# Chapter 10
# Differential Equations

**Abstract** The laws of the exact sciences are often formulated as differential equations, that is, equations connecting functions and their derivatives. In this chapter, we present examples from three different fields: mechanical vibrations, population growth, and chemical reactions.

## 10.1 Using Calculus to Model Vibrations

Most people realize that *sound*—its generation, transmission, and perception—is a vibration. For this reason alone, vibration is a very important subject. But vibrations are more general and pervasive than mere sound, and they constitute one of the fundamental phenomena of physics. The reason is the mechanical stability of everyday objects from bells and horns to the basic constituents of matter. Mechanical stability means that when an object is distorted by an outside force, it springs back into its original shape when released. This is accomplished by restoring forces inherent in any object. Restoring forces work in a peculiar way: they not only bring the object back to its original shape, but they tend to overcorrect and distort it in the opposite direction. This is again overcorrected, and so ad infinitum, leading to a vibration around an equilibrium. In this section, we shall explain this process in simple situations, as an example of the application of calculus.

### *10.1a Vibrations of a Mechanical System*

The fundamental concepts of one-dimensional mechanics are particle, mass, position, velocity, acceleration, and force.

The *position* of a particle along a line is specified by a single real number $x$. Since the position of the particle changes in time, it is a function of the time $t$. The derivative of position with respect to $t$ is the velocity of the particle, denoted usually

P.D. Lax and M.S. Terrell, *Calculus With Applications*, Undergraduate Texts in Mathematics, 375
DOI 10.1007/978-1-4614-7946-8_10, 

by $v(t)$. The derivative of velocity with respect to time is called *acceleration*, and is denoted by $a(t)$:

$$x' = v, \quad v' = x'' = a.$$

The mass of the particle, denoted by $m$, does not change throughout the motion. Newton's law of motion says that

$$f = ma,$$

where $f$ is the total force acting on the particle, $m$ the mass, and $a$ the acceleration. To put teeth into Newton's law, we have to be able to calculate the total force acting on the particle. According to Newton, the total force acting on a particle (in the direction of increasing $x$) is the sum of all the various forces acting on it. In this section, we shall deal with two kinds of forces: *restoring forces* and *frictional forces*. We shall describe them in the following specific context.

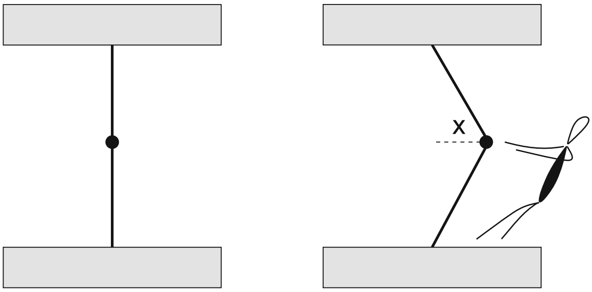


**Fig. 10.1** *Left*: position of static equilibrium. *Right*: mass displaced by distance $x$

Imagine a piece of elastic string (rubber band, elastic wire) placed in a vertical position with its endpoints fastened and a mass attached to its middle. In this position the mass is at rest. Now displace the mass to one side (Fig. 10.1). In this position, the elastic string exerts a force on the mass. It is clear, to anyone who ever shot paper clips with a rubber band, that the restoring force depends on position, and

(a) *the force acts in the direction opposite to the displacement, tending to restore the mass to its previous position;*

(b) *the greater the magnitude of the displacement, the greater the magnitude of the force.*

A force with these two properties is called a *restoring force*, written $f_{\text{re}}$. The graph of a typical restoring force is shown in Fig. 10.2. Many restoring forces, such as the one exerted by a rubber band, have yet a third property, symmetry:

(c) *Displacements by the same magnitude but in opposite directions generate restoring forces $f_{\text{re}}$ that are equal in magnitude but opposite in direction.*

We turn next to describing the force of friction. Friction can be caused by various mechanisms, one of which is air resistance. As anyone who has ever bicycled at high speed knows, the frictional force depends on velocity, and

(a) *the force of air resistance acts in the direction opposite to the direction of motion*;
(b) *the greater the velocity, the greater the force of resistance*.

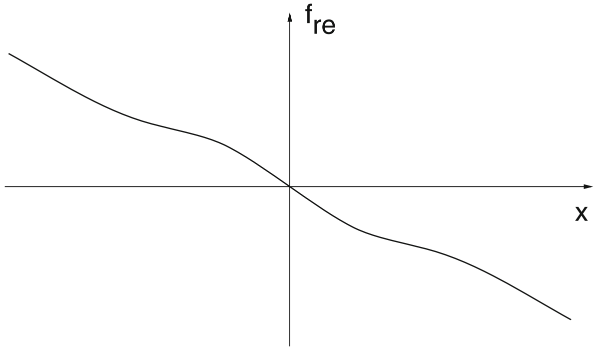


**Fig. 10.2** A restoring force $f_{\mathrm{re}}$, as a function of displacement

Any force with these two properties is called a *frictional force*, written $f_{\mathrm{fr}}$. The graph of a typical frictional force is shown in Fig. 10.3. This graph displays yet another property common to most frictional forces, their symmetry:

(c) *The magnitude of the frictional force $f_{\mathrm{fr}}$ depends only on the magnitude of the velocity*.

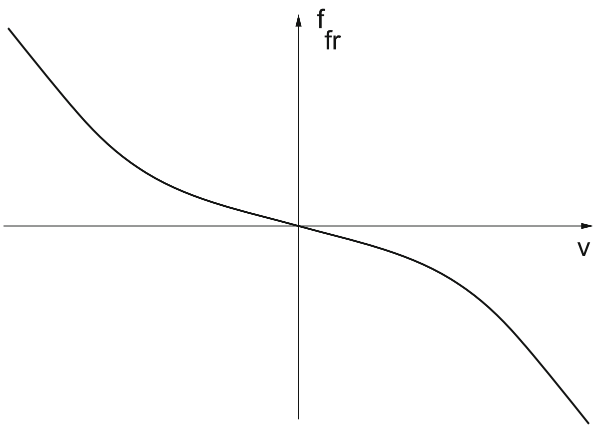


**Fig. 10.3** A frictional force $f_{\mathrm{fr}}$, as a function of velocity

In order to turn the verbal descriptions of these two kinds of forces into mathematical descriptions, we regard the restoring force $f_{\text{re}}$ as a function of the position $x$. The properties (a)–(c) can be expressed as follows in the language of functions:

$$f_{\text{re}}(x) = \begin{cases} < 0 & \text{for } x > 0, \\ = 0 & \text{for } x = 0, \\ > 0 & \text{for } x < 0, \end{cases}$$

and

$$f_{\text{re}}(x) \text{ is a decreasing function of } x.$$

The symmetry property of $f_{\text{re}}(x)$ can be expressed in the following way: $f_{\text{re}}$ is an *odd* function of $x$, i.e.,

$$f_{\text{re}}(-x) = -f_{\text{re}}(x).$$

We regard the frictional force $f_{\text{fr}}$ as a function of velocity. The properties of a frictional force can be expressed as follows:

$$f_{\text{fr}}(v) = \begin{cases} < 0 & \text{for } v > 0, \\ = 0 & \text{for } v = 0, \\ > 0 & \text{for } v < 0, \end{cases}$$

and

$$f_{\text{fr}}(v) \text{ is a decreasing function of } v.$$

Assumption (c), that the magnitude of friction depends only on the magnitude of velocity, together with assumption (a) implies that $f_{\text{fr}}$ is an odd function, i.e.,

$$f_{\text{fr}}(-v) = -f_{\text{fr}}(v).$$

The total force $f$ is the sum of the individual forces:

$$f = f_{\text{fr}} + f_{\text{re}}.$$

The additivity of forces is an experimental fact. With this decomposition of force, Newton's law may be written as

$$ma = f_{\text{fr}}(v) + f_{\text{re}}(x). \tag{10.1}$$

Since velocity $v$ and acceleration $a$ are the first and second derivatives of $x$, this is the differential equation

$$mx'' - f_{\text{fr}}(x') - f_{\text{re}}(x) = 0 \tag{10.2}$$

for $x$ as function of $t$. Solutions of this differential equation describe all possible motions of a particle subject to a restoring force and a frictional force.

In the rest of Sect. 10.1, we shall study the behavior of solutions of Eq. (10.2) for various kinds of restoring force and frictional force. The basic fact is that if we prescribe the initial position and velocity of a particle, then the motion of the particle is completely determined for all time by the differential equation. We call this basic result the *uniqueness theorem*:

**Theorem 10.1. Uniqueness.** *Denote by $x(t)$ and $y(t)$ two solutions of the differential equation (10.2) that are equal at some time $s$ and whose first derivatives are equal at time $s$:*

$$x(s) = y(s), \qquad x'(s) = y'(s).$$

*Then $x(t)$ and $y(t)$ are equal for all $t \geq s$.*

The steps of the proof of this theorem are outlined in Problem 10.21 at the end of Sect. 10.1, where we ask you to justify each step.

*Example 10.1.* We have encountered the following differential equations in previous chapters:

$$\text{(a) } x' = x, \quad \text{(b) } x'' - x = 0\,, \quad \text{(c) } x'' + x = 0.$$

Which of these are examples of Eq. (10.2) governing vibrations of a simple mechanical system?

(a) The equation $x' = x$ can be rewritten $0 + x' - x = 0$, but it is not an example of Eq. (10.2), because there is no second-order term $mx''$.
(b) $x'' - x = 0$ looks promising. Take mass $m = 1$, friction $f_{\text{fr}}(x') = 0$, and restoring force $f_{\text{re}}(x) = x$. This seems to fit the form. However, in our model we made the assumption that $f_{\text{re}}$ is decreasing and odd. Since this $f_{\text{re}}(x)$ is not decreasing, $x'' - x = 0$ is not an example of Eq. (10.2).
(c) $x'' + x = 0$ looks very much like case (b), except that now the restoring force $f_{\text{re}}(x) = -x$ is decreasing and odd. Therefore, the equation $x'' + x = 0$ is an example of the equation $mx'' - f_{\text{fr}}(x') - f_{\text{re}}(x) = 0$ governing vibrations of a simple mechanical system.

## *10.1b Dissipation and Conservation of Energy*

This section will be devoted to the mathematics of extracting information about solutions of the differential equation $mx'' - f_{\text{fr}}(x') - f_{\text{re}}(x) = 0$. It is remarkable how much we can deduce about the solutions without knowing the frictional force $f_{\text{fr}}(v)$ or the restoring force $f_{\text{re}}(x)$ explicitly, but knowing only that they both are decreasing odd functions.

We start with a trick. Multiply the equation by $v$, obtaining

$$mva - vf_{\text{fr}}(v) - vf_{\text{re}}(x) = 0.$$

Since the sign of $f_{\rm fr}(v)$ is opposite to that of $v$, it follows that $-vf_{\rm fr}(v)$ is positive except when $v=0$. By dropping this positive term, we convert the equality into the inequality

$$mva - vf_{\rm re}(x) \le 0. \tag{10.3}$$

Recalling that acceleration is the derivative of velocity, we can rewrite the term $mva$ as $mvv'$. We recognize this as the derivative of $\frac{1}{2}mv^2$:

$$mva = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}mv^2\right). \tag{10.4}$$

Recalling that $v$ is the derivative of $x$, we can rewrite the term $vf_{\rm re}(x)$ as $x'f_{\rm re}(x)$.

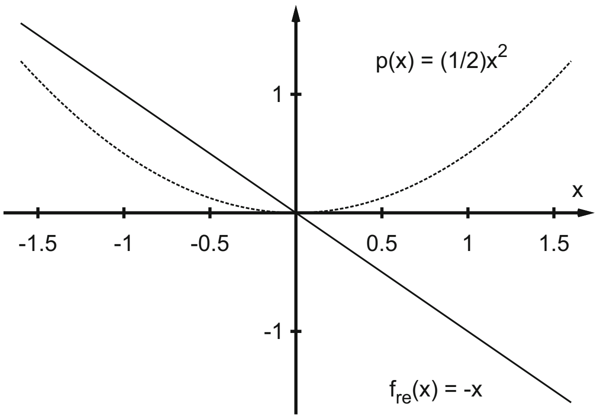


**Fig. 10.4** Graphs of the restoring force $f_{\rm re}(x) = -x$ and the potential energy $p(x) = \frac{1}{2}x^2$ for the equation $x'' + x = 0$ in Example 10.1

Let us introduce the function $p(x)$ as the integral of $-f_{\rm re}$,

$$p(x) = -\int_0^x f_{\rm re}(y)\,\mathrm{d}y.$$

By the fundamental theorem of calculus, the derivative of $p$ is $-f_{\rm re}$,

$$\frac{\mathrm{d}}{\mathrm{d}x}p(x) = -f_{\rm re}(x), \tag{10.5}$$

and by definition,

$$p(0) = 0.$$

The derivative of $p(x)$ with respect to $x$ is $-f_{\rm re}(x)$, which is positive for $x$ positive, and negative for $x$ negative. (See Figs. 10.2 and 10.4.) According to the monotonicity criterion, this means that $p$ is increasing for $x > 0$ and decreasing for $x < 0$. Since $p(0) = 0$, it follows that $p(x)$ is positive for all $x \neq 0$. Using the chain rule and Eq. (10.5), we can express the derivative of $p(x(t))$ with respect to $t$ as

$$\frac{\mathrm{d}}{\mathrm{d}t}p(x(t)) = x'(t)\frac{\mathrm{d}p}{\mathrm{d}x} = -x'(t)f_{\mathrm{re}}(x(t)) = -vf_{\mathrm{re}}(x). \tag{10.6}$$

Substituting this result and expression (10.4) for the first and second terms into inequality (10.3) we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}mv^2 + p(x)\right) \le 0. \tag{10.7}$$

According to the monotonicity criterion, a function whose derivative is less than or equal to zero is decreasing. We use "decreasing" to mean "nonincreasing." So we conclude that the function

$$\frac{1}{2}mv^2 + p(x)$$

decreases with time. This function, and both terms appearing in it, have physical meaning: the quantity $\frac{1}{2}mv^2$ is called *kinetic energy*, and the quantity $p(x)$ is called *potential energy*. The sum of kinetic and potential energies is called the *total energy*. In this terminology, we have derived the following.

**Law of Decrease of Energy.** *The total energy of a particle moving under the influence of a restoring force and a frictional force decreases with time.*

Suppose there is no frictional force, i.e., $f_{\mathrm{fr}}$ is zero. Then the energy inequality (10.7) becomes an equality: $\dfrac{\mathrm{d}}{\mathrm{d}t}\left(\dfrac{1}{2}mv^2 + p(x)\right) = 0$. A function whose derivative is zero for all $t$ is a constant, $E$, so we have derived the following.

**Law of Conservation of Energy.** *In the absence of friction, the total energy of a particle moving under the influence of a restoring force does not change with time.*

$$\frac{1}{2}mv^2 + p(x) = E.$$

From Eq. (10.2) governing vibrations of a simple mechanical system, we have derived energy laws for the particle in the presence, and in the absence, of friction. When there is no friction, the total mechanical energy does not change with time. When there is friction, the total *mechanical* energy decreases. That energy is not lost but is turned into heat energy.

### *10.1c Vibration Without Friction*

Next, we turn our attention to the study of the motion of a particle subject to a restoring force in the absence of friction. That is, the forces satisfy

$$mx'' - f_{\mathrm{re}}(x) = 0. \tag{10.8}$$

In Example 10.1, we saw that the differential equation $x'' - (-x) = x'' + x = 0$ is an example of Eq. (10.8), and we showed in Sect. 3.4b that all solutions of $x'' + x = 0$ are of the form $x(t) = u\cos t + v\sin t$, where $u$ and $v$ are arbitrary constants. These functions have period $2\pi$:

$$x(t+2\pi) = x(t).$$

Now we show that *every* function $x(t)$ satisfying Eq. (10.8) is periodic. We start with a qualitative description of the motion determined by Eq. (10.8) and the law of conservation of energy that we derived from it in the last section:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}mv^2 + p(x)\right) = 0, \quad p(x) = -\int_0^x f_{\mathrm{re}}(y)\,\mathrm{d}y, \quad \frac{1}{2}mv^2 + p(x) = E.$$

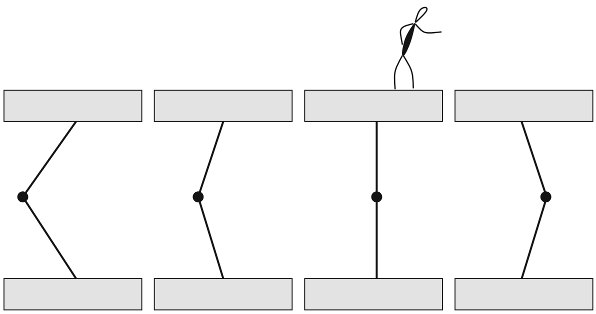

**Fig. 10.5** At *left*, $x = -b$, $v = 0$, and $p = E$. Then $x$ is shown between $-b$ and 0, with $v$ positive and $p < E$. Then $x = 0$, $p = 0$, and $\frac{1}{2}mv^2 = E$. At *right*, the particle has almost reached the point $x = b$ halfway through the cycle

Suppose we start the motion at time $t = 0$ by displacing the particle to the position $x = -b$, $b > 0$, and holding it there until we let it go, so that initially its velocity is zero. See Fig. 10.5. The total energy imparted thereby to the system is $p(-b) = E$. On being released, the restoring force starts moving the particle toward the position $x = 0$. For negative $x$, $p(x)$ decreases with $x$. It follows then from the law of conservation of energy that the kinetic energy, $\frac{1}{2}mv^2$, increases. Since $v^2$ is increasing, the particle gains speed during this phase of the motion. The potential energy reaches its minimum at $x = 0$. As soon as the particle swings past $x = 0$, its potential energy starts increasing, and its kinetic energy decreases accordingly. This state of affairs persists until the particle reaches the position $x = b$. At this point, its potential energy equals $p(b)$. Since $p$ is the integral of an odd function, $p$ is an even function, and $p(b) = p(-b) = E$ is the total energy. Therefore, at point $b$, the kinetic energy $\frac{1}{2}mv^2$ is zero, and $b$ is the right endpoint of the interval through which the particle moves. On reaching $x = b$, the particle is turned around by the restoring force and describes a similar motion from right to left until it returns to its original position $x = -b$. Its velocity at this time $t = T$ is zero, so everything is just as it was at the beginning of the motion. Therefore, according to Theorem 10.1 in the previous section, the *same* pattern is repeated all over again. Such motion is called *periodic*, and the time $T$ taken by the particle to return to its original position is called the *period* of the motion. The mathematical expression of periodicity is

$$x(t+T) = x(t),$$

and the graph of such a period-$T$ function is shown in Fig. 10.6. In fact, due to the assumption that $f_{\mathrm{re}}$ is odd, the position $x = b$ occurs at exactly $t = \frac{1}{2}T$, for the motions to left and right are mirror images of each other.

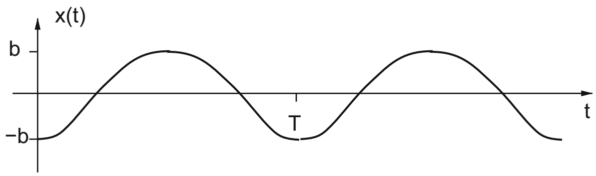


**Fig. 10.6** The motion repeats with period $T$

We now turn from this qualitative description of motion to a quantitative description, which we also shall deduce from the law of conservation of energy. Using the energy equation $\frac{1}{2}mv^2 + p(x) = E$, we can express $v$ as function of $x$:

$$v = \sqrt{\frac{2}{m}(E - p(x))}.$$

In the first phase of the motion, $0 \le t \le \frac{1}{2}T$, $x$ is an increasing function of time. Therefore, $v = x'$ is positive, so that the positive square root is to be taken. Since $x(t)$ is strictly monotonic during this interval, we can express $t$ as a function of $x$. According to the rule for differentiating the inverse of a function, the derivative of $t$ with respect to $x$ is

$$\frac{\mathrm{d}t}{\mathrm{d}x} = \frac{1}{\frac{\mathrm{d}x}{\mathrm{d}t}} = \frac{1}{v}.$$

Using the formula above for $v$, we deduce that

$$\frac{\mathrm{d}t}{\mathrm{d}x} = \sqrt{\frac{m}{2(E - p(x))}}.$$

According to the fundamental theorem of calculus, $t$ is the integral with respect to $x$ of $\frac{\mathrm{d}t}{\mathrm{d}x}$:

$$t(y_2) - t(y_1) = \int_{y_1}^{y_2} \sqrt{\frac{m}{2(E - p(x))}}\, \mathrm{d}x. \tag{10.9}$$

The integral on the right expresses the time it takes for the particle to move from position $y_1$ to position $y_2$ during the first phase of the motion. Take, in particular, $y_1 = -b$ and $y_2 = b$. These positions are reached at $t = 0$ and $t = \frac{1}{2}T$, respectively. Therefore, $\frac{1}{2}T - 0 = \int_{-b}^{b} \sqrt{\frac{m}{2(E-p(x))}}\,\mathrm{d}x$, and multiplying by 2, we get

$$T = \int_{-b}^{b} \sqrt{\frac{2m}{E-p(x)}}\,\mathrm{d}x. \tag{10.10}$$

We have seen that the energy conservation $\frac{1}{2}mv^2 + p(x) = E$ at times $t = 0$ and $t = \frac{1}{2}T$ gives $E = p(-b) = p(b)$. This shows that as $x$ approaches $-b$ or $b$, the difference $E - p(x)$ tends to zero. This makes the integrand tend to infinity as $x$ approaches the endpoints. In the terminology of Sect. 7.3, this integral is improper, and therefore is defined by evaluating the integral over a subinterval and taking the limit as the subinterval approaches the original interval.

We show now that the improper integral (10.10) for the period $T$ converges. According to the mean value theorem, the function in the denominator of Eq. (10.10) is

$$E - p(x) = E - p(b) - p'(c)(x-b)$$

for some $c$ between $b$ and $x$. Since $E = p(b)$, this gives

$$E - p(x) = -f_{\mathrm{re}}(c)(b-x).$$

For $x$ slightly less than $b$, that is, near the upper limit in integral (10.10), this is a positive multiple of $(b-x)$, because $-f_{\mathrm{re}}(c)$ is nearly $-f_{\mathrm{re}}(b) > 0$. Therefore,

$$\sqrt{\frac{2m}{E-p(x)}} \le \frac{\text{const}}{\sqrt{b-x}}.$$

The integrand is similarly bounded near $-b$, the lower bound of integration. As we have seen in Example 7.27 of Sect. 7.3, such a function is integrable. In other words, the period $T$ is well defined by the integral (10.10).

We have been able to deduce quite a bit about the function $x(t)$ from the fact that it satisfies $mx'' - f_{\mathrm{re}}(x) = 0$. First, we showed that $x(t)$ is periodic. Second, we showed that the period is a number $T = \int_{-b}^{b} \sqrt{\frac{2m}{E-p(x)}}\,\mathrm{d}x$ that depends on the initial displacement $b$ of the particle and on the restoring force $f_{\mathrm{re}}$, since $p(x) = -\int_0^x f_{\mathrm{re}}(y)\,\mathrm{d}y$. Next, we look at the specific cases in which the restoring forces are linear functions.

## *10.1d Linear Vibrations Without Friction*

Suppose that the restoring force is a *differentiable* function of $x$. According to the basic tenet of differential calculus, *a differentiable function can be well approximated over a short interval by a linear function*. We have seen earlier that the motion is confined to the interval $-b \le x \le b$, where $-b$ is the initial displacement. For small $b$, $f_{\text{re}}(x)$ can be well approximated over $[-b, b]$ by a linear function (Fig. 10.7). It is reasonable to expect that if we replace the true restoring force by its linear approximation over the small interval $[-b, b]$, the characteristic properties of motions with small displacements will not change drastically.

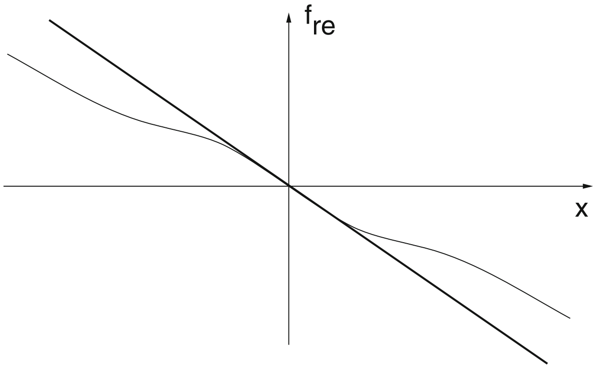


**Fig. 10.7** Linearized restoring force

In this section, we study vibration under a *linear* restoring force

$$f_{\text{re}}(x) = -kx.$$

The positive constant $k$ measures the stiffness of the elastic medium exerting the force, i.e., the larger $k$ is, the greater the resistance to the displacement. For this reason, $k$ is called the *stiffness* constant. The corresponding potential

$$p(x) = -\int_0^x f_{\text{re}}(s)\,\mathrm{d}s = -\int_0^x -ks\,\mathrm{d}s = \frac{1}{2}kx^2$$

is quadratic. Let us substitute it into formula (10.10) for the period of the motion. Using the fact that $E = p(b)$, we get

$$T = \int_{-b}^{b} \sqrt{\frac{2m}{E - p(x)}}\,\mathrm{d}x = \int_{-b}^{b} \sqrt{\frac{4m}{kb^2 - kx^2}}\,\mathrm{d}x.$$

Performing the change of variable $x = by$, we get

$$T = \int_{-b}^{b} \sqrt{\frac{4m}{kb^2 - kx^2}}\,\mathrm{d}x = \int_{-1}^{1} \sqrt{\frac{4m}{kb^2 - k(by)^2}}\,b\,\mathrm{d}y = 2\sqrt{\frac{m}{k}} \int_{-1}^{1} \frac{\mathrm{d}y}{\sqrt{1 - y^2}}.$$

We recall that the function $\frac{1}{\sqrt{1-y^2}}$ is the derivative of $\sin^{-1} y$, so

$$\int_{-1}^{1} \frac{1}{\sqrt{1-y^2}}\, \mathrm{d}y = \sin^{-1} 1 - \sin^{-1}(-1) = \frac{\pi}{2} - (-\frac{\pi}{2}) = \pi.$$

Substituting this into the above formula for $T$, we obtain

$$T = 2\pi\sqrt{\frac{m}{k}}. \tag{10.11}$$

This remarkable formula shows how the period of the motion depends on the data:

(a) The period is independent of the size of the initial displacement, provided that the initial displacement is small enough to warrant approximating $f_{\text{re}}$ by a linear function.
(b) The period is proportional to $\sqrt{\frac{m}{k}}$.

What does our physical intuition tell us? Increasing the mass $m$ slows down the motion, and tightening the elastic string, which is the same as increasing the stiffness constant $k$, speeds up the motion. Therefore, the period is an increasing function of $m$ and a decreasing function of $k$; this is evident from formula (10.11).

We show now how to derive formula (10.11) using dimensional analysis. In a linear restoring force $f_{\text{re}}(x) = -kx$, the dimension of the number $k$ is force per length, which is equal to

$$\frac{(\text{mass})(\text{acceleration})}{\text{length}} = \frac{(\text{mass})\frac{\text{length}}{(\text{time})^2}}{\text{length}} = \frac{\text{mass}}{(\text{time})^2}.$$

The only way to build a number whose dimension is time out of the two numbers $m$ and $k$ is $\sqrt{\frac{m}{k}}$. Therefore, the period $T$ must be a constant multiple of $\sqrt{\frac{m}{k}}$. Calculus is needed only to nail down that constant as $2\pi$.

A periodic motion is often called a vibration. Any portion of such motion lasting a full period is called a *cycle*. The number of cycles per unit time is called *frequency*, i.e.,

$$\text{frequency} = \frac{1}{\text{period}} = \frac{1}{2\pi}\sqrt{\frac{k}{m}}.$$

The most striking manifestation of vibration is caused by the pressure waves transmitted through the air to the ears of a nearby auditor who perceives them as sound. The pitch of the sound is determined by the number of pressure pulses per unit time reaching the eardrum, and this number is the frequency of the vibrating source of the sound. When struck by a hammer, piece of metal vibrates. We know from everyday observation that the pitch of the sound generated does *not* depend on how hard the metal has been struck, although the loudness of the sound does. On the other hand, the sound generated by a plucked rubber band has a twangy quality, indicating that the pitch changes as the displacement changes. We conclude that the elastic force that acts in metal when slightly displaced from equilibrium is a linear function of displacement, while the force exerted by a rubber band is a nonlinear function of displacement.

### *10.1e Linear Vibrations with Friction*

We now turn to the study of motion with friction. We shall restrict our study to motions for which displacement $x$ and velocity $v$ are relatively small, so small that both $f_{\text{re}}$ and $f_{\text{fr}}$ are so well approximated by linear functions that we might as well take them to be linear, i.e., $f_{\text{re}} = -kx$ and $f_{\text{fr}} = -hv$, where $k > 0$ and $h > 0$. Newton's equation becomes

$$mx'' + hx' + kx = 0, \tag{10.12}$$

where the constant $h$ is called the *friction* constant. Such a differential equation whose coefficients $m$, $k$, and $h$ are constants has a solution of the form $\mathrm{e}^{rt}$. Substituting $\mathrm{e}^{rt}$ and its first and second derivatives $r\mathrm{e}^{rt}$, $r^2\mathrm{e}^{rt}$ into Eq. (10.12), we get $mr^2\mathrm{e}^{rt} + hr\mathrm{e}^{rt} + k\mathrm{e}^{rt} = 0$, and factoring out the exponential yields $(mr^2 + hr + k)\mathrm{e}^{rt} = 0$. Since the exponential factor is never zero, the sum in the parentheses must be zero:

$$mr^2 + hr + k = 0. \tag{10.13}$$

Our efforts have led to a solution $\mathrm{e}^{rt}$ of Eq. (10.12) for each root of Eq. (10.13).

This is a quadratic equation for $r$, whose solutions are

$$r_{\pm} = -\frac{h}{2m} \pm \frac{\sqrt{h^2 - 4mk}}{2m}.$$

There are two cases, depending on the sign of the quantity under the square root.

- Case I: $h^2 - 4mk$ negative, or $h < 2\sqrt{mk}$.
- Case II: $h^2 - 4mk$ nonnegative, or $2\sqrt{mk} \le h$.

In Case I, the roots are complex, while in Case II, they are real. We first consider Case I.

**Case I, $h < 2\sqrt{mk}$.** Denote by $w$ the real quantity

$$\frac{1}{2m}\sqrt{4mk - h^2} = w.$$

Then the roots can be written as

$$r_{\pm} = -\frac{h}{2m} \pm \mathrm{i}w.$$

This gives two complex-valued solutions,

$$x_-(t) = \mathrm{e}^{r_- t} = \mathrm{e}^{(-\frac{h}{2m} - \mathrm{i}w)t} \quad \text{and} \quad x_+(t) = \mathrm{e}^{r_+ t} = \mathrm{e}^{(-\frac{h}{2m} + \mathrm{i}w)t}.$$

We have seen in Chap. 9 that $\mathrm{e}^{a+\mathrm{i}b} = \mathrm{e}^a(\cos b + \mathrm{i}\sin b)$. Using this with $a = -\dfrac{h}{2m}$ and $b = w$, and again with $b = -w$, we can express

$$x_\pm(t) = \mathrm{e}^{r_\pm t} = \mathrm{e}^{-\frac{h}{2m}t}(\cos wt \pm \mathrm{i}\sin wt).$$

We have shown in Theorem 9.3 that complex exponentials satisfy $(\mathrm{e}^{rt})' = r\mathrm{e}^{rt}$, so the functions $x_+$ and $x_-$ are solutions of Eq. (10.12). We ask you to verify in Problem 10.11 that sums and complex multiples of these solutions are also solutions. As a result,

$$\frac{1}{2}\big(x_+(t) + x_-(t)\big) = \mathrm{e}^{-\frac{h}{2m}t}\cos wt \quad \text{and} \quad \frac{1}{2\mathrm{i}}\big(x_+(t) - x_-(t)\big) = \mathrm{e}^{-\frac{h}{2m}t}\sin wt$$

are solutions. These functions are the product of a trigonometric and an exponential function. The trigonometric function is periodic, with period $\dfrac{2\pi}{w}$, and the exponential function tends to 0 as $t$ tends to infinity. The exponential function diminishes by the factor $\mathrm{e}^{-\frac{h}{2m}}$ per unit time. This is called the *decay rate* of $x(t)$. Such motion is called a *damped vibration*.

By applying the linearity principle for real solutions of a differential equation (see Problem 10.9), we see that every combination of the form

$$x(t) = \mathrm{e}^{-\frac{h}{2m}t}(A\cos wt + B\sin wt),$$

where $A$ and $B$ are constants, is also a solution.

*Example 10.2.* Consider the equation

$$x'' + \frac{1}{2}x' + \frac{17}{16}x = 0. \tag{10.14}$$

Solving $r^2 + \dfrac{1}{2}r + \dfrac{17}{16} = 0$, we obtain $r = -\dfrac{1}{4} + \mathrm{i}$, so $\mathrm{e}^{-\frac{1}{4}t}\big(\cos t + \mathrm{i}\sin t\big)$ is a complex solution. The functions $\mathrm{e}^{-\frac{1}{4}t}\cos t$ and $\mathrm{e}^{-\frac{1}{4}t}\sin t$ are both real solutions to Eq. (10.14), and so are all linear combinations of them. The particular linear combination

$$x(t) = -\mathrm{e}^{-\frac{1}{4}t}\left(\cos t + \frac{1}{4}\sin t\right)$$

is graphed in Fig. 10.8. This solution has initial values $x(0) = -1$ and $x'(0) = 0$.

**Case II, $2\sqrt{mk} \le h$.** The special case of equal roots, $h = 2\sqrt{mk}$, will be discussed in Problem 10.13. For $h > 2\sqrt{mk}$, the roots

$$r_\pm = -\frac{h}{2m} \pm \frac{\sqrt{h^2 - 4mk}}{2m}$$

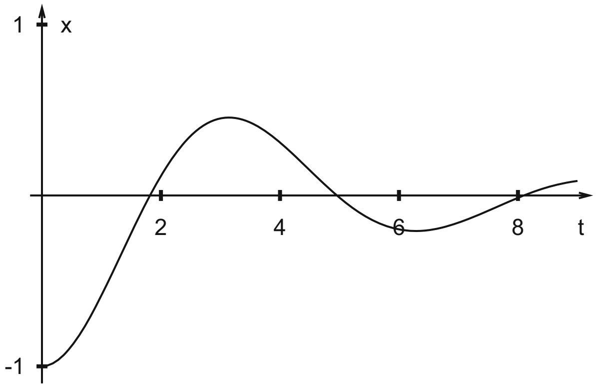


**Fig. 10.8** A graph of the damped vibration $x(t) = -\mathrm{e}^{-t/4}(\cos t + \frac{1}{4}\sin t)$ in Example 10.2

are both real, and they furnish two distinct real exponential solutions, $\mathrm{e}^{r_+t}$ and $\mathrm{e}^{r_-t}$. According to the principle of linearity, every combination of them,

$$x(t) = A_+\mathrm{e}^{r_+t} + A_-\mathrm{e}^{r_-t},$$

is also a solution. We would like to choose the constants $A_+$ and $A_-$ so that the initial displacement is $x(0) = -b$ and the initial velocity is $x'(0) = v(0) = 0$. The desired values of $A_+$ and $A_-$ have to satisfy

$$\begin{aligned} x(0) &= A_+ + A_- = -b, \\ v(0) &= r_+A_+ + r_-A_- = \ 0. \end{aligned}$$

Since $r_+$ and $r_-$ are unequal, $A_+$ and $A_-$ are easily determined from these relations.

*Example 10.3.* We consider the equation

$$x'' + \frac{3}{2}x' + \frac{1}{2}x = 0$$

with initial displacement $x(0) = -1$. The roots of $r^2 + \frac{3}{2}r + \frac{1}{2} = 0$ are $r_- = -1$ and $r_+ = -\frac{1}{2}$. We need to solve $x(0) = A_+ + A_- = -1$ and $-\frac{1}{2}A_+ - A_- = 0$. Adding, we obtain $A_+ = -2$, then $A_- = 1$. Our solution is

$$x(t) = -2\mathrm{e}^{-\frac{1}{2}t} + \mathrm{e}^{-t}.$$

This is graphed for $t > 0$ in Fig. 10.9.

Both roots $r_+$ and $r_-$ in Case II are negative. Consequently, both exponentials tend to zero as $t$ tends to infinity. Of the two negative roots, $r_-$ has the greater magnitude:

$$|r_-| > |r_+|.$$

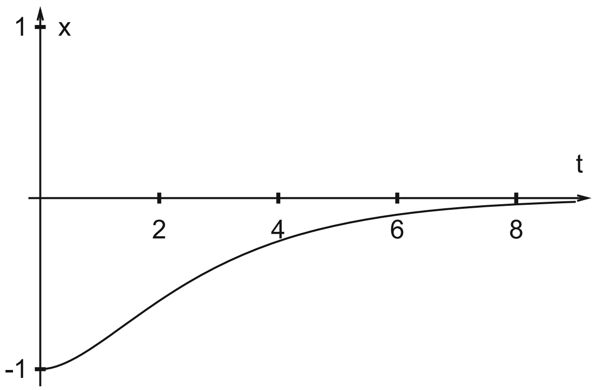


**Fig. 10.9** The overdamped vibration $x(t) = \mathrm{e}^{-t} - 2\mathrm{e}^{-t/2}$ in Example 10.3

It is also true that $|A_+| > |A_-|$, and that for $t > 0$, $\mathrm{e}^{r_+t} > \mathrm{e}^{r_-t}$. Hence $|A_+\mathrm{e}^{r_+t}|$ is always greater than $|A_-\mathrm{e}^{r_-t}|$. As $t$ tends to infinity, the first term becomes very much greater than the second. This shows that the decay of $x(t)$ is governed by the decay rate of the first term. That decay rate is $\mathrm{e}^{r_+}$.

**Rates of Decay.** The difference between Case I and Case II is that in Case I, the force of friction is not strong enough to prevent the particle from swinging back and forth, although it does diminish the magnitude of successive swings. In Case II, friction is so strong compared to the restoring force that it slows down the particle to such an extent that it never swings over to the other side (except possibly in the rare case where $h = 2\sqrt{mk}$). This motion is called *overdamped*.

In both Case I and Case II, motion decays to zero as $t$ tends to infinity. We now investigate the rates of this decay, respectively $\mathrm{e}^{-\frac{h}{2m}}$ and $\mathrm{e}^{r_+}$. The logarithms of these decay rates are called *coefficients of decay* and are denoted by the symbol $\ell$. We have the following formula for $\ell$:

$$\ell(h) = \begin{cases} -\frac{h}{2m} & \text{for } h < 2\sqrt{mk}, \text{ Case I damped,} \\ \frac{-h+\sqrt{h^2-4mk}}{2m} & \text{for } 2\sqrt{mk} < h, \text{ Case II overdamped.} \end{cases}$$

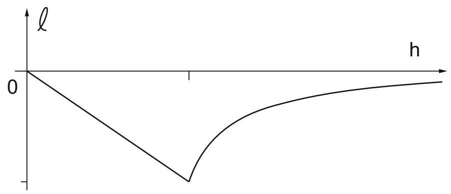


**Fig. 10.10** Coefficient of decay $\ell$ is minimal at $h = 2\sqrt{mk}$

We next study how $\ell$ varies as the friction constant $h$ changes while $m$ and $k$ remain fixed. Properties of the coefficient of decay $\ell$:

(a) $\ell(h)$ is a continuous function for $0 \le h$.
This is true because at the point $h = 2\sqrt{mk}$ where Case I joins Case II, the two formulas for $\ell$ furnish the same value.

(b) $\ell(h)$ is a decreasing function of $h$ for $0 \le h < 2\sqrt{mk}$.
This is true because for $0 \le h < 2\sqrt{mk}$, the derivative of $\ell$ is $-\frac{1}{2m} < 0$.

(c) $\ell(h)$ is an increasing function of $h$ for $2\sqrt{mk} < h$.
This is true because for $h > 2\sqrt{mk}$, the derivative of $\ell$ is positive. To see this, note in $\ell'(h)$ that since $h > \sqrt{h^2 - 4mk}$, the fraction in the large parentheses below is greater than 1:

$$\ell'(h) = \frac{1}{2m}\left(-1 + \frac{h}{\sqrt{h^2 - 4mk}}\right).$$

(d) $\ell(h)$ reaches its minimum value at the *critical damping* $h = 2\sqrt{mk}$.
This is a consequence of the first three items.

Note that the function $\ell(h)$ is continuous and its absolute value is largest at $h = 2\sqrt{mk}$. It is not differentiable at $h = 2\sqrt{mk}$, as can be seen from Fig. 10.10. As the graph indicates, $\ell(h)$ tends to zero as $h$ tends to infinity. Knowing the value of $h$ that maximizes $|\ell|$ is important. For example, in an automobile bouncing after hitting a pothole, the springs provide a restoring force and the shock absorbers provide frictional damping. In fact, the shock is absorbed by the springs; the role of the shock absorbers is to dissipate the energy resulting from a sudden displacement.

## *10.1f Linear Systems Driven by an External Force*

Next, we study the motion of particles under the influence of a restoring frictional force and a *driving force* $f_d$ presented as a known function of time. This is a frequently occurring situation; examples of it are

(a) the motion of the eardrum driven by pressure pulses in the air,
(b) the motion of a magnetic diaphragm under an electromagnetic force,
(c) the motion of air in the resonating cavity of a violin under the force exerted by a vibrating violin string, and
(d) the motion of a building under the force exerted by wind or tremors in the Earth.

Of course, these examples are much more complicated than the case of a single particle that we shall investigate.

Newton's law of motion governing a single particle says that

$$mx'' = f_{\text{re}}(x) + f_{\text{fr}}(v) + f_d(t).$$

We shall discuss the case in which the restoring force and the frictional force are linear functions of their arguments and the driving force is the simple periodic function

$$f_d(t) = F\cos(qt),$$

where $F$ is a positive constant. Substituting these forces into Newton's law, we get the equation

$$mx'' + hx' + kx = F\cos(qt). \tag{10.15}$$

See Fig. 10.11.

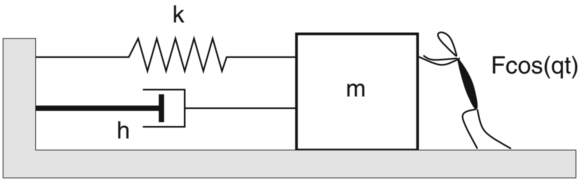


**Fig. 10.11** Forces are applied to a mass $m$ with position $x(t)$ in Eq. (10.15): the linear spring restoring force $-kx$, the linear frictional force $-hx'$, and the applied force $F\cos(qt)$

We begin by establishing a simple relation between any two solutions of this equation. Let $x_0$ be another solution:

$$mx_0'' + hx_0' + kx_0 = F\cos(qt).$$

Subtracting from Eq. (10.15), we get

$$m(x-x_0)'' + h(x-x_0)' + k(x-x_0) = 0, \tag{10.16}$$

i.e., the difference of *any* two solutions of Eq. (10.15) is a solution of Eq. (10.16). But this is the equation governing the motion of particles subject only to a restoring force and a frictional force. In the previous section, we showed that all solutions of Eq. (10.16) tend to zero as $t$ tends to infinity (see Figs. 10.8 and 10.9). This shows that for large $t$, *any two solutions of* Eq. (10.15) *differ by very little*. Thus we may study the large-time behavior of any one solution.

We shall find a solution of Eq. (10.15) by the following trick. We look for complex-valued solutions $z$ of the complex equation

$$mz'' + hz' + kz = Fe^{iqt}. \tag{10.17}$$

We ask you in Problem 10.16 to verify that if $z$ is a complex-valued solution of Eq. (10.17), then $x = \operatorname{Re} z$ is a real solution of Eq. (10.15). The advantage of $z$ is the ease with which we can calculate with exponentials.

We take $z$ of the same form as the driving force, because it is reasonable to guess that the mass oscillates at the same frequency at which it is pushed:

$$z(t) = Ae^{iqt}.$$

Then by Theorem 9.5,

$$z' = Aiq\mathrm{e}^{\mathrm{i}qt} \quad \text{and} \quad z'' = -Aq^2\mathrm{e}^{\mathrm{i}qt}.$$

Substituting these into Eq. (10.17), we get, after division by $\mathrm{e}^{\mathrm{i}qt}$,

$$A(-mq^2 + \mathrm{i}hq + k) = F.$$

Solving for $A$ from this equation, we get that

$$z(t) = \frac{F}{-mq^2 + \mathrm{i}hq + k}\mathrm{e}^{\mathrm{i}qt}$$

is a solution to Eq. (10.17). The real part $x$ of $z$ is a solution of the real part of the complex equation, which is Eq. (10.15), the equation we originally wanted to solve.

**The Response Curve.** The absolute value of the complex solution $z(t)$ is

$$\frac{F}{|-mq^2 + \mathrm{i}hq + k|}$$

for all $t$. This is the maximum of the absolute value of its real part $x(t)$, reached for those values of $t$ for which $z$ is real. This maximum is called the *amplitude* of the vibration. Furthermore, $F$ is the maximum of the absolute value of the imposed force; it is called the amplitude of the force. The ratio of the two amplitudes is

$$R(q) = \frac{\max|x|}{F} = \frac{1}{|-mq^2 + \mathrm{i}hq + k|}.$$

In many ways, the most interesting question is this: for what value of $q$ is $R(q)$ the largest? Clearly, $R(q)$ tends to zero as $q$ tends to infinity, so $R(q)$ has a maximum. We shall calculate the value of the maximum. It occurs at the same frequency $q$ at which the reciprocal of $R(q)$ is minimized:

$$\frac{1}{\big(R(q)\big)^2} = |-mq^2 + \mathrm{i}hq + k|^2 = (k - mq^2)^2 + h^2q^2.$$

The derivative of this with respect to $q$ is

$$4mq(mq^2 - k) + 2h^2q = 2q(2m^2q^2 - 2mk + h^2),$$

and it is zero at $q = 0$. To find other possible zeros, we set the remaining factor equal to zero: $2m^2q^2 - 2mk + h^2 = 0$. After rearrangement, we get

$$q^2 = \frac{2mk - h^2}{2m^2} = \frac{k}{m} - \frac{h^2}{2m^2}.$$

If the quantity on the right is negative, which is the overdamped Case II, $h > \sqrt{2mk}$, the equation cannot be satisfied, and we conclude that the maximum of $R$ is reached at $q = 0$. If, however, the quantity on the right is positive, in Case I, then

$$q_r = \sqrt{\frac{k}{m} - \frac{h^2}{2m^2}}$$

is a possible candidate for the value for which $R(q)$ achieves its maximum. A direct calculation shows that

$$R(q_r) = \frac{1}{h} \frac{1}{\sqrt{\frac{k}{m} - \frac{h^2}{4m^2}}}.$$

In Problem 10.19, we ask you to show that $R(q_r)$ is greater than $R(0) = \dfrac{1}{k}$. So for $h < \sqrt{2mk}$, the graph of $R(q)$ looks qualitatively like the example in Fig. 10.12. The graph of $R$ is called the *response curve* of the vibrating system.

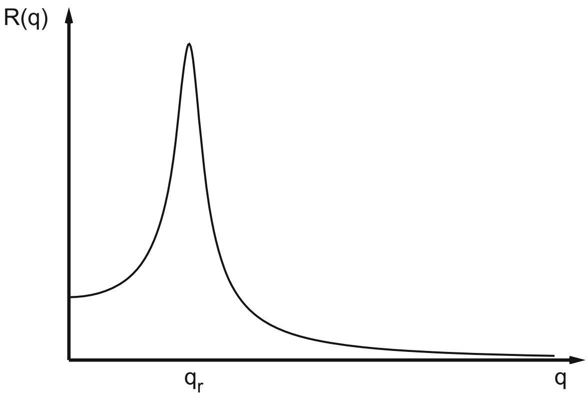


**Fig. 10.12** Graph of the response function $R(q) = \dfrac{1}{\sqrt{(-q^2+1)^2 + \left(\frac{1}{5}q\right)^2}}$ for the damped equation $x'' + \dfrac{1}{5}x' + x = F\cos(qt)$. The maximum is $R(0.989\ldots) = 5.025\ldots$

The significance of the maximum at $q_r$ is that *among all driving forces of the form $F\cos qt$, the one with $q = q_r$ causes the motion with the largest amplitude*. This phenomenon is called *resonance*, and $\dfrac{q_r}{2\pi}$ is called the *resonant frequency*. Resonance is particularly striking if friction is small, i.e., if $h$ is small, for then $R(q_r)$ is so large that at the resonant frequency, even a low-amplitude driving force will cause a motion of large amplitude. A known dramatic example of this kind of resonance is the shattering of a wine glass by a musical note pitched at the resonant frequency of the glass.

We conclude this section with a summary:

*For motion under a restoring force without friction:*

(a) Total energy is conserved.
(b) All motions are periodic.
(c) All motions with relatively small amplitude have approximately the same period.

*For motion under a restoring force with friction:*

(d) Total energy is decreasing.
(e) All motion decays to zero at an exponential rate.
(f) There is a critical value of the coefficient of friction that maximizes the rate at which solutions decay to zero.

We have proved (e) and (f) only for a linear restoring force and linear friction.

*For motions under a linear restoring force, linear friction, and a sinusoidal driving force:*

(g) All motions tend toward a sinusoidal motion with the same frequency as the driving force.
(h) If friction is not too large, there is a resonant frequency.

## Problems

**10.1.** Which of the following differential equations are examples of the model (10.2) that we developed for vibrations of a mechanical system? Be sure to check the required properties of the frictional and restoring forces.

(a) $2x'' - x = 0$
(b) $x'' + x' + x + x^3 = 0$
(c) $x'' + x' = 0$
(d) $x'' - x^2 = 0$
(e) $x'' - 0.07x' - 3x = 0$

**10.2.** Verify that since we assumed that the restoring force $f_{\text{re}}(x)$ is an odd function, the potential energy $p(x) = -\int_0^x f_{\text{re}}(y)\,\mathrm{d}y$ is an even function.

**10.3.** Solve $x'' + x' = 0$, which has no restoring force, by trying a combination of exponential solutions $x(t) = \mathrm{e}^{rt}$ for the two cases

(a) $x(0) = 5,\ x'(0) = 7,$
(b) $x(0) = 5,\ x'(0) = -7.$

Do the solutions have limits as $t$ tends to infinity?

**10.4.** Find an equation $mx'' + hx' + kx = 0$ if the roots to $mr^2 + hr + k = 0$ are $r_\pm = -\frac{1}{10} \pm \mathrm{i}$.

**10.5.** Find exponential solutions $\mathrm{e}^{rt}$ of $2x'' + 7x' + 3x = 0$.

**10.6.** Find trigonometric solutions of $2y'' + 3y = 0$.

**10.7.** As indicated by the graph in Fig. 10.10 of the coefficient of decay $\ell$, some solutions $x(t)$ of $mx'' + hx' + kx = 0$ decay toward zero very slowly if $h$ is either very small or very large. Sketch typical solutions for both cases.

**10.8.** Find a complex exponential solution $z(t)$ of the equation $z'' + 4z' + 5z = 0$ and verify that the real part $x(t) = \operatorname{Re} z(t)$ is a solution of $x'' + 4x' + 5x = 0$.

**10.9.** Let $x_1(t)$ and $x_2(t)$ be real-valued functions that are solutions of the $n$th-order differential equation

$$A_n x^{(n)}(t) + \cdots + A_2 x''(t) + A_1 x'(t) + A_0 x(t) = 0,$$

where the $A_i$ are real constants.

(a) Show that if $c$ is any real constant, then $cx_1(t)$ is a solution.
(b) Show that $y(t) = x_1(t) + x_2(t)$ is a solution.

Combining these observations, we observe that $c_1 x_1(t) + c_2 x_2(t)$ is a solution whenever $x_1$ and $x_2$ are solutions and the $c$'s are constant. This is an example of the *linearity* of this differential equation.

**10.10.** Suppose that the coefficients $A_0, A_1, \ldots, A_n$ in the differential equation of Problem 10.9 are functions of $t$. Are the assertions made there still valid? Are the assertions true if instead, the equation is modified to

$$A_n x^{(n)}(t) + \cdots + A_2 x''(t) + A_1 x'(t) + A_0 x(t) = \cos t\ ?$$

**10.11.** Suppose $x_1(t) = p_1(t) + \mathrm{i} q_1(t)$ and $x_2(t) = p_2(t) + \mathrm{i} q_2(t)$ are complex-valued solutions of

$$mx'' + hx' + kx = 0$$

and that $c_1$ and $c_2$ are complex numbers. Show that $c_1 x_1(t) + c_2 x_2(t)$ is a solution.

**10.12.** The function $x(t) = \mathrm{e}^{-bt}\cos(wt)$ represents a motion under a linear restoring force and linear friction.

(a) Show that the interval between successive times when $x(t) = 0$ has length $\dfrac{\pi}{w}$.

(b) Show that the time interval between successive local maxima is $\dfrac{2\pi}{w}$.

**10.13.** Consider the equation of motion $mx'' + hx' + kx = 0$, and suppose that $h$ has the critical value $2\sqrt{mk}$.

(a) Show that the only solution of the form $\mathrm{e}^{rt}$ has $r = -\sqrt{\dfrac{k}{m}}$.

(b) Show that $t\mathrm{e}^{-\sqrt{\frac{k}{m}}t}$ is a solution.

**10.14.** Find all solutions $x(t)$ of the equation of motion $x'' + x' + x = 0$.

**10.15.** Find a complex exponential solution $z(t)$ of the equation $z'' + z' + 6z = 52e^{6it}$, and verify that the real part $x(t) = \mathrm{Re}\, z(t)$ is a solution of $x'' + x' + 6x = 52\cos(6t)$.

**10.16.** Show that if $z(t)$ is any solution of $mz'' + hz' + kz = Fe^{iqt}$, then the real part $x(t) = \mathrm{Re}\, z(t)$ is a solution of $mx'' + hx' + kx = F\cos(qt)$.

**10.17.** Find a solution $x_1(t)$ of the equation $x'' + x' + x = \cos t$. Verify that you may add any solution of $y'' + y' + y = 0$ to your solution to get another solution $x_2 = y + x_1$ of $x'' + x' + x = \cos t$.

**10.18.** A heavy motor runs at 1800 rpm, causing the floor to vibrate with small vertical displacements $y(t) = A\cos(\omega t)$. Find $\omega$ if $t$ is measured in minutes.

**10.19.** Prove that in the damped case $h < \sqrt{2mk}$, the response maximum $R(q_r)$ is greater than $R(0) = \frac{1}{k}$, i.e., that $\frac{1}{h}\frac{1}{\sqrt{\frac{k}{m} - \frac{h^2}{4m^2}}} > \frac{1}{k}$.

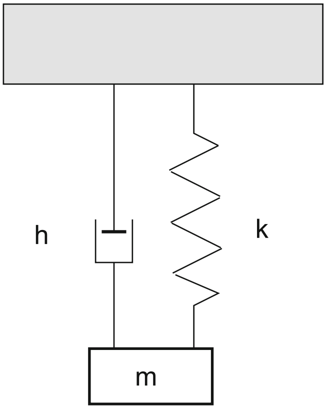


**Fig. 10.13** A spring and mass with friction. Before gravity is applied, $y = 0$ is the equilibrium. $f_{\mathrm{re}}(y) = -ky$ $f_{\mathrm{fr}}(v) = -hv$. See Problem 10.20

**10.20.** Newton's equation of motion for a particle at the end of a vertical spring (see Fig. 10.13) under the influence of the restoring force of the spring, friction, and the applied force of gravity is

$$my'' + hy' + ky = mg.$$

Here the displacement $y$ is measured as positive downward, and $m$, $h$, $k$, and $g$ are positive constants.

(a) Show that the difference of any two solutions solves the equation for the case of no gravity, $mx'' + hx' + kx = 0$.
(b) Find a constant solution $y$.
(c) Show that every solution is of the form $y(t) = \frac{gm}{k} + x(t)$, where $x$ solves the case of no gravity.
(d) Show that as $t$ tends to infinity, every solution $y(t)$ tends to the constant solution.

**10.21.** Justify the following items, which prove the uniqueness theorem, Theorem 10.1, stated at the end of Sect. 10.1a.

(a) If $mx'' - f_{\text{fr}}(x') - f_{\text{re}}(x) = 0$ and $my'' - f_{\text{fr}}(y') - f_{\text{re}}(y) = 0$, denote by $w$ the difference $w(t) = y(t) - x(t)$. Then

$$mw'' - \big(f_{\text{fr}}(w' + x') - f_{\text{fr}}(x')\big) - \big(f_{\text{re}}(w + x) - f_{\text{re}}(x)\big) = 0.$$

(b) For each $t$, there is $u$ between $x'(t)$ and $y'(t)$ and $v$ between $x(t)$ and $y(t)$ such that

$$mw'' - f'_{\text{fr}}(u)w' - f'_{\text{re}}(v)w = 0.$$

(c) Therefore, $mw''w' - f'_{\text{re}}(v)ww' \le 0$.
(d) $f'_{\text{re}}$ is bounded above by some constant $-k \le 0$. Therefore, $mw''w' + kww' \le 0$, and

$$\frac{1}{2}m(w')^2 + \frac{1}{2}kw^2$$

is nonincreasing.
(e) A nonnegative nonincreasing function that is 0 at time $s$ must be 0 for all times $t > s$. Explain why this implies $w(t) = 0$ for all $t > s$.

## 10.2 Population Dynamics

In this section, calculus is used to study the evolution of populations—animal, vegetable, or mineral. About half the material is devoted to formulating the laws governing population changes in the form of differential equations, and the other half to studying their solutions. Only in the simplest cases can this be accomplished by obtaining explicit formulas for solutions. When explicit solutions are not available, relevant qualitative and quantitative properties of solutions can nevertheless be deduced directly from the equations, as our examples will show. Using numerical methods that extend those we mention in Sect. 10.4, one can generate extremely accurate approximations to any specific solution of a differential equation. These may lead to the answers we seek or suggest trends that once perceived, can often be deduced logically from the differential equations.

Theoretical population models have become more and more useful in such diverse fields as the study of epidemics and the distribution of inherited traits. Yet the most important application, the one about which the public needs to be informed in order to make intelligent decisions, is to demography, the study of human populations. Indeed, as Alexander Pope put it, "The proper study of Mankind is Man."

In Sect. 10.2a, we develop a theory of differential equations needed for the study of population growth. In Sect. 10.2b, we describe the dynamics of a population consisting of a single species, and in Sect. 10.2c, the dynamics of a population consisting of two species.

## *10.2a The Differential Equation* $\frac{dN}{dt} = R(N)$

In this section, we analyze the type of differential equations that govern both population growth and chemical reactions, but without reference to these applications. We consider the equation

$$\frac{dN}{dt} = R(N), \tag{10.18}$$

where $R(N)$ is a known rate depending on $N$. In Sect. 3.3, we solved one equation of this type, $\frac{dN}{dt} = kN$. We saw that the solutions $N(t) = N(0)e^{kt}$ are exponential functions, including the constant function $N(t) = 0$. In Fig. 10.14, we plot the solutions of $\frac{dN}{dt} = -N$ for five different initial conditions $N_0 = N(0) = 3, 1, 0, -1, -2$. We know that as $t$ tends to infinity, each solution shown, $N(t) = N(0)e^{-t}$, regardless of the initial condition, tends to the constant solution $N = 0$. In the context of populations, $N_0 < 0$ may not make sense, but the differential equation has solutions for those initial conditions, so we include them in our analysis.

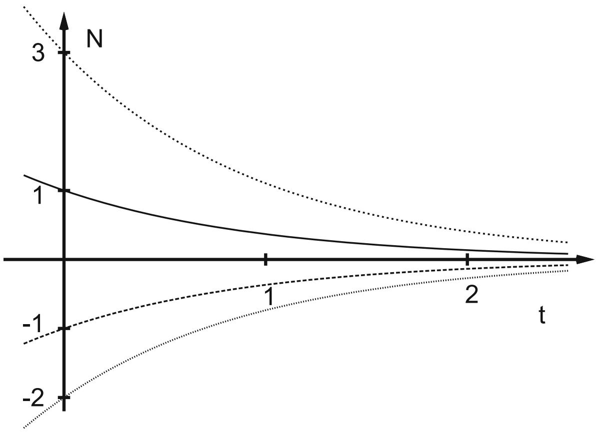


**Fig. 10.14** Graphs of five solutions to $\frac{dN}{dt} = -N$

The equation $\frac{dN}{dt} = 2N - N^2 = N(2-N)$ is another example of an equation of the form $\frac{dN}{dt} = R(N)$. In Sect. 10.2b, you can see that we will find explicit solution formulas for this equation, some of which we have graphed on the right side of Fig. 10.16.

Both of these differential equations are statements about the relative growth rate $\frac{\frac{dN}{dt}}{N}$ of the population. In the first equation, the relative growth rate $\frac{\frac{dN}{dt}}{N} = -1$ is constant. In the second, the relative growth rate $\frac{\frac{dN}{dt}}{N} = 2 - N$ decreases as $N$ increases from 0 to 2, perhaps due to a lack of resources.

A third equation of the form (10.18) is $\frac{dN}{dt} = -N(N-1)(N-2)$. Solutions of this equation are plotted in Fig. 10.15. A fourth such equation is given in Example 10.7, whose solutions are plotted in Fig. 10.22.

You may have begun to perceive a pattern. It appears that the constant solutions, i.e., the places where $R(N) = 0$, play a key role in describing the long-term behavior of the other solutions. Our first task is to determine conditions under which solutions exist and are determined uniquely by the initial condition. Then we will show how the zeros of $R(N)$ are related to the long-term behavior of solutions.

To begin, assume that $R(N)$ is a continuous function of $N$, different from zero. Then we can divide both sides of

$$\frac{dN}{dt} = R(N)$$

by $R(N)$, obtaining

$$\frac{1}{R(N)} \frac{dN}{dt} = 1.$$

By the fundamental theorem of calculus, since $\frac{1}{R(N)}$ is a continuous function of $N$, there is a function $Q(N)$ whose derivative is

$$\frac{dQ}{dN} = \frac{1}{R(N)}. \tag{10.19}$$

If $N(t)$ satisfies $\frac{dN}{dt} = R(N)$, then by the chain rule,

$$\frac{dQ}{dt} = \frac{dQ}{dN}\frac{dN}{dt} = \frac{1}{R(N)}\frac{dN}{dt} = 1.$$

A function with constant derivative is linear, so

$$Q(N(t)) = t + c, \quad c \text{ a constant.}$$

It follows from $\frac{dQ}{dN} = \frac{1}{R(N)}$ that $\frac{dQ}{dN}$ is not zero, and from the continuity of $R$ that $\frac{dQ}{dN}$ does not change sign. Therefore, $Q(N)$ is strictly monotonic and hence invertible. This means that $Q(N(t)) = t + c$ can be solved for

$$N(t) = Q^{-1}(t+c).$$

The constant $c$ can be related to the initial value $N(0) = N_0$ by setting $t = 0$:

$$N(0) = Q^{-1}(c) \quad \text{or} \quad Q(N_0) = c.$$

With this determination of $c$, the function

$$N(t) = Q^{-1}(t + Q(N_0)) \tag{10.20}$$

is the solution of $\frac{dN}{dt} = R(N)$ with initial value $N_0$. Thus we have proved the following theorem.

**Theorem 10.2. Existence.** *If $R(N)$ is a continuous function of $N$ that is never 0, then the differential equation*

$$\frac{dN}{dt} = R(N), \quad \text{with} \quad N(0) = N_0,$$

*has a unique solution on a possibly infinite $t$-interval $(r,s)$. If one of the endpoints $r$ or $s$ is finite, the solution $N(t)$ approaches plus or minus infinity as $t$ approaches the endpoint.*

It is instructive to look at the example $\frac{dN}{dt} = N^2 + 1$, with initial value $N(0) = 0$. We divide this equation by $N^2 + 1$ and get $\frac{1}{N^2+1}\frac{dN}{dt} = 1$. The left side is the derivative of $\tan^{-1} N$, so integration gives $\tan^{-1} N = t + c$. Since we specified that $N(0) = 0$, it follows that $c = 0$, and so $N(t) = \tan t$, defined on the interval $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$. As $t$ approaches the left or right endpoint of this interval, $N(t)$ tends to minus or plus infinity.

We now turn to the more interesting case that $R(N)$ vanishes at some points. The derivation of formula (10.20) for the solution of the initial value problem shows more than what is stated in Theorem 10.2. It shows that even if $R(N)$ vanishes at some points, if it is not zero for $R(N_0)$, then the method used to solve the initial value problem yields a solution on a short time interval $(-d, d)$.

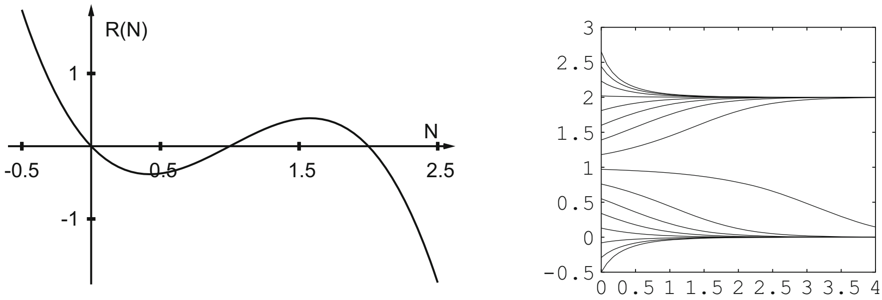


**Fig. 10.15** *Left*: graph of the right-hand side $R(N) = -N(N-1)(N-2)$. *Right*: graphs of some computed solutions to $\dfrac{dN}{dt} = -N(N-1)(N-2)$

**Theorem 10.3.** *Suppose that the function $R(N)$ is differentiable and $N(t)$ is a solution of*

$$\frac{dN}{dt} = R(N).$$

*Suppose that $N(0)$ is not a zero of the function $R(N)$, that is, $R(N(0)) \neq 0$.*

*Then $N(t)$ is not a zero of $R(N)$ for any value of $t$.*

*Proof.* We shall argue indirectly. Suppose, to the contrary, that at some $s$, $N(s)$ is a zero of R(N). Denote this zero by $Z$:

$$N(s) = Z, \qquad R(Z) = 0.$$

We shall show that then $R(N(t))$ is zero for all $t$.

Since $R(Z) = 0$, we have $R(N) = R(N) - R(Z)$. Using the mean value theorem, we obtain $R(N) = k(N-Z)$, where $k$ is the value of the derivative of the function $R$ at some point between $N$ and $Z$. Since $Z$ is a constant, we can rewrite the differential equation governing $N$ as

$$\frac{d(N-Z)}{dt} = R(N) = k(N-Z).$$

Denote the function $N(t) - Z$ by $M(t)$ and write the differential equation for $N - Z$ as $\dfrac{dM}{dt} = kM$. Multiply this equation by $2M$. We get $2M\dfrac{dM}{dt} = 2kM^2$. Denote the function $M^2$ by $P$ and rewrite the equation above as

$$\frac{dP}{dt} = kP.$$

Denote by $m$ an upper bound for the function $k(N)$. We deduce from this differential equation the inequality

$$\frac{dP}{dt} \le mP.$$

We write this inequality as $\frac{dP}{dt} - mP \le 0$, and multiply it by $e^{-mt}$. We get

$$e^{-mt}\frac{dP}{dt} - me^{-mt}P \le 0.$$

The left side is the derivative of $e^{-mt}P$, and since it is nonpositive, $e^{-mt}P$ is a non-increasing function of $t$. The function $P$ was defined as $M^2$, and $M$ as $N - Z$. Since $Z$ is the value of the function $N$ at $s$, $M(s)$ is zero, and so is $P(s)$. Since the function $P$ is a square, its values are nonnegative, and so are the values of $e^{-mt}P(t)$. We have shown before that $e^{-mt}P(t)$ is a nonincreasing function of $t$. But since $e^{-mt}P(t)$ is zero at $s$, it follows that $e^{-mt}P(t)$, and therefore $P(t)$, is zero for $t$ greater than $s$. Using a similar argument but with a lower bound for the function $k(N)$, we show similarly that $P(t)$ is zero for all $t$ less than $s$.

Since $P(t)$ is the square of $M(t)$ and $M(t)$ is $N(t) - Z$, this proves that $N(t) = Z$ for all $t$. But this contradicts the assumption that $N(0)$ is not a zero of $R(N)$. Since we got into this contradiction by denying Theorem 10.3, this proves the theorem.

□

Next we see how the zeros of $R(N)$ are related to the long-term behavior of the solutions of $\frac{dN}{dt} = R(N)$. From Theorem 10.3, we shall deduce the following property of solutions of Eq. (10.18).

**Theorem 10.4.** *Denote by $N(t)$ a solution of $\frac{dN}{dt} = R(N)$, and its value $N(0)$ by $N_0$. Suppose that $R(N)$ is differentiable, its derivative bounded, and assume that $R(N)$ is positive for $N$ large negative and that it is negative for $N$ large positive.*

*(a) If $R(N_0)$ is negative, then the solution $N(t)$ decreases as $t$ increases, and as $t$ tends to infinity, $N(t)$ tends to the largest zero of $R(N)$ that is less than $N_0$.*
*(b) Similarly, if $R(N_0)$ is positive, then $N(t)$ is an increasing function of $t$, and as $t$ tends to infinity, $N(t)$ tends to the smallest zero of $R(N)$ that is larger than $N_0$.*

Before we write the proof of Theorem 10.4, let us see what the theorem tells us about our two examples:

- $\frac{dN}{dt} = N(2 - N)$. If $0 < N_0 < 2$, then $R(N_0)$ is positive. According to Theorem 10.4, as $t$ tends to infinity, the solution $N(t)$ increases to $N = 2$. If $2 < N_0$, $R(N_0)$ is negative, and the solution decreases to 2, the largest zero of $R(N)$ that is less than $N_0$.
- $\frac{dN}{dt} = -N(N-1)(N-2)$. The graph of $R(N)$ on the left side of Fig. 10.15 will help us tell where $R(N_0)$ is positive or negative. If $N_0 < 0$, $R(N_0)$ is positive, so

according to Theorem 10.4, $N(t)$ increases to $N = 0$. If $0 < N_0 < 1$, then $R(N_0)$ is negative, so $N(t)$ decreases to $N = 0$. If $1 < N_0 < 2$, then $R(N_0)$ is positive, so $N(t)$ increases to the smallest zero that is larger than $N_0$, i.e., $N = 2$. If $2 < N_0$, then $R(N_0)$ is negative, and so $N(t)$ decreases to $N = 2$ as $t$ tends to infinity. This agrees with the approximate solutions computed in Fig. 10.15.

Now we will prove the theorem that makes so much qualitative information about the solutions readily available.

*Proof.* We prove part (a); assume that $R(N_0)$ is negative. Since $R(N)$ is positive for $N$ large negative, it has a zero less than $N_0$. Denote by $M$ the largest zero of $R(N)$ less than $N_0$. When $R(N_0)$ is negative, then according to $\frac{\mathrm{d}N}{\mathrm{d}t} = R(N)$, the derivative of $N(t)$ is negative at $t = 0$ and remains negative as long as $N(t)$ is greater than $M$, because $R(N)$ is negative for all values of $N$ between $M$ and $N_0$. It follows that $N(t)$ is a decreasing function of $t$ and keeps decreasing as long as $N(t)$ is greater than $M$. According to Theorem 10.3, $N(t)$ is not equal to a zero of $R(N)$. Therefore, $N(t)$ is greater than $M$ for all positive $t$.

We show now that as $t$ tends to infinity, $N(t)$ tends to $M$. We again argue indirectly and assume to the contrary that for all $t$, $N(t)$ is greater than $M + p$, $p$ some positive number. The function $R(N)$ is negative on the interval $[M + p, N_0]$. Denote by $m$ the maximum of $R(N)$ on this interval; $m$ is a negative number. We apply the mean value theorem to the function $N(t)$:

$$\frac{N(t) - N(0)}{t} = \frac{\mathrm{d}N}{\mathrm{d}t}(c),$$

where $c$ is some number between 0 and $t$. Since $N(t)$ is a solution of the differential equation, this gives

$$\frac{N(t) - N(0)}{t} = \frac{\mathrm{d}N}{\mathrm{d}t}(c) = R(N(c)) \le m.$$

We deduce that $N(t) \le N(0) + mt$ for all positive $t$. Since $m$ is negative, this would imply that $N(t)$ tends to minus infinity as $t$ tends to infinity. This is contrary to our previous demonstration that $N(t)$ is greater than $M$ for all $t$. Therefore, our assumption that $N(t)$ is greater than $M + p$ for all $t$ must be false.

This completes the proof in the case that $R(N_0)$ is negative. The proof for $R(N_0)$ positive is analogous. □

*Remark.* Take the case that the zeros $M$ of $R(N)$ are *simple* in the sense that $\frac{\mathrm{d}R}{\mathrm{d}N}(M)$ is not zero. Theorem 10.4 can be expressed in this way: *The zeros of* $R(N)$ *where the derivative* $\frac{\mathrm{d}R}{\mathrm{d}N}$ *is negative attract solutions of* $\frac{\mathrm{d}N}{\mathrm{d}t} = R(N)$. Here is why: the proof shows that solutions less than a root $M$ will increase toward $M$ where $R$ is positive, and solutions greater than $M$ decrease toward $M$ when $R$ is negative. But if $R$ is positive below $M$ and negative above, then $\frac{\mathrm{d}R}{\mathrm{d}N}(M)$ must be either negative or 0.

The zeros of $R(N)$ are called equilibrium solutions. A zero that attracts solutions which that nearby it is called a *stable* equilibrium. A zero that repels some near solutions, such as $N = 1$ in Fig. 10.15, is called unstable.

## 10.2b Growth and Fluctuation of Population

*The Arithmetic of Population and Development*

> The rate of population growth is itself impeding efforts at social and economic development. Take the case, for example, of a developing country which has achieved an annual increase in its gross national product of five percent–a very respectable effort indeed, and one which few countries have been able to maintain on a continuing basis. Its population is increasing at 3 percent annually. Thus, its per capita income is increasing by 2 percent each year, and will take 35 years to double, say from \$100 per year to \$200. In the meantime, its population will have almost tripled, so that greatly increased numbers of people will be living at what is still only a subsistence level. Reduction of the rate of population growth is not a *sufficient* condition for social and economic development–other means, such as industrialization, must proceed along with such reduction–but it is clear that it is a *necessary* condition without which the development process is seriously handicapped.

Dr. John Maier
Director for Health Sciences
The Rockefeller Foundation

In this section we shall study the growth of population of a single species and of several species living in a shared environment. The growth rate of a population is related to the birth and death rates. The basic equation governing the growth in time $t$ of a single population of size $N(t)$ is

$$\frac{\mathrm{d}N}{\mathrm{d}t} = B - D,$$

where $B$ is the *birth rate* and $D$ is the *death rate* for the total population. What do $B$ and $D$ depend on? They certainly depend on the age distribution within the population; a population with a high percentage of old members will have a higher death rate and lower birth rate than a population of the same size that has a low percentage of old members. Yet in this section we shall disregard this dependence of birth and death rates on age distribution. The results we shall derive are quantitatively relevant in situations in which the age distribution turns out to change fairly little over time.

If we assume, in addition, that the basic biological functions of the individuals in the population are unaffected by the population size, then it follows that *both birth rate and death rate are proportional to the population size*. The mathematical expression of this idea is

$$B = cN, \quad D = dN,$$

where $c,\ d$ are constants. Substituting this into the differential equation leads to

$$\frac{\mathrm{d}N}{\mathrm{d}t} = aN,$$

where $a = c - d$. The solution of this equation is

$$N(t) = N_0 e^{at},$$

where $N_0 = N(0)$ is the initial population size. For positive $a$ this is the celebrated—and lamented—Malthusian law of population explosion.

**The Verhulst Model.** If the population grows beyond a certain size, the sheer size of the population will *depress* the birth rate and *increase* the death rate. We summarize this as

$$\frac{\mathrm{d}N}{\mathrm{d}t} = aN - \text{effect of overpopulation.}$$

How can we quantify the effect of overpopulation? Let us assume that the *effect of overpopulation is proportional to the number of encounters between members of the population* and that these encounters are *by chance*, i.e., are due to individuals bumping into each other without premeditation. For each individual, the number of encounters is proportional to the population size. In Chap. 11, we find that probabilities of independent events need to be multiplied, and therefore, the total number of such encounters is proportional to the *square* of the population. So the effect of overpopulation is to depress the rate of population growth by $bN^2$, for $b$ some positive number. The resulting growth equation is

$$\frac{\mathrm{d}N}{\mathrm{d}t} = aN - bN^2, \quad a, b > 0. \tag{10.21}$$

This equation was introduced into the theory of population growth by Verhulst. It is a special instance of the equation $\frac{\mathrm{d}N}{\mathrm{d}t} = R(N)$, discussed in Sect. 10.2a.

*Example 10.4.* Consider $\frac{\mathrm{d}N}{\mathrm{d}t} = 2N - N^2 = N(2-N)$ in Fig. 10.16. Note that when the population $N$ is between 0 and 2, it must increase, because the rate of change $N(2-N)$ is positive then.

Now we find a solution formula for the Verhulst model (10.21). Suppose that the right side of Eq. (10.21) is not zero. Then dividing by $aN - bN^2$, we get

$$1 = \frac{1}{aN - bN^2}\frac{\mathrm{d}N}{\mathrm{d}t}.$$

We write the right side as a derivative:

$$1 = \frac{1}{\frac{a}{N} - b}\frac{1}{N^2}\frac{\mathrm{d}N}{\mathrm{d}t} = \frac{1}{\frac{a}{N} - b}\frac{\mathrm{d}}{\mathrm{d}t}\left(-\frac{1}{N}\right) = -\frac{1}{a}\frac{\mathrm{d}}{\mathrm{d}t}\left(\log\left(\frac{a}{N} - b\right)\right).$$

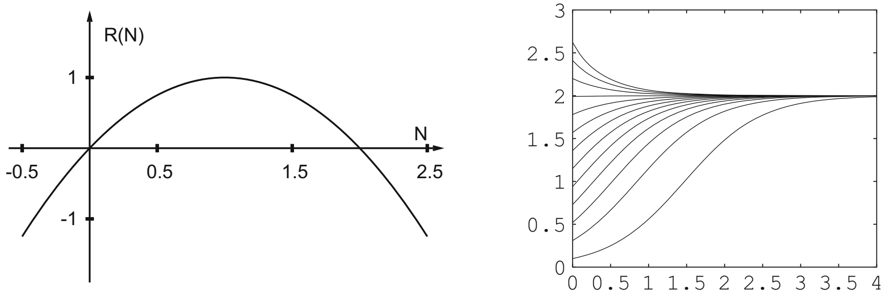


**Fig. 10.16** *Left*: graph of the right-hand side $R(N) = N(2-N)$ in Example 10.4. *Right*: graphs of some computed solutions to $\frac{dN}{dt} = N(2-N)$

Integrating yields $\log\left(\frac{a}{N} - b\right) = c - at$ for some number $c$. If $N_0$ denotes the initial value of $N$, then $\log\left(\frac{a}{N_0} - b\right) = c$. Therefore,

$$\log\left(\frac{a}{N} - b\right) = \log\left(\frac{a}{N_0} - b\right) - at.$$

Combining the logarithms gives

$$\log\left(\frac{\frac{a}{N} - b}{\frac{a}{N_0} - b}\right) = -at.$$

Applying the exponential function gives $\frac{\frac{a}{N} - b}{\frac{a}{N_0} - b} = e^{-at}$, which can be solved for $N(t)$ as

$$N(t) = \frac{\frac{a}{b}N_0}{N_0 - \left(N_0 - \frac{a}{b}\right)e^{-at}}.$$

Many interesting properties of $N(t)$ can be deduced from this formula by inspection.

**Theorem 10.5.** *Assume that the initial value $N(0) = N_0$ is positive in the Verhulst model*

$$\frac{dN}{dt} = aN - bN^2.$$

*(a) If $N_0 > \frac{a}{b}$, then $N(t) > \frac{a}{b}$ for all $t$ and decreases as time increases.*
*(b) If $N_0 = \frac{a}{b}$, then $N(t) = \frac{a}{b}$ for all $t$.*
*(c) If $N_0 < \frac{a}{b}$, then $N(t) < \frac{a}{b}$ for all $t$ and increases as time increases.*

*In all cases, $N(t)$ tends to $\frac{a}{b}$ as $t$ tends to infinity.*

These findings are in complete agreement with Theorem 10.4. For according to that theorem, every solution of $\frac{\mathrm{d}N}{\mathrm{d}t} = R(N)$ tends to the nearest stable steady state. For the equation at hand, $R(N) = aN - bN^2 = bN\left(\frac{a}{b} - N\right)$. The steady states are the zeros of $R$, in this case $N = 0$ and $N = \frac{a}{b}$. The derivative of $R$ is $\frac{\mathrm{d}R}{\mathrm{d}N} = a - 2bN$, so its values at the zeros of $R$ are

$$\frac{\mathrm{d}R}{\mathrm{d}N}(0) = a \quad \text{and} \quad \frac{\mathrm{d}R}{\mathrm{d}N}\left(\frac{a}{b}\right) = -a.$$

Since $a$ is positive, we conclude that both zeros are simple, and that $\frac{\mathrm{d}R}{\mathrm{d}N}$ is positive at $N = 0$, negative at $N = \frac{a}{b}$. Therefore, the zero 0 is unstable and the zero $\frac{a}{b}$ is stable, and all solutions with initial value $N_0 > 0$ tend to the stable steady state $\frac{a}{b}$ as $t$ tends to infinity. This is exactly what we found by studying the explicit formula for all solutions. It is gratifying that properties of solutions can be deduced directly from the differential equation that they satisfy without help from an explicit formula for solutions. Indeed, there are very few differential equations whose solutions can be described by an explicit formula.

The result we have just obtained, that all solutions of the Verhulst model (10.21) tend to $\frac{a}{b}$ as $t$ tends to infinity, has great demographic significance, for it predicts the eventual steady state of any population that can reasonably be said to be governed by an equation of that form.

**An Extinction Model.** We now return to the basic equation $\frac{\mathrm{d}N}{\mathrm{d}t} = R(N)$ of population growth and again we assume that the *death rate is proportional to the population size*. This amounts to assuming that death is due to "natural" causes, and not due to one member of the population eating the food needed by another member, or due to one member eating another. On the other hand, we *challenge the assumption that birth rate is proportional to population size*. This assumption holds for extremely primitive organisms, such as amoebas, which reproduce by dividing. It is also true of well-organized species, such as human beings, who seek out a partner and proceed to produce a biologically or socially determined number of offspring. But there are important classes of organisms whose reproductive sophistication falls between those of the amoeba and humans, who need a partner for reproduction but must rely on chance encounters for meeting a mate. The expected number of encounters is proportional to the *product* of the numbers of males and females. If these are equally distributed in the population, the number of encounters—and so the birth rate—is proportional to $N^2$. The death rate, on the other hand, is proportional to the population size $N$. Since the rate of population growth is the difference between birth rate and death rate, the equation governing the growth of such populations is

$$\frac{\mathrm{d}N}{\mathrm{d}t} = bN^2 - aN, \ a, b > 0.$$

This equation is of the form $\frac{dN}{dt} = R(N)$ with

$$R(N) = bN^2 - aN = bN\left(N - \frac{a}{b}\right), \quad a, b > 0.$$

This function has two zeros, 0 and $\frac{a}{b}$. The derivative is $\frac{dR}{dN} = 2bN - a$, so its values at the zeros of $R$ are

$$\frac{dR}{dN}(0) = -a \quad \text{and} \quad \frac{dR}{dN}\left(\frac{a}{b}\right) = a.$$

Since $a$ is positive, it follows that both zeros are simple, and that $R'(N)$ is negative at $N = 0$, positive at $N = \frac{a}{b}$. Therefore, the zero 0 is stable, and the zero $\frac{a}{b}$ is unstable; *all solutions with initial value* $N_0 < \frac{a}{b}$ tend to 0 as $t$ tends to infinity.

*Example 10.5.* The case $\frac{dN}{dt} = N^2 - 2N$ can be viewed in Fig. 10.16 by time reversal, where we imagine $t$ increasing from right to left along the horizontal axis. We ask you to explore this idea in Problem 10.24.

This stability of 0 is the stability of death; what we have discovered by our analysis is a very interesting and highly significant threshold effect. *Once the population size* $N_0$ *drops below the critical size* $\frac{a}{b}$, *the population tends to extinction*. This notion of a critical size is important for the preservation of a species. A species is classified *endangered* if its current size is perilously close to its critical size.

## 10.2c Two Species

We now turn to a situation involving *two* species, where one species feeds on nourishment whose supply is ample, and the other species feeds on the first species. We denote the population sizes of the two species by $N$ and $P$, $N$ denoting the *prey*, $P$ denoting the *predators*. Both $N$ and $P$ are functions of $t$, and their growth is governed by differential equations of the form $\frac{dN}{dt} = B - D$. The initial task is to choose suitable functions $B$ and $D$ describing the birth rates and death rates of each species.

We assume that the two species encounter each other by chance, at a rate proportional to the product of the size of the two populations. If we assume that the principal cause of death among the first species is due to being eaten by a member of the second species, then the death rate for $N$ is proportional to the product $NP$. We assume that the birth rate for the predator is proportional to the population size $P$, and that the portion of the young that survive is proportional to the available food supply $N$. Thus the effective birth rate is proportional to $NP$. Finally, we assume that the birth rate of the prey and the death rate of the predator are proportional to the size of their respective populations. So the equations governing the growth of these species are called the Lotka-Volterra equation, Table 10.1.

**Table 10.1** The Lotka-Volterra equation

| Species | Growth rate | | Birth rate | | Death rate |
|---|---|---|---|---|---|
| Prey | $\frac{dN}{dt}$ | $=$ | $aN$ | $-$ | $bNP$ |
| Predator | $\frac{dP}{dt}$ | $=$ | $hNP$ | $-$ | $cP$ |

In the Lotka-Volterra equations, $a,\ b,\ c$, and $h$ are all positive constants. These equations were first set down and analyzed, independently, by Volterra and by Lotka. Lotka's work on this and other population models is described in his book *Elements of Physical Biology*, originally published in 1925 and reprinted by Dover, New York, in 1956. The work of Volterra, inspired by the fluctuations in the size and composition of the catch of fish in the Adriatic, appeared in *Cahier Scientifique*, vol. VII, Gauthier-Villars, Paris, 1931, under the romantic title "Leçons sur la théorie mathématique de la lutte pour la vie" (Lessons on the mathematical theory of the fight for survival). It is reprinted in his collected works published by Accademia dei Lincei, Rome. We first give an example in which there is no predation.

*Example 10.6.* Consider the case in which there is no interaction between the predator and prey, so $b = h = 0$ and $a = 2,\ c = 3$. Then the system reads

| Species | Growth rate | | Birth rate | | Death rate |
|---|---|---|---|---|---|
| Prey | $\frac{dN}{dt}$ | $=$ | $2N$ | | |
| Predator | $\frac{dP}{dt}$ | $=$ | | $-$ | $3P$ |

The solutions are exponential:

$$N(t) = N_0 e^{2t}, \qquad P(t) = P_0 e^{-3t}.$$

Note that by properties of exponents, $(e^{2t})^{-3/2} = e^{-3t}$, so

$$\frac{P(t)}{P_0} = \left(\frac{N(t)}{N_0}\right)^{-3/2}.$$

We plot two such relations in the $(N,P)$-plane in Fig. 10.17.

Next we consider the general case that all of $a$, $b$, $c$, and $h$ are positive. The first order of business is to show that these laws of growth, and knowledge of the initial population size, are sufficient to determine the size of both populations at all future times. We formulate this as a uniqueness theorem.

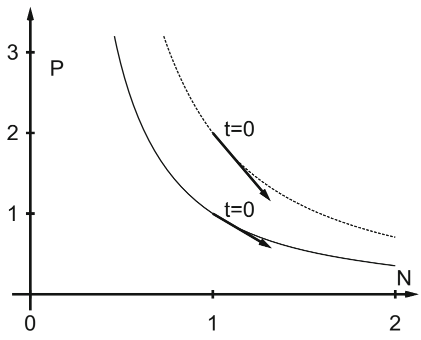


**Fig. 10.17** The $(N,P)$-plane in a case of no predation, $\frac{dN}{dt} = 2N$, $\frac{dP}{dt} = -3P$. Two time histories are shown, one starting from $(N_0,P_0) = (1,1)$ and the other from $(N_0,P_0) = (1,2)$. See Example 10.6

**Theorem 10.6. Uniqueness.** *A solution of the Lotka-Volterra equations (Table 10.1) is uniquely determined for all time by the specification of the initial values* $N_0$, $P_0$ *of* $N$ *and* $P$. *That is, if* $N, P$ *and* $n, p$ *are solutions with the same initial values, then* $N(t) = n(t)$ *and* $P(t) = p(t)$ *for all* $t$.

We ask you to fill in steps of a proof of this theorem in Problem 10.26. We omit a proof of existence of solutions, and instead investigate the properties of solutions. We take the case that both species are present, i.e., $P > 0$, $N > 0$. We divide the Lotka–Volterra equations by $N$ and $P$ respectively, obtaining

$$\frac{1}{N}\frac{dN}{dt} = a - bP,$$
$$\frac{1}{P}\frac{dP}{dt} = hN - c.$$

It follows that $P = \frac{a}{b}$ and $N = \frac{c}{h}$ are steady-state solutions. That is, if the initial values are $N_0 = \frac{c}{h}$ and $P_0 = \frac{a}{b}$, then $P = \frac{a}{b}$ and $N = \frac{c}{h}$ for all $t$. To study the non-steady-state solutions, we multiply the first equation by $hN - c$, the second equation by $bP - a$, and add. The sum of the right sides is 0, so we get the relation

$$\left(h - \frac{c}{N}\right)\frac{dN}{dt} + \left(b - \frac{a}{P}\right)\frac{dP}{dt} = 0.$$

Using the chain rule, we rewrite this relation as

$$\frac{d}{dt}\left(hN - c\log N + bP - a\log P\right) = 0.$$

We introduce the abbreviations $H$ and $K$ through (Fig. 10.18)

$$H(N) = hN - c\log N, \quad K(P) = bP - a\log P.$$

Then $\dfrac{\mathrm{d}(H+K)}{\mathrm{d}t} = 0$. We conclude from the fundamental theorem of calculus the following.

**Theorem 10.7.** *For any solution of the Lotka-Volterra equations*

$$\frac{\mathrm{d}N}{\mathrm{d}t} = aN - bNP,$$

$$\frac{\mathrm{d}P}{\mathrm{d}t} = hNP - cP,$$

*the quantity*

$$H(N) + K(P) = hN - c\log N + bP - a\log P$$

*is independent of $t$.*

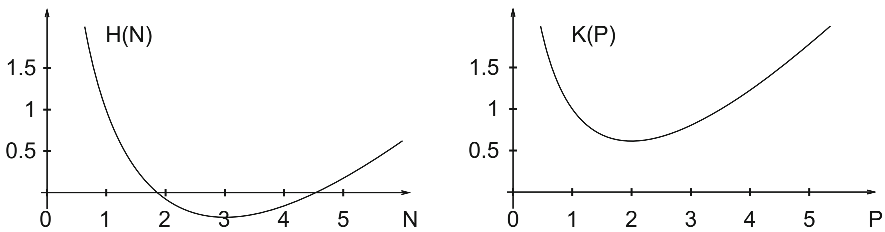


**Fig. 10.18** Graphs of the functions $H(N) = N - 3\log N$ and $K(P) = P - 2\log P$ for the system $\dfrac{\mathrm{d}N}{\mathrm{d}t} = 2N - NP$, $\dfrac{\mathrm{d}P}{\mathrm{d}t} = -3P + NP$

The constancy of the sum $H+K$ is strongly reminiscent of the law of conservation of energy in mechanics and can be used, like the law of conservation of energy, to gain qualitative and quantitative information about solutions. For that purpose, we note the following properties of functions the $H$ and $K$. Inspection of their definitions shows that the functions $H(N)$ and $K(P)$ tend to infinity as $N$ and $P$ tend to zero or to infinity. Since the functions $H(N)$ and $K(P)$ are continuous, they have minimum values, which we now locate. The derivatives of these functions are

$$\frac{\mathrm{d}H}{\mathrm{d}N} = h - \frac{c}{N}, \qquad \frac{\mathrm{d}K}{\mathrm{d}P} = b - \frac{a}{P},$$

and these are zero at

$$N_m = \frac{c}{h} \quad \text{and} \quad P_m = \frac{a}{b}.$$

Notice that these are the steady-state values for the Lotka–Volterra equations.

**Theorem 10.8.** *Consider species N and P that satisfy the Lotka-Volterra equations (Table 10.1)*

*(a) Neither species can become extinct, i.e., there is a positive lower bound for each population, throughout the whole time history.*
*(b) Neither species can proliferate ad infinitum, i.e., there is an upper bound for each population throughout its time history.*
*(c) The steady state is neutrally stable in the following sense: if the initial state $N_0$, $P_0$ is near the steady state, then $N(t)$, $P(t)$ stays near the steady state throughout the whole time history.*

*Proof.* All three results follow from the conservation law: The sum $H(N)+K(P)$ tends to infinity if either $N$ or $P$ tends to 0 or infinity. Since this is incompatible with the constancy of $H+K$, we conclude that neither $N(t)$ nor $P(t)$ can approach 0 or infinity. This proves parts (a) and (b).

For part (c), we specify the meaning of "near the steady state" by describing sets $G_s$ of points as follows. The minimum of $H(N)$ is reached at $N_m$, and the minimum of $K(P)$ is reached at $P_m$. Let $s$ be a small positive number. Denote by $G_s$ the set of points in the $(N,P)$-plane where $H(N)+K(P)$ is less than $H(N_m)+K(P_m)+s$. For $s$ small, $G_s$ is a small region around the point $(N_m,P_m)$; see Fig. 10.19. Choose $(N(0),P(0))$ in $G_s$. Then since $H(N)+K(P)$ has the same value for all $t$, $H(N)+K(P)$ is less than $H(N_m)+K(P_m)+s$ for all $t$. This shows that $(N(t),P(t))$ remains in $G_s$ for all $t$. This completes the proof of part (c). □

The next result, Theorem 10.9, is both interesting and surprising. It is suggested by some computed solutions that are shown in Fig. 10.20.

**Theorem 10.9.** *Every time history is periodic, i.e., for every solution $N(t),P(t)$ of the Lotka-Volterra equations*

$$\frac{\mathrm{d}N}{\mathrm{d}t} = aN - bNP,$$
$$\frac{\mathrm{d}P}{\mathrm{d}t} = hNP - cP,$$

*there is a time $T$ such that*

$$N(T) = N(0), \qquad P(T) = P(0).$$

The number $T$ is called the *period* of this particular time history. Different time histories have different periods. It is instructive to picture the time histories

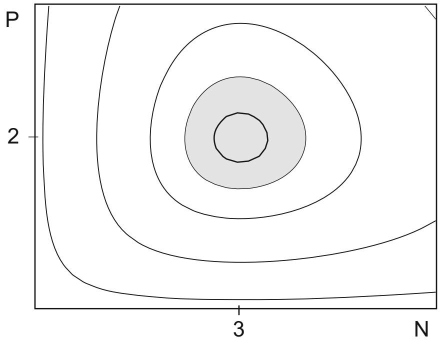


**Fig. 10.19** Stability: solutions starting in the region $G_s$ remain in that region, as shown in the proof of Theorem 10.8. The gray region here is $G_{1/10}$, where $N - 3\log N + P - 2\log P$ exceeds its minimum by less than $\frac{1}{10}$

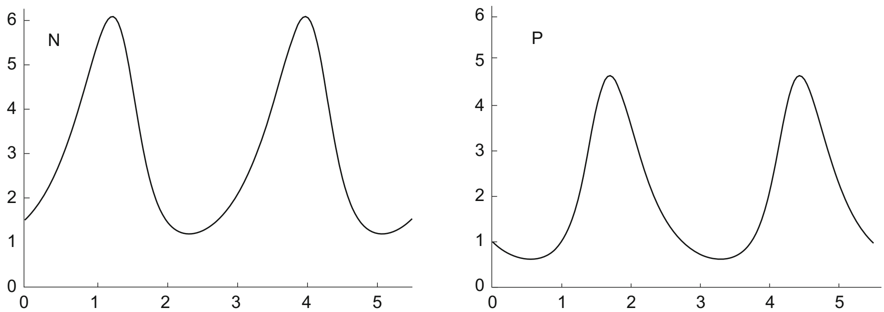


**Fig. 10.20** Computed graphs of $N(t)$ and $P(t)$ for the system $\dfrac{\mathrm{d}N}{\mathrm{d}t} = 2N - NP$, $\dfrac{\mathrm{d}P}{\mathrm{d}t} = -3P + NP$ with $N(0) = 1.5$, $P(0) = 1$. Two cycles are shown

$N(t)$, $P(t)$ graphed in Fig. 10.20 as curves in the $(P,N)$-plane, in Fig. 10.21. Periodicity means that these curves close. Compare to Fig. 10.17, where the species do not interact.

*Proof.* We write the Lotka–Volterra equations in terms of the steady-state values as

$$\frac{\mathrm{d}N}{\mathrm{d}t} = aN - bNP = bN\Big(\frac{a}{b} - P\Big) = bN(P_m - P)$$

$$\frac{\mathrm{d}P}{\mathrm{d}t} = hNP - cP = hP\Big(N - \frac{c}{h}\Big) = hP(N - N_m)$$

and conclude from the monotonicity criterion that $N$ and $P$ are increasing or decreasing functions of $t$, depending on whether the right sides are positive or negative. Therefore,

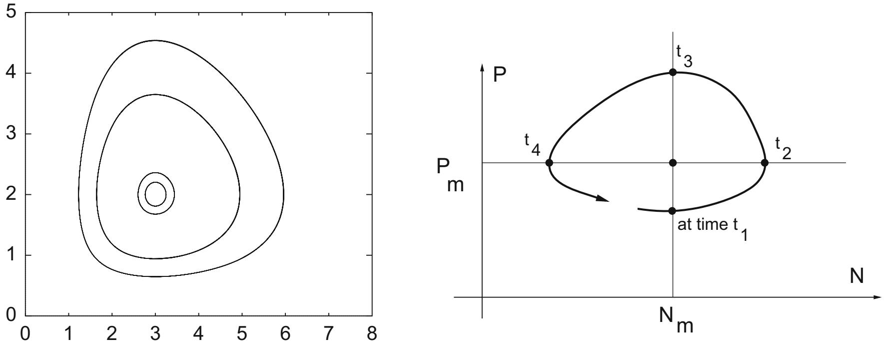


**Fig. 10.21** *Left*: four computed time histories folded into the $(N,P)$-plane, for the same system as in Fig. 10.20. *Right*: a sketch showing the notation used in the proof of Theorem 10.9

$$N(t) \begin{cases} \text{increases when} & P < P_m, \\ \text{decreases when} & P > P_m, \end{cases}$$

$$P(t) \begin{cases} \text{decreases when} & N < N_m, \\ \text{increases when} & N > N_m. \end{cases}$$

We shall use these relations to trace qualitatively the time histories of $N(t)$ and $P(t)$. See Fig. 10.21. The initial values $N_0$, $P_0$ may be chosen arbitrarily. For sake of definiteness, we choose $N_0 < N_m$, $P_0 < P_m$. Then $N$ starts to increase, $P$ to decrease, until time $t_1$, when $N$ reaches $N_m$. At time $t_1$, $P$ starts to increase, and $N$ continues to increase, until time $t_2$, when $P$ reaches $P_m$. Then $N$ starts to decrease, and $P$ continues to increase, until time $t_3$, when $N$ reaches $N_m$. Then $P$ starts to decrease, and $N$ continues to decrease, until time $t_4$, when $P$ reaches $P_m$. Then $N$ starts to increase, and $P$ continues to decrease, until time $t_5$, when $N$ reaches $N_m$.

We claim that at time $t_5$, the value of $P$ is the same as at time $t_1$. To convince you of this, we appeal to the conservation law. Denoting the values of $N$ and $P$ at $t_1$ and $t_5$ by subscripts 1 and 5, we conclude from the conservation law that

$$H(N_5) + K(P_5) = H(N_1) + K(P_1).$$

The times $t_1$ and $t_5$ were chosen so that both $N_1$ and $N_5$ are equal to $N_m$. It follows then from the conservation law that

$$K(P_5) = K(P_1).$$

The value of $K(P)$ decreases for $P < P_m$. Recall that each value of a decreasing function occurs only once. Since $t_1$ and $t_5$ were chosen so that both $P_1$ and $P_5$ are less than $P_m$, it follows therefore from $K(P_5) = K(P_1)$ that

$$P_5 = P_1.$$

It follows from the uniqueness theorem, Theorem 10.6, that the time history of $N(t)$, $P(t)$ after $t_5$ is a repetition of its time history after $t_1$. Therefore, the periodicity claim is established, the period being $T = t_5 - t_1$. □

The closed curves in the $(N,P)$-plane can be determined without solving the differential equations. According to the conservation law, on each curve, the function

$$H(N) + K(P) = hN - c\log N + bP - a\log P$$

is constant, because the function is independent of $t$. The value of the constant can be determined from the initial condition

$$H(N_0) + K(P_0) = \text{constant}.$$

As remarked before, each solution is periodic, but different solutions have different periods. It is quite remarkable that the following quantities are the same for all solutions.

**Theorem 10.10.** *The average values of P and N over a period are the same for all solutions of the Lotka-Volterra equations (Table 10.1), and they equal their steady-state values* $P_m = \dfrac{a}{b}$ *and* $N_M = \dfrac{c}{h}$*. That is,*

$$\frac{1}{T}\int_0^T N(t)\,\mathrm{d}t = N_m, \qquad \frac{1}{T}\int_0^T P(t)\,\mathrm{d}t = P_m,$$

*where T is the period of N and P.*

*Proof.* Write the Lotka–Volterra equations as

$$\frac{1}{N}\frac{\mathrm{d}N}{\mathrm{d}t} = a - bP, \qquad \frac{1}{P}\frac{\mathrm{d}P}{\mathrm{d}t} = hN - c.$$

We integrate both equations from 0 to $T$, where $T$ is the period of the solution in question. Using the chain rule, we get

$$\log N(T) - \log N(0) = \int_0^T \frac{1}{N}\frac{\mathrm{d}N}{\mathrm{d}t}\,\mathrm{d}t = \int_0^T (a - bP)\,\mathrm{d}t,$$

$$\log P(T) - \log P(0) = \int_0^T \frac{1}{P}\frac{\mathrm{d}P}{\mathrm{d}t}\,\mathrm{d}t = \int_0^T (hN - c)\,\mathrm{d}t.$$

Since $T$ is the period of $N$ and $P$, the left sides are zero. So we obtain the relations

$$0 = aT - b\int_0^T P(t)\,\mathrm{d}t, \qquad 0 = h\int_0^T N(t)\,\mathrm{d}t - cT.$$

Dividing the first equation by $bT$, the second by $hT$, we get

$$\frac{1}{T}\int_0^T P(t)\,\mathrm{d}t = \frac{a}{b}, \quad \text{and} \quad \frac{1}{T}\int_0^T N(t)\,\mathrm{d}t = \frac{c}{h}.$$

The expressions on the left are the average values of $P$ and $N$ over a period, while those on the right are their steady-state values. This concludes the proof. □

This result contains several interesting features; we mention two. The constants $a$, $b$, $c$, $h$ that determine the steady state have nothing to do with the initial values $P_0$, $N_0$ of our populations. So it follows that the *average values of P and N are independent of the initial values*. Thus, if we were to increase the initial population $N_0$, for example by stocking a lake with fish, this would not affect the average size of $N(t)$ over a period, but would only lead to different oscillations in the size of $N(t)$. We ask you in Problem 10.28 to explore cases in which stocking the lake may result in either larger or smaller oscillations in $N(t)$.

For another application, suppose we introduce fishing into the model. Assuming that the catch of predator and prey is proportional to the number of each, fishing diminishes each population at a rate proportional to the size of that population. Denoting by $f$ the constant of proportionality, we have the following modification of the equations:

$$\frac{dN}{dt} = aN - bNP - fN, \qquad \frac{dP}{dt} = -cP + hNP - fP.$$

We may write these equations in the form

$$\begin{aligned}\frac{dN}{dt} &= (a-f)N - bPN,\\ \frac{dP}{dt} &= -(c+f)P + hNP,\end{aligned}$$

and observe that they differ from the original system only in that the coefficient $a$ of $N$ in the first has been replaced by $a-f$, and the coefficient $-c$ of $P$ in the second has been replaced by $-(c+f)$. According to Theorem 10.10, the average values of $P$ and $N$ are $\frac{a-f}{b}$ and $\frac{c+f}{h}$, respectively. In other words, *increased fishing depresses the average population of predators, but increases the average population of edible fish*. During the First World War, the Italian fishing industry reported a marked increase in the ratio of sharks to edible fish. Since less fishing was done during that war than before, this observation is consistent with Volterra's surprising result.

In more complicated models, numerical computations are indispensable. They not only provide numerical answers that cannot be found in any other way, but often reveal patterns of behavior amenable to mathematical analysis. For example, the computed solutions in Fig. 10.20 suggested that solutions are periodic, and we proved that they are.

To conclude, we point out simplifications that were made in the models presented in this section:

- We have neglected to take into account the *age distribution* of the population. Since birth rate and death rate are sensitive to this, our models are deficient and would not describe correctly population changes accompanied by shifts in the population in and out of childbearing age. This phenomenon is particularly important in demography, the study of human populations.

- We have assumed that the population is *homogeneously distributed* in its environment. In many cases this is not so; the population distribution changes from location to location.

In problems such as the geographic spread of epidemics and the invasion of the territory belonging to one species by another, the interesting phenomenon is precisely the change in population as a function of time and location. Population sizes that depend on age and location as well as time are prototypes of functions of several variables. The calculus of functions of several variables is the natural language for the formulation of laws governing the growth of such populations.

You probably noticed throughout this section that we have treated population size as a differentiable function of $t$, whereas in fact, population changes by whole numbers, and so is not even a continuous function of $t$. Our defense is that these models are just models, i.e., approximations to reality, where some less-essential features are sacrificed for the sake of simplicity. The point we are making is that the *continuous is sometimes simpler than the discrete, since it allows us to use the powerful notions and tools of calculus*. Analogous simplifications are made in dealing with matter, e.g., in applying calculus to such physical quantities as pressure or density as functions of time or space. After all, according to the atomic theory of matter, these functions, too, change discontinuously.

## Problems

**10.22.** Take the case of Eq. (10.19) where $Q(N) = N^{4/3}$ and $N_0 = 1000$.

(a) Solve $Q(N) = t + c$ for $N$ as a function of $t$.
(b) Evaluate $c$.
(c) Find $R(N)$ and verify that your answer $N(t)$ is a solution to $\dfrac{dN}{dt} = R(N)$.

**10.23.** Show that the differential equation

$$\frac{dN}{dt} = \sqrt{N}$$

is satisfied by both functions $N(t) = 0$ and $N(t) = \dfrac{1}{4}t^2$ for $t \geq 0$. Since both functions are 0 at $t = 0$, does this contradict Theorem 10.3, according to which two solutions with the same initial value agree for all $t$?

**10.24.** Verify that the change of variables $n(t) = N(-t)$ converts the Verhulst model $\frac{dN}{dt} = 2N - N^2$ of Example 10.4 into the extinction model $\frac{dn}{dt} = n^2 - 2n$ of Example 10.5.

**10.25.** Consider the differential equation

$$\frac{dN}{dt} = N^2 - N$$

for values $0 < N < 1$. Derive a formula for the solutions. Use your formula to verify that if the initial value $N_0$ is between 0 and 1, then $N(t)$ tends to 0 as $t$ tends to infinity.

**10.26.** Let $p$ and $n$ be functions of $t$ that satisfy the following differential equations:

$$n' = f(p), \quad p' = g(n),$$

where the prime denotes differentiation with respect to $t$, and $f$ and $g$ are differentiable functions.

(a) Let $n_1,\ p_1$ and $n_2,\ p_2$ be two pairs of solutions. Show that the differences

$$n_1 - n_2 = m, \quad p_1 - p_2 = q$$

satisfy the inequalities

$$|m'| \le k|q|, \quad |q'| \le k|m|,$$

where $k$ is an upper bound for the absolute value of the derivatives of the functions $f$ and $g$.

(b) Deduce that

$$mm' + qq' \le 2k|m||q| \le k(m^2 + q^2).$$

(c) Define $E = \frac{1}{2}m^2 + \frac{1}{2}q^2$. Prove that $E' \le 2kE$.

(d) Deduce that $e^{-2kt}E$ is a nonincreasing function of $t$. Deduce from this that if $E(0) = 0$, then $E(t) = 0$ for all $t > 0$. Show that this implies that two solutions $n_1,\ p_1$ and $n_2,\ p_2$ that are equal at $t = 0$ are equal forever after.

**10.27.** Consider the relation between numbers $N$ and $P$ given by the equation

$$H(N) + K(P) = \text{constant}, \tag{10.22}$$

where $H$ and $K$ are convex functions, and suppose that $K(P)$ is a decreasing function of $P$ for $P$ less than some number $P_m$, and an increasing function for $P > P_m$.

(a) Show that solutions of Eq. (10.22) where $P > P_m$ can be described by expressing $P$ as a function of $N$. Show that solutions where $P < P_m$ can be described similarly. Denote these functions by $P_+(N)$ and $P_-(N)$.

(b) Let $P(N)$ be either of the functions $P_+(N)$, $P_-(N)$. Show by differentiating Eq. (10.22) twice that

$$\frac{\mathrm{d}H}{\mathrm{d}N} + \frac{\mathrm{d}K}{\mathrm{d}P}\frac{\mathrm{d}P}{\mathrm{d}N} = 0$$

and

$$\frac{\mathrm{d}^2H}{\mathrm{d}N^2} + \frac{\mathrm{d}^2K}{\mathrm{d}P^2}\left(\frac{\mathrm{d}P}{\mathrm{d}N}\right)^2 + \frac{\mathrm{d}K}{\mathrm{d}P}\frac{\mathrm{d}^2P}{\mathrm{d}N^2} = 0.$$

(c) Use the result of part (b) to express the second derivative as

$$\frac{\mathrm{d}^2P}{\mathrm{d}N^2} = -\frac{\frac{\mathrm{d}^2H}{\mathrm{d}N^2} + \frac{\mathrm{d}^2K}{\mathrm{d}P^2}\left(\frac{\mathrm{d}P}{\mathrm{d}N}\right)^2}{\frac{\mathrm{d}K}{\mathrm{d}P}}.$$

Deduce from this formula and the information given about $H$ and $K$ that $P_+(N)$ is a concave function and $P_-(N)$ is convex.

*Remark.* This confirms that the oval shapes computed in Fig. 10.21 are qualitatively correct.

**10.28.** What is the common point contained in all the ovals in Fig. 10.21?

## 10.3 Chemical Reactions

We give an elementary introduction to the theory of chemical reactions. This subject is of enormous interest to chemical engineers and to theoretical chemists. It also plays a central role in two topics that have recently been at the center of public controversy: emission by automobile engines and the deleterious effect on ozone of the accumulation of fluorocarbon compounds in the stratosphere.

In high-school chemistry, we studied the concept of a *chemical reaction*: it is the formation of one or several compounds called the *products* of the reaction out of one or several compounds or elements called *reactants*. Here is a familiar example:

$$2H_2 + O_2 \rightarrow 2H_2O$$

In words: two molecules of hydrogen and one molecule of oxygen form two molecules of water. Another example is

$$H_2 + I_2 \rightarrow 2HI.$$

In words: one molecule of hydrogen and one molecule of iodine form two molecules of hydrogen iodide.

A chemical reaction may require energy or may release energy in the form of heat; the technical terms are *endothermic* and *exothermic*. Familiar examples of reactions that release energy are the burning of coal or oil, and, more spectacularly, the burning of an explosive. In fact, the whole purpose of these chemical reactions is to garner the energy they release; the products of these reactions are uninteresting. In fact, they can be a severe nuisance, namely pollution. On the other hand, in the chemical industry, the desired commodity is the end product of the reactions or of a series of reactions.

The above description of chemical reactions deals with the phenomenon entirely in terms of its initial and final states. In this section, we shall study *time histories* of chemical reactions. This branch of chemistry is called *reaction kinetics*. An understanding of kinetics is essential in the chemical industry, because many reactions necessary in certain production processes must be set up so that they occur in the right order within specified time intervals. Similarly, the kinetics of burning must be understood in order to know what the end products are, for when released into the atmosphere, these chemicals affect global warming. The effect of fluorocarbons on depletion of ozone in the stratosphere must be judged by computing the rates at which various reactions involving these molecules occur. Last but not least, reaction kinetics is a valuable experimental tool for studying the structure of molecules.

In this section we shall describe the kinetics of fairly simple reactions, in particular those in which both reactants and products appear as gases. Furthermore, we shall assume that all components are homogeneously distributed in the vessel in which the reaction takes place. That is, we assume that the concentration, temperature, and pressure of all components at any given time are the same at all points in the vessel.

The *concentration* of a reactant measures the number of molecules of that reactant present per unit volume. Note that if two components in a vessel have the same concentration, then that vessel contains the same number of molecules of each component.

In what follows, we shall denote different molecules as well as atoms, ions, and radicals that play important roles in chemical reactions by different capital letters such as $A,B,C$, and we shall denote their concentrations by the corresponding lowercase letters such as $a,b,c$. (In the chemical literature, the concentration of molecule A would be denoted by [A].) These concentrations change with time. The *rates* at which they change, i.e., the derivatives of the concentrations with respect to time, are called the *reaction rates*. A basic principle of reaction kinetics says that the reaction rates are completely determined by pressure, temperature, and the concentrations of all components present. Mathematically, this can be expressed by specifying the rates as functions of pressure, temperature, and concentrations; then the laws of reaction kinetics take the form of differential equations:

$$\frac{\mathrm{d}a}{\mathrm{d}t} = f(a,b;T,p), \quad \frac{\mathrm{d}b}{\mathrm{d}t} = g(a,b;T,p),$$

where $f,g$ are functions specific to each particular reaction. In the simple reactions considered here, we suppress the dependence of $f$, $g$, on the temperature $T$ and the pressure $p$. The determination of these functions is the task of the theorist and experimenter. We shall start with some theoretical observations; of course, the last word belongs to the experimentalist.

The products of a chemical reaction are built out of the same basic components as the reactants, i.e., the same nuclei and the same number of electrons, but the components are now arranged differently. In other words, the chemical reaction is the process by which the rearrangement of the basic components occurs. One can think of this process of rearrangement as a continuous distortion, starting with the original component configuration and ending up with the final one. There is an *energy* associated with each transient configuration; the initial and the final states are *stable*, which means that energy is at a local minimum in those configurations. It follows that during a continuous distortion of one state into the other, energy increases until it reaches a peak and then decreases as the final configuration is reached. There are many paths along which this distortion can take place; the reaction is channeled mainly along the path where the peak value is minimum. The difference between this minimum peak value of energy and the energy of the initial configuration is called the *activation energy*. It is an energy barrier that has to be surmounted for the reaction to take place.

This description of a chemical reaction as rearrangement in one step is an oversimplification; it is applicable to only a minority of cases, called *elementary reactions*. In the great majority of cases, the reaction is *complex*, meaning that it takes place in a number of stages that lead to the formation of a number of intermediate states. The intermediate states—atoms, free radicals, and activated states—disappear when the reaction is completed. The transitions from the initial state to an intermediate state, from one intermediate state to another, and from an intermediate state to the final state are all elementary reactions. So a complex reaction may be thought of as a network of elementary reactions.

We now study the rate of an elementary reaction of form

$$A_2 + B_2 \to 2AB,$$

where one $A_2$ molecule consisting of two $A$ atoms and one $B_2$ molecule consisting of two $B$ atoms combine to form two molecules of the compound $AB$. The reaction takes place only if the two molecules collide and are energetic enough. The kinetic energies of the molecules in a vessel are not uniform, but are distributed according to a Maxwellian probability distribution (see Chap. 11). Therefore, some molecules always have sufficient kinetic energy to react when they collide, to supply the activation energy needed for the reaction. The frequency with which this happens is proportional to the *product* of the concentrations of $A_2$ and $B_2$ molecules, i.e., is equal to

$$kab, \qquad k \quad \text{a positive number.}$$

Here $a$ and $b$ denote the concentrations of $A_2$ and $B_2$, and $k$ is the *rate constant*. This is called the *law of mass action*. Denote the concentration of the reaction product $AB$ at time $t$ by $x(t)$. By the law of mass action, $x$ satisfies the differential equation

$$\frac{\mathrm{d}x}{\mathrm{d}t} = kab.$$

Denote by $a_0$ and $b_0$ the initial concentrations of $A_2$ and $B_2$. Since each molecule of $A_2$ and $B_2$ make two molecules of $AB$, the concentrations at time $t$ are

$$a(t) = a_0 - \frac{x(t)}{2}, \qquad b(t) = b_0 - \frac{x(t)}{2}.$$

Substituting this into the differential equation yields

$$\frac{\mathrm{d}x}{\mathrm{d}t} = k\left(a_0 - \frac{x}{2}\right)\left(b_0 - \frac{x}{2}\right).$$

This equation is of the form of our population model $\frac{\mathrm{d}N}{\mathrm{d}t} = R(N)$ in Eq. (10.18), with $x$ in place of $N$:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = R(x), \qquad R(x) = k\left(a_0 - \frac{x}{2}\right)\left(b_0 - \frac{x}{2}\right).$$

If the initial concentration of $AB$ is zero, then

$$x(0) = x_0 = 0.$$

Since $R(0) = ka_0b_0$ is positive, the solution $x(t)$ with initial value $x(0) = 0$ starts to increase. According to Theorem 10.4, this solution tends to the zero of $R(x)$ to the right of $x = 0$ that is nearest to $x = 0$. The zeros of $R(x)$ are $x = 2a_0$ and $x = 2b_0$. The one nearest to zero is the smaller of the two. We denote it by $x_\infty$:

$$x_\infty = \min\{2a_0, 2b_0\}.$$

It follows, then, that as $t$ tends to infinity, $x(t)$ tends to $x_\infty$. Observe that the quantity $x_\infty$ is the largest amount of $AB$ that can be made out of the given amounts $a_0$ and $b_0$ of $A_2$ and $B_2$. Therefore, our result shows that as $t$ tends to infinity, one or the other of the reactants gets completely used up.

*Example 10.7.* Consider

$$\frac{\mathrm{d}x}{\mathrm{d}t} = (1-x)(3-x).$$

The smaller root is $x_\infty = 1$. Some computed solutions are plotted in Fig. 10.22. We see that solutions starting between 0 and 3 tend to 1.

Our second observation is that although $x(t)$ tends to $x_\infty$, $x(t)$ never reaches $x_\infty$. So strictly speaking, the reaction goes on forever. However, when the difference between $x(t)$ and $x_\infty$ is so small that it makes no practical difference, the reaction is practically over. We show how to estimate the time required for the practical completion of the reaction, using a linear rate instead of the quadratic rate $R(x)$, as follows.

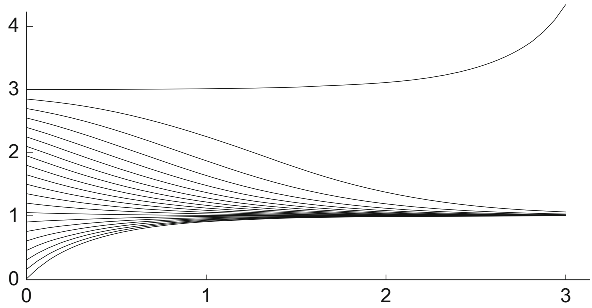


**Fig. 10.22** Graphs of several solutions to $x' = R(x) = (1-x)(3-x)$ of Example 10.7. The function $R$ is graphed in Fig. 10.23

The function $R(x) = k\left(a_0 - \frac{x}{2}\right)\left(b_0 - \frac{x}{2}\right)$ is quadratic. Therefore, its graph is a parabola. The second derivative of $R(x)$ is $\frac{1}{2}k$, a positive quantity. Therefore, as explained in Sect. 4.2b, its graph is a convex curve. This means that the points of the curve lie above its tangent lines. In particular, they lie above the tangent line at the point $x_\infty$, as illustrated in Fig. 10.23. Denote the slope of the tangent line by $n = \frac{dR}{dx}(x_\infty)$. So we deduce that

$$R(x) > n(x - x_\infty) \qquad \text{when} \quad x \neq x_\infty.$$

Since $x_\infty$ is the smaller zero of $R(x)$, we deduce from Fig. 10.23 that $n$ is negative. We set this inequality into the rate equation and get

$$\frac{dx}{dt} = k\left(a_0 - \frac{x}{2}\right)\left(b_0 - \frac{x}{2}\right) \geq n(x - x_\infty).$$

Since $x_\infty$ is a constant, the relation $\frac{dx}{dt} \geq n(x - x_\infty)$ can be further simplified if we write $\frac{d(x - x_\infty)}{dt} \geq n(x - x_\infty)$: denote the difference $x_\infty - x$ by $y$. Then the inequality can be expressed as

$$0 \geq \frac{dy}{dt} - ny.$$

Keep in mind that we want to estimate how long it takes for $x(t)$ to approach $x_\infty$, i.e., for $y(t)$ to approach 0. We multiply this inequality by $e^{-nt}$:

$$0 \geq e^{-nt}\frac{dy}{dt} - e^{-nt}ny.$$

We recognize the function on the right as the derivative of $y(t)e^{-nt}$,

$$0 \geq \frac{d}{dt}\left(y(t)e^{-nt}\right).$$

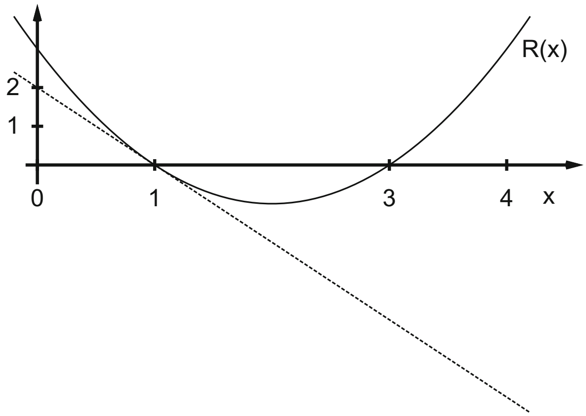


**Fig. 10.23** The graph of the convex function $R(x) = (1-x)(3-x)$ lies above its tangent line at $x_\infty = 1$: $(1-x)(3-x) \geq 2-2x$. See Example 10.7

Since the derivative is nonpositive, the function $y(t)\mathrm{e}^{-nt}$ is nonincreasing. Therefore, for $t$ positive, $y(t)\mathrm{e}^{-nt} \leq y(0)$. Multiplying both sides by $\mathrm{e}^{nt}$ gives

$$y(t) \leq y(0)\mathrm{e}^{nt}.$$

Since $n$ is a negative number, this shows that $y(t)$ tends to zero at an exponential rate.

*Example 10.8.* For the equation in Example 10.7, we have

$$\frac{\mathrm{d}x}{\mathrm{d}t} = (1-x)(3-x) \geq 2-2x,$$

$x_\infty = 1,\; n = -2$, and $y(t) = 1-x(t)$ approaches 0 at the rate

$$|1-x(t)| \leq |1-x(0)|\mathrm{e}^{-2t}.$$

We now calculate the decay rate $n$ for any quadratic reaction rate. We have

$$R(x) = \frac{k}{4}x^2 - \frac{k}{2}(a_0+b_0)x + ka_0b_0.$$

Differentiate:

$$\frac{\mathrm{d}R}{\mathrm{d}x} = \frac{k}{2}x - \frac{k}{2}(a_0+b_0).$$

Suppose $a_0$ is less than $b_0$. Then $x_\infty = 2a_0$, and so

$$n = \frac{\mathrm{d}R}{\mathrm{d}x}(2a_0) = ka_0 - \frac{k}{2}(a_0+b_0) = \frac{k}{2}(a_0-b_0).$$

Notice that when $a_0$ and $b_0$ are nearly equal, $n$ is very small. Therefore, in this case $x(t)$ approaches $x_\infty$ rather slowly. When $a_0$ and $b_0$ are equal, $n = 0$, and our argument tells us nothing about $x(t)$ approaching $x_\infty$.

We show that in case $a_0$ equals $b_0$, $x(t)$ tends to $x_\infty$, but not very fast. In this case, $R(x) = \frac{k}{4}(x - x_\infty)^2$, so the differential equation says that

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{k}{4}(x - x_\infty)^2.$$

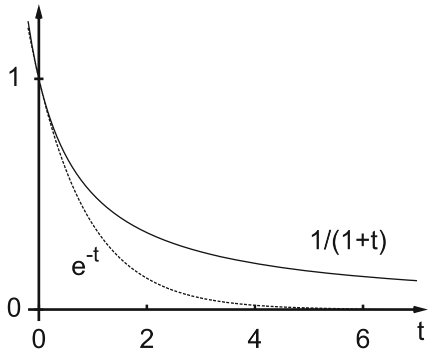


**Fig. 10.24** Graphs of the functions $\frac{1}{1+t}$ and $\mathrm{e}^{-t}$ to illustrate differences in completion rates typical for reactions starting with equal or unequal concentrations of $A_2$ and $B_2$

Introducing as before $y = x_\infty - x > 0$ as a new variable, we can rewrite the differential equation as

$$\frac{\mathrm{d}y}{\mathrm{d}t} = -\frac{k}{4}y^2.$$

Divide both sides by $y^2$. The resulting equation

$$\frac{1}{y^2}\frac{\mathrm{d}y}{\mathrm{d}t} = -\frac{k}{4}$$

can be written as

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{y}\right) = \frac{k}{4}.$$

Integrate from 0 to $t$; it follows that $\frac{1}{y(t)} = \frac{1}{y(0)} + \frac{k}{4}t$. Taking reciprocals, we get

$$y(t) = \frac{y(0)}{1 + \frac{1}{4}ky(0)t}.$$

Since $k$ and $y(0)$ are positive, $y(t)$ is defined for all $t \geq 0$, and tends to 0 as $t$ tends to infinity. This proves that $x(t) = x_\infty - y(t)$ tends to $x_\infty$, but at a very slow rate. See Fig. 10.24.

There is a good chemical reason why the reaction proceeds to completion much more slowly when the ingredients $a_0$ and $b_0$ are so perfectly balanced that they get used up simultaneously. If there is a shortage of both kinds of molecules, a collision leading to a reaction is much less likely than when there is a scarcity of only one kind of molecule but an ample supply of the other.

**Complex Reactions.** We consider a typical complex reaction such as the spontaneous decomposition of some molecule $A$, for example $N_2H_4$. The decomposition occurs in two stages; the first stage is the formation of a population of *activated molecules* $B$ followed by the spontaneous splitting of the activated molecules. The mechanism for the formation of activated molecules $B$ is through collision of two sufficiently energetic A molecules. See Fig. 10.25. The number of these collisions per unit time in a unit volume is proportional to $a^2$, the square of the concentration of $A$. There is also a reverse process of *deactivation*, due to collisions of activated and nonactivated molecules; the number of these per unit time in a unit volume is proportional to the product $ab$ of the concentrations of $A$ and $B$. There is, finally, a spontaneous decomposition of $B$ molecules into the end products $C$. The number of these decompositions per unit time in a unit volume is proportional to the concentration of $B$. If we denote the rate constant of the formation of activated molecules by $k$, that of the reverse process by $r$, and that of the spontaneous decomposition by $d$, we get the following rate equations:

$$\frac{\mathrm{d}a}{\mathrm{d}t} = -ka^2 + rab,$$
$$\frac{\mathrm{d}b}{\mathrm{d}t} = ka^2 - rab - db, \tag{10.23}$$

and $\dfrac{\mathrm{d}c}{\mathrm{d}t} = db$, which may be used after $b(t)$ is determined. It can be shown that $a(t)$ and $b(t)$ tend to zero as $t$ tends to infinity, but the argument goes beyond the scope of this chapter.

Finally, we again call attention to the striking similarity between the differential equations governing the evolution of concentrations of chemical compounds during reaction and the laws governing the evolution of animal species interacting with each other. This illustrates the universality of mathematical ideas.

## Problems

**10.29.** Consider the differential equations

(a) $\dfrac{\mathrm{d}y}{\mathrm{d}t} = -y^2,$

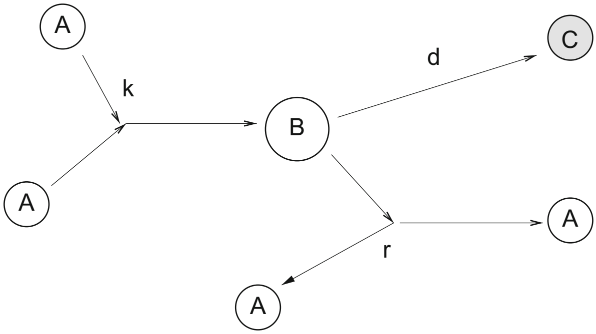


**Fig. 10.25** An illustration for a complex decomposition of $A$ to $C$. Here $k$ is the rate of formation of activated $B$ from two $A$, $r$ the reverse rate, and $d$ the rate of spontaneous decomposition of $B$ to $C$

(b) $\dfrac{\mathrm{d}y}{\mathrm{d}t} = -y.$

Take $y(0)$ positive. For which equation does $y(t)$ tend to zero faster as $t$ tends to infinity?

**10.30.** Show that if $a(t)$ and $b(t)$ are positive functions that satisfy the differential equations (10.23), than $a+b$ is a decreasing function of $t$.

**10.31.** In Eq. (10.23) let $p(t) = rb(t) - ka(t)$, so that

$$\frac{\mathrm{d}a}{\mathrm{d}t} = ap, \qquad \frac{\mathrm{d}b}{\mathrm{d}t} = -ap - db.$$

Divide the positive quadrant of the $(a,b)$-plane into two parts along the ray $rb = ka$. Show that on the side where $p(t) < 0$, $a(t)$ is a decreasing function, and that on the side where $p(t) > 0$, $b(t)$ is decreasing.

## 10.4 Numerical Solution of Differential Equations

In Sect. 10.2a we showed how integration and inverting a function can be used to find solutions of an equation linking a function $N$ to its first derivative. Such methods no longer work for finding solutions of equations involving functions and their higher derivatives, or for systems of equations relating the derivatives of several functions. Such equations can be solved only by numerical methods. In this section, we show how this is done with a very simple example.

The basis of most numerical methods for finding approximate solutions of a differential equation

$$\frac{\mathrm{d}N}{\mathrm{d}t} = N' = R(N,t), \qquad N(0) = N_0,$$

is to replace the derivative $\frac{dN}{dt}$ in the differential equation by a difference quotient

$$\frac{N(t+h)-N(t)}{h}.$$

Instead of the differential equation, we solve the *difference* equation

$$\frac{N(t+h)-N(t)}{h} = R(N(t),t), \qquad N(0) = N_0,$$

which we rearrange as

$$N(t+h) = N(t) + hR(N(t),t).$$

This is called Euler's method for approximating solutions.

We denote solutions of the difference equation by $N_h(t)$, defined for values of $t$ that are integer multiples $nh$ of $h$. One of the results of the theory of approximations by difference equations is that $N_h(nh)$ tends to $N(t)$ as $h$ tends to zero and $nh$ tends to $t$. The proof goes beyond the scope of this book. However, we will show that the method converges in the case $N' = N$. In the process, we will encounter some familiar sequences from Chap. 1.

**The Equation $N' = N$.** The solution of the differential equation

$$N'(t) = N(t), \qquad \text{with initial condition } N(0) = 1, \tag{10.24}$$

is our old friend $N(t) = e^t$. Denote by $e_h(t)$ a function that satisfies the equation obtained by replacing the derivative on the left in Eq. (10.24) by a difference quotient:

$$\frac{e_h(t+h)-e_h(t)}{h} = e_h(t), \qquad e_h(0) = 1. \tag{10.25}$$

For $h$ small, the difference quotient doesn't differ too much from the derivative, so it is reasonable to expect that the solution $e_h$ of Eq. (10.25) does not differ too much from the solution $e^t$ of Eq. (10.24). From Eq. (10.25), we can express $e_h(t+h)$ in terms of $e_h(t)$ as follows:

$$e_h(t+h) = (1+h)e_h(t). \tag{10.26}$$

Set $t = 0$ in Eq. (10.26). Since $e_h(0) = 1$, we get $e_h(h) = 1+h$. Set $t = h$ in Eq. (10.26), and we get $e_h(2h) = (1+h)e_h(h) = (1+h)^2$. Continuing in this fashion, we get for any positive integer $k$, $e_h(kh) = (1+h)^k$. When $h = \frac{1}{n}$ and $t = kh$, $k = nt$, we have (Fig. 10.26)

$$e_h(t) = e_{1/n}(k\frac{1}{n}) = \left(1+\frac{1}{n}\right)^k = \left(1+\frac{1}{n}\right)^{nt} = \left(\left(1+\frac{1}{n}\right)^n\right)^t.$$

Recall that $\left(1+\frac{1}{n}\right)^n$ is the number we called $e_n$ in Sect. 1.4. So we have

$$e_{1/n}(t) = (e_n)^t$$

for every positive rational number $t$.

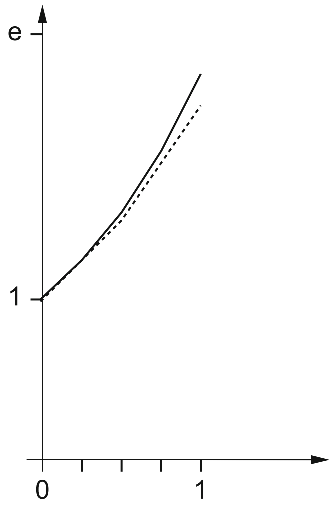


**Fig. 10.26** Approximations $e_h(kh)$ using Eq. (10.26) with $h = \frac{1}{n}$. Values are connected by segments. *Dotted line*: $n = 2$ and $h = 0.5$. *Solid line*: $n = 4$ and $h = 0.25$. The highest points are $\left(1+\frac{1}{2}\right)^2$ and $\left(1+\frac{1}{4}\right)^4$

We investigate now another way of replacing the differential equation (10.24) by a difference equation. We denote this approximate solution by $f_h(t)$. As before, we replace the derivative by the same difference quotient, but we set this equal to $f_h(t+h)$:

$$\frac{f_h(t+h) - f_h(t)}{h} = f_h(t+h), \quad f_h(0) = 1. \tag{10.27}$$

This equation can be used to express $f_h(t+h)$ in terms of $f_h(t)$:

$$f_h(t+h) = \frac{1}{1-h} f_h(t).$$

Since $f_h(0) = 1$, we deduce from this equation that $f_h(h) = \frac{1}{1-h}$. Arguing as before, where $t = kh$, we deduce that for every positive integer $k$, $f_h(kh) = \left(\frac{1}{1-h}\right)^k$.

Set $h = \frac{1}{n+1}$. Then

$$f_h(kh) = \left(\frac{1}{1-\frac{1}{n+1}}\right)^k = \left(\frac{n+1}{n}\right)^k = \left(1+\frac{1}{n}\right)^{t(n+1)}.$$

Recall that $\left(1+\frac{1}{n}\right)^{n+1}$ is the number we called $f_n$ in Sect. 1.4. So $f_{1/(n+1)}(t) = (f_n)^t$.

Next we use calculus to compare the solutions to the difference equations with the exact solution of the differential equation (10.24), for the case $t = 1$.

**Theorem 10.11.** *For every positive integer n,*

$$\left(1+\frac{1}{n}\right)^n < \mathrm{e} < \left(1+\frac{1}{n}\right)^{n+1}. \tag{10.28}$$

*Proof.* We use the mean value theorem to express the difference quotient

$$\frac{\mathrm{e}^{t+h}-\mathrm{e}^t}{h} = \mathrm{e}^c,$$

where $c$ lies between $t$ and $t+h$. Since $\mathrm{e}^t$ is an increasing function, we get that if $h > 0$, then

$$\mathrm{e}^t < \mathrm{e}^c = \frac{\mathrm{e}^{t+h}-\mathrm{e}^t}{h} < \mathrm{e}^{t+h}.$$

Multiply by $h$ and rearrange terms to get

$$(1+h)\mathrm{e}^t < \mathrm{e}^{t+h} < \frac{1}{1-h}\mathrm{e}^t. \tag{10.29}$$

Choosing first $t = 0$ and then $t = h$, we deduce from the inequality on the left in Eq. (10.29) that

$$1+h < \mathrm{e}^h, \quad (1+h)\mathrm{e}^h < \mathrm{e}^{2h}.$$

Multiply the first of these by $(1+h)$ and use the second to see that $(1+h)^2 < \mathrm{e}^{2h}$. Similarly, for every positive integer $n$, $(1+h)^n < \mathrm{e}^{nh}$. Choosing $h = \dfrac{1}{n}$, we get

$$\left(1+\frac{1}{n}\right)^n < \mathrm{e}. \tag{10.30}$$

Using similarly the inequality on the right in Eq. (10.29), we get $\mathrm{e}^h < \dfrac{1}{1-h}$. Then

$$\mathrm{e}^{2h} < \frac{1}{1-h}\mathrm{e}^h < \left(\frac{1}{1-h}\right)^2,$$

and so forth. Taking $n+1$ steps and $h = \dfrac{1}{n+1}$ gives

$$\mathrm{e} = \mathrm{e}^{(n+1)\frac{1}{n+1}} < \left(\frac{1}{1-\frac{1}{n+1}}\right)^{n+1} = \left(1+\frac{1}{n}\right)^{n+1}.$$

This completes the proof. □

In Sect. 1.4 we proved inequality (10.28) using the A-G inequality; here we have given an entirely different proof.

From Theorem 10.11, we can easily deduce that both $\left(1+\frac{1}{n}\right)^n$ and $\left(1+\frac{1}{n}\right)^{n+1}$ tend to e as $n$ tends to infinity. Take their difference:

$$\left(1+\frac{1}{n}\right)^{n+1} - \left(1+\frac{1}{n}\right)^{n} = \left(1+\frac{1}{n}\right)^{n}\frac{1}{n} < \frac{\mathrm{e}}{n}.$$

The inequality we have just proved shows that the difference tends to zero. Since e lies between these two numbers, it follows that their difference from e also tends to zero. This proves the convergence of both difference schemes (10.25) and (10.27).

**The Rate of Convergence.** In Sect. 1.4, we saw that the convergence of $e_n$ and $f_n$ to the limit e is very slow. For example, with $n = 1000$ we had $e_{1000} = 2.717\ldots$ and $f_{1000} = 2.719\ldots$, only two correct digits after the decimal point. Now we shed some light on why these approximations to e are so crude. Both $e_n$ and $f_n$ are derived from one-sided approximations to the derivative. We saw in Sect. 4.4 that for a twice differentiable function $g$, the error in using the asymmetric difference quotient

$$g'(t) - \frac{g(t+h)-g(t)}{h}$$

tends to zero with $h$, while the error in using the symmetric difference quotient

$$g'(t) - \frac{g(t+h)-g(t-h)}{2h}$$

is equal to $sh$, where $s$ tends to zero with $h$, which gives a better approximation. We can take advantage of this observation to improve the approximate solution to $y' = y$. Use the equation

$$\frac{g(t+h)-g(t-h)}{2h} = \frac{g(t+h)+g(t-h)}{2},$$

in which we use the symmetric difference quotient on the left-hand side to approximate $y'$, and on the right hand-side we average the values of $g$ at $t+h$ and $t-h$ as an approximation of $y(t)$. Solving for $g(t+h)$, we obtain $g(t+h) = \dfrac{1+h}{1-h}g(t-h)$. Replace $t$ by $t+h$ to get

$$g(t+2h) = \frac{1+h}{1-h}g(t).$$

Taking $g(0) = 1$, this gives

$$g(2h) = \frac{1+h}{1-h}, \quad g(4h) = \frac{1+h}{1-h}g(2h) = \left(\frac{1+h}{1-h}\right)^2, \quad \ldots \quad g(2nh) = \left(\frac{1+h}{1-h}\right)^n.$$

Take $h = \dfrac{1}{2n}$. We obtain

$$g(1) = \left(\frac{1+\frac{1}{2n}}{1-\frac{1}{2n}}\right)^n.$$

For $n = 10$ and 20, this gives the estimates

$$\left(\frac{1.05}{0.95}\right)^{10} = 2.7205\ldots \quad \text{and} \quad \left(\frac{1.025}{0.975}\right)^{20} = 2.7188\ldots,$$

much closer to $e$ than the numbers $e_n$ and $f_n$.

## Problems

**10.32.** Use Euler's numerical method with $h = 0.1$ to approximate the solution to

$$\frac{dy}{dt} = -1 - t, \quad y(0) = 1$$

for several steps, sufficient to estimate the time $t$ at which $y$ becomes 0. Compare to the exact solution.

**10.33.** Verify that for any differential equation $y' = f(t)$ with $y(0) = 0$, Euler's numerical method with $n$ subdivisions gives exactly the approximate integral

$$y_n = I_{\text{left}}(f, [0, nh]).$$

**10.34.** Consider the differential equation $y' = a - y$, where $a$ is a constant.

(a) Verify that the constant function $y(t) = a$ is a solution.
(b) Suppose $y$ is a solution, and for some interval of $t$ we have $y(t) > a$. Is $y$ increasing or decreasing?
(c) Consider two numerical methods. For the first, we use Euler's method to produce a sequence $y_n$ according to

$$y_{n+1} = y_n + h(a - y_n).$$

For the second, we use a method similar to that in Eq. (10.27) to produce a sequence $Y_n$ according to

$$Y_{n+1} = Y_n + h(a - Y_{n+1}).$$

Show that if some $y_n$ is equal to $a$, then $y_{n+1} = a$, and if some $Y_n$ is equal to $a$, then $Y_{n+1} = a$.

(d) Show that if some $Y_n$ is greater than $a$, then $Y_{n+1} > a$.

(e) Find a value of $h$ such that the sequence $y_n$ alternates between numbers less than $a$ and greater than $a$.