3. Find:

(a) $\int \frac{3dx}{x} \quad (x \neq 0)$

(b) $\int \frac{dx}{x-2} \quad (x \neq 2)$

(c) $\int \frac{2x}{x^2+3}\, dx$

(d) $\int \frac{x}{3x^2+5}\, dx$

4. Find:

(a) $\int (x+3)(x+1)^{1/2}\, dx$

(b) $\int x \ln x\, dx \quad (x > 0)$

5. Given n constants k_i (with $i = 1, 2, \ldots, n$) and n functions $f_i(x)$, deduce from Rules IV and V that

$$\int \sum_{i=1}^{n} k_i f_i(x)\, dx = \sum_{i=1}^{n} k_i \int f_i(x)\, dx$$

14.3 Definite Integrals

Meaning of Definite Integrals

All the integrals cited in Sec. 14.2 are of the *indefinite* variety: each is a function of a variable and, hence, possesses no definite numerical value. Now, for a given indefinite integral of a continuous function $f(x)$,

$$\int f(x)\, dx = F(x) + c$$

if we choose two values of x in the domain, say, a and b ($a < b$), substitute them successively into the right side of the equation, and form the difference

$$[F(b) + c] - [F(a) + c] = F(b) - F(a)$$

we get a specific numerical value, free of the variable x as well as the arbitrary constant c. This value is called the *definite integral* of $f(x)$ from a to b. We refer to a as the *lower limit of integration* and to b as the *upper limit of integration*.

In order to indicate the limits of integration, we now modify the integral sign to the form $\int_a^b$. The evaluation of the definite integral is then symbolized in the following steps:

$$\int_a^b f(x)\, dx = F(x)\Big]_a^b = F(b) - F(a) \tag{14.6}$$

where the symbol $]_a^b$ (also written $|_a^b$ or $[\cdots]_a^b$) is an instruction to substitute b and a, successively, for x in the result of integration to get $F(b)$ and $F(a)$, and then take their difference, as indicated on the right of (14.6). As the first step, however, we must find the indefinite integral, although we may omit the constant c, since the latter will drop out in the process of difference-taking anyway.

Example 1 Evaluate $\int_1^5 3x^2\, dx$. Since the indefinite integral is $x^3 + c$, this definite integral has the value

$$\int_1^5 3x^2\, dx = x^3\Big]_1^5 = (5)^3 - (1)^3 = 125 - 1 = 124$$

Example 2 Evaluate $\int_a^b ke^x\,dx$. Here, the limits of integration are given in symbols; consequently, the result of integration is also in terms of those symbols:

$$\int_a^b ke^x\,dx = ke^x\Big]_a^b = k(e^b - e^a)$$

Example 3 Evaluate $\int_0^4 \left(\frac{1}{1+x} + 2x\right) dx,\ (x \neq -1)$. The indefinite integral is $\ln|1+x| + x^2 + c$; thus the answer is

$$\begin{aligned}\int_0^4 \left(\frac{1}{1+x} + 2x\right) dx &= \Big[\ln|1+x| + x^2\Big]_0^4 \\ &= (\ln 5 + 16) - (\ln 1 + 0) \\ &= \ln 5 + 16 \qquad [\text{since} \ln 1 = 0]\end{aligned}$$

It is important to realize that the limits of integration a and b both refer to values of the variable x. Were we to use the substitution-of-variables technique (Rules VI and VII) during integration and introduce a variable u, care should be taken *not* to consider a and b as the limits of u. Example 4 will illustrate this point.

Example 4 Evaluate $\int_1^2 (2x^3-1)^2(6x^2)\,dx$. Let $u = 2x^3 - 1$; then $du/dx = 6x^2$, or $du = 6x^2\,dx$. Now notice that, when $x = 1$, u will be 1 but that, when $x = 2$, u will be 15; in other words, the limits of integration in terms of the variable u should be 1 (lower) and 15 (upper). Rewriting the given integral in u will therefore give us not $\int_1^2 u^2\,du$ but

$$\int_1^{15} u^2\,du = \frac{1}{3}u^3\Big]_1^{15} = \frac{1}{3}(15^3 - 1^3) = 1{,}124\tfrac{2}{3}$$

Alternatively, we may first convert u back to x and then use the original limits of 1 and 2 to get the identical answer:

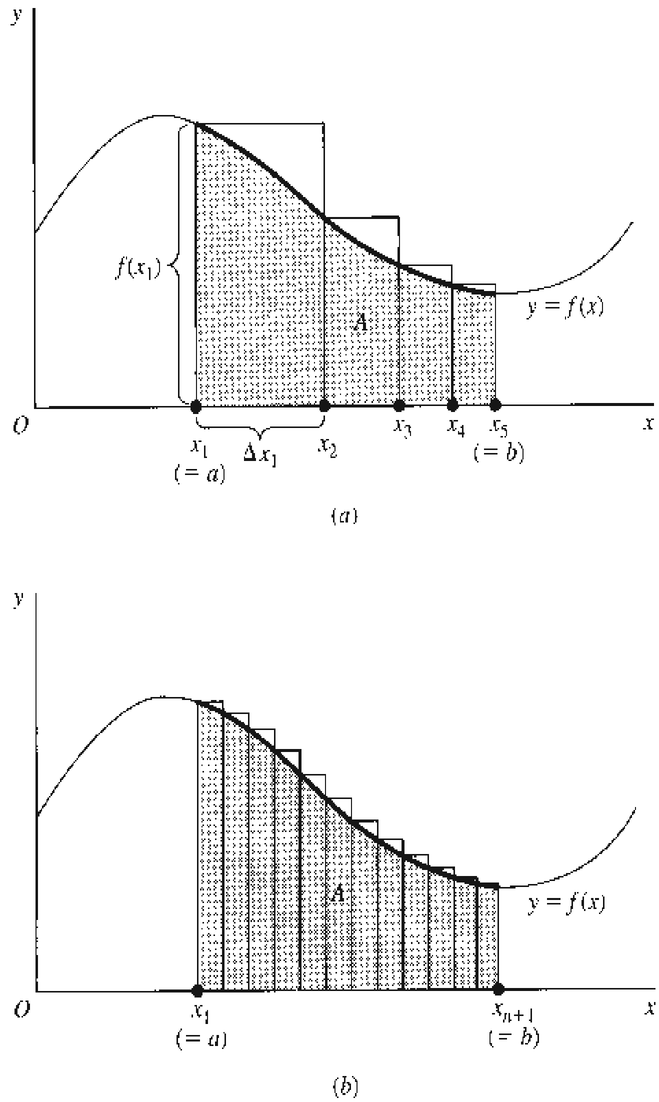
$$\left[\frac{1}{3}u^3\right]_{u=1}^{u=15} = \left[\frac{1}{3}(2x^3-1)^3\right]_{x=1}^{x=2} = \frac{1}{3}(15^3 - 1^3) = 1{,}124\tfrac{2}{3}$$

A Definite Integral as an Area under a Curve

Every definite integral has a definite value. That value may be interpreted geometrically to be a particular area under a given curve.

The graph of a continuous function $y = f(x)$ is drawn in Fig. 14.1. If we seek to measure the (shaded) area A enclosed by the curve and the x axis between the two points a and b in the domain, we may proceed in the following manner. First, we divide the interval $[a, b]$ into n subintervals (not necessarily equal in length). Four of these are drawn in Fig. 14.1a—that is, $n = 4$—the first being $[x_1, x_2]$ and the last, $[x_4, x_5]$. Since each of these represents a change in x, we may refer to them as $\Delta x_1, \ldots, \Delta x_4$, respectively. Now, on the subintervals let us construct four rectangular blocks such that the height of each block is equal to the highest value of the function attained in that block (which happens to occur at the left-side boundary of each rectangle here). The first block thus has the height $f(x_1)$ and

FIGURE 14.1



the width Δx_1, and, in general, the ith block has the height $f(x_i)$ and the width Δx_i. The total area A^* of this set of blocks is the sum

$$A^* = \sum_{i=1}^{n} f(x_i)\,\Delta x_i \qquad (n = 4 \text{ in Fig. 14.1}a)$$

This, though, is obviously *not* the area under the curve we seek, but only a very rough approximation thereof.

What makes A^* deviate from the true value of A is the unshaded portion of the rectangular blocks; these make A^* an *over*estimate of A. If the unshaded portion can be shrunk in size and be made to approach zero, however, the approximation value A^* will correspondingly approach the true value A. This result will materialize when we try a finer and finer segmentation of the interval $[a, b]$, so that n is increased and Δx_i is shortened indefinitely. Then the blocks will become more slender (if more numerous), and the protrusion beyond the curve will diminish, as can be seen in Fig. 14.1b. Carried to the limit, this "slenderizing" operation yields

$$\lim_{n\to\infty} \sum_{i=1}^{n} f(x_i)\,\Delta x_i = \lim_{n\to\infty} A^* = \text{area } A \tag{14.7}$$

provided this limit exists. (It does in the present case.) This equation, indeed, constitutes the formal definition of an area under a curve.

The summation expression in (14.7), $\sum_{i=1}^{n} f(x_i)\,\Delta x_i$, bears a certain resemblance to the definite integral expression $\int_a^b f(x)\,dx$. Indeed, the latter is based on the former. The replacement of Δx_i by the differential dx is done in the same spirit as in our earlier discussion of "approximation" in Sec. 8.1. Thus, we rewrite $f(x_i)\,\Delta x_i$ into $f(x)\,dx$. What about the summation sign? The $\sum_{i=1}^{n}$ notation represents the sum of a *finite* number of terms. When we let $n \to \infty$, and take the limit of that sum, the regular notation for such an operation is rather cumbersome. Thus a simpler substitute is needed. That substitute is $\int_a^b$, where the elongated S symbol also indicates a sum, and where a and b (just as $i = 1$ and n) serve to specify the lower and upper limits of this sum. In short the definite integral is a shorthand for the limit-of-a-sum expression in (14.7). That is,

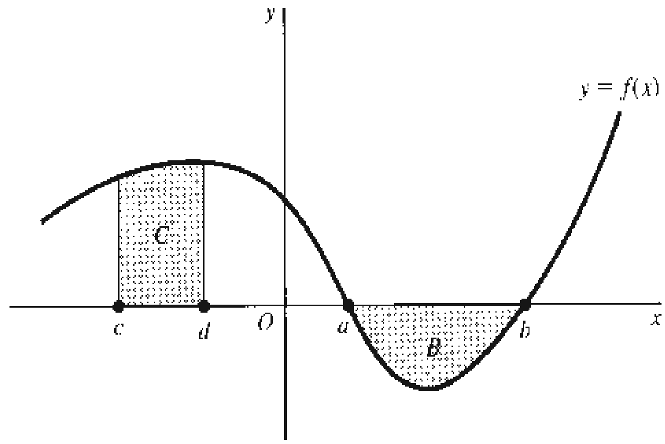
$$\int_a^b f(x)\,dx \equiv \lim_{n\to\infty} \sum_{i=1}^{n} f(x_i)\,\Delta x_i = \text{area } A$$

Thus the said definite integral (referred to as a *Riemann integral*) now has an *area* connotation as well as a *sum* connotation, because $\int_a^b$ is the continuous counterpart of the discrete concept of $\sum_{i=1}^{n}$.

In Fig. 14.1, we attempted to approximate area A by systematically reducing an *over*-estimate A^* by finer segmentation of the interval $[a, b]$. The resulting limit of the sum of block areas is called the *upper integral*—an approximation from above. We could also have approximated area A from below by forming rectangular blocks inscribed by the curve rather than protruding beyond it (see Exercise 14.3-3). The total area A^{**} of this new set of blocks will *under*estimate A, but as the segmentation of $[a, b]$ becomes finer and finer, we shall again find $\lim_{n\to\infty} A^{**} = A$. The last-cited limit of the sum of block areas is called the *lower integral*. If, and only if, the upper integral and lower integral are equal in value, then the Riemann integral $\int_a^b f(x)\,dx$ is defined, and the function $f(x)$ is said to be *Riemann integrable*. There exist theorems specifying the conditions under which a function $f(x)$ is integrable. According to the fundamental theorem of calculus, a function is integrable in $[a, b]$ if it is continuous in that interval. As long as we are working with continuous functions, therefore, we should have no worries in this regard.

Another point may be noted. Although the area A in Fig. 14.1 happens to lie entirely under a decreasing portion of the curve $y = f(x)$, the conceptual equating of a definite integral with an area is valid also for upward-sloping portions of the curve. In fact, both types of slope may be present simultaneously; e.g., we can calculate $\int_0^b f(x)\,dx$ as the area under the curve in Fig. 14.1 above the line Ob.

FIGURE 14.2



Note that, if we calculate the area B in Fig. 14.2 by the definite integral $\int_a^b f(x)\,dx$, the answer will come out negative, because the height of each rectangular block involved in this area is negative. This gives rise to the notion of a *negative area,* an area that lies *below* the x axis and *above* a given curve. In case we are interested in the numerical rather than the algebraic value of such an area, therefore, we should take the absolute value of the relevant definite integral. The area $C = \int_c^d f(x)\,dx$, on the other hand, has a positive sign even though it lies in the negative region of the x axis; this is because each rectangular block has a positive height as well as a positive width when we are moving from c to d. From this, the implication is clear that interchange of the two limits of integration would, by reversing the direction of movement, alter the sign of Δx_i and of the definite integral. Applied to area B, we see that the definite integral $\int_b^a f(x)\,dx$ (from b to a) will give the negative of the area B; this will measure the numerical value of this area.

Some Properties of Definite Integrals

The discussion in the preceding paragraph leads us to the following property of definite integrals.

Property I The interchange of the limits of integration changes the sign of the definite integral:

$$\int_b^a f(x)\,dx = -\int_a^b f(x)\,dx$$

This can be proved as follows:

$$\int_b^a f(x)\,dx = F(a) - F(b) = -[F(b) - F(a)] = -\int_a^b f(x)\,dx$$

Definite integrals also possess some other interesting properties.

Property II A definite integral has a value of zero when the two limits of integration are identical:

$$\int_a^a f(x)\,dx = F(a) - F(a) = 0$$

Under the "area" interpretation, this means that the area (under a curve) above any single *point* in the domain is nil. This is as it should be, because on top of a point on the x axis, we can draw only a (one-dimensional) *line*, never a (two-dimensional) *area*.

Property III A definite integral can be expressed as a sum of a finite number of definite subintegrals as follows:

$$\int_a^d f(x)\,dx = \int_a^b f(x)\,dx + \int_b^c f(x)\,dx + \int_c^d f(x)\,dx \qquad (a < b < c < d)$$

Only three subintegrals are shown in this equation, but the extension to the case of n subintegrals is also valid. This property is sometimes described as the *additivity property*.

In terms of area, this means that the area (under the curve) lying above the interval $[a, d]$ on the x axis can be obtained by summing the areas lying above the subintervals in the set $\{[a, b], [b, c], [c, d]\}$. Note that, since we are dealing with closed intervals, the border points b and c have each been included in *two* areas. Is this not double counting? It indeed is. But fortunately no damage is done, because by Property II the area above a single point is zero, so that the double counting produces no effect on the calculation. But, needless to say, the double counting of any *interval* is never permitted.

Earlier, it was mentioned that all continuous functions are Riemann integrable. Now, by Property III, we can also find the definite integrals (areas) of certain discontinuous functions. Consider the step function in Fig. 14.3*a*. In spite of the discontinuity at point b in the interval $[a, c]$, we can find the shaded area from the sum

$$\int_a^b f(x)\,dx + \int_b^c f(x)\,dx$$

The same also applies to the curve in Fig. 14.3*b*.

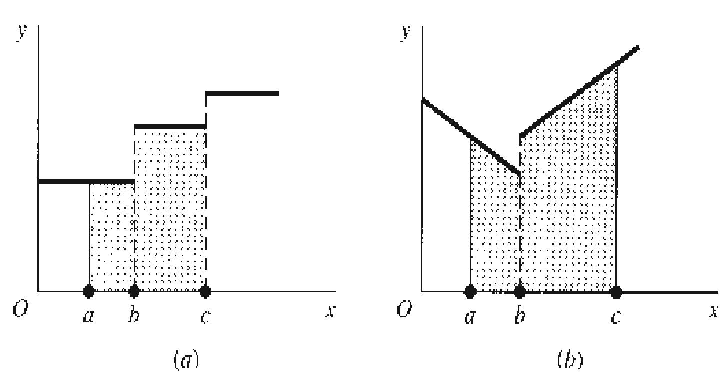
Property IV

$$\int_a^b -f(x)\,dx = -\int_a^b f(x)\,dx$$

Property V

$$\int_a^b kf(x)\,dx = k\int_a^b f(x)\,dx$$

FIGURE 14.3



Property VI

$$\int_a^b [f(x) + g(x)]\, dx = \int_a^b f(x)\, dx + \int_a^b g(x)\, dx$$

Property VII (integration by parts) Given $u(x)$ and $v(x)$,

$$\int_{x=a}^{x=b} v\, du = uv\Big|_{x=a}^{x=b} - \int_{x=a}^{x=b} u\, dv$$

These last four properties, all borrowed from the rules of indefinite integration, should require no further explanation.

Another Look at the Indefinite Integral

We introduced the definite integral by way of attaching two limits of integration to an indefinite integral. Now that we know the meaning of the definite integral, let us see how we can revert from the latter to the indefinite integral.

Suppose that, instead of fixing the upper limit of integration at b, we allow it to be a variable, designated simply as x. Then the integral will take the form

$$\int_a^x f(x)\, dx = F(x) - F(a)$$

which, now being a function of x, denotes a *variable* area under the curve of $f(x)$. But since the last term on the right is a constant, this integral must be a member of the family of primitive functions of $f(x)$, which we denoted earlier as $F(x) + c$. If we set $c = -F(a)$, then the above integral becomes exactly the indefinite integral $\int f(x)\, dx$.

From this point of view, therefore, we may consider the $\int$ symbol to mean the same as $\int_a^x$, provided it is understood that in the latter version of the symbol the lower limit of integration is related to the constant of integration by the equation $c = -F(a)$.

EXERCISE 14.3

1. Evaluate the following:

   (a) $\int_1^3 \frac{1}{2}x^2\, dx$ (d) $\int_2^4 (x^3 - 6x^2)\, dx$

   (b) $\int_0^1 x(x^2 + 6)\, dx$ (e) $\int_{-1}^1 (ax^2 + bx + c)\, dx$

   (c) $\int_1^3 3\sqrt{x}\, dx$ (f) $\int_4^2 x^2 \left(\frac{1}{3}x^3 + 1\right) dx$

2. Evaluate the following:

   (a) $\int_1^2 e^{-2x}\, dx$ (c) $\int_2^3 (e^{2x} + e^x)\, dx$

   (b) $\int_1^{e-2} \frac{dx}{x+2}$ (d) $\int_e^6 \left(\frac{1}{x} + \frac{1}{1+x}\right) dx$

3. In Fig. 14.1*a*, take the lowest value of the function attained in each subinterval as the height of the rectangular block, i.e., take $f(x_2)$ instead of $f(x_1)$ as the height of the first block, though still retaining Δx_1 as its width, and do likewise for the other blocks.
   (*a*) Write a summation expression for the total area A^{**} of the new rectangles.
   (*b*) Does A^{**} overestimate or underestimate the desired area A?
   (*c*) Would A^{**} tend to approach or to deviate further from A if a finer segmentation of $[a, b]$ were introduced? (*Hint:* Try a diagram.)
   (*d*) In the limit, when the number n of subintervals approaches ∞, would the approximation value A^{**} approach the true value A, just as the approximation value A^* did?
   (*e*) What can you conclude from (*a*) to (*d*) about the Riemann integrability of the function $f(x)$ in Fig. 14.1a?
4. The definite integral $\int_a^b f(x)\,dx$ is said to represent an area under a curve. Does this curve refer to the graph of the integrand $f(x)$, or of the primitive function $F(x)$? If we plot the graph of the $F(x)$ function, how can we show the given definite integral on it—by an area, a line segment, or a point?
5. Verify that a constant c can be equivalently expressed as a definite integral:
   (*a*) $c \equiv \int_0^b \frac{c}{b}\,dx$ (*b*) $c \equiv \int_0^c 1\,dt$

14.4 Improper Integrals

Certain integrals are said to be "improper." We shall briefly discuss two varieties thereof.

Infinite Limits of Integration

When we have definite integrals of the form

$$\int_a^\infty f(x)\,dx \qquad \text{and} \qquad \int_{-\infty}^b f(x)\,dx$$

with one limit of integration being infinite, we refer to them as *improper integrals*. In these cases, it is not possible to evaluate the integrals as, respectively,

$$F(\infty) - F(a) \qquad \text{and} \qquad F(b) - F(-\infty)$$

because ∞ is not a number, and therefore it cannot be substituted for x in the function $F(x)$. Instead, we must resort once more to the concept of limits.

The first improper integral we cited can be defined to be the limit of another (proper) integral as the latter's upper limit of integration tends to ∞; that is,

$$\int_a^\infty f(x)\,dx \equiv \lim_{b\to\infty} \int_a^b f(x)\,dx \tag{14.8}$$

If this limit exists, the improper integral is said to be convergent (or to converge), and the limiting process will yield the value of the integral. If the limit does not exist, the improper integral is said to be divergent and is in fact meaningless. By the same token, we can define

$$\int_{-\infty}^b f(x)\,dx \equiv \lim_{a\to-\infty} \int_a^b f(x)\,dx \tag{14.8'}$$

with the same criterion of convergence and divergence.

Example 1 Evaluate $\int_1^\infty \frac{dx}{x^2}$. First we note that

$$\int_1^b \frac{dx}{x^2} = \frac{-1}{x}\Big]_1^b = \frac{-1}{b} + 1$$

Hence, in line with (14.8), the desired integral is

$$\int_1^\infty \frac{dx}{x^2} = \lim_{b\to\infty} \int_1^b \frac{dx}{x^2} = \lim_{b\to\infty} \left(\frac{-1}{b} + 1\right) = 1$$

This improper integral does converge, and it has a value of 1.

Since the limit expression is cumbersome to write, some people prefer to omit the "lim" notation and write simply

$$\int_1^\infty \frac{dx}{x^2} = \frac{-1}{x}\Big]_1^\infty = 0 + 1 = 1$$

Even when written in this form, however, the improper integral should nevertheless be interpreted with the limit concept in mind.

Graphically, this improper integral still has the connotation of an area. But since the upper limit of integration is allowed to take on increasingly larger values in this case, the right-side boundary must be extended eastward indefinitely, as shown in Fig. 14.4*a*. Despite this, we are able to consider the area to have the definite (limit) value of 1.

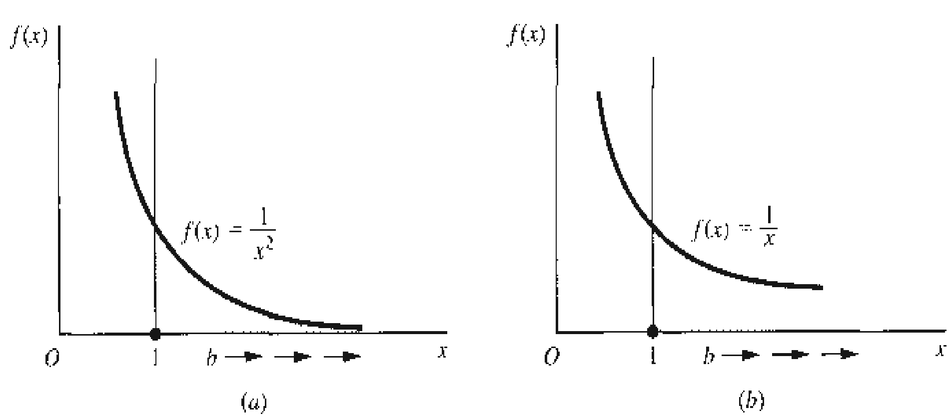
Example 2 Evaluate $\int_1^\infty \frac{dx}{x}$. As before, we first find

$$\int_1^b \frac{dx}{x} = \ln x\Big]_1^b = \ln b - \ln 1 = \ln b$$

When we let $b \to \infty$, by (10.16′) we have $\ln b \to \infty$. Thus the given improper integral is divergent.

Figure 14.4*b* shows the graph of the function $1/x$, as well as the area corresponding to the given integral. The indefinite eastward extension of the right-side boundary will result this time in an infinite area, even though the shape of the graph displays a superficial similarity to that of Fig. 14.4*a*.

FIGURE 14.4



What if both limits of integration are infinite? A direct extension of (14.8) and (14.8′) would suggest the definition

$$\int_{-\infty}^{\infty} f(x)\,dx = \lim_{\substack{b\to+\infty \\ a\to-\infty}} \int_a^b f(x)\,dx \qquad \textbf{(14.8'')}$$

Again, this improper integral is said to converge if and only if the limit in question exists.

Infinite Integrand

Even with finite limits of integration, an integral can still be improper if the integrand becomes infinite somewhere in the interval of integration $[a, b]$. To evaluate such an integral, we must again rely upon the concept of a limit.

Example 3 Evaluate $\int_0^1 \frac{1}{x}\,dx$. This integral is improper because, as Fig. 14.4*b* shows, the integrand is infinite at the lower limit of integration ($1/x \to \infty$ as $x \to 0^-$). Therefore we should first find the integral

$$\int_a^1 \frac{1}{x}\,dx = \ln x\Big]_a^1 = \ln 1 - \ln a = -\ln a \qquad [\text{for } a > 0]$$

and then evaluate its limit as $a \to 0^+$:

$$\int_0^1 \frac{1}{x} dx \equiv \lim_{a\to0^+} \int_a^1 \frac{1}{x}\,dx = \lim_{a\to0^+} (-\ln a)$$

Since this limit does not exist (as $a \to 0^+$, $\ln a \to -\infty$), the given integral is divergent.

Example 4 Evaluate $\int_0^9 x^{-1/2}\,dx$. When $x \to 0^+$, the integrand $1/\sqrt{x}$ becomes infinite; the integral is improper. Again, we can first find

$$\int_a^9 x^{-1/2}\,dx = 2x^{1/2}\Big]_a^9 = 6 - 2\sqrt{a}$$

The limit of this expression as $a \to 0^+$ is $6 - 0 = 6$. Thus the given integral is convergent (to 6).

The situation where the integrand becomes infinite at the *upper* limit of integration is perfectly similar. It is an altogether different proposition, however, when an infinite value of the integrand occurs in the open interval (a, b) rather than at a or b. In this eventuality, it is necessary to take advantage of the additivity of definite integrals and first decompose the given integral into subintegrals. Assume that $f(x) \to \infty$ as $x \to p$, where p is a point in the interval (a, b); then, by the additivity property, we have

$$\int_a^b f(x)\,dx = \int_a^p f(x)\,dx + \int_p^b f(x)\,dx$$

The given integral on the left can be considered as convergent if and only if each subintegral has a limit.

Example 5 Evaluate $\int_{-1}^{1} \frac{1}{x^3}\,dx$. The integrand tends to infinity when x approaches zero; thus we must write the given integral as the sum

$$\int_{-1}^{1} x^{-3}\,dx = \int_{-1}^{0} x^{-3}\,dx + \int_{0}^{1} x^{-3}\,dx \qquad (\text{say}, \equiv I_1 + I_2)$$

The integral I_1 is divergent, because

$$\lim_{b\to 0^-} \int_{-1}^{b} x^{-3}\,dx = \lim_{b\to 0^-} \left[\frac{-1}{2}x^{-2}\right]_{-1}^{b} = \lim_{b\to 0^-} \left(-\frac{1}{2b^2} + \frac{1}{2}\right) = -\infty$$

Thus, we can conclude immediately, without having to evaluate I_2, that the given integral is divergent.

EXERCISE 14.4

1. Check the definite integrals given in Exercises 14.3-1 and 14.3-2 to determine whether any of them is improper. If improper, indicate which variety of improper integral each one is.
2. Which of the following integrals are improper, and why?

   (*a*) $\int_{0}^{\infty} e^{-rt}\,dt$ (*d*) $\int_{-\infty}^{0} e^{rt}\,dt$

   (*b*) $\int_{2}^{3} x^4\,dx$ (*e*) $\int_{1}^{5} \frac{dx}{x-2}$

   (*c*) $\int_{0}^{1} x^{-2/3}\,dx$ (*f*) $\int_{-3}^{4} 6\,dx$

3. Evaluate all the *improper* integrals in Prob. 2.
4. Evaluate the integral I_2 of Example 5, and show that it is also divergent.
5. (*a*) Graph the function $y = ce^{-t}$ for nonnegative t, $(c > 0)$, and shade the area under the curve.

   (*b*) Write a mathematical expression for this area, and determine whether it is a finite area.

14.5 Some Economic Applications of Integrals

Integrals are used in economic analysis in various ways. We shall illustrate a few simple applications in the present section and then show the application to the Domar growth model in Sec. 14.6.

From a Marginal Function to a Total Function

Given a total function (e.g., a total-cost function), the process of differentiation can yield the marginal function (e.g., the marginal-cost function). Because the process of integration is the opposite of differentiation, it should enable us, conversely, to infer the total function from a given marginal function.

Example 1

If the marginal cost (MC) of a firm is the following function of output, $C'(Q) = 2e^{0.2Q}$, and if the fixed cost is $C_F = 90$, find the total-cost function $C(Q)$. By integrating $C'(Q)$ with respect to Q, we find that

$$\int 2e^{0.2Q}\, dQ = 2\frac{1}{0.2}e^{0.2Q} + c = 10e^{0.2Q} + c \tag{14.9}$$

This result may be taken as the desired $C(Q)$ function except that, in view of the arbitrary constant c, the answer appears indeterminate. Fortunately, the information that $C_F = 90$ can be used as an initial condition to definitize the constant. When $Q = 0$, total cost C will consist solely of C_F. Setting $Q = 0$ in the result of (14.9), therefore, we should get a value of 90; that is, $10e^0 + c = 90$. But this would imply that $c = 90 - 10 = 80$. Hence, the total-cost function is

$$C(Q) = 10e^{0.2Q} + 80$$

Note that, unlike the case of (14.2), where the arbitrary constant c has the same value as the initial value of the variable $H(0)$, in the present example we have $c = 80$ but $C(0) \equiv C_F = 90$, so that the two take different values. In general, it should *not* be assumed that the arbitrary constant c will always be equal to the initial value of the total function.

Example 2

If the marginal propensity to save (MPS) is the following function of income, $S'(Y) = 0.3 - 0.1Y^{-1/2}$, and if the aggregate savings S is nil when income Y is 81, find the saving function $S(Y)$. As the MPS is the derivative of the S function, the problem now calls for the integration of $S'(Y)$:

$$S(Y) = \int (0.3 - 0.1Y^{-1/2})\, dY = 0.3Y - 0.2Y^{1/2} + c$$

The specific value of the constant c can be found from the fact that $S = 0$ when $Y = 81$. Even though, strictly speaking, this is not an *initial* condition (not relating to $Y = 0$), substitution of this information into the preceding integral will nevertheless serve to definitize c. Since

$$0 = 0.3(81) - 0.2(9) + c \qquad \Rightarrow \qquad c = -22.5$$

the desired saving function is

$$S(Y) = 0.3Y - 0.2Y^{1/2} - 22.5$$

The technique illustrated in Examples 1 and 2 can be extended directly to other problems involving the search for total functions (such as total revenue, total consumption) from given marginal functions. It may also be reiterated that in problems of this type the validity of the answer (an integral) can always be checked by differentiation.

Investment and Capital Formation

Capital formation is the process of adding to a given stock of capital. Regarding this process as continuous over time, we may express capital stock as a function of time, $K(t)$, and use the derivative dK/dt to denote the rate of capital formation.† But the rate of capital

† As a matter of notation, the derivative of a variable with respect to *time* often is also denoted by a dot placed over the variable, such as $\dot{K} \equiv dK/dt$. In dynamic analysis, where derivatives with respect to *time* occur in abundance, this more concise symbol can contribute substantially to notational simplicity. However, a dot, being such a tiny mark, is easily lost sight of or misplaced; thus, great care is required in using this symbol.

formation at time t is identical with the rate of *net investment* flow at time t, denoted by $I(t)$. Thus, capital stock K and net investment I are related by the following two equations:

$$\frac{dK}{dt} \equiv I(t)$$

and
$$K(t) = \int I(t)\, dt = \int \frac{dK}{dt}\, dt = \int dK$$

The first of the preceding equations is an identity; it shows the synonymity between net investment and the increment of capital. Since $I(t)$ is the derivative of $K(t)$, it stands to reason that $K(t)$ is the integral or antiderivative of $I(t)$, as shown in the second equation. The transformation of the integrand in the latter equation is also easy to comprehend: The switch from I to dK/dt is by definition, and the next transformation is by cancellation of two identical differentials, i.e., by the substitution rule.

Sometimes the concept of *gross investment* is used together with that of net investment in a model. Denoting gross investment by I_g and net investment by I, we can relate them to each other by the equation

$$I_g = I + \delta K$$

where δ represents the rate of depreciation of capital and δK, the rate of *replacement investment*.

Example 3

Suppose that the net investment flow is described by the equation $I(t) = 3t^{1/2}$ and that the initial capital stock, at time $t = 0$, is $K(0)$. What is the time path of capital K? By integrating $I(t)$ with respect to t, we obtain

$$K(t) = \int I(t)\, dt = \int 3t^{1/2}\, dt = 2t^{3/2} + c$$

Next, letting $t = 0$ in the leftmost and rightmost expressions, we find $K(0) = c$. Therefore, the time path of K is

$$K(t) = 2t^{3/2} + K(0) \tag{14.10}$$

Observe the basic similarity between the results in (14.10) and in (14.2″).

The concept of definite integral enters into the picture when one desires to find the amount of capital formation during some interval of time (rather than the time path of K). Since $\int I(t)\, dt = K(t)$, we may write the definite integral

$$\int_a^b I(t)\, dt = K(t)\Big]_a^b = K(b) - K(a)$$

to indicate the total capital accumulation during the time interval $[a, b]$. Of course, this also represents an area under the $I(t)$ curve. It should be noted, however, that in the graph of the $K(t)$ function, this definite integral would appear instead as a vertical distance—more specifically, as the difference between the two vertical distances $K(b)$ and $K(a)$. (cf. Exercise 14.3-4.)

To appreciate this distinction between $K(t)$ and $I(t)$ more fully, let us emphasize that capital K is a *stock* concept, whereas investment I is a *flow* concept. Accordingly, while $K(t)$ tells us the *amount* of K existing at each point of time, $I(t)$ gives us the information about the *rate* of (net) investment per year (or per period of time) which is prevailing at each point of time. Thus, in order to calculate the *amount* of net investment undertaken (capital accumulation), we must first specify the length of the interval involved. This fact can also be seen when we rewrite the identity $dK/dt \equiv I(t)$ as $dK \equiv I(t)\,dt$, which states that dK, the increment in K, is based not only on $I(t)$, the rate of flow, but also on dt, the time that elapsed. It is this need to specify the time interval in the expression $I(t)\,dt$ that brings the definite integral into the picture, and gives rise to the *area* representation under the $I(t)$—as against the $K(t)$ curve.

Example 4

If net investment is a constant flow at $I(t) = 1{,}000$ (dollars per year), what will be the total net investment (capital formation) during a year, from $t = 0$ to $t = 1$? Obviously, the answer is \$1,000; this can be obtained formally as follows:

$$\int_0^1 I(t)\,dt = \int_0^1 1{,}000\,dt = 1{,}000t\Big]_0^1 = 1{,}000$$

You can verify that the same answer will emerge if, instead, the year involved is from $t = 1$ to $t = 2$.

Example 5

If $I(t) = 3t^{1/2}$ (thousands of dollars per year)—a nonconstant flow—what will be the capital formation during the time interval [1, 4], that is, during the second, third, and fourth years? The answer lies in the definite integral

$$\int_1^4 3t^{1/2}\,dt = 2t^{3/2}\Big]_1^4 = 16 - 2 = 14$$

On the basis of the preceding examples, we may express the amount of capital accumulation during the time interval $[0, t]$, for any investment rate $I(t)$, by the definite integral

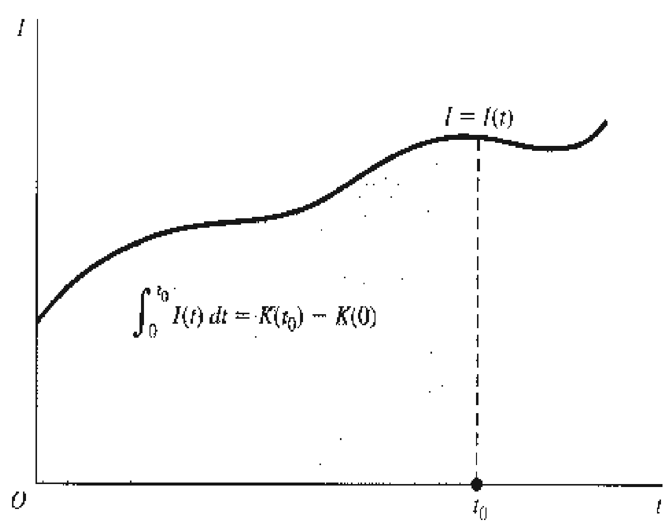
$$\int_0^t I(t)\,dt = K(t)\Big]_0^t = K(t) - K(0)$$

Figure 14.5 illustrates the case of the time interval $[0, t_0]$. Viewed differently, the preceding equation yields the following expression for the time path $K(t)$:

$$K(t) = K(0) + \int_0^t I(t)\,dt$$

The amount of K at any time t is the initial capital plus the total capital accumulation that has occurred since.

FIGURE 14.5



Present Value of a Cash Flow

Our earlier discussion of discounting and present value, limited to the case of a *single future value* V, led us to the discounting formulas

$$A = V(1+i)^{-t} \qquad [\textit{discrete} \text{ case}]$$

and

$$A = Ve^{-rt} \qquad [\textit{continuous} \text{ case}]$$

Now suppose that we have a stream or flow of future values—a series of revenues receivable at various times or of cost outlays payable at various times. How do we compute the present value of the entire cash stream, or cash flow?

In the *discrete* case, if we assume three future revenue figures R_t $(t = 1, 2, 3)$ available at the end of the tth year and also assume an interest rate of i per annum, the present values of R_t will be, respectively,

$$R_1(1+i)^{-1} \qquad R_2(1+i)^{-2} \qquad R_3(1+i)^{-3}$$

It follows that the total present value is the sum

$$\Pi = \sum_{t=1}^{3} R_t(1+i)^{-t} \tag{14.11}$$

(Π is the uppercase Greek letter pi, here signifying *present*.) This differs from the single-value formula only in the replacement of V by R_t and in the insertion of the Σ sign.

The idea of the sum readily carries over to the case of a continuous cash flow, but in the latter context the Σ symbol must give way, of course, to the definite integral sign. Consider a continuous revenue stream at the rate of $R(t)$ dollars per year. This means that at $t = t_1$ the rate of flow is $R(t_1)$ dollars per year, but at another point of time $t = t_2$ the rate will be $R(t_2)$ dollars per year—with t taken as a continuous variable. At any point of time, the amount of revenue during the interval $[t, t + dt]$ can be written as $R(t)\,dt$ [cf. the previous discussion of $dK \equiv I(t)\,dt$]. When continuously discounted at the rate of r per year, its present value should be $R(t)e^{-rt}\,dt$. If we let our problem be that of finding the total present value of a 3-year stream, our answer is to be found in the following definite integral:

$$\Pi = \int_0^3 R(t)e^{-rt}\,dt \tag{14.11'}$$

This expression, the continuous version of the sum in (14.11), differs from the single-value formula only in the replacement of V by $R(t)$ and in the appending of the definite integral sign.[†]

[†] It may be noted that, whereas the upper summation index and the upper limit of integration are identical at 3, the lower summation index 1 differs from the lower limit of integration 0. This is because the first revenue in the discrete stream, by assumption, will not be forthcoming until $t = 1$ (end of first year), but the revenue flow in the continuous case is assumed to commence immediately after $t = 0$.

Example 6

What is the present value of a continuous revenue flow lasting for y years at the constant rate of D dollars per year and discounted at the rate of r per year? According to (14.11′), we have

$$\Pi = \int_0^y De^{-rt}\,dt = D\int_0^y e^{-rt}\,dt = D\left[\frac{-1}{r}e^{-rt}\right]_0^y$$

$$= \frac{-D}{r}e^{-rt}\Big]_{t=0}^{t=y} = \frac{-D}{r}(e^{-ry}-1) = \frac{D}{r}(1-e^{-ry}) \qquad \textbf{(14.12)}$$

Thus, Π depends on D, r and y. If $D = \$3{,}000$, $r = 0.06$, and $y = 2$, for instance, we have

$$\Pi = \frac{3{,}000}{0.06}(1-e^{-0.12}) = 50{,}000(1-0.8869) = \$5{,}655 \qquad \text{[approximately]}$$

The value of Π naturally is always positive; this follows from the positivity of D and r, as well as $(1-e^{-ry})$. (The number e raised to any negative power will always give a positive fractional value, as can be seen from the second quadrant of Fig. 10.3*a*.)

Example 7

In the wine-storage problem of Sec. 10.6, we assumed zero storage cost. That simplifying assumption was necessitated by our ignorance of a way to compute the present value of a cost flow. With this ignorance behind us, we are now ready to permit the wine dealer to incur storage costs.

Let the purchase cost of the case of wine be an amount C, incurred at the present time. Its (future) sale value, which varies with time, may be generally denoted as $V(t)$—its present value being $V(t)e^{-rt}$. Whereas the sale value represents a single future value (there can be only one sale transaction on this case of wine), the storage cost is a stream. Assuming this cost to be a constant stream at the rate of s dollars per year, the total present value of the storage cost incurred in a total of t years will amount to

$$\int_0^t se^{-rt}\,dt = \frac{s}{r}(1-e^{-rt}) \qquad \text{[cf. (14.12)]}$$

Thus the *net* present value—what the dealer would seek to maximize—can be expressed as

$$N(t) = V(t)e^{-rt} - \frac{s}{r}(1-e^{-rt}) - C = \left[V(t)+\frac{s}{r}\right]e^{-rt} - \frac{s}{r} - C$$

which is an objective function in a single choice variable t.

To maximize $N(t)$, the value of t must be chosen such that $N'(t) = 0$. This first derivative is

$$N'(t) = V'(t)e^{-rt} - r\left[V(t)+\frac{s}{r}\right]e^{-rt} \qquad \text{[product rule]}$$

$$= [V'(t) - rV(t) - s]e^{-rt}$$

and it will be zero if and only if

$$V'(t) = rV(t) + s$$

Thus, this last equation may be taken as the necessary optimization condition for the choice of the time of sale t^*.

The economic interpretation of this condition appeals easily to intuitive reasoning: $V'(t)$ represents the rate of change of the sale value, or the increment in V, if sale is postponed for a year, while the two terms on the right indicate, respectively, the increments in the interest cost and the storage cost entailed by such a postponement of sale (revenue and cost are both reckoned at time t^*). So, the idea of the equating of the two sides is to us just some "old wine in a new bottle," for it is nothing but the same MC = MR condition in a different guise!

Present Value of a Perpetual Flow

If a cash flow were to persist forever—a situation exemplified by the interest from a perpetual bond or the revenue from an indestructible capital asset such as land—the present value of the flow would be

$$\Pi = \int_0^{\infty} R(t)e^{-rt}\, dt$$

which is an improper integral.

Example 8 Find the present value of a perpetual income stream flowing at the uniform rate of D dollars per year, if the continuous rate of discount is r. Since, in evaluating an improper integral, we simply take the limit of a proper integral, the result in (14.12) can still be of help. Specifically, we can write

$$\Pi = \int_0^{\infty} De^{-rt}\, dt = \lim_{y\to\infty}\int_0^{y} De^{-rt}\, dt = \lim_{y\to\infty}\frac{D}{r}(1 - e^{-ry}) = \frac{D}{r}$$

Note that the y parameter (number of years) has disappeared from the final answer. This is as it should be, for here we are dealing with a *perpetual* flow. You may also observe that our result (present value = rate of revenue flow ÷ rate of discount) corresponds precisely to the familiar formula for the so-called capitalization of an asset with a perpetual yield.

EXERCISE 14.5

1. Given the following marginal-revenue functions:
   (*a*) $R'(Q) = 28Q - e^{0.3Q}$ (*b*) $R'(Q) = 10(1 + Q)^{-2}$
   find in each case the total-revenue function $R(Q)$. What initial condition can you introduce to definitize the constant of integration?
2. (*a*) Given the marginal propensity to import $M'(Y) = 0.1$ and the information that $M = 20$ when $Y = 0$, find the import function $M(Y)$.
   (*b*) Given the marginal propensity to consume $C'(Y) = 0.8 + 0.1Y^{-1/2}$ and the information that $C = Y$ when $Y = 100$, find the consumption function $C(Y)$.
3. Assume that the rate of investment is described by the function $I(t) = 12t^{1/3}$ and that $K(0) = 25$:
   (*a*) Find the time path of capital stock K.
   (*b*) Find the amount of capital accumulation during the time intervals [0, 1] and [1, 3], respectively.
4. *Given a continuous income stream at the constant rate of $1,000 per year:*
   (*a*) What will be the present value Π if the income stream lasts for 2 years and the continuous discount rate is 0.05 per year?
   (*b*) What will be the present value Π if the income stream terminates after exactly 3 years and the discount rate is 0.04?
5. What is the present value of a perpetual cash flow of:
   (*a*) $1,450 per year, discounted at $r = 5\%$?
   (*b*) $2,460 per year, discounted at $r = 8\%$?

14.6 Domar Growth Model

In the population-growth problem of (14.1) and (14.2) and the capital-formation problem of (14.10), the common objective is to delineate a time path on the basis of some given pattern of change of a variable. In the classic growth model of Professor Domar,[†] on the other hand, the idea is to stipulate the type of time path required to prevail if a certain equilibrium condition of the economy is to be satisfied.

The Framework

The basic premises of the Domar model are as follows:

1. Any change in the rate of investment flow per year $I(t)$ will produce a dual effect: it will affect the aggregate demand as well as the productive capacity of the economy.
2. The demand effect of a change in $I(t)$ operates through the multiplier process, assumed to work instantaneously. Thus an increase in $I(t)$ will raise the rate of income flow per year $Y(t)$ by a multiple of the increment in $I(t)$. The multiplier is $k = 1/s$, where s stands for the given (constant) marginal propensity to save. On the assumption that $I(t)$ is the only (parametric) expenditure flow that influences the rate of income flow, we can then state that

$$\frac{dY}{dt} = \frac{dI}{dt}\frac{1}{s} \tag{14.13}$$

3. The capacity effect of investment is to be measured by the change in the rate of *potential* output the economy is capable of producing. Assuming a constant capacity-capital ratio, we can write

$$\frac{\kappa}{K} \equiv \rho \qquad (= \text{a constant})$$

where κ (the Greek letter kappa) stands for capacity or potential output flow per year, and ρ (the Greek letter rho) denotes the given capacity-capital ratio. This implies, of course, that with a capital stock $K(t)$ the economy is potentially capable of producing an annual product, or income, amounting to $\kappa \equiv \rho K$ dollars. Note that, from $\kappa \equiv \rho K$ (the production function), it follows that $d\kappa = \rho\, dK$, and

$$\frac{d\kappa}{dt} = \rho\frac{dK}{dt} = \rho I \tag{14.14}$$

In Domar's model, equilibrium is defined to be a situation in which productive capacity is fully utilized. To have equilibrium is, therefore, to require the aggregate demand to be exactly equal to the potential output producible in a year; that is, $Y = \kappa$. If we start initially from an equilibrium situation, however, the requirement will reduce to the balancing of the respective *changes* in capacity and in aggregate demand; that is,

$$\frac{dY}{dt} = \frac{d\kappa}{dt} \tag{14.15}$$

[†] Evsey D. Domar, "Capital Expansion, Rate of Growth, and Employment," *Econometrica*, April 1946, pp. 137–147; reprinted in Domar, *Essays in the Theory of Economic Growth*, Oxford University Press, Fair Lawn, N.J., 1957, pp. 70–82.

What kind of time path of investment $I(t)$ can satisfy this equilibrium condition at all times?

Finding the Solution

To answer this question, we first substitute (14.13) and (14.14) into the equilibrium condition (14.15). The result is the following differential equation:

$$\frac{dI}{dt}\frac{1}{s} = \rho I \qquad \text{or} \qquad \frac{1}{I}\frac{dI}{dt} = \rho s \tag{14.16}$$

Since (14.16) specifies a definite pattern of change for I, we should be able to find the equilibrium (or required) investment path from it.

In this simple case, the solution is obtainable by directly integrating both sides of the second equation in (14.16) with respect to t. The fact that the two sides are identical in equilibrium assures the equality of their integrals. Thus,

$$\int \frac{1}{I}\frac{dI}{dt}\,dt = \int \rho s\,dt$$

By the substitution rule and the log rule, the left side gives us

$$\int \frac{dI}{I} = \ln|I| + c_1 \qquad (I \neq 0)$$

whereas the right side yields (ρs being a constant)

$$\int \rho s\,dt = \rho s t + c_2$$

Equating the two results and combining the two constants, we have

$$\ln|I| = \rho s t + c \tag{14.17}$$

To obtain $|I|$ from $\ln|I|$, we perform an operation known as "taking the antilog of $\ln|I|$," which utilizes the fact that $e^{\ln x} = x$. Thus, letting each side of (14.17) become the exponent of the constant e, we obtain

$$e^{\ln|I|} = e^{(\rho s t + c)}$$

or

$$|I| = e^{\rho s t}e^{c} = Ae^{\rho s t} \qquad \text{where } A \equiv e^{c}$$

If we take investment to be positive, then $|I| = I$, so that the preceding result becomes $I(t) = Ae^{\rho s t}$, where A is arbitrary. To get rid of this arbitrary constant, we set $t = 0$ in the equation $I(t) = Ae^{\rho s t}$, to get $I(0) = Ae^{0} = A$. This definitizes the constant A, and enables us to express the solution—the required investment path—as

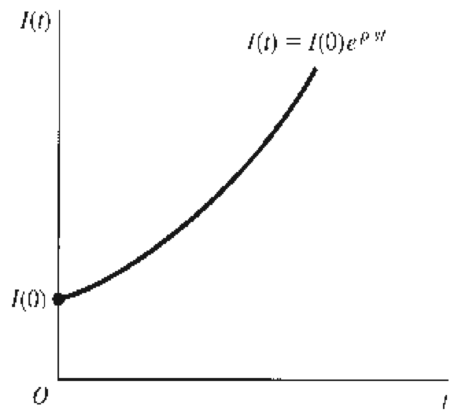
$$I(t) = I(0)e^{\rho s t} \tag{14.18}$$

where $I(0)$ denotes the initial rate of investment.†

This result has a disquieting economic meaning. In order to maintain the balance between capacity and demand over time, the rate of investment flow must grow precisely

† The solution (14.18) will remain valid even if we let investment be negative in the result $|I| = Ae^{\rho s t}$. See Exercise 14.6-3.

FIGURE 14.6



at the exponential rate of ρs, along a path such as illustrated in Fig. 14.6. Obviously, the larger the capacity-capital ratio or the marginal propensity to save, the larger the required rate of growth will be. But at any rate, once the values of ρ and s are known, the required growth path of investment becomes very rigidly set.

The Razor's Edge

It now becomes relevant to ask what will happen if the *actual* rate of growth of investment—call that rate r—differs from the *required* rate ρs.

Domar's approach is to define a *coefficient of utilization*

$$u = \lim_{t\to\infty} \frac{Y(t)}{\kappa(t)} \qquad [u = 1 \text{ means full utilization of capacity}]$$

and show that $u = r/\rho s$, so that $u \gtreqless 1$ as $r \gtreqless \rho s$. In other words, if there is a discrepancy between the actual and required rates ($r \neq \rho s$), we will find in the end (as $t \to \infty$) either a shortage of capacity ($u > 1$) or a surplus of capacity ($u < 1$), depending on whether r is greater or less than ρs.

We can show, however, that the conclusion about capacity shortage and surplus really applies at any time t, not only as $t \to \infty$. For a given growth rate r implies that

$$I(t) = I(0)e^{rt} \qquad \text{and} \qquad \frac{dI}{dt} = rI(0)e^{rt}$$

Therefore, by (14.13) and (14.14), we have

$$\frac{dY}{dt} = \frac{1}{s}\frac{dI}{dt} = \frac{r}{s}I(0)e^{rt}$$

$$\frac{d\kappa}{dt} = \rho I(t) = \rho I(0)e^{rt}$$

The ratio between these two derivatives,

$$\frac{dY/dt}{d\kappa/dt} = \frac{r}{\rho s}$$

should tell us the relative magnitudes of the demand-creating effect and the capacity-generating effect of investment at any time t, under the actual growth rate of r. If r (the actual rate) exceeds ρs (the required rate), then $dY/dt > d\kappa/dt$, and the demand effect will outstrip the capacity effect, causing a shortage of capacity. Conversely, if $r < \rho s$, there will be a deficiency in aggregate demand and, hence, a surplus of capacity.

The curious thing about this conclusion is that if investment actually grows at a *faster* rate than required ($r > \rho s$), the end result will be a *shortage* rather than a surplus of capacity. It is equally curious that if the actual growth of investment lags behind the required rate ($r < \rho s$), we will encounter a capacity *surplus* rather than a shortage. Indeed, because of such paradoxical results, if we now allow the entrepreneurs to adjust the actual growth rate r (hitherto taken to be a constant) according to the prevailing capacity situation, they will most certainly make the "wrong" kind of adjustment. In the case of $r > \rho s$, for instance, the emergent capacity shortage will motivate an even faster rate of investment. But this would mean an increase in r, instead of the reduction called for under the circumstances. Consequently, the discrepancy between the two rates of growth would be intensified rather than reduced.

The upshot is that, given the parametric constants ρ and s, the only way to avoid both shortage and surplus of productive capacity is to guide the investment flow ever so carefully along the equilibrium path with a growth rate $r^* = \rho s$. And, as we have shown, any deviation from such a "razor's edge" time path will bring about a *persistent failure* to satisfy the norm of full utilization which Domar envisaged in this model. This is perhaps not too cheerful a prospect to contemplate. Fortunately, more flexible results become possible when certain assumptions of the Domar model are modified, as we shall see from the growth model of Professor Solow, to be discussed in Chap. 15.

EXERCISE 14.6

1. How many factors of production are explicitly considered in the Domar model? What does this fact imply with regard to the capital-labor ratio in production?
2. We learned in Sec. 10.2 that the constant r in the exponential function Ae^{rt} represents *the rate of growth of the function*. Apply this to (14.16), and deduce (14.18) without going through integration.
3. Show that even if we let investment be negative in the equation $|I| = Ae^{\rho st}$, upon definitizing the arbitrary constant A we will still end up with the solution (14.18).
4. Show that the result in (14.18) can be obtained alternatively by finding—and equating—the *definite* integrals of both sides of (14.16),

   $$\frac{1}{I}\frac{dI}{dt} = \rho s$$

   with respect to the variable t, with limits of integration $t = 0$ and $t = t$. Remember that when we change the variable of integration from t to I, the limits of integration will change from $t = 0$ and $t = t$, respectively, to $I = I(0)$ and $I = I(t)$.