which equations (2) and (3) are defined is not specified, it is to be understood to be the largest one for which the equations have meaning. Sometimes, in the case of (2), this interval is determined by the variable coefficients $a_r(x)$, whereas in applications it is often determined by the nature of the problem that restricts x to a specific interval.

nonlinear equation and degree

An ordinary differential equation that is not linear is said to be **nonlinear**. Nonlinearity arises in ordinary differential equations because of the occurrence of a nonlinear function of the dependent variable y that sometimes occurs in the form of a power or a radical. The terms homogeneous and nonhomogeneous have no meaning for nonlinear equations.

A term that is also in use, mainly as an indication of the complexity to be expected of a solution, is the *degree* of an equation. The **degree** is the greatest power to which the highest order derivative in the differential equation is raised after the radicals have been cleared from expressions involving the dependent variable y.

EXAMPLE 5.1

(a) The ODE

$$\frac{dy}{dx} + 2xy = \sin x$$

is a linear variable coefficient nonhomogeneous first order equation.

(b) The ODE

$$(1 - x^2)\frac{d^2 y}{dx^2} - 2x\frac{dy}{dx} + 6y = 0, \quad \text{with } -1 < x < 1,$$

is a linear variable coefficient homogeneous second order equation.

(c) The ODE

$$\frac{d^2 y}{dx^2} + a\frac{dy}{dx} + by = \sin \omega x, \quad \text{with } \omega = \text{constant},$$

is a linear constant coefficient nonhomogeneous second order equation.

(d) The ODE

$$\frac{d^2\theta}{dt^2} + k \sin\theta = 0, \quad \text{with } k = \text{constant}$$

is a nonlinear second order equation because θ occurs nonlinearly in the function $\sin\theta$.

(e) The ODE

$$k\frac{d^2y}{dx^2} = f(x)[1 + (dy/dx)^2]^{3/2}, \quad \text{with } k > 0 \text{ a constant}$$

is a nonlinear second order equation of degree 2 involving a power and a radical.

general and particular solutions, and integral curves

A **solution** of an ordinary differential equation is a function $y = \Phi(x)$ that, when substituted into the equation, makes it identically zero over the interval on which the equation is defined. A solution of an nth order equation that contains n arbitrary constants is called the **general solution** of the equation. If the arbitrary constants in the general solution are assigned specific values, the result is called a **particular solution** of the equation.

singular solution

For obvious reasons the solution of an ordinary differential equation is also called an **integral curve**. A solution that cannot be obtained from the general solution for any choice of its arbitrary constants is called a **singular solution**. In the case of linear equations *all* possible solutions of the equation can be obtained from the general solution, so linear equations have no singular solutions. Nonlinear equations possess a more complicated structure that often allows the existence of one or more singular solutions.

EXAMPLE 5.2

(a) The general solution of the linear constant coefficient nonhomogeneous equation

$$\frac{d^2y}{dx^2} - 4y = x$$

is $y = Ae^{2x} + Be^{-2x} - x/4$, where A and B are arbitrary constants. This is easily checked, because substituting for y in the equation leads to the identity $x \equiv x$.

(b) The nonlinear equation

$$\left(\frac{dy}{dx}\right)^2 + y^2 = 1$$

has the general solution $y = \sin(x + A)$. However, $y = \pm 1$ are also seen to be solutions, though as these cannot be obtained from the general solution for any choice of A, they are *singular solutions*.

The linear equation (2) is often written in the more compact form

$$L[y] = f(x), \tag{4}$$

linear operator

where L is the **linear operator**

$$L[\cdot] \equiv a_0(x)\frac{d^n}{dx^n} + a_1(x)\frac{d^{n-1}}{dx^{n-1}} + \cdots + a_{n-1}(x)\frac{d}{dx} + a_n(x), \tag{5}$$

with coefficients that may or may not be functions of x. Only when $L[\cdot]$ acts on an n times differentiable function does it produce a function.

Equation (2) is called **linear** because if y_1 and y_2 are any two solutions of the homogeneous form of the equation $L[y] = 0$, the linear combination $y = C_1 y_1 + C_2 y_2$ where C_1 and C_2 are constants is also a solution. In terms of the differential operator $L[\cdot]$ this property becomes $L[C_1 y_1 + C_2 y_2] = C_1 L[y_1] + C_2 L[y_2]$, and it follows directly from the linearity of the differentiation operation, because

$$\frac{d^m}{dx^m}(y_1 + y_2) = \frac{d^m y_1}{dx^m} + \frac{d^m y_2}{dx^m},$$

for $m = 0, 1, \ldots, n$, with $d^0 y/dx^0 \equiv y$.

If $y_1(x), y_2(x), \ldots, y_m(x)$ are solutions of the nth order homogeneous equation $L[y] = 0$, with $m \leq n$ and $C_1, C_2, \ldots, C_m$ arbitrary constants, the linear combination

$$\boxed{y(x) = C_1 y_1(x) + C_2 y(x) + \cdots + C_m y_m(x)}$$

linear superposition

is called a **linear superposition** of the m solutions, and it is also a solution of the homogeneous equation.

Later we will define the linear independence of a set of functions over an interval and show that the homogeneous form of (2) has precisely n linearly independent solutions $y_1(x), y_2(x), \ldots, y_n(x)$, and that its general solution is

$$y_c(x) = C_1 y_1(x) + C_2 y(x) + \cdots + C_n y_n(x), \tag{6}$$

complementary solution, particular integral, and complete solution

where $C_1, C_2, \ldots, C_n$ are arbitrary constants. This general solution of the homogeneous form of equation (2) is called the **complementary function** or the **complementary solution** of (2). A function $y_p(x)$ that is a solution of the nonhomogeneous equation (2) but contains *no* arbitrary constants is called a **particular integral** of (2). The **complete solution** $y(x)$ of equation (2) is

$$y(x) = y_c(x) + y_p(x). \tag{7}$$

In applications of ordinary differential equations the values of the arbitrary constants in specific problems are obtained by choosing them so the solution satisfies auxiliary conditions that identify a particular problem.

Auxiliary conditions specified at a single point $x = a$, say, are called **initial conditions**, because x often represents the time so that conditions of this type describe how the solution starts. An **initial value problem (i.v.p.)** involves finding a solution of a differential equation that satisfies prescribed initial conditions.

boundary and initial conditions

A different type of problem arises when the auxiliary conditions are specified at two different points $x = a$ and $x = b$, say. Conditions of this type are called **boundary conditions**, because in such problems x usually represents a space variable, and the solution is required to be determined between two boundaries located at $x = a$ and $x = b$ where boundary conditions are prescribed. A **boundary value problem (b.v.p.)** involves finding a solution of a differential equation that satisfies prescribed boundary conditions.

EXAMPLE 5.3

(a) The linear nonhomogeneous ordinary differential equation

$$\frac{d^2y}{dx^2} + y = x$$

has the general solution $y = A\cos x + B\sin x + x$. This equation together with the initial conditions $y(0) = 0$, $y'(0) = 0$ specified at the point $x = 0$ constitutes an *initial value problem* for y. Choosing A and B to satisfy these initial conditions shows the unique solution of this i.v.p. to be $y = x - \sin x$ for $x \geq 0$.

(b) The linear homogeneous ordinary differential equation

$$\frac{d^2y}{dx^2} + y = 0$$

has the general solution $y = A\cos x + B\sin x$. This equation together with the conditions $y(0) = 0$, $y'(\pi/3) = 3$ specified at the two different points $x = 0$ and $x = \pi/3$ constitutes a *boundary value problem* for y. Choosing A and B to satisfy these conditions shows that this b.v.p. has the unique solution $y = 6\sin x$ for $0 < x < \pi/3$.

(c) Consider the linear homogeneous ordinary differential equation

$$\frac{d^2y}{dx^2} - y = 0 \quad \text{defined for } x \geq 0,$$

which is easily seen to have the general solution $y = Ae^x + Be^{-x}$. Imposing the boundary conditions $y(0) = 1$ and $y(+\infty) = 0$ constitutes a boundary value problem for y in which one condition is at $x = 0$ and the other is at plus infinity. The condition at infinity can only be satisfied if $A = 0$, so matching the solution $y = Be^{-x}$ to the condition $y(0) = 1$ shows that this b.v.p. has the **unique** (only) solution $y = e^{-x}$.

unique and nonunique solutions

(d) It is possible for a boundary value problem to have a unique solution as in (b), more than one solution, or no solution at all. More will be said about this later, but for the moment we give a simple example that shows why a boundary value problem may have many solutions or no solution.

The general solution of (b) is $y = A\cos x + B\sin x$, so if the boundary conditions $y(0) = 0$ and $y(\pi) = 0$ are imposed we find that $A = 0$ and B is indeterminate, so it may be assigned any value. In this case a solution certainly exists, as it is given by $y = B\sin x$, but B is arbitrary, so there is more than one solution. When more than one solution can be found that satisfies the auxiliary conditions, the solution is said to be **nonunique**.

If, in this example, the boundary conditions are replaced by $y(0) = 0$ and $y(\pi) = 1$, no choice of constants A and B can make the general solution satisfy the boundary conditions, so in this case there is no solution. ■

Summary

This section introduced the concept of an nth order ordinary differential equation, and the initial and boundary conditions that such equations are often required to satisfy. Emphasis was placed on linear equations and, in particular, on the structure of the solution of a linear first order equation, because the structure of the solution of this fundamental type of equation is shared by the solutions of all higher order linear equations.

EXERCISES 5.1

In Exercises 1 through 10, determine the order and degree of the equation and classify it as homogeneous linear, nonhomogeneous linear, or nonlinear.

1. $y''' + 3y'' + 4y' - y = 0$.
2. $y'' + 4y' + y = x \sin x$.
3. $y'' + x(y')^2 = \cosh x$.
4. $(y'')^{3/2} + xy' = [(1 + x)y']$.
5. $y'' + 3y' + 2y = x^2 \sin y$.
6. $y^{(4)} + x^2\sqrt{y} = 3 + x^3$.
7. $y' + 3xy = 1 + x^2$.
8. $y'' + y = \tan(y')$.
9. $(2 + x^2)y' + x(1 - y^2) = 0$.
10. $y'/y + \sin x = 3$.

5.2 Some Problems Leading to Ordinary Differential Equations

Before we develop methods for the solution of ordinary differential equations, it will be helpful to examine some simple geometrical and physical problems that lead to ODEs. There are many such problems, so we only consider some representative examples.

(a) A Geometrical Problem: Orthogonal Trajectories

The equation

$$F(x, y, c) = 0,$$

where the real variable c is a *parameter*, defines a **one-parameter** family of curves in the (x, y)-plane. This means that assigning a specific value to c determines a particular curve in the (x, y)-plane, and a different value of c will determine a different curve. It often happens that the equation $F(x, y, c) = 0$ defines y implicitly in terms of x, so that the equation cannot be solved explicitly as $y = f(x, c)$.

orthogonal trajectory

A curve that intersects every member of a one-parameter family of curves orthogonally (at right angles) is called an **orthogonal trajectory** of the family. A geometrical problem that often occurs is how to find a *family* of curves that form orthogonal trajectories to a given family.

isotherms, heat flow, streamlines, equipotentials, and flux lines

When some applications of conformal mapping to two-dimensional physical problems are considered in Chapter 17, it will be seen that orthogonal trajectories arise in the study of steady state heat conduction, fluid dynamics, and electromagnetic theory. In heat conduction (see Chapter 18), one family of curves represents lines of constant temperature called **isotherms**, and their orthogonal trajectories then represent **heat flow** lines. In two-dimensional fluid dynamics, orthogonal trajectories express the relationship between the curves followed by fluid particles called **streamlines**, and the associated **equipotential lines** along which a function called the **fluid potential** is constant. In two-dimensional electromagnetic theory an analogous situation arises where one family of curves describes lines of constant electric potential, again called **equipotential lines**, and the family of orthogonal trajectories that describes what are then called **flux lines**.

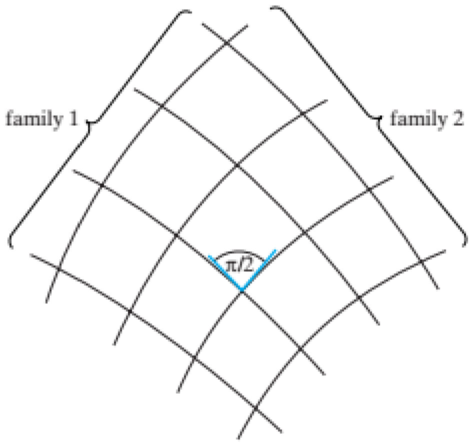


FIGURE 5.1 Two typical families of orthogonal trajectories.

Two typical families of orthogonal trajectories are illustrated in Fig. 5.1, and if these curves are related to steady state heat flow, family 1 could represent the isotherms and family 2 the heat flow lines.

Two specific examples of families of orthogonal trajectories are shown in Fig. 5.2, where in case (a) the curves are given by

$$x^2 + y^2 = c^2 \quad \text{and} \quad y = kx \qquad (\text{with } c \text{ and } k \text{ real}).$$

The first equation describes a family of concentric circles centered on the origin, and the second family that forms their orthogonal trajectories comprises all the straight lines that pass through the origin.

In case (b) the curves are given by

$$x^2 - y^2 = c \quad \text{and} \quad xy = k \qquad (\text{with } c \text{ and } k \text{ real}),$$

where the two families of curves are families of mutually orthogonal rectangular hyperbolas.

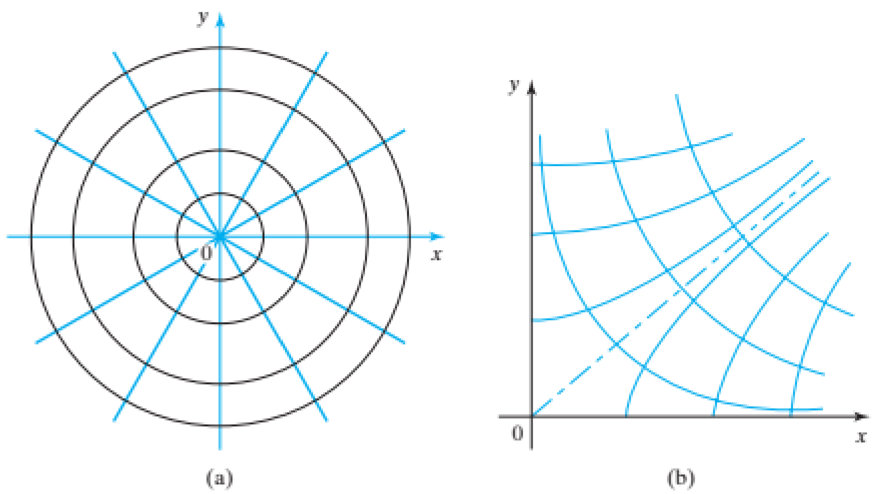


FIGURE 5.2 Specific examples of orthogonal trajectories.

In general the equation

$$F(x, y, c) = 0, \tag{8}$$

with c a parameter, describes a family of curves. To find their orthogonal trajectories we first need to obtain the differential equation for the family of curves determined by (8). This can be done by differentiating (8) with respect to x and then eliminating c between (8) and the equation with dy/dx to arrive at a differential equation of the form

$$\frac{dy}{dx} = f(x, y). \tag{9}$$

If the family of curves described by this differential equation is to be orthogonal to another family, the products of the gradients of every pair of intersecting curves must equal -1. So the gradient dy/dx of the family of curves that are mutually orthogonal to those of (9) must be such that

$$\frac{dy}{dx} = -\frac{1}{f(x, y)}. \tag{10}$$

This is the differential equation of the required family of orthogonal trajectories. In general (10) can often be solved by the method of separation of variables that will be discussed later.

(b) Chemical Reaction Rates and Radioactive Decay

In many circumstances, for a limited period of time, the rate of reaction of a chemical process can be considered to be proportional only to the amount Q of the chemical that is present at a given time t. The differential equation governing such a process then has the form

$$\frac{dQ}{dt} = kQ, \tag{11}$$

where $k \geq 0$ is a constant of proportionality. This is a homogeneous linear first order differential equation.

An analogous situation applies to the radioactive decay of an isotope for which the decay takes place at a rate proportional to the amount of radioactive isotope that is present at any given instant of time. The equation governing the amount Q of the isotope as a function of time t is also of the form shown in (11), but instead of the amount growing as in the previous case, it is decreasing, so as in this case the constant of proportionality is usually denoted by a positive number λ, the equation for radioactive decay takes the form

$$\frac{dQ}{dt} = -\lambda Q. \tag{12}$$

It is not difficult to see by inspection that the general solution of (12) is

$$Q = Q_0 e^{-\lambda t},$$

half-life

where Q_0 is the amount of the isotope present at the start when $t = 0$. The so-called **half-life** T_h of an isotope is the time taken for half of it to decay away, so setting $Q = (1/2)Q_0$ in the above result shows the half-life to be given by $T_h = (1/\lambda)\ln 2$.

(c) The Logistic Equation: Population Growth

In the study of phenomena involving the rate of increase of a quantity of interest, it often happens that the rate is influenced both by the amount of the quantity that is present at any given instant of time and by the limitation of a resource that is necessary to enable an increase to occur. Such a situation arises in a population of animals that compete for limited food resources, leading to the so-called *predator–prey* situations where an animal (the predator) feeds on another species (the prey) with the effect that overfeeding leads to starvation. This in turn leads to a reduction in the number of predators that in turn can lead to a recovery of the food stock. Similar situations arise in manufacturing when there is competition for scarce resources, and in a variety of similar situations.

To model the situation we let P represent the amount of the quantity of interest present at a given time t, and M represent the amount of resources available at the start. Then a simple model for this process is provided by the differential equation

$$\frac{dP}{dt} = kP(M - P), \tag{13}$$

logistic equation

in which k is a constant of proportionality. When constructing this equation the assumption has been made that the rate of increase dP/dt is proportional to both the amount P that is present at time t and to the amount $M - P$ that remains. Equation (13) is called the **logistic equation**, and it is nonlinear because of the presence of the term $-kP^2$ on the right, though it is easily integrated by the method of separation of variables to be described later.

(d) A Differential Equation that Models Damped Oscillations

damping

Mechanical and electrical systems, and control systems in general, can exhibit oscillatory behavior that after an initial disturbance slowly decays to zero. The process producing the decay is a *dissipative* one that removes energy from the system, and it is called **damping**. To see the prototype equation that exhibits this phenomenon we need only consider the following very simple mechanical model. A mass M rests on a rough horizontal surface and is attached by a spring of negligible mass to a fixed point. The mass–spring system is caused to oscillate along the line of the spring by being displaced from its equilibrium position by a small amount and then released. Figure 5.3a shows the system in its equilibrium configuration, and Fig. 5.3b shows it when the mass has been displaced through a distance x from its rest position.

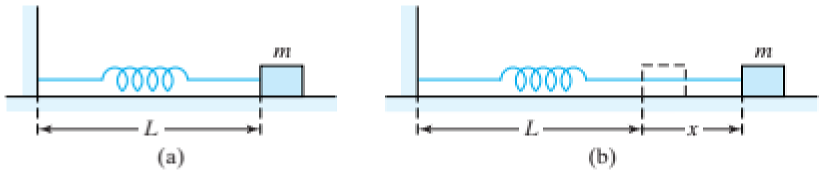


FIGURE 5.3 Mass–spring system.

If t is the time, the acceleration of the mass is d^2x/dt^2, so the force acting due to the motion is Md^2x/dt^2. The forces opposing the motion are the spring force, assumed to be proportional to the displacement x from the equilibrium position, and the frictional force, assumed to be proportional to the velocity dx/dt of the mass M. If the spring constant of proportionality is p and the frictional constant of proportionality is k, the two opposing forces are kdx/dt due to friction and px due to the spring. Equating the forces acting along the line of the spring and taking account of the fact that the spring and frictional forces oppose the force due to the acceleration shows the equation of motion to be the homogeneous second order linear equation

$$M\frac{d^2x}{dt^2} = -k\frac{dx}{dt} - px,$$

or

$$\frac{d^2x}{dt^2} + a\frac{dx}{dt} + bx = 0, \tag{14}$$

where $a = k/M$ and $b = p/M$.

If an external force $Mf(t)$ is applied to the spring, the equation governing the damped oscillations becomes the linear nonhomogeneous second order equation

$$\frac{d^2x}{dt^2} + a\frac{dx}{dt} + bx = f(t).$$

An equation of the same form as (14) governs the oscillation of the charge q in the R–L–C electric circuit shown in Fig. 5.4. The open circuit is shown in Fig. 5.4a with the plates of the capacitor C carrying initial charges Q and $-Q$, while Fig. 5.4b shows the circuit when the switch S has been closed, causing a current i to flow due to a charge is q at time t.

The respective potential drops in the direction of the arrow across the resistor R, the inductance L, and the capacitor C are $V = iR$, where $i = dq/dt$, Ldi/dt,

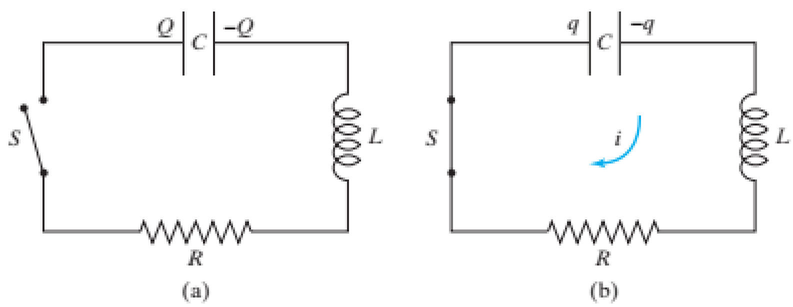


FIGURE 5.4 An R–L–C circuit.

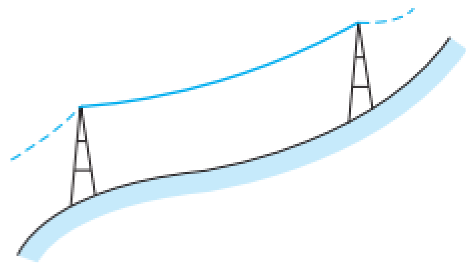


FIGURE 5.5 Suspended cable.

and q/C. Applying Kirchhoff's law, which requires the sum of the potential drops around the circuit to be zero, gives

$$L\frac{di}{dt} + Ri + \frac{q}{C} = 0.$$

Eliminating i by using the result $i = dq/dt$ leads to the following homogeneous linear second order equation for q:

$$LC\frac{d^2q}{dt^2} + RC\frac{dq}{dt} + q = 0.$$

This ODE is of the same form as (14) with $a = R/L$ and $b = 1/LC$.

(e) The Shape of a Suspended Power Line: The Catenary

An analysis of the forces acting on a power line attached to pylons as shown in Fig. 5.5, or on the suspension cable of a cable car, shows the shape of the cable to be determined by the solution $y(x)$ of the nonlinear differential equation

$$\frac{d^2y}{dx^2} = a\sqrt{1 + (dy/dx)^2}.$$

The shape taken by the cable is called a **catenary**, after the Latin word *catena*, meaning chain. Although this equation will not be solved here, it is not difficult to show that its solution is a hyperbolic cosine curve.

(f) Bending of Beams

An analysis of the forces and moments acting on a horizontal beam of uniform construction made from a material with Young's modulus E and supported at its two end points, with the moment of inertia of its cross-section about the central horizontal axis of the beam equal to I, leads to the following equation for the vertical

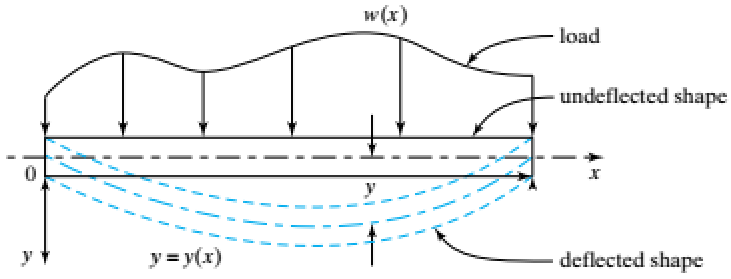


FIGURE 5.6 Deflection of a loaded beam.

deflection y caused by the weight of the beam and any loads it is supporting:

$$\frac{EI\, d^2y/dx^2}{[1+(dy/dx)^2]^{3/2}} = M(x). \tag{15}$$

Here $M(x)$ is the bending moment that acts to one side of a point x in the beam. If a distributed load of line density $w(x)$ acts along the beam creating a load $\int_a^b w(x)dx$ on the segment from $x = a$ to $x = b$, as represented in Fig. 5.6, it can be shown that $M(x)$ and $w(x)$ are related by the result

$$\frac{d^2M}{dx^2} = -w(x). \tag{16}$$

Using this result in (15) shows that the deflection $y(x)$ is determined by the solution of the nonlinear fourth order equation

$$\frac{d^2}{dx^2}\left\{\frac{EId^2y/dx^2}{[1+(dy/dx)^2]^{3/2}}\right\} = w(x), \tag{17}$$

flexural rigidity

in which the product EI is called the **flexural rigidity** of the beam. If the bending is small and the term $(dy/dx)^2$ can be neglected, (17) simplifies to the linear fourth order constant coefficient equation

$$\frac{d^4y}{dx^4} = \frac{w(x)}{EI},$$

which can be solved by direct integration.

Many applications of ordinary differential equations to physical problems are to be found in reference [3.6].

Summary

This section has provided mathematical and physical examples of problems that give rise to ordinary differential equations, some with initial conditions and others with boundary conditions. The logistic equation was seen to be nonlinear and first order, whereas others such as the equation governing radioactive decay and the equation describing damped oscillations were seen to be linear and of first and second order, respectively. The beam equation is nonlinear, though when the bending is small it was seen to reduce to a simple linear fourth order equation that could be solved by direct integration.

EXERCISES 5.2

1. Derive the differential equation that describes the families of circles that are tangent to both the x- and y-axes.
2. Derive the differential equation satisfied by all curves such that the magnitude of the area under the curve between any two ordinates at $x = a$ and $x = b$ is proportional to the magnitude of the arc length of the curve from $x = a$ to $x = b$. Verify that the *catenary* $y(x) = k\cosh(x/k - K)$ is such a curve, with k and K parameters.
3.* A launch travels along the y-axis a constant speed U, starting from the origin, and a police launch starting from a point $a > 0$ on the x-axis pursues it at a constant speed $V > U$. If t is the time measured from the start of the pursuit, write down the differential equation that describes the pursuit path. At all times the police launch steers toward the first launch.

5.3 Direction Fields

In certain applications of mathematics it is necessary to know the qualitative behavior of solutions of a general first order equation

$$\frac{dy}{dx} = f(x, y) \tag{18}$$

over the entire (x, y)-plane, when either no analytical solution is available or, if one exists, it is too complicated to be useful. General properties of solutions of (18) that are known throughout the (x, y)-plane are called **global properties**. A typical global property might be that the solutions are known to be bounded for all x.

global properties

A numerical solution of (18) can always be obtained for any given initial condition (see Chapter 19), but it is impracticable to obtain such solutions for a large enough set of initial conditions simply to enable general the behavior of solutions all over the (x, y)-plane to be understood.

A convenient answer to this problem involves constructing a graphical representation of what is called the *direction field* of (18) at a conveniently chosen mesh of points covering a region R of interest in the (x, y)-plane.

The idea involved is simple and starts by dividing the interval $a \le x \le b$ into m subintervals of equal length $\Delta x = (b - a)/m$, and the interval $c \le y \le d$ into n subintervals of equal length $\Delta y = (d - c)/n$. The mesh of points to be used to cover R are then located at the points (x_r, y_s), where $x_r = a + r\,\Delta x$ and $y_s = c + s\,\Delta y$ with $r = 0, 1, \ldots, m$ and $s = 0, 1, \ldots, n$.

Once the mesh has been chosen, the function $f(x, y)$ is evaluated at each of the points (x_r, y_s). It follows directly that the number $f(x_r, y_s)$ associated with the point (x_r, y_s) is the *gradient* (slope) of the integral curve (solution curve) that passes through that point. Accordingly, the next step is to construct through each point (x_r, y_s), a small straight line segment making an angle $\theta_{rs} = \text{Arctan}\ f(x_r, y_s)$ with the x-axis, as in Fig. 5.7a.

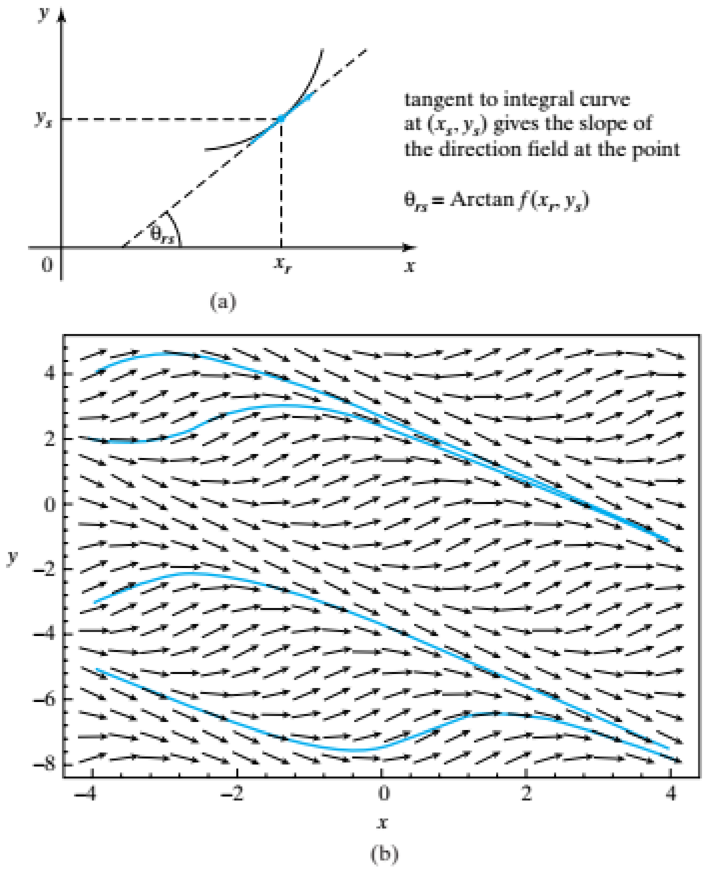


FIGURE 5.7 (a) The construction of a direction field vector at the point (x_r, y_s). (b) The direction field and integral curves for $dy/dx = \cos(x + y)$.

By the nature of their construction, each line segment that is drawn in this manner is tangent to the integral curve that passes through the point through which the segment is drawn. An examination of the pattern of the line segments indicates the overall pattern of behavior of all of the integral curves passing through region R. The assignment of a gradient $f(x, y)$ to each point of R is said to define the **direction field** of the ODE in (18) over R, and the method just described is its geometrical interpretation at a finite number of points of R.

direction field

The graphical interpretation of a direction field can be used to obtain an approximation to the integral curve that passes through an initial point (x_0, y_0) in R. This is accomplished by starting with the line segment through the point (x_0, y_0) and then joining up successive line segments as they intersect one another. As the construction of a direction field over a large region involves many calculations, it is usual to construct them with the aid of a computer.

The direction field for the nonlinear first order equation

$$\frac{dy}{dx} = \cos(x + y)$$

over the region $-4 \leq x \leq 4$ and $-8 \leq y \leq 5$ is shown in Fig. 5.7b, to which have been added some integral curves to show their relationship to the direction field.

Summary

The concept of a direction field of a first order differential equation $dy/dx = f(x, y)$ was introduced in this section. It is a graphical representation of the slope (gradient) of solution curves of the differential equation where they pass through a rectangular mesh of points inside a region of the (x, y)-plane where the solution of the differential equation is of interest. It involves plotting at each mesh point (x_i, y_i) a short segment of the tangent to the solution curve with slope $f(x_i, y_i)$ that passes through that point, to which is added an arrow showing the direction in which the solution is changing as x increases. A direction field provides a geometrical representation of the global nature of the solution inside the region of interest, and tracing successive line segments from one to another, starting from any mesh point, provides a rough picture of the solution curve that originates from the initial condition represented by that mesh point.

EXERCISES 5.3

In each of the following exercises, with the aid of a computer algebra package: (a) Construct the direction field for the given equation at a suitable number of mesh points, (b) use the results of (a) to sketch some representative integral curves, and (c) compare an approximate integral curve through a chosen initial point (x_0, y_0) with the exact solution found by requiring the given general solution to pass through that point.

1. $dy/dx = y + 2x$; $y = Ce^x - 2 - 2x$.
2. $dy/dx = y + 2\cos x$; $y = Ce^x - \cos x + \sin x$.
3. $dy/dx = 2x - y$; $y = Ce^{-x} - 2 + 2x$.
4. $dy/dx = x(1 + y/2)$; $y = C\exp(x^2/4) - 2$.
5. $dy/dx = y + x^2$; $y = Ce^x - 2 - 2x - x^2$.

5.4 Separable Equations

Sometimes the function $f(x, y)$ in the first order differential equation

$$\frac{dy}{dx} = f(x, y) \tag{19}$$

can be written as the product of a function $F(x)$ depending only on x and a function $G(y)$ depending only on y, so that $f(x, y) = F(x)G(y)$, allowing (19) to be written

$$\frac{dy}{dx} = F(x)G(y). \tag{20}$$

two forms of a separable equation

When (19) can be expressed in this simple form, its variables x and y are said to be **separable**, and the equation itself to be of **variables separable** type. If we use differential notation, (20) becomes

$$\frac{1}{G(y)}dy = F(x)dx, \tag{21}$$

so provided $G(y) \neq 0$, equation (21) can be solved by routine integration of the left side with respect to y and of the right side with respect to x. Thus, in principle, the solution of a first order differential equation in which the variables are separable can always be found, though in practice the integrals involved may be difficult or sometimes impossible to evaluate analytically.

Separable first order equations

The differential equation

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

is said to be **separable** if it can be written in the form

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

or, in differential form,

$$\frac{1}{G(y)}dy = F(x)dx.$$

EXAMPLE 5.4

examples of separable equations

Solve the logistic equation

$$\frac{dP}{dt} = kP(M - P)$$

given in equation (13) of Section 5.2(c), assuming $k > 0$ and $0 \leq P \leq M$. Find the solution of the initial value problem in which $P = P_0$ when $t = 0$, and draw some typical integral curves.

Solution The equation is separable and can be written in the differential form

$$\frac{dP}{P(M - P)} = kdt.$$

If we write the left-hand side in partial fraction form, the equation becomes

$$\frac{dP}{P} + \frac{dP}{(M - P)} = Mkdt,$$

and after integration we find that

$$\ln\left|\frac{P}{M - P}\right| = Mkt + C,$$

where C is an arbitrary constant of integration. As the solution for P must lie in the interval $0 \leq P \leq M$, this result simplifies to

$$P = \frac{MA}{A + \exp(-Mkt)},$$

where A is an arbitrary constant.

The arbitrary constant A is related to C by $A = e^C$, but as C is arbitrary, the constant A is also arbitrary, so for simplicity we denote the arbitrary constant in this last result by A without mentioning how it is related to C. In general, arithmetic is not usually performed on arbitrary constants, so after algebraic manipulations, either constants are renamed or the same symbol is used for a related constant.

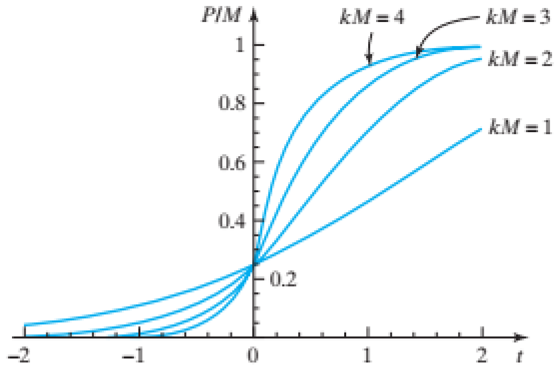


FIGURE 5.8 Integral curves for the logistic equation.

To solve the initial value problem we must find A such that $P = P_0$ when $t = 0$, from which it is easily seen that $A = P_0/(M - P_0)$. The required particular solution is thus

$$P = \frac{MP_0}{P_0 + (M - P_0)\exp(-Mkt)}.$$

Representative integral curves of $P(t)/M$ obtained from this expression using $P_0/M = 1/4$ and $kM = 1, 2, 3$, and 4 are shown in Fig. 5.8 for $-2 \le t \le 2$. ■

EXAMPLE 5.5 Solve the initial value problem for the equation expressed in differential form

$$x^2y^2dx - (1 + x^2)dy = 0, \quad \text{given that } y(0) = 1.$$

Solution The equation is separable because it can be written

$$\frac{dy}{y^2} = \frac{x^2}{(1 + x^2)}dx.$$

Integration gives

$$\int \frac{dy}{y^2} = \int \frac{x^2}{(1 + x^2)}dx,$$

and after the integrations have been performed this becomes

$$-1/y = x - \text{Arctan}\, x + C,$$

where C is an arbitrary constant of integration. This general solution will satisfy the initial condition $y(0) = 1$ if $C = -1$, so the required solution is seen to be

$$y = 1/(\text{Arctan}\, x - x + 1).$$ ■

EXAMPLE 5.6 Derive the differential equation that determines the orthogonal trajectories of the one parameter family of curves $y = Cxe^x$, and solve it to find the equation of these trajectories.

Solution The differential equation describing the family of curves $y = Cxe^x$ is found by first calculating $y'(x)$, and then using the original equation to eliminate C

from the result. We have

$$y'(x) = Ce^x(1 + x),$$

but from the original equation $C = y/xe^x$, so eliminating C between these two results shows that the required differential is

$$y'(x) = y(1 + x)/x.$$

The product of the gradient $y'(x)$ of curves belonging to this family and the gradient of the family of orthogonal trajectories must equal -1 (see Section 5.2(a)), so the differential equation of the orthogonal trajectories is the separable equation

$$\frac{dy}{dx} = -\frac{x}{y(1 + x)}.$$

After separation of the variables and integration, this becomes

$$\int y dy = -\int \frac{x}{1 + x} dx,$$

so that

$$y^2 = \ln(1 + x)^2 - 2x + C. \qquad \blacksquare$$

EXAMPLE 5.7 A circular metal radiator pipe has inner radius R_1 and outer radius R_2 ($R_2 > R_1$). When operating under steady conditions the radial temperature distribution $T(r)$ in the metal wall of the pipe is known to be a solution of the ordinary differential equation (see the heat equation in cylindrical polar coordinates in Section 18.5)

$$r\frac{d^2T}{dr^2} + \frac{dT}{dr} = 0.$$

(i) Find the radial temperature distribution in the pipe wall when the inner surface is maintained at a constant temperature T_1 and the outer surface is maintained at a constant temperature T_2.

(ii) Find the radial temperature distribution in the pipe wall when the inner surface is maintained at a constant temperature T_1 and heat is lost by radiation from the outer surface according to Newton's law of cooling that requires the heat flux across the outer surface to be proportional to the difference in temperature between the surface and the surrounding air at a temperature T_2.

Solution

(i) Setting $u = dT/dr$ the equation becomes the separable equation

$$r\frac{du}{dr} + u = 0 \quad \text{and so} \quad \frac{du}{u} = -\frac{dr}{r},$$

from which it follows that

$$\ln u = -\ln r + \ln A,$$

where for convenience the arbitrary integration constant has been written $\ln A$. Thus $ur = A$, so after substituting for u and again separating variables we have

$$\frac{dT}{dr} = \frac{A}{r}.$$

A final integration gives the general solution

$$T(r) = A \ln r + B,$$

where B is another arbitrary integration constant.

Matching the arbitrary constants A and B to the required conditions $T(R_1) = T_1$ and $T(R_2) = T_2$ then gives the required solution

$$T(r) = \frac{T_1 \ln(R_2/r) + T_2 \ln(r/R_1)}{\ln(R_2/R_1)}.$$

(ii) The heat flux across the surface $r = R_2$ is proportional to dT/dr at $r = R_2$, and this in turn is proportional to the temperature difference $T(R_2) - T_2$, so the required boundary condition on the outer surface of the pipe is of the form

$$\left(\frac{dT}{dr}\right)_{r=R_2} = -h[T(R_2) - T_2],$$

where the negative sign is necessary because heat is being lost across the surface $r = R_2$, and h is a constant depending on the metal in the pipe and the heat transfer condition at its surface.

The general solution is still $T(r) = A \ln r + b$, but now the arbitrary constants A and B must be matched to the condition $T(R_1) = T_1$ on the inside wall of the pipe, and to the above condition derived from Newton's law of cooling. When this is done the temperature distribution in the pipe is found to be

$$T(r) = T_1 + \frac{hR_2(T_2 - T_1)}{1 + hR_2 \ln(R_2/R_1)} \ln\left(\frac{r}{R_1}\right).$$

Summary

This section introduced the important class of separable differential equations $dy/dx = F(x)G(y)$, so called because when written in the form $dy/G(y) = F(x)dx$ the variables are separated by the $=$ sign; they can be integrated immediately provided antiderivatives (indefinite integrals) of $1/G(y)$ and $F(x)$ can be found. This method was used to integrate the nonlinear logistic equation and to obtain the equation of some orthogonal trajectories.

EXERCISES 5.4

In Exercises 1 through 4 solve the given differential equation by hand and confirm the result by using computer algebra.

1. $2yy' = x(1 - 2y)$ with $y(1) = 1$.
2. $2x^2y^2y' + y^4 = 4$ with $y(1) = 3$.
3. $(x^2 - 4)y' = x(1 - 2y)$ with $y(\sqrt{5}) = 1$.
4. $2\sqrt{(1 + x^2)}y' = \sqrt{(1 - y^2)}$ with $y(1) = 1$.

In Exercises 5 through 14 find the general solution of the given differential equation.

5. $\sqrt{(1 + x^2)}y' - 3x\sqrt{(y^2 - 1)} = 0$.
6. $e^{-3x}y' + x \sin 2y = 0$.
7. $2(1 + x)(1 + y)y' + (y + 2)^2 = 0$.
8. $2(x - 1)y' + (x^2 - 2x + 3)\cos^2 y = 0$.
9. $(1 + 3y^2)y' + 2y \ln|1 + x| = 0$.
10. $2(1 - \cos x)y' + 3 \sin y = 0$.
11. $(1 + x^2)yy' - x(y^2 + y + 1) = 0$.
12. $(x^2 + 9)y^2y' - \sqrt{(4 - y^2)} = 0$.
13. $y' \operatorname{ctg} x + 2y = 4$.
14. $(x + 1)y^2y' = x(y^2 + 4)$.

In Exercises 15 through 17 derive and then solve the differential equation that determines the orthogonal trajectories to the given one parameter family of curves.

15. $y = b + k(x - a)$ with a and b constants and k a parameter.

16. $x^2 - 4y^2 + y = c$ with c a parameter.

17. $y = Cx^2e^{2x}$ with C a parameter.

18. A snowball of radius 2 inches is brought into a warm room at a constant temperature above freezing point, and it is found that after 6 hours it has melted to a radius of 1.5 inches. Assuming the melting occurs at a rate proportional to the surface area, write down the differential equation determining the radius as a function of time t in hours, and find the general expression for the radius as a function of time. Comment on any deficiency exhibited by this mathematical model.

19. A simple model called *Malthus' law* for the change in a bacterial population $N(t)$ as a function of time t involves assuming the rate of change is proportional to the population present at time t. Write down the differential equation governing $N(t)$ if the constant of proportionality is $\lambda > 0$, and find an expression for $N(t)$ given that initially $N(0) = N_0$. Find λ if $N(t_1) = N_1$ when $t = t_1$ and $N(t_2) = N_2$ when $t = t_2$, with $N_1 > N_2$ and $t_2 > t_1$. Give a reason why this model is unrealistic when t is large.

20. When a beam of light enters a parallel slab of transparent material at right angles to its plane surface, its intensity I decreases at a rate proportional to the intensity $I(x)$ at a perpendicular distance x into the material. Given a slab of material where the intensity at a distance h into the slab is 40% of the initial intensity, write down the differential equation for $I(x)$. Solve the equation for $I(x)$ and find the distance at which the intensity is 10% of its initial value.

21. The dating of a fossilized bone is based on the amount of radioactive isotope carbon-14 present in the bone. The method uses the fact that the isotope is produced in the atmosphere at a steady rate by bombardment of nitrogen by cosmic radiation when it is absorbed into the living bone. The process stops when the bone is dead, after which the C-14 present in the bone decays exponentially. Assuming the half-life of C-14 is 5600 years, and a bone is found to contain 1/500*th* of the original amount of C-14 that was present originally, determine its age. This approach is called **radioactive carbon dating**.

22. A cylindrical tank of cross-sectional area A standing in a vertical position is filled with water to a depth h. At time $t = 0$ a circular hole of radius a in the bottom of the tank is opened and water is allowed to drain away under gravity. It is known from *Torricelli's law* that the speed of flow of the water through the hole when the water in the tank has depth x is equal to $\sqrt{2gx}$, this being the speed attained by a particle falling freely from rest under gravity through a distance x, where g is the acceleration due to gravity. Write down the differential equation determining the water height $x(t)$ in the tank when $t > 0$, and solve the equation for $x(t)$. If water is added to the tank at a rate $V(t)$, write down the modified equation governing the water height. If $V(t) = V_0$ is constant, and the flow into and out of the tank reaches equilibrium, find the equilibrium height of the water in the tank. Remark: In applications the expression $\sqrt{2gx}$ is replaced by $k\sqrt{2gx}$, with $0 < k < 1$ a constant. The factor k allows for the contraction of the jet after leaving the hole. In the case of water $k \approx 0.6$.

5.5 Homogeneous Equations

homogeneous equation of degree *n*

A function $f(x, y)$ is said to be **algebraically homogeneous of degree *n***, or simply **homogeneous of degree *n***, if $f(tx, ty) = t^n f(x, y)$ for some real number n and all $t > 0$, for $(x, y) \neq (0, 0)$.

EXAMPLE 5.8

(a) If $f(x, y) = x^2 + 3xy + 4y^2$, then $f(tx, ty) = t^2(x^2 + 3xy + 4y^2) = t^2 f(x, y)$, so $f(x, y)$ is homogeneous of degree 2.

(b) If $f(x, y) = \ln|y| - \ln|x|$ for $(x, y) \neq (0, 0)$, then $f(x, y) = \ln|y/x|$, so $f(tx, ty) = f(x, y)$, showing that $f(x, y)$ is homogeneous of degree 0.

(c) If

$$f(x, y) = \frac{x^{3/2} + x^{1/2}y + 3y^{3/2}}{2x^{3/2} - xy^{1/2}}, \text{ then } f(tx, ty) = t^0 f(x, y),$$

showing that $f(x, y)$ is homogeneous of degree 0.

(d) If

$$f(x, y) = x^2 + 4y^2 + \sin(x/y), \text{ then } f(tx, ty) = t^2(x^2 + 4y^2) + \sin(x/y),$$

so $f(x, y)$ is *not* homogeneous, because although both the first group of terms and the last term are homogeneous functions of x and y, they are not both homogeneous of the same degree.

(e) If $f(x, y) = \tan(xy + 1)$, then $f(tx, ty) = \tan(t^2xy + 1)$, so $f(x, y)$ is *not* homogeneous. ■

Homogeneous differential equations

The first order ODE in differential form

$$P(x, y)dx + Q(x, y)dy = 0$$

is called **homogeneous** if P and Q are homogeneous functions of the same degree or, equivalently, if when written in the form

$\dfrac{dy}{dx} = f(x, y)$, the function $f(x, y)$ can be written as $f(x, y) = g(y/x)$.

The substitution $y = ux$ will reduce either form of the homogeneous equation to an equation involving the independent variable x and the new dependent variable u in which the variables are separable. As with most separable equations the solution can be complicated, and it is often the case that y is determined implicitly in terms of x.

EXAMPLE 5.9 Solve

$$(y^2 + 2xy)dx - x^2\,dy = 0.$$

Solution Both terms in the differential equation are homogeneous of degree 2, so the equation itself is homogeneous. Differentiating the substitution $y = ux$ gives

$$\frac{dy}{dx} = u + x\frac{du}{dx}, \quad \text{or} \quad dy = udx + xdu.$$

After substituting for y and dy in the differential equation and cancelling x^2, we obtain the variables separable equation

$$u(u + 1)dx = xdu, \quad \text{or} \quad \frac{du}{u(u + 1)} = \frac{dx}{x}.$$

This has the general solution

$$u = \frac{Cx}{1 - Cx}, \quad \text{but} \quad y = ux \quad \text{and so } y = \frac{Cx^2}{1 - Cx},$$

where C is an arbitrary constant. In this case the general solution is simple and y is determined explicitly in terms of x. ■

EXAMPLE 5.10 Solve

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

Solution The equation is homogeneous because it can be written

$$\frac{dy}{dx} = \frac{(y/x)^2}{(y/x) - 1}.$$

Making the substitution $y = ux$, and again using the result $dy/dx = u + xdu/dx$, reduces this to the separable equation

$$u + x\frac{du}{dx} = \frac{u^2}{u-1}, \text{ or } \left(1 - \frac{1}{u}\right)du = \frac{dx}{x}.$$

Integration gives

$$u - \ln|u| = \ln|x| + \ln|C|,$$

where C is an arbitrary integration constant. Finally, substituting $u = y/x$ and simplifying the result we arrive at the following implicit solution for y:

$$y = Ce^{y/x}.$$

An equation of the form

$$\frac{dy}{dx} = \frac{ax + by + c}{px + qy + r}$$

near-homogeneous

is called **near-homogeneous**, because it can be transformed into a homogeneous equation by means of a variable change that shifts the origin to the point of intersection of the two lines

$$ax + by + c = 0 \quad \text{and} \quad px + qy + r = 0.$$

EXAMPLE 5.11 Solve the initial value problem

$$\frac{dy}{dx} = \frac{y+1}{x+2y} \quad \text{with } y(2) = 0.$$

Solution The equation is near-homogeneous and the lines $y + 1 = 0$ and $x + 2y = 0$ intersect at the point $x = 2$ and $y = -1$, so we make the variable change $x = X + 2$ and $y = Y - 1$, as a result of which the equation becomes the homogeneous equation

$$\frac{dY}{dX} = \frac{Y}{X + 2Y}.$$

Solving this as in Example 5.9 by setting $Y = uX$ leads to the equation

$$-\left(\frac{1+2u}{2u^2}\right)du = \frac{dX}{X},$$

with the solution

$$1/u = 2\ln|CuX|,$$

where C is an arbitrary integration constant. If we set $u = Y/X$, this becomes

$$X = 2Y \ln |CY|,$$

where C is an arbitrary constant. Returning to the original variables by substituting $X = x - 2$, $Y = y + 1$, we arrive at the required general solution

$$x = 2 + 2(y + 1) \ln |C(y + 1)|.$$

Although this is an implicit solution for y, if we regard y as the independent variable and x as the dependent variable, solution curves (integral curves) are easily graphed. Substituting the initial condition $y = 0$ when $x = 2$ in the general solution shows that $C = 1$, so the solution of the initial value problem is

$$x = 2 + 2(y + 1) \ln |y + 1|.$$ ■

Summary

This section introduced the special type of first order ordinary differential equation known as an algebraically homogeneous equation. This name is frequently shortened to the term homogeneous equation, though this must not be confused with the sense in which the term homogeneous is used in Section 5.1. After showing how such equations can be solved, it was shown how a simple linear change of variables changes a near-homogeneous equation to a homogeneous equation that can then be solved.

EXERCISES 5.5

In Exercises 1 through 14 find by hand calculation the general solution of the given homogeneous or near-homogeneous equations and confirm the result by using computer algebra.

1. $y' = y/(2x + y)$.
2. $y' = (2xy + y^2)/(3x^2)$.
3. $y' = (2x^2 + y^2)/xy$.
4. $y' = (2xy + y^2)/x^2$.
5. $y' = (x - y)/(x + 2y)$.
6. $y' = (x + 4y)/x$.
7. $y' = (2x + y\cos^2(y/x))/(x\cos^2(y/x))$.
8. $y' = 3y^2/(1 + x^2)$.
9. $y' = (x + y\sin^2(y/x))/(x\sin^2(y/x))$.
10. $y' = 3x\exp(x + 2y)/y$.
11. $y' = (y + 2)/(x + y + 2)$.
12. $y' = (y + 1)/(x + 2y + 2)$.
13. $y' = (x + y + 1)/(x - y + 1)$.
14. $y' = (x - y + 1)/(x + y)$.

5.6 Exact Equations

The so-called *exact* equations have a simple structure, and they arise in many important applications as, for example, in the study of thermodynamics. After definition of an exact equation, a test for exactness will be derived and the general solution of such an equation will be found.

definition of an exact equation

Exact equations

The first order ODE

$$M(x, y)dx + N(x, y)dy = 0$$

is said to be **exact** if a function $F(x, y)$ exists such that the total differential

$$d[F(x, y)] = M(x, y)dx + N(x, y)dy.$$

It follows directly that if

$$M(x, y)dx + N(x, y)dy = 0 \tag{22}$$

is exact, then the total differential

$$d[F(x, y)] = 0,$$

so the general solution of (22) must be

$$F(x, y) = \text{constant}. \tag{23}$$

EXAMPLE 5.12

The total differential of $F(x, y) = 3x^3 + 2xy^2 + 4y^3 + 2x$ is

$$\begin{aligned} d[F(x, y)] &= (\partial F/\partial x)dx + (\partial F/\partial y)dy \\ &= (9x^2 + 2y^2 + 2)dx + (4xy + 12y^2)dy, \end{aligned}$$

so the exact differential equation

$$(9x^2 + 2y^2 + 2)dx + (4xy + 12y^2)dy = 0$$

has the general solution

$$3x^3 + 2xy^2 + 4y^3 + 2x = \text{constant}.$$

Three questions now arise:

(i) Is there a test for exactness?
(ii) If an equation is exact, is it possible to find its general solution?
(iii) If an equation is not exact, is it possible to modify it to make it exact?

There are satisfactory answers to the first two questions, and a less satisfactory answer to the third question. We deal with the last question first.

It can be shown that an equation of the form (21) that is *not* exact can always be made exact if it is multiplied by a suitable factor $\mu(x, y)$, called an **integrating factor**, though there is no general method by which such an integrating factor can be found. Fortunately, however, an integrating factor can always be found for a variable coefficient linear first order ODE, and in the next section the integrating factor will be derived for such an ODE and then used to find its general solution.

We now turn our attention to the first question. If $F(x, y) = \text{constant}$ is a solution of the exact differential equation

$$M(x, y)dx + N(x, y)dy = 0, \tag{24}$$

then $M(x, y) = \partial F/\partial x$ and $N(x, y) = \partial F/\partial y$. So, provided the derivatives $\partial F/\partial x$, $\partial F/\partial y$, $\partial^2 F/\partial x\partial y$, and $\partial^2 F/\partial y\partial x$ are defined and continuous in the region within which the differential equation is defined, the mixed derivatives will be equal so that $\partial^2 F/\partial x\partial y = \partial^2 F/\partial y\partial x$. This last result is equivalent to requiring that $\partial M/\partial y = \partial N/\partial x$ in order that (24) is exact, so this provides the required test for exactness.

THEOREM 5.1

a simple test for exactness

Test for exactness The differential equation

$$M(x, y)dx + N(x, y)dy = 0$$

is exact if and only if $\partial M/\partial y = \partial N/\partial x$. ■

EXAMPLE 5.13

Test for exactness the differential equations

(a) $\{\sin(xy + 1) + xy\cos(xy + 1)\}dx + x^2\cos(xy + 1)dy = 0$.
(b) $(2x + \sin y)dx + (2x\cos y + y)dy = 0$.

Solution In case (a) $M(x, y) = \sin(xy + 1) + xy\cos(xy + 1)$ and $N(x, y) = x^2\cos(xy + 1)$, and $\partial M/\partial y = \partial N/\partial x$, so the equation is exact.

In case (b) $M(x, y) = 2x + \sin y$ and $N(x, y) = 2x\cos y + y$ but $\partial M/\partial y \neq \partial N/\partial x$, so the equation is not exact. ■

Having established a test for exactness, it remains for us to determine how the general solution of an exact equation can be found. The starting point is the fact that if $F(x, y) =$ constant is a solution of the exact equation

$$M(x, y)dx + N(x, y)dy = 0,$$

then $\partial F/\partial x = M(x, y)$ and $\partial F/\partial y = N(x, y)$.

Two expressions for $F(x, y)$ can be obtained from these results by integrating M with respect to x while regarding y as a constant, and integrating N with respect to y while regarding x as a constant, because this reverses the process of partial differentiation by which M and N were obtained. However, after integrating M it will be necessary to add not only an arbitrary constant, but also an arbitrary function $f(y)$ of y, because this will behave like a constant when F is differentiated partially with respect to x to obtain M. Similarly, after integrating N it will be necessary to add not only an arbitrary constant, but also an arbitrary function $g(x)$ of x, because this will behave like a constant when F is differentiated partially with respect to y to obtain N.

These two expressions for F will look different but must, of course, be identical. The arbitrary function $f(y)$ can be found by identifying it with any function only of y that occurs in the expression for F obtained by integrating N, while the arbitrary function $g(x)$ can be found by identifying it with any function only of x that occurs in the expression for F found by integrating M, where, of course, the true constants introduced after each integration must be identical.

EXAMPLE 5.14

Show the following equation is exact and find its general solution:

$$\{3x^2 + 2y + 2\cosh(2x + 3y)\}dx + \{2x + 2y + 3\cosh(2x + 3y)\}dy = 0.$$

Solution In this equation $M(x, y) = 3x^2 + 2y + 2\cosh(2x + 3y)$, and $N(x, y) = 2x + 2y + 3\cosh(2x + 3y)$, so as $M_y = N_x = 2 + 6\sinh(2x + 3y)$ the equation is exact:

$$F(x, y) = \int M(x, y)dx = \int \{3x^2 + 2y + 2\cosh(2x + 3y)\}dx$$
$$= x^3 + 2xy + \sinh(2x + 3y) + f(y) + C,$$

and

$$F(x, y) = \int N(x, y)dy = \int \{2x + 2y + 3\cosh(2x + 3y)\}dy$$
$$= 2xy + y^2 + \sinh(2x + 3y) + g(x) + D.$$

For these two expressions to be identical, we must set $f(y) \equiv y^2, g(x) \equiv x^3$, and $D = C$, so $F(x, y)$ is seen to be

$$F(x, y) = x^3 + 2xy + y^2 + \sinh(2x + 3y) + C,$$

and so the general solution is

$$x^3 + 2xy + y^2 + \sinh(2x + 3y) = C,$$

where as C is an arbitrary constant we have chosen to write C rather than $-C$ on the right of the solution. ■

Summary

This section introduced the class of first order ordinary differential equations known as exact equations that arise in many different applications. It was then shown how the equality of mixed derivatives yields a simple test for exactness.

EXERCISES 5.6

In Exercises 1 through 8 test the equation for exactness, and when an equation is exact, find its general solution.

1. (a) $\{\sin(3y) + 4x^2y\}dx + \{3x\cos(3y) + y + 2x^3\}dy = 0$;
   (b) $\{4x^3 + 3y^2 + \cos x\}dx + \{6xy + 2\}dy = 0$.
2. (a) $\{(2x + 3y^2)^{-1/2} + 4y^3 + 2x\}dx + \{3y/(2x + 3y^2) + 12xy^2\}dy = 0$;
   (b) $\{\cos(x + 3y^2) + 4xy^3\}dx + \{6y\cos(x + 3y^2) + 3x^2y^2 + 2y\}dy = 0$.
3. (a) $\{\sin x + x\cos x + \cosh(x + 2y)\}dx + \{3y^2 + 2\cosh(x + 2y)\}dy = 0$;
   (b) $\{6x(2x^2 + y^2)^{1/2} + x^2\}dx + 2y(2x^2 + y^2)^{1/2}dy = 0$.
4. (a) $\{6x/(3x^2 + y) + 4xy^3\}dx + \{1/(3x^2 + y) + 6x^2y^2 + 3y^2\}dy = 0$;
   (b) $\{\sin(xy) + xy\cos(xy) + y^2\sin(xy)\}dx + \{x^2\cos(xy) + \cos(xy) - xy\sin(xy)\}dy = 0$.
5. (a) $\dfrac{3x^2}{2\sqrt{x^3 + y^2}}dx + \left\{\dfrac{y}{\sqrt{x^3 + y^2}} + 6y\right\}dy = 0$;
   (b) $\{y/x + 2x\sinh(y^2)\}dx + \{\ln x + 2x^2y\cosh(y^2)\}dy = 0$.
6. (a) $\{4xy + 1/x\}dx + \{2x^2 - 1/y\}dy = 0$;
   (b) $\{6xy - 2/(x^2y)\}dx + \{3x^2 - 2/(xy^2)\}dy = 0$.
7. (a) $\{2xy + 6/x\}dx + \{x^2 + 4/y\}dy = 0$;
   (b) $\{2x/(2x + 3y^2) - 2x^2/(2x + 3y^2)^2 + 2\}dx - 6x^2y/(2x + 3y^2)^2dy = 0$.
8. (a) $\{(5/2)x^{3/2} + 14y^3\}dx + \{(3/2)\sqrt{y} + 42xy^2\}dy = 0$;
   (b) $(y/x^2)\cos(y/x)dx + \{(1/x)\cos(y/x) + 6y\exp(y^2)\}dy = 0$.

5.7 Linear First Order Equations

The **standard form** of the **linear first order differential equation** is

standard form of linear first order equation

$$\frac{dy}{dx} + P(x)y = Q(x), \tag{25}$$

where $P(x)$ and $Q(x)$ are known functions. An **initial value problem (i.v.p)** for a linear first order ODE involves the specification of an initial condition

$$y(x_0) = y_0, \tag{26}$$

where this last condition means that $y = y_0$ when $x = x_0$. Thus, the solution of the initial value problem will evolve away from the point (x_0, y_0) in the (x, y)-plane as x increases from x_0.

To find the general solution of (25) we multiply the equation by a function $\mu(x)$, still to be determined, to obtain

$$\mu\frac{dy}{dx} + \mu P(x)y = \mu Q(x), \tag{27}$$

and seek a choice for μ that allows the left-hand side of (26) to be written as $d(\mu y)/dx$.

With this choice of μ, equation (27) becomes

$$\frac{d(\mu y)}{dx} = \mu Q(x), \tag{28}$$

so integrating with respect to x and dividing by μ shows the general solution of (25) to be

$$y(x) = \frac{C}{\mu(x)} + \frac{1}{\mu(x)}\int \mu(x)Q(x)dx, \tag{29}$$

where C is an arbitrary integration constant. Notice that it is essential to include the arbitrary integration constant *immediately* after the integration $\int \mu(x)Q(x)dx$ has been performed, and *before* dividing by $\mu(x)$; otherwise, the form of the general solution will be incorrect.

integrating factor

To make use of (29) it is necessary to determine the function $\mu(x)$ called the **integrating factor** for the linear first order ODE in (24). By definition

$$\frac{d(\mu y)}{dx} = \mu\frac{dy}{dx} + \mu P(x)y,$$

so after expanding the left-hand side this becomes

$$\mu\frac{dy}{dx} + y\frac{d\mu}{dx} = \mu\frac{dy}{dx} + \mu P(x)y.$$

Cancelling the terms $\mu dy/dx$ and dividing by y gives the following variables separable equation for the integrating factor $\mu(x)$:

$$\frac{d\mu}{dx} = \mu P(x).$$

This has the solution

$$\mu(x) = A\exp\left\{\int P(x)dx\right\},$$

where A is an arbitrary integration constant. As μ multiplies the entire equation (27), the choice of A is immaterial, so for simplicity we will always set $A = 1$ and take the **integrating factor** to be

finding the integrating factor

$$\mu(x) = \exp\left\{\int P(x)dx\right\}. \tag{30}$$

Inserting (30) into (29) shows the **general solution** of (25) to be

$$y(x) = C\exp\left\{-\int P(x)dx\right\} + \exp\left\{-\int P(x)dx\right\}\int \exp\left\{\int P(x)dx\right\}Q(x)dx. \tag{31}$$

If an initial value problem is involved in which the solution of (25) is required subject to the initial condition $y(x_0) = y_0$, the value of the arbitrary constant C in (31) must be chosen accordingly.

The form of the general solution in (31) is mainly of importance for theoretical reasons, because it shows that the general solution is the sum of a **complementary function**

complementary function, particular integral, and general solution

$$y_c(x) = C\exp\left\{-\int P(x)dx\right\} \tag{32}$$

that contains the arbitrary constant belonging to the general solution of (25), and a **particular integral**

$$y_p(x) = \exp\left\{-\int P(x)dx\right\}\int \exp\left\{\int P(x)dx\right\}Q(x)dx \tag{33}$$

that contains no arbitrary constant and is determined by the nonhomogeneous term $Q(x)$.

Substitution of $y_c(x)$ into the homogeneous form of (25) given by

$$\frac{dy}{dx} + P(x)y = 0$$

shows that $y_c(x)$ is its general solution. The general solution of the nonhomogeneous equation (25) is now seen to be the sum of the general solution of the homogeneous form of the equation, and a particular integral determined by the nonhomogeneous term. It will be shown later that this is the pattern of the general solution for all linear nonhomogeneous differential equations, no matter what their order.

Rather than trying to remember the form of general solution given in (31), it is better to obtain the solution by starting from the integrating factor $\mu(x)$ in (30) and integrating result (28), while not forgetting to include the arbitrary constant immediately after the integration before dividing by $\mu(x)$. For convenience, the steps in the determination of the general solution of (25) can be listed as follows.

steps used when solving a linear first order equation

Rule for solving linear first order equations

STEP 1 If the equation is not in standard form and is written

$$a(x)\frac{dy}{dx} + b(x)y = c(x),$$

divide by $a(x)$ to bring it to the standard form

$$\frac{dy}{dx} + P(x)y = Q(x),$$

with $P(x) = b(x)/a(x)$ and $Q(x) = c(x)/a(x)$

STEP 2 Find the integrating factor

$$\mu(x) = \exp\left\{\int P(x)dx\right\}.$$

STEP 3 Rewrite the original differential equation in the form

$$\frac{d(\mu y)}{dx} = \mu Q(x).$$

STEP 4 Integrate the equation in Step 3 to obtain

$$\mu(x)y(x) = \int \mu(x)Q(x)dx + C.$$

STEP 5 Divide the result of Step 4 by $\mu(x)$ to obtain the required general solution of the linear first order differential equation in Step 1.

STEP 6 If an initial condition $y(x_0) = y_0$ is given, the required solution of the i.v.p. is obtained by choosing the arbitrary constant C in the general solution found in Step 5 so that $y = y_0$ when $x = x_0$.

EXAMPLE 5.15 Solve the initial value problem

$$\cos x\frac{dy}{dx} + y = \sin x, \text{ subject to the initial condition } y(0) = 2.$$

Solution We follow the steps in the above rule.

STEP 1 When written in standard form the equation becomes

$$\frac{dy}{dx} + \frac{1}{\cos x}y = \tan x,$$

so $P(x) = 1/\cos x$ and $Q(x) = \tan x$.

STEP 2 The integrating factor

$$\mu(x) = \exp\left\{\int \frac{dx}{\cos x}\right\} = \exp\{\ln|\sec x + \tan x|\}$$
$$= \sec x + \tan x = \frac{1 + \sin x}{\cos x}.$$

STEP 3 The original differential equation can now be written

$$\frac{d}{dx}\left[\left(\frac{1 + \sin x}{\cos x}\right)y(x)\right] = \left(\frac{1 + \sin x}{\cos x}\right)\tan x.$$

STEP 4 Integrating the result of Step 3 gives

$$\left(\frac{1 + \sin x}{\cos x}\right)y(x) = \int \left(\frac{1 + \sin x}{\cos x}\right)\tan x dx + C$$
$$= \int \sec x \tan x dx + \int \tan^2 x dx + C$$
$$= \sec x + \tan x - x + C = \frac{1 + \sin x}{\cos x} - x + C.$$

STEP 5 Dividing the result of Step 4 by the integrating factor $\mu(x) = (1 + \sin x)/\cos x$ shows that the required general solution is

$$y(x) = \frac{C \cos x}{1 + \sin x} + 1 - \frac{x \cos x}{1 + \sin x},$$

for x such that $1 + \sin x \neq 0$.

The *complementary function* is seen to be

$$y_c(x) = \frac{C \cos x}{1 + \sin x},$$

and the *particular integral* is

$$y_p(x) = 1 - \frac{x \cos x}{1 + \sin x}.$$

STEP 6 The initial condition requires that $y = 2$ when $x = 0$, and the general solution is seen to satisfy this condition if $C = 1$, so the solution of the i.v.p. is

$$y(x) = 1 + \frac{(1 - x) \cos x}{1 + \sin x}.$$

■

EXAMPLE 5.16

An R–L circuit contains an inductor and resistor in series, and a current is made to flow through them by applying a voltage across the ends of the circuit. If the inductance varies linearly with time in such a way that $L(t) = L_0(1 + kt)$, find the current $i(t)$ flowing in the circuit when $t > 0$, given that a constant voltage V_0 is applied at time $t = 0$ when $i(t) = 0$.

Solution The voltage change due to a current $i(t)$ flowing through the inductance is $d(L(t)i)/dt$, and from Ohm's law the corresponding voltage change across the resistance R is Ri, so as the sum of the these voltage changes must equal the imposed constant voltage V_0, the differential equation determining the current becomes

$$\frac{d}{dt}(L(t)i) + Ri = V_0 \quad \text{for} \quad t > 0.$$

Substituting for $L(t)$ and rearranging terms we arrive at the following linear first order variable coefficient nonhomogeneous equation for $i(t)$

$$\frac{di}{dt} + \left(\frac{kL_0 + R}{L_0(1+kt)}\right) i = \frac{V_0}{L_0(1+kt)},$$

subject to the initial condition $i(0) = 0$.

In the notation of this section $P(t) = \left(\dfrac{kL_0 + R}{L_0(1+kt)}\right)$ and $Q(t) = \dfrac{V_0}{L_0(1+kt)}$, so the integrating factor in Step 2 becomes

$$\mu(t) = \exp\left\{\int P(t)dt\right\} = (1+kt)^{[kL_0+R]/kL_0}.$$

Using $\mu(t)$ and $Q(t)$ in Step 4 and applying the initial condition $i(0) = 0$ then shows that the current $i(t)$ at a time $t > 0$ is determined by

$$i(t) = \left(\frac{V_0}{kL_0 + R}\right)\left(1 - (1+kt)^{\left(\frac{kL_0+R}{kL_0}\right)}\right).$$

■

Summary

The study of the linear first order differential equation considered in this section is important in its own right, and it also provides the key to understanding the nature of the solution of linear higher order differential equations. It was shown how, after an equation is written in standard form, it can be solved by means of an integrating factor that can be found directly from the coefficient of y in the equation.

EXERCISES 5.7

In Exercises 1 through 10 find the general solution for the linear first order differential equation, and check your result by using computer algebra.

1. $dy/dx + 2y = 1$.
2. $dy/dx + (1/x)y = x$.
3. $(x+1)dy/dx + y = 2x(x+1)$.
4. $x^2\,dy/dx + xy = x^2 \sin x$.
5. $x^2\,dy/dx - 2xy = 1 + x$.
6. $\sin x\,dy/dx - y \cos x = 2 \sin^2 x$.
7. $x\,dy/dx + 2y = x^2$.
8. $(x+3)dy/dx - 2y = x + 3$.
9. $\sin x\,dy/dx - y = 2 \sin x$.
10. $\sin x\,dy/dx + y = \sin x$.

In Exercises 11 through 16 solve the initial value problem for the linear first order differential equation, and check your result by using a computer algebra package.

11. $x\,dy/dx - y = x^2 \cos x$, with $y(\pi/2) = \pi$.
12. $x^2\,dy/dx + 2xy = 2 + x$, with $y(1) = 1$.
13. $x\,dy/dx - 2y = 2 + x$, with $y(1) = 0$.
14. $x\,dy/dx + 2y = 2x^4$, with $y(1) = 1$.
15. $\sin x\,dy/dx + y \cos x = 2 \sin^2 x$, with $y(\pi/2) = 0$.
16. $2\,dy/dx + y = x^2$, with $y(0) = 1$.
17. A 25-liter gas cylinder contains 80% oxygen and 20% helium. If helium is added at a rate of 0.2 liters a second, and the mixture is drawn off at the same rate, how long will it be before the cylinder contains 80% helium?
18. If in Exercise 17 the volume of the gas cylinder is 20 liters and initially it contains 90% oxygen and 10% helium, and the rate of supply of helium is q liters a second, what must be the value of q if the cylinder is to become 80% full of helium in 1 minute?

19. A particle of unit mass moves horizontally in a resisting medium with velocity $v(t)$ at time t with a resistance opposing the motion given by $kv(t)$, with $k > 0$. If the particle is also subject to an additional resisting force kt, write down the differential equation for $v(t)$, and hence find the value of k if the motion starts with $v(0) = v_0$, and at time $t = 1/k$ its velocity is $v(1/k) = \frac{1}{4}v_0$.

5.8 The Bernoulli Equation

The **Bernoulli equation** is a nonlinear first order differential equation with the standard form

standard form of the Bernoulli equation

$$\frac{dy}{dx} + P(x)y = Q(x)y^n, \quad (n \neq 1). \tag{34}$$

The substitution

$$u = y^{1-n} \tag{35}$$

reduces (34) to the linear first order ODE

$$\frac{1}{(1-n)}\frac{du}{dx} + P(x)u = Q(x), \tag{36}$$

and this can be solved by the method described in Section 5.7. Once the general solution $u(x)$ of (36) has been found, the general solution $y(x)$ of (34) follows by returning to the original dependent variable by making the substitution $u = y^{1-n}$.

When using the general solution in (36) it is important to write the Bernoulli equation in standard form before identifying $P(x)$, $Q(x)$, and n. However, if the form of the equation corresponding to (36) is derived directly, starting from the substitution $u = y^{1-n}$, there is no need for the equation to be in standard form.

The Bernoulli equation occurs in various applications of mathematics that involve some form of nonlinearity. It occurs, for example, in solid and fluid mechanics, where it is found to describe an important characteristic of special types of wave that propagate through space as time increases. To appreciate how this ODE enters into these problems, we consider a simple application to solid mechanics involving a long bar made of a composite material or a polymer whose properties are such that the extension caused by a force does not obey Hooke's law, and so is *not* proportional to the force. Materials of this type are said to be **nonlinearly elastic**. If such a bar receives a blow at one end a disturbance will propagate along it at a finite speed, so that at any instant of time there will be a region in the bar through which the disturbance has passed, and a region ahead of the disturbance through which it has still to pass. When the blow is not large, the propagating boundary between these two regions is called a **wavefront** and t the function representing the displacement at position x at any given time t will be continuous along the bar, though its derivative with respect to x will be discontinuous across the wavefront. The propagating jump in the derivative of the displacement with respect to x at the wavefront as a function of time is called an **acceleration wave**, and we will denote it

wavefront and acceleration wave

by $a(t)$. For many nonlinear materials the magnitude $a(t)$ of the acceleration wave obeys a Bernoulli equation of the form

$$\frac{da}{dt} + \mu(t)a = \beta(t)a^2. \tag{37}$$

It was shown by P. J. Chen (*Selected Topics in Wave Propagation*, Noordhoff, Leyden, 1976, p. 29) that $\mu(t)$ depends on the material properties of the medium through which the disturbance propagates and also the geometry involved, which in a one-dimensional case may be plane, cylindrically, or spherically symmetric, but that the function $\beta(t)$ depends only on the material properties of the medium. This same equation governs the behavior of acceleration waves in three space dimensions and time.

Because of the effects of nonlinearity, in many materials it is possible for the acceleration wave to strengthen as it propagates to the point at which the continuity of the displacement function breaks down and what is called a **shock wave** forms. When this occurs, the speed of propagation of disturbances and other physical quantities become discontinuous across the shock wave, and this in turn can lead to the fracture of the material. Once the material properties of such a medium are specified together with the nature of the initial disturbance, the Bernoulli equation in (37) can be used to determine whether or not a shock wave will form and, if it does, the point along the bar where this occurs.

EXAMPLE 5.17

examples of the Bernoulli equation

Solve the Bernoulli equation

$$\frac{da}{dt} + a = ta^2,$$

and find a condition that determines when the solution becomes unbounded.

Solution The equation is in standard form with $P(t) = 1$, $Q(t) = t$, and $n = 2$. Making the substitution $u = 1/a$ corresponding to (35) and substituting into (36) leads to the linear first order equation

$$\frac{du}{dt} - u = -t.$$

Solving this by the method described in Section 5.7 gives

$$u(t) = Ce^t + 1 + t,$$

so transforming back to the variable $a(t)$, we find that

$$a(t) = 1/(Ce^t + 1 + t).$$

The solution $a(t)$ of the Bernoulli equation will become unbounded at $t = t_c$ if t_c is a solution of the equation $C\exp(t_c) + 1 + t_c = 0$. This result shows that an acceleration wave starting at time $t = 0$ will decay instead of evolving into a shock wave if $C > 0$, because then the equation for t_c has no positive solution, whereas a shock wave will always form if $C < 0$.

Had $a(t)$ represented the magnitude of an acceleration wave, the development of an infinite gradient in the displacement corresponding to $a(t_c) = \infty$ would indicate shock formation. ■

EXAMPLE 5.18 Find the general solution of

$$\frac{dy}{dx} - 2y = xy^{1/2}.$$

Solution In terms of the standard form of the Bernoulli equation given in (34), $P(x) = -2$, $Q(x) = x$, and $n = 1/2$. However, rather than substituting into equation (36) to obtain a linear differential equation for $u(x)$, we will derive it directly starting from the substitution $u = y^{1/2}$, and differentiating it to find du/dx in terms of dy/dx. We have

$$\frac{du}{dx} = \frac{1}{2}y^{-1/2}\frac{dy}{dx} = \frac{1}{2u}\frac{dy}{dx}, \quad \text{so} \quad \frac{dy}{dx} = 2u\frac{du}{dx}.$$

Substituting for y and dy/dx in the Bernoulli equation and cancelling a factor $2u$ gives the following linear equation (compare it with (36) after substituting for $P(x)$, $Q(x)$ and n):

$$\frac{du}{dx} - u = \frac{1}{2}x.$$

The method of Section 5.6 shows this equation to have the general solution

$$u(x) = Ce^{x} - (1/2)(1 + x),$$

so as $u = y^{1/2}$, the required general solution of the Bernoulli equation is

$$y(x) = [Ce^{x} - (1/2)(1 + x)]^2.$$ ■

JACOB BERNOULLI (1654–1705)
A Swiss mathematician born in Basel where he was professor of mathematics until his death. He was a member of one of the most distinguished families of mathematicians in all of the history of mathematics. His most important contributions were to the theory of probability and the calculus and theory of elasticity. Other members of the family contributed to many different parts of mathematics including hydrodynamics and the calculus of variations.

Summary

In a sense, the Bernoulli equation, which is a nonlinear first order differential equation, stands on the boundary between linear and nonlinear first order differential equations, so for this and other reasons it is important in applications. It arises in different applications, many of which themselves arise from problems bordering on linear and nonlinear regimes. This section showed how a straightforward change of variable transforms a Bernoulli equation into a linear first order differential equation that can then be solved by the method of Section 5.6.

EXERCISES 5.8

In Exercises 1 through 8 find the general solution of the Bernoulli equation.

1. $dy/dx + 2y = 2xy^{1/2}$.
2. $dy/dx + y = 3y^2$.
3. $dy/dx - y = 2xy^{3/2}$.
4. $x\,dy/dx + y = xy^2$.
5. $dy/dx + 2y \sin x = 2y^2 \sin x$.
6. $x\,dy/dx + y = 2xy^{1/2}$.
7. $x\,dy/dx - 2y = xy^{3/2}$.
8. $dy/dx + 4xy = xy^3$.

9. A model for the variation of a finite amount of stock $n(t)$ in a warehouse as a function of the time t caused by the supply of fresh stock and its removal by demand is

$$\frac{dn}{dt} = (a - bn)n \quad \text{with the constants } a, b > 0,$$

where $n(0) = n_0$. Find $n(t)$ and discuss the nature of the change in the stock level as a function of time according as n_0 is less than a/b, equal to a/b, or greater than a/b.

10.* This exercise concerns water in a canal of variable depth with the x-axis taken along the canal in the equilibrium surface of the water, and the y-axis vertically downwards. Let the equilibrium depth of water in a channel be $h(x)$, and the cross-sectional area of water in the canal be a slowly varying function $W(x)$. When a water wave advances along the channel into water at rest there will be a change of acceleration across the advancing line (wavefront) that separates the disturbed water from the undisturbed water. Such an advancing disturbance is called an **acceleration wave**. If the change in acceleration across the wavefront at point x along the channel is $a(x)$, it can be shown that the strength $a(x)$ of the acceleration wave obeys the Bernoulli equation

$$\frac{da}{dx} + \left(\frac{3h'}{4h} + \frac{W'}{2W}\right)a + \frac{3a^2}{2h} = 0.$$

If the initial condition for $a(x)$ is $a(0) = a_0$, then a wave of **elevation** wave is one for which $a_0 < 0$, and a wave of **depression** is one for which $a_0 > 0$. In this approximation the wave will **break**, due to the water surface becoming vertical at the wavefront if, after propagating a critical distance x_c along the channel, the strength of the acceleration $a(x_c) = \infty$.

(i) Find $a(x)$ in terms of $a_0 = a(0)$, $h_0 = h(0)$ and $W_0 = W(0)$.

(ii) Discuss the breaking and non-breaking of waves of elevation and depression.

(iii) If the water shelves to zero at $x = l$, so that $h(l) = 0$, find a condition that ensures the wave breaks before $x = l$.

5.9 The Riccati Equation

The **Riccati equation** is an important nonlinear equation with the standard form

standard form of the Riccati equation

$$\frac{dy}{dx} + P(x)y + R(x)y^2 = Q(x). \tag{38}$$

Its significance derives from the fact that it stands at the boundary between linear and nonlinear equations, and it occurs in various applications of mathematics that involve nonlinear problems. The Riccati equation reduces to a linear first order equation when $R(x) \equiv 0$, and to a Bernoulli equation when $Q(x) \equiv 0$.

Obtaining the general solution of a Riccati equation is difficult, but the task is simplified if a particular solution is known, or can be found by inspection. If a particular solution is $y_1(x)$ is known, then

substitutions that simplify the Riccati equation

(i) The substitution $y = y_1 + 1/u$ reduces the equation to a linear first order equation.

(ii) The substitution $y = y_1 + u$ reduces the equation to a Bernoulli equation.

(iii) The general substitution

$$y = \frac{1}{R(x)z}\frac{dz}{dx}$$

reduces the Riccati equation to the linear homogeneous second order ODE

$$\frac{d^2z}{dx^2} + \left\{P(x) - \frac{R'(x)}{R(x)}\right\}\frac{dz}{dx} - R(x)Q(x)z = 0$$

discussed in Chapters 6 and 8.

Substitution (i) is often the most convenient one to use, as will be seen from the next example.

EXAMPLE 5.19 Find the general solution of the Riccati equation

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

Solution Inspection shows that $y_1(x) = x$ is a particular solution, so we make the substitution $y = x + 1/u$, from which it follows that

$$\frac{dy}{dx} = 1 - \frac{1}{u^2}\frac{du}{dx},$$

and after substitution for y and dy/dx in the Riccati equation it reduces to the linear ODE

$$\frac{du}{dx} + x^2u = -x.$$

Solving this by the method of Section 5.6 gives

$$u(x) = C\exp(-x^3/3) - \exp(-x^3/3)\int x\exp(x^3/3)dx,$$

where the integral in the last term cannot be expressed in terms of elementary functions. Transforming back to the variable $y(x)$ shows the general solution of the Riccati equation to be

$$y(x) = x + \frac{\exp(x^3/3)}{C - \int x\exp(x^3/3)dx}.$$

It is not unusual for solutions of ODEs to give rise to functions such as $\int x\exp(x^3/3)dx$ that have no representation in terms of known functions, because not all functions have antiderivatives that are expressible in terms of elementary functions. ■

Jacopo Francesco (Count) Riccati (1676–1754)
An Italian mathematician whose main contributions to mathematics were in the field of differential equations, though he also contributed to geometry and the study of acoustics.

Additional information relevant to the material in Sections 5.4 to 5.9 is to be found in the appropriate chapters of any one of references [3.3] to [3.5], [3.15], [3.16], and [3.19]. A sophisticated and extremely enlightening discussion of ordinary differential equations is to be found in reference [3.1] that considers not only first order equations, but also higher order equations and systems.

Summary This section introduced the Riccati equation, of which the Bernoulli equation is a special case. Solving the Riccati equation is difficult, but some substitutions were given that simplify this task when one solution of the Riccati equation is already known, possibly by inspection.

EXERCISES 5.9

1. Show that the substitution $y = y_1 + 1/u$ reduces the Riccati equation in (38) to a linear first order equation.

2. Show that the substitution $y = y_1 + u$ reduces the Riccati equation in (38) to a Bernoulli equation.

In Exercises 3 through 6 verify that $y_1(x)$ is a solution of the Riccati equation and use it to find the general solution of the equation.

3. $dy/dx + 2x^2y - 2xy^2 = 1$, with $y_1(x) = x$.

4. $dy/dx + 2y^2 - y = 1$, with $y_1(x) \equiv 1$.

5. $dy/dx - 2y^2 + 3y = 1$, with $y_1(x) \equiv 1$.

6. $dy/dx - 3x^2y + 3xy^2 = 1$, with $y_1(x) = x$.

7. Verify that the substitution

$$y = \frac{1}{R(x)z}\frac{dz}{dx}$$

reduces the Riccati equation (38) to the linear homogeneous second order ODE

$$\frac{d^2z}{dx^2} + \left\{P(x) - \frac{R'(x)}{R(x)}\right\}\frac{dz}{dx} - R(x)Q(x)z = 0.$$

5.10 Existence and Uniqueness of Solutions

existence and uniqueness

The questions of whether a solution to an initial value problem for a first order differential equation can be found and, when a solution does exist, whether it is the only solution are of fundamental importance in the theory of differential equations, and also in their applications. Establishing that a solution to an initial value problem can be found is called the **existence** problem, while ensuring that when a solution exists it is the only one is called the **uniqueness** problem. To show that the questions of existence and uniqueness arise even with very simple initial value problems we examine the following two examples.

Let us consider the initial value problem

$$\frac{dy}{dx} = \frac{4}{3}y^{1/4}, \quad \text{with } y(0) = -1,$$

involving a variables separable equation. Integration shows the general solution to be

$$y^3 = (x + C)^4,$$

from which it can be seen that y is essentially nonnegative. Clearly there can be no solution to this equation such that $y = -1$ when $x = 0$, so this is an example of an initial value problem that has *no* solution. Had the initial condition been $y(0) = 1$ the unique solution would have been

$$y^3 = (x + 1)^4.$$

In fact this equation has a solution for any initial condition in which $y(x)$ is *positive*, but no solution when it is *negative*. This is hardly surprising, because had we examined the function $y^{1/4}$ carefully before proceeding with the integration we would have seen that it is a complex number whenever y is negative. Sometimes, as here, an inspection of the initial condition and the equation can show in advance whether or not the condition is appropriate, but more frequently constraints on an initial condition that allow a solution to the differential equation to exist only emerge when the form of the solution is known.

To illustrate nonuniqueness, we need only consider the differential equation

$$\frac{dy}{dx} = 3y^{2/3}, \text{ subject to the initial condition } y(0) = 0.$$

The equation is variables separable, and integration shows it has the solution $y = x^3$, but this is not the only solution because it also has the singular, though somewhat uninteresting, solution $y = 0$.

However, these are not the only two solutions, because for any $a > 0$ the function

$$y(x) = \begin{cases} 0, & x < a \\ (x-a)^3, & x \geq a \end{cases}$$

is continuous, has a continuous first derivative, and satisfies both the differential equation and the initial condition, showing that it also is a solution. As $a > 0$ is arbitrary, we see that $y(x)$ is a one-parameter family of solutions, so clearly this initial value problem does not have a unique solution.

The following theorem on existence and uniqueness is stated without proof (see, for example, references [3.1],[3.3],[3.4],[3.10] and [3.12]). It is important to appreciate that though the conditions in the theorem are *sufficient* to ensure existence and uniqueness, they are not *necessary* conditions, as examples can be constructed that fail to satisfy the conditions of the theorem, but nevertheless have a unique solution.

THEOREM 5.2

conditions that definitely ensure existence and uniqueness

Existence and uniqueness of solutions Let $f(x, y)$ be a continuous and bounded function of x and y in a rectangular region R of the (x, y)-plane that contains a given point (x_0, y_0). Then for some suitably small positive number h the initial value problem

$$\frac{dy}{dx} = f(x, y), \quad \text{with} \quad y(x_0) = y_0$$

has at least one solution within the open interval $x_0 - h < x < x_0 + h$. If, in addition, $\partial f/\partial y$ is continuous and bounded in R, the solution is unique in an open interval centered on x_0 that may lie within the interval $x_0 - h < x < x_0 + h$. ■

Let us apply this theorem to the initial value problem

$$\frac{dy}{dx} = 3y^{2/3}, \quad \text{with } y(0) = 0,$$

that we have just shown does not have a unique solution. The function $f(x, y) = 3y^{2/3}$ is continuous in any neighborhood of the origin where the initial condition is given, but $\partial f/\partial y = 2y^{-1/3}$ is unbounded at the origin. So the first condition of Theorem 5.2 is satisfied but the second is not, showing that although this initial value problem has a solution, it is not unique.

Summary

This section described what is meant by the existence of a solution of a differential equation, and the uniqueness of a solution that is usually expected in applications to physical problems. A theorem, stated without proof, was given that guarantees both the existence and uniqueness of a solution. However, the conditions of the theorem are more restrictive than necessary, so equations can be found that while not satisfying the conditions of the theorem nevertheless have a solution, and it is unique.

EXERCISES 5.10

In Exercises 1 through 6, find any points at which the imposition of initial conditions will not lead to a unique solution.

1. $dy/dx = (1 - x)^{1/2}$.

2. $dy/dx = xy + 1$.

3. $dy/dx = x^2 + y^2$.

4. $dy/dx = (x^2 + y^2 - 1)^{-1/2}$.

5. $dy/dx = -y/x$.

6. $dy/dx = x \ln|1 - y^2|$.