Chapter 17

Discrete Time: First-Order Difference Equations

In the continuous-time context, the pattern of change of a variable y is embodied in the derivatives $y'(t)$, $y''(t)$, etc. The time change involved in these is occurring continuously. When time is, instead, taken to be a *discrete* variable, so that the variable t is allowed to take integer values only, the concept of the derivative obviously will no longer be appropriate. Then, as we shall see, the pattern of change of the variable y must be described by so-called differences, rather than by derivatives or differentials, of $y(t)$. Accordingly, the techniques of differential equations will give way to those of *difference equations*.

When we are dealing with discrete time, the value of variable y will change only when the variable t changes from one integer value to the next, such as from $t = 1$ to $t = 2$. Meanwhile, nothing is supposed to happen to y. In this light, it becomes more convenient to interpret the values of t as referring to *periods*—rather than *points*—of time, with $t = 1$ denoting period 1 and $t = 2$ denoting period 2, and so forth. Then we may simply regard y as having one unique value in each time period. In view of this interpretation, the discrete-time version of economic dynamics is often referred to as *period analysis*. It should be emphasized, however, that "period" is being used here not in the calendar sense but in the analytical sense. Hence, a period may involve one extent of calendar time in a particular economic model, but an altogether different one in another. Even in the same model, moreover, each successive period should not necessarily be construed as meaning equal calendar time. In the analytical sense, a period is merely a length of time that elapses before the variable y undergoes a change.

17.1 Discrete Time, Differences, and Difference Equations

The change from continuous time to discrete time produces no effect on the fundamental nature of dynamic analysis, although the formulation of the problem must be altered. Basically, our dynamic problem is still to find a time path from some given pattern of change of a variable y over time. But the pattern of change should now be represented by the difference quotient $\Delta y/\Delta t$, which is the discrete-time counterpart of the derivative dy/dt. Recall, however, that t can now take only integer values; thus, when we are comparing the values of y in two consecutive periods, we must have $\Delta t = 1$. For this reason, the difference quotient $\Delta y/\Delta t$ can be simplified to the expression Δy; this is called the *first difference* of y. The symbol Δ, meaning difference, can accordingly be interpreted as a directive to take the first difference of (y). As such, it constitutes the discrete-time counterpart of the operator symbol d/dt.

The expression Δy can take various values, of course, depending on which two consecutive time periods are involved in the difference-taking (or "differencing"). To avoid ambiguity, let us add a time subscript to y and define the first difference more specifically, as follows:

$$\Delta y_t \equiv y_{t+1} - y_t \tag{17.1}$$

where y_t means the value of y in the tth period, and y_{t+1} is its value in the period immediately following the tth period. With this symbology, we may describe the pattern of change of y by an equation such as

$$\Delta y_t = 2 \tag{17.2}$$

or

$$\Delta y_t = -0.1 y_t \tag{17.3}$$

Equations of this type are called *difference equations*. Note the striking resemblance between the last two equations, on the one hand, and the differential equations $dy/dt = 2$ and $dy/dt = -0.1y$ on the other.

Even though difference equations derive their name from difference expressions such as Δy_t, there are alternate equivalent forms of such equations which are completely free of Δ expressions and which are more convenient to use. By virtue of (17.1), we can rewrite (17.2) as

$$y_{t+1} - y_t = 2 \tag{17.2'}$$

or

$$y_{t+1} = y_t + 2 \tag{17.2''}$$

For (17.3), the corresponding alternate equivalent forms are

$$y_{t+1} - 0.9 y_t = 0 \tag{17.3'}$$

or

$$y_{t+1} = 0.9 y_t \tag{17.3''}$$

The double-prime-numbered versions will prove convenient when we are calculating a y value from a known y value of the preceding period. In later discussions, however, we shall employ mostly the single-prime-numbered versions, i.e., those of (17.2′) and (17.3′).

It is important to note that the choice of time subscripts in a difference equation is somewhat arbitrary. For instance, without any change in meaning, (17.2′) can be rewritten as $y_t - y_{t-1} = 2$, where $(t-1)$ refers to the period which immediately precedes the tth. Or, we may express it equivalently as $y_{t+2} - y_{t+1} = 2$.

Also, it may be pointed out that, although we have consistently used subscripted y symbols, it is also acceptable to use $y(t)$, $y(t+1)$, and $y(t-1)$ in their stead. In order to avoid using the notation $y(t)$ for both continuous-time and discrete-time cases, however, we shall, in the discussion of period analysis, adhere to the subscript device.

Analogous to differential equations, difference equations can be either linear or nonlinear, homogeneous or nonhomogeneous, and of the first or second (or higher) orders. Take (17.2′) for instance. It can be classified as: (1) linear, for no y term (of any period) is raised to the second (or higher) power or is multiplied by a y term of another period; (2) nonhomogeneous, since the right-hand side (where there is no y term) is nonzero; and (3) of the first order, because there exists only a *first difference* Δy_t, involving a one-period time lag only. (In contrast, a second-order difference equation, to be discussed in Chap. 18, involves a two-period lag and thus entails three y terms: y_{t+2}, y_{t+1}, as well as y_t.)

Actually, (17.2′) can also be characterized as having constant coefficients and a constant term (= 2). Since the constant-coefficient case is the only one we shall consider, this characterization will henceforth be implicitly assumed. Throughout the present chapter, the constant-term feature will also be retained, although a method of dealing with the variable-term case will be discussed in Chap. 18.

Check that the equation (17.3′) is also linear and of the first order; but unlike (17.2′), it is homogeneous.

17.2 Solving a First-Order Difference Equation

In solving a differential equation, our objective was to find a time path $y(t)$. As we know, such a time path is a function of time which is totally free from any derivative (or differential) expressions and which is perfectly consistent with the given differential equation as well as with its initial conditions. The time path we seek from a difference equation is similar in nature. Again, it should be a function of t—a formula defining the values of y in every time period—which is consistent with the given difference equation as well as with its initial conditions. Besides, it must not contain any difference expressions such as Δy_t (or expressions like $y_{t+1} - y_t$).

Solving differential equations is, in the final analysis, a matter of integration. How do we solve a difference equation?

Iterative Method

Before developing a general method of attack, let us first explain a relatively pedestrian method, the *iterative method*—which, though crude, will prove immensely revealing of the essential nature of a so-called solution.

In this chapter we are concerned only with the first-order case; thus the difference equation describes the pattern of change of y between *two* consecutive periods only. Once such a pattern is specified, such as by (17.2″), and once we are given an initial value y_0, it is no problem to find y_1 from the equation. Similarly, once y_1 is found, y_2 will be immediately obtainable, and so forth, by repeated application (iteration) of the pattern of change specified in the difference equation. The results of iteration will then permit us to infer a time path.

Example 1

Find the solution of the difference equation (17.2), assuming an initial value of $y_0 = 15$. To carry out the iterative process, it is more convenient to use the alternative form of the difference equation (17.2″), namely, $y_{t+1} = y_t + 2$, with $y_0 = 15$. From this equation, we can deduce step-by-step that

$$y_1 = y_0 + 2$$
$$y_2 = y_1 + 2 = (y_0 + 2) + 2 = y_0 + 2(2)$$
$$y_3 = y_2 + 2 = [y_0 + 2(2)] + 2 = y_0 + 3(2)$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$

and, in general, for any period t,

$$y_t = y_0 + t(2) = 15 + 2t \tag{17.4}$$

This last equation indicates the y value of any time period (including the initial period $t = 0$); it therefore constitutes the solution of (17.2).

The process of iteration is crude- it corresponds roughly to solving simple differential equations by straight integration—but it serves to point out clearly the manner in which a time path is generated. In general, the value of y_t will depend in a specified way on the value of y in the immediately preceding period (y_{t-1}); thus a given initial value y_0 will successively lead to $y_1, y_2, \ldots,$ via the prescribed pattern of change.

Example 2

Solve the difference equation (17.3); this time, let the initial value be unspecified and denoted simply by y_0. Again it is more convenient to work with the alternative version in (17.3″), namely, $y_{t+1} = 0.9y_t$. By iteration, we have

$$y_1 = 0.9y_0$$
$$y_2 = 0.9y_1 = 0.9(0.9y_0) = (0.9)^2 y_0$$
$$y_3 = 0.9y_2 = 0.9(0.9)^2 y_0 = (0.9)^3 y_0$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$

These can be summarized into the solution

$$y_t = (0.9)^t y_0 \tag{17.5}$$

To heighten interest, we can lend some economic content to this example. In the simple multiplier analysis, a single investment expenditure in period 0 will call forth successive rounds of spending, which in turn will bring about varying amounts of income increment in succeeding time periods. Using y to denote *income increment,* we have y_0 = the amount of investment in period 0; but the subsequent income increments will depend on the marginal propensity to consume (MPC). If MPC = 0.9 and if the income of each period is consumed only in the next period, then 90 percent of y_0 will be consumed in period 1, resulting in an income increment in period 1 of $y_1 = 0.9y_0$. By similar reasoning, we can find $y_2 = 0.9y_1$, etc. These, we see, are precisely the results of the iterative process cited previously. In other words, the multiplier process of income generation can be described by a difference equation such as (17.3″), and a solution like (17.5) will tell us what the magnitude of income increment is to be in any time period t.

Example 3

Solve the homogeneous difference equation

$$my_{t+1} - ny_t = 0$$

Upon normalizing and transposing, this may be written as

$$y_{t+1} = \left(\frac{n}{m}\right) y_t$$

which is the same as (17.3″) in Example 2 except for the replacement of 0.9 by n/m. Hence, by analogy, the solution should be

$$y_t = \left(\frac{n}{m}\right)^t y_0$$

Watch the term $\left(\frac{n}{m}\right)^t$. It is through this term that various values of t will lead to their corresponding values of y. It therefore corresponds to the expression e^{rt} in the solutions to differential equations. If we write it more generally as b^t (b for base) and attach the more general multiplicative constant A (instead of y_0), we see that the solution of the general homogeneous difference equation of Example 3 will be in the form

$$y_t = Ab^t$$

We shall find that this expression Ab^t will play the same important role in difference equations as the expression Ae^{rt} did in differential equations.† However, even though both are exponential expressions, the former is to the base b, whereas the latter is to the base e. It stands to reason that, just as the type of the continuous-time path $y(t)$ depends heavily on the value of r, the discrete-time path y_t hinges principally on the value of b.

General Method

By this time, you must have become quite impressed with the various similarities between differential and difference equations. As might be conjectured, the general method of solution presently to be explained will parallel that for differential equations.

Suppose that we are seeking the solution to the first-order difference equation

$$y_{t+1} + ay_t = c \tag{17.6}$$

where a and c are two constants. The general solution will consist of the sum of two components: a *particular solution* y_p, which is *any* solution of the complete nonhomogeneous equation (17.6), and a *complementary function* y_c, which is the general solution of the reduced equation of (17.6):

$$y_{t+1} + ay_t = 0 \tag{17.7}$$

The y_p component again represents the intertemporal equilibrium level of y, and the y_c component, the deviations of the time path from that equilibrium. The sum of y_c and y_p constitutes the *general* solution, because of the presence of an arbitrary constant. As before, in order to definitize the solution, an initial condition is needed.

Let us first deal with the complementary function. Our experience with Example 3 suggests that we may try a solution of the form $y_t = Ab^t$ (with $Ab^t \neq 0$, for otherwise y_t will turn out simply to be a horizontal straight line lying on the t axis); in that case, we also

† You may object to this statement by pointing out that the solution (17.4) in Example 1 does not contain a term in the form of Ab^t. This latter fact, however, arises only because in Example 1 we have $b = n/m = 1/1 = 1$, so that the term Ab^t reduces to a constant.

have $y_{t+1} = Ab^{t+1}$. If these values of y_t and y_{t+1} hold, the homogeneous equation (17.7) will become

$$Ab^{t+1} + aAb^t = 0$$

which, upon canceling the nonzero common factor Ab^t, yields

$$b + a = 0 \qquad \text{or} \qquad b = -a$$

This means that, for the trial solution to work, we must set $b = -a$; then the complementary function should be written as

$$y_c(= Ab^t) = A(-a)^t$$

Now let us search for the particular solution, which has to do with the complete equation (17.6). In this regard, Example 3 is of no help at all, because that example relates only to a homogeneous equation. However, we note that for y_p we can choose *any* solution of (17.6); thus if a trial solution of the simplest form $y_t = k$ (a constant) can work out, no real difficulty will be encountered. Now, if $y_t = k$, then y will maintain the same constant value over time, and we must have $y_{t+1} = k$ also. Substitution of these values into (17.6) yields

$$k + ak = c \qquad \text{and} \qquad k = \frac{c}{1+a}$$

Since this particular k value satisfies the equation, the particular integral can be written as

$$y_p(= k) = \frac{c}{1+a} \qquad (a \neq -1)$$

This being a constant, a stationary equilibrium is indicated in this case.

If it happens that $a = -1$, as in Example 1, however, the particular solution $c/(1+a)$ is not defined, and some other solution of the nonhomogeneous equation (17.6) must be sought. In this event, we employ the now-familiar trick of trying a solution of the form $y_t = kt$. This implies, of course, that $y_{t+1} = k(t+1)$. Substituting these into (17.6), we find

$$k(t+1) + akt = c \qquad \text{and} \qquad k = \frac{c}{t+1+at} = c \qquad [\text{because } a = -1]$$

thus $$y_p(= kt) = ct$$

This form of the particular solution is a nonconstant function of t; it therefore represents a moving equilibrium.

Adding y_c and y_p together, we may now write the general solution in one of the two following forms:

$$y_t = A(-a)^t + \frac{c}{1+a} \qquad [\text{general solution, case of } a \neq -1] \qquad \textbf{(17.8)}$$

$$y_t = A(-a)^t + ct = A + ct \qquad [\text{general solution, case of } a = -1] \qquad \textbf{(17.9)}$$

Neither of these is completely determinate, in view of the arbitrary constant A. To eliminate this arbitrary constant, we resort to the initial condition that $y_t = y_0$ when $t = 0$. Letting $t = 0$ in (17.8), we have

$$y_0 = A + \frac{c}{1+a} \qquad \text{and} \qquad A = y_0 - \frac{c}{1+a}$$

Consequently, the definite version of (17.8) is

$$y_t = \left(y_0 - \frac{c}{1+a}\right)(-a)^t + \frac{c}{1+a} \qquad \text{[definite solution, case of } a \neq -1\text{]} \qquad \textbf{(17.8')}$$

Letting $t = 0$ in (17.9), on the other hand, we find $y_0 = A$, so the definite version of (17.9) is

$$y_t = y_0 + ct \qquad \text{[definite solution, case of } a = -1\text{]} \qquad \textbf{(17.9')}$$

If this last result is applied to Example 1, the solution that emerges is exactly the same as the iterative solution (17.4).

You can check the validity of each of these solutions by the following two steps. First, by letting $t = 0$ in (17.8′), see that the latter equation reduces to the identity $y_0 = y_0$, signifying the satisfaction of the initial condition. Second, by substituting the y_t formula (17.8′) and a similar y_{t+1} formula—obtained by replacing t with $(t + 1)$ in (17.8′)—into (17.6), see that the latter reduces to the identity $c = c$, signifying that the time path is consistent with the given difference equation. The check on the validity of solution (17.9′) is analogous.

Example 4

Solve the first-order difference equation

$$y_{t+1} - 5y_t = 1 \qquad \left(y_0 = \tfrac{7}{4}\right)$$

Following the procedure used in deriving (17.8′), we can find y_c by trying a solution $y_t = Ab^t$ (which implies $y_{t+1} = Ab^{t-1}$). Substituting these values into the homogeneous version $y_{t-1} - 5y_t = 0$ and canceling the common factor Ab^t, we get $b = 5$. Thus

$$y_c = A(5)^t$$

To find y_p, try the solution $y_t = k$, which implies $y_{t+1} = k$. Substituting these into the complete difference equation, we find $k = -\frac{1}{4}$. Hence

$$y_p = -\tfrac{1}{4}$$

It follows that the general solution is

$$y_t = y_c + y_p = A(5)^t - \tfrac{1}{4}$$

Letting $t = 0$ here and utilizing the initial condition $y_0 = \frac{7}{4}$, we obtain $A = 2$. Thus the definite solution may finally be written as

$$y_t = 2(5)^t - \tfrac{1}{4}$$

Since the given difference equation of this example is a special case of (17.6), with $a = -5$, $c = 1$, and $y_0 = \frac{7}{4}$, and since (17.8′) is the solution "formula" for this type of difference equation, we could have found our solution by inserting the specific parameter values into (17.8′), with the result that

$$y_t = \left(\frac{7}{4} - \frac{1}{1-5}\right)(5)^t + \frac{1}{1-5} = 2(5)^t - \frac{1}{4}$$

which checks perfectly with the earlier answer.

Note that the y_{t+1} term in (17.6) has a unit coefficient. If a given difference equation has a nonunit coefficient for this term, it must be normalized before using the solution formula (17.8′).

EXERCISE 17.2

1. Convert the following difference equations into the form of (17.2″):
   (*a*) $\Delta y_t = 7$
   (*b*) $\Delta y_t = 0.3y_t$
   (*c*) $\Delta y_t = 2y_t - 9$
2. Solve the following difference equations by iteration:
   (*a*) $y_{t+1} = y_t - 1 \qquad (y_0 = 10)$
   (*b*) $y_{t+1} = \alpha y_t \qquad (y_0 = \beta)$
   (*c*) $y_{t+1} = \alpha y_t - \beta \qquad (y_t = y_0 \text{ when } t = 0)$
3. Rewrite the equations in Prob. 2 in the form of (17.6), and solve by applying formula (17.8′) or (17.9′), whichever is appropriate. Do your answers check with those obtained by the iterative method?
4. For each of the following difference equations, use the procedure illustrated in the derivation of (17.8′) and (17.9′) to find y_c, y_p, and the definite solution:
   (*a*) $y_{t+1} + 3y_t = 4 \qquad (y_0 = 4)$
   (*b*) $2y_{t+1} - y_t = 6 \qquad (y_0 = 7)$
   (*c*) $y_{t+1} = 0.2y_t + 4 \qquad (y_0 = 4)$

17.3 The Dynamic Stability of Equilibrium

In the continuous-time case, the dynamic stability of equilibrium depends on the Ae^{rt} term in the complementary function. In period analysis, the corresponding role is played by the Ab^t term in the complementary function. Since its interpretation is somewhat more complicated than Ae^{rt}, let us try to clarify it before proceeding further.

The Significance of *b*

Whether the equilibrium is dynamically stable is a question of whether or not the complementary function will tend to zero as $t \to \infty$. Basically, we must analyze the path of the term Ab^t as t is increased indefinitely. Obviously, the value of b (the base of this exponential term) is of crucial importance in this regard. Let us first consider its significance alone, by disregarding the coefficient A (by assuming $A = 1$).

For analytical purposes, we can divide the range of possible values of b, $(-\infty, +\infty)$, into seven distinct regions, as set forth in the first two columns of Table 17.1, arranged in descending order of magnitude of b. These regions are also marked off in Fig. 17.1 on a vertical b scale, with the points $+1$, 0, and -1 as the demarcation points. In fact, these latter three points in themselves constitute the regions II, IV, and VI. Regions III and V, on the other hand, correspond to the set of all positive fractions and the set of all negative fractions, respectively. The remaining two regions, I and VII, are where the numerical value of b exceeds unity.

In each region, the exponential expression b^t generates a different type of time path. These are exemplified in Table 17.1 and illustrated in Fig. 17.1. In region I (where $b > 1$), b^t must increase with t at an increasing pace. The general configuration of the time path will therefore assume the shape of the top graph in Fig. 17.1. Note that this graph is shown

TABLE 17.1
A Classification of the Values of b

				Value of b^t in Different Time Periods				
Region	Value of b		Value of b^t	$t=0$	$t=1$	$t=2$	$t=3$	$t=4\cdots$
I	$b>1$	$(\lvert b\rvert>1)$	e.g., $(2)^t$	1	2	4	8	16
II	$b=1$	$(\lvert b\rvert=1)$	$(1)^t$	1	1	1	1	1
III	$0<b<1$	$(\lvert b\rvert<1)$	e.g., $\left(\frac{1}{2}\right)^t$	1	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$
IV	$b=0$	$(\lvert b\rvert=0)$	$(0)^t$	0	0	0	0	0
V	$-1<b<0$	$(\lvert b\rvert<1)$	e.g., $\left(-\frac{1}{2}\right)^t$	1	$-\frac{1}{2}$	$\frac{1}{4}$	$-\frac{1}{8}$	$\frac{1}{16}$
VI	$b=-1$	$(\lvert b\rvert=1)$	$(-1)^t$	1	−1	1	−1	1
VII	$b<-1$	$(\lvert b\rvert>1)$	e.g., $(-2)^t$	1	−2	4	−8	16

as a step function rather than as a smooth curve; this is because we are dealing with period analysis. In region II ($b=1$), b^t will remain at unity for all values of t. Its graph will thus be a horizontal straight line. Next, in region III, b^t represents a positive fraction raised to integer powers. As the power is increased, b^t must decrease, though it will always remain positive. The next case, that of $b=0$ in region IV, is quite similar to the case of $b=1$; but here we have $b^t=0$ rather than $b^t=1$, so its graph will coincide with the horizontal axis. However, this case is of peripheral interest only, since we have earlier adopted the assumption that $Ab^t\neq 0$.

When we move into the negative regions, an interesting new phenomenon occurs: The value of b^t will *alternate* between positive and negative values from period to period! This fact is clearly brought out in the last three rows of Table 17.1 and in the last three graphs of Fig. 17.1. In region V, where b is a negative fraction, the alternating time path tends to get closer and closer to the horizontal axis (cf. the positive-fraction region, III). In contrast, when $b=-1$ (region VI), a perpetual alternation between -1 and -1 results. And finally, when $b<-1$ (region VII), the alternating time path will deviate farther and farther from the horizontal axis.

What is striking is that, whereas the phenomenon of a fluctuating time path cannot possibly arise from a single Ae^{rt} term (the complex-root case of the second-order differential equation requires a *pair* of complex roots), fluctuation can be generated by a single b^t (or Ab^t) term. Note, however, that the character of the fluctuation is somewhat different; unlike the circular-function pattern, the fluctuation depicted in Fig. 17.1 is nonsmooth. For this reason, we shall employ the word *oscillation* to denote the new, nonsmooth type of fluctuation, even though many writers do use the terms fluctuation and oscillation interchangeably.

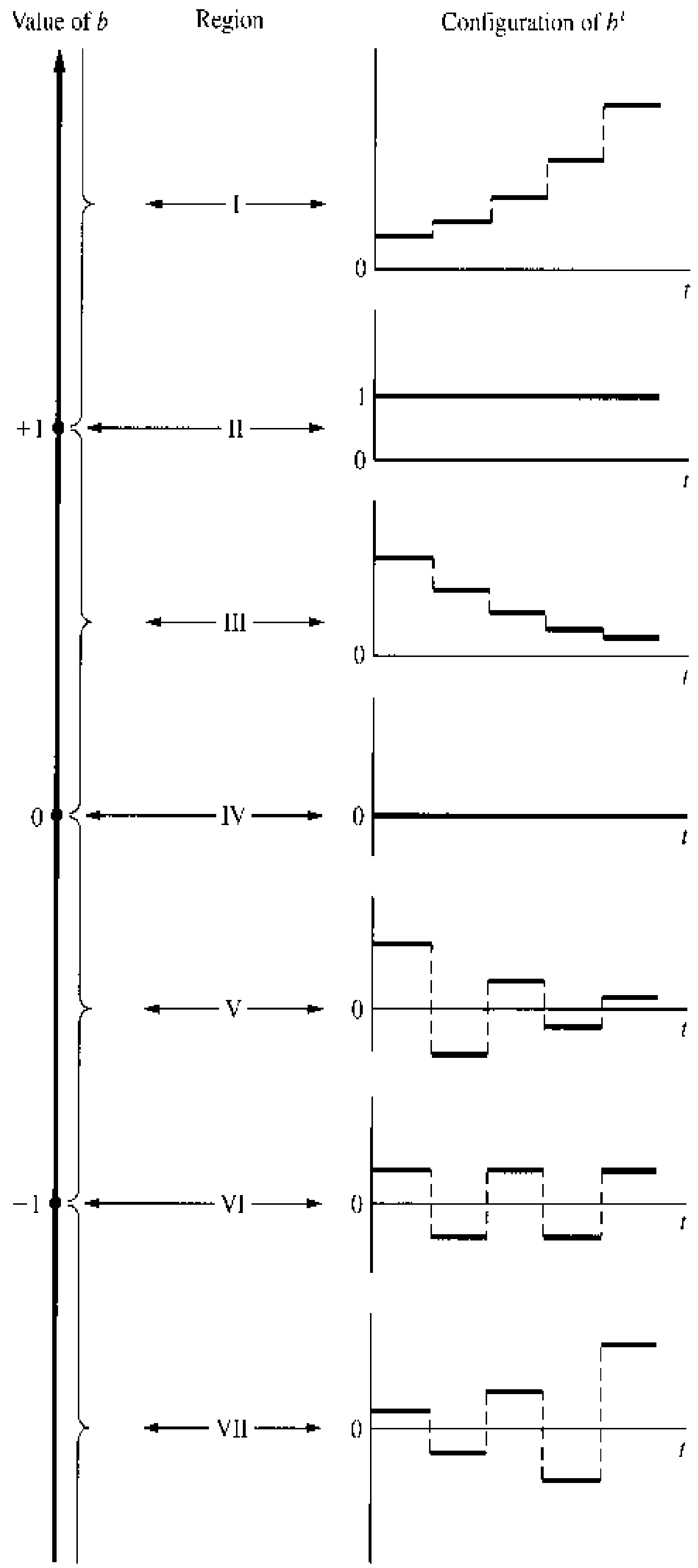
The essence of the preceding discussion can be conveyed in the following general statement: The time path of b^t $(b\neq 0)$ will be

$$\left.\begin{array}{l}\text{Nonoscillatory}\\ \text{Oscillatory}\end{array}\right\}\quad\text{if}\quad\left\{\begin{array}{l}b>0\\ b<0\end{array}\right.$$

$$\left.\begin{array}{l}\text{Divergent}\\ \text{Convergent}\end{array}\right\}\quad\text{if}\quad\left\{\begin{array}{l}|b|>1\\ |b|<1\end{array}\right.$$

It is important to note that, whereas the convergence of the expression e^{rt} depends on the *sign* of r, the convergence of the b^t expression hinges, instead, on the *absolute value* of b.

FIGURE 17.1



The Role of A

So far we have deliberately left out the multiplicative constant A. But its effects—of which there are two—are relatively easy to take into account. First, the *magnitude* of A can serve to "blow up" (if, say, $A = 3$) or "pare down" (if, say, $A = \frac{1}{5}$) the values of b^t. That is, it can produce a *scale effect* without changing the basic configuration of the time path. The *sign* of A, on the other hand, does materially affect the shape of the path because, if b^t is multiplied by $A = -1$, then each time path shown in Fig. 17.1 will be replaced by its own mirror image with reference to the horizontal axis. Thus, a negative A can produce a *mirror effect* as well as a scale effect.

Convergence to Equilibrium

The preceding discussion presents the interpretation of the Ab^t term in the complementary function, which, as we recall, represents the deviations from some intertemporal equilibrium level. If a term (say) $y_p = 5$ is added to the Ab^t term, the time path must be shifted up vertically by a constant value of 5. This will in no way affect the convergence or divergence of the time path, but it will alter the level with reference to which convergence or divergence in gauged. What Fig. 17.1 pictures is the convergence (or lack of it) of the Ab^t expression to zero. When the y_p is included, it becomes a question of the convergence of the time path $y_t = y_c + y_p$ to the equilibrium level y_p.

In this connection, let us add a word of explanation for the special case of $b = 1$ (region II). A time path such as

$$y_t = A(1)^t + y_p = A + y_p$$

gives the impression that it converges, because the multiplicative term $(1)^t = 1$ produces no explosive effect. Observe, however, that y_t will now take the value $(A + y_p)$ rather than the equilibrium value y_p; in fact, it can never reach y_p (unless $A = 0$). As an illustration of this type of situation, we can cite the time path in (17.9), in which a moving equilibrium $y_p = ct$ is involved. This time path is to be considered divergent, not because of the appearance of t in the particular solution but because, with a nonzero A, there will be a constant deviation from the moving equilibrium. Thus, in stipulating the condition for convergence of time path y_t to the equilibrium y_p, we must rule out the case of $b = 1$.

In sum, the solution

$$y_t = Ab^t + y_p$$

is a convergent path if and only if $|b| < 1$.

Example 1 What kind of time path is represented by $y_t = 2(-\frac{4}{5})^t + 9$? Since $b = -\frac{4}{5} < 0$, the time path is oscillatory. But since $|b| = \frac{4}{5} < 1$, the oscillation is damped, and the time path converges to the equilibrium level of 9.

You should exercise care not to confuse $2(-\frac{4}{5})^t$ with $-2(\frac{4}{5})^t$; they represent entirely different time-path configurations.

Example 2 How do you characterize the time path $y_t = 3(2)^t + 4$? Since $b = 2 > 0$, no oscillation will occur. But since $|b| = 2 > 1$, the time path will diverge from the equilibrium level of 4.

EXERCISE 17.3

1. Discuss the nature of the following time paths:

   (a) $y_t = 3^t + 1$ (c) $y_t = 5\left(-\frac{1}{10}\right)^t + 3$

   (b) $y_t = 2\left(\frac{1}{3}\right)^t$ (d) $y_t = -3\left(\frac{1}{4}\right)^t + 2$

2. What is the nature of the time path obtained from each of the difference equations in Exercise 17.2-4?
3. Find the solutions of the following, and determine whether the time paths are oscillatory and convergent:

   (*a*) $y_{t+1} - \frac{1}{3}y_t = 6 \qquad (y_0 = 1)$

   (*b*) $y_{t+1} + 2y_t = 9 \qquad (y_0 = 4)$

   (*c*) $y_{t+1} + \frac{1}{4}y_t = 5 \qquad (y_0 = 2)$

   (*d*) $y_{t+1} - y_t = 3 \qquad (y_0 = 5)$

17.4 The Cobweb Model

To illustrate the use of first-order difference equations in economic analysis, we shall cite two variants of the market model for a single commodity. The first variant, known as the *cobweb model,* differs from our earlier market models in that it treats Q_s as a function not of the current price but of the price of the preceding time period.

The Model

Consider a situation in which the producer's output decision must be made one period in advance of the actual sale—such as in agricultural production, where planting must precede by an appreciable length of time the harvesting and sale of the output. Let us assume that the output decision in period t is based on the then-prevailing price P_t. Since this output will not be available for the sale until period $(t+1)$, however, P_t will determine not Q_{st}, but $Q_{s,t+1}$. Thus we now have a "lagged" supply function.[†]

$$Q_{s,t+1} = S(P_t)$$

or, equivalently, by shifting back the time subscripts by one period,

$$Q_{st} = S(P_{t-1})$$

When such a supply function interacts with a demand function of the form

$$Q_{dt} = D(P_t)$$

interesting dynamic price patterns will result.

Taking the linear versions of these (lagged) supply and (unlagged) demand functions, and assuming that in each time period the market price is always set at a level which clears the market, we have a market model with the following three equations:

$$\begin{aligned} Q_{dt} &= Q_{st} \\ Q_{dt} &= \alpha - \beta P_t \qquad (\alpha, \beta > 0) \\ Q_{st} &= -\gamma + \delta P_{t-1} \qquad (\gamma, \delta > 0) \end{aligned} \tag{17.10}$$

† We are making the implicit assumption here that the entire output of a period will be placed on the market, with no part of it held in storage. Such an assumption is appropriate when the commodity in question is perishable or when no inventory is ever kept. A model with inventory will be considered in Sec. 17.5.

By substituting the last two equations into the first, however, the model can be reduced to a single first-order difference equation as follows:

$$\beta P_t + \delta P_{t-1} = \alpha + \gamma$$

In order to solve this equation, it is desirable first to normalize it and shift the time subscripts ahead by one period [alter t to $(t+1)$, etc.]. The result,

$$P_{t+1} + \frac{\delta}{\beta} P_t = \frac{\alpha + \gamma}{\beta} \tag{17.11}$$

will then be a replica of (17.6), with the substitutions

$$y = P \qquad a = \frac{\delta}{\beta} \qquad \text{and} \qquad c = \frac{\alpha + \gamma}{\beta}$$

Inasmuch as δ and β are both positive, it follows that $a \neq -1$. Consequently, we can apply formula (17.8′), to get the time path

$$P_t = \left(P_0 - \frac{\alpha + \gamma}{\beta + \delta}\right)\left(-\frac{\delta}{\beta}\right)^t + \frac{\alpha + \gamma}{\beta + \delta} \tag{17.12}$$

where P_0 represents the initial price.

The Cobwebs

Three points may be observed in regard to this time path. In the first place, the expression $(\alpha + \gamma)/(\beta + \delta)$, which constitutes the particular integral of the difference equation, can be taken as the intertemporal equilibrium price of the model:[†]

$$\overline{P} = \frac{\alpha + \gamma}{\beta + \delta}$$

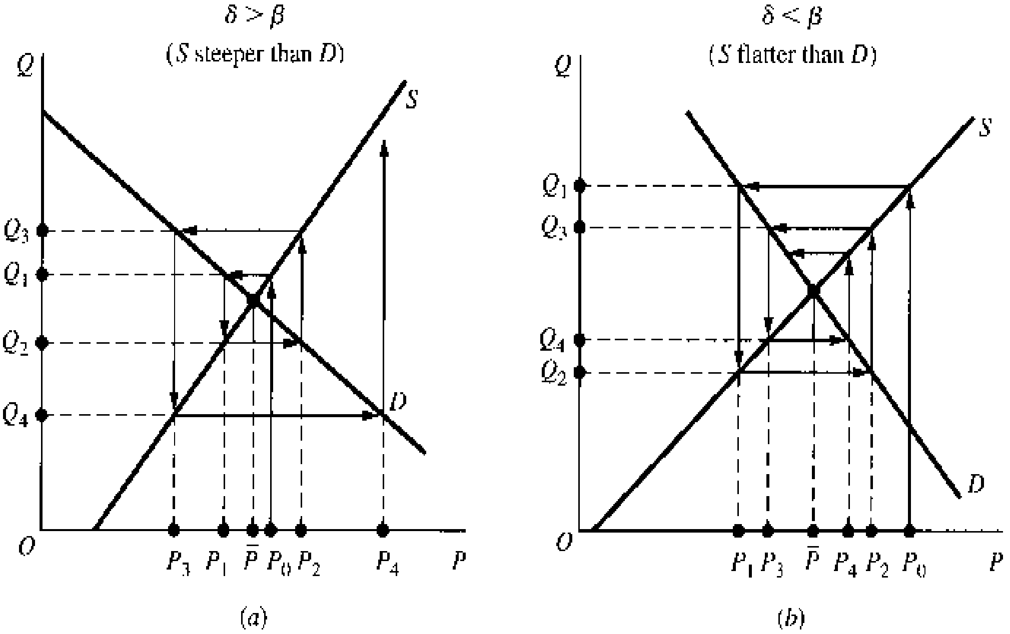
Because this is a constant, it is a stationary equilibrium. Substituting $\overline{P}$ into our solution, we can express the time path P_t alternatively in the form

$$P_t = (P_0 - \overline{P})\left(-\frac{\delta}{\beta}\right)^t + \overline{P} \tag{17.12′}$$

This leads us to the second point, namely, the significance of the expression $(P_0 - \overline{P})$. Since this corresponds to the constant A in the Ab^t term, its sign will bear on the question of whether the time path will commence above or below the equilibrium (mirror effect), whereas its magnitude will decide how far above or below (scale effect). Lastly, there is the expression $(-\delta/\beta)$, which corresponds to the b component of Ab^t. From our model specification that $\beta, \delta > 0$, we can deduce an oscillatory time path. It is this fact which gives rise to the cobweb phenomenon, as we shall presently see. There can, of course, arise *three*

[†] As far as the market-clearing sense of equilibrium is concerned, the price reached in each period is an equilibrium price, because we have assumed that $Q_{dt} = Q_{st}$ for every t.

FIGURE 17.2



possible varieties of oscillation patterns in the model. According to Table 17.1 or Fig. 17.1, the oscillation will be

$$\left.\begin{matrix}\text{Explosive}\\ \text{Uniform}\\ \text{Damped}\end{matrix}\right\} \quad \text{if} \quad \delta \gtreqless \beta$$

where the term *uniform oscillation* refers to the type of path in region VI.

In order to visualize the cobwebs, let us depict the model (17.10) in Fig. 17.2. The second equation of (17.10) plots as a downward-sloping linear demand curve, with its slope numerically equal to β. Similarly, a linear supply curve with a slope equal to δ can be drawn from the third equation, if we let the Q axis represent in this instance a *lagged* quantity supplied. The case of $\delta > \beta$ (S steeper than D) and the case of $\delta < \beta$ (S flatter than D) are illustrated in Fig. 17.2*a* and *b*, respectively. In either case, however, the intersection of D and S will yield the intertemporal equilibrium price $\overline{P}$.

When $\delta > \beta$, as in Fig. 17.2*a*, the interaction of demand and supply will produce an explosive oscillation as follows. Given an initial price P_0 (here assumed above $\overline{P}$), we can follow the arrowhead and read off on the S curve that the quantity supplied in the next period (period 1) will be Q_1. In order to clear the market, the quantity demanded in period 1 must also be Q_1, which is possible if and only if price is set at the level of P_1 (see downward arrow). Now, via the S curve, the price P_1 will lead to Q_2 as the quantity supplied in period 2, and to clear the market in the latter period, price must be set at the level of P_2 according to the demand curve. Repeating this reasoning, we can trace out the prices and quantities in subsequent periods by simply following the arrowheads in the diagram, thereby spinning a "cobweb" around the demand and supply curves. By comparing the price levels, $P_0, P_1, P_2, \ldots,$ we observe in this case not only an oscillatory pattern of change but also a tendency for price to widen its deviation from $\overline{P}$ as time goes by. With the cobweb being spun from inside out, the time path is divergent and the oscillation explosive.

By way of contrast, in the case of Fig. 17.2*b*, where $\delta < \beta$, the spinning process will create a cobweb which is centripetal. From P_0, if we follow the arrowheads, we shall be led ever closer to the intersection of the demand and supply curves, where $\overline{P}$ is. While still oscillatory, this price path is convergent.

In Fig. 17.2 we have not shown a third possibility, namely, that of $\delta = \beta$. The procedure of graphical analysis involved, however, is perfectly analogous to the other two cases. It is therefore left to you as an exercise.

The preceding discussion has dealt only with the time path of P (that is, P_t); after P_t is found, however, it takes but a short step to get to the time path of Q. The second equation of (17.10) relates Q_{dt} to P_t, so if (17.12) or (17.12′) is substituted into the demand equation, the time path of Q_{dt} can be obtained immediately. Moreover, since Q_{dt} must be equal to Q_{st} in each time period (clearance of market), we can simply refer to the time path as Q_t rather than Q_{dt}. On the basis of Fig. 17.2, the rationale of this substitution is easily seen. Each point on the D curve relates a P_t to a Q_t pertaining to the same time period; therefore, the demand function can serve to map the time path of price into the time path of quantity.

You should note that the graphical technique of Fig. 17.2 is applicable even when the D and S curves are nonlinear.

EXERCISE 17.4

1. On the basis of (17.10), find the time path of Q, and analyze the condition for its convergence.
2. Draw a diagram similar to those of Fig. 17.2 to show that, for the case of $\delta = \beta$, the price will oscillate uniformly with neither damping nor explosion.
3. Given demand and supply for the cobweb model as follows, find the intertemporal equilibrium price, and determine whether the equilibrium is stable:
   (*a*) $Q_{dt} = 18 - 3P_t \qquad Q_{st} = -3 + 4P_{t-1}$
   (*b*) $Q_{dt} = 22 - 3P_t \qquad Q_{st} = -2 + P_{t-1}$
   (*c*) $Q_{dt} = 19 - 6P_t \qquad Q_{st} = 6P_{t-1} - 5$
4. In model (17.10), let the $Q_{dt} = Q_{st}$ condition and the demand function remain as they are, but change the supply function to
   $Q_{st} = -\gamma - \delta P_t^*$
   where P_t^* denotes the *expected price* for period t. Furthermore, suppose that sellers have the "adaptive" type of price expectation:†
   $P_t^* = P_{t-1}^* + \eta(P_{t-1} - P_{t-1}^*) \qquad (0 < \eta \le 1)$
   where η (the Greek letter eta) is an expectation-adjustment coefficient.
   (*a*) Give an economic interpretation to the preceding equation. In what respects is it similar to, and different from, the adaptive expectations equation (16.34)?
   (*b*) What happens if η takes its maximum value? Can we consider the cobweb model as a special case of the present model?

† See Marc Nerlove, "Adaptive Expectations and Cobweb Phenomena," *Quarterly Journal of Economics*, May 1958, pp. 227–240.

(*c*) Show that the new model can be represented by the first-order difference equation

$$P_{t+1} - \left(1 - \eta - \frac{\eta\delta}{\beta}\right) P_t = \frac{\eta(\alpha + \gamma)}{\beta}$$

(*Hint:* Solve the supply function for P_t^*, and then use the information that $Q_{st} = Q_{dt} = \alpha - \beta P_t$.)

(*d*) Find the time path of price. Is this path necessarily oscillatory? Can it be oscillatory? Under what circumstances?

(*e*) Show that the time path P_t, if oscillatory, will converge only if $1 - 2/\eta < -\delta/\beta$. As compared with the cobweb solution (17.12) or (17.12′), does the new model have a wider or narrower range for the stability-inducing values of $-\delta/\beta$?

5. The cobweb model, like the previously encountered dynamic market models, is essentially based on the static market model presented in Sec. 3.2. What economic assumption is the dynamizing agent in the present case? Explain.

17.5 A Market Model with Inventory

In the preceding model, price is assumed to be set in such a way as to clear the current output of every time period. The implication of that assumption is either that the commodity is a perishable which cannot be stocked or that, though it is stockable, no inventory is ever kept. Now we shall construct a model in which sellers do keep an inventory of the commodity.

The Model

Let us assume the following:

1. Both the quantity demanded, Q_{dt}, and the quantity currently produced, Q_{st}, are unlagged linear functions of price P_t.
2. The adjustment of price is effected not through market clearance in every period, but through a process of price-setting by the sellers: At the beginning of each period, the sellers set a price for that period after taking into consideration the inventory situation. If, as a result of the preceding-period price, inventory accumulated, the current-period price is set at a lower level than before, in order to "move" the merchandise; but if inventory decumulated instead, the current price is set higher than before.
3. The price adjustment made from period to period is inversely proportional to the observed change in the inventory (stock).

With these assumptions, we can write the following equations:

$$\begin{aligned} Q_{dt} &= \alpha - \beta P_t && (\alpha, \beta > 0) \\ Q_{st} &= -\gamma + \delta P_t && (\gamma, \delta > 0) \\ P_{t+1} &= P_t - \sigma(Q_{st} - Q_{dt}) && (\sigma > 0) \end{aligned} \tag{17.13}$$

where σ denotes the *stock-induced-price-adjustment* coefficient. Note that (17.13) is really nothing but the discrete-time counterpart of the market model of Sec. 15.2, although we have now couched the price-adjustment process in terms of *inventory* ($Q_{st} - Q_{dt}$) rather than *excess demand* $(Q_{dt} - Q_{st})$. Nevertheless, the analytical results will turn out to be much different; for one thing, with discrete time, we may encounter the phenomenon of oscillations. Let us derive and analyze the time path P_t.

The Time Path

By substituting the first two equations into the third, the model can be condensed into a single difference equation:

$$P_{t-1} - [1 - \sigma(\beta + \delta)]P_t = \sigma(\alpha - \gamma) \tag{17.14}$$

and its solution is given by (17.8′):

$$\begin{aligned} P_t &= \left(P_0 - \frac{\alpha + \gamma}{\beta + \delta}\right)[1 - \sigma(\beta + \delta)]^t + \frac{\alpha + \gamma}{\beta - \delta} \\ &= (P_0 - \overline{P})[1 - \sigma(\beta + \delta)]^t + \overline{P} \end{aligned} \tag{17.15}$$

Obviously, therefore, the dynamic stability of the model will hinge on the expression $1 - \sigma(\beta + \delta)$; for convenience, let us refer to this expression as b.

With reference to Table 17.1, we see that, in analyzing the exponential expression b^t, seven distinct regions of b values may be defined. However, since our model specifications $(\sigma, \beta, \delta > 0)$ have effectually ruled out the first two regions, there remain only five possible cases, as listed in Table 17.2. For each of these regions, the b specification of the second column can be translated into an equivalent σ specification, as shown in the third column. For instance, for region III, the b specification is $0 < b < 1$; therefore, we can write

$$0 < 1 - \sigma(\beta + \delta) < 1$$

$$-1 < -\sigma(\beta + \delta) < 0 \qquad \text{[subtracting 1 from all three parts]}$$

and $$\frac{1}{\beta + \delta} > \sigma > 0 \qquad \text{[dividing through by } -(\beta + \delta)\text{]}$$

TABLE 17.2
Types of Time Path

Region	Value of $b = 1 - \sigma(\beta + \delta)$	Value of σ	Nature of Time Path P_t
III	$0 < b < 1$	$0 < \sigma < \frac{1}{\beta + \delta}$	Nonoscillatory and convergent
IV	$b = 0$	$\sigma = \frac{1}{\beta + \delta}$	Remaining in equilibrium†
V	$-1 < b < 0$	$\frac{1}{\beta + \delta} < \sigma < \frac{2}{\beta + \delta}$	With damped oscillation
VI	$b = -1$	$\sigma = \frac{2}{\beta + \delta}$	With uniform oscillation
VII	$b < -1$	$\sigma > \frac{2}{\beta + \delta}$	With explosive oscillation

† The fact that price will be remaining in equilibrium in this case can also be seen directly from (17.14). With $\sigma = 1/(\beta + \delta)$, the coefficient of P_t becomes zero, and (17.14) reduces to $P_{t+1} = \sigma(\alpha + \gamma) = (\alpha + \gamma)/(\beta + \delta) = \overline{P}$.

This last gives us the desired equivalent σ specification for region III. The translation for the other regions may be carried out analogously. Since the type of time path pertaining to each region is already known from Fig. 17.1, the σ specification enables us to tell from given values of σ, β, and δ the general nature of the time path P_t, as outlined in the last column of Table 17.2.

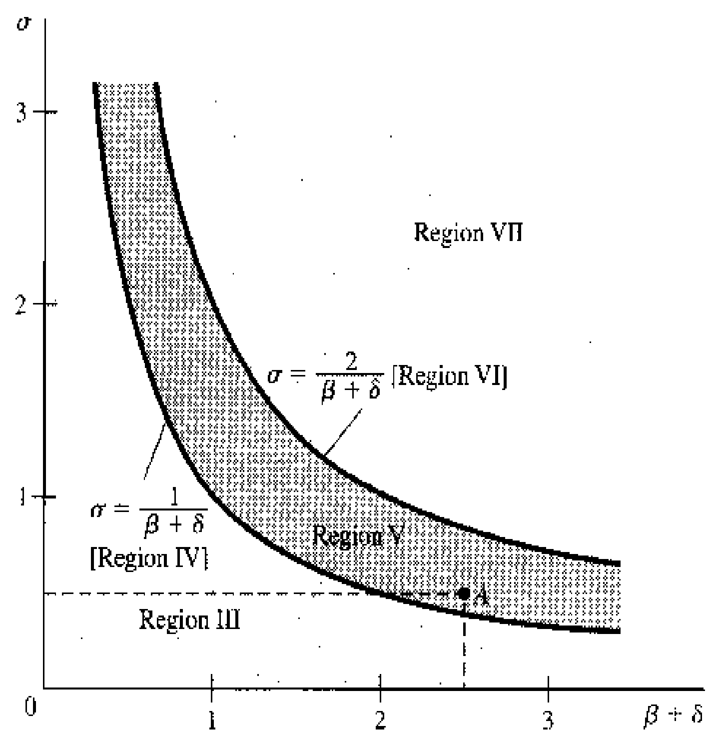
Example 1 If the sellers in our model always increase (decrease) the price by 10 percent of the amount of the decrease (increase) in inventory, and if the demand curve has a slope of −1 and the supply curve a slope of 15 (both slopes with respect to the price axis), what type of time path P_t will we find?

Here, we have $\sigma = 0.1$, $\beta = 1$, and $\delta = 15$. Since $1/(\beta + \delta) = \frac{1}{16}$ and $2/(\beta + \delta) = \frac{1}{8}$, the value of σ $(= \frac{1}{10})$ lies between the former two values; it is thus a case of region V. The time path P_t will be characterized by damped oscillation.

Graphical Summary of the Results

The substance of Table 17.2, which contains as many as five different possible cases of σ specification, can be made much easier to grasp if the results are presented graphically. Inasmuch as the σ specification involves essentially a comparison of the relative magnitudes of the parameters σ and $(\beta + \delta)$, let us plot σ against $(\beta + \delta)$, as in Fig. 17.3. Note that we need only concern ourselves with the positive quadrant because, by model specification, σ and $(\beta + \delta)$ are both positive. From Table 17.2, it is clear that regions IV and VI are specified by the equations $\sigma = 1/(\beta + \delta)$ and $\sigma = 2/(\beta + \delta)$, respectively. Since each of these plots as a rectangular hyperbola, the two regions are graphically represented by the two hyperbolic curves in Fig. 17.3. Once we have the two hyperbolas, moreover, the other three regions immediately fall into place. Region III, for instance, is merely the set of points lying below the lower hyperbola, where we have σ less than $1/(\beta + \delta)$. Similarly, region V is represented by the set of points falling between the two hyperbolas, whereas all the points located above the higher hyperbola pertain to region VII.

FIGURE 17.3



Example 2 If $\sigma = \frac{1}{2}$, $\beta = 1$, and $\delta = \frac{3}{2}$, will our model (17.13) yield a convergent time path P_t? The given parametric values correspond to point A in Fig. 17.3. Since it falls within region V, the time path is convergent, though oscillatory.

You will note that, in the two models just presented, our analytical results are in each instance stated as a set of alternative possible cases—three types of oscillatory path for the cobwebs, and five types of time path in the inventory model. This richness of analytical results stems, of course, from the parametric formulation of the models. The fact that our result cannot be stated in a single unequivocal answer is, of course, a merit rather than a weakness.

EXERCISE 17.5

1. In solving (17.14), why should formula (17.8′) be used instead of (17.9′)?
2. On the basis of Table 17.2, check the validity of the translation from the b specification to the σ specification for regions IV through VII.
3. If model (17.13) has the following numerical form:

   $$Q_{dt} = 21 - 2P_t$$
   $$Q_{st} = -3 + 6P_t$$
   $$P_{t+1} = P_t - 0.3(Q_{st} - Q_{dt})$$

   find the time path P_t and determine whether it is convergent.
4. Suppose that, in model (17.13), the supply in each period is a fixed quantity, say, $Q_{st} = k$, instead of a function of price. Analyze the behavior of price over time. What restriction should be imposed on k to make the solution economically meaningful?

17.6 Nonlinear Difference Equations—The Qualitative-Graphic Approach

Thus far we have only utilized *linear* difference equations in our models; but the facts of economic life may not always acquiesce to the convenience of linearity. Fortunately, when nonlinearity occurs in the case of first-order difference-equation models, there exists an easy method of analysis that is applicable under fairly general conditions. This method, graphic in nature, closely resembles that of the qualitative analysis of first-order differential equations presented in Sec. 15.6.

Phase Diagram

Nonlinear difference equations in which only the variables y_{t+1} and y_t appear, such as

$$y_{t+1} + y_t^3 = 5 \qquad \text{or} \qquad y_{t+1} + \sin y_t - \ln y_t = 3$$

can be categorically represented by the equation

$$y_{t+1} = f(y_t) \tag{17.16}$$

where f can be a function of any degree of complexity, as long as it is a function of y_t alone without t as another argument. When the two variables y_{t+1} and y_t are plotted against each