As may be expected, the omission of the constants of integration serves to simplify the procedure substantially.

The differential equation $\frac{dy}{dt} + uy = w$ in (15.12) is more general than the equation $\frac{dy}{dt} + ay = b$ in (15.4), since u and w are not necessarily constant, as are a and b. Accordingly, solution formula (15.15) is also more general than solution formula (15.5). In fact, when we set $u = a$ and $w = b$, (15.15) should reduce to (15.5). This is indeed the case. For when we have

$$u = a \qquad w = b \qquad \text{and} \qquad \int u\,dt = at \qquad \text{[constant omitted]}$$

then (15.15) becomes

$$y(t) = e^{-at}\left(A + \int be^{at}\,dt\right) = e^{-at}\left(A + \frac{b}{a}e^{at}\right) \qquad \text{[constant omitted]}$$

$$= Ae^{-at} + \frac{b}{a}$$

which is identical with (15.5).

EXERCISE 15.3

Solve the following first-order linear differential equations; if an initial condition is given, definitize the arbitrary constant:

1. $\frac{dy}{dt} + 5y = 15$
2. $\frac{dy}{dt} + 2ty = 0$
3. $\frac{dy}{dt} + 2ty = t;\ y(0) = \frac{3}{2}$
4. $\frac{dy}{dt} + t^2y = 5t^2;\ y(0) = 6$
5. $2\frac{dy}{dt} + 12y + 2e^t = 0;\ y(0) = \frac{6}{7}$
6. $\frac{dy}{dt} + y = t$

15.4 Exact Differential Equations

We shall now introduce the concept of exact differential equations and use the solution method pertaining thereto to obtain the solution formula (15.15) previously cited for the differential equation (15.12). Even though our immediate purpose is to use it to solve a *linear* differential equation, an exact differential equation can be either linear or nonlinear by itself.

Exact Differential Equations

Given a function of two variables $F(y, t)$, its total differential is

$$dF(y, t) = \frac{\partial F}{\partial y}\,dy + \frac{\partial F}{\partial t}\,dt$$

When this differential is set equal to zero, the resulting equation

$$\frac{\partial F}{\partial y}\,dy + \frac{\partial F}{\partial t}\,dt = 0$$

is known as an *exact differential equation*, because its left side is exactly the differential of the function $F(y, t)$. For instance, given

$$F(y, t) = y^2 t + k \qquad (k \text{ a constant})$$

the total differential is

$$dF = 2yt\,dy + y^2\,dt$$

thus the differential equation

$$2yt\,dy + y^2\,dt = 0 \qquad \text{or} \qquad \frac{dy}{dt} + \frac{y^2}{2yt} = 0 \tag{15.16}$$

is exact.

In general, a differential equation

$$M\,dy + N\,dt = 0 \tag{15.17}$$

is exact if and only if there exists a function $F(y, t)$ such that $M = \partial F/\partial y$ and $N = \partial F/\partial t$. By Young's theorem, which states that $\partial^2 F/\partial t\,\partial y = \partial^2 F/\partial y\,\partial t$, however, we can also state that (15.17) is exact if and only if

$$\frac{\partial M}{\partial t} = \frac{\partial N}{\partial y} \tag{15.18}$$

This last equation gives us a simple test for the exactness of a differential equation. Applied to (15.16), where $M = 2yt$ and $N = y^2$, this test yields $\partial M/\partial t = 2y = \partial N/\partial y$; thus the exactness of the said differential equation is duly verified.

Note that no restrictions have been placed on the terms M and N with regard to the manner in which the variable y occurs. Thus an exact differential equation may very well be *nonlinear* (in y). Nevertheless, it will always be of the first order and the first degree.

Being exact, the differential equation merely says

$$dF(y, t) = 0$$

Thus its general solution should clearly be in the form

$$F(y, t) = c$$

To solve an exact differential equation is basically, therefore, to search for the (primitive) function $F(y, t)$ and then set it equal to an arbitrary constant. Let us outline a method of finding this for the equation $M\,dy + N\,dt = 0$.

Method of Solution

To begin with, since $M = \partial F/\partial y$, the function F must contain the integral of M with respect to the variable y; hence we can write out a preliminary result—in a yet indeterminate form—as follows:

$$F(y, t) = \int M\,dy + \psi(t) \tag{15.19}$$

Here M, a *partial* derivative, is to be integrated with respect to y only; that is, t is to be treated as a constant in the integration process, just as it was treated as a constant in the partial differentiation of $F(y, t)$ that resulted in $M = \partial F/\partial y$.[†] Since, in differentiating $F(y, t)$ partially with respect to y, any additive term containing only the variable t and/or some constants (but with no y) would drop out, we must now take care to reinstate such terms in the integration process. This explains why we have introduced in (15.19) a general term $\psi(t)$, which, though not exactly the same as a constant of integration, has a precisely identical role to play as the latter. It is relatively easy to get $\int M\,dy$; but how do we pin down the exact form of this $\psi(t)$ term?

The trick is to utilize the fact that $N = \partial F/\partial t$. But the procedure is best explained with the help of specific examples.

Example 1 Solve the exact differential equation

$$2yt\,dy + y^2\,dt = 0 \qquad \text{[reproduced from (15.16)]}$$

In this equation, we have

$$M = 2yt \qquad \text{and} \qquad N = y^2$$

STEP i By (15.19), we can first write the preliminary result

$$F(y, t) = \int 2yt\,dy + \psi(t) = y^2t + \psi(t)$$

Note that we have omitted the constant of integration, because it can automatically be merged into the expression $\psi(t)$.

STEP ii If we differentiate the result from Step i partially with respect to t, we can obtain

$$\frac{\partial F}{\partial t} = y^2 + \psi'(t)$$

But since $N = \partial F/\partial t$, we can equate $N = y^2$ and $\partial F/\partial t = y^2 + \psi'(t)$, to get

$$\psi'(t) = 0$$

STEP iii Integration of the last result gives us

$$\psi(t) = \int \psi'(t)\,dt = \int 0\,dt = k$$

and now we have a specific form of $\psi(t)$. It happens in the present case that $\psi(t)$ is simply a constant; more generally, it can be a nonconstant function of t.

STEP iv The results of Steps i and iii can be combined to yield

$$F(y, t) = y^2t + k$$

The solution of the exact differential equation should then be $F(y, t) = c$. But since the constant k can be merged into c, we may write the solution simply as

$$y^2t = c \qquad \text{or} \qquad y(t) = ct^{-1/2}$$

where c is arbitrary.

[†] Some writers employ the operator symbol $\int(\cdots)\,\partial y$ to emphasize that the integration is with respect to y only. We shall still use the symbol $\int(\cdots)\,dy$ here, since there is little possibility of confusion.

Example 2 Solve the equation $(t+2y)\,dy+(y+3t^2)\,dt=0$. First let us check whether this is an exact differential equation. Setting $M=t+2y$ and $N=y+3t^2$, we find that $\partial M/\partial t = 1 = \partial N/\partial y$. Thus the equation passes the exactness test. To find its solution, we again follow the procedure outlined in Example 1.

Step i Apply (15.19) and write

$$F(y,t)=\int (t+2y)\,dy+\psi(t)=yt+y^2+\psi(t) \qquad \text{[constant merged into } \psi(t)\text{]}$$

Step ii Differentiate this result with respect to t, to get

$$\frac{\partial F}{\partial t}=y+\psi'(t)$$

Then, equating this to $N=y+3t^2$, we find that

$$\psi'(t)=3t^2$$

Step iii Integrate this last result to get

$$\psi(t)=\int 3t^2\,dt=t^3 \qquad \text{[constant may be omitted]}$$

Step iv Combine the results of Steps i and iii to get the complete form of the function $F(y,t)$:

$$F(y,t)=yt+y^2+t^3$$

which implies that the solution of the given differential equation is

$$yt+y^2+t^3=c$$

You should verify that setting the total differential of this equation equal to zero will indeed produce the given differential equation.

This four-step procedure can be used to solve any exact differential equation. Interestingly, it may even be applicable when the given equation is *not* exact. To see this, however, we must first introduce the concept of integrating factor.

Integrating Factor

Sometimes an inexact differential equation can be made exact by multiplying every term of the equation by a particular common factor. Such a factor is called an *integrating factor.*

Example 3 The differential equation

$$2t\,dy+y\,dt=0$$

is not exact, because it does not satisfy (15.18):

$$\frac{\partial M}{\partial t}=\frac{\partial}{\partial t}(2t)=2\neq\frac{\partial N}{\partial y}=\frac{\partial}{\partial y}(y)=1$$

However, if we multiply each term by y, the given equation will turn into (15.16), which has been established to be exact. Thus y is an integrating factor for the differential equation in the present example.

When an integrating factor can be found for an inexact differential equation, it is always possible to render it exact, and then the four-step solution procedure can be readily put to use.

Solution of First-Order Linear Differential Equations

The general first-order linear differential equation

$$\frac{dy}{dt} + uy = w$$

which, in the format of (15.17), can be expressed as

$$dy + (uy - w)\,dt = 0 \tag{15.20}$$

has the integrating factor

$$e^{\int u\,dt} \equiv \exp\left(\int u\,dt\right)$$

This integrating factor, whose form is by no means intuitively obvious, can be "discovered" as follows. Let I be the (yet unknown) integrating factor. Multiplication of (15.20) through by I should convert it into an exact differential equation

$$\underbrace{I}_{M}\,dy + \underbrace{I(uy - w)}_{N}\,dt = 0 \tag{15.20'}$$

The exactness test dictates that $\partial M/\partial t = \partial N/\partial y$. Visual inspection of the M and N expressions suggests that, since M consists of I only, and since u and w are functions of t alone, the exactness test will reduce to a very simple condition if I is also a function of t alone. For then the test $\partial M/\partial t = \partial N/\partial y$ becomes

$$\frac{dI}{dt} = Iu \qquad \text{or} \qquad \frac{dI/dt}{I} = u$$

Thus the special form $I = I(t)$ can indeed work, provided it has a rate of growth equal to u, or more explicitly, $u(t)$. Accordingly, $I(t)$ should take the specific form

$$I(t) = Ae^{\int u\,dt} \qquad \text{[cf. (15.13) and (15.14)]}$$

As can be easily verified, however, the constant A can be set equal to 1 without affecting the ability of $I(t)$ to meet the exactness test. Thus we can use the simpler form $e^{\int u\,dt}$ as the integrating factor.

Substitution of this integrating factor into (15.20′) yields the exact differential equation

$$e^{\int u\,dt}\,dy + e^{\int u\,dt}(uy - w)\,dt = 0 \tag{15.20''}$$

which can then be solved by the four-step procedure.

Step i First, we apply (15.19) to obtain

$$F(y, t) = \int e^{\int u\,dt}\,dy + \psi(t) = ye^{\int u\,dt} + \psi(t)$$

The result of integration emerges in this simple form because the integrand is independent of the variable y.

STEP ii Next, we differentiate the result from Step i with respect to t, to get

$$\frac{\partial F}{\partial t} = yue^{\int u\,dt} + \psi'(t) \qquad \text{[chain rule]}$$

And, since this can be equated to $N = e^{\int u\,dt}(uy - w)$, we have

$$\psi'(t) = -we^{\int u\,dt}$$

STEP iii Straight integration now yields

$$\psi(t) = -\int we^{\int u\,dt}\,dt$$

Inasmuch as the functions $u = u(t)$ and $w = w(t)$ have not been given specific forms, nothing further can be done about this integral, and we must be contented with this rather general expression for $\psi(t)$.

STEP iv Substituting this $\psi(t)$ expression into the result of Step i, we find that

$$F(y, t) = ye^{\int u\,dt} - \int we^{\int u\,dt}\,dt$$

So the general solution of the exact differential equation (15.20″)—and of the equivalent, though inexact, first-order linear differential equation (15.20)—is

$$ye^{\int u\,dt} - \int we^{\int u\,dt}\,dt = c$$

Upon rearrangement and substitution of the (arbitrary constant) symbol c by A, this can be written as

$$y(t) = e^{-\int u\,dt}\left(A + \int we^{\int u\,dt}\,dt\right) \qquad \textbf{(15.21)}$$

which is exactly the result given earlier in (15.15).

EXERCISE 15.4

1. Verify that each of the following differential equations is exact, and solve by the four-step procedure:
   (*a*) $2yt^3\,dy + 3y^2t^2\,dt = 0$
   (*b*) $3y^2t\,dy + (y^3 + 2t)\,dt = 0$
   (*c*) $t(1 + 2y)\,dy + y(1 + y)\,dt = 0$
   (*d*) $\dfrac{dy}{dt} + \dfrac{2y^4t + 3t^2}{4y^3t^2} = 0$ [*Hint:* First convert to the form of (15.17).]
2. Are the following differential equations exact? If not, try t, y, and y^2 as possible integrating factors.
   (*a*) $2(t^3 + 1)\,dy + 3yt^2\,dt = 0$
   (*b*) $4y^3t\,dy + (2y^4 + 3t)\,dt = 0$
3. By applying the four-step procedure to the general exact differential equation $M\,dy + N\,dt = 0$, derive the following formula for the general solution of an exact differential equation:

$$\int M\,dy + \int N\,dt - \int\left(\frac{\partial}{\partial t}\int M\,dy\right)dt = c$$

15.5 Nonlinear Differential Equations of the First Order and First Degree

In a *linear* differential equation, we restrict to the *first degree* not only the derivative dy/dt, but also the dependent variable y, and we do not allow the product $y(dy/dt)$ to appear. When y appears in a power higher than one, the equation becomes *nonlinear* even if it only contains the derivative dy/dt in the first degree. In general, an equation in the form

$$f(y, t)\,dy + g(y, t)\,dt = 0 \tag{15.22}$$

or

$$\frac{dy}{dt} = h(y, t) \tag{15.22'}$$

where there is no restriction on the powers of y and t, constitutes a first-order first-degree nonlinear differential equation because dy/dt is a first-order derivative in the first power. Certain varieties of such equations can be solved with relative ease by more or less routine procedures. We shall briefly discuss three cases.

Exact Differential Equations

The first is the now-familiar case of exact differential equations. As was pointed out earlier, the y variable can appear in an exact equation in a high power, as in (15.16) $2yt\,dy + y^2\,dt = 0$—which you should compare with (15.22). True, the cancellation of the common factor y from both terms on the left will reduce the equation to a linear form, but the exactness property will be lost in that event. As an *exact* differential equation, therefore, it must be regarded as nonlinear.

Since the solution method for exact differential equations has already been discussed, no further comment is necessary here.

Separable Variables

The differential equation in (15.22)

$$f(y, t)\,dy + g(y, t)\,dt = 0$$

may happen to possess the convenient property that the function f is in the variable y alone, while the function g involves only the variable t, so that the equation reduces to the special form

$$f(y)\,dy + g(t)\,dt = 0 \tag{15.23}$$

In such an event, the variables are said to be *separable*, because the terms involving y—consolidated into $f(y)$ can be mathematically separated from the terms involving t, which are collected under $g(t)$. To solve this special type of equation, only simple integration techniques are required.

Example 1 Solve the equation $3y^2\,dy - t\,dt = 0$. First let us rewrite the equation as

$$3y^2\,dy = t\,dt$$

Integrating the two sides (each of which is a differential) and equating the results, we get

$$\int 3y^2\,dy = \int t\,dt \qquad \text{or} \qquad y^3 + c_1 = \frac{1}{2}t^2 + c_2$$

Thus the general solution can be written as

$$y^3 = \frac{1}{2}t^2 + c \qquad \text{or} \qquad y(t) = \left(\frac{1}{2}t^2 + c\right)^{1/3}$$

The notable point here is that the integration of each term is performed with respect to a different variable; it is this which makes the separable-variable equation comparatively easy to handle.

Example 2 Solve the equation $2t\,dy + y\,dt = 0$. At first glance, this differential equation does not seem to belong in this spot, because it fails to conform to the general form of (15.23). To be specific, the coefficients of dy and dt are seen to involve the "wrong" variables. However, a simple transformation—dividing through by $2yt$ ($\neq 0$)—will reduce the equation to the separable-variable form

$$\frac{1}{y}\,dy + \frac{1}{2t}\,dt = 0$$

From our experience with Example 1, we can work toward the solution (without first transposing a term) as follows:[†]

$$\int \frac{1}{y}\,dy + \int \frac{1}{2t}\,dt = c$$

so
$$\ln y + \frac{1}{2}\ln t = c \qquad \text{or} \qquad \ln(yt^{1/2}) = c$$

Thus the solution is

$$yt^{1/2} = e^c = k \qquad \text{or} \qquad y(t) = kt^{-1/2}$$

where k is an arbitrary constant, as are the symbols c and A employed elsewhere.

Note that, instead of solving the equation in Example 2 as we did, we could also have transformed it first into an exact differential equation (by the integrating factor y) and then solved it as such. The solution, already given in Example 1 of Sec. 15.4, must of course be identical with the one just obtained by separation of variables. The point is that a given differential equation can often be solvable in more than one way, and therefore one may have a choice of the method to be used. In other cases, a differential equation that is not amenable to a particular method may nonetheless become so after an appropriate transformation.

Equations Reducible to the Linear Form

If the differential equation $dy/dt = h(y, t)$ happens to take the specific nonlinear form

$$\frac{dy}{dt} + Ry = Ty^m \tag{15.24}$$

where R and T are two functions of t, and m is any number other than 0 and 1 (what if $m = 0$ or $m = 1$?), then the equation—referred to as a *Bernoulli equation*—can always be reduced to a linear differential equation and be solved as such.

[†] In the integration result, we should, strictly speaking, have written $\ln|y|$ and $\frac{1}{2}\ln|t|$. If y and t can be assumed to be positive, as is appropriate in the majority of economic contexts, then the result given in the text will occur.

The reduction procedure is relatively simple. First, we can divide (15.24) by y^m, to get

$$y^{-m}\frac{dy}{dt} + Ry^{1-m} = T$$

If we adopt a shorthand variable z as follows:

$$z = y^{1-m} \qquad \left[\text{so that } \frac{dz}{dt} = \frac{dz}{dy}\frac{dy}{dt} = (1-m)y^{-m}\frac{dy}{dt}\right]$$

then the preceding equation can be written as

$$\frac{1}{1-m}\frac{dz}{dt} + Rz = T$$

Moreover, after multiplying through by $(1-m)\,dt$ and rearranging, we can transform the equation into

$$dz + [(1-m)Rz - (1-m)T]\,dt = 0 \qquad \textbf{(15.24')}$$

This is seen to be a first-order linear differential equation of the form (15.20), in which the variable z has taken the place of y.

Clearly, we can apply formula (15.21) to find its solution $z(t)$. Then, as a final step, we can translate z back to y by reverse substitution.

Example 3

Solve the equation $dy/dt + ty = 3ty^2$. This is a Bernoulli equation, with $m = 2$ (giving us $z = y^{1-m} = y^{-1}$), $R = t$, and $T = 3t$. Thus, by (15.24'), we can write the linearized differential equation as

$$dz + (-tz + 3t)\,dt = 0$$

By applying formula (15.21), the solution can be found to be

$$z(t) = A\exp\left(\tfrac{1}{2}t^2\right) + 3$$

(As an exercise, trace out the steps leading to this solution.)

Since our primary interest lies in the solution $y(t)$ rather than $z(t)$, we must perform a reverse transformation using the equation $z = y^{-1}$, or $y = z^{-1}$. By taking the reciprocal of $z(t)$, therefore, we get

$$y(t) = \frac{1}{A\exp\left(\tfrac{1}{2}t^2\right) + 3}$$

as the desired solution. This is a general solution, because an arbitrary constant A is present.

Example 4

Solve the equation $dy/dt + (1/t)y = y^3$. Here, we have $m = 3$ (thus $z = y^{-2}$), $R = 1/t$, and $T = 1$; thus the equation can be linearized into the form

$$dz + \left(\frac{-2}{t}z + 2\right)dt = 0$$

As you can verify, by the use of formula (15.21), the solution of this differential equation is

$$z(t) = At^2 + 2t$$

It then follows, by the reverse transformation $y = z^{-1/2}$, that the general solution in the original variable is to be written as

$$y(t) = (At^2 + 2t)^{-1/2}$$

As an exercise, check the validity of the solutions of these last two examples by differentiation.

EXERCISE 15.5

1. Determine, for each of the following, (1) whether the variables are separable and (2) whether the equation is linear or else can be linearized:

   (*a*) $2t\,dy + 2y\,dt = 0$ (*c*) $\frac{dy}{dt} = -\frac{t}{y}$

   (*b*) $\frac{y}{y+t}\,dy + \frac{2t}{y+t}\,dt = 0$ (*d*) $\frac{dy}{dt} = 3y^2t$

2. Solve (*a*) and (*b*) in Prob. 1 by separation of variables, taking y and t to be positive. Check your answers by differentiation.
3. Solve (*c*) in Prob. 1 as a separable-variable equation and, also, as a Bernoulli equation.
4. Solve (*d*) in Prob. 1 as a separable-variable equation and, also, as a Bernoulli equation.
5. Verify the correctness of the intermediate solution $z(t) = At^2 + 2t$ in Example 4 by showing that its derivative dz/dt is consistent with the linearized differential equation.

15.6 The Qualitative-Graphic Approach

The several cases of nonlinear differential equations previously discussed (exact differential equations, separable-variable equations, and Bernoulli equations) have all been solved *quantitatively*. That is, we have in every case sought and found a time path $y(t)$ which, for each value of t, tells the specific corresponding value of the variable y.

At times, we may not be able to find a quantitative solution from a given differential equation. Yet, in such cases, it may nonetheless be possible to ascertain the *qualitative* properties of the time path—primarily, whether $y(t)$ converges—by directly observing the differential equation itself or by analyzing its graph. Even when quantitative solutions are available, moreover, we may still employ the techniques of qualitative analysis if the qualitative aspect of the time path is our principal or exclusive concern.

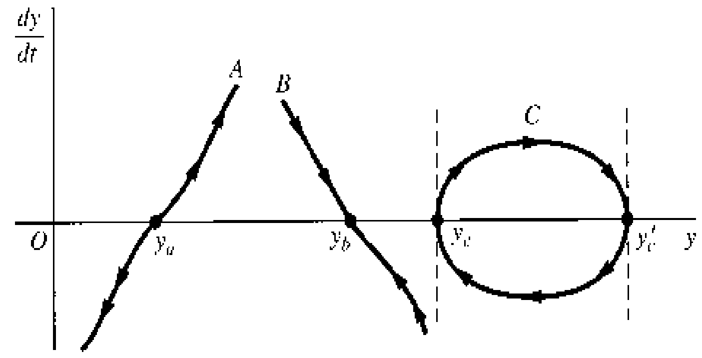
The Phase Diagram

Given a first-order differential equation in the general form

$$\frac{dy}{dt} = f(y)$$

either linear or nonlinear in the variable y, we can plot dy/dt against y as in Fig. 15.3. Such a geometric representation, feasible whenever dy/dt is a function of y alone, is called a *phase diagram*, and the graph representing the function f, a *phase line*. (A differential equation of this form—in which the time variable t does not appear as a separate argument of

FIGURE 15.3



the function *f*—is said to be an *autonomous* differential equation.) Once a phase line is known, its configuration will impart significant qualitative information regarding the time path $y(t)$. The clue to this lies in the following two general remarks:

1. Anywhere *above* the horizontal axis (where $dy/dt > 0$), y must be increasing over time and, as far as the y axis is concerned, must be moving from left to right. By analogous reasoning, any point *below* the horizontal axis must be associated with a leftward movement in the variable y, because the negativity of dy/dt means that y decreases over time. These directional tendencies explain why the arrowheads on the illustrative phase lines in Fig. 15.3 are drawn as they are. Above the horizontal axis, the arrows are uniformly pointed toward the right—toward the northeast or southeast or due east, as the case may be. The opposite is true below the y axis. Moreover, these results are independent of the algebraic sign of y; even if phase line A (or any other) is transplanted to the left of the vertical axis, the direction of the arrows will not be affected.
2. An equilibrium level of y—in the intertemporal sense of the term if it exists, can occur only on the horizontal axis, where $dy/dt = 0$ (y stationary over time). To find an equilibrium, therefore, it is necessary only to consider the intersection of the phase line with the y axis.† To test the dynamic stability of equilibrium, on the other hand, we should also check whether, regardless of the initial position of y, the phase line will always guide it toward the equilibrium position at the said intersection.

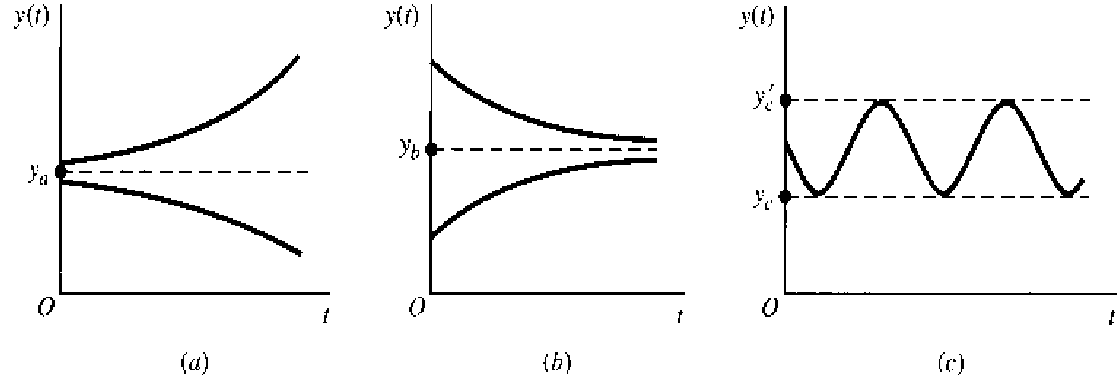
Types of Time Path

On the basis of the preceding general remarks, we may observe three different types of time path from the illustrative phase lines in Fig. 15.3.

Phase line A has an equilibrium at point y_a; but *above* as well as *below* that point, the arrowheads consistently lead away from equilibrium. Thus, although equilibrium can be attained if it happens that $y(0) = y_a$, the more usual case of $y(0) \neq y_a$ will result in y being ever-increasing [if $y(0) > y_a$] or ever-decreasing [if $y(0) < y_a$]. Besides, in this case the deviation of y from y_a tends to grow at an increasing pace because, as we follow the arrowheads on the phase line, we deviate farther from the y axis, thereby encountering ever-increasing numerical values of dy/dt as well. The time path $y(t)$ implied by phase line A can therefore be represented by the curves shown in Fig. 15.4a, where y is plotted against t (rather than dy/dt against y). The equilibrium y_a is dynamically unstable.

† However, not all intersections represent equilibrium positions. We shall see this when we discuss phase line C in Fig. 15.3.

FIGURE 15.4



In contrast, phase line B implies a stable equilibrium at y_b. If $y(0) = y_b$, equilibrium prevails at once. But the important feature of phase line B is that, even if $y(0) \neq y_b$, the movement along the phase line will guide y toward the level of y_b. The time path $y(t)$ corresponding to this type of phase line should therefore be of the form shown in Fig. 15.4*b*, which is reminiscent of the dynamic market model.

The preceding discussion suggests that, in general, it is the slope of the phase line at its intersection point which holds the key to the dynamic stability of equilibrium or the convergence of the time path. A (finite) *positive* slope, such as at point y_a, makes for dynamic *instability;* whereas a (finite) *negative* slope, such as at y_b, implies dynamic *stability.*

This generalization can help us to draw qualitative inferences about given differential equations without even plotting their phase lines. Take the linear differential equation in (15.4), for instance:

$$\frac{dy}{dt} + ay = b \qquad \text{or} \qquad \frac{dy}{dt} = -ay + b$$

Since the phase line will obviously have the (constant) slope $-a$, here assumed nonzero, we may immediately infer (without drawing the line) that

$$a \gtrless 0 \quad \Leftrightarrow \quad y(t) \begin{Bmatrix} \text{converges to} \\ \text{diverges from} \end{Bmatrix} \text{equilibrium}$$

As we may expect, this result coincides perfectly with what the quantitative solution of this equation tells us:

$$y(t) = \left[y(0) - \frac{b}{a}\right]e^{-at} + \frac{b}{a} \qquad \text{[from (15.5')]}$$

We have learned that, starting from a nonequilibrium position, the convergence of $y(t)$ hinges on the prospect that $e^{-at} \to 0$ as $t \to \infty$. This can happen if and only if $a > 0$; if $a < 0$, then $e^{-at} \to \infty$ as $t \to \infty$, and $y(t)$ cannot converge. Thus, our conclusion is one and the same, whether it is arrived at quantitatively or qualitatively.

It remains to discuss phase line C, which, being a closed loop sitting across the horizontal axis, does not qualify as a *function* but shows instead a *relation* between dy/dt and y.[†] The interesting new element that emerges in this case is the possibility of a periodically fluctuating time path. The way that phase line C is drawn, we shall find y fluctuating between the two values y_c and y_c' in a perpetual motion. In order to generate the periodic

[†] This can arise from a second-degree differential equation $(dy/dt)^2 = f(y)$.

fluctuation, the loop must, of course, straddle the horizontal axis in such a manner that dy/dt can alternately be positive and negative. Besides, at the two intersection points y_c and y_c', the phase line should have an infinite slope; otherwise the intersection will resemble either y_a or y_b, neither of which permits a continual flow of arrowheads. The type of time path $y(t)$ corresponding to this looped phase line is illustrated in Fig. 15.4c. Note that, whenever $y(t)$ hits the upper bound y_c' or the lower bound y_c, we have $dy/dt = 0$ (local extrema); but these values certainly do not represent equilibrium values of y. In terms of Fig. 15.3, this means that not all intersections between a phase line and the y axis are equilibrium positions.

In sum, for the study of the dynamic stability of equilibrium (or the convergence of the time path), one has the alternative either of finding the time path itself or else of simply drawing the inference from its phase line. We shall illustrate the application of the latter approach with the Solow growth model. Henceforth, we shall denote the intertemporal equilibrium value of y by $\bar{y}$, as distinct from y^*.

EXERCISE 15.6

1. Plot the phase line for each of the following, and discuss its qualitative implications:
   (a) $\dfrac{dy}{dt} = y - 7$ (c) $\dfrac{dy}{dt} = 4 - \dfrac{y}{2}$
   (b) $\dfrac{dy}{dt} = 1 - 5y$ (d) $\dfrac{dy}{dt} = 9y - 11$
2. Plot the phase line for each of the following and interpret:
   (a) $\dfrac{dy}{dt} = (y+1)^2 - 16 \qquad (y \geq 0)$
   (b) $\dfrac{dy}{dt} = \dfrac{1}{2}y - y^2 \qquad (y \geq 0)$
3. Given $dy/dt = (y-3)(y-5) = y^2 - 8y + 15$:
   (a) Deduce that there are two possible equilibrium levels of y, one at $y = 3$ and the other at $y = 5$.
   (b) Find the sign of $\dfrac{d}{dy}\left(\dfrac{dy}{dt}\right)$ at $y = 3$ and $y = 5$, respectively. What can you infer from these?

15.7 Solow Growth Model

The growth model of Professor Robert Solow,[†] a Nobel laureate, is purported to show, among other things, that the razor's-edge growth path of the Domar model is primarily a result of the particular production-function assumption adopted therein and that, under alternative circumstances, the need for delicate balancing may not arise.

The Framework

In the Domar model, output is explicitly stated as a function of capital alone: $\kappa = \rho K$ (the productive capacity, or potential output, is a constant multiple of the stock of capital). The

[†] Robert M. Solow, "A Contribution to the Theory of Economic Growth," *Quarterly Journal of Economics*, February 1956, pp. 65–94.

absence of a labor input in the production function carries the implication that labor is always combined with capital in a *fixed* proportion, so that it is feasible to consider explicitly only one of these factors of production. Solow, in contrast, seeks to analyze the case where capital and labor can be combined in *varying* proportions. Thus his production function appears in the form

$$Q = f(K, L) \qquad (K, L > 0)$$

where Q is output (net of depreciation), K is capital, and L is labor—all being used in the *macro* sense. It is assumed that f_K and f_L are positive (positive marginal products), and f_{KK} and f_{LL} are negative (diminishing returns to each input). Furthermore, the production function f is taken to be linearly homogeneous (constant returns to scale). Consequently, it is possible to write

$$Q = Lf\left(\frac{K}{L}, 1\right) = L\phi(k) \qquad \text{where } k \equiv \frac{K}{L} \tag{15.25}$$

In view of the assumed signs of f_K and f_{KK}, the newly introduced ϕ function (which, be it noted, has only a single argument, k) must be characterized by a positive first derivative and a negative second derivative. To verify this claim, we first recall from (12.49) that

$$f_K \equiv \text{MPP}_K = \phi'(k)$$

hence $f_K > 0$ automatically means $\phi'(k) > 0$. Then, since

$$f_{KK} = \frac{\partial}{\partial K}\phi'(k) = \frac{d\phi'(k)}{dk}\frac{\partial k}{\partial K} = \phi''(k)\frac{1}{L} \qquad \text{[see (12.48)]}$$

the assumption $f_{KK} < 0$ leads directly to the result $\phi''(k) < 0$. Thus the ϕ function—which, according to (12.46), gives the APP_L for every capital–labor ratio—is one that increases with k at a decreasing rate.

Given that Q depends on K and L, it is necessary now to stipulate how the latter two variables themselves are determined. Solow's assumptions are:

$$\dot{K}\left(\equiv \frac{dK}{dt}\right) = sQ \qquad \text{[constant proportion of } Q \text{ is invested]} \tag{15.26}$$

$$\frac{\dot{L}}{L}\left(\equiv \frac{dL/dt}{L}\right) = \lambda \qquad (\lambda > 0) \qquad \text{[labor force grows exponentially]} \tag{15.27}$$

The symbol s represents a (constant) marginal propensity to save, and λ, a (constant) rate of growth of labor. Note the dynamic nature of these assumptions; they specify not how the *levels* of K and L are determined, but how their *rates of change* are.

Equations (15.25) through (15.27) constitute a complete model. To solve this model, we shall first condense it into a single equation in one variable. To begin with, substitute (15.25) into (15.26) to get

$$\dot{K} = sL\phi(k) \tag{15.28}$$

Since $k \equiv K/L$, and $K \equiv kL$, however, we can obtain another expression for $\dot{K}$ by differentiating the latter identity:

$$\begin{aligned} \dot{K} &= L\dot{k} + k\dot{L} \qquad &&\text{[product rule]} \\ &= L\dot{k} + k\lambda L &&\text{[by (15.27)]} \end{aligned} \tag{15.29}$$

When (15.29) is equated to (15.28) and the common factor L eliminated, the result emerges that

$$\dot{k} = s\phi(k) - \lambda k \tag{15.30}$$

This equation—a differential equation in the variable k, with two parameters s and λ—is the fundamental equation of the Solow growth model.

A Qualitative-Graphic Analysis

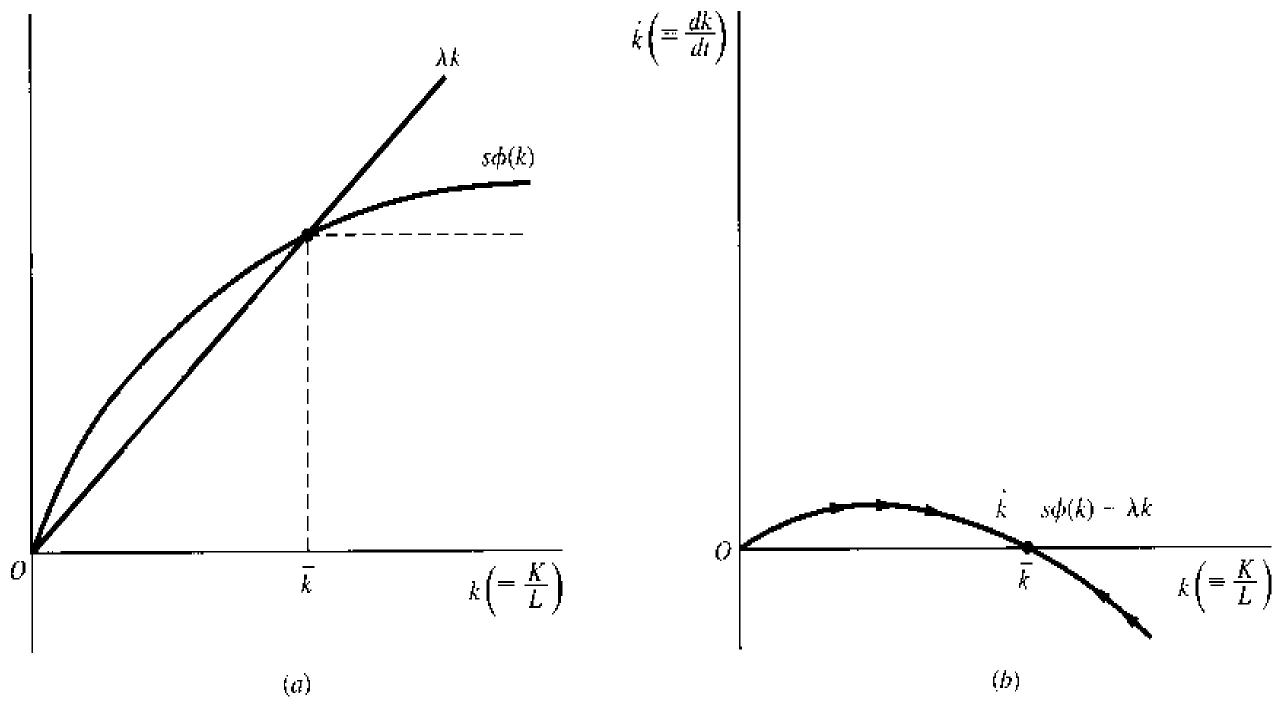
Because (15.30) is stated in a general-function form, no specific quantitative solution is available. Nevertheless, we can analyze it qualitatively. To this end, we should plot a phase line, with $\dot{k}$ on the vertical axis and k on the horizontal.

Since (15.30) contains two terms on the right, however, let us first plot these as two separate curves. The λk term, a linear function of k, will obviously show up in Fig. 15.5*a* as a straight line, with a zero vertical intercept and a slope equal to λ. The $s\phi(k)$ term, on the other hand, plots as a curve that increases at a decreasing rate, like $\phi(k)$, since $s\phi(k)$ is merely a constant fraction of the $\phi(k)$ curve. If we consider K to be an indispensable factor of production, we must start the $s\phi(k)$ curve from the point of origin; this is because if $K = 0$ and thus $k = 0$, Q must also be zero, as will be $\phi(k)$ and $s\phi(k)$. The way the curve is actually drawn also reflects the implicit assumption that there exists a set of k values for which $s\phi(k)$ exceeds λk, so that the two curves intersect at some positive value of k, namely $\bar{k}$.

Based upon these two curves, the value of $\dot{k}$ for each value of k can be measured by the vertical distance between the two curves. Plotting the values of $\dot{k}$ against k, as in Fig. 15.5*b*, will then yield the phase line we need. Note that, since the two curves in Fig. 15.5*a* intersect when the capital–labor ratio is $\bar{k}$, the phase line in Fig. 15.5*b* must cross the horizontal axis at $\bar{k}$. This marks $\bar{k}$ as the intertemporal equilibrium capital–labor ratio.

Inasmuch as the phase line has a negative slope at $\bar{k}$, the equilibrium is readily identified as a stable one; given any (positive) initial value of k, the dynamic movement of the model

FIGURE 15.5



must lead us convergently to the equilibrium level $\bar{k}$. The significant point is that once this equilibrium is attained—and thus the capital–labor ratio is (by definition) unvarying over time—capital must thereafter grow apace with labor, at the identical rate λ. This will imply, in turn, that net investment must grow at the rate λ (see Exercise 15.7-2). Note, however, that the word *must* is used here not in the sense of requirement, but with the implication of automaticity. Thus, what the Solow model serves to show is that, given a rate of growth of labor λ, the economy by itself, and without the delicate balancing à la Domar, can eventually reach a state of steady growth in which investment will grow at the rate λ, the same as K and L. Moreover, in order to satisfy (15.25), Q must grow at the same rate as well because $\phi(k)$ is a constant when the capital–labor ratio remains unvarying at the level $\bar{k}$. Such a situation, in which the relevant variables all grow at an identical rate, is called a *steady state*—a generalization of the concept of *stationary state* (in which the relevant variables all remain constant, or in other words all grow at the zero rate).

Note that, in the preceding analysis, the production function is assumed for convenience to be invariant over time. If the state of technology is allowed to improve, on the other hand, the production function will have to be duly modified. For instance, it may be written instead in the form

$$Q = T(t)f(K, L) \qquad \left(\frac{dT}{dt} > 0\right)$$

where T, some measure of technology, is an increasing function of time. Because of the increasing multiplicative term $T(t)$, a fixed amount of K and L will turn out a larger output at a future date than at present. In this event, the $s\phi(k)$ curve in Fig. 15.5 will be subject to a secular upward shift, resulting in successively higher intersections with the λk ray and also in larger values of $\bar{k}$. With technological improvement, therefore, it will become possible, in a succession of steady states, to have a larger and larger amount of capital equipment available to each representative worker in the economy, with a concomitant rise in productivity.

A Quantitative Illustration

The preceding analysis had to be qualitative, owing to the presence of a general function $\phi(k)$ in the model. But if we specify the production function to be a linearly homogeneous Cobb-Douglas function, for instance, then a quantitative solution can be found as well.

Let us write the production function as

$$Q = K^{\alpha}L^{1-\alpha} = L\left(\frac{K}{L}\right)^{\alpha} = Lk^{\alpha}$$

so that $\phi(k) = k^{\alpha}$. Then (15.30) becomes

$$\dot{k} = sk^{\alpha} - \lambda k \qquad \text{or} \qquad \dot{k} + \lambda k = sk^{\alpha}$$

which is a Bernoulli equation in the variable k [see (15.24)], with $R = \lambda$, $T = s$, and $m = \alpha$. Letting $z = k^{1-\alpha}$, we obtain its linearized version

$$dz + [(1-\alpha)\lambda z - (1-\alpha)s]\,dt = 0$$

or

$$\frac{dz}{dt} + \underbrace{(1-\alpha)\lambda}_{a} z = \underbrace{(1-\alpha)s}_{b}$$

This is a linear differential equation with a constant coefficient a and a constant term b. Thus, by formula (15.5′), we have

$$z(t) = \left[z(0) - \frac{s}{\lambda}\right]e^{-(1-\alpha)\lambda t} + \frac{s}{\lambda}$$

The substitution of $z = k^{1-\alpha}$ will then yield the final solution

$$k^{1-\alpha} = \left[k(0)^{1-\alpha} - \frac{s}{\lambda}\right]e^{-(1-\alpha)\lambda t} + \frac{s}{\lambda}$$

where $k(0)$ is the initial value of the capital–labor ratio k.

This solution is what determines the time path of k. Recalling that $(1-\alpha)$ and λ are both positive, we see that as $t \to \infty$ the exponential expression will approach zero; consequently,

$$k^{1-\alpha} \to \frac{s}{\lambda} \qquad \text{or} \qquad k \to \left(\frac{s}{\lambda}\right)^{1/(1-\alpha)} \qquad \text{as } t \to \infty$$

Therefore, the capital–labor ratio will approach a constant as its equilibrium value. This equilibrium or steady-state value, $(s/\lambda)^{1/(1-\alpha)}$, varies directly with the propensity to save s, and inversely with the rate of growth of labor λ.

EXERCISE 15.7

1. Divide (15.30) through by k, and interpret the resulting equation in terms of the growth rates of k, K, and L.
2. Show that, if capital is growing at the rate λ (that is, $K = Ae^{\lambda t}$), net investment I must also be growing at the rate λ.
3. The original input variables of the Solow model are K and L, but the fundamental equation (15.30) focuses on the capital–labor ratio k instead. What assumption(s) in the model is(are) responsible for (and make possible) this shift of focus? Explain.
4. Draw a phase diagram for each of the following, and discuss the qualitative aspects of the time path $y(t)$:

   (a) $\dot{y} = 3 - y - \ln y$ (b) $\dot{y} = e^y - (y + 2)$